1995 American Control Conference,
Seattle, Washington, June 21-23, 1995.

Improved Force Control Through Visual Servoing

Bradley J. Nelson*
J. Daniel Morrow*
Pradeep K. Khosla**

*The Robotics Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh PA 15213-3890

**Dept. of Elec. and Comp. Eng.
and The Robotics Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh PA 15213-3890

WWW URL http://www.cs.cmu.edu:8001/afs/cs/user/bnelson/www/home.html

Abstract

Force controlled manipulation is a common technique for compliantly contacting and manipulating uncertain environments. Visual servoing is effective for reducing alignment uncertainties between objects using imprecisely calibrated camera-lens-manipulator systems. These two types of manipulator feedback, force and vision, represent complementary sensing modalities; visual feedback provides information over a relatively large area of the workspace without requiring contact with the environment, and force feedback provides highly localized and precise information upon contact. This paper presents three different strategies which combine force and vision within the feedback loop of a manipulator, traded control, hybrid control, and shared control. A discussion of the types of tasks that benefit from the strategies is included, as well as experimental results which show that the use of visual servoing to stably guide a manipulator simplifies the force control problem by allowing the effective use of low gain force control with relatively large stability margins.

1. Introduction

Roboticists have traditionally considered force feedback to be the sensing modality most relevant to manipulation. This is because of the need for highly accurate information on the relative positions of objects and on the nature of contact forces between objects being manipulated. More recently, many researchers have realized the benefits of using visual servoing techniques to reduce alignment uncertainties between objects using imprecisely calibrated camera-lens-manipulator systems. By properly combining these two sensors within the same high bandwidth feedback loop, the benefits of each sensing mode can be simultaneously realized throughout a manipulation task.

Combining high bandwidth force and vision feedback within a manipulator feedback loop presents many challenges. The camera-lens system, the manipulator, and the force sensor must all be properly modeled. Because of the different nature of force and vision feedback, traditional sensor integration techniques are not appropriate, and separate control strategies specific to particular classes of tasks must be derived. The most common method of implementing force control is through a wrist force sensor and a stiff manipulator position control loop. Because the force sensor measures all forces (inertial, gravitational, and tactile), the inertial coupling of the end-effector mass beyond the sensor introduces a fundamental problem when sharing control between force and vision sensors. When the vision system commands motions, the resulting accelerations can result in inertial forces which cause unstable excitations of the force control system. In order to compensate for unstable excitations, it is necessary to develop robust strategies for avoiding the excitations.

In this paper, we propose and experimentally investigate three different control strategies which combine force and vision in the same manipulator feedback loop. Traded and hybrid control avoid force instabilities due to inertial coupling by specifying that force and visual servoing do not occur simultaneously along the same direction. A novel shared control strategy has been developed that does not require this assumption. Our proposed controller effectively compensates for inertial coupling while responding to both vision and force feedback along the same direction. Depending on the type of manipulation task and the accuracy of the visual feedback, these three strategies may all be used during the task, or a single strategy may be capable of performing an entire task. Experimental results from the three control strategies are presented, along with a discussion of the applicability of each of the strategies to different types of tasks. Results show that significant performance improvements in task execution can be realized by our proposed controllers. In particular, the use of visual servoing simplifies the force control problem by allowing the effective use of low gain force control with relatively large stability margins. The resulting system has the capability of approaching a rigid surface with a high velocity and stably impacting the surface while minimizing impact forces and avoiding bounce between surfaces.

2. Previous Work

2.1. Force Servoing

Robotic force control has been studied since the 1950's. A survey on the history of force control can be found in [32]. Active impedance control has been suggested as the most general form of force control [11], however, difficulties in programming impedance controlled manipulators have resulted in very limited use of this strategy. Hybrid control [26] separates position control and force control into two separate control loops that operate in orthogonal directions and is the most common force control implementation.

One of the most important issues in force control is maintaining manipulator stability [34]. Force controllers must be properly formulated and tuned in order to maintain stability. This can be difficult, particularly during initial contact between stiff surfaces. During impact, another important issue is the generation of large impact forces. In [29], an effective impact strategy is presented based on a proportional gain explicit force controller with a feed-forward signal and negative gains. The gains for the controller were chosen using a fourth order model of the arm, sensor, and environment in which a frictionless arm was assumed (experiments were conducted with the CMU Direct-Drive Arm II). Although high impact velocities were achieved (75cm/s), large impact forces were also generated (90N). This illustrates a typical problem exhibited by all force control strategies during impact with rigid objects, for example [17][2][5][12][16][25][14][34]—high impact velocities, manipulator stability, low impact forces, and quickly achieving the desired force are all contradictory system requirements.

2.2. Visual Servoing

Although previous researchers had considered fast visual feedback for guiding manipulator motion, for example [27], the visual servoing field was first well defined by Weiss [30]. Since this work, 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][4][7][23][6][8][10][33]. Static camera visual servoing guides manipulator motion based on feedback from a camera observing the end-effector [18][20][3][9]. Most of this past work has been with monocular systems, though recently stereo systems have been used for visual servoing [19][9][13]. 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.

