5-31-1991

# Direct recovering of composite surface using UOFF

Zhesheng Huang
*New Jersey Institute of Technology*

# ABSTRACT

Title of Thesis: Direct Recovering of Composite Surface
Using UOFF

Name: Zhesheng Huang
Master of Science in Electrical Engineering, 1991
Department of Electrical and Computer Engineering

Thesis directed by: Dr. Y.Q. Shi

The research work reported in the thesis is motivated by the problem raised in the computer vision area. It utilizes direct method to recover surface structure of objects in 3-D space from a pair of stereo images. The direct method, compared with optical flow-based approach and feature-based approach, is robust and computationally efficient. However it can only recover planar surface structure so far. A new approach to recover structure from a pair of stereo images based on unified optical flow field ($UOFF$) is developed recently. It can recover curved surface structure.

In this thesis, the problem of recovering surface structure of composite objects has been studied. That is, the surface is characterized by a factorable polynomial equation. Furthermore, instead of surfaces having finite large area (finite components in size), the infinite large surfaces (infinite components in size) are considered. The successful simulation results are presented. They show that the new approach is capable of recovering composite structure.

# Direct Recovering of Composite Surface Using UOFF

by

**Zhesheng Huang**

Thesis submitted to the Faculty of the Graduate School of the New Jersey Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering

1991

# APPROVAL SHEET

**Title of Thesis:** Direct Recovering of Composite Surface Using UOFF

**Candidate:** Zhesheng Huang
Master of Science in Electrical Engineering, 1991

**Thesis and Abstract Approved by the Examining Committee:**

Dr. Y.Q. Shi, Advisor — Date
Assistant Professor
Department of Electrical and Computer Engineerig

Dr. E.S.H.Hou — Date
Assistant Professor
Department of Electrical and Computer Engineerig

Dr. C.Q. Shu — Date
Research Associate
Department of Electrical and Computer Engineerig

New Jersey Institute of Technology, Newark, New Jersey.

# VITA

## Zhesheng Huang

**Date of Birth:**

**Place of Birth:**

| | |
|---|---|
| Education: | *1989-1991* |
| | *New Jersey Institute of Technology, NJ: MSEE* |
| | *1984-1987* |
| | *Shanghai U. of Science & Technology, China: MSEE* |
| | *1978-1982* |
| | *Shanghai U. of Science & Technology, China: BSEE* |
| Experience: | *1987-1989* |
| | *Shanghai U. of Science & Technology, China* |
| | *Instructor of EE Dept.* |
| | *1982-1984* |
| | *Ji-Xing Radio Equipment Co.Inc., China* |
| | *Electronic Engineer* |

## ACKNOWLEDGMENTS

I wish to express my sincere thanks to Dr. Y. Q. Shi for his valuable guidance and encouragement throughout the entire course of this thesis. I am specially indebted to him for his insightful and constructive suggestions at every stage of this research without which this thesis could never have been completed. I also wish to express my indebtedness to Dr. C. Q. Shu for his support in this research. Appreciation is also extended to the committee member. Dr. Edwin S. H. Hou for his evaluation of this research. The last but not the least I would like to thank all those who have helped me in many ways during the whole course of this research.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Problem Formation

The ability to discern objects, ascertain their motion, and navigate in three-dimensional space through the use of vision is almost universal among animals. The ease with which humans detect structure and motion around objects and the difficulty of duplicating these capabilities in machines have recently led to major efforts by computer engineers and scientist to develop the vision systems by computer.

The main task in the area of computer vision is recovering body motion and surface structure from a sequence of stereo images. Two types of approaches have been pursued: the feature correspondence approach and the optical flow field approach [1].

In much of the work on recovering surface structure and motion by two approaches mentioned above, it is assumed that either a correspondence between a sufficient number of feature points in successive frames has been established for the feature-based approach or that a reasonable estimate of the full optical flow field is available for the optical flow approach [1-3]. In general, it is difficult to extract and establish feature correspondence. On other hand, to computer the optical flow field, one needs additional constraints such as the smoothness constraint. This, in some cases, leads to an estimated optical flow field that is not the same as the true motion field.

In 1985, a new approach, direct method, was proposed by *Shahriar* and *Horn* [4]. It is more robust since information over the whole image is employed. And it requires less computation since readily computable data (image brightness gradients and time derivatives) are used directly to extract structure and motion information. It does not need the intermediate steps (feature detection or optical flow computation) which are computationally complicated.

Since then, several developmental and fundamental works on the direct method were introduced and contributed by *Horn* and *Weldom* (1988) [5], *Heel* and *Negahdaripour* (1990) [6], *Negahderipour et al.* (1989) [7], *Aloimonos et al.* [8], *Hayashi* and *Negahdaripour* (1990) [9] and so on.

*Negahderipour* and *Horn* showed how to recover the structure and motion of an observer relative to a planar surface from image brightness derivatives in their experiment. The direction of translational motion and depth of object are recovered using a procedure which involves minimizing the sum of the squared error of a linear constraint equation over the image by *Hayashi* and *Negahdaripour*. The contributions of *Horn* and *Weldon* are recovering the structure and motion of an object in a static-environment in the case of pure rotation and pure translation, where a planar object is assumed.

However this method only solves the planar surface successfully. So far there is little in the literature about how to recover the higher order surface structure of combinational object by using the direct approach. This is the motivation for pursuing a new approach to solve this problem.

## 1.2 A Solution to the Problem

Recently, a new concept of the unified optical flow field is established which is an extension of the classical optical flow determined by *Horn* and *Schunck*. Two main aspects of the *UOFF* are discussed in [10-13]. First, the brightness function of an image is considered as not only a function of time but also a function of the camera's spatial position. The concept of imaging space is presented as an accurate description of the set of all possible brightness function. Secondly, the brightness invariance is recognized not only for the time variation but also for the camera's space variation so that the brightness invariance equations for both time domain and space domain are established.

Based on the *UOFF*, a new approach for recovering surface structure characterized by an $Nth$ degree polynomial equation has been recently developed [14]. In this thesis work, this approach is utilized to recover composite surface structure, *i.e.*, surface characterized by an $Nth$ degree factorable polynomial equation.

If the relative positions between two cameras are set and known at any specific moment, we can determine the surface structure of a combinational object directly from the image brightness gradient without the need to computer the optical flow as an intermediate step. Under the assumption that the geometrical relation between two cameras is known, we will first derive the image brightness constraint equation and an $Nth$ order polynomial equation. Then from these equations, a least squares formulation allows us to derive a linear matrix equation with the structure parameters as unknows. By solving this linear matrix equations, we can recover the structure of the object.

## 1.3 Outline

The rest of this thesis is organized as follows. In Chapter 2, the concept of the *UOFF* is introduced. We derive the brightness invariant equation in spatial domain. Some basic image geometry, such as coordinate system transformation and perspective transformation, are examined in Chapter 3. With these basic image geometry we establish some relationships between $3-D$ space and image plane in Chapter 4. In Chapter 4, we also present polynomial equations which describe surface structures under consideration. Using the spatial brightness invariant equation, a direct method, which involves minimizing the sum of the squared error of this surface polynomial equation over the whole image, is derived to estimate all the coefficients of the polynomial equations characterizing the combinational structures.

Simulation experiments which involve recovering combinational structure of two planes and three planes, respectively, and the respective results are presented in Chapter 5. Finally, some conclusions and the possible further researches are discussed in Chapter 6.

# Chapter 2

# Background

## 2.1 Optical Flow

### 2.1.1 Brightness Invariant Equation for Temporal Optical Flow

Optical flow is the distribution of apparent velocities of movement of brightness pattern in an image. Optical flow arises from the relative movement between object and the viewer. Consequently, optical flow can give important information about the spatial arrangement of the object viewed and the rate of change of this arrangement.

We will derive an equation that relates the change in image brightness at a point to the motion of brightness pattern [2]. Let the image brightness at the point $(x, y)$ in the image plane at the moment $t$ be denoted $g(x, y, t)$ and it is displaced a distance $\Delta x$ in the X-direction and $\Delta y$ in the Y-direction in the time interval $\Delta t$. The brightness of this image point is assumed to remain constant so that:

$$g(x, y, t) = g(x + \Delta x, y + \Delta y, t + \Delta t) \tag{2.1}$$

Expanding the right-hand side about the point $(x, y, t)$ by Taylor series, we get:

$$g(x,y,t) = g(x,y,t) + \Delta x \frac{\partial g}{\partial x} + \Delta y \frac{\partial g}{\partial y} + \Delta t \frac{\partial g}{\partial t} + \epsilon \tag{2.2}$$

where $\epsilon$ contains second and higher order terms in $\Delta x, \Delta y$ and $\Delta t$. After subtracting $g(x,y,t)$ from both sides and dividing through by $\Delta t$ and evaluating the limit as $\Delta t \to 0$. we have:

$$\frac{\Delta x}{\Delta t}\frac{\partial g}{\partial x} + \frac{\Delta y}{\Delta t}\frac{\partial g}{\partial y} + \frac{\partial g}{\partial t} + O(\Delta t) = 0 \tag{2.3}$$

where $O(\Delta t)$ is a term of order $\Delta t$ (we assume that $\Delta x$ and $\Delta y$ vary as $\Delta t$). In the limit as $\Delta t \to 0$, this becomes:

$$\frac{\partial g}{\partial x}\frac{ds}{dt} + \frac{\partial g}{\partial y}\frac{dy}{dt} + \frac{\partial g}{\partial t} = 0 \tag{2.4}$$

If we let:

$$\begin{aligned} u &\triangleq \frac{dx}{dt} \\ v &\triangleq \frac{dy}{dt} \\ g_x &\triangleq \frac{\partial g}{\partial x} \\ g_y &\triangleq \frac{\partial g}{\partial y} \\ g_t &\triangleq \frac{\partial g}{\partial t} \end{aligned} \tag{2.5}$$

Then we get the brightness invariant equation in the time domain:

$$g_x u + g_y v + g_t = 0 \tag{2.6}$$

### 2.1.2 The Smoothness Constraints

We can't solve two unknown flow velocities $u$ and $v$ from one brightness invariant equation (2.6). In order to recover optical flow we must introduce additional constraints.

It is evident that if every point of the brightness pattern can move independent, there is little hope of recovering the optical velocities. So we assume that the neighboring points on the objects have similar velocities and the velocity field of the brightness pattern in the image varies smoothly almost everywhere.

One way to meet this smoothness constraint is to minimize the sum of the square of the magnitude of the gradient of the optical flow:

$$(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2 + (\frac{\partial v}{\partial x})^2 + (\frac{\partial v}{\partial y})^2 \tag{2.7}$$

That is, the sum of the square of the magnitude of the gradient is used as the smoothness measure.

