Optimal Pixel Aspect Ratio for Enhanced 3D TV Visualization*

Hossein Azari[1], Irene Cheng[1], Kostas Daniilidis[2] and Anup Basu[1]

[1]Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada

[2]Dept. of Computer and Information Science, University of Pennsylvania, Philadelphia, USA

EMAIL: hazari@ualberta.ca, locheng@ualberta.ca, kostas@cis.upenn.edu and basu@ualberta.ca

Phone: 780-492-3330, FAX: 780-492-1071

* Parts of this work have been presented at the IEEE 3D TV Conference, Istanbul, Turkey, 2008.

ABSTRACT

In multiview 3D TV, a pair of corresponding pixels in adjacent 2D views contributes to the reconstruction of voxels (3D pixels) in the 3D scene. We analyze this reconstruction process and determine the optimal pixel aspect ratio based on which the estimated object position can be improved given specific imaging or viewing configurations and constraints. By applying mathematical modeling, we deduce the optimal solutions for two general stereo configurations: parallel and with vergence. We theoretically show that a non-uniform finer horizontal resolution, compared to the usual uniform pixel distribution, in general provides a better 3D visual experience for both configurations. The optimal value may vary depending on different configuration parameter values. We validate our theoretical results by conducting subjective studies using a set of simulated non-uniformly discretized red-blue stereo pairs and show that human observers indeed have a better 3D viewing experience with an optimized vs. a non-optimized representation of 3D-models.

Keywords: Aspect ratio, vergence, 3D TV, perceptual quality, sense of depth, stereo display

1. INTRODUCTION

Over the past few years, 3D TV technologies have received increasing public attention with wide applications in IMAX and entertainment centres. Research in this area covers issues relating to 3D video capture, transmission, and display [1],[2],[3],[4],[5]. Extensive research has also been conducted on designing 3D displays including studies on display screen aspect ratio [6]. These studies show that screen dimension with width:height ratios between 5:3 and 6:3 are more visually pleasant to the viewers [7]. In this regard, wider displays provide a better sensation of depth even with 2D images. This observations have led to designing wider screens for HDTV displays compared to previous generation TVs [6]. Different from the screen aspect ratio, a square pixel component with an aspect ratio of 1:1 is usually adopted by 2D digital media devices, although research suggested that non-uniform pixel aspect ratios can create better visual experience [9]. Despite this observation, there is not much work addressing the problem of determining an appropriate aspect ratio of pixels with respect to the parameters affecting the 3D picture formation process on (auto)stereoscopic displays. By default, similar to conventional 2D displays, square pixels have been adopted to compose different views of 3D displays; whereas, such an arbitrary decision may not be consistent with the underlying depth perception inherent in the human visual system.

Differing from conventional 2D displays, the output of a stereoscopic 3D display is constructed based on two or more different views of the same scene. That is, the 3D points of the reconstructed virtual 3D scene are computed based on the corresponding pixels of adjacent 2D views. Since discrete pixels are involved in 3D point reconstruction, *discretization error* [8], [9], [10] is expected in the 3D point estimation process. The magnitude of this error depends on the inter-pixel distance in the 2D views and the viewing distance. For 3D displays, discretization is closely related to the 3D resolution or *stereoscopic resolution* which is defined in terms of the precision in locating 3D points within the comfortable viewing range of the 3D display [11]. The magnitude of this error depends on the inter-pixel distance in the 2D views and is proportional to the viewing distance. In other words, the discretization error, or equivalently

stereoscopic resolution on each 3D coordinate component (horizontal, vertical, and depth resolution), is directly related to the size of pixels; and depending on the stereo configuration parameters the relative width and height of a pixel can affect the amount of these errors differently. Since the horizontal and vertical discretizations contribute separately to the 3D point estimation error, given a total display resolution the challenge is to compute an optimal horizontal *vs.* vertical discretization in a unit area (pixel aspect ratio) that can minimize the error, in order to maximize the 3D-display viewing quality without changing the total resolution. This optimization issue is important especially for multi-view 3D displays, which provide several different 3D views (currently 6-9 views) and require several 2D images as input (7-10 images for 6-9 3D views).

In this paper, we study the abovementioned optimization problem for different stereo configurations. We extend our earlier research [8, 9, 10, 12, 13] and conduct subjective studies using two stereo configurations: parallel and with vergence. We show that the aspect ratio of pixels for optimizing 3D reconstruction can vary significantly in different configurations. Moreover, based on the physical pixel grouping technique, we present methods that simulate 3D viewing with different pixel aspect ratios on conventional 2D displays. These methods enable us to conduct human observer evaluations on both simple and complex 3D objects. Our evaluation results, which are consistent with our theoretical derivations, show that given a constant total resolution it is possible to improve the 3D visual experience by choosing a finer horizontal resolution relative to the vertical display resolution. The optimal value may vary depending on the parameter settings. In our implementation, we follow the usual viewing conditions to specify these parameters.

The remainder of this paper is organized as follows: Section 2 formally defines the problem and notation used in this paper. Section 3 discusses optimal discretization for a single 3D point with and without vergence. Results obtained in Section 3 are extended in Section 4 to determine the optimal aspect ratio within the viewing volume inside a camera's field of view. Experimental results and subjective studies are described in detail in Section 5. Finally, concluding remarks and suggestions for future work are outlined in Section 6.

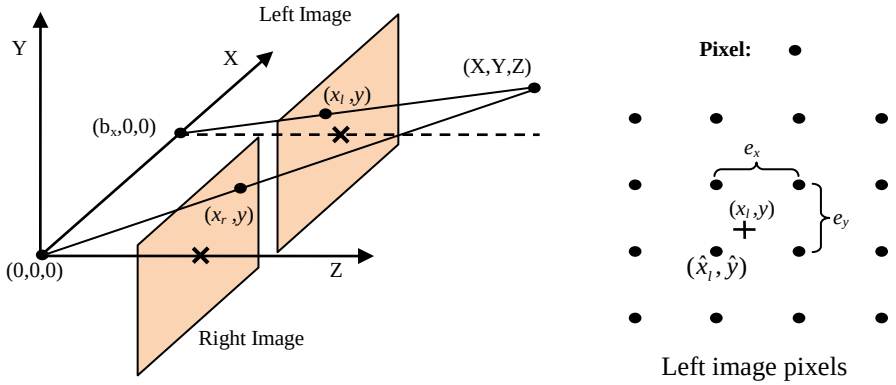

Figure 1: Left: parallel-camera stereo projection with baseline b_x. Right: discretization error.

2. PROBLEM DEFINITION AND NOTATION

We use the same notation and definitions as in our earlier publications (see Figures 1 and 2):

- f: focal length of both cameras.
- (X,Y,Z): a 3D point, and $(\hat{X},\hat{Y},\hat{Z})$: estimated 3D position.
- (x_l, y_l)*:* projection of a 3D point on the left image.
- (x_r, y_r)*:* projection of a 3D point on the right image.
- e_x (e_y): distance between two neighboring pixels (picture elements or viewing elements) in the x (y) direction.
- b_x*:* baseline of a stereo imaging setup.
- *R:* total resolution, *i.e.*, the total number of pixels or viewing elements being used for stereo capture or viewing.

Using triangulation rules, the image projections of a 3D point on the left and right image screens in parallel geometry (Figure 1) are given by:

$$x_r = \frac{fX}{Z}, \quad x_l = \frac{f(X - b_x)}{Z}, \quad y = y_l = y_r = \frac{fY}{Z} \qquad (1)$$

Thus, a 3D point location (X,Y,Z) can be computed as:

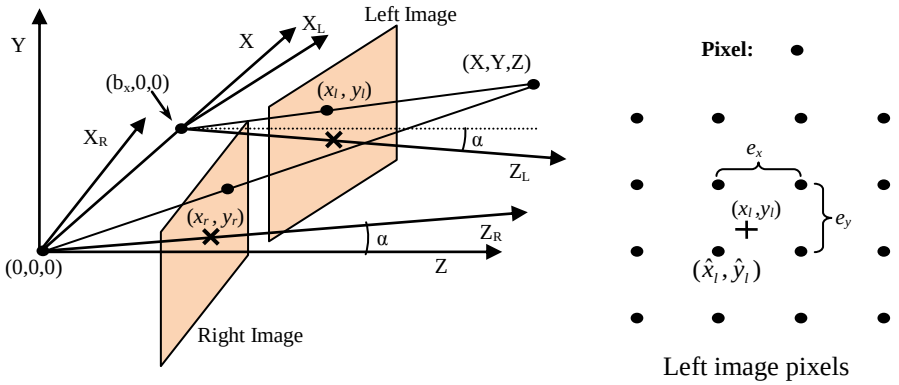

Figure 2: Left: Stereo camera projection with vergence and baseline b_x. Right: discretization error.

$$X = Z\frac{x_r}{f}, \quad Y = Z\frac{y}{f}, \quad Z = \frac{fb_x}{x_r - x_l} \tag{2}$$

In practice, since the image projections are discretized to the nearest pixel, location of a 3D-point is determined using points $(\hat{x}_r, \hat{y})$ and $(\hat{x}_l, \hat{y})$. Therefore, the location of a 3D-point (*X*,*Y*,*Z*) is estimated as:

$$\hat{X} = \hat{Z}\frac{\hat{x}_r}{f}, \quad \hat{Y} = \hat{Z}\frac{\hat{y}}{f}, \quad \hat{Z} = \frac{fb_x}{\hat{x}_r - \hat{x}_l} \tag{3}$$

For cameras with vergence (Figure 2), the location of a 3D point (X, Y, Z) can be computed using following formulae (see [10] for details):

$$\begin{aligned} X &= \frac{Z(f\sin\alpha + x_r\cos\alpha)}{(f\cos\alpha - x_r\sin\alpha)} \\ Y &= \frac{y_r(X\sin\alpha + Z\cos\alpha)}{f} \\ Z &= \frac{b_x(f\cos\alpha + x_l\sin\alpha)(f\cos\alpha - x_r\sin\alpha)}{(f\cos\alpha + x_l\sin\alpha)(f\sin\alpha + x_r\cos\alpha) + (f\sin\alpha - x_l\cos\alpha)(f\cos\alpha - x_r\sin\alpha)} \end{aligned} \tag{4}$$

Again, because of discretization error $(\hat{x}_r, \hat{y}_r)$ and $(\hat{x}_l, \hat{y}_l)$ will be used in the computations resulting in an estimation error with respect to the original 3D point.

Figure 3 shows the error pattern for two camera configurations (left: parallel geometry and middle: vergence geometry). Figure 3 (right) illustrates a 3D-view of an error generated by two corresponding pixels. The quadrilateral volume corresponds to a voxel (3D pixel). Obviously, reducing the pixel size and thus increasing the total resolution will reduce the voxel size as well as the 3D point estimation error.

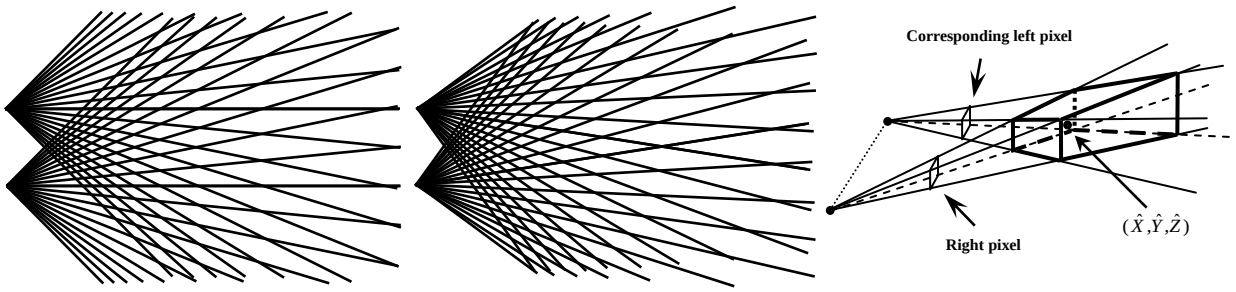

Figure 3: Discretization error patterns. Left: parallel-axes stereo. Middle: stereo with vergence. Right: 3D-view of error pattern for two corresponding pixels.

However, assuming a constant total resolution *R*, it is reasonable to look for an optimal pixel aspect ratio which minimizes the size of voxels or equivalently the discretization error inside the viewing region. Such optimization can improve the output image quality of a typical 3D TV system while the size of the input data is unchanged. The optimal solution can also be used in solving the stereo based depth estimation problem, which has been a research focus in the past decades [14], [15].

