A

Active Stereo Head Calibration

This appendix accompanies the discussion on active vision in Section 2.2.3, and describes a calibration procedure for the Biclops stereo head. Biclops is a commercially available robotic pan-tilt-verge camera platform that enables the viewpoint of the cameras to be controlled for active tracking. Biclops provides active vision for the experimental service robot detailed in Chapter 7 (see Figure 7.1).

When implementing model-based vision on an active vision head, accurate feedback of the pan, tilt and verge angles (see Figure 2.5) are important to determine the current camera projection matrices. The Biclops head reports these quantities in their measured units of encoder counts. Manual calibration is thus required to transform these values into meaningful units such as radians or degrees. Fortunately, a linear relationship exists between encoder count and angular position for the pan and tilt axes, and the scale factor can be calculated by recording the encoder value at the limits of motion for each axis and dividing by the factory specified angular range. Conversely, the verge mechanism of the Biclops head does not possess such as linear relationship between encoder counts and angular travel. As illustrated in Figure A.1, the Biclops verges the cameras from a single motor via a worm gear, nut and transmission arm with leaf spring hinges. As the nut travels along the worm gear, the cameras rotate with a symmetrical, non-linear change in verge angle and baseline. Manufacturer specifications for these relationships are not provided, but the verge angle can be calibrated manually using the image-based procedure described below.

The calibration procedure is based on placing a coloured target at varying known distances from the head and controlling the verge angle to drive the measured colour centroid to the centre of the left and right image planes. The smooth pursuit control laws given by equations (2.31)-(2.33), with ${}^{L}\mathbf{x} = ({}^{L}x, {}^{L}y)^{\top}$ and ${}^{R}\mathbf{x} = ({}^{R}x, {}^{R}y)^{\top}$ representing the measured colour centroid, are used to centre the target at each known distance d_i, and the verge encoder value at the completion of each trial is also recorded. Using a manually measured stereo baseline $2b$, the verge angle v_i at the ith target position can be calculated as

$$v_i = \tan^{-1}(b/d_i) \tag{A.1}$$

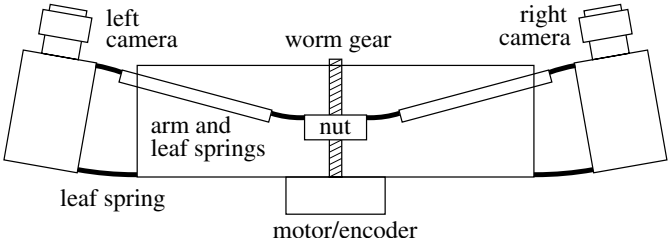


Fig. A.1. Verge mechanism for Biclops head (top view).

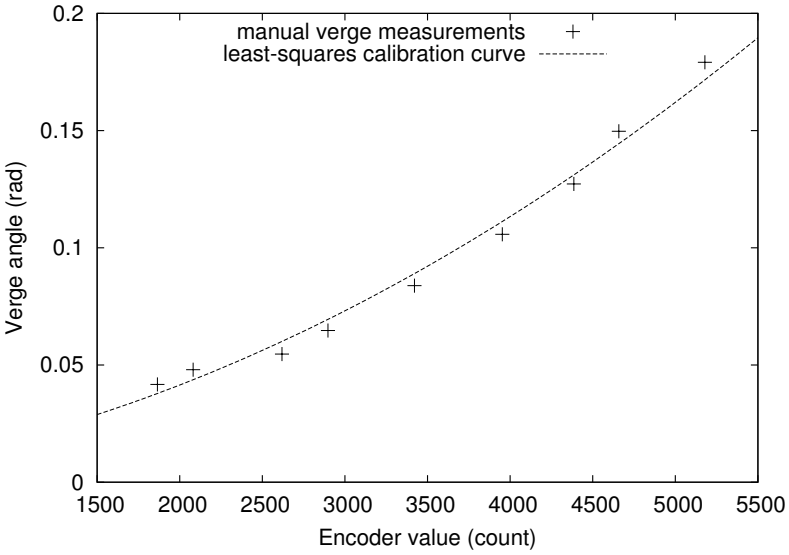


Fig. A.2. Calibration of verge angle.

Figure A.2 plots the measured verge angle (from equation (A.1)) against the corresponding encoder value, when this procedure was performed on the experimental Biclops head. Finally, an approximate quadratic calibration curve is calculated using a least squares fit, as shown by the dashed line in Figure A.2. While this approach only provides approximate verge and baseline calibration, Chapter 6 shows how the uncertainty in these parameters can be compensated during hybrid position-based visual servoing.

B

Light Stripe Validation and Reconstruction

This appendix presents detailed calculations of the theoretical results for the robust stereoscopic light stripe scanner described in Chapter 3.

B.1 Optimization of Light Plane Error Function

In Section 3.2.3, the following cost function is proposed to determine whether a pair of stereo measurements, ${}^{L}\mathbf{x}$ and ${}^{R}\mathbf{x}$, correspond to a point $\widehat{\mathbf{X}}$ on the light plane:

$$E = d^2({}^{L}\mathbf{x}, {}^{L}\mathrm{P}\widehat{\mathbf{X}}) + d^2({}^{R}\mathbf{x}, {}^{R}\mathrm{P}\widehat{\mathbf{X}}) \tag{B.1}$$

where $d(\mathbf{x}_1, \mathbf{x}_2)$ is the Euclidean distance between $\mathbf{x}_1$ and $\mathbf{x}_2$ and ${}^{L,R}\mathrm{P}$ are the projection matrices of the stereo cameras. The reconstruction $\widehat{\mathbf{X}}$ is constrained to the light plane $\mathbf{\Omega}$ by:

$$\mathbf{\Omega}^{\top}\widehat{\mathbf{X}} = 0 \tag{B.2}$$

The minimum error E^* and optimal reconstruction $\widehat{\mathbf{X}}$ for a given measurement pair is determined by the constrained optimization of equation (B.1) with respect to the constraint in (B.2). As noted, a direct optimization is analytically cumbersome, but the problem can be reduced to an unconstrained optimization by determining the direct relationship between projections ${}^{L}\hat{\mathbf{x}}$ and ${}^{R}\hat{\mathbf{x}}$ for points on the light plane.

The following formulation is based on the observation that $\widehat{\mathbf{X}}$ lies at the intersection of the light plane with the image ray back-projected through ${}^{R}\hat{\mathbf{x}}$ (or ${}^{L}\hat{\mathbf{x}}$). Plücker matrices provide a concise notation for the intersection of planes and lines (see [52]). If $\mathbf{A}$ and $\mathbf{B}$ represent the homogeneous vectors of two points on a line, the Plücker matrix L describing the line is

$$\mathrm{L} = \mathbf{A}\mathbf{B}^{\top} - \mathbf{B}\mathbf{A}^{\top} \tag{B.3}$$

Then, the intersection $\mathbf{X}$ of a plane $\mathbf{\Omega}$ and the line described by L is simply

$$\mathbf{X} = \mathrm{L}\mathbf{\Omega} = (\mathbf{A}\mathbf{B}^{\top} - \mathbf{B}\mathbf{A}^{\top})\mathbf{\Omega} \tag{B.4}$$

Now, the Plücker matrix $\mathbf{L}_R$ for the back-projection of ${}^R\hat{\mathbf{x}}$ can be constructed from two known points on the ray: $\mathbf{C}_R$ and ${}^R\mathrm{P}^+{}^R\hat{\mathbf{x}}$, where ${}^R\mathrm{P}^+$ is the pseudo-inverse of the camera projection matrix ${}^R\mathrm{P}$, given by equation (2.35). Applying these to equation (B.3), the Plücker matrix L_R for the back-projection of ${}^R\hat{\mathbf{x}}$ is

$$\mathrm{L}_R = \mathbf{C}_R({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})^\top - ({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})\mathbf{C}_R^\top \tag{B.5}$$

The intersection of L_R with the laser plane $\boldsymbol{\Omega}$, can now be expressed using equation (B.4) as:

$$\widehat{\mathbf{X}} = \mathrm{L}_R\boldsymbol{\Omega} = [\mathbf{C}_R({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})^\top - ({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})\mathbf{C}_R^\top]\boldsymbol{\Omega} \tag{B.6}$$

Equation (B.6) is the result quoted in equation (3.9) of Section 3.2.3. Finally, the left projection ${}^L\hat{\mathbf{x}}$ corresponding to ${}^R\hat{\mathbf{x}}$ is obtained by projecting $\widehat{\mathbf{X}}$, given by equation (B.6), via ${}^L\mathrm{P}$:

$$\begin{aligned} {}^L\hat{\mathbf{x}} &= {}^L\mathrm{P}\widehat{\mathbf{X}} \\ &= {}^L\mathrm{P}[\mathbf{C}_R({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})^\top - ({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})\mathbf{C}_R^\top]\boldsymbol{\Omega} \\ &= {}^L\mathrm{P}\mathbf{C}_R({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})^\top\boldsymbol{\Omega} - {}^L\mathrm{P}({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})\mathbf{C}_R^\top\boldsymbol{\Omega} \end{aligned} \tag{B.7}$$

Using the identity $({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})^\top\boldsymbol{\Omega} = \boldsymbol{\Omega}^\top({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}})$ (since both sides are scalar), and noting that $(\mathbf{C}_R^\top\boldsymbol{\Omega})$ is scalar, the common factors are collected to simplify the above expression to

$$\begin{aligned} {}^L\mathbf{x} &= {}^L\mathrm{P}(\mathbf{C}_R\boldsymbol{\Omega}^\top)({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}}) - {}^L\mathrm{P}(\mathbf{C}_R^\top\boldsymbol{\Omega})({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}}) \\ &= \left({}^L\mathrm{P}[\mathbf{C}_R\boldsymbol{\Omega}^\top - (\mathbf{C}_R^\top\boldsymbol{\Omega})\mathrm{I}]{}^R\mathrm{P}^+\right)\hat{\mathbf{x}}_R \end{aligned} \tag{B.8}$$

Equation (B.8) is the desired relationship between projections ${}^L\hat{\mathbf{x}}$ and ${}^R\hat{\mathbf{x}}$ of points on the light plane $\boldsymbol{\Omega}$, and is the result quoted in equation (3.10) of Section 3.2.3. Finally, the error function in equation (B.1) can be written as

$$E = d^2({}^L\mathbf{x}, \mathrm{H}{}^R\hat{\mathbf{x}}) + d^2({}^R\mathbf{x}, {}^R\hat{\mathbf{x}}) \tag{B.9}$$

with $\mathrm{H} = {}^L\mathrm{P}[\mathbf{C}_R\boldsymbol{\Omega}^\top - (\mathbf{C}_R^\top\boldsymbol{\Omega})\mathrm{I}]{}^R\mathrm{P}^+$ and the minimum error can be found by an unconstrained optimization of equation (B.9) with respect to the projection ${}^R\hat{\mathbf{x}}$.

B.2 Optimal Reconstruction for Rectilinear Stereo and Pin-Hole Cameras

Equation (B.8) will now be evaluated for the case of pin-hole cameras in a rectilinear stereo configuration. The camera centres $\mathbf{C}_{L,R}$ and projection matrices ${}^{L,R}\mathrm{P}$ for rectilinear pin-hole cameras are given by (see Section 2.2.3)

$$\mathbf{C}_{L,R} = (\mp b,\ 0,0,\ 1)^\top \tag{B.10}$$

$${}^{L,R}\mathrm{P} = \begin{pmatrix} f & 0 & 0 & \pm fb \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{B.11}$$

where the top sign is taken for L and the bottom sign for R. Substituting ${}^{R}\mathrm{P}$ into equation (2.35), the pseudo-inverse for the right projection matrix is

$$\begin{aligned} {}^{R}\mathrm{P}^{+} &= {}^{R}\mathrm{P}^{\top}({}^{R}\mathrm{P}\,{}^{R}\mathrm{P}^{\top})^{-1} \\ &= \begin{pmatrix} 1/[f(1+b^2)] & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \\ -b/[f(1+b^2)] & 0 & 0 \end{pmatrix} \end{aligned} \tag{B.12}$$

Next, the term $[\mathbf{C}_R\boldsymbol{\Omega}^{\top} - (\mathbf{C}_R^{\top}\boldsymbol{\Omega})\mathrm{I}]$ in equation (B.8) evaluates to:

$$\mathbf{C}_R\boldsymbol{\Omega}^{\top} - (\mathbf{C}_R^{\top}\boldsymbol{\Omega})\mathrm{I} = \begin{pmatrix} -D & bB & bC & bD \\ 0 & -(Ab+D) & 0 & 0 \\ 0 & 0 & -(Ab+D) & 0 \\ A & B & C & -Ab \end{pmatrix} \tag{B.13}$$

Finally, multiplying the matrices in equations (B.11), (B.13) and (B.12), equation (B.8) is evaluated as:

$${}^{L}\hat{\mathbf{x}} = \begin{pmatrix} Ab-D & 2Bb & 2Cbf \\ 0 & -(Ab+D) & 0 \\ 0 & 0 & -(Ab+D) \end{pmatrix} {}^{R}\hat{\mathbf{x}} \tag{B.14}$$

which is the result quoted in equation (3.12) in Section 3.2.4. Evaluating equation (B.14) with ${}^{L}\hat{\mathbf{x}} = ({}^{L}\hat{x}, {}^{L}\hat{y}, {}^{L}\hat{w})^{\top}$ and ${}^{R}\hat{\mathbf{x}} = ({}^{R}\hat{x}, {}^{R}\hat{y}, 1)^{\top}$ gives

$$\begin{aligned} \begin{pmatrix} {}^{L}\hat{x} \\ {}^{L}\hat{y} \\ {}^{L}\hat{w} \end{pmatrix} &= \begin{pmatrix} Ab-D & 2Bb & 2Cbf \\ 0 & -(Ab+D) & 0 \\ 0 & 0 & -(Ab+D) \end{pmatrix} \begin{pmatrix} {}^{R}\hat{x} \\ {}^{R}\hat{y} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} (Ab-D){}^{R}\hat{x} + 2Bb{}^{R}\hat{y} + 2Cbf \\ -(Ab+D){}^{R}\hat{y} \\ -(Ab+D) \end{pmatrix} \end{aligned} \tag{B.15}$$

Expressed in inhomogeneous coordinates, the relationship between ${}^{L}\hat{\mathbf{x}}$ and ${}^{R}\hat{\mathbf{x}}$ is

$${}^{L}\hat{x} = -\frac{(Ab-D){}^{R}\hat{x} + 2Bb{}^{R}\hat{y} + 2Cbf}{Ab+D} \tag{B.16}$$

$${}^{L}\hat{y} = {}^{R}\hat{y} \tag{B.17}$$

which is the result quoted in equations (3.13)-(3.14) in Section 3.2.4.