2.3. Force and Visual Servoing

One of the first papers to mention the benefits of combining high bandwidth visual and force feedback is [27]. They implemented a 0.1Hz visual servoing scheme and mentioned force servoing, but a lack of computational resources hampered their effort, and many of the issues of combining the two sensing modalities went unnoticed. In [15], visual servoing of 2 Hz was used to align a wrench with a bolt before a compliant wrenching operation is performed. Again, vision and force were not explicitly combined, and the issues concerning their integration remained unaddressed.

In this paper we consider the integration of high bandwidth visual feedback with high bandwidth force feedback within the same manipulator feedback loop. An important contribution of this work is that we show how vision can be used to greatly simplify the stability problem by allowing the effective use of low gain force control with high-friction manipulators (a Puma 560). Since the stability of low gain force control is much easier to maintain, the use of force feedback during manipulator fine motion is more easily realized because simple force control strategies can be used without the need for high order models of the arm, sensor, and environment for choosing stable controller gains.

3. Visual Servoing

The visual servoing component of our system is an extension of the controlled active vision paradigm [23], originally created for monocular eye-in-hand visual servoing. In this section we extend controlled active vision to static camera visual servoing. A model of the camera-lens-manipulator system is presented, followed by a derivation of the visual servoing control law.

3.1. System Model

In formulating the visual servoing component of our system, the Jacobian mapping from task space to sensor space is needed. We desire a Jacobian for the camera-lens system of the form

$$\dot{\boldsymbol{x}}_S = \boldsymbol{J}_v(\phi)\dot{\boldsymbol{X}}_T \qquad (1)$$

where $\dot{\boldsymbol{x}}_S$ is a velocity vector in sensor space; $\boldsymbol{J}_v(\phi)$ 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; and $\dot{\boldsymbol{X}}_T$ is a velocity vector in task space.

A pinhole camera model is used to model the camera-lens system. The camera frame {C} is placed at the focal point of the lens as shown in Figure 1. A feature on an object at ${}^C\boldsymbol{P}$ in the camera frame projects onto the camera's image plane at

$$x_S = \frac{fX_C}{s_x Z_C} \qquad y_S = \frac{fY_C}{s_y Z_C} \qquad (2)$$

where (x_S, y_S) are the image coordinates of the feature, 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 CCD array. This model assumes that $|Z_C| \gg |f|$.

A feature on a manipulated object is observed on the image plane with respect to the camera frame at ${}^C\boldsymbol{P}$, but is controlled with respect to the manipulator's task frame at ${}^T\boldsymbol{P}$. The rotation of the task frame {T} with respect to {C} ${}^C_T\boldsymbol{R}$ is assumed to be approximately known. The relationship between the velocity of ${}^C\boldsymbol{P}$ with respect to task frame velocities is written as

$${}^C\dot{\boldsymbol{P}} = {}^C_T\boldsymbol{R}\left({}^T\boldsymbol{V} + {}^T\dot{\boldsymbol{P}} + {}^T\Omega \times {}^T\boldsymbol{P}\right) \qquad (3)$$

where ${}^T\boldsymbol{V} = \left[\dot{X}_T\ \dot{Y}_T\ \dot{Z}_T\right]^{\mathrm{T}}$ and ${}^T\Omega = \left[\omega_{X_T}\ \omega_{Y_T}\ \omega_{Z_T}\right]^{\mathrm{T}}$ are the translational and rotational velocities of the task frame with respect to itself. By differentiating (2) and combining this result with (3) the Jacobian transformation for a single feature from task space to sensor space can be written in the form

$$\begin{bmatrix}\dot{x}_S\\ \dot{y}_S\end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_C} & 0 & -\frac{x_S}{Z_C} & -\frac{Y_T x_S}{Z_C} & \left[\frac{fZ_T}{s_x Z_C} + \frac{X_T x_S}{Z_C}\right] & -\frac{fY_T}{s_x Z_C} \\ 0 & \frac{f}{s_y Z_C} & -\frac{y_S}{Z_C} & -\left[\frac{fZ_T}{s_y Z_C} + \frac{Y_T y_S}{Z_C}\right] & \frac{X_T y_S}{Z_C} & \frac{fX_T}{s_y Z_C} \end{bmatrix} \begin{bmatrix}\dot{X}_T\\ \dot{Y}_T\\ \dot{Z}_T\\ \omega_{X_T}\\ \omega_{Y_T}\\ \omega_{Z_T}\end{bmatrix} \qquad (4)$$

For the above form of the Jacobian, the parameters of the Jacobian are given by $\phi = (f, s_x, s_y, x_S, y_S, Z_C, X_T, Y_T, Z_T)$. A complete derivation of (4) can be found in [21].

At least three features on an object must be tracked in order to control both the position and orientation of the object in three dimensions. For n feature points, the Jacobian is of the form

$$\boldsymbol{J}_v(\phi) = \left[\boldsymbol{J}_1(\phi)\ \dots\ \boldsymbol{J}_n(\phi)\right]^{\mathrm{T}} \qquad (5)$$

where $\boldsymbol{J}_i(\phi)$ is the Jacobian matrix for each feature given by the 2x6 matrix in (4).

