Direct visual servoing using ZNCC criterion

Amaury Dame, Eric Marchand

***Abstract*— This paper proposes a direct visual scheme. In direct visual servoing approaches, the goal is to consider all the image as a whole. Therefore we do no rely on a feature extraction or tracking process. In nominal condition, these approaches have proved to be robust and very precise. In this paper, we proposed to consider a new metric that is the Zero mean Normalized Cross-Correlation function (ZNCC). This correlation criterion is known to be more robust that the classical SSD to linear brightness variations. This paper gives the expressions required to perform a ZNCC-based visual servoing task. Experimental results demonstrates the efficiency of the proposed method.**

I. Overview

This paper relates to the direct visual servoing scheme. Visual servoing which consists in using data provided by a vision sensor for controlling the motions of a robot [2], [11] has proved to a very approach for a long time.

A visual servoing problem can be written as an optimization problem [15]. The goal of visual servoing is that, from an initial arbitrary pose, the camera reaches the desired pose $\mathbf{r}^*$ that better satisfies some properties measured in or from the images. If we note f, the function that measures the positioning error, then the visual servoing task can be written as:

$$\widehat{\mathbf{r}} \;=\; \arg\min_{\mathbf{r}} f(\mathbf{r}, \mathbf{r}^*). \tag{1}$$

The visual servoing problem can therefore be considered as an optimization of the function f where $\mathbf{r}$ is incrementally updated to reach an optimum of f at $\widehat{\mathbf{r}}$ (if f is correctly chosen at the end of the minimization we should have $\widehat{\mathbf{r}} = \mathbf{r}^*$). The pose update is performed by applying a velocity $\mathbf{v}$, corresponding to the direction of descent, to the camera that is mounted on a robot: $\mathbf{r}_{k+1} = \mathbf{r}_k \oplus \mathbf{v}$ where "$\oplus$" is the operator that updates the pose and which is "implemented" through the robot controller.

Classical visual servoing approaches [2] consider a function f based on distance between geometrical features extracted from the image. The visual feature $\mathbf{s}$ can be 2D feature leading to image-based visual servoing approach (IBVS) or 3D features (such as the camera pose) leading to position-based visual servoing approaches (PBVS). These visual features (points, lines, moments, contours, pose, etc) have thus to be selected of extracted from the images to control the desired degrees of freedom of the robot. The minimization process has to be designed so that these visual features $\mathbf{s}(\mathbf{r})$ reach a desired value $\mathbf{s}^*$ (that is function of the desired posed $\mathbf{r}^*$), leading to a correct realization of the task. The optimization problem can thus be written:

$$\widehat{\mathbf{r}} \;=\; \arg\min_{\mathbf{r}} \; (\mathbf{s}(\mathbf{r}) - \mathbf{s}^*)\,. \tag{2}$$

Although very efficient, these approaches have some drawbacks. First, features (points, lines, moments, contours, pose, etc) have to be chosen depending on the scene characteristics. Second, the current features $\mathbf{s}(\mathbf{r})$ have to be tracked in real-time over frame and matched with the desired ones $\mathbf{s}^*$. Despite progressions in computer vision, this tracking issue is far from being solved. Finally, the tracking and matching tasks are prone to some measurement errors that cause the visual servoing task to be less accurate than it could be.

For some year now, more global approaches have been proposed. Not considering the specific image features but the image as a whole. These "direct" approaches have many advantages [4], [8], [13], [17]. In this context, visual servoing task is defined as an alignment between the current image $\mathbf{I}(\mathbf{r})$ and the image acquired at the desired camera pose $\mathbf{I}^*$. The main characteristic of these method is to closely tie the image to the control. Although this basic formulation is quite simple, it raises many modeling issue. First, despite the fact that all these techniques consider the image as a whole, it does not mean that some image transformations cannot be done. Second, whereas in the features based approach the alignment function is usually a simple Euclidian distances, here one can consider other criterion. In the general case, the optimization problem can be rewritten by:

$$\widehat{\mathbf{r}} \;=\; \arg\min_{\mathbf{r}} \; f\left(g(\mathbf{I}(\mathbf{r})), g(\mathbf{I}^*)\right). \tag{3}$$

where $g(.)$ is a transformation of the image and $f(.)$ is the alignment function. Let us note that in any case to build the control, one has to determine the interaction matrix that relies variation of the cost function to the camera velocity. When g is the identity, this lead to the photometric visual servoing approach [4], [5]. More complex visual information can be considered such as color invariant [3], a subspace computed from the image using a PCA approach [8], [18], the image gradient [17], the image entropy [6], Kernel information [13], mixture of Gaussian [1]. Dealing with the alignment function f, it can be as simple as the Euclidian distance [3], [4], [17] or, depending of the transformation f, more complex. When $g(.)$ measure the entropy of the image one has to consider for $f(.)$ the mutual information [6] that measures the quantity of information shared by two images.

