Survey of vision-based robot control

Ezio Malis

INRIA, Sophia Antipolis, France, Ezio.Malis@sophia.inria.fr

Abstract

In this paper, a short survey of vision-based robot control (generally called visual servoing) is presented. Visual servoing concerns several field of research including vision systems, robotics and automatic control. Visual servoing can be useful for a wide range of applications and it can be used to control many different dynamic systems (manipulator arms, mobile robots, aircraft, etc.). Visual servoing systems are generally classified depending on the number of cameras, on the position of the camera with respect to the robot, on the design of the error function to minimize in order to reposition the robot. In this paper, we describe the main visual servoing approaches proposed in the literature. For simplicity, the examples in the survey focuses on manipulator arms with a single camera mounted on the end-effector. Examples are taken from work made at the University of Cambridge for the European Long Term Research Project VIGOR (Visually guided robots using uncalibrated cameras).

1 Introduction

Vision feedback control loops have been introduced in order to increase the flexibility and the accuracy of robotic systems. The aim of the visual servoing approach is to control a robot using the information provided by a vision system. More generally, vision can be used to control disparate dynamic systems like for example vehicles, aircrafts and submarines. Vision systems are generally classified depending on the number of cameras and on their positions. Single camera vision systems are generally used since they are cheaper and easier to build than multi-camera vision systems. On the other hand, using two cameras in a stereo configuration [26, 22, 27] (i.e. the two cameras have a common field of view) make easier several computer vision problems. If the camera(s) are mounted on the robot we call the system "in-hand". In contrast, if the camera observe the robot from we can call the system "out-hand" (the term "stand-alone" is generally used in the literature). There exist hybrid systems where one camera is in-hand and another camera stand-alone observing the scene [19].

A fundamental classification of visual servoing approaches depends on the design of the control scheme. Two different control schemes are generally used for the visual servoing of a dynamic system [48, 30]. The first control scheme is called "*direct visual servoing*" since the vision-based controller directly compute the input of the dynamic systems (see Figure 1) [33, 47]. The visual servoing is carried out at a very fast (at least 100 Hz, with a rate of 10 ms). The second

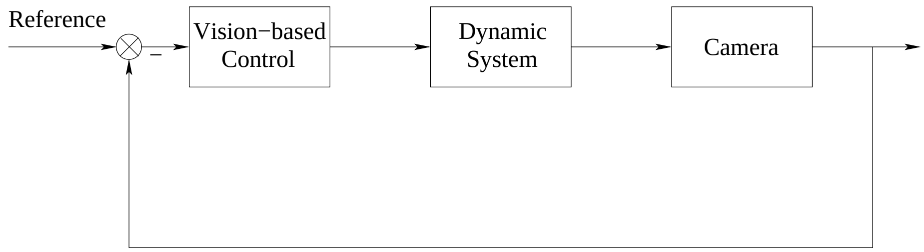

Figure 1: Direct visual servoing system

control scheme can be called, contrary to the first one, "*indirect visual servoing*" since the vision-based control compute a reference control law which is sent to the low-level controller of the dynamic system (see Figure 2). Most of the visual servoing proposed in the literature follows an indirect control scheme which is called "dynamic look-and-move" [30]. In this case the servoing of the inner loop (generally the rate is 10 ms) must be faster than the visual servoing (generally the rate is 50 ms) [8]. For simplicity, in this paper we consider examples of

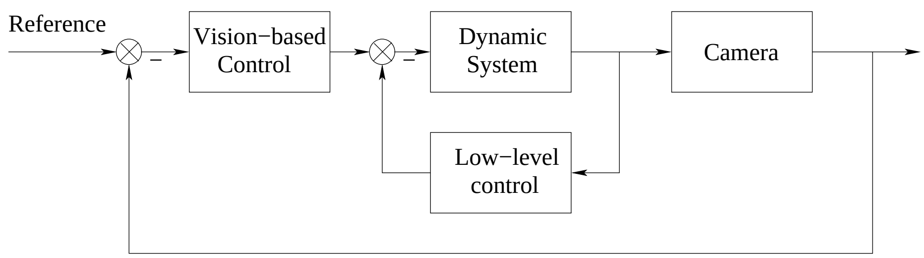

Figure 2: Indirect visual servoing system

positioning tasks using a manipulator with a single camera mounted on the end-effector. Examples are taken from work made at the University of Cambridge for the European Long Term Research Project VIGOR (Visually guided robots using uncalibrated cameras). The VIGOR project, leaded by INRIA Rhones-Alpes, produced a successful application of visual servoing to a welding task for shipbuilding industry at Odense Steel Shipyard. Several other examples of visual servoing systems can be found in [24] and in the special issues on visual servoing which have been published in international journals. The first one, published in the IEEE Transaction on Robotics and Automation, October 1996, contains a detailed tutorial [30]. The second one has been published in the International Journal of Computer Vision, June 2000. A third special issue on visual servoing should appear fall 2002 in the International Journal of Robotics Research.