### 2.1.3 Solving the Optical Flow

Now we have two equations: the brightness invariant equation and the smoothness constraint equation. The problem then is to minimize the sum of the errors in the brightness invariant equation:

$$\xi_b = g_s u + g_y v + g_t \tag{2.8}$$

and the measure of the departure from smoothness in the optical flow:

$$\xi_c^2 = (\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2 + (\frac{\partial v}{\partial x})^2 + (\frac{\partial v}{\partial y})^2 \tag{2.9}$$

Let the total error function is:

$$\xi^2 = \int\int(\alpha^2\xi_c^2 + \xi_b^2)dxdy \tag{2.10}$$

The minimization of this error function is to be accomplished by finding suitable value for the optical flow $(u, v)$. We can get:

$$\begin{aligned} g_x^2 u + g_x g_y v &= \alpha^2 \nabla^2 - g_x g_t \\ g_x g_y u + g_y^2 v &= \alpha^2 \nabla^2 - g_y g_t \end{aligned} \tag{2.11}$$

Then we can use iterative method to solve $(u, v)$ from these equations.

From the above description, it is evident that the recovering of the optical flow, the intermediate step for recovering of structure and motion of object, involves large amount of computation. Moreover it is necessary to introduce the smoothness constraint for determining the optical flow. The problem is sometime this constraint is not realistic.

## 2.2 Unified Optical Flow Field

### 2.2.1 General Concept

The current interest is in the development of robust and computationally efficient approach to recover motion and structure. Recently, a new concept, *UOFF*, has been developed [12][13]. It is an extension of classic optical flow determined by *Horn* and *Schunck*.

According to the concept of *UOFF*, the brightness function of an image is considered as not only a function of time but also a function of camera's spatial position and the brightness invariance is recognized not only for the time variation but also for the space variation so that the brightness invariant equation for both time and space domain are established.

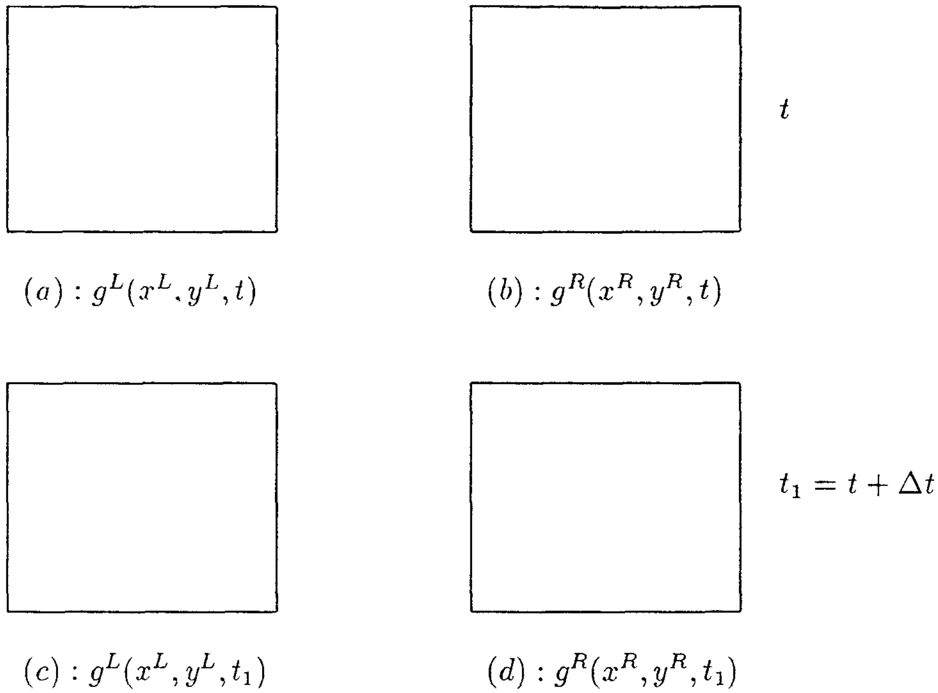


Figure 2.1: Four Frame Model

By extending the optical flow to the *UOFF*, more information for recovering of structure and motion of object in *3-D* space can be provided. We can use direct approach, which will be discussed in next section, to recover the structure and motion of object directly and efficiently.

### 2.2.2 Spatial Brightness Invariant Equation

The four frame models of image shown in Figure 2.1 are chosen from a stereo image sequence where images (a) and (c) are taken by the left camera at moment $t$ and $t_1 = t + \Delta t$, respectively, images (b) and (d) by left camera at $t$ and $t_1$, respectively. Images (a) and (b) are a pair of stereo image at $t$, images (c) and (d) a pair at $t_1$. In our derivations, the image brightness at the point $(x, y)$ on the image plane at time $t$ is denoted by $g(x, y, t)$ with superscript indicating which camera is associated with.

The classic optical flow only considers the temporal sequence of images with the camera fixed in the space domain, for example from image (a) to image (c) or from image (b) to image(d). From image (a) to image (b) can be viewed as a spatial sequence of images with the moment "*fixed*" in the time domain. Under the assumption made above, we have:

$$g^L(x^L, y^L, t) = g^R(x^R, y^R, t) \tag{2.12}$$

where $(x^L, y^L)$ and $(x^R, y^R)$ are the coordinates of the image points on image (a) and image (b) respectively, such that they are corresponding to the same world point in $3-D$ space. It is the counterpart of the invariance of the brightness of a pair of image pixels on two images took at two different moment, respectively associated with the same world point in $3-D$ space.

Define:

$$\delta x \triangleq x^R - x^L \qquad \delta y \triangleq y^R - y^L \tag{2.13}$$

Hence:

$$g^L(x^L, y^L, t) = g^L(x^R - \delta x, y^R - \delta y, t) \tag{2.14}$$

The right hand side of the above equation can be expanded in the $Taylor Series$:

$$g^L(x^R - \delta x, y^R - \delta y, t) = g^L(x^R, y^R, t) - \frac{\partial g^L}{\partial x}\delta x - \frac{\partial g^L}{\partial y}\delta y + \epsilon \tag{2.15}$$

Where $\epsilon$ contains second and higher order terms in $\delta x$ and $\delta y$. From equation (2.12), (2.14), (2.15), it follows that:

$$\frac{\partial g^L}{\partial x}\delta x + \frac{\partial g^L}{\partial y}\delta y + [g^R(x^R, y^R, t) - g^L(x^R, y^R, t)] + \epsilon = 0 \tag{2.16}$$

Dividing both sides of Equation (2.16) by $\delta s$ leads to:

$$\frac{\partial g^L}{\partial x}\frac{\delta x}{\delta s} + \frac{\partial g^L}{\partial y}\frac{\delta y}{\delta s} + \frac{[g^R(x^R, y^R, t) - g^L(x^R, y^R, t)]}{\delta s} + O(\delta s) = 0 \qquad (2.17)$$

Where $O(\delta s)$ is a term of order $\delta s$ ($\delta x$ and $\delta y$ vary as $\delta s$). In the limit as $\delta s \to 0$ Equation (2.17) becomes:

$$\frac{\partial g^L}{\partial x}u^s + \frac{\partial g^L}{\partial y}v^s + g_s = 0 \qquad (2.18)$$

where:

$$g_s \doteq \frac{g^R(x^R, y^R, t) - g^L(x^R, y^R, t)}{\delta s} \qquad (2.19)$$

$$u^s \triangleq \lim_{\delta s \to 0} \frac{\delta x}{\delta s} \qquad (2.20)$$

$$v^s \triangleq \lim_{\delta s \to 0} \frac{\delta y}{\delta s} \qquad (2.21)$$

The above defined $u^s$ and $v^s$ are, respectively, the spatial variation rates of $\delta x$ and $\delta y$ with respect to $\delta s$. These two quantities generated from the spatial sequence of images can be viewed as the counterpart of $u$ and $v$ generated from the temporal sequence of images in Equation (2.6).

## 2.3 Direct Approach

The direct method is one of the methods for the recovering of structure and motion of object. In this method, we first derive a linear equation by some constraints, so-called the linear constraint equation, for estimating the structure and motion of object. Then we can get the optimal results by minimizing the integral of the square of the error difference, i.e.,the between the estimated and the actual values. The integral is taken over the image region of interest which is usually taken to be the whole image [4-9].

The first advantage of this direct approach is that certain computational difficulties inherent in the calculation of optical flow are avoided. In optical flow approach, it is necessary to make an extra assumption, an often utilized one is that the optical flow field is smooth, which sometimes is not realistic.

Secondly, this direct approach is more robust. Image brightness value are distorted with camera noise and quantization error. These inaccuracies are further accentuated by methods used for estimating the brightness gradient. Thus it is not advisable to base a method on measurement at just a few points. Instead we propose to minimize the error in the brightness constraint equation over the whole region $I$ in the image plane.

The current interest is in the development of robust and computationally efficient approach to recover structure and motion information. In this thesis, by using $UOFF$ equations and direct approach, a new recovering algorithm based on the direct structure constraint equation is discussed and used to estimate the combinational structure.

# Chapter 3

# Image Geometry

## 3.1 Coordinate System Transformations

In order to deal with the generation of stereo images in Chapter 5, we will discuss some basic transformation of coordinate [18]. A point in *3-D* space has coordinates $(X'_L, Y'_L, Z'_L)$ with respect to the Cartesian coordinate system $O_L - X_L Y_L Z_L$ associated with the left camera, and $(X'_R, Y'_R, Z'_R)$ for the Cartesian coordinate system $O_L - X_L Y_L Z_L$ associated with the right camera.

### 3.1.1 Translation

Suppose that the Cartesian coordinate system of the right camera is generated by translating the Cartesian coordinate system of the left camera to a new location with displacement $(l, h, k)$. The relation between the coordinates of the two Cartesian coordinate systems for a same point can be described by using the following equation:

$$\begin{aligned} X'_L &= X'_R + l \\ Y'_L &= Y'_R + h \\ Z'_L &= Z'_R + k \end{aligned} \tag{3.1}$$

Equation (3.1) may be expressed in matrix form by writing:

$$\begin{pmatrix} X'_L \\ Y'_L \\ Z'_L \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & l \\ 0 & 0 & 1 & h \\ 0 & 0 & 1 & k \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X'_R \\ Y'_R \\ Z'_R \\ 1 \end{pmatrix} \tag{3.2}$$

Denoted the translation matrix as:

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & l \\ 0 & 1 & 0 & h \\ 0 & 0 & 1 & k \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.3}$$

Using this transform, we can calculate the coordinates of a point in the Cartesian coordinate system of the left camera from the coordinates of the point in the Cartesian coordinate system of the right camera and *vice versa*. Since the matrix $\mathbf{T}$ is nonsingular.

### 3.1.2 Rotation

The transformation used for rotation in *3-D* space is inherently more complex than the transformations discussed above.

