Stereoscopic Correspondence by Applying Physical Constraints and Statisitical Observations to Dissimilarity Map

T.-Y. Chao[a], S.-J. Wang[b] and H.-M. Hang[c]

[abc] Dept. of Electronics Engineering,
National Chiao Tung University,
Hsinchu, Taiwan 30010, ROC

ABSTRACT

To deal with the correspondence problem in stereo imaging, a new approach is presented to find the disparity path on a newly defined dissimilarity map (DSMP). Based on an image formation model of stereo images and some statistical observations, two constraints and four assumptions are adopted. In addition, some heuristic criteria are developed to define a unique solution. All these constraints, assumptions and criteria are applied to the DSMP to do correspondence. At the beginning, Epipolar Constraint, Valid Pairing Consraint and Lambertian Surface Assumption are applied to DSMP to find the Low Dissimilarity Zones (LDZ's). Then, the Opaque Assumption and Minimum Occlusion Assumption are applied to LDZ's to obtain the admissible LDZ sets. Finally, the Depth Smoothness Assumption and some other criteria are applied to the admissible LDZ sets to choose the final solution. The focus of this paper is to find reasonable constraints and assumptions and then properly convert these constraints and assumptions into executible descriptions on DSMP. In addition to its capability in estimating occlusion accurately, this approach may work well even when the commonly used monotonic-ordering assumption is violated. The simulation results show that occlusions can be properly handled and the disparity map can be calculated with a fairly high degree of accuracy.

Keywords: Epipolar Constraint, Valid Pairing constraint, Lambertian Surface Model, Opaque Assumption, low dissimilarity zone, disparity, depth

1. INTRODUCTION

The correspondence of stereo images is an important but difficult key technique in geometric stereo. Many papers[1–5] have discussed the correspondence issue and their methods can be roughly classified into two classes: feature-based methods and intensity-based methods. Feature-based methods utilize the features in the image pair to generate a more accurate but looser disparity map. Intensity-based methods utilize the intensity information in the images to produce a denser disparity map. Since feature-based methods only utilize partial information of the image pair, we adopt an intensity-based approach in this paper to fully utilize all the information of the image pair. Our method is called a dissimilarity map (DSMP) based method. DSMP is a newly defined map to measure the dissimilarity between stereo image pairs. As shown in Figure 1, this DSMP-based method applies geometic constraints, photometic assumptions, and statistical assumptions on DSMP to do correspondence. These geometric constraints contain Epipolar Constraint and Valid Pairing Constraint; these photometric assumptions contain Lambertian Surface Model Assumption and Opaque Assumption; and these statistical assumptions contain Minimum Occlusion Assumption and Depth Smoothness Assumption. In this paper, the above constraints and assumptions are converted into related operations on DSMP. The basic theory is first developed assuming the image has infinite resolution and noise is absent. Then the noise interference and the projective distortion due to the finite resolution of a real image is considered and some modifications on these related operations are made.

Further author information:
Tsi Yi Chao: E-mail: u8411824@cc.nctu.edu.tw
S.-J. Wang: E-mail: shengjyh@cc.nctu.edu.tw
H.-M. Hang: E-mail: hmhang@cc.nctu.edu.tw

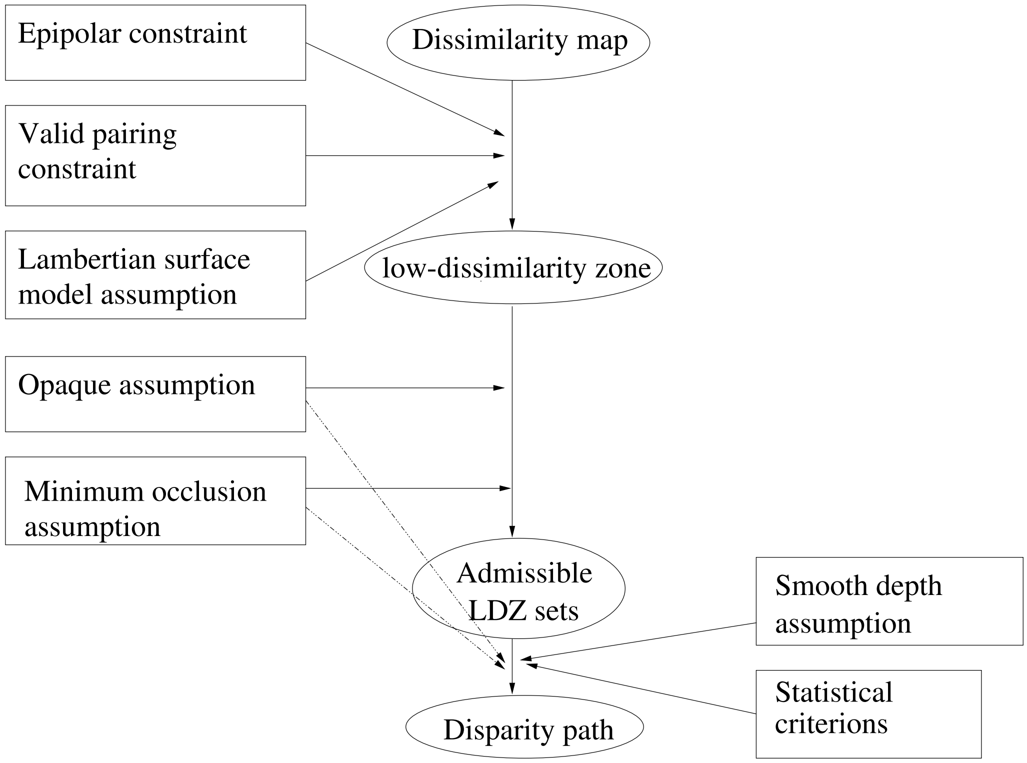


Figure 1. The flow diagram of DSMP method

2. GEOMETRIC MODEL

2.1. Projection Model

The projection mechanism of a camera can be modeled as a perspective projection as shown in Figure 2. In this figure, there are two coordinate systems, (Xi, Yi) and (x, y, z). (Xi, Yi) represents the coordinate of a projected point on the image plane. By using triangular similarity, we have

$$\frac{Xi}{f} = -\frac{x}{z} \quad and \quad \frac{Yi}{f} = -\frac{y}{z}.$$

If we define the new coordinates to be:$Xm = -Xi$ and $Xm = -Yi$, then

$$\frac{Xm}{f} = \frac{x}{z} \quad and \quad \frac{Ym}{f} = \frac{y}{z}. \tag{1}$$

The above transformation simply translates the original image to its mirror image with respect to the lens center (from (Xi, Yi) to (Xm, Ym) as shown in Figure 2) . That is, we deal with the mirror image, instead of the image plane, to simplify the transformation form.

