ISSN 2176-5480

22nd International Congress of Mechanical Engineering (COBEM 2013)
November 3-7, 2013, Ribeirão Preto, SP, Brazil

COMPARATIVE PERFORMANCE ANALYSIS OF IMAGE-BASED AND POSITION –BASED VISUAL SERVOING IN A 6 DOF MANIPULATOR

Gloria Liliana Lopez Muñoz
José Mauricio Santos Torres da Motta
University of Brasília, Department of Mechanical Engineering, Brasília, D.F., 70.910-900, Brazil.
glorialiliana@unb.br, jmmotta@unb.br

Abstract. *The development of control systems based on visual information has increased in recent years, especially the control of robotic manipulators in locating and grasping tasks. There are two main approaches of Visual Servoing Control: Image-Based (IBVS) and Position-Based (PBVS). This article presents the implementation of these two approaches in order to make a comparative analysis, mostly based on stability, flexibility and robustness issues on the systems performance. In this work a monocular camera was placed on the 6 DOF manipulator end-effector in an eye-in-hand configuration, and target objects was supposed to be plane. The results obtained from the system for each type of control are achieved by implementing both control strategies in a real environment using a IRB-140 ABB Robot to comparatively quantify the system performance in terms of processing speed, tracking and position accuracy. Furthermore, the results can bring about optimized conditions in which each of both control strategies could be the best choice for a sort of applications.*

Keywords: *visual servoing, 6dof manipulator, eye-in-hand configuration.*

1. INTRODUCTION

Visual Servoing (VS) or visual servo control is a currently well known robot control technique whose aim is to provide the robot with an active vision system. This kind of robot is capable to achieve autonomy and robustness against environment variations that allows its use not only to highly structured environments but to other situations like dynamic settings, outdoors and unknown environments.

As shown in Fig. 1 VS is essentially a close loop motion control for robot systems to achieve a target pose (relative to an external frame), the command is based on the visual information (Hutchinson, 1996) acquired at each control cycle, hence, the technique relies on the relevant data extracted from the scene images, either measured or inferred, in real time (camera acquisition time). The nature of this data and its posterior use are the key to define the visual control type, as will be seen bellow, and the image processing stage is designed and adapted for each specific task always looking for efficiency and reliability. Once the data are obtained from the image, is the control loop that actually finds the desirable robot state and depends on it the effectiveness of the entire system. Because of all this is clearly seen that each stage plays an important role in the visual servoing system, even for a simple robot, and it is necessary to test different techniques to validate them seeking the best performance and the detection of their weakness in common tasks.

In order to prove the efficiency of visual servoing techniques, much work has be done until now, the first reports are from the early 70's when for the first time was integrated to the robot control loop the visual information from a camera (Shirai and Inoue, 1973). In 1980 Sanderson presented the first taxonomy of classification of visual servoing based on the kind of the control signal; many other classifications exists, based either on the configuration or the nature of the components, refer to them in Corke (1993) or Kragic and Christensen (2002), between others. Hutchinson, et al. (1996) and Chaumette and Hutchinson (2006, 2007) give a comprehensive review of visual servoing fundamentals, showing the image feature based control as an effective technique in manipulator robot control and describing their two major classes a) the image based visual servoing IBVS and b) the position based visual servoing PBVS, both of them described in detail further. Sanderson and Weiss (1987) also gave a great advance in this technique by introducing the interaction matrix to finding the direct relation between image features and camera move. Regarding to stability there is a general work given in Chaumette (1998) but rather specific tests must be performed to gain the needed knowledge in the choosing process for defined tasks.

In this article are presented the two most popular visual servoing approaches IBVS and PBVS both of them under a unique image processing technique that shows well suited results for the specific purpose. The tests were carry out in an IRB-140, industrial robot ABB with 6-DOF with an eye-in-hand configuration, meaning a single camera rigidly attached to the robot end-effector. These two visual servoing methodologies were probe under similar environment parameters and even showing a well response in almost all situations, they exhibited particular characteristics that allow to delimit their scope; this work intends to validate the two techniques in the Robotics and Computer Vision Lab in the Brasilia University as the first real application in the visual servoing area and provide a knowledge base for future researches.

