

Realistic Stereo Error Models and Finite Optimal Stereo Baselines

Tao Zhang*

Vision and Technology Lab

University of Colorado at Colorado Springs

zhangt@vast.uccs.edu

Terry Boulton†

VAST Lab UCCS and

Securics Inc.

tboulton@vast.uccs.edu

Abstract

Stereo reconstruction is an important research and application area, both for general 3D reconstruction and for operations like robotic navigation and remote sensing. This paper addresses the determination of parameters for a stereo system to optimize/minimize 3D reconstruction errors. Previous work on error analysis in stereo reconstruction optimized error in disparity space which led to the erroneous conclusion that, ignoring matching errors, errors decrease when the baseline goes to infinity.

In this paper, we derive the first formal error model based on the more realistic “point-of-closest-approach” ray model used in modern stereo systems. We then show this results in finite optimal baseline that minimizes reconstruction errors in all three world directions. We also show why previous oversimplified error analysis results in infinite baselines. We derive the mathematical relationship between the error variances and the stereo system parameters. In our analysis, we consider the situations where errors exist in only one camera as well as errors in both cameras. We have derived the results for both parallel and verged systems, though only the simpler models are presented algebraically herein. The paper includes simulations to highlight the results and validate the approximations in the error propagation. The results should allow stereo system designers, or those using motion-stereo, to improve their system.

1 Introduction

An important task in stereo reconstruction is how to obtain accurate three-dimensional reconstruction positions even when errors exist in matching/locating of points. Prevailing wisdom, based on decade’s old work, is that larger baselines are better, at least until the gross matching errors and 3D geometry-variations dominate the error the baseline induces from localization error. As part of a project looking for motion-stereo, using a single-camera UAV-based measurements, we sought to determine the optimal baseline to balance these errors. While it turned out it was too difficult to model the probability and impact of gross matching errors, our effort led us to question the validity of existing models which result suggest “bigger is better”. This paper may appear is somewhat theoretical, but the underlying equations

*Initial work done while at UCCS. Now with MediaTek, San Jose,CA

†This work was supported by NSF PFI Award #065025 and ONR STTR ONR MURI N00014-08-1-0638

for optimal designs are applicable to design/construction of stereo systems and for parameter selection for motion-stereo systems. The new error models could also be useful in stereo algorithms that attempt to optimize reconstruction with respect to model errors. In the limited space we could not present both the useful theory and details of the application.

In practical applications such as stereo navigations [9], object tracking [5], biometrics, remote sensing, motion-based stereo and wave height monitoring [2], it is often desirable to know how to determine the extrinsic parameters such as the baseline and the rotation angle between cameras or intrinsic parameters such as focal length, so that the operational errors will be minimized. We noted that the error of interest needs not be the traditional error along the viewing direction, but may be the error in the world space, e.g. “height” of a potential obstacle, where both in-image localization errors and the depth computation error, impacts the overall error in the height measurement. For example, if the camera were almost parallel to the floor, the height of the obstacle is measured within the image plane, and errors in depth mostly impact where the obstacle is, not its height.

This two observations lead to new questions/models for stereo errors and for camera positioning. In order to determine the optimal parameters of a stereo system to minimize the reconstruction errors, we need to know the relationship between these parameters and the reconstruction errors. In the last two decades, many papers analyzed this problem. By the errors modeled, these methods can be roughly classified into two groups:

The first group uses scalar errors such as bounded quantization errors or worst case analysis [1, 13, 3, 8, 4, 15] and derives upper bounds or lower bounds of errors or the probability of errors less than a given tolerance. As Matthies pointed out in [9], “the uncertainty induced by triangulation is not a simple scalar function of distance to the point; it is also skewed and oriented. Scalar error measures do not capture these distinctions in shape.” In addition, many of these papers use relative errors in their analysis, the parameters in which we are interested, such as the baseline, are canceled out from the relative errors and cannot be optimized via the relative error formula.

The second group of stereo error analysis papers use Gaussian error models [9, 11, 12] or simply analyze the absolute error [16].

Both these groups of papers share two important common features. For those papers that have analytical error results on the relation between the systems parameters and the reconstruction errors, the results show that the depth reconstruction errors decrease as the baseline increases, and thus are minimized when the baseline goes to infinity.

The second is how they use triangulation in their analysis. It is well known from triangulation that the reconstruction point is the intersection point of the two view rays, *if there is no noise*, but that in general the two viewing rays will not intersect in the space if noise exists in 3D [14]. In operational stereo systems, people use the point of closest approach of the two noisy rays. However, all the error analysis papers simply use noise free version for simplicity or even analyze errors in 2D case when dealing with 3D errors [10, 16, 15, 9].