Now, with ${}^{L,R}\hat{y} = y$ where $y = {}^{L}y = {}^{R}y$ is the height of the scan-line on which ${}^{L}\mathbf{x}$ and ${}^{R}\mathbf{x}$ are measured (see the discussion in Section 3.2.4), the error function in equation (B.9) can be evaluated from equations (B.16)-(B.17) as:

$$\begin{aligned} E &= d^2({}^{L}\mathbf{x}, {}^{L}\hat{\mathbf{x}}) + d^2({}^{R}\mathbf{x}, {}^{R}\hat{\mathbf{x}}) \\ &= ({}^{L}x - {}^{L}\hat{x})^2 + ({}^{L}y - {}^{L}\hat{y})^2 + ({}^{R}x - {}^{R}\hat{x})^2 + ({}^{R}y - {}^{R}\hat{y})^2 \\ &= \left({}^{L}x + \frac{Ab-D}{Ab+D}{}^{R}\hat{x} + \frac{2Bb}{Ab+D}y + \frac{2Cb}{Ab+D}f\right)^2 + \left({}^{R}x - {}^{R}\hat{x}\right)^2 \\ &= ({}^{L}x + \alpha{}^{R}\hat{x} + \beta y + \gamma f)^2 + ({}^{R}x - {}^{R}\hat{x})^2 \end{aligned} \tag{B.18}$$

where the last line makes the change of variables

$$\alpha = (Ab - D)/(Ab + D) \tag{B.19}$$
$$\beta = 2Bb/(Ab + D) \tag{B.20}$$
$$\gamma = 2Cb/(Ab + D) \tag{B.21}$$

Equations (B.18)-(B.21) are the result quoted in equations (3.15)-(3.18) in Section 3.2.4.

Equation (B.18) expresses the image plane error for points on the light plane as a function of a *single* variable, ${}^R\hat{x}$. Optimization now proceeds using standard techniques, setting $\frac{dE}{d^R\hat{x}}$ to zero:

$$\begin{aligned}\frac{dE}{d^R\hat{x}} &= 2\alpha({}^Lx + \alpha{}^R\hat{x} + \beta y + \gamma f) - 2({}^Rx - {}^R\hat{x}) \\ &= 2{}^R\hat{x}(\alpha^2 + 1) + 2[\alpha({}^Lx + \beta y + \gamma f) - {}^Rx] \\ &= 0\end{aligned}$$

Solving for ${}^R\hat{x}$ gives the optimal projection ${}^R\hat{x}^*$:

$${}^R\hat{x}^* = \frac{{}^Rx - \alpha({}^Lx + \beta y + \gamma f)}{\alpha^2 + 1} \tag{B.22}$$

Substituting equation (B.22) into (B.18) gives the minimum error E^* for the optimal reconstruction:

$$\begin{aligned}E^* &= \left(\frac{\alpha{}^Rx + ({}^Lx + \beta y + \gamma f)}{\alpha^2 + 1}\right)^2 + \left(\frac{\alpha^2{}^R\hat{x} + \alpha({}^Lx + \beta y + \gamma f)}{\alpha^2 + 1}\right)^2 \\ &= \frac{1}{(\alpha^2+1)^2}(\alpha{}^Rx + {}^Lx + \beta y + \gamma f)^2 + \frac{\alpha^2}{(\alpha^2+1)^2}(\alpha{}^R\hat{x} + {}^Lx + \beta y + \gamma f)^2 \\ &= \frac{({}^Lx + \alpha{}^Rx + \beta y + \gamma f)^2}{\alpha^2 + 1}\end{aligned} \tag{B.23}$$

Equations (B.22) and (B.23) are the results quoted in equations (3.25) and (3.26) in Section 3.2.4. Finally, the optimal 3D reconstruction $\widehat{\mathbf{X}}^*$ is recovered by substituting ${}^R\hat{\mathbf{x}}^*$ into equation (B.6). The Plücker matrix L_R for the back-projected ray through ${}^R\hat{\mathbf{x}}^* = ({}^R\hat{x}^*, y)^\top$ is evaluated by substituting equations (B.10)-(B.12) into (B.5):

$$\begin{aligned}\mathrm{L}_R &= \mathbf{C}_R({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}}^*)^\top - ({}^R\mathrm{P}^+{}^R\hat{\mathbf{x}}^*)\mathbf{C}_R^\top \\ &= \begin{pmatrix} 0 & by/f & b & -{}^R\hat{x}^*/f \\ -by/f & 0 & 0 & -y/f \\ -b & 0 & 0 & -1 \\ {}^R\hat{x}^*/f & y/f & 1 & 0 \end{pmatrix}\end{aligned} \tag{B.24}$$

Multiplying L_R by the plane parameters $\mathbf{\Omega} = (A, B, C, D)^\top$ and substituting the result in equation (B.22) gives the optimal reconstruction $\widehat{\mathbf{X}}^*$ in homogeneous coordinates:

$$\begin{aligned}\widehat{\mathbf{X}}^* &= \mathrm{L}_R\boldsymbol{\Omega} \\ &= \begin{pmatrix} Bby/f + Cb - D^R\hat{x}^*/f \\ -(Ab+D)y/f \\ -(Ab+D) \\ A^R\hat{x}^*/f + By/f + C \end{pmatrix}\end{aligned} \tag{B.25}$$

The x-coordinate of $\widehat{\mathbf{X}}^*$ in inhomogeneous notation is calculated by dividing the first row of equation (B.25) by the fourth row:

$$\widehat{X}^* = \frac{Bby + Cbf - D^R\hat{x}^*}{A^R\hat{x}^* + By + Cf} \tag{B.26}$$

Multiplying the top and bottom row by $2b/(Ab+D)$, and then making the change of variables in equations (B.19)-(B.21), noting that $(\alpha+1) = 2Ab/(Ab+D)$ and $(\alpha-1) = -2D/(Ab+D)$, gives $\widehat{X}^*$ in terms of α, β and γ:

$$\widehat{X}^* = \frac{b[\beta y + \gamma f + (\alpha-1)^R\hat{x}^*]}{(\alpha+1)^R\hat{x}^* + \beta y + \gamma f} \tag{B.27}$$

The expressions for $\widehat{Y}^*$ and $\widehat{Z}^*$ follow similarly

$$\widehat{Y}^* = \frac{-2by}{(\alpha+1)^R\hat{x}^* + \beta y + \gamma f} \tag{B.28}$$

$$\widehat{Z}^* = \frac{-2bf}{(\alpha+1)^R\hat{x}^* + \beta y + \gamma f} \tag{B.29}$$

Equation (B.29) is used to derive the limits of valid image plane measurements in Section 3.2.6. Finally, substituting ${}^R\hat{x}^*$ from equation (B.22) into equations (B.27)-(B.29), the inhomogeneous coordinates of the optimal reconstruction $\widehat{\mathbf{X}}^*$ can be written as a function of the image plane measurements ${}^L\mathbf{x} = ({}^Lx, y)^\top$ and ${}^R\mathbf{x} = ({}^Rx, y)^\top$, and system parameters α, β, γ, b and f:

$$\widehat{X}^* = \frac{[(\alpha-1)(\alpha{}^Lx - {}^Rx) - (\alpha+1)(\beta y + \gamma f)]b}{(\alpha+1)(\alpha{}^Lx - {}^Rx) + (\alpha-1)(\beta y + \gamma f)} \tag{B.30}$$

$$\widehat{Y}^* = \frac{2by(\alpha^2+1)}{(\alpha+1)(\alpha{}^Lx - {}^Rx) + (\alpha-1)(\beta y + \gamma f)} \tag{B.31}$$

$$\widehat{Z}^* = \frac{2bf(\alpha^2+1)}{(\alpha+1)(\alpha{}^Lx - {}^Rx) + (\alpha-1)(\beta y + \gamma f)} \tag{B.32}$$

which is the result quoted in equations (3.28)-(3.30) in Section 3.2.4.

C

Iterated Extended Kalman Filter

This Appendix presents an overview of the *Iterated Extended Kalman Filter* (IEKF) equations, which are used in Chapters 5 and 6 for model-based object tracking. A detailed treatment of Kalman filter theory and the IEKF can be found in [8] and [74]. The purpose of the Kalman filter is to estimate the state and error covariance matrix of a linear dynamic system from measurements with additive noise, and the IEKF provides a near-optimal solution when the measurement model is non-linear. The Kalman filter assumes that the measurement and system noise are Gaussian distributed, zero mean and white with known covariance. When these assumptions are not satisfied, such as for biased measurements, the filter equations under-estimate the true state covariance.

Let $\mathbf{x}(k)$ represent the state vector (the variables we wish to estimate), $\mathrm{P}(k)$ represent the covariance matrix describing the uncertainty in the state, and $\mathbf{y}(k)$ represent the measurements on a linear dynamic system at sample time k. Assuming the system has no inputs, the evolution of the state can be described by a discrete time *state transition equation* (also known as the *dynamic model*):

$$\mathbf{x}(k+1) = \mathrm{F}\mathbf{x}(k) + \mathbf{v}(k) \tag{C.1}$$

where F is the state transition matrix and $\mathbf{v}(k)$ is the process noise, which absorbs unmodelled dynamics. The measurements $\mathbf{y}(k)$ are related to the state by a *measurement model* of the form:

$$\mathbf{y}(k+1) = \mathrm{H}(k+1)\mathbf{x}(k+1) + \mathbf{w}(k+1) \tag{C.2}$$

where H represents the measurement function, and $\mathbf{w}(k)$ is the measurements noise. For non-linear sensors, such as the pin-hole camera used throughout this book, the measurement model takes the form

$$\mathbf{y}(k+1) = \mathbf{h}(k+1, \mathbf{x}(k+1)) + \mathbf{w}(k+1) \tag{C.3}$$

where $\mathbf{h}(k, \mathbf{x})$ is the vector-valued non-linear measurement function describing the sensor model. As already noted, the IEKF assumes $\mathbf{v}(k)$ and $\mathbf{w}(k)$ are white, zero

mean, Gaussian distributed noise sources. The covariance matrices for the process noise $\mathrm{Q}(k)$ and measurement noise $\mathrm{R}(k)$ are assumed to be known, and are given by

$$\mathrm{Q}(k) = E[\mathbf{v}(k)\mathbf{v}^\top(k)] \tag{C.4}$$
$$\mathrm{R}(k) = E[\mathbf{w}(k)\mathbf{w}^\top(k)] \tag{C.5}$$

where $E[\cdot]$ denotes expectation. When the state and measurement noise components are assumed to be independent, as is the case for the systems in this book, R(k) and Q(k) are diagonal.

