Technical Report

The Resolvability Ellipsoid for

Sensor Based Manipulation

Brad Nelson

Pradeep K. Khosla

CMU-RI-TR-93-28

The Robotics Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891

October 1993

This research was supported in part by the U.S. Army Research Office thorugh Grant Number DAAL03-91-G-0272 and by Sandia National Laboratories through Contract Number AC-3752D.

Abstract

This paper presents a new sensor placement measure called *resolvability*. This measure provides a technique for estimating the relative ability of various sensor systems, including single camera systems, stereo pairs, multi-baseline stereo systems, and 3D rangefinders, to accurately control visually manipulated objects. The measure also indicates the capability of a visual sensor to provide spatially accurate data on objects of interest. The term *resolvability* refers to the ability of a visual sensor to resolve object positions and orientations. Our main interest in *resolvability* is in determining the accuracy with which a manipulator being observed by a camera can visually servo an object to a goal position and orientation. The *resolvability ellipsoid* is introduced to illustrate the directional nature of *resolvability*, and can be used to direct camera motion and adjust camera intrinsic parameters in real-time so that the servoing accuracy of the visual servoing system improves with camera-lens motion. The Jacobian mapping from task space to sensor space is derived for a single camera system, a stereo pair with parallel optical axes, and a stereo pair with perpendicular optical axes. *Resolvability ellipsoids* based on these mappings for various sensor configurations are presented. Visual servoing experiments demonstrate that *resolvability* can be used to direct camera-lens motion in order to increase the ability of a visually servoed manipulator to precisely servo objects.

Table of Contents

List of Figures

1. Introduction

In order to effectively use visual feedback to perform robotic tasks, many researchers have recognized that the placement of the sensor relative to the task is an important consideration [2], [12], and [13]. Active cameras-lens systems that possess several extrinsic as well as intrinsic variable parameters provide a high degree of flexibility in providing information about the task by allowing real-time control over sensor placement. This flexibility, however, provides challenges in effective real-time camera-lens control. A framework for stably controlling various extrinsic camera parameters in real-time in order to optimize various sensor placement criteria has been previously proposed and experimentally verified [8]. This paper presents a new sensor placement measure called *resolvability*, which provides a technique for estimating the relative ability of various visual sensor systems, including single camera systems, stereo pairs, multi-baseline stereo systems, and 3D rangefinders, to accurately control visually manipulated objects and to provide spatially accurate data on objects of interest.

The term *resolvability* refers to the ability of a visual sensor to resolve object positions and orientations. For example, a typical single camera system has the ability to accurately resolve object locations that lie in a plane parallel to the image plane, but can less accurately resolve object depth based on the projection of object features on the image plane. Similarly, rotations within planes parallel to the image plane can be more accurately resolved than rotations in planes perpendicular to the image plane. The degree of *resolvability* is dependent on many factors. For example, depth, focal length, number of features tracked and their image plane coordinates, position and orientation of the camera, and relative positions and orientations of multiple cameras, all effect the magnitudes and directions of *resolvability*. Due to the difficulty in understanding the multi-dimensional nature of *resolvability*, we propose the *resolvability ellipsoid* as a geometrical representation of the ability of different visual sensor configurations to resolve object positions and orientations.

Resolvability can be applied to both visual servoing and sensor planning. Our main interest is in determining the accuracy with which a manipulator being observed by a camera can visually servo an object to a goal position and orientation. The *resolvability ellipsoid* can be used to direct camera motion and adjust camera intrinsic parameters in real-time so that the servoing accuracy of the visual servoing system improves with camera-lens motion. A second use of *resolvability* is as an aid in determining static camera placement for either object recognition or for visual servoing.

Sensor resolution has been considered in the past as a criterion for sensor planning [1], [2], and [11]. These efforts concern static camera systems in which a required spatial resolution is known and a single camera placement is desired. In [3], a study of stereo, vergence, and focus cues for determining range is described in which the performance of each cue for determining range accuracy is characterized. This characterization can be used to control camera parameters in order to improve the accuracy of range estimates. Our *resolvability* approach can be used for determining the ability of a visually servoed manipulator to accurately resolve positions and orientations of objects along all six degrees of freedom. Camera-lens intrinsic and extrinsic parameters can be actively controlled using a *resolvability* measure in conjunction with other sensor placement criteria so that the accuracy of visual control can be improved. The concept can also be used for static sensor placement for either object recognition or visual servoing. This application parallels the use of manipulability in determining the optimal placement of assembly tasks in a manipulator's workspace [7].

There are many parallels between *resolvability* and the concept of *manipulability* as proposed in [14]. Both concepts use an elliptical representation of the singular value decomposition of the Jacobian of the mapping between sensor space and task space. An important difference is that a high *manipulability* means that relatively small joint motions translate into a relatively large endeffector motions. A high *resolvability*, on the other hand, means that a small object displacement translates into a relatively large displacement in sensor space. A high *resolvability* also indicates that a small object velocity projects a large optical flow onto the image plane. In fact, our concept of *resolvability* more closely parallels Ghosal and Roth's *transmission ratio* [4]. *Resolvability* can also be viewed as a measure of the sensitivity of the sensor to displacements of the object of interest along particular directions in the task space.

In this paper, we first derive a mapping from task space to sensor space, which is the first step in determining the *resolvability* of a camera-lens system. We then briefly describe the singular value decomposition and its ellipsoidal representation for *resolvability*. Next, different *resolvability ellipsoids* for various camera-lens-task configurations are compared. The *resolvability ellipsoids* of a stereo pair with parallel optical axes and a stereo pair with perpendicular optical axes are presented as well. A technique for using the *resolvability ellipsoid* to direct camera motion is proposed. A section on experimental results briefly describes our visual servoing hardware, our visual tracking strategy, and results from visual servoing using different camera configurations. These results demonstrate that *resolvability* predicts the ability of the system to accurately servo under different camera-lens configurations.

2. Task Space → Sensor Space

Resolvability depends on the Jacobian of the mapping from task space to sensor space. We desire a matrix form of the Jacobian which contains both intrinsic and extrinsic sensor parameters, in order to analyze the effects of changing these parameters on the structure of the Jacobian. For any visual sensor system, we desire an equation of the form

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

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

2.1 Camera Model

The mapping from task space to sensor space for any system using a camera as the visual sensor requires a camera-lens model in order to represent the projection of task objects onto the CCD image plan. For visual servoing, a simple pin hole camera model has proven adequate for visual tracking using our experimental setup [8]. If we place the camera coordinate frame {C} at the focal point of the lens as shown in Figure 1, a feature on an object at ${}^{C}\boldsymbol{P}$ with coordinates (X_C, Y_C, Z_C) in the camera frame projects onto the camera's image plane at

$$x_i = \frac{fX_C}{s_x Z_C} + x_p \tag{2}$$

$$y_i = \frac{fY_C}{s_y Z_C} + y_p \tag{3}$$

where (x_i, y_i) are the image coordinates of the feature, f is the focal length of the lens, s_x and s_y are the horizontal and vertical dimensions of the pixels on the CCD array, and (x_p, y_p) is the piercing point of the optical axis on the CCD. This model assumes that $Z_C >> f$.

The mapping from camera frame feature velocity to image plane optical flow, or sensor space velocity, can be obtained simply by differentiating (2) and (3). This yields the following equations

$$\dot{x}_S = \frac{f\dot{X}_C}{s_x Z_C} - \frac{fX_C \dot{Z}_C}{s_x Z_C^2} = \frac{f\dot{X}_C}{s_x Z_C} - x_S \frac{\dot{Z}_C}{Z_C} \tag{4}$$

$$\dot{y}_S = \frac{f\dot{Y}_C}{s_y Z_C} - \frac{fY_C \dot{Z}_C}{s_y Z_C^2} = \frac{f\dot{Y}_C}{s_y Z_C} - y_S \frac{\dot{Z}_C}{Z_C} \tag{5}$$

where $x_s = x_i - x_p$ and $y_s = y_i - y_p$. The mapping from the camera frame onto the image plane is now defined. The next step is to transform task space velocities into the camera frame, and then project these camera frame velocities onto the sensor space to obtain the mapping from task space velocity to sensor space velocity.

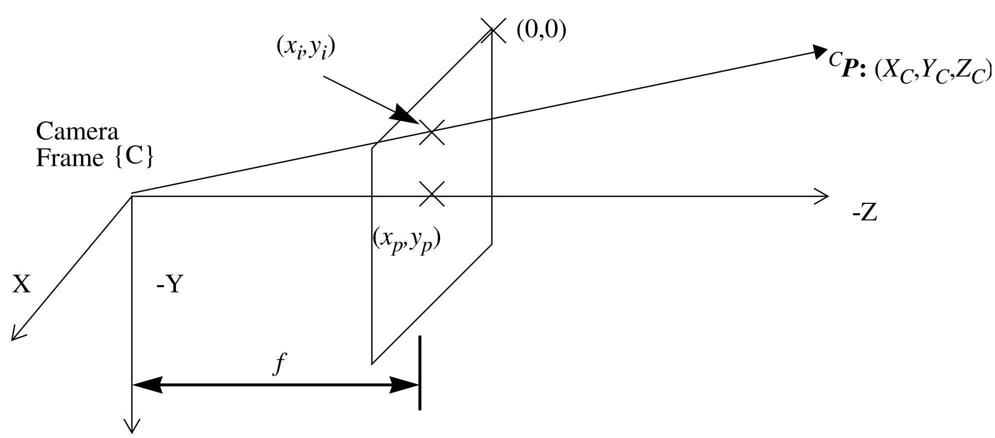

Figure 1: Pinhole camera model.

2.2 Objects Defined in a Task Frame

For visually servoing a manipulator holding an object, the objective is to move the image coordinates of ${}^{C}\boldsymbol{P}$ to some location on the image plane by controlling the motion of ${}^{C}\boldsymbol{P}$. Typically, ${}^{C}\boldsymbol{P}$ is some feature on an object being held by a manipulator. Thus, the motion of ${}^{C}\boldsymbol{P}$ is induced relative to the tool frame of the manipulator being observed. Figure 2 shows the coordinate systems used to define the mapping from task space to sensor space for ${}^{T}\boldsymbol{P}$ with coordinates in the task frame of (X_T, Y_T, Z_T). For now, we assume that the rotation of the task frame {T} with respect to {C} ${}^{C}_{T}\boldsymbol{R}$ is known. The velocity of ${}^{C}\boldsymbol{P}$ can be written as

$$ {}^{C}\dot{\boldsymbol{P}} = {}^{C}_{T}\boldsymbol{R}\,({}^{T}\boldsymbol{V} + {}^{T}\dot{\boldsymbol{P}} + {}^{T}\Omega \times {}^{T}\boldsymbol{P}) \tag{6} $$

where ${}^{T}\boldsymbol{V} = \begin{bmatrix} \dot{x}_T & \dot{y}_T & \dot{z}_T \end{bmatrix}$ and ${}^{T}\Omega = \begin{bmatrix} \omega_{x_T} & \omega_{y_T} & \omega_{z_T} \end{bmatrix}^T$ are the translational and rotational velocities of the task frame with respect to itself, respectively. These are manipulator endeffector velocities that can be commanded. Since the object being servoed or observed is rigidly attached to the task frame, ${}^{T}\dot{\boldsymbol{P}} = 0$, and (6) becomes

$$ {}^{C}\dot{\boldsymbol{P}} = {}^{C}_{T}\boldsymbol{R}\,({}^{T}\boldsymbol{V} + {}^{T}\Omega \times {}^{T}\boldsymbol{P}) \tag{7} $$

