2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

A Model-free Vision-based Robot Control for Minimally Invasive Surgery using ESM Tracking and Pixels Color Selection

Frédéric Bourger, Christophe Doignon, Philippe Zanne and Michel de Mathelin

Abstract— **This paper deals with the visual servoing of textured surfaces inside the human abdomen with a laparoscope for the robot-assisted minimally invasive surgery (MIS). The well-known image-based visual servoing (IBVS) is one of the most common approach used for model-based servoing. When no CAD model is available, the Efficient Second-order Minimization (ESM) tracking developed by Malis ([14], [16]) for grey-level images is one of the powerful recent techniques which is extended here to color images so as to handle occluded parts of the region of interest (ROI). Firstly, the ROI is splitted into small areas and an histogram-based color feature comparison of image areas is presented. For each frame and for each area, a metric based on the Bhattacharyya criterion is used to select the contributing areas for the computation of the planar homography between views. Secondly, since for any MIS technique, the endoscopic lens is passing through an insertion point on the abdominal wall, a specific control strategy is developed to perform the ESM tracking with a 4-DOF surgical robot. The method presented in this paper has been validated with several video sequences. Experimental results show that the tracking method is efficient even with more than 75 % of the tracked ROI occluded. Finally, the model-free visual servoing has been performed with the AESOP surgical robot and a training box. Even if the convergence rate is a little bit slow, the desired region is always reached.**

I. INTRODUCTION

For three decades, the field of vision-based robotics has been widely grown, and more and more complex 3D scenes are within robot reach due to deeper understandings of the scene perception, the increase of computer capabilities and control theory. Although the binary machine vision and "block's world" were the ingredients of pioneered works, real-time color images processing of natural 3D scene is becoming a more and more common tool to offer a reliable solution to a wide spectrum of applications. For application areas like medical robotics, mobile robotics, micro-robotic manipulation, agricultural automation for example, the achievement needs the integration of several research areas in computer vision and automatic control ([10], [17]).

The main challenge in visual tracking for eye-in-hand robotic purposes is to catch the necessary video information from images of a target region of the scene for determining the camera velocity. When a CAD model of a viewed rigid object in the scene is available, the position-based visual servoing (PBVS) is an appropriate family of vision-based control methods for which the 3D pose of the target region is computed at each frame. Geometric features extraction and matching are needed and usually bring high computational costs. Image-based visual servoing (IBVS) is generally more accurate since the target region is directly expressed in the image plane. It allows to compute the relative 3D motion from image feature variabilities [10].In the case of a rigid object of interest, relations which occur between the velocity screw and the position variations of image features lead to the computation of a Jacobian matrix called interaction matrix [5]. This matrix may be on-line estimated all along the servoing or only at the desired localization of image features. Contrarily to the PBVS approach, the pose may be partially estimated, and only the depth is required to update the interaction matrix components. However, the depth estimation is often noisy and a servoing which uses a pre-computed depth works quite well provided the convergence is ensured. Hybrid or 2D 1/2 visual servoing methods are based on both 2D video information and 3D projective reconstruction. They need less knowledge of the 3D scene and they solve problems that may appear with IBVS when the initial position is far from the desired one [15]. The 3-D features (often a depth ratio) can be retrieved, thus the robustness of the servoing w.r.t. calibration errors is improved. Moreover, the block-triangular shape of the Jacobian matrix tends this latter approach to be well-adapted to the decoupling of position and orientation controls.

The efficiency of most of model-based visual servoing approaches relies on correspondences between the position of tracked visual features in the current image and their position in the reference image. If these correspondences contain errors then the servoing usually fails or converges upon a wrong position. Overcoming these errors is often achieved by improving the quality of tracking algorithms and features selection methods ([12], [18]).

Model-free (object) methods do not need any 3D geometric cue, but rather, similarly to the visual tracking, the goal is to determine the movement w.r.t. an observed target over a long sequence of images. However, although most of visual tracking methods are able to compute the apparent motion only, the design of motion models allow to recover the 3D motion between views and leads to more efficient tracking algorithms [7].

