

Visual Servoing Based on Shifted Moments

Omar Tahri, Aurélien Yeremou Tamtsia, Youcef Mezouar, Cédric Demonceaux

▶ **To cite this version:**

Omar Tahri, Aurélien Yeremou Tamtsia, Youcef Mezouar, Cédric Demonceaux. Visual Servoing Based on Shifted Moments. IEEE Transactions on Robotics, 2015, 31, pp.798 - 804. 10.1109/TRO.2015.2412771 . hal-01485458

HAL Id: hal-01485458
https://hal.science/hal-01485458

Submitted on 8 Mar 2017

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.

Visual Servoing based on Shifted Moments

Omar Tahri, Aurélien Yeremou Tamtsia, Youcef Mezouar and Cédric Demonceaux

***Abstract*—In the last decade, image-moments have been exploited in several visual servoing schemes for their ability to represent object regions, objects defined by contours or a set of discrete points. Moments have also been useful to achieve control decoupling properties and to choose minimal number of features to control the whole degrees of freedom of a camera. However the choice of moment-based features to control the rotational motions around the x-axis and y-axis simultaneously with the translational motions along the same axis remains a key issue. In this paper, we introduce new visual features computed from low order 'shifted moments invariants'. Importantly, they allow (1) to define a unique combination of visual features to control the whole 6 DOF of a eye-in-hand camera independently from the object shape and (2) to significantly enlarge the convergence domain of the closed loop system.**

***Index Terms*—Visual servoing, shifted moments.**

I. Introduction

In image-based visual servoing (IBVS), the control loop is directly closed in the image space of the vision sensor. Compared to position-based visual servoing [25] and hybrid visual servoing [14], IBVS schemes get rid of 3D reconstruction steps to compute the visual features. This yields a high robustness to disturbances and to modeling errors [4], [15]. However, IBVS has been usually considered as suitable only for reasonable displacements because of the non-linearities in the relationship between the image space and workspace and the coupling between degrees of freedom (DOF) [2]. This turns the analysis of the global asymptotic stability of IBVS as an almost unfeasible task except for some special cases [10], [18]. An other important issue relies on complexity of the required image processing step to extract the features used as input of the visual servoing process.

A vision sensor provides a large spectrum of potential visual features. In practice, the choice of visual features influences the performance of the closed loop system and the ability to analyze the system's dynamics. Several global image features have been exploited in the context of visual servoing to avoid complex image processing steps (extraction, matching and tracking): Fourier descriptors in [5], light intensity in [12] or luminance signal in [19], [8], [6], [7]. However, one important issue when exploiting global image features is the reduced convergence domain due to strong non-linearities in the system dynamics. The work presented in this paper exploits image moments which also belong to the family of globlal image features. Several types of moments can be used: geometric moments [11], Zernike moments [13] and Legendre moments [23], etc. Moment-based visual servoing has led to interesting results (refer to [3], [22], [17]). Moments have also been used for pose estimation [21], and for task sequencing [16]. In this paper, selection of minimal combinations of visual features from image moments is discussed. More precisely, based on shifted moments [26], the selection of a unique feature vector independently of the object shape is proposed and exploited in an image-based control scheme.

The remaining of this paper is organized as follows. Section II discusses several state-of-the art related methods and gives some preliminary materials. Section III formally presents our contributions which exploit shifted moments to design a decoupled image-based control law. Section IV evaluates the performances of our control law.

Parts of this paper were presented at ICRA'13, Karlsruhe, Germany, May 2013.

O. Tahri and C. Demonceaux are with Le2i Lab, University of Burgundy 71200 LE CREUSOT, France (email: omartahri@isr.uc.pt, cedric.demonceaux@u-bourgogne.fr).

A. Yeremou Tamtsia is with University of Douala, Cameroun (email: ayeremou@yahoo.fr).

Y. Mezouar is with Clermont Université, IFMA, Institut Pascal, BP 10448, F-63000 Clermont-Ferrand, France youcef.mezouar@ifma.fr.

II. Mathematical background

A. Moments

Let $f(x,y) \geq 0$ be a real bounded function with support on a compact region R. The $(p+q)$-order moments of f are defined as [20]:

$$m_{pq} = \int\int_R x^p y^q f(x,y) dx dy \tag{1}$$

The centered moments μ_{pq} are computed with respect to the object centroid (x_g, y_g)

$$\mu_{pq} = \int\int_R (x-x_g)^p (y-y_g)^q f(x,y)\ dx dy \tag{2}$$

and

$$x_g = \frac{m_{10}}{m_{00}}; \qquad y_g = \frac{m_{01}}{m_{00}}$$

where m_{00} represents the area of the object image. In the following, objects defined by a binary region or closed contour are considered. Therefore, we assume that $f(x,y) = 1$ in all the region that defines the object. Definitions necessary to the subsequent development are recalled in the next section.