Furthermore, if we assume that {C} and {T} are aligned, as shown in Figure 2, then ${}^{C}_{T}\boldsymbol{R} = \boldsymbol{I}$ and the elements of ${}^{C}\dot{\boldsymbol{P}}$ can be written as

$$ \begin{aligned} \frac{dX_C}{dt} &= \dot{x}_T + Z_T\omega_{y_T} - Y_T\omega_{z_t} \\ \frac{dY_C}{dt} &= \dot{y}_T - Z_T\omega_{x_T} + X_T\omega_{z_T} \\ \frac{dZ_C}{dt} &= \dot{z}_T + Y_T\omega_{x_T} - X_T\omega_{y_t} \end{aligned} \tag{8} $$

The assumption that {C} and {T} are aligned is only used in formulating the Jacobian from task space to sensor space. If the transformation from task space to sensor space is initially known, and the commanded task frame velocity is known, then the coordinates (X_T, Y_T, Z_T) can be appropriately updated while visual servoing. It will also be necessary to account for task frame rotations when determining the velocity to command the task frame based on ${}^{T}\boldsymbol{V} = \begin{bmatrix} \dot{x}_T & \dot{y}_T & \dot{z}_T \end{bmatrix}$ and

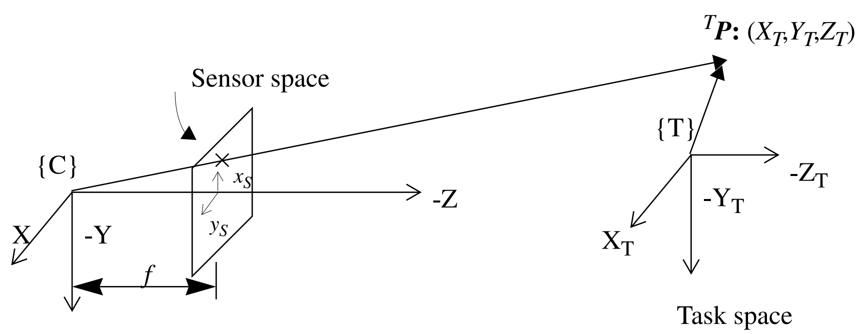

Figure 2: Task frame and camera frame definitions.

${}^T\Omega = \begin{bmatrix} \omega_{x_T} & \omega_{y_T} & \omega_{z_T} \end{bmatrix}^T$. It would have been possible to include the terms of ${}^C_T\boldsymbol{R}$ in (8), however, the assumption made simplifies the derivation and does not affect the end result.

By combining (8) with (4) and (5), the entire Jacobian transformation for a single feature from task space to sensor space can now be written in the form

$$\begin{bmatrix} \dot{x}_S \\ \dot{y}_S \end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_C} & 0 & -\frac{x_S}{Z_C} & -\frac{Y_T x_S}{Z_C} & \left[\frac{fZ_T}{s_x Z_C} + \frac{X_T x_S}{Z_C}\right] & -\frac{fY_T}{s_x Z_C} \\ 0 & \frac{f}{s_y Z_C} & -\frac{y_S}{Z_C} & -\left[\frac{fZ_T}{s_y Z_C} + \frac{Y_T y_S}{Z_C}\right] & \frac{X_T y_S}{Z_C} & \frac{fX_T}{s_y Z_C} \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \\ \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \tag{9}$$

For the above form of the Jacobian, the parameters of the Jacobian are given by $\phi = (f, s_x, s_y, x_S, y_S, Z_C, X_T, Y_T, Z_T)$. Alternatively, the sensor coordinates may be omitted and replaced with camera frame coordinates to arrive at a Jacobian of the form

$$\begin{bmatrix} \dot{x}_S \\ \dot{y}_S \end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_C} & 0 & -\frac{fX_C}{s_x Z_C^2} & -\frac{fX_C Y_T}{s_x Z^2{}_C} & \left[\frac{fZ_T}{s_x Z_C} + \frac{fX_C X_T}{s_x Z^2{}_C}\right] & -\frac{fY_T}{s_x Z_C} \\ 0 & \frac{f}{s_y Z_C} & -\frac{fY_C}{s_y Z_C^2} & -\left[\frac{fZ_T}{s_y Z_C} + \frac{fY_C Y_T}{s_y Z^2{}_C}\right] & \frac{fY_C X_T}{s_y Z^2{}_C} & \frac{fX_T}{s_y Z_C} \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \\ \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \tag{10}$$

where the parameters are now $\phi = (f, s_x, s_y, X_C, Y_C, Z_C, X_T, Y_T, Z_T)$. Either form may be desirable depending on the design parameters desired for determining sensor placement.

Generally, several features on an object are tracked. For n feature points, the Jacobian is of the form

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_1 \\ \dots \\ \boldsymbol{J}_n \end{bmatrix} \tag{11}$$

where $\boldsymbol{J}_i$ is the Jacobian matrix for each feature given by the 2x6 matrix in (9) or (10).

2.3 Objects Defined in the Camera Frame

For eye-in-hand tracking, it may be preferable to define the task frame to be equivalent to the camera frame. This simplifies the derivation of the task frame to sensor frame Jacobian. The point being observed ${}^C\boldsymbol{P}$ is observed by translating and rotating the camera. Therefore, (6) becomes

$${}^C\dot{\boldsymbol{P}} = -{}^C\boldsymbol{V} - {}^C\Omega \times {}^C\boldsymbol{P} \tag{12}$$

where ${}^{C}\boldsymbol{V} = \begin{bmatrix}\dot{x}_C & \dot{y}_C & \dot{z}_C\end{bmatrix}$ and ${}^{C}\Omega = \begin{bmatrix}\omega_{x_C} & \omega_{y_C} & \omega_{z_C}\end{bmatrix}^T$ are the translational and rotational velocities of the camera with respect to the current camera frame. By combining (4) and (5) with (12) as in the previous derivation, one can arrive at the following form for the Jacobian

$$\begin{bmatrix}\dot{x}_S \\ \dot{y}_S\end{bmatrix} = \begin{bmatrix} -\frac{f}{s_x Z_C} & 0 & \frac{x_S}{Z_C} & \frac{s_y}{f} x_S y_S & -\left[\frac{f}{s_x} + \frac{s_x}{f} x_S^2\right] & \frac{s_y}{s_x} y_S \\ 0 & -\frac{f}{s_y Z_C} & \frac{y_S}{Z_C} & \left[\frac{f}{s_y} + \frac{s_y}{f} y_S^2\right] & -\frac{s_x}{f} x_S y_S & -\frac{s_x}{s_y} x_S \end{bmatrix} \begin{bmatrix}\dot{x}_C \\ \dot{y}_C \\ \dot{z}_C \\ \omega_{x_C} \\ \omega_{y_C} \\ \omega_{z_C}\end{bmatrix} \tag{13}$$

3. Ellipsoidal Representations for *Resolvability*

The ability of a visual sensor to resolve task positions and orientations is, of course, directionally dependent. By performing a singular value decomposition on the task space to sensor space Jacobian, and analyzing the singular values and the eigenvectors of $\boldsymbol{J}^T\boldsymbol{J}$ which result from the decomposition, the directional properties of the ability of the sensor to resolve positions and orientations becomes apparent.

Singular value decomposition is a common technique used for analyzing the range and null space of a matrix transformation. Details concerning the SVD can be found in [5]. The SVD of a matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^T \tag{14}$$

where Σ is a diagonal matrix containing the square roots of the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^T$, also called the singular values of $\boldsymbol{A}$, $\boldsymbol{U}$ contains the eigenvectors of $\boldsymbol{A}\boldsymbol{A}^T$, and $\boldsymbol{V}$ contains the eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$. For *resolvability* the eigenvectors of $\boldsymbol{J}^T\boldsymbol{J}$ are those in which we are interested, because these eigenvectors give us a set of basis vectors for the row space of $\boldsymbol{J}$, which is also the vector space described by $\Re(\boldsymbol{J}^T)$, the range of $\boldsymbol{J}^T$. These basis vectors combined with their corresponding singular values indicate the directionality of the ability of the camera-lens system to resolve positions and orientations in task space. Conversely, the eigenvectors of $\boldsymbol{J}\boldsymbol{J}^T$ tell us the effect of task space object motion in sensor space.

In order to use an ellipsoidal representation of *resolvability*, we first assume that the object of interest has an equal ability to translate and rotate about all of its cartesian axes. Furthermore, we assume that the velocity of the object is constrained to fall within some six dimensional spheroid, such that

$$\|\dot{\boldsymbol{x}}_T\| = \left(\dot{x}_T^2 + \dot{y}_T^2 + \dot{z}_T^2 + \omega_{x_T}^2 + \omega_{y_T}^2 + \omega_{z_T}^2\right)^{\frac{1}{2}} \tag{15}$$

and $\|\dot{\boldsymbol{x}}_T\| \leq 1$. Under these assumptions the principal axes of the ellipsoid representing the ability of $\boldsymbol{J}$ to resolve positions and orientations in task space are given by $\sigma_1\boldsymbol{v}_1$, $\sigma_2\boldsymbol{v}_2$, $\sigma_3\boldsymbol{v}_3$, $\sigma_4\boldsymbol{v}_4$, $\sigma_5\boldsymbol{v}_5$, and $\sigma_6 \boldsymbol{v}_6$, where σ_i is the ith singular value of $\boldsymbol{J}$ and $\boldsymbol{v}_i$ is the ith eigenvector of $\boldsymbol{J}^T\boldsymbol{J}$. In the next section we present ellipsoids for various camera-lens-object configurations.

4. *Resolvability Ellipsoids* for a Single Camera System

The *resolvability ellipsoid* makes the directional nature of a visual sensor's ability to resolve positions and orientations apparent. In this section, the *resolvability ellipsoid* for a simple single camera system is illustrated. The extrinsic parameters that can be varied include the six degrees of freedom of the camera position and orientation. The intrinsic parameter that can vary is the focal length. Changing the position or orientation of the camera also has the effect of changing the coordinates of the features being tracked in sensor space. The effect of including more features in the tracking system is also shown.