With reference to Figure 3.1, because of the rotation of a Cartesian coordinate system about Y coordinate axes by an angle $\phi$, the transformation between coordinates of the two Cartesian coordinate systems for a same point can be described by using the following matrix:

$$\mathbf{R_y} = \begin{pmatrix} cos\phi & 0 & -sin\phi & 0 \\ 0 & 1 & 0 & 0 \\ sin\phi & 0 & cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.4}$$

where the rotation angle $\phi$ is measured clockwise when looking at the origin from a point on the $+Y$ axis. It is noted that this transformation affects only

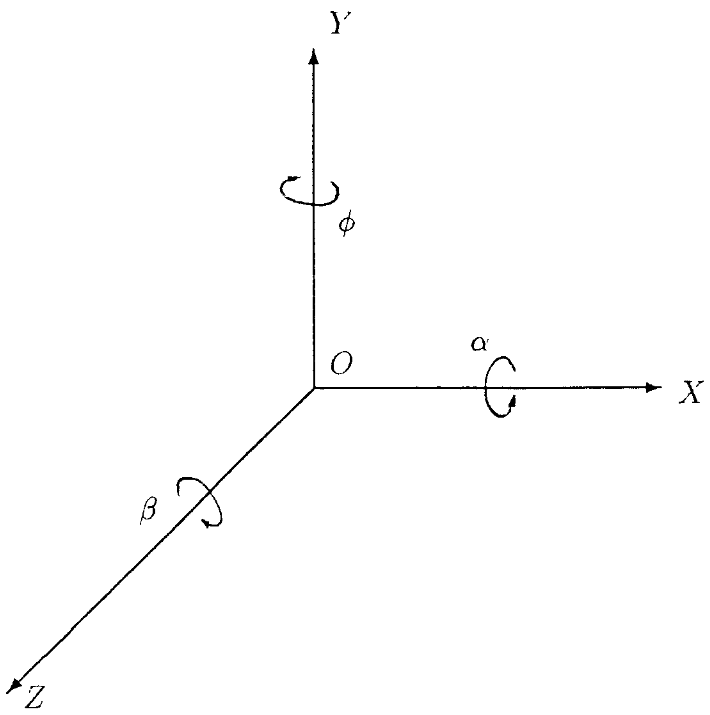


Figure 3.1: Rotation of Cartesian Coordinate System

the values of $X$ and $Z$ coordinates.

Rotation of a Cartesian coordinate system about $X$ axis by angle $\alpha$ is performed by using the transformation

$$\mathbf{R_x} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & cos\alpha & sin\alpha & 0 \\ 0 & -sin\alpha & cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.5}$$

and rotation of a Cartesian coordinate system about $Z$ axis by angle $\beta$ is achieved by using transformation:

$$\mathbf{R_z} = \begin{pmatrix} cos\beta & sin\beta & 0 & 0 \\ -sin\beta & cos\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.6}$$

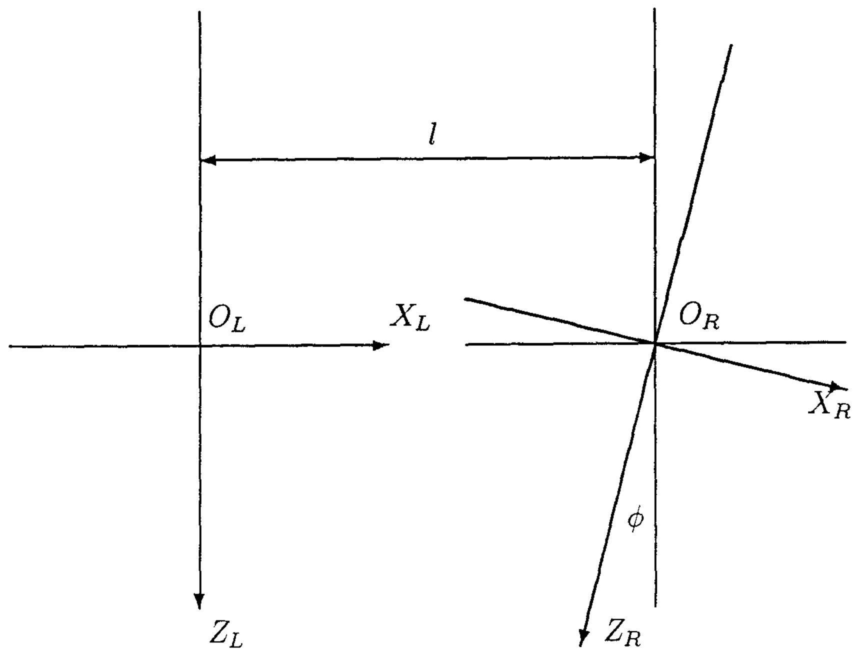


Figure 3.2: Two Cameras Setting

### 3.1.3 Camera Setting

The setting of the two cameras is shown in Figure 3.2. There $O_L - X_LY_LZ_L$ and $O_R - X_RY_RZ_R$ are two Cartesian coordinate systems such that $O_L$ and $O_R$ are the centers of the left lens and the right lens, respectively, $O_LZ_L$ and $O_RZ_R$ are the optical axes of the left and right lenses, respectively. It is assumed in this thesis that the two optical axes $O_LZ_L$ and $O_RZ_R$ are on the same plane, *i.e.*, the four axes: $O_LZ_L$, $O_RZ_R$, $O_LX_L$ and $O_RX_R$ are coplanar. The axis $O_LY_L$ is not drawn in Figure 3.2 and is understood as being perpendicular to $X_LO_LZ_L$ plane, $O_RY_R$ is the corresponding component associated with $O_R - X_RY_RZ_R$. The distance between two lens centers is $O_LO_R$ denoted by $l$. The angle between $O_LX_L$ and $O_RX_R$ is denoted by $\phi$.

Considering a point $P$ in *3-D* space. Its coordinates in the two coordinate system are $(X'_L, Y'_L, Z'_L)$ and $(X'_R, Y_R,' Z'_R)$, respectively. According to the transformation of coordinate system described above, the relationship between the two coordinates can be expressed as follows:

$$\begin{pmatrix} X'_L \\ Y'_L \\ Z'_L \\ 1 \end{pmatrix} = \mathbf{R_y T} \begin{pmatrix} X'_R \\ Y'_R \\ Z'_R \\ 1 \end{pmatrix} \tag{3.7}$$

where $\mathbf{R_y}$ is rotation matrix about $Y$ axis and $\mathbf{T}$ is translation matrix. So we can get:

$$\begin{aligned} X'_L &= cos\phi X'_R - sin\phi Z'_R + l \\ Y'_L &= Y'_R \\ Z'_L &= sin\phi X'_R + cos\phi Z'_R \end{aligned} \tag{3.8}$$

This is a basic relation formula between two Cartesian coordinate systems of camera in Chapter 5.

## 3.2 Image Perspective Transformations

### 3.2.1 Perspective Transformation

A perspective transformation (also called an imaging transformation) projects *3-D* points onto a image plane. Perspective transformation play a central role in image processing because they provide an approximation to the manner in which an image is formed by viewing a *3-D* world.

A model of the image perspective transformation is shown in Figure 3.3. We define camera coordinate system as having the image plane coincident with the $XY$ plane, and assume the optical axis of the camera aligns with the $Z$ axis and $f$ is the focal length of the lens.

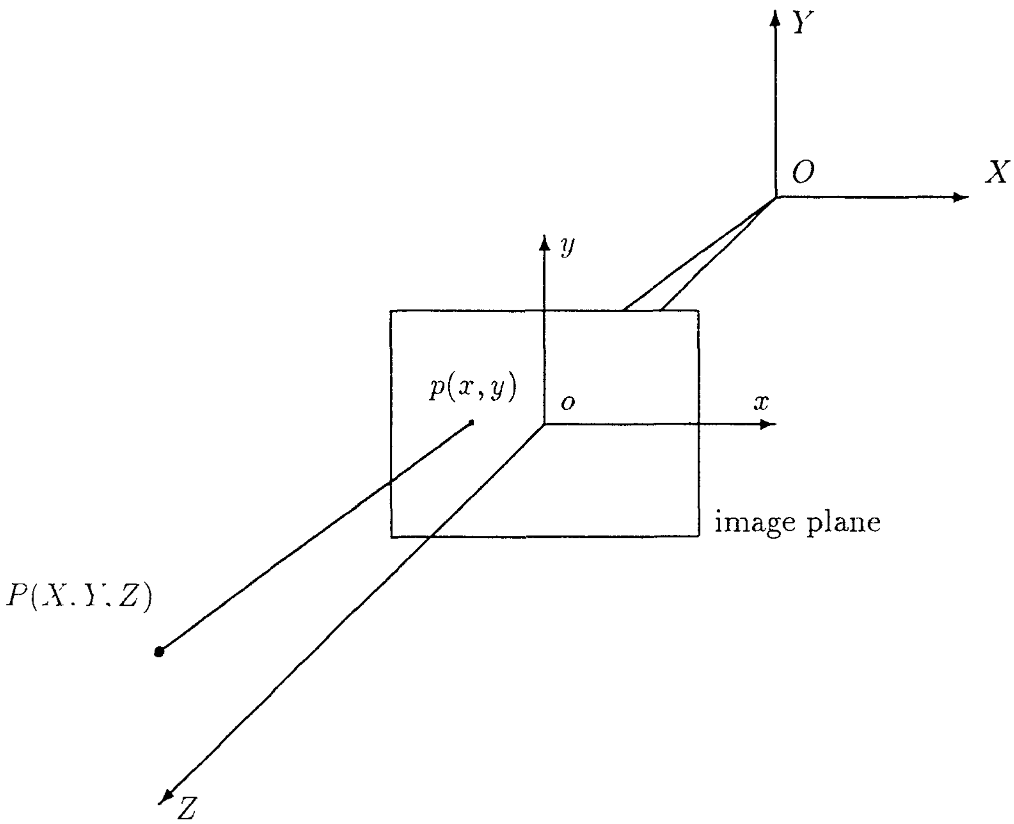


Figure 3.3: Image Perspective Transformation

Let $P(X, Y, Z)$ be the coordinates of any point in a $3-D$ scene. What we wish to do first is to obtain a relationship that gives the coordinates $p(x, y)$ of the projection of the point $P(X, Y, Z)$ onto the image plane. This is easily accomplished by the use of similar triangles. With reference to Figure 3.3, it follows that:

$$\frac{x}{f} = \frac{X}{Z} \tag{3.9}$$

$$\frac{y}{f} = \frac{Y}{Z} \tag{3.10}$$

In this thesis, $f$ is set to unit, $f = 1$. So we get:

$$x = \frac{X}{Z} \qquad y = \frac{Y}{Z} \tag{3.11}$$

# Chapter 4

# Direct Structure Stereo

## 4.1 Optical Flow of Stereo Images

According to the concept of imaging space introduced in Section 1.2, though an object does not move in *3-D* space, it looks as if it would have experienced certain movement from the different cameras' view. These pseudo-movement can be treated in a manner similar to the treatment of the relative motion between the camera and the object. This type of movement is characterized by:

$$\vec{T}_s = (U_s, V_s, W_s)^T \tag{4.1}$$

$$\vec{\omega}_s = (A_s, B_s, C_s)^T \tag{4.2}$$

where $\vec{T}_s$ is the translational component and the $\vec{\omega}_s$ the rotational component. We define:

$$\vec{V}_s = (\frac{dX}{ds}, \frac{dY}{ds}, \frac{dZ}{ds})^T \tag{4.3}$$

Then we have:

$$\vec{V_s} = -\vec{T_s} - \vec{\omega_s} \times \vec{r_s} \tag{4.4}$$

where $\vec{r_s}$ is a vector of $(X, Y, Z)^T$ representing a point in *3-D* space. By extending Equation (4.4), we get:

$$\begin{aligned} \frac{dX}{ds} &= -U_s - B_s Z + C_s Y \\ \frac{dY}{ds} &= -V_s - C_s X + A_s Z \\ \frac{dZ}{ds} &= -W_s - A_s Y + B_s X \end{aligned} \tag{4.5}$$