In [4] the simplest transformation function f has been considered: the SSD or sum of squared differences (f is then the identity). Although the approach based on the luminance is very efficient in constrained environments, the cost function may be affected by some illumination variations

Amaury Dame is with CNRS, IRISA, INRIA Rennes Bretagne-Atlantique, Lagadic research group, France, email: Amaury.Dame@irisa.fr. Eric Marchand is with Université de Rennes 1, IRISA UMR 6074, INRIA Rennes-Bretagne Atlantique, Lagadic research group, France, email: Eric.Marchand@irisa.fr. This work is supported by DGA under contribution to student grant.

or occlusion. In [6], it has been demonstrated that mutual information (to be maximized) is a very robust cost function which leads to an efficient visual servoing scheme. In this paper we propose a good trade-off between the SSD and the mutual information: In a tracking context [12], this approach has proved to be very robust to linear brightness variation [9] thanks to the normalization embodied into the ZNCC. While considering only image intensity, this correlation criterion is more robust that the simple SSD (used in [4]) and less complex to compute that the mutual information (used in [6]).

The next section presents the ZNCC computation along with the derivation of the associated control law. Results will then be presented.

II. Zero-mean Normalized Cross Correlation

Along with the SSD, other classical correlation functions include the Normalized Cross Correlation (NCC) and the Zero-mean Normalized Cross Correlation (ZNCC).

The NCC is given by:

$$NCC(\mathbf{I},\mathbf{I}^*) = \frac{\sum_{\mathbf{x}} \mathbf{I}(\mathbf{x})\mathbf{I}^*(\mathbf{x})}{\sigma_{\mathbf{I}}\sigma_{\mathbf{I}^*}} \tag{4}$$

where $\sigma_{\mathbf{I}}$ and $\sigma_{\mathbf{I}^*}$ are the standard deviation of the two images defined by:

$$\sigma_{\mathbf{I}} = \sqrt{\sum_{\mathbf{x}} \left(\mathbf{I}(\mathbf{x}) - \overline{\mathbf{I}}\right)^2} \tag{5}$$

This normalization embodied into the NCC allows for tolerating linear brightness variations.

The Zero-mean Normalized Cross Correlation between two images $\mathbf{I}$ and $\mathbf{I}^*$ is given by:

$$ZNCC(\mathbf{I},\mathbf{I}^*) = \frac{\sum_{\mathbf{x}} \left(\mathbf{I}(\mathbf{x}) - \overline{\mathbf{I}}\right)\left(\mathbf{I}^*(\mathbf{x}) - \overline{\mathbf{I}^*}\right)}{\sigma_{\mathbf{I}}\sigma_{\mathbf{I}^*}} \tag{6}$$

where $\overline{\mathbf{I}}$ and $\overline{\mathbf{I}^*}$ are respectively the average intensities of the images $\mathbf{I}$ and $\mathbf{I}^*$:

$$\overline{\mathbf{I}} = \frac{1}{N_c}\sum_{\mathbf{x}} \mathbf{I}(\mathbf{x}) \tag{7}$$

where N_c is the number of pixels considered in the summation. Thanks to the subtraction of the image mean, ZNCC can tolerate uniform brightness variations and, thus, provides better robustness than the NCC [9], [10]. It is invariant to linear radiometric changes. In the reminder of the paper we will note the zero-mean images $\hat{\mathbf{I}}$ and $\hat{\mathbf{I}}^*$ with $\hat{\mathbf{I}}(\mathbf{x}) = \mathbf{I}(\mathbf{x}) - \overline{\mathbf{I}}$.

Figures 1 and 2 shows the differences between SDD, ZNCC for local and global brightness intensity of the images.

Global illuminations affect every pixel of the image in a consistent way. We limit our study to variations that affect the intensities with a simple affine relationship. We compute the cost function using the same reference image and a current image acquired under a more intense light. The results are presented in Figure 1. Dealing with the SSD, we mainly observe a shift of the values of SSD. Indeed, the optimum of the SSD function corresponding to the position where the smallest (respectively the highest) intensities of the current image are matched with the smallest (respectively the highest) intensities of the reference image, and this is typically enough to be robust with respect to a linear change of the pixel intensities. ZNCC, which is by definition robust to linear relationships between the intensities of the images pixels gives the correct alignment position.