2 Vision systems

A "pinhole" camera perform the perspective projection of a 3D point to the image plane. The image plane is a matrix of light sensitive cells. The resolution of the image is the size of the matrix. The single cell is called a "pixel". For each pixel of coordinates (u, v), the camera measures the intensity of the light. For example, a 3D point, with homogeneous coordinates $\boldsymbol{\mathcal{X}} = (X, Y, Z, 1)$ project to an image point with homogeneous coordinates $\mathbf{p} = (u, v, 1)$ (see Figure 3):

$$\mathbf{p} \propto \left[\begin{array}{cc} \mathbf{K} & 0 \end{array} \right] \boldsymbol{\mathcal{X}} \tag{1}$$

where $\mathbf{K}$ is a matrix containing the intrinsic parameters matrix of the camera:

$$\mathbf{K} = \left[\begin{array}{ccc} f k_u & f k_u \cot(\phi) & u_0 \\ 0 & \dfrac{f k_v}{\sin(\phi)} & v_0 \\ 0 & 0 & 1 \end{array} \right] \tag{2}$$

where u_0 and v_0 are the pixels coordinates of the principal point, k_u and k_v are the scaling factors along the $\overrightarrow{u}$ and $\overrightarrow{v}$ axes (in pixels/meters), ϕ is the angle between these axes and f is the focal length. For most of commercial cameras, it is a reasonable approximation to suppose square pixels (i.e. $\phi = \frac{\pi}{2}$ and $k_u = k_v$).

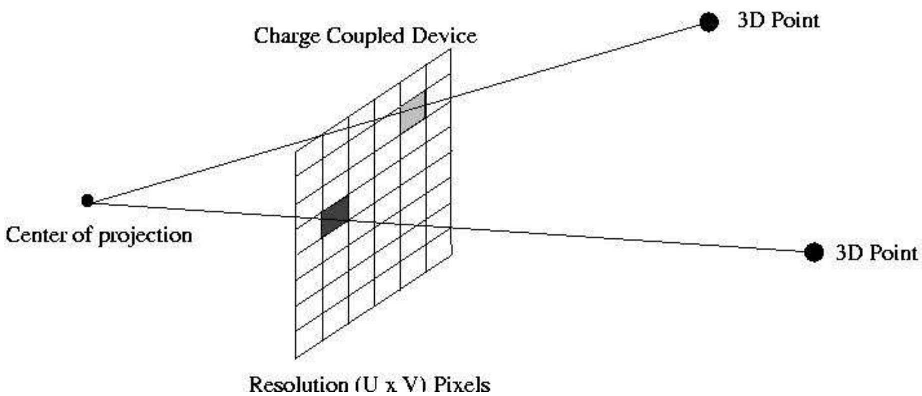

Figure 3: Camera model

The intrinsic parameters of the camera are often only roughly known. Precise calibration of the parameters is a tedious procedure which needs a specific calibration grid [17]. It is thus preferable to estimate the intrinsic parameters without knowing the model of the observed object. If several images of any rigid object are available it is possible to use a self-calibration algorithm [16] to estimate the camera intrinsic parameters.

2.1 Features extraction

Vision-based control approaches generally use points as visual features. One of the most known algorithm used to extracts interest points is the Harris detector [23]. However, several other features (straight lines, ellipses, contours, etc.) can be extracted from the images and used in the control scheme. One of the most known algorithm used to extracts contours from the image has been proposed by Canny [3].

2.2 Matching features

The problem of matching features, common to all visual servoing techniques, has been investigated in the literature but it is not yet a solved problem. For the model-based approach, we need to match the model to the current image [31]. With the model-free approach, we need to match feature points [52] or curves [49] between the initial and reference views. Finally, when the camera is zooming we need to match images with different resolutions [12]. Figure 4 shows an example of matching features between two views of the same object. The matching problem consist in finding the features in the left image which corresponds to the features in the right image. The problem is particularly difficult when the displacement of the camera between the two images is big and when light conditions change.

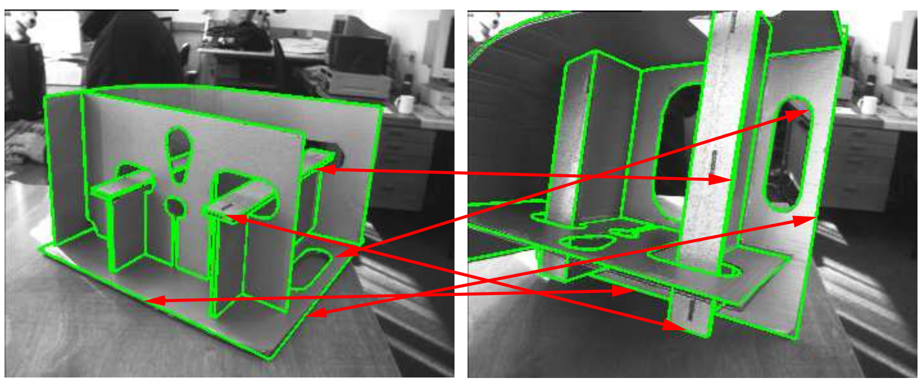

Figure 4: Matching features between two images

2.3 Tracking features

Tracking features is a similar problem to matching. However, in this case the displacement between the two images is generally smaller. Several tracking algorithms have been proposed in the literature. They can be classified depending on the a priori knowledge on the target used in the algorithm. If the model of the target is known see [18, 44, 45] and if the model of the target is known see [21, 11, 50, 41].

