1995 IEEE International Conference on Robotics and Automation, Nagoya, Japan, May 21-27, 1995.

An Extendable Framework for Expectation-Based Visual Servoing Using Environment Models

Bradley J. Nelson   Pradeep K. Khosla

The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213-3890
WWW URL http://www.cs.cmu.edu:8001/afs/cs/user/bnelson/www/home.html

Abstract

Visual servoing is a manipulation control strategy that precisely positions objects using imprecisely calibrated camera-lens-manipulator systems. In order to quickly and easily integrate sensor-based manipulation strategies such as visual servoing into robotic systems, a system framework and a task representation must exist which facilitates this integration. The framework must also be extendable so that obsolete sensor systems can be easily replaced or extended as new technologies become available. In this paper, we present a framework for expectation-based visual servoing which visually guides tasks based on the expected visual appearance of the task. The appearance of the task is generated by a model of the environment that uses texture-mapped geometric models to represent objects. A system structure which facilitates the integration of various configurations of visual servoing systems is presented, as well as a hardware implementation of the proposed system and experimental results using a stereo camera system.

1. Introduction

Visual servoing is a robust technique for guiding a manipulator through an uncertain environment using poorly calibrated camera-lens-manipulator systems. The technique employs a differential view of the world in which only small scene changes between image frames are assumed and compensated for by any of a number of proposed control techniques.

The most important disadvantage of visual servoing is the complexity it adds to robotic systems that are often already too difficult and expensive to effectively use. Visual servoing is highly dependent on specialized hardware that is constantly evolving into faster, cheaper, and more robust sensing systems, resulting in systems that are difficult to extend. The implication is that any visual servoing implementation quickly becomes obsolete as new technology surpasses the capabilities of the current hardware implementation.

These facts point to the need for a general framework within which visually servoed manipulators can be quickly and easily integrated. This framework must be able to take advantage of the benefits visual servoing provides, specifically the ability to control manipulator motion using high bandwidth visual feedback, while avoiding the pitfall of being dependent on a particular visual sensor hardware configuration. The proper framework will allow the integration of any of a number of visual sensor hardware configurations, such as monocular camera systems, stereo pairs with varying baselines and verging angles, multi-baseline stereo systems, and laser rangefinders. The framework should also facilitate the integration of visual servoing with other sensor-based manipulation strategies.

In its basic form, image-based visual servoing is a purely reactive manipulator motion strategy. The objective of the control system is simply to make some set of measured feature states defined in the camera's sensor space match some set of desired states defined in the same space by moving the manipulator. In order for the execution of a robotic task to benefit from the advantages visual servoing provides, a representation of the task based on some model of the environment must exist. This representation must be capable of effectively directing a visually servoed manipulator.

In this paper, we propose a framework for visually servoed manipulation that uses an expectation-based approach to task execution. Visual servoing is used to aid in the manipulation of known objects existing within an imprecisely calibrated and dynamically varying environment. The camera-lens system as well as the manipulator affecting the manipulation task are assumed to be imprecisely calibrated. The objects being manipulated may be subject to some unknown disturbance. Because we are concerned with manipulation, it is important that strong geometrical representations of the objects being manipulated exist. Because most visual servoing systems use visual tracking algorithms based on grayscale features, it is important that the representation of the task also contains knowledge on the visual appearance of the task. Therefore, we propose the use of texture-mapped geometric models as the basis of our environment model. From this environment model we can then infer manipulation strategies using assembly planners which operate on the geometrical information provided by the models. We can subsequently guide the manipulator by guiding visual servoing strategies based on the current and desired visual appearance of the task. By adhering to principles of reconfigurability, and by using a low-level reconfigurable real-time architecture, we have developed an extendable framework for visual servoing based on an expectation-based approach to sensor-based manipulation.