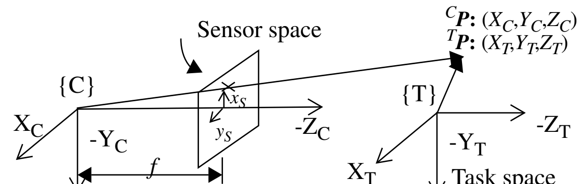

Figure 1. Task frame-camera frame definitions.

3.2. Visual Servoing Controller

The state equation for the visual servoing system is created by discretizing (1) and rewriting the discretized equation as

$$\boldsymbol{x}(k+1) = \boldsymbol{x}(k) + T\boldsymbol{J}_v(k)\boldsymbol{u}(k) \qquad (6)$$

where M is the number of features being tracked; $\boldsymbol{x}(k) \in R^{2M}$; T is the sampling period of the vision system; and $\boldsymbol{u}(k) = \left[\dot{X}_T\ \dot{Y}_T\ \dot{Z}_T\ \omega_{X_T}\ \omega_{Y_T}\ \omega_{Z_T}\right]^{\mathrm{T}}$ the manipulator end-effector velocity. The Jacobian $\boldsymbol{J}_v(\phi)$ is written as $\boldsymbol{J}_v(k)$ in order to emphasize its time varying nature due to the changing feature coordinates on the image plane. The intrinsic parameters of the camera-lens system are constant for the experimental results to be presented.

The control objective of the visual tracking system is to control end-effector motion in order to place the image plane coordinates of features on the target at some desired position. The desired image plane coordinates could be constant or changing with time. The control strategy used to achieve the control objective is based on the minimization of an objective function that places a cost on errors in feature positions and a cost on providing control energy.

$$F(k+1) = [\boldsymbol{x}(k+1) - \boldsymbol{x}_D(k+1)]^{\mathrm{T}} \boldsymbol{Q}\, [\boldsymbol{x}(k+1) - \boldsymbol{x}_D(k+1)] + \boldsymbol{u}^{\mathrm{T}}(k)\boldsymbol{L}\boldsymbol{u}(k) \qquad (7)$$

This expression is minimized with respect to the current control input $\boldsymbol{u}(k)$. The end result yields the following expression for the control input

$$\boldsymbol{u}(k) = -\left(T\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}T\boldsymbol{J}_v(k) + \boldsymbol{L}\right)^{-1} T\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\, [\boldsymbol{x}(k) - \boldsymbol{x}_D(k+1)] \qquad (8)$$

The weighting matrices $\boldsymbol{Q}$ and $\boldsymbol{L}$ allow the user to place more or less emphasis on the feature error and the control input. Their selection effects the stability and response of the tracking system. The $\boldsymbol{Q}$ matrix must be positive semi-definite, and $\boldsymbol{L}$ must be positive definite for a bounded response. Although no standard procedure exists for choosing the elements of $\boldsymbol{Q}$ and $\boldsymbol{L}$, general guidelines can be found in [24], along with system models and control derivations that account for system delays, modeling and control inaccuracies, and measurement noise.

3.3. Feature Tracking

The measurement of the motion of the features on the image plane must be done continuously and quickly. The method used to measure this motion is based on an optical flow technique called Sum-of-Squares-Differences (SSD). A more complete description of the algorithm and its implementation can be found in [20].

4. Visual and Force Servoing

Visual servoing can be combined with force control for many different types of robotic tasks. For autonomous manipulation, visual servoing brings objects near one another and into proper alignment, and force servoing ensures compliant contact and accomplishes the final parts mating. For teleoperation, a mouse cursor can be used to guide the manipulator by a teleoperator observing live video images of the robot. Visual servoing ensures that the teleoperator's desired manipulator trajectory is accurately followed by controlling the manipulator's end-effector position based on reference inputs from the cursor position. Force sensing is used to contact the environment compliantly. Feedback from other devices, for example a trackball, may also be used to guide the manipulator. Three different types of controllers $\mathbf{C}(t)$ have been implemented and tested—a shared, a traded, and a hybrid controller. A block diagram of the generalized force/vision control system is shown in Figure 2.

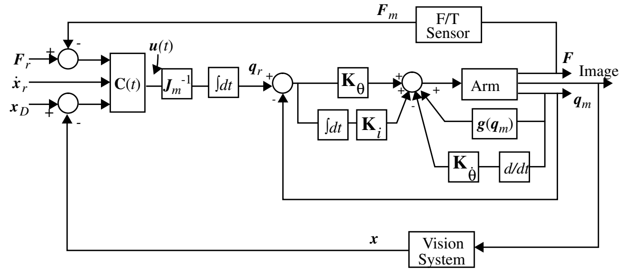

Figure 2. Generalized force/vision controller with inner loop PID controllers and gravity compensation.

The force control portion of our proposed visual/force servoing strategies is based on past work in hybrid force control. The implemented force control scheme is a combination of hybrid force/position control [26] and damping force control [31], resulting in a hybrid force/velocity control scheme. Because the dynamics, particularly friction, of the laboratory robot (a Puma 560) are difficult to accurately model, a simple cartesian control scheme is used in which a manipulator Jacobian inverse converts cartesian velocities to joint velocities, which are then integrated to joint reference positions. High servo rate (500Hz) PID controllers are implemented for each joint in order to follow joint trajectories which achieve the desired cartesian motion.