Let $\hat{\mathbf{x}}(k|k)$ represent the estimated state at time k, and $\mathbf{y}(k+1)$ represent a new set of measurements observed at time $k+1$. The IEKF algorithm estimates the new state $\hat{\mathbf{x}}(k+1|k+1)$ as a weighted mean of the state predicted by the system dynamics and the new measurements, with the filter weight calculated from the state error and measurement error covariances. In the first step of the algorithm, a predicted state vector $\hat{\mathbf{x}}(k+1|k)$ and covariance matrix $\mathrm{P}(k+1|k)$ are calculated as

$$\hat{\mathbf{x}}(k+1|k) = \mathrm{F}\hat{\mathbf{x}}(k|k) \tag{C.6}$$
$$\mathrm{P}(k+1|k) = \mathrm{FP}(k|k)\mathrm{F}^\top + \mathrm{Q}(k) \tag{C.7}$$

The predicted state is then updated using the new measurements, which is an iterated process in the IEKF to solve the non-linear measurement equations. Let $\mathbf{n}_i$ represent the updated state estimate at the ith iteration, with $\mathbf{n}_0 = \hat{\mathbf{x}}(k+1|k)$. To calculate the filter weight, the measurement function $\mathbf{h}(k+1)$ is replaced by a linear approximation $\mathrm{M}_i(k+1)$, operating about the current updated state:

$$\mathrm{M}_i(k+1,\mathbf{n}_i) = \left.\frac{\partial \mathbf{h}(k+1,\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{n}_i} \tag{C.8}$$

Then, the filter weight $\mathrm{K}_i(k+1)$ at the ith iteration is given by

$$\mathrm{K}_i(k+1) = \mathrm{P}(k+1|k)\mathrm{M}_i^\top(k+1)[\mathrm{M}_i(k+1)\mathrm{P}(k+1|k)\mathrm{M}_i^\top(k+1) + \mathrm{R}(k+1)]^{-1} \tag{C.9}$$

Using the non-linear sensor model, the measurements associated with the current state estimate are predicted as $\mathbf{h}(k+1,\mathbf{n}_i)$ and compared to the actual measurements $\mathbf{y}(k+1)$ to form an observation error $\mathbf{e}(k+1)$:

$$\mathbf{e}(k+1) = \mathbf{y}(k+1) - \mathbf{h}(k+1,\mathbf{n}_i) \tag{C.10}$$

Finally, a new estimate $\mathbf{n}_{i+1}$ of the state is recovered as the weighted mean of the predicted state $\hat{\mathbf{x}}(k+1|k)$ and the observation error $\mathbf{e}(k+1)$ ([74], page 279):

$$\mathbf{n}_{i+1} = \hat{\mathbf{x}}(k+1|k) + \mathrm{K}_i(k+1)\{\mathbf{e}(k+1) - \mathrm{M}_i(k+1)[\hat{\mathbf{x}}(k+1|k) - \mathbf{n}_i]\} \tag{C.11}$$

Equations (C.8)-(C.11) are iterated until successive state estimates $\mathbf{n}_i$ converge according to suitable criteria. In the current implementation, the system is deemed to have converged when the maximum absolute error between successive estimates of all state variables is below a threshold c_{th}:

$$\max_j |n_{j,i+1} - n_{j,i}| < c_{th} \tag{C.12}$$

where $n_{j,i}$ is the jth element of vector $\mathbf{n}_i$. The final result is taken as the updated state estimate $\hat{\mathbf{x}}(k+1|k+1)$. Finally, the state covariance is updated from the predicted covariance (equation (C.7)) by

$$\mathrm{P}(k+1|k+1) = [\mathrm{I} - \mathrm{K}(k+1)\mathrm{M}(k+1)]\mathrm{P}(k+1|k) \tag{C.13}$$

In practice, the Joseph form covariance update is used:

$$\mathrm{P}(k+1|k+1) = \mathrm{A}(k+1)\mathrm{P}(k+1|k)\mathrm{A}^\top(k+1) + \mathrm{K}(k+1)\mathrm{R}(k+1)\mathrm{K}^\top(k+1) \tag{C.14}$$

where

$$\mathrm{A}(k+1) = \mathrm{I} - \mathrm{K}(k+1)\mathrm{M}(k+1) \tag{C.15}$$

which is less sensitive to numerical round-off errors ([8], pg. 216).

D

Stereo Reconstruction Error Models

This appendix presents error models describing the uncertainty in reconstructing a single point from stereo measurement, and similarly estimating the pose of an object modelled by multiple points. The error models are used in Chapter 6 for hybrid position-based visual servoing.

D.1 Optimal Reconstruction of a Single Point

This section derives the optimal stereo reconstruction $\widehat{\mathbf{X}}$ of a point from measurement ${}^{L}\mathbf{x}$ and ${}^{R}\mathbf{x}$ on the left and right image planes. Assuming pin-hole cameras in a rectilinear configuration, the projections of $\widehat{\mathbf{X}}$ on the stereo image planes are given by

$$ {}^{L,R}\hat{\mathbf{x}} = {}^{L,R}\mathrm{P}\widehat{\mathbf{X}} \tag{D.1} $$

where ${}^{L,R}\mathrm{P}$ are the projection matrices of the left and right cameras, given by equation (2.29). The above transformation can be expressed in inhomogeneous coordinates as

$$ \begin{pmatrix} {}^{L,R}\hat{x} \\ {}^{L,R}\hat{y} \end{pmatrix} = \frac{f}{\widehat{Z}} \begin{pmatrix} \widehat{X} \pm b \\ \widehat{Y} \end{pmatrix} \tag{D.2} $$

taking the positive sign for L and the negative for R. Now, the optimal estimate is the point $\widehat{\mathbf{X}}$ that minimizes the image plane error between the measurements ${}^{L,R}\mathbf{x}$ and the projections ${}^{L,R}\hat{\mathbf{x}}$, calculated (in inhomogeneous coordinates) from equation (6.12) as:

$$ D^2 = |{}^{L}\hat{\mathbf{x}} - {}^{L}\mathbf{x}|^2 + |{}^{R}\hat{\mathbf{x}} - {}^{R}\mathbf{x}|^2 \tag{D.3} $$

Substituting equation (D.2) into equation (D.3):

$$ D^2 = \left(\frac{f(\widehat{X}+b)}{\widehat{Z}} - {}^{L}x \right)^2 + \left(\frac{f\widehat{Y}}{\widehat{Z}} - {}^{L}y \right)^2 + \left(\frac{f(\widehat{X}-b)}{\widehat{Z}} - {}^{R}x \right)^2 + \left(\frac{f\widehat{Y}}{\widehat{Z}} - {}^{R}y \right)^2 \tag{D.4} $$

Minimization of D^2 with respect to $\widehat{\mathbf{X}}$ proceeds in the usual manner by setting partial derivatives of D^2 to zero and solving for the optimal reconstruction. The relevant partial derivatives are:

$$\frac{\partial D^2}{\partial \widehat{X}} = \frac{2f}{\widehat{Z}}\left(\frac{f(\widehat{X}+b)}{\widehat{Z}} - {}^{L}x\right) + \frac{2f}{\widehat{Z}}\left(\frac{f(\widehat{X}-b)}{\widehat{Z}} - {}^{R}x\right) \tag{D.5}$$

(D.6)

$$\frac{\partial D^2}{\partial \widehat{Y}} = \frac{2f}{\widehat{Z}}\left(\frac{f\widehat{Y}}{\widehat{Z}} - {}^{L}y\right) + \frac{2f}{\widehat{Z}}\left(\frac{f\widehat{Y}}{\widehat{Z}} - {}^{R}y\right) \tag{D.7}$$

(D.8)

$$\begin{aligned}\frac{\partial D^2}{\partial \widehat{Z}} = &-\frac{2f(\widehat{X}+b)}{\widehat{Z}^2}\left(\frac{f(\widehat{X}+b)}{\widehat{Z}} - {}^{L}x\right) - \frac{2f\widehat{Y}}{\widehat{Z}^2}\left(\frac{f\widehat{Y}}{\widehat{Z}} - {}^{L}y\right) \\ &-\frac{2f(\widehat{X}-b)}{\widehat{Z}^2}\left(\frac{f(\widehat{X}-b)}{\widehat{Z}} - {}^{R}x\right) - \frac{2f\widehat{Y}}{\widehat{Z}^2}\left(\frac{f\widehat{Y}}{\widehat{Z}} - {}^{R}y\right)\end{aligned} \tag{D.9}$$

Setting the right hand side of equations (D.5) and (D.6) to zero and solving for $\widehat{X}$ and $\widehat{Y}$ gives:

$$\widehat{X} = \frac{\widehat{Z}}{2f}({}^{L}x + {}^{R}x) \tag{D.10}$$

$$\widehat{Y} = \frac{\widehat{Z}}{2f}({}^{L}y + {}^{R}y) \tag{D.11}$$

Now, setting $\frac{\partial D^2}{\partial \widehat{Z}} = 0$ in equation (D.8) and collecting common factors yields:

$$\begin{aligned}0 &= (\widehat{X}+b)(f(\widehat{X}+b) - {}^{L}x\widehat{Z}) + \widehat{Y}(f\widehat{Y} - {}^{L}y\widehat{Z}) \\ &\quad + (\widehat{X}+b)(f(\widehat{X}-b) - {}^{R}x\widehat{Z}) + \widehat{Y}(f\widehat{Y} - {}^{R}y\widehat{Z}) \\ &= 2f\widehat{X}^2 + 2f\widehat{Y}^2 + 2fb^2 - \widehat{Z}\widehat{X}({}^{L}x + {}^{R}x) - \widehat{Z}\widehat{Y}({}^{L}y + {}^{R}y) - \widehat{Z}b({}^{L}x - {}^{R}x)\end{aligned} \tag{D.12}$$

Then, substituting equations (D.10)-(D.11) into equation (D.12):

$$\begin{aligned}0 &= \frac{\widehat{Z}^2}{2f}({}^{L}x + {}^{R}x)^2 + \frac{\widehat{Z}^2}{2f}({}^{L}y + {}^{R}y)^2 + 2fb^2 \\ &\quad - \frac{\widehat{Z}^2}{2f}({}^{L}x + {}^{R}x)^2 - \frac{\widehat{Z}^2}{2f}({}^{L}y + {}^{R}y)^2 - \widehat{Z}b({}^{L}x - {}^{R}x) \\ &= 2fb - \widehat{Z}({}^{L}x - {}^{R}x)\end{aligned} \tag{D.13}$$

Finally, equation (D.13) can be solved for $\widehat{Z}$ and the result substituted into equations (D.10)-(D.11). Finally, the components of the optimal reconstruction $\widehat{\mathbf{X}}$ can be expressed as a function of the image plane measurements:

$$\widehat{X} = \frac{b}{{}^{L}x - {}^{R}x}({}^{L}x + {}^{R}x) \tag{D.14}$$

$$\widehat{Y} = \frac{b}{{}^{L}x - {}^{R}x}({}^{L}y + {}^{R}y) \tag{D.15}$$

$$\widehat{Z} = \frac{2bf}{{}^{L}x - {}^{R}x} \tag{D.16}$$

The above result can be compactly represented in vector form as:

$$\widehat{\mathbf{X}} = \frac{b}{{}^{L}x - {}^{R}x}\left({}^{L}x + {}^{R}x,\ {}^{L}y + {}^{R}y,\ 2f\right)^{\top} \tag{D.17}$$

which is the result quoted in equation (6.14) in Section 6.3.

D.2 Error Model for Reconstruction of a Single Point

This section now considers the error introduced in the reconstruction $\widehat{\mathbf{X}}$ (given by equation (D.17)) due to uncertainties in the parameters of the stereo camera model. Let b^*, f^* and v^* represent the *actual* baseline, focal length and verge angle of the stereo rig, and let b, f and v represent the *calibrated* parameters of the camera model. If the camera parameters are perfectly calibrated, the measurements ${}^{L,R}\mathbf{x}$ are related to the *actual* point, $\mathbf{X} = (X,Y,Z)^{\top}$, by (see equation(D.2)):

$${}^{L,R}\mathbf{x} = \frac{f^*}{Z}(X \pm 2b^*, Y)^{\top} \tag{D.18}$$

In practice, features are observed in non-rectilinear stereo cameras, and then corrected using projective rectification (see Section 2.2.4). However, any error in the calibrated verge angle will cause projective rectification to over or under-compensate by angle $(v - v^*)$ when correcting the raw measurements for rectilinear stereo. Thus, in the presence of verge angle error, the *ideal* rectilinear measurements in equation (D.18) are offset by angular error $(v - v^*)$ (see equations (2.37)-(2.38)) to give the *actual* rectilinear measurements:

$${}^{L,R}\mathbf{x} = \frac{f^*\left((X \pm 2b^*)\cos(v - v^*) \mp Z\sin(v - v^*),\ Y\right)^{\top}}{Z\cos(v - v^*) \pm (X \pm 2b^*)\sin(v - v^*)} \tag{D.19}$$

Assuming the verge angle error is not severe, the small angle approximations $\cos(v - v^*) \approx 1$ and $\sin(v - v^*) \approx (v - v^*)$ are introduced to reduce the above to

$${}^{L,R}\mathbf{x} = \frac{f^*\left((X \pm 2b^*) \mp Z(v - v^*),\ Y\right)^{\top}}{Z \pm (X \pm 2b^*)(v - v^*)} \tag{D.20}$$

Based on these measurements and *calibrated* camera parameters f, b and v, the optimal reconstruction $\widehat{\mathbf{X}}$ is found by substituting equation (D.20) into equation (D.17). Evaluating the intermediate terms in equation (D.17):

$$
\begin{aligned}
{}^{L}x - {}^{R}x &= f^* \frac{(X+2b^*) - Z(\nu-\nu^*)}{Z+(X+2b^*)(\nu-\nu^*)} - f^* \frac{(X-2b^*)+Z(\nu-\nu^*)}{Z-(X-2b^*)(\nu-\nu^*)} \\
&= 2f^* \frac{2Zb^* - (X^2+Z^2-4(b^*)^2)(\nu-\nu^*) - 2Zb^*(\nu-\nu^*)^2}{Z^2+4Zb^*(\nu-\nu^*) - (X^2-4(b^*)^2)(\nu-\nu^*)^2} \qquad \text{(D.21)}
\end{aligned}
$$

