Using image gradient as a visual feature for visual servoing

E. Marchand, Christophe Collewet

Using image gradient as a visual feature for visual servoing

Eric Marchand, Christophe Collewet

Abstract— **Direct photometric visual servoing has proved to be an efficient approach for robot positioning. Instead of using classical geometric features such as points, straight lines, pose or an homography, as it is usually done, information provided by all pixels in the image are considered. In the past mainly luminance information has been considered. In this paper, considering that most of the useful information in an image is located in its high frequency areas (that are contours), we have consider various possible combinations of global visual feature based on luminance and gradient. Experimental results are presented to show the behavior of such features.**

I. Overview

Visual servoing consists in using data provided by a vision sensor for controlling the motions of a robot [1]. Classically, to achieve a visual servoing task, a set of visual features has to be selected from the image allowing to control the desired degrees of freedom. For year these features have mainly been geometric features such as points, straight lines, pose [1]. Robust extraction and real-time spatio-temporal tracking of visual cues is then usually one of the keys to success of a visual servoing task.

More global information can be considered such as 2D motion model [2]–[5]. Nevertheless, such approaches require a complex image processing step. Removing the entire matching process is only possible when closely tie the image to the control. Considering the whole image as a feature has previously been studied [6]–[10]. In [9], the visual features are directly the image luminance. The control is then achieve by minimizing the error between the current and the desired image. The key point of this approach relies on the analytic computation of the interaction matrix that links the time variation of the luminance to the camera motions. In [6], [7] an eigenspace decomposition is performed to reduce the dimensionality of image data. The control is then performed in the eigenspace and not directly with the image intensity. In that case the interaction matrix related to the eigenspace is not computed analytically but learned during an off-line step. In [10], visual servoing uses directly the information (as defined by Shannon) contained in the image. A metric derived from information theory (mutual information) is used as a visual feature for visual servoing us to build the control law. [8] also consider the pixels intensity. This approach is based on the use of kernel methods that lead to a high decoupled control law but only four degrees of freedom are considered.

Eric Marchand is with Université de Rennes 1, IRISA UMR 6074, INRIA Rennes-Bretagne Atlantique, Lagadic research group, email: Eric.Marchand@irisa.fr, Christophe Collewet is with Cemagref, INRIA Rennes-Bretagne Atlantique, Fluminance research group, email: Christophe.Collewet@cemagref.fr

In visual servoing, a control law is designed so that visual features $\mathbf{s}$ reach a desired value $\mathbf{s}^*$, leading to a correct realization of the task. The control principle is thus to regulate to zero the error vector $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$. To build the control law, the knowledge of the interaction matrix $\mathbf{L_s}$ is usually required. For eye-in-hand systems and a static environment, this matrix links the time variation of $\mathbf{s}$ to the camera instantaneous velocity $\mathbf{v}$

$$\dot{\mathbf{s}} = \mathbf{L_s}\,\mathbf{v} \tag{1}$$

with $\mathbf{v} = (\boldsymbol{v}, \boldsymbol{\omega})$ where $\boldsymbol{v}$ is the linear camera velocity and $\boldsymbol{\omega}$ its angular velocity. Considering direct visual servoing scheme, the difficulty is to find a visual feature $\mathbf{s}$ that is computed on the whole image and for which the interaction matrix can be analytically computed.

The contribution of this paper are the definition of a new visual feature $\mathbf{s}(\mathbf{r})$ (where $\mathbf{r}$ is the camera pose): the norm of the image gradient, the interaction matrix that links the variation of norm of the gradient to camera motion, and an experimental analysis of the behavior of the system using various way to integrate gradient information in the control laws.

The next section, Section II, recalls the basic way to achieve a direct photometric visual servoing process. The definition of visual feature as introduced in [9] is considered. Section III presents the new visual features proposed in this paper based on the norm of the image gradient. Control laws are presented in Section IV along with various choices for the visual feature that consider luminance, gradient or various combinations of these two features. Section V shows the behavior of these various control laws with results obtained on a six degrees of freedom robot.

II. Luminance as a visual feature

The visual features considered in this paper are the luminance I of each point of the image. In fact we have

$$\mathbf{s}(\mathbf{r}) = \mathbf{I}(\mathbf{r}) = (\mathbf{I}_{1\bullet}, \mathbf{I}_{2\bullet}, \cdots, \mathbf{I}_{N\bullet}) \tag{2}$$

where $\mathbf{I}_{k\bullet}$ is nothing but the k-th line of the image. $\mathbf{I}(\mathbf{r})$ is then a vector of size $N \times M$ where $N \times M$ is the size of the image. As mentioned in the previous Section, any control law requires an estimation of the interaction matrix. In our case, as already stated, we are looking for the interaction matrix related to the luminance of a pixel in the image. It is fully described in [9] and is recalled now.

The basic hypothesis assumes the temporal constancy of the brightness for a physical point between two successive images. This hypothesis leads to the so-called *optical flow constraint equation* (OFCE) that links the temporal variation

of the luminance I to the image motion at a given pixel $\mathbf{x}$ [11].