3. OPTIMAL DISCRETIZATION FOR A SINGLE 3D POINT ESTIMATION

As mentioned in the previous section, the discrete nature of a display device (or viewing screen) leads to rounding. The rounding error is bounded by half a pixel. Thus, in the worst case:

$$\begin{aligned} \hat{x}_r &= x_r \pm (e_x/2) \\ \hat{x}_l &= x_l \pm (e_x/2) \\ \hat{y} = \hat{y}_r &= \hat{y}_l = y \pm (e_y/2) \end{aligned} \qquad (5)$$

Considering the stereo setup with parallel configuration (shown in Figure 1), and the worst-case error in 3D estimation, which will be our default assumption in the rest of this paper, we obtain $\hat{Z}$ from Equations 3 and 5 (as explained in [8]):

$$\hat{Z} = \frac{fb_x}{(x_r - x_l) \pm e_x} = Z\left(1 \pm e_x \frac{Z}{fb_x}\right)^{-1} \qquad (6)$$

The Taylor series expansion of Equation 6 about $Z = 0$ is given by:

$$\hat{Z} = Z\left(1 \pm e_x \frac{Z}{fb_x}\right)^{-1} = Z\left(1 \pm \frac{e_x}{fb_x} Z + \left(\frac{e_x}{fb_x}\right)^2 Z^2 \pm \left(\frac{e_x}{fb_x}\right)^3 Z^3 + \cdots\right)$$

In practice f and b_x values are much greater than values taken by e_x. Therefore, we can assume that the higher order terms in the above expansion are negligible. As result, $\hat{Z}$ can be properly approximated as:

$$\hat{Z} \cong Z\left(1 \pm \frac{e_x Z}{f b_x}\right) \tag{7}$$

Bounds on error in estimating X can be obtained as follow:

$$\begin{aligned} \hat{X} &= \frac{\hat{Z}}{f}\hat{x}_r \cong \frac{Z}{f}\left(1 \pm \frac{e_x Z}{f b_x}\right)\left(x_r \pm \frac{e_x}{2}\right) \\ &= X\left(1 \pm \frac{e_x Z}{f b_x} \pm \frac{e_x}{2x_r} \pm \frac{e_x^2 Z}{2 f b_x x_r}\right) \end{aligned} \tag{8}$$

Similarly, for Y:

$$\hat{Y} = Y\left(1 \pm \frac{e_x Z}{f b_x} \pm \frac{e_y}{2y} \pm \frac{e_x e_y Z}{2 f b_x y}\right)$$

or,

$$\left|\frac{\hat{Y} - Y}{Y}\right| \le f_p(e_x, e_y) = \left\{\frac{e_x Z}{f b_x} + \frac{e_y}{2|y|} + \frac{e_x e_y Z}{2 f |y| b_x}\right\} \tag{9}$$

Considering a unit viewing or image capture area:

$$\left(\frac{1}{e_x}\right)\left(\frac{1}{e_y}\right) = R \text{ or } e_y = \frac{1}{e_x R} \tag{10}$$

From Equations 9 and 10 we have:

$$f_p(e_x) = \left\{\left(\frac{Z}{f b_x}\right)e_x + \left(\frac{1}{2R|y|}\right)\frac{1}{e_x} + \frac{Z}{2 f |y| R b_x}\right\} \tag{11}$$

From Equation 8 it is obvious that the best solution which minimizes the estimation error of X or Z is to have e_x as small as possible. Unfortunately, it is also the worst possible choice for estimating the Y component. As a compromise, Equation 11, which is the upper bound of the relative estimation error in Y as a function of e_x constrained by the total resolution R through Equation 10, can be used to find the best vertical *vs.* horizontal resolution tradeoff for estimating Y. The advantage of using relative versus absolute estimation error in the optimization process is to offset the monotonic increase of absolute error in the vertical direction from the center of the image plane, and allows the discretization optimization problem

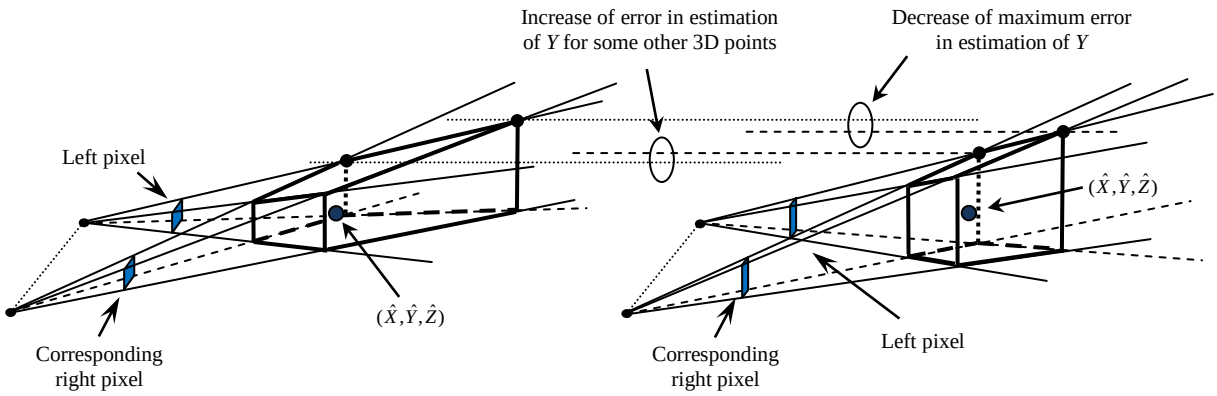

Figure 4: Comparing discretization error pattern and changes of error in estimation of Y for regular square pixels, $e_x = e_y$, (left) *vs*. vertically rectangular pixels, $e_x < e_y$, with optimal pixel aspect ratio (right).

to be studied independent of the pixel location on the image plane. Our analysis motivates the derivations of a Mathematical model given below.

<u>Result 1</u> – Optimal discretization for a single 3D point and parallel configuration: Considering the parallel stereo configuration in Figure 1, the optimal discretization for estimating Y for a single 3D point is:

$$e_x = \frac{1}{\sqrt{R}}\sqrt{\frac{fb_x}{2|y|Z}}, \quad e_y = \frac{1}{\sqrt{R}}\sqrt{\frac{2|y|Z}{fb_x}} \tag{12}$$

Derivation: The results are obtained by equating the derivative of $f_p(e_x)$ in Equation 11 to zero to minimize the error:

$$f_p'(e_x) = \left\{\frac{Z}{fb_x} - \frac{1}{e_x^2}\left(\frac{1}{2R|y|}\right)\right\} = 0 \tag{13}$$

Solving Equation 13 in terms of e_x gives the value of e_x in Equation 12. To show that the function f_p achieves the minimum value at this point we need to perform the second derivative test. From Equation 13 the second derivative of f_p in terms of e_x is calculated as:

$$f_p''(e_x) = \left(\frac{1}{R|y|}\right)\frac{1}{e_x^3}$$

On the other hand, from Equation 10 we can see that e_x changes over the open interval $(0,\infty)$ which means that e_x always takes positive values. Since the other two parameters involved in the above equation are also positive, the second derivative of f_p is always positive over the domain of possible values for e_x. As a result, f_p is strongly convex over this domain and takes its minimum value at e_x given by Equation 12. The value of e_y in Equation 12 can be obtained from e_x by applying Equation 10.

The solution that is obtained in this way is not only optimal in terms of the maximum relative estimation error on *Y*, but often yields better estimates for the *X* and *Z* components as well. From 12 the ratio of e_x to e_y can be calculated as:

$$\frac{e_x}{e_y} = \frac{fb_x}{2|y|Z} = \frac{(x_r - x_l)}{2|y|} \tag{14}$$

In Equation 14, $(x_r - x_l)$ is the amount of the disparity between the corresponding points in the left and right image planes. It should be clarified that x_r and x_l in above-mentioned equations are measured with respect to the corresponding camera coordinate system. As result, and as can be also deducted from Equation 2, $(x_r - x_l)$ is always greater than or equal to zero.

The amount of disparity $(x_r - x_l)$ in most (practical) cases is smaller than $2|y|$. This means that in practice the ratio in Equation 14 is usually less than one or equivalently e_x is often smaller than e_y. As a result, the optimal solution in Result 1 often improves the estimates of the *X* and *Z* components. However, when $e_x < e_y$ the optimal solution changes the error pattern so that the *Y* estimation error may be increased for some 3D points estimated by $(\hat{X},\hat{Y},\hat{Z})$. This phenomenon is pictorially illustrated in Figure 4 for a typical error pattern. As shown in Figure 4, the optimal pixel aspect ratio less than one (*i.e.* $e_x < e_y$) results in improvement in estimation of *X* and *Z* plus reduction of the maximum error in estimation of *Y*, which in turn means improvement of estimation of *Y* for some 3D points. But generally this is achieved in the cost of degradation in the estimation of *Y* for some other 3D points.

Now we focus on the stereo setup with vergence, shown in Figure 2. We start with some auxiliary derivations as follow.

Auxiliary Result 1 (AR1): Assuming a small vergence angle α, the estimated depth $\hat{Z}$ of a 3D point for the vergence configuration (Figure 2) can be related to the true depth Z by:

$$\hat{Z} \cong Z\left(1 \pm \frac{e_x Z \cos\alpha}{2b_x(f\cos\alpha - x_r \sin\alpha)} \pm \frac{e_x Z \cos\alpha}{2b_x(f\cos\alpha + x_l \sin\alpha)}\right)^{-1} \tag{15}$$

Derivation: From Equation 4, Z is estimated as:

$$\hat{Z} = \frac{b_x(f\cos\alpha + \hat{x}_l \sin\alpha)(f\cos\alpha - \hat{x}_r \sin\alpha)}{(f\cos\alpha + \hat{x}_l \sin\alpha)(f\sin\alpha + \hat{x}_r \cos\alpha) + (f\sin\alpha - \hat{x}_l \cos\alpha)(f\cos\alpha - \hat{x}_r \sin\alpha)}$$

Using the same approach as the case without vergence:

$$\begin{aligned}\hat{Z} &= \left(\frac{f\sin\alpha + \hat{x}_r \cos\alpha}{b_x(f\cos\alpha - \hat{x}_r \sin\alpha)} + \frac{f\sin\alpha - \hat{x}_l \cos\alpha}{b_x(f\cos\alpha + \hat{x}_l \sin\alpha)}\right)^{-1} \\ &= \left(\frac{f\sin\alpha + (x_r \pm (e_x/2))\cos\alpha}{b_x(f\cos\alpha - (x_r \pm (e_x/2))\sin\alpha)} + \frac{f\sin\alpha - (x_l \pm (e_x/2))\cos\alpha)}{b_x(f\cos\alpha + (x_l \pm (e_x/2))\sin\alpha)}\right)^{-1}\end{aligned}$$

Assuming that $(e_x/2)\sin\alpha \cong 0$ for a small vergence angle α:

$$\begin{aligned}\hat{Z} &\cong \left(\frac{2f\sin\alpha + 2x_r \cos\alpha \pm e_x \cos\alpha}{2b_x(f\cos\alpha - x_r \sin\alpha)} + \frac{2f\sin\alpha - 2x_l \cos\alpha \pm e_x \cos\alpha}{2b_x(f\cos\alpha + x_l \sin\alpha)}\right)^{-1} \\ &= \left(Z^{-1} \pm \frac{e_x \cos\alpha}{2b_x(f\cos\alpha - x_r \sin\alpha)} \pm \frac{e_x \cos\alpha}{2b_x(f\cos\alpha + x_l \sin\alpha)}\right)^{-1} \\ &= Z\left(1 \pm \frac{e_x Z \cos\alpha}{2b_x(f\cos\alpha - x_r \sin\alpha)} \pm \frac{e_x Z \cos\alpha}{2b_x(f\cos\alpha + x_l \sin\alpha)}\right)^{-1}\end{aligned}$$

AR2: Assuming a small vergence angle α, for the vergence configuration the estimated depth of a 3D point is bounded by:

$$\hat{Z} \le Z\left(1 + \frac{e_x Z}{f b_x} K\right) \text{ where } K = \left(1 + \frac{x_{\max}}{f}\tan\alpha\right) \tag{16}$$

Derivation: Similar to the parallel configuration case (Equation 6) the Taylor expansion of Equation 15 about $Z = 0$ can be calculated. Assuming a small vergence angle α the higher order terms in the Taylor series expansion can be ignored and $\hat{Z}$ can be properly approximated as:

$$\hat{Z} \cong Z\left(1 \pm \frac{e_x Z}{2b_x} \frac{\cos\alpha}{(f\cos\alpha - x_r \sin\alpha)} \pm \frac{e_x Z}{2b_x} \frac{\cos\alpha}{(f\cos\alpha + x_l \sin\alpha)}\right) \tag{17}$$

Expanding the second term of Equation 17 and ignoring higher order terms, we have:

$$\frac{\cos\alpha}{(f\cos\alpha - x_r \sin\alpha)} = \frac{1}{f}\left(1 - \frac{x_r \sin\alpha}{f\cos\alpha}\right)^{-1} \cong \frac{1}{f}\left(1 + \frac{x_r}{f}\tan\alpha\right) \tag{18}$$

Similarly, for the third term:

$$\frac{\cos\alpha}{(f\cos\alpha + x_l \sin\alpha)} = \frac{1}{f}\left(1 + \frac{x_l \sin\alpha}{f\cos\alpha}\right)^{-1} \cong \frac{1}{f}\left(1 - \frac{x_l}{f}\tan\alpha\right) \tag{19}$$

From Equations 17, 18, and 19:

$$\begin{aligned}
\hat{Z} &\cong Z\left(1 \pm \frac{e_x Z}{2b_x}\frac{1}{f}\left(1 + \frac{x_r}{f}\tan\alpha\right) \pm \frac{e_x Z}{2b_x}\frac{1}{f}\left(1 - \frac{x_l}{f}\tan\alpha\right)\right) \\
&= Z\left(1 \pm \frac{e_x Z}{2fb_x}\left(2 + \frac{(x_r - x_l)}{f}\tan\alpha\right)\right) \\
&\le Z\left(1 + \frac{e_x Z}{fb_x}\left(1 + \frac{x_{\max}}{f}\tan\alpha\right)\right)
\end{aligned} \tag{20}$$

where $x_{\max} = \max\left(\frac{x_r - x_l}{2}\right)$ is half of the maximum possible horizontal disparity, which is in fact equal to the maximum possible value of *x* or equivalently half of the image plane (or display) width.