2.4 Motion estimation

The use of geometric features in visual servoing supposes the presence of these features on the target. Often, textured objects do not have any evident geometric feature. Thus, in this case, a different approach to tracking and visual servoing can be obtained by estimating the motion of the target between two consecutive images [43]. The velocity of the target in the image (see Figure 5) can be measured without any a priori knowledge of the model of the target.

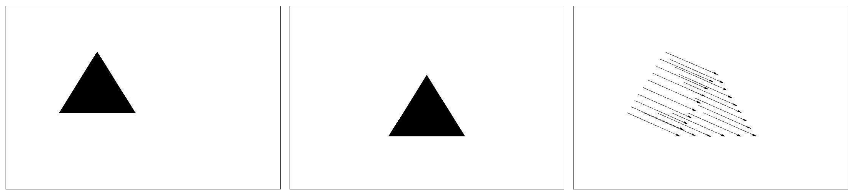

Figure 5: Motion estimation between two consecutive images.

3 Visual servoing approaches

Visual servoing schemes can be classified on the basis of the knowledge we have on the target and on the camera parameters. If the camera parameters are known we can use a "*calibrated visual servoing*" approach, while if they are only roughly known we must use an "*uncalibrated visual servoing*" approach. If a 3D model of the target is available we can use a "*model-based visual servoing*" approach, while if the 3D model of the target is unknown we must use a "*model-free visual servoing*" approach.

3.1 Model-based visual servoing

Let $\mathcal{F}_0$ be the coordinate frame attached to the target, $\mathcal{F}^*$ and $\mathcal{F}$ be the coordinate frames attached to the camera in its desired and current position respectively. Knowing the coordinates, expressed in $\mathcal{F}_0$, of at least four points of the target [10] (i.e. the 3D model of the target is supposed to be perfectly known), it is possible from their projection to compute the desired camera pose and the current camera pose (thus, the robot can be servoed to the reference pose). In this case, the camera parameters must be perfectly known and we are in the presence of a calibrated visual servoing scheme. If more than four points of the target are available, it is possible to compute the pose of the camera without knowing the camera parameters [15] and we are in the presence of an uncalibrated visual servoing scheme.

3.2 Model-free visual servoing

If the model of the target is unknown, a positioning task can still be achieved using a "teaching-by-showing" approach. With this approach, the robot is moved to a goal position, the camera is shown the target view and a "reference image" of the object is stored (i.e. a set of features of the reference image). The position of the camera with respect to the object will be called the "reference position". After the image corresponding to the desired camera position has been learned, and after the camera and/or the target has been moved, an error control vector can be extracted from the two views of the target. A zero error implies that the robot end-effector has reached its desired position with an accuracy regardless of calibration errors.

4 Vision-based control

Vision-based robot control can be classified, depending on the error used to compute the control law, into four groups: *position-based*, *image-based*, *hybrid* and *motion-based* control systems. In a *position-based* control system, the error is computed in the 3D Cartesian space [1, 51, 42, 20] (for this reason, this approach can be called *3D visual servoing*). In an *image-based* control system, the error is computed in the 2D image space (for this reason, this approach can be called *2D visual servoing*) [25, 14]. Recently, a new class of *hybrid visual servoing* approaches has been proposed. For example, in [38] I proposed a new approach which is called *2 1/2 D visual servoing* since the error is expressed in part in the 3D Cartesian space and in part in the 2D image space. Finally, a *motion-based* control system compute the error as a function of the optical flow measured in the image and a reference optical flow which should be obtained by the system [9].

4.1 3D Visual Servoing (Position-based Visual Servoing)

The 3D visual servoing is also called "position-based" visual-servoing since the control law uses directly the error on the position of the camera. Depending on the number of visual features available in the image one can compute the position error using or not the model of the target. The main advantage of this approach is that it directly controls the camera trajectory in Cartesian space. However, since there is no control in the image, the image features used in the pose estimation may leave the image (especially if the robot or the camera are coarsely calibrated), which thus leads to servoing failure. Also note that, if the camera is coarse calibrated, or if errors exist in the 3D model of the target, the current and desired camera poses will not be accurately estimated.

4.1.1 Model-based 3D Visual Servoing

The 3D visual servoing has originally been proposed as a model-based control approach since the pose of the camera can be computed using the knowledge of the 3D model of the target. Once we have the desired camera and the current camera pose, the camera displacement to reach the desired position is thus easily obtained, and the control of the robot end-effector can be performed either in open loop or, more robustly, in closed-loop as shown in Figure 6.

