Identification of stereo rig alignment error based on vertical disparity map

Sandeep Katragadda
KiwiSecurity Software GmbH,
Guglgasse 15, 1110, Vienna, Austria
s.katragadda@kiwisecurity.com

ABSTRACT

Poor quality of 3D video content can lead to headache, blurry vision and overall exhaustive experience for the viewer. To ensure quality and comfortable 3D experience for the end consumer, common production errors must be detected and corrected. Vertical disparity is one of these distortions and is caused by improper stereo-camera setup. This paper aims at identifying the possible rotational and placement errors that cause the vertical disparity. An estimation of these errors is necessary to produce good quality 3D content. According to my knowledge, there exists no method to identify rig alignment errors without the knowledge of camera setup parameters and this work is the first step in that direction. Feature point detection has proven to be an interesting approach to find vertical disparity present in the given stereo image pair. In this work feature extraction techniques such as SIFT, SURF and Harris features are efficiently used to compute reliable and robust vertical disparity patterns. This paper classifies vertical disparity patterns according to rig errors. If the vertical disparity values are too small or ambiguous to be identified by pattern analysis, this paper uses graphical analysis that highlights the relationship between the total vertical disparity and the contribution of each possible error to the total. Experimental results show that the proposed approach identifies the reason behind the presence of vertical disparity of a stereo image pair.

Keywords
Stereoscopic 3D, feature matching, vertical disparity, epipolar constraint, RANSAC.

1 INTRODUCTION

Experiencing 3D is one of the most important perception features of human kind. Several depth cues are used by human visual system to perceive depth and to experience 3D. Stereoscopic video systems simulate the human binocular vision [Sammons92]. Binocular depth cues are observed by both eyes. They include retinal disparity in the two views of a scene (one from each eye) resulted from inter-ocular distance, and convergence due to the inward movement of both eyes to focus at a point of interest. If the cues are not reproduced properly in the stereo content, it results in a bad user experience [Perek16, Tam11, Hodges91]. Poor quality of stereoscopic video content results in visual fatigue [IJsselsteijn00, ATSC11]. Consequences of visual fatigue on human 3D perception might be eye strain, watery eyes, nausea, head aches, focusing difficulty and blurred vision. Therefore, stereographers and camera teams have to take into account a variety of conditions, guidelines and rules right from the beginning of the production chain to create good stereoscopic video content [Zilly11, Knorr12]. It includes accurate rigging and calibration of the stereo cameras, good adjustment and matching of electronic and optical camera parameters as well as the adaption of the stereo baseline to the depth structure of the scene content.

When there is an error in the alignment of the stereo rig setup, vertical disparity can be observed in the image pair. Vertical disparity is defined as the vertical displacement between corresponding pixels in the left and right images. Vertical disparity in the stereo content is result of distortions. There are other distortions such as keystone distortion, depth plane curvature, barrel distortion and pin cushion distortion caused by other factors. One of the results of vertical disparity is vertical parallax which can cause eye strain, visual fatigue, confusion and loss of stereopsis. The alignment errors in the stereo rig setup include rotational errors, translational errors or both. The rotational errors include relative roll, relative pan and relative tilt between the cameras. The translational errors include relative vertical shift and relative forward or backward shift between the cameras. One more error is zoom difference between the two cameras. It is caused when the lenses used in the two cameras have different focal length. Converged
(toed-in) camera setup and parallel camera setup are the
two stereoscopic camera configurations used in 3D cap-
turing [Woods93]. In the toed-in setup, the two cameras
are angled towards each other resulting in a finite con-
vergence distance (the distance at which the two cam-
era images coincide in the centre of the stereoscopic
display). The setup causes keystone distortion and ver-
tical disparities which can be corrected using stereo im-
age rectification [Faugeras93] during post processing.

In the parallel setup, the optical axes of the two cam-
eras are parallel so the convergence distance is infinity.