A region tracking with linear 2D motion models has been proposed by Hager and Belhumeur [7]. Their method is based on the image constancy constraint with a basis of training images and their work integrates both geometric and photometric image changes. To take care of occlusions, they also developed a regression technique based on M-

F. Bourger, C. Doignon, Ph. Zanne and M. de Mathelin are with the LSIIT (UMR ULP-CNRS 7005), University of Strasbourg, Boulevard Brant, 67412 Illkirch, France. e-mail: {name}@lsiit.u-strasbg.fr

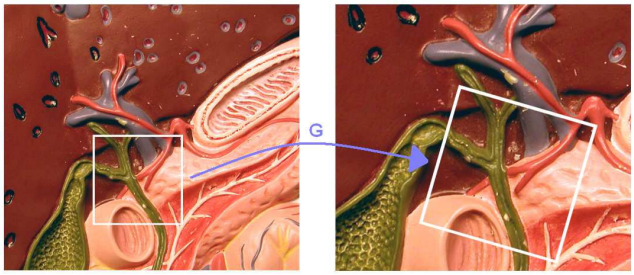


Fig. 1. The geometric transformation G - a planar homography - between the region of interest in the reference image (white square, left) and in the current image (right).

estimators that is occlusions are viewed as outliers by the algorithm. The Efficient Second-order Minimization visual tracking proposed by Malis ([14], [16]) extends the region tracking to the 3D rigid motion models of planar surfaces. Visual tracking and servoing are fused in a unifying framework which does not need any measurement of the 3D structure of the observed target. The ESM method uses both reference and current image regions of interest (see Fig. 1) and their gradients in order to estimate the velocity screw of a moving camera. A judicious decomposition of the planar homography between the images with the Lie algebra (firstly introduced by Drummond and Cipolla [4]) allows to minimize a cost function and avoids the costly computation of the Hessian matrix. The cost function is the sum of squared differences (SSD) of the intensity distributions.

There are many sources of image variabilities, like the deformation of the tracked object or changes in illumination which can cause the 3D motion to be misestimated. Although the results obtained by Malis are good even in presence of the above mentioned disturbances, the SSD is a global comparison which is very sensitive to occlusions in any way. For the application of concern, occlusions of organs by other internal parts or by surgical instruments are very harmful while performing an accurate instruments positioning and they sometimes lead the target to be lose. It is then of prime importance to tackle this problem in order to carry out a reliable tracking inside the human body. Based on these considerations, the objectives of the paper are twofold:

- To enhance the visual servoing through the automatic detection of irrelevant pixels and to reject them prior to the tracking process,
- To integrate the motion constraint of the minimally invasive surgery while controlling a robot thanks to a model-free visual servoing.

The paper is organized as follows. In the next section, the proposed method for pixels color selection is described. In section III, the model-free visual servoing control approach with a 4-DOF robot is presented. We provide in section IV a set of results from several video sequences and primary results of the visual servoing with a surgical robot. In the final section, we conclude the paper.

II. A joint Hue Saturation-based Statistical Pixels Selection

The idea we have developed in this section consists in splitting the target region of interest (ROI) both in the reference image and in the current image into a set of small pixel areas so as to select those which can contribute to the homography computation in the ESM tracking. Hence, every area in the ROI of the reference image is compared to its corresponding one in the current image in order to detect an occlusion. Thanks to the on-line computation of the planar homography in the ESM, the ROI in the current image is warped so as to be aligned with the ROI in the reference image. By this way, the original ESM is not modified, it's only working with some parts of ROI and various methods for the selection of pixels can be investigated.

Statistical methods are commonly used in the motion detection for video surveillance applications or data compression. Some of them involve a low-level processing with real-time capabilities, an undoubtedly factor of success while controlling the robot motion. With the geo-pixel statistical method [9], each pair of pixel areas (i, j) in two images (reference and current images) is characterized with a likehood ratio L_{ij}. It's a metric based on second-order statistics. This method has been carried out on intensity images. Its major drawbacks are its computational middle-level, the presence of a user-defined threshold and an inaccurate value of the likelihood ratio when the variances are very low.

