

Optical flow estimation based on 3-D gradient constraint

Tang, Tianyu

2010

Tang, T. (2010). Optical flow estimation based on 3-D gradient constraint. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/35247>

<https://doi.org/10.32657/10356/35247>

Optical Flow Estimation Based On 3-D Gradient Constraint

Tang Tianyu

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Master of Engineering

2010

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

.....

Date

.....

Tang Tianyu

Acknowledgments

First and foremost, I wish to express my sincere and utmost gratitude to my supervisor, Professor Kai-Kuang Ma, for his earnest instructions, inspiring suggestions and kindly encouragement through my graduate study in NTU. His assiduous endeavor and persistent pursuing in the academic career, as well as his rigorous scholarship not only brings me enormous advantage in my research work, but also sets an excellent example that deserves my lifelong respect.

In addition, it is my great pleasure to thank and acknowledge my fellow graduate students and staffs in Media Technology Laboratory and Temasek Laboratories. I have been benefited a great deal from the genuine collaboration and the in-depth discussion with them, which did tremendous help to my professional growth. Besides, the sincere friendship would be a great treasure in my life as well.

Lastly but most important, this thesis is dedicated to my parents, whose love and support are always the greatest inspiration for me in my life.

Contents

Acknowledgments	i
List of Figures	iv
List of Tables	vii
Summary	viii
1 Introduction	1
1.1 Problems and Major Contributions	2
1.2 Organization of The Thesis	4
2 Basics of Optical Flow Estimation	5
2.1 Motion Field	5
2.2 Optical Flow	7
2.3 Optical Flow Estimation Techniques	9
2.3.1 Gradient-based Regularization Methods	9
2.3.2 Region-based Matching Approaches	15
2.3.3 Phase-based Differential Algorithms	16
3 Three-dimensional (3-D) Gradient Constraint	18
3.1 Optical Flow for 1-D Time-variant Signal	19

CONTENTS

CONTENTS

3.1.1	The Velocity Estimator for 1-D Time-variant Signal	19
3.1.2	The Gradient Constraint for 1-D Time-variant Signal	20
3.2	The 3-D Gradient Constraint	23
3.2.1	The Spatio-temporal Representation of Video	24
3.2.2	Definition of the Constraint Plane	25
3.2.3	Construct the Constraint Plane	26
3.2.4	The Direction of Constant Brightness	28
3.3	The Standard Aperture Problem	31
3.3.1	The Spatio-temporal Image on the Constraint Plane	31
3.3.2	Estimate the Direction of Image Gradient for Spatio-temporal Image	32
3.3.3	Condition for Solving The Standard Aperture Problem	35
3.4	The Optical Flow Estimation Scheme	36
4	Experimental Results	37
4.1	Implementation	37
4.2	Results	38
5	Conclusions	47

List of Figures

1.1	An example of optical flow field. (a) The 8th frame of the <i>Sphere</i> sequence. (b) The ground truth of the optical flow obtained from frames 8 and 9.	2
1.2	The optical flow solution (u, v) is restricted on a constraint line. . .	3
2.1	The projection of a 3-D motion on a 2-D translation on the image plane.	6
2.2	The optical flow is not always equal to the motion field. (a) A smooth sphere <i>rotating</i> under constant illumination. (b) A <i>stationary</i> smooth sphere illuminated by a moving light source.	7
2.3	The optical flow and the motion field of a rotating barber's pole. . .	8
2.4	Analysis of optical flow components. The full flow \mathbf{v} is orthogonally decomposed into the normal flow \mathbf{v}_{NF} and the tangent flow \mathbf{v}_{TF} . Only the normal flow could be determined directly by the constraint line.	10
3.1	The linear approximation of the translation for 1-D signals. (a) Estimate the translation for a linear signal. (b) Estimate the translation for a nonlinear signal.	20

LIST OF FIGURES

LIST OF FIGURES

3.2	Estimate optical flow from spatio-temporal image of the 1-D shift sine pattern. (a) A shifting 1-D sine-wave pattern. (b) The spatio-temporal image of (a). (c) The gradient of the spatio-temporal image. (d) The perpendicular relationship between the image gradient (I_x, I_t) and the direction of constant brightness $(v, 1)$	22
3.3	The 3-D spatio-temporal cube for 2-D video sequence.	23
3.4	Sample the spatio-temporal cube. (a) Sample the cube with a plane parallel to the x - t plane. (b) The sample result of (a).	24
3.5	The constant plane is perpendicular to the local gradient (I_x, I_y, I_t)	26
3.6	Sample the constraint plane. (a) The intersection lines on the constraint plane. (b) Sample the intersection lines.	27
3.7	The distance between intersection lines in spatio-temporal space.	28
3.8	The direction of constant brightness on the constraint plane. (a) The nine points on the constraint plane. (b) Normalize the distance between the intersection lines.	29
3.9	The decomposition of the direction of constant brightness.	30
3.10	The derivative would be zero at local extremum.	32
3.11	Estimate local gradient of a pixel by its 8-adjacent. (a) The 3×3 sample from the spatio-temporal image. (b) The direction of maximum intensity change and the corresponding velocity.	33
3.12	Estimate the direction of image gradient based on the absolute intensity difference along the 4 directions.	34