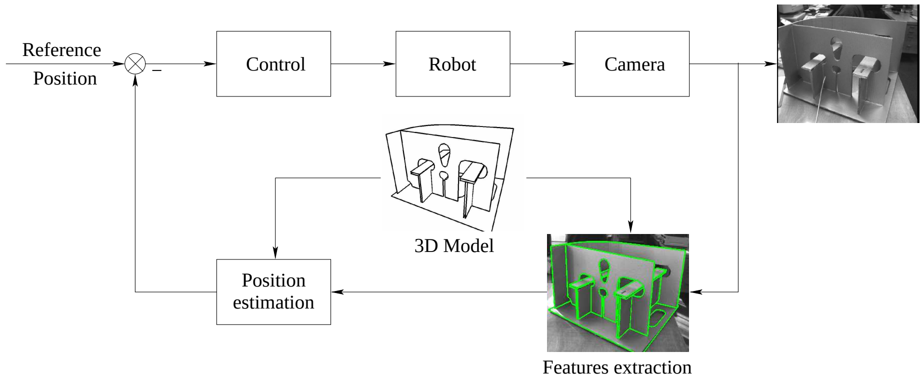

Figure 6: Model-based 3D Visual Servoing

4.1.2 Model-free 3D Visual Servoing

Recently, a 3D visual servoing scheme, which can be performed without knowing the 3D structure of the target, has been proposed in [2] (see Figure 7). The rotation and the direction of translation are obtained from the essential matrix [29]. The essential matrix is estimated from the current and reference images of the target [15]. However, as for the previous 3D visual servoing, such a control vector does not ensure that the considered object will always remain in the camera field of view, particularly in the presence of important camera or robot calibration errors.

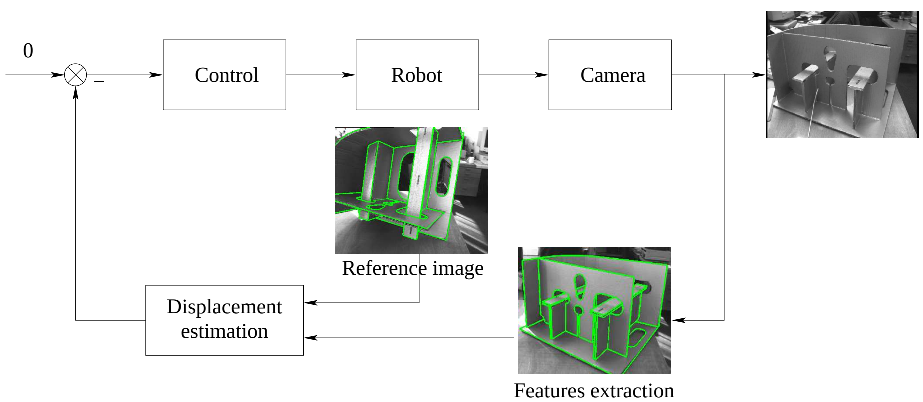

Figure 7: Model-free 3D Visual Servoing

4.2 2D Visual Servoing (Image-based Visual Servoing)

The 2D visual servoing is a model-free control approach since it does not need the knowledge of the 3D model of the target. The control error function is now expressed directly in the 2D image space (see Figure 8). Thus, the 2D visual servoing is also called "image-based" visual-servoing. In general, image-based visual servoing is known to be robust not only with respect to camera but also to robot calibration errors [13]. However, its convergence is theoretically ensured only in a region (quite difficult to determine analytically) around the desired position. Except in very simple cases, the analysis of the stability with respect to calibration errors seems to be impossible, since the system is coupled and non-linear.

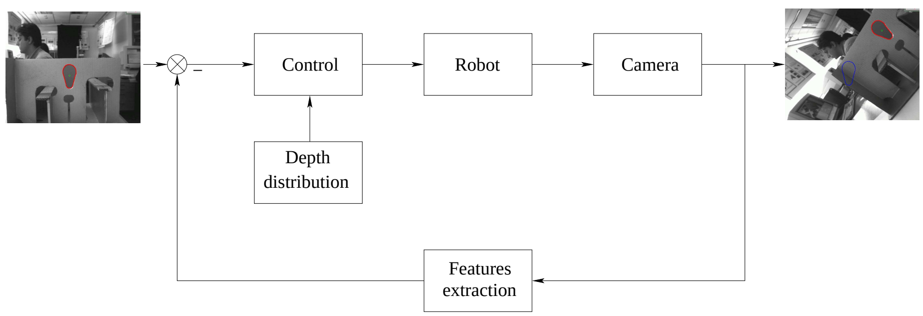

Figure 8: 2D visual servoing

Let $\mathbf{s}$ be the current value of visual features observed by the camera and $\mathbf{s}^*$ be the desired value of $\mathbf{s}$ to be reached in the image. For simplicity, we consider points as visual features. The interaction matrix for a large range of image features (straight lines, ellipses, etc.) can be found in [5]. The time variation of $\mathbf{s}$ is related to camera velocity $\mathbf{v}$ by:

$$\dot{\mathbf{s}} = \mathbf{L}(\mathbf{s}, Z)\mathbf{v} \tag{3}$$

where $\mathbf{L}(\mathbf{s}, Z)$ is the interaction matrix (also called the image Jacobian matrix) related to $\mathbf{s}$. Note that $\mathbf{L}$ depends on the depth Z of each selected feature. Thus, even if the 2D visual servoing is "model-free" we still need some knowledge on the depths of the points. Several solutions have been proposed in order to solve the problem of the estimation of the depth.