B. Moment invariants

Moment invariants are important shape descriptors in computer vision. They are special functions of image moments that are insensitive to particular transformations [9]. They were first introduced to the pattern recognition community in 1962 [11]. Several types of transformations can be considered: translation, rotation, scaling, affine, projective, and elastic. The simplest transformations in the two dimensional space are Translation, Rotation and Scaling (TRS):

$$\mathbf{x}' = c \begin{bmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{bmatrix} \mathbf{x} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3}$$

where c is a scale factor change, α is a rotation angle and (t_x, t_y) are translation parameters. The centered moments defined from (2) are known to be invariant to 2D translational motions. The first rotation invariants computed from moments have been proposed by Hu in 1962 [11]. Invariants to both translation and rotation can be obtained using centered moments, for instance using the following polynomials:

$$i_0 = \mu_{20} + \mu_{02} \tag{4}$$

$$i_1 = \mu_{20}\mu_{02} - \mu_{11}\mu_{11} \tag{5}$$

$$i_2 = -\mu_{30}\mu_{12} + \mu_{21}\mu_{21} - \mu_{03}\mu_{21} + \mu_{12}\mu_{12} \tag{6}$$

$$i_3 = 3\mu_{30}\mu_{12} + \mu_{30}\mu_{30} + 3\mu_{03}\mu_{21} + \mu_{03}\mu_{03} \tag{7}$$

Invariants to translation and scaling can be computed as follows [11]:

$$i_s = \frac{\mu_{pq}}{m_{00}^{(p+q+2)/2}} \tag{8}$$

Besides, invariants to TRS can be obtained from (5), (6) and (7):

$$\begin{aligned} i_{n1} &= \frac{i_2}{i_1^{8/10}},\ i_{n2} = \frac{i_3}{i_1^{8/10}},\ i_{n3} = \frac{i_3}{i_2}, \\ i_{n4} &= \frac{i_3}{m_{00}^5},\ i_{n5} = \frac{i_2}{m_{00}^5},\ i_{n6} = \frac{i_1}{m_{00}^4} \end{aligned} \tag{9}$$

A more general transformation than TRS is the affine transformation:

$$\mathbf{x}' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mathbf{x} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{10}$$

Invariants to affine transformation can also be obtained (refer to [9]). For instance the ratio:

$$i_a = \frac{\mu_{20}\mu_{02} - \mu_{11}\mu_{11}}{m_{00}^4} \tag{11}$$

is an invariant to (10). Other invariants to affine transformations computed from moments of higher order can be found in [9]. In the next paragraph, we recall some previous results and explain how invariants are useful for visual servoing.

C. Moments invariant for visual servoing

1) Visual servoing: Let $\mathbf{s}$ be a set of N visual features used as input of the control scheme. The time variation $\dot{\mathbf{s}}$ of the visual features $\mathbf{s}$ can be linearly expressed with respect to the relative camera-object kinematics screw:

$$\dot{\mathbf{s}} = \mathbf{L_s}\boldsymbol{\tau}, \tag{12}$$

where $\mathbf{L_s}$ is the interaction matrix ($N \times 6$ matrix if 6 DOF are considered) related to $\mathbf{s}$. Usually, the control scheme is designed to reach an exponential decoupled decrease of the current visual features $\mathbf{s}$ to their goal value $\mathbf{s}^*$. If we consider an eye-in-hand system observing a static object, the corresponding control law is:

$$\boldsymbol{\tau_c} = -\lambda \widehat{\mathbf{L_s}}^{+}(\mathbf{s} - \mathbf{s}^*), \tag{13}$$

where $\widehat{\mathbf{L_s}}$ is a model or an approximation of $\mathbf{L_s}$, $\widehat{\mathbf{L_s}}^{+}$ the pseudo-inverse of $\widehat{\mathbf{L_s}}$, λ a positive gain tuning the time to convergence, and $\boldsymbol{\tau_c} = (\boldsymbol{\upsilon_c}, \boldsymbol{\omega}_c)$ the camera velocity sent to the low-level robot controller.

Invariants built from combinations of moments are useful (1) to select the minimal number m of visual features to control m DOF of a camera and (2) to select features providing a sparse interaction matrix (refer to [22]). The following paragraph presents how moments can be used to select good image features.

2) Features selection: In order to control the six DOF of a camera, a vector of six visual features $\mathbf{s} = [x_n,\ y_n,\ a_n,\ s_4,\ s_5,\ \theta]$ was proposed in [22], where:

- $a_n = Z^* \sqrt{\frac{m_{00}^*}{m_{00}}}$ with m_{00}^* is the desired area in the image, and Z^* the desired depth between the camera and the object.
- $x_n = a_n x_g$, $y_n = a_n y_g$, (x_g, y_g) are the object center coordinates in the image.
- θ is the orientation of the principal axes in the image.
- s_4 and s_5 are two different invariants to TRS transformations (as those given by (9)). They are used to control the rotation velocities w_x and w_y around $x - axis$ and $y - axis$.