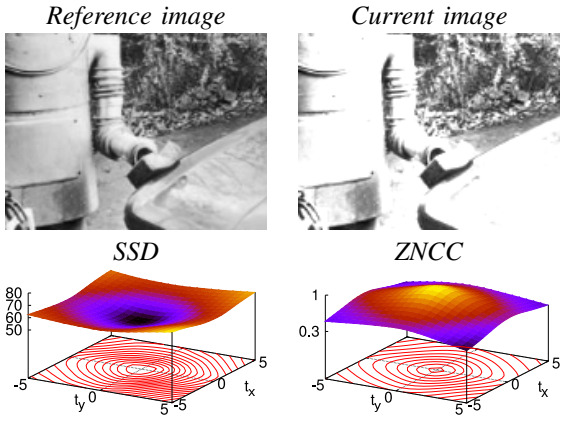


Fig. 1. Robustness of the cost functions with respect to global illumination variations.

Local illumination variations partially change the illumination of the scene. They appear in two major conditions: when surrounding objects partially increase the illumination (as a focused light would do) or decrease the illumination (as an object occluding the light directed to the scene) or when the object is non Lambertian, which does not diffuse an isotropic light (same light in all directions). This time (see Figure 2), the relationship between the pixel intensities is different from one part of the image to another part, so the minimum of SSD is affected by the illumination variation and gives a wrong estimation of the correct alignment position.

Since the relation between the pixels intensities is different from one part of the image to another, ZNCC is affected by the variation. The modification is visible on the shapes of the cost function, nevertheless the current and learned images still share a lot of common information, so the maximum is still located on the correct alignment position. ZNCC can therefore be considered as strongly robust to global and local illumination variations.

III. ZNCC based control law

The solution, proposed in this paper, is to define the alignment function f as the ZNCC between the two images. During the visual servoing task the current image is varying with respect to the camera pose $\mathbf{r}$ and can therefore be rewritten as $\mathbf{I}(\mathbf{r})$. The ZNCC becomes:

$$ZNCC(\mathbf{I}(\mathbf{r}),\mathbf{I}^*)) = \frac{\sum_{\mathbf{x}} \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})\hat{\mathbf{I}}^*(\mathbf{x})}{\sigma_{\mathbf{I}}(\mathbf{r})\sigma_{\mathbf{I}^*}} \tag{8}$$

A. Control law

To reach the desired pose, the goal is to maximize the ZNCC between the desired and current image. The ZNCC function is a quasi-concave function. The problem is therefore similar as the maximization of the MI in [7] and as in the other direct approaches, we still use the entire information

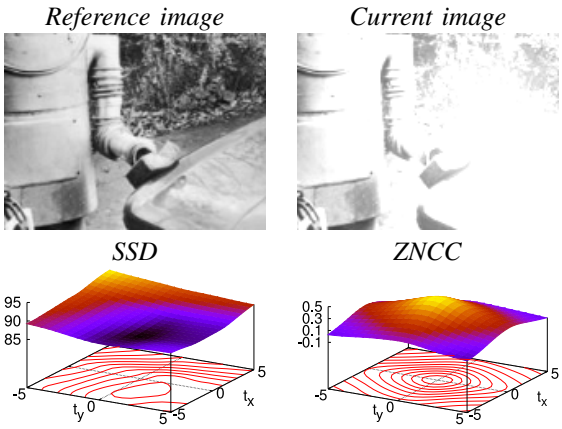


Fig. 2. Robustness of the cost functions with respect to global illumination variations.

provided by the images $\mathbf{I}$ and $\mathbf{I}^*$ by computing the following optimization:

$$\widehat{\mathbf{r}} = \arg\max_{\mathbf{r}} ZNCC(\mathbf{I}(\mathbf{r}), \mathbf{I}^*). \tag{9}$$

The problem of finding the camera pose for which the ZNCC will be maximal can be reformulated as finding the velocity that bring the ZNCC derivatives to a null value. The control law is then given by:

$$\mathbf{v} = \mathbf{H}^{-1}\mathbf{L} \tag{10}$$

where $\mathbf{L}$ is the interaction matrix of the ZNCC with respect to the camera pose and $\mathbf{H}$ is the Hessian matrix and $\mathbf{v} = (\boldsymbol{v}, \boldsymbol{\omega})$ is the computed camera velocity. In practice, and as usual in visual servoing [2], we use the Hessian $\mathbf{H}^*$ computed as if the camera was at convergence in the nominal conditions (computed using $\mathbf{I} = \mathbf{I}^*$).