In an attempt to reduce the latter drawbacks, we incorporate an histogram-based matching method to the pixel selection algorithm. The two discrete distributions associated to area a_i in the reference image and area a_j in the current image are compared through a similarity metric function based on the Bhattacharyya coefficient B_{ij} ([3], [11]). As noticed by Comaniciu *et al.* [2], this coefficient is derived from the Bhattacharyya coefficient which is not a metric in its original form. To do this, we have chosen the histogram as a usually satisfactory nonparametric density estimate. Thus, for a pixel area a_i, the discrete density is

$$\hat{\mathbf{a}}_i = \{\hat{a}_i^{(l)}\}_{l=1,\ldots,m} \quad \text{with} \quad \sum_{l=1}^{m} \hat{a}_i^{(l)} = 1 \qquad (1)$$

where $\hat{a}_i^{(l)}$ is the l^{th} bin of the m-bins histogram of pixel area a_i. The statistical distance between the distributions of two areas a_i and a_j is then defined by

$$d_{ij} = \sqrt{1 - B_{ij}} \qquad (2)$$

with

$$B_{ij} = \sum_{l=1}^{m} \sqrt{\hat{a}_i^{(l)}\ \hat{a}_j^{(l)}}\ . \qquad (3)$$

Minimizing the distance (2) is equivalent to maximizing the Bhattacharyya coefficient (3) and the area in the current ROI is assumed to be occluded if d exceeds a threshold value τ_b.

When these two methods are carried out on intensity images, they are inexorably sensitive to global changes in brightness or non-uniform illumination, and it is hard to distinguish an occluded part from brightness changes or the appearance of specularities. The image gradient is a better choice since it brings texture information. It's what the ESM does in the original form but for the whole target region of interest. Nowadays, many vision sensors integrate well-coded video color information, and chromatic attributes are salient frequency markers and have often been used in the past for vision-based robot control, especially in natural environment with mobile robot but also in laparoscopiy with endoscope-holder robot. Wei *et al.* [21] have used a stereo-laparoscope mounted on a robot arm and have designed a color marker to realize a tracking task. By means of a color histogram, the color bin with the lowest value was selected to mark the instrument. This spectral mark was then utilized to control the robot motion at a sampling rate of 15 Hz. An interesting feature of this technique is the choice of the color space, HSV, for the segmentation leading to a good robustness w.r.t. lighting variations. Wang *et al.* [20] and Uecker *et al.* [19] have conceived a general framework that uses visual modelling and servoing methods to assist the surgeon in manipulating a laparoscope mounted on a robot end-effector. Color signatures are used in a Bayesian classifier to segment images into two classes (organ and markerless instrument). Finally, this has been applied to the instrument localization (the 2D position of the imaged tip of instrument) and 2D tracking with 3 DOFs of the AESOP robot in a way to follow the laparoscope.

Following this purpose, color hue and saturation seem to be appropriate discriminant attributes. A low saturation value indicates a low colored pixel and a high value corresponds to a purely colored pixel. However, it's well-known that the color saturation is affected by surface reflectance and the main drawback of hue is its undefinedness for achromatic pixels and small changes round the gray axis result in large changes of the hue value. With regard to these works, the following joint hue saturation color feature we used for the histogram-based matching of pixel areas is defined as

$$\mathcal{H} = S_0 \cos(H) \tag{4}$$

$$\text{with} \quad H = \arccos\left(\frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right) \tag{5}$$

and S_0 is the modified version of the color saturation $S_0 = \max\{R,G,B\} - \min\{R,G,B\}$ due to Hanbury and Serra [8] which is invariant to any shift of an arbitrary RGB vector parallel to the achromatic axis.

III. Robot Control with the MIS constraint

As a MIS technique, the laparoscopic surgery involves motions of surgical instruments which are constrained to by insertion points in the abdominal wall, thus reducing the mobility down to 4 DOF [1] [6]. For visual feedback, an endoscopic vision system is used for achieving the intra-operative guidance. It includes an optical lens which is a laparoscope. It's a long pipe with monovision or (short base-line) stereo-vision capabilities and light transportation/emission. For the application of interest, the laparoscope (E) is mounted on the robot end-effector (7) and such configuration is named "eye-in-hand" (see Fig. 2). In the sequel, points O_A and O_C are respectively the insertion point and the origin of the camera frame (E). O_A is assumed to be motionless although its position may be disturbed by the patient breathing.

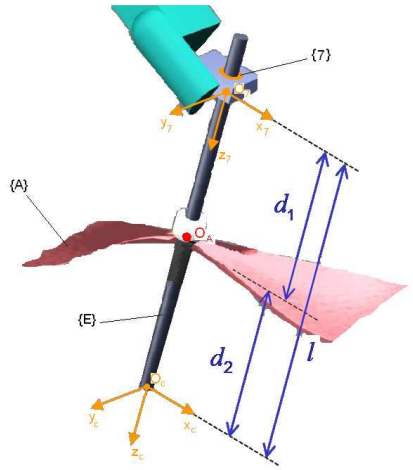


Fig. 2. Geometric modeling and motion constraint in laparoscopic surgery with a "eye-in-hand" robot vision system.