More precisely, assuming that the point has a displacement $\mathbf{dx}$ in the time interval dt, the previous hypothesis leads to

$$I(\mathbf{x}+\mathbf{dx}, t+dt) = I(\mathbf{x}, t). \tag{3}$$

Written under a differential form, a first order Taylor series expansion of this equation about $\mathbf{x}$ gives

$$\nabla I^\top \dot{\mathbf{x}} + \dot{I} = 0 \tag{4}$$

with $\dot{I} = \partial I/\partial t$. It becomes then straightforward to compute the interaction matrix $\mathbf{L}_I$ related to I by plugging the interaction matrix $\mathbf{L}_\mathbf{x}$ related to $\mathbf{x}$ into (4). We obtain

$$\dot{I} = -\nabla I^\top \mathbf{L}_\mathbf{x} \mathbf{v}. \tag{5}$$

Finally, if we introduce the interaction matrices $\mathbf{L}_x$ and $\mathbf{L}_y$ related to the coordinates x and y of $\mathbf{x}$, such as:

$$\dot{x} = \begin{pmatrix} -1/Z & 0 & x/Z & xy & -(1+x^2) & y \end{pmatrix} \mathbf{v} \tag{6}$$

that we can rewrite $\dot{x} = \mathbf{L}_x \mathbf{v}$ and

$$\dot{y} = \begin{pmatrix} 0 & -1/Z & y/Z & 1+y^2 & -xy & -x \end{pmatrix} \mathbf{v} \tag{7}$$

that we rewrite $\dot{y} = \mathbf{L}_y \mathbf{v}$ and we obtain

$$\mathbf{L}_I = -\left(\nabla I_x \mathbf{L}_x + \nabla I_y \mathbf{L}_y\right) \tag{8}$$

where ∇I_x and ∇I_y are the components along x and y of the spatial image gradient ∇I. Note that it is actually the only image processing step necessary to implement the presented method.

III. Norm of the gradient as a visual feature

Although the luminance is a very rich information, it can be considered as unstable when large and non-affine lighting variation occurred. Therefore, rather than considering the luminance we propose here to use the norm of the image gradient as a feature. Like the luminance we have a global feature. The resulting visual servoing scheme will still then not require any matching nor tracking step. Here again, the main difficulty, is to derive the interaction matrix related to this new visual feature.

The square of the norm of the gradient in a point $\mathbf{x} = (x, y)$ is given by:

$$\| \nabla \mathbf{I}(x,y) \|^2 = \nabla I_x^2(x,y) + \nabla I_y^2(x,y). \tag{9}$$

This measure is classically used to determine regions of rapid intensity changes. This usually corresponds to edge in the image.

Considering the square of the norm of the image gradient (see Fig. 1) as a visual feature, vector $\mathbf{s}(\mathbf{r})$ is given by:

$$\mathbf{s}(\mathbf{r}) = \| \nabla \mathbf{I}(\mathbf{r}) \|^2 = \nabla I_x^2(\mathbf{r}) + \nabla I_y^2(\mathbf{r}) \tag{10}$$

where $\| \nabla \mathbf{I}(\mathbf{r}) \|^2$ is the square of the norm of the image gradient.

As in Section II we therefore need to compute the interaction matrix related to the norm of the gradient. This matrix

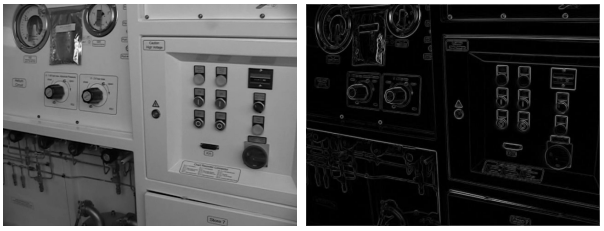

Fig. 1. An image (a) $\mathbf{I}(\mathbf{r})$ and (b) the norm of its gradient $\| \nabla \mathbf{I}(\mathbf{r}) \|^2$

$\mathbf{L}_\mathbf{G}$ relies the variation of the norm of the image gradient to the camera velocity. We have

$$\mathbf{L}_\mathbf{G} = \frac{\partial \mathbf{s}}{\partial x}\mathbf{L}_x + \frac{\partial \mathbf{s}}{\partial y}\mathbf{L}_y \tag{11}$$

with

$$\frac{\partial \mathbf{s}}{\partial x} = 2\left(\frac{\partial^2 I}{\partial x^2}\frac{\partial I}{\partial x} + \frac{\partial^2 I}{\partial x \partial y}\frac{\partial I}{\partial y}\right) \tag{12}$$

and

$$\frac{\partial \mathbf{s}}{\partial y} = 2\left(\frac{\partial^2 I}{\partial x \partial y}\frac{\partial I}{\partial x} + \frac{\partial^2 I}{\partial y^2}\frac{\partial I}{\partial y}\right). \tag{13}$$

After some rewriting, we finally get:

$$\begin{aligned} \mathbf{L}_\mathbf{G} = 2\Bigg[&\left(\nabla I_x \frac{\partial \nabla I_x}{\partial x} + \nabla I_y \frac{\partial \nabla I_y}{\partial x}\right)\mathbf{L}_x \\ &+ \left(\nabla I_x \frac{\partial \nabla I_x}{\partial y} + \nabla I_y \frac{\partial \nabla I_y}{\partial y}\right)\mathbf{L}_y\Bigg]. \end{aligned} \tag{14}$$