AR3: Assuming a small vergence angle α:

$$\left|\frac{\hat{Y} - Y}{Y}\right| \le f_v(e_x) = \left\{\left(\frac{ZK}{fb_x}\right)e_x + \left(\frac{1}{2R|y_r|}\right)\frac{1}{e_x} + \frac{ZK}{2f|y_r|Rb_x}\right\} \tag{21}$$

Derivation: For a small vergence angle α, the parallel configuration formulae can be used to approximate the Y component. Therefore:

$$\begin{aligned}
\hat{Y} &\cong \frac{\hat{Z}\hat{y}_r}{f} \le \frac{Z}{f}\left(1 + \frac{e_x Z}{fb_x}K\right)\left(y_r \pm \frac{e_y}{2}\right) \\
&\cong Y\left(1 + \frac{e_x ZK}{fb_x} \pm \frac{e_y}{2y_r} \pm \frac{e_x e_y ZK}{2fb_x y_l}\right)
\end{aligned} \tag{22}$$

Considering a total resolution of *R*, Equation 21 can be obtained from Equations 10 and 22.

Summarizing the abovementioned derivations, the following result can be stated for a stereo configuration with vergence.

<u>Result 2</u> – Optimal discretization for a single 3D point and stereo setup with vergence: Considering a stereo system with a small vergence angle α (Figure 2), the optimal discretization for a single 3D point can be obtained as:

$$e_x = \frac{1}{\sqrt{R}}\sqrt{\frac{fb_x}{2|y_r|ZK}}, \quad e_y = \frac{1}{\sqrt{R}}\sqrt{\frac{2|y_r|ZK}{fb_x}} \tag{23}$$

Derivation: Similar to the parallel configuration, from Equation 21:

$$f_v'(e_x) = \left(\frac{ZK}{fb_x}\right) - \left(\frac{1}{2R|y_r|}\right)\frac{1}{e_x^2} \tag{24}$$

The results can be obtained by equating 24 to zero, and solving it in terms of e_x. The same argument as Result 1 can be applied to show that f_v indeed takes its minimum value at this point, considering the domain of permissible values for e_x.

4. OPTIMIZING WITHIN A VIEWING VOLUME

The derivations in the last section consider optimization with respect to a single 3D point. Instead, we need to consider a region in 3D defined by a set of constraints on the range of values in depth, height and the stereo Field-Of-View (FOV). Obviously, the particular values obtained for e_x (or e_y) in the previous section will not be optimum for all points satisfying these constraints. A solution to this problem could be calculating the integral of the error function f_p or f_v over the entire specified region. However, this approach leads to logarithmic terms which complicate further steps in the optimization process. Instead, we consider minimizing an appropriate error metric subject to the restrictions imposed.

For the parallel camera configuration (Figure 1), we use Equation 13 to obtain such an error metric. In this case, from Equation 13 for the optimal discretization e_x for a single 3D point we have:

$$\frac{Z}{fb_x} = \frac{1}{e_x^2}\left(\frac{1}{2R|y|}\right)$$
$$\frac{fb_x}{Z} = |y|2e_x^2 R$$

or equivalently,

$$\frac{fb_x}{Z} - |y|2e_x^2 R = 0 \tag{25}$$

In other words, for a single 3D point, the derivative of f_p or equivalently the right side of Equation 25 is zero at the optimal discretization and non-zero in its neighbourhood. This, in fact, gives a distance criterion that can be used as an error metric for optimizing the discretization error over the entire region. We define the square of this distance as the error metric, as follow:

$$E_p = \left(\frac{fb_x}{Z} - |y|2e_x^2 R\right)^2 \tag{26}$$

On the other hand, for parallel configuration the range of values of X varies with changes of depth Z and values of the cameras' field of view by the following simple relationship (see Figure 5-left):

$$X \in \left\{X_0 \pm \frac{\Delta X}{2}\frac{(Z - Z_0)}{(Z_{\max} - Z_0)}\right\} \tag{27}$$

Thus, we need to optimize the following function w.r.t. e_x :

$$F_p(e_x) = \int_{Z_{\min}}^{Z_{\max}} \int_{-y_{\max}}^{y_{\max}} \int_{X_0 - \frac{\Delta X}{2}\frac{(Z-Z_0)}{(Z_{\max}-Z_0)}}^{X_0 + \frac{\Delta X}{2}\frac{(Z-Z_0)}{(Z_{\max}-Z_0)}} \left(\frac{fb_x}{Z} - |y|2e_x^2 R\right)^2 dX dy dZ \tag{28}$$

Calculating the integral in 28 and equating its derivative to zero gives us the following result:

<u>Result 3</u> – Optimal discretization for a viewing volume and parallel configuration: Considering the abovementioned notation and the parallel stereo configuration given in Figure 5-left, and assuming a reasonably large viewing volume across the Z dimension, the discretization in x for optimizing the average error in estimation of Y is given by:

$$e_x = \left(-\frac{fb_x}{2R}\frac{Z_0 I_1[Z_{\min}, Z_{\max}] - I_2[Z_{\min}, Z_{\max}]}{I_3[Z_{\min}, Z_{\max}] - Z_0 I_4[Z_{\min}, Z_{\max}]}\right)^{\frac{1}{2}} \tag{29}$$

where,

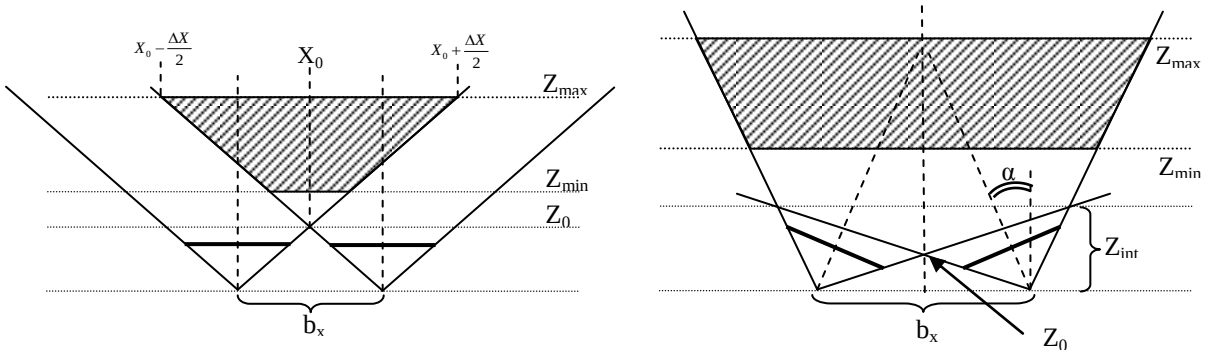

Figure 5: Viewing region without (left) and with vergence (right).

$$I_1[Z_1, Z_2] = \int_{Z_1}^{Z_2} \int_{-y_{\max}}^{y_{\max}} \frac{|y|}{Z} dydZ = y_{\max}^2 \ln\left(\frac{Z_2}{Z_1}\right)$$

$$I_2[Z_1, Z_2] = \int_{Z_1}^{Z_2} \int_{-y_{\max}}^{y_{\max}} |y| dydZ = y_{\max}^2 (Z_2 - Z_1)$$

$$I_3[Z_1, Z_2] = \int_{Z_1}^{Z_2} \int_{-y_{\max}}^{y_{\max}} Z y^2 dydZ = \frac{y_{\max}^3}{3}\left(Z_2^2 - Z_1^2\right)$$

$$I_4[Z_1, Z_2] = \int_{Z_1}^{Z_2} \int_{-y_{\max}}^{y_{\max}} y^2 dydZ = \frac{2y_{\max}^3}{3}\left(Z_2 - Z_1\right)$$

Derivation: See Appendix A.

Determining the range of values of *X* when Z varies over a specific domain is a little more complicated for the vergence configuration (see Figure 5-right). This is because between Z_0 to Z_{int} the values of *X* vary in a certain way, and for $Z > Z_{\text{int}}$ the values of *X* vary in a different way. The values of Z_0 and Z_{int} are given through the following equations.

AR4: Given a field of view θ and a vergence angle α:

$$Z_0 = \frac{b_x}{2} \tan\left(\frac{\pi}{2} - \alpha - \frac{\theta}{2}\right) \quad (30)$$

$$Z_{\text{int}} = \frac{b_x}{2} \frac{\cos(2\alpha) + \cos\theta}{\sin(2\alpha)} \quad (31)$$

Derivation: Follows from Figures 5 and 6, using simple geometric rules, such as the sine rule. Details are skipped here.

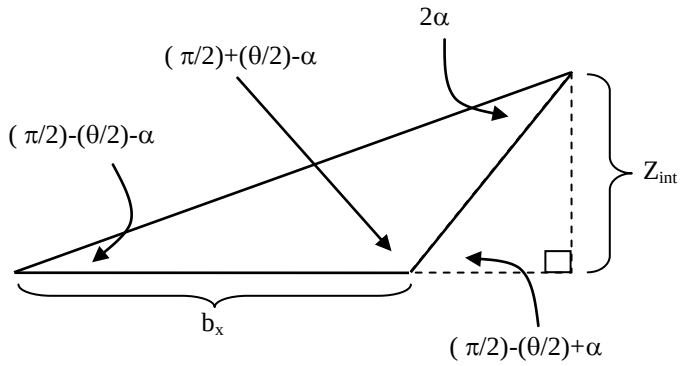

Figure 6: A close-up of Figure 5 (bottom right) illustrating its geometric constraints.

AR5: Considering the setup with vergence in Figure 5-right, the range of values of *X*, depending on depth *Z* and the field of view of the cameras, varies based on the following rules:

$$\Delta X_{\text{int}} = \frac{b_x}{2}\frac{\sin\theta}{\sin(2\alpha)} \quad \text{for} \quad Z = Z_{\text{int}} \tag{32}$$

$$\Delta X = \Delta X_{\text{int}}\frac{(Z - Z_0)}{(Z_{\text{int}} - Z_0)} \quad \text{for} \quad Z_0 < Z < Z_{\text{int}} \tag{33}$$

$$\Delta X = (\Delta X_{\text{int}} - b_x)\frac{Z}{Z_{\text{int}}} + b_x \quad \text{for} \quad Z_{\text{int}} \le Z \tag{34}$$

(It is assumed that the *X*-axis is the dashed line joining the centers' of the two cameras.)

Derivation: Follows from trigonometric rules and the earlier derivations. Details are skipped here.

Here again, we need to consider an appropriate error metric to minimize over the range of possible *X*, *Y*, and *Z* values. Similar to the parallel configuration scenario, we use the following error metric based on the derivative in Equation 24:

$$E_v = \left(\frac{fb_x}{Z} - |y_r|2RKe_x^2\right)^2 \tag{35}$$

We thus need to optimize the following function w.r.t. e_x (from this point forward, for simplicity we ignore subscript *r* in y_r):

$$F_v(e_x) = \int_Z\int_y\int_X \left(\frac{fb_x}{Z} - 2RKe_x^2|y|\right)^2 dXdydZ \tag{36}$$

Unlike the multiple integral in Equation 28, the computation of the integral above may need to be divided into two parts depending on the values of Z_{min} and Z_{max} relative to Z_{int}. If $Z_{int} \leq Z_{min}$ or $Z_{max} \leq Z_{int}$, we need to consider either Equation 33 or Equation 34 in computing this integral. However, if $Z_{min} < Z_{int} < Z_{max}$, then we must consider both Equations 33 and 34.

By calculating $F_v(e_x)$ in these three possible cases and equating corresponding derivatives to zero, we can obtain optimal discretization for all cases (see Appendix A for details). Thus, we can state the following general result.