2. BACKGROUND OF VISUAL SERVOING

As any other closed loop control, VS seeks to minimize an error function $\boldsymbol{e}$ between a reference $\boldsymbol{s}^*$and a feedback signal $\boldsymbol{s}$ (Eq. (1)), based on visual information. The data provide for the visual system defines, in turn, the specific approach between IBVS or PBVS. The former one uses directly the image features in order to calculate the error function and the control is performed in the image coordinates (Fig. 1(a)). The last one, uses the relative pose between the camera at the target and at the current position, inferred from its images, and the control is performed in the Cartesian space (Fig. 1(b)) (Corke, 2011). Perhaps, the error function in a visual servoing system is defined in a general way by

$$\boldsymbol{e} = (\boldsymbol{s}(\boldsymbol{r}, t) - \boldsymbol{s}^*) \quad (1)$$

here, $\boldsymbol{r}$ is the relative robot pose in the world frame at the time t, and $\boldsymbol{s}$ and $\boldsymbol{s}^*$ are visual characteristic vectors at the current and target position respectively. With the aim to minimize this error it could be used a proportional control law (Chaumette, 1990, Comport et al, 2003) in the form:

$$\boldsymbol{e} = -\lambda \boldsymbol{e} \quad (2)$$

where λ is a positive scalar. If the visual features vector is well chosen, there is only one camera final position that allows to achieve the error minimization. Then the control law can be seen as a relationship between the image characteristics and the camera motion (as result of robot motion):

$$\boldsymbol{v_c} = -\lambda \widehat{\boldsymbol{L}^+} \boldsymbol{e} \quad (3)$$

In Eq. (3) $\boldsymbol{v_c}$ is the camera velocity and $\widehat{\boldsymbol{L}^+}$ is the estimated pseudo-inverse of the *interaction matrix* $\boldsymbol{L}$ which relates the change in the image features with the camera movement (Corke, 1993). The input to the control law is thus the image feature, expressed by $\widehat{\boldsymbol{L}^+}$ and the output is the velocity command sent to the control of robot axis.

2.1 IBVS and PBVS

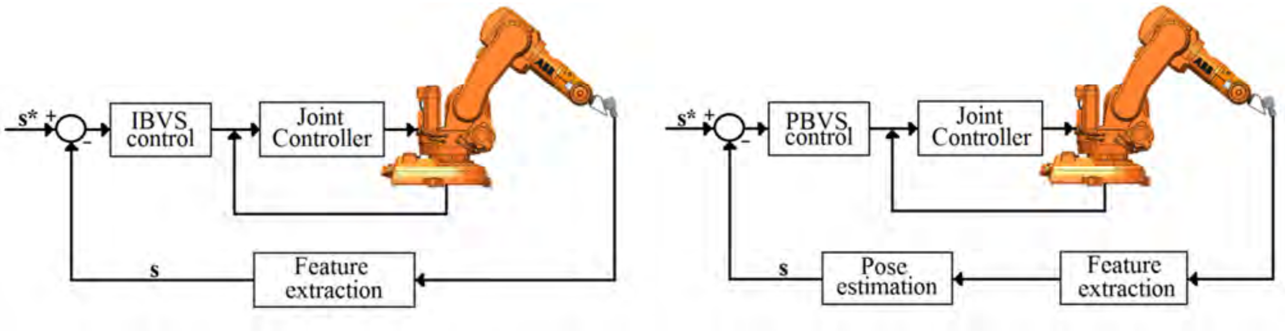

Figure 1. Visual Servoing scheme (a) Image Based Visual Servoing (b) Position Based Visual Servoing, (Corke, 2011)

The IBVS, Also known as 2D visual servoing because the bi-dimensionality of the image coordinates, uses directly the image features in the error function as mentioned before (Fig.1a), here, were used images of real planar objects as optical targets (Fig.3), and the features are the coordinates of the extracted corners. In order to implement the control of Eq. (3), the Eq. (1) is derived in time as:

$$\dot{\boldsymbol{e}} = \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{r}} \frac{\partial \boldsymbol{r}}{\partial t} \quad (4)$$

According to Corke (1993), this non-linear relationship can be linearized including the *Image Interaction Matrix* or Image Jacobian, $\boldsymbol{L_S}$, as

$$\dot{\boldsymbol{e}} = \boldsymbol{L_S} \dot{\boldsymbol{r}} \quad (5)$$

with: $$\boldsymbol{L}_s = \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{r}} = \begin{bmatrix} \frac{\partial s_1}{\partial r_1} & \cdots & \frac{\partial s_1}{\partial r_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_k}{\partial r_1} & \cdots & \frac{\partial s_k}{\partial r_m} \end{bmatrix} \quad (6)$$

In an eye-in-hand configuration the control variable can be chosen as the velocity of the robot end-effector obtaining the control law in Eq. (3), where $\dot{\boldsymbol{r}} = \boldsymbol{v_c}$.