Again assuming the verge angle error is small, the approximation $(\nu-\nu^*)^2 \ll 1$ is introduced to reduce equation (D.21) to

$$
{}^{L}x - {}^{R}x = 2f^* \frac{2Zb^* - (X^2+Z^2-4(b^*)^2)(\nu-\nu^*)}{Z^2+4Zb^*(\nu-\nu^*)} \qquad \text{(D.22)}
$$

The remaining intermediate terms follow similarly:

$$
{}^{L}x + {}^{R}x = 2f^* \frac{XZ}{Z^2+4Zb^*(\nu-\nu^*)} \qquad \text{(D.23)}
$$

$$
{}^{L}y + {}^{R}y = 2f^* \frac{Y(Z+2b^*(\nu-\nu^*))}{Z^2+4Zb^*(\nu-\nu^*)} \qquad \text{(D.24)}
$$

Finally, the reconstruction $\widehat{\mathbf{X}} = (\widehat{X},\widehat{Y},\widehat{Z})^\top$ from equation (D.17) is

$$
\widehat{X} = \frac{XZb}{2Zb^* - (X^2+Z^2-4(b^*)^2)(\nu-\nu^*)} \qquad \text{(D.25)}
$$

$$
\widehat{Y} = \frac{Yb(Z+2b^*(\nu-\nu^*))}{2Zb^* - (X^2+Z^2-4(b^*)^2)(\nu-\nu^*)} \qquad \text{(D.26)}
$$

$$
\widehat{Z} = \frac{2bf(Z^2+4Zb^*(\nu-\nu^*))}{2Zb^* - (X^2+Z^2-4(b^*)^2)(\nu-\nu^*)} \qquad \text{(D.27)}
$$

Again assuming small errors, the effect of calibration errors on $\widehat{\mathbf{X}}$ is most easily examined by taking a Taylor series expansion of equations (D.25)-(D.27) with respect to f, b and ν about the operating point $f=f^*$, $b=b^*$ and $\nu=\nu^*$. Taking the first order expansion for each component:

$$
\widehat{X}_i(f,b,\nu) = X\left[1 + \frac{b-b^*}{2b^*} + \frac{X^2+Z^2-4(b^*)^2}{2Zb^*}(\nu-\nu^*)\right] \qquad \text{(D.28)}
$$

$$
\widehat{Y}_i(f,b,\nu) = Y\left[1 + \frac{b-b^*}{2b^*} + \frac{X^2+Z^2}{2Zb^*}(\nu-\nu^*)\right] \qquad \text{(D.29)}
$$

$$
\widehat{Z}_i(f,b,\nu) = Z\left[1 + \frac{b-b^*}{2b^*} + \frac{f-f^*}{f^*} + \frac{X^2+Z^2+4(b^*)^2}{2Zb^*}(\nu-\nu^*)\right] \qquad \text{(D.30)}
$$

Rearrange the above and expressing in vector notation, the relationship between a real point $\mathbf{X} = (X,Y,Z)^\top$ and its reconstruction $\widehat{\mathbf{X}}$ in the presence of camera calibration errors can be written as a linear function of f, b and ν:

$$
\widehat{\mathbf{X}} = \left(1 + \frac{b-b^*}{2b^*} + \frac{X^2+Z^2}{2Zb^*}(\nu-\nu^*)\right)\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \frac{f-f^*}{f^*}\begin{pmatrix} 0 \\ 0 \\ Z \end{pmatrix} + (\nu-\nu^*)\begin{pmatrix} 2Xb^*/Z \\ 0 \\ 2b^* \end{pmatrix} \qquad \text{(D.31)}
$$

which is the final error model quoted in equation (6.17) in Section 6.3.

D.3 Error Model for Pose Estimation

The previous section considered the effect of calibration errors on the reconstruction of a single point. This section now examines the effect of errors on estimating the pose of an object modelled by multiple points. For simplicity and without loss of generality, the object is assumed to undergo pure translation. Let $\mathbf{G}_i$, $i = 1,\ldots,N$ represent the N points of the model, with $\sum \mathbf{G}_i = \mathbf{0}$ as discussed in Section 6.3, and let $\mathbf{T}_E = (X_E, Y_E, Z_E)^\top$ represent the *actual* position of the object. Now, let $\mathbf{G}_i^*$ represent the *visually reconstructed* points in the model. As discussed in Section 6.3, the effect of calibration errors is to scale each visually reconstructed point by K_1. Thus, the measured 3D points (after biasing by calibration errors) are effectively given by:

$$\mathbf{G}_i^* = K_1(\mathbf{G}_i + \mathbf{T}_E) \tag{D.32}$$

and the *actual* image plane measurements are given by the projection in equation (D.2):

$${}^{L,R}\mathbf{g}_i = \frac{f}{K_1(Z_i + Z_E)}(K_1(X_i + X_E) \pm 2b,\ K_1(Y_i + Y_E))^\top \tag{D.33}$$

Now, let $\widehat{\mathbf{T}}_E = (\widehat{X}_E, \widehat{Y}_E, \widehat{Z}_E)^\top$ represent the *estimated* pose of the object. The points in the model are estimated as $\widehat{\mathbf{G}}_i = \mathbf{G}_i + \widehat{\mathbf{T}}_E$ (ie. without the unknown bias K_1) and the corresponding *predicted* measurements for the given pose are (from equation (D.2)):

$${}^{L,R}\hat{\mathbf{g}}_i = \frac{f}{Z_i + \widehat{Z}_E}(X_i + \widehat{X}_E \pm 2b,\ Y_i + \widehat{Y}_E)^\top \tag{D.34}$$

An optimal estimate of the translation $\widehat{\mathbf{T}}_E$ is obtained by minimizing the reprojection error $D^2(\widehat{\mathbf{T}}_E)$ in equation (6.12) between the actual and predicted measurements. Substituting equations (D.33)-(D.34) into (6.12) gives the reprojection error as:

$$\begin{aligned} D^2(\widehat{\mathbf{T}}_E) &= \sum_i \left[({}^L\hat{x}_i - {}^Lx_i)^2 + ({}^L\hat{y}_i - {}^Ly_i)^2 + ({}^R\hat{x}_i - {}^Rx_i)^2 + ({}^R\hat{y}_i - {}^Ry_i)^2 \right] \\ &= f^2 \sum_i \left[\left(\frac{X_i + \widehat{X}_E + 2b}{Z_i + \widehat{Z}_E} - \frac{K_1(X_i + X_E) + 2b}{K_1(Z_i + Z_E)} \right)^2 \right. \\ &\qquad + \left(\frac{Y_i + \widehat{Y}_E}{Z_i + \widehat{Z}_E} - \frac{K_1(Y_i + Y_E)}{K_1(Z_i + Z_E)} \right)^2 \\ &\qquad + \left(\frac{X_i + \widehat{X}_E - 2b}{Z_i + \widehat{Z}_E} - \frac{K_1(X_i + X_E) - 2b}{K_1(Z_i + Z_E)} \right)^2 \\ &\qquad \left. + \left(\frac{Y_i + \widehat{Y}_E}{Z_i + \widehat{Z}_E} - \frac{K_1(Y_i + Y_E)}{K_1(Z_i + Z_E)} \right)^2 \right] \end{aligned}$$

The optimal estimated pose can be found analytically by setting the partial derivatives of $D^2(\widehat{\mathbf{T}}_E)$ to zero and solving for $\widehat{\mathbf{T}}_E$. Taking the partial derivative with respect to $\widehat{X}_E$ yields:

$$\frac{\partial D^2}{\partial \widehat{X}_E} = f^2 \sum_i \left[2\left(\frac{X_i + \widehat{X}_E + 2b}{Z_i + \widehat{Z}_E} - \frac{K_1(X_i + X_E) + 2b}{K_1(Z_i + Z_E)} \right) \frac{1}{Z_i + \widehat{Z}_E} + 2\left(\frac{X_i + \widehat{X}_E - 2b}{Z_i + \widehat{Z}_E} - \frac{K_1(X_i + X_E) - 2b}{K_1(Z_i + Z_E)} \right) \frac{1}{Z_i + \widehat{Z}_E} \right] \tag{D.35}$$

Expressing equation (D.35) over a common denominator and equating the numerator to zero gives

$$\begin{aligned} 0 &= \sum_i \Big[K_1(Z_i + Z_E)(X_i + \widehat{X}_E + 2b) - (Z_i + \widehat{Z}_E)[K_1(X_i + X_E) + 2b] \\ &\qquad + K_1(Z_i + Z_E)(X_i + \widehat{X}_E - 2b) - (Z_i + \widehat{Z}_E)[K_1(X_i + X_E) - 2b] \Big] \\ &= K_1 \sum (X_i Z_i) + K_1(\widehat{X}_E + 2b) \sum Z_i + K_1 Z_E \sum X_i + N K_1 Z_E (\widehat{X}_E + 2b) \\ &\qquad - K_1 \sum (X_i Z_i) - K_1(X_E + 2b) \sum Z_i - K_1 \widehat{Z}_E \sum X_i - N K_1 \widehat{Z}_E (X_E + 2b) \\ &\qquad + K_1 \sum (X_i Z_i) + K_1(\widehat{X}_E - 2b) \sum Z_i + K_1 Z_E \sum X_i + N K_1 Z_E (\widehat{X}_E - 2b) \\ &\qquad - K_1 \sum (X_i Z_i) - K_1(X_E - 2b) \sum Z_i - K_1 \widehat{Z}_E \sum X_i - N K_1 \widehat{Z}_E (X_E - 2b) \\ &= Z_E \widehat{X}_E - \widehat{Z}_E X_E \end{aligned}$$

where the last line is obtained after substituting $\sum \mathbf{G}_i = \mathbf{0}$. Solving for $\widehat{X}_E$:

$$\widehat{X}_E = \frac{X_E \widehat{Z}_E}{Z_E} \tag{D.36}$$

Taking the partial derivatives of $D^2(\widehat{\mathbf{T}}_E)$ with respect to $\widehat{Y}_E$, and following similar algebraic manipulation leads to

$$\widehat{Y}_E = \frac{Y_E \widehat{Z}_E}{Z_E} \tag{D.37}$$

and finally for $\widehat{Z}_E$:

$$\widehat{Z}_E = \frac{K_1[(\widehat{X}_E - X_E)\sum X_i Z_i + (\widehat{Y}_E - Y_E)\sum Y_i Z_i + Z_E(\sum X_i^2 + \sum Y_i^2 + N\widehat{X}_E^2 + N\widehat{Y}_E^2 + 4Nb^2)]}{K_1(\sum X_i^2 + \sum Y_i^2 + N X_E \widehat{X}_E + N Y_E \widehat{Y}_E) + 4Nb^2} \tag{D.38}$$

Substituting equations (D.36) and (D.37) into equation (D.38) allows $\widehat{X}_E$ and $\widehat{Y}_E$ to be eliminated from the expression for $\widehat{Z}_E$ (left as an exercise for the reader), and the result is substituted back into equations (D.36) and (D.37) to express the optimal estimated pose as a function of the real pose, the model points and the scale error K_1. The final solution can be expressed in vector notation as

$$\widehat{\mathbf{T}}_E = \frac{X_E \sum X_i Z_i + Y_E \sum Y_i Z_i - Z_E(\sum X_i^2 + \sum Y_i^2 + 4Nb^2)}{K_1(X_E \sum X_i Z_i + Y_E \sum Y_i Z_i) - Z_E[K_1(\sum X_i^2 + \sum Y_i^2) + 4Nb^2]} K_1 \mathbf{T}_E \tag{D.39}$$

which is the result quoted in equation (6.26) in Section 6.3.

E

Calibration of System Latencies

A service robot is a complex system requiring the coordination of multiple sensors and actuators. Complications arise when independent sensors sample measurements at different times, and actuators have an unknown delay between receiving and executing commands. Without compensation, these various latencies can have a negative impact on the stability and reliability of dynamic processes such as visual object tracking and feedback control. One way to avoid these effects is to ensure that the dynamics of the robot and environment remain within the region of stability. For the experiments in Chapters 6 and 7, for example, stability was ensured by setting a deliberately low gain on the visual servo controller. Obviously, the drawback of this approach is an increase in the total execution time. Other methods have been proposed to deal with latencies directly (for example, see [27]) for cases in which this side effect cannot be tolerated.

Clearly, methods that compensate for latency must possess good estimates of the relevant delays. This appendix presents a simple calibration procedure that can be used to estimate the latency between sensors and actuators. We illustrate this method by applying it to the light stripe sensor (see Chapter 3); specifically, estimating the delay between exposing the CCDs to a new image of the light stripe and sampling the angle encoder for the position of the light stripe. However, the same method also applies to the delay between sending a motion command to the Puma and observation of the subsequent motion on the CCD, or the delay between exposing the left and right CCDs (negligible on our experimental platform since the cameras are synchronized in hardware).