4.2.1 Estimating the interaction matrix on-line

The interaction matrix and the Jacobian of the robot can be computed on-line using a Broyden Jacobian estimation [28, 32]. However, the approximation of the interaction matrix is valid only in a neighborhood of the starting point.

4.2.2 Approximating the interaction matrix

In [14], the regulation of $\mathbf{s}$ to $\mathbf{s}^*$ corresponds to the regulation of a task function $\mathbf{e}$ (to be regulated to 0) defined by:

$$\mathbf{e} = \mathbf{C}(\mathbf{s} - \mathbf{s}^*) \tag{4}$$

where $\mathbf{C}$ is a matrix which has to be selected such that $\mathbf{CL}(\mathbf{s}, Z) > 0$ in order to ensure the global stability of the control law. The optimal choice is to consider $\mathbf{C}$ as the pseudo-inverse $\mathbf{L}(\mathbf{s}, Z)^+$ of the interaction matrix. The matrix $\mathbf{C}$ thus depends on the depth Z of each target point used in visual servoing. In order to avoid the estimation of Z at each iteration of the control law, one can choose $\mathbf{C}$ as a constant matrix equal to $\mathbf{L}(\mathbf{s}^*, Z^*)^+$, the pseudo-inverse of the interaction matrix computed for $\mathbf{s} = \mathbf{s}^*$ and $Z = Z^*$, where Z^* is an approximate value of Z at the desired camera position. In this simple case, the condition for convergence is satisfied only in the neighborhood of the desired position, which means that the convergence may not be ensured if the initial camera position is too far away from the desired one [4].

4.2.3 Estimating the depth distribution

An estimation of the depth can be obtained using, as in 3D visual servoing, a pose determination algorithm (if a 3D target model is available), or using a structure from known motion algorithm (if the camera motion can be measured). However, using this choice may lead the system close to, or even reach, a singularity of the interaction matrix. Furthermore, the convergence may also not be attained due to local minima reached because of the computation by the control law of unrealizable motions in the image [4].

4.3 2 $\frac{1}{2}$ D Visual Servoing (Hybrid visual servoing)

The main drawback of 3D visual servoing is that there is no control in the image which implies that the target may leave the camera field of view. Furthermore, a model of the target is needed to compute the pose of the camera. 2D visual servoing does not explicitly need this model. However, a depth estimation or approximation is necessary in the design of the control law. Furthermore, the main drawback of this approach is that the convergence is ensured only in a neighborhood of the desired position (whose domain seems to be impossible to determine analytically). In [39] I proposed a hybrid control scheme (called 2 1/2 D visual servoing) avoiding these drawbacks. Contrary to the position-based visual servoing, the 2 1/2 D control scheme does not need any geometric 3D model of the object. Furthermore and contrary to image-based visual servoing, the approach ensures the convergence of the control law in the whole task space. 2 1/2 D visual servoing is based on the estimation of the partial camera displacement from the current to the desired camera poses at each iteration of the control law [40]. Visual features [6] and data extracted from the partial displacement allow us to design a decoupled control law controlling the six camera d.o.f. (see Figure 9). The robustness of the visual servoing scheme with respect to camera calibration errors has also been analyzed. The necessary and sufficient conditions for local asymptotic stability are given in [39]. Finally, experimental results with an eye-in-hand robotic system confirm the improvement in the stability and convergence domain of the 2 1/2 D visual servoing with respect to classical position-based and image-based visual servoing.

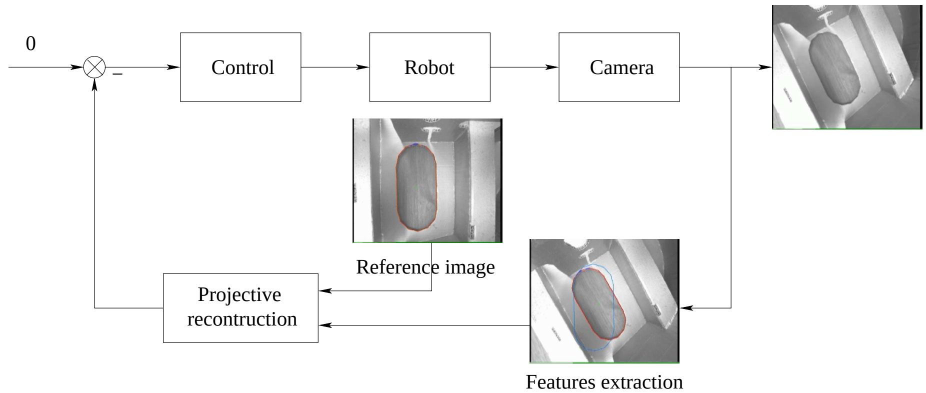

Figure 9: $2\frac{1}{2}$ D visual servoing

4.4 $\frac{d2D}{dt}$ Visual Servoing (Motion-based visual servoing)

Motion-based visual servoing is a model-free vision-based control technique since the optical flow in the image can be measured without having any a priori knowledge of the target. By specifying a reference motion field that should be obtained in the image, it is possible, for example, to control an eye-in-hand system in order to position the image plane parallel to a target plane [9]. The plane is unmarked, which means that no geometric visual features, such as points or straight lines, can be easily extracted. Other possible task are contour tracking [7], camera self orientation and docking [46]. The major problem with motion-based visual servoing is the high servoing rate (1.2 Hz in [9]) which imposes a small robot velocity. However, improvements on the motion estimation algorithms will make possible a faster visual control loop.

