



Colorimetry-based visual servoing

Christophe Collewet, Eric Marchand

► **To cite this version:**

Christophe Collewet, Eric Marchand. Colorimetry-based visual servoing. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Oct 2009, St Louis, United States. pp.5438-5443, 2009. <inria-00629746>

HAL Id: inria-00629746

<https://hal.inria.fr/inria-00629746>

Submitted on 6 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Colorimetry-based visual servoing

Christophe Collewet and Eric Marchand

Abstract—The goal of this paper is to present a way to perform visual servoing tasks from color attributes. This approach can be seen as an extension of our previous papers based on the luminance [1], [2]. Indeed, as we did for the luminance, color attributes are directly used in the control law avoiding therefore any complex images processing as features extraction or matching. We propose in this paper several potential color features and then a way to select a priori the best choice among them with respect to the scene being observed. Experimental results validate as well the interest of using color attributes as visual features as our selection process.

I. INTRODUCTION

Visual servoing is a widely used technique in robot control [3]. It is based on visual features extracted from a vision sensor. More precisely, the control law is designed so that the visual features s extracted from the current image at the current pose \mathbf{r} , reach a desired value s^* acquired at the desired pose \mathbf{r}^* leading to a correct realization of the task. The control principle relies on the regulation to zero of the following error vector $\mathbf{e} = s - s^*$. To design this control law, the knowledge of the interaction matrix \mathbf{L}_s is usually required. For eye-in-hand systems, it links the time variation of s to the camera instantaneous velocity \mathbf{v} :

$$\dot{s} = \mathbf{L}_s \mathbf{v} \quad (1)$$

with $\mathbf{v} = (v, \omega)$ where v is the linear velocity and ω is the angular velocity.

In this approach, the choice of the visual features as well as its related interaction matrix play an essential role [3]. Most often, geometric visual features as points, straight lines, pose or homography are chosen to control the robot. However, we recently introduced non geometric visual features [1], [2]. Indeed, this new approach is very interesting since it strongly limits the image processing. More precisely, it avoids the robust extraction, matching (between $\mathbf{x}(\mathbf{r}_k)$ and $\mathbf{x}(\mathbf{r}^*)$, where \mathbf{x} denotes the visual measurements required to design and compute s) and real-time spatio-temporal tracking (between $\mathbf{x}(\mathbf{r}_{k-1})$ and $\mathbf{x}(\mathbf{r}_k)$) which are well known to be difficult tasks. Supplementary advantages are that this approach is not sensitive to partial occlusions and to coarse approximations of the depths required to compute the interaction matrix. In these papers, instead of using geometric visual features we directly used the luminance. In that case, the visual feature vector s is nothing but the image itself while s^* is the desired image. The error \mathbf{e} is then simply the difference between the current and desired images. The main issue of these works

was to exhibit the analytical form of the interaction matrix related to the luminance. This was therefore a very different approach than previous works based on luminance [4]–[7] (see [1] for more details).

In this paper we investigate the use of color images instead of using the luminance as we did before. Indeed, using color attributes are more discriminant than using simply the luminance since two image points with a same luminance value can be differentiated from their color attributes. That is why color is widely used in image processing as for example in image segmentation. In contrast, to our knowledge, color has not been used in visual servoing. Of course, it has been used in robot control to segment the object of interest and then to apply the control law (see for example [8]). However, let us point out that our approach is radically different. Our goal is to use features based on the R , G and B components of the image directly in the control law. As we did for the luminance, we also here want to avoid any complex image processing as features extraction and matching. Another contribution of this work is to propose a criterion able to select good attributes from a set of interesting potential color attributes. As we shall see, this way to proceed allows to greatly improve the behavior of the control law, especially the 3D trajectory of the camera.

This paper is organized as follows. First, we recall in Section II the way we used the luminance in our previous works [1]. Then, we present in Section III potential color attributes and compute their related interaction matrix. In Section IV, a way to select them is discussed. Experimental results on positioning tasks are shown in Section V.

II. LUMINANCE AS A VISUAL FEATURE

To make the paper more readable, before investigated color visual features, we first recall how the luminance has been used in [1].

In this work, we have considered the luminance I of each point of the image. Thus, we have

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

where $\mathbf{I}_{k\bullet}$ is 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.

The main problem with such visual features is, of course, to derive the analytical form of the interaction matrix related to the luminance. It has been performed in [1] for Lambertian scenes while the general case has been performed in [2]. However, for the sake of clarity, we only present here the former case.