IV. Control law and choice of visual feature

A. Control law as an optimization problem

To derive the control, we choose to consider visual servoing as an optimization problem [12]. The goal is to minimize the following cost function

$$\mathcal{C}(\mathbf{r}) = \left(\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*)\right)^\top \left(\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*)\right) \tag{15}$$

where $\mathbf{r}$ describes the current pose of the camera with respect to the object (it is an element of $\mathbb{R}^3 \times SO(3)$) and where $\mathbf{r}^*$ is the desired pose. In that case, a step of the minimization scheme can be written as follows

$$\mathbf{r}_{k+1} = \mathbf{r}_k \oplus \mathbf{vd}\left(\mathbf{r_k}\right) \tag{16}$$

where "$\oplus$" is the operator that updates the pose $\mathbf{r}_k$ and which is "implemented" through the robot controller and $\mathbf{d}\left(\mathbf{r_k}\right)$ a direction of descent. In that case, the following velocity control law can be derived considering that λ_k is small enough

$$\mathbf{v} = \lambda_k \mathbf{d}\left(\mathbf{r_k}\right) \tag{17}$$

λ_k is often chosen as a constant value. In the remainder of the paper we will omit the subscript k for the sake of clarity.

When $\mathbf{r_k}$ lies in a neighborhood of $\mathbf{r}^*$, $\mathbf{s}(\mathbf{r})$ can be linearized around $\mathbf{s}(\mathbf{r_k})$ and plugged into (15). Then, after having zeroed its gradient, we obtain

$$\mathbf{d}\left(\mathbf{r}\right) = -\left(\mathbf{L_s}^\top \mathbf{L_s}\right)^{-1} \nabla \mathcal{C}\left(\mathbf{r}\right) \tag{18}$$

that becomes:

$$\mathbf{v} = -\lambda \mathbf{L}_{\mathbf{s}}^{+} \left(\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*)\right). \quad (19)$$

This is the control law usually used in visual servoing [1].

In our case we consider a Levenberg-Maquardt like approach. This method considers the following direction

$$\mathbf{d}(\mathbf{r}) = -\left(\mathbf{G} + \mu \, \mathrm{diag}(\mathbf{G})\right)^{-1} \nabla \mathcal{C}(\mathbf{r}) \quad (20)$$

where $\mathbf{G}$ is usually chosen as $\nabla^2 \mathcal{C}(\mathbf{r})$ or more simply as $\mathbf{L}_{\mathbf{s}}^{\top} \mathbf{L}_{\mathbf{s}}$ leading in that last case to

$$\mathbf{v} = -\lambda \left(\mathbf{H} + \mu \, \mathrm{diag}(\mathbf{H})\right)^{-1} \mathbf{L}_{\mathbf{s}}^{\top} \left(\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*)\right) \quad (21)$$

with $\mathbf{H} = \mathbf{L}_{\mathbf{s}}^{\top} \mathbf{L}_{\mathbf{s}}$. The parameter μ makes possible to switch from a steepest descent like approach to a Gauss-Newton one [9]. For all the experiments reported in the next section, the following value has been used: $\mu = 0.01$.

B. Choice of the visual features

In this paper, for comparison issue, we will consider various visual features:

1) Luminance: Considering direct photometric visual servoing the basic choice is to consider the luminance as the visual feature. In that case, and denoting $\mathbf{I}(\mathbf{r})$ the image (represented here by a vector), we have $\mathbf{s} = \mathbf{I}(\mathbf{r})$. This corresponds to the visual feature presented in Section II and used in [9]. In this case the control law is given by:

$$\mathbf{v} = -\lambda \left(\mathbf{H}_{\mathbf{I}} + \mu \, \mathrm{diag}(\mathbf{H}_{\mathbf{I}})\right)^{-1} \mathbf{L}_{\mathbf{I}}^{\top} \left(\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{r}^*)\right) \quad (22)$$

with $\mathbf{H}_{\mathbf{I}} = \mathbf{L}_{\mathbf{I}}^{\top} \mathbf{L}_{\mathbf{I}}$

2) Weighted luminance: The basic way to introduce the square of the norm of the gradient is to weight, the luminance with a scalar $\alpha, \alpha \in [0, 1]$ which will be high when the norm of the gradient is large (ie, where there is useful information) and which decreases when gradient magnitude decreases (ie, in area with no significant information). Let us denote $\| \nabla I_{max} \|^2$ the maximum of the norm of the gradient over all the pixels of the image. The coefficient $\alpha(x, y)$ for pixel (x, y) is then given by

$$\alpha(x, y) = \frac{\| \nabla I(x, y) \|^2}{\| \nabla I_{max} \|^2}$$

In that case, the visual feature is given by $\mathbf{s} = \mathbf{D}\mathbf{I}(\mathbf{r})$ where $\mathbf{D}$ is a diagonal matrix that contains the weights α associated to each pixel. Formally, when computing the interaction matrix related to $\mathbf{s}$ we should consider that $\mathbf{D}$ is modified when the camera is moving. Nevertheless we will consider here that it is a constant. The control law here may be consider as an iterative reweighted least square (IRLS) as presented in [13].

