

An Improved Modelling Scheme for Photometric Moments with Inclusion of Spatial Weights for Visual Servoing with Partial Appearance/Disappearance

Manikandan Bakthavatchalam, François Chaumette, Omar Tahri

Abstract—Visual servoing based on photometric data is of great interest since it does not necessitate any image processing or visual tracking steps. A vital issue in such methods is the change in the image resulting from the appearance and disappearance of portions of the scene from the camera field-of-view during the visual servo. In this paper, we propose a spatial weighting scheme to counter this problem. A general model for photometric moments that allows for inclusion of spatial weights is presented. A custom weighting function is then chosen and specific properties of its derivative are exploited in order to develop the interaction matrix in analytical form. The resultant effects on the invariance properties are discussed. Experiments have been performed to validate the visual servoing with the proposed weighting scheme on a 4 dof robot arm.

I. INTRODUCTION

Visual Servoing(VS) encompasses a whole spectrum of methods wherein the objective is to control the motions of a robot with a stream of images acquired from a camera sensor [1]. In Image-based Visual Servoing (IBVS), the error to be regulated ($e = s - s^*$) is defined in the image space [2] [3]. The s are the visual features built from image measurements from the projections of geometric (points, straight lines) or volumetric (spheres, cylinders) entities in the scene. In [4] and [5], visual features derived from projective homography are used. This is estimated using a photo-geometric image registration algorithm, which by itself is a tracking problem. All these methods therefore require reliable visual tracking and/or matching processes.

Despite decades of research in VS, its potential for real-world applications is still not maximal. Among several factors that impede this expansion, dependence of VS on visual tracking procedures has been identified as a major factor [6] [7]. Methods like [8] used Principal Component Analysis (PCA) to transform image intensities to a lower dimensional representation. However, it was computationally prohibitive to expand them to several degrees of freedom [9]. Pure photometric visual servoing [10] directly utilised the luminance information in the image. It offers an excellent accuracy due to its usage of highly redundant features. It is also highly non-linear which reduces its convergence domain and produces non-intuitive robot trajectories.

Image moments are attractive for IBVS and shown to be better than geometric features because of their decoupling properties [11] [12]. But their applications remain limited because of the restriction that well-segmented regions of interest or a discrete point set should always be available in the image. [13] tried to remove this restriction with photometric

image moments defined from the image plane intensities. While no visual tracking or matching were required, all the decoupling properties of moments were retained. However, the modelling did not take into account the resultant effects induced by portions of the image which are not common between the current image and the next (including the image learnt from the desired robot pose). Since extraneous regions can disturb the stability and convergence of the VS, it is important to solve this vital issue. Interestingly, this issue arising from the uncommon image portions was pointed out earlier in [11].

Image moments [11] [13] use polynomial kernels and are a specific case of KBVS [14]. Kernels of finite support were used in [15] perhaps to deal with this issue. The non-availability of analytical models for the interaction matrix in KBVS (and in learning-based approaches like [8]) make conducting further analyses impossible.

Works like [16] and [17] deal with feature visibility and signal loss in IBVS, problems of somewhat similar nature. But image points are used as features and so not applicable to dense photometric methods. To the best of our knowledge, solutions to this problem do not currently exist.

The main idea in this paper is to introduce spatial weights in image moments so that the areas where the appearance and disappearance occur (at the periphery) have as less influence as possible when compared to the central parts of the image during the visual servoing.

Our paper is organized as follows: Sec II presents the modelling aspects to obtain the interaction matrix of weighted photometric moments and the selection of the spatial weighting function. Sec III discusses the properties for the most commonly used visual features. Finally, validation results are presented in Sec IV.

II. MODELLING OF WEIGHTED PHOTOMETRIC MOMENTS

We define a weighted photometric moment (WPM) of order $(p + q)$ as follows

$$m_{pq} = \iint_{\pi} x^p y^q w(\mathbf{x}) I(\mathbf{x}, t) dx dy \quad (1)$$

where $\mathbf{x} = (x, y)$ is a spatial point where the intensity $I(\mathbf{x}, t)$ is measured at time t , π denotes the image support and $w(\mathbf{x})$ is a spatial weighting function.

A. General case

To maintain generality, we first develop an analytical form of the interaction matrix for an abstract spatial weighting function $w(\mathbf{x})$ and then for a specific family of functions in Sec II-B. In the following derivation, we follow the same line of reasoning as in [13], only emphasizing the changes introduced by the weighting function, wherever appropriate. We recall that $\dot{m}_{pq} = \mathbf{L}_{m_{pq}} \mathbf{v}_c$ where $\mathbf{L}_{m_{pq}}$ is the interaction matrix of the moments defined in (1) and $\mathbf{v}_c = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)$ is the camera velocity screw. By taking the derivative of (1), we have

$$\dot{m}_{pq} = \iint_{\pi} x^p y^q w(\mathbf{x}) \dot{I}(\mathbf{x}, t) dx dy \quad (2)$$

where $\dot{I}(\mathbf{x}, t) = \mathbf{L}_I \mathbf{v}_c$, \mathbf{L}_I being the interaction matrix of the luminance feature [10]. Let us note that this result is obtained from the optic flow constraint equation (OFCE) with the brightness constancy hypothesis [18]. After following the same developments as in [13], we obtain