2.2. Stereoscopic Camera System Model

A stereoscopic camera system[1] is shown in Figure 3. The line B connecting two projection centers of projection is called a **baseline**. In addition, f represents the focal length and θ represents the convergence angle $(0^0 \leq \theta < 180^0)$. In Figure 3, we have three coordinates: left, center, and right. Equations (2) is derived based on these coordinates to relate the left coordinates with right coordinates.[1]

$$\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \times \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} + \begin{bmatrix} -B\cos\frac{\theta}{2} \\ 0 \\ B\sin\frac{\theta}{2} \end{bmatrix}. \tag{2}$$

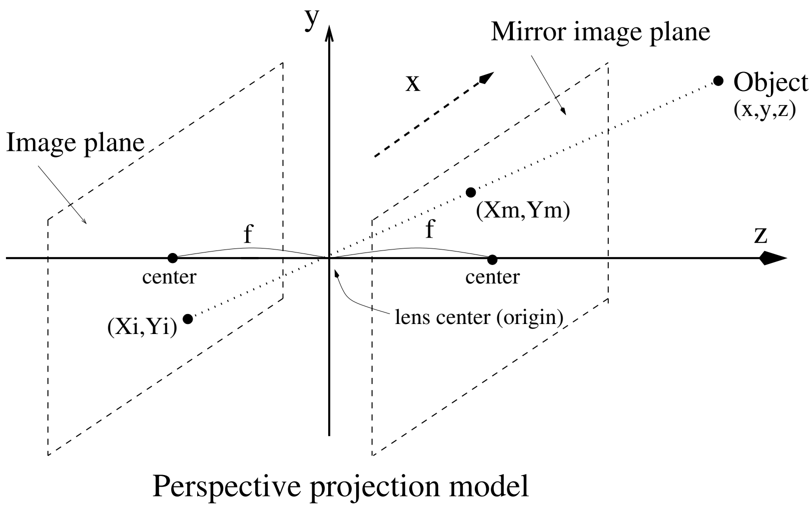


Figure 2. Persepective Projection model and It's Coordinate Translation

Combining (1) and (2), we can derive

$$z_l = \frac{B(f\cos\frac{\theta}{2} + Xm_r \sin\frac{\theta}{2})}{f\sin\theta + \frac{\sin\theta}{f}Xm_l Xm_r + \cos\theta(Xm_l - Xm_r)}. \tag{3}$$

$$z_r = \frac{B(f\cos\frac{\theta}{2} - Xm_l \sin\frac{\theta}{2})}{f\sin\theta + \frac{\sin\theta}{f}Xm_l Xm_r + \cos\theta(Xm_l - Xm_r)}. \tag{4}$$

$$\tag{5}$$

Therefore, if the (Xm_l, Xm_r) pair in the conjugate epipolar lines is known, z_l and z_r (depth) can be found by using (3) and (4). Equation (5) represents the epipolar line constraint which will be described later. To find depth, we match (Xm_l, Xm_r) first and then substitute (Xm_l, Xm_r) into Equations (3) to calculate the depth information. Usually, we use z_l to represent the depth information. In addition, the disparity is defined as $D_m = (Xm_l - Xm_r)$. As $\theta = 0$, Equation (3) degenerates to

$$z_l = \frac{Bf}{Xm_l - Xm_r} = \frac{Bf}{D_m}. \tag{6}$$

2.3. Camera's Coordinate

The unit of the camera's coordinate is pixel unit length. This unit is usually different from the unit used to measure the depth or the baseline width. Hence, we need a scaling factor to match the coordinates between the camera's coordinate (Xc, Yc) and the coordinate (Xm, Ym). Suppose the camera's resolution is S $(pixel/unit)$, then we can get

$$Xc = \frac{Xm}{S}, \quad Yc = \frac{Ym}{S}. \tag{7}$$

3. GEOMETRIC CONSTRAINTS

The geometric constraints include Epipolar Constraint and Valid Pairing Constraint. The geometric constraints are derived based on the calebration of the left and right cameras. Any matching point must obey the geometic constraint.

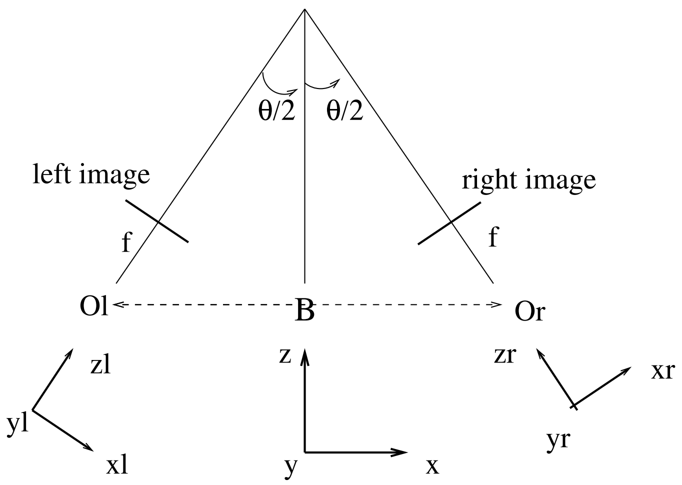


Figure 3. The stereoscopic camera geometry

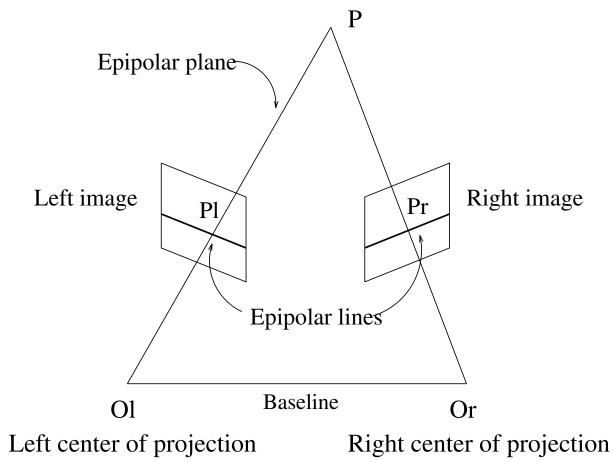


Figure 4. Epipolar plane and epipolar lines

3.1. Epipolar line constrarint

Consider a pair of images. Let Pl and Pr denote the projections of a real world 3-D point p onto the left and right image planes, respectively. Let Ol and Or denote the corresponding centers of the pinhole projection model. A plane that passes through these two centers of projection is called an **epipolar plane**. Each epipolar plane intersects with both mirror image planes and these intersections are called conjugate (or corresponding) **epipolar lines**[6] (Figure 4).