4.1. Traded Control

There are two main benefits to using visual servoing and force servoing together. First, visual servoing is effective in significantly improving manipulator performance upon impact by allowing the effective use of low gain force controllers. Second, visual servoing is effective in reducing alignment uncertainties between objects using imprecisely calibrated camera-lens-manipulator systems. A traded control strategy realizes both of these benefits. For traded control, the control law given by $\mathbf{C}(t)$ in Figure 2 is written as

$$\boldsymbol{u}(k) = \begin{cases} -\left(\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\boldsymbol{J}_v(k) + \boldsymbol{L}\right)^{-1}\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\,[\boldsymbol{x}(k) - \boldsymbol{x}_D(k+1)]\,, & \|\boldsymbol{x}_D(k) - \boldsymbol{x}(k)\| > \varepsilon \\ \boldsymbol{S}_F\boldsymbol{G}_F(\boldsymbol{F}_r(k) - \boldsymbol{F}_m(k))\,, & \|\boldsymbol{x}_D(k) - \boldsymbol{x}(k)\| \le \varepsilon \end{cases} \qquad (9)$$

where $T\boldsymbol{J}_v$ is rewritten as $\boldsymbol{J}_v$ to simplify equations without loss of generality. During traded control, manipulator motion is first controlled by visual feedback. The controller then switches to force control after visual servoing approaches sufficiently near a surface. The desired state of the manipulator on the image plane $\boldsymbol{x}_D$ represents a state near the surface to be contacted. $\boldsymbol{S}_F$ is a matrix that selects axes along which force control will be applied, $\boldsymbol{G}_F$ is a matrix of force control gains, and $\boldsymbol{F}_r$ and $\boldsymbol{F}_m$ represent reference and measured forces. The benefit of this strategy is that the surface can be quickly approached under visual servoing. Manipulator control is switched to low gain force control which can be used effectively when the end-effector is near the surface. Stable impact with a surface can then be achieved, large contact forces can be minimized, and bounce can be avoided.

With simple force damping control alone, a manipulator can easily become unstable during impact, unless force gains are tuned to extremely low values resulting in unacceptably slow motion during the approach phase of the task. Because of this, most manipulation strategies use a guarded move to initiate contact with a surface. During a guarded move, surfaces are approached under position control while the force sensor is monitored. If the sensed force exceeds a threshold, motion is immediately stopped and a force control strategy can then be invoked. The main limitation of this strategy is that high contact forces can result unless the effective mass of the manipulator is low so that the end-effector can be quickly stopped before contact forces increase significantly. In Section 6.1, experimental results are presented which compare these three impact strategies, their stability, and the impact forces created by each strategy.

4.2. Hybrid Control

An obvious technique for combining visual servoing with other types of feedback, for example position feedback [3], is in a hybrid control scheme. This type of scheme, however, is limited because it ignores much of the information provided by visual feedback. The proper use of vision in the feedback loop provides information relevant to motion along all six degrees of freedom with which an object can move. Hybrid control allows visual servoing only in those directions that are orthogonal to directions along which other types of feedback is used. Despite this drawback there are specific classes of tasks which fall naturally into a hybrid control scheme. One useful application is the grasping of electrical lines through teleoperation [22]. For this task, hybrid control could maintain the alignment of the gripper with the wire to be grasped, and a teleoperator or other sensor, such as a force sensor or an optical switch, could signal when the line had been reached, much as a guarded move is used to indicate when contact is made under position control.

We have used this scheme to demonstrate remote teleoperated grasping of objects in which a supervisor is presented with live video input from a single camera mounted on the manipulator's gripper (Figure 3). Visual servoing maintains gripper alignment, while the supervisor controls the velocity with which the object is approached, thus a hybrid control scheme is used. Initial experimental results demonstrate that the task can be completed much more quickly and with a higher probability of success when using hybrid visual servoing than when the teleoperator must control all degrees of manipulator motion. An important advantage of this strategy is that the camera-lens-manipulator system need not be precisely calibrated.

The hybrid control law we use is of the form

$$\boldsymbol{u}(k) = -\left(\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\boldsymbol{J}_v(k)+\boldsymbol{L}\right)^{-1}\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\,[\boldsymbol{x}(k)-\boldsymbol{x}_D(k+1)] + \boldsymbol{S}_r\dot{\boldsymbol{x}}_r + \boldsymbol{S}_F\boldsymbol{G}_F(\boldsymbol{F}_r(k)-\boldsymbol{F}_m(k)) \qquad (10)$$

$\boldsymbol{S}_v$, $\boldsymbol{S}_r$, and $\boldsymbol{S}_F$ are diagonal selection matrices, the elements of which are 0 or 1, and which must sum to the identity matrix, *i.e.* $\boldsymbol{S}_v+\boldsymbol{S}_r+\boldsymbol{S}_F = \boldsymbol{I}$. The vector $\dot{\boldsymbol{x}}_r$ represents some desired end-effector reference velocity due, for example, to trackball input from a teleoperator.