Horizontal Image Translation (HIT) is done during post
processing to place the point of convergence wherever
desired. In the ideal case, the parallel setup does not
produce vertical disparity.

This paper mainly focuses on vertical disparity caused
by stereo rig alignment errors in the parallel setup and
proposes a method to identify rig alignment errors
based on vertical disparity in the given stereo image
pair. These errors include relative rotation, relative
translation and relative zoom difference. It is assumed
that no lens distortion is present in the given image
pair and the image pair is not rectified after capturing.

The steps include reliable estimation of the vertical
disparity and finding what kind of stereo rig error is
causiing this vertical disparity. The main contributions
are (1) classification of possible vertical disparity patterns
according to the stereo rig errors, and (2) identification of rig alignment error using vertical

disparity map.

The paper is organised as follows. Section 2 briefly
presents the state of the art on stereo rig alignment error
identification. Section 3 presents the required mathe-
matical background of stereo capturing. Section 4 pro-
poses the vertical disparity based alignment error iden-
tification method. Section 5 presents the experimental
results. Finally, Section 6 concludes the paper.

2 STATE OF THE ART

For a calibrated system, there is a lot of work pre-
sented in the literature to identify rig alignment er-
ors. Behrooz Kamgar-Parsi et al. [Kamgar-Parsi88]
computed small rotational errors caused by mechani-
cal difficulties in the stereo rig setup. They proved
that an error of 0.5 degrees in pan angle can result in
an error of 4 meters in depth estimation which empha-
sizes the importance of accurate alignment. They
assumed that both cameras are calibrated, i.e., the focal
length, spatial digitisation of image plane and lens dis-
tortion are known. Behrooz’s work also emphasizes
that the points selected from a specific region cannot
cover all properties of relative rotations. Hence, it is
important to select point correspondences from differ-
ent regions of the image. In case of uncalibrated camera
systems, identification of stereo rig error is still a chal-
lenge. Richard I. Hartley [Hartley92] and Frederik Zilly
et al. [Zilly10] proposed methods to compute the align-
ment errors from an uncalibrated stereo by using some
assumptions on intrinsic parameters. Hartley did a lot
of research in multiple view geometry [Hartley04]. He
proposed two algorithms to calculate the relative place-
ment of the cameras in the stereo rig [Hartley92]. One
algorithm deals with the calibrated camera system and
the other deals with the uncalibrated camera system. In
case of calibrated system, it is assumed that everything
is known except relative positions and relative ori-
entations. In case of uncalibrated system, it is assumed
that the two cameras are of the same type (having the
same intrinsic parameters and the same lenses) and the
principle point is at the image centre, the focal lengths
of the two cameras, their relative placement and ori-
entation are determined. This approach results in imagi-
nary values when the assumptions fail to satisfy. Hence
the result is not guaranteed in all cases. Frederik Zilly
et al. proposed a new method [Zilly10] that computes
rectifying homographies using sparse stereo correspon-
dences from an uncalibrated stereo. This paper uses
this approach to classify the disparity patterns caused
by different alignment errors and to identify the stereo
rig errors.

3 MATHEMATICAL BACKGROUND

Consider a parallel stereo setup shown in Figure 1 and
2. The line joining the optical centres \(C_l\) and \(C_r\)
of the left and right cameras respectively is called base-
line. Consider a point \(P\) in space whose perspective
projection is at point \(p_l = (u_l, v_l)\) on the image plane
\(I_l\) of left camera and at point \(p_r = (u_r, v_r)\) on the
image plane \(I_r\) of right camera. The plane connecting \(P, \)
\(C_l\) and \(C_r\) is called epipolar plane. The plane inter-
sects the image planes \(I_l\) and \(I_r\) in lines called epipolar
lines. Let \(P_x = [X, Y, Z]\) represent the same point \(P\)
in the right camera coordinate system. The perspective
projection of the point \(P\) on the image plane \(I_l\) is at
\((f_lX_l/Z_l, f_lY_l/Z_l)\) where \(f_l\) is the focal length of the
right camera. To obtain this image location in terms of
pixel values, intrinsic parameters of the camera have to

