

Flexible Self-Calibrated Visual Servoing for a Humanoid Robot

Geoffrey Taylor and Lindsay Kleeman

Intelligent Robotics Research Centre

Department of Electrical and Computer Systems Engineering,

Monash University, Clayton, VIC 3168, Australia

{Geoffrey.Taylor,Lindsay.Kleeman}@eng.monash.edu.au

Abstract

This paper develops a flexible position-based visual servo framework to enable a humanoid robot to perform a variety of visually controlled manipulation tasks. The system overcomes classical drawbacks of position-based visual servoing, including hand-eye calibration and handling large pose errors, while catering for complex motion planning techniques. Active vision is used to track the gripper during servoing, which reduces the need for accurate camera calibration and allows the system to handle very large pose errors without losing vital visual information. Robust, continuous hand-eye calibration is achieved using a Kalman filter to estimate the pose of the gripper. Experimental work demonstrates the robustness and flexibility of the system in performing a complex task requiring grasping and assembling objects.

1 Introduction

In recent years, visual servoing has attracted a strong interest for a variety of applications [Corke and Hutchinson, 2000]. However, proposed systems usually focus on a single task, particularly *reach-to-grasp* control of robot manipulators [Chaumette and Malis, 2000; Wilson et al, 1996; Shen et al, 2000]. The motivation for this work was to develop a flexible visual servo framework for a humanoid robot, capable of performing a wide variety of tasks. The experimental upper-torso humanoid platform is shown in Figure 1.

Humanoid robots face a number of challenges when performing general manipulations in an unstructured environment. The possibility of large pose errors must be accommodated for *ad hoc* task specifications. Calculation of the optimal arm and gripper pose must take into account arm joint limits and mechanical constraints, such as those imposed when twisting a screw. Furthermore, motion planning in a cluttered human environment may involve obstacle avoidance in addition to reaching a target pose. This paper addresses the problem of large pose errors, but the other issues must also be considered when

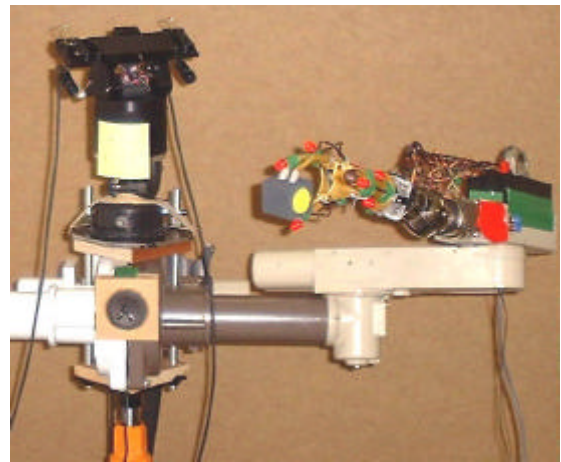


Figure 1. An anthropomorphic platform for visual servoing experiments.

developing a visual servoing framework so that high level grasp planning techniques [Cho et al, 1997] can be employed if necessary.

Visual servoing systems are commonly classed as using *image-based* [Hosoda and Asada, 1994; Kim et al, 2000; Rives, 2000], *position-based* [Wilson et al, 1996] or hybrid [Chaumette and Malis, 2000] techniques. In image-based visual servoing, control errors are measured on the image plane without any transformation to real space. The main attraction arises from avoiding complex 3D scene reconstruction, but this also introduces the possibility that resulting robot trajectories in Cartesian space are unpredictable. As the latter effect increases with pose error, obstacle avoidance planning is hindered. Steps towards addressing this issue are reported in recent work [Kim et al, 2000; Rives, 2000], including hybrid schemes [Chaumette and Malis, 2000]. However, observing mechanical constraints and obstacle avoidance are tasks most easily addressed in Cartesian space, making 3D reconstruction desirable. Furthermore, recent studies in applying biological reach-to-grasp to robotic systems suggest that humans use 3D structural cues rather than projected image features [Hu et al, 1999], and that human motions are planned in Cartesian space rather than joint space [Hauck et al,

1999]. Both properties are characteristic of position-based visual servoing, in which control errors are formed in a visually reconstructed Cartesian space. In this work, position-based servoing is adopted as a more flexible framework for addressing various planning and control issues encountered in humanoid manipulation tasks.