Displaying a six-dimensional ellipsoid is somewhat problematic, so the mapping described by (10) has been decomposed into two mappings, one representing translational components and one representing rotational components. The mappings are described by

$$\begin{bmatrix} \dot{x}_S \\ \dot{y}_S \end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_C} & 0 & -\frac{fX_C}{s_x Z_C^2} \\ 0 & \frac{f}{s_y Z_C} & -\frac{fY_C}{s_y Z_C^2} \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \end{bmatrix} \tag{16}$$

$$\begin{bmatrix} \dot{x}_S \\ \dot{y}_S \end{bmatrix} = \begin{bmatrix} -\frac{fX_C Y_T}{s_x Z^2{}_C} & \left[\frac{fZ_T}{s_x Z_C} + \frac{fX_C X_T}{s_x Z^2{}_C}\right] & -\frac{fY_T}{s_x Z_C} \\ -\left[\frac{fZ_T}{s_y Z_C} + \frac{fY_C Y_T}{s_y Z^2{}_C}\right] & \frac{fY_C X_T}{s_y Z^2{}_C} & \frac{fX_T}{s_y Z_C} \end{bmatrix} \begin{bmatrix} \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \tag{17}$$

Figure 3 shows the components of the *resolvability ellipsoid* when a feature on an object is tracked. The translational components from (16) are shown by the ellipsoid on the left, and the rotational components from (17) are shown to the right. The object is 1m from the camera frame, f is 12mm, and the feature is located in the task frame at (0.1m,0,0). The pixel dimensions used are s_x=11µm and s_y=13µm. The two-dimensional ellipse that lies in a plane approximately parallel to the image plane indicates that depth cannot be resolved since $\Re(\boldsymbol{J}^T)$ has no basis vector which can describe motion along that direction. Since the feature lies on the x_S axis of the image plane, rotations about X_T cannot be resolved, however, rotational motion about Y_T and Z_T can be observed. In Figure 4 a second features is added at (-0.1m,0,0). The ability to resolve positions in X_T and Y_T increases with another independent data point, and differences in depth can now be resolved as the three dimensional shape of the ellipsoid indicates, although, depth resolution is much poorer than X_T and Y_T resolution. Rotations about X_T can still not be resolved, however, in Figure 5 the coordinates are moved to (0.1m,0.1m,0) and (-0.1m,0.1m,0), and the rotational ellipsoid becomes three dimensional, indicating that rotations about all axes can be resolved. It is important to note that translations

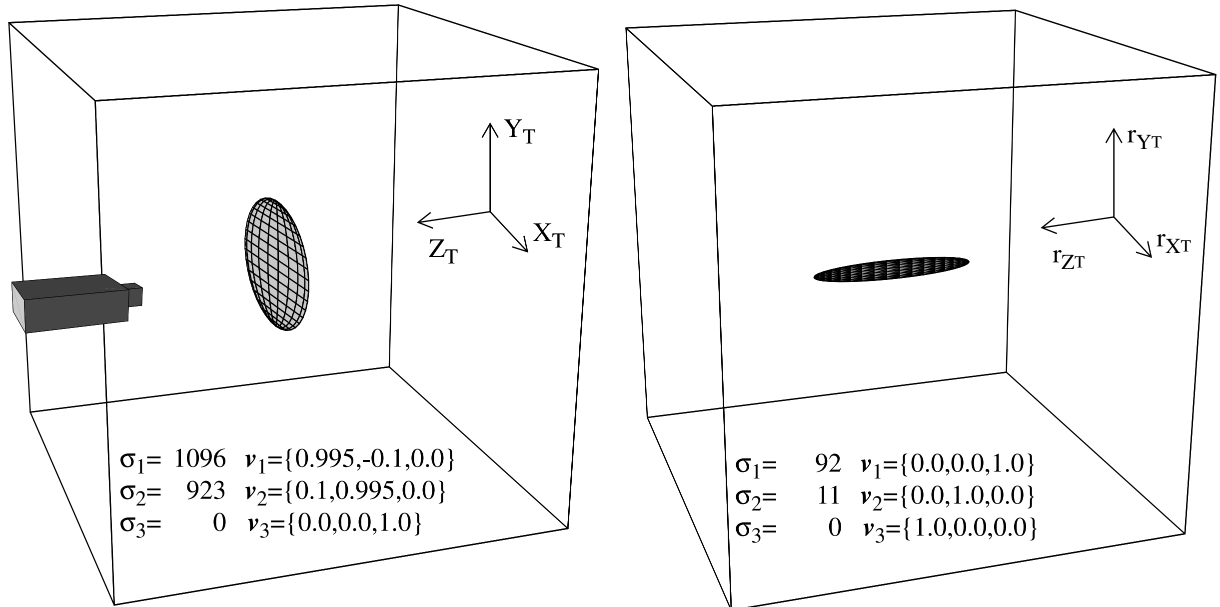

Figure 3: ***Resolvability Ellipsoids*****: single camera system, *f*=12mm, depth=1.0m, 1 feature located in the task frame at (0.1m,0,0).**

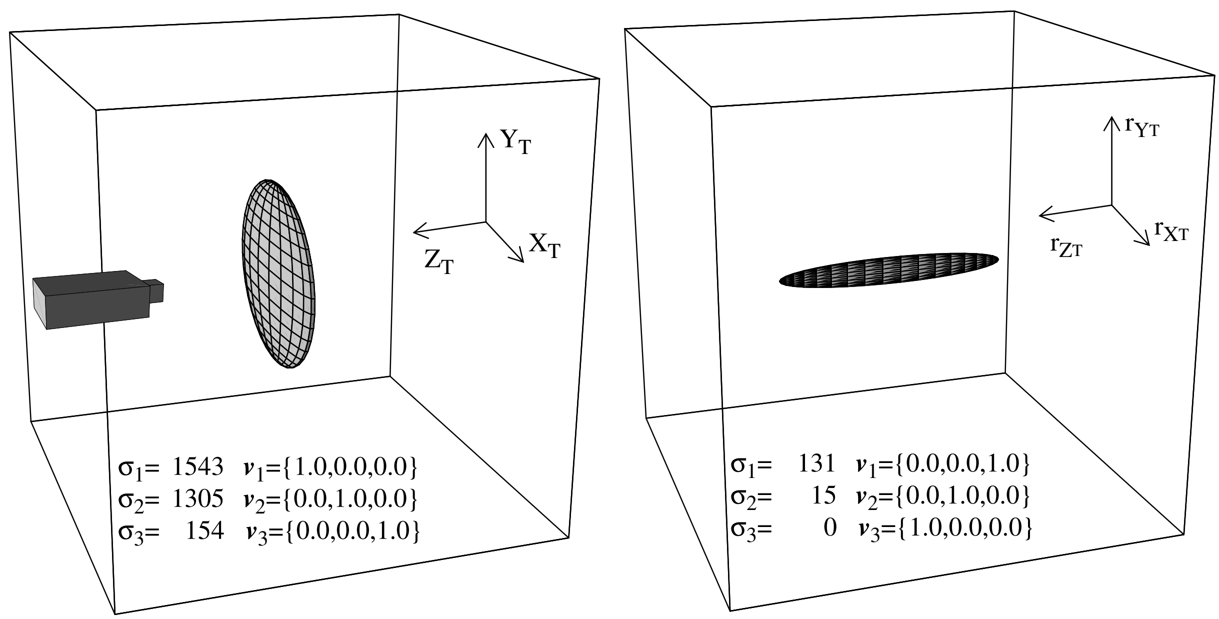

Figure 4: ***Resolvability Ellipsoids*****: single camera system, *f*=12mm, depth=1.0m, 2 features located in the task frame at (0.1m,0,0) and (-0.1m,0,0).**

and rotations are to be resolved independently using (16) and (17). If an SVD of (10) is performed, only four singular values would be non-zero indicating that translational and/or rotational motion about two axes cannot be resolved independent of the other axes.

In Figure 6, the depth to the object is 0.5m and f is 12mm. In Figure 7, the depth is 1.0m, and f is 24mm. It is interesting to note that the image plane coordinates for both ellipsoids are the same, but the smaller depth results in a *resolvability* that is approximately twice the *resolvability* of the example in which the focal length is doubled. This indicates that for a given magnification of the object (f/Z_C) reducing depth is preferable to increasing focal length when trying to resolve depth using a single camera.

Figure 8 is a plot of *resolvability* in depth versus focal length and depth of the object in the camera frame. The plot also shows the boundary of the image plane, beyond which the feature projections do not fall on the CCD of the camera. Pixel coordinates may vary along x_S from -255 to 256, and along y_S from -239 to 240. For the graph, an object with two features is observed. The task frame coordinates of the features are (0.05m,0,0) and (-0.05m,0,0). The graph shows the relationship be-

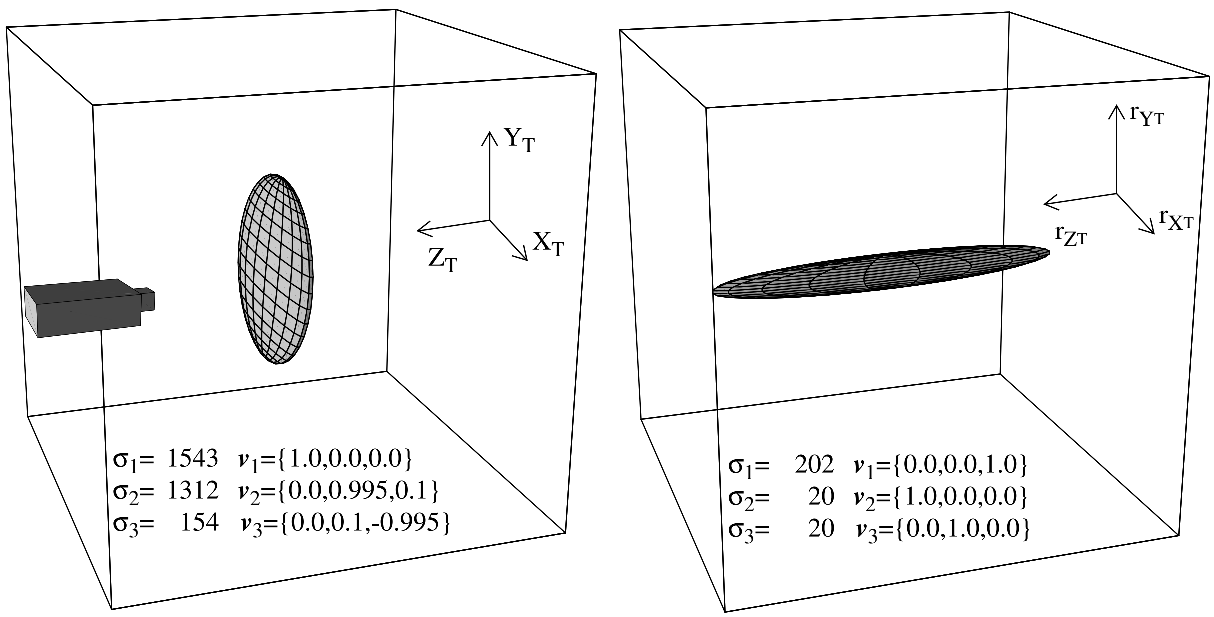

Figure 5: *Resolvability Ellipsoids*: single camera system, *f*=12mm, depth=1.0m, 2 features located in the task frame at (0.1m,0.1m,0), (-0.1m,0.1m,0).

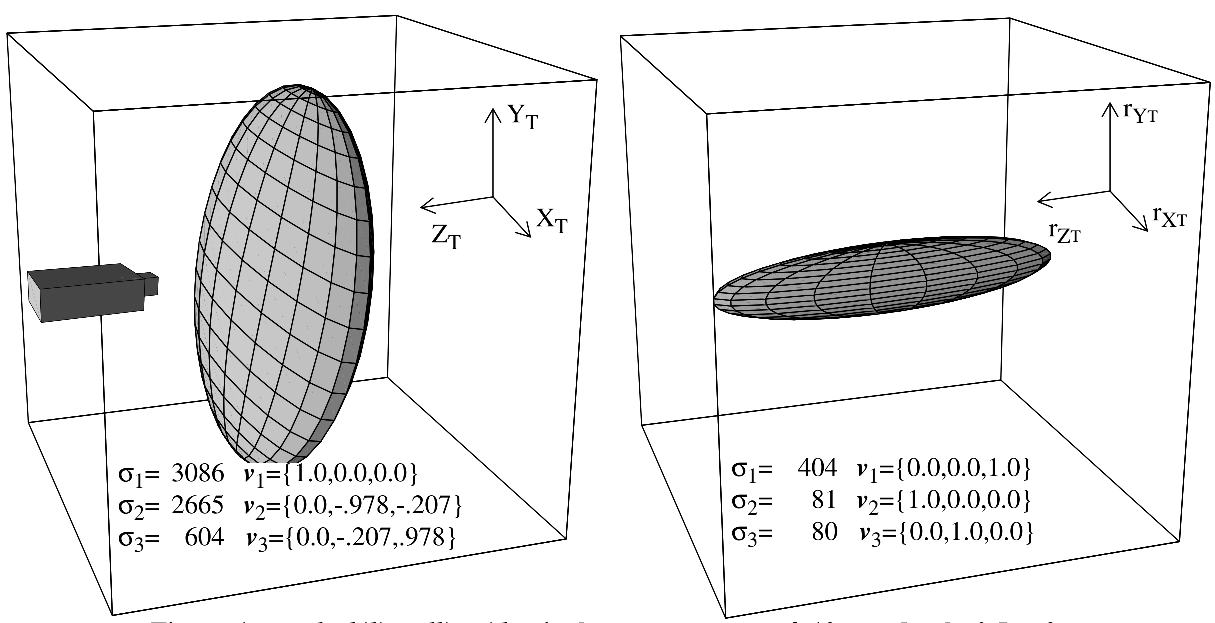

Figure 6: *Resolvability Ellipsoids*: single camera system, *f*=12mm, depth=0.5m, 2 features located in the task frame at (0.1m,0.1m,0) and (-0.1m,0.1m,0).