When the object is parallel to the image plane, the interaction matrix related to $\mathbf{s}$ is [22]:

$$\begin{bmatrix} \mathbf{L}_{x_n}^{\parallel} \\ \mathbf{L}_{y_n}^{\parallel} \\ \mathbf{L}_{a_n}^{\parallel} \\ \mathbf{L}_{s_4}^{\parallel} \\ \mathbf{L}_{s_5}^{\parallel} \\ \mathbf{L}_{\theta}^{\parallel} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & a_n\varepsilon_{11} & -a_n(1+\varepsilon_{12}) & y_n \\ 0 & -1 & 0 & a_n(1+\varepsilon_{21}) & -a_n\varepsilon_{11} & -x_n \\ 0 & 0 & -1 & -3y_n/2 & 3x_n/2 & 0 \\ 0 & 0 & 0 & s_{4wx} & s_{4wy} & 0 \\ 0 & 0 & 0 & s_{5wx} & s_{5wy} & 0 \\ 0 & 0 & 0 & \theta_{wx} & \theta_{wy} & -1 \end{bmatrix} \tag{14}$$

where $\varepsilon_{11} = n_{11} - x_g y_g/2$, $\varepsilon_{12} = n_{20} - x_g^2/2$, $\varepsilon_{21} = n_{02} - y_g^2/2 \varepsilon_{11} = n_{11} - x_g y_g/2$, $\varepsilon_{12} = n_{20} - x_g^2/2$, $\varepsilon_{21} = n_{02} - y_g^2/2$ and $n_{ij} = \frac{\mu_{ij}}{m_{00}}$. The analytical forms of s_{4wx}, s_{4wy}, s_{5wx} and s_{5wy} depend on the chosen invariants, and θ_{w_x} , θ_{w_y} can be found in [22].

Note first that the translational part of the interaction matrices related to x_n, y_n and a_n form a block diagonal matrix. Moreover, thanks to the invariance to TRS, the features used to control rotations s_4, s_5 and θ are invariant to translational motions when the object and the camera plane are parallel. These three features are used to control the rotational DOF. The resulting interaction matrix for the features vector $\mathbf{s}$ is sparse which simplifies the singularities analysis of $\mathbf{L}_{\mathbf{s}}^{\parallel}$. Let us now define $\mathbf{L}_{s_{4,5}}^{\parallel}$ as the matrix that links the variation s_4 and s_5 to rotational velocities around the $x-axis$ and $y-axis$:

$$\mathbf{L}_{s_{4,5}}^{\parallel} = \begin{bmatrix} s_{4wx} & s_{4wy} \\ s_{5wx} & s_{5wx} \end{bmatrix}$$

If $\mathbf{L}_{s_{4,5}}^{\parallel}$ is non singular, then the interaction matrix $\mathbf{L}_{\mathbf{s}}^{\parallel}$ related to the features $\mathbf{s} = \begin{bmatrix} x_n & y_n & a_n & s_4 & s_5 & \theta \end{bmatrix}$ will be non-singular as well. This can be easily checked since we have $det(\mathbf{L}_{\mathbf{s}}^{\parallel}) = (-1)^4 det(\mathbf{L}_{s_{4,5}}^{\parallel})$. Unfortunately, the combination (s_4, s_5), as presented in a previous work [22], depends on the shape of the object and it is computed from moments with

orders ranging from 3 to 5 [22]. Therefore, despite the sparse form of the resulting interaction matrix, since the choice of the combination (s_4, s_5) depends of the object, the analysis of the closed loop system can not be carried out analytically. Additionally, since the sensitivity to data noise increases with the moment orders [24], we should use moments of orders as low as possible for selecting (s_4, s_5). In the next section, we introduce a new class of moments called 'shifted moments' from which it is possible to define a unique couple (s_4, s_5) whatever is the shape of the target object.

III. Shifted moments for visual servoing

A. Shifted moments

Let us first define a shifted point as $\mathbf{P^{sh}} = [x_g - x_{sh},\ y_g - y_{sh}]$ where (x_g, y_g) is the object center and x_{sh} and y_{sh} are shift parameters. Shifted moments can then be defined as follow:

$$\mu^{sh}_{pq} = \int\int_R (x - x_g + x_{sh})^p (y - y_g + y_{sh})^q \ dxdy \quad (15)$$

Note that for $x_{sh} = y_{sh} = 0$, the shifted moments are nothing but equal to the classical centered ones. Furthermore, as it will be shown in the following, for suitable choices of x_{sh} and y_{sh}, the invariance properties obtained using centered moments can still hold for shifted ones. The results proposed in this paper can then be seen as generalization of those previously obtained using classical centered moments with the improvements exhibited below.

To underline the advantage of shifted moments with respect to centered moments, let us consider a symmetric object. In this case, invariants computed from odd orders centered moments are zero and invariants to TRS (as for instance i_{n3} given by (9)) become useless because their denominators are zero. This is an important issue when dealing with moment-based visual servoing. Shifted moments naturally overcome this problem. This can be intuitively explained by the fact that the object symmetry is broken if its shape is parametrized with respect to a well chosen shifted point. In addition and as it will be shown in the results section, computing moments with respect to a shifted point has also beneficial effect on interaction matrix values.