Position-based visual servoing is attributed with the following shortcomings: the accuracy of 3D scene reconstruction relies on camera calibration, servoing accuracy depends on calibration of the *hand-eye* transform between camera and robot frames, and objects may leave the field of view during servoing since motions are not planned on the image plane. In this work, the first two issues are addressed using active vision to track objects while servoing. Keeping image features near the centre of the image plane where non-linear effects (such as radial distortion) and sensitivity to projected depth are lowest reduces the reliance on camera calibration. Also, very large pose errors are accommodated without losing vital image features, using a robust strategy of tracking the gripper as it moves and measuring other objects when visible.

Various approaches to hand-eye calibration have been proposed, including the use of custom 3D calibration targets and on-line estimation. A calibration procedure using a custom target [Tsai and Lenz, 1989] is clearly not feasible in a humanoid system, which must provide continuous operation and therefore self-calibration. However, reported on-line calibration techniques based on optimization [Shen et al, 2000; Heikkela et al, 2000] require an initial estimate of the hand-eye transform. This paper describes model-based on-line calibration using a Kalman filter to estimate the pose of the gripper, which acts as a permanent calibration target. The technique is shown to provide continuous, robust calibration, without requiring an initial estimate. Other work [Wilson et al, 1996] has applied Kalman filtering to pose estimation in position-based visual servoing, but not in the context of hand-eye calibration.

Section 2 describes the overall architecture of the proposed position-based visual servoing system, and the various components in detail. In Section 3, image processing implementation issues are discussed, followed in Section 4 by a description of the real experiments performed using the platform shown in Figure 1. Section 5 draws conclusions and indicates future work.

2 System Overview

The system described in this work has been implemented experimentally on the upper-torso humanoid robot shown in Figure 1. Hardware consists of two Puma 260 arms with 1-DOF Otto Bock prosthetic hands, and off-the-shelf PAL cameras on a Biclops robotic head. Components are mounted anthropomorphically and are approximately to human scale. All processing is performed on a 450 MHz Pentium II desktop PC, running Linux.

Figure 2 illustrates the basic arrangement of coordinate frames relevant to visual servoing. Image features are measured on the stereo image planes I_L and I_R for reconstructing the Cartesian position of point targets p in the

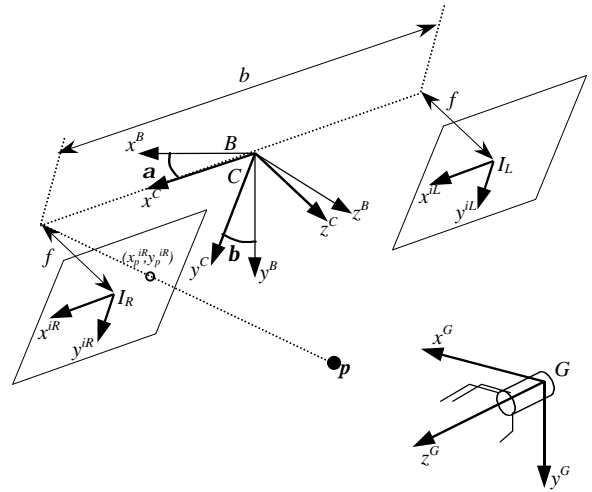


Figure 2. Coordinate frames.

camera frame C . Active vision allows C to pan and tilt by angles a and b with respect to a base frame B , so targets can be tracked and constrained within image boundaries. The gripper is described by frame G , and the basic task of visual servoing is to control the manipulator so G has some relative pose (position and orientation) to p .