$$\begin{cases} L_{m_{pq}}^{v_x} = A m_{p+1,q}^{\nabla x} + B m_{p,q+1}^{\nabla x} + C m_{p,q}^{\nabla x} \\ L_{m_{pq}}^{v_y} = A m_{p+1,q}^{\nabla y} + B m_{p,q+1}^{\nabla y} + C m_{p,q}^{\nabla y} \\ L_{m_{pq}}^{v_z} = -A m_{p+2,q}^{\nabla x} - B m_{p+1,q+1}^{\nabla x} - C m_{p+1,q}^{\nabla x} \\ \quad - A m_{p+1,q+1}^{\nabla y} - B m_{p,q+2}^{\nabla y} - C m_{p,q+1}^{\nabla y} \\ L_{m_{pq}}^{\omega_x} = -m_{p+1,q+1}^{\nabla x} - m_{p,q}^{\nabla y} - m_{p,q+2}^{\nabla y} \\ L_{m_{pq}}^{\omega_y} = m_{p,q}^{\nabla x} + m_{p+2,q}^{\nabla x} + m_{p+1,q+1}^{\nabla y} \\ L_{m_{pq}}^{\omega_z} = -m_{p,q+1}^{\nabla x} + m_{p+1,q}^{\nabla y} \end{cases} \quad (3)$$

with

$$m_{pq}^{\nabla x} = \iint_{\pi} x^p y^q w(\mathbf{x}) \frac{\partial I}{\partial x} dx dy \quad (4a)$$

$$m_{pq}^{\nabla y} = \iint_{\pi} x^p y^q w(\mathbf{x}) \frac{\partial I}{\partial y} dx dy \quad (4b)$$

where $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ are the gradients in the image along x and y . A , B and C are the plane parameters to which the imaged scene belongs (only planar scenes have been considered in this paper for simplicity). These terms can be simplified using Green's theorem:

$$\iint_{\pi} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint_{\partial\pi} P dx + \oint_{\partial\pi} Q dy \quad (5)$$

If we let $Q = x^p y^q w(\mathbf{x}) I(\mathbf{x})$ and $P = 0$ in (5), we have

$$\begin{aligned} \iint_{\pi} \left[x^p y^q w(\mathbf{x}) \frac{\partial I}{\partial x} + x^p y^q \frac{\partial w}{\partial x} I(\mathbf{x}) \right. \\ \left. + p x^{p-1} y^q w(\mathbf{x}) I(\mathbf{x}) \right] dx dy = \oint_{\partial\pi} x^p y^q w(\mathbf{x}) I(\mathbf{x}) dy \end{aligned} \quad (6)$$

So, equation (4a) becomes

$$\begin{aligned} m_{pq}^{\nabla x} = - \iint_{\pi} \left(x^p y^q \frac{\partial w}{\partial x} I(\mathbf{x}) + p x^{p-1} y^q w(\mathbf{x}) I(\mathbf{x}) \right) dx dy \\ + \oint_{\partial\pi} x^p y^q w(x, y) I(x, y) dy \end{aligned} \quad (7)$$

Using the definition of the weighted moments in (1), we subsequently have

$$\begin{aligned} m_{pq}^{\nabla x} = -p m_{p-1,q} - \iint_{\pi} x^p y^q \frac{\partial w}{\partial x} I(\mathbf{x}) dx dy \\ + \oint_{\partial\pi} x^p y^q w(x, y) I(x, y) dy \end{aligned} \quad (8)$$

The last term has to be evaluated over the image border $\partial\pi$. In [13], this term was removed with the assumption that the image borders are of zero intensity. This hypothesis is invalid when there are dynamic image regions which enter and leave during the visual servo. Here, we define the weighting function so that $w(x, y) = 0 \forall (x, y) \in \partial\pi$ and so, the last term vanishes again but without the zero border assumption. Further, if we take $P = x^p y^q w(x, y) I(x, y)$ and $Q = 0$ in Green's theorem and make the same series of developments as before, $m_{pq}^{\nabla y}$ can be obtained. The terms in (4) related to the interaction matrix of the WPM become

$$m_{pq}^{\nabla x} = -p m_{p-1,q} - \iint_{\pi} x^p y^q \frac{\partial w}{\partial x} I(x, y) dx dy \quad (9a)$$

$$m_{pq}^{\nabla y} = -q m_{p,q-1} - \iint_{\pi} x^p y^q \frac{\partial w}{\partial y} I(x, y) dx dy \quad (9b)$$

Substituting (9) in Eq(3), we see that the interaction matrix is dependent not only on weighted moments m_{pq} , but also has additional terms composed of $\iint_{\pi} x^p y^q \frac{\partial w}{\partial \mathbf{x}} I(\mathbf{x})$. For this, obviously the first order spatial derivatives should exist for $w(\mathbf{x})$ ¹. Next, we select an appropriate weighting function such that a closed-form expression for the interaction matrix is obtained as functions of the WPM.

B. Weighting scheme

The weighting function should attribute maximal importance to the pixels in the area around the image centre, smoothly reducing it outwards towards the image edges. As pointed out earlier, the first order spatial derivatives of $w(\mathbf{x})$ should exist $\forall (x, y) \in \pi$. The standard logistic function $l(x) = \frac{1}{1+\exp^{-x}}$ was initially considered for the weighting scheme. With this choice however, it was not possible to obtain the interaction matrix in an analytical form as functions of the weighted photometric moments. That is why in this paper, we propose to use functions with

¹Note that if $w(x, y) = 1 \forall (x, y)$, the gradients of w would vanish. Both $m_{pq}^{\nabla x}$ and $m_{pq}^{\nabla y}$ would then reduce to a single term (which was the case in [13] as well). That is, they reduce exactly to the case of photometric moments with no weight

the general structure given in (10), where the exponential is raised to the negative power of a polynomial function.