One of our key contributions is the derivation of stereo error models using the point of closest approach for noisy rays. Given this more realistic model, we derive the error formulas for all three world dimensions. Our second contribution is to use these models to define optimal baseline, which in general is finite. From the model, we can optimize system parameters, e.g. optimal viewing angle for obstacle height measurements, or the optimal baseline for a verged stereo system with targets at 1-3 meters distance. The initially surprising result, that stereo systems have finite optimal baselines, follows directly from the model.

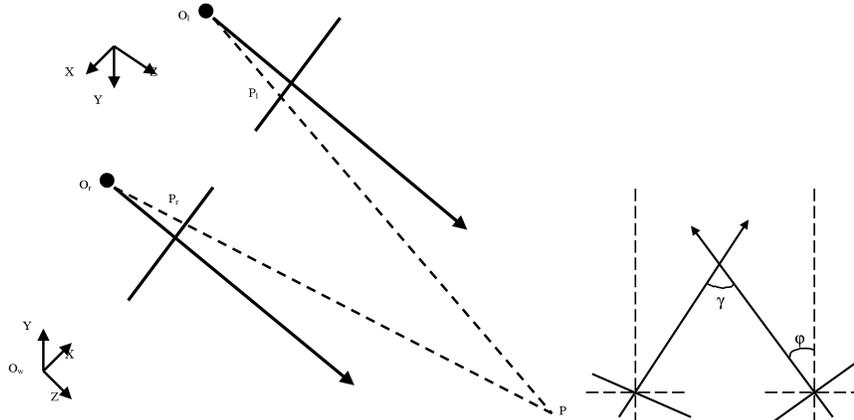


Figure 1. Stereo system setup and the relationship between the verge angle φ and the rotation angle γ

Section 2 describes the stereo system model used in this paper. Section 3 shows how to determine the reconstruction point. Section 4 analyzes the error propagation from the image frame to the camera frame and to the world frame, gives the formulas for finite optimal baselines when errors are modeled in only one camera. Section 5 shows the error propagation when errors are modeled in both cameras. Section 6 shows simulation results and section 7 summarizes the paper.

2 Description of Stereo System

The stereo system used in this paper is shown in Figure 1. The basic assumption used in this paper is that the cameras used are pinhole cameras or well calibrated optically. Therefore, we do not consider the lens distortion problems in this paper. For simplicity, further assume the two cameras in the stereo system have the same focal length f . Now, let us define the relationships between different reference frames.

Let P_l and P_r be the coordinates of a point P in the left and right camera frame respectively; (x_l, y_l) and (x_r, y_r) be the projection of P onto the left and right camera plane.

The relation between P_l and P_r is:

$$P_r = R_c(P_l - T_c)$$

where R_c is the rotation matrix given by:

$$R_c = \begin{pmatrix} \cos(\gamma) & 0 & \sin(\gamma) \\ 0 & 1 & 0 \\ -\sin(\gamma) & 0 & \cos(\gamma) \end{pmatrix}$$

Without loss of generality, let $T_c = [B \ 0 \ 0]^T$. B is the baseline. As shown in Figure 1, the verge angle $\varphi = \frac{\gamma}{2}$. In the analysis of this paper, we will assume $R_c = I_{3 \times 3}$. That is, the principle axes of the two cameras are parallel. The analysis steps for the cases where two cameras are not parallel are same. However, the analytical results for those cases are too long to include in the paper, and will appear in a yet to be released technical report.

In this paper, we generally use the left camera frame as the reference frame. The relation between the left camera frame P_l and the world frame P_w is given by:

$$P_w = R_w(P_l - T_w) \quad (1)$$

where R_w is the rotation matrix and T_w is the translation vector. R_w is given by:

$$R_w = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\cos \theta & -\sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \quad (2)$$

where θ is the rotation angle. Because of the definition of our coordinate system as shown in Figure 1, R_w is obtained by rotation along the x axis followed by sign changing in the x and y directions.

The rotation angle θ is also the view angle. We consider the view angle case because it is commonly used in applications.

3 Determination of Reconstruction Point

When both the intrinsic and extrinsic parameters of the stereo system are known, the reconstruction can be easily done by triangulation: the intersection of ray $O_l P_l$ and $O_r P_r$ is the reconstruction point. However, it is well known that these two rays may not intersect in space if P_l and P_r are not accurate.

Two popular methods are used in practice. One method takes the point whose distances to both rays are minimized [14]. The other method seeks a point which gets a new pair of projections onto the two cameras, and the distance between the new projection points and P_l, P_r is minimized [7]. We will use the former in our analysis because it can be used analytically while the latter cannot. We briefly summarize the method below.