The determination of $\widehat{\boldsymbol{L}_s}^{+}$, the approximate value of the pseudo-inverse of the *Image Interaction Matrix* $\boldsymbol{L}_s$ (Chaumette, et al., 2006), depends on the type of image features considered. As previously stated, the image features in this work are points, so the *Image Interaction Matrix* is given by:

$$\boldsymbol{L}_s = \begin{bmatrix} 1/Z & 0 & x/Z & -xy & (1+x^2) & -y \\ 0 & 1/Z & y/Z & -(1+y^2) & xy & x \end{bmatrix} \quad (7)$$

where, $(\boldsymbol{x}, \boldsymbol{y})$ are the coordinates of the image point feature and $\boldsymbol{Z}$ is the depth component of the point feature in the camera coordinate system to be estimated from its actual value in the scene. A detailed description of the Image Interaction Matrix can be found in Hutchinson et al. (2006).

For a 6DOF robot control it is necessary at least three points to define the matrix $\boldsymbol{L_s}$. To avoid singularities and minimal local here were used six points for this computation.

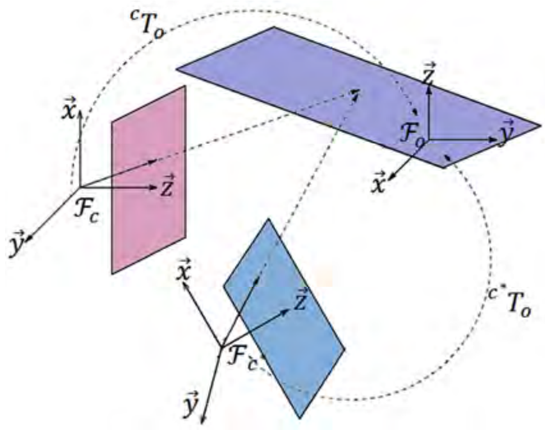

Figure 2. Camera movement scheme, (Malis, 1998)

In the PBVS or 3D, the system retrieves the relative pose between the actual camera location and the desired one, the control acts directly in the camera pose (Fig, 1b), and to define the error function it is necessary to take account the frames illustrated in Fig.2

It is important to notice that the camera pose can be found in relation to a fixed target object by using only a single image, once one has a prior knowledge of the object geometry and the camera intrinsic parameters (Hartley and Zisserman, 2003).

To define the control variable, the geometric relationship between the system components is established as shown in Fig. 2. Considering $\mathcal{F}_o$ the object frame, $\mathcal{F}_c$ the camera frame at the current position and $\mathcal{F}_{c^*}$ the camera frame in the desired position, the error can be calculated as a function of these relative positions. The coordinate transform between them is given by the homogeneous transformation

$${}^{c^*}T_c = \left[{}^{c^*}\boldsymbol{R}_c \middle| {}^{c^*}\boldsymbol{t}_c\right] \quad (8)$$

In the desired position this transformation will be null or zero, so this coordinate transform becomes the cost function to minimize, and accordingly to Ma, et al.(2001) the Eq. (8) can be rewrite as

$$({}^{c^*}\boldsymbol{t}_c, \mathbf{u}\theta) \quad (9)$$

,where θ is the rotation angle around a fixed axis $\mathbf{u}$ equivalent to ${}^{c^*}\boldsymbol{R}_c$, and using the Eq. (1) the error can be expressed as

$$\boldsymbol{e} = (\boldsymbol{s} - \boldsymbol{s}^{*}) = ({}^{c^*}\boldsymbol{t}_c, \mathbf{u}\theta) \tag{10}$$

Following the same strategy as in the IBVS control to formulate the control law of Eq. (2), the error derivative is

$$\dot{\boldsymbol{e}} = \left(\frac{\partial {}^{c^*}\boldsymbol{t}_c}{\partial \mathrm{r}} + \frac{\partial \theta}{\partial \mathrm{r}}\right)\frac{\partial \mathrm{r}}{\partial \mathrm{t}} = \boldsymbol{L}_{sr}\dot{\boldsymbol{r}} \tag{11}$$