3.2. Valid pairing constraints

Valid pairing constraint describes the constraint on Xm_l and Dm under the condition that the object at Xm_l and Dm can be seen by both the left and right cameras with a complete viewing angle.

Substitute $D_m = (Xm_l - Xm_r)$ into Equations (3) and (4), we get

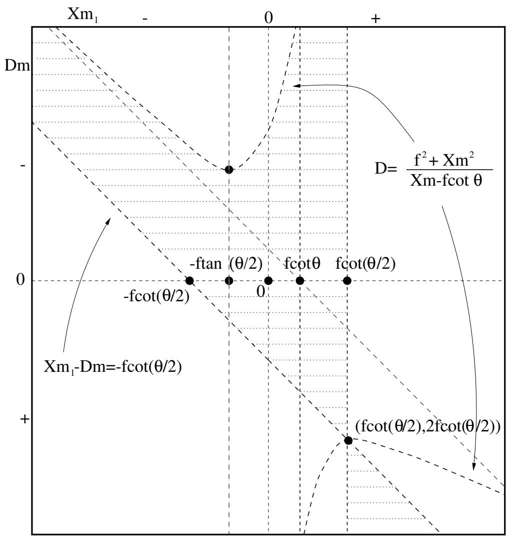


Figure 5. Valid pairing constraint

$$z_l = \frac{B(f\cos\frac{\theta}{2} + (Xm_l - D_m)\sin\frac{\theta}{2})}{f\sin\theta + \frac{\sin\theta}{f}Xm_l(Xm_l - D_m) + D_m\cos\theta}, \tag{8}$$

$$z_r = \frac{B(f\cos\frac{\theta}{2} - Xm_l\sin\frac{\theta}{2})}{f\sin\theta + \frac{\sin\theta}{f}Xm_l(Xm_l - D_m) + D_m\cos\theta}, \tag{9}$$

where z_l and z_r must be positive due to the camera's structure. Hence, we can derive the restriction on Dm as

$$\begin{cases} (Xm_l - Dm) > -f\cot\frac{\theta}{2} \;\; and \;\; Xm_l < f\cot\frac{\theta}{2} & \Rightarrow \;\; (f\sin\theta + \frac{\sin\theta}{f}Xm_l(Xm_l - D_m) + D_m\cos\theta) > 0 \\ (Xm_l - Dm) < -f\cot\frac{\theta}{2} \;\; and \;\; Xm_l > f\cot\frac{\theta}{2}, & \Rightarrow \;\; (f\sin\theta + \frac{\sin\theta}{f}Xm_l(Xm_l - D_m) + D_m\cos\theta) < 0. \end{cases}$$

Reorganize the above constraints on Dm, we get

$$\begin{cases} (Xm_l - Dm) > -f\cot\frac{\theta}{2} \;\; and \;\; Xm_l < f\cot\theta, & \Rightarrow \;\; Dm > \frac{Xm_l^2+f^2}{Xml-f\cot\theta} \\ (Xm_l - Dm) > -f\cot\frac{\theta}{2} \;\; and \;\; f\cot\theta < Xm_l < f\cot\frac{\theta}{2}, & \Rightarrow \;\; Dm < \frac{Xm_l^2+f^2}{Xml-f\cot\theta} \\ (Xm_l - Dm) < -f\cot\frac{\theta}{2} \;\; and \;\; Xm_l > f\cot\frac{\theta}{2}, & \Rightarrow \;\; Dm > \frac{Xm_l^2+f^2}{Xml-f\cot\theta}, \end{cases} \tag{10}$$

where $0^0 < \theta < 180^0$ and these equations are drawn in Figure 5. When $\theta = 0^0$, the Valid Pairing Constraint will degenerate to $Dm > 0$. In short, all the 3-D points correspond to the points must follow Valid Pairing Constraint in SMP.

4. PHOTOMETRIC ASSUMPTIONS

The photometric assumptions include Lambertian Surface Assumption and Opaque Assumption. Lambertian Surface Assumption is based on the simple light model of lambertian surface. Opaque assumption is a very popular assumption and it holds in most cases.

4.1. Lambertian surface model assumption

The light model is very complex in nature and is difficult to model. Here, we use a simple light model called Lambertian surface model, which has the property that its radiance depends only on the incident illumination but not the viewing direction. Based on Lambertian surface model assumption, the reflected illumination from the same incident point should be the same. With noise interference, the reflected illumination from the same incident point should be about the same.

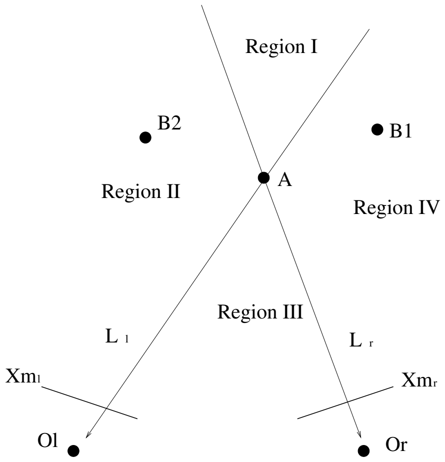


Figure 6. The slope analysis on the 3-D point under monotonic order

4.2. Opaque Assumption

4.2.1. Slope constraints under monotonic ordering

Here, we define Region I, II, III, and IV to be the regions seperated by lines L_l and L_r as shown in Figure 6. Point A is the crossing point of L_l and L_r. The line L_l passing through point both A and the left optical center, while the line L_r passing through both point A and the right optical center. Point $B1$ or $B2$ represents a neighborly unoccluded point which follows monotonic order points with A. The Opaque Constraint implies that points $B1$ and $B2$ cannot locate at lines L_l and L_r. If point $B1$ is in the right side of L_l (that is the next matching point), then this point should be in the rightside of L_r under monotonic ordering assumption. Hence, the possible region for $B1$ should be in Region IV. In the same way, if point $B2$ is in the left side of L_l (that is the previous matching point), then this point should be in the leftside of L_r under the monotonic ordering assumption (Region II). Then we know the possible region of $B1$ or $B2$ should be in the same side of lines L_l and L_r (Region II and IV).