LIST OF FIGURES

LIST OF FIGURES

4.1	Estimated optical flow field for sequence <i>Sine</i> . (a) The 9th frame of <i>Sine</i> . (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.	42
4.2	Estimated optical flow field for sequence <i>Translating Tree</i> . (a) The 9th frame of <i>Translating Tree</i> . (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.	43
4.3	Estimated optical flow field for sequence <i>Diverging Tree</i> . (a) The 9th frame of <i>Diverging Tree</i> . (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.	44
4.4	Estimated optical flow field for sequence <i>Yosemite</i> . (a) The 9th frame of <i>Yosemite</i> . (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter.. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.	45
4.5	Estimated optical flow field for sequence <i>Taxi</i> . (a) The 9th frame of <i>Taxi</i> . (b) Horn and Schunck 100 iterations. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.	46

List of Tables

4.1	Precision comparison in AAE and STD.	39
4.2	Precision comparison in error distribution (%).	39
4.3	Precision comparison of the sequence <i>Yosemite</i>	40
4.4	Execution time.	41

Summary

The research work presented in this thesis is particularly focused on a fundamental investigation and development of the *optical flow estimation*, which is considered as one of the fundamental computation tasks often required in computer vision. The goal of the optical flow computation is to extract the two-dimensional velocity information at each pixel on the individual frame of a video sequence.

The fundamental constraint equation for *gradient*-based optical flow estimation is the gradient constraint which restricts the optical flow solution by the spatial and temporal derivatives of the pixel intensity on video frames. However, only the *normal flow*, which is one of the two optical flow components that is parallel to the image gradient, can be computed. On the other hand, the *tangent flow*, which is another optical flow component that is perpendicular to the image gradient, cannot be calculated from the gradient constraint. This is known as the *standard aperture problem*, and the tangent flow has to be estimated by making an additional smoothness assumption over the motion field. However, such assumption might not be effective, especially for the boundaries of moving objects or non-rigid objects, besides the computational complexity issue.

To address the above-mentioned issues, a *three-dimensional gradient constraint* is proposed in this research to calculate the optical flow without imposing any hypothesis on the smoothness of the motion field. By representing a video as a spatio-temporal cube, the constraint *line* of the conventional gradient constraint

Summary

is replaced by the constraint *plane* which is perpendicular to the local gradient of the spatio-temporal cube. It is further shown that the tangent flow could be calculated on the constraint plane by finding the *direction of constant brightness* which is a 3-D vector from the local pixel to its counterpart on the next frame in the spatio-temporal cube.

As the tangent flow could be obtained based on the proposed 3-D gradient constraint, the standard aperture problem could be effectively and efficiently solved. The condition for solving this problem is derived by analyzing the spatio-temporal image of the constraint plane. Based on this analysis, an estimator for estimating the direction of image gradient is developed to replace the regular image gradient estimators, such as the Sobel estimator, which may fail to extract the variational information from the spatio-temporal image on the constraint plane.

In addition, because the optical flow are computed from our derived closed-form equations, no iterative computations are involved. This yields a relatively low computational complexity and has a good potential to provide real-time performance on computing the optical flows.

The performance of the proposed optical flow estimation method is evaluated in terms of the optical flow precision and the runtime. Simulation results clearly demonstrate that the developed optical flow estimation method is able to obtain a fairly precise optical flow field with very low computational complexity.

Chapter 1

Introduction

Motion is the most salient temporal information perceived by the human visual system to detect any movement that might be incurred in our surroundings. In digital video processing, motion information can be extracted from the intensity change at each pixel based on the consecutive frames. The obtained motion field could be further utilized in many video processing and computer vision tasks, including video coding, video object segmentation, 3-D shape acquisition and scene understanding, etc.

Generally, the motion estimation methods could be divided into two categories: *optical flow estimation* (e.g., [1]) and *block-matching algorithm* (e.g., [2]). In this thesis, the research work is only focused on the former, from which the obtained optical flow presents the apparent motion of brightness pattern derived from two consecutive video frames [3] (see Figure 1.1 for a demonstration). In general, there are three categories of optical flow estimation techniques that have been developed: *gradient*-based regularization methods (e.g., [1, 4]), *region*-based matching approaches (e.g., [5]) and *phase*-based differential algorithms (e.g., [6]). They will be succinctly discussed in Chapter 2 as the essential background. The gradient-based methods will be further studied for the development of a new optical flow estimation scheme.

1.1 Problems and Major Contributions

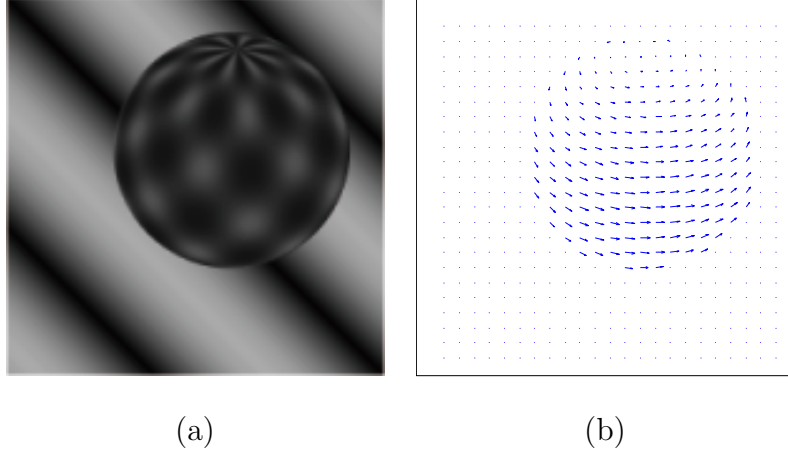


Figure 1.1. An example of optical flow field. (a) The 8th frame of the *Sphere* sequence. (b) The ground truth of the optical flow obtained from frames 8 and 9.