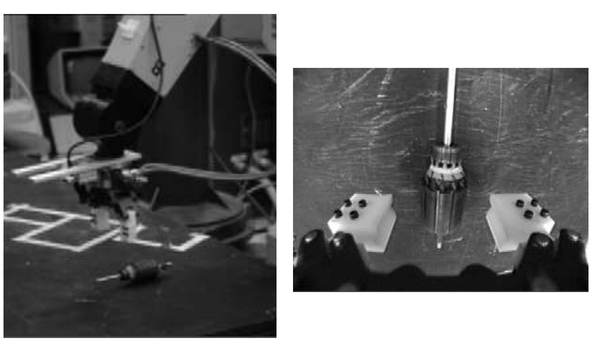

Figure 3. Visual servoing maintains gripper alignment, while a user remotely controls the gripper approach velocity using a hybrid visual servoing control scheme.

4.3. Shared Control

Shared control allows the use of visual and force control along the same directions. This strategy is useful if surfaces can not be reliably detected with vision and it is necessary that the manipulator reacts compliantly to stiff environments along directions of visual servoing. One application of visual servoing for which shared control is particularly useful is when visual feedback is used in teleoperation. Experiments have been performed in which a user guides a manipulator under visual control with a mouse and cursor on a live video image. When guiding the manipulator to grasp an object, the user may sometimes command the manipulator to contact a surface. This can damage the object and often causes the manipulator to become unstable. With shared control, this is easily avoided because force as well as visual feedback is used throughout the task.

A fundamental problem when sharing control between force and vision sensors occurs due to end-effector inertial effects. Because force sensors measure all forces (inertial, gravitational, and tactile), the inertial coupling of the end-effector mass beyond the sensor introduces inertial forces into force sensor readings. When the vision system commands motions, the resulting accelerations cause unstable excitations of the force control system. In order to compensate for the unstable excitations, it is necessary to develop robust strategies for avoiding the excitations. Thresholding of force readings is not feasible, since inertial effects can often be as large as desired contact forces. Figure 4 shows the magnitude of experimentally determined inertial forces and the associated measured cartesian accelerations which cause these forces.

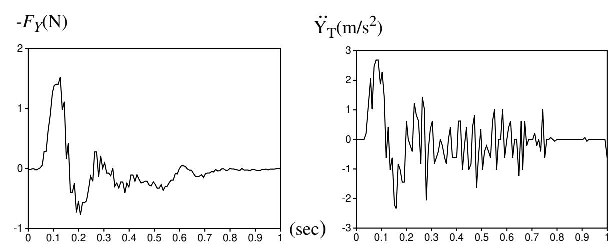

Figure 4. Inertial forces measured by the force sensor and corresponding measured cartesian accelerations that induce these forces.

We have developed a robust shared strategy that compensates for large accelerations induced by inertial forces. If visual servoing results in measurable end-effector accelerations of sufficient magnitude, then force readings in directions opposite to these accelerations are being induced. Because measured cartesian accelerations are derived from joint encoder readings, thus requiring two differentiations of measured joint values and a transformation from joint space to cartesian space, measured cartesian accelerations are noisy. Therefore, we also consider the measured velocity of end-effector motion. If measured cartesian accelerations have been induced by visual servoing, and if a measurable cartesian velocity exists, then sensed forces must be due to inertial coupling, and force control commands should be ignored.

This strategy can be written as

$$\dot{\boldsymbol{x}}_{ref_v} = -\left(\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\boldsymbol{J}_v(k) + \boldsymbol{L}\right)^{-1}\boldsymbol{J}_v^{\mathrm{T}}(k)\boldsymbol{Q}\,[\boldsymbol{x}(k) - \boldsymbol{x}_D(k+1)]$$

$$\dot{\boldsymbol{x}}_{ref_f} = \boldsymbol{S}_F\boldsymbol{G}_F\,(\boldsymbol{F}_r(k) - \boldsymbol{F}_m(k))$$

for each axis, i {

if $\left(\left(\left(\left|\ddot{x}_{m_i}\right| > \varepsilon_a\right) \wedge \left(\dot{x}_{m_i}\operatorname{sgn}F_{m_i} < \varepsilon_v\right)\right) \vee \left(\dot{x}_{ref_{vi}}F_{m_i} > 0.0\right) \vee \left(\left|F_{m_i}\right| < F_T\right)\right)$

$\boldsymbol{S}_{\boldsymbol{v}}[i, i] = 1.0 \quad \boldsymbol{S}_F[i, i] = 0.0$

else $\boldsymbol{S}_{\boldsymbol{v}}[i, i] = 0.0 \quad \boldsymbol{S}_F[i, i] = 1.0$

}

$$\boldsymbol{u}(k) = \boldsymbol{S}_v\dot{\boldsymbol{x}}_{ref_v} + \boldsymbol{S}_r\dot{\boldsymbol{x}}_r + \boldsymbol{S}_F\dot{\boldsymbol{x}}_{ref_f} \quad (11)$$

where ε_a, ε_v, and F_T threshold sensor noise; $\dot{x}_{m_i}$ and $\ddot{x}_{m_i}$ represent measured cartesian velocities and accelerations along particular task space directions, and F_{m_i} represents measured forces along particular directions.