The basic hypothesis assumes the temporal constancy of the brightness for a physical point between two successive

images. Therefore, assuming that the image point \mathbf{x} has a displacement $d\mathbf{x}$ in the time interval dt , the previous hypothesis leads to

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

Written with a differential form, a first order Taylor series expansion of this equation around \mathbf{x} gives

$$\nabla I^\top \dot{\mathbf{x}} + \dot{I} = 0. \quad (4)$$

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

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

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

$$\mathbf{L}_I = -(\nabla I_x \mathbf{L}_x + \nabla I_y \mathbf{L}_y) \quad (6)$$

where ∇I_x and ∇I_y are the components along x and y of the image gradient ∇I . Let us point out that it is actually the only image processing step necessary to implement this approach.

III. POTENTIAL COLOR VISUAL FEATURES

Instead of considering the luminance, we consider here color attributes. Similarly to the previous section, the simpler visual features could be one of the three R, G, B components of the color vector $\mathbf{C} = (R, G, B)$. Note that another choices could be any combinations of 2 of these color planes or even all the three components. In these cases, the interaction matrix is simply the one obtained by stacking all the interaction matrices related to the color planes that have been chosen. Note also that the interaction matrix related to one color plane is obtained from (6) when I has to be substituted by R, G , or B . However, more complex visual features can be used as non-linear functions of \mathbf{C} . Interesting non-linear functions are color invariants [9]. Indeed, such functions are invariant to any changes in the scene geometry (that is the relative positions observer / lighting source / surface being observed). They are also invariant to intensity changes of the lighting.

Assuming a white lighting, two classes of color invariants exist: invariants dedicated to Lambertian materials and invariants dedicated to specular materials (see for example [10]). For Lambertian materials, the following invariants are defined [9]:

$$L_i^\varphi = \frac{\varphi}{N_i} \quad (7)$$

where $\varphi \in \{R, G, B\}$ and $i = 1$ or 2 . N_1 is defined as

$$N_1 = R + G + B \quad (8)$$

while N_2 is defined as

$$N_2 = \sqrt{R^2 + G^2 + B^2}. \quad (9)$$

This invariants are known as L_1 or L_2 invariants since they performed a L_1 or L_2 normalization of the (color) image.

For specular materials, the following invariants, known as $l_1 l_2 l_3$, are used [10]:

$$\begin{cases} l_1 &= (R - G)^2 / N_3 \\ l_2 &= (R - B)^2 / N_3 \\ l_3 &= (G - B)^2 / N_3 \end{cases} \quad (10)$$

where $N_3 = (R - G)^2 + (R - B)^2 + (G - B)^2$. Note that this invariant is also normalized. However, this latter invariant may be very noisy or even undefined at any point where the color is not saturated¹. That is for the set of points belonging to the achromatic axis (the grey points). In contrast, this problem occurs with the L_1 invariants only for $R \approx G \approx B \approx 0$. That is for black points only.

In addition, let us point out that the components of all of these invariants (7) or (10) are not independent, given two of them the third one can be computed; therefore two components are sufficient to describe completely the color at a point \mathbf{x} . That is why we never combine all the components of a normalized invariant.

To derive a control law directly based on one of a component of (7) or (10), or on different combinations of two of their components, the related interaction matrix has to be determined. This computation is, however, simple. Considering a component I_{inv} of an invariant given by (7) or (10), the temporal luminance constancy hypothesis (3) holds, leading similarly as (6) to

$$\mathbf{L}_{I_{inv}} = -\nabla I_{inv}^\top \mathbf{L}_\mathbf{x}. \quad (11)$$

Note that the spatial gradient of I_{inv} can be easily computed from (7) or (10) from the spatial gradients of R, G and B . For example, considering (7), we simply have

$$\mathbf{L}_{L_i^\varphi} = -\frac{1}{N_i} (\nabla \varphi - L_i^\varphi \nabla N_i)^\top \mathbf{L}_\mathbf{x}. \quad (12)$$

Recall that to remove the temporal luminance constancy hypothesis the approach described in [2] has to be used leading, of course, to a different and more complicated expression than (11).

A more complicated problem is rather to choose the best color attributes to use from a given color image of the scene being observed. For example, let us consider the simple case where an image only contains the blue and green colors, it is obvious that using the R color plane as the only visual features will lead to a bad behavior of the control law. This problem is the subject of the next section.