To control the robot with visual feedback, a task function $\mathbf{e} = (\mathbf{e}_\nu, \mathbf{e}_\omega) \in \mathbb{R}^6$ is built similarly to [1]. With a distortion-corrected laparoscope, the task function may be expressed as

$$\mathbf{e}_\nu = (\mathtt{H} \; - \; \mathtt{I}_{3\times3}) \; \mathbf{m}^\star \qquad \text{and} \qquad [\mathbf{e}_\omega]_\times = \mathtt{H} - \mathtt{H}^\mathsf{T} \tag{6}$$

where $\mathbf{m}^\star$ and $\mathtt{H}\,\mathbf{m}^\star$ are corresponding points respectively in the ROI of the reference image and in the current one and $[\mathbf{v}]_\times$ is the skew-symmetric matrix of vector $\mathbf{v}$. With a calibrated device characterized by the intrinsic parameters matrix $\mathtt{K}$, $\mathtt{H}$ is computed through the ESM tracking and is related to the planar homography $\mathtt{G} = \mathtt{K}\,\mathtt{H}\,\mathtt{K}^{-1}$. The decoupling between the position error $\mathbf{e}_\nu$ and the orientation error $\mathbf{e}_\omega$ allows to design a control law with independent exponential decrease behaviours when considering the relation between the derivative of the task function and the velocity screw $\boldsymbol{\tau}_c = \left[\boldsymbol{\nu}^\mathsf{T} = (v_x \; v_y \; v_z) \; , \; \boldsymbol{\omega}^\mathsf{T} = (w_x \; w_y \; w_z)\right]^\mathsf{T}$

$$(\boldsymbol{\tau}_c)^{\mathsf{esm}}_{/O_C} = \begin{bmatrix} \boldsymbol{\nu} \\ \boldsymbol{\omega} \end{bmatrix} = - \begin{bmatrix} \lambda_\nu \; \mathtt{I}_{3\times3} & \mathtt{0}_{3\times3} \\ \mathtt{0}_{3\times3} & \lambda_\omega \; \mathtt{I}_{3\times3} \end{bmatrix} \begin{bmatrix} \mathbf{e}_\nu \\ \mathbf{e}_\omega \end{bmatrix} \tag{7}$$

λ_ν and λ_ω are positive real scalars. Malis and Benhimane have given the proof of the local stability of this law in [1]. To ensure no translational motion occurs in the tangent plane

[1] In MIS, some laparoscopic instruments are provided with extensive range of articulations designed at the tip (microwrist) to increase the dexterity.

at the insertion point, the velocity screw expressed at that point should be limited to 4 DOF as follows

$$(\boldsymbol{\tau}_c)^{4\ \text{dof}}_{/O_A} = \begin{bmatrix} 0 & 0 & V_z & \Omega_x & \Omega_y & \Omega_z \end{bmatrix}^{\mathsf{T}} . \quad (8)$$

This means that only pan/tilt/spin and insertion are the allowed motions. Considering the computation of the velocity screw given by (7) at the insertion point O_A, it comes

$$(\boldsymbol{\tau}_c)^{\text{esm}}_{/O_A} = \begin{bmatrix} v_x - d_2\omega_y & v_y + d_2\omega_x & v_z & \omega_x & \omega_y & \omega_z \end{bmatrix}^{\mathsf{T}} \quad (9)$$

where the scalar d_2 is the penetration depth (see Fig. 2). By identification of equations (8) and (9), one has

$$\begin{cases} v_x - d_2\omega_x = 0 & \omega_x \equiv \Omega_x \\ v_y + d_2\omega_x = 0 & \omega_y \equiv \Omega_y \\ v_z \equiv V_z & \omega_z \equiv \Omega_z \end{cases} \quad (10)$$