Let the point determined from (x_l, y_l) and (x_r, y_r) be P_2 in the left camera frame. Let $l = [x_l, y_l, f]^T$, $r = T_c + \beta R_c^T [x_r, y_r, f]^T$, and $w = l \times R_c^T [x_r, y_r, f]^T$. In the following sections, we consider the case where $\gamma = 0$, then α, β and λ can be determined by the following equation.

$$\alpha l + \lambda w = r \quad (3)$$

P_2 is given by:

$$P_2 = \alpha l + \frac{\lambda}{2} w \quad (4)$$

where

$$\begin{aligned} \alpha &= \frac{x_r y_l y_r - x_l y_r^2 + (x_r - x_l) f^2}{\Delta} B \\ \lambda &= \frac{B(y_r - y_l) f}{\Delta} \\ \Delta &= \begin{vmatrix} x_l & -x_r & (y_l - y_r) f \\ y_l & -y_r & (x_r - x_l) f \\ f & -f & x_l y_r - x_r y_l \end{vmatrix} \end{aligned} \quad (5)$$

In noise free cases, P_2 is reduced to

$$P_2 = \frac{B}{d} \begin{pmatrix} x_l \\ y_l \\ f \end{pmatrix} \quad (6)$$

where B is the baseline, $d = x_l - x_r$ is the disparity.

4 Errors in One Camera

We now turn to the error propagation in the stereo system when errors are modeled only in one camera, e.g. assuming the point in the left image is the definition of the “correct” feature. This is a common approach for a fixed and rigid stereo system. Assume the left camera coordinate (x_l, y_l) for a point is accurate, and the errors in the right camera can be modeled by zero mean Gaussian distributions with covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \tag{7}$$

The mapping defined in equation (4) is nonlinear but in general quite smooth. If we model this mapping as approximately affine in the vicinity of the mean of the distribution, then the errors of P_2 in the camera frame is also a Gaussian distribution with the following parameters [6]:

$$\mu_{p_2} = \frac{B}{d} [x_l, y_l, f]^T \tag{8}$$

$$\Sigma_c = J_c \Sigma J_c^T \tag{9}$$

$$J_c = \frac{B}{d^2} \begin{bmatrix} x_l & \frac{-x_l x_r y_l}{y_l^2 + f^2} \\ y_l & \frac{-x_r y_l^2 - 0.5(x_r - x_l)f^2}{y_l^2 + f^2} \\ f & \frac{-0.5y_l f(x_l + x_r)}{y_l^2 + f^2} \end{bmatrix} \tag{10}$$

J_c is the Jacobian matrix computed from the mapping (4).

In the remaining part of the paper, we use following abbreviations to save space. Let $J_2 = x_r y_l^2 + 0.5(x_r - x_l)f^2$, $J_3 = 0.5y_l f(x_l + x_r)$, $J_4 = x_l y_l^2 + 0.5(x_l - x_r)f^2$ and $b = y_l^2 + f^2$.

4.1 Optimal Finite Baselines

From Σ_c obtained by equation (9), we can obtain the variances in three directions as following:

$$\sigma_{cx}^2 = \frac{x_l^2 B^2}{d^4} \left(\sigma_x^2 + x_r^2 y_l^2 \frac{\sigma_y^2}{b^2} \right) \tag{11}$$

$$\sigma_{cy}^2 = \frac{B^2}{d^4} \left(y_l^2 \sigma_x^2 + J_2^2 \frac{\sigma_y^2}{b^2} \right) \tag{12}$$

$$\sigma_{cz}^2 = \frac{B^2 f^2}{d^4} \left(\sigma_x^2 + \frac{y_l^2 (x_l + x_r)^2}{4} \frac{\sigma_y^2}{b^2} \right) \tag{13}$$

These variances are expressed in the accurate projection coordinates (x_l, y_l) and (x_r, y_r) . Note $y_l = y_r$. Using equation (14), these variances are then represented by the projection coordinate (x_l, y_l) and the Z coordinate of a point in the left camera frame. Finite optimal baselines can be obtained by minimizing these variances.

- Optimal baseline to minimize the depth error Let us first determine the finite optimal baseline for the most popular depth errors. To obtain this optimal baseline, we fix the left camera and move the right camera to change the baseline. So x_r will change with the baseline. Let $P_l = [X \ Y \ Z]^T$ be a point represented in left camera frame. The relation between the baseline B and x_r is given by:

$$x_r = x_l - \frac{B}{Z}f \quad (14)$$

substitute equation (14) into equation (13), then

solve the equation $\frac{d\sigma_{cz}^2}{dB} = 0$, we get the optimal baseline to minimize the depth error.

$$B_{cz} = \frac{2Z}{f} \left(x_l + \frac{b^2\sigma_x^2}{x_l y_l^2 \sigma_y^2} \right) \quad (15)$$

- Optimal baseline to minimize width and height errors

Following similar steps, we can determine the optimal baselines for minimizing the width and height errors:

$$B_{cx} = \frac{Z}{f} \left(x_l + \frac{b^2\sigma_x^2}{x_l y_l^2 \sigma_y^2} \right) \quad (16)$$

$$B_{cy} = \frac{2Z}{f(f^2 + 2y_l^2)} \left(x_l y_l^2 + \frac{b^2\sigma_x^2}{x_l \sigma_y^2} \right) \quad (17)$$

Note that in this case, we have $B_{cz} = 2B_{cx}$.

