Fast Stable Contact Transitions with a Stiff Manipulator Using Force and Vision Feedback

Bradley J. Nelson† J. Daniel Morrow† Pradeep K. Khosla‡

†The Robotics Institute
‡Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh PA 15213

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

Abstract

Ideal manipulator end-effector transitions from non-contact to contact states should be fast and stable with minimal impact forces and without bounce. These specifications, however, are difficult to simultaneously achieve, especially for the most common manipulator force control configuration employing a wrist force sensor and a stiff manipulator position loop. In this paper we present a control strategy that uses high bandwidth vision feedback (30Hz) in addition to force feedback (100Hz) for contact transient control. A nonlinear control strategy is proposed that considers force and vision feedback simultaneously and then switches to pure force control when the camera-lens system becomes unable to accurately resolve the location of the end-effector relative to the surface to be contacted. Experimental results are presented which demonstrate that a stiff manipulator can quickly contact a stiff surface stably without bounce. Results also show that the magnitude of the impact force spike is directly related to the accuracy with which contact surfaces are visually observed.

1 Introduction

To quickly and efficiently perform robotic manipulation tasks in uncertain environments, a robotic end-effector must be able to successfully approach and contact objects rapidly using sensor feedback. In a rigid environment this is difficult. With a stiff manipulator this becomes even more difficult because neither the surface nor the manipulator are able to easily dissipate excess energy upon contact. However, the most common form of force control uses a proportional-derivative strategy with high proportional gains and low damping, which is an inherently stiff system. The reason this type of force control is popular is because it is simple to implement, choosing gains is easy, and it achieves a relatively high bandwidth once contact is successfully made. The problem with this strategy, however, is in achieving initial contact quickly and stably while maintaining low impact forces. Many researchers have studied this problem, and various impact strategies have been proposed. However, the fundamental problem of using force feedback alone to minimize impact forces while quickly achieving contact stably within imprecisely calibrated environments still exists.

Visual servioing research has shown that by placing vision feedback within a manipulator feedback loop, manipulators can be robustly and accurately guided through imprecisely calibrated and dynamically varying environments, even when imprecisely calibrated camera-lens systems and manipulators are used. Vision feedback, however, has a limited spatial resolution that makes its use inappropriate during contact stages of manipulation tasks. Force feedback, on the other hand, provides accurate but highly localized contact information. These characteristics of vision and force feedback illustrate their complementary nature: vision allows accurate part alignment within imprecisely calibrated and dynamically varying environments without requiring contact; force allows the task to be controlled with a precision beyond the capability of vision alone during contact phases of the task.

In this paper we show how force and vision feedback can be integrated in a manipulator feedback loop in order to achieve fast and stable contact transitions between a stiff manipulator and a rigid environment. Our proposed controller effectively compensates for inertial coupling between the force sensor and the end-effector mass beyond the sensor while responding to both vision and force feedback along the same direction. Experimental results show that the use of visual servoing to stably guide a manipulator simplifies the force control problem by allowing 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.

Brad Nelson and Dan Morrow are with the Robotics Ph.D. Program at Carnegie Mellon University.
2 Previous Work

2.1 Force Servoing

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 [15], an effective impact strategy is presented based on a proportional gain explicit force controller with a feedforward 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 [1][3][4][5][7][8][12][18]—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 vision feedback for guiding manipulator motion, for example [14], the visual servoing field was first well defined by Weiss [16]. An excellent survey of recent work in visual servoing can be found in [2]. 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 Vision Servoing

One of the first papers to mention the benefits of combining high bandwidth vision and force feedback is [14]. They implemented a 0.1Hz visual servoing scheme and mentioned the advantages of 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 [6], 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 vision 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

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

\[ \dot{x}_s = J_\phi \dot{X}_T \]  

(1)

where \( \dot{x}_s \) is a velocity vector in sensor space; \( J_\phi \) is the Jacobian matrix and is a function of the extrinsic and intrinsic parameters of the vision sensor as well as the number of features tracked and their locations on the image plane; and \( \dot{X}_T \) is a velocity vector in task space.