While the spin Ω_z and penetration V_z velocities control are straightforward, there is no any guarantee that both equations $v_x - d_2\omega_x = 0$ and $v_y + d_2\omega_x = 0$ are verified. Then, the goal is to control the two remainder parameters Ω_x and Ω_y so as to avoid translational motions at the insertion point and to simultaneously track the target region. There is no way to solve this problem unless a trade-off can be planned since a target region (like the organ surface) can be moved without constraint. The solution we propose comes from practical and safety considerations. Due to the high level of scaling with laparoscopes and the lever effect, target translational motions induce more significant displacements in the image than rotational ones (the latter, rather, lead to perspective distortions). This means that the tracking of the target centre is of prime importance when one wish to keep the target in the field of view. Following this purpose, the solution we propose is to express the velocity screw components Ω_x and Ω_y as functions of v_x and v_y, that is

$$\begin{bmatrix} V_z \\ \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & k_d & 0 & 0 & 0 \\ 0 & -\frac{k_p}{d_2} & 0 & 0 & 0 & 0 \\ \frac{k_t}{d_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_s \end{bmatrix} (\boldsymbol{\tau}_c)^{\text{esm}}_{/O_C} . \quad (11)$$

Finally, by expressing the velocity screw at the end-effector frame origin O_7, one obtains

$$(\boldsymbol{\tau}_c)^{4\ \text{dof}}_{/O_7} = \begin{bmatrix} \frac{-k_t}{l-d_1}\, d_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{k_p}{l-d_1}\, d_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_d & 0 & 0 & 0 \\ 0 & \frac{-k_p}{l-d_1} & 0 & 0 & 0 & 0 \\ \frac{k_t}{l-d_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_s \end{bmatrix} (\boldsymbol{\tau}_c)^{\text{esm}}_{/O_C} \quad (12)$$

with $l = d_1 + d_2$ is a (constant) distance. In [13], two methods for the estimation of d_1 has been provided.

IV. Results

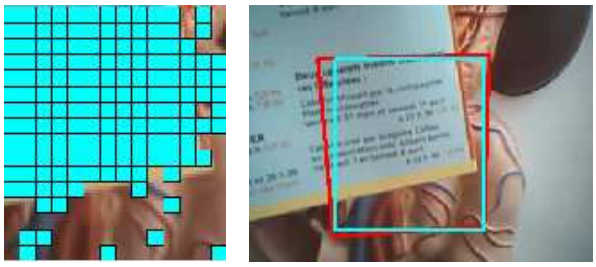
A. ESM tracking with pixels color selection

Concerning the detection of occlusions in the region of interest, the proposed method has been tested by means of a set of video sequences of a motionless target (a maquette of abdominal organs) while a moving sheet is bringing in front of. In figure 5, frames 10, 15, 20, 25 and 30 of a such video sequences sample have been displayed. Images in the first row correspond to the ESM tracking without the detection of occlusions while those of the second and third rows correspond to the tracking with the pixels color selection algorithm, for a threshold $\tau_b = 0.5$. The occlusions occur between frames 10 and 30. In the third row, blue filled squares mark the detected occlusions inside the ROI. Although the algorithm performs well on most images, one can observe that sometimes some areas have been selected as occlusions while there is no occluding part (see, e.g., frame 30). This is probably due to the poor-coded (YUV) color webcam used for experiments, then these areas are viewed as "outliers". Nevertheless, the unoccluded areas give satisfactory results while computing the collineation $\mathtt{H}$ since the current warped ROI (in red) is still well-aligned with the ROI in the reference image (in blue).

A numerical comparison is provided in order to assess the detection of occlusions. To do so, we have defined the algebraic error $e_{occl}(t)$ based on the Frobenius norm so as to compare the current estimate of the homography (between the current ROI and that of the reference image) while the occlusions may occur, $\mathtt{G}(t)$, and without occlusions, $\mathtt{G}^\star(t)$ ($\mathtt{G}^\star(t) = \mathtt{G}^\star$ for a motionless target), as follows

$$e_{occl}(t) = \|\mathtt{G}(t)\ \mathtt{G}^\star(t)^{-1} - \mathtt{I}_{3\times 3}\| . \quad (13)$$

The algebraic error has been reported in Fig. 4 both with the detection of occlusions (red curve) and without the



(a) (b)

Fig. 3. Detection of occluded areas for frame 23. (a) 155 of the 196 areas are occluded. (b) One may visually evaluate the homography accuracy (differences between shapes of the quadrilaterals) for this frame.