1.1 Problems and Major Contributions

The fundamental constraint for gradient-based optical flow estimation is the gradient constraint proposed by Horn and Schunck [1], from which the optical flow is only restricted on a *constraint line* (Figure 1.2). As the constraint line is a linear equation in two variables, the optical flow solution could not be determined based on the gradient constraint. This is known as the *standard aperture problem* [3, 7]. Due to this problem, the optical flow can only be estimated by imposing a smoothness assumption on the motion field. For that, various smoothness constraints for regularization [1, 8] as well as the least-squares or the tensor-based computational methods [4, 9] are developed to estimate the optical flow field. In addition, to make the estimation more robust against image noise, illumination change and object occlusion, different energy functionals [10, 11] and probabilistic models [12, 13] have been attempted. To preserve more accurate boundary of the moving objects in the motion field, parametric models [14, 15] and piecewise-smooth hypothesis [16, 17] are incorporated, on the expense of increased computational complexity. Besides, these methods still fail to reasonably handle non-rigid moving object deformation.

1.1 Problems and Major Contributions

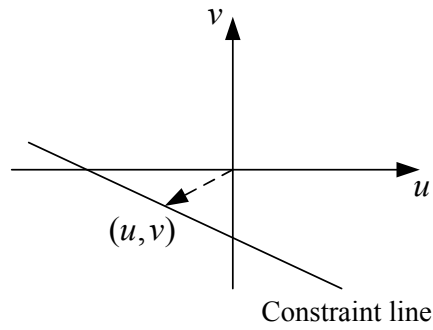


Figure 1.2. The optical flow solution (u, v) is restricted on a constraint line.

This is due to the fact that the estimation of optical flow based on the gradient constraint relies heavily on the smoothness assumption of the motion field.

To address these issues, the *3-D gradient constraint* is proposed to calculate the optical flow without imposing any hypothesis on the smoothness of the motion field. In this constraint, a constraint *plane* is introduced to constrain the optical flow components at each pixel position in the spatio-temporal cube of a video, which replaces the constraint *line* of Horn and Schunck. And it is derived that the spatio-temporal image on the constraint plane could be utilized to calculate the *tangent flow*, which can not be estimated by the conventional gradient constraint. This implies that both the normal flow and the tangent flow could be computed at each pixel position, and the standard aperture problem could be solved. The condition for solving the standard aperture problem is derived by analyzing the spatio-temporal image of the constraint plane. In addition, a new image gradient estimator is introduced to replace the regular gradient estimator which may fail to estimate the image gradient of the spatio-temporal image on the constraint plane. Moreover, the proposed optical flow estimation method has a good potential to provide real-time performance as the optical flow is calculated from a set of closed-form equations.

1.2 Organization of The Thesis

The rest of this thesis is organized as follows. Chapter 2 provides the fundamental knowledge of optical flow and a comprehensive review of the conventional optical flow estimation techniques. In Chapter 3, a new *three-dimensional gradient constraint* is proposed, which extends the conventionally exploited constraint *line* to the constraint *plane*, from which the tangent flow could be estimated. The condition for solving the standard aperture problem is derived, and a image gradient estimator is introduced in this chapter. In Chapter 4, the experimental simulation results are presented in terms of the optical flow precision and the runtime to demonstrate the accuracy of computed optical flow by using the proposed approach. Conclusions of this thesis are drawn in Chapter 5.

Chapter 2

Basics of Optical Flow Estimation

In this chapter, the fundamentals of the optical flow is presented as the essential background for the understanding and development of this thesis. Three categories of optical flow estimation methods are succinctly reviewed; they are: gradient-based regularization methods, region-based matching approaches and phase-based differential algorithms.

2.1 Motion Field

When a video camera is used to take a video sequence of the 3-D scene in the real world, the 3-D motion of an object in this scene is projected onto the camera sensor and appears as the 2-D position change of the object's image between consecutive frames of the video. Based on the principle of optics, this 2-D position change is the perspective projection of the object's 3-D motion relative to the camera in the real world.

In computer vision, *motion field* is defined as the ideal representation of a 3-D motion when it is projected onto the image plane of a video camera [3]. As shown in Figure 2.1, at a particular time instant t , the point P_i acquired by the camera

2.1 Motion Field

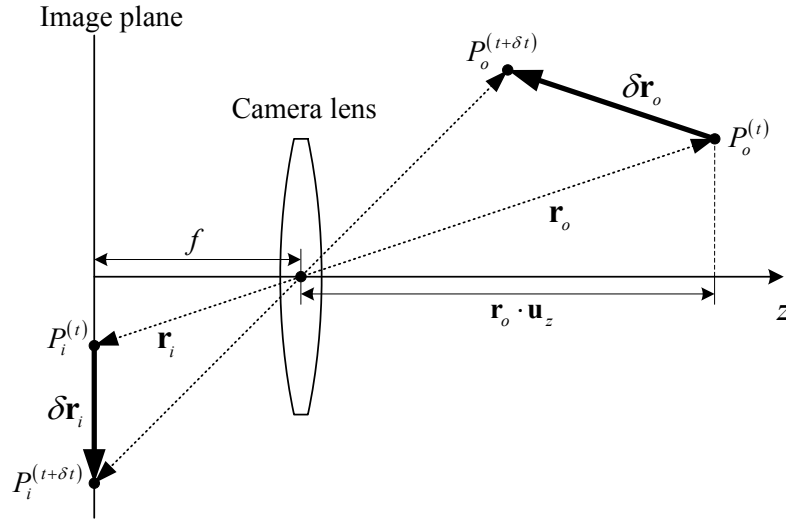


Figure 2.1. The projection of a 3-D motion on a 2-D translation on the image plane.