The calibration technique is based on minimizing the hysteresis induced by the delay for some cyclic motion. For the light stripe scanner, the laser is scanned backward and forward across a planar surface and the encoder value $e(t)$ and laser position on the image plane $x(t)$ (averaged over a few scan-lines) are recorded along with a timestamp t. Figure E.1 shows the image plane position plotted against encoder value, with a visible hysteresis loop resulting from the lag between these two variables. Now, the loop can be closed by phase shifting the encoder values by the acquisition delay δt. Thus, the latency δt may be calculated as the phase shift that minimizes the residual error from a linear regression applied to $x(t)$ and $e(t+\delta t)$. It

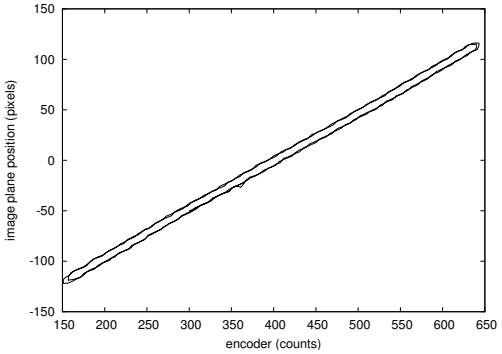


Fig. E.1. Hysteresis due to acquisition delay between corresponding images and encoder measurements of the light stripe.

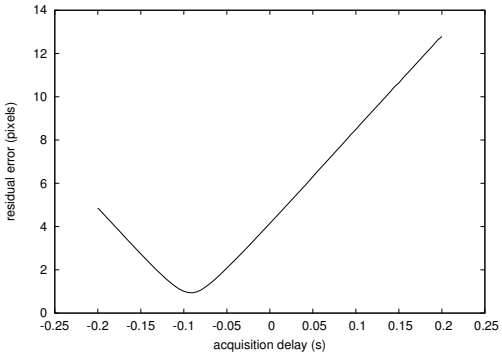


Fig. E.2. Residual error for estimation of time delay from linear regression.

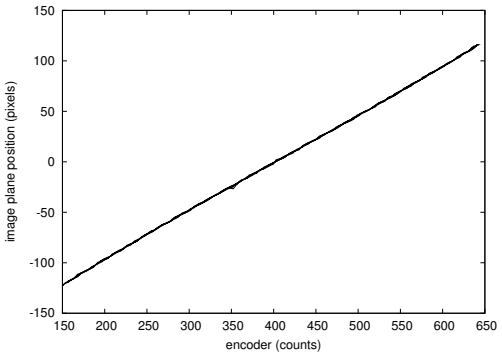


Fig. E.3. Relationship between captured images and encoder measurements of the light stripe after compensating for latency.

should be noted that the image plane position is not exactly a linear function of the encoder value, but the approximation is valid over a small range of motion.

The phase shift is applied using linear interpolation to calculate the encoder value $e(t+\delta t)$ corresponding to each image sample $x(t)$. Figure E.2 shows the residual error from a linear regression applied to $x(t)$ and $e(t+\delta t)$ for phase shifts in 5 ms increments between ± 200 ms. The plot exhibits a minima at about -90 ms (to the nearest 5 ms), indicating that the shaft encoder is sampled about 90 ms *after* the associated image was exposed. Finally, the relationship between encoder and CCD measurements after compensating for the delay is shown in Figure E.3, and the hysteresis has disappeared.

The calibration procedure was also applied to the Puma arm by moving the end-effector in a periodic motion and observing the image plane position of an active marker. The analysis revealed that the effect of a particular motion command was observed by the cameras about 170 ms after the command was issued to the Puma controller.

F

Task Planning

This appendix details the task planning required guide the motion of the end-effector for the experiments in Chapter 7. Task planning is divided into two phases: *grasp planning* and *trajectory planning*. Grasp planning is the process of determining the best contact points between the fingers and the target object to stably grasp the object. We simplify this process by formulating a specific algorithm for each known object, rather than using a general solution as in [87]. Based on established principles [37, 140], grasp planners for rectangular prisms (boxes) and upright cylinders (cups) are developed in Sections F.1 and F.2. Following grasp planning, a trajectory planner generates a collision-free path to guide the gripper from the current pose to the planned grasp. Details of the trajectory planner are presented in Section F.3. Finally, Section F.4 describes a simple inverse kinematic model for the Puma to calculate the angle of the wrist joint, which is required for grasp planning, and determine whether the a planned grasp is reachable.

F.1 Grasping a Box

While a box may be grasped in numerous ways, the algorithm described here simplifies the problem by considering only the precision grasp shown in the Figure F.1. This grasp is widely applicable and minimizes the possibility of collisions with other objects. Optimal grasp planning requires knowledge of surface properties such as friction the distribution of mass, both of which are typically unknown in service applications. However, general principles dictate that maximum stability is achieved when the force applied by the fingers is normal to the surface and the object is grasped near or above the centre of mass to minimize load torque when lifted. These rules are easily applied to a box, assuming a uniformly distributed mass.

Optimal grasp planning algorithms typically determine a set of candidate contact points on the surface of the object using heuristics and then select the best set based on a suitable cost function [38,140]. The fast but sub-optimal planner described here considers only the two candidate contact pairs on orthogonal faces shown in Figure F.1: $(\mathbf{C}_1, \mathbf{C}_2)$ and $(\mathbf{C}_3, \mathbf{C}_4)$, where $\mathbf{C}_1$ and $\mathbf{C}_3$ are on the faces nearest the robot. The

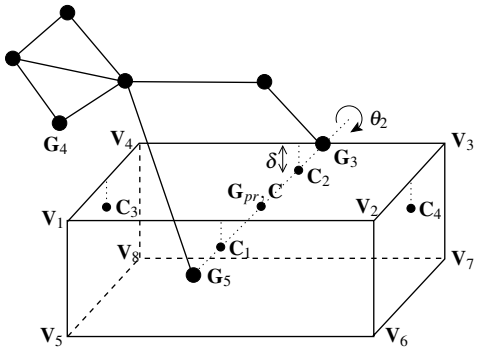


Fig. F.1. Grasp planning for a box.

contacts are placed a small distance δ below the top surface and above the midpoint of opposite faces to minimize torque and slippage. The contacts are calculated from the vertices $\mathbf{V}_1$ to $\mathbf{V}_8$ of the polygonal box model. For example, $\mathbf{C}_1$ is calculated as

$$\mathbf{C}_1 = \frac{1}{2}(\mathbf{V}_1 + \mathbf{V}_2) + \delta \cdot \frac{\mathbf{V}_5 - \mathbf{V}_1}{|\mathbf{V}_5 - \mathbf{V}_1|} \tag{F.1}$$

For clarity, the remaining discussion considers only $(\mathbf{C}_1, \mathbf{C}_2)$, and the equivalent calculations for $(\mathbf{C}_3, \mathbf{C}_4)$ follow similarly.

The first step in the planning process is to determine whether the contacts are reachable. Let $\mathbf{G}_5$ and $\mathbf{G}_3$ describe the location of the thumb and index fingertips when the gripper is fully opened. Candidates $(\mathbf{C}_1, \mathbf{C}_2)$ are considered to fit within the grasp when $|\mathbf{C}_1 - \mathbf{C}_2| < |\mathbf{G}_5 - \mathbf{G}_3|$. If the contact points violate this condition, the grasp candidate is considered no further. If successful, the algorithm calculates the transformation that aligns the fingertips with the contacts, by placing $\mathbf{G}_{pr}$ (defined in Figure 7.3) at the midpoint $\mathbf{C} = \frac{1}{2}(\mathbf{C}_1 + \mathbf{C}_2)$. The desired rotation aligns the pair of lines joining $(\mathbf{G}_5, \mathbf{G}_3)$ and $(\mathbf{C}_1, \mathbf{C}_2)$, and is calculated as an angle θ_1 about axis $\mathbf{A}_1$, given by:

$$\mathbf{A}_1 = \frac{({}^E\mathbf{G}_5 - {}^E\mathbf{G}_3) \times (\mathbf{C}_1 - \mathbf{C}_2)}{|{}^E\mathbf{G}_5 - {}^E\mathbf{G}_3||\mathbf{C}_1 - \mathbf{C}_2|} \tag{F.2}$$

$$\theta_1 = \cos^{-1}\left(\frac{({}^E\mathbf{G}_5 - {}^E\mathbf{G}_3)^\top (\mathbf{C}_1 - \mathbf{C}_2)}{|{}^E\mathbf{G}_5 - {}^E\mathbf{G}_3||\mathbf{C}_1 - \mathbf{C}_2|}\right) \tag{F.3}$$

The translation of the end-effector that aligns $\mathbf{G}_{pr}$ with $\mathbf{C}$ after rotation is

$$\mathbf{T}_1 = \mathbf{C} - \mathrm{R}_1(\theta_1, \mathbf{A}_1)\,{}^E\mathbf{G}_{pr} \tag{F.4}$$

where $\mathrm{R}_1(\theta_1, \mathbf{A}_1)$ is the rotation matrix corresponding to the axis/angle given in equations (F.2)-(F.3).

While the above transformation aligns the fingertips and contacts, the orientation of the hand about the line through $(\mathbf{C}_1, \mathbf{C}_2)$, represented as angle θ_2 in Figure F.1,

remains unconstrained. To avoid collisions, the grasp planner chooses θ_2 such that the palm is above the top surface of the box. The lowest point on the gripper is $\mathbf{G}_4 = (X_4, Y_4, Z_4)^\top$, and the grasp is planned to satisfy $Y_4 > (Y_C + \delta + Y_{th})$, where Y_C is the height of $\mathbf{C}$ and error threshold Y_{th} ensures that $\mathbf{G}_4$ is well above the box. The location of $\mathbf{G}_4$ as a function of the orientation of the hand is:

$$^W\mathbf{G}_4(\theta_2) = \mathrm{R}_2(\theta_2, \mathbf{A}_2)\mathrm{R}_1(\theta_1, \mathbf{A}_1)(^E\mathbf{G}_4 - {}^E\mathbf{G}_{pr}) + \mathbf{C} \tag{F.5}$$

where the rotation axis is $\mathbf{A}_2 = (\mathbf{C}_2 - \mathbf{C}_1)/|\mathbf{C}_2 - \mathbf{C}_1|$. In the current implementation, θ_2 is calculated numerically as

$$\theta_2 = \mathrm{argmin}_\theta |Y_C + \delta + Y_{th} - Y_4(\theta)| \tag{F.6}$$

where $Y_4(\theta)$ is calculated from equation (F.5) for all angles in steps of one degree.

Finally, the two transformations are combined to give the planned orientation $^W\mathrm{R}_E$ and translation $^W\mathbf{T}_E$ of end-effector in the planned grasp:

$$^W\mathrm{R}_E = \mathrm{R}_2\mathrm{R}_1 \tag{F.7}$$

$$^W\mathbf{T}_E = \mathbf{C} - \mathrm{R}_2\mathrm{R}_1\,{}^E\mathbf{G}_{pr} \tag{F.8}$$

The condition developed in Section F.4 is used to determine whether the planned grasp is within reach of the robot, and the grasp is discarded if this condition is violated.

As noted earlier, a candidate grasp is calculated for both pairs of contact points, $(\mathbf{C}_1, \mathbf{C}_2)$ and $(\mathbf{C}_3, \mathbf{C}_4)$. If more than one grasp is realizable, the algorithm arbitrarily chooses the grasp that minimizes the wrist angle (Section F.4 describes the inverse kinematic model for calculating wrist angle of the Puma robot). Finally, the planned grasp is transformed to the frame of the end-effector as

$$^E\mathrm{H}_G = {}^W\mathrm{H}_E^{-1}\,{}^W\mathrm{H}_O \tag{F.9}$$

where $^W\mathrm{H}_O$ is the pose of the object, $^W\mathrm{H}_E$ is the planned pose of the end-effector given by equations (F.7)-(F.8), and $^E\mathrm{H}_G$ is defined in Figure 6.1. The rotational and translational components of $^E\mathrm{H}_G$ are:

$$^E\mathrm{R}_G = (\mathrm{R}_2\mathrm{R}_1)^{-1}\,{}^W\mathrm{R}_O \tag{F.10}$$

$$^E\mathbf{T}_G = {}^E\mathbf{G}_{pr} - (\mathrm{R}_2\mathrm{R}_1)^{-1}({}^W\mathbf{T}_O + \mathbf{C}) \tag{F.11}$$

F.2 Grasping a Cup

Grasping for a cylindrical or conical cup follows a similar development to the previous section, except that a power grasp is employed since a precision grasp is difficult to stabilize on a curved surface. Only one candidate power grasp is possible for a cup, as shown in Figure F.2. The first step in planning is to determine whether the cup fits within the grasp, ie. $2r < |\mathbf{G}_5 - \mathbf{G}_3|$, where r is the radius of the cup and $\mathbf{G}_5$

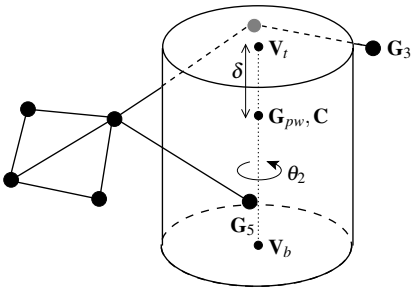


Fig. F.2. Grasp planning for an upright cup.

and $\mathbf{G}_3$ are the locations of the fingertips when the grasp is fully open. Furthermore, to ensure the hand does not collide with the floor, the height of the cup is required to satisfy $|\mathbf{V}_t - \mathbf{V}_b| > H_{th}$, where $\mathbf{V}_t$ and $\mathbf{V}_b$ (shown in Figure F.2) are the centroids of the top and bottom faces and H_{th} is the minimum height. For cups with sufficient height, the grasp centre $\mathbf{C}$ is positioned a small distance δ below the rim:

$$\mathbf{C} = \mathbf{V}_t + \delta \cdot \frac{\mathbf{V}_b - \mathbf{V}_t}{|\mathbf{V}_b - \mathbf{V}_t|} \tag{F.12}$$

To ensure the fingers exert a force approximately perpendicular to the surface, the hand is oriented to align the x-axis of the end-effector frame with the line joining $\mathbf{V}_t$ and $\mathbf{V}_b$. This alignment is achieved by rotating the end-effector by angle θ_1 about axis $\mathbf{A}_1$, given by:

$$\mathbf{A}_1 = \widehat{\mathbf{X}} \times (\mathbf{V}_t - \mathbf{V}_b)/|\mathbf{V}_t - \mathbf{V}_b| \tag{F.13}$$

$$\theta_1 = \cos(\widehat{\mathbf{X}}^\top (\mathbf{V}_t - \mathbf{V}_b)/|\mathbf{V}_t - \mathbf{V}_b|) \tag{F.14}$$

where $\widehat{\mathbf{X}} = (1,0,0)^\top$ is a unit vector in the direction of the x-axis.