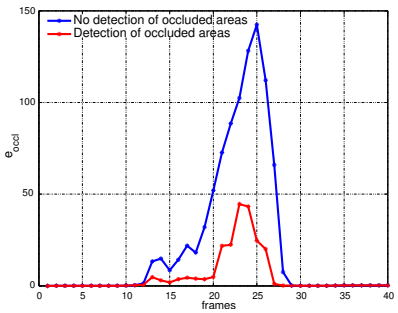


Fig. 4. The algebraic error e_{occl} during a video sequence with (red curve) and without (blue curve) the detection of occluded areas.

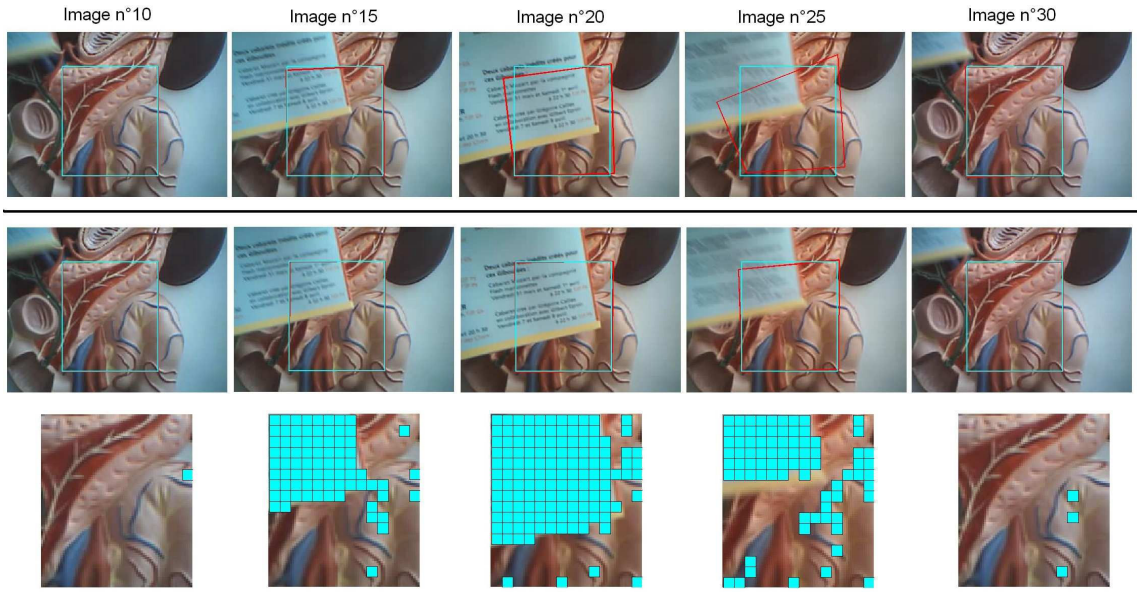


Fig. 5. A comparison of the ESM tracking without the detection of occlusions (top row) and with the pixels color selection we propose (middle row) to detect occluded parts (blue filled square in the bottom row) for a recorded video sequence of a motionless region of interest.

detection of occlusions (blue curve). The maximum error rate is observed for frame 23 where 155 of the 196 areas of the ROI (79 %) are occluded (see Fig. 3). One can observe the benefits of the proposed detection of occlusions prior to the homography computation since the algebraic error found without the detection of occlusions is three times that of with the detection of occlusions. Further ahead an accurate homography computation, we have noticed a significant improvement of tracking performances. The main drawback remains the still high computational cost we expect to reduce for real-time issues thanks to GPU capabilities.

B. Visual Servoing with a Surgical Robot

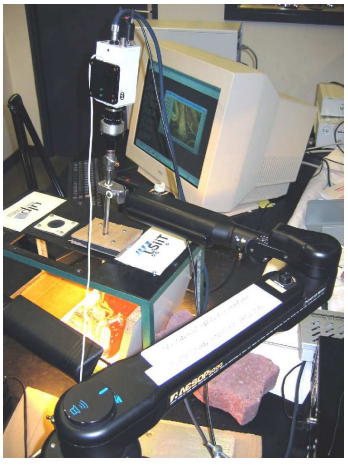
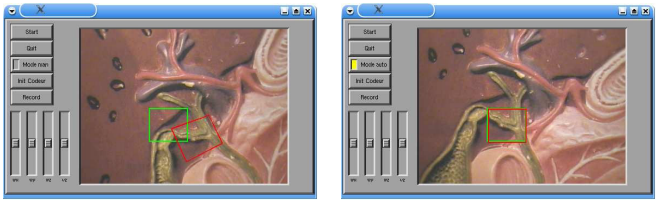


Fig. 6. The laparoscope-holder AESOP robot and the training box used for experiments with the model-free visual servoing.