In order to preserve the invariance to TRS, the shifted points have to be carefully chosen. A first option consists in selecting the shifted points from points of the object projection in the image. A second option consists in computing the shifted points from the object moments as we will show in the following.

B. Physical points as shifted points

Let us consider $\mathbf{P^{sh}} = [x_o\ y_o]^\top$ the coordinates of a point which is rigidly attached to the object. The moments computed with respect to $\mathbf{x_o}$ can be expressed as:

$$\mu^{sh}_{pq} = \int\int_R (x - x_o)^p (y - y_o)^q \ dxdy \quad (16)$$

As for the classical centered moments [3], [22], the time variation of the shifted moments can be obtained by:

$$\begin{aligned}\dot{\mu}^{sh}_{pq} = &\int\int_R p(x - x_o)^{p-1}(y - y_o)^q(\dot{x} - \dot{x}_o)dxdy \\ &+ \int\int_R q(x - x_o)^p(y - y_o)^{q-1}(\dot{y} - \dot{y}_o)dxdy \\ &+ \int\int_R (x - x_o)^p(y - y_o)^q(\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y})dxdy\end{aligned} \quad (17)$$

where $\dot{x}_0$ and $\dot{y}_0$ are linked to the camera screw through the interaction matrix related to the shifted point. They can be obtained by using the formula of the interaction matrix related to a point in the image: using the equation defining the object plane to compute the point depth (since the point belongs to the object) and the coordinates of the point in the image. If the current value of the interaction matrix is used in the control law, then to compute $\dot{x}_0$ and $\dot{y}_0$, the parameters of the object plane will be needed. Otherwise, if only the desired value of the interaction matrix is used, then only the object plane parameters for the desired position are needed.

For any point on a planar object, we have:

$$\frac{1}{Z} = \mathrm{A}x + \mathrm{B}y + C$$

where Z is the depth of the observed point and A, B, C are the plane parameters. Using similar calculation to those presented in [22], [3] for centered moments, the interaction matrix related to μ^{sh}_{pq} can be written as:

$$\mathbf{L}_{\mu^{sh}_{pq}} = \begin{bmatrix} \mu^{sh}_{v_x} & \mu^{sh}_{v_y} & \mu^{sh}_{v_z} & \mu^{sh}_{w_x} & \mu^{sh}_{w_y} & \mu^{sh}_{w_z} \end{bmatrix} \quad (18)$$

Where:

$$\begin{cases} \mu^{sh}_{v_x} = -(p+1)\,A\mu^{sh}_{pq} - pB\mu^{sh}_{p-1,p+1} \\ \mu^{sh}_{v_y} = -qA\mu^{sh}_{p+1,q-1} - (q+1)\,B\mu^{sh}_{pq} \\ \mu^{sh}_{v_z} = -A\mu^{sh}_{w_y} + B\mu^{sh}_{w_x} + (p+q+2)\,C\mu^{sh}_{pq} \\ \mu^{sh}_{w_x} = (p+q+3)\,\mu^{sh}_{p,q+1} + (p+2q+3)y_o\mu^{sh}_{p,q} + p\,x_o\mu^{sh}_{p-1,q+1} \\ \mu^{sh}_{w_y} = -(p+q+3)\,\mu^{sh}_{p+1,q} - (2p+q+3)\,x_o\mu^{sh}_{p,q} - q\,y_o\mu^{sh}_{p+1,q-1} \\ \mu^{sh}_{w_z} = p\mu^{sh}_{p-1,q+1} - q\mu^{sh}_{p+1,q-1} \end{cases}$$

Note that if $A = B = 0$, we have $\mu^{sh}_{v_x} = \mu^{sh}_{v_y} = 0$ which confirms the invariance to the corresponding translational motions. In addition, $\mu^{sh}_{w_z}$ and $\mu^{sh}_{v_z}$ have the same formulas than the one obtained with centered moments [22],[3]. Therefore using points from the object allows preserving the invariance to TRS obtained with functions of centered moments.

C. Shifted point computed from the object moments

Shifted points that preserve the invariance to TRS can be obtained from the object moments as follows:

$$\mathbf{P^{sh}} = [x_g,\ y_g] + a\,[\cos(\theta + \Delta),\ \sin(\theta + \Delta)] \quad (19)$$

where:

- a is an invariant to rotation and changes linearly with the object scale,
- θ is the object orientation in the image,
- Δ is a constant.

Let us show that any shifted point that holds the conditions mentioned above undergoes the same TRS transformation as the object. First, let us consider a translational motion in the image. In this case, since the object orientation θ and the parameter a are invariant to translation, then the value of $a\,[\cos(\theta+\Delta),\ \sin(\theta+\Delta)]$ does not change. Therefore, the point $\mathbf{P^{sh}}$ and the object center undergoes the same translational motion.