2. Previous Work in Visual Servoing

The visual servoing field was first well defined by Weiss [30], though previous researchers have considered fast visual feedback for guiding manipulator motion, for example [26]. Since the work by Weiss, two types of visual servoing configurations have emerged, eye-in-hand configurations and static camera configurations. Eye-in-hand visual servoing tracks objects of interest with a camera mounted on a manipulator's end-effector [1][5][10][11][23][8][13][31]. Static camera visual servoing guides manipulator motion based on feedback from a camera observing the end-effector [18][20][4][12]. Most of this past work has been with monocular systems, though stereo systems have recently been used for visual servoing [19][12][15].

A typical visual servoing feedback loop is shown in Figure 1. Differences between the various approaches to visual servoing include the space in which reference inputs are provided, the dimensionality of the control space, the structure of the controller, the physical configuration of the system, the derivation of the control law, and the feature tracking algorithms used. An excellent survey of recent work in visual servoing can be found in [6].

Some researchers have considered the use of high bandwidth visual feedback to aid in the manipulation of objects, for example, [27][16][2]. These efforts have been more concerned with the visual tracking and calibration aspects of the problem rather than task representations and the extendability of system frameworks.

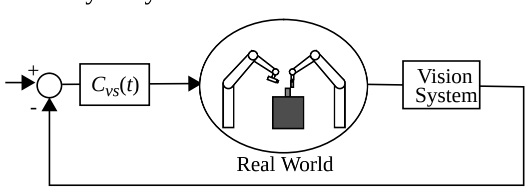

Figure 1. Visual servoing feedback loop.

3. Expectation/Verification Approaches

An expectation-based approach to scene understanding was first explicitly proposed by Dickmanns [7]. His work is concerned with guiding autonomous mobile systems in rapidly changing environments, particularly autonomous vehicles and aircraft. Roth and Jain propose a "verification-based" approach to navigation in the world [25]. A key point of both the expectation and verification-based approaches is that strong internal models of the recent world state are maintained. This significantly reduces the number of hypotheses that must be considered when determining the current state of the world based on sensory data. Neisser's view of the human "perceptual cycle" [22], as Jain points out [17], is similar in many ways to a verification or expectation based approach. Figure 2 shows a modified representation of Neisser's "perceptual cycle." This figure illustrates our view of the relationship between the model of the world, the real world, and where the visually servoed manipulator exists within this scheme. The counter-clockwise flow of information represents the cyclical nature of the system; sensory data updates the environment model, which in turn guides the visually servoed manipulator, which provides sensory data obtained from the real world to the environment model. This cycle illustrates the interaction between perception of the world, actions taken within this world, and plans made about the world.

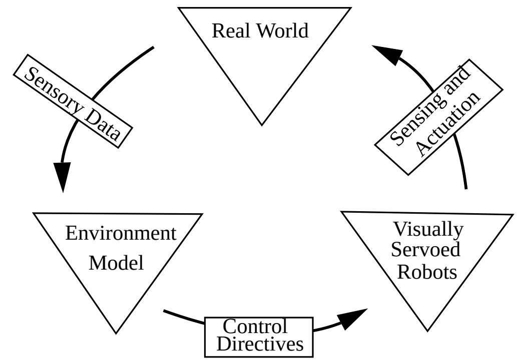

Figure 2. A modified "perceptual cycle" for visually servoed manipulators.

Our proposed integration of visually servoed robots into systems capable of performing manipulation tasks in the real world uses this cycle to clearly delineate the flow of information from the environment model to the visually servoed robot. By clearly separating the system in this way, different types of sensor-based manipulation strategies can be more easily integrated into the system, because the internal representation of the task becomes clearly separate from the hardware and low-level behavior-based control systems which perform actions and sensing in the real world. The aim is toward plug-and-play kinds of actuation/sensor components. In this paper, our interest is in the integration of a visually servoed actuation/sensing component, and what the structure of the internal representation of the task must be to facilitate quick and easy integration of visual servoing into the task.