Result 4 – Optimal discretization for a viewing volume and stereo setup with vergence: Considering the abovementioned notation and the stereo system framework in Figure 5-right, and assuming a reasonably thick viewing volume across the Z dimension, the discretization in x for optimizing the average error in estimation of Y is given by:

I) $Z_{int} \leq Z_{min}$

$$e_x = \left(\frac{fb_x}{2RK} \frac{A[Z_{min}, Z_{max}]}{B[Z_{min}, Z_{max}]} \right)^{\frac{1}{2}} \tag{37}$$

II) $Z_{max} \leq Z_{int}$

$$e_x = \left(\frac{fb_x}{2RK} \frac{C[Z_{min}, Z_{max}]}{D[Z_{min}, Z_{max}]} \right)^{\frac{1}{2}} \tag{38}$$

III) $Z_{min} < Z_{int} < Z_{max}$

$$e_x = \left(\frac{fb_x}{2RK} \frac{A[Z_{min}, Z_{int}] + C[Z_{int}, Z_{max}]}{B[Z_{min}, Z_{int}] + D[Z_{int}, Z_{max}]} \right)^{\frac{1}{2}} \tag{39}$$

where,

$$A[Z_1, Z_2] = b_x I_1[Z_1, Z_2] + \frac{(\Delta X_{int} - b_x)}{Z_{int}} I_2[Z_1, Z_2]$$

$$B[Z_1, Z_2] = b_x I_4[Z_1, Z_2] + \frac{(\Delta X_{int} - b_x)}{Z_{int}} I_3[Z_1, Z_2]$$

$$C[Z_1, Z_2] = I_2[Z_1, Z_2] - Z_0 I_1[Z_1, Z_2]$$

$$D[Z_1, Z_2] = I_3[Z_1, Z_2] - Z_0 I_4[Z_1, Z_2]$$

I_1 to I_4 are equations defined in Result 3.

Table 1: Parameters' values used for calculating optimal pixel aspect ratio.

Parameter	Value	Parameter	Value
f	17 mm	Z_{min}	From 200 mm to 600 mm
b_x	65 mm	Z_{range} (Z_{max}- Z_{min})	From 1 mm to 2000 mm
x_{max}	3.15 mm	Resolution (R)	1280x800 (40635 p/mm^2)
y_{max}	2 mm		

Derivation: See Appendix A.

5. EXPERIMENTAL RESULTS

5.1. NUMERICAL RESULTS

Table 1 shows a typical set of parameter values used for calculating optimal pixel aspect ratios (e_x/e_y) for with and without vergence stereo setups. Values of x_{max}, y_{max}, and R are set to 3.15 mm, 2 mm and 40635 pixels per mm^2, respectively. These values are obtained assuming that a CCD of size 6.3 x 4.0 mm and a resolution of 1280 x 800 (around 1 mega pixel) is used in the capturing process. Focal length and stereo baseline are respectively set to 17 and 65 mm, which are close to the human visual system parameters. Table 1 also includes the range of values used for Z_{min} and Z_{max} in our calculations. Figure 7 shows the results of our calculations for parameter values in Table 1 based on Result 3 and Result 4 and other related formulae. Figure 7-left shows how optimal aspect ratio changes with different viewing volumes at different distances in the parallel configuration. Figure 7-right presents corresponding results in the vergence configuration using three different small vergence angles, α = 1, 3, and 5 degrees, over the same depth ranges and viewing volumes. The corner of the original graph is magnified in order to highlight the effect of applying different vergence angles. It can be seen that for small vergence angles there is no significant difference either between the left and right graphs or between results obtained for different small vergence angles on the right graph. Based on this observation, we can say that at least for small vergence angles the effect of vergence in determining the optimal pixel aspect ratio is negligible. It should be noted that for a stereo baseline of 65 mm, a 5 degree vergence angle is equivalent to focusing on a point about 371 mm from cameras/viewing point, which is a good representation of a practical

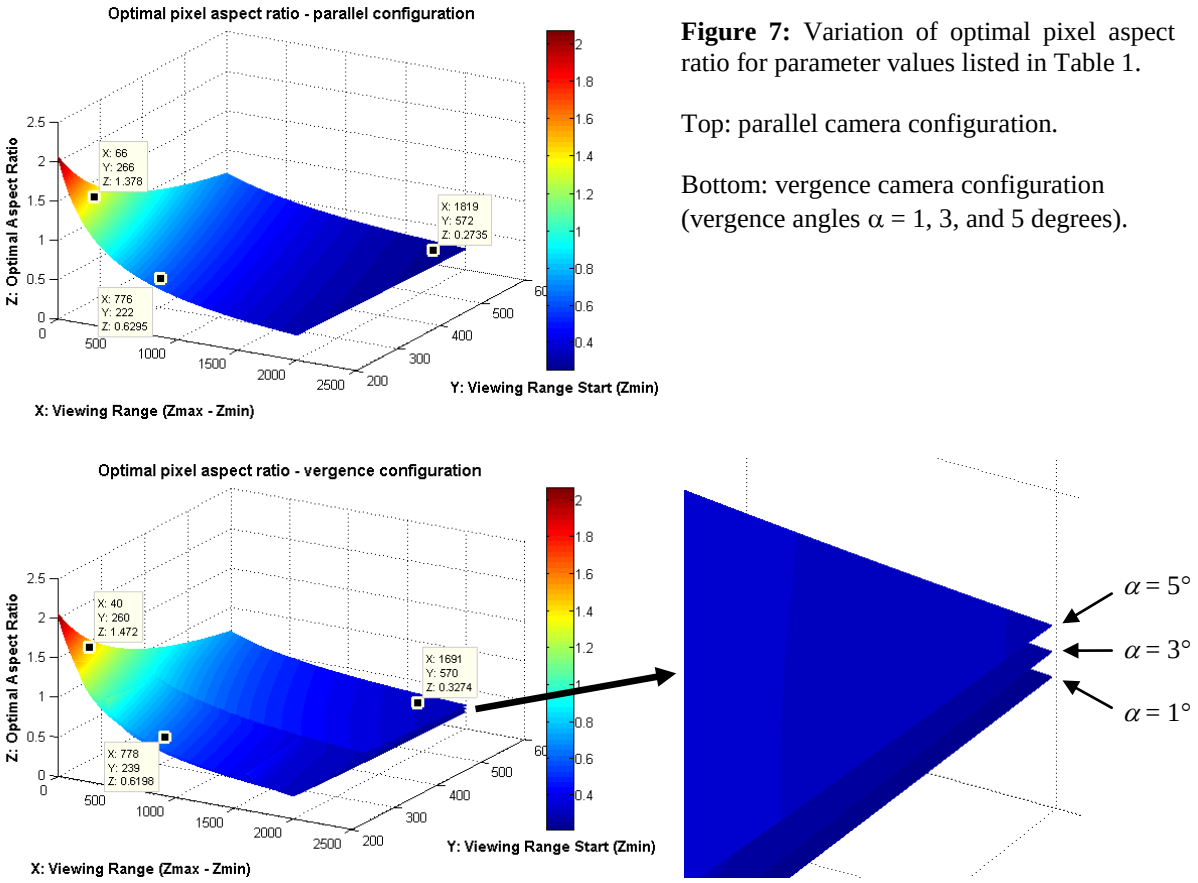

Figure 7: Variation of optimal pixel aspect ratio for parameter values listed in Table 1.

Top: parallel camera configuration.

Bottom: vergence camera configuration (vergence angles α = 1, 3, and 5 degrees).

situation.

Another observation from these graphs is how the optimal aspect ratio changes relative to the viewing range start point (Z_{min}) and length of the viewing range (Z_{max} - Z_{min}). For closer start points with smaller viewing ranges, a larger optimal ratio is obtained; but, if the start point is far enough or the viewing range

Table 2: Some optimal aspect ratios calculated for parameter values mentioned in Table 1, for both with and without vergence configurations.

Row	$\mathbf{Z_{min}}$ (mm)	$\mathbf{Z_{max}}$ (mm)	$\boldsymbol{\alpha}$ (Degree)	$\mathbf{Z_0}$ (mm)	$\mathbf{Z_{int}}$ (mm)	$\boldsymbol{K}$	$\boldsymbol{e_x/e_y}$
1	200	1000	0	175.3968	-	-	0.6315
2	180	650	0	175.3968	-	-	0.9007
3	500	2500	0	175.3968	-	-	0.2687
4	200	1000	3	135.3982	599.4954	1.0097	0.5844
5	200	940	5	117.2110	359.0514	1.0162	0.6153
6	180	650	3	135.3982	599.4954	1.0097	0.9538
7	180	650	5	117.2110	359.0514	1.0162	0.8747
8	500	2000	3	135.3982	599.4954	1.0097	0.3143
9	500	2000	5	117.2110	359.0514	1.0162	0.3718

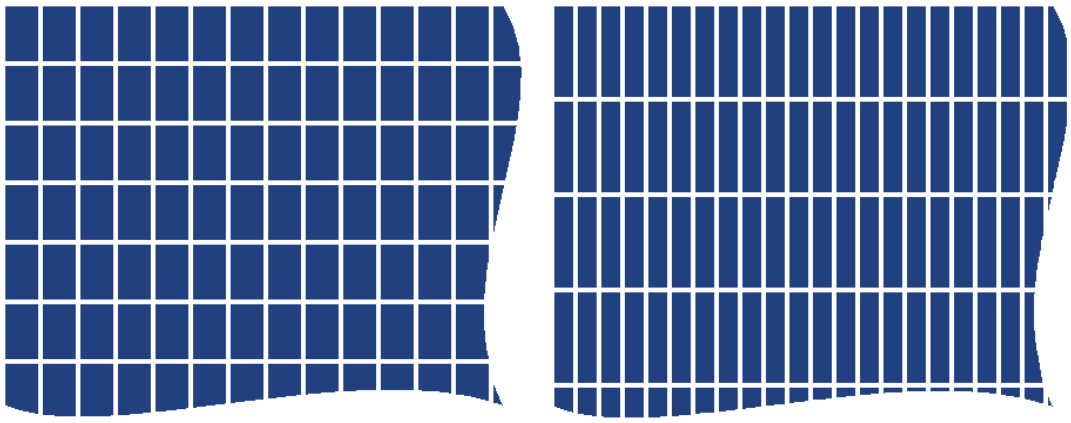

Figure 8: Non-uniform pixel distribution. Left: 2:3 discretization *i.e.* three horizontal pixels *vs.* two vertical pixels. Right: 1:4 discretization.

is large enough, a smaller aspect ratio is obtained. Moreover, from these graphs we can see that even though the optimal points are obtained by optimization of the relative error in estimating the *Y* component, the optimal aspect ratio is smaller than one in most configurations (and in fact in more practical configurations), which means that the *X* and *Z* estimates are improved as well, compared to a uniform pixel distribution over the *x* and *y* axes.

Table 2 presents a set of optimal aspect ratios and some other parameters calculated for a number of typical start points and viewing ranges. As mentioned before, the optimal ratio may vary significantly depending on viewing distance and volume. We therefore focus on closer objects in the optimization process. For example, if we optimize for an object appearing between 200 to 1000 mm, *i.e.* within one meter of the viewer/camera (Table 2 – row 1), then the optimal pixel aspect ratio will be calculated as 0.6315 which after rounding off is approximately equivalent to a non-uniform 2:3 pixel distribution ([pixel-width]:[pixel-height] or two vertical pixels versus three horizontal pixels). If an object appears between 500 to 2500 mm (Table 2 – row 3) the optimal aspect ratio will be 0.2687, which is close to a 1:4 pixel discretization. Figure 8 shows these two non-uniform pixel distributions. It is well known that the human visual system is less sensitive to depth discrimination for distant objects. Thus, for a 3D scene composed of multiple objects at various distances, the quality of closer objects has bigger impact on the overall visual impression.

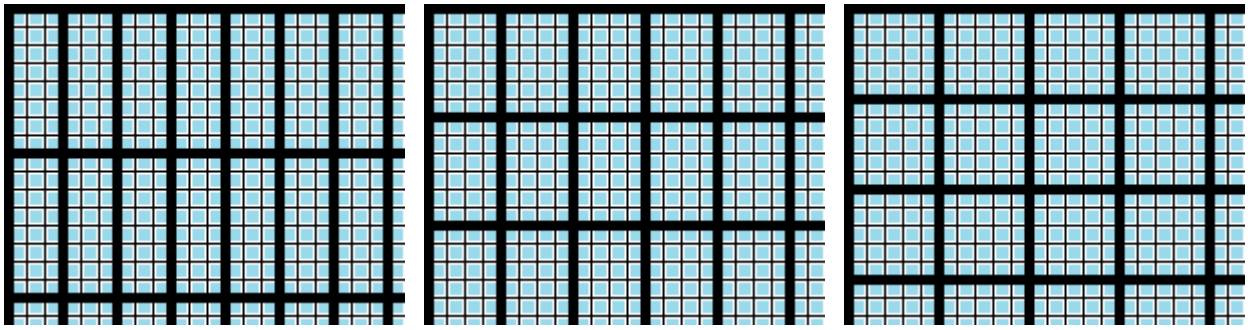

Figure 9: Simulating different pixel aspect ratios by grouping neighboring pixels as a single pixel while preserving the same total resolution by devoting almost the same number of pixels to each virtual pixel in all three configurations. Left: 3:8, Middle: 2:3, and Right: 1:1 pixel aspect ratio.