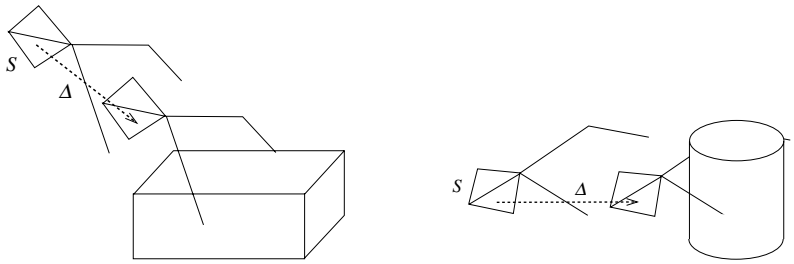
To constrain the orientation of the hand about the axis of the cup (angle θ_2 in Figure F.2), the planning algorithm minimizes the angle of the wrist. Let θ_2 and $\mathbf{A}_2 = (\mathbf{V}_t - \mathbf{V}_b)/|\mathbf{V}_t - \mathbf{V}_b|$ represent the angle/axis of rotation about the axis of the cup. After rotation, the translation of the end-effector (aligning $\mathbf{C}$ and $\mathbf{G}_{pw}$) is:

$${}^W\mathbf{T}_E(\theta_2) = \mathbf{C} - \mathrm{R}_2(\theta_2, \mathbf{A}_2)\mathrm{R}_1(\theta_1, \mathbf{A}_1)\,{}^E\mathbf{G}_{pw} \tag{F.15}$$

Now, let $\theta_5({}^W\mathbf{T}_E)$ represent the angle of the wrist when the end-effector is positioned at ${}^W\mathbf{T}_E$ (details of wrist angle calculations are provided in Section F.4). The desired rotation that minimizes the wrist angle is given by:

$$\theta_2 = \mathrm{argmin}_\theta |\theta_5({}^W\mathbf{T}_E(\theta))| \tag{F.16}$$

As in the previous section, the minimization is performed as a numerical search over all angles in steps of one degree. Finally, the transformations are combined using equations (F.7)-(F.8) (with ${}^E\mathbf{G}_{pw}$ in place of ${}^E\mathbf{G}_{pr}$) to obtain the optimal pose of



(a) Approach set-point for box. (b) Approach set-point for cup.

Fig. F.3. Trajectory planning for approaching the grasp of a box and cup.

the hand, ${}^{W}\mathrm{H}_{E}$. Finally, the reachability of the grasp is tested using the condition developed in Section (F.4). If this test is satisfied, the planned grasp is transformed to the end-effector frame ${}^{E}\mathrm{H}_{G}$ using equations (F.10)-(F.11). Otherwise, the cup is considered ungraspable.

F.3 Trajectory Planning

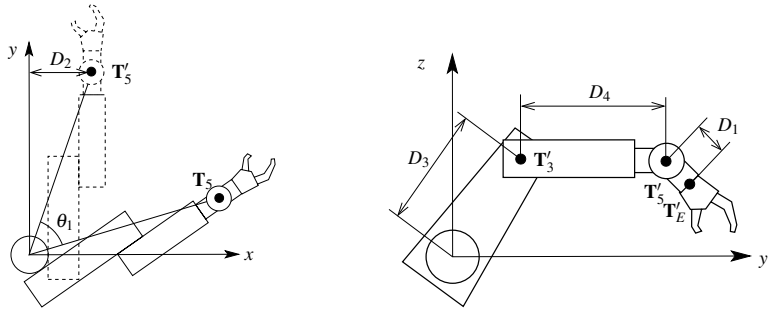
In general, a planned grasp cannot be approached from an arbitrary initial pose. Trajectory planning generates set-points to guide the motion of the end-effector along a suitable path to minimize the likelihood of collisions. A complete collision-free path planner should consider general obstacles, the target object itself and self-collisions. The planner developed in this section simplifies task planning by considering only collisions between the hand (or grasped object) and the target object. In this case, only a single set-point is required to plan a path from an arbitrary initial pose to the desired grasp (see Figure 7.4).

Figure F.3 shows the set-points generated for grasping a box and cup with the left hand. The target pose is approached away from the direction in which the fingers are pointing, which minimizes the chance of collision with the object. Let ${}^{W}\mathrm{R}_{S}$ and ${}^{W}\mathbf{T}_{S}$ represent the rotation and translation of the set-point S, and ${}^{W}\mathrm{R}_{E}$ and ${}^{W}\mathbf{T}_{E}$ represent the pose of the planned grasp. Then, the set-point (for a box or cup) is calculated as

$$ {}^{W}\mathbf{R}_{S} = {}^{W}\mathrm{R}_{E} \tag{F.17} $$

$$ {}^{W}\mathbf{T}_{S} = {}^{W}\mathbf{T}_{E} - \Delta\,{}^{W}\mathrm{R}_{E}\,{}^{E}\mathbf{F} \tag{F.18} $$

where ${}^{E}\mathbf{F}$ is the direction in which the fingers are pointing (in the end-effector frame E), and Δ is the distance between S and the final grasp. Using the planned set-point, the grasp is performed by visually servoing the hand first to S and then to the final pose.



(a) Calculation of Puma waist angle θ_1 (b) Calculation of elbow position $\mathbf{T}_3$

Fig. F.4. Puma kinematic model.

F.4 Puma Kinematic Model

This section presents a simple inverse kinematics model for the Puma arm to recover the wrist angle (Puma joint 5) for a given pose of the end-effector, and also determine whether a planned pose is within the configuration space of the arm. These parameters are required by the grasp planning algorithms developed above.

Let $\mathbf{T}_E$ and R_E represent the position and orientation of the end-effector in the robot base frame, and let $\mathbf{T}_5$ represent the position of the wrist, which is distance D_1 from the origin of the end-effector frame (see Figure F.4). The location of the wrist can be calculated as

$$\mathbf{T}_5 = \mathbf{T}_E - D_1 \widehat{\mathbf{Z}}_E \tag{F.19}$$

where $\widehat{\mathbf{Z}}_E$ is a unit vector in the direction of the Z-axis of E, obtained from the third column of R_E. To calculate the position of the elbow (joint 3), the robot base frame is first rotated about its X-axis by angle θ_1 so that the arm is parallel to the YZ-plane, as shown in Figure F.4(a). The required angle of rotation is calculated from $\mathbf{T}_5 = (X_5, Y_5, Z_5)^\top$ as

$$\theta_1 = \tan^{-1}(Y_5/X_5) - \tan^{-1}\left(\sqrt{X_5^2 + Y_5^2 - D_2^2}/D_2\right) \tag{F.20}$$

After rotation, the position of the wrist is $\mathbf{T}'_5 = (D_2, Y'_5, Z'_5)^\top$. Then, the position of the elbow $\mathbf{T}'_3 = (D_2, Y'_3, Z'_3)^\top$ is at the intersection of two circles in the YZ-plane of radius D_3 and D_4, centred at the origin and $(Y'_E, Z'_E)^\top$ respectively (see Figure F.4(b)). The point of intersection has Y and Z-coordinates:

$$Z'_3 = \frac{Z'_5(Y'^2_5 + Z'^2_5 + D_3^2 - D_4^2) + Y'_5\sqrt{4D_3^2(Y'^2_5 + Z'^2_5) - (Y'^2_5 + Z'^2_5 + D_3^2 - D_4^2)^2}}{2(Y'^2_5 + Z'^2_5)} \tag{F.21}$$

$$Y'_3 = \sqrt{D_3^2 - Z'^2_3} \tag{F.22}$$

where the positive solutions are taken for the elbow pointing away from the body. Finally, the pose of the end-effector is beyond the reach of the Puma when the solution to equation (F.21) is imaginary, and the planned grasp in this case is unreachable. When the pose is reachable, the elbow position $\mathbf{T}_3'$ is rotated by $-\theta_1$ about the x-axis to give $\mathbf{T}_3$ in the original frame. Finally, the wrist angle $\theta_5(\mathbf{T}_E)$ for the desired pose is calculated as the angle between the lines joining $\mathbf{T}_5$ to $\mathbf{T}_3$ and $\mathbf{T}_5$ to $\mathbf{T}_E$:

$$\theta_5(\mathbf{T}_E) = \text{acos}\left(\frac{(\mathbf{T}_3 - \mathbf{T}_5)^\top (\mathbf{T}_E - \mathbf{T}_5)}{|\mathbf{T}_3 - \mathbf{T}_5||\mathbf{T}_E - \mathbf{T}_5|}\right) \tag{F.23}$$

The distances D_1 through to D_4 are taken from manufacturer specifications.

References

1. B. Adams, C. Breazeal, R. A. Brooks, and B. Scassellati. Humanoid robots: A new kind of tool. *IEEE Intelligent Systems*, 15(4):25–31, 2000.
2. J. I. Agbinya and D. Rees. Multi-object tracking in video. *Real-Time Imaging*, 5:295–304, 1999.
3. G. J. Agin and T. O. Binford. Computer description of curved objects. In *Proc. 3rd Int. Joint Conf. on Artifical Intelligence*, pages 629–640, 1973.
4. M. J. Aldon and L. Le Bris. Mobile robot localization using a light stripe sensor. In *Proc. of the Intelligent Vehicles '94 Symposium*, pages 255–259, 1994.
5. H. M. M. Alshawish and C. R. Allen. 3D object recognition using coded light projection for robot assembly applications. In *Proc. 21st International Conference on Industrial Electronics, Control and Instrumentation*, volume 2, pages 1420–1247, 1995.
6. N. Andreff, R. Horaud, and B. Espiau. Robot hand-eye calibration using structure from motion. *International Journal of Robotics Research*, 20(3):228–248, 2001.
7. G. C. Atkeson, J. G. Hale, F. Pollick, M. Riley, S. Kotosaka, S. Schaal, T. Shibata, G. Tevatia, A. Ude, S. Vijayakumar, E. Kawato, and M. Kawato. Using humanoid robots to study human behaviour. *IEEE Intelligent Systems*, 15(4):45–56, 2000.
8. Y. Bar-Shalom and X-R. Li. *Estimation and Tracking: Principles, Techniques and Software*. Artech House, 1993.
9. C. M. Bastuscheck. Techniques for real-time generation of range images. In *Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, pages 262–268, 1989.
10. M. Becker, E. Kefalea, E. Maël, C. von der Malsburg, M. Pagel, J. Triesch, J. C. Vorbrüggen, R. P. Würtz, and S. Zadel. GripSee: A gesture-controlled robot for object perception and manipulation. *Autonomous Robots*, 6:203–21, 1999.
11. D.C. Bentivegna, A. Ude, C.G. Atkeson, and G. Cheng. Humanoid robot learning and game playing using PC-based vision. In *Proc. IEEE/RSJ 2002 International Conference on Intelligent Robots and Systems*, volume 3, pages 2449–2454, 2002.
12. P. J. Besl and R. C. Jain. Segmentation through variable-order surface fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2):167–192, March 1988.
13. I. Bloch. Information combination operators for data fusion: A comparative review with classification. *IEEE Transactions on Systems, Man and Cybernetics–Part A: Systems and Humans*, 26(1):52–67, 1996.
14. C. Breazeal. Socially intelligent robots: research, development, and applications. In *Proc. 2001 IEEE International Conference on Systems, Man, and Cybernetics*, volume 4, pages 2121–2126, 2001.