In the case where a rotation is considered in the image then the object center undergoes this rotation, a keeps the same value since it is an invariant to rotation, and the value of the angle θ changes by the angle β. This implies that $a\,[\cos(\theta+\Delta),\ \sin(\theta+\Delta)]$ undergoes the rotation defined by β. Since the object center $[x_g,\ y_g]$ and $a\,[\cos(\theta+\Delta),\ \sin(\theta+\Delta)]$ rotates as the object, $\mathbf{P^{sh}}$ will do the same.

Finally, if we consider a scale change in the image then the angle θ keeps the same value since it is independent from the scale. Furthermore, since the parameter a is assumed to change linearly with respect to the scale, this implies that $a\,[\cos(\theta+\Delta),\ \sin(\theta+\Delta)]$ and the object undergoes the same scale change. This means that the point $\mathbf{P^{sh}}$ follows the same scale change.

In practice, there exist at least as many choices for $\mathbf{P^{sh}}$ using the formula given by (19) as there are for the parameter a. In [26], we have proposed $a=(\mu_{20}+\mu_{02})^{\frac{1}{4}}$ to compute the points $\mathbf{P^{sh}}$. Other choices can be made using for instance moments of order 4 by setting $a=(\mu_{40}+2\mu_{22}+\mu_{04})^{\frac{1}{6}}$. The power $\frac{1}{4}$ and $\frac{1}{6}$ are used to make the parameter a changes linearly with respect to scale. The adequate power is obtained using the fact that a scale change in the image by a scalar equal to γ induces a change defined by a scalar γ^{p+q+2} on a moment μ_{pq}. For instance, the polynomial $\mu_{40}+2\mu_{22}+\mu_{04}$ is multiplied by γ^6 after a scale change defined by a a scalar γ. For this reason, the chosen power is $\frac{1}{6}$ in $a=(\mu_{40}+2\mu_{22}+\mu_{04})^{\frac{1}{6}}$ to obtain a linear change with respect to scale. In this paper, for more robustness to noise, we use $a=\sqrt{m_{00}/2}$ since m_{00} is the lowest order moment. All these choices of shifted points preserve the invariance obtained with functions of centered moments. One can also show this based on interaction matrix formula of the shifted moments. Indeed, when the object is parallel to the image plane, the form of the interaction matrix related to $\mathbf{P^{sh}}$ is identical to the one related to the center of gravity of the object.

D. Visual features selection

The features vector $\mathbf{s}=[x_n,\ y_n,\ a_n,\ s_4,\ s_5,\ \theta]$ is used to control the movements of the camera. The four features $(x_n,\ y_n,\ a_n,\ \theta)$ are defined as described in section II-C2. These features ensure nice decoupling and linearizing properties as shown by (14). The features s_4 and s_5 are obtained using an invariant to TRS (as one of those given by (9)) computed with respect to two different shifted points. The shifted points are chosen as previously detailed: a rigidly attached point to the object or from the moments.

IV. Results

In this section, we evaluate our control scheme. We consider that for the desired pose, the image and object planes are parallel. Remind that the non parallel case can be transformed to the parallel one as explained in [22].

A. Results using shifted points computed from moments

Three different object shapes a 'whale', an 'octopus' and a rectangle (refer to Fig. 1.a, Fig. 3.a and Fig. 5.a) are considered. A motion composed of a translation defined by the vector $\mathbf{T}=(-0.1,\ -0.2,\ 0.8)$ meter and a rotation defined by an angle equal to 30 degrees around the unitary vector $\mathbf{u}=(-1,\ -1,\ 0.5)/\sqrt{2.25}$ has been considered between the desired and the initial camera poses. The couple of features (s_4, s_5) is obtained by computing the invariant to TRS (refer to equations (6), (7) and (9)):

$$i^{sh}_{n3}=\frac{3\mu^{sh}_{30}\mu^{sh}_{12}+\mu^{sh}_{30}\mu^{sh}_{30}+3\mu^{sh}_{03}\mu^{sh}_{21}+\mu^{sh}_{03}\mu^{sh}_{03}}{-\mu^{sh}_{30}\mu^{sh}_{12}+\mu^{sh}_{21}\mu^{sh}_{21}-\mu^{sh}_{03}\mu^{sh}_{21}+\mu^{sh}_{12}\mu^{sh}_{12}}$$

with respect to two shifted points:

$$\mathbf{P_1}=[x_g+a\cos(\theta),\ y_g+a\sin(\theta)] \tag{20}$$

and

$$\mathbf{P_2}=\left[x_g+a\cos(\theta+\frac{\pi}{2}),\ y_g+a\sin(\theta+\frac{\pi}{2})\right] \tag{21}$$