We showed if we use the mapping (4), optimal baselines exist. Next, we will explain why using mapping (6) does not yield finite optimal baselines.

4.2 What If Noise-free Triangulation Is Used

As we mentioned in the introduction, prior papers on stereo error analysis assume ideal ray-intersection for stereo reconstruction. Even though most are aware of the fact that with noise the two rays $O_l P_l$ and $O_r P_r$ will no longer intersect in space, they use noise-free triangulation or analyze the 3D errors in a reduced 2D space for simplicity. We now show why such simplification leads to infinite baseline conclusion in those papers.

When the ideal triangulation is used, the mapping is reduced to mapping (6). The corresponding Jacobian matrix is shown in equation (18). Note the item $x_l - x_r$ in this Jacobian matrix can only contribute $\frac{1}{B}$ in the propagated covariances.

$$\tilde{J}_c = \frac{B}{d^2} \begin{bmatrix} x_l & 0 \\ y_l & x_l - x_r \\ f & 0 \end{bmatrix} \quad (18)$$

The computed variances in three directions are:

$$\tilde{\sigma}_{cx}^2 = \frac{x_l^2 B^2 \sigma_x^2}{d^4} = \frac{x_l^2 Z^4 \sigma_x^2}{B^2 f^4} \quad (19)$$

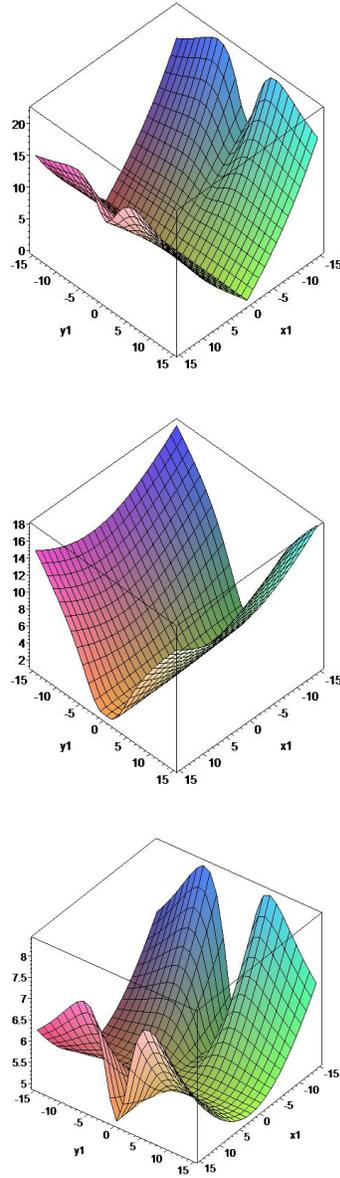


Figure 2. Error variances in width, height and depth (top to bottom) with all errors in single-camera. Though the magnitude are different, the overall shapes are very similar for the dual-camera error model.

$$\tilde{\sigma}_{cy}^2 = \frac{(y_t^2 Z^2 \sigma_x^2 + B^2 f^2 \sigma_y^2) Z^2}{B^2 f^4} \quad (20)$$

$$\tilde{\sigma}_{cz}^2 = \frac{B^2 f^2 \sigma_x^2}{d^4} = \frac{Z^4 \sigma_x^2}{B^2 f^2} \quad (21)$$

In order to minimize these errors , the baseline B has to be as large as possible and tends to infinity.

4.3 Error Propagation to World Frame

As stated, there are applications where the goal is not optimal range estimates, but optimal measurements in the world coordinate systems. In this section, we consider the errors propagated to the world frame. Similarly, the error of mapping (4) propagated from the camera frame to the world frame may also be modeled by Gaussian distributions with the following parameters:

$$\mu_{p_w} = R_w(\mu_{p_c} - T_w) \quad (22)$$

$$\Sigma_w = R_w \Sigma_c R_w^T \quad (23)$$

From the covariance matrix (23), the variances for width, height and depth (in x, y and z directions respectively) are:

$$\begin{aligned} \sigma_{wx}^2 &= \frac{B^2 x_l^2}{d^4} \left(\sigma_x^2 + x_r^2 y_l^2 \frac{\sigma_y^2}{b^2} \right) \\ \sigma_{wy}^2 &= \frac{B^2}{d^4} \left((y_l \cos \theta + f \sin \theta)^2 \sigma_x^2 + (J_2 \cos \theta + J_3 \sin \theta)^2 \frac{\sigma_y^2}{b^2} \right) \\ \sigma_{wz}^2 &= \frac{B^2}{d^4} \left((y_l \sin \theta - f \cos \theta)^2 \sigma_x^2 + (J_3 \cos \theta - J_2 \sin \theta)^2 \frac{\sigma_y^2}{b^2} \right) \end{aligned}$$