Region II and IV can be represented by

$$\begin{cases} region\ II: & Xm_{Bl} < Xm_{Al} \ \ and \ \ Xm_{Br} < Xm_{Ar} \\ region\ IV: & Xm_{Bl} > Xm_{Al} \ \ and \ \ Xm_{Br} > Xm_{Ar}. \end{cases} \tag{11}$$

The corresponding region of Region II and IV in the (Xm_l,Dm) domain can be represented by

$$\begin{aligned} &\begin{cases} Dm_B = Xm_{Bl} - Xm_{Br} \\ Dm_A = Xm_{Al} - Xm_{Ar} \end{cases} \\ \Rightarrow \frac{Dm_B - Dm_A}{Xm_{Bl} - Xm_{Al}} &= \frac{(Xm_{Bl} - Xm_{Al}) - (Xm_{Br} - Xm_{Ar})}{Xm_{Bl} - Xm_{Al}} \\ &= 1 - \frac{Xm_{Br} - Xm_{Ar}}{Xm_{Bl} - Xm_{Al}} < 1. \end{aligned} \tag{12}$$

In addition, the unmatched parts under monotonic ordering assumption are not "true occlusion" and are called "pseudo occlusion" here.

4.2.2. Multi-layer description for 3-D points's disparity paths in (Xc_l, Dc_l) domain

Here, the opaque assumption is converted into an multi-layer description for the 3-D points's disparity paths in (Xc_l, Dc_l) domain and this description representing opaque assumption may be applied on dissimilarity map described later to find solutions.

Assume a stereo imaging system is set up. Given a set of 3-D points, we may find the corresponding 2-D points of these 3-D points on the dissmilarity map. We call these 2-D points the "disparity points". That is , the corresponding positions of the unoccluded 3-D points in the (Xc_l, Dc_l) domain are defined as the disparity points. Since we are not interested in describing these disparity points that have been occluded. To have an efficient description of the disparity information of these unoccluded disparity points, we adopt a multi-layer approach. At first, several sets of disparity points that follows the monotonic ordering assumption are extracted. Then, the disparity points corresponding to these 3-D points is described by a multi-layer architecture.

In Section 4.2.1, we know the neighboring unoccluded 3-D points under the monotonic ordering assumption fall in the range of Region II and IV (that is, the slope of the line from the current unoccluded disparity point to any neighoring unoccluded disparity point is smaller than 1). First, we pick up the leftmost disparity point in the dissimilarity map and name it $P1$. Then we search for the point $P2$, which is the leftmost point among all the disparity points that can follow the monotonic ordering assumption with $P1$. We repeat the same procedure to group more and more disparity points into the set such that any added disparity point follows the monotonic ordering assumption with all the disparity points in the set. After finishing the whole procedure, we call this set "*layer* 1 disparity path".

Under the Opaque Assumption, the residual unoccluded 3-D points must be on the intersection of the left, right pseudo occlusion in *layer* 1 disparity path and the region violating monotonic order with the *layer* 1 disparity path shown in Figure 7. We call these intersections $matching\ region_1$. In $matching\ region_1$, the *layer* 2 disparity path can be found and the residual disparity points must be on the intersection of the left, right pseudo occlusion in *layer* 1 and *layer* 2 disparity path, and the region violating monotonic order with the *layer* 1, 2 disparity path. By the same manner, we can finally describe the real disparity path by muti-layer monotonic ordering disparity path. Here, $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$ is defined as the intersection of the left, right pseudo occlusion in the disparity path *layer* $1,\ 2,\ 3....,\ (k-1)$ and the region violating monotonic order with the *layer* $1,\ 2,\ 3....,\ (k-1)$ disparity path. The *layer* k disparity path can be found on $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$ until no residual disparity path left.

In short, the above descriptions can be described as repeating the following three steps to caculate each-layer disparity path until there are no disparity points left.

- **step1:** Find the leftmost point (Xc_l, D) in the $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$.
- **step2:** Find the next leftmost point in region VI of the former matching point and repeat this process until the right boundary is reached in the $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$.
- **step3:** Finding $matching\ region_{1,\ 2,\ 3....,\ k}$.

, where $matching\ region_0$ is defined as the original stereo image pairs. In the above procedure, we know that the multi-layer disparity path and real disparity path is a one-to-one mapping. Hence, once this procedure is used to find disparity path, the opaque assumption is applied simultaneous. Therefore, this multi-layer structure can be used to substitute for the opaque assumption. The proof for that the $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$ are below the *layer* $(k-1),\ (k-2)$,...., 1 disparity path is proved as follows. This phenomenon can be used to reduce the searching area for each layer disparity path.

As illustrated in Figure 7, the set of $\{P1,\ P2,\ P3,\ P4,\ P5\}$ represents the *layer* 1 disparity path. Since $P2$ is the leftmost disparity point that follows the monotonic ordering assumption with $P1$, no disparity point exists in region $A1$, that can follow the monotonic ordering assumption with $P1$. Similarly, no disparity point exists in region $A2$, $A3$, and $A4$ that can follows the monotonic ordering assumption with respect to $P1$, $P2$, $P3$, $P4$, and $P5$. Hence, after having found the *layer* 1 disparity path, we repeat the same procedure for searching *layer* 1 path in the shaded region below *layer* 1 disparity path ($match\ regions_1$) in Figure 7 to find *layer* 2 disparity path. Then, $later-2$ disparity path is below *layer* 1 disparity path. By the same derivation for each layer disparity path, the $matching\ regions_{1,\ 2,\ 3....,\ (k-1)}$ should be below the *layer* $k-1$ disparity path. Hence, the *layer* k disparity path found in $matching\ regions_{1,\ 2,\ 3....,\ (k-1)}$ should be below the *layer* $k-1$ disparity path. Then, we can conclude that the $matching\ regions_{1,\ 2,\ 3....,\ (k-1)}$ are below *layer* $(k-1)$ disparity path (below *layer* $k-2$, below....below *layer* 2 below *layer* 1), that is below *layer* $(k-1),\ (k-2)$,...., 1 disparity path.