B. Computing the interaction matrix

Given equation (8) and using the chain rules, the interaction matrix $\mathbf{L}$ is given by:

$$\mathbf{L} = \frac{\sum_{\mathbf{x}} \frac{\partial \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}} \hat{\mathbf{I}}^*(\mathbf{x})}{\sigma_{\mathbf{I}}(\mathbf{r})\sigma_{\mathbf{I}^*}} - \frac{1}{\sigma_{\mathbf{I}}(\mathbf{r})} \frac{\partial \sigma_{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}} ZNCC(\mathbf{r}) \tag{11}$$

To compute the derivative to the ZNCC we have first to compute the derivative of the zero-mean image intensities and the covariance of the current image with

$$\frac{\partial \hat{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}} = \frac{\partial \mathbf{I}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}} - \frac{1}{N_c} \sum_{\mathbf{x}} \frac{\partial \mathbf{I}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}} \tag{12}$$

The derivative of the covariance $\sigma_{\mathbf{I}}(\mathbf{r})$ is also obtained by the derivative chain rules:

$$\frac{\partial \sigma_{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}} = \frac{\sum_{\mathbf{x}} \frac{\partial \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}} \hat{\mathbf{I}}(\mathbf{r})}{\sigma_{\mathbf{I}}(\mathbf{r})} \tag{13}$$

Using the OFCE the derivative of a pixel intensity with respect to the camera position is given by [4], [16]:

$$\frac{\partial \mathbf{I}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}} = \nabla \mathbf{I}(\mathbf{x}) \mathbf{L}_{\mathbf{x}} \tag{14}$$

where $\nabla \mathbf{I}(\mathbf{x})$ are the gradients of the image expressed in the meter space at the point $\mathbf{x}$ and $\mathbf{L}_{\mathbf{x}}$ is the interaction matrix that links the displacement of the point in the image space with respect to the camera velocity. The interaction matrix is given by [2]:

$$\mathbf{L}_{\mathbf{x}} = \begin{bmatrix} -1/Z & 0 & x/Z & xy & -(1+x^2) & y \\ 0 & -1/Z & y/Z & 1+y^2 & -xy & -x \end{bmatrix}$$

where (x, y) are the coordinates of the point expressed in meters in the image plan and Z is its depth relatively to the camera. In this work we consider that the depth of the scene is unknown and thus we simply set the depth of each point to a constant.

C. Computing the Hessian

The second order derivatives of the ZNCC is obtained by computing the derivative of the equation (11) and performing some simplifications:

$$\mathbf{H} = \frac{1}{\sigma_{\mathbf{I}}(\mathbf{r})\sigma_{\mathbf{I}^*}} \sum_{\mathbf{x}} \frac{\partial^2 \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}^2} \hat{\mathbf{I}}^*(\mathbf{x}) + \frac{1}{\sigma_{\mathbf{I}}(\mathbf{r})} \frac{\partial^2 \sigma_{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}^2} ZNCC(\mathbf{r}) \tag{15}$$

where the second order derivative of the covariance $\sigma_{\mathbf{I}}(\mathbf{r})$ is

$$\frac{\partial^2 \sigma_{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}^2} = \left(\sum_{\mathbf{x}} \left(\frac{\partial^2 \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}^2} \hat{\mathbf{I}}^*(\mathbf{x}) - \frac{\partial \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}}^\top \frac{\partial \hat{\mathbf{I}}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}} \right) - \frac{\partial \sigma_{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}}^\top \frac{\partial \sigma_{\mathbf{I}}(\mathbf{r})}{\partial \mathbf{r}} \right) \frac{1}{\sigma_{\mathbf{I}}(\mathbf{r})^2} \tag{16}$$

and

$$\frac{\partial^2 \mathbf{I}(\mathbf{r},\mathbf{x})}{\partial \mathbf{r}^2} = \nabla \mathbf{I}_x \mathbf{H}_{\mathbf{x}} + \nabla \mathbf{I}_y \mathbf{H}_{\mathbf{y}} + \mathbf{L}_{\mathbf{x}}^\top \nabla^2 \bar{\mathbf{I}} \, \mathbf{L}_{\mathbf{x}} \tag{17}$$

where $\nabla^2 \bar{\mathbf{I}} \in \mathbb{R}^{2\times 2}$ is the gradient of $\nabla \bar{\mathbf{I}}$ in the metric space and $\mathbf{H}_{\mathbf{x}}$ and $\mathbf{H}_{\mathbf{y}}$ are respectively the derivatives of the first and second line of the interaction matrix $\mathbf{L}_{\mathbf{x}}$ (see [14] for the computation of the two Hessian matrices).