Let $\mathbf{r}^R = [d^R, \mathbf{s}^R]^T$ represent the pose of the gripper with respect to the robot base frame R , where position $d^R = [X, Y, Z]^T$ describes translations along respective axes, and orientation $\mathbf{s}^R = [f, \mathbf{q}, \mathbf{y}]^T$ describes *roll*, *pitch* and *yaw Euler angles*, and let \mathbf{j} be the manipulator joint angles. For velocity control, the *Jacobian* matrix $\mathbf{J}(\mathbf{j})$ relates joint and pose velocities:

$$\dot{\mathbf{r}}^R = \mathbf{J}(\mathbf{j}) \dot{\mathbf{j}} \quad (1)$$

In classical position-based visual servoing, control signals $\dot{\mathbf{r}}^C$ are calculated in the camera frame and the resulting joint velocities are found using:

$$\dot{\mathbf{r}}^C = \mathbf{H}_R^C \mathbf{J}(\mathbf{j}) \dot{\mathbf{j}} \quad (2)$$

where \mathbf{H}_R^C is the homogeneous transform from R to C , determined by *hand-eye* calibration. However, by calculating control signals $\dot{\mathbf{r}}^G$ in the gripper frame, manipulator joint velocities can be obtained using:

$$\dot{\mathbf{r}}^G = \mathbf{H}_C^G \dot{\mathbf{r}}^C = [\mathbf{H}_G^R(\mathbf{j})]^{-1} \mathbf{J}(\mathbf{j}) \dot{\mathbf{j}} \quad (3)$$

where $\mathbf{H}_G^R(\mathbf{j})$ is the forward kinematic matrix for the manipulator. The right hand side of (3) is now completely specified by joint angles and known kinematics. Hand-eye calibration reduces to measuring \mathbf{H}_C^G , equivalent to the pose of the gripper in the camera frame. In essence, the gripper acts as a calibration target. Later sections develop a model-based framework for continuous estimation of the gripper pose and calculation of control errors.

2.1 Active Vision

Active vision allows the cameras to scan the workspace for a particular target and then track its motion. Thus, an initial estimate of the target location is not required, and losing sight of features during servoing is less likely to occur. Furthermore, active tracking allows features to be

driven to the centre of the image plane where calibration errors have least effect. This allows us to use a simple, linear camera model for stereo localization. Large pose errors are also accommodated; if the gripper and target are not simultaneously observable, a *look-then-move* strategy is adopted. In this case, the target is first measured, and the stored location is used to generate control errors while tracking and measuring only the gripper. Target measurements are dynamically resumed when the target again becomes visible.

The pan/tilt axes are driven using proportional velocity control. Consider tracking \mathbf{p} in Figure 2; let (x_p^{iR}, y_p^{iR}) and (x_p^{iL}, y_p^{iL}) be the coordinates of \mathbf{p} in the right and left image planes. Noting that the x and y axis of I_L, I_R and C are parallel, pan and tilt angular velocities, $\dot{\mathbf{a}}$ and $\dot{\mathbf{b}}$, are set proportional to the average position of \mathbf{p} :

$$\begin{aligned}\dot{\mathbf{a}} &= -\frac{1}{2}k_a(x_p^{iL} + x_p^{iR}) \\ \dot{\mathbf{b}} &= -\frac{1}{2}k_b(y_p^{iL} + y_p^{iR})\end{aligned}\quad (4)$$

where k_a and k_b are appropriately selected constants. This drives the average position of \mathbf{p} to the centre of the image.

2.2 Feature Localization

In this system, stereo reconstruction is used to recover the 3D position of point targets. As shown in Figure 2, the cameras are separated by baseline b , with parallel optical axes. Let $\mathbf{p}^C = (x_p^C, y_p^C, z_p^C)$ be the position of a point target in the camera frame, and let $(x_p^{iL}, y_p^{iL}), (x_p^{iR}, y_p^{iR})$ be the measured image coordinates of \mathbf{p}^C , as before. From the pin-hole camera model with focal length f , we have:

$$x_p^{iL} = \frac{f \cdot (x_p^C + \frac{b}{2})}{z_p^C}, \quad x_p^{iR} = \frac{f \cdot (x_p^C - \frac{b}{2})}{z_p^C} \quad (5)$$

$$y_p^{iL} = y_p^{iR} = \frac{f \cdot y_p^C}{z_p^C} \quad (6)$$

Measurement of y_p^{iL} and y_p^{iR} are likely to produce different values due to camera misalignment. To yield a unique solution, we calculate the average coordinate y_p^i as:

$$y_p^i = \frac{1}{2}(y_p^{iL} + y_p^{iR}) = \frac{f \cdot y_p^C}{z_p^C} \quad (7)$$