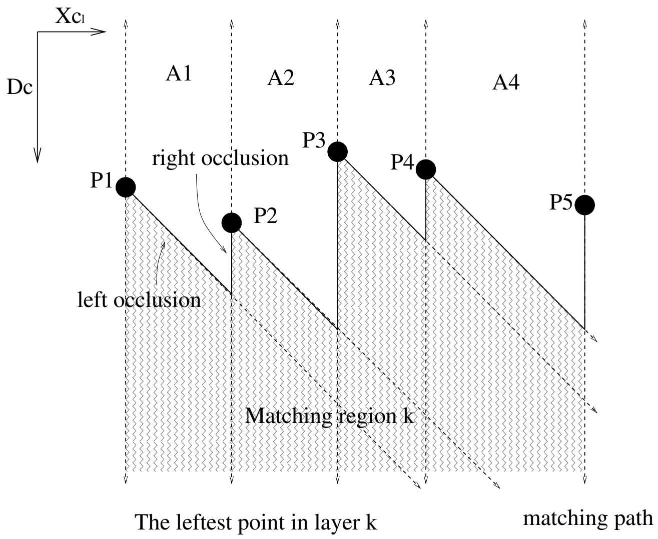


Figure 7. The matching on the layer 2

5. OCCLUSION ANALYSIS AND STATISTICAL ASSUMPTIONS

The matching points must have followed all the geometric constraints and photometric assumptions. Hence, all the possible matching paths follow the geometric constraints and photometric assumptions. Epilolar Constraint is used when conjugate epipolar lines are used to do correspodence. Valid Pairing Constraint is used to make sure that the object at (Xc_l, D_c) can be seen by both the left and right cameras. Lambertian Surface Assumption is used to suggest that the the corresponding intensity may come from the same 3-D point. Opaque Constraint is used to make sure that the uniqueness of the matching points along the line of sight. From the above description, we know that given the conjugate epipolar line, all the matching points following the geometric constraints and photometric assumptions could be the possible matching points. Thus, to find all the matching paths is just to apply the geometric constraints and photometric assumptions on the conjugate epipolar lines. Some statistical assumptions are used to choose the solution which has the highest probability in real world. These constraints and assumptions come from fundamental principles and they are crucial when we want to find the correct solution.

5.1. Occlusion analysis

Here, we define Region I, II, III, and IV as the regions seperated by lines L_l and L_r again, as shown in Figure 8. Point A is the crossing point of L_l and L_r. The line L_l passes through point A and the left optical center, and the line L_r passes through point A and the right optical center. Point B_i represents the possible neighboring matching point. Point B_i may fall in Regions I, II, III or IV. The location of point B_i can be classified into four categories.

- **Point A and B_i is continuous**
  This case means that both the Xm_l and Xm_r coordinates of A and B_i are continuous; there is no occlusion. Point B_1 in Figure 8 belongs to this case.
- **Point A and B_i is discontinuous, but the distance of B_i to L_l is infinitely small**
  This case means that only the Xm_l coordinate of A and B_i is continous; there is a "right occlusion". Point B_2 in Figure 8 belongs to this case.
- **Point A and B_i is discontinuous, but the distance of B_i to L_r is infinitely small**
  This case means that only the Xm_r coordinate of A and B_i is continuous, there is a "left occlusion". Point B_3 in Figure 8 belongs to this case.
- **Point A and B_i is discontinuous**
  This case means that both the Xm_l and Xm_r coordinates of A and B_i are discontinuous. Both a left occlusion

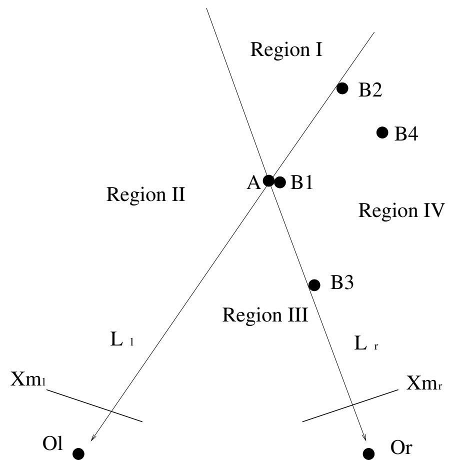


Figure 8. The occlusion analysis on the 3-D point

and a right occlusion occurs at the same time. We call this "cross-over occlusion". Point B_4 in Figure 8 belongs to this case.

5.2. Statistical assumptions

Statistical Assumptions contain Depth Smoothness Assumption and Minimum Occlusion Assumption. Depth Smoothness Assumption assumes that the depth of the scene which has a smooth intensity variation should be smooth. Minimum Occlusion Assumption assumes that a real matching path should have minimum amount of occlusion under geometric constraints and photometric assumptions. This Minimum Occlusion Assumption implies that there is no cross-over occlusion whenever there is a matching.

6. DISSIMILARITY MAP AND ITS APPLICATIONS

6.1. Dissimilarity map (DSMP)

We define the intensity value in the left scanline as $I_L(Xc_l)$ and the right scanline as $I_R(Xc_r)$. Then, the dissimilarity map (DSMP) is defined as

$$DSMP(Xc_l, D) = F(I_L(Xc_l), I_R(Xc_l - D)), \tag{13}$$

where Xc_l is the coordinate in the left scanline and D is the position shift between the left and right scanline coordinates. When noise is absent, the $F(\cdot)$ should be zero, at $D = D_c$.

6.2. Lambertian surface model assumption on dissimilarity map

Suppose the corresponding matching point of Xc_l in the right scanline is Xc_r. From the Lambertian Surface Assumption, we know

$$F(I_L(Xc_l), I_R(Xc_l - D)) \approx 0, \tag{14}$$

Now, we define the disparity as $Xc_l - Xc_r$. When $D = D_c = Xc_l - Xc_r = disparity$,

$$DSMP(Xc_l, D_c) = F(I_L(Xc_l), I_R(Xc_l - D_c)) = F(I_L(Xc_l), I_R(Xc_r)) \approx 0 \tag{15}$$

From the above derivation, we know that the (Xc_l, Dc) coordinates of the unoccluded 3-D point locate in the low dissimilarity zone (LDZ) in DSMP. In addition, if there is no noise, the LDZ becomes zero-valued zones.