4. The Environment Model

A key component of our modified perceptual cycle is the structure and content of the environment model. In order to provide the system with the capability to reason about manipulation strategies for objects, the internal representation must contain accurate three dimensional geometric knowledge of the objects being manipulated. The internal representation must also be capable of being correlated with visual information. This implies that some visual representation of the task must exist internally. Texture mapped CAD models are used to provide this visual representation. This same representation can be used to provide feedback to a remote teleoperator as in [9] and [14].

5. System Structure

Our previous work in defining the sensor placement measure *resolvability* [21] compared various visual sensor systems from a control standpoint in terms of the accuracy of control. This measure quantifies the ability of various monocular and stereo camera configurations to resolve the position and orientation of visually servoed objects. As discussed in [21], the measure is easily extendable to other visual sensors including multi-baseline stereo and laser rangefinders. A key component of this measure, the Jacobian mapping from task space to sensor space, is also a critical component of our visual servoing control strategy. Resolvability provides a shared ontology, that is a scheme allowing us to understand the relationships among various visual sensor configurations used for visual control. This shared ontology provided by resolvability is a key component of our expectation-based approach to visual servoing.

Figure 3 shows a block diagram structure of the expectation-based visually servoed system. The goal of the system is to maintain a temporal coherence between the environment model and the real world using visual feedback and visual servoing. We assume that if this coherence is maintained the task is being properly executed, within the limits of sensor resolution. Two feedback loops exist in the system. The top loop maintains a dynamic model of the environment by servoing the positions and orientations of objects within the environment model so that feedback provided by the vision system agrees with the expected

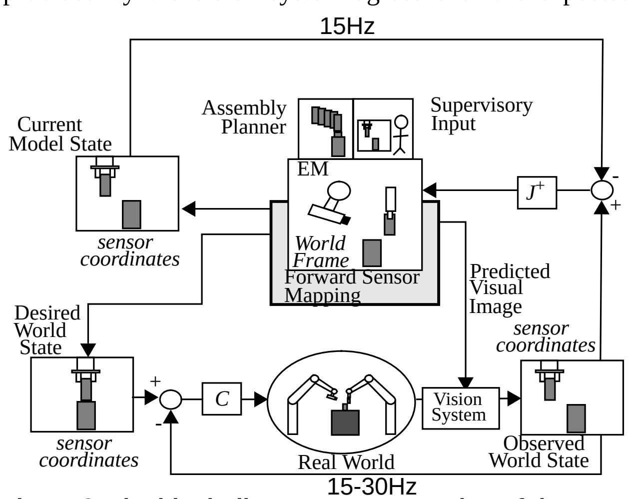

Figure 3. The block-diagram representation of the proposed framework for visually servoed manipulation

appearance of the model. We refer to the dynamically updated geometric models of objects as “servoed models.”

The bottom loop is a typical visual servoing feedback loop as shown in Figure 1. The reference inputs to the feedback loop are provided by planning algorithms that operate within the internal representation of the world. Instead of using an autonomous planner, a supervisor may guide objects within the dynamically varying environment model. Like most visual servoing researchers, we prefer an image-based visual servoing approach, rather than a position-based approach, to avoid calculating the inverse perspective mapping of the scene at each sampling period. Therefore, we must provide reference inputs to our visual servoing system in feature coordinates. To do this, desired object positions existing in the 3D internal representation must be mapped into image coordinates using a simple perspective projection model of the particular visual sensor.

5.1. Task space to sensor space mapping

Both servo loops require knowledge of the mapping between motion in the task space and motion in the sensor space. For any visual sensor configuration, we desire an equation of the form

$$\dot{\boldsymbol{x}}_S = \boldsymbol{J}(\phi)\dot{\boldsymbol{x}}_T \quad (1)$$

where $\dot{\boldsymbol{x}}_S$ is a velocity vector in sensor space and $\dot{\boldsymbol{x}}_T$ is a velocity vector in task space. $\boldsymbol{J}(\phi)$ is the Jacobian matrix and is a function of the extrinsic and intrinsic parameters of the visual sensor as well as the number of features tracked and their locations on the image plane.