The overall error $\sigma_w^2 = \sigma_{wx}^2 + \sigma_{wy}^2 + \sigma_{wz}^2$ is:

$$\sigma_w^2 = \frac{B^2}{d^4} \left((x_l^2 + y_l^2 + f^2) \sigma_x^2 + ((x_l x_r y_l)^2 + J_2^2 + J_3^2) \frac{\sigma_y^2}{b^2} \right)$$

The variances of errors in three directions over a small region are shown in Figure 2. Due to the view angle θ , the variances of height and depth are influenced by θ as shown in Figure 4, and as expected, the variances of width and overall errors are not. Similarly, the optimal baselines can be computed:

- Optimal baseline to minimize the depth error

$$B_{wz} = \frac{2Z(y_l \sin \theta - f \cos \theta)(x_l^2 y_l^2 \sigma_y^2 + b^2 \sigma_x^2)}{x_l y_l f \sigma_y^2 (2y_l^2 \sin \theta + f^2 \sin \theta - y_l \cos \theta)}$$

- Optimal baseline to minimize the height error

$$B_{wy} = \frac{2Z(y_l \cos \theta + f \sin \theta)(b^2 \sigma_x^2 + x_l^2 y_l^2 \sigma_y^2)}{x_l y_l f \sigma_y^2 (2y_l^2 \cos \theta + f^2 \cos \theta + y_l f \sin \theta)}$$

- Optimal baseline to minimize the width error

$$B_{wx} = B_{cx}$$

- Optimal baseline to minimize the overall error

$$B_{wa} = \frac{Z}{f} \left(\frac{b^2 \sigma_x^2}{x_l y_l^2 \sigma_y^2} + x_l \right) \tag{24}$$

Note the optimal baseline value may be negative for some points. For a single point, the negative baseline means changing the right camera to the left camera. However, if we require the baseline to be positive, this negative optimal baseline means there is no finite positive optimal baseline for that point, so at that point, the optimal positive baseline is infinite. In practical applications, an area is generally considered instead of a single point, given an optimization criteria, a finite optimal baseline can be obtain for that area using the single point optimal baseline formulas over that area.

The variances of errors and optimal baselines in the camera frame as shown in section 4.1 can be obtained from the corresponding values in the world frame by setting θ to zero.

4.4 Consider Image Frame Directly

In above analysis, the projection of a point to the image plane is expressed in terms of the camera frame coordinates. If we know all of the intrinsic parameters of the cameras, the error propagation between the image frame and the camera frame is given below:

$$\mu_{im} = \begin{bmatrix} s_x^{-1} & 0 \\ 0 & s_y^{-1} \end{bmatrix} (\mu - [o_x \ o_y]^T) \tag{25}$$

$$\Sigma_{im} = \begin{bmatrix} \frac{\sigma_x^2}{s_x^2} & 0 \\ 0 & \frac{\sigma_y^2}{s_y^2} \end{bmatrix} \tag{26}$$

where o_x, o_y, s_x, s_y are the intrinsic parameters of the cameras. In our error model, the influence of quantification errors is assumed to be absorbed into the covariance matrix Σ .

5 Errors in Two Cameras

Now we consider the case where errors are modeled in both cameras. This is particularly useful when doing verging stereo systems or motion-based stereo. Due to the similarity in the analysis steps, we do not repeat the whole process, but give the results in the final steps directly. As showed before, the variances of errors and optimal baselines in the camera frame can be computed from corresponding values in the world frame, so we only give the results in the world frame.

For simplicity of presentation, we assume the errors in both cameras share the same covariance matrix. The combined covariance matrix Σ_2 is:

$$\Sigma_2 = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_x^2 & 0 \\ 0 & 0 & 0 & \sigma_y^2 \end{bmatrix} \tag{27}$$

The mean values are computed the same way as shown in equation (22). The corresponding Jacobian matrix J_{c2} is:

$$J_{c2} = \frac{B}{d^2} \begin{bmatrix} x_l & \frac{-x_l y_l x_r}{b} & -x_r & \frac{x_l y_l x_r}{b} \\ y_l & \frac{-J_2}{b} & -y_l & \frac{J_4}{b} \\ f & \frac{-J_3}{b} & -f & \frac{J_3}{b} \end{bmatrix} \tag{28}$$