3) Norm of the image gradient: The other solution to build a control law based on gradient information is to directly consider a vector of visual feature build using the square of the norm of the gradient. In that case we have $\mathbf{s} = \| \nabla \mathbf{I}(\mathbf{r}) \|^2$ which corresponds to the visual feature presented in Section III.

The control law is then given by:

$$\mathbf{v} = -\lambda \left(\mathbf{H}_{\mathbf{G}} + \mu \, \mathrm{diag}(\mathbf{H}_{\mathbf{G}})\right)^{-1} \mathbf{L}_{\mathbf{G}}^{\top} \mathbf{e} \quad (23)$$

with

$$\mathbf{e} = \| \nabla \mathbf{I}(\mathbf{r}) \|^2 - \| \nabla \mathbf{I}(\mathbf{r}^*) \|^2$$

and with $\mathbf{H}_{\mathbf{G}} = \mathbf{L}_{\mathbf{G}}^{\top} \mathbf{L}_{\mathbf{G}}$ and $\mathbf{L}_{\mathbf{G}}$ given by equation (14).

4) Luminance and norm of the gradient: Finally, one can consider both luminance and norm of the gradient within the same visual feature vector. In that case $\mathbf{s} = (\mathbf{I}(\mathbf{r}), \| \nabla \mathbf{I}(\mathbf{r}) \|^2)^{\top}$ that stack the two former visual features. The goal is here to take advantage of both luminance and gradient information.

V. Experimental results

In all the experiments reported here, the camera is mounted on a 6 degrees of freedom gantry robot. Control law is computed on a Core 2 Duo 3Gz PC. The image processing time (Gaussian filter and the 6 gradient images) along with the interaction matrix computation required 60ms (for 320×240 images). Let us emphasize the fact that for all these experiments, the six degrees of freedom of our robot are controlled.

A. Comparison of the various visual features

For the first experiment a planar object has been used (see Fig. 2). The initial error pose was $\mathbf{\Delta r_{init}} = (3$ cm, -18cm, 8cm, 17°, 2°, -9°$)$. The desired pose was so that the object and CCD planes are parallel. The interaction matrix has been computed at each iteration but assuming that all the depths are constant and equal to $Z^* = 80$ cm, which is a coarse approximation. We conduct a full set of experiment with all the visual feature presented in Section IV-B. In each case we report the evolution of the value $\mathcal{C}(\mathbf{r})$ of the cost function (see equation (15)), the camera velocity $\mathbf{v} = (\boldsymbol{v}, \boldsymbol{\omega})$ in meter/s and radian/s and the positioning error (between $\mathbf{r}$ and $\mathbf{r}^*$) in meter and radian.

We assume in this section that the temporal luminance constancy hypothesis is valid, i.e. we use the interaction matrix given in (8). In order to make this assumption as valid as possible, a diffuse lighting as been used. Moreover, the lighting is also motionless wrt the scene being observed.

In the first experiment we consider luminance as the visual feature $\mathbf{s}(\mathbf{r}) = \mathbf{I}(\mathbf{r})$. Fig. 2a and 2b show the initial and desired image. Fig. 2c and 2d show the error images at the beginning and at the end of the positioning task. These later images can be seen as the error vector $\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{r}^*)$ used in the control law. In that case, as expected, the robot converges toward the correct pose (see Fig. 3) with a huge precision (see Table I third row).

In a second experiment, we consider the square of the norm of the gradient as visual feature $\mathbf{s}(\mathbf{r}) = \| \nabla \mathbf{I}(\mathbf{r}) \|^2$. Fig. 4a and 4b show the scene viewed from the initial and desired pose. Fig. 4c and 4d shows the norm of the gradient images at the beginning and at the end of the positioning task. This can be seen as the current (resp. desired) visual feature $\mathbf{s}$ (resp. $\mathbf{s}^*$).

Image 4e shows the error in the luminance space (not used in the control law) whereas Image 4f shows the error between the two image gradient. Both images are acquired from the initial pose. This later image can be seen as the

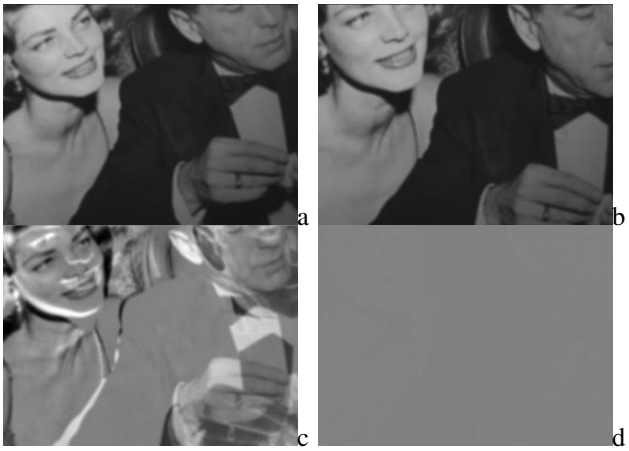

Fig. 2. Luminance as a visual feature [9] (a) Scene viewed from the initial pose (b) scene viewed from the desired pose $\mathbf{r}^*$ (c) initial error $(\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{r}^*))$ (d) final error.