From the spatial optical flow $u^s$ and $v^s$ defined in Equations (2.20) and (2.21), the perspective projection Equation (3.11) and the Equation (4.5), we can obtain the following equations:

$$\begin{aligned} u^s &= \frac{\frac{dx}{ds}}{Z} - \frac{x\frac{dZ}{ds}}{Z^2} \\ &= (-\frac{U_s}{Z} - B_s + C_s y) - x(-\frac{W_s}{Z} - A_s y + B_s x) \end{aligned} \tag{4.6}$$

$$\begin{aligned} v^s &= \frac{\frac{dy}{ds}}{Z} - \frac{y\frac{dZ}{ds}}{Z^2} \\ &= (-\frac{V_s}{Z} - C_s x + A_s) - y(-\frac{W_s}{Z} - A_s y + B_s x) \end{aligned} \tag{4.7}$$

These are important equations for recovering the structure of combinational object in this thesis work.

## 4.2 Surface Structure Polynomial Equation

In this thesis, we consider a combinational surface that can be characterized by an $Nth$ order polynomial equation [14]. That is:

$$\sum_{j=0}^{K-1} \lambda(j) X^{\alpha(j)} Y^{\beta(j)} Z^{\gamma(j)} = 0 \tag{4.8}$$

where $0 \leq \alpha(j) + \beta(j) + \gamma(j) \leq N$, $K$ is the number of coefficients, that are not identically vanishing, in the $Nth$ order polynomial. The coefficients $\lambda(j)$, $j = 0, 1, \ldots K-1$ is an arbitrary but fixed index sequence in which all $K$ coefficients are arranged. Obviously, there is a problem in choosing $K-1$ independent coefficients from the total $K$ dependent coefficients. So that Equation (4.8) can be rewritten as:

$$\sum_{j=0, j\neq r}^{K-2} [\lambda(j) X^{\alpha(j)} Y^{\beta(j)} Z^{\gamma(j)}] + X^{\alpha(r)} Y^{\beta(r)} Z^{\gamma(r)} = 0 \tag{4.9}$$

The use of (3.11) leads to

$$\sum_{j=0, j\neq r}^{K-2} [\lambda(j) (xZ)^{\alpha(j)} (yZ)^{\beta(j)} Z^{\gamma(j)}] + (xZ)^{\alpha(r)} (yZ)^{\beta(r)} Z^{\gamma(r)} = 0 \tag{4.10}$$

or

$$\sum_{j=0, j\neq r}^{K-2} [\lambda(j) x^{\alpha(j)} y^{\beta(j)} Z^{\alpha(j)+\beta(j)+\gamma(j)}] + x^{\alpha(r)} y^{\beta(r)} Z^{\alpha(r)+\beta(r)+\gamma(r)} = 0 \tag{4.11}$$

## 4.3 Direct Structure Stereo Constraint Equation

We have gotten spatial brightness invariant equation and the surface constraint equation of combinational structure:

$$g_x u^s + g_y v^s + g_s = 0 \tag{4.12}$$

$$\sum_{j=0,j\neq r}^{K-2}[\lambda(j)x^{\alpha(j)}y^{\beta(j)}Z^{\alpha(j)+\beta(j)+\gamma(j)}] + x^{\alpha(r)}y^{\beta(r)}Z^{\alpha(r)+\beta(r)+\gamma(r)} = 0 \qquad (4.13)$$

From Equation (4.13), it is obvious that if we can solve $Z$, we can estimate the coefficients $\lambda(j)$ of the constraint Equation (4.11) so that the structure of the combination could be recovered.

When the camera translational component $\vec{T_s}$ and rotation component $\vec{\omega_s}$ are known in advance, we can solve $Z$ from spatial invariant equation (4.12). Substituting Equations (4.6) and (4.7) into (4.12) we obtain:

$$\begin{aligned}
\frac{1}{Z} &= \frac{(-A - sy + B_s x)(xg_x + yg_y) - g_x(-B_s + C_s y) - g_y(-C_s x + A_s) - g_s}{xW_sG_x + yW_sg_y - U_sg_x - V_sg_y)} \\
&= Q(x, y, g_x, g_y, g_s, A_s, B_s, C_s, U_s, V_s, W_s) \\
&= \frac{q}{p} \qquad (4.14)
\end{aligned}$$

where:

$$q = (-A - sy + B_s x)(xg_x + yg_y) - g_x(-B_s + C_s y) - g_y(-C_s x + A_s) - g_s \quad (4.15)$$

$$p = xW_sg_x + yW_sg_y - U_sg_x - V_sg_y \qquad (4.16)$$

It is noted that $A_s, B_s, C_s, U_s, V_s$ and $W_s$ can be determined once the relative positions of the two cameras in the stereo system are known. The $g_x, g_y$ and $g_s$ can be determined from the given image data. The $(x, y)$ is coordinate on the image plane. However instead of explicitly solving $Z^{-1}$, we will apply direct approach to the structure equation in a least squares formulation below.

According to (4.14), Equation (4.13) can be converted to:

$$\sum_{j=0,j\neq r}^{K-2}[\lambda(j)x^{\alpha(j)}y^{\beta(j)}Q^{-(\alpha(j)+\beta(j)+\gamma(j))}]+x^{\alpha(r)}y^{\beta(r)}Q^{-(\alpha(r)+\beta(r)+\gamma(r))}=0 \quad (4.17)$$

Define the cost function $J$ as:

$$J=\int\int_I\{\sum_{j=0,j\neq r}^{K-2}[\lambda(j)x^{\alpha(j)}y^{\beta(j)}Q^{-(\alpha(j)+\beta(j)+\gamma(j))}]+x^{\alpha(r)}y^{\beta(r)}Q^{-(\alpha(r)+\beta(r)+\gamma(r))}\}^2dxdy \quad (4.18)$$

where $I$ is the region on the image plane associated with thee concerned surface in *3-D* space. The task is to find a set of coefficients $\lambda(j)$ so that the cost function $J$ is minimized.

It is well known that the following linear equations are the necessary conditions for the minimization of the cost function $J$:

$$\frac{\partial J}{\partial\lambda(i)}=0 \quad (4.19)$$

where $i=0,1,\ldots,K-2$. We can get:

$$\frac{\partial J}{\partial\lambda(i)}=2\int\int_I\{\sum_{j=0,j\neq r}^{k-2}[\lambda(j)x^{\alpha(j)}y^{\beta(j)}Q^{-(\alpha(j)+\beta(j)+\gamma(j))}]+ x^{\alpha(r)}y^{\beta(r)}Q^{-(\alpha(r)+\beta(r)+\gamma(r))}\}\{x^{\alpha(i)}y^{\beta(i)}Q^{-(\alpha(i)+\beta(i)+\gamma(i))}\}dxdy=0 \quad (4.20)$$

and

$$\sum_{j=0}^{K-2}[\int\int_I(x^{\alpha(j)+\alpha(i)}y^{\beta(j)+\beta(i)}Q^{-(\alpha(j)+\beta(j)+\gamma(j)+\alpha(i)+\beta(i)+\gamma(i))})dxdy]\lambda(j)= -\int\int_I x^{\alpha(r)+\alpha(i)}y^{\beta(r)+\beta(i)}Q^{-(\alpha(r)+\alpha(i)+\beta(r)+\beta(i)+\gamma(r)+\gamma(i))}dxdy \quad (4.21)$$

In matrix vector form, we have:

$$\begin{bmatrix} M_{0,0} & \dots & M_{0,K-2} \\ \vdots & \vdots & \vdots \\ M_{K-2,0} & \dots & M_{K-2,K-2} \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \vdots \\ \lambda_{K-2} \end{bmatrix} = \begin{bmatrix} D_0 \\ \vdots \\ D_{K-2} \end{bmatrix} \tag{4.22}$$

where:

$$M_{i,j} = \int\int_I (x^{\alpha(j)+\alpha(i)} y^{\beta(j)+\beta(i)} Q^{-(\alpha(j)+\alpha(i)+\beta(j)+\beta(i)+\gamma(j)+\gamma(i))}) dx\,dy \tag{4.23}$$

$$D_i = -\int\int_I x^{\alpha(r)+\alpha(i)} y^{\beta(r)+\beta(i)} Q^{-(\alpha(r)+\alpha(i)+\beta(r)+\beta(i)+\gamma(r)+\gamma(i))} dx\,dy \tag{4.24}$$

with $i, j = 0, 1, \dots, K-2$.

In this set of linear equations all of coefficients of the $N$ the order polynomial, $i.e.$,$\lambda(0), \dots, \lambda(K-2)$ are unknown. all of the entries in the matrix, $i.e.$,$M_{i,j}$ and all of the entries in the vector,$i.e.$, $D_i$ can be computed from the given stereo image data and the known imaging setting. The unknown coefficients $\lambda(0), \dots, \lambda(K-2)$ can thus be solved. In other words, we can recover the surface structure of combination: both the shape of the combination and the position of the combination in $3-D$ space because the polynomial equation characterizing the combinational structure has been completely determined.

From the above description, the problem of recovering complicated and even insolvable high order surface structure in time domain is solved in space domain successfully. This is made possible by the introduction of the $UOFF$. The setup parameters of the camera in space simplify our estimation.