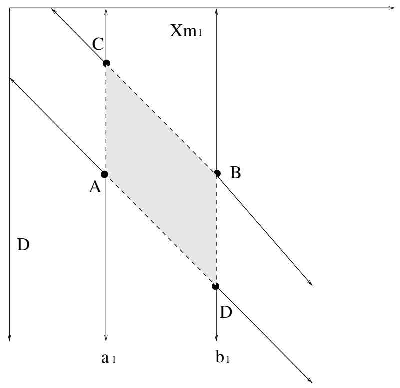


Figure 9. The LDZ element

6.2.1. The shape of LDZ in DSMP

The shape of LDZ is highly related to the possible location of the disparity path. Here we discuss the shape of the LDZs under the condition that the $F(\cdot) = 0$ for these matching pairs. We partition the shapes into four types.

Suppose $I_L(Xc_l) = C$ for $a_l < Xc_l < b_l$ with $a_l \neq C$ and $b_l \neq C$, and $I_R(Xc_r) = C$ for $a_r < Xc_r < b_r$ with $a_r \neq C$ and $b_r \neq C$. Then

$$DSMP(Xc_l, D) = F(I_L(Xc_l), I_R(Xc_l - D)) = 0 \tag{16}$$

when $a_l < Xc_l < b_l$ and $a_r < Xc_l - D < b_r$. This parallelogram, called the LDZ element, is plotted in Figure 9, where A, B, C and D are the possible neighboring zero points. The equations $DSMP(A) = 0$, $DSMP(B) = 0$, $DSMP(C) = 0$, and $DSMP(D) = 0$ correspond to the situation $I_L(a_l) = I_r(a_R)$, $I_L(b_l) = I_r(b_R)$, $I_L(a_l) = I_r(b_R)$, and $I_L(b_l) = I_r(a_R)$, respectively.

- **Parallelogram:** When $b_l - a_l$ and $b_r - a_r$ are not infinitely small, the correspoding zero region in SMP is a parallelogram.
- **Line segment perpendicular to Xc_l axis:** When $b_l - a_l$ is infinitely small but $b_r - a_r$ is not, the parallelogram degenerates to a line perpendicular to the Xc_l axis.
- **Line segment with slope=1:** When $b_l - a_l$ is not infinitely small but $b_r - a_r$, the parallelogram degenerates to a line with slope=1.
- **Point:** When both $b_l - a_l$ and $b_r - a_r$ are infinitely small, the parallelogram degenerates to a point.

All the LDZs includes all the above four types of LDZ elements with A, B, C and D as the connection points. The two cases of "Line segment perpendicular to Xc_l axis" and "Line segment with slope=1" hardly occur in the real world.

6.3. Geometric Constraints and Photometric Assumptions on dissimilarity map

When all the above geometric constraints and photometric assumptions are applied to the dissimilarity map, the allowable disparity path can be found by using the following description. The matching procedure for *layer k* matching path is performed between the conjugate epipolar lines and is restricted by Valid Pairing Constraint.

- **Step1:** Find the leftmost point (Xc_l, D) in $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$.
- **Step2:** Find the next leftmost point in region VI of the former point and repeat this process until the right boundary is reached in the $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$.
- **Step3:** Finding $matching\ region_{1,\ 2,\ 3....,\ k}$.

The above procedure leads to infinite disparity paths.

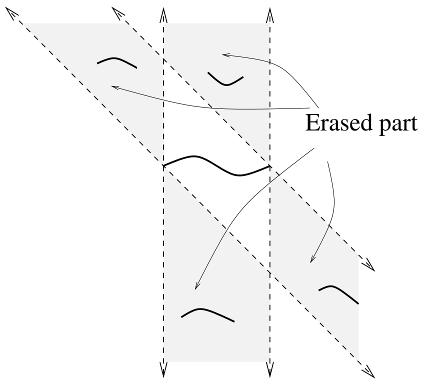


Figure 10. Largest LDZ assumption

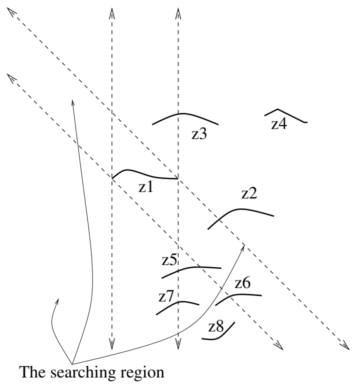


Figure 11. All combinations of two LDZs

6.4. Geometric Constraints, Photometric Assumptions and Minimum Occlusion Assumption on the dissimilarity map

The Minimum Occlusion Assumption implies that there is no cross-over occlusion whenever there is a possible matching. In real stereo images, the Minimum Occlusion Assumption also implies the use of non-maximum-LDZ suppression.

The non-maximum-LDZ suppression is to eliminate the LDZs that are contained in another LDZ when viewed from the left camera or the right camera, as shown in Figure 10. Based on this operation, the repeated texture in a surface can be eliminated. Under this assumption, all the combinations of two survived LDZs are illustrated in Figure 11. Suppose that $z1$ is the LDZ at the left side of the other LDZ, and $z2$, $z3$, $z4$, $z5$, $z6$, $z7$, $z8$ are the possible LDZ's to the right of $z1$.

From the above derivation of the disparity paths, we know that the number of possible solutions is infinite. Hence, we try to search for all the possible solutions based on these extracted LDZs. We use a LDZ set to represent all the disparity paths in this LDZ set. Because any disparity path must belongs to one of the LDZ sets, we can use numerable LDZ sets to contain all the possible disparity paths. Then we apply Minimum Occlusion Assumption in these LDZ sets to find the admissible LDZ set.

When all Geometric Constraints, Photometric Assumptions and Minimum Occlusion Assumptions are applied to the dissimilarity map, the procedure for finding all the possible disparity paths can be converted to the finding of all the admissible LDZ sets.

The multi-layer disparity path is in the LDZ sets. Corresponding to the description of multi-layer disparity path, we use the multi-layer LDZ to represent the LDZ sets and contain the multi-layer disparity paths.