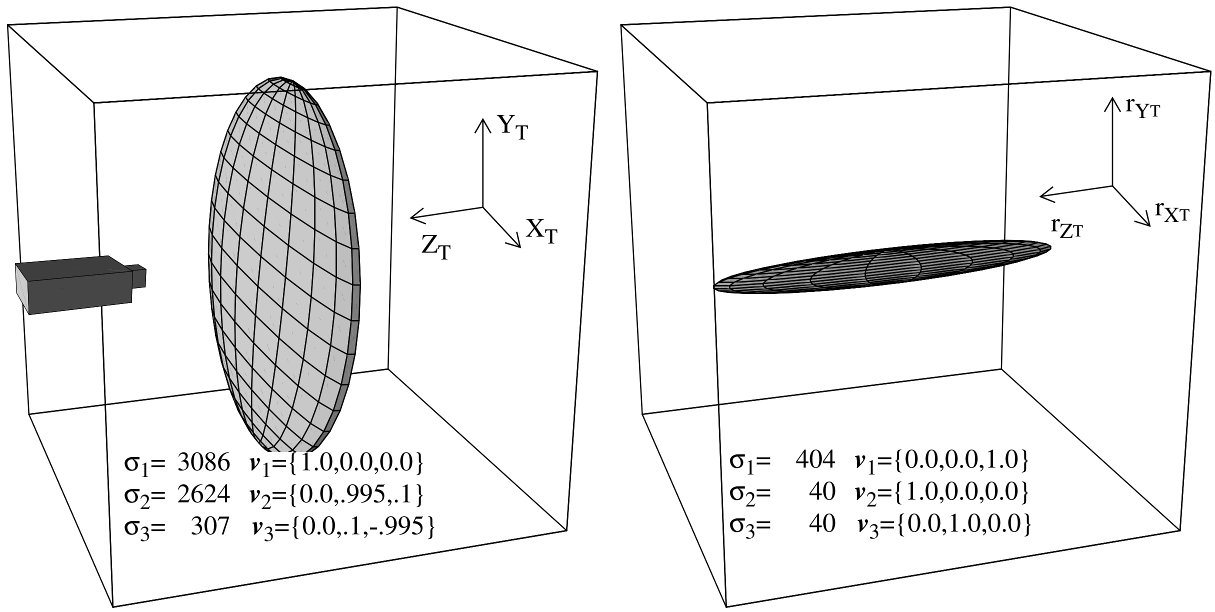

Figure 7: ***Resolvability Ellipsoids*: single camera system, *f*=24mm, depth=1.0m, 2 features located in the task frame at (0.1m,0.1m,0) and (-0.1m,0.1m,0).**

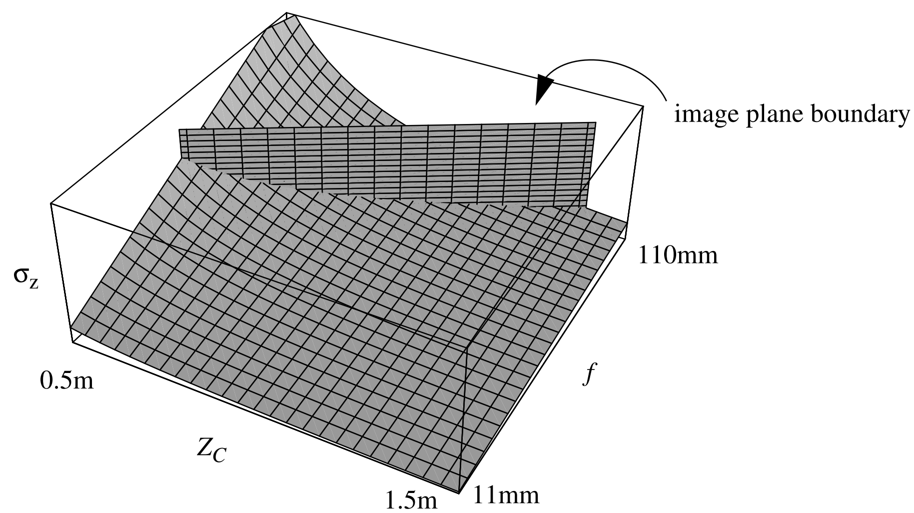

Figure 8: ***Resolvability* of depth versus depth of object and focal length for two features located in the task frame at (0.05m,0,0) and (-0.05m,0,0).**

tween depth, focal length, and *resolvability* in depth. From the plot one can observe that progressively smaller depths have progressively larger effects on *resolvability* in depth, while focal length tends to effect depth *resolvability* more linearly. In practice, depth becomes limited by the depth-of-field of the lens, and a trade-off must be made between focal length, depth, depth-of-field, and field-of-view.

The position of object features on the image plane effect *resolvability* as well. In Figure 9, a planar object with four features is moved across the image plane and the ability of the camera-lens system to resolve different orientations of the object are plotted. The plot illustrates that the ability to re-

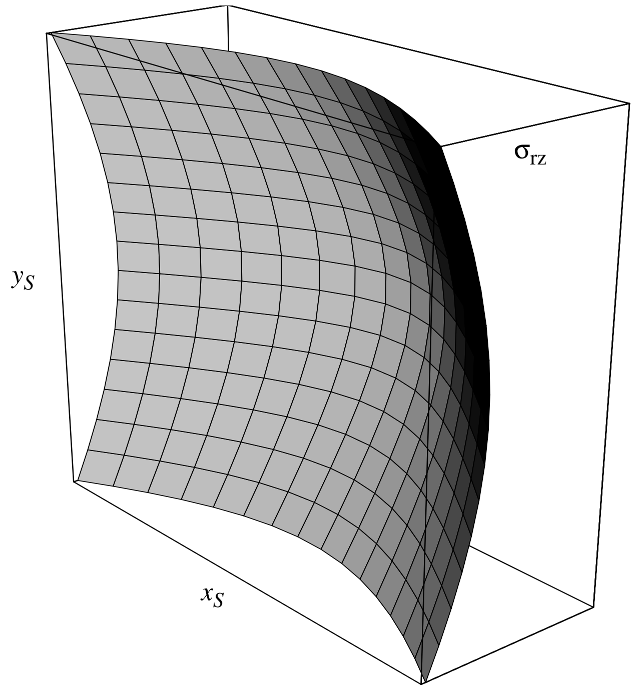

Figure 9: ***Resolvability*** **in orientation about** ***Z*** **versus center of object projection onto the image plane.**

solve orientations about the optical axis increases with the distance of the object from the optical axis, as long as the features projections remain on the image plane.

5. *Resolvability Ellipsoids* for Multiple Camera Systems

5.1 Stereo Pair - Parallel Optical Axes

In this section, the Jacobian for a stereo pair with parallel optical axes observing an object described relative to a task frame is derived. The derivation is based on equations for a stereo eye-in-hand system given in [6]. The term b represents the length of the baseline of the cameras, which is the line segment between camera focal points. The origin of the camera frame lies on the baseline midway between focal points, with the -Z axis pointing towards the object task frame, as shown in Figure 10. A feature on an object at ${}^{C}\boldsymbol{P}$ with coordinates (X_C, Y_C, Z_C) in the camera frame projects onto each image plane at

$$x_{Sl} = \frac{f\left(X_C + \frac{b}{2}\right)}{s_x Z_C} \qquad y_{Sl} = \frac{fY_C}{s_y Z_C} \tag{18}$$

$$x_{Sr} = \frac{f\left(X_C - \frac{b}{2}\right)}{s_x Z_C} \qquad y_{Sr} = \frac{fY_C}{s_y Z_C} \tag{19}$$

where it is assumed that f, s_x, and s_y are the same for both cameras.

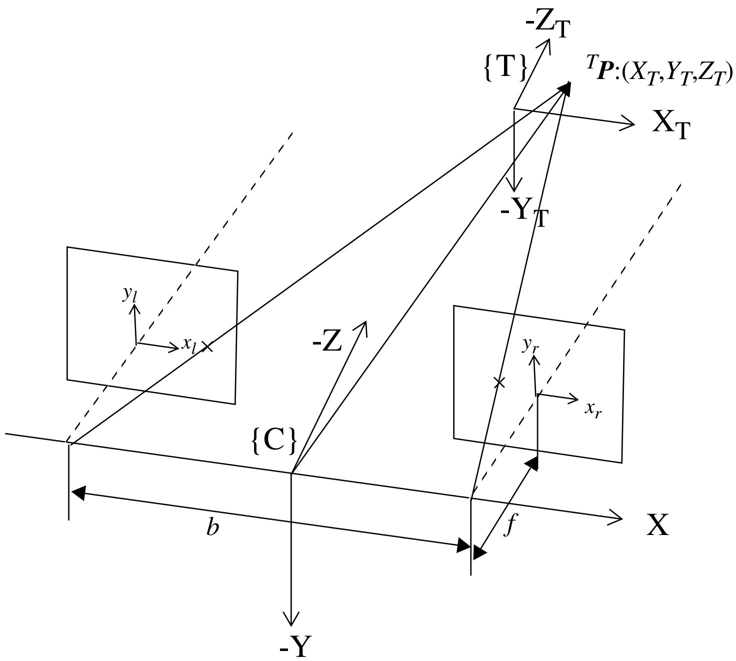

Figure 10: Coordinate frames for a stereo pair with parallel axes.