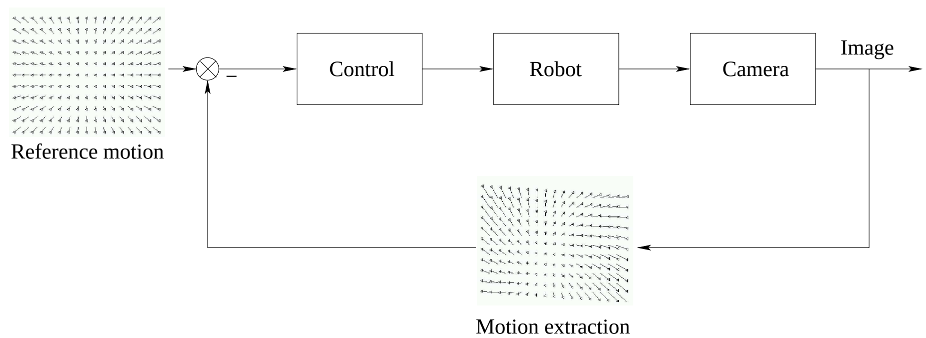

Figure 10: $\frac{d2D}{dt}$ visual servoing

5 Intrinsics-free visual servoing

As already mentioned in Section 3, standard visual servoing schemes can be classified into two groups: model-based and model-free visual servoing. Model-based visual servoing is used when a 3D model of the observed object is available. Obviously, if the 3D structure of the target is completely unknown model-based visual servoing can not be used. In that case, robot positioning can still be achieved using a teaching-by-showing approach. This model-free technique, completely different from the previous one, needs a preliminary learning step during which a reference image of the scene is stored. When the current image observed by the camera is identical to the reference image the robot is back to the reference position. Whatever visual servoing method is used, the robot can be correctly positioned with respect to the target if and only if the camera intrinsic parameters at the convergence are the same parameters of the camera used for learning. Indeed, if the camera intrinsic parameters change during the servoing (or the camera used during the servoing is different from the camera used to learn the reference image), the position of the camera with respect to the object will be completely different from the reference position.

Both model-based and model-free vision-based control approaches are useful but, depending on the "a priori" knowledge we have of the scene, we must switch between them. A new unified approach to vision-based control which can be used with a zooming camera whether the model of the object is known or not has been proposed in [36]. The key idea of the unified approach is to build a reference in a projective space invariant to camera intrinsic parameters [35]which can be computed if the model is known or if an image of the object is available. Thus, only one low level visual servoing technique must be implemented at once (see Figure 11). The strength of this approach is to keep the advantages of model-based and model-free methods and, at the same time, to avoid some of their drawbacks [34, 37]. The unified visual servoing scheme will be useful especially when a zooming camera is mounted on the end-effector of the robot. In that case, model-free visual servoing techniques cannot be used. Using the zoom during servoing is very important. The zoom can be used to enlarge the field of view of the camera if the object is getting out of the image and to bound the size of the object observed in the image (this can improve the robustness of features extraction).