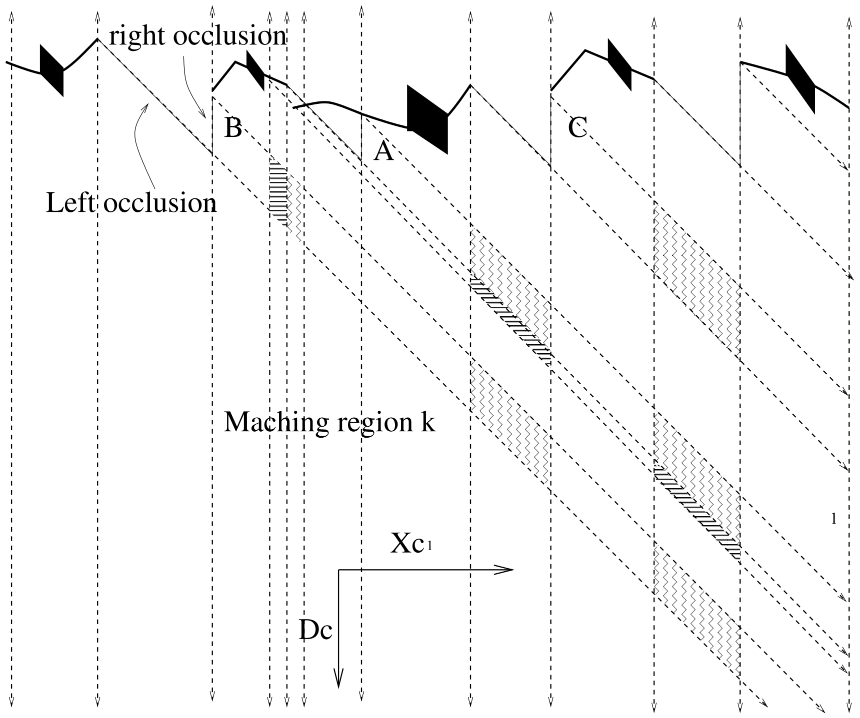


Figure 12. The layer k matching under picewise continuous curve with finite width

The definition for *layer* $1,\ 2,\ 3....,\ k$ LDZs are almost the same as the definition for disparity path. But the definition of the corresponding $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$ is modified to the intersection of the the left and right possible pseudo occlusion in the disparity path passing through the LDZs in *layer* $1,\ 2,\ 3....,\ (k-1)$ and the region violating monotonic order with the *layer* $1,\ 2,\ 3....,\ (k-1)$ shown in Figure 12. The procedure to find multi-layer LDZ is described as

- **Step1:** Find the LDZ with the leftmost point (Xc_l, D) in the $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$. The leftmost point in the first LDZ is the first reference point.
- **Step2:** Find the next LDZ with the leftmost point in the resudual LDZ in region VI with respect to the left reference point of the former LDZ. The choosen leftmost point in the next LDZ is the new reference point. The above process is equal to choose $z2$, $z3$, $z4$, $z5$, and $z6$ as shown in Figure 11. Repeat this process until the right boundary is reached in the $matching\ region_{1,\ 2,\ 3....,\ (k-1)}$.
- **Step3:** Find the $matching\ region_{1,\ 2,\ 3....,\ k}$

6.5. Depth Smoothness Assumption on the admissible LDZ sets

The Depth Smoothness Assumption assumes that the depth (disparity) should be smooth over the region of smooth intensity. The region with smooth intensity corresponds to an LDZ element of paralellogram shape. The disparity path in the paralellograms which obey the opaque assumption may have infinite solutions. The Depth Smoothness Assumption implies that the disparity path should be a line or a well fit curve across the paralellogram of LDZ.

6.6. Statistical criterions on admissible LDZ sets

When there are crossing parts between two neighboring LDZs, there may have infinite possible transition points. The selection of the transition point between two neighboring LDZs may depend on the criteria based on some statistical observations.

7. THE DSMP-BASED METHOD AND SIMULATION RESULTS

The DSMP-based method applies geometric constraints, photometric assumptions, and statistical assumptions to DSMP. The main operation is bounded by the Valid Pairing Constraint and is stated as below.

- **Step0: DSMP calculation**
  We use the mean square difference for $F(\cdot)$ on a 5×5 window to calculate the DSMP.
- **Step1: LDZ Growing**
  A threshold is used to detect the seeds of the LDZs, the areas which has small dissimilarity values. The growing technique is then applied.
- **Step2: LDZ Merging**
  The LDZ may be broken due to the thresholding operation. Hence, the broken LDZ need to be merged together.
- **Step3: LDZ modeling**
  Seperate the LDZ into paralellogram and curve.
- **Step4: Splitting of LDZ**
  A detected LDZ may contains more than two LDZs. These merged LDZs should be split.
- **Step5: Non-maximum-LDZ suppression**
  Eliminate the LDZs that are contained in another LDZ when viewed from one of the cameras.
- **Step6: Formation of the admissible LDZ sets**
  find the multi-layer LDZs.
- **Step7: Detection of occlusion**
  Detect the accurate boundary of occlusion.

Actually, in our simulation, Step5 is executed right after Step1 to reduce the computation complexity.

7.1. LDZ modeling

An LDZ is modeled as the combination of paralellograms and curves. In our simulation, the LDZ elements with their disparity range smaller than three are treated as curve. The remaining LDZ elements are treated as paralellograms.

7.2. Splitting of LDZ

When the disparities of the closest two end-points in two neighboring curves within a grown LDZ are different, this LDZ should be split.

7.3. Detection of occlusion

Occlusions occur between LDZs. The crossing points between the neighboring LDZs indicate the position of occlusion. The disparity difference between the neighboring LDZs is the width of the corresponding occlusion. The crossing point is decided by choosing the minimum cost solution along the disparity path between two neighboring LDZs.

7.4. Simulation results

The simulation results are shown in Figure 13. The bottom figure is the original left image. The top figure is the DSMP corresponding to the epipolar line shown in the left image. The second figure shows the extraction of LDZs. The third figure shows the disparity path. We can find that the disparity paths shown in Figure 13 are quite accurate and can detect the occlusions quite well. Another example by using the dynamic programming method[7] on the same image pairs is shown in Figure 14. The disparity path in Figure 14 includes many fake occlusions. The disparity maps generated by these two methods are compared in Figure 15. The DSMP-based method has demonstrated its potential to detect the occlusion more accurately.