5. Hardware Implementation

The visual tracking algorithms previously described have been implemented on a robotic assembly system consisting of three Puma 560's. The Pumas are controlled from a VME bus with two Ironics IV-3230 (68030 CPU) processors, an IV-3220 (68020 CPU) processor which also communicates with a trackball, a Mercury floating point processor, and a Xycom parallel I/O board communicating with three Lord force sensors mounted on the Pumas' wrists. All processors on the controller VME run the Chimera 3.0 reconfigurable real-time operating system [28]. An Adept robot is also used for providing accurate target motion. A Datacube Maxtower Vision System calculates the optical flow of the features using the SSD algorithm discussed in Section 3.3. A high performance floating point processor on the Datacube is used to calculate the optical flow of features, and a 68030 board, also on the vision system, computes the control input. An image can be grabbed and displacements for up to five 16x16 features in the scene can be determined at 30Hz. The force sensor provides force and torque values for each cartesian axis at 100Hz.

6. Experimental Results

6.1. Traded Control

Throughout this section, experimental results given will be referenced to the coordinate frames shown in Figure 1. Results from a traded control strategy are shown in Figure 5 in which visual servoing is first used to servo the end-effector to a surface 5.9cm from the initial end-effector position. The controller then switches to force control and achieves stable contact with the surface, maintaining a force of -2N between the end-effector and the surface. This strategy achieves contact after 1.43s, and achieves a stable 2N contact force after approximately 4.5s. With simple damping force control alone, the manipulator travels 5.9cm in 3.1s before reaching the surface. When the surface is contacted, the manipulator becomes unstable, as Figure 6 shows. The only way to achieve stable contact using damping control alone given the force control implementation used, is to reduce the force gains to extremely low values, resulting in unacceptably slow motion. Because of this, most manipulation strategies use a guarded move to initiate contact with a surface. During a guarded move, surfaces are approached under position control while the force sensor is monitored. If the sensed force exceeds a threshold, motion is immediately stopped and a force control strategy can then be invoked. Figure 6 also shows a force plot of a guarded move in which the force sensor is monitored at 100Hz. High contact forces are created because of the finite amount of time required to stop the end-effector after contact is sensed illustrating the main limitation of a guarded move strategy.

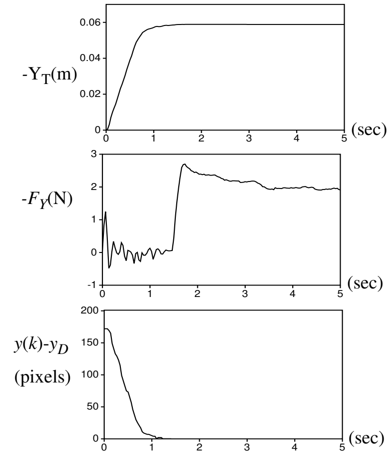

Figure 5. Vertical position of end-effector in cartesian space, force measured along the vertical direction, and pixel error versus time for traded control.

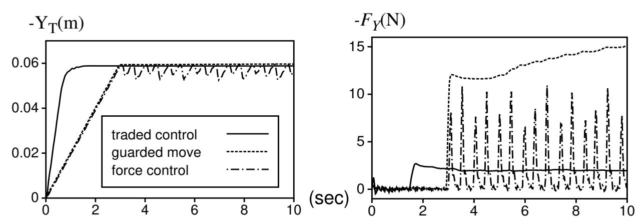

Figure 6. Vertical position and measured force for traded control, a guarded move, and damping force control.

From Figure 6, the advantages of combining force and vision feedback in a traded strategy are evident. Traded visual/force servoing uses visual servoing to quickly approach the surface before switching to a stable low gain (0.001 (m/s)/N) damping force controller. For force control alone, higher force gains (0.005 (m/s)/N) had to be chosen in order to induce end-effector motion of a reasonable speed in free space, but this gain, while resulting in less than half the speed of visual servoing, still proved to be highly unstable. The guarded move strategy also allowed only moderate speeds (0.02m/s), and still resulted in high impact forces. At higher speeds, extremely high impact forces would result which could have easily damaged the manipulator. Using visual servoing to bring the manipulator near the surface provides a simple technique for slowing the end-effector before contact is imminent. These results clearly show that visual servoing greatly simplifies the impact problem by providing low-level feedback on the proximity of the surface to the end-effector. The result is a fast approach velocity that generates low impact forces with no bounce.

6.2. Hybrid Control

To illustrate hybrid visual/force servoing, the manipulator is servoed to maintain a force of 2N normal to a moving plate while maintaining the point of contact in the center of the plate. The edges of a plate are visually tracked, and the manipulator is visually servoed along X_T to maintain the proper contact point. As the plate moves in X_T, force servoing along Y_T maintains the 2N contact force. Plots of the varying motion of the end-effector in X_T and the applied force in Y_T are shown in Figure 7. Motion along the visual servoing direction is precise, but the force servoing data shows the effects of force sensor noise introduced by the motion along X_T. Even so, contact is maintained and is repeatable. Figure 8 shows visual servoing performance on the image plane. The plate moves with a maximum speed of approximately 5cm/s and with extremely high accelerations when switching directions, but the manipulator very accurately tracks this motion despite the large accelerations required.