The positions of these points in the image for the three objects are shown in Figs 1.a, 3.a and 5.a. The desired values of the interaction matrix obtained for the rectangle, the 'whale' and the 'octopus' are given respectively by (22), (23) and (24). From these values, it can be seen that in addition of making the invariant i^{sh}_{n3} computed from odd moment exploitable for symmetric object as a rectangle, using shifted moments also allowed to obtain well conditioned interaction matrices in the three cases. Indeed, changing the position of the shifted point allows to vary the value of the interaction matrix. Recent works have studied optimal choice of the two shifted points to be used in order to enlarge the convergence performance when using photometric moments [1].

$$\mathbf{L_s}=\begin{bmatrix} -1. & 0 & 0 & -0. & -1.04 & 0 \\ 0 & -1. & 0 & 1.08 & 0. & 0 \\ 0 & 0 & -1. & 0 & 0 & 0 \\ 0 & 0 & 0. & -2.21 & -0 & 0. \\ 0 & 0 & 0. & 0. & -0.14 & -0. \\ 0 & 0 & 0 & 0. & -0. & -1. \end{bmatrix} \tag{22}$$

$$\mathbf{L_s}=\begin{bmatrix} -1. & 0 & 0 & 0.01 & -1.10 & 0.08 \\ 0 & -1. & 0 & 1.03 & -0.01 & 0 \\ 0 & 0 & -1. & -0.13 & -0. & 0 \\ 0 & 0 & 0. & -0.36 & 5.71 & 0. \\ 0 & 0 & 0. & 0.48 & 2.04 & -0. \\ 0 & 0 & 0 & -0.07 & -0.14 & -1. \end{bmatrix} \tag{23}$$

$$\mathbf{L_s}=\begin{bmatrix} -1. & 0 & 0. & -0.03 & -1.02 & -0.10 \\ 0 & -1. & 0 & 1.10 & 0. & 0 \\ 0 & 0 & -1. & 0.15 & 0.03 & 0 \\ 0 & 0 & -0. & 18.05 & 3.11 & -0. \\ 0 & 0 & -0. & 3.45 & 1.24 & 0. \\ 0 & 0 & 0 & -0.09 & 0.04 & -1. \end{bmatrix} \tag{24}$$

The results obtained are depicted in Fig. 1, Fig. 3 and in Fig. 5. The servoing achieved using only the desired value of the interaction matrix shows convergence of the features errors (Figs 1.b, 3.b and 5.b) and velocities (Figs 1.c, 3.c and 5.c) for the three cases. The same simulations have been considered with a white noise (standard deviation equal to 1 pixel) on the coordinates of contour points. The behavior in the image for the three simulations with and without noise is shown in the video attached to this paper. The results show nice performances in presence of measurement noise.

B. Results using an affine invariant and two tracked points

In this part, the same objects are considered ('whale', 'octopus' and rectangle). The couple of features (s_4, s_5) is obtained by computing the affine invariant (11):

$$i_a = \frac{\mu^{sh}_{20}\mu^{sh}_{02} - \mu^{sh}_{11}\mu^{sh}_{11}}{m^4_{00}}$$

with respect to two points from the object contour (as shown for each object in Figs 2.a, 4.a and 6.a). The two points have been chosen from the object contour and are tracked during the servoing. Other types of shifted points can be chosen from the object image as Harris or SIFT points for instance as long as they can be tracked.

The motion to be achieved is composed by a translation defined by the vector $\mathbf{T} = (-0.5, 0.8, 1.3)$ (meter) and a rotation defined by an angle equal to 80 degrees around the unitary vector $\mathbf{u} = (\sqrt{1/2}, \sqrt{1/2}, 0)$. Note the huge and unusual rotational motion considered around the axes x and y of the camera. The results obtained for the three objects are depicted in Fig. 2, Fig. 4 and in Fig. 6. These results are obtained using only the desired value of the interaction matrix. They show convergence of the features errors (Figs 2.b, 4.b and 6.b) and velocities (Figs 2.c, 4.c and 6.c) despite the huge motion to be achieved. As for the previous simulations, a white noise (standard deviation equal to 1 pixel) has been added to the coordinates of contour points to evaluate the effect of noisy data. The behavior in the image for the three simulations with and without noise is also shown in the video attached to this paper. Once again, the obtained results show that noisy data do not disturb the behavior of the servoing.

V. Conclusion

In this paper, a class of invariants computed from shifted centered moments have been exploited to select two features to control efficiently the rotational motion around the $x-axis$ and $y-axis$. Two cases of shifted moments have been considered. The first case considers shifted parameters computed from object moments in the image and the second case considers shifted points from object contour. The second option uses affine invariants and requires extracting and tracking two points, but ensure higher performances since it allows convergence for huge motion (exceptional for an IBVS scheme). First, differently from invariants to TRS, affine invariants provide invariance to translations along the x-axis and y-axis whatever is the orientation of the object with respect to the camera plane. Second, the affine invariants introduced in