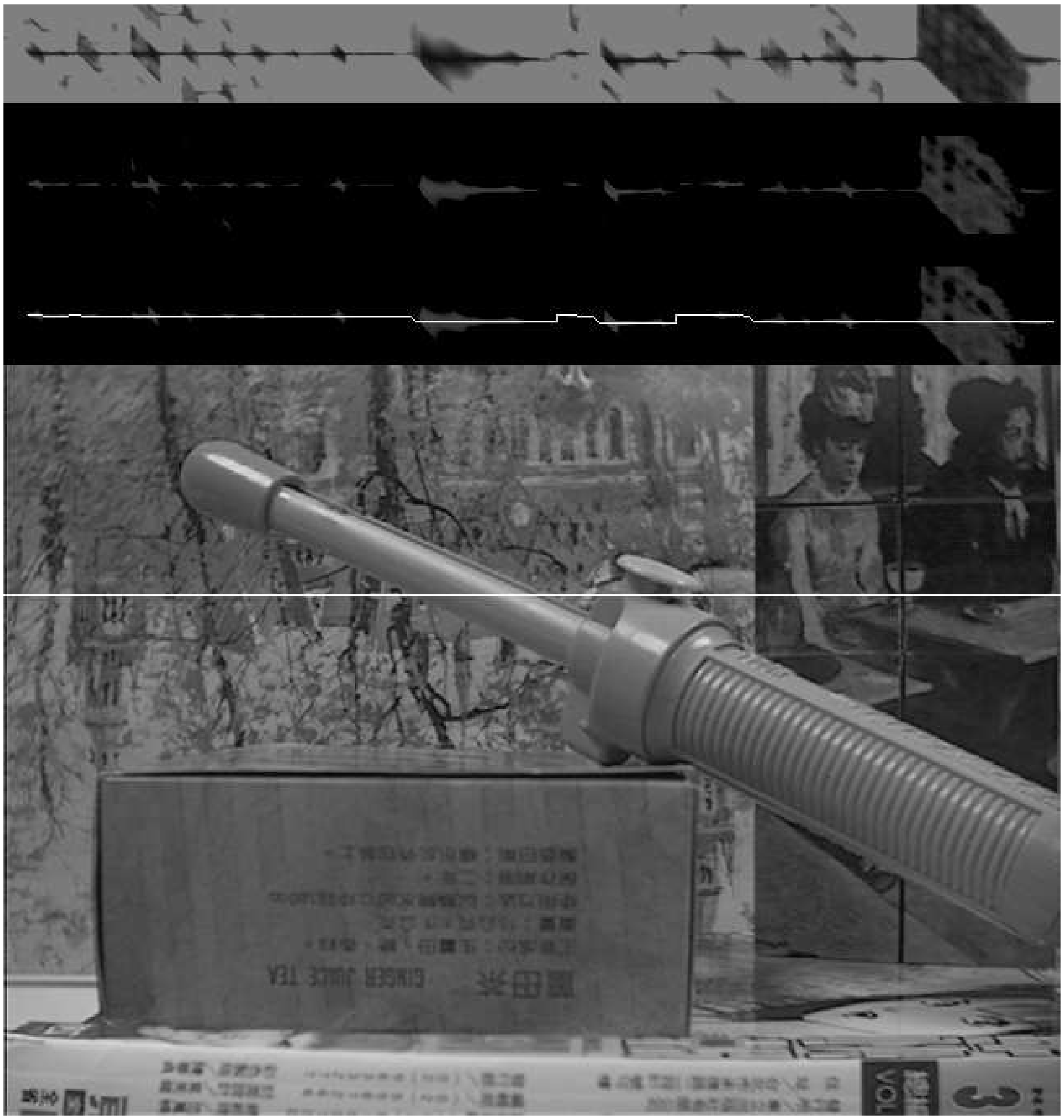
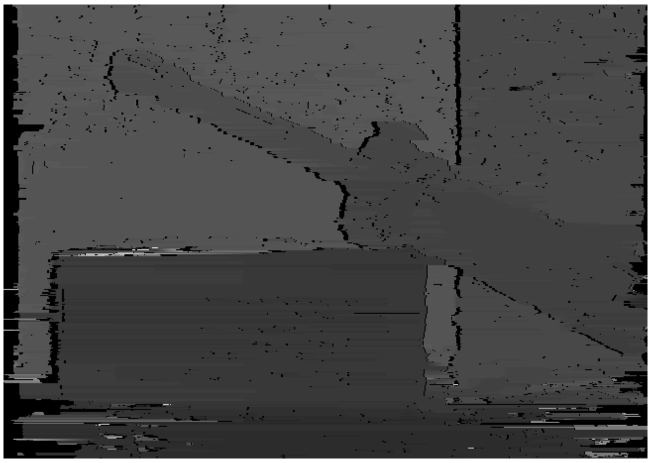


Figure 13. The DSMP, disparity path and original left image

Figure 14. The disparity path of dynamic programming



DSMP-based method

dynamic programming

Figure 15. The disparity maps produced by using DSMP-based method and dynamic programming

REFERENCES

1. E. Krotkov, K. Henriksen and R. Kories, "Stereo Ranging with Verging Cameras," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1200-1205, vol. 12, no. 12, December 1990.
2. T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment," *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pp. 1088-1094, Sacramento, California, April 1991.
3. T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1628-1634, vol. 16, no. 9, Sep. 1994.
4. Z. N. Le and G. H. Hu, "Analysis of disparity Gradient Based Cooperative Stereo," *IEEE Trans. on Image Processing*, pp. 1493-1506, vol. 5, no. 11, Nov. 1996.
5. T. Y. Chen, A. C. Bovik and L. K. Cormack, "Stereoscopic Ranging by Matching Image Modulations," *IEEE Trans. on Image Processing*, pp 785-797, vol. 8, no. 6, June, 1999.
6. V. S. Nalwa, "A Guided Tour of Computer Vision," *Addison-Wesley Publishing Company*, 1987.
7. I. J. Cox, S. L. Hingorani and S. B. Rao, "A Maximum Likelihood Stereo Algorithm," *Computer Vision and Image Understanding*, Vol. 63, No. 3, pp. 542-567, May, 1996.
8. S. Birchfied and C. Tomasi, "Depth Discontinuities by Pixel-to-Pixel Stereo," *Computer Science Department Stanford Unerversity.*