$$\mathcal{F}(x) = \exp^{-p(x)} \text{ with} \quad (10)$$

$$p(x) = a_0 + a_1x + \frac{1}{2}a_2x^2 + \frac{1}{3}a_3x^3 + \dots + \frac{1}{n}a_nx^n$$

Functions of this structure possess the interesting property that their derivatives can be expressed in terms of the function itself. The derivative of $\mathcal{F}(x)$ is given by

$$\mathcal{F}'(x) = \exp^{-p(x)} p'(x) \text{ with} \quad (11)$$

$$p'(x) = a_1 + a_2x + a_3x^2 + \dots + a_nx^{n-1}$$

This in turn allows to obtain the interaction matrix as a function of the weighted moments (see Eq15). We propose the following weighting function with the form as in (10)

$$w(x, y) = K \exp^{-a(x^2+y^2)^2} \quad (12)$$

where K is the maximum value of the weighting function and a can be used to vary the areas which receive maximum and 0 weights respectively (See Fig.1). The polynomial function $(x^2 + y^2)^2$ to which the exponential is raised preserves the invariance property of the original image moments (i.e., moments without weights) to rotations around the optic axis. This can be confirmed from $L_{m_{pq}}^{\omega_z}$ in (15). We now continue

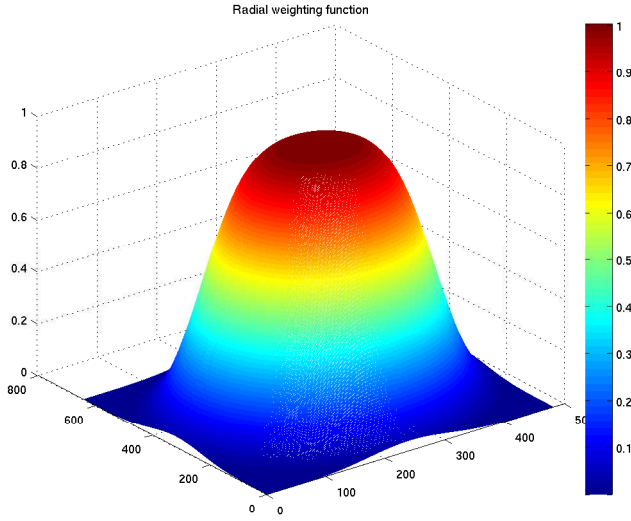


Fig. 1. Surface of the weighting function in (12) with $K = 1$, $a = 650$ with our developments with (12) as the weighting function. Its spatial derivatives are as follows:

$$\frac{\partial w}{\partial x} = -4ax(x^2 + y^2) w(\mathbf{x}) \quad (13a)$$

$$\frac{\partial w}{\partial y} = -4ay(x^2 + y^2) w(\mathbf{x}) \quad (13b)$$

Substituting (13) into (9), we obtain

$$m_{pq}^{\nabla x} = -p m_{p-1,q} + 4a(m_{p+3,q} + m_{p+1,q+2}) \quad (14a)$$

$$m_{pq}^{\nabla y} = -q m_{p,q-1} + 4a(m_{p,q+3} + m_{p+2,q+1}) \quad (14b)$$

By insertion of (14) into (3), the interaction matrix of photometric moments $\mathbf{L}_{m_{pq}}$ weighted with the radial function in (12) can be obtained:

$$\mathbf{L}_{m_{pq}} = [L_{m_{pq}}^{v_x} \quad L_{m_{pq}}^{v_y} \quad L_{m_{pq}}^{v_z} \quad L_{m_{pq}}^{\omega_x} \quad L_{m_{pq}}^{\omega_y} \quad L_{m_{pq}}^{\omega_z}] \quad (15)$$

with:

$$L_{m_{pq}}^{v_x} = A(-(p+1)m_{p,q} + 4a(m_{p+4,q} + m_{p+2,q+2}))$$

$$+ B(-pm_{p-1,q+1} + 4a(m_{p+3,q+1} + m_{p+1,q+3}))$$

$$+ C(-pm_{p-1,q} + 4a(m_{p+3,q} + m_{p+1,q+2}))$$

$$L_{m_{pq}}^{v_y} = A(-qm_{p+1,q-1} + 4a(m_{p+3,q+1} + m_{p+1,q+3}))$$

$$+ B(-(q+1)m_{p,q} + 4a(m_{p,q+4} + m_{p+2,q+2}))$$

$$+ C(-qm_{p,q-1} + 4a(m_{p,q+3} + m_{p+2,q+1}))$$

$$L_{m_{pq}}^{v_z} = A[(p+q+3)m_{p+1,q}$$

$$- 4a(m_{p+5,q} + 2m_{p+3,q+2} + m_{p+1,q+4})]$$

$$+ B[(p+q+3)m_{p,q+1}$$

$$- 4a(m_{p+4,q+1} + 2m_{p+2,q+3} + m_{p,q+5})]$$

$$+ C[(p+q+2)m_{p,q}$$

$$- 4a(m_{p+4,q} + 2m_{p+2,q+2} + m_{p,q+4})]$$

$$L_{m_{pq}}^{\omega_x} = (p+q+3)m_{p,q+1} + qm_{p,q-1}$$

$$- 4a(m_{p+4,q+1} + 2m_{p+2,q+3} + m_{p,q+3}$$

$$+ m_{p+2,q+1} + m_{p,q+5})$$

$$L_{m_{pq}}^{\omega_y} = -pm_{p-1,q} - (p+q+3)m_{p+1,q}$$

$$+ 4a(m_{p+3,q} + m_{p+1,q+2} + m_{p+5,q}$$

$$+ 2m_{p+3,q+2} + m_{p+1,q+4})$$

$$L_{m_{pq}}^{\omega_z} = pm_{p-1,q+1} - qm_{p+1,q-1}$$

This interaction matrix is expressed as a matrix sum

$$\mathbf{L}_{m_{pq}} = \mathbf{L}_1 + 4a\mathbf{L}_2 \quad (16)$$

where \mathbf{L}_1 has the same form as the non-weighted moments in [13] (but the moments are now weighted as in (1)). \mathbf{L}_2 is tied directly to the weighting function. To compute $\mathbf{L}_{m_{pq}}$, moments m_{tu} of order $t+u \leq p+q+5$ are required. So, the choice of the weighting function dictates the maximum moment order required for computing the interaction matrix. In the case of a uniform weighting function, the above interaction matrix reduces² exactly to the case in [13]. In $\mathbf{L}_{m_{pq}}$, it is to be noted that $L_{m_{pq}}^{\omega_z}$ has no new term contributed by the weighting function. Thus, the invariance of the weighted moments to optic axis rotations is retained. This outcome was of course desired from the structure of the weighting function in (12). Considering the other components in $\mathbf{L}_{m_{pq}}$, we observe additional terms induced by the weighting function. As a result, invariants to scale and 2D translations built from the classical moments are not invariant anymore when using the WPM. This loss of invariance is an issue for selecting visual features to control the full 6 dof of a system.

²This can be easily verified since $w(x, y) = 1$ would imply $a = 0$ and therefore $\mathbf{L}_{m_{pq}} = \mathbf{L}_1$ and only moments of order $(p+q+1)$ are sufficient to compute it.

III. ANALYSIS OF THE LOWER ORDER WEIGHTED MOMENTS

A. Weighted photometric area moment

With WPM, the zeroth order moment can be interpreted as the weighted photometric area. Its interaction matrix can be derived from (15) simply by taking $p = q = 0$ and by using $K' = 4a$ for brevity. We obtain

$$\begin{cases} L_{m_{00}}^{v_x} = -m_{0,0}A + K'[A\alpha_1 + B\alpha_2 + C\lambda_1] \\ L_{m_{00}}^{v_y} = -m_{0,0}B + K'[A\alpha_2 + B\alpha_3 + C\lambda_2] \\ L_{m_{00}}^{v_z} = m_{00}(3/Z_g - C) - K'[A\lambda_3 + B\lambda_4 + C(\alpha_1 + \alpha_3)] \\ L_{m_{00}}^{\omega_x} = 3m_{00}y_g - K'(\lambda_2 + \lambda_5) \\ L_{m_{00}}^{\omega_y} = -3m_{00}x_g + K'(\lambda_1 + \lambda_3) \\ L_{m_{00}}^{\omega_z} = 0 \end{cases} \quad (17)$$

$$\text{with } \begin{cases} \alpha_1 = m_{4,0} + m_{2,2} & \lambda_1 = m_{3,0} + m_{1,2} \\ \alpha_2 = m_{3,1} + m_{1,3} & \alpha_3 = m_{0,4} + m_{2,2} \\ \lambda_2 = m_{0,3} + m_{2,1} \\ \lambda_3 = m_{1,4} + 2m_{3,2} + m_{5,0} \\ \lambda_4 = m_{0,5} + 2m_{2,3} + m_{4,1} \end{cases}$$

For the original non-weighted moments, the area m_{00} is a very interesting visual feature mainly related to the translation motions in z . It is invariant to (i) rotations around the optic axis and (ii) to planar translations when the camera and target planes are parallel ($A = B = 0$). With WPM, from the terms in (17), we indeed observe that the invariance to optic axis rotations is preserved thanks to our choice of the weighting function. The influence on planar translations can be studied by taking $A = B = 0$ in the corresponding terms of the interaction matrix. Then, we will have $L_{m_{00}}^{v_x} = K' C(m_{3,0} + m_{1,2})$ and $L_{m_{00}}^{v_y} = K' C(m_{0,3} + m_{2,1})$. Clearly, third order moments appear, which means that the invariance property is generally lost. From the experimental results however, this loss did not have any strong effect on the convergence of the servo. On the contrary, it is to be noted that if these terms are neglected, it will introduce a gross error into the modelling and the visual servoing might not converge at all.

B. Weighted Centre of gravity

The ratio between the first and zeroth order moments gives the centre of gravity coordinates in the image $x_g = m_{10}/m_{00}$ and $y_g = m_{01}/m_{00}$. The interaction matrix of x_g can be obtained analytically using $\mathbf{L}_{x_g} = (\mathbf{L}_{m_{1,0}} - x_g \mathbf{L}_{m_{0,0}}) / m_{0,0}$.