lens and presented in the image plane corresponds to the point P_o on the surface of an object by the perspective projection. Let \mathbf{r}_o and \mathbf{r}_i be the vector from the optical center of the lens to P_o and P_i , respectively. In a short time interval δt , the motion of P_o and P_i is represented as $\delta \mathbf{r}_o$ and $\delta \mathbf{r}_i$, respectively (Figure 2.1). As the time interval δt is very short, the velocities of P_o and P_i could be given as

$$\mathbf{v}_o = \frac{d\mathbf{r}_o}{dt} \quad \text{and} \quad \mathbf{v}_i = \frac{d\mathbf{r}_i}{dt}, \quad (2.1)$$

where \mathbf{r}_o and \mathbf{r}_i are related by

$$\frac{\mathbf{r}_i}{f} = -\frac{\mathbf{r}_o}{\mathbf{r}_o \cdot \mathbf{u}_z}. \quad (2.2)$$

Here, f is the focal length of the camera lens, and \mathbf{u}_z corresponds to the unit vector on z -axis which is perpendicular to the image plane. By differentiating this perspective projection equation with respect to time t , it can be arrived at

$$\frac{\mathbf{v}_i}{f} = -\frac{(\mathbf{r}_o \cdot \mathbf{u}_z)\mathbf{v}_o - (\mathbf{v}_o \cdot \mathbf{u}_z)\mathbf{r}_o}{(\mathbf{r}_o \cdot \mathbf{u}_z)^2} = -\frac{(\mathbf{r}_o \times \mathbf{v}_o) \times \mathbf{u}_z}{(\mathbf{r}_o \cdot \mathbf{u}_z)^2}. \quad (2.3)$$

Equation (2.3) reveals the relation of perspective projection between the motion \mathbf{v}_o in the 3-D scene and its projected counterpart \mathbf{v}_i on the image plane.

2.2 Optical Flow

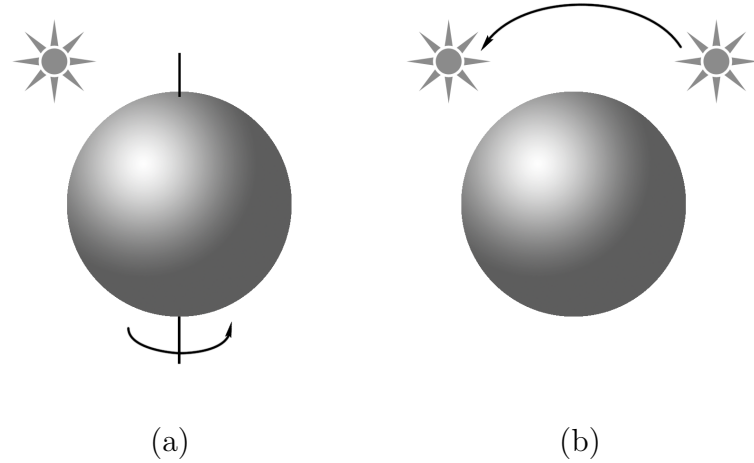


Figure 2.2. The optical flow is not always equal to the motion field. (a) A smooth sphere *rotating* under constant illumination. (b) A *stationary* smooth sphere illuminated by a moving light source.

The motion field is composed of the velocity vectors projected onto the image plane as described above. Here, the perspective projection is shown to illustrate the scheme that the motion vectors are obtained to construct the motion field. In the following part of this thesis, we will focus on the analysis of the 2-D motion in video sequences.

2.2 Optical Flow

Unlike the motion field, which is the projection of the 3-D *real* motion in the scene, the so-called *optical flow* is the *apparent* motion of the brightness pattern (spatial variation of image intensity, such as textures and edges) between consecutive frames of a video sequence [3]. The difference between the motion field and the optical flow could be demonstrated by the following two examples.

In the first example, consider a uniform sphere with smooth surface rotating about an axis across its spherical center under a fixed illumination (Figure 2.2(a)). Because of its uniform shape and textureless surface, the brightness pattern on

2.2 Optical Flow

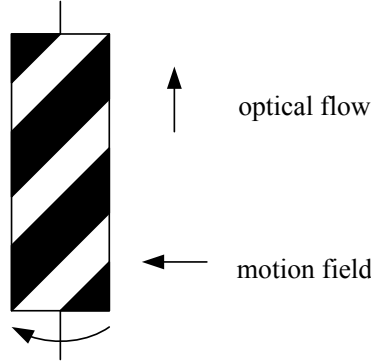


Figure 2.3. The optical flow and the motion field of a rotating barber's pole.

the sphere will not change accordingly with the rotation of the sphere. Therefore, the optical flow is *zero* everywhere, while the motion field is *nonzero* owing to the rotation of the sphere. In contrast to the setup of the above-mentioned experiment, let us consider the same sphere without any movement illuminated by a moving light source (Figure 2.2(b)). Along with the movement of the light source, the position of the shading on the sphere surface will be changed, and this will yield the movement of brightness patterns across image frames. The corresponding optical flow is obviously *nonzero*, although the motion field of the sphere is in fact *zero* everywhere. Another famous example is the *barber's pole illusion* [18] (Figure 2.3), in which the motion field (i.e., real motion) and the optical flow (i.e., apparent motion) are perpendicular to each other.

Although the motion field and the optical flow are different in some occasions, only the optical flow could be extracted from a video sequence. Therefore, the optical flow would still be investigated as an approximation of the motion field to obtain the motion of the objects. The basic formulation of optical flow is given as follows.