, where

$$\boldsymbol{L}_s = \begin{bmatrix} {}^{c^*}\boldsymbol{R}_c & 0 \\ 0 & \boldsymbol{L}_{\theta \boldsymbol{u}} \end{bmatrix} \quad , \qquad \boldsymbol{L}_{\theta \boldsymbol{u}}{}^{-1}(\theta \boldsymbol{u}) = \theta \boldsymbol{u} \tag{12}$$

hence, the Interaction matrix is defined as a decoupled function of the rotation and translation of the camera and can be deduced by the relationship between the two poses involved (Chaumette e Hutchinson, 2006). From the Eq. (10), Eq.(11) and Eq. (12), the control law becomes:

$$\dot{\boldsymbol{r}} = -\lambda \boldsymbol{L}_{sr}{}^{-1}({}^{c^*}\boldsymbol{t}_c, \theta \boldsymbol{u}) \tag{13}$$

$$\dot{\boldsymbol{r}} = \begin{bmatrix} -\lambda\ \boldsymbol{R}^T\ {}^{c^*}\boldsymbol{t}_c \\ -\lambda\ \theta \boldsymbol{u} \end{bmatrix} \tag{14}$$

This expression defines the control variable $\dot{\boldsymbol{r}}$ as the camera velocity $\boldsymbol{v_c}$ that, in this particular case, is the velocity of the manipulator end-effector.

3. IMAGE PROCESSING

As stated before, a critical stage of the VS is the image analysis process, where are defined the relevant points in the images to ensure the correct movement of the robot. For the two approaches studied here, there are two different situations; in IBVS are compared two image feature vectors of a target in two location respect to the camera, in PBVS the compare is directly between the poses of the target and camera in those two positions. Even though the information is different in each case the principle used is the same, which is extract image points in different camera locations for later comparison.

For this work were used real planar objects images (Fig. 3), then applied a Harris corner extractor (Harris and Stephens, 1988), and once having the points at each image it was applied a robust algorithm of matching. The first stage of the matching was a variance normalized correlation (Vincent and Laganière, 2001). The Eq. (15) shows the correlation between two points p and p' in two images I e I' respectively.

$$VNC(p,p') = \frac{\sum_{i,i'}^{\mathrm{K}} (\boldsymbol{I}(i) - \overline{\boldsymbol{I}(\boldsymbol{p})}(\boldsymbol{I'}(i') - \overline{\boldsymbol{I'}(\boldsymbol{p'})}}{N\sqrt{\sigma_I(\boldsymbol{p})^2 \sigma_{I'}(\boldsymbol{p'})^2}} \tag{15}$$

here, k is the correlation window or neighborhood, $i \in k(p)$ and $i' \in k(p'), N$ is the pixel number in k, and $\sigma_I\ (p)\ ^2$ and $\overline{\boldsymbol{I}(p)}$ are the variance and mean of intensity in the neighborhood respectively. Once it is obtained a first match, it is applied a refinement algorithm by RANSAC using the *Homography* matrix as the random variable. The *Homography* is defined as a type of perspective projection between points of planes. A point on a plane in the 3D space $\boldsymbol{P}$ = *[X, Y, 0, 1]*, is projected at a point $\boldsymbol{p}$ = *[x, y, 1]* on the camera image plane by its *Homography* $\boldsymbol{H}$ (Hartley and Zisserman 2003), as

$$\boldsymbol{p}^T = \boldsymbol{H}\boldsymbol{P}^T \tag{16}$$

with $\boldsymbol{H}$ being a linear transformation represented by a non-singular 3 x 3 matrix, Each correspondence between the points $\boldsymbol{p} \leftrightarrow \boldsymbol{P}$ produces two independent linear equations as a function of the $\boldsymbol{H}$ components, h_{ij}:

$$\begin{gathered} h_{11}X + h_{12}Y + h_{13} - h_{31}xX - h_{32}xY - h_{33}x = 0 \\ h_{21}X + h_{22}Y + h_{23} - h_{31}xX - h_{32}xY - h_{33}x = 0 \end{gathered} \tag{17}$$

ISSN 2176-5480

A single solution of $\boldsymbol{H}$ is obtained with four correspondences between non-collinear points, thus, one can find the geometric relationship between two object images or between a plane object and its image. Once the *Homography* is estimated it is possible to implement the RANSAC algorithm to reject the matching outliers produced by the VNC.