It is straightforward to show that the location of \mathbf{p}^C can be recovered from (5) and (7) as:

$$\mathbf{p}^C = \left(\frac{b(x_p^{iL} + x_p^{iR})}{2(x_p^{iL} - x_p^{iR})}, \frac{b(y_p^{iL} + y_p^{iR})}{2(x_p^{iL} - x_p^{iR})}, \frac{bf}{x_p^{iL} - x_p^{iR}} \right) \quad (8)$$

If we could ensure that the target and the gripper are always simultaneously observable, (8) provides sufficient information for position-based servoing. For large pose errors, however, the look-then-move strategy described earlier must be employed. Unfortunately, positions measured with disparate camera poses cannot be directly compared in the camera frame. The solution used here is to transform measurements from \mathbf{p}^C in camera frame to \mathbf{p}^B in a stationary base frame B . The homogeneous transform \mathbf{H}_C^B is parameterized by \mathbf{a} and \mathbf{b} , the pan and tilt angles of the active vision head:

$$\begin{aligned}\mathbf{p}^B &= \mathbf{H}_C^B \mathbf{p}^C \\ \mathbf{H}_C^B &= \mathbf{H}_C^B(\mathbf{a}, \mathbf{b}) = \begin{bmatrix} C_a & S_a S_b & S_a C_b & 0 \\ 0 & C_b & -S_b & 0 \\ -S_a & C_a S_b & C_a C_b & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\end{aligned}\quad (9)$$

where $S_x = \sin(x)$, $C_x = \cos(x)$.

2.3 Model-based Gripper Pose Estimation using an Extended Kalman Filter

The pose of the gripper with respect to the base frame is described by the homogeneous transform matrix \mathbf{H}_G^B , parameterized by the six spatial degree of freedom:

$$\begin{aligned}\mathbf{H}_G^B &= \mathbf{H}_G^B(\mathbf{r}) = \mathbf{H}_G^B(X, Y, Z, \mathbf{f}, \mathbf{q}, \mathbf{y}) \\ &= \begin{bmatrix} C_q C_y & S_f S_q C_y - C_q S_y & C_f S_q C_y + S_f S_y & X \\ C_q S_y & S_f S_q S_y + C_q C_y & C_f S_q S_y - S_f C_y & Y \\ -S_q & S_f C_q & C_f C_q & Z \\ 0 & 0 & 0 & 1 \end{bmatrix}\end{aligned}\quad (10)$$

The gripper pose is estimated by matching a model of the gripper to observed images. The model is specified by a set of n_g point features \mathbf{g}_i^G , $i=1,2,\dots,n_g$, in the gripper frame. Importantly, the matching process must handle the possibility that one or more model features are unobservable due to occlusion (by the hand or otherwise). Figure 3 shows the model currently used, with $n_g=7$. The model includes the grasp width \mathbf{e} as a single degree of freedom, which can be read from the gripper encoder.

Kalman filtering provides an optimal framework for matching the gripper model to observed features in the presence of occlusions. The Kalman filter is a well established technique for estimating the state of a known linear system from noisy measurements. If the noise is normally distributed (a reasonable assumption in this case), the state estimate is optimal with respect to error variance. The *extended* Kalman filter applies when measurements are non-linearly related to the system state, which is the case here. A detailed development of Kalman filter theory, and relevant equations, may be found in [Shalom and Li, 1993].

Assuming the motion of the gripper to be smooth, a constant velocity dynamic model is used to describe system dynamics. Thus, the state vector \mathbf{x}_k at time t_k is composed of each degree of freedom and its first derivative:

$$\mathbf{x}_k = [\mathbf{r}_k, \dot{\mathbf{r}}_k]^T \quad (11)$$

and the state dynamic equations are of the form:

$$\begin{aligned}\mathbf{r}_k &= \mathbf{r}_{k-1} + \mathbf{D}t_k \cdot \dot{\mathbf{r}}_{k-1} \\ \dot{\mathbf{r}}_k &= \dot{\mathbf{r}}_{k-1}\end{aligned}\quad (12)$$

where $\mathbf{D}t_k = (t_k - t_{k-1})$ is the inter-sample period. The state error covariance matrix \mathbf{Q} , which compensates for unmodeled dynamics, is a diagonal matrix (assuming independent pose parameters), with elements fixed to a constant estimate of the variance of random system dynamics.