$$\begin{cases} L_{x_g}^{v_x} = -1/Z_g \\ \quad + \frac{K'}{m_{0,0}} [A\epsilon_1 + B\epsilon_2 + C(\alpha_1 - x_g\lambda_1)] \\ L_{x_g}^{v_y} = \frac{K'}{m_{0,0}} [A\epsilon_2 + B\epsilon_1 + C(\alpha_2 - x_g\lambda_2)] \\ L_{x_g}^{v_z} = x_g/Z_g + 4(An_{20} + Bn_{11}) \\ \quad - \frac{K'}{m_{0,0}} [A(\phi_1 - x_g\lambda_3) + B(\phi_2 - x_g\lambda_4) \\ \quad + C(\lambda_5 - x_g(\alpha_1 + \alpha_3))] \end{cases}$$

$$\begin{cases} L_{x_g}^{\omega_x} = (x_g y_g + 4n_{1,1}) \\ \quad + \frac{K'}{m_{0,0}} [(\phi_2 + \alpha_2) - x_g(\lambda_2 + \lambda_4)] \\ L_{x_g}^{\omega_y} = -(1 + x_g^2 + 4n_{20}) \\ \quad + \frac{K'}{m_{0,0}} [(\phi_1 + \alpha_1) - x_g(\lambda_1 + \lambda_5)] \\ L_{x_g}^{\omega_z} = y_g \end{cases} \quad (18)$$

$$\text{with } \begin{cases} \epsilon_i = \delta_i - x_g \alpha_i \\ \delta_1 = m_{5,0} + m_{3,2} \\ \delta_2 = m_{4,1} + m_{2,3} \\ \phi_1 = m_{2,4} + 2m_{4,2} + m_{6,0} \\ \phi_2 = m_{5,1} + 2m_{3,3} + m_{1,5} \\ \lambda_5 = m_{5,0} + 2m_{3,2} + m_{1,4} \end{cases}$$

and $n_{i,j} = \mu_{i,j}/m_{0,0}$ where $\mu_{i,j} = m_{i,j} - m_{0,0} x_g^i y_g^j$. From the second term $L_{x_g}^{v_y}$ in the interaction matrix, we can notice that translational velocities in y will modify the value of x_g even when the camera and target planes are parallel. Hence x_g is not invariant to motions in y . This loss of invariance to translations was unfortunately expected from the spatial weighting. We can notice the weighting function did not alter the behaviour of the centre of gravity w.r.t optic axis rotations. In a similar manner, the interaction matrix of the other coordinate y_g can also be obtained. But since it does not present any particularly interesting characteristics, it is omitted here.

C. Photometric orientation feature

The orientation α is an interesting feature that can be calculated from the centred second order moments.

$$\alpha = \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \quad (19)$$

where μ_{pq} are the weighted centred moments. The exact analytical form of \mathbf{L}_α is not given here due to space constraints. It is straight-forward to compute however as shown in [11] [12]. As with the case of the centre of gravity and area, the behaviour with respect to optic axis rotations is not modified. That is, the last term of the interaction matrix is equal to -1 , indicating that it is linearly related to rotational motion around the optic axis. This linear relation shows α is well-suited to control motion along this dof.

IV. VALIDATION RESULTS

We present in this section both simulation and experimental results that validate the proposed approach. Software for this research was built on top of the ViSP [19] library. A calibrated camera is mounted on the end-effector in an eye-in-hand configuration and a motionless, planar target is considered. For simulations as well as the experiments with the robot, a SCARA-type (3T+1R) actuation with only the 3D translations and rotation around the optical axis was considered. This is the same subset of dof considered in KBVS [14]. The partial velocity screw $\mathbf{v}_c = (v_x, v_y, v_z, \omega_z)$ is computed using the classical IBVS control law [2]:

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L}}_s^{-1} (\mathbf{s} - \mathbf{s}^*) \quad (20)$$

with $\mathbf{s} = (x_n, y_n, a_n, \alpha)$, where $(x_n = x_g \cdot a_n, y_n = y_g \cdot a_n)$ is the normalized weighted centre of gravity, $a_n = Z^* \sqrt{\frac{m_{00}^*}{m_{00}}}$ is the normalized area moment, (x_g, y_g) the centre of gravity and α is the orientation as defined in (19). These features are used in several works including [12] and [13] (but from moments without weighting). x_n, y_n, a_n are normalized versions of the features described in Sec III and λ is the gain for the control law. The weighting function (12) with $K = 1$ and $a = 1000$ was satisfactory (no precise tuning was required). Further, the spatial weights are computed only once in a single offline step and kept in a convenient matrix form.

A. Simulation Results

For the simulations, the desired pose is chosen $1.0m$ vertically above the target such that the camera and target planes are parallel. A gain of $\lambda = 2.0$ was used for faster convergence (not affected by lower values). The initial pose is chosen so as to obtain the image shown in Fig.2b. From Figs.2a and 2b, we can observe that there is no distinguishable background and the ship structure remains obstructed in the initial image. The displacements required for convergence are a translation of ${}^c \mathbf{t}_c = [2.41cm, 6.65cm, 10cm]$ and a rotation of $\mathbf{R}_z(25^\circ)$ around the optic axis.

1) Using $\hat{\mathbf{L}}_s = \mathbf{L}_{s^*}$: The control law in (20) was used with the interaction matrix computed once at the desired pose $\hat{\mathbf{L}}_s(\mathbf{s} = \mathbf{s}^*)$. Hence, the control action is linear in $(\mathbf{s} - \mathbf{s}^*)$ since \mathbf{L}_{s^*} is constant and is simpler than in [10]. The interaction matrix computed at the desired pose has a condition number of 2365.2 and is given by

$$\mathbf{L}_{s^*} = \begin{bmatrix} 0.1417 & -0.1134 & -0.0042 & -0.0069 \\ -0.1145 & -0.0015 & -0.0153 & 0.0032 \\ 0.1030 & 0.1024 & 0.0716 & 0 \\ -0.3779 & -8.4665 & -0.0879 & -1.0000 \end{bmatrix} \quad (21)$$

Discussion : From Eq.(21), several inferences can be made. The feature a_n is invariant to optic axis rotations and α has retained its property with respect to optic axis rotations (-1 in last column). But x_n (resp. y_n) is no more invariant to translations in y (resp. x). This invariance loss was discussed in Sec.II-B. Also, a_n and α in (21) are more reactive to errors in degrees of freedom different from the ones they are intended to control (translations along and around the optic axis resp). This consequence is due to the weighting function, which implies new features would have to be determined for a more optimal design. At this juncture, we would like to stress the importance of explicit analytical modelling which facilitates such analyses. Despite the loss in invariance, the visual servoing converged satisfactorily, especially with respect to the translation motions. The rotation error increases briefly before the final convergence. This is due to the extraneous regions appearing in the image during the initial stages of the servo. The 3D camera trajectory is not close to a straight line (See Fig.2f) but an overall satisfactory behaviour was obtained. The camera traversed a distance of $14.11cm$ in this experiment.