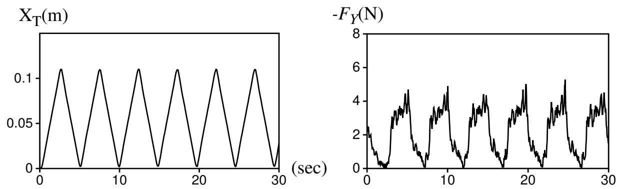

Figure 7. Translation along the surface and measured force applied to the surface for hybrid control.

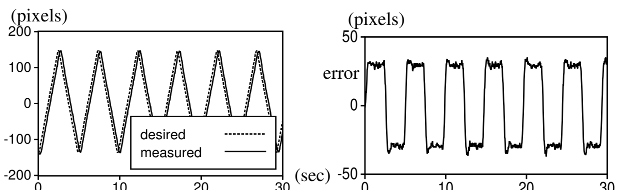

Figure 8. Horizontal target motion (desired) and end-effector motion (measured) on the image plane and errors between target motion and end-effector motion.

6.3. Shared Control

To demonstrate shared control, the desired goal position for the visual servoing strategy is purposely chosen to have differing magnitudes of error with respect to the true location of the surface. This allows us to evaluate the ability of the shared control strategy (11) to operate under conditions in which force and vision information significantly disagree. Figure 9 shows the motion of the end-effector on the image plane for the three trials. For trials 2 and 3 the desired image plane position of the end-effector y_D actually falls beneath the true surface. In trial 2 the error in surface position is 15 pixels, and in trial 3 the error is 45 pixels. For trial 1 the estimate of the surface and the true location are in close agreement, as would normally be the case. In trials 2 and 3, the end-effector impacts the surface after approximately 0.3s, when motion of the end-effector on the image plane abruptly stops.

For trial 1, the surface is not contacted until after approximately 0.5s, because the manipulator purposely slows down before impact. The force plot in Figure 10 shows that this results in significantly reduced impact forces, and a much quicker transition to the desired contact force of -2N. When visual feedback incorrectly estimates the location of the surface, as in the case of the second and third trials, high contact forces occur. If the error in the estimate falls within the surface, as in trials 2 and 3, then the poorer the estimate of the surface, the higher the contact force because the higher the commanded visual servoing velocity at impact. If the error in the surface location estimate is in the other direction, then the time in which it takes to initiate contact would increase directly with the magnitude of the error. The impact force, however, would be on the order of trial 1's impact force.

The commanded end-effector velocity for all three trials is shown in Figure 11. The solid lines correspond to $\boldsymbol{u}(k)$ in (11), the dashed lines to the visual servoing velocity $\dot{\boldsymbol{x}}_{ref_v}$, and the dotted/dashed lines to the force servoing velocity $\dot{\boldsymbol{x}}_{ref_f}$. Visual servoing brings the end-effector quickly towards the surface, and upon contact force servoing takes over. From the force plot in Figure 10, one can observe measurable inertial forces before contact actually occurs. These forces are of a magnitude greater than 1.5N, however, our proposed control strategy (11) successfully rejects these observed forces since they are not the result of contact. From Figure 11, one can see that end-effector velocities have been clipped at 0.10m/s. This is because the feature tracker can only track objects with a limited optical flow. Thus, the trial in which the surface location is in error of 45 pixels represents the worst case impact force, because the manipulator is traveling at approximately 0.10m/s at the time of impact. For these experimental results, force gains of 0.001 (m/s)/N were used, the diagonal elements of $\boldsymbol{Q}$ were chosen to be 2.0×10^{-6}, and the diagonal elements of $\boldsymbol{L}$ were chosen to be 10.0. Thresholds were chosen to be $\varepsilon_a = 0.01 \text{m/s}^2$, $\varepsilon_v = 0.001 \text{m/s}$, and $F_T = 1.5 \text{N}$.

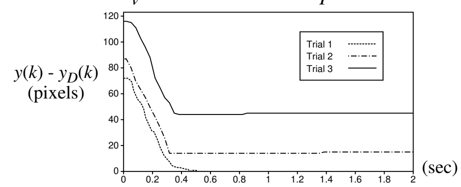

Figure 9. Image plane error between desired and measured end-effector position for the three impact trials.

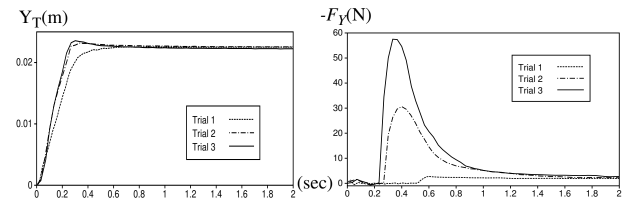

Figure 10. Vertical position of end-effector in cartesian space and measured force for the three impact trials.