Let $I(x, y, t)$ be the pixel intensity at position (x, y) in the image plane at time t , which reflects the brightness of a point in the real-world scene. After a lapse of time δt , the same point is reflected at the position $(x + \delta x, y + \delta y)$ in the image plane

2.3 Optical Flow Estimation Techniques

at time $t + \delta t$. In optical flow estimation, the intensity of this point is expected to be unchanged during such a short time interval. That is

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t), \quad (2.4)$$

Let (u, v) denote the optical flow vector representing the 2-D velocity in the image plane; here, $u = \delta x / \delta t$ and $v = \delta y / \delta t$. Equation (2.4) could be represented as

$$I(x + u\delta t, y + v\delta t, t + \delta t) = I(x, y, t). \quad (2.5)$$

As the time interval δt is very short, (2.5) can further be derived as

$$\frac{d}{dt}I(x, y, t) = 0. \quad (2.6)$$

Equation (2.5) is the so-called *constant intensity constraint* [3] used for optical flow computation. That is, the velocity vector $\mathbf{v} = (u, v)$ satisfying (2.5) is the optical flow at point (x, y) . Equivalently, the optical flow can be computed by minimizing $dI(x, y, t)/dt$. And the corresponding motion vector is known as the velocity with minimum intensity change.

2.3 Optical Flow Estimation Techniques

In general, the optical flow estimation techniques can be divided into three categories: *gradient*-based regularization methods (e.g., [1, 4]), *region*-based matching approaches (e.g., [5]) and *phase*-based differential algorithms (e.g., [6]). In these methods, different *shift-invariant* properties of pixel intensity, block pattern or image phase are utilized respectively, with more details as follows.

2.3.1 Gradient-based Regularization Methods

Known as the *differential techniques* or the *variational methods*, the gradient-based methods have been extensively investigated since the novel work done by Horn

2.3 Optical Flow Estimation Techniques

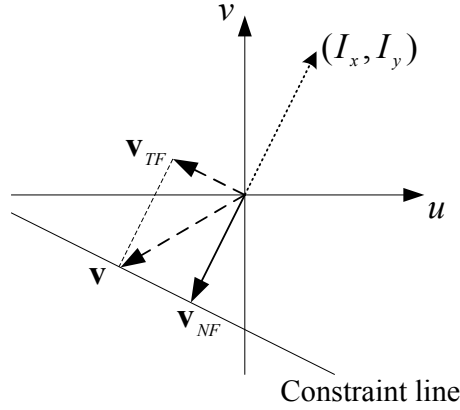


Figure 2.4. Analysis of optical flow components. The full flow \mathbf{v} is orthogonally decomposed into the normal flow \mathbf{v}_{NF} and the tangent flow \mathbf{v}_{TF} . Only the normal flow could be determined directly by the constraint line.

and Schunck [1], and have been proved to provide very precise optical flow field [19, 20, 21].

The gradient-based methods are based on the *gradient constraint*, which is first introduced by Horn and Schunck [1], nearly three decades ago. Serving as the fundamental constraint in optical flow estimation, the gradient constraint relates the optical flow components with spatial and temporal derivatives of the image intensity on consecutive frames of a video sequence.

The *gradient constraint* is derived from (2.6) by using the chain rule of differentiation with respect to t , which is

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0. \quad (2.7)$$

As the two components of optical flow $u = dx/dt$ and $v = dy/dt$, (2.7) can be written as a linear equation in the two variables u and v ,

$$I_x u + I_y v + I_t = 0, \quad (2.8)$$

where the abbreviations I_x , I_y and I_t are the partial derivatives for image intensity with respect to x , y and t . As shown in Figure 2.4, (2.8) defines the *constraint line*,

2.3 Optical Flow Estimation Techniques

on which the optical flow is confined. However, it's impossible to determine the two unknowns u and v uniquely with merely one linear equation.

As shown in Figure 2.4, the so-called *full flow* \mathbf{v} is composed of two orthogonal components: the *normal flow* \mathbf{v}_{NF} and the *tangent flow* \mathbf{v}_{TF} , where the normal flow is parallel to the image gradient (I_x, I_y) and the tangent flow is perpendicular to the normal flow. In (2.8), if we move I_t to the right side and rewrite the left side as the inner product of two vectors, it can be represented as

$$(I_x, I_y) \cdot (u, v) = -I_t. \quad (2.9)$$

Because the normal flow is parallel to the image gradient, the normal flow $\mathbf{v}_{NF} = (u_{NF}, v_{NF})$ could be obtained by representing (u_{NF}, v_{NF}) as

$$(u_{NF}, v_{NF}) = \alpha \cdot \frac{(I_x, I_y)}{\|(I_x, I_y)\|}, \quad (2.10)$$

where α is a scalar, and $(I_x, I_y)/\|(I_x, I_y)\|$ is the unit vector parallel to the image gradient. By substituting (2.10) into (2.9), it can be calculated that $\alpha = -I_t/\|(I_x, I_y)\|$, which yields

$$(u_{NF}, v_{NF}) = -\frac{I_t}{\|(I_x, I_y)\|} \cdot \frac{(I_x, I_y)}{\|(I_x, I_y)\|}. \quad (2.11)$$

However, the tangent flow \mathbf{v}_{TF} can not be estimated directly based on the gradient constraint, which is referred to as the *standard aperture problem* [3, 16]. Due to such limitation of gradient constraint, two branches of methods, the *global smoothness constraint* and the *least-squares minimization* are proposed to further constrain the optical flow field.

i) The Global Smoothness Constraint

In this category of methods, the gradient constraint (2.8) is combined with a smoothness term to establish an energy functional $E(u, v)$, subjected to be minimized globally. Therefore, these methods are also known as the *global methods*.