IV. SELECTION OF VISUAL FEATURES

The answer to that question must be seen from the visual servoing problem, that is which color attributes will lead to a better behavior of the control law ? that essentially means a convergence from a larger domain and a better 3D camera trajectory. Usually, the condition number (CN) of the interaction matrix is used as a criterion able to compare visual features [11], [12]. Using the L_2 norm, it writes as

$$CN_s = \frac{\max \sigma_i}{\min \sigma_i} \quad (13)$$

¹The color at a given image point is not saturated if $R \approx G \approx B$.

with $i = 1, \dots, 6$ for a 6 ddl robot. The vector σ is obtained from the singular value decomposition of the interaction matrix related to \mathbf{s} . A low value of this criterion leads to a better behavior of the closed-loop system with respect to the noise. It means that for small changes in the pose or in the interaction matrix, changes in the visual features remain small too.

Let us introduce another criterion. To do that, contrary to the seminal paper [13] where the control law was designed so that a desired behavior was required for \mathbf{e} , we propose to see the design of the control law as an optimization problem as proposed in [1], [14], [15]. In this case, the goal is to minimize the following cost function (see Section I).

$$\mathcal{C}(\mathbf{r}) = \frac{1}{2} \|\mathbf{e}(\mathbf{r})\|^2. \quad (14)$$

Of course, the graal would be to find some visual features so that $\mathcal{C}(\mathbf{r})$ would be an hypersphere. Indeed, only a global minimum would exist and a simple steepest descent method would ensure to reach this minimum. Therefore, we are interested in studying the Hessian of (14). It is given by

$$\nabla^2 \mathcal{C}(\mathbf{r}) = \left(\frac{\partial \mathbf{s}}{\partial \mathbf{r}} \right)^\top \left(\frac{\partial \mathbf{s}}{\partial \mathbf{r}} \right) + \sum_{i=1}^{i=\dim \mathbf{s}} \nabla^2 s_i (s_i(\mathbf{r}) - s_i(\mathbf{r}^*)). \quad (15)$$

However, this expression is far too complex to derive some useful results. Thus, we study it around the desired position \mathbf{r}^* , leading to

$$\nabla^2 \mathcal{C}(\mathbf{r}^*) = \left(\frac{\partial \mathbf{s}}{\partial \mathbf{r}} \right)^\top \left(\frac{\partial \mathbf{s}}{\partial \mathbf{r}} \right). \quad (16)$$

Moreover, since we have $\dot{\mathbf{s}} = \frac{\partial \mathbf{s}}{\partial \mathbf{r}} \dot{\mathbf{r}} = \mathbf{L}_s \mathbf{v}$, we are interested in practice in the following matrix

$$\mathbf{H}^* = \mathbf{L}_{s^*}^\top \mathbf{L}_{s^*}. \quad (17)$$

This matrix allows us to estimate the cost function around \mathbf{r}^* . Indeed, a first order Taylor series expansion of the visual features $\mathbf{s}(\mathbf{r})$ around \mathbf{r}^* gives

$$\mathbf{s}(\mathbf{r}) = \mathbf{s}(\mathbf{r}^*) + \mathbf{L}_{s^*} \Delta \mathbf{r} \quad (18)$$

where $\Delta \mathbf{r}$ denotes the relative pose between \mathbf{r} and \mathbf{r}^* . Therefore, by plugging (18) into (14), we obtain an approximation of the cost function in a neighborhood of \mathbf{r}^*

$$\widehat{\mathcal{C}}(\mathbf{r}) = \frac{1}{2} \Delta \mathbf{r}^\top \mathbf{H}^* \Delta \mathbf{r}. \quad (19)$$

Consequently, another criterion to select visual features is to choose them so that the eigenvalues of \mathbf{H}^* are the most similar as possible, therefore we introduce the following criterion

$$c_s = \frac{\max \lambda_i}{\min \lambda_i} \quad (20)$$

where the λ_i 's are the eigenvalues of \mathbf{H}^* . Of course, the optimal choice for \mathbf{s} occurs when $c_s = 1$ since in that case the cost function is an hypersphere. In contrast, if c_s is a high value, then the shape of the cost function presents a narrow valley, which is well known to be a complex

optimization problem. Therefore, dedicated control laws have to be designed as the one we have proposed in [1].