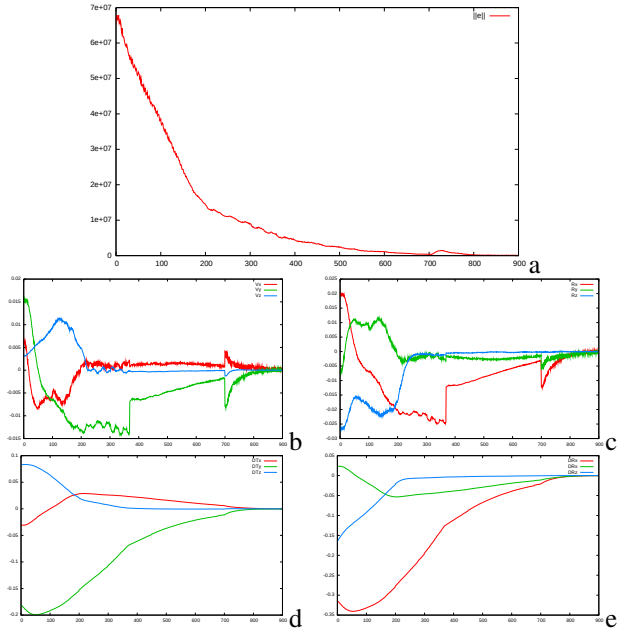

Fig. 3. Luminance as a visual feature [9] (a) Cost function $\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{r}^*)$, (b-c) Camera velocities (m/s or rad/s), (d) Translational part of $\boldsymbol{\Delta}\mathbf{r}$ (m), (e) Rotational part of $\boldsymbol{\Delta}\mathbf{r}$ (rad.)

error vector $\| \nabla\mathbf{I}(\mathbf{r}) \|^2 - \| \nabla\mathbf{I}(\mathbf{r}^*) \|^2$ used in the control law. The last row show similar error images but for the final pose reached by the robot. Let us note that image gradient and image gradient difference are normalized for visualization issue which artificially emphasizes noise when gradient difference is very small (as can be seen on Fig. 4i).

When considering the norm of the gradient, although the robot behavior (3D trajectory and velocity profile) are different (see Fig. 3 vs Fig. 5), the camera reaches its desired pose with a similar accuracy (see Table I, 4th row). In both cases, the precision accuracy is less than 0.2mm in translation and less than 0.02 degree in rotation. Let us note that is very difficult to reach so low positioning errors when using geometric visual features as it is usually done.

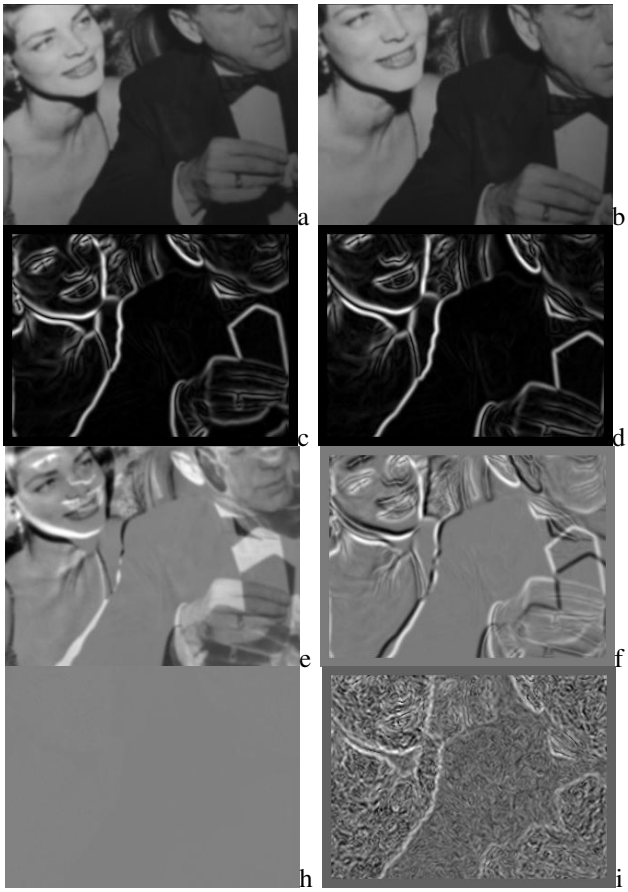

Fig. 4. Square of the norm of the gradient as a visual feature (a) Scene viewed from the initial pose (b) scene viewed from the desired pose $\mathbf{r}^*$ (c) initial visual feature $\mathbf{s}(\mathbf{r}) = \| \nabla\mathbf{I}(\mathbf{r}) \|^2$ (d) desired visual feature $\mathbf{s}^* = \nabla^2\mathbf{I}(\mathbf{r}^*)$ (e) initial error in the intensity space $(\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{r}^*)$ and (f) error on the square of the norm of the gradient used in the control law $(\| \nabla\mathbf{I}(\mathbf{r}) \|^2 - \| \nabla\mathbf{I}(\mathbf{r}^*) \|^2)$ (g) final error in the intensity space and (h) final error on the square of the norm of the gradient used in the control law (error is normalized)

approach	Δt_x	Δt_y	Δt_z	ΔR_x	ΔR_y	ΔR_z
initial error	-30.58	-180.22	82.98	-17.92	1.32	-9.5
Luminance	0	-0.16	-0.02	-0.01	-0.01	0
Gradient	0.2	-0.11	-0.04	-0.01	0	0.02
Lum. & gradient	0.32	0.23	-0.04	0.05	-0.04	0.02
Weighted lum.	0.64	-0.46	-0.04	-0.05	-0.02	-0.08
3D initial error	167.88	0.26	2.19	0.02	-18	0.28
3D final error	0.24	-0.08	0.1	-0.01	0.02	0.1