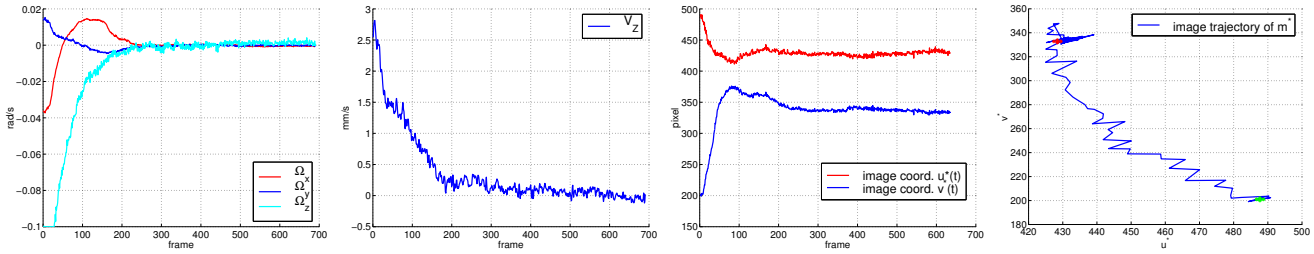
In order to validate the control strategy approach, experiments have been conducted with the AESOP surgical robot designed either for assisting surgeons in manipulating the surgical instruments and as a remote moving laparoscope-holder (see Fig. 6). The experimental setup is composed on a CCD camera connected to the laparoscope, a training box with two insertion points and a maquette of internal organs. The desired ROI (green rectangle) is selected by means of a graphical user interface depicted in Fig. 7. The ESM tracking (without servoing) is activated and the maquette is moved by hand. Finally, the robot is moved (thanks to the 4 slides at lower left side) to another position (red warped quadrilateral) and the visual servoing is executed (see Fig. 7b).



(a) (b)

Fig. 7. (a) The desired ROI is selected (green rectangle), then the ESM tracking (without servoing) is activated and both the maquette and the robot are moved (red rectangle). (b) The convergence of the visual servoing.

Fig. 8a and 8b show time responses of the rotational velocities $(\Omega_x, \Omega_y, \Omega_z)$ and the penetration velocity V_z during the servoing. The gains of the closed loop are set to $(k_p, k_t, k_s, k_d) = (0.3, 0.3, 0.3, 20)$ (subscripts p, t, s, and d stand for pan, tilt, spin and descent). The coordinates of the ROI center ($\mathbf{m}^\star = (u^\star, v^\star)$) and the image trajectory are reported in Fig. 8c and Fig. 8d respectively. Despite the significant amount of noise, the robot position converges to the desired one in approximately 400 images (16 seconds) with the above-mentioned gains.



(a) (b) (c) (d)

Fig. 8. Dynamic behavior of the model-free servoing with the 4-DOF laparoscopic-holder AESOP robot. (a-b) Time responses for the parameters control values. (c) Time responses for the ROI centre coordinates $\mathbf{m}^\star = (u^\star(t), v^\star(t))$. (d) Image trajectory of $\mathbf{m}^\star$.

It seems hard to improve the dynamic behaviour of the global system, since a limit-cycle phenomenon is observed close to the desired position. Non-linearities such as friction or deadzones induce abrupt displacements in the image, then bring inaccuracies in the homography computation.
Another limiting factor of performances is the large time delay of 160 ms with the AESOP robot controller. Furthermore, the experiments were accomplished with a constant value for the distance between the end-effector and the insertion point onto the abdominal wall, d_1, while it is rather a varying parameter. Finally, one has to keep in mind that only 4 DOF are allowed to compensate both the robot motion and the unconstrained moving target.