The problem is now to compare the criterion given by (13) and the one given by (20). First, let us write the interaction matrix from its singular value decomposition:

$$\mathbf{L}_s = \mathbf{U} \Sigma \mathbf{V}^\top \quad (21)$$

where Σ is a diagonal matrix which diagonal is σ . By multiplying each side of (21) by $\mathbf{L}_{s^*}^\top$, this relation becomes

$$\mathbf{L}_{s^*}^\top \mathbf{L}_{s^*} = \mathbf{V} \Sigma^2 \mathbf{V}^\top. \quad (22)$$

However, since $\mathbf{L}_{s^*}^\top \mathbf{L}_{s^*}$ is a symmetric matrix, it is diagonalizable as

$$\mathbf{L}_{s^*}^\top \mathbf{L}_{s^*} = \mathbf{W} \Lambda \mathbf{W}^\top. \quad (23)$$

where \mathbf{W} is an orthonormal matrix. Consequently, we have $\Lambda = \Sigma^2$ leading finally from the definitions of CN_s and c_s to

$$CN_s = \sqrt{c_s}. \quad (24)$$

This last relation means that the condition number is not only an interesting criterion related to robustness and sensitivity to noise but also related to the shape of the cost function at the desired position.

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 running Linux. Image are acquired at 66Hz using an IEEE 1394 camera with a resolution of 320×240 . Moreover, we used specular and planar objects (some posters) as shown on Fig. 1. The experiments consist in achieving positioning tasks by visual servoing using the following control law

$$\mathbf{v} = -\lambda (\mathbf{H} + \mu \text{diag}(\mathbf{H}))^{-1} \mathbf{L}_{s^*}^\top (\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*)) \quad (25)$$

with $\mathbf{H} = \mathbf{L}_s^\top \mathbf{L}_s$ and where λ and μ are positive scalars (see [1] where the way to set these parameters is described).

For all the experiments the desired positions was so that objects and CCD planes are parallel. We compared the use of the following different sets of visual features from the values of the criterion (24):

- set 1:** the color planes R , G , B or the luminance;
- set 2:** all the combinations of the color planes R , G or B ;
- set 3:** one component of each invariants;
- set 4:** the combinations of two components of a given invariant.

For the first experiment, the initial pose was $\mathbf{r}_{init} = (14 \text{ cm}, -18 \text{ cm}, 2.5 \text{ cm}, -16^\circ, -10^\circ, -1.3^\circ)$. The initial and desired images are reported respectively on Fig. 2a and Fig. 1a. The worse results have been obtained when using the $l_1 l_2 l_3$ invariant as well for the set of visual features 3 as for the set of visual features 4. In those cases, the condition numbers were very bad, between 195 and 274. Therefore, either the control law diverged or converged very slowly. The control law also diverged when using the luminance (the condition number was 148). Better results were obtained

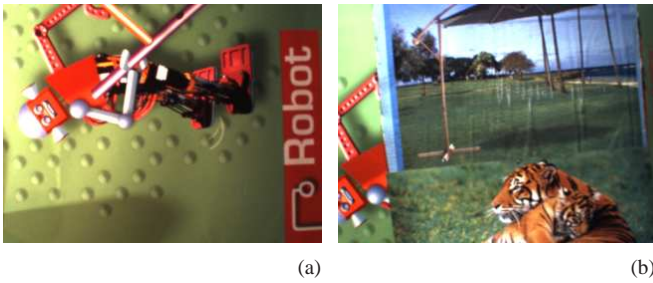


Fig. 1. Desired image for the objects used in the experiments. (a) First and second experiment. (b) Third and fourth experiment.

using the sets of visual features 1 or 2 (the condition numbers were around 140). However, the best results were obtained for the sets of visual features 3 or 4 when using either the L_1 or the L_2 invariants as can be seen on Fig. 3. More precisely, Fig. 3a depicts the behavior of the different normalized cost functions (a cost function evaluated at time k divided by the cost function evaluated at the first frame) while Fig. 3b depicts the trajectories of the camera in a fixed frame. As can be seen using a L_1 or L_2 invariant leads to a better trajectory than when using visual features from the set 1, it also leads to better decreases of cost functions. Fig. 3 also shows that the differences between the L_1 and L_2 invariants are very low (for both the condition numbers were very similar, around 105). These results validate that a bad condition number leads to a bad behavior of the convergence of the control law (a slow convergence or even a divergence) and to a bad 3D trajectory of the camera. This also validates that the condition number is also related to the form of the cost function.