TABLE I

POSITIONING PRECISION: WE COMPARED THE DIFFERENT APPROACH PRESENTED IN THE PAPER. FINAL ERROR IN mm AND DEGREES.

B. Hybrid visual feature: luminance and gradient

We also experiment a weighted combination of luminance and gradient within the same control law as defined in Section IV-B.2. Although efficient, the camera converges toward the correct pose, it has to be noted that the positioning error is higher than with the two previous experiments (see Fig. 6 and Table I). This is mainly due to the fact that less information are considered (that is the area with low gradients).

In the next experiment, we followed the standard practice

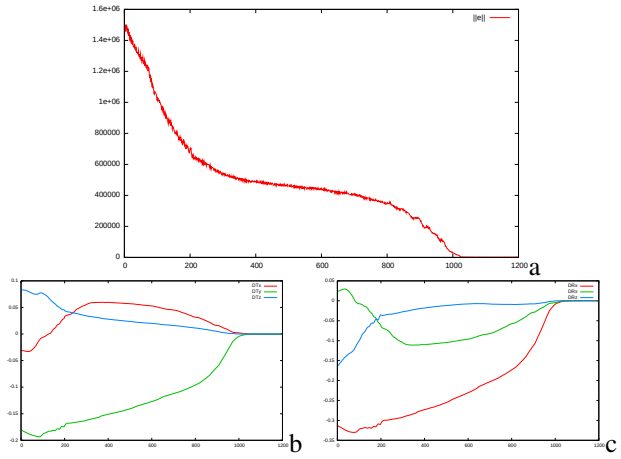

Fig. 5. Square of the norm of the gradient as visual feature (a) Cost function $\| \nabla \mathbf{I}(\mathbf{r}) \|^2 - \| \nabla \mathbf{I}(\mathbf{r}^*) \|^2$, (b) Translational part of $\boldsymbol{\Delta}\mathbf{r}$ (m), (c) Rotational part of $\boldsymbol{\Delta}\mathbf{r}$ (rad.)

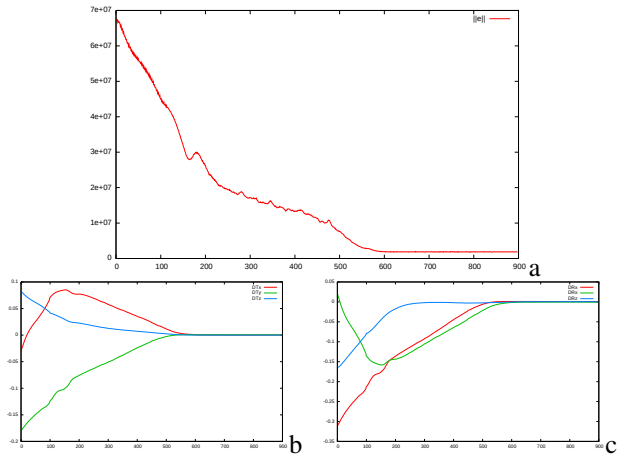

Fig. 6. Using gradient based weighted luminance $\mathbf{s} = \mathbf{D}\ \mathbf{I}(\mathbf{r})$ (a) Cost function, (b) Translational part of $\boldsymbol{\Delta}\mathbf{r}$ (m), (c) Rotational part of $\boldsymbol{\Delta}\mathbf{r}$ (rad.)

in visual servoing by stacking luminance and the norm of the gradient for each image pixel, as defined in Section IV-B.4, in order to create the visual features and its interaction matrix. As expected, in that case we have a good positioning accuracy with a higher convergence rate (see Fig. 7 and Table I).

C. Servo on a 3D object

The goal of the last experiment is to show the robustness of the control law presented in Section IV-B.3 wrt the depths. For this purpose, a non planar scene has been used as shown on Fig. 8. It shows that large errors in the depth are introduced. Let us recall that since depths are not known (and can hardly be recovered on-line) assuming a constant depth in the interaction matrix introduce a modeling error in the control law. Nevertheless, considering luminance as a visual feature, it has already been demonstrated that it is robust to depth variations. Here we consider norm of the gradient as the visual feature. Although the minimization of the cost function is more complex, the positioning error always