5.2. USER TESTS WITH 3D MODELS AND ANALYSIS

We conducted a set of user tests to compare our theoretical results with subjective user evaluations. For simplicity, we restricted ourselves to a parallel stereo setup, but the same approach can be applied for conducting user tests under a stereo setup with vergence. Since we used conventional 2D displays in our subjective studies and it was impossible to adjust the actual pixel aspect ratio on these screens, we needed to devise some techniques to simulate different non-uniform discretizations along the horizontal and vertical axes. For this purpose we established a virtual stereo imaging system with arbitrary sized rectangular pixels. This virtual stereo imaging system accepts 3D structure descriptions such as 3D triangular meshes or 3D point clouds and generates different stereo projections. On the display side, we used a group of neighbouring pixels on a conventional display to simulate different aspect ratios. For example, a set of 5x5, 6x4, or 8x3 neighbouring pixels can be grouped together as a single pixel to simulate 1:1, 2:3, and 3:8 pixel aspect ratios while the total resolution is kept almost the same for these three combinations (see Figure 9 for a pictorial description of this process). Although this method is a bit restrictive in term of selecting any arbitrary pixel aspect ratio, it is a very simple, inexpensive technique for simulating different pixel aspect ratios (PARs). We also used this technique to create test sets from real stereo images. We used pure red and blue colors for generating the left and right views, respectively. A pair of simple anaglyph red-blue glasses was used by the viewers to conduct the right and left views to the corresponding eye.

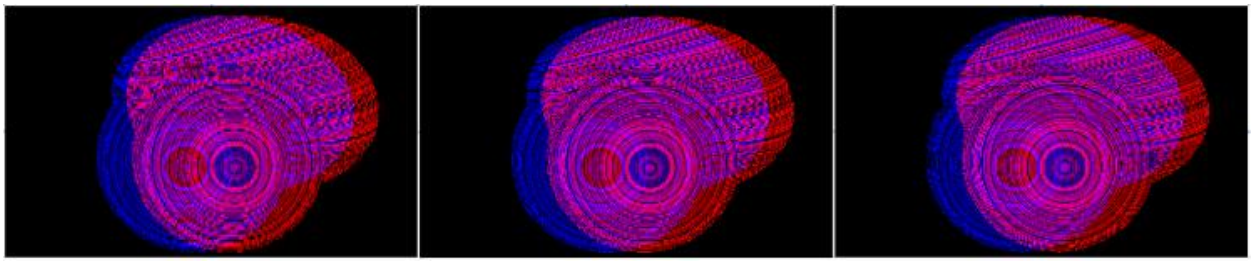

Figure 10: Red-blue stereo pairs generated from a synthetic 3D scene with different aspect ratios. Left 3:8 (pixel width: pixel height), middle 2:3, and right 1:1.

Experiment 1 - User Tests with Synthetic Stereo Images - Figure 10 shows a set of three red-blue images generated from a synthetic 3D model with three different pixel aspect ratios ─ 3:8, 2:3 and 1:1. These images are generated using our virtual stereo imaging system. The advantage of using synthetic scenes is that the properties of the scene can be defined according to the targeted evaluation criteria. Here, the model is composed of a set of 3D points which collectively describe a sphere with a radius of 19 cm and an ellipsoid with semi-minors of 50, 15, and 35 cm. The centers of the sphere and ellipsoid are located at a distance of 40 and 80 centimetres respectively from the stereo camera. Two arbitrary rotations are also applied to the ellipsoid. The colors (intensities) are randomly assigned to the 3D scene points but geometric shape of the objects is preserved by assigning the same color to points located on the same orbit. For this test set, the imaging system is configured using the values shown in Table 1. Considering the depth range filled by these two objects, the theoretical optimal aspect ratio is given by the first row in Table 2, *i.e.* 0.6315, which is almost equivalent to an aspect ratio of 2:3 – the ratio that is used to generate the middle image in Figure 10.

We used the images shown in Figure 10 to conduct our first set of user tests. These images were numbered from 1 to 3 from left to right. We followed the 3D visual experience model proposed in [16], and asked the viewers to rank the images according to three criteria: sense of depth, picture quality and overall sense. Sense of depth represents how objects are distinguishable and well-shaped across the depth dimension. Picture quality refers to the quality of the perceived image which may be influenced by several factors such as blockiness, brightness, noise and blurriness. Here, the picture quality is affected

Sense of Depth

Groups	*Count*	*Sum*	*Avg*	*Var*
Image 1	90	147	1.63	0.64
Image 2	90	195	2.17	0.54
Image 3	90	198	2.20	0.63

ANOVA: Single Factor

Source of Variation	*SS*	*df*	*MS*
Between Groups	18.2	2	9.1
Within Groups	161.8	267	0.61

F = 15.02 **$F_{critical}$ = 3.03** **P-value = 6.6E-07**

Picture Quality

Groups	*Count*	*Sum*	*Avg*	*Var*
Image 1	90	204	2.27	0.83
Image 2	90	180	2.00	0.36
Image 3	90	156	1.73	0.69

ANOVA: Single Factor

Source of Variation	*SS*	*df*	*MS*
Between Groups	12.8	2	6.4
Within Groups	167.2	267	0.63

F = 10.22 **$F_{critical}$ = 3.03** **P-value = 5.29E-05**

Overall Sense Preference

Groups	*Count*	*Sum*	*Avg*	*Var*
Image 1	90	180	2.00	0.72
Image 2	90	159	1.77	0.49
Image 3	90	201	2.23	0.70

ANOVA: Single Factor

Source of Variation	*SS*	*df*	*MS*
Between Groups	9.8	2	4.9
Within Groups	170.2	267	0.64

F = 7.69 **$F_{critical}$ = 3.03** **P-value = 0.000568**

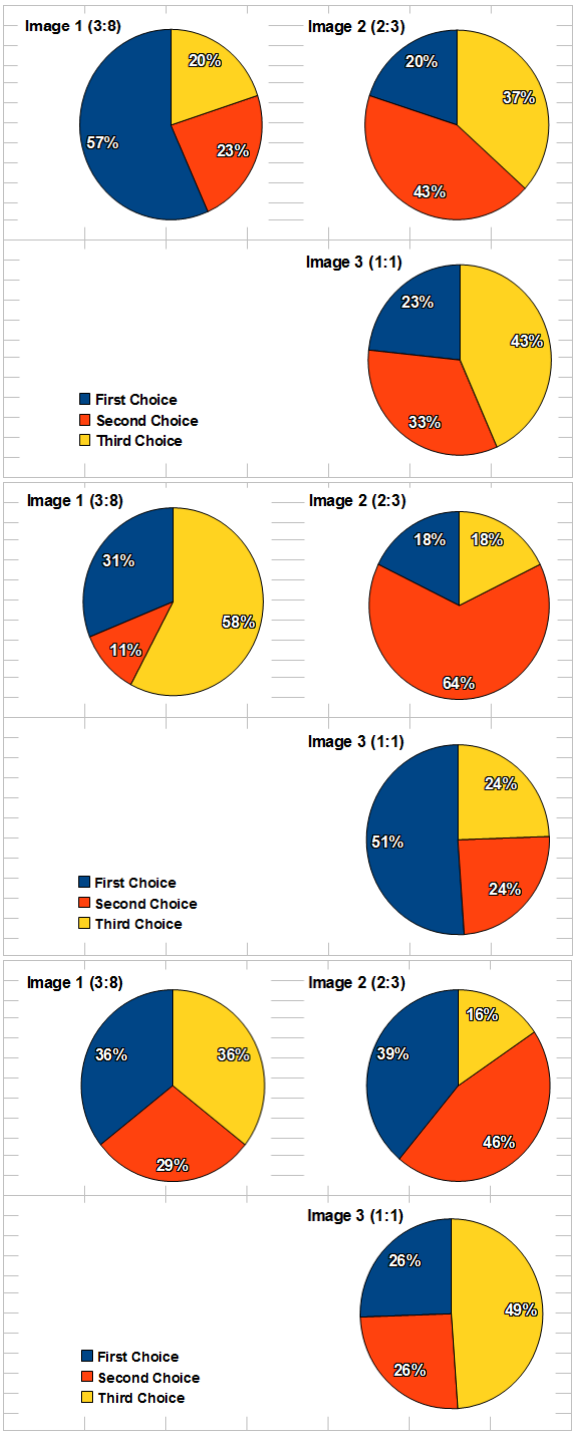

Figure 11: Results of conducting subjective tests with images shown in Figure 10.

mainly by the non-similar discretization across the vertical and horizontal dimensions. Overall sense measures a viewer's overall satisfaction with each of these three 3D representations.

Fifteen subjects (university graduate students) participated in our experiments. None of them were aware of the judgments made by other participants and did not know the parameters used to create these three images. Before ranking, we asked all participants to describe the reconstructed 3D scene to make sure that all of them were able to clearly perceive the 3D scene composed of the ellipsoid and the sphere in front it. Then we asked them to look at the 3D pictures as many times as they needed to decide on the rankings. Since the entire screen was needed to display an image, the viewers were asked to switch forward and backward to view all three images for comparison. A 14.1" display with 1280x800 resolution is used for this experiment.

The pie charts in Figure 11 illustrate the result of this subjective study. The figure also presents some statistics for each evaluation criterion including mean and variance in each group (image) and results of analysis of variance (ANOVA test) for comparing means of groups. These results suggest a correlation between the sense-of-depth and pixel aspect ratio. As depicted in the top-left graph, 57 percent of viewers have the best sense of depth with Image 1 which has a pixel aspect ratio of 3:8, 20 percent have the best sense of depth in Image 2 with pixel aspect ratio 2:3, and 23 percent selected Image 3 with a regular 1:1 pixel aspect ratio. In other words, a majority of the viewers preferred Image 1, which has the finest horizontal resolution, as their first choice. As suggested by ANOVA test it is very unlikely (*P*-value = 6.6E-07) that the results obtained for Image 1, Image 2, and Image 3 to be belong to the same distribution. In other words, it is very unlikely that users have selected images by chance. On the average, Image 1 is selected as the first choice with significant difference and Image 2 as the second preference, even though the difference between Image 2 and Image 3 is not statistically significant. This result, in general, is consistent with our theoretical results that images with horizontally finer resolution provide better sensation of depth. Some environmental and individual factors may explain why some people (the minority) preferred Image 3. Since the human visual system varies from one person to another, we could not fix a viewing distance for the best stereo effect. During the experiments, the viewers' body positions and movements could have affected the visual effect. Also, the capability of perceiving stereo effect varies among the subjects. The small visual differences between the images might not be detectable to all

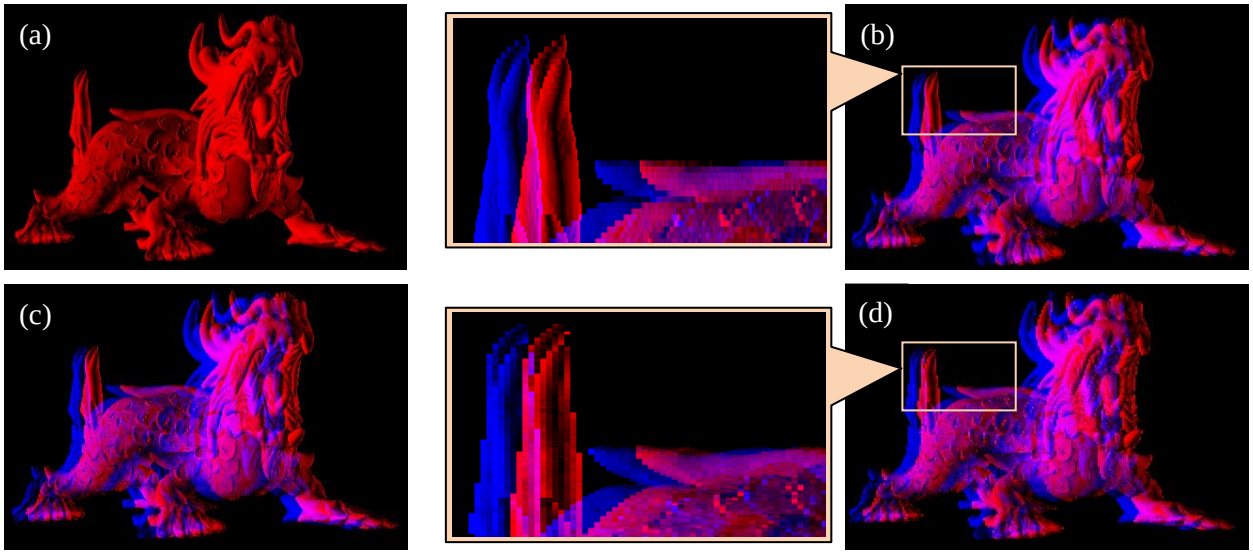

Figure 12: Samples generated from the Chinese Dragon 3D point cloud. (a) Original 3D object (red channel projection), (b) 1:4, (c) 1:1, and (d) 4:1 pixel aspect ratios.

viewers. For small visual differences, it is worth mentioning that there were circumstances where the viewers could not distinguish the sense of depth between two images. In these cases, we assigned equal ranking to both images.