The second experiment still concerns the same object, only the initial pose has been changed. Thus, the desired image is still the one reported on Fig. 1a. The initial image is represented on Fig. 2b. For this experiment we had $\mathbf{r}_{init} = (4.3 \text{ cm}, -18 \text{ cm}, 0 \text{ cm}, -16^\circ, 2.5^\circ, -12.5^\circ)$. As during the first experiment, using visual features from first and second sets of visual features leads to either a divergence, a slow convergence or a bad camera trajectory. For example, using the R or B color plane in the first experiment yields a very slow convergence (not described in Fig. 3) while in the second experiment their use leads to a divergence. It is also the case for the G color plane where as well during the first as during the second experiment a bad behavior have been obtained. Besides, here again using the luminance failed. Moreover, as during the first experiment, the best results have been obtained using either the invariants L_1 or L_2 (those based on $l_1 l_2 l_3$ failed); in addition, since the values of their criterion are similar, a similar behavior with respect to the convergence of the control law or with respect to the camera trajectory is observed.

Fig. 5 reports the L_1 and L_2 invariants at the desired position for the experiments 1 and 2. As can be seen, the differences between a L_1^φ and L_2^φ are small. Moreover, except considering $\varphi = B$, these images are quite similar. Indeed, for $\varphi = B$, the images are dark which means that these invariants are not discriminant at all. That explains why

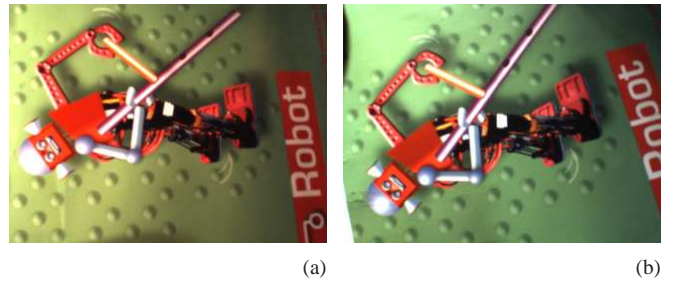


Fig. 2. Initial images. (a) For the first experiment. (b) For the second experiment.

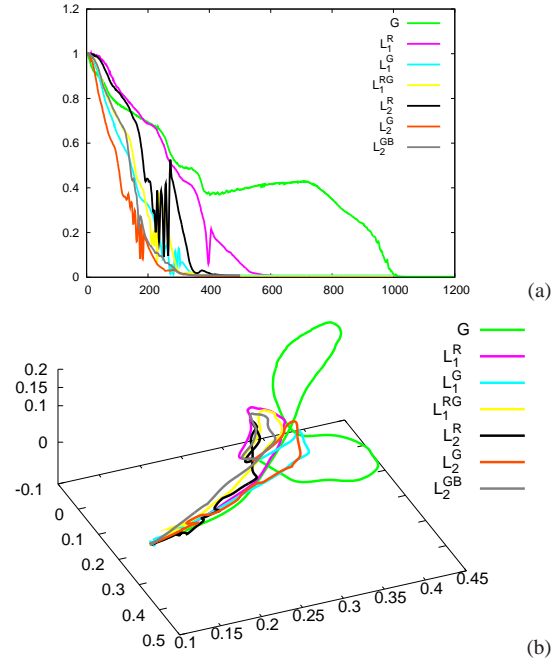


Fig. 3. First experiment : comparison of potential color visual features. (a) Normalized cost functions $\| \mathbf{e}_k \|^2 / \| \mathbf{e}_{k=0} \|^2$ vs frame iteration (denoted by k). (b) Trajectories of the camera in a fixed frame.

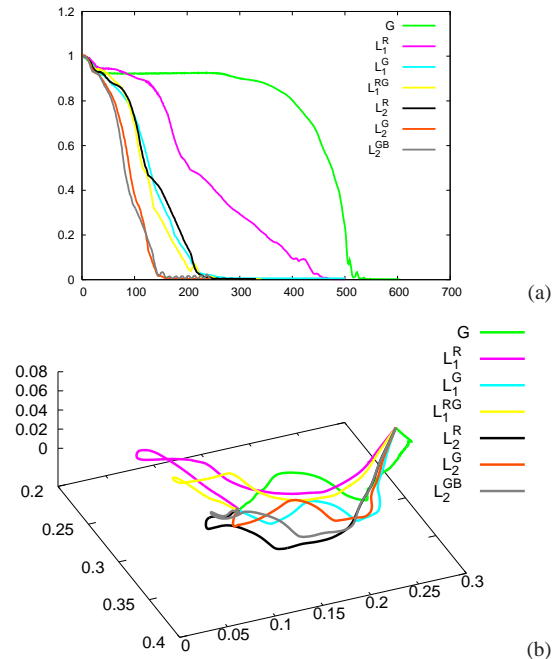


Fig. 4. Second experiment : comparison of potential color visual features.