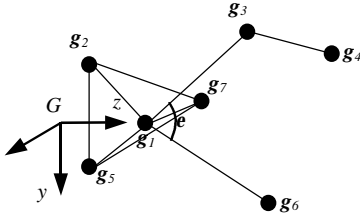


Figure 3. Articulated hand model.

The measurement vector \mathbf{m}_k contains the measured position $\hat{\mathbf{g}}_{i,k}^B$ of each model point \mathbf{g}_i^B in B , using (8) and (9):

$$\mathbf{m}_k = [\hat{\mathbf{g}}_{i,k}^B]^T, \quad i=1,2,\dots,n_g \quad (13)$$

The measurement error covariance matrix \mathbf{R}_k is a diagonal matrix with elements set to an estimate of the position measurement error variance. If the position of any model point is not measurable due to occlusion, the corresponding measurement error in \mathbf{R}_k is set to a high value, which effectively removes that point from the state estimate.

The final element of the extended Kalman filter is the measurement function $\mathbf{M}(\mathbf{x}_k)$ which relates the state vector to the measurement vector. In this case, $\mathbf{M}(\mathbf{x}_k)$ transforms the model points \mathbf{g}^G from the gripper frame to the base frame according to the estimated pose at t_k and (10):

$$\mathbf{m}_k = \mathbf{M}(\mathbf{x}_k) = [\mathbf{H}_G^B(\mathbf{r})\mathbf{g}_i^G]^T, \quad i=1,2,\dots,n_g \quad (14)$$

During servoing, the state vector \mathbf{x}_0 is first initialized by setting translation parameters to the measured location of any gripper point, with remaining elements set to zero. For the k^{th} captured frame, visible gripper features are measured in the base frame and passed to the filter as a new measurement vector \mathbf{m}_k and error covariance matrix \mathbf{R}_k . The Kalman filter equations are used to update the state vector \mathbf{x}_k and a new pose estimate \mathbf{r}_k is recovered.

2.4 Visual Servo Control Law

The desired pose of the gripper for a particular manipulation is task-specific; to grasp an object using a *precision grip*, the object must be positioned between the tips of the thumb and forefinger, while mating a grasped object to another object requires positioning the gripper so the objects meet. Once the desired pose of the gripper has been determined, the position and orientation are controlled independently. Considering position first, let \mathbf{p}^G and \mathbf{r}^G be the current and desired target position of some point in the gripper frame. Using (9) and (10), the gripper position error \mathbf{Dd} is:

$$\mathbf{Dd} = \mathbf{p}^G - \mathbf{r}^G = (\mathbf{H}_G^B)^{-1} \mathbf{H}_C^B \mathbf{p}^C - \mathbf{r}^G \quad (15)$$

Now, let \mathbf{s}^B and \mathbf{d}^B represent Euler angles describing the current and desired gripper orientation in the base frame. Note that the desired orientation may be specified in any frame (such as a frame attached to an object); B is used here for convenience. The orientation error in the gripper frame is calculated by determining the angle of rotation \mathbf{g} about an axis \mathbf{u}^G that rotates the gripper to the desired pose. Rotations represented in this way are most

efficiently manipulated using *quaternions*, as described in [Taylor, 1982]. The relationship between quaternion \mathbf{q} and (\mathbf{g}, \mathbf{u}) is:

$$\mathbf{q} = (s, \mathbf{v}) = \left(\cos \frac{\mathbf{g}}{2}, \mathbf{u} \sin \frac{\mathbf{g}}{2} \right) \quad (16)$$

A quaternion can also be constructed from Euler angles:

$$\mathbf{q} = (C_f C_q C_y + S_f S_q S_y, S_f C_q C_y - C_f S_q S_y, C_f S_q C_y + S_f C_q S_y, C_f C_q S_y - S_f S_q C_y) \quad (17)$$

where $S_x = \sin(x)$, $C_x = \cos(x)$. Using (17), let ${}^B \mathbf{q}_d$ and ${}^B \mathbf{q}_s$ be the quaternion equivalent of \mathbf{s}^B and \mathbf{d}^B . These rotations are concatenated by complex multiplication to give the desired orientation relative to the current gripper pose \mathbf{q}_d^s :

$$\mathbf{q}_d^s = (\mathbf{q}_s^B)^{-1} \mathbf{q}_d^B = (s_s s_d - (-\mathbf{v}_s) \cdot \mathbf{v}_d, s_s \mathbf{v}_d + s_d (-\mathbf{v}_s) + (-\mathbf{v}_s) \times \mathbf{v}_d) \quad (18)$$