Through a similar derivation as the one presented in Section 2.2 (and with the same assumption that the camera and task frames are aligned), the relation between task space velocity and sensor space velocity can be written as

$$\begin{bmatrix} \dot{x}_{Sl} \\ \dot{y}_{Sl} \\ \dot{x}_{Sr} \\ \dot{y}_{Sr} \end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_C} & 0 & -\frac{f\left(X_C + \frac{b}{2}\right)}{s_x Z_C^2} & -\frac{f\left(X_C + \frac{b}{2}\right)Y_T}{s_x {Z^2}_C} & \left[\frac{fZ_T}{s_x Z_C} + \frac{f\left(X_C + \frac{b}{2}\right)\left(X_T + \frac{b}{2}\right)}{s_x {Z^2}_C}\right] & -\frac{fY_T}{s_x Z_C} \\ 0 & \frac{f}{s_y Z_C} & -\frac{fY_C}{s_y Z_C^2} & -\left[\frac{fZ_T}{s_y Z_C} + \frac{fY_C Y_T}{s_y {Z^2}_C}\right] & \frac{fY_C\left(X_T + \frac{b}{2}\right)}{s_y {Z^2}_C} & \frac{f\left(X_T + \frac{b}{2}\right)}{s_y Z_C} \\ \frac{f}{s_x Z_C} & 0 & -\frac{f\left(X_C - \frac{b}{2}\right)}{s_x Z_C^2} & -\frac{f\left(X_C - \frac{b}{2}\right)Y_T}{s_x {Z^2}_C} & \left[\frac{fZ_T}{s_x Z_C} + \frac{f\left(X_C - \frac{b}{2}\right)\left(X_T - \frac{b}{2}\right)}{s_x {Z^2}_C}\right] & -\frac{fY_T}{s_x Z_C} \\ 0 & \frac{f}{s_y Z_C} & -\frac{fY_C}{s_y Z_C^2} & -\left[\frac{fZ_T}{s_y Z_C} + \frac{fY_C Y_T}{s_y {Z^2}_C}\right] & \frac{fY_C\left(X_T - \frac{b}{2}\right)}{s_y {Z^2}_C} & \frac{f\left(X_T - \frac{b}{2}\right)}{s_y Z_C} \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \\ \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \qquad (20)$$

Equations (18)-(19) can be used to obtain this Jacobian in terms of the sensor coordinates and task coordinates, alone. Camera coordinates, including the depth Z_C, are not needed. We assume task coordinates are known for visual servoing, since we assume the object we are servoing on is of a known size and shape. The form for the Jacobian is

$$\begin{bmatrix} \dot{x}_{Sl} \\ \dot{y}_{Sl} \\ \dot{x}_{Sr} \\ \dot{y}_{Sr} \end{bmatrix} = \begin{bmatrix} \frac{d}{b} & 0 & -x_{Sl}\frac{s_x d}{fb} & -x_{Sl}\frac{s_x d}{fb}Y_T & \left[\frac{d}{b}Z_T + x_{Sl}\frac{s_x d}{fb}\left(X_T + \frac{b}{2}\right)\right] & -\frac{d}{b}Y_T \\ 0 & \frac{s_x d}{s_y b} & -y_{Sl}\frac{s_x d}{fb} & -\left[\frac{d}{s_y b}Z_T + y_{Sl}\frac{s_x d}{fb}Y_T\right] & y_{Sl}\frac{s_x d}{fb}\left(X_T + \frac{b}{2}\right) & \frac{s_x d}{s_y b}\left(X_T + \frac{b}{2}\right) \\ \frac{d}{b} & 0 & -x_{Sr}\frac{s_x d}{fb} & -x_{Sr}\frac{s_x d}{fb}Y_T & \left[\frac{d}{b}Z_T + x_{Sr}\frac{s_x d}{fb}\left(X_T - \frac{b}{2}\right)\right] & -\frac{d}{b}Y_T \\ 0 & \frac{s_x d}{s_y b} & -y_{Sr}\frac{s_x d}{fb} & -\left[\frac{s_x d}{s_y b}Z_T + y_{Sr}\frac{s_x d}{fb}Y_T\right] & y_{Sr}\frac{s_x d}{fb}\left(X_T - \frac{b}{2}\right) & \frac{s_x d}{s_y b}\left(X_T - \frac{b}{2}\right) \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \\ \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \tag{21}$$

where $d = x_{Sl} - x_{Sr}$ is the disparity of each corresponding feature point. This form does not require an explicit estimate of the depth Z_C.

The *resolvability ellipsoid* for this system is very similar to the previous ellipsoids for the monocular system in Section 4. The *resolvability* in depth for a stereo system with a baseline b is similar to the depth *resolvability* for a monocular system tracking two features separated in the task frame X_T-Y_T plane by a length b. A comparison of Figure 11 with Figure 4 illustrates this. Figure 12 shows the ellipsoids when tracking two features in each image. Figure 13 is a plot of *resolvability* in depth for the stereo system versus b and depth when tracking a single feature in both images. This plot shows that the effect of baseline length b on depth *resolvability* is very similar to the effect of focal length on depth *resolvability* as shown in Figure 8. As one would expect, the plot shows that a high depth *resolvability* is more easily achieved with small depths, rather than large baselines, subject to the boundary of the image plane.

5.2 Stereo Pair - Orthogonal Optical Axes

An orthogonal stereo pair is shown in Figure 14. If the axes are aligned as shown in the figure, the Jacobian mapping from task space to sensor space can be written as

$$\begin{bmatrix} \dot{x}_{Sl} \\ \dot{y}_{Sl} \\ \dot{x}_{Sr} \\ \dot{y}_{Sr} \end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_{Cl}} & 0 & -\frac{fX_{Cl}}{s_x Z_{Cl}^2} & -\frac{fX_{Cl}Y_T}{s_x Z_{Cl}^2} & \left[\frac{fZ_T}{s_x Z_{Cl}} + \frac{fX_{Cl}X_T}{s_x Z_{Cl}^2}\right] & -\frac{fY_T}{s_x Z_{Cl}} \\ 0 & \frac{f}{s_y Z_{Cl}} & -\frac{fY_{Cl}}{s_y Z_{Cl}^2} & -\left[\frac{fZ_T}{s_y Z_{Cl}} + \frac{fY_{Cl}Y_T}{s_y Z_{Cl}^2}\right] & \frac{fY_C X_T}{s_y Z_{Cl}^2} & \frac{fX_T}{s_y Z_{Cl}} \\ \frac{fX_{Cr}}{s_x Z_{Cr}^2} & 0 & \frac{f}{s_x Z_{Cr}} & \frac{fY_T}{s_x Z_{Cr}} & \frac{fX_{Cr}Z_T}{s_x Z_{Cr}^2} - \frac{fX_T}{s_x Z_{Cr}} & \frac{fX_{Cr}Y_T}{s_x Z_{Cr}^2} \\ \frac{fY_{Cr}}{s_y Z_{Cr}^2} & \frac{f}{s_y Z_{Cr}} & 0 & -\frac{fZ_T}{s_y Z_{Cr}} & \frac{fY_{Cr}Z_T}{s_y Z_{Cr}^2} & \frac{fY_{Cr}Y_T}{s_y Z_{Cr}^2} - \frac{fX_T}{s_y Z_{Cr}} \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \\ \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \tag{22}$$

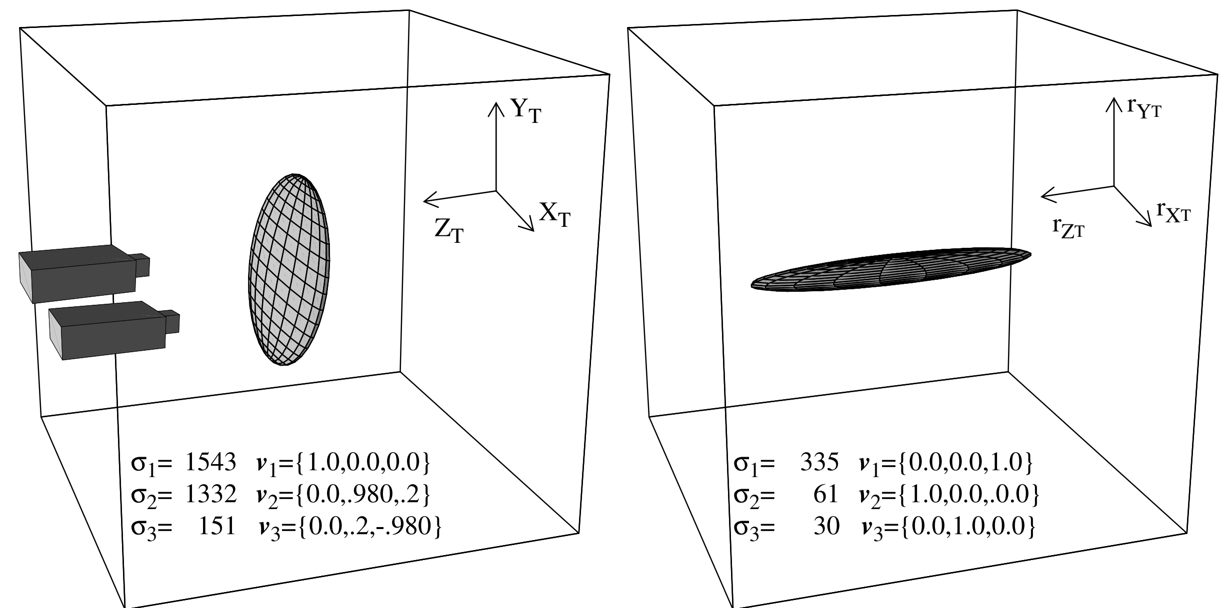

Figure 11: *Resolvability Ellipsoids*: stereo pair-parallel optical axes, *f*=12mm, *b*=20cm, depth=1.0m, 1 feature located in the task frame at (0,0.2m,0).

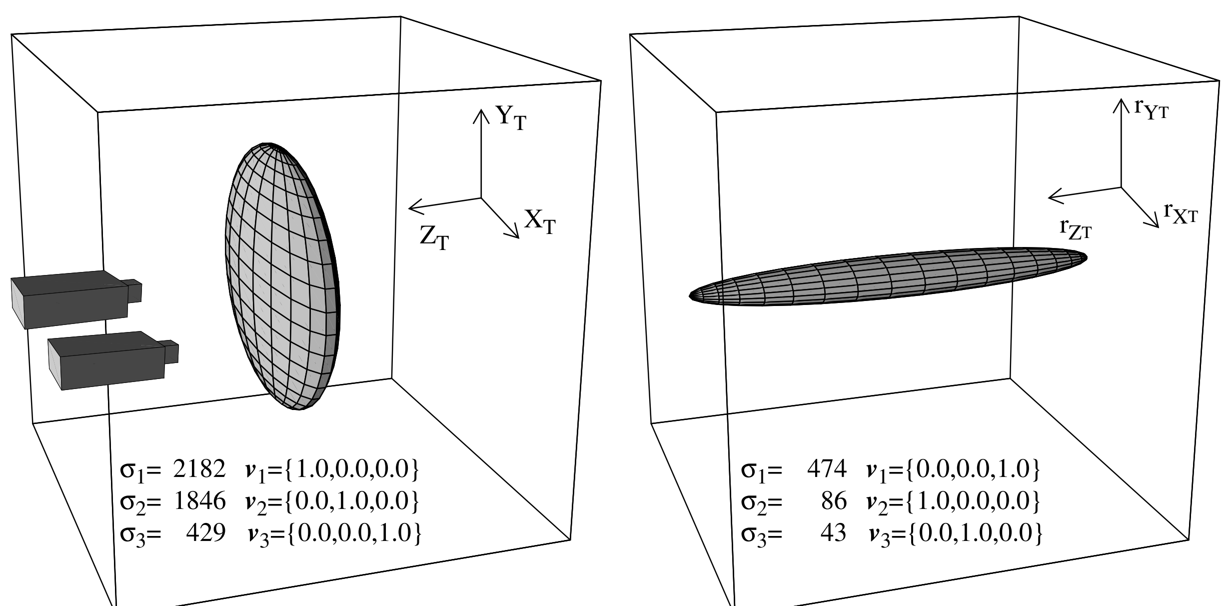

Figure 12: *Resolvability Ellipsoids*: stereo pair-parallel optical axes, *f*=12mm, *b*=20cm, depth=1.0m, 2 features located in the task frame at (0,0.2m,0) and (0,-0.2m,0).