For ranking against picture quality (Figure 11 second row), viewers were asked to rank images regardless of any perception of depth. Again, the results show meaningful but reverse relationship between pixel aspect ratio and picture quality. More than 50 percent selected Image 3 for the best picture quality. As discussed earlier in Section 3, this may be the result of image degradation that happens in the vertical direction when the discretization becomes coarser in that direction, and as a result the *Y* component estimation error might be increased. However, this trade-off of obtaining finer horizontal discretization by sacrificing some quality in the *Y* component improves both the *X* and *Z* components, and therefore the viewers have better overall sense in Image 1 and Image 2 which have non-uniform discretization (Figure 11 third row). We observe in Figure 11 that in terms of overall sense, Image 2 with aspect ratio 2:3 was the first or second choice for most of the viewers. This result illustrates certain compromise in the viewers' decisions towards the trade-off between sense of depth and picture quality. It should be noted that the ANOVA test again rejects the null hypothesis of obtaining these results by

(a) Aloe (b) Bowling (c) Rocks (d) VP (e) Chinese dragon

Figure 13: Set of images used in experiment 2 – user evaluation with real images and objects.

chance for both picture quality (*P*-value = 5.29E-05) and overall sense preference (*P*-value =0.000568) criteria.

Figure 12 represents more clearly the image quality degradation in the vertical or horizontal direction for a coarser discretization. These images are generated from Chinese Dragon 3D point cloud using our virtual stereo imaging system with 1:4, 1:1, and 4:1 pixel aspect ratios, respectively, with the same total resolution. For the 1:4 image (Figure 12-b) the discretization is more apparent along horizontal lines. For example, observe the horizontal supplementary tail over the vertebral line of the dragon. On the other hand, for the 4:1 image (Figure 12-c) the degradation is more apparent along vertical lines. Observe the dragon's vertical tail as an example. Both these images reveal lower image quality in some parts compare to the uniform pixel discretization in the 1:1 image (Figure 12-d). In general, for complex scenes composed of a set of edges, line, or textural features, depending on the feature directions, non-uniform pixel discretization may degrade the picture quality in parts while quality may be improved in some other areas. However, since a finer resolution on the horizontal direction also improves 3D estimation (sense of depth), in overall, a better 3D picture is experienced when an appropriate aspect ratio is selected.

Experiment 2 – User Tests with Real Stereo Images – We conducted a second round of user tests to study the effects of pixel aspect ratio on 3D perception using real stereo scenes and objects. We simplified evaluation criteria in this experiment and asked the users to compare the images just in term of their overall personal preference. Instead, we used several scenes in this experiment. As already mentioned, we used pixel grouping technique to create images with different PARs from real stereo images. For each scene, stereo pairs with different pixel aspect ratios but the same total resolution were generated and the

Overall Personal Preference

Groups	*Count*	*Sum*	*Avg*	*Var*
PAR 3:8	49	133	2.71	0.29
PAR 2:3	49	71	1.45	0.34
PAR 1:1	49	90	1.84	0.56

ANOVA: Single Factor

Source of Variation	*SS*	*df*	*MS*
Between Groups	41.18	2	20.59
Within Groups	56.82	144	0.39

F = 52.19 **$F_{critical}$ = 3.06** **P-value = 8.99E-18**

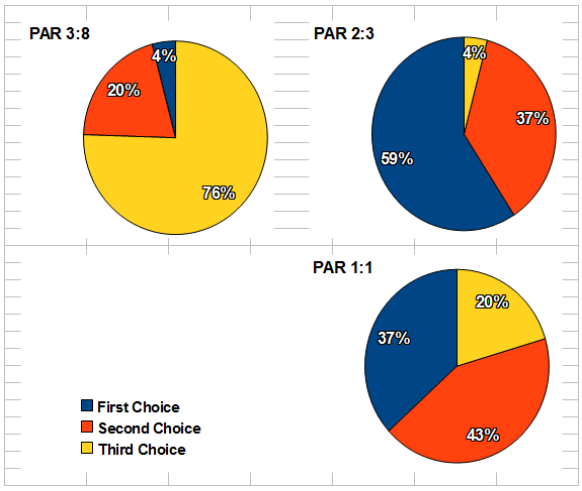

Figure 14: Results of subjective tests using real 3D scenes.

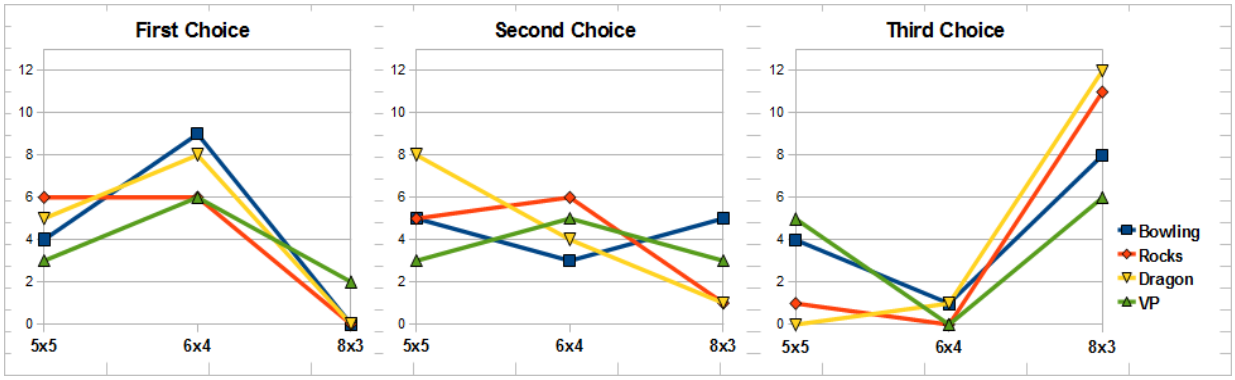

Figure 15: Results of subjective tests on real 3D scenes for different scenes used in experiment.

users were asked to choose their preferred ones in a pair-wise manner. More precisely, assuming that I_1, I_2, and I_3 are the three stereo images generated for scene S with three different PARs 1:1 (5x5), 2:3 (6x4), and 3:8 (8x3), we created three pairs I_1I_2 (or I_2I_1), I_2I_3 (or I_3I_2), and I_1I_3 (or I_3I_1) of these images (the order of images are picked by random). Images (b), (c), (d), and (e) of Figure 13 are used for generating these test cases. We put all these pairs together and shuffled them to form the test set. We enriched the test set by randomly interleaving some other cases that were generated with some finer total resolutions to examine how users will treat the images when the noise introduced by pixel grouping technique is less apparent. These finer test cases were generated from images (a), (b), (c), and (d) of Figure 13. These test cases were generated by grouping 4x4 and 5x3 pixels to examine 1:1 PAR versus 3:5 (~2:3) PAR, and by grouping 3x3 and 4x2 pixels to compare 1:1 PAR versus 1:2 PAR.

Thirteen subjects are participated in this experiment, majority of them university graduate students. During the test, the users are asked to show their preference toward one of the stereo images in each pair. Similar to Experiment 1 we did not enforce any time limit but all subjects spent almost the same amount of time (between 25 to 30 minutes) to evaluate all cases. We also did not enforce a specific viewing distance but we recommended the viewing distance of 70-80 cm considering that we were using a wider display screen (19" display) in this experiment.

Figure 14 shows the user evaluation results for 5x5, 6x4, and 8x3 groupings with some associated statistics. As can be seen from the pie charts, the results are greatly different for these three groups in favour of the images with pixel aspect ratio 2:3 (6x4 pixel grouping). The outcome of the ANOVA test (F = 52.19, P-value = 8.99E-18) also confirms the substantial difference between groups. PAR 2:3 not only has been selected as the first choice by many viewers (59%) but also rarely has been the third choice of viewers (4%). In other words, in 96% of cases images with PAR 2:3 are selected either as first or second preferred image.

In contrast to the results of Experiment 1 (synthetic scene), here images with PAR 1:1 (*i.e.* images with uniform pixel distribution) have mainly owned the second place (43%) and images with PAR 3:8 have often been selected as the least favourite ones (76%). This can be explained by the way that pixel grouping technique treats image features. In fact, since the real images we have used in this experiment possess more textural or structural features, comparing to the simple synthetic scene of Experiment 1 that is composed of poorly featured geometrical objects, they are more prone to the noise introduced by sharper simulated PARs.

Figure 15 shows a more detailed representation of the results presented in Figure 14 pie charts. The line charts of this figure show how different scenes are treated by viewers. In general, the drawings show a consistent behaviour toward all four different scenes that are used in this experiment. However, the results seem to be a little affected by the scene structures. For example, the users have had tendency to select the Rocks scene with more horizontally oriented edges and textural features, as their first choice than the VP

scene with more vertically aligned structures. However, the differences are not that much that contradicts each other and the expected theoretical results.

We separately analysed the results of test cases with finer total resolution. For 3x3 verses 4x2 test cases, 54% of viewers were in favour of images with uniform pixel distribution (3x3, 1:1 PAR). We did not find statistically significant difference between these two groups. However, at least we can say that the images with 1:2 PAR (4x2) with less apparent pixel grouping noise have received much more attention when they are compared with the images of corresponding uniform pixel distribution than images with 3:8 PAR (8x3). For 4x4 versus 5x3 cases we found that 63% of viewers were in favour of images with non-uniform pixel distribution (5x3, 3:5 PAR) with statistically significant difference between two groups (paired T-Test *P*-value = 0.01). These results are consistent with the results obtained for images with the same PARs but coarser total resolution (see pie charts in Figure 14).

5.3. EFFECTS OF VERGENCE ON 3D VIEWING

In order to study the benefit of vergence in 3D viewing we compared the effect of visualization at close range with and without vergence. Figure 17 shows a test set generated for the Bunny 3D mesh under parallel camera configuration. As depicted in this figure the amount of disparity for this small model is quite large. This makes it difficult for some people to easily merge the left and right views and causes eye stress after a short time. Vergence can help control the amount of disparity in the stereo output. Even with a small vergence angle the amount of disparity is significantly affected. Figure 17 shows the images generated for the same model as in Figure 16, but incorporating a vergence angle of around 3 degrees. These images are much easier to fuse together for a human observer. However, vergence influences the 3D output in several ways. In fact, the amount of disparity affects the extent of “popping-out-of” or “sinking-into” the screen. In other words, focusing in front of the object moves it farther away and focusing at the back of the object brings it closer to the 3D display viewer. Although these effects may

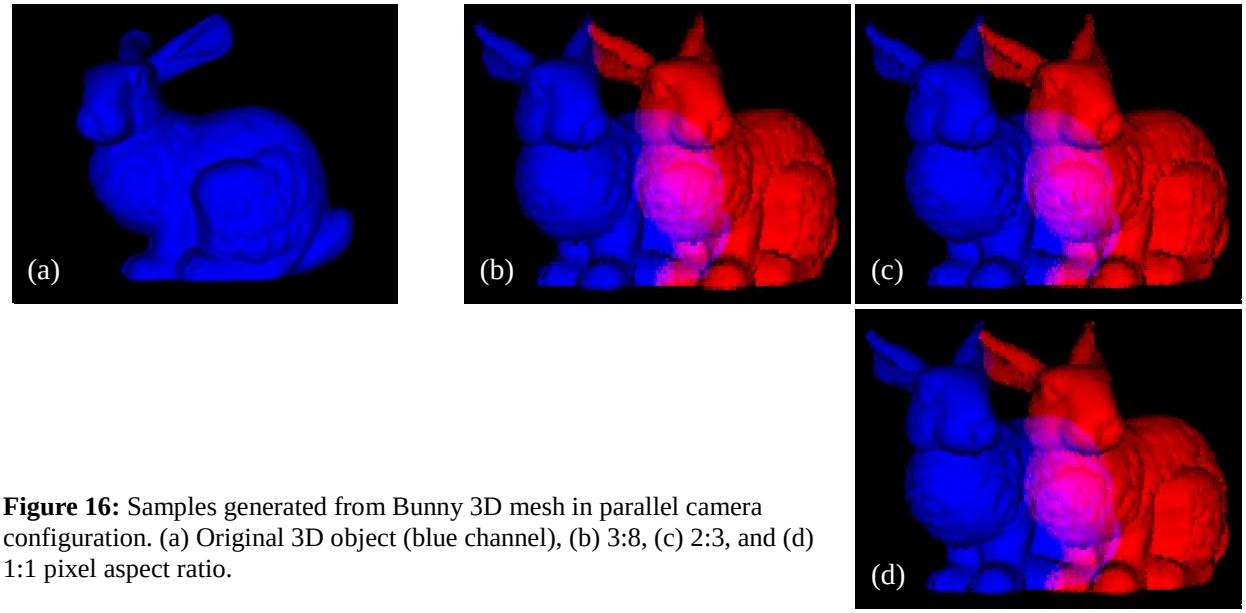

Figure 16: Samples generated from Bunny 3D mesh in parallel camera configuration. (a) Original 3D object (blue channel), (b) 3:8, (c) 2:3, and (d) 1:1 pixel aspect ratio.