# Chapter 5

# Simulation Experiment

## 5.1 Combinational Structure of Multiplane in Space

### 5.1.1 Geometry Function of Combinational Structure

To simulate our recovering approach derived above, we set up a simulated combinational structure depicted in Figure 5.1, the combination of three planes is the object in *3-D* space in our experiment. We denoted plane $aO'c$ as plane 1 ($p1$), plane $aO'b$ as plane 2 ($p2$) and plane $bO'c$ as plane 3 ($p3$). Each plane is defined by its normal vector $(A, B, C)$ and one point $(X_0, Y_0, Z_0)$ on this plane and can be expressed as:

$$\begin{aligned} p1 : A_1(X - X_{01}) + B_1(Y - Y_{01}) + C_1(Z - Z_{01}) &= 0 \\ p2 : A_2(X - X_{02}) + B_2(Y - Y_{02}) + C_2(Z - Z_{02}) &= 0 \\ p3 : A_3(X - X_{03}) + B_3(Y - Y_{03}) + C_3(Z - Z_{03}) &= 0 \end{aligned} \tag{5.1}$$

In this experiment, we set:

$$\begin{aligned} X_{01} &= X_{02} = X_{03} = 0 \\ Y_{01} &= Y_{02} = Y_{03} = 0 \\ Z_{01} &= Z_{02} = Z_{03} = d \end{aligned} \tag{5.2}$$

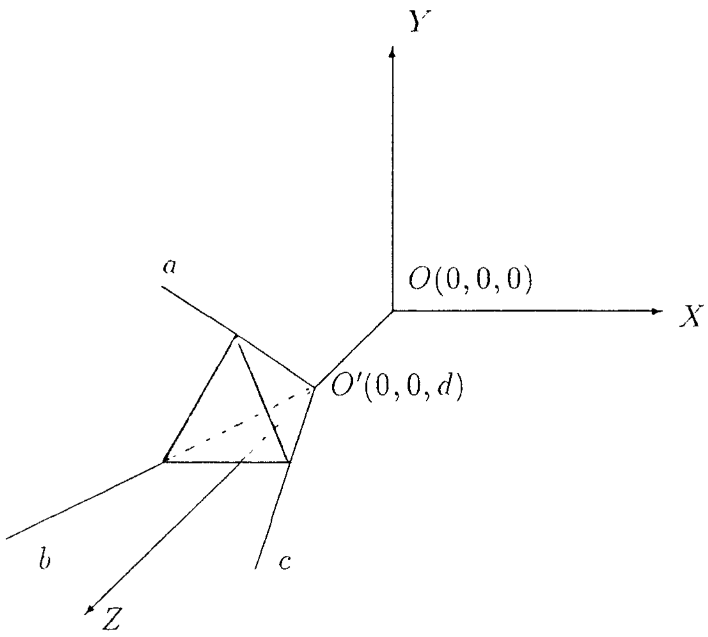


Figure 5.1: Combinational Structure by Three Planes

By substituting these parameters into (5.1), we can rewrite Equation (5.1) as:

$$\begin{aligned} p1: A_1X + B_1Y + C_1(Z-d) &= 0 \\ p2: A_2X + B_2Y + C_2(Z-d) &= 0 \\ P3: A_3X + B_3Y + C_3(Z-d) &= 0 \end{aligned} \tag{5.3}$$

or

$$\begin{aligned} p1: A_1X + B_1Y + C_1Z + D_1 &= 0 \\ p2: A_2X + B_2Y + C_2Z + D_2 &= 0 \\ p3: A_3X + B_3Y + C_3Z + D_3 &= 0 \end{aligned} \tag{5 4}$$

where:

$$D_1 = -C_1d$$

$$D_2 = -C_2 d$$
$$D_3 = -C_3 d \tag{5.5}$$

Thus the equation characterizing the composite structure consisting of three planes can be expressed as

$$(A_1X+B_1Y+C_1Z+D_1)(A_2X+B_2Y+C_2Z+D_2)(A_3X+B_3Y+C_3Z+D_3)=0 \tag{5.6}$$

### 5.1.2 Brightness Function of Combinational Structure

For each plane, the brightness function is defined as a sinusoidal function

$$g_i = k_1 sin(k_2 X_i') sin(k_2 Y_i') sin(k_2 Z_i') \tag{5.7}$$

where $i = 1, 2, 3$ and $(X_i', Y_i', Z_i')$ is the coordinate of a point on plane $i$ within the $O - XYZ$ Cartesian coordinate system. And $k_1 = 1000$ and $k_2 = 4$, which are used to adjust the brightness pattern on the planes.

## 5.2 Simulation Data in Stereo Image

### 5.2.1 Setup of Stereo Images

With respect to the combinational structure of three planes in space, we can obtain a pair of stereo images by using a stereo imaging described in Figure 5.2, which provides the basic information for recovering this combinational structure.

Figure 5.2 shows a stereo imaging system for this experiment. The center of the left camera $O_L$ is located on the origin of the reference coordinate system $O_L - X_L Y_L Z_L$. The optical axis of the left camera is aligned with the $O_L Z_L$. The distance between the vertex of the combination of three planes and the origin $O_L$ equals $d$.

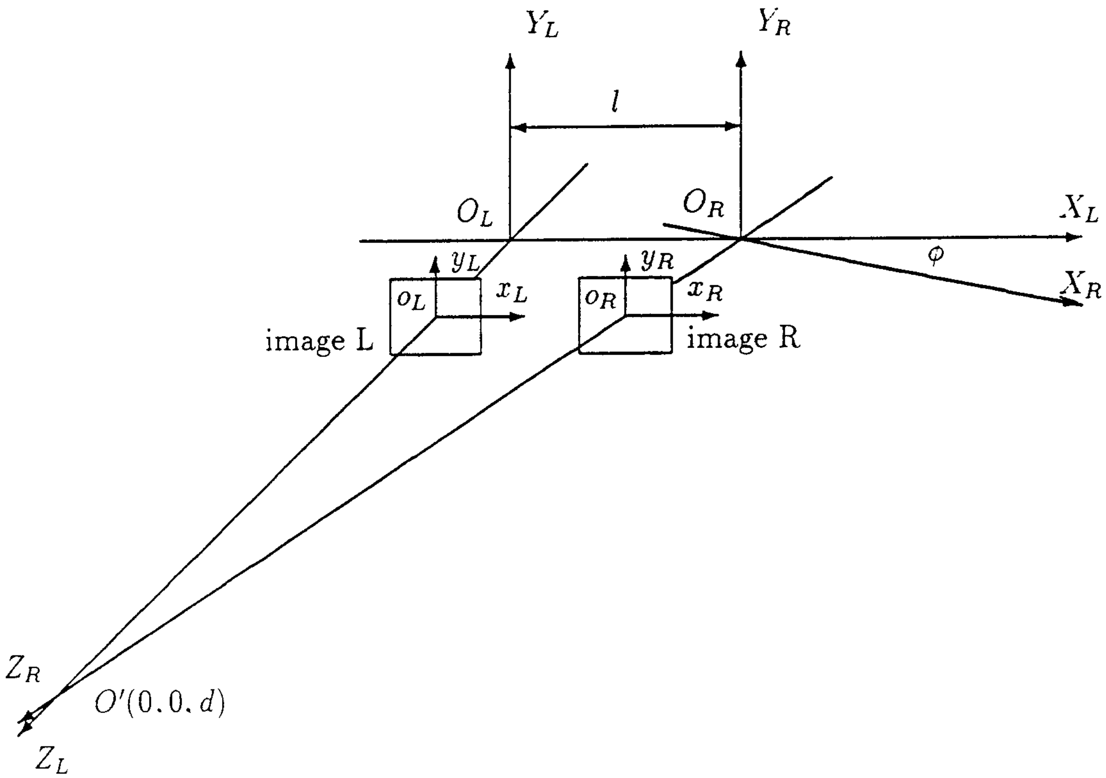


Figure 5.2: Stereo System Configuration

The Cartesian coordinate system $O_R - X_RY_RZ_R$ of the right camera can be obtained by translating the $O_L - X_LY_LZ_L$ along $X_L$ axis by $l$ and followed by rotating the $O_L - X_LY_LZ_L$ around $Y_L$ axis by angle $\phi$. The optical axis of the right camera is aligned with the $O_RZ_R$.

In this experiment, we assume that the two cameras are identical with focal length being equal to 1 and that the two cameras are **far from** the object. We should choose small $\phi$ and large $d$, so that

$$O_RO' \doteq O_LO' \tag{5.8}$$

The left and right image planes are perpendicular to $O_LZ_L$ and $O_RZ_R$, respectively. And $x_Lo_Ly_L$ and $x_Ro_Ry_R$ are the Cartesian coordinate systems of left image plane and right image plane respectively.

From this stereo image system, we can get a pair of images. Figure 5.3 shows that a pair of stereo images for the composite surface of three planes.

### 5.2.2 The Generation of Simulation Data

For recovering the combinational structure in *3-D* space, we should first get some information from a pair of stereo images. They are the corresponding point $(X', Y', Z')$ in *3-D* space of a point $(x, y)$ on the image plane, the gradients of image brightness, *i.e.*, $g_x$ and $g_y$ and $g_s$, which are introduced in Chapter 2.

#### Determine the corresponding point in *3-D* space

In this experiment, we choose the lengths of object in *3-D* space as $2R$, so the whole visual field of object in *3-D* space is a $2R \times 2R$ square. According to similar triangle principle, shown in Figure 5.4, if $f = 1$ then a square on the image plane with size $2R/d \times 2R/d$ will contain the whole picture of the object.

We convert this square on the image plane into $128 \times 128$ pixels. Then for each pixel in the image plane, we should find its corresponding point on the object in *3-D* space. That means for a specific pixel $(x, y)$, the corresponding point $(X', Y', Z')$ must be calculated firstly.

Starting from the plane Equation (5.4)