Finally, the inverse of (16) is applied to \mathbf{q}_d^s to calculate $(\mathbf{g}, \mathbf{u}^G)$, the error angle and axis of rotation between current and desired gripper orientations, expressed in the gripper frame. The control error \mathbf{Ds} is calculated as:

$$\mathbf{Ds} = \mathbf{g} \mathbf{u}^G \quad (19)$$

The current implementation uses proportional velocity control to drive the gripper to a target pose, but other control laws are easily accommodated. The control signal in (3) is calculated as $\dot{\mathbf{r}}^G = (c_1 \mathbf{Dd}, c_2 \mathbf{Ds})$, for constants c_1 and c_2 , and passed to the PUMA controller which determines joint velocities and provides motion control.

3 Implementation and Image Processing

As described earlier, all processing is performed on a desktop PC. Frame-rate vision (25Hz PAL) is achieved by restricting images to 320×240 pixels with 16-bit colour. Coloured point features on target objects and the gripper and extracted using predefined filters implemented as lookup tables. Conventional binary image processing and sub-pixel centroid calculations are used to locate features. Performance is improved by restricting processing to *windows of interest*, generated by projecting estimated point feature locations onto the image planes using (5) and (6). Additionally, the Pentium MMX architecture is exploited where possible to perform parallel pixel processing. Combining these techniques, the system measures image features and generates control commands at video frame-rate.

The point features \mathbf{g}_i on the gripper are implemented as individually controlled red LEDs. During initial pose estimation, each LED is flashed and measured in sequence. This allows the gripper model points to be associated unambiguously with measured features, but requires greater than n_g frames to capture all LEDs. After Kalman filter initialization, the LEDs are activated and measured simultaneously, providing a new pose estimate every frame. The association problem is solved by projecting the estimated gripper pose onto the image planes (using (5) and (6)), and associating the projected points with measured features using robust pattern matching. Unmatched model points are assumed to be occluded by the gripper.

4 Experimental Results

Real experiments were performed on the humanoid robot to demonstrate the robustness and flexibility of the proposed system. Since gripper pose estimation is central to the visual servoing approach developed here, the first experiment was designed to demonstrate the accuracy and robustness of model-based pose estimation to track large changes in pose. The gripper was manually positioned with the z axis aligned to the x axis of the base frame and driven in a screw trajectory, ie. translated along the gripper z -axis at about 45 mm/s, and rotated around the same axis at about 0.35 rad/s. Such a trajectory caused each model point on the gripper to suffer occlusion at some time. With active vision used to track the gripper, the Kalman filter produced pose estimates over a 7 s period.

Figure 4 shows the a stereo view from the robot during the tracking sequence, with a wire-frame overlay of the estimated gripper pose. Figure 5 shows the pose parameters estimated in the base frame, with a linear ramp in one translational and one rotational degree of freedom as expected for a screw trajectory. While it was not possible to make absolute pose measurements, a least-squares linear fit to the ramping components provides a quantitative approximation of tracking accuracy. The linear fit gives a standard error of about 2 mm and 0.04 rad from the expected trajectory, which indicates that model-based Kalman filtering, with active vision, can accurately and robustly track rapid changes in the pose of the gripper.

The second experiment employs all the elements of the system to perform a manipulation task in the presence of a large initial pose error. The task is to locate and stack three randomly placed wooden blocks. Each block has an identical spot painted on the top and bottom face, with the three blocks coloured green, blue and yellow. The steps carried out to complete the task are as follows:

1. Locate and grasp the blue block by visually servoing to the spot on the top face.
2. Rotate the block so the bottom spot is visible, and measure the position of this spot in the gripper frame.
3. Locate the green block, and visually servo the gripper so the bottom of the blue block is placed on top of the green block. Release the blue block.
4. Repeat for the yellow block.