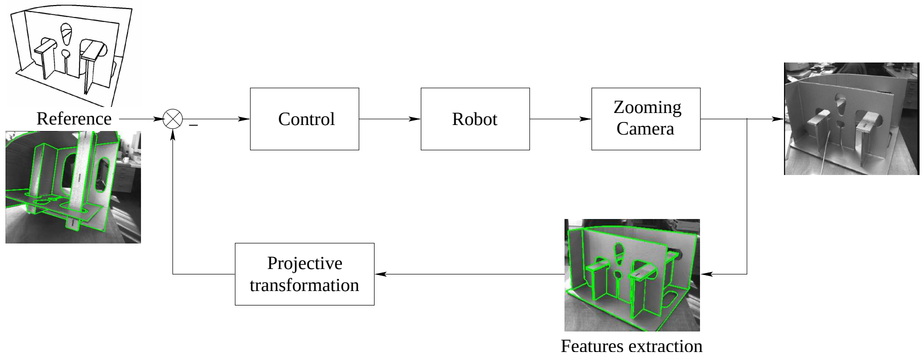

Figure 11: Intrinsics-free visual servoing

6 Conclusion

Visual Servoing is a very flexible technique for controlling autonomous dynamic systems. A wide range of applications, from object grasping to mobile robot navigation, are now possible. In the beginning, visual servoing techniques used a 3D model of the target and the positioning tasks were performed with position-based visual servoing. However, the 3D model of the target is not always easy to obtain. In that case, model-free visual servoing techniques allows to position the robot with respect to an unknown object. Image-based visual servoing is robust to noise in the image and calibration errors. However, it can have problem of convergence if the initial camera displacement is big. Moreover, care must be taken in order to provide a reasonable estimation of the depth distribution of 3D points. Hybrid visual servoing, and in particular 2 1/2 D visual servoing, is a possible solution to improve image-based and position-based visual servoing. Finally, an unified approach to model-based and model-free visual servoing, invariant to camera intrinsic parameters, is a new promising direction of research since it allows to use a zooming camera in the vision-based control loop. The main limitations of modern vision systems, common to all visual servoing techniques, are the image processing of natural images, the extraction of robust features for visual servoing. From a control point of view, the robustness to calibration errors is still an issue as well as keeping the target in the field of view of the camera during the servoing.

References

[1] P. K. Allen, A. Timcenko, Y. B., and P. Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Trans. on Robotics and Automation*, 9(2):152–165, April 1993.

[2] R. Basri, E. Rivlin, and I. Shimshoni. Visual homing: surfing on the epipole. *International Journal of Computer Vision*, 33(2):22–39, 1999.

[3] J. F. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[4] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In D. Kriegman, G. Hager, and A. Morse, editors, *The confluence of vision and control*, volume 237 of *LNCIS Series*, pages 66–78. Springer Verlag, 1998.

[5] F. Chaumette and E. Malis. 2 1/2 d visual servoing: a possible solution to improve image-based and position-based visual servoings. In *IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 630–635, San Francisco, USA, April 2000.

[6] G. Chesi, E. Malis, and R. Cipolla. Collineation estimation from two unmatched views of an unknown planar contour for visual servoing. In *British Machine Vision Conference*, volume 1, pages 224–233, Nottingham, United Kingdom, September 1999.

[7] C. Colombo, B. Allotta, and P. Dario. Affine visual servoing : A framework for relative positioning with a robot. In *International Conference on Robotics and Automation*, pages 464–471, Nagoya, Japan, May 1995.

[8] P. I. Corke and M. C. Good. Dynamic effects in visual closed-loop systems. *IEEE Trans. on Robotics and Automation*, 12(5):671–683, October 1996.

[9] A. Crétual and F. Chaumette. Positionning a camera parallel to a plane using dynamic visual servoing. In *IEEE Int. Conf. on Intelligent Robots and Systems*, volume 1, pages 43–48, Grenoble, France, September 1997.

[10] D. Dementhon and L. S. Davis. Model-based object pose in 25 lines of code. *International Journal of Computer Vision*, 15(1/2):123–141, June 1995.

[11] T. Drummond and R. Cipolla. Visual tracking and control using Lie algebras. In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, volume 2, pages 652–657, Fort Collins, Colorado, USA, June 1999.

[12] Y. Dufournaud, C. Schmid, and R. Horaud. Matching images with different resolutions. In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 618–618, Hilton Head Island, South Carolina, USA, June 2000.

[13] B. Espiau. Effect of camera calibration errors on visual servoing in robotics. In *3rd International Symposium on Experimental Robotics*, Kyoto, Japan, October 1993.

[14] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(3):313–326, June 1992.

[15] O. Faugeras. *Three-dimensionnal computer vision: a geometric viewpoint*. MIT Press, Cambridge, MA, 1993.

[16] O. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera self-calibration: Theory and experiments. In G. Sandini, editor, *Computer Vision - ECCV '92*, volume 588 of *Lecture Notes in Computer Science*, pages 321–334, Santa Margherita Ligure, Italia, May 1992. Springer-Verlag.

[17] O. Faugeras and G. Toscani. The calibration problem for stereo. In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 15–20, June 1986.

[18] J. T. Feddema and C. George Lee. Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera. *IEEE Trans. on Systems, Man, and Cybernetics*, 20(5):1172–1183, October 1990.

[19] J. T. Feddema, C. S. George Lee, and O. R. Mitchell. Weighted selection of image features for resolved rate visual feedback control. *IEEE Trans. on Robotics and Automation*, 7(1):31–47, February 1991.

[20] J. Gangloff, M. de Mathelin, and G. Abba. 6 d.o.f. high speed dynamic visual servoing using GPC controllers. In *IEEE Int. Conf. on Robotics and Automation*, pages 2008–2013, Leuven, Belgium, May 1998.

[21] D. Gennery. Visual tracking of known three-dimensional objects. *International Journal of Computer Vision*, 7(3):243–270, April 1992.

[22] G. D. Hager, W.-C. Chang, and A. S. Morse. Robot feedback control based on stereo vision: Towards calibration-free hand-eye coordination. In *IEEE Int. Conf. on Robotics and Automation*, volume 4, pages 2850–2856, San Diego, California, USA, May 1994.

[23] C. Harris and M. Stephens. A colbined corner and edge detector. In *4th Alvey Vision conference*, pages 147–151, 1988.