Figure 4 shows the result of applying the entire robust matching to a pair of pictures of the same object in the different locations showed in Fig. 3.

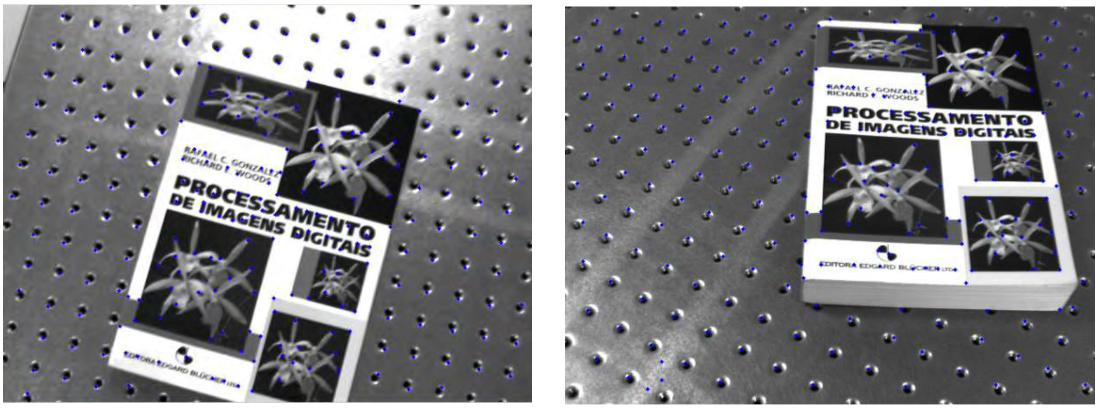

Figure 3. Harris corner extractor in two images

The entire image processing algorithm response shown to be well suited for these applications in visual servoing, it was able to process more than 15 fps that was our acquisition rate. From this point was possible apply the IBVS technique directly in the matching algorithm output that is the feature vector needed to compute $\boldsymbol{L_s}$ in the eq. (7), for this purpose were randomly taken 6 of the matches found .

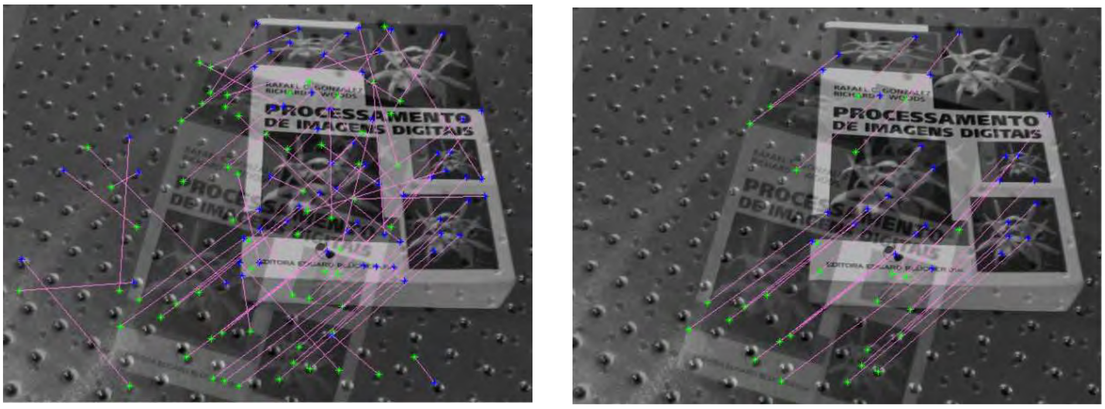

Figure 4. Robust Matching in two images, (a) with only VNC algorithm, (b) after RANSAC refiner

Computing the relative pose between the camera and the object is straightforward from the camera calibration matrix (Bougets toolbox, 2010) and the Homography, and then the PBVS error function can be obtained from the eq.(15)

4. EXPERIMENTAL SETUP AND RESULTS

For the experimental setup was used an industrial robot ABB IRB 140 with 6 DOF, a camera PULNIX TM 120 attached at the robot end effector, and a set of real planar objects as targets (Fig. 5). The reference image was taken in advance, and then the manipulator was set to an arbitrary initial position in which the target was fully or partially in the camera visual field. The image processing program was developed in C++ and the robot control in *RAPID* language on RobotStudio platform, the frame grabber used was a NI PCI-1428 and the scene illumination was not controlled. The Robot had a 40% reduction from its original velocity to allow the image acquisition, mainly because in very fast move the pictures were not suitable for the processing stage.