For the experimental results to be presented, an orthogonal stereo pair is used. Figure 4 shows the coordinate frame definitions for this type of camera-lens configuration. If the axes are aligned as shown in the figure, the Jacobian mapping from task space to sensor space can be written as

$$J = \begin{bmatrix} \frac{f}{s_x Z_{Cl}} & 0 & -\frac{x_{Sl}}{Z_{Cl}} & -\frac{x_{Sl}Y_T}{Z_{Cl}} & \left[\frac{fZ_T}{s_x Z_{Cl}} + \frac{x_{Sl}X_T}{Z_{Cl}}\right] & -\frac{fY_T}{s_x Z_{Cl}} \\ 0 & \frac{f}{s_y Z_{Cl}} & -\frac{y_{sl}}{Z_{Cl}} & -\left[\frac{fZ_T}{s_y Z_{Cl}} + \frac{y_{Sl}Y_T}{Z_{Cl}}\right] & \frac{y_{Sl}X_T}{Z_{Cl}} & \frac{fX_T}{s_y Z_{Cl}} \\ \frac{x_{Sr}}{Z_{Cr}} & 0 & \frac{f}{s_x Z_{Cr}} & \frac{fY_T}{s_x Z_{Cr}} & \frac{x_{Sr}Z_T}{Z_{Cr}} - \frac{fX_T}{s_x Z_{Cr}} & \frac{x_{Sr}Y_T}{Z_{Cr}} \\ \frac{y_{Sr}}{Z_{Cr}} & \frac{f}{s_y Z_{Cr}} & 0 & -\frac{fZ_T}{s_y Z_{Cr}} & \frac{y_{Sr}Z_T}{Z_{Cr}} & \frac{y_{Sr}Y_T}{Z_{Cr}} - \frac{fX_T}{s_y Z_{Cr}} \end{bmatrix} \quad (2)$$

where f is the focal length of the lens and s_x and s_y are the horizontal and vertical dimensions of the pixels on the

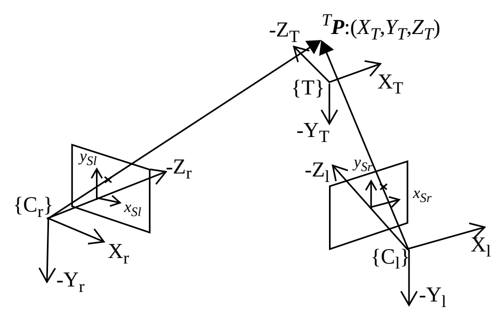

Figure 4. Task frame-camera frame definitions.

CCD array. In (2), we assume these parameters are identical for both cameras. The other terms in (2) correspond to Figure 4.

Any of a number of camera configurations could be used. The difference from a geometrical standpoint comes only with different Jacobian matrices. In [21], resolvability is used to analyze the effects of different camera configurations on the accuracy of control.

5.2. Servoed modeling

The servoed model loop updates models in the environment model through a recursive least-squares algorithm in which errors between the observed feature locations and the model feature locations are multiplied by a pseudoinverse of the camera Jacobian (2). Model locations are then adjusted in the internal representation to reduce errors between the internal representation and sensory data. A key aspect of servoed modeling is that different sensor configurations can be quickly and easily integrated into the system simply by using the appropriate Jacobian matrix.

5.3. Visual servoing control strategy

The control strategy for the bottom loop is obtained using controlled active vision[23] and is of the form

$$\boldsymbol{u}(k) = -\left(\boldsymbol{J}^T(k)\boldsymbol{Q}\boldsymbol{J}(k) + \boldsymbol{R}\right)^{-1}\boldsymbol{J}^T(k)\boldsymbol{Q}\,[\boldsymbol{x}(k) - \boldsymbol{x}_D(k+1)] \qquad (3)$$

where $\boldsymbol{x}(k) \in R^{2M}$, T is the sampling period of the vision system, $\boldsymbol{u}(k) = \begin{bmatrix}\dot{x}_T & \dot{y}_T & \dot{z}_T & \omega_{x_T} & \omega_{y_T} & \omega_{z_T}\end{bmatrix}^T$ is the commanded manipulator end-effector velocity, and M is the number of features being tracked. $\boldsymbol{Q}$ and $\boldsymbol{R}$ are weighting matrices and allow the user to place a varying emphasis on the feature error and the control input. Extensions to this control strategy and guidelines for choosing the matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ can be found in [24]. Again, new sensor configurations can be quickly and easily added to the system by substituting the correct $\boldsymbol{J}$ and by adjusting $\boldsymbol{Q}$ and $\boldsymbol{R}$ accordingly.