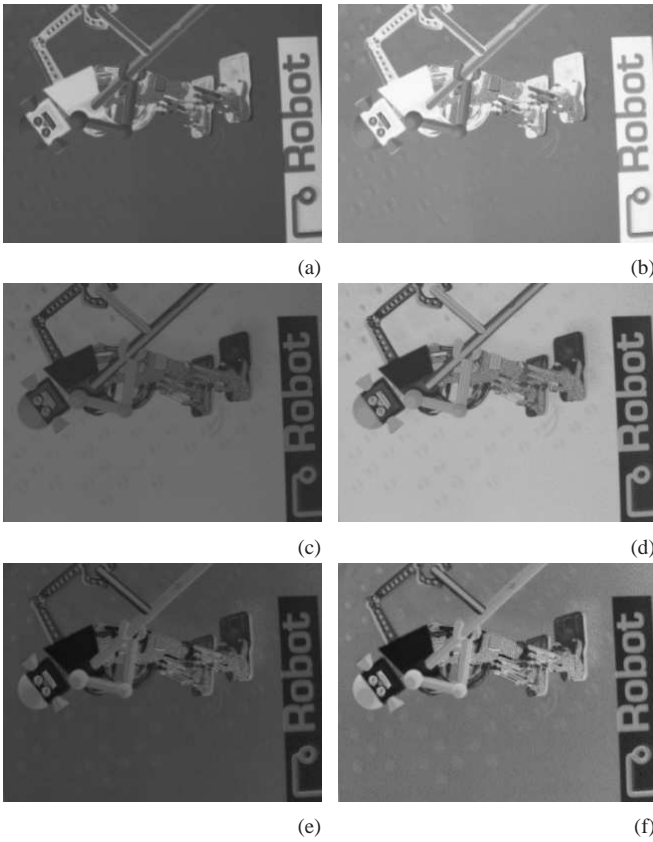


Fig. 5. Images of the L_i^φ invariants for the first object. (a) L_1^R . (b) L_2^R . (c) L_1^G . (d) L_2^G . (e) L_1^B . (f) L_2^B .

we obtained among the L_i^φ invariants the worse value for the criterion (120). For $\varphi = B$, the control law has converged very slowly during the first experiment and has diverged during the second one.

The third experiment concerned a different object. The initial and desired images are reported respectively on Fig. 6a and Fig. 1b. The initial pose was the same as for the first experiment. As for the previous experiments, the $l_1l_2l_3$ invariants have very bad values for their criterion (around 114) and consequently have failed. The sets of visual features 1 and 2, and the luminance, have also some bad values, but slight better (around 109). Thus, as can be seen on Fig. 7, using the R color plane or the luminance have led to a slow convergence and to a bad camera trajectory, but they have converged. As previously, Fig. 8 reports the L_1 and L_2 invariants at the desired position for the experiments 3 and 4. As can be seen, these images are much more different than for the first and second experiments, especially for L_1^R and L_2^R , the L_2^R image is reached than the L_1^R one. It is also confirmed by the value of the criterion, respectively 105 and 96. Therefore, as expected, using L_1^R leads to a divergence while L_2^R leads to a convergence. On the other hand, as can be observed on Fig. 8c and d, these images are quite homogeneous and thus these invariants are not interesting (recall that the interaction matrix requires non null values for the spatial image gradients, see (11)). This is validated

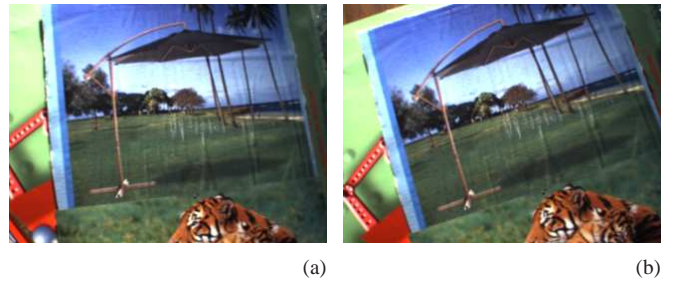


Fig. 6. Initial images. (a) For the third experiment. (b) For the fourth experiment.