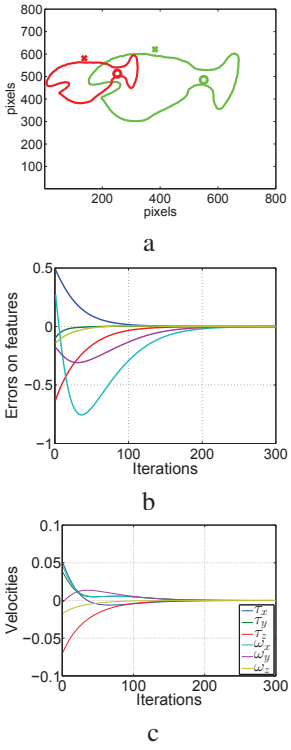


Fig. 1. Results obtained using a 'whale' contour: a) initial (red) and the desired (green) images, b) errors on features, c) velocities

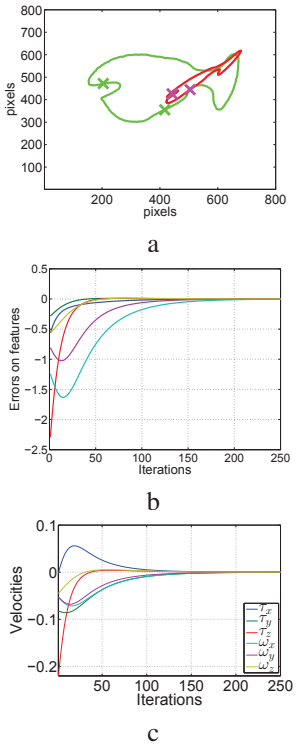


Fig. 2. Results obtained using a 'whale' contour: : a) initial (red) and the desired (green) images, b) errors on features, c) velocities

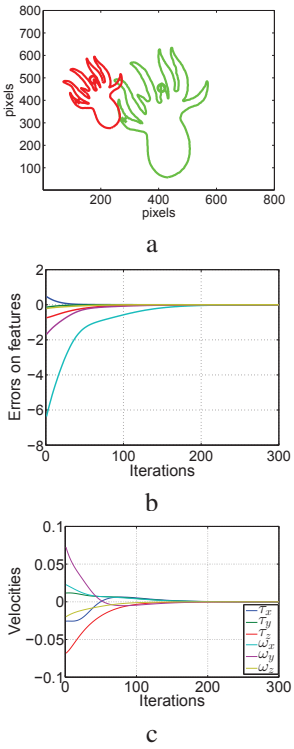


Fig. 3. Results obtained using a 'octopus' contour: a) initial (red) and the desired (green) images, b) errors on features, c) velocities

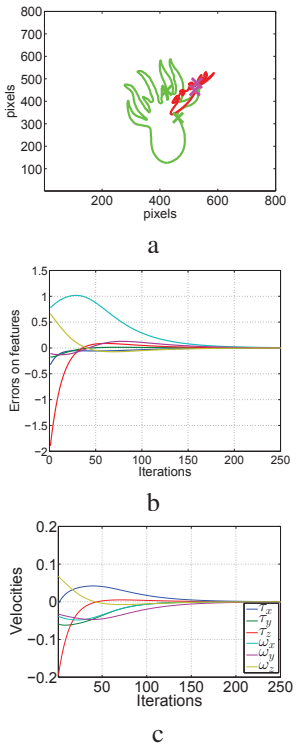


Fig. 4. Results obtained using a 'Octopus' contour: : a) initial (red) and the desired (green) images, b) errors on features, c) velocities

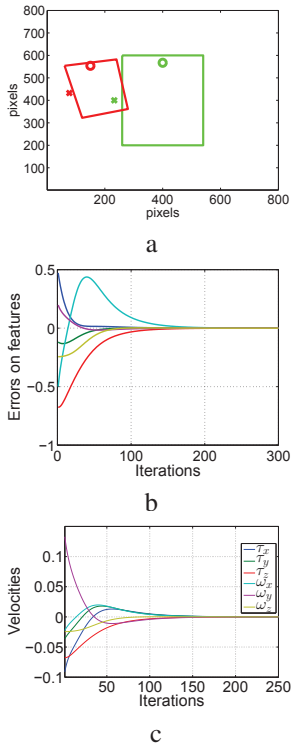


Fig. 5. Results obtained using a rectangle: a) initial (red) and the desired (green) images, b) errors on features, c) velocities

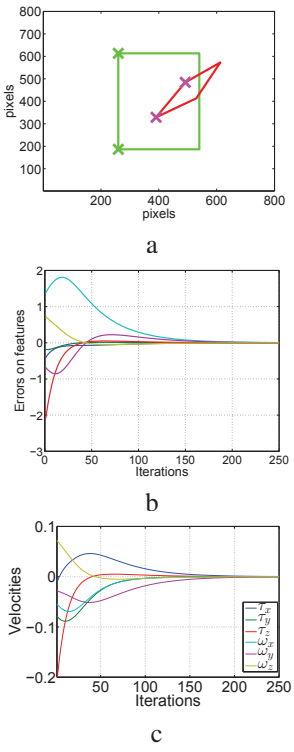


Fig. 6. Results obtained using a rectangle: a) initial (red) and the desired (green) images, b) errors on features, c) velocities

this paper are computed from moments of order 2 only. This leads to weaker non-linearities and more robustness to noise. Importantly, in both cases (shifted parameters from moments or shifted points from the object contour), we have obtained unique combinations of features for any object. However, the optimum choices for shifted points still remain an issue.