\[
\begin{align*}
P &= [X, Y, Z]\\
P_l &= (u_l, v_l)\\
P_r &= (u_r, v_r)\\
\end{align*}
\]
be considered. They include principle point [Clarke98], pixel dimensions [Lenz88] and skew coefficient describing non-rectangular pixels. If \((c_{1x}, c_{1y})\) is the pixel coordinate of the principle point, \(s_x^l\) and \(s_y^l\) represent effective pixel sizes along the horizontal and vertical directions of the image respectively [Lenz88], the 3D point \(P_r\) to 2D point \(p_r\) transformation is [Hartley04]

\[
\begin{align*}
   u_r &= (f_x x_r / z_r) / s_x^l + c_{1x}, \\
   v_r &= (f_y y_r / z_r) / s_y^l + c_{1y}.
\end{align*}
\]

(1)

An upper triangular matrix describing these intrinsic properties of the right camera holding the parameters such as principle point \((c_{1x}^r, c_{1y}^r)\), pixel dimensions \(f_x^r, f_y^r\) and \(s_x^r, s_y^r\), and skew \(s_r\) is defined as

\[
K_r = \begin{bmatrix}
   f_x^r / s_x^r & s_r & c_{1x}^r \\
   0 & f_y^r / s_y^r & c_{1y}^r \\
   0 & 0 & 1
\end{bmatrix}.
\]

(2)

Assuming zero skew, equation 1 can be written as

\[
\begin{bmatrix}
   p_r \\
   1
\end{bmatrix} = K_r \hat{P}_r,
\]

(3)

where \(\hat{P}_r = [X_r / Z_r, Y_r / Z_r, 1]^T\) represents the projection on the right image plane (the superscript \(\hat{\cdot}\) refers to the transpose operation). Similarly, \(p_l\) is image point \((u_l, v_l)\) of the same world point \(P\) whose 3D location is \(P_l\) in the left camera coordinate system and \(\hat{P}_l\) represents the projection on the left image plane.

Epipolar geometry describes the geometric relationship between a pair of stereo images [Boufama95]. Let the left camera coordinate system be the reference coordinate system, \(R\) be the \(3 \times 3\) matrix (the rotation matrix) defining the relative rotation of the right camera with respect to the left camera and \(T\) be the \(3 \times 1\) vector (the translation vector) defining the relative translation of the right camera with respect to the left camera. Then the relation between \(\hat{P}_r\) and \(\hat{P}_l\) is \(\hat{P}_r = R (\hat{P}_l - T)[\text{Andrea00}]\). From the plane equation of two vectors, the equation of epipolar plane is

\[
\begin{align*}
   (\hat{P}_l - T)^T (R^T \hat{P}_r) &= 0, \\
   (R^T \hat{P}_r)^T [T]_x \hat{P}_l &= 0.
\end{align*}
\]

(4)

Here \([T]_x\) is the skew-symmetric matrix of vector \(T\). From equation 3 and equation 4,

\[
\begin{bmatrix}
   p_r \\
   1
\end{bmatrix} F \begin{bmatrix}
   p_l \\
   1
\end{bmatrix} = 0,
\]

(5)

where \(F = (K_r^{-1})^T R [T]_x K_l^{-1}\) is called fundamental matrix between the pair of stereo images. Equation 5 is called Epipolar constraint. This maps points from one image to corresponding epipolar lines in the other image. The fundamental matrix \(F\) encodes the information of both extrinsic (\(R\) and \(T\)) and intrinsic parameters (\(K_l\) and \(K_r\)). A stereo camera system is said to be calibrated if its intrinsic parameters and the extrinsic parameters are known. If some of the parameters are known it is called semi calibrated or partially calibrated. If nothing about these parameters is known the system is an uncalibrated system.