V. Conclusion

Most of model-based visual servoing methods have considered the problem of outliers rejection either in the image processing step or in the control law. While robust estimations have been designed for both levels, the existing techniques need geometric features to extract so as to match an object model. For model-free visual servoing methods, like the ESM tracking algorithm, we presented in this paper a pixel color selection method based on histograms matching, prior to the tracking itself in order to handle occlusions. . The experiments for the detection of occlusions show the efficiency of the proposed method, we expect to apply for the tracking of locally planar surfaces inside the human body. Finally, a vision-based robot control strategy accounting of the motion constraint in minimally invasive surgery and based on the original ESM has been presented and carried out with a 4-DOF surgical robot. For future work, we plan to integrate the detection of occlusions in the vision-based robotized laparoscopy framework although the ESM is already a high-computational demanding algorithm. This will be a new step towards the simultaneous tracking of surgical instruments and living tissues respectively with model-based and model-free visual servoing approaches.

Acknowledgments
The authors would like to thank Ezio Malis and Selim Benhimane from INRIA for their valuable advices and for providing data and code of the ESM tracking algorithm.

References

[1] S. Benhimane and E. Malis. Homography-based 2d visual servoing. In *IEEE ICRA*, Orlando, USA, May 2006.
[2] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Trans. PAMI*, 25(5), May 2003.
[3] A. Djouadi, O. Snorrason, and F. Garber. The quality of training-sample estimates of the bhattacharyya coefficient. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12:92–97, 1990.
[4] T. Drummond and R. Cipolla. Application of lie algebras to visual servoing. *Int'l Journal of Computer Vision*, 37(1):21–41, June 2000.
[5] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. Robotics and Automation*, 8(3):313–326, 1992.
[6] J. Funda, R.H. Taylor, B. Eldridge, S. Gomory, and K.G. Gruben. Constrained cartesian motion control for teleoperated surgical robots. *IEEE Trans. on Robotics and Automation*, 12(3), June 1996.
[7] G.D. Hager and Peter N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(10):1025–1039, 1998.
[8] A. Hanbury and J. Serra. Colour image analysis in 3d-polar coordinates. In *DAGM congress*, Magdeburg, Germany, 2003.
[9] Y.Z. Hsu, H.H. Nagel, and G. Refers. New likelihood test method for change detection in image sequences. *Computer Vision, Graphics and Image Processing*, 26:73–106, 1984.
[10] S. Hutchinson, G.D. Hager, and P.I. Corke. A tutorial on visual servo control. *IEEE Trans. Robotics and Automation*, 12(5):651–670, 1996.
[11] T. Kailah. The divergence and bhattacharyya distance measures in signal selection. *IEEE Trans. Comm. Technology*, 15:52–60, 1967.
[12] D. Kragic and C. Christensen. Cue integration for visual servoing. *IEEE Trans. on Robotics and Autom.*, 17(1):19–26, 2001.
[13] A. Krupa. Commande par vision d'un robot de chirurgie laparoscopique. *Thèse de doctorat, INP Lorraine, France*, 2003.
[14] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *Proceedings of ICRA*, volume 2, pages 1843– 1848, New Orleans, USA, April 2004.
[15] E. Malis and F. Chaumette. 2 1/2 d visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. *Int'l. Journal of Computer Vision*, 37(1):79–97, 2000.
[16] E. Malis and F. Chaumette. Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. *IEEE Trans. Robotics and Automation*, 18(2):176–186, 2002.
[17] N. Papanikolopoulos, P. Khosla, and T. Kanade. Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision. *IEEE Trans. Rob. and Autom.*, 9(1):14–35, 1993.
[18] T. Tommasini, A. Fusiello, E. Trucco, and V. Roberto. Making good features track better. In *IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pages 178–183, Santa Barbara, USA, June 1998.
[19] D. R. Uecker, C. Lee, Y. F. Wang, and Y. Wang. Automated instrument tracking in robotically-assisted laparoscopic surgery. *Journal of Image Guided Surgery*, 1(6), 1996.
[20] Y. F. Wang, D. R. Uecker, and Y. Wang. A new framework for vision-enabled and robotically assisted minimally invasive surgery. *Journal of Computerized Medical Imaging and Graphics*, 22:429–437, 1998.
[21] G.-Q. Wei, K. Arbter, and G. Hirzinger. Automatic tracking of laparoscopic instruments by color-coding. In Springer Verlag, editor, *Proc. First Int. Joint Conf. CRVMed-MRCAS'97*, pages 357–366, Grenoble, France, March 1997.