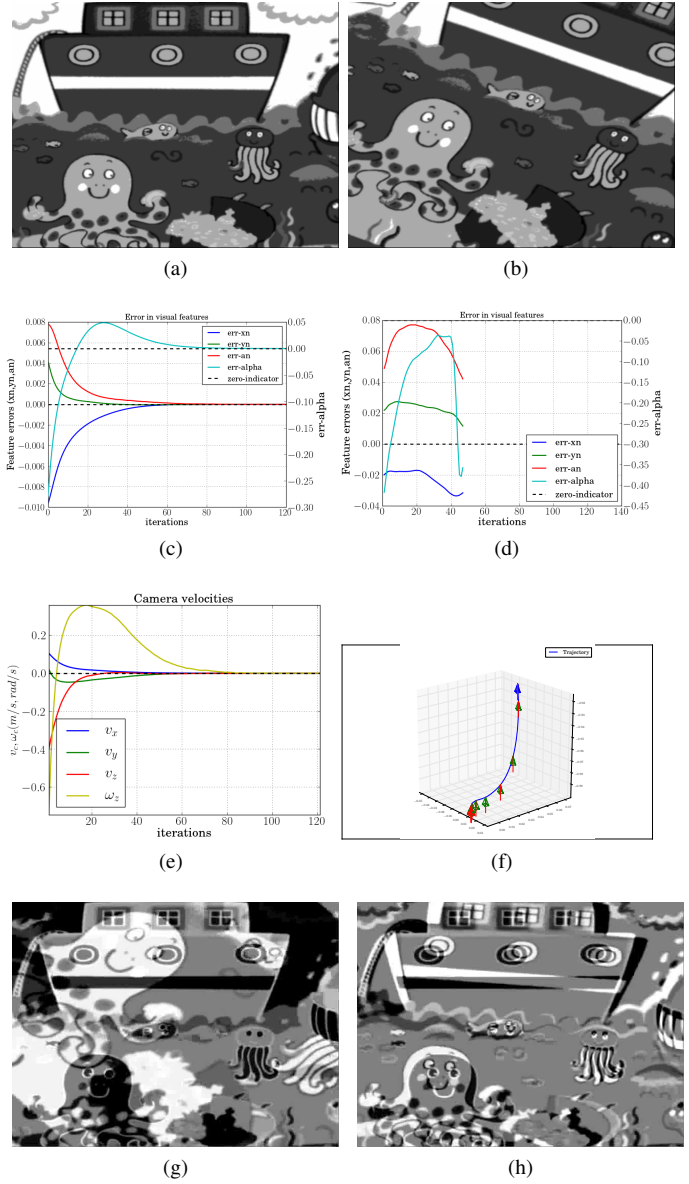
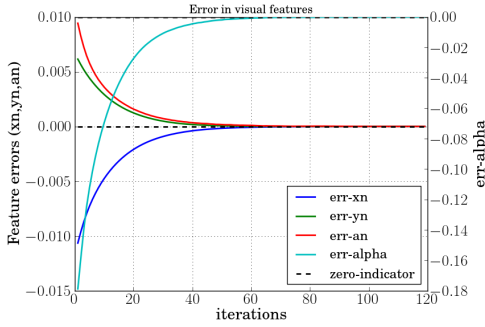


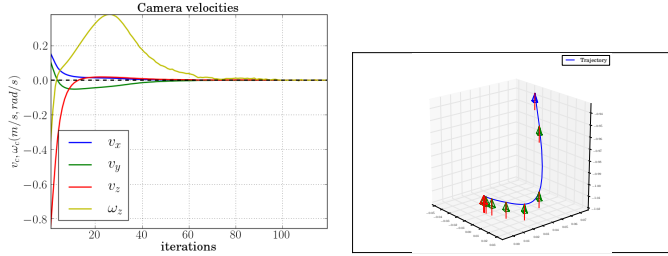
Fig. 2. Servo Results for Sec.IV-A.1 (a) Desired image, (b) Initial image, Visual feature errors for (c) Weighted Photometric Moments and for (d) non-weighted moments [13], both shown in Multi-scale y-axis (x_n, y_n, a_n on left and α on right) (e) Camera velocities and (e) Camera 3-D trajectory for Weighted Photometric Moments, Difference image after 39 iterations (f) using non-weighted moments and (g) using Weighted Photometric Moments

Comparison to non-weighted moments : With the same desired and initial poses, the visual servoing with non-weighted photometric moments proposed in [13] was tested. As seen from Fig.2d, the decrease in the errors for the non-weighted moments is not satisfactory at all and the servoing diverged. After about 40 iterations, the algorithm using the proposed WPM is driven near to convergence whereas for the non-weighted moments in [13], there is a large error as seen from the difference image in Fig.2g. The experiment was stopped when the error norm increased rapidly, otherwise leading to a jitter. It is evident that the control law in this case cannot handle the extraneous regions.