The experiment is performed with a large initial pose error, such that the blocks and the gripper are not visible in the same frame. During the initial period of visual servoing, the system tracks the gripper pose and uses the stored block position to generate control errors. When the gripper is sufficiently close to the block, both the gripper and block are measured in the control loop. Manipulator speed is kept sufficiently low that dynamic effects can be ignored (this will be the subject of future work).

The sequence of images in Figure 6 shows the humanoid executing the stacking task, with approximate progress times indicated. The first image illustrates the initial pose error, and the final image shows the stacked blocks after completion. This successful result demonstrates a level of flexibility suited to humanoid tasks; grasping, examining

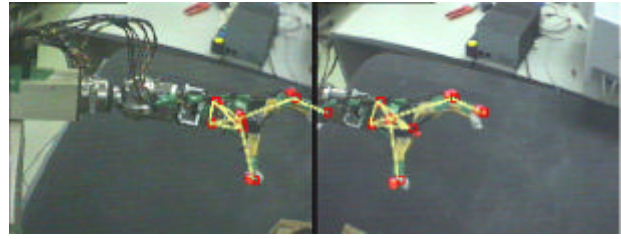


Figure 4. Stereo view from the robot showing gripper and estimated gripper pose overlaid as a wireframe model.

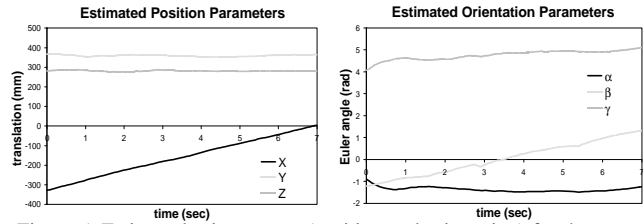


Figure 5. Estimated gripper pose (position and orientation) for the screw trajectory tracking experiment.

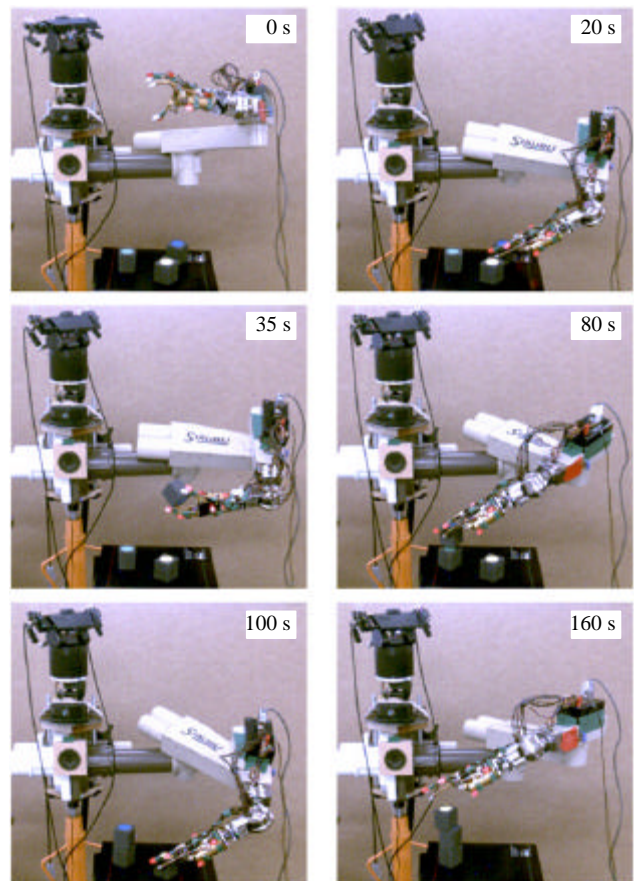


Figure 6. Sequence showing the humanoid perform a stacking task in the presence of a large initial pose error, with progress times indicated. A video sequence of this demonstration may be found at: <http://www.ecse.monash.edu.au/centres/IRRC/ServoDemoQCIF.mpg>.

and assembling objects are all performed in the presence of a large initial pose error.