[24] K. Hashimoto. *Visual Servoing: Real Time Control of Robot manipulators based on visual sensory feedback*, volume 7 of *World Scientific Series in Robotics and Automated Systems*. World Scientific Press, Singapore, 1993.

[25] K. Hashimoto, T. Kimoto, T. Ebine, and H. Kimura. Manipulator control with image-based visual servo. In *IEEE Int. Conf. on Robotics and Automation*, volume 3, pages 2267–2271, Sacramento, California, USA, April 1991.

[26] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand eye coordination. *Image and Vision Computing*, 12(3):187–192, 1994.

[27] R. Horaud, F. Dornaika, and B. Espiau. Visually guided object grasping. *IEEE Transactions on Robotics and Automation*, 14(4):525–532, 1998.

[28] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true jacobian. In *IEEE Int. Conf. on Intelligent Robots and Systems*, pages 186–193, September 1994.

[29] T. S. Huang and O. Faugeras. Some properties of the E matrix in two-view motion estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(12):1310–1312, December 1989.

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

[31] D. Jacobs and R. Basri. 3-d to 2-d pose determination with regions. *International Journal of Computer Vision*, 34(2-3):123–145, 1999.

[32] M. Jgersand, O. Fuentes, and R. Nelson. Experimental evaluation of uncalibrated visual servoing for precision manipulation. In *IEEE Int. Conf. on Robotics and Automation*, volume 3, pages 2874–2880, Albuquerque, New Mexico, April 1997.

[33] R. Kelly. Robust asymptotically stable visual servoing of planar robots. *IEEE Trans. on Robotics and Automation*, 12(5):759–766, October 1996.

[34] E. Malis. Vision-based control using different cameras for learning the reference image and for servoing. In *IEEE/RSJ International Conference on Intelligent Robots Systems*, volume 3, pages 1428–1433, Maui, Hawaii, November 2001.

[35] E. Malis. Visual servoing invariant to changes in camera intrinsic parameters. In *International Conference on Computer Vision*, volume 1, pages 704–709, Vancouver, Canada, July 2001.

[36] E. Malis. An unified approach to model-based and model-free visual servoing. In *European Conference of Computer Vision*, Copenhaguen, Denmark, May 2002.

[37] E. Malis. Vision-based control invariant to camera intrinsic parameters: stability analysis and path tracking. In *IEEE International Conference on Robotics and Automation*, Washington, D.C., USA, May 2002.

[38] E. Malis, F. Chaumette, and S. Boudet. Positioning a coarse-calibrated camera with respect to an unknown planar object by 2D 1/2 visual servoing. In *Proc. 5th IFAC Symposium on Robot Control (SYROCO'97)*, volume 2, pages 517–523, Nantes, France, September 1997.

[39] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):234–246, April 1999.

[40] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. *International Journal of Computer Vision*, 37(1):79–97, June 2000.

[41] E. Marchand, P. Bouthemy, and F. Chaumette. A 2d-3d model-based approach to real-time visual tracking. *Image and Vision Computing*, 19(13):941–955, November 2001.

[42] P. Martinet, N. Daucher, J. Gallice, and M. Dhome. Robot control using monocular pose estimation. In *Workshop on New Trends In Image-Based Robot Servoing (IROS'97)*, pages 1–12, Grenoble, France, September 1997.

[43] J. Odobez and P. Bouthemy. Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4):348–365, December 1995.

[44] N. P. Papanikolopoulos, B. Nelson, and P. K. Khosla. Six degree-of-freedom hand/eye visual tracking with uncertain parameters. In *IEEE Int. Conf. on Robotics and Automation*, pages 174–179, San Diego, California, May 1994.

[45] J. A. Piepmeier, G. V. McMurray, and H. Lipkin. Tracking a moving target with model independent visual servoing: a predictive estimation approach. In *IEEE Int. Conf. on Robotics and Automation*, volume 3, pages 2652–2657, Leuven, Belgium, May 1998.

[46] P. Questa, E. Grossmann, and G. Sandin. Camera self orientation and docking maneuver using normal flow. In *SPIE AeroSense*, Orlando, Florida, April 1995.

[47] F. Reyes and R. Kelly. Experimental evaluation of fixed-camera direct visual controllers on a direct-drive robot. In *IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 2327–2332, Leuven, Belgium, May 1998.

[48] A. C. Sanderson and L. E. Weiss. Image based visual servo control using relational graph error signal. In *Proc. of the Int. Conf. on Cybernetics and Society*, pages 1074–1077, Cambridge, MA, October 1980.

[49] C. Schmid and A. Zisserman. The geometry and matching of lines and curves over multiple views. *International Journal of Computer Vision*, 40(3):199, 234 2000.

[50] H.-H. Tonko, M.and Nagel. Model-based stereo-tracking of non-polyhedral objects for automatic disassembly experiments. *International Journal of Computer Vision*, 37(1):99–118, June 2000.

[51] W. J. Wilson, C. C. W. Hulls, and G. S. Bell. Relative end-effector control using cartesian position-based visual servoing. *IEEE Trans. on Robotics and Automation*, 12(5):684–696, October 1996.

[52] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Technical Report 2273, INRIA, May 1994.