A pinhole camera model is used to model the camera-lens system, as shown in Figure 1. A feature on a manipulated object is observed on the image plane with respect to the camera frame \( \{C\} \) at \( \{P \} \), but is controlled with respect to the manipulator's task frame \( \{T \} \) at \( \{P \} \). The Jacobian transformation for a single feature from task space to sensor space can be written in the form

\[
\begin{bmatrix}
\dot{x}_s \\
\dot{y}_s \\
\dot{z}_s
\end{bmatrix} = \begin{bmatrix}
\frac{f}{x_C} & 0 & \frac{y_P z_C}{x_C} \\
\frac{f}{y_C} & \frac{x_P z_C}{y_C} & 0 \\
0 & \frac{x_P y_C}{y_C} & \frac{y_P z_C}{y_C}
\end{bmatrix} \begin{bmatrix}
\dot{x}_T \\
\dot{y}_T \\
\dot{z}_T
\end{bmatrix}
\]

(2)

For the above form of the Jacobian, the parameters of the Jacobian are given by \( \phi = (f, x_C, y_C, x_P, y_P, z_C, x_T, y_T, z_T) \).

A complete derivation of (2) can be found in [10].

![Figure 1: Task frame-camera frame definitions.](image-url)
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

\[
J_{f}(\phi) = \begin{bmatrix} J_{f1}(\phi) & \cdots & J_{f}(\phi) \end{bmatrix}^T
\]

where \( J_{f}(\phi) \) is the Jacobian matrix for each feature given by the 2x6 matrix in (2).

### 3.2 Visual Servoing Controller

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

\[
x(k + 1) = x(k) + T J_{f}(k) u(k)
\]

where \( M \) is the number of features being tracked; \( x(k) \in \mathbb{R}^{2M} \); \( T \) is the sampling period of the vision system; and \( u(k) = \begin{bmatrix} x_{p}^T & \tau_{p}^T & \omega_{p}^T & \omega_{p}^T & \alpha_{p}^T \end{bmatrix}^T \) is the manipulator end-effector velocity. The Jacobian \( J_{f}(\phi) \) is written as \( J_{f}(\phi) \) 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 visual servoing controller is formulated based on controlled active vision[11]. The control objective of the 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.

\[
f(k + 1) = (x(k + 1) - x(k + 1))^T Q (x(k + 1) - x(k + 1) + u^T(k) L u(k))
\]

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

\[
u(k) = \left( T J_{f}^T(k) Q T J_{f}(k) + L \right)^{-1} T J_{f}^T(k) Q (x(k) - x(k + 1))
\]

The weighting matrices \( Q \) and \( L \) 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 \( Q \) matrix must be positive semi-definite, and \( L \) must be positive definite for a bounded response. Although no standard procedure exists for choosing the elements of \( Q \) and \( L \), general guidelines can be found in [11], along with system models and control derivations that account for system delays, modeling and control inaccuracies, and measurement noise.

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 [9].

### 4 Vision and Force Servoing

The force control portion of our proposed vision/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 [13] and damping force control [17], 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.

If simple force damping control is used to impact surfaces, a manipulator can easily become unstable 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 initialize 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.

The proper use of vision feedback can overcome the problems exhibited by guarded move and pure force control strategies upon impact. Visual servoing improves manipulator performance during contact transitions by incorporating information regarding the proximity of the surface to be contacted in the manipulator feedback loop. When the end-effector is far from a surface, visual servoing commands fast end-effector motion. The speed of the approach decreases as the end-effector approaches nearer the surface. Contact can then be stably initiated through the use of low gain force controllers. A generic control framework for vision/force servoing is shown in Figure 2.

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

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 asdesired contact forces.

We have developed a robust vision/force control strategy based on the fact that large accelerations induce 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, this requiring two differentiation of measured joint values and a transformation from joint space to cartesian space, measured cartesian accelerations are noisy. Therefore, we also consider the measured direction 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

\[
\begin{align*}
\dot{\bar{x}}_{\text{ref},i} &= \left( J_i^T(k) Q J_i(k) + L_i \right)^{-1} J_i^T(k) Q \left[ \bar{x}(k) - \bar{x}_p(k + 1) \right] \\
\dot{\bar{x}}_{\text{ref},i} &= S_{\text{p}} G_p(F_p(k) - F_m(k))
\end{align*}
\]

for each axis, \(i\) \{ 

\[
\begin{align*}
\text{if} \left( \left( \left[ F_{\text{mid}} \right] > \varepsilon_u \right) \land \left( \dot{x}_{\text{ref},i} F_{\text{mid}} > 0 \right) \right) \lor \left( \left[ F_{\text{mid}} \right] < F_T \right) \\
S_{p}[i,i] = 1.0 \\
S_{f}[i,i] = 0.0
\end{align*}
\]

\[
\begin{align*}
\text{else} \ S_{p}[i,i] = 0.0 \\
S_{f}[i,i] = 1.0
\end{align*}
\]