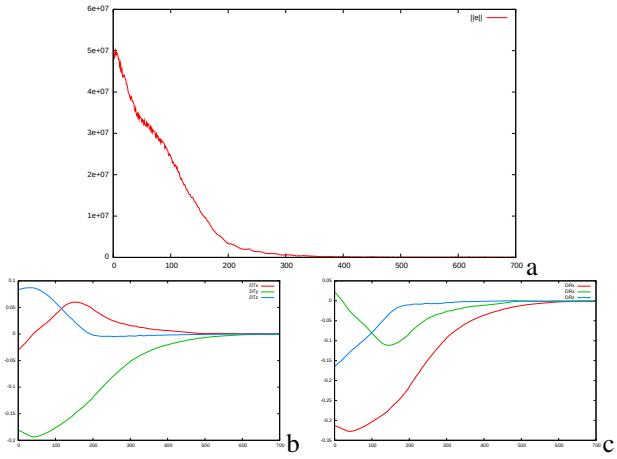

Fig. 7. Stacking luminance and norm of the gradient $\mathbf{s} = (\mathbf{I}(\mathbf{r}), \| \nabla \mathbf{I}(\mathbf{r}) \|^2)^\top$ (a) Cost function, (b) Translational part of $\boldsymbol{\Delta}\mathbf{r}$ (m), (c) Rotational part of $\boldsymbol{\Delta}\mathbf{r}$ (rad.)

decreases leading to positioning accuracy of $\Delta\mathbf{r}$ =(0.24mm, -0.07mm, 0.09mm, -0.014deg, 0.022deg, 0.05deg).

D. Sensibility to light variation

Finally, one of the advantage of the new feature presented in Section IV-B.3 is to be more robust to light variation than luminance. One can see on Fig. 11 that despite an important residual on the cost function due to the modification of lighting conditions during the positioning task (wrt to, eg, Fig. 5a) the final error of the positioning task is quite negligible (less than 0.4mm in translation and 0.12 degrees in rotation to be compared to an error of 1mm and 0.2 degrees when considering luminance feature).

VI. Conclusion

This study was part on research that seek to define direct visual servoing strategy. The goal is to proposed a system that does not require any matching or tracking process which have proved to be one of the bottleneck for the development of visual servoing based systems. Pure photometric visual servoing have first been introduced in [6], [7], [9]. Considering that most of the information in an image are located in its high frequency areas (that are contours), we have consider a new visual feature based on gradient information. We have shown in this paper that it is possible to use the square norm of the gradient of all the pixels in an image as visual features in visual servoing.

References

[1] F. Chaumette, S. Hutchinson, "Visual servo control, Part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, Dec. 2006.

[2] V. Sundareswaran, P. Bouthemy, F. Chaumette, "Exploiting image motion for active vision in a visual servoing framework," *Int. J. of Robotics Research*, vol. 15, no. 6, pp. 629–645, Jun. 1996.

[3] J. Santos-Victor, G. Sandini, "Visual behaviors for docking," *Computer Vision and Image Understanding*, vol. 67, no. 3, Sep. 1997.

[4] A. Crétual, F. Chaumette, "Visual servoing based on image motion," *Int. J. of Robotics Research*, vol. 20, no. 11, pp. 857–877, Nov. 2001.

[5] S. Benhimane, E. Malis, "Homography-based 2d visual tracking and servoing," *Int. J. of Robotics Research*, vol. 26, no. 7, Jul. 2007.

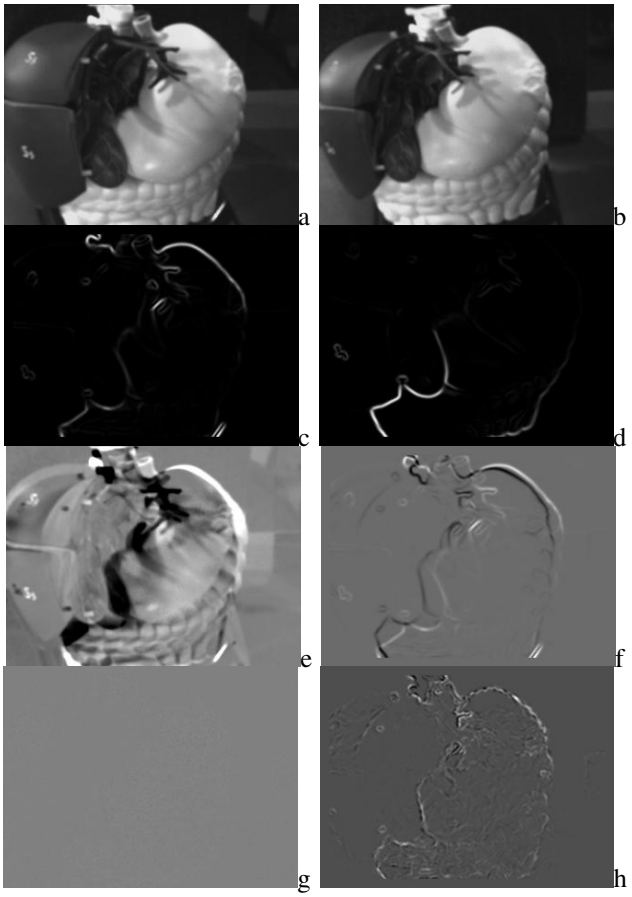

Fig. 8. 3D scene experiment(a) Scene viewed from the initial pose (b) scene viewed from the desired pose $\mathbf{r}^*$ (c) initial visual feature $\mathbf{s}(\mathbf{r}) = \nabla^2\mathbf{I}(\mathbf{r})$ (d) desired visual feature $\mathbf{s}^* = \| \nabla\mathbf{I}(\mathbf{r}) \|^2$ (e) initial error in the intensity space $(\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{r}^*)$ and (f) initial error on the norm of the gradient used in the control law $(\| \nabla\mathbf{I}(\mathbf{r}) \|^2 - \| \nabla\mathbf{I}(\mathbf{r}^*) \|^2)$ (g) final error in the intensity space and (h) final error on the norm of the gradient used in the control law.