$$p_i : A_i X + B_i Y + C_i Z + D_i = 0 \tag{5.9}$$

using perspective projection Equation (3.11), *i.e.*

$$X = xZ \tag{5.10}$$

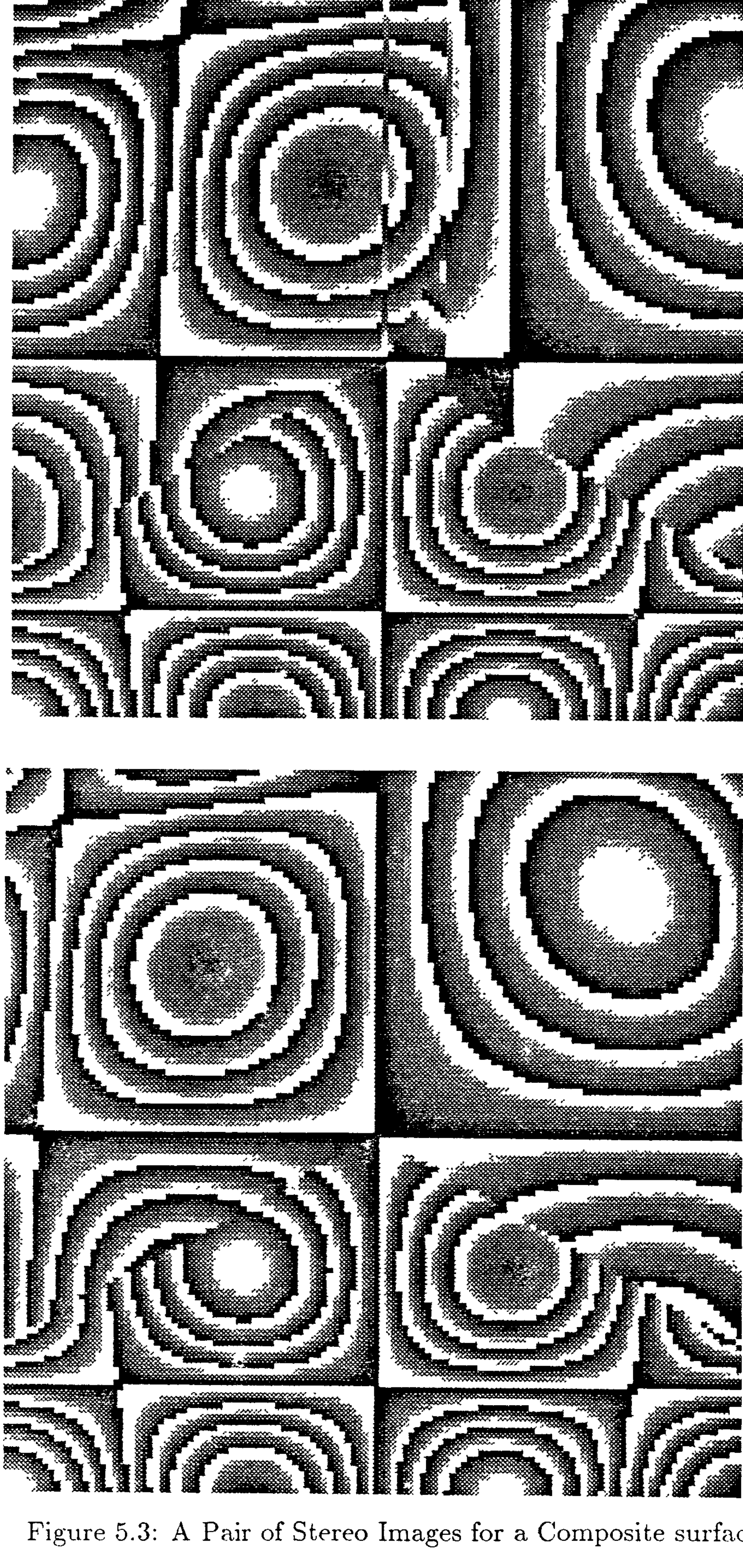

Figure 5.3: A Pair of Stereo Images for a Composite surface

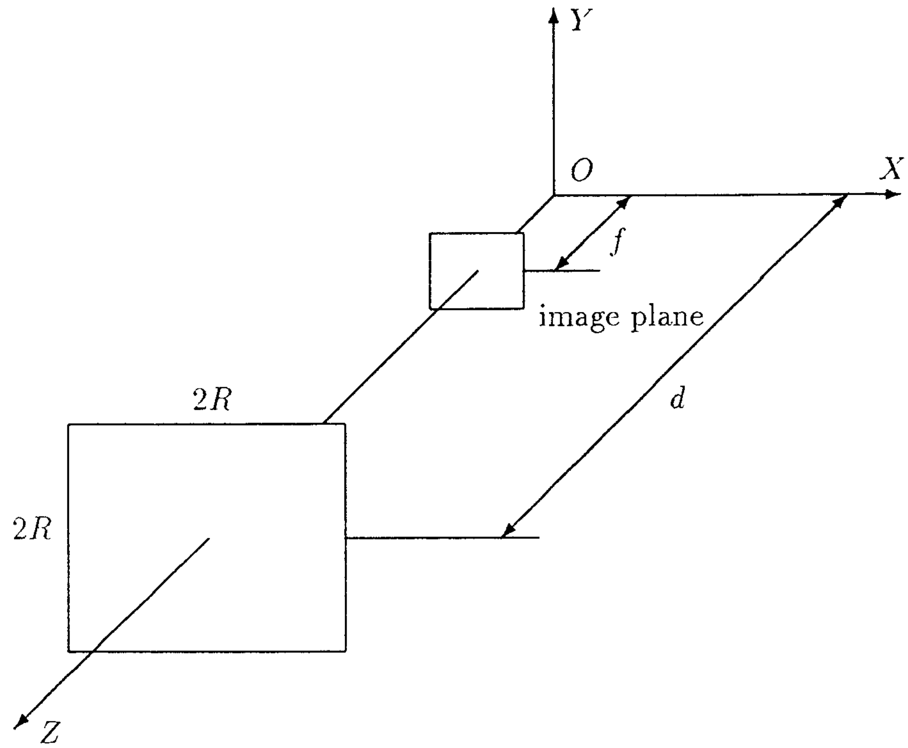


Figure 5.4: The Region of Image Plane

$$Y = yZ \tag{5.11}$$

we can have

$$A_i xZ + B_i yZ + C_i Z + D_i = 0 \tag{5.12}$$

So the corresponding coordinates in *3-D* space will be determined according to the next three equations.

$$\begin{aligned} Z' &= \frac{-D_i}{A_i x + B_i y + C_i} \\ &= \frac{C_i d}{A_i x + B_i y + C_i} \\ X' &= xZ' \\ Y' &= yZ' \end{aligned} \tag{5.13}$$

There are two planes (for the case of a combinational surface consisting of two planes) or three planes (for the case of a combinational surface consisting of three planes). So for a specific pixel $(x, y)$, there will be two or three corresponding points in *3-D* space. Which one is truly needed is depended on the structure of these combinational surface. In our experiment, we select the corresponding point which has farther distance from the origin of its coordinate system. For example, in the case of a combinational surface of two planes in *3-D* space, *i.e.*, plane 1 and plane 2, there will be two corresponding points $(X'_1, Y'_1, Z'_1)$ and $(X'_2, Y'_2, Z'_2)$ on the plane 1 and plane 2, respectively. If $Z'_1 > Z'_2$. we will select $(X'_1, Y'_1, Z'_1)$ as the corresponding point of the pixel, otherwise, the $(X'_2, Y'_2, Z'_2)$ as the corresponding point. This criterion is also applied in the case of a combinational surface of three planes.

### Calculate $g_x$ and $g_y$

We can rewrite brightness function Equation (5.7) for each plane equation as follows:

$$\begin{aligned} g_\imath(X', Y', Z') &= k_1 sin(k_2 x Z') sin(k_2 y Z') sin(k_2 Z') \\ &= k_1 sin(k_2 \frac{xC_\imath d}{A_\imath x + B_\imath y + C_\imath}) sin(k_2 \frac{yC_\imath d}{A_\imath x + B_\imath y + C_\imath}) \\ & \quad sin(k_2 \frac{C_\imath d}{A_\imath x + B_\imath y + C_\imath}) \end{aligned} \tag{5.14}$$

Hence

$$\begin{aligned} g_{x,\imath} &= \frac{\partial g_\imath}{\partial x} \\ &= k_1 [\frac{k_2 (Z')^2 (B_\imath y + C_\imath)}{C_\imath d} cos(k_2 x Z') sin(k_2 y Z') sin(k_2 Z') \\ & - \frac{k_2 A_\imath y (Z')^2}{C_\imath d} sin(k_2 x Z') cos(k_2 y Z') sin(k_2 Z') \end{aligned}$$

$$- \frac{k_2 A_\imath (Z')^2}{C_\imath d} sin(k_2 x Z') sin(k_2 y Z') cos(k_2 Z')] \tag{5.15}$$

$$\begin{aligned} g_{y,\imath} &= \frac{\partial g_\imath}{\partial y} \\ &= k_1 [\frac{-k_2 B_\imath x (Z')^2}{C_\imath d} cos(k_2 x Z') sin(k_2 y Z') sin(k_2 Z') \\ &+ \frac{k_2 (Z')^2 (A_\imath x + C_\imath)}{C_\imath d} sin(k_2 x Z') cos(k_2 y Z') sin(k_2 Z') \\ &- \frac{k_2 B_\imath (Z')^2}{C_\imath d} sin(k_2 x Z') sin(k_2 y Z') cos(k_2 Z')] \end{aligned} \tag{5.16}$$

For each given pixel $(x, y)$ on the image plane, if the plane equation, *i.e.*, $A_\imath, B_\imath, C_\imath$ and the distance $d$ are known, we can calculate its $g_x$ and $g_y$ by using Equations (5.15) and (5.16).

### Calculate $g_s$

According to Equation (2.19):

$$g_s \doteq \frac{g^R(x^R, y^R, t) - g^L(x^R, y^R, t)}{\delta s} \tag{5.17}$$

Or

$$g_s \doteq \frac{g^R(x^R, y^R) - g^L(x^L, y^L)}{\delta s} \tag{5.18}$$

It means that $g_s$ can be approximated by the ratio of the difference between the brightness of pixels in both right image plane and left image plane at the same location and the $\delta s$.
We can calculate the $g_s$ through the following steps.

**STEP 1**: Obtain the planar equations in the coordinate systems associated with the left camera and the right camera, respectively.

We express a planar structure equation in the Cartesian coordinate system of associated with the left camera as

$$A_L X_L + B_L Y_L + C_L Z_L + D_L = 0 \tag{5.19}$$

Using coordinate system transformation discussed in Chapter 3, we then get the planar structure equation for the same plane in the Cartesian coordinate system of the right camera. Using Equation (3.8), Equation (5.19) can be converted to:

$$A_L(cos\phi X_R - sin\phi Z_R + l) + B_L Y_R + C_L(sin\phi X_R + cos\phi Z_R) + D_L = 0 \tag{5.20}$$

That is equivalent to:

$$(A_L cos\phi + C_L sin\phi)X_R + B_L Y_R + (-A_L sin\phi + C_L cos\phi)Z_R + A_L l + D_L = 0 \tag{5.21}$$

We define:

$$\begin{aligned} A_R &= A_L cos\phi + C_L sin\phi \\ B_R &= B_L \\ C_R &= -A_L sin\phi + C_L cos\phi \\ D_R &= A_L l + D_L \end{aligned} \tag{5.22}$$

Thus Equation (5.21) becomes

$$A_R X_R + B_R Y_R + C_R Z_R + D_R = 0 \tag{5.23}$$

This is the planar structure equation for the same plane in the Cartesian coordinate system of the right camera.

**STEP 2**: Calculate the coordinates of the corresponding point for the same pixel in the left image plane and right image plane respectively.

Denote the pixels having the same location on the left image plane and right image plane as $(x_L, y_L)$ and $(x_R, y_R)$, respectively. Using Equation (5.13), we have

$$\begin{aligned} Z'_L &= \frac{-D_L}{A_L x_L + B_L y_L + C_L} \\ X'_L &= x_L Z'_L \\ Y'_L &= y_L Z'_L \end{aligned} \tag{5.24}$$

and

$$\begin{aligned} Z'_R &= \frac{-D_R}{A_R x_R + B_R y_R + C_R} \\ X'_R &= x_R Z'_R \\ Y'_R &= y_R Z'_R \end{aligned} \tag{5.25}$$

**STEP 3**: Calculate the difference of brightness between the world point $(X'_L, Y'_L, Z'_L)$ and $(X'_R, Y'_R, Z'_R)$.