ACKNOWLEDGMENTS

This work was partly founded by the French government research program Investissements d'avenir through the RobotEx Equipment of Excellence (ANR-10-EQPX-44) and the LABEX IMOBS3 (ANR-7107-LABX-716701).

REFERENCES

[1] M. Bakthavatchalam, O. Tahri, and F. Chaumette, "Improving moments-based visual servoing with tunable visual features," in *IEEE Int. Conf. on Robotics and Automation, ICRA'14*, Hong Kong, China, June 2014, pp. 6186–6191.

[2] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The Confluence of Vision and Control*, D. Kriegman, G. . Hager, and A. Morse, Eds. LNCIS Series, No 237, Springer-Verlag, 1998, pp. 66–78.

[3] ——, "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.

[4] 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.

[5] C. Collewet and F. Chaumette, "Positioning a camera with respect to planar objects of unknown shape by coupling 2d visual servoing and 3d estimations," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 3, pp. 322–333, June 2002.

[6] 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, pp. 5438–5443.

[7] ——, "Photometric visual servoing," *IEEE Trans. on Robotics*, vol. 27, no. 4, pp. 828–834, August 2011.

[8] K. Deguchi, "A direct interpretation of dynamic images with camera and object motions for vision guided robot control," *International Journal of Computer Vision*, pp. 7–20, 2000.

[9] J. Flusser, T. Suk, and B. Zitov, *Moments and Moment Invariants in Pattern Recognition*. John Wiley and Sons, 2009.

[10] R. T. Fomena, O. Tahri, and F. Chaumette, "Distance-based and orientation-based visual servoing from three points," *Robotics, IEEE Transactions on*, vol. 27, no. 2, pp. 256–267, 2011.

[11] M.-K. Hu, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 179–187, 1962.

[12] V. Kallem, M. Dewan, J. P. Swensen, G. D. Hager, and N. J. Cowan, "Kernel-based visual servoing," in *IROS*, 2007, pp. 1975–1980.

[13] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments." *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 489–497, 1990.

[14] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 d visual servoing," *IEEE Trans. on Robotics and Automation*, vol. 15, no. 2, pp. 238–250, 1999.

[15] E. Malis, Y. Mezouar, and P. Rives, "Robustness of image-based visual servoing with a calibrated camera in the presence of uncertainties in the 3d structure," *IEEE Transactions on RObotics*, vol. 26, no. 1, pp. 112–120, February 2010.

[16] N. Mansard and F. Chaumette, "Task sequencing for high level sensor-based control," *IEEE Trans. on Robotics*, vol. 23, no. 1, pp. 60–72, February 2007.

[17] R. Mebarki, A. Krupa, and F. Chaumette, "2d ultrasound probe complete guidance by visual servoing using image moments," *IEEE Trans. on Robotics*, vol. 26, no. 2, pp. 296–306, April 2010.

[18] H. Michel and P. Rives, "Singularities in the determination of the situation of a robot effector from the perspective view of 3 points," Institut National de Recherche en Informatique et en Automatique (INRIA), 06-Valbonne (France), Internal report, 1993.

[19] S. Nayar, S. Nene, and H. Murase, "Subspace Methods for Robot Vision," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 750–758, Oct 1996.

[20] K. R.Mukundan, "Moment functions in image analysis: Theory and applications," in *World Scientific Publishing Co Pte Ltd., Singapore*, september 1998, p. 165.

[21] O. Tahri and F. Chaumette, "Complex objects pose estimation based on image moment invariants," in *IEEE Int. Conf. on Robotics and Automation, ICRA'05*, Barcelona, Spain, April 2005, pp. 438–443.

[22] ——, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. on Robotics*, vol. 21, no. 6, pp. 1116–1127, December 2005.

[23] M. R. Teague, "Image analysis via the general theory of moments," *Journal of the Optical Society of America*, vol. 70 Issue 8, pp. 920–930, 1980.

[24] C. Teh and R. T. Chin, "On image analysis by the method of moments," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496–513, 1988.

[25] W. Wilson, C. Hulls, and G. Bell, "Relative end-effector control using cartesian position-based visual servoing," *IEEE Trans. on Robotics and Automation*, vol. 12, no. 5, pp. 684–696, October 1996.

[26] A. Yeremou Tamtsia, O. Tahri, Y. Mezouar, H. Djalo, and E. Tonye, "New results in image moments-based visual servoing," in *IEEE International Conference on Robotics and Automation*, Karlsruhe, Germany, May 2013, pp. 5271–5276.

View publication stats