In the ideal parallel stereo setup (parallel Z-axis of the cameras) case, the X-axis of each camera is colinear with the base line and hence projects the point \(P\) on to the two image points \(p_l\) and \(p_r\) on the same scan line making the vertical disparity \(v_r - v_l\) zero. The distance between the point \(P\) and the stereo camera is proportional to the horizontal disparity \(u_r - u_l\).

4 PROPOSED METHOD

Let \(\alpha_x\), \(\alpha_y\) and \(\alpha_z\) are the relative rotation angles of the right camera with respect to the left camera around \(X\), \(Y\) and \(Z\) axes respectively, and \(t_x\), \(t_y\) and \(t_z\) are the relative translation of right camera with respect to the left camera along \(X\), \(Y\) and \(Z\) axes respectively. This work assumes that (1) the relative rotation angles are small, i.e. \(\alpha_x << 1\), \(\alpha_y << 1\) and \(\alpha_z << 1\), and the rotation matrix \(R\) can be approximated as

\[
\hat{R} = \begin{bmatrix}
   1 & -\alpha_x & \alpha_y \\
   \alpha_x & 1 & -\alpha_z \\
   -\alpha_y & \alpha_z & 1
\end{bmatrix},
\]

(6)

(2) relative translation in \(Y\) and \(Z\) directions are small compared to that in \(X\)-direction (baseline length), i.e. \(\hat{t}_y = t_y / t_z << 1\) and \(\hat{t}_z = t_z / t_y << 1\), so without affecting Equation 5, \(T\) can be taken as \(t = [t_x, t_z]^T\) and

\[
[\hat{t}]_x = \begin{bmatrix}
   0 & \hat{t}_z \\
   \hat{t}_z & 0 & \hat{t}_y \\
   -\hat{t}_y & 1 & 0
\end{bmatrix},
\]

(7)
4.1 Classification of disparity patterns

It has been shown that without the knowledge of the exact focal length of each camera in the stereo rig, computation of exact relative angles and relative positions is not possible. Moreover, the focal length changes with changes in zoom. However, irrespective of the focal length value used to capture a scene, properties of the vertical disparity map are unique for each kind of error. Using equation 8, different stereo rig errors are classified according to the properties of the vertical disparity map. These properties can be used to identify the error sources based on the vertical disparity pattern present in the given stereo image pair. This section discusses the vertical disparity map properties for all possible stereo rig errors.

4.1.1 Relative translation in Y-direction

If the two cameras in a stereo setup are placed at different heights (X-axes of the cameras are separated vertically as shown in Figure 4), one image is shifted upward or downward relative to the other depending on the vertical shift direction. If the right camera is placed above the left camera, all left image features are moved downwards in the right image. See Figure 5(a) where the dashed border shows the right image and the thick border shows the left image. The vertical disparity is represented using arrows. Each arrow’s head and tail locate at the row value (y-coordinates) of corresponding feature pixel in the right and the left images respectively. Similarly when the right camera is placed below the left camera, all left image features are moved upwards in the right image as shown in the Figure 5(b).

If vertical translation is the only error present in a stereo rig, equation 8 can be reduced to

\[ v_r - v_l = \tilde{f}_r (u_r - u_l), \]

where \( u_r - u_l \) is horizontal disparity representing the depth. The properties of this error are: (1) vertical disparity is directly proportional to the horizontal disparity (depth) magnitude and (2) the value of vertical disparity is the same for all points located at a particular depth. These vertical disparity patterns can be seen in Figure 6.

Figure 3: Block diagram showing the steps in the proposed classification based rig alignment error identification approach.