\[
u(k) = S_{p} \dot{\bar{x}}_{\text{ref},i} + S_{f} \dot{x} + S_{f} \dot{\bar{x}}_{\text{ref},i}
\]

where \(x_p\) represents a state in feature space on the surface to be contacted; \(S_p\) and \(S_f\) are selection matrices; \(G_p\) is a matrix of force control gains; \(F_p\) and \(F_m\) represent reference and measured forces; \(\dot{x}_{\text{ref},i}\) and \(\dot{x}_{i}\) represent measured cartesian velocities and accelerations along particular task space directions; \(\varepsilon_u\), \(\varepsilon_f\), and \(F_T\) threshold sensor noise; and \(\dot{x}_\text{r}\) represents some desired end-effector reference velocity due, for example, to trackball input from a teleoperator.

For teleoperation tasks guided by visual servoing, compliant contact with the environment will occur if (7) is used alone assuming the teleoperator adjusts the desired visual feature state to be at or below the surface to be contacted. For autonomous manipulation, however, this strategy does not ensure contact will occur if the actual location of the surface is beyond the visual estimate of the surface. During autonomous manipulation, the strategy given by (7) must be rewritten as

\[
u(k) = \begin{cases} 
(7), & \left| \bar{x}_p(k) - \bar{x}(k) \right| > \varepsilon \\
S_{p} G_p(F_p(k) - F_m(k)), & \left| \bar{x}_p(k) - \bar{x}(k) \right| \leq \varepsilon 
\end{cases}
\]

in order to ensure that contact will occur. Manipulator motion is first controlled by the strategy given in (7). The controller then switches to pure force control if the error between desired and measured visual state converges to within some threshold. Stable impact with a surface can then be achieved, large contact forces can be minimized, and bounce can be avoided.

5 Hardware Implementation

The vision 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. The force sensor provides force and torque values for each cartesian axis at 100Hz. A Datacube Maxtower Vision System calculates the optical flow of the features. An image can be grabbed and displacements for up to five 16x16 features in the scene can be determined at 30Hz.

6 Experimental Results

For the initial set of experiments, the results of three trials are shown in which 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. A final contact force of -2N is desired. This allows us to evaluate the ability of our force/vision control strategy to operate under conditions in which force and vision information significantly disagree. Figure 3 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 \(y_p\) 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 4 shows that this results in significantly reduced impact forces, and a much quicker transition to the desired contact force of -2N. When vision 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 5. The solid lines correspond to $u(k)$ in (8), the dashed lines to the visual servoing velocity $x_{ref}$, and the dotted/dashed lines to the force servoing velocity $x_{ref}$. Visual servoing brings the end-effector quickly towards the surface, and upon contact force servoing takes over. From the force plot in Figure 4, 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 (8) successfully rejects these observed forces because they are not the result of contact. From Figure 5, one can see that end-effector velocities have been clipped to 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 0.001m/s (sN) were used, the diagonal elements of $Q$ were chosen to be 2.0x$10^{-6}$, and the diagonal elements of $L$ were chosen to be 10.0. Thresholds were chosen to be $\varepsilon_u$=0.01m/s$^2$, $\varepsilon_f$=0.001m/s, and $F_f$=1.5N.

A second set of experimental results was collected in order to illustrate the advantages of our proposed force/vision strategy over two other common impact strategies, the guarded move and damping force control. Figure 6 compares the three strategies. The solid line shows the performance of our proposed force/vision servoing algorithm in which (8) is used to servo the end-effector to a surface 5.9cm from the initial end-effector position. A force of -2N between the end-effector and the surface is maintained after contact. This strategy achieves contact after 1.43s, and achieves a stable -2N contact force after approximately 4.5s. With simple damping force control alone (dash-dotted line), the manipulator travels 5.9cm in 3.1s before reaching the surface. As soon as the surface is contacted, the manipulator becomes unstable. 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. Figure 6 also shows a force plot of a guarded move in which the force sensor is monitored at 100Hz. 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. 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.
7 Conclusion

Properly combining force and vision feedback within a manipulator feedback loop gives a robotic manipulation system that ability to take advantage of the complementary nature of the two disparate sensing modalities. An important 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. Several issues must be addressed when integrating force and vision, including the inertial effects vision feedback may induce in the force control loop. We have shown how these effects can be eliminated, and how stable and robust control using vision and force feedback simultaneously can be achieved.

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. The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the funding agencies.

9 References