2) Using $\widehat{\mathbf{L}}_s = \mathbf{L}_s$: The same initial and desired poses as in Sec.IV-A.1 are considered. For the interaction matrix, it is computed at the current pose $\widehat{\mathbf{L}}_s(s)$ for use in the control law (20). As seen from the results in Fig.3a, a perfect



(a) Perfect exponential decrease of the visual feature errors (Multiscale y-axis: x_n, y_n, a_n on left and α on right)



(b) Camera velocities

(c) Camera spatial trajectory

Fig. 3. Visual Servoing Results with the proposed Weighted Photometric Moments using $\widehat{\mathbf{L}}_s = \mathbf{L}_s$

exponential convergence of the errors is achieved in this case. This concretely validates the correctness of all the modelling steps proposed in this paper. The camera traversed a distance of about $18.32cm$, greater than the distance of $14.11cm$ traversed in the earlier experiment in Sec.IV-A.1. The departure from a straight line behaviour is more in this case (see Fig.3c) than the previous one (see Fig.2f). This exponential error decrease and more deviation of the camera trajectory were both expected. This is because usage of the current interaction matrix ensures a nice exponential regulation of the errors but does not ensure the same for the Cartesian trajectory [11].

B. Experimental Results

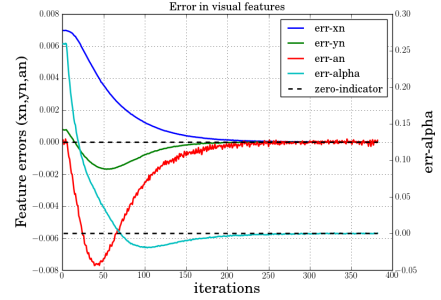
For the experiments, the Viper850 platform was configured for a SCARA-type actuation. The desired camera pose is chosen at approximately $40cm$ in front of the target. The displacement required between the initial and desired poses is ${}^c\mathbf{t}_c = [-2.32cm, 0.97cm, -8.3cm]$ in translation and $\mathbf{R}_z(-11^\circ)$ rotation around the optic axis. Like in the simulations, these poses are chosen such that there is no distinction between what is the foreground and what constitutes the background in the initial and desired images. As for the control law, a gain of $\lambda = 0.3$ and the interaction matrix \mathbf{L}_{s*} at the desired configuration is used.

1) *Experiment I*: The image learnt from the desired pose is shown in Figure 4a. The image acquired from the initial

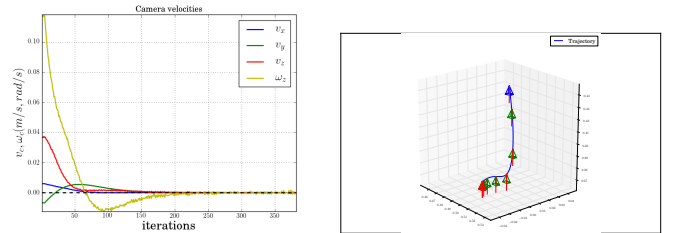
pose is shown in Figure 4b. Clearly, the actress who is present in the left portion of the initial image is not present at all in the learnt image. This is a voluntary choice as explained previously. The interaction matrix at the desired position had a condition number of 818.39 with same invariance properties discussed in IV-A. The extraneous image regions have induced large variations in the error in a_n during the initial iterations as seen from Fig.4c. Nevertheless, the system is driven to convergence by the generated control velocities shown in Fig.4d. The final error in the achieved robot pose is $[-0.29mm, -2.21mm, -0.96mm]$ in translation and 0.1° around the optic axis respectively.



(a) Camera view at desired robot pose (b) Camera view at initial robot pose



(c) Errors in visual features (y-axis Multi-scale) x_n, y_n, a_n on left and α on right



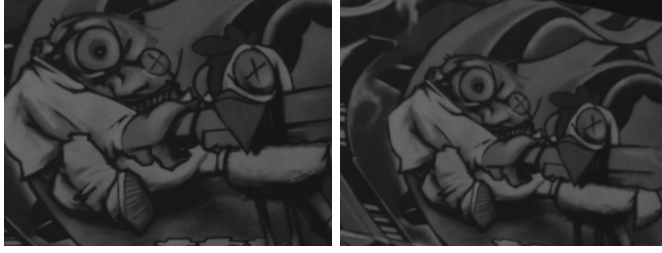
(d) Camera velocities

(e) Camera 3-D trajectory

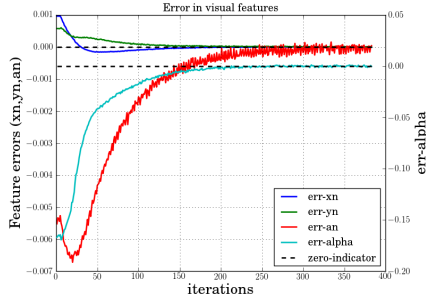
Fig. 4. Experiment I : Visual Servoing on the Viper850 robot with weighted photometric moments

2) *Experiment II*: In this experiment, the same initial and desired poses as Experiment IV-B.1 is chosen, but with a different photometric target. The images acquired at the desired and initial poses are shown in Figures 5a and 5b respectively.