A set of tests were performed in order to make a realistic compare between IVBS and PBVS approaches. The Fig. 6 and Fig. 7 show the velocity responses in the IBVS control with misleading camera calibration parameters. The responses at these same conditions of PBVS are shown in Fig. 8 and Fig. 9. In Fig.10 it is possible to see the path followed by the camera from an initial position to a target at this experiment.

Figure 5. Experimental setup

As can be seen (Fig. 6 and Fig. 7) the results presented a good performance in the IBVS control, instead, for the PBVS the error in the camera calibration prevented the system to find the stability in most cases. It must be said that in order to achieve a proper out coming it was necessary to reject some singular responses in the control loop, the IBVS had a poor answer in the lack of this stage, basically in the compute for the interaction matrix and this fail was unavoidable in some situations like poses too near to the object. Another test was ran with regard to the model object, the model was provided to the PBVS with errors from (0.1% to 10%), the results in Tab. 1 show the high dependency of this approach to a well-defined model objet unlike the IBVS, which does not involve this parameter. The graphical responses of both controls are shown in Fig 11 and Fig 12.

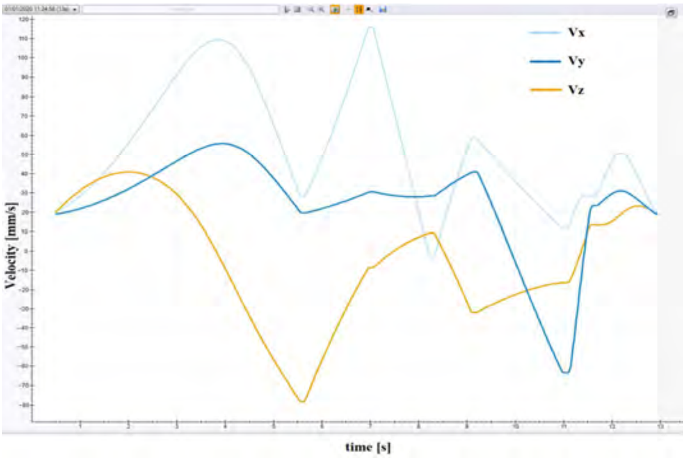

Figure. 6 Linear velocity of TCP in IBVS

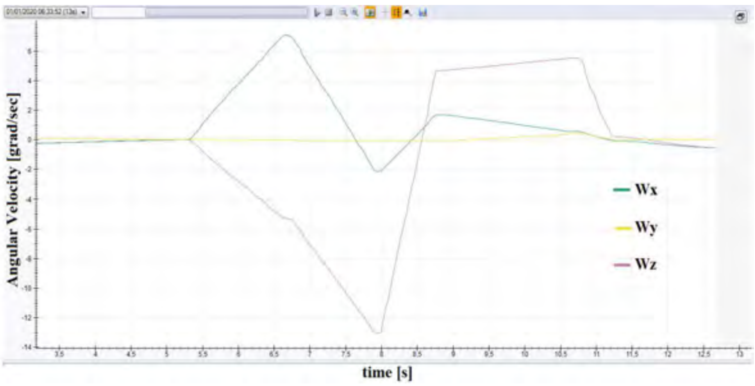

Figure. 7 Angular velocity of TCP in IBVS

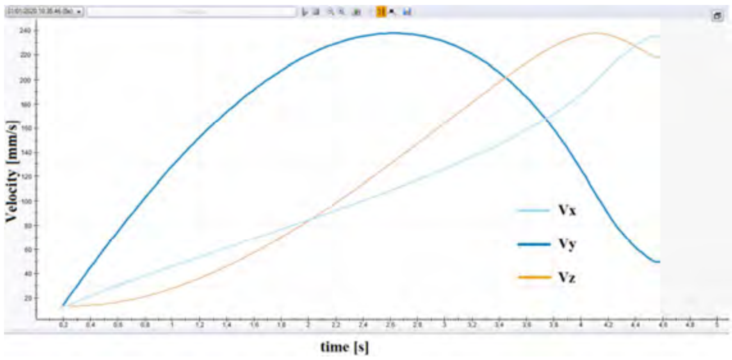

Figure. 8. Linear Velocity of TCP in PBVS

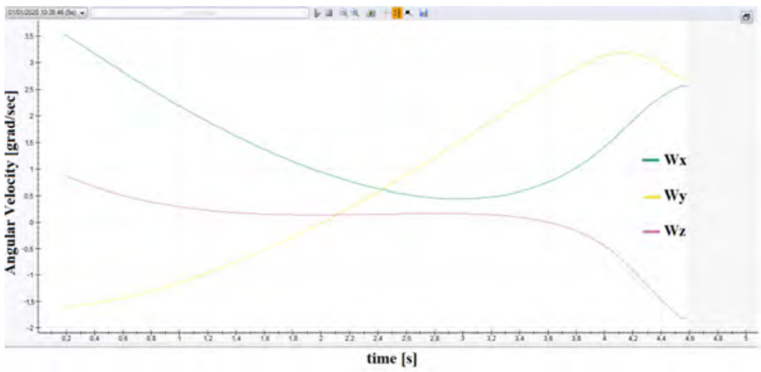

Figure. 9. Angular Velocity of TCP in PBVS

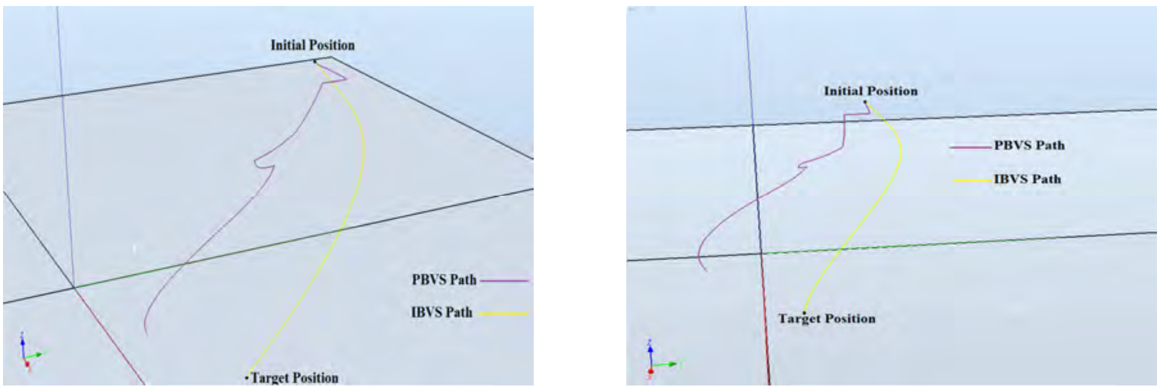

Figure. 10 Visual Servoing path with mistaken calibration parameters

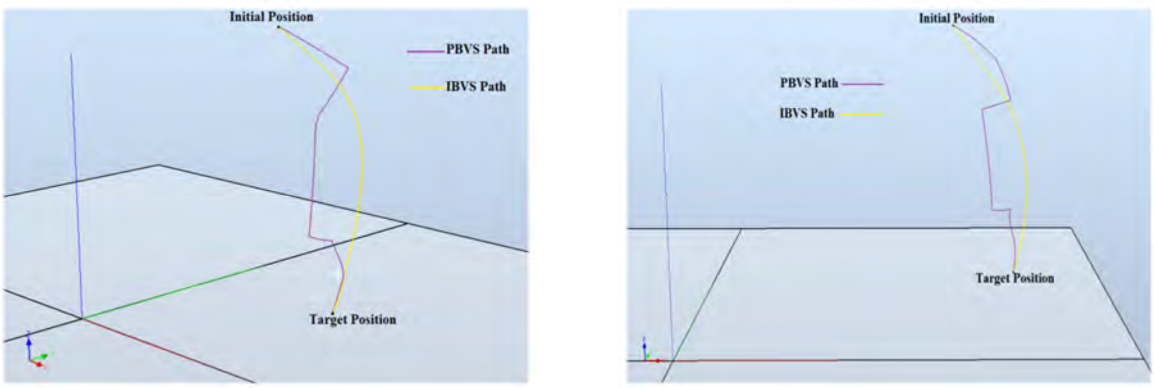

Figure 11. Visual Servoing path with correct model object

In the absence of errors, meaning, calibrated camera, well-defined object model, object always in the image plane, and workspace limitation. The two control approaches presented accurate localization response as shown in Fig.11, here are shown both the initial and the target position as the path followed by the camera in each approach.