15. T. J. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):90–99, Jan 1986.
16. R. A. Brooks. Intelligence without representation. *Artificial Intelligence Journal*, 47:139–159, 1991.
17. D. R. Butenhof. *Programming with POSIX threads*. Addison Wesley, 1997.
18. E. Cervera and P. Martinet. Visual servoing with indirect image control and a predictable camera trajectory. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 381–386, 1999.
19. T. Chaperon and F. Goulette. Extracting cylinders in full 3D data using a random sampling method and the Gaussian image. In *Proc. Vision, Modeling and Visualization 2001*, pages 35–42, 2001.
20. F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In D. Kriegman, G. Hager, and A. Morse, editors, *The Confluence on Vision and Control*, volume 237 of *Lecture Notes in Control and Information Systems*, pages 66–78. Springer-Verlag, New York, 1988.
21. J. Chen and A. Zelinsky. Programming by demonstration: Coping with suboptimal teaching actions. *International Journal of Robotics Research*, 22(5):299–319, May 2003.
22. K. S. Chong. *Simultaneous Mapping and Localization for a Mobile Robot using Sonar Sensing*. PhD thesis, Monash University, Australia, 1997.
23. J-M. Chung and T. Nagata. Reasoning simplified volumetric shapes for robotic grasping. In *Proc. 1995 IEEE International Conference on Intelligent Robotics and Systems*, pages 348–353, 1995.
24. R. Cipolla and N. Hollinghurst. Visually guided grasping in unstructured environments. *Robotics and Autonomous Systems*, 19:337–346, 1997.
25. J. Clark, E. Trucco, and H-F. Cheung. Improving laser triangulation sensors using polarization. In *Proc. 5th Intnational Conference on Computer Vision*, pages 981–986, 1995.
26. F. S. Cohen and R. D. Rimey. A maximum likelihood approach to segmenting range data. In *Proc. 1988 IEEE International Conference on Robotics and Automation*, volume 3, pages 1696–1701, 1988.
27. P. I. Corke and M. C. Good. Dynamic effects in visual closed-loop systems. *IEEE Transactions on Robotics and Automation*, 12(5):671–683, 1996.
28. C. Dane and R. Bajcsy. Three-dimensional segmentation using the Gaussian image and spatial information. In *Proc. IEEE Computer Society Conference on Pattern Recognition and Image Processing*, pages 54–56, 1981.
29. T. Darrell, G. Gordon, M. Harville, and J. Woodfill. Integrated person tracking using stereo, color and pattern detection. *International Journal of Computer Vision*, 37(2):175–185, 2000.
30. A. Davison. *Mobile Robot Navigation Using Active Vision*. PhD thesis, University of Oxford, 1998.
31. L. Deng, W. J. Wilson, and F. Janabi-Sharifi. Characteristics of robot visual servoing methods and target model estimation. In *Proc. IEEE International Symposium on Intelligent Control*, pages 684–689, 2002.
32. M. Ehrenmann, R. Zollner, O. Rogalla, and R. Dillman. Programming service tasks in household environments by human demonstration. In *Proc. 11th IEEE International Workshop on Robot and Human Interactive Communication*, pages 460–467, 2002.

33. S. Ekvall and F. Hoffmann . Kragic. Object recognition and pose estimation for robotic manipulation using color cooccurrence histograms. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1284–1289, 2003.
34. B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3):313–326, 1992.
35. T-J. Fan, G. Medioni, and R. Nevatia. Segmented descriptions of 3-D surfaces. *IEEE Journal of Robotics and Automation*, 3(6):527–538, Dec 1987.
36. O. Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, 1993.
37. C. Ferrari and J. Canny. Planning optimal grasps. In *Proc. 1992 IEEE International Conference on Robotics and Automation*, pages 2290–95, 1992.
38. M. Fischer and G. Hirzinger. Fast planning of precision grasps of 3D objects. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 120–126, 1997.
39. D. A. Forsyth and J. Ponce. *Computer vision: a modern approach*. Prentice Hall, 2003.
40. H. Fujimoto, L.-C. Zhu, and K. Abdel-Malek. Image-based visual servoing for grasping unknown objects. In *Proc. 26th Annual Conference of the IEEE Industrial Electronics Society*, volume 2, pages 876–881, 2000.
41. V. Garric and M. Devy. Evaluation of calibration and localization methods for visually guided grasping. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 387–393, 1995.
42. C. Gaskett, P. Brown, G. Cheng, and A. Zelinsky. Learning implicit models during target pursuit. In *Proc. IEEE International Conference on Robotics and Automation*, volume 3, pages 4122–4129, 2003.
43. Chris Gaskett and Gordon Cheng. Online learning of a motor map for humanoid robot reaching. In *Proceedings of the 2nd International Conference on Computational Intelligence, Robotics and Autonomous Systems*, 2003.
44. J. R. Goldschneider and A. Q. Li. Variational segmentation by peicewise facet models with application to range imagery. In *Proc. 2001 IEEE International Conference on Image Processing*, pages 812–815, 2001.
45. E. Grosso, G. Metta, A. Oddera, and G. Sandini. Robust visual servoing in 3-D reaching tasks. *IEEE Transactions on Robotics and Automation*, 12(5):732–742, 1996.
46. E. Grosso and G. Vercelli. Grasping strategies for reconstructed unknown 3D objects. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 70–75, 1991.
47. A. Gruss, L. R. Carley, and T. Kanade. Integrated sensor and range-finding analog signal processor. *IEEE Journal of Solid-State Circuits*, 26(3):184–191, Mar 1991.
48. R. M. Haralick, J. Joo, C. Lee, X. Zhuang, V. G. Vaidya, and M. B. Kim. Pose estimation from corresponding point data. *IEEE Transactions on Systems, Man and Cybernetics*, 19(6):1426–1445, 1989.
49. R. M. Haralick, C. Lee, K. Ottenberg, and M. Nölle. Analysis and solutions of the three point perspective pose estimation problem. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 592–598, 1991.
50. R. M. Haralick, L. T. Watson, and T. J. Laffey. The topographic primal sketch. *The International Journal of Robotics Research*, 2(1):50–72, Spring 1983.
51. H. H. Harman. *Modern Factor Analysis*, chapter 8. The University of Chicago Press, 1969.
52. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

53. R. I. Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision*, 35(2):115–127, 1999.
54. R. I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997.
55. A. Hauck, J. Rüttinger, M. Sorg, and G. Färber. Visual determination of 3D grasping points on unknown objects with a binocular camera system. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 272–278, 1999.
56. A. Hauck, M. Sorg, G. Faber, and T. Schenk. What can be learned from human reach-to-grasp movements from the design of robotic hand-eye systems. In *Proc. IEEE International Conference on Robotics and Automation*, pages 2521–2526, 1999.
57. J. Haverinen and J. Röning. An obstacle detection system using a light stripe identification based method. In *Proc. IEEE International Joint Symposium on Intelligence and Systems*, pages 232–236, 1998.
58. J. Haverinen and J. Röning. A 3-D scanner capturing range and colour for the robotics applications. In *24th Workshop of the AAPR*, May 25-26, Austria 2000.
59. M. Hebert. Active and passive range sensing for robotics. In *Proc. IEEE International Conference on Robotics and Automation*, pages 102–110, 2000.
60. D. D. Hoffman. *Visual Intelligence*, chapter 2. W. W. Norton and Company, 1998.
61. N. J. Hollinghurst. *Uncalibrated Stereo and Hand-Eye Coordination*. PhD thesis, University of Cambridge, Cambridge, 1997.
62. A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7):673–689, July 1996.
63. R. Horaud and F. Dornaika. Visually guided object grasping. *IEEE Transactions on Robotics and Automation*, 14(4):525–532, 1998.
64. B. K. P Horn. Extended Gaussian images. *Proc. of the IEEE*, 72(12):1671–86, Dec 1984.
65. K. Hosoda and M. Asada. Versatile visual servoing without knowledge of the true Jacobian. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 186–191, 1994.
66. C. Hu, M. Q. Meng, P. X. Liu, and X. Wang. Visual gesture recognition for human-machine interface of robot teleoperation. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1560–1565, 2003.
67. Y. Hu, R. Eagleson, and M. A. Goodale. Human visual servoing for reaching and grasping: The role of 3-D geometric features. In *Proc. IEEE International Conference on Robotics and Automation*, pages 3209–3216, 1999.
68. S. Hutchinson, G. D. Hager, and P. I. Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5):651–670, 1996.
69. D. Q. Huynh, R. A. Owens, and P. E. Hartmann. Calibrating a structured light stripe system: A novel approach. *International Journal of Computer Vision*, 33(1):73–86, 1999.
70. K. Ikeuchi, M. Kawade, and T. Suehiro. Towards assembly plan from observation. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2294–2301, 1993.
71. *IA-32 Intel Architecture Software Developers Manual*. Intel Corporation, 2002.
72. M. Isard and A. Blake. ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework. In *Proc. 5th European Conference on Computer Vision*, volume 1, pages 893–908, 1998.
73. R. A. Jarvis. A perspective on range finding techniques for computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):122–139, 1983.

74. A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Academic Press, New York, 1970.
75. R. Joshi and A. C. Sanderson. Application of feature-based multi-view servoing for lamp filament alignment. *IEEE Robotics and Automation Magazine*, pages 25–31, December 1998.
76. R. E. Kahn, M. J. Swain, P. N. Prokopowicz, and R. J. Firby. Gesture recognition using the Perseus system. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 734–741, 1996.
77. T. Kanade and D. D. Morris. Factorization methods for structure from motion. *Philosophical Transactions of the Royal Society of London, Series A*, 356(1740):1153–73, 1998.
78. T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu. Development and evaluation of an interactive humanoid robot "Robovie". In *Proc. 2002 IEEE International Conference on Robotics and Automation*, pages 1848–1855, 2002.
79. S. B. Kang and K. Ikeuchi. Determining 3-D object pose using the complex extended Gaussian image. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 580–585, 1991.
80. E. Kefalea, E. Maël, and R. P. Würtz. An integrated object representation for recognition and grasping. In *Proc. 1999 Third International Conference on Knowledge-Based Intelligent Information Engineering Systems*, pages 423–426, 1999.
81. R. Kelly, R. Carelli, O. Nasisi, B. Kuchen, and F. Reyes. Stable visual servoing of camera-in-hand robotic systems. *IEEE Transactions on Mechatronics*, 5(1):39–48, 2000.
82. D. Khadraoui, G. Motyl, P. Martinet, J. Gallice, and F. Chaumette. Visual servoing in robotics scheme using a camera/laser-stripe sensor. *IEEE Transactions on Robotics and Automation*, 12(5):743–750, 1996.
83. H. Kim, J. Cho, and I. Kweon. A novel image-based control-law for the visual servoing system under large pose error. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 263–267, 2000.
84. L. Kleeman. Optimal estimation of position and heading for mobile robots using ultrasonic beacons and dead-reckoning. In *IEEE International Conference on Robotics and Automation*, pages 2582–2587, 1992.
85. H. Kobayashi, K. Igawa, T. Bito, and K. Kikuchi. 3D object recognition and grasping for human support robotic systems. In *Proc. 2000 IEEE International Workshop on Robot and Human Interactive Communication*, pages 430–435, 2000.
86. A. Konno, T. Yoshiike, K. Nagashima, M. Inaba, and H. Inoue. Preliminary experiments in motion programming of humanoid robot by human demonstration. *JSME International Journal*, 43(2):401–407, 2000.
87. D. Kragić. *Visual Servoing for Manipulation: Robustness and Integration Issues*. PhD thesis, Royal Institute of Technology, Stockholm, 2001.
88. D. Kragic and H. I. Christensen. Model based techniques for robotic servoing and grasping. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 299–304, 2002.
89. D. Kragić and H. I. Christensen. Survey on visual servoing for manipulation. Technical Report ISRN KTH/NA/P-02/01-SE, KTH, 2002.
90. P. Krsek, G. Lukács, and R. R. Martin. Algorithms for computing curvatures from range data. In *The Mathematics of Surfaces VIII*, pages 1–16, 1998.
91. A. Krupa, C. Doignon, J. Gangloff, and M. De Mathelin. Combining image-based and depth visual servoing applied to robotized laparoscopic surgery. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 323–329, 2002.