The covariance matrix propagated into the world frame can be computed as $\Sigma_{w2} = R_w J_{c2} \Sigma_2 J_{c2}^T R_w^T$. From Σ_{w2} , the variances of errors in all three directions are:

$$\sigma_{wx2}^2 = \frac{B^2}{d^4} \left((x_l^2 + x_r^2) \sigma_x^2 + \frac{2x_l^2 y_l^2 x_r^2}{b^2} \sigma_y^2 \right) \quad (29)$$

$$\sigma_{wy2}^2 = \frac{B^2}{d^4} \left(2(f \sin \theta + y_l \cos \theta)^2 \sigma_x^2 + \frac{\sigma_y^2}{b^2} \left((J_2 \cos \theta + J_3 \sin \theta)^2 + (J_4 \cos \theta + J_3 \sin \theta)^2 \right) \right) \quad (30)$$

$$\sigma_{wz2}^2 = \frac{B^2}{d^4} \left(2(y_l \sin \theta - f \cos \theta)^2 \sigma_x^2 + \frac{\sigma_y^2}{b^2} \left((J_3 \cos \theta - J_2 \sin \theta)^2 + (J_4 \sin \theta - J_3 \cos \theta)^2 \right) \right) \quad (31)$$

The variance of the overall error σ_{w2} is defined as:

$$\sigma_{w2}^2 = \sigma_{wx2}^2 + \sigma_{wy2}^2 + \sigma_{wz2}^2$$

The optimization of variances in terms of B results in the optimal baselines for x, y, z directions and overall errors:

$$B_{wx2} = \frac{2x_l Z (b^2 \sigma_x^2 + x_l^2 y_l^2 \sigma_y^2)}{f (b^2 \sigma_x^2 + 2x_l^2 y_l^2 \sigma_y^2)} \quad x_l \neq 0 \quad (32)$$

$$B_{wy2} = B_{wz2} = \frac{2Z}{f} \left(x_l + \frac{b^2 \sigma_x^2}{x_l y_l^2 \sigma_y^2} \right) \quad (33)$$

$$B_{w2} = \frac{2Z (x_l^2 + y_l^2 + f^2) (b^2 \sigma_x^2 + x_l^2 y_l^2 \sigma_y^2)}{x_l f (b^2 \sigma_x^2 + y_l^2 (2x_l^2 + y_l^2 + f^2) \sigma_y^2)} \quad (34)$$

Note that optimal baselines for y and z directions do not depend on θ even though the variances of errors in these directions are affected by θ . Importantly, we can optimize the baseline for overall error and don't need to modify it for different world measurements. Not surprisingly, optimal stereo system height and viewing angle, do depend on the measurement desired, but are too complex to present herein.

While not presented in detail herein, as the equations start to become pages long, The above propagation analysis can be directly, and with Maple easily, applied to the case of verged cameras with errors modeled in one or both cameras. Again, they result in finite optimal parameters, with in general smaller optimal baselines.

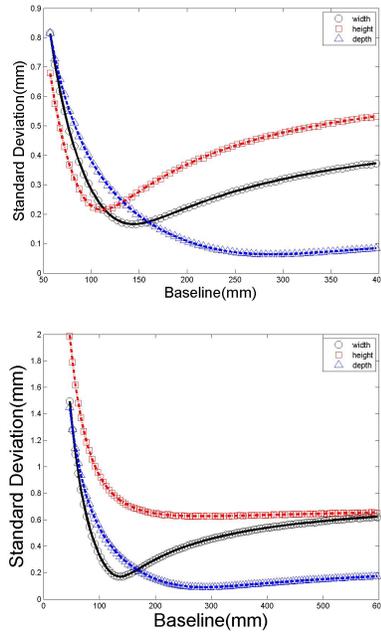


Figure 3. Error in world coordinates for various baselines B when $x_l = y_l = 150$ pixels. Top graph shows for single camera model, bottom is for two camera error model. Note the differences in vertical scale. Each graph shows world error in each of 3 dimensions. The optimal baseline value is obvious in the graphs.

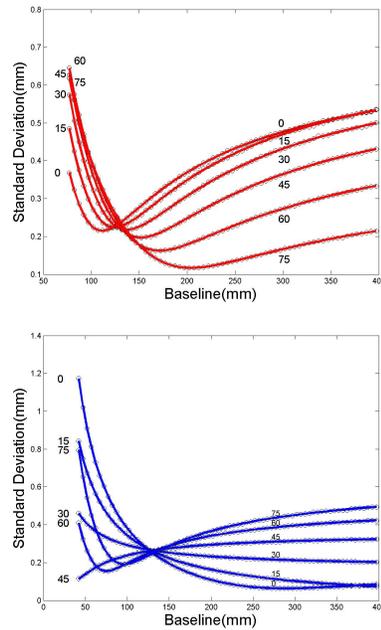


Figure 4. Error in world coordinates for various view angles. Top graph shows the height error, bottom is for the depth error. The optimal baseline value changes with the view angle.