Table 1. PBVS Effectiveness finding target pose with erroneous object model

% Error model (four points)	***PBVS*** ***% effectiveness (25 trials)***
1	87
2	85
3	79
4	85
5	78
6	75
7	64
8	70
9	62
10	60

5. CONCLUTIONS AND FUTURE WORK

The two major approaches of visual servoing, IBVS and PBVS, have been implemented and tested in a in a real platform composed by a 6DOF manipulator robot and a monocular camera attached to its end effector. They shared a unique image processing technique to identify the feature points in real planar objects that shows well suited response for the specific application of corner detection and point matching.

The both control techniques performed the positioning task with an overall good accuracy in the tested cases, however, the IBVS prove more robustness in lack of precisely calibration parameters as well in the absence of model objects, accordingly with the literature, but, a rejection stage was needed to prevent singularities mainly in the Image Based approach. One characteristic problem presented in both controls was a fail in sudden illumination changes because no control in this parameter was made.

The instability of the systems was improve significantly using more than 6 matching points in the images which is useful when time consuming is not a relevant issue, but, some drawbacks like occlusion in the processed image were not addressed in this work and overcome with this kind of problem in a real situation will increase the flexibility of VS. Another important improve to be made is the robustness against lighting changes and the use of non-planar target objects for the same aim.

The general results obtained from this first approach seem quite encouraging for the incoming research, once the basic aspects of visual servoing have been successfully achieve.

6. REFERENCES

Bouguet, J. Y. 2010. Camera calibration toolbox for MATLAB www.vision.caltech.edu/bouguetj/calib_doc/index.html.

Chaumette, F. 1990. La relation vision-commande: théorie et application à des tâches robotiques. Thesis. L'Université de Rennes I, Institut de Formation Supérieure en Informatique et Communication (IFSIC). Rennes.

Chaumette F., 1998. Potential problems of stability and convergence in image-based and position-based visual servoing. In: Kriegman D.J., Hager G.D., Morse A.S. (ed). Lecture notes in control and information sciences. The confluence of vision and control. Vol. 237. Springer-Verlag.

Chaumette F., Hutchinson S., 2006. Visual servo control I: Basic approaches. IEEE Transactions on Robotics and Automation. Vol. 13, No. 4, p. 82.

Chaumette F., Hutchinson S., 2007. Visual servo control II: Advanced approaches. IEEE Transactions on Robotics and Automation. Vol. 14, No. 1, p. 109.

Comport, A., Pressigout, M., Marchand, É., Chaumette, F., 2003. A visual servoing law that is robust to image outliers. Proceedings of the 2003 IEEE/RSJ, Conference on Intelligent Robots and Systems. Las Vegas: IEEE.

Corke, P. I., 1993. Visual control of robot manipulators: a review. In: Visual Servoing. Singapore: World Sci. Vol. 7.

Corke P. I., 2011. Robotics, Vision and Control. Fundamental algorithms in MATLAB. Springer.

Hartley R., Zisserman A. 2003. Multiple View Geometry in computer vision. Cambridge University Press, New York.

Hutchinson S., Hager G. D. and Corke P.I., 2006. A tutorial in visual servo control. IEEE Transactions on robotics and automation. Vol. 12, No. 5, p. 65.

ISSN 2176-5480

Kragic, D., Christensen, H. I., 2002. Survey on visual servoing for manipulation. Royal Institute of Technology, Computational Vision and Active Perception Laboratory. Stockholm.
Ma, Y., Kosecká, J. , Soatto, S. and Sastry, S., 2001. “An Invitation to 3-D Vision, From Images to Models”.
anderson, A. and Weiss, L. 1980. Image-based visual servo control using relational graph error signals. ProcIEEE, p. 1074.
Sanderson, A., Weiss, L. and Newman C., 1987. Dynamic sensor-based control of robots with visual feedback. IEEE Transactions on Robotics and Automation. Vol.3, No. 5. p:404.
Shirai Y., and Inoue H., 1973. Guiding a robot by visual feedback in assembling tasks. Patern recognition. Vol. 5, No. 2, p:99.
Vincent, É., Laganière, R. 2001. An empirical study of some feature matching strategies. Machine Graphics Vision Vol. 10, No. 3, p. 237.

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.