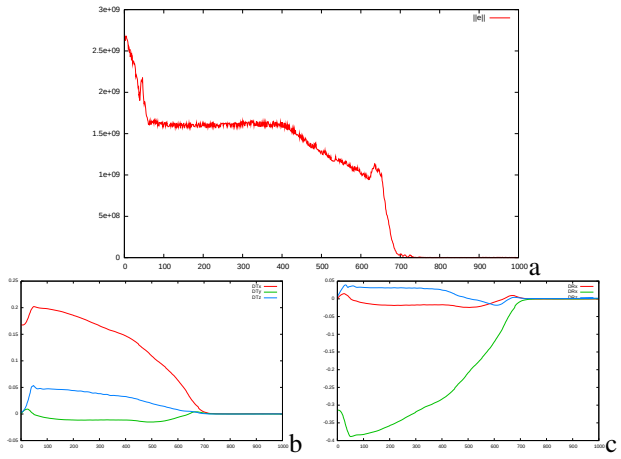

Fig. 9. 3D scene experiment (a) Cost function $\| \nabla\mathbf{I}(\mathbf{r}) \|^2 - \| \nabla\mathbf{I}(\mathbf{r}^*) \|^2$, (b) Translational part of $\mathbf{\Delta r}$ (m), (c) Rotational part of $\mathbf{\Delta r}$ (rad.)

[6] S. Nayar, S. Nene, H. Murase, "Subspace methods for robot vision," *IEEE Trans. on Robotics*, vol. 12, no. 5, pp. 750 – 758, Oct. 1996.

[7] K. Deguchi, "A direct interpretation of dynamic images with camera and object motions for vision guided robot control," *Int. J. of Computer Vision*, vol. 37, no. 1, pp. 7–20, Jun. 2000.

Fig. 10. Influence of luminance variation (a) desired image (b) final image $\mathbf{r}^*$ (c) initial error in the intensity space $(\mathbf{I}(\mathbf{r}) - \mathbf{I}(\mathbf{s}(\mathbf{r}))$ and (e) initial error on the norm of the gradient used in the control law $(\| \nabla\mathbf{I}(\mathbf{r}) \|^2 - \| \nabla\mathbf{I}(\mathbf{r}^*) \|^2)$ (e) final error in the intensity space and (f) final error on the norm of the gradient used in the control law.

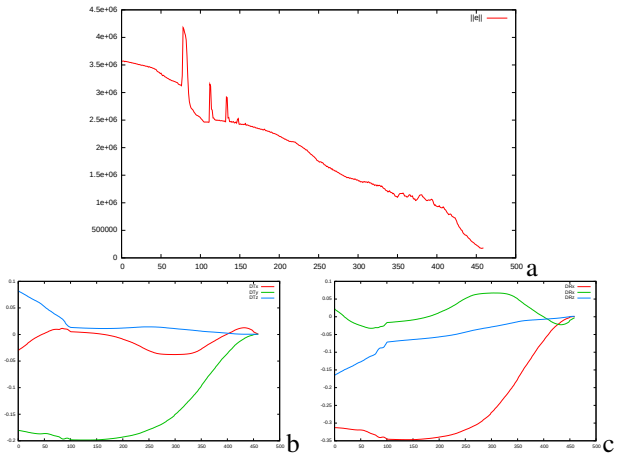

Fig. 11. 3D scene experiment (a) Cost function $\nabla^2\mathbf{I}(\mathbf{r}) - \nabla^2\mathbf{I}(\mathbf{r}^*)$, (b) Translational part of $\mathbf{\Delta r}$ (m), (c) Rotational part of $\mathbf{\Delta r}$ (rad.)

[8] V. Kallem, M. Dewan, J. Swensen, G. Hager, N. Cowan, "Kernel-based visual servoing," in *IEEE/RSJ IROS'07*, San Diego, Oct. 2007.

[9] C. Collewet, E. Marchand, F. Chaumette, "Visual servoing set free from image processing," in *IEEE Int. Conf. on Robotics and Automation, ICRA'08*, Pasadena, CA, May 2008, pp. 81–86.

[10] A. Dame, E. Marchand, "Entropy-based visual servoing," in *IEEE ICRA'09*, Kobe, Japan, May 2009, pp. 707–713.

[11] B. Horn, B. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1-3, pp. 185–203, Aug. 1981.

[12] E. Malis, "Improving vision-based control using efficient second-order minimization techniques," in *IEEE ICRA'04*, New Orleans, Apr. 2004.

[13] A. Comport, E. Marchand, F. Chaumette, "Statistically robust 2d visual servoing," *IEEE T. on Robotics*, vol. 22, no. 2, pp. 415–421, apr 2006.