2.3 Optical Flow Estimation Techniques

Taking the Horn and Schunck method [1] as an example, the optical flow is calculated by minimizing the energy functional defined over a domain D , as

$$E(u, v) = \iint_D (I_x u + I_y v + I_t)^2 + \lambda (\|\nabla u\|_2^2 + \|\nabla v\|_2^2) dx dy, \quad (2.12)$$

where the regularization parameter λ is determined experimentally and used to reflect the influence of the smoothness term $(\|\nabla u\|_2^2 + \|\nabla v\|_2^2)$. Normally, the optical flow is calculated by Gauss-Seidel iterative equations, which yield very high computational load. Although the multigrid algorithm is developed to achieve fast calculation [22, 23], the optical flow estimation procedure is still getting more and more complicated as more models are incorporated.

Due to the different motion of the moving objects and the background, this smoothness assumption of the motion field would be violated at the boundary of the moving object. In order to handle such problem, Nagel and Enkelmann [8] introduced the *oriented-smoothness constraint* in which smoothness is not imposed across steep intensity gradients (edges). In order to achieve piecewise smoothness of the optical flow field, another approach was proposed by Black and Jepson [14], in which the parametric flow model was used to estimate the motion in segmented and planar hypothesized regions with piecewise smooth brightness. Weickert and Schnörr [24] introduced a nonlinear spatio-temporal regularization method based on non-quadratic convex variational regularization which yields discontinuity-preserving computation of optical flow. For the same purpose, a multi-cue driven adaptive bilateral filter is utilized to regularize the flow field in the method proposed by Xiao *et al.* [17].

In order to cope with large displacement as well as reduce the risk to be trapped by irrelevant minimum, the multi-resolution technique and the coarse-to-fine strategy were widely adopted [25, 26, 27]. Brox *et al.* [10] used an energy function combining the brightness constancy, the gradient constancy and the discontinuity-

2.3 Optical Flow Estimation Techniques

preserving spatio-temporal smoothness constraint. The coarse-to-fine warping strategy is also incorporated to make the algorithm insensitive to parameter variation and relatively more robust under noise.

An additional approach to acquire robust optical flow estimation is to combine the global method (i.e., the global smoothness constraint) and the local method (i.e., the least-squares minimization), which was proposed by Bruhn *et al.* [11]. It is not only more robust under noise, but also yields dense flow fields, which takes the advantage of the global and local methods.

Although various smoothness constraints, regularizers and computational approaches are designed, incorporating parametric models as well as probability models [13], accurate motion vectors for non-rigid objects with deformed motion boundaries are still difficult to obtain. This is due to the fact that the variational information is only extracted in the direction of the image gradient. Furthermore, the motion component perpendicular to the image gradient has to be estimated from the local region of the pixel by imposing the smoothness assumption, from which, however, such smoothness often gets violated especially in the motion boundary, with illumination change and in occlusion areas.

ii) The Least-squares Minimization

The least-squares minimization approaches are also known as the *local methods*. A weighted least-squares of the first-order gradient constraints (2.8) is used in this category of methods to calculate the optical flow [4, 9]. In each spatial neighborhood Ω , the optical flow is obtained by minimizing

$$E(u, v) = \sum_{\mathbf{x} \in \Omega} W(\mathbf{x})(I_x u + I_y v + I_t)^2, \quad (2.13)$$

where $\mathbf{x} = (x, y)$, and $W(\mathbf{x})$ denotes a weighting function. For that, a Gaussian function is commonly used, giving more influence to the center of the neighborhood.

2.3 Optical Flow Estimation Techniques

In the approach of Lucas and Kanade [4], the solution to the minimization of (2.13) is obtained by closed-form equations using the *standard* least-squares estimation. Unlike this approach, the *structure tensor technique* [28] solves the minimization of (2.13) by using the *total* least-squares estimation which was first introduced by Bigün *et al.* [9]. In the approach proposed by Liu *et al.* [15], the parametric flow model was combined with the 3-D structure tensor to further improve the accuracy of the optical flow.

In these approaches, because the motion of each pixel within the neighborhood Ω is assumed to be the same, which might easily be violated, the robust model designed by Black and Anandan [16] was widely utilized to remove the outliers from the neighborhood. Probabilistic models [12, 29] were also incorporated in the estimation of optical flow to act against the uncertainty posed by the image noise and illumination change. However, the estimation of full flow field still relies on the smoothness assumption of the motion field.

iii) Second-order Approaches

Nagel [30] and Uras *et al.* [31] used the second-order derivatives to constrain the optical flow, which assumes $d\nabla I(x, y, t)/dt = 0$ as a stronger restriction than the gradient constraint (2.8). That is,

$$\begin{cases} I_{xx}u + I_{xy}v + I_{xt} = 0; \\ I_{xy}u + I_{yy}v + I_{yt} = 0. \end{cases} \quad (2.14)$$

However, higher-order derivatives are often extremely noisy, and the conservation of $d\nabla I(x, y, t)/dt = 0$ implies that the motion field has no first order deformation, which might also be violated easily.