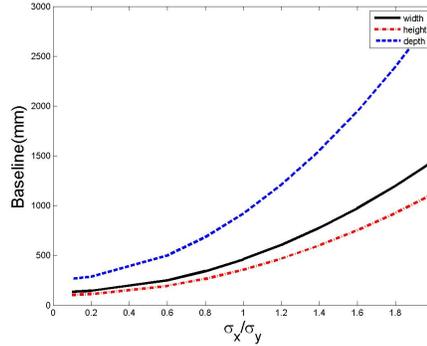


Figure 5. The relationship between $\frac{\sigma_x}{\sigma_y}$ and the optimal baseline. Smaller ratio leads to shorter optimal baseline.

6 Simulations

The simulations are carried out using Matlab. The parameters for the stereo system and the error models are defined as follow. The camera focus length is set to 17mm. The cameras have 1/3 inch CCDs, that is, one pixel is around 0.148mm. The standard derivation of the error in y direction is assume to be $\sigma_y = 1 \text{ pixel}$. The standard derivation in x direction is set to $\sigma_x = 0.2\sigma_y$ unless otherwise specified. The simulation is done on one point with $x_l = 150 \text{ pixel}$, $y_l = 150 \text{ pixel}$ and $Z = 100 \text{ mm}$. In the simulation, the sample size for a point is 50000.

Figure 3 shows the error in width, height and depth for both the single error model and the two camera error model. The simulated results as shown in markers match the theoretical results as shown in lines well. The optimal baselines for three directions are obvious in the graphs. Figure 4 shows the error in height and depth for the single error model. The graph clearly shows that the reconstruction error and the optimal baseline are affected by the view angle. This result suggests that a carefully selected view angle may reduce the reconstruction error or the optimal baseline.

Figure 5 shows the relationship between the ratio $\frac{\sigma_x}{\sigma_y}$ and the optimal baseline in three directions for the single error model. The optimal baseline is affected by the ratio of variances instead of the variances. Smaller ratio leads to shorter optimal baseline. The two camera error model has similar results, which are not shown.

Figure 6 shows how the verge angle affects the reconstruction error and the optimal baseline in the single error model. The relationship between the verge angel φ and the rotation angle γ is shown in Figure 1. The distance Z has an influence on the shape of these curves. However, these curves share some common characteristics. From the graph, we observe the optimal baselines for three directions decrease, gradually converge to same values and increase again as the verge angle increases. The error quickly decreases to a very low level, and may increase again as the verge angle increases. From the simulation results, we can see for a properly selected verge angle, a short optimal baseline and a very low error level can be obtained.

Figure 7 shows how the verge angle affects the optimal baseline and the reconstruction error in the case of two camera error model. Unlike the single error model, the optimal baselines do not converge to same values. However, they decrease to some minimal values, then increase again, respectively. As shown in the figure, the influence of the verge angle on the reconstruction error is different in three directions. The influence of distance Z on the shape of the curve is not significant. The graph obtained when $Z = 1000 \text{ mm}$ (not shown) is almost same to the one obtained when $Z = 100 \text{ mm}$ except the

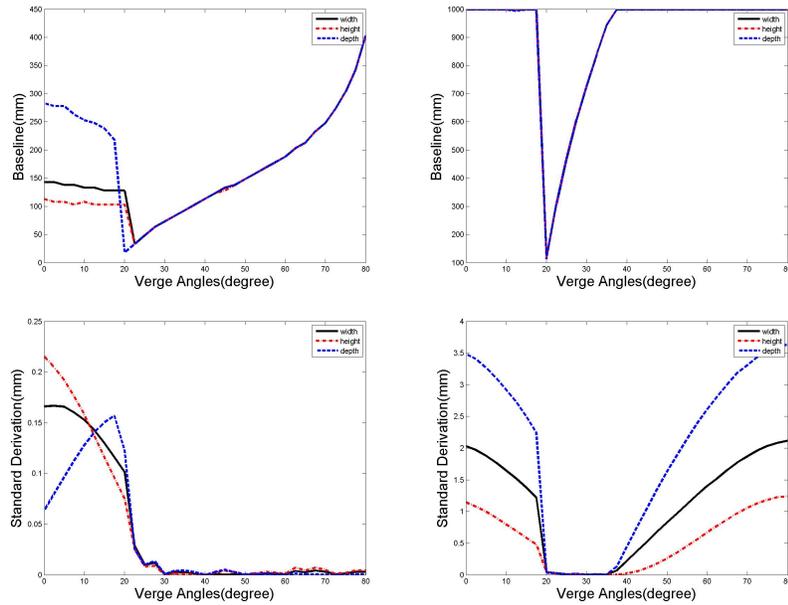


Figure 6. The verge angle affects the optimal baseline and the reconstruction error in the single error model. $Z = 100\text{mm}$ for the top two graphs column and $Z = 1000\text{mm}$ for the bottom two. In both cases, a small optimal baseline and a low error level can be obtained.