According to Equation(5.7), we have

$$g^L(x_L, y_L) = k_1 sin(k_2 X'_L) sin(k_2 Y'_L) sin(k_2 Z'_L) \tag{5.26}$$

In order to follow the principle that the brightness of every point on object surface must remain the same for the right image, we should calculate the brightness of world point $(X'_R, Y'_R, Z'_R)$ with the **same brightness function**. So before assigning the brightness value to $g^R(x_R, y_R)$, we first should transform the coordinate $(X'_R, Y'_R, Z'_R)$ to the coordinate $(X''_L, Y''_L, Z''_L)$ in the Cartesian coordinate system $O_L - X_L Y_L Z_L$. Using Equation (3.8), we have

$$\begin{aligned} X''_L &= cos\phi X'_R - sin\phi Z'_R + l \\ Y''_L &= Y'_R \\ Z''_L &= sin\phi X'_R + cos\phi Z'_R \end{aligned} \tag{5.27}$$

Then we get

$$g^R(x_R, y_R) = k_1 sin(k_2 X''_L) sin(k_2 Y''_L) sin(k_2 Z''_L) \tag{5.28}$$

Final we obtain

$$\begin{aligned} g_s &= g^R(x_R, y_R) - g^L(x_L, y_L) \\ &= k_1[sin(k_2 X''_L) sin(k_2 Y''_L) sin(k_2 Z''_L) \\ &\quad - sin(k_2 X'_L) sin(k_2 Y'_L) sin(k_2 Z'_L)) \end{aligned} \tag{5.29}$$

We emphasize here that the correspondence problem of a stereo system has, unfortunately, been found to be a difficult one. Because of combination of multiplanes, there are several world points (correspondences) in *3-D* space corresponding to one point on the image plane. On the other hand, due to the transformation of position between two cameras, the pixels at the same location of the two image planes would possibly be associated with different planes in *3-D* space.

In order to recover the combinational structure, before computation of $g_s$, we should first determine that which plane in *3-D* space the pixel at same location of the two image planes will correspond to. Let

$$f(X_i, Y_i, Z_i) \triangleq k_1 sin(k_2 X_i) sin(k_2 Y_i) sin(k_2 Z_i) \tag{5.30}$$

then Equation (5.29) can be rewritten as

$$g_s = f(X''_{Li}, Y''_{Li}, Z''_{Li}) - f(X'_{Lj}, Y'_{Lj}, Z'_{Lj}) \tag{5.31}$$

where $(X''_{Li}, Y''_{Li}, Z''_{Li})$ and $(X'_{Lj}, Y'_{Lj}, Z'_{Lj})$ are discussed above.

For example, for combinational structure of two planes, there are four cases for calculating $g_s$:

- The correspondence of $(x_l, y_L)$ on plane 1

  The correspondence of $(x_R, y_R)$ on plane 1

  $g_s = f(X''_{L1}, Y''_{L1}, Z''_{L1}) - f(X'_{L1}, Y'_{L1}, Z'_{L1})$

- The correspondence of $(x_L, y_L)$ on plane 1

  The correspondence of $(x_R, y_R)$ on plane 2

  $g_s = f(X''_{L2}, Y''_{L2}, Z''_{L2}) - f(X'_{L1}, Y'_{L1}, Z'_{L1})$

- The correspondence of $(x_L, y_L)$ on plane 2

  The correspondence of $(x_R, y_R)$ on plane 1

  $g_s = f(X''_{L1}, Y''_{L1}, Z''_{L1}) - f(X'_{L2}, Y'_{L2}, Z'_{L2})$

- The correspondence of $(x_L, y_L)$ on plane 2

  The correspondence of $(x_R, y_R)$ on plane 2

  $g_s = f(X''_{L2}, Y''_{L2}, Z''_{L2}) - f(X'_{L2}, Y'_{L2}, Z'_{L2})$

There are 9 cases for combinational object of three planes which will not be discussed here in detail.

## 5.3 The Result of Experiment

### 5.3.1 Choice of Coordinate System for Estimation

Because the assumption of *far-field* in the stereo system is made in this experiment and in reality which usually is the case in stereo imagery, the depth $Z$ in the simulation is too large compared with the other quantities involved. To avoid computational error, we shift the coordinate system from the reference coordinate system $O - XYZ$ with its origin being $(0,0,0)$ to the coordinate system $\tilde{O} - \tilde{X}\tilde{Y}\tilde{Z}$ with its new origin bring $(0,0,d)$. Obviously, $\tilde{X} = X, \tilde{Y} = Y$ and $\tilde{Z} = Z - d$. By using this new coordinate system, for combinational structure of two planes, we get:

$$\begin{aligned} p1 : A_1\tilde{X} + B_1\tilde{Y} + C_1\tilde{Z} &= 0 \\ p2 : A_2\tilde{X} + B_2\tilde{Y} + C_2\tilde{Z} &= 0 \end{aligned} \tag{5.32}$$

The final combinational equation of two planes in space with new coordinate system can be written as:

$$(A_1\tilde{X} + B_1\tilde{Y} + C_1\tilde{Z})(A_2\tilde{X} + B_2\tilde{Y} + C_2\tilde{Z}) = 0 \tag{5.33}$$

and

$$\begin{aligned} A_1A_2\tilde{X}^2 + B_1B_2\tilde{Y}^2 + C_1C_2\tilde{Z}^2 + (A_1B_2 + B_1A_2)\tilde{X}\tilde{Y} + \\ (A_1C_2 + C_1A_2)\tilde{X}\tilde{Z} + (B_1C_2 + C_1B_2)\tilde{Y}\tilde{Z} = 0 \end{aligned} \tag{5.34}$$

Also, we have a structure equation for the composite surface of three planes:

$$(A_1\tilde{X} + B_1\tilde{Y} + C_1\tilde{Z})(A_2\tilde{X} + B_2\tilde{Y} + C_2\tilde{Z})$$
$$(A_3\tilde{X} + B_3\tilde{Y} + C_3\tilde{Z}) = 0 \tag{5.35}$$

where

$$\begin{aligned}
A_1 &= cos(35.2^o) \\
&= 0.816 \\
B_1 &= cos(61.8^o) \\
&= 0.472 \\
C_1 &= cos(109.47^o) \\
&= -0.333 \\
A_2 &= cos(144.7^o) \\
&= -0.816 \\
B_2 &= cos(61.8^o) \\
&= 0.472 \\
C_2 &= cos(109.47^o) \\
&= -0.333 \\
A_3 &= cos(90.0^o) \\
&= 0.0 \\
B_3 &= cos(160.7^o) \\
&= -0.944 \\
C_3 &= cos(109.47^o) \\
&= -0.333
\end{aligned}$$

We will estimate the coefficients of the combinational structure of two planes and three planes, respectively.

### 5.3.2 Estimation of Coefficient

Given $N = 2$, the coefficient of $\tilde{Z}^2$ can be chosen as the normalized coefficient. We can use a second degree polynomial equation to estimate the combinational structure of two planes. That is:

$$\begin{aligned}\lambda(0)\tilde{X}^2 + \lambda(1)\tilde{Y}^2 + \lambda(2)\tilde{X}\tilde{Y} + \lambda(3)\tilde{X}\tilde{Z} + \lambda(4)\tilde{Y}\tilde{Z} + \\ \lambda(5)\tilde{X} + \lambda(6)\tilde{Y} + \lambda(7)\tilde{Z} + \lambda(8) + \tilde{Z}^2 = 0\end{aligned} \tag{5.36}$$

where

$$\begin{aligned}\tilde{X} &= X \\ \tilde{Y} &= Y \\ \tilde{Z} &= Z - d \\ &= \frac{1}{Q} - d \\ &= \frac{q}{p} - d\end{aligned} \tag{5.37}$$

According Equations (4.17) and (3.10), we can get

$$\begin{aligned}\lambda(0)x^2q^2 + \lambda(1)y^2q^2 + \lambda(2)xyq^2 + \lambda(3)xq(q - dp) + \lambda(4)yq(q - dp) + \\ \lambda(5)xqp + \lambda(6)yqp + \lambda(7)(q - dp)p + \lambda(8)p^2 + (q - dp)^2 = 0\end{aligned} \tag{5.38}$$

We use direct approach here to minimize sum error over all pixels of image. As discussed in Chapter 4, we have:

$$\mathbf{J} = \int\int_I [\sum_{j=0}^{8} \lambda(j)G(j) + G(9)]^2 dxdy \tag{5.39}$$

where

$$\begin{aligned} G(0) &= x^2q^2 \\ G(1) &= y^2q^2 \\ G(2) &= xyq^2 \\ G(3) &= xq(q-dp) \\ G(4) &= yq(q-dp) \\ G(5) &= xqp \\ G(6) &= yqp \\ G(7) &= (q-dp)p \\ G(8) &= p^2 \\ G(9) &= (q-dp)^2 \end{aligned} \tag{5.40}$$

By minimizing the cost function $\mathbf{J}$ as derived in Chapter 4, finally, we obtain the linear matrix equations for recovering combinational structure:

$$\begin{aligned} \frac{\partial \mathbf{J}}{\partial \lambda(i)} &= \int\int_I [\sum_{j=0}^{8} \lambda(j)G(j) + G(9)]G(i)dxdy \\ &= 0 \end{aligned} \tag{5.41}$$

When $N = 3$ and the coefficient of $\tilde{Z}^3$ chosen as the normalized term, we can also get the linear matrix equations for recovering combinational structure of three planes:

$$\int\int_I [\sum_{j=0}^{18} \lambda(j)G(j) + G(19)]G(i)dxdy = 0 \tag{5.42}$$

where

$$G(0) = x^3q^3$$

$$
\begin{aligned}
G(1) &= y^3q^3 \\
G(2) &= x^2yq^3 \\
G(3) &= x^2q^2(q-dp) \\
G(4) &= xy^2q^3 \\
G(5) &= y^2q^2(q-dp) \\
G(6) &= xq(q-dp) \\
G(7) &= yq(q-dp) \\
G(8) &= xyq^2(q-dp) \\
G(9) &= x^2q^2p \\
G(10) &= y^2q^2p \\
G(11) &= (q-dp)^2p \\
G(12) &= xyq^2p \\
G(13) &= xq(q-dp)p \\
G(14) &= yq(q-dp)p \\
G(15) &= xqp^2 \\
G(16) &= yqp^2 \\
G(17) &= (q-dp)p^2 \\
G(18) &= p^3 \\
G(19) &= (q-dp)^3 \qquad (5.43)
\end{aligned}
$$

The results of experiment, the estimation of coefficients of $\lambda(j)$ for combinational structure of two planes, are shown in Tables 5.1 to 5.3 and the results for combinational structure of three planes are shown in Table 5.4.