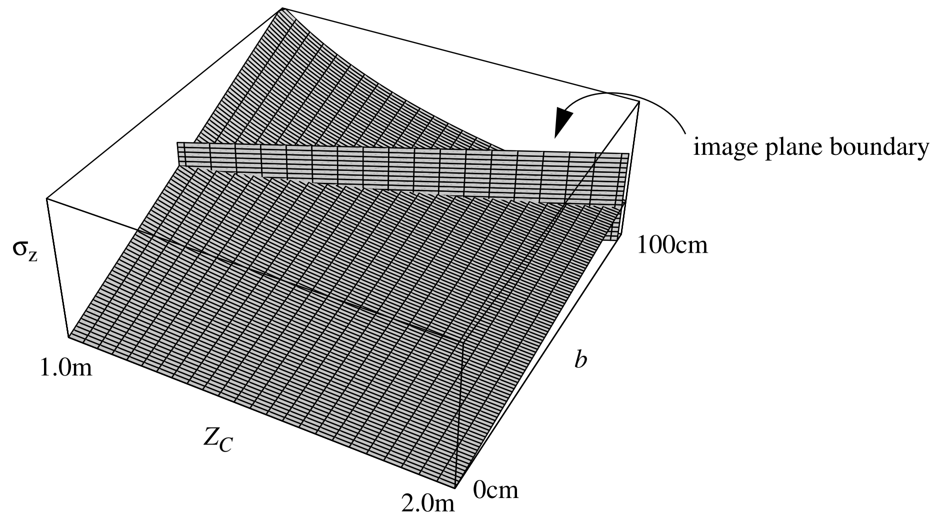

Figure 13: *Resolvability* in depth versus baseline length and depth of object for a stereo pair, parallel optical axes, *f*=12mm, and a single feature located at the origin of the task frame.

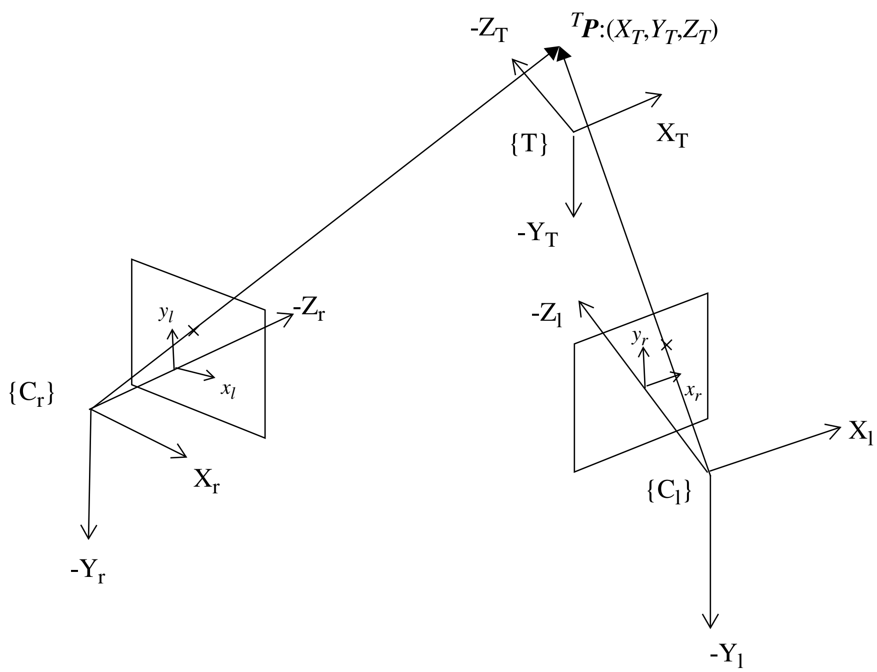

Figure 14: Coordinate frames for a stereo pair with orthogonal optical axes.

By rewriting this Jacobian in a form which uses sensor coordinates, the Jacobian becomes

$$\begin{bmatrix} \dot{x}_{Sl} \\ \dot{y}_{Sl} \\ \dot{x}_{Sr} \\ \dot{y}_{Sr} \end{bmatrix} = \begin{bmatrix} \frac{f}{s_x Z_{Cl}} & 0 & -\frac{x_{Sl}}{Z_{Cl}} & -\frac{x_{Sl} Y_T}{Z_{Cl}} & \left[\frac{f Z_T}{s_x Z_{Cl}} + \frac{x_{Sl} X_T}{Z_{Cl}}\right] & -\frac{f Y_T}{s_x Z_{Cl}} \\ 0 & \frac{f}{s_y Z_{Cl}} & -\frac{y_{sl}}{Z_{Cl}} & -\left[\frac{f Z_T}{s_y Z_{Cl}} + \frac{y_{Sl} Y_T}{Z_{Cl}}\right] & \frac{y_{Sl} X_T}{Z_{Cl}} & \frac{f X_T}{s_y Z_{Cl}} \\ \frac{x_{Sr}}{Z_{Cr}} & 0 & \frac{f}{s_x Z_{Cr}} & \frac{f Y_T}{s_x Z_{Cr}} & \frac{x_{Sr} Z_T}{Z_{Cr}} - \frac{f X_T}{s_x Z_{Cr}} & \frac{x_{Sr} Y_T}{Z_{Cr}} \\ \frac{y_{Sr}}{Z_{Cr}} & \frac{f}{s_y Z_{Cr}} & 0 & -\frac{f Z_T}{s_y Z_{Cr}} & \frac{y_{Sr} Z_T}{Z_{Cr}} & \frac{y_{Sr} Y_T}{Z_{Cr}} - \frac{f X_T}{s_y Z_{Cr}} \end{bmatrix} \begin{bmatrix} \dot{x}_T \\ \dot{y}_T \\ \dot{z}_T \\ \omega_{x_T} \\ \omega_{y_T} \\ \omega_{z_T} \end{bmatrix} \tag{23}$$

A feature tracked in each image gives a very well conditioned translational ellipsoid, as shown in Figure 15. When tracking two features in each camera, the rotational ellipsoid shown in Figure 16 is representative of the *resolvability* of the camera-lens arrangement.

6. Directing Camera-Lens Motion Using the *Resolvability Ellipsoid*

Our primary interest in *resolvability* is in using it to actively guide camera-lens motion for a single camera while performing visually servoed manipulation tasks. Static sensor placement algorithms can also use the directional qualities of the *resolvability ellipsoid* to direct the search for an optimal placement of the visual sensor. In order to guide camera-lens motion, it must be determined which camera-lens parameters can be freely changed. A gradient in parameter space can be determined based on these free parameters which directs camera-lens motion towards configurations with im-

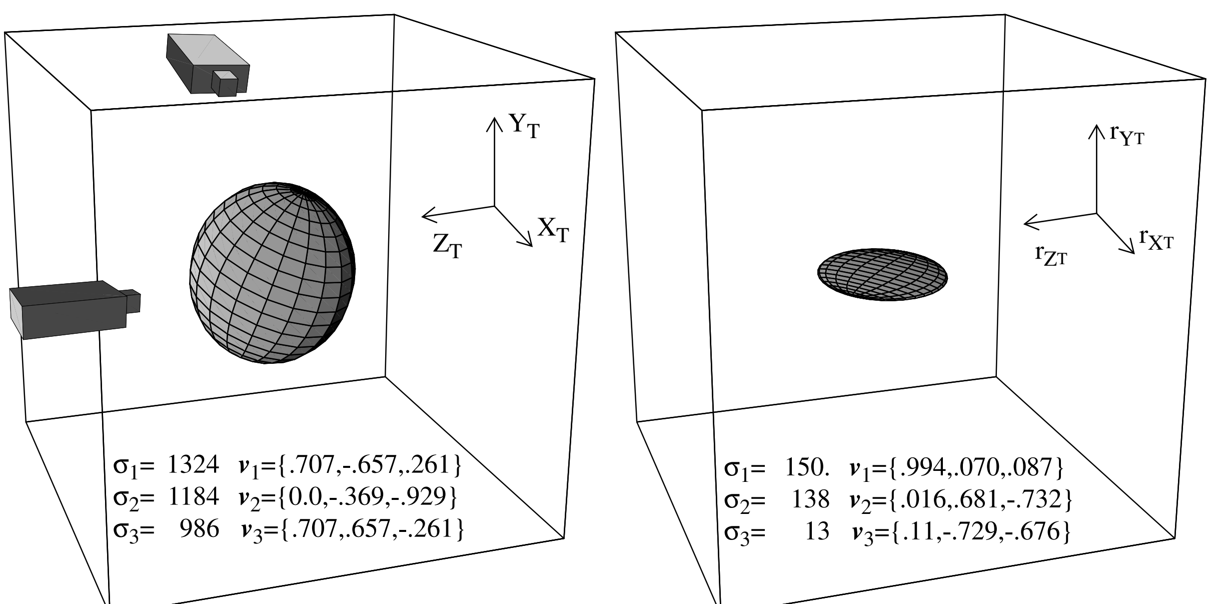

Figure 15: *Resolvability Ellipsoids*: stereo pair-perpendicular optical axes, *f*=12mm, depth=1.0m, 1 features located in the task frame at (0.1m,0.1m,0).

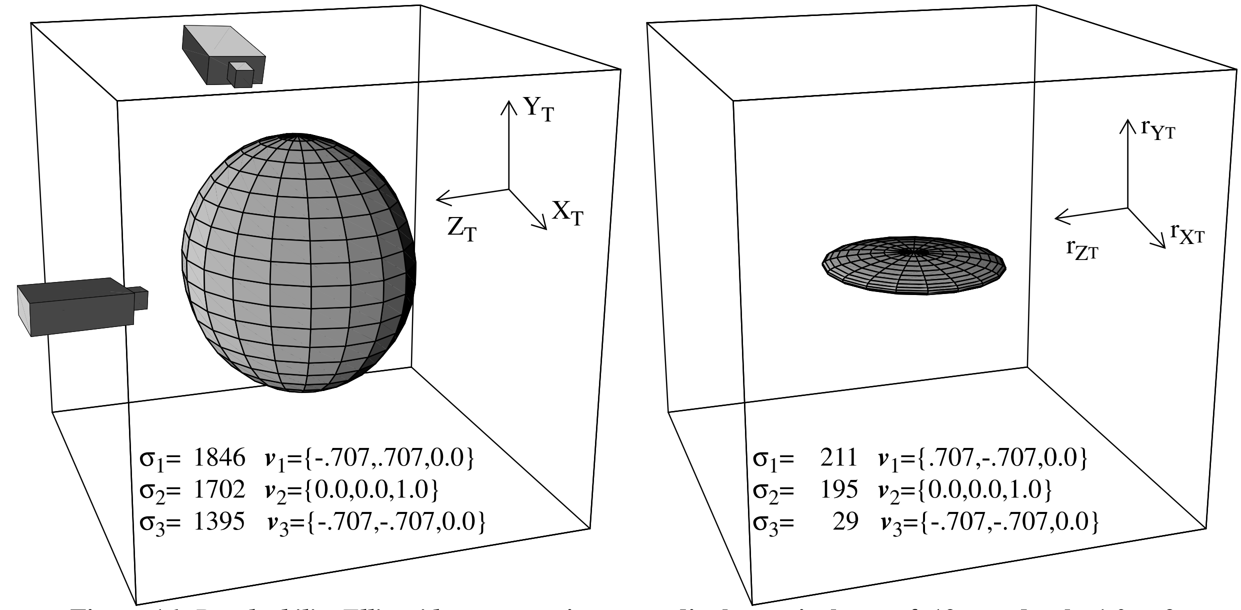

Figure 16: *Resolvability Ellipsoids*: stereo pair-perpendicular optical axes, *f*=12mm, depth=1.0m, 2 features located in the task frame at (-0.1m,0.1m,0), and (0.1m,-0.1m,-0.1m).