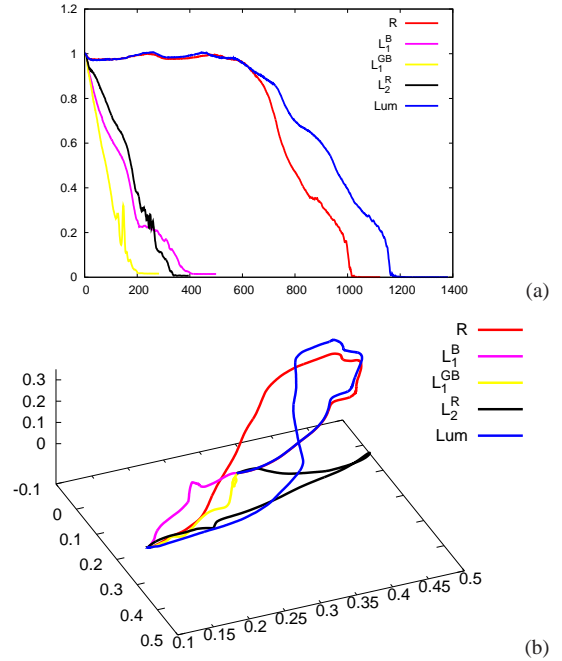


Fig. 7. Third experiment : comparison of potential color visual features.

by bad criterion values (around 107), thus these invariants have led to a divergence. In contrast, using L_i^B has led to a convergence (the criterion value was 103).

The fourth experiment concerned the same object as for the third experiment but the initial pose was the one of the second experiment. This experiment validates the results of the third experiment, the visual features with a bad criterion value have also led to a divergence or to a bad camera trajectory during this experiment (see Fig. 9).

Remark 1: The reader may wonder why we do not present any experiment concerning robustness against changes in the scene geometry as mentioned in section III. In fact, the L_1 and L_2 invariants only perform on true Lambertian materials (such as cloths for example) and are not robust at all (concerning changes in the geometry) if the object is not Lambertian as the objects we used in our experiments. In addition, we have tried to use invariants to specular objects (the $l_1l_2l_3$ ones), but these invariants have a very bad criterion value and, consequently, have always led to a bad behavior of the control law.

Remark 2: Even if none result are given here, as in [1], using

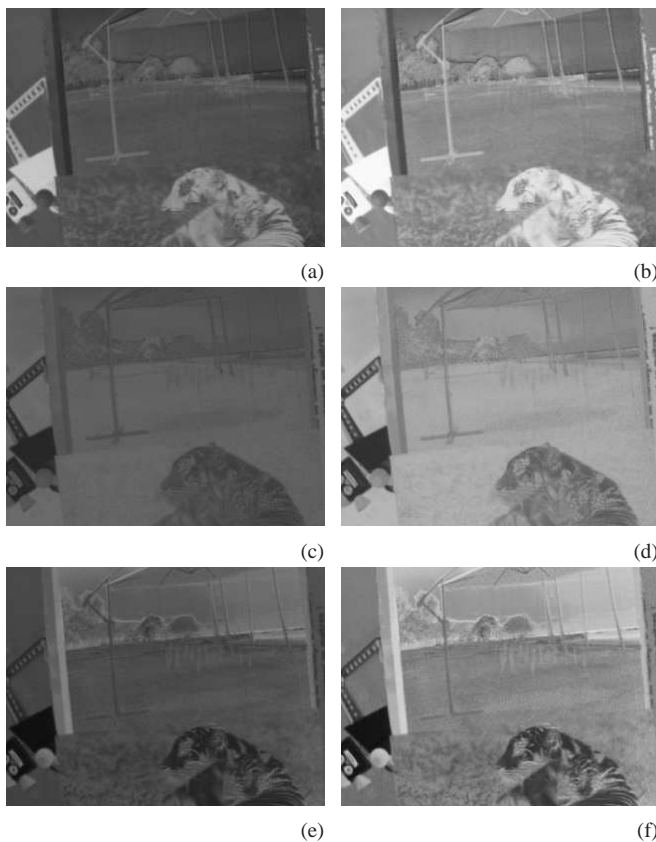


Fig. 8. Images of the L_i^ϕ invariants for the second object. (a) L_1^R . (b) L_2^R . (c) L_1^G . (d) L_2^G . (e) L_1^B . (f) L_2^B .

color attributes with a low criterion value also leads to a very low positioning error (typically 0.1 mm for the translations and 0.01° for the rotations). It is because $s - s^*$ is very sensitive to the pose r . Such visual features also lead to an approach that is not sensitive to partial occlusions (this nice behavior is due to the high redundancy of the visual features we use) and also not sensitive to coarse approximations of the depths required to compute the interaction matrix.

VI. CONCLUSION

It has been shown in this paper that using color attributes greatly improved simply the use of the luminance. As in the luminance case, the complex problems of features extraction and matching are avoided. However, this paper has also shown that a selection of these color attributes is required since their values highly influence the convergence of the control law and the 3D camera trajectory. Especially, the visual features obtained from the L_1 and L_2 invariants have led to the best behaviors.