Figure 4: Cameras are at different heights

(3) pixels are square shaped meaning the skew of the cameras is zero (i.e. \( s_r = s_l = 0 \), \( s_r^2 = s_l^2 \) and \( s_r^4 = s_l^4 \), (4) focal lengths’ difference \( \alpha_f \) is very small such that \( \tilde{f}_r = (1 - \alpha_f) / \tilde{f}_l \) where \( \tilde{f}_l = f_l / s_l^2 \) and \( \tilde{f}_r = f_r / s_r^2 \) are the focal lengths of left and right cameras (in pixels) respectively.

By substituting the assumptions in Epipolar constraint (equation 5) and ignoring the second and higher order terms, the total vertical disparity \( v_r - v_l \) corresponding to the point \( P \) can be written as a function of \( (\tilde{f}_l, \alpha_f, \alpha_x, \alpha_y, \alpha_z, \tilde{t}_y, \tilde{t}_z) \) as [Zilly10]

\[
v_r - v_l = \tilde{f}_l (u_r - u_l) + \frac{\tilde{f}_l}{\tilde{f}_r} (u_l v_r - u_r v_l) + \alpha_f v_r \quad \text{Y-shift} \]

\[
\alpha_f \tilde{f}_r - \frac{\alpha_x}{\tilde{f}_r} v_l v_r + \frac{\alpha_x}{\tilde{f}_l} u_l v_l + \alpha_z \tilde{t}_z \quad \text{roll} \]

\[
-\frac{\alpha_f}{\tilde{f}_r} v_r + \frac{\alpha_f}{\tilde{f}_r} v_l \quad \text{pan} \]

(8)

It can be seen from the equation that different errors produce different vertical disparity patterns in the image. The following section shows the disparity patterns caused by some errors. If stereographers or quality analysts are familiar with these patterns, they can identify the error causing the pattern in a given stereo image pair. Vertical disparity map can be obtained by using sparse features matching algorithms such as SIFT [Lowe04], SURF [Bay08] and Harris features. Figure 3 shows the block diagram of the proposed pattern classification based error identification approach.
Figure 5: Placement error along Y-axis of left camera (thick rectangle) and right camera (dashed rectangle), and the corresponding vertical disparity pattern. The distance between dashed lines in pattern represents the magnitude of disparity.

Figure 6: Vertical disparity patterns at different depths ($u_r - u_l$) when the cameras are at different heights ($\hat{t}_y = 0.1$).

4.1.2 Relative translation in Z-direction

If one camera is placed a little forward or backward (along Z-axis) to the other camera, it results in vertical disparity. Figure 7 helps to visualise this type of setup error. If focal lengths of the two cameras are the same, a disparity pattern as shown in Figure 8(a) and Figure 8(b) can be observed. If translation in Z-direction is the only error present in a stereo rig, equation 8 can be reduced to

$$v_r - v_l = \frac{\hat{t}_z}{f_l} (u_r v_r - u_l v_l). \quad (10)$$

The properties of this error are: (1) as the vertical disparity depends on $u_r$ and $u_l$ values, the disparity map varies with the depth, (2) the value of vertical parallax increases or decreases from the image top edge to the bottom edge, (3) the vertical disparity decreases as the depth increases and (4) vertical parallax is symmetrical about the row at the image centre. These vertical disparity patterns can be seen in Figure 9. Disparity due to translation in Z-direction decreases with increasing depth and becomes zero at infinity (i.e. at depth=0).

4.1.3 Zoom difference

Another possible alignment error is focal length error. If two cameras use slightly different focal length values (different zoom), it shows similar vertical disparity pattern as previous (shift in Z-direction). See Figure 10(a) and Figure 10(b) to visualise the effect. If relative zoom is the only error present in a stereo rig, equation 8 can be reduced to

$$v_r - v_l = \alpha_f v_r. \quad (11)$$

The properties of relative zoom error are: (1) the disparity map is independent of depth information. This is the key point that differentiates this error from translational error in Z-direction where the vertical parallax varies with depth, (2) the amount of vertical parallax increases or decreases from the image top to the bottom and (3) The magnitude of vertical disparity is constant in a row. These vertical disparity patterns can be seen in Figure 11.