proved *resolvability*. For monocular and stereo tracking, resolution along the optical axis often has the poorest *resolvability*. In order to determine motions which will increase *resolvability* along the optical axis, the gradient of σ_Z with respect to the camera-lens parameter space can be calculated, where σ_Z represents the singular value which corresponds to the eigenvector along the optical axis. The gradient is a function of the intrinsic and extrinsic parameters of the camera that can vary, as well as the location of the features on the image plane. The gradient can be written as

$$\nabla_{\phi}\sigma_Z = \left[\frac{\partial\sigma_Z}{\partial x_C}\ \frac{\partial\sigma_Z}{\partial y_C}\ \frac{\partial\sigma_Z}{\partial z_C}\ \frac{\partial\sigma_Z}{\partial r_{XC}}\ \frac{\partial\sigma_Z}{\partial r_{YC}}\ \frac{\partial\sigma_Z}{\partial r_{ZC}}\ \frac{\partial\sigma_Z}{\partial f}\right] \tag{24}$$

The individual components of $\nabla_{\phi}\sigma_Z$ can be calculated numerically. Camera-lens motion can then be directed along $\nabla_{\phi}\sigma_Z$ in order to increase σ_Z subject to other sensor placement criteria such as depth-of-field and field-of-view constraints which simultaneously effect camera-lens motion. A technique for integrating this gradient into the visual tracking control law can be found in [8].

7. Results

7.1 Experimental Setup

To experimentally demonstrate the implications of *resolvability*, a manipulator was visually servoed by an active camera-lens system at various depths and focal lengths. Comparisons of the performance of the controller at different camera-lens configurations are presented. The experimental setup is shown in Figure 17. Visual servoing algorithms have been implemented on a robotic assembly system consisting of three Puma 560's called the Troikabot. One of the Pumas has a Sony XC-77RR camera with a zoom lens mounted at its endeffector. The camera is connected to a Datacube Maxtower Vision System.The Pumas are controlled from a VME bus with two Ironics

Figure 17: Experimental Setup.

IV-3230 (68030 CPU) processors, an IV-3220 (68020 CPU) 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 Chimera3 real-time operating system [10]. A diagram of the hardware setup is given in Figure 18. The vision system VME communicates with the controller VME using BIT3 VME-to-VME adapters. The Datacube Maxtower Vision System calculates the optical flow of the features using a Sum-of-Squares Differences algorithm. A special high performance floating- point processor on the Datacube is used to calculate the optical flow of the feature, and a 68030 board, also on the vision system, computes the control inputs. An image can be grabbed and displacements for up to ten 16x16 feature templates in the scene can be determined at 30Hz. The control input for tracking is sent to the Mercury floating point processor at 30Hz, where a cartesian controller calculates the proper joint control commands.

7.2 Visual Tracking Controller

The state equation for the visual servoing system is created by discretizing (9) and rewriting the discretized equation as

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + T\boldsymbol{J}(k)\boldsymbol{u}(k) \tag{25}$$

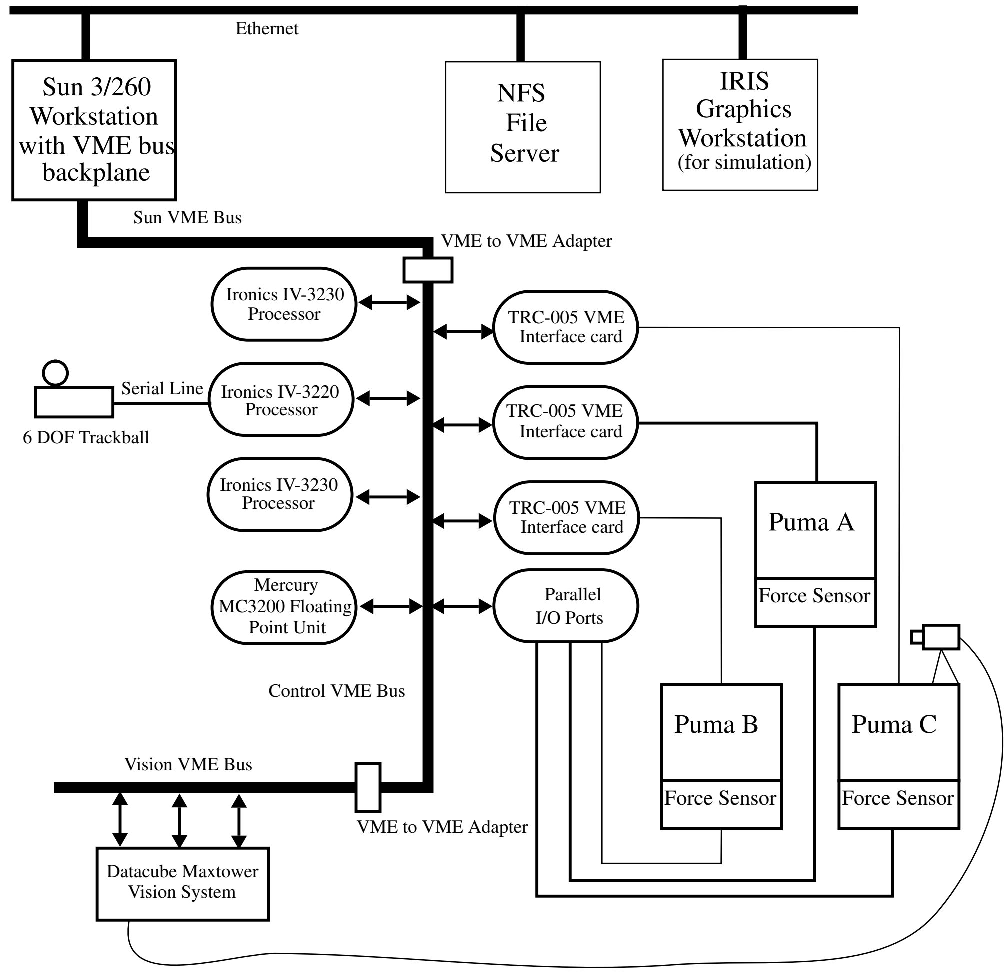

Figure 18: The Troikabot System architecture.

where M is the number of features being tracked, $\boldsymbol{A}=\mathbf{I}_{2M}$, $\boldsymbol{x}(k)\in R^{2M}$, T is the sampling period of the vision system, and $\boldsymbol{u}(k) = \left[\dot{x}_T\ \dot{y}_T\ \dot{z}_T\ \omega_{x_T}\ \omega_{y_T}\ \omega_{z_T}\right]^T$, the manipulator endeffector velocity.

A control strategy can be derived using the controlled active vision paradigm [9]. The control objective of the visual tracking system is to control endeffector 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 at each time instant. The objective function places a cost on differences in feature positions from desired positions, as well as a cost on providing control input, and is of the form

$$F(k+1) = [\boldsymbol{x}(k+1) - \boldsymbol{x}_D(k+1)]^T \boldsymbol{Q}\,[\boldsymbol{x}(k+1) - \boldsymbol{x}_D(k+1)] + \boldsymbol{u}^T(k)\boldsymbol{R}\boldsymbol{u}(k) \tag{26}$$

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

$$\boldsymbol{u}(k) = -(\boldsymbol{J}^T(k)\boldsymbol{Q}\boldsymbol{J}(k) + \boldsymbol{R})^{-1}\boldsymbol{J}^T(k)\boldsymbol{Q}\,[\boldsymbol{x}(k) - \boldsymbol{x}_D(k+1)] \quad (27)$$

The weighting matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ allow the user to place more or less emphasis on the feature error and the control input. Their selection effects the response and stability of the tracking system. The $\boldsymbol{Q}$ matrix must be positive semi-definite, and $\boldsymbol{R}$ must be positive definite for a bounded response. Although no standard procedure exists for choosing the elements of $\boldsymbol{Q}$ and $\boldsymbol{R}$, general guidelines can be found in [9].

For the experiments performed, only the translational *resolvability* components were considered, therefore, the Jacobian $\boldsymbol{J}$ given by (16) is used in the control law (27) and $\boldsymbol{u}(k) = \begin{bmatrix}\dot{x}_T & \dot{y}_T & \dot{z}_T\end{bmatrix}^T$.

7.3 Experimental Results

In order to demonstrate the implications of *resolvability* for visual servoing, an object was placed in the gripper of the manipulator such that two features on the object located approximately 4cm apart fell in a plane parallel to the image plane of the camera. The desired coordinates of the image plane projection of the features were commanded such that the distance between the features should decrease by 10 pixels, but the center of mass of the features should remain constant. Figure 19 illustrates how the object might initially appear on the image plane and how the object should appear after the desired locations of the features coordinates have been achieved. In order

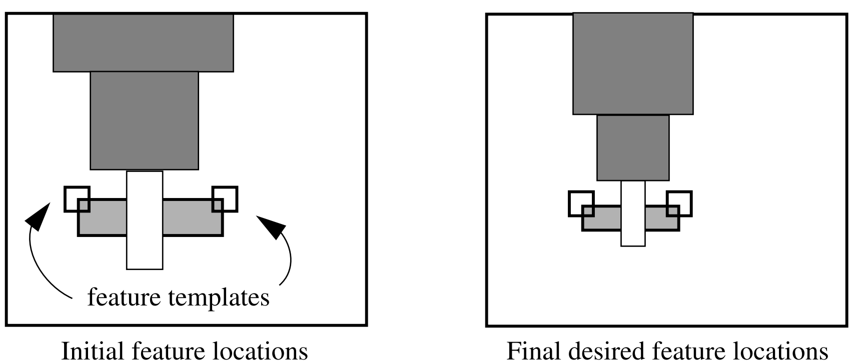

Figure 19: Images showing initial and desired locations of feature template locations for the experiments performed.

for the tracking controller to command the manipulator to move the image plane coordinates of the features to their desired locations, the manipulator must move parallel to the optical axis of the camera. This is the poorest direction of translational *resolvability* for a single camera system. The change in the depth of the object with respect to the camera frame was recorded as a function of time and results are plotted in Figures 20 and 21. The greater the depth *resolvability* of a given camera-lens configuration, the smaller the required translation of the object to effect a 10 pixel change in object length on the image plane.