In these tables, the left-most column contains all of the monomial in the polynomial equation for characterizing the composite surface structure of two or three planes. For each row, the right three items are coefficients associated with the monomial listed in the left-most item of the row. Among these right three items, the left one is the actual value and the right two are the estimated values.

Table 5.1: Estimation of Combinational Structure of Two Planes by Varying $\phi$

with $R = 2$ and $d = 100$

| Terms | Coef. of object | Coef.:$\phi = 0.05^o$ | Coef.:$\phi = 0.01^o$ |
|---|---|---|---|
| $\tilde{X}^2$ | -6.1100 | -5.5782 | -5.9989 |
| $\tilde{Y}^2$ | 2.0458 | 2.0380 | 2.0199 |
| $\tilde{X}\tilde{Y}$ | 0.0 | -0.0054 | 0.0015 |
| $\tilde{X}\tilde{Z}$ | 0.0 | 0.0097 | 0.0005 |
| $\tilde{Y}\tilde{Z}$ | -2.8623 | -2.8553 | -2.8420 |
| $\tilde{X}$ | 0.0 | 0.1241 | 0.0080 |
| $\tilde{Y}$ | 0.0 | 0.6370 | 0.0308 |
| $\tilde{Z}$ | 0.0 | -0.4185 | -0.0206 |
| constant | 0.0 | 0.4331 | 0.0209 |
| depth | 100.0 | 102.6254 | 102.61 |

Table 5.2: Estimation of Combinational Structure of Two Planes by Varying $R$

with $\phi = 0.01^o$ and $d = 300$

| Terms | Coef. of object | Coef.:$R = 4$ | Coef.:$R = 2$ |
|---|---|---|---|
| $\tilde{X}^2$ | -6.1100 | -6.0031 | -6.0017 |
| $\tilde{Y}^2$ | 2.0458 | 2.0241 | 2.0188 |
| $\tilde{X}\tilde{Y}$ | 0.0 | -0.0082 | -0.0029 |
| $\tilde{X}\tilde{Z}$ | 0.0 | 0.0072 | 0.0062 |
| $\tilde{Y}\tilde{Z}$ | -2.8623 | -2.8460 | -2.8414 |
| $\tilde{X}$ | 0.0 | 0.0351 | 0.0102 |
| $\tilde{Y}$ | 0.0 | 0.0752 | 0.0241 |
| $\tilde{Z}$ | 0.0 | -0.0392 | -0.0169 |
| constant | 0.0 | 0.0943 | 0.0153 |
| depth | 300.0 | 305.29 | 302.53 |

Table 5.3: Estimation of Combinational Structure of Two Planes by Varying $d$

with $\phi = 0.01^o$ and $R = 4$

| Terms | Coef. of object | Coef.:$d = 300$ | Coef.:$d = 500$ |
|---|---|---|---|
| $\tilde{X}^2$ | -6.1100 | -6.0031 | -6.0054 |
| $\tilde{Y}^2$ | 2.0458 | 2.0241 | 2.0228 |
| $\tilde{X}\tilde{Y}$ | 0.0 | -0.0082 | 0.0001 |
| $\tilde{X}\tilde{Z}$ | 0.0 | 0.0072 | 0.0041 |
| $\tilde{Y}\tilde{Z}$ | -2.8623 | -2.8460 | -2.8426 |
| $\tilde{X}$ | 0.0 | -0.0351 | -0.0104 |
| $\tilde{Y}$ | 0.0 | 0.0752 | 0.0633 |
| $\tilde{Z}$ | 0.0 | -0.0392 | -0.0336 |
| constant | 0.0 | 0.0943 | 0.0809 |
| depth | 300/500 | 305.29 | 505.09 |

Table 5.4: Estimation of Combinational Structure of Three Planes with $\phi = 0.001^o$, $d = 1000$

| Terms | Coef. of object | Coef.:$R = 4$ | Coef.:$R = 2$ |
|---|---|---|---|
| $\tilde{X}^3$ | 0.0 | - 0.0018 | -0.0015 |
| $\tilde{Y}^3$ | 5.6991 | 5.7275 | 5.7617 |
| $\tilde{X}^2\tilde{Y}$ | -17.0232 | - 17.054 | -17.176 |
| $\tilde{X}^2\tilde{Z}$ | -6.0040 | - 6.0401 | -6.0134 |
| $\tilde{X}\tilde{Y}^2$ | 0.0 | - 0.0023 | 0.0009 |
| $\tilde{Y}^2\tilde{Z}$ | -6.0314 | - 6.0419 | -6.0901 |
| $\tilde{X}\tilde{Z}^2$ | 0.0 | - 0.0033 | 0.0034 |
| $\tilde{Y}\tilde{Z}^2$ | 0.0 | - 0.0054 | 0.0150 |
| $\tilde{X}\tilde{Y}\tilde{Z}$ | 0.0 | 0.0069 | 0.0010 |
| $\tilde{X}^2$ | 0.0 | 0.0582 | -0.0110 |
| $\tilde{Y}^2$ | 0.0 | - 0.1046 | 0.0014 |
| $\tilde{Z}^2$ | 0.0 | - 0.0508 | -0.0112 |
| $\tilde{X}\tilde{Y}$ | 0.0 | - 0.0411 | -0.0076 |
| $\tilde{X}\tilde{Z}$ | 0.0 | 0.0509 | -0.0213 |
| $\tilde{Y}\tilde{Z}$ | 0.0 | 0.1019 | -0.0110 |
| $\tilde{X}$ | 0.0 | - 0.1534 | 0.0350 |
| $\tilde{Y}$ | 0.0 | 0.0329 | 0.0129 |
| $\tilde{Z}$ | 0.0 | - 0.1916 | 0.0376 |
| constant | 0.0 | 0.2072 | -0.0298 |
| depth | 1000.0 | 1006.82 | 1003.39 |

# Chapter 6

# Conclusion and Discussion

1. The optical flow determined by *Horn* and *Schunck* has been extended to spatial sequences of images resulting in unified optical flow field ( *UOFF*) [13]. The brightness invariant equations in both temporal domain and spatial domain are presented, which can provide more information to recover structure and motion of object efficiently.

2. Based on the *UOFF*, a direct recovering algorithm [14], which involves minimizing the sum of the squared error of the surface polynomial equation over the whole image plane, is presented to estimate all the coefficients of the polynomial equation of a combinational structure. In the thesis work, this method is utilized to tackle the problem of recovering the combination of multiplanar surface.

3. The associated experimental works are conducted to verify the feasibility, efficiency of this new algorithm for recovering the composite multiplanar surface structure. This is a significant progress made compared with the cases where only one planar structure can be recovered [4].

4. In the generation of the simulation stereo images for composite surfaces, care has to be taken to deal with possible multiple correspondence prop-

erly.

5. As shown by the experiments, the setting of two cameras has a great influence upon the result of estimation with this new method. Generally, the choice of the rotation angle $\phi$ between two cameras, the distance $d$ between the camera and object and the radius of the visual field of camera $R$ should be chosen such that the assumption of *far-field* is satisfied. In general, $\phi$ and $R$ should be small and $d$ large properly. This assumption is accorded with the practice of stereo imagery.

6. The estimation about motion of combinational structure is under investigation.

# Bibliography


[1] J.K.Aggarwel and N.Nandhakumar, "On the Computation of Motion from Sequences of Images - A Review," *Proceedings of the IEEE* , Vol.76, No.8, pp.917-935, August, 1988.

[2] B.K.P.Horn and B.G.Schunck, "Determining Optical Flow," *Artificial Intelligence*, pp.185-203, 17(1981).

[3] A.R.Bruss and B.K.P.Horn, "Passive Navigation," *Computer Vision, Graphics, and Image Processing*, Vol.21, No.3-29 pp.3-20, 1983.

[4] S.Negahdaripour and B.K.P.Horn, "Direct Passive Navigation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.PAMI-9, No.1, pp.168-176, Jan.,1987.

[5] B.K.P.Horn and E.J.Weldon Jr., "Direct Methods for Recovering Motion," *Journal of computer Vision*, No.2, pp.51-76, 1988.

[6] J.Heel and S.Negahdaripour, "Time-sequential Structure and Motion Estimation without Optical Flow," *Proceedings of SPIE*, Vol.1260, Sensing and Reconstruction of Three-dimensional objects and Scenes, pp.50-61, Feb., 1990.

[7] S.Negandaripour, B.Hayashi and J.Aloimonos, "Direct Motion stereo for

a Translation Observer," *Proceedings of IEEE conference on Image Processing*, Singapore, Sept., 1989.

[8] J.Aloimonos and J.Herve, " Stereo and Motion: Planar Surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.12, No.5, pp.504-510, May, 1990.

[9] B.Hayashi and S.Negahdaripour, "Direct Motion stereo," *Proceedings of SPIE*, Vol.1260, Sensing and Reconstruction of Three-Dimensional Objects and scenes, pp.78-85, Feb., 1990.

[10] C.Q.Shu and Y.Q.Shi. "A New Approach to Motion Analysis from a Sequence of Stereo Images," *Technical report* No.18, Electronic Imaging Laboratory of electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, 1990.

[11] C.Q.Shu and Y.Q.Shi, "Computation of Motion from Stereo Image sequence using Unified Optical Flow Field," *SPIE's 1990 International Symposium on Optical and Optoelectronic Applied Science and Engineering*, San Diego, CA, July 1990.

[12] C.Q.Shu and Y.Q.Shi, "Unified Optical Flow Field," *Proceedings of the 1990 Conference on Sciences and Systems*, pp.445, Princeton University, NJ, March, 1990.

[13] C.Q.Shu and Y.Q.Shi, "On Unified Optical Flow Field," *Pattern Recognition* (Accepted).

[14] C.Q.Shu, Y.Zhu, Y.Q.Shi and C.H.Lu, "Recovering Surface Structure Characterized by an *Nth* Degree Polynomial Equation," IEEE Seventh

Workshop on Multidimension Signal Processing, Lake Placid, NY, Sept. 1991 (Accpted)

[15] A.M.Warman and J.W.Duncan, "Binocular Image Flow: Steps toward Stereo-motion Fusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.8, No.6, pp.715-729, Nev., 1986.

[16] D.H.Ballard and O.A.Kimball, "Rigid Body Motion from Depth and Optical Flow," *Computer Vision, Graphics and Image Processing*, Vol.12, pp.95-115, 1983.

[17] J.Weng, T.S.Huang and N.Ahuja, "Motion and Structure from Two Perspective Views: Algorithm, Error analysis and error Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.11, No.5, pp.451-476, May 1989

[18] R.C.Conzalez and P.Wintz, Digital Image Processing 2nd Edition, *Addison-Wesley Publishing Company*, New York, 1987.