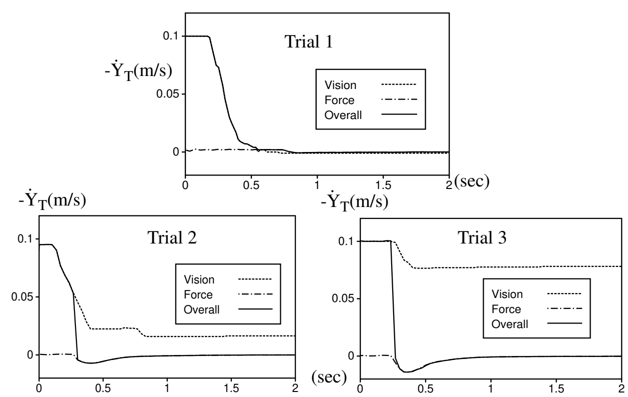

Figure 11. Commanded end-effector velocities for shared control for the three impact trials.

7. Conclusion

Force and vision sensors provide complementary information, yet they are fundamentally different sensing modalities. This implies that traditional sensor integration techniques are not appropriate for combining the feedback from these two disparate sensors. In this paper, we have presented and experimentally verified three force/vision servoing strategies. Experimental results demonstrate significant performance increases over traditional manipulation strategies. One benefit of combining vision and force is illustrated by the system's ability to approach rigid surfaces at high velocities while initiating stable contact with low impact forces and no bounce.

8. 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. Dan Morrow was supported in part by the Computational Science Graduate Fellowship (CSFG) Program of the Office of Scientific Computing in the Department of Energy.

9. References

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

[2] An, C. and Hollerbach, J. 1987. Dynamic stability issues in force control of manipulators. *Proc. 1987 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.860-896.

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

[4] Corke, P.I. and Paul, R.P. 1989. 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.

[5] Eppinger, S. and Seering, W. 1987. Understanding bandwidth limitations on robot force control. *Proc. 1987 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.904-909.

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

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

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

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

[10] Hashimoto, K. and Kimura, H. 1993. 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.

[11] Hogan, N. 1985. Impedance control: an approach to manipulation, Parts I-III. *ASME J. Dyn. Sys., Measurements, and Control*. 107(1):1-24.

[12] Hogan, N. 1987. Stable execution of contact tasks using impedance control. *Proc. 1987 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.1047-1054.

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

[14] Hyde, J.M. and Cutkosky, M.R. 1993. Contact transition: an experimental study. *Proc. 1993 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.363-368.

[15] Ishikawa, J., Kosuge, K. and Furuta, K. 1990. Intelligent control of assembling robot using vision sensor. *Proc. 1985 Int. Conf. on Robotics and Autom*. New York:IEEE, pp. 1904-1909.

[16] Kazerooni, H. 1987. Robust, non-linear impedance control for robot manipulators. *Proc. 1987 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.741-750.

[17] Khatib, O. and Burdick, J. 1986. Motion and force control for robot manipulators. *Proc. 1986 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.1381-1386.

[18] Koivo, A.J. and Houshangi, N. 1991. 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.

[19] Maru, N., Kase, H., Yamada, S., Nishikawa, A. and Miyazaki, F. 1993. 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.

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

[21] Nelson, B. and Khosla, P.K. 1993. The Resolvability Ellipsoid for Visually Guided Manipulation. Technical Report CMU-RI-TR-93-28, Robotics Institute, Carnegie Mellon University.

[22] Nio, S. and Maruyama Y. 1993. Remote-operated robotic system for live-line maintenance work. *Proc. ESMO '93, IEEE 6th Int. Conf. on Transmission and Distribution Construction and Live-Line Maintenance*. New York:IEEE, pp.425-435.

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

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

[25] Qian, H.P. and De Schutter, J. 1992. Introducing active linear and nonlinear damping to enable stable high gain force control in case of stiff contact. *Proc. 1992 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.1374-1379.

[26] Raibert, M.H. and Craig, J.J. 1981. Hybrid position/force control of manipulators. *ASME J. of Dyn. Sys., Measurement, and Control*. 103(2):126-133.

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

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

[29] Volpe, R.A. and Khosla, P.K. 1991. Experimental verification of a strategy for impact control. *Proc. 1991 IEEE Int. Conf. on Robotics and Automation*. New York:IEEE. pp. 1854-1860.

[30] Weiss. L.E. 1984. 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.

[31] Whitney, D.E. 1977. Force feedback control of manipulator fine motions. *ASME J. of Dyn. Sys. Measurement, and Control*. June:91-97.

[32] Whitney, D.E. 1985. Historical perspective and state of the art on robot force control. *Proc. 1985 Int. Conf. on Robotics and Automation*. New York:IEEE, pp. 262-268.

[33] Wilson, W.J. 1993. 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.

[34] Xu, Y., Hollerbach, J.M., and Ma, D. 1994. Force and Contact Transient Control Using Nonlinear PD Control. *Proc. 1994 Int. Conf. on Robotics and Automation*. New York:IEEE, pp.924-930.