IV. 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 along with the interaction matrix computation required 80ms (for 320×240 images). Let us emphasize the fact that for all these experiments, the six degrees of freedom of our robot are controlled.

We conduct a full set of experiment considering various conditions (planar and non planar scenes, local and global lighting variation, comparison with photometric visual servoing [4],...). In each case we report the evolution of the value $ZNCC$ of the cost function (see equation (6)), 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 centimeter and radian.

A. *Positioning tasks under nominal condition*

For the first set of experiments the initial error pose was $\mathbf{\Delta r_{init}} = (17$ cm, -2, -3cm, $13.7°$, $5°$, $28°)$. This is a quite large displacement as can be seen on the first row of Figure 3 which features the initial and desired images. Only around 60% of the image pixels are common to both images. From this position,

In the first experiment the scene is planar and 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 $Z^* = 70$ cm, which is a coarse approximation. As can be seen on Figure 3.e the ZNCC increases rapidly and reach it maximal value in 50 iterations. Considering that image processing and control law computation is achieved in 80ms, convergence is reached in 4 seconds. The robot reaches a pose error $\Delta\mathbf{r}$ below 0.2mm in translation on each axis and $0.01°$ in rotation despite a distance to the scene of approximately 0.7 meter.

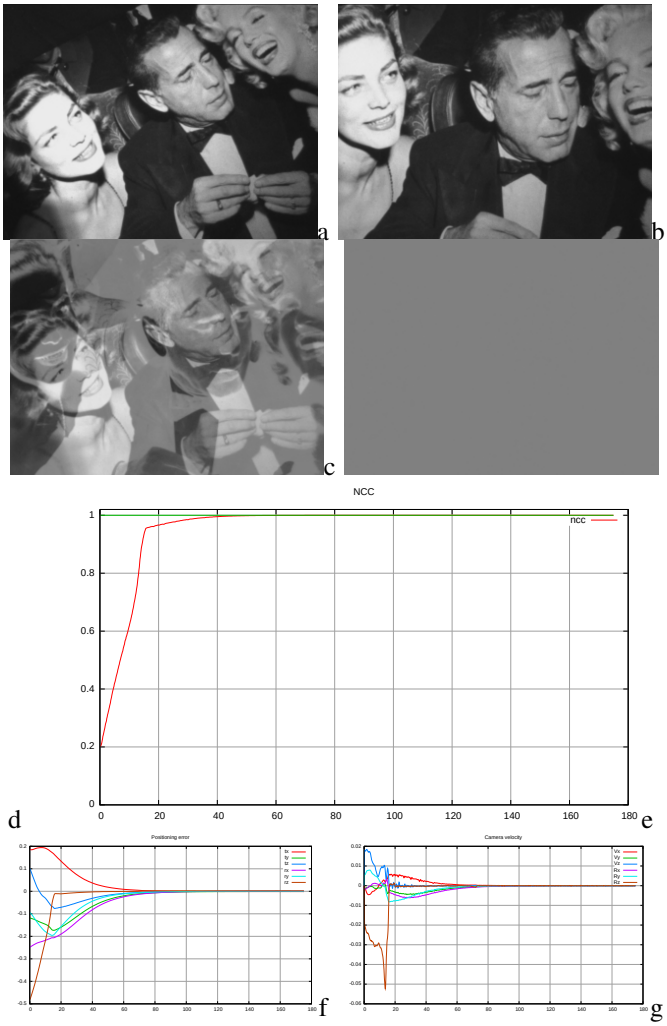


Fig. 3. ZNCC based visual servoing under nominal condition. (a-b) initial and desired image acquired by the camera. (c-d) initial and final error (that is $\mathbf{I} - \mathbf{I}^*$) (e) ZNCC criterion (f) positioning error (in cm and rad) (g) camera velocities in cm/s and rad rad/s

For comparison issue we have also considered the photometric visual servoing as described in [4]. Under nominal condition that is no lighting are considered reached precision and camera trajectory are similar (see Figure 4).