2.3 Optical Flow Estimation Techniques

2.3.2 Region-based Matching Approaches

In this category of methods, the optical flow \mathbf{v} is defined as the shift \mathbf{d} that corresponds to the best fit between image regions in neighboring frames. We can utilize the minimization of a distance measure, such as the sum-of-squared difference (SSD) to find the best match, which is,

$$\begin{aligned} SSD_{1,2}(\mathbf{x}; \mathbf{d}) &= \sum_{j=-n}^n \sum_{i=-n}^n W^2(i, j) [(I_1(\mathbf{x} + (i, j)) - I_2(\mathbf{x} + \mathbf{d} + (i, j)))^2] \\ &= W(\mathbf{x}) * [(I_1(\mathbf{x}) - I_2(\mathbf{x} + \mathbf{d}))^2] \end{aligned} \quad (2.15)$$

where $\mathbf{x} = (x, y)$, $W(\mathbf{x})$ denotes a 2-D window function, and $\mathbf{d} = (d_x, d_y)$ takes only integer values. It is not too different to use the cross-correlation similarity measure and differential techniques: minimizing the SSD distance is equal to maximizing the integral of the term $I_1(\mathbf{x})I_2(\mathbf{x}+\mathbf{d})$; the SSD in (2.15) can also be viewed as the weighted least-squares first-order gradient constraint, except the integer results.

Using only SSD minimization might be enough for video coding tasks, but for optical flow computation, it's only adequate. To make the matching algorithm produce a better flow field, Anandan [5] introduced a technique based on the Laplacian pyramid and the coarse-to-fine SSD-based matching strategy. Anandan also introduced a smoothness constraint on the estimated velocities to produce smooth flow fields. Another SSD-based matching method proposed by Singh [32] uses two stages to get the optical flow: the first stage computes the SSD_0 , which is,

$$SSD_0(\mathbf{x}, \mathbf{d}) = SSD_{0,1}(\mathbf{x}, \mathbf{d}) + SSD_{0,-1}(\mathbf{x}, \mathbf{d}) \quad (2.16)$$

on band-pass filtered images, and the SSD_0 is converted into a probability distribution to acquire sub-pixel velocity; the second stage propagates the velocity using neighborhood constraint, which assumes that a weighted least-squares velocity estimation could be obtained from the velocities of its neighboring pixels.

2.3 Optical Flow Estimation Techniques

2.3.3 Phase-based Differential Algorithms

This category of techniques are defined in terms of the phase behavior of band-pass filter outputs. Fleet and Jepson [6, 33] assumed that the phase would conserve in each band, into which the image was decomposed. The output of band-pass filter is complex-valued and can be written as

$$R(\mathbf{x}, t) = \rho(\mathbf{x}, t) \exp[i\phi(\mathbf{x}, t)], \quad (2.17)$$

where $\rho(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$ are the amplitude part and the phase part of $R(\mathbf{x}, t)$. The optical flow is calculated by applying the phase-based gradient constraint to the phase component of $R(\mathbf{x}, t)$, which is quite similar to the gradient-based approach:

$$\phi_x(\mathbf{x}, t)u + \phi_y(\mathbf{x}, t)v + \phi_t(\mathbf{x}, t) = 0. \quad (2.18)$$

However, because the phase is a multi-function, only uniquely defined on intervals of width 2π , it's difficult to make explicit differentiation. Instead, it's convenient to exploit the following equations to compute the derivatives

$$\phi_x(\mathbf{x}, t) = \frac{\text{Im}[R^*(\mathbf{x}, t)R_x(\mathbf{x}, t)]}{|R(\mathbf{x}, t)|^2}, \quad (2.19)$$

where R^* is the complex conjugate of R , $\text{Im}[R]$ denotes the imaginary part of R , and $R_x = \partial R / \partial x$.

The use of phase is motivated by the fact that the stability of the phase component of band-pass filter output is much higher than the amplitude component [34]. However, the phase would not be so stable around the neighborhood of the phase singularities.

As the development of the wavelet theory, wavelet-based transforms are also used as band-pass filter to extract the phase which would be used in the computation of optical flow. Magarey and Kingsbury [35] proposed a *complex discrete wavelet transform* (CDWT) based method, which builds a multi-resolution pyramid

2.3 Optical Flow Estimation Techniques

to achieve the coarse-to-fine strategy. Bayro-Corrochano [36] introduced a method based on *quaternion wavelet transform* (QWT), which serves as an extension of the discrete wavelet transform and combines the advantage of the wavelet and the quaternion algebra.

Chapter 3

Three-dimensional (3-D) Gradient Constraint

In this chapter, the velocity estimator for 1-D time-variant signal is introduced as the foundation for deriving the proposed constraint. By extending the conventional gradient constraint, the *3-D gradient constraint* is proposed by formulating the constraint *plane*, on which the *tangent flow* could be uniquely determined. Consequently, the *standard aperture problem* could be solved based on the 3-D gradient constraint, and the solvability condition for this problem is given as well. Based on the analysis of the spatio-temporal image of the constraint plane, a gradient estimator is developed to estimate the direction of image gradient on the spatio-temporal image of the constraint plane. The optical flow estimation scheme based on the proposed 3-D gradient constraint is shown in the end.

3.1 Optical Flow for 1-D Time-variant Signal

3.1 Optical Flow for 1-D Time-variant Signal

3.1.1 The Velocity Estimator for 1-D Time-variant Signal

As defined in (2.5), the optical flow estimation is based on the assumption of brightness constancy. Although such assumption might be violated by illumination change, object occlusion and specular reflection, it still performs well in most cases. In this section, we would still utilize this brightness constancy constraint as our basic hypothesis to derive the optical flow formulas.