5.4. Visual tracking of features

The measurement of the motion of the features on the image plane must be done continuously and quickly. Our method for measuring this motion is based on optical flow techniques and is a modification of the method proposed in [3]. This technique is known as Sum-of-Squared-Differences (SSD) optical flow, and is based on the assumption that the intensities around a feature point remain constant as that point moves across the image plane. A more complete description of the algorithm and its original implementation can be found in [24].

By integrating texture mapped environment models with our visual tracker, we are able to use context-based vision strategies to take advantage of a large amount of previously unavailable information. We use texture-mapped models in two ways: 1. during initial feature selection, and 2. during feature tracking.

Our original feature selector was based on a technique proposed in [29]. This feature selector assumes that the best features to track are features which have the strongest omni-directional gradients, for example well defined corners or small circles. One problem we have encountered during feature selection is that exterior features are often selected as the best features because exterior boundaries often exhibit strong gradients. Tracking algorithms are more prone to failure if feature templates are located near the exterior boundaries of objects due to background "noise." Because texture mapped models inherently maintain the correct segmentation of objects, this segmentation is used to avoid choosing features which lie near the exterior edges of the object.

During feature tracking, environment models are also used to aid in tracking. Although the feature template used for tracking is obtained from actual image data rather than from the texture mapped model, correlations are periodically performed with the current feature template on the model in order to determine if a particular feature template is unable to successfully track an object because of scale changes or occlusions. If tracking performance decreases, a new feature is selected using the most recent image.

6. Hardware Implementation

The expectation-based visually servoing framework has been implemented on a robotic assembly system consisting of three Puma 560's called the Troikabot. The Pumas are controlled using the Chimera 3.0 reconfigurable real-time operating system[28]. An Adept robot is also used for providing accurate target motion for experimental purposes.

A Datacube Maxtower Vision System calculates the optical flow of the features using the SSD algorithm discussed in Section 5.4. An image can be grabbed and displacements for up to five 16x16 features in the scene can be determined at 30Hz. Stereo system implementations result in half the sampling frequency because only a single digitizer exists on the Datacube. For the stereo system with which experimental results were obtained, visual servoing is performed at 15Hz while tracking ten features. The vision system VME communicates with the robot controller VME using BIT3 VME-to-VME adapters.

The environment model exists within a Silicon Graphics Indigo2 Extreme. The robot simulation package Telegrip™ is used for modeling and texture mapping objects. Unix sockets and the Chimera Enet protocol are used to communicate between the Silicon Graphics machine and the robot controller and vision system. The rate of communication between the Silicon Graphics and the robot controller is approximately 15Hz.

7. Experimental Results

In order to illustrate how our proposed framework operates, an example of servoed modeling is shown, followed by an example of an insertion task guided by a supervisor within the environment model. Figure 5 shows images of the real scene and the visual internal representation of the same scene from a single camera. The fingers of a gripper mounted at the manipulator flange can be seen, as well as a texture-mapped rotor-stator motor pair. An Adept robot provides periodic motion to the surface on which the rotor-stator pair rest. The servoed modeling technique treats the motion as a disturbance and updates the position of the model accordingly within the environment model. Figure 6 shows the 3D translational motion of the object. Coordinate axes correspond to those shown in Figure 4. Maximum object speeds of approximately 6cm/s were induced and successfully tracked. Figure 7 shows the motion of a feature in the live image and the position of the same feature in the visual representation. As can be seen in the plot, a significant time delay is introduced due to the use of an ethernet link as a communication mechanism between the environment model that exists on the Silicon Graphics machine and the controller VME running Chimera. The link introduces a delay of approximately 0.6sec into the servoed modeling feedback loop. The two plots show that relatively fast tracking speeds can be achieved, but latency causes significant error between the magnitude of feature errors on the image plane for the live image and the corresponding expected feature error obtained from the internal visual representation of the task.