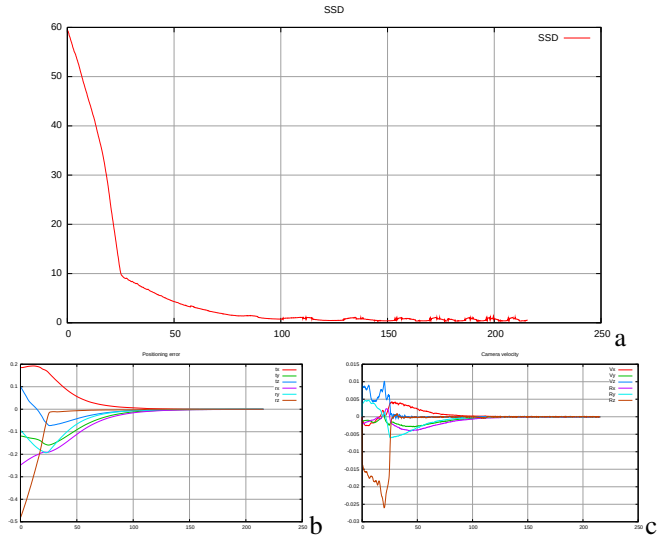


Fig. 4. SSD based visual servoing under nominal condition [4]. (a) SSD criterion (b) positioning error (in cm and rad) (c) camera velocities in cm/s and rad rad/s

B. *Robustness to non-planar scene*

The goal of the last experiment is to show the robustness of the proposed control law wrt depth variations. For this purpose, a non planar scene has been used as shown on Figure 5. Large errors in the depth are introduced (some object are 20cm high whereas the camera is at 70cm from the scene). Let us recall that since depths are not known (and can hardly be recovered on-line), assuming a constant depth in the interaction matrix introduces a modeling error in the control law. Despite these modeling errors and the fact that cost function to be maximized is more non-linear, ZNCC increases and the positioning error always decreases leading to similar positioning accuracy (that is less that 0.2mm and 0.1deg).

C. *Behavior with lighting variations*

As already mentioned the normalization embodied into the NCC allows for tolerating linear brightness variations. Furthermore thanks to the subtraction of the image mean, ZNCC can tolerate uniform brightness variations and thus provides better robustness than the SSD. In this experiment we radically modified the lighting condition during the realization of the positioning task. Figure 6a shows the first image acquired by the camera while Figure 6b show the desired one. Figure 6c show images acquired during the positioning task (at iteration 1, 75 and at convergence). One can see that most of the light has been shut down at the beginning of the experiment. Although ZNCC is maximized (see Figure 7a), its value at convergence is no longer 1. Despite the light modification and the fact that the SSD $\mathbf{I}-\mathbf{I}^*$ is obviously no longer null at convergence as can be seen on Figure 6d, ZNCC, as expected, allows a precise repositioning process. Precision is of 0.8mm in translation and 0.3 deg