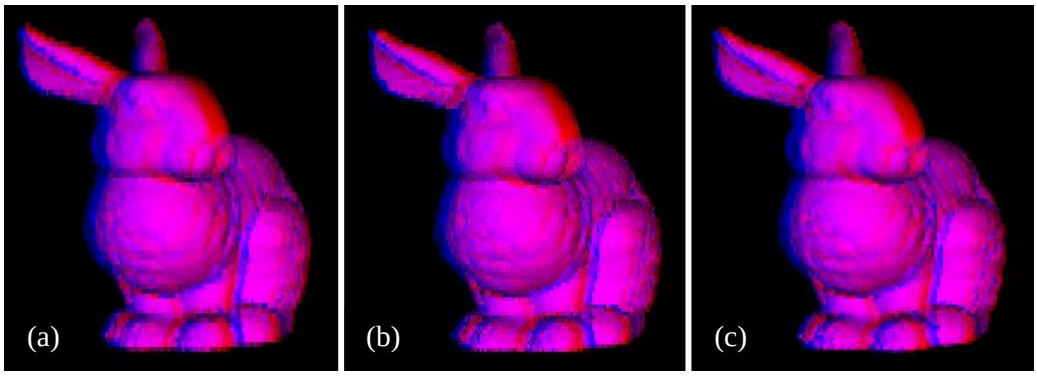

Figure 17: Samples generated from Bunny 3D mesh, with vergence being included. (a) 3:8, (b) 2:3, and (c) 1:1 pixel aspect ratio.

help provide enhanced impressions for some 3D content, they may also cause some undesired side-effects such as improper scale or unnatural shape [6].

6. CONCLUSION AND FUTURE WORK

We described a theoretical model for determining the trade-off between vertical and horizontal resolutions when viewing a 3D scene in stereo. The optimization procedure discussed in this paper is based on mathematical modeling. We inferred optimal PARs by considering two commonly used stereo setups: parallel and with vergence, and showed how optimal PAR is related to the corresponding stereo configuration parameters. Using the formulations derived, we showed that although the optimal PAR may

considerably vary with different parameter values, in general, for a given total resolution a more dense distribution of pixels along the horizontal direction often improves the sense of depth and overall output quality of the 3D display. Applying practical parameter values into inferred equations, we theoretically suggested that a pixel aspect ratio of 2:3 ([pixel width]:[pixel height]) will lead to a better stereo-based 3D viewing experience. We validated the theoretical results through conducting subjective user experiments on both synthetic and real stereo content with different simulated pixel aspect ratios. Our experiments showed that from a subjective point of view the suggested optimal PAR indeed improves the 3D output.

Our method of simulating different PARs using pixel grouping introduces noticeable blockiness into images. In this regard, it will be interesting to incorporate the concept of Just Noticeable Difference (JND) for 3D perception [22] to see how the results may change. We expect that the whole effect of applying optimal PAR on a high-resolution display lead to a noticeable improvement on the overall 3D output quality, even though, the user may not be able to separately distinguish the pixels of such high-resolution displays (similar to 2D displays that overall quality improves by increasing the resolution even if the pixels are not distinguishable by viewers). It should also be clarified that the optimal PAR greatly improves the resolution across the depth dimension. This is very important for many interactive stereo-based 3D applications that usually suffer from low pointing accuracy across the depth. Optimal PAR facilitates more accurate 3D pointing and hence more accurate 3D interaction and manipulation [27].

It will be also interesting to incorporate other factors such as different viewing conditions, human perceptual factors [17][18][19][20], motion factors [21], and error measures produced by cross-talk, correlation, and epipolar constraints [23][24], in order to derive more advanced models. Such models can also take into account the relative importance of vertical vs. horizontal disparities for the human visual system, and incorporate an image quality metric to reflect the picture quality changes caused by non-uniform discretization. Moreover, our brief discussion on various aspects of introducing vergence into the system can also be extended especially for large vergence angles and for scenes composed of multiple objects with cameras focused on one of these objects.

The approach developed in this study could be extended to understand the cause of the anisotropy [25][26] present between horizontal vs. vertical depth perception through stereo (or motion induced stereo) in humans. Our analysis shows that considering practical 3D viewing distances, uniform distribution of pixels on a sensor (display) produces better reconstruction results for the horizontal axis, when the stereo baseline is along the horizontal axis. However, the analysis needs to be extended taking into account the spatially varying distribution of cones in the human eye and possibly issues like "depth consistency" in human perception. Finally, the analysis in this paper could be made more geometric by considering "volumes of uncertainty" extending the approach discussed in [10].

Acknowledgments

The authors thank iCORE, Alberta and NSERC, Canada, for their research support. The authors also thank the Stanford Computer Graphics Laboratory for providing the Bunny and the Chinese Dragon 3D models (the Chinese Dragon is originally provided by XYZ RGB Inc.), and the Middlebury College for providing some of the real 3D scenes that are used in user evaluations.

References

[1] D. F. McAllister, "Stereo computer graphics and other true 3D technologies", Princeton University Press, 1993.

[2] K. Satoh and Y. Ohta, "Passive depth acquisition for 3D image displays", IEICE transactions on Information and System, September 1994.

[3] O. Schreer, P. Kauff and T. Sikora (Eds.), "3D Video Communication", Wiley, 2005.

[4] W. Matusik, and H. Pfister, "3D TV: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes", ACM Transactions on Graphics, 2004, volume23, pp 814-824.

[5] S. Baker, R. Szeliski and P. Anandan, "A Layered Approach to Stereo Reconstruction", IEEE CVPR, 1998.

[6] B. Javidi, F. Okano, "Three-dimensional Television, Video, and Display Technologies", Chapter 1, Springer, 2002.

[7] T. Mitsuhashi, “A study of the relationship between scanning specifications and picture quality,” NHK, NHK Laboratories Note, no. 256, Oct. 1980.

[8] I. Cheng and A. Basu, “Optimal aspect ratio for 3D TV,” IEEE 3D TV Conference, 4 pages, KOS, Greece, May 2007.

[9] A. Basu, “Optimal discretization for stereo reconstruction”, Pattern Recognition Letters, vol. 13, no. 11, pp 813-820, 1992.

[10] H. Sahabi and A. Basu, “Analysis of error in depth perception with vergence and spatially varying sensing”, Computer Vision and Image Understanding, vol. 63, no. 3, May, pp. 447-461, 1996.

[11] N. Holliman, “3D display systems”, Department of Computer Science, University of Durham, Science Laboratories, South Road, Durham, DH1 3LE; Feb 2, 2005; http://www.dur.ac.uk/n.s.holliman/ Presentations/3dv3-0.pdf.

[12] I. Cheng, K. Daniilidis, and A Basu, “Optimal Aspect Ratio under Vergence for 3D TV”, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008, pp.209-212, 28-30 May 2008.

[13] A. Basu, and H. Sahabi, “Optimal Non-uniform discretization for Stereo Reconstruction”, Pattern Recognition, 1996, Proceedings of the 13th International Conference on, vol.1, pp.755-759 vol.1, 25-29 Aug 1996.

[14] M. Z. Brown, D. Burschka, and G. D. Hager, “Advances in Computational Stereo”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 8, August 2003.

[15] D. Scharstein, and R. Szeliski, “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms”, International Journal of Computer Vision”, vol. 47, pp. 7-42, 2002.

[16] P. J. H. Seuntins, “Visual experience of 3D TV”, Eindhoven: Technische Universiteit Eindhoven, 2006, Proefschrift. http://alexandria.tue.nl/extra2/200610884.pdf.

[17] B.T. Backus, D.J. Fleet, A.J. Parker and D.J. Heeger, “Human Cortical Activity Correlates With Stereoscopic Depth Perception,” The Journal of Neurophysiology, 2054-2068, October 2001.

[18] I. Cheng and A. Basu, “Perceptually Optimized 3D Transmission over Wireless Networks”, IEEE Transactions on Multimedia, 386-396, 2007.

[19] L.-F. Cheong, “Depth perception under motion and stereo with implications for 3D TV”, 3D TV Conference, Greece, 4 pages, May 2007.

[20] Y. Pan, I. Cheng and A. Basu, "Quality Metric for Approximating Subjective Evaluation of 3D Objects", IEEE Trans. on Multimedia, 269-279, April 2005.

[21] L.-F. Cheong, C. Fermüller and Y. Aloimonos, "Effects of errors in the viewing geometry on shape estimation," Computer Vision and Image Understanding, 356-372, Sep. 1998.

[22] I. Cheng and P. Boulanger, "A 3D Perceptual Metric using Just-Noticeable-Difference", EUROGRAPHICS, 2005, Dublin, Ireland.

[23] T. Brodsky, C. Fermüller and Y. Aloimonos, "Structure from motion: Beyond the epipolar constraint," Int. Journal of Computer Vision, vol. 37, no. 3, 231-258, 2000.

[24] N. Ayache, "Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception," MIT Press, 1991.

[25] B.J. Rogers and M.E. Graham, "Anisotropies in the perception of three-dimensional surfaces," Science, 221, 1409-1411, 1983.

[26] V. Cornilleau-Pérès and J. Droulez, "Visual perception of surface curvature: psychophysics of curvature detection induced by motion parallax," Perception and Psychophysics, 46: 351-364. 1989.

[27] Hossein Azari, Irene Cheng, and Anup Basu, "Stereo 3D Mouse Cursor: A Method for Interaction with 3D Objects in a Stereoscopic Virtual 3D Space," International Journal of Digital Multimedia Broadcasting, vol. 2010, Article ID 419493, 11 pages, 2010.

Appendix A

Derivation for Result 3:

$$F_p(e_x)=\int_{Z_{\min}}^{Z_{\max}}\int_{-y_{\max}}^{y_{\max}}\int_{X_0-\frac{\Delta X}{2}\frac{(Z-Z_0)}{(Z_{\max}-Z_0)}}^{X_0+\frac{\Delta X}{2}\frac{(Z-Z_0)}{(Z_{\max}-Z_0)}}\left(\frac{fb_x}{Z}-|y|2e_x^2R\right)^2 dXdydZ$$

$$=\int_{Z_{\min}}^{Z_{\max}}\int_{-y_{\max}}^{y_{\max}}\Delta X\frac{(Z-Z_0)}{(Z_{\max}-Z_0)}\left(\frac{f^2b_x^2}{Z^2}-4Re_x^2fb_x\frac{|y|}{Z}+4R^2e_x^4y^2\right)dydZ$$

$$=e_x^2\left[\frac{4\Delta XZ_0Rfb_x}{(Z_{\max}-Z_0)}\int_{Z_{\min}}^{Z_{\max}}\int_{-y_{\max}}^{y_{\max}}\frac{|y|}{Z}dydZ-\frac{4\Delta XRfb_x}{(Z_{\max}-Z_0)}\int_{Z_{\min}}^{Z_{\max}}\int_{-y_{\max}}^{y_{\max}}|y|dydZ\right]$$

$$+e_x^4\left[\frac{4\Delta XR^2}{(Z_{\max}-Z_0)}\int_{Z_{\min}}^{Z_{\max}}\int_{-y_{\max}}^{y_{\max}}Zy^2dydZ-\frac{4\Delta XR^2Z_0}{(Z_{\max}-Z_0)}\int_{Z_{\min}}^{Z_{\max}}\int_{-y_{\max}}^{y_{\max}}y^2dydZ\right]+I_0$$

where I_0 is independent of e_x, and thus has no influence on the minimization of $F_p(e_x)$.

$$F_p(e_x)=e_x^2\left[\frac{4\Delta XZ_0Rfb_x}{(Z_{\max}-Z_0)}I_1[Z_{\min},Z_{\max}]-\frac{4\Delta XRfb_x}{(Z_{\max}-Z_0)}I_2[Z_{\min},Z_{\max}]\right]$$

$$+e_x^4\left[\frac{4\Delta XR^2}{(Z_{\max}-Z_0)}I_3[Z_{\min},Z_{\max}]-\frac{4\Delta XR^2Z_0}{(Z_{\max}-Z_0)}I_4[Z_{\min},Z_{\max}]\right]+I_0$$

where I_1 to I_4 are integrals defined in Result 3.