REFERENCES

- [1] C. Collewet, E. Marchand, and F. Chaumette, "Visual servoing set free from image processing," in *IEEE Int. Conf. on Robotics and Automation, ICRA'08*, Pasadena, California, May 2008, pp. 81–86.
- [2] C. Collewet and E. Marchand, "Photometry-based visual servoing using light reflexion models," in *IEEE Int. Conf. on Robotics and Automation, ICRA'09*, Kobe, Japan, May 2009.

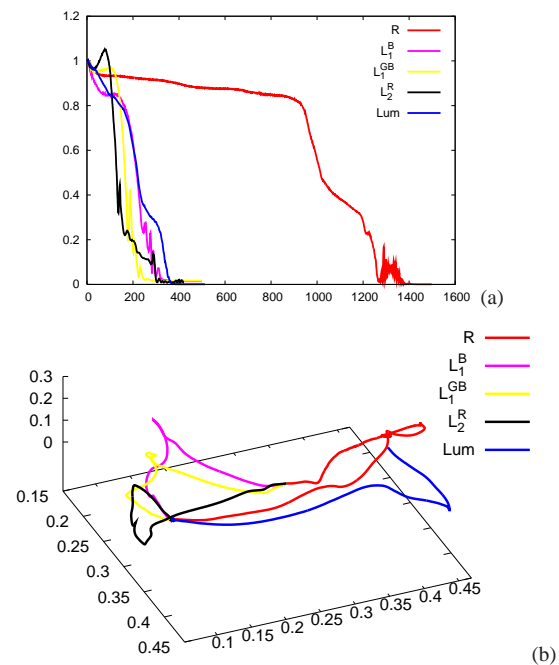


Fig. 9. Fourth experiment : comparison of potential color visual features.

- [3] F. Chaumette and S. Hutchinson, "Visual servo control, Part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, December 2006.
- [4] K. Deguchi, "A direct interpretation of dynamic images with camera and object motions for vision guided robot control," *Int. Journal of Computer Vision*, vol. 37, no. 1, pp. 7–20, June 2000.
- [5] S. Nayar, S. Nene, and H. Murase, "Subspace methods for robot vision," *IEEE Trans. on Robotics*, vol. 12, no. 5, pp. 750–758, October 1996.
- [6] V. Kallem, M. Dewan, J. Swensen, G. Hager, and N. Cowan, "Kernel-based visual servoing," in *IEEE/RSJ Int. Conf. on Intelligent Robots and System, IROS'07*, San Diego, USA, October 2007.
- [7] G. Silveira and E. Malis, "Direct visual servoing with respect to rigid objects," in *IEEE/RSJ Int. Conf. on Intelligent Robots and System, IROS'07*, San Diego, USA, October 2007, pp. 1963–1968.
- [8] G. De Cubber, S. Berrabah, and H. Sahli, "Color-based visual servoing under varying illumination conditions," *Robotics and Autonomous Systems*, vol. 47, no. 4, pp. 225–249, 2004.
- [9] G. Healey, "Segmenting images using normalized color," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 22, no. 1, pp. 64–73, Jan/Feb 1992.
- [10] T. Gevers and A. Smeulders, "Object recognition based on photometric color invariants," in *Scandinavian Conference on Image Analysis, Lappeenranta, Finland, June 1997*.
- [11] J. Feddema, C. Lee, and O. Mitchell, "Automatic selection of image features for visual servoing of a robot manipulator," in *IEEE Int. Conf. on Robotics and Automation, ICRA'89*, vol. 2, Scottsdale, Arizona, May 1989, pp. 832–837.
- [12] F. Chaumette, "Image moments: a general and useful set of features for visual servoing," *IEEE Trans. on Robotics*, vol. 20, no. 4, pp. 713–723, August 2004.
- [13] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. on Robotics and Automation*, vol. 8, no. 3, pp. 313–326, June 1992.
- [14] K. Miura, J. Gangloff, and M. De Mathelin, "Robust and uncalibrated visual servoing without Jacobian using a simplex method," in *IEEE/RSJ Int. Conf. on Intelligent Robots and System, IROS'02*, Lausanne, Switzerland, October 2002.
- [15] E. Malis, "Improving vision-based control using efficient second-order minimization techniques," in *IEEE Int. Conf. on Robotics and Automation, ICRA'04*, vol. 2, New Orleans, April 2004, pp. 1843–1848.