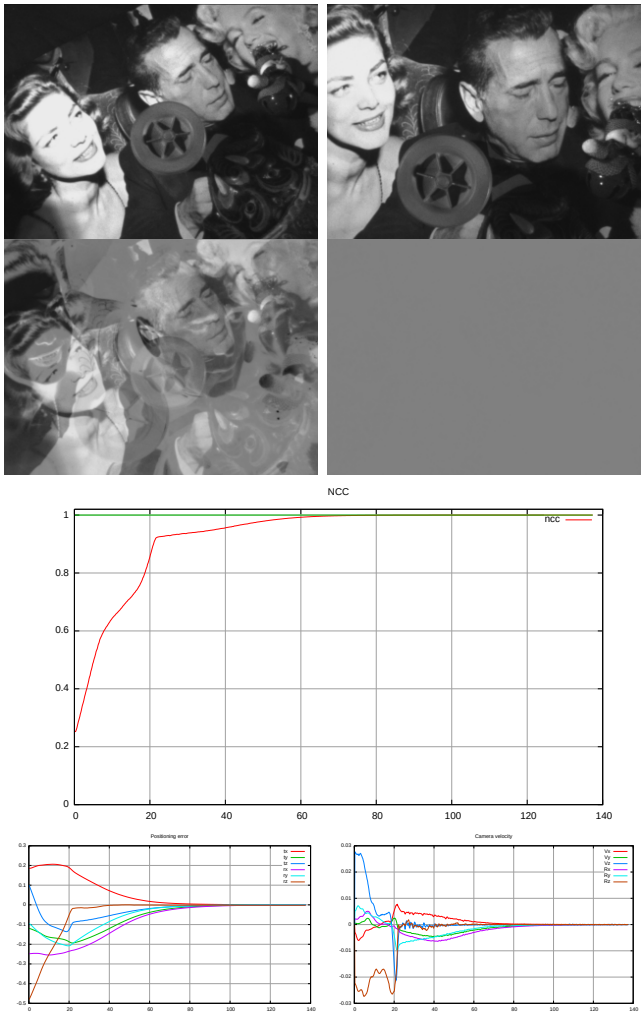


Fig. 5. ZNCC based visual servoing on a 3D scene (a-b) initial and desired image acquired by the camera. (c-d) initial and final error (that is $\mathbf{I}-\mathbf{I}^*$) (e) ZNCC criterion (f) positioning error (in cm and rad) (g) camera velocities in cm/s and rad rad/s

in rotation. Let us note that SSD based photometric visual servoing failed in this case.

The same experiment is considered on a 3D scene. With respect to the previous experiment, along with global lighting variation, shadows due to 3D objects have to be handled (see Figure 8).

Finally, we considered a scene with various metallic tools and a ring-light mounted around the camera. This configuration leads to a constant but moving light source. Furthermore considering that object are metallic, we have moving specularities. Despite these difficulties the proposed control law allows a fast positioning task with a good final accuracy.

V. Conclusion

In this paper we focused on direct visual servoing. We present a pure photometric approach. Considering that the registration criterion using in [4] is not robust to important lighting variation, we have considered another similarity criterion: the ZNCC. The new proposed criterion shows interesting properties since it is robust to illumination variations



Fig. 6. ZNCC based visual servoing with important lighting variations (a-b) initial and desired image acquired by the camera. (c-d) images acquired during the positioning task (iteration 1, 75 and 1000) and the corresponding error

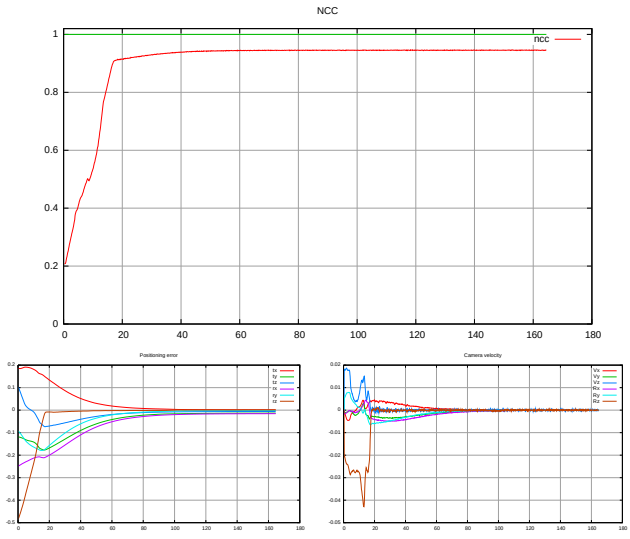


Fig. 7. ZNCC based visual servoing with important lighting variations (a) ZNCC criterion (c) positioning error (in cm and rad) (d) camera velocities in cm/s and rad rad/s



Fig. 8. ZNCC based visual servoing with important lighting variations on a 3D scene (a-b) initial and desired image acquired by the camera. (c-d) images acquired during the positioning task (iteration 1, 60 and 1300) and the corresponding error (e) ZNCC criterion

between the reference and the current image. Furthermore let us emphasize that, as for previous direct visual approaches, the method does not require any feature extraction or matching/tracking process and thanks to the use of highly redundant information it allows to achieved very accurate positioning task.