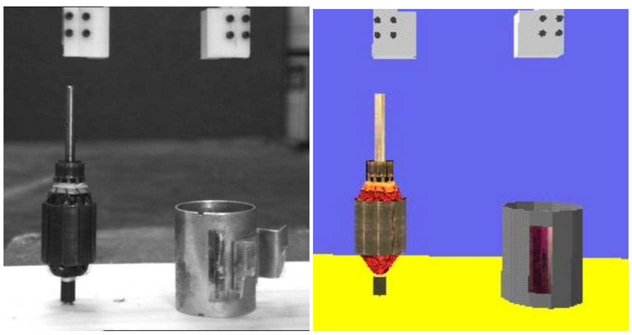

Figure 5. A live image of the environment is shown of the left, and the corresponding visual representation of the same scene as derived from the texture mapped environment model is shown to the right.

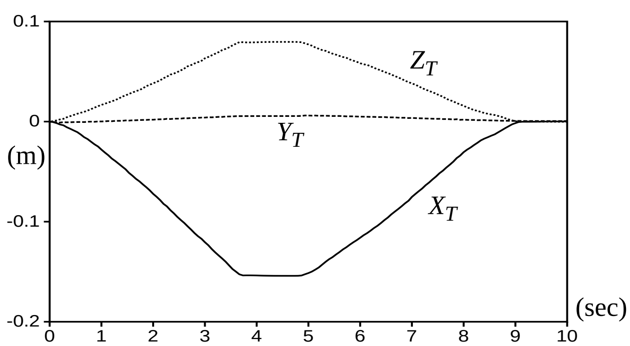

Figure 6. Translational motion of rotor in 3D versus time.

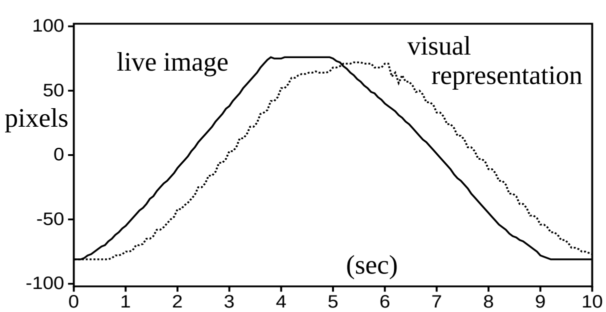

Figure 7. Feature trajectories in the live image and in the visual representation of the image.

In Figure 8, feature trajectories are shown in which a supervisor guides the rotor into the stator after the rotor is grasped by a manipulator. The supervisor guides the rotor by clicking on its environment model with a mouse and dragging the model in a plane within the visual representation of the environment. Visual servoing is used to maintain a temporal coherence between the environment model and the real world. Two feature states are shown. The solid line corresponds to the feature locations within the visual representation. The dashed line corresponds to the same feature in the live image. Again, the effects of latency due to the ethernet link are evident, but relatively fast tracking speeds are still obtained. The main limitation of the latency inherent in the ethernet link is that the speed at which moving objects can be grasped is reduced. We are currently investigating techniques for reducing this latency.

Finally, in Figure 9 feature templates from the rotor are shown. The left template is the selected feature, and the right template is the corresponding feature matched on the texture mapped environment model that contains information concerning the location of the feature on the object and possible occlusions.