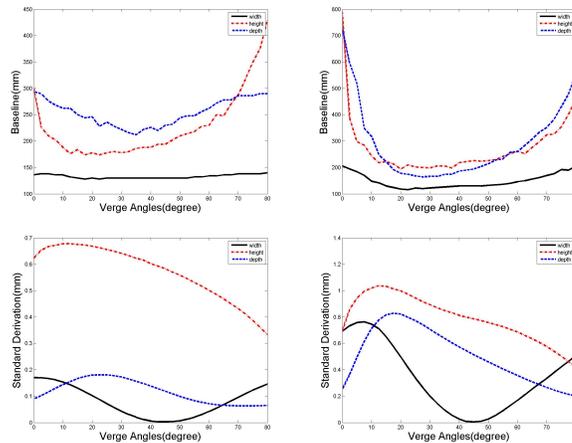


Figure 7. The verge angle affects the optimal baseline and the reconstruction error in the two camera error model. $\frac{\sigma_x}{\sigma_y} = 0.2$ for the left column; $\frac{\sigma_x}{\sigma_y} = 1$ for the right column.

scale is different. The ratio $\frac{\sigma_x}{\sigma_y}$ does affect the shape of the curves, however, the basic curve trend is similar.

7 Conclusion

This work started off to define a more realistic stereo error model for use in stereo measurement and Kalman-filtering of motion-stereo data, and to optimize parameters with respect to world coordinate measurements rather than simply depth measurements. Due to that, we defined a more general error

model, using the realistic assumptions from noisy ray intersection computations as they are used in stereo. The resulting models, and process for generating new ones, are directly useful for system design and for data fusion.

For stereo systems, especially verged systems, we show that the baseline that optimizes the localization error is finite. It is a complex function of distance to target, expected error and viewing angles, but is in general finite and for verged systems, it is reasonably small. For parallel cameras the baseline is still finite, though larger. The results also provide a new model for errors and optimal viewing angles for measuring height of “obstacles” in robotic navigation and objects in remote sensing.

References

- [1] N. Alvertos. Resolution limitations and error analysis for stereo camera models. In *IEEE Conf. Southeastcon '88*, pages 220–224, 1988. 3
- [2] A. Benetazzo. Practical use of 3-d image analysis for water wave measurement. In *the 30th Italian Conf. on Hydraulic Engineering*, 2006. 3
- [3] S. Blostein and T. Huang. Error analysis on stereo determination of 3-d point positions. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(6):752–765, 1987. 3
- [4] C. Chang, S. Chatterjee, and P. R. Kube. A quantization error analysis for convergent stereo. In *IEEE Int. Conf. on Image Processing*, pages 735–739, 1994. 3
- [5] J. Denzler, M. Zobel, and H. Niemann. Information theoretic focal length selection for real-time active 3d object tracking. In *International conf. on Computer Vision*, pages 400–407, 2003. 3
- [6] R. Hartley and A. Zisserman. Multiple view geometry in computer vision, 2003. Cambridge University Press. 6
- [7] R. I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997. 5
- [8] B. Kamgar-Parsi and B. Kamgar-Parsi. Evaluation of quantization error in computer vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(9):929–940, 1989. 3
- [9] L. Matthies and S. A. Shafer. Error modeling in stereo navigation. *IEEE Journal of Robotics and Automation*, AR-3(3):239–248, 1987. 3
- [10] J. Mulligan, V. Isler, and K. Daniilidis. Performance evaluation of stereo for tele-presence. In *International Conf. on Computer Vision*, pages II: 558–565, 2001. 3
- [11] J. S. Reed. Error propagation in two-sensor 3d position estimation. *Optical Engineering*, 40(4), April 2001. 3
- [12] A. H. Rivera-Rios, F.-L. Shih, and M. Marefat. Stereo camera pose determination with error reduction and tolerance satisfaction for dimensional measurements. In *Proceedings of the 2005 IEEE Int. Conf. on Robotics and Automation*, pages 423–428, 2005. 3
- [13] J. J. Rodriguez and J. Aggarwal. Stochastic analysis of stereo quantization error. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(5):467–470, 1990. 3
- [14] E. Trucco and A. Verri. Introductory techniques for 3-d computer vision, 1998. Prentice Hall. 3, 5
- [15] S. Wenhardt, J. Denzler, and H. Niemann. On minimizing errors in 3d reconstruction for stereo camera systems. *Pattern Recognition and Image Analysis: Advanced in Mathematical Theory and Applications*, 15(1):468–471, 2005. 3
- [16] W. Zhao and N. Nandhakumar. Effects of camera alignment errors on stereoscopic depth estimates. *Pattern Recognition*, 29(12):2115–2126, 1996. 3