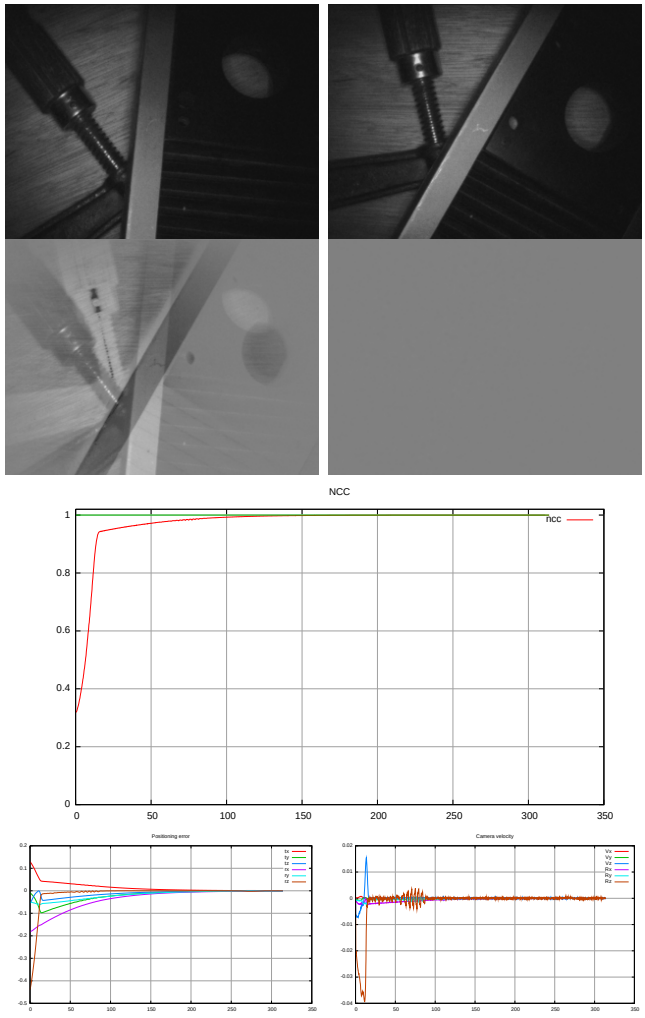


Fig. 9. Positioning task with respect with various metallic tools. A ring-light is mounted around the camera. Light is then moving.

REFERENCES

[1] A.H. Abdul Hafez, S. Achar, and C.V. Jawahar. Visual servoing based on gaussian mixture models. In *IEEE Int. Conf. on Robotics and Automation, ICRA'08*, pages 3225–3230, Pasadena, California, May 2008.

[2] F. Chaumette and S. Hutchinson. Visual servo control, Part I: Basic approaches. *IEEE Robotics and Automation Magazine*, 13(4):82–90, December 2006.

[3] C. Collewet and E. Marchand. Colorimetry-based visual servoing. In *IEEE Int. Conf. on Intelligent Robots and Systems, IROS'09*, St Louis, USA, October 2009.

[4] C. Collewet and E. Marchand. Photometric visual servoing. *IEEE Trans. on Robotics*, 27(4):828–834, August 2011.

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

[6] A. Dame and E. Marchand. Entropy-based visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'09*, pages 707–713, Kobe, Japan, May 2009.

[7] A. Dame and E. Marchand. Improving mutual information based visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'10*, pages 5531–5536, Anchorage, Alaska, May 2010.

[8] K. Deguchi. A direct interpretation of dynamic images with camera and object motions for vision guided robot control. *Int. Journal of Computer Vision*, 37(1):7–20, June 2000.

[9] L. Di Stephano, S. Mattoccia, and F. Tombari. ZNCC-based template matching using bounded partial correlation. *Pattern Recognition Letters*, 26:2129–2134, 2005.

[10] O. Faugeras, T. Viéville, and et al. Real-time correlation-based stereo: algorithm, implementations and applications. Technical Report 2013, INRIA, 1993.

[11] S. Hutchinson, G. Hager, and P. Corke. A tutorial on visual servo control. *IEEE Trans. on Robotics and Automation*, 12(5):651–670, October 1996.

[12] M. Irani and P. Anandan. Robust multi-sensor image alignment. In *IEEE Int. Conf. on Computer Vision, ICCV'98*, pages 959–966, Bombay, India, 1998.

[13] V. Kallem, M. Dewan, J.P. Swensen, G.D. Hager, and N.J. Cowan. Kernel-based visual servoing. In *IEEE/RSJ Int. Conf. on Intelligent Robots and System, IROS'07*, pages 1975–1980, San Diego, USA, October 2007.

[14] J.-T. Lapresté and Y. Mezouar. A Hessian approach to visual servoing. In *IEEE/RSJ Int. Conf. on Intelligent Robots and System, IROS'04*, pages 998–1003, Sendai, Japan, October 2004.

[15] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE Int. Conf. on Robotics and Automation, ICRA'04*, volume 2, pages 1843–1848, New Orleans, April 2004.

[16] E. Marchand. Control camera and light source positions using image gradient information. In *IEEE Int. Conf. on Robotics and Automation, ICRA'07*, pages 417–422, Roma, Italia, April 2007.

[17] E. Marchand and C. Collewet. Using image gradient as a visual feature for visual servoing. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'10*, Taipei, Taiwan, October 2010.

[18] S.K. Nayar, S.A. Nene, and H. Murase. Subspace methods for robot vision. *IEEE Trans. on Robotics*, 12(5):750 – 758, October 1996.