5 Conclusions and Future Directions

This paper presents a flexible position-based visual servoing framework to address the issues involved in performing general manipulation tasks by a humanoid robot. Solutions to classical drawbacks of position-based visual servoing are proposed and experimentally validated. Active vision removes the need for accurate camera calibration and handles large pose errors (even when the gripper and target are not simultaneously observable) without losing visual feedback during servoing. Model-based gripper pose estimation using a Kalman filter provides robust hand-eye self-calibration, without requiring an initial estimate. Experimental work verifies the robustness and flexibility of the system, allowing complex tasks involving grasping and assembling objects to be performed.

Future work can proceed in a number of directions. Modifying the control law to account for robot dynamics will lead to performance improvements, as will adding vergence control to active vision. Incorporating more sophisticated scene analysis and motion planning will allow the humanoid to operate reliably in a cluttered environment. More complex manipulations may be possible by cooperatively servoing both arms on the humanoid platform. Finally, a long term goal is to have task specifications interpreted visually through human interaction.

Acknowledgement

This project was funded by the Special Monash University Research Fund as part of the *Humanoid Robotics: Perception, Intelligence and Control* project at IRRC.

References

- [Chaumette and Malis, 2000] F. Chaumette and E. Malis. 2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoing. In *Proc. IEEE Int. Conf. on Robotics & Automation*, pp. 630-635, 2000.
- [Cho et al, 1997] K.R. Cho, Y.K. Hwang, M.S. Kim, C.W. Lee and J.B. Song. Determination of Optimal Arm and Hand Configurations for Grasping Objects by Humanoid Robots and Avatars. In *Proc. IEEE Conf. on Comput. Cybernetics & Simulation*, pp. 3279-3284, 1997.
- [Corke and Hutchinson, 2000] P.I. Corke and S.A. Hutchinson. Real-Time Vision, Tracking and Control. In *Proc. IEEE Conf. on Robotics & Automation*, pp. 622-629, 2000.
- [Hauck et al, 1999] A. Hauck, M. Sorg, G. Faber, T. Schenk. What Can Be Learned from Human Reach-To-Grasp Movements from the Design of Robotic Hand-Eye Systems? In *Proc. IEEE Int. Conf. on Robotics & Automation*, pp. 2521-2526, 1999.
- [Heikkela et al, 2000] T. Heikkela, M. Sallinen, T. Matsu-shita, F. Tomita. Flexible Hand-Eye Calibration for Multi-Camera Systems. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, pp. 2292-2297, 2000.
- [Hosoda and Asada, 1994] K. Hosoda and M. Asada. Versatile Visual Servoing without Knowledge of True Jacobian. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, pp. 186-191, 1994.
- [Hu et al, 1999] Y. Hu, R. Eagleson, M.A. Goodale. Human Visual Servoing for Reaching and Grasping: The Role of 3-D Geometric Features. In *Proc. IEEE Int. Conf. on Robotics & Automation*, pp. 3209-3216, 1999.
- [Kim et al, 2000] H. Kim, J. Cho and I. Kkweon. A novel image-based control-law for the visual servoing system under large pose error. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, 263-267, 2000.
- [Rives, 2000] P. Rives. Visual Servoing Based on Epipolar Geometry. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, 2000, pp. 602-607.
- [Shalom and Li, 1993] Y. Bar-Shalom, X-R. Li. *Estimation and Tracking: Principles, Techniques and Software*. Artech House, Boston, 1993.
- [Shen et al, 2000] Y. Shen, Y. Lui and K. Li. Asymptotic Position Control of Robot Manipulators Using Uncalibrated Visual Feedback. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, pp.435-440, 2000.
- [Taylor, 1982] R.H. Taylor. Planning and Execution of Straight Line Manipulator Trajectories. In *Robot Motion: Planning and Control*, M. Brady et al, ed., Cambridge, Mass: The MIT Press, pp. 265-286, 1982.
- [Tsai and Lenz, 1989] R.Y. Tsai, R.K. Lenz. A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration. *IEEE Trans. on Robotics & Automation*, 5(3):345-358, 1989.
- [Wilson et al, 1996] W. Wilson, C. Williams Hulls and G. Bell. Relative End-Effector Control Using Cartesian Position Based Visual Servoing. In *IEEE Trans. on Robotics & Automation*, vol. 12, no. 5, pp. 684-696, 1996.