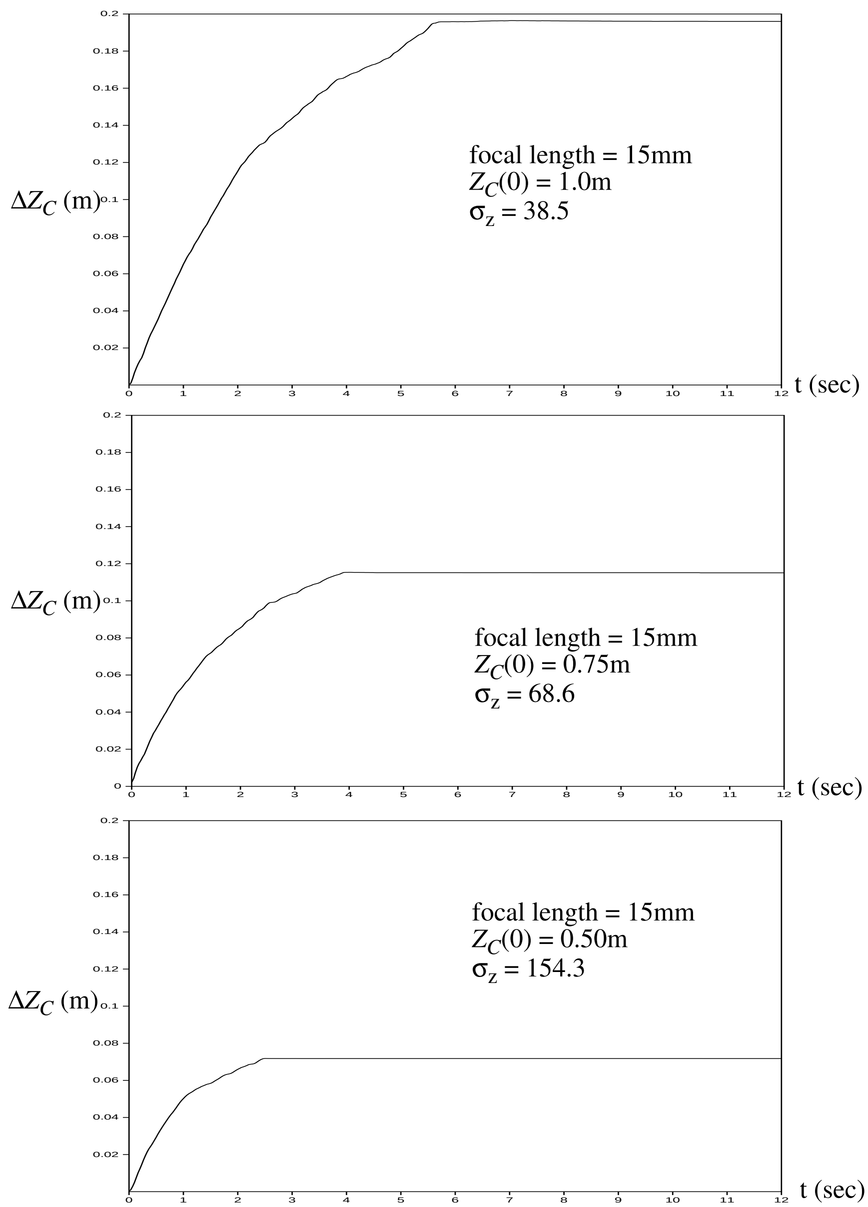

Figure 20: Change in object depth required to effect a 10 pixel change in projected object length for three different initial depths and a focal length of 15mm.

In Figure 20 the change in depth of the object is plotted versus time for three trials in which the focal length of the camera is held constant at 15mm and the initial depth of the object with respect to the camera varies from 50cm to 1m. The increase in *resolvability* with decreasing depth is made apparent by the smaller changes in depth required to move the object so that the length between features on the image plane decreases by 10 pixels. At 1m, the object must translate 19.6cm in order to effect a 10 pixel change on the image plane. At 75cm, only 11.5cm of translational motion is required, while at 50cm the object moves 7.2cm. Also shown on the figure is the estimated *resolvability* along Z based on the singular value along Z of (16). Since *resolvability* varies continuously across camera intrinsic and extrinsic parameters, the trials will not give a completely accurate

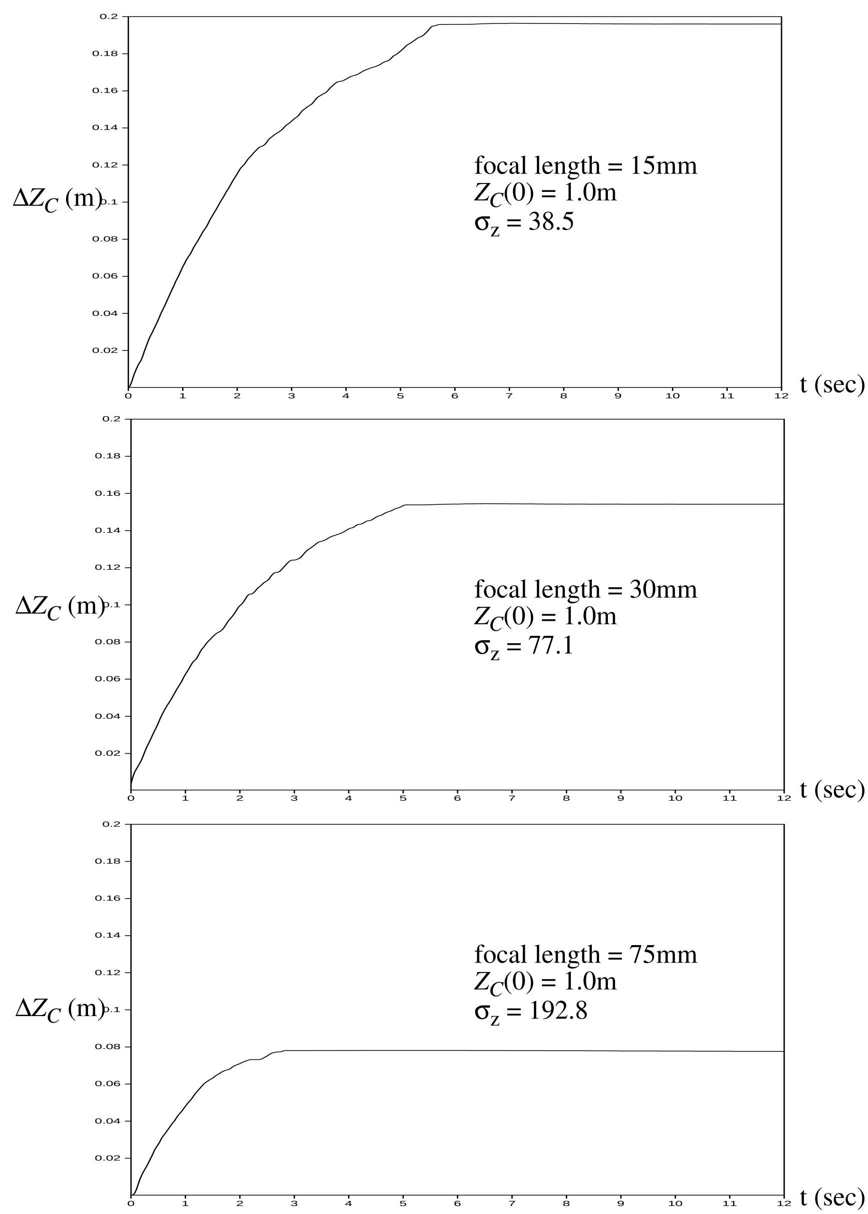

Figure 21: Change in object depth required to effect a 10 pixel change in projected object length for three different focal lengths at an initial depth of 1.0m.

picture of the instantaneous *resolvability* for each of the three initial camera configurations based on change in object depth. It can be seen, however, that there is general agreement in the relative change in the calculated *resolvability* and the relative change in the difference in translation along Z.

Figure 21 shows translation motion along Z when the object servos form the same initial depth but with a different focal length. All trials were run from an initial object depth of 1m. For a focal length of 15mm, the depth change was 19.6cm, for 30mm the depth change was 15.7cm, and for 75mm the depth change was 7.7cm. It should be noted that the zoom lens was not calibrated. The focal lengths are estimates obtained from the scale inscribed by the manufacturer on the lens body, and is undoubtedly highly inaccurate. Nevertheless, the trends one would expect in *resolvability* at various focal lengths versus required changes in depth to effect a 10pixel change in feature distance are apparent.

8. Conclusion

The directional nature of the *resolvability ellipsoid* makes this sensor placement criterion particularly useful for guiding visual sensor motion in real-time, or as an aid in determining the placement of static sensors. In this paper we have shown that *resolvability* can effectively represent the ability of several different sensor configurations to resolve translational and rotational positions in objects being observed. For visual servoing, this concept directly relates to the accuracy with which a manipulator can move an object to some desired goal position and orientation. Experimental results using an uncalibrated zoom lens demonstrate that *resolvability* can be used to direct camera-lens motion in particular directions in order to increase the ability of a visually servoed manipulator to perform precise manipulation tasks.

9. References

[1] C.K. Cowan and P.D. Kovesi, "Automatic Sensor Placement from Vision Task Requirements," In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5), 407-416, 1988.

[2] C.K. Cowan and A. Bergman, "Determining the Camera and Light Source Location for a Visual Task," In: *Proc. of the 1989 IEEE Int. Conf. on Robotics and Automation*, 509-514, 1989.

[3] S. Das and N. Ahuja, "A Comparative Study of Stereo, Vergence, and Focus as Depth Cues for Active Vision," In: *Proc. of the 1992 Computer Vision and Pattern Recognition Conference (CVPR-92)*, 194-199, 1993.

[4] A. Ghosal and B. Roth, "A New Approach for Kinematic Resolution of Redundancy," in *Int. J. of Robotics Res.*, 7(2), 22-35, 1988.

[5] V.C. Klema and A.J. Laub, "The Singular Value Decomposition: Its Computation and Some Applications," In: *IEEE Transactions on Automatic Control*, 25(2), 164-176, 1980.

[6] N. Maru, H. Kase, S. Yamada, A. Nishikawa, and F. Miyazaki, "Manipulator Control by Visual Servoing with the Stereo Vision," In: *Proc. of the 1993 IEEE/RSJ Int. Conf. on Intelligent Robots and System (IROS-93)*, 1866-1870, 1993.

[7] B. Nelson and M. Donath, "Optimizing the Location of Assembly Tasks in a Manipulator's Workspace," In: *Journal of Robotics Systems*, 7(6), 791-811, 1990.

[8] B. Nelson and P.K. Khosla, "Integrating Sensor Placement and Visual Tracking Strategies," In: *EXPERIMENTAL ROBOTICS III: The Third International Symposium*, *Kyoto, October 28-30 1993*, eds. T. Yoshikawa and F. Miyazaki, Springer-Verlag London Ltd., London, 1993.

[9] N.P. Papanikolopoulos, B. Nelson, and P.K. Khosla, "Monocular 3-D Visual Tracking of a Moving Target by an Eye-in-Hand Robotic System," In: *Proc. of the 31st IEEE Conf. on Decision and Control (31stCDC)*, 3805-3810, 1992.

[10] D.B. Stewart, D.E. Schmitz, and P.K. Khosla, "The Chimera II Real-Time Operating System for Advanced Sensor-Based Control System," In: *IEEE Trans. Sys., Man and Cyb.*, **22**, 1282-1295, 1992.

[11] K. Tarabanis, R.Y. Tsai, and P.K. Allen, "Satisfying the Resolution Constraint in the "MVP" Machine Vision Planning System," In: *Proc. of the 1990 Darpa Image Understanding Workshop*, 850-860, 1990.

[12] K. Tarabanis, R.Y. Tsai, and P.K. Allen, "Automated Sensor Planning for Robotic Vision Tasks," In: *Proc. of the 1991 IEEE Int. Conf. on Robotics and Automation*, 76-82, 1991.

[13] S. Yi, R.M. Haralick, and L.G. Shapiro, "Automatic Sensor and Light Positioning for Machine Vision," In: *Proc. of the 10th Int. Conf. on Pattern Recognition*, 55-59, 1990.

[14] T. Yoshikawa, "Manipulability of Robotic Mechanisms," In: *Robotics Research 2*, eds. H. Hanafusa and H. Inoue, 439-446, MIT Press, Cambridge, MA, 1985.