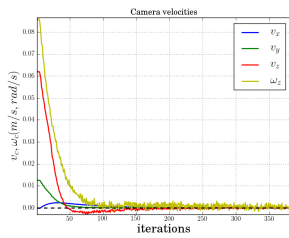
The form of the interaction matrix is similar to that in (21). Its condition number is 4435.2, higher than in the previous experiment. Indeed, the exact values in the interaction matrix depends on the spatial distribution of the intensities in the image. Further, from Figure 5c, we find that the error in a_n is more noisy than in the previous experiment. This is due to the fact that the noise effect depends on the area



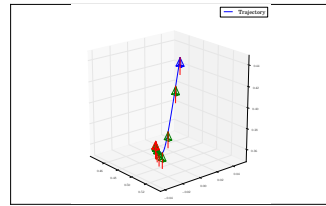
(a) Camera view at desired robot pose (b) Camera view at initial robot pose



(c) Errors in visual features (y-axis Multi-scale) x_n, y_n, a_n on left and α on right



(d) Camera velocities



(e) Camera 3-D trajectory

Fig. 5. Experiment II : Visual Servoing on the Viper850 robot with weighted photometric moments

under the influence of the weighting function. This can also be attributed in part to the uncontrolled lighting conditions which can contribute to noise in the pixels. The accuracy in this experiment is $[0.84mm, -1.72mm, -0.95mm]$ in translations which is acceptable and an error of only 0.07° around the optic axis.

The proposed scheme was thus validated by several successful experiments despite the loss in invariance. We recall that our scheme does not require any image matching or visual tracking processes but only the computation of a set of weighted moments on the image.

The hypothesis about the useful information being present in the central image parts might not always hold. If the shared area between the current and learnt image is very low, the method might fail (and this is true for all VS methods in general). This approach is also affected by illumination variations, which is indeed a limitation.

V. CONCLUSIONS AND FUTURE WORK

In this paper, an enhanced modelling scheme for photometric moments that accommodates spatial weights was presented. This improved model took into account scenario of extraneous image regions. A carefully selected weighting

function was introduced into the photometric moments formulation to obtain the interaction matrix in analytical form. Simulations and experiments with a real robot configured for SCARA-type actuation validate the proposed method. The weighted photometric moments preserved their behaviour with respect to rotations around the optic axis while behaviour with respect to other degrees of freedom were not retained. This affects the invariance properties of moment invariants, on which the features that control rotations around the image axes [12] are based. Establishment of this lost invariance is therefore an important problem and requires a considerable advance in theory, which is left to future work.

REFERENCES

- [1] J. Hill and W. Park, "Real time control of a robot with a mobile camera," in *ISIR*, March 1979.
- [2] F. Chaumette and S. Hutchinson, "Visual servo control, part i: Basic approaches," *IEEE RAM*, vol. 13, no. 4, pp. 82–90, Dec. 2006.
- [3] M. Iwatsuki and N. Okiyama, "A new formulation of visual servoing based on cylindrical coordinate system," *IEEE Trans. on Robotics*, vol. 21, no. 2, pp. 266–273, 2005.
- [4] G. Silveira and E. Malis, "Direct visual servoing: Vision-based estimation and control using only nonmetric information," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 974–980, 2012.
- [5] S. Benhimane and E. Malis, "Homography-based 2d visual tracking and servoing," *The International Journal of Robotics Research*, *IJRR'07*, vol. 26, no. 7, pp. 661–676, 2007.
- [6] M. Gridseth, C. Quintero, R. Fomena, O. Ramirez, and M. Jagersand, "Bringing visual servoing into real world applications," in *Human Robot Collaboration Workshop, RSS'13*, Germany, June 2013.
- [7] 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.
- [8] S. Nayar, S. Nene, and H. Murase, "Subspace methods for robot vision," *Robotics and Automation, IEEE Transactions on*, vol. 12, no. 5, pp. 750–758, Oct 1996.
- [9] K. Deguchi, "A direct interpretation of dynamic images with camera and object motions for vision guided robot control," *International Journal of Computer Vision*, vol. 37, no. 1, pp. 7–20, 2000.
- [10] C. Collewet and E. Marchand, "Photometric visual servoing," *IEEE Trans. on Robotics*, vol. 27, no. 4, pp. 828–834, Aug. 2011.
- [11] 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, Aug. 2004.
- [12] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. on Robotics*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.
- [13] M. Bakhavatchalam, F. Chaumette, and E. Marchand, "Photometric moments: New promising candidates for visual servoing," in *IEEE Int. Conf. on Robotics and Automation, ICRA'13*, Karlsruhe, Germany.
- [14] V. Kallem, M. Dewan, J. Swensen, G. Hager, and N. Cowan, "Kernel-based visual servoing," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'07*, San Deigo, California.
- [15] J. Swensen, V. Kallem, and N. Cowan, "Empirical characterization of convergence properties for kernel-based visual servoing," in *Visual Servoing via Advanced Numerical Methods*, ser. LNCIS, Springer London, 2010, vol. 401, pp. 23–38.
- [16] N. García-Aracil, E. Malis, R. Aracil-Santonja, and C. Pérez-Vidal, "Continuous visual servoing despite the changes of visibility in image features," *Robotics, IEEE Trans. on*, vol. 21, no. 6, pp. 1214–1220, 2005.
- [17] P. A. Durand, S. Durola, V. Cadenat, and M. Courdesses, "Management of visual signal loss during image based visual servoing," in *Control Conference (ECC), 2013 European*, July 2013, pp. 2305–2310.
- [18] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *ARTIFICIAL INTELLIGENCE*, vol. 17, pp. 185–203, 1981.
- [19] E. Marchand, F. Spindler, and F. Chaumette, "ViSP for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robot. Autom. Mag.*, vol. 12, no. 4, pp. 40–52, Dec. 2005.