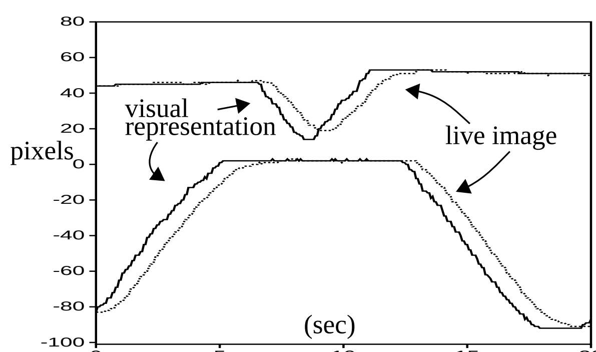

Figure 8. Feature trajectories during an insertion.

Figure 9. Example templates derived from a live image on the left and the visual representation of the environment on the right.

8. Conclusion

A framework and task representation for integrating visually servoed manipulation strategies with sensor-based robotic systems has been described and initial experimental results presented. Our goal is to develop a framework which will allow the use of plug-and-play types of sensor-based manipulation components. A key component of the framework is the use of an expectation-based approach in which texture mapped geometric models are used to describe and reason about the environment. A technique called servoed modeling is used to insure that the current state of the environment model agrees with current sensory input. Visual servoing is used to ensure that desired actions which are described within the environment model are carried out in the real world. Resolvability provides a shared ontology among various visual sensor configurations and allows the quick and easy incorporation of different visual sensor configurations into the servoed model and visual servo control loops. Ongoing work aims to demonstrate system extendability by integrating other visual sensors into the system as well as feedback from other types of sensors such as force.

Acknowledgments

Brad Nelson was supported in part by a National Defense Science and Engineering Graduate Fellowship through the U.S. Army Research Office through Grant Number DAAL03-91-G-0272 and by Sandia National Laboratories through Contract Number AC-3752D.

References

[1] P.K. Allen, "Real-time motion tracking using spatio-temporal filters," *Proc. Image Understanding Workshop*. San Mateo, CA:Morgan Kaufmann, pp. 695-701, 1989.

[2] P.K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman, "Automated Tracking and Grasping of a Moving Object with a Robotic Hand-Eye System," in *IEEE Trans. on Robotics and Automation*, 9(2), pp.152-165, 1993.

[3] P. Anandan, Measuring visual motion from image sequences. Tech. Rept. COINS-TR-87-21, Amherst, Mass.:University of Massachusetts COINS Department, 1987.

[4] A. Castaño and S. Hutchinson, "Visual compliance: task-directed visual servo control," *IEEE Trans. on Robotics and Automation*. 10(3):334-342, 1994.

[5] P.I. Corke and Paul, R.P., "Video-rate visual servoing for robots," *Lecture notes in control and information science*. eds. V. Hayward and O. Khatib. London: Springer-Verlag, pp. 429-451, 1989.

[6] P.I. Corke, "Visual control of robot manipulators-a review," *Visual Servoing: Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*, ed. K. Hashimoto. London:World Scientific, pp. 1-31, 1993.

[7] E.D. Dickmanns, "Expectation-based Dynamic Scene Understanding," in *Active Vision*, eds. A. Blake and A. Yuille, 303-335, The MIT Press, Cambridge, 1992.

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

[9] C. Fagerer, D. Dickmanns, and E.D. Dickmanns, "Visual Grasping with Long Delay Time of a Free Floating Object in Orbit," *Autonomous Robots*, 1(1):53-68, 1994.

[10] J.T. Feddema and Lee, C.S.G., "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, 1990.

[11] B.K. Ghosh, Jankovic, M. and Wu, Y.T., "Some problems in perspective system theory and its application to machine vision," *Proc. 1992 IEEE/RSJ Int. Conf. on Intelligent Robots and Sys. (IROS92)*, pp. 139-146, 1992.

[12] G.D. Hager, "Robot feedback control based on stereo vision: towards calibration-free hand-eye coordination," *Proc. 1994 Int. Conf. on Robotics and Automation*, pp. 2850-2856, 1994.

[13] K. Hashimoto and Kimura, H., "LQ optimal and nonlinear approaches to visual servoing," *Visual Servoing: Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*, ed. K. Hashimoto. London:World Scientific, pp. 165-198, 1993.