For the convenience of demonstration, we would first show the derivation of the velocity estimator for a 1-D time-variant signal, which will be used in the following sections. Let $f_1(x)$ and $f_2(x)$ be the intensity of a 1-D ‘video’ at two time instants, where $f_2(x)$ is the translated version of $f_1(x)$. That is

$$f_2(x) = f_1(x - d), \quad (3.1)$$

where d denotes the translation. Expanded in the form of a Taylor series, $f_1(x - d)$ is given by

$$f_1(x - d) = f_1(x) - df'_1(x) + O(d^2), \quad (3.2)$$

which can be rewritten as the difference between the two signals at location x :

$$f_1(x) - f_2(x) = df'_1(x) + O(d^2). \quad (3.3)$$

Ignoring the second and the higher order terms, the linear approximation to d could be obtained as

$$\hat{d} = \frac{f_1(x) - f_2(x)}{f'_1(x)}. \quad (3.4)$$

As shown in Figure 3.1, equation (3.4) gives exactly the displacement for a linear signal as the ignored higher order component $O(d^2)$ is equal to 0. For a nonlinear signal, however, what (3.4) gives is an approximation of the translation [33], which ignores the higher order terms of corresponding Taylor series. Assuming

3.1 Optical Flow for 1-D Time-variant Signal

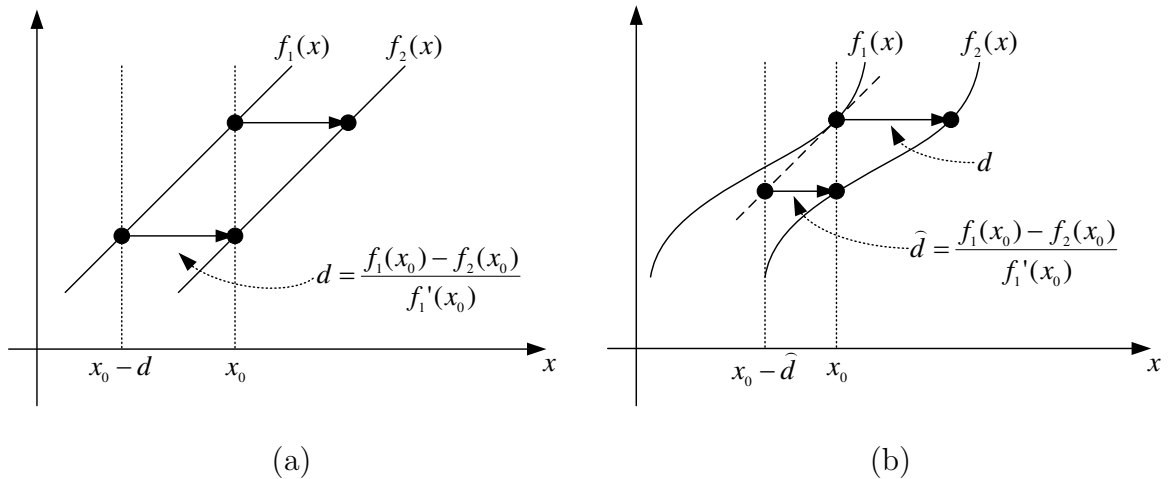


Figure 3.1. The linear approximation of the translation for 1-D signals. (a) Estimate the translation for a linear signal. (b) Estimate the translation for a nonlinear signal.

the displacement is well approximated by (3.4), if we replace $f_1(x)$ and $f_2(x)$ with $I(x, t)$ and $I(x, t + \delta t)$, respectively, and divide both sides of the equation by δt , (3.4) can be written as

$$\frac{\hat{d}}{\delta t} = \frac{I(x, t) - I(x, t + \delta t)}{I_x(x, t)\delta t}. \quad (3.5)$$

Let $v = \hat{d}/\delta t$ and $\delta t \rightarrow 0$, it becomes

$$v = -\frac{I_t(x, t)}{I_x(x, t)}, \quad (3.6)$$

which gives the linear estimation of the optical flow for the 1-D ‘video’. It’s obvious that such estimation approximates the local area of the signal as linear. Besides, the estimation of the 1-D velocity is not valid when $I_x(x, t) = 0$.

3.1.2 The Gradient Constraint for 1-D Time-variant Signal

In this part, the fundamental perpendicular relationship between the image gradient and the direction of constant brightness on the *spatio-temporal image* of the 1-D time-variant signal would be demonstrated in deriving the gradient constraint for the 1-D signal.

3.1 Optical Flow for 1-D Time-variant Signal

Similar with the derivation of the gradient constraint for 2-D video sequences, the velocity for the 1-D ‘video’ can be derived from $dI(x, t)/dt = 0$ by using the chain rule, which is

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial t} = 0. \quad (3.7)$$

As the 1-D velocity $v = dy/dt$, (3.7) can be written as a linear equation in the variable v ,

$$I_x v + I_t = 0, \quad (3.8)$$

where the abbreviations I_x and I_t are the partial derivatives for image intensity with respect to x and t . The solution of the 1-D velocity v to (3.8) is

$$v = -\frac{I_t}{I_x}, \quad (3.9)$$

which is the same as the linear approximation of the velocity (3.6). Unlike the situation in the gradient constraint for 2-D video sequences, the gradient constraint for the 1-D ‘video’ has a unique solution.

By parallel putting the intensity of a 1-D time-variant signal one after another according to its time label, the 1-D signal can be represented as a 2-D image, which is named as the *spatio-temporal image* with one spatial coordinate x and one temporal coordinate t . Because $I(x, t)$ is equal to $I(x + v, t + 1)$ under the constant intensity assumption, the vector $(v, 1)$ points to the *direction of constant brightness* on the spatio-temporal image. As (3.8) can be represented as the inner product of two vectors

$$(I_x, I_t) \cdot (v, 1) = 0, \quad (3.10)$$

the image gradient of the spatio-temporal image is perpendicular to the direction of constant brightness, which can be illustrated by the following example.

Figure 3.2(b) shows the spatio-temporal image of the shift 1-D sine pattern (Figure 3.2(a)), in which the sine pattern is moving 0.4 pixel to the right by each

3.1 Optical Flow for 1-D Time-variant Signal