4.1.4 Relative rotation about X-axis (Tilt)

If one camera is rotated about X-axis relative to the other (as shown in Figure 12), features move up or down in the other camera’s view. See Figure 13(b) and
Figure 10: Zoom error representation with left camera (thick rectangle) and right camera (dashed rectangle), and the corresponding vertical disparity pattern. The distance between dashed lines in pattern represents the magnitude of disparity.

(a) Positive zoom difference  
(b) Negative zoom difference

Figure 11: Vertical disparity patterns at different depths \((u_r - u_l)\) when there is different zoom in camera pair \(\alpha_f = 0.1\).

Figure 13(a) to visualise the effect. If relative tilt angle is the only error present in a stereo rig, equation 8 can be reduced to

\[
v_r - v_l = -\alpha_f \hat{f}_l - \frac{\alpha_x}{\hat{f}_l} v_l v_r. \tag{12}
\]

The properties of the relative tilt error are: (1) the disparity map is independent of depth information, (2) the amount of vertical parallax increases or decreases from the image centre to the upper and to the lower edges and (3) if the focal length is very large, the component \(-\frac{\alpha_x}{\hat{f}_l} v_l v_r\) can be ignored. In that case vertical disparity is uniform throughout the image. These vertical disparity patterns can be seen in Figure 14.

4.1.5 Relative rotation about \(Y\)-axis (Pan)

If one camera is rotated about \(Y\)-axis relative to the other (as shown in Figure 15), it is called toed-in / toed-out setup. It results in key stone distortion. In this case, the vertical disparity can be observed clearly at the corners than at the image centre. Opposite corners exhibit similar disparity pattern while adjacent corners show opposite directions. See Figure 16(b) and Figure 16(a) to visualise the effect. If relative pan angle is the only error present in a stereo rig, equation 8 can be reduced to

\[
v_r - v_l = \frac{\alpha_y}{\hat{f}_l} u_r v_r. \tag{13}
\]

The properties of the relative pan angle are: (1) the disparity map is independent of depth information, (2) the direction and magnitude of the vertical parallax is the same on the opposite corners, (3) neighboring corners
Figure 16: Relative rotation along Y-axis error of left camera (thick rectangle) and right camera (dashed rectangle) and the corresponding vertical disparity pattern. The distance between dashed lines in pattern represents the magnitude of disparity.

Figure 17: Vertical disparity patterns at different depths when the relative pan angle between the cameras $\alpha_p = 0.1$.

The properties of the relative pan angle are: (1) the disparity map is independent of depth, (2) the value of vertical parallax increases or decreases from the left edge to the right edge of the image (in a row) and (3) the vertical disparity pattern is symmetrical about central column of the image. These vertical disparity patterns can be seen in Figure 17.

4.1.6 Relative rotation about Z-axis (Roll)

If one camera is rotated about Z-axis relative to the other (as shown in Figure 18), the features on one side of the central column move up and on the other side move down. See Figure 19(a) and Figure 19(b) to visualise the effect. If relative roll angle is the only error present in a stereo rig, equation 8 can be reduced to

$$v_r - v_l = \alpha_z u_r.$$  \hfill (14)

Figure 18: Cameras with relative roll angle.

Figure 19: Relative rotation along Z-axis error of left camera (thick rectangle) and right camera (dashed rectangle) and the corresponding vertical disparity pattern. The distance between dashed lines in pattern represents the magnitude of disparity.

Figure 20: Vertical disparity patterns at different depths when the relative roll angle between the cameras $\alpha_z = 0.1$.

The properties of the relative roll angle are: (1) the disparity map is independent of depth information, (2) the value of vertical parallax increases or decreases from the left edge to the right edge of the image (in a row) and (3) the vertical disparity pattern is symmetrical about central column of the image. These vertical disparity patterns can be seen in Figure 20.