92. J. J. Kuffner and S. M. LaValle. RRT-connect: An efficient approach to single-query path planning. In *Proc. IEEE International Conference on Robotics and Automation*, volume 2, pages 995–1001, 2000.
93. R. Langea, P. Seitza, A. Bibera, and R. Schwarteb. Time-of-flight range imaging with a custom solid-state image sensor. In *Proc. EOS/SPIEConference on Laser Metrology and Inspection*, volume SPIE 3823, pages 180–191, 1999.
94. J. M. Lavest, G. Rives, and M. Dhome. Three-dimensional reconstruction by zooming. *Proc. IEEE Transactions on Robotics and Automation*, 9(2):196–207, 1993.
95. C. S. Lovchik and M. A. Diftler. The Robonaut hand: A dextrous robot hand for space. In *Proc. IEEE International Conference on Robotics and Automation*, pages 907–912, 1999.
96. M. Magee, R. Weniger, and E. A. Franke. Location of features of known height in the presence of reflective and refractive noise using a stereoscopic light-striping approach. *Optical Engineering*, 33(4):1092–1098, April 1994.
97. E. Malis, F. Chaumette, and S. Boudet. 2-1/2-D visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2):238–250, 1999.
98. E. Marchand, P. Bouthemy, F. Chaumette, and V. Moreau. Robust real-time visual tracking using a 2D-3D model-based approach. In *Proc. 7th IEEE International Conference on Computer Vision*, pages 262–268, 1999.
99. M. Marjanovic, B. Scassellati, and M. Williamson. Self-taught visually-guided pointing for a humanoid robot. In *Proc. Fourth International Conference on Simulation of Adaptive Behaviour*, pages 35–44, 1996.
100. D. Marshall, G. Lukacs, and R. Martin. Robust segmentation of primitives from range data in the presence of geometric degeneracy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):304–314, March 2001.
101. S.J. McKenna, Y. Raja, and S. Gong. Object tracking using adaptive colour mixture models. In *Proc. Third Asian Conference on Computer Vision*, pages 615–622, 1997.
102. P. J. McKerrow. *Introduction to Robotics*. Addison-Wesley, 1991.
103. P. Michelman and P. Allen. Forming complex dextrous manipulations from task primitives. In *Proc. IEEE International Conference on Robotics and Automation*, volume 4, pages 3383–3388, 1994.
104. A. Morales, E. Chinellato, A. H. Fagg, and A. P. del Pobil. Experimental prediction of the performance of grasp tasks from visual features. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3423–3428, 2003.
105. J. J. Moré, B. S. Garbow, and K. E. Hillstrom. Users' guide for MINPACK-1. Technical Report ANL-80-74, Applied Math. Div., Argonne National Laboratory, 1980.
106. E. Mouaddib, J. Batllea, and J. Salvia. Recent progress in structured light in order to solve the correspondence problem in stereovision. In *Proc. IEEE International Conference on Robotics and Automation*, pages 130–136, 1997.
107. M. C. Moy. Gesture-based interaction with a pet robot. In *Proc. of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence*, pages 628–633, 1999.
108. M. Müller and H. Wörn. Planning of rapid grasp operations in unstructured scenes. In *Proc. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 1975–80, 2000.
109. Kouichi Nakano, Yasuo Watanabe, and Sukeyasu Kanno. Extraction and recognition of 3-dimensional information by projecting a pair of slit-ray beams. In *Proceedings of the 9th International Conference on Pattern Recognition*, pages 736–738, 1988.

110. K. Namba and N. Maru. 3D linear visual servoing for humanoid robot. In *28th Annual Conference of the IEEE Industrial Electronics Society*, volume 3, pages 2225–2230, 2002.
111. S. K. Nayar, A. C. Sanderson, L. E. Weiss, and D. A. Simon. Specular surface inspection using structured highlight and Gaussian image. *IEEE Transactions on Robotics and Automation*, 6(2):208–218, April 1990.
112. J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.
113. B.J. Nelson and P.K. Khosla. An extendable framework for expectation-based visual servoing using environment models. In *Proc. IEEE International Conference on Robotics and Automation*, pages 184–189, 1995.
114. K. Nickels and S. Hutchinson. Model-based tracking of complex articulated objects. *IEEE Transactions on Robotics and Automation*, 17(1):28–36, 2001.
115. H. Nomura and T. Naito. Integrated visual servoing system to grasp industrial parts moving on conveyer by controlling 6DOF arm. In *Proc. IEEE/RSJ International Conference on Systems, Man and Cybernetics*, pages 1768–1775, 2000.
116. J. Nygards, T. Högström, and Å. Wernersson. Docking to pallets with feedback from a sheet-of-light range camera. In *Proc. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1853–1859, 2000.
117. J. Nygards and Å. Wernersson. Specular objects in range cameras: Reducing ambiguities by motion. In *Proc. of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 320–328, 1994.
118. K. Okada, M. Inaba, and H. Inoue. Integration of real-time binocular stereo vision and whole body information for dynamic walking navigation of humanoid robot. In *Proc. IEEE Conference on Multisonsor Fusion and Integration for Intelligent Systems*, pages 131–136, 2003.
119. M. Okada, Y. Nakamura, and S. Ban. Design of programmable passive compliance shoulder mechanism. In *Proc. 2001 IEEE International Conference on Robotics and Automation*, volume 1, pages 348–353, 2001.
120. T. Okada, M. Sano, and H. Kaneko. Three-dimensional object recognition usind spherical correlation. In *Proc. 11th IAPR International Conference on Pattern Recognition*, volume 1, pages 250–254, 1992.
121. T. Okatani and K. Deguchi. Computation of the sign of the Gaussian curvature of a surface from multiple unknown illumination images without knowledge of the reflectance property. *Computer Vision and Image Understanding*, 76(2):125–134, November 1999.
122. V. I. Pavlovic, R. Sharma, and T. S. Huang. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):677–695, 1997.
123. M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual modeling with a hand-held camera. *International Journal of Computer Vision*, 59(3):207–232, 2004.
124. M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. IEEE 6th International Conference on Computer Vision*, pages 90–95, 1998.
125. P. N. Prokopowicz, M. J. Swain, and R. E. Kahn. Task and environment-sensitive tracking. In *Proc. IARP/IEEE Workshop on Visual Behaviours*, pages 73–78, 1994.
126. K. Rao, G. Medioni, H. Liu, and G. A. Bekey. Shape description and grasping for robot hand-eye coordination. *IEEE Control Systems Magazine*, pages 22–29, February 1989.
127. R. P. N. Rao and D. H. Ballard. Learning saccadic eye movements using multiscale spatial filters. *Advances in Neural Information Processing Systems*, 7:893–900, 1995.

128. H. J. Ritter, T. M. Martinetz, and K. J. Schulten. Topology-conserving maps for learning visuo-motor-coordination. *Neural Networks*, 2(3):159–168, 1989.
129. P. Rives. Visual servoing based on epipolar geometry. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 602–607, 2000.
130. M. A. Rodrigues, Y. F. Li, M. H. Lee, J. J. Rowland, and C. King. Robotic grasping of complex objects without full geometrical knowledge of shape. In *Proc. IEEE International Conference on Robotics and Automation*, pages 737–742, 1995.
131. O. Rogalla, M. Ehrenmann, R. Zollner, R. Becher, and R. Dillman. Using gesture and speech control for commanding a robot assistant. In *Proc. 11th IEEE International Workshop on Robot and Human Interactive Communication*, pages 454–459, 2002.
132. R. A. Russell and A. H. Purnamadjaja. Odor and airflow: complementary senses for a humanoid robot. In *Proc. IEEE International Conference on Robotics and Automation*, volume 2, pages 1842–1847, 2002.
133. R. A. Russell, G. Taylor, L. Kleeman, and A. H. Purnamadjaja. Multi-sensory synergies in humanoid robotics. *International Journal of Humanoid Robotics*, 1(2):289–314, 2004.
134. Y. K. Ryu and H. S. Cho. A neural network approach to extended Gaussian image based solder joint inspection. *Mechatronics*, 7(2):159–184, 1997.
135. A. C. Sanderson and L. E. Weiss. Image-based visual servo control using relational graph error signals. In *Proc. IEEE Conference on Cybernetics and Society*, pages 1074–1077, 1980.
136. B. Scassellati. A binocular, foveated active vision system. Technical Report 1628, MIT Artificial Intelligence Lab, 1998.
137. S. Schaal. Is imitation learning the route to humanoid robots. *Trends in Cognitive Sciences*, 3(6):233–242, 1999.
138. J. Shi and C. Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
139. Y. Shirai and M. Suwa. Recognition of polyhedrons with a range finder. In *Proc. 2nd Int. Joint Conf. on Artifical Intelligence*, pages 80–87, 1971.
140. G. Smith, E. Lee, K. Goldberg, K. Böhringer, and J. Craig. Computing parallel-jaw grips. In *Proc. 1999 IEEE International Conference on Robots and Automation*, pages 1897–1903, 1999.
141. M. Spengler and B. Schiele. Towards robust multi-cue integration for visual tracking. *Machine Vision and Applications*, 14:50–158, 2003.
142. J. Stavnitsky and D. Capson. Multiple camera model-based 3-D visual servo. *IEEE Transactions on Robotics and Automation*, 16(6):732–739, Dec 2000.
143. C. Sun and J. Sherrah. 3D symmetry detection using the extended Gaussian image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):164–169, February 1997.
144. A. Takanishi, M. Ishida, Y. Yamazaki, and I. Kato. The realization of dynamic walking by the biped walking robot wl-10rd. In *Proc. '85 International Conference on Advanced Robotics*, pages 459–466, 1985.
145. A. Takanishi, H. Lim, M. Tsuda, and I. Kato. Realization of dynamic biped walking stabilized by trunk motion on a sagittally uneven surface. In *Proc. IEEE International Workshop on Intelligent Robots and Systems*, pages 323–330, 1990.
146. M. Takizawa, Y. Makihara, N. Shimada, J. Miura, and Y. Shirai. A service robot with interactive vision – object recognition using dialog with user –. In *Fist International Workshop on Language Understanding and Agents for Real World Interaction*, pages 16–23, 2003.

147. G. Taylor and L. Kleeman. Grasping unknown objects with a humanoid robot. In *Proc. 2002 Australiasian Conference on Robotics and Automation*, pages 191–196, 2002.
148. G. Taylor and L. Kleeman. Robust range data segmentation using geometric primitives for robotic applications. In *Proc. 9th IASTED International Conference on Signal and Image Processing*, pages 467–472, 2003.
149. G. Taylor and L. Kleeman. Hybrid position-based visual servoing with online calibration for a humanoid robot. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 686–691, 2004.
150. G. Taylor and L. Kleeman. Integration of robust visual perception and control for a domestic humanoid robot. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1010–1015, 2004.
151. G. Taylor and L. Kleeman. Stereoscopic light stripe scanning: Interference rejection, error minimization and calibration. *International Journal of Robotics Research*, 23(12):1141–1156, December 2004.
152. M. Tonko, K. Schäfer, F. Heimes, and H.-H. Nagel. Towards visually servoed manipulation of car engine parts. In *Proc. IEEE International Conference on Robotics and Automation*, pages 1366–1371, 1997.
153. K. Toyama and G.D. Hager. Incremental focus of attention for robust vision-based tracking. *International Journal of Computer Vision*, 35(1):45–63, 1999.
154. E. Trucco and R. B. Fisher. Computing surface-based representations rrom range images. In *Proc. IEEE International Symposium on Intelligent Control*, pages 275–280, 1992.
155. E. Trucco and R. B. Fisher. Acquisition of consistent range data using local calibration. In *Proc. of the 1994 IEEE International Conference on Robotics and Automation*, volume 4, pages 3410–3415, 1994.
156. R. Y. Tsai and R. K. Lenz. A new technique for fully autonomous and efficient 3D robotic hand/eye calibration. *IEEE Transactions on Robotics and Automation*, 5(3):345–358, 1989.
157. T. Uhlin, P. Nordlund, A. Maki, and J.-O. Eklundh. Towards an active visual observer. In *Proc. of the International Conference on Computer Vision*, pages 679–686, Cambridge, MA, 1995.
158. K. Usher, P. Ridley, and P. Corke. Visual servoing for a car-like vehicle - an application in omnidirectional vision. In *Proc. 2002 Australasian Conference on Robotics and Automation*, pages 37–42, 2002.
159. S. van der Zwann, A. Bernardino, and J. Santos-Victor. Vision based station keeping and docking for an aerial blimp. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 614–619, 2000.
160. P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.
161. G. Welch and G. Bishop. SCAAT: Incremental tracking with incomplete information. *Computer Graphics*, 31:333–344, 1997.
162. D. S. Wheeler, A. H. Fagg, and R. A. Grupen. Learning prospective pick and place behaviour. In *Proc. The 2nd International Conference on Development and Learning*, pages 197–202, 2002.
163. W. Wilson, C. Williams Hulls, and G. Bell. Relative end-effector control using cartesian position based visual servoing. *IEEE Transactions on Robotics and Automation*, 12(5):684–696, 1996.
164. P. Wira and J. P. Urban. A new adaptive Kalman filter applied to visual servoing tasks. In *Fourth International Conference on Knowledge-Based Intelligent Engineering Systems and Applied Technologies*, pages 267–270, 2000.

165. P. Wunsch and G. Hirzinger. Real-time visual tracking of 3-D objects with dynamic handling of occlusion. In *Proc. IEEE International Conference on Robotics and Automation*, pages 2868–2873, 1997.
166. J. Vanden Wyngaerd and L. Van Gool. Coarse registration of surface patches with local symmetries. In *Proc. 7th European Conference on Computer Vision*, pages 572–586, 2002.
167. H. S. Yang and A. C. Kak. Determination of the identity, position and orientation of the topmost object in a pile: Some futher experiments. In *Proc. 1986 IEEE International Conference on Robotics and Automation*, volume 1, pages 293–298, 1986.
168. Y. Yokokohji, M. Sakamoto, and T. Yoshikawa. Vision-aided object manipulation by a multifingered hand with soft fingertips. In *Proc. IEEE International Conference on Robotics and Automation*, pages 3201–3208, 1999.

Index