[14] G. Hirzinger, "ROTEX-the first space robot technology experiment," *Experimental Robotics III: The Third Int. Symp., Kyoto, Japan, Oct. 28-30, 1993*, eds. T. Yoshikawa and F. Miyazaki, Springer-Verlag, pp.579-598, 1994.

[15] K. Hosoda and Asada, M., "Versatile visual servoing without knowledge of true jacobian," *Proc. 1994 IEEE/RSJ Int. Conf. on Intelligent Robots and Sys. (IROS94)*, 1994.

[16] J. Ishikawa, Kosuge, K. and Furuta, K., "Intelligent control of assembling robot using vision sensor," *Proc. 1985 Int. Conf. on Robotics and Automation*, pp. 1904-1909, 1990.

[17] R. Jain, "Environment Models and Information Assimilation," Technical Report RJ 6866(65692), IBM-Yorktown Heights, 1989.

[18] A.J. Koivo and Houshangi, N., "Real-time vision feedback for servoing of a robotic manipulator with self-tuning controller," *IEEE Trans. on Sys, Man, and Cybernetics*, 21(1):134-142, 1991.

[19] N. Maru, Kase, H., Yamada, S., Nishikawa, A. and Miyazaki, F., "Manipulator control by visual servoing with the stereo vision," *Proc. 1993 IEEE/RSJ Int. Conf. on Intelligent Robots and System (IROS-93)*. pp. 1866-1870, 1993.

[20] B. Nelson, N.P. Papanikolopoulos, and P.K. Khosla, "Visual servoing for robotic assembly," *Visual Servoing—Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*. ed. K. Hashimoto. River Edge, NJ:World Scientific Publishing Co. Pte. Ltd. pp. 139-164, 1993.

[21] B.J. Nelson and P.K. Khosla, "The Resolvability Ellipsoid for Visual Servoing," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR94)*, pp. 829-832, 1994.

[22] U. Neisser, *Cognition and Reality*, W.H. Freeman and Co., New York, 1976.

[23] N.P. Papanikolopoulos, Khosla, P.K. and Kanade, T., "Adaptive robotic visual tracking," *Proc. of the American Control Conference*. Evanston, IL.:American Automation Control Council, pp. 962-967, 1991.

[24] N.P. Papanikolopoulos, Nelson, B. and Khosla, P.K., "Full 3-d tracking using the controlled active vision paradigm," *Proc. 1992 IEEE Int. Symp. on Intelligent Control (ISIC-92)*, pp. 267-274, 1992.

[25] Y. Roth and R. Jain, "Verification versus Discovery in Vision-Based Systems," Technical Report CSE-TR-110-91, The University of Michigan, 1991.

[26] S. Shirai and Inoue H., "Guiding a robot by visual feedback in assembling tasks," *Pattern Recognition*. 5:99-108, 1973.

[27] S.B. Skaar, W.H. Brockman, and R. Hanson, "Camera-space Manipulation," in *Int. J. of Robotics Research*, **6**(4), 20-32, 1987.

[28] D.B. Stewart, Schmitz, D.E. and Khosla, P.K., "The Chimera II real-time operating system for advanced sensor-based control systems," *IEEE Trans. Sys., Man Cyber.* 22(6):1282-1295., 1992.

[29] C. Tomasi and Kanade, T., "Detection and tracking of point features," Tech. Rept. CMU-CS-91-132. Pittsburgh:Carnegie Mellon University School of Computer Science, 1991.

[30] L.E. Weiss, "Dynamic visual servo control of robots: an adaptive image-based approach," Ph.D Thesis CMU-RI-TR-84-16, Pittsburgh, PA:The Robotics Institute Carnegie Mellon University, 1984.

[31] W.J. Wilson, "Visual servo control of robots using kalman filter estimates of robot pose relative to work-pieces," *Visual Servoing: Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*, ed. K. Hashimoto. London:World Scientific, pp. 71-104, 1993.