4.1.7 Multiple errors (Tilt and Roll)

The patterns can be classified also for combination of the errors. Here, only the tilt and roll combination is presented. If relative tilt and relative roll are present in a stereo rig, it results in disparity patterns as shown in Figure 21. Equation 8 can be reduced to

$$v_r - v_l = \alpha_z u_r - \hat{f}_l \alpha_x - \frac{\alpha_x}{\hat{f}_l} v_l v_r.$$  \hfill (15)

The properties of the resulting vertical disparity map are: (1) the disparity map is independent of depth (prop-

Figure 21: Resulting vertical disparity pattern when relative rotations along X-axis (tilt) and Z-axis (roll) are present simultaneously. The distance between dashed lines in pattern represents the magnitude of disparity.
If there are multiple error sources present simultaneously, it is difficult to identify them from the classified patterns using the proposed classification based approach (shown in Figure 3). Moreover, there can be false matches that create false disparity at certain locations in the map. To overcome these challenges we propose a mathematical approach that solves the unknowns in equation 8. As there are 7 unknowns, namely \( \alpha_x, \alpha_y, l_z, f_l, \alpha_z, r_t, z_0 \) are needed to solve for the unknowns. To increase the reliability of the solution in the presence of false feature matches, RANSAC algorithm [Fischler81] is used. In this work, RANSAC uses a threshold on Sampson distance to identify false matches. After filtering the false matches, the vertical disparity contributions of each error source are computed for each true match. The error sources that highly influence the total vertical disparity are identified and are considered as the major rig-alignment errors. Figure 22 presents the block diagram of the approach.

5 RESULTS

This section presents the experimental results of the proposed approach. According to my knowledge, this is the first work in this direction and there exists no other work to compare the performance. Consider a stereo image pair (shown in Figure 24) captured using an uncalibrated stereo rig i.e. unknown intrinsic and extrinsic parameters of the cameras. Sparse feature matches are computed using SIFT, SURF and Harris features. Figure 25 show the results of stereo matching. It can be observed that there exists false matches. After applying the RANSAC on equation 8, the false matches are filtered and the unknowns \((\alpha_x, l_z, f_l, \alpha_y, r_t, \alpha_z, z_0)\) are calculated. Figure 26 shows the results of stereo matching after removing false matches. Figure 27 shows the vertical disparity at the true matched features drawn on the left image. It can be observed that the disparity increases from right to left and in the entire image the disparity is upwards. Based on the error to vertical disparity pattern classification, the pattern in the results is close to the one caused by tilt and roll (see Figure 21). To confirm this, a graph that shows the individual contributions of the possible error sources is drawn and the errors that highly influence the total vertical disparity are identified. Figure 28 shows the contributions of each possible error source. It can be observed that tilt and roll are the major alignment errors influencing the total vertical disparity.

\[
\begin{align*}
\alpha_x & : \text{relative shift in Z-direction}, \\
\alpha_y & : \text{focal length of left camera in pixels}, \\
r_t & : \text{relative shift in Y-direction}, \\
\alpha_z & : \text{relative tilt}, \\
r_f & : \text{relative pan}, \\
z_0 & : \text{relative roll}, \\
f_l & : \text{focal length of left camera in pixels}, \\
n & : \text{unknown intrinsic and extrinsic parameters of the cameras}. \\
\end{align*}
\]
6 CONCLUSION

Depending on the type of rig alignment error the vertical disparity is different in different regions of the stereo images. This paper presents a classification of possible disparity patterns to make stereographers or quality analysts identify the error source through visual inspection of the patterns in the vertical disparity map. SIFT, SURF and Harris feature detection algorithms are used to produce the map. This paper also proposed a mathematical approach that computes the contribution of each possible error to the vertical disparity. The advantage is that the approaches do not require any calibration information about the stereo rig. The possible future works include automatic adjustment of rig alignment based on the computed vertical disparity contributions and evaluation of the results using controlled experiments on a real stereo rig setup.

7 REFERENCES