Considering the derivative of $F_p(e_x)$ in terms of e_x and equating it to zero, we have:

$$F_p'(e_x)=2e_x\left[\frac{4\Delta XZ_0Rfb_x}{(Z_{\max}-Z_0)}I_1[Z_{\min},Z_{\max}]-\frac{4\Delta XRfb_x}{(Z_{\max}-Z_0)}I_2[Z_{\min},Z_{\max}]\right]+$$

$$4e_x^3\left[\frac{4\Delta XR^2}{(Z_{\max}-Z_0)}I_3[Z_{\min},Z_{\max}]-\frac{4\Delta XR^2Z_0}{(Z_{\max}-Z_0)}I_4[Z_{\min},Z_{\max}]\right]\quad\cdot$$

$$=0$$

Since we cannot take $e_x = 0$ as the optimal point, we should have:

$$fb_x\left(Z_0I_1[Z_{\min},Z_{\max}]-I_2[Z_{\min},Z_{\max}]\right)+2\,\mathrm{Re}_x^2\left(I_3[Z_{\min},Z_{\max}]-Z_0I_4[Z_{\min},Z_{\max}]\right)=0$$

Solving this equation in terms of e_x, and again considering that we cannot take negative values for e_x as the optimal point, we obtain the following equation (or Equation 29) as the possible optimal point:

$$e_x=\left(-\frac{fb_x}{2R}\frac{Z_0I_1[Z_{\min},Z_{\max}]-I_2[Z_{\min},Z_{\max}]}{I_3[Z_{\min},Z_{\max}]-Z_0I_4[Z_{\min},Z_{\max}]}\right)^{\frac{1}{2}}=e_{xOpt}\,.$$

For ease of reference in the rest of the argument, we have renamed the value of e_x in above equation as e_{xOpt}. Now, we need to show that $F_p(e_{xOpt})$ is the only minimum value of $F_p(e_x)$ over the domain of possible values for e_x (as we discussed earlier in Result 1, the domain of possible values for e_x is the open interval $(0,\infty)$). For this purpose, first we show that $F_p(e_x)$ has a local minimum at e_{xOpt} under the specified assumptions, and then we show that it does not take any other minimum over the $(0,\infty)$ interval. We use the second derivative test to show that $F_p(e_x)$ has a local minimum at e_{xOpt} :

$$F_p''(e_x) = \frac{8\Delta XR}{Z_{\max} - Z_0}\left[fb_x\left(Z_0 I_1[Z_{\min}, Z_{\max}] - I_2[Z_{\min}, Z_{\max}]\right) + 6\,\mathrm{Re}_x^2\left(I_3[Z_{\min}, Z_{\max}] - Z_0 I_4[Z_{\min}, Z_{\max}]\right)\right].$$

Therefore,

$$\begin{aligned} F_p''(e_{xOpt}) &= \frac{8\Delta XR}{Z_{\max} - Z_0}\left[fb_x\left(Z_0 I_1[Z_{\min}, Z_{\max}] - I_2[Z_{\min}, Z_{\max}]\right) + \right. \\ &\quad \left. 6R\left(-\frac{fb_x}{2R}\right)\left(\frac{\left(Z_0 I_1[Z_{\min}, Z_{\max}] - I_2[Z_{\min}, Z_{\max}]\right)}{I_3[Z_{\min}, Z_{\max}] - Z_0 I_4[Z_{\min}, Z_{\max}]}\right)\left(I_3[Z_{\min}, Z_{\max}] - Z_0 I_4[Z_{\min}, Z_{\max}]\right)\right] \\ &= \frac{8\Delta XR}{Z_{\max} - Z_0}\left[fb_x\left(Z_0 I_1[Z_{\min}, Z_{\max}] - I_2[Z_{\min}, Z_{\max}]\right) - 3\,fb_x\left(Z_0 I_1[Z_{\min}, Z_{\max}] - I_2[Z_{\min}, Z_{\max}]\right)\right] \\ &= \frac{8\Delta XR}{Z_{\max} - Z_0}\left[fb_x\left(I_2[Z_{\min}, Z_{\max}] - Z_0 I_1[Z_{\min}, Z_{\max}]\right)\right] \end{aligned}$$

Replacing I_1 and I_2 with their definitions we have:

$$F_p''(e_{xOpt}) = \left(\frac{8\Delta XR}{Z_{\max} - Z_0}\, fb_x y_{\max}^2\right)\left((Z_{\max} - Z_{\min}) - Z_0\left(\ln Z_{\max} - \ln Z_{\min}\right)\right).$$

If $F_p''(e_{xOpt}) > 0$ then $F_p(e_x)$ has a local minimum at e_{xOpt} . Since the first factor in the above equation is positive this condition implies that the second factor is positive. In other words, we should have:

$$(Z_{\max} - Z_{\min}) - Z_0\left(\ln Z_{\max} - \ln Z_{\min}\right) > 0$$

or equivalently,

$$\frac{Z_{\max} - Z_{\min}}{\ln Z_{\max} - \ln Z_{\min}} > Z_0.$$

This condition is not hard to achieve since in practice Z_0 is a small value comparing to $Z_{\min}$ and $Z_{\max}$ and the denominator of the left-side fraction grows much slowly than its numerator. As result, a reasonably

large viewing volume in the direction of the Z axis easily satisfies this condition and so $F_p(e_{xOpt})$ will be a local minimum under this assumption.

Following the same procedure as above for the other two extrema (*i.e.*, $e_x = 0$ and $e_x = -e_{xOpt}$), it is not difficult to show that under the same assumption $F_p(0)$ and $F_p(-e_{xOpt})$ are local maximum and local minimum, respectively. Therefore, considering that $F_p(-e_{xOpt})$, $F_p(0)$, and $F_p(e_{xOpt})$ are the only local extrema of function $F_p(e_x)$, $F_p(e_{xOpt})$ should be the only minimum (or equivalently a global minimum) over open interval $(0,\infty)$. □

Derivation for Result 4:

$$F_v(e_x) = \int_Z \int_y \int_X \left(\frac{fb_x}{Z} - 2RKe_x^2 |y| \right)^2 dXdydZ$$

Case I) $Z_{int} \le Z_{min}$

Considering the equations in AR5 (page 16), the integral in Equation 34 should be calculated according to the following range of values:

$$F_v(e_x) = \int_{Z_{\min}}^{Z_{\max}} \int_{-y_{\max}}^{y_{\max}} \int_{-\frac{1}{2}\left[(\Delta X_{\text{int}} - b_x)\frac{Z}{Z_{\text{int}}} + b_x\right]}^{+\frac{1}{2}\left[(\Delta X_{\text{int}} - b_x)\frac{Z}{Z_{\text{int}}} + b_x\right]} \left(\frac{fb_x}{Z} - 2RKe_x^2 |y| \right)^2 dXdydZ$$

Thus,

$$\begin{aligned} F_v(e_x) &= \int_{Z_{\min}}^{Z_{\max}} \int_{-y_{\max}}^{y_{\max}} \left((\Delta X_{\text{int}} - b_x)\frac{Z}{Z_{\text{int}}} + b_x \right) \left(\frac{f^2 b_x^2}{Z^2} - 4RKe_x^2 fb_x \frac{|y|}{Z} + 4R^2K^2e_x^4 y^2 \right) dydZ \\ &= e_x^2 \left[-4RKfb_x^2 I_1[Z_{\min}, Z_{\max}] - \frac{4(\Delta X_{\text{int}} - b_x)RKfb_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}] \right] \\ &\quad + e_x^4 \left[\frac{4(\Delta X_{\text{int}} - b_x)R^2K^2}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + 4R^2K^2 b_x I_4[Z_{\min}, Z_{\max}] \right] + I_0 \end{aligned}$$

where again I_0 is independent of e_x and has no effect on optimization of $F_v(e_x)$, and I_1 to I_4 are defined as in Result 3.

$$F_v'(e_x) = 2e_x\left[-4RKfb_x^2 I_1[Z_{\min}, Z_{\max}] - \frac{4(\Delta X_{\text{int}} - b_x)RKfb_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]\right] + 4e_x^3\left[\frac{4(\Delta X_{\text{int}} - b_x)R^2K^2}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + 4R^2K^2 b_x I_4[Z_{\min}, Z_{\max}]\right] = 0$$

or

$$\left[-fb_x^2 I_1[Z_{\min}, Z_{\max}] - \frac{(\Delta X_{\text{int}} - b_x)fb_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]\right] + 2e_x^2\left[\frac{4(\Delta X_{\text{int}} - b_x)RK}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + RKb_x I_4[Z_{\min}, Z_{\max}]\right] = 0 \cdot$$

Solving the last equation in terms of e_x and again considering that we cannot pick negative values as the optimal point, the value given in Equation 37 is obtained. To show that the function $F_v(e_x)$ takes a minimum value at this point and this is its global minimum value over the allowable values for e_x, we calculate the second derivative of $F_p(e_x)$ at this point:

$$F_v''(e_x) = 2\left[-4RKfb_x^2 I_1[Z_{\min}, Z_{\max}] - \frac{4(\Delta X_{\text{int}} - b_x)RKfb_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]\right] + 12e_x^2\left[\frac{4(\Delta X_{\text{int}} - b_x)R^2K^2}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + 4R^2K^2 b_x I_4[Z_{\min}, Z_{\max}]\right] = 8RK\left[-fb_x\left(b_x I_1[Z_{\min}, Z_{\max}] + \frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]\right) + 12e_x^2 RK\left(\frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + b_x I_4[Z_{\min}, Z_{\max}]\right)\right]$$

Recalling the expression in Equation 37 as e_{xOpt}, the value of second derivative of $F_v(e_x)$ at e_{xOpt} is calculated as:

$$F_v''(e_{xOpt}) = 8RK\left[-fb_x\left(b_x I_1[Z_{\min}, Z_{\max}] + \frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]\right) + 12RK\left(\frac{fb_x}{2RK}\right)\left(\frac{b_x I_1[Z_{\min}, Z_{\max}] + \frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]}{\frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + b_x I_4[Z_{\min}, Z_{\max}]}\right)\left(\frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_3[Z_{\min}, Z_{\max}] + b_x I_4[Z_{\min}, Z_{\max}]\right)\right] = 40RKfb_x\left(b_x I_1[Z_{\min}, Z_{\max}] + \frac{\Delta X_{\text{int}} - b_x}{Z_{\text{int}}} I_2[Z_{\min}, Z_{\max}]\right)$$

Considering the definitions of I_1 and I_2 and also noting that $(\Delta X_{\text{int}} - b_x)$ is greater than zero, we can see that $F_v''(e_{xOpt})$ is greater than zero. As a result, $F_v(e_{xOpt})$ is a local minimum for $F_v(e_x)$. Following the same

procedure for the other two extrema, it is not difficult to show that $F_v(0)$ and $F_v(-e_{xOpt})$ are local maximum and local maximum, respectively. Since $F_v(-e_{xOpt})$, $F_v(0)$, and $F_v(e_{xOpt})$ are the only extrema of $F_v(e_x)$, we can say that $F_v(e_{xOpt})$ is a global minimum over the interval $(0,\infty)$.

Case II) $Z_{max} \leq Z_{int}$

In this case following integral should be calculated:

$$F_v(e_x) = \int_{Z_{\min}}^{Z_{\max}} \int_{-y_{\max}}^{y_{\max}} \int_{-\frac{1}{2}\left[\Delta X_{\text{int}} \frac{Z-Z_0}{Z_{\text{int}}-Z_0}\right]}^{+\frac{1}{2}\left[\Delta X_{\text{int}} \frac{Z-Z_0}{Z_{\text{int}}-Z_0}\right]} \left(\frac{fb_x}{Z} - 2RKe_x^2 |y| \right)^2 dXdydZ$$

$$= \int_{Z_{\min}}^{Z_{\max}} \int_{-y_{\max}}^{y_{\max}} \left(\Delta X_{\text{int}} \frac{Z-Z_0}{Z_{\text{int}}-Z_0} \right) \left(\frac{f^2 b_x^2}{Z^2} - 4RKe_x^2 fb_x \frac{|y|}{Z} + 4R^2K^2e_x^4 y^2 \right) dydZ$$

Similar to Case I, calculating the derivative of $F_v(e_x)$ and equating it to zero gives the result. To show that the expression obtained, *i.e.* Equation 38, minimizes F_v over the domain of allowable values for e_x, the same argument as in Result 3 should be applied. The assumption of choosing a large enough volume along with the Z dimension is also needed to be considered for this case. Because of many similarities with the derivation for Result 3 the details are skipped here.

Case III) $Z_{min} < Z_{int} < Z_{max}$

In this case the integral in Equation 34 should be divided into two parts as follow. The other steps are similar to the Cases I and II and are skipped here.

$$F_v(e_x) = \int_{Z_{\min}}^{Z_{\text{int}}} \int_{-y_{\max}}^{y_{\max}} \int_{-\frac{1}{2}\left[\Delta X_{\text{int}} \frac{Z-Z_0}{Z_{\text{int}}-Z_0}\right]}^{+\frac{1}{2}\left[\Delta X_{\text{int}} \frac{Z-Z_0}{Z_{\text{int}}-Z_0}\right]} \left(\frac{fb_x}{Z} - 2RKe_x^2 |y| \right)^2 dXdydZ + \int_{Z_{\text{int}}}^{Z_{\max}} \int_{-y_{\max}}^{y_{\max}} \int_{-\frac{1}{2}\left[(\Delta X_{\text{int}} - b_x)\frac{Z}{Z_{\text{int}}} + b_x\right]}^{+\frac{1}{2}\left[(\Delta X_{\text{int}} - b_x)\frac{Z}{Z_{\text{int}}} + b_x\right]} \left(\frac{fb_x}{Z} - 2RKe_x^2 |y| \right)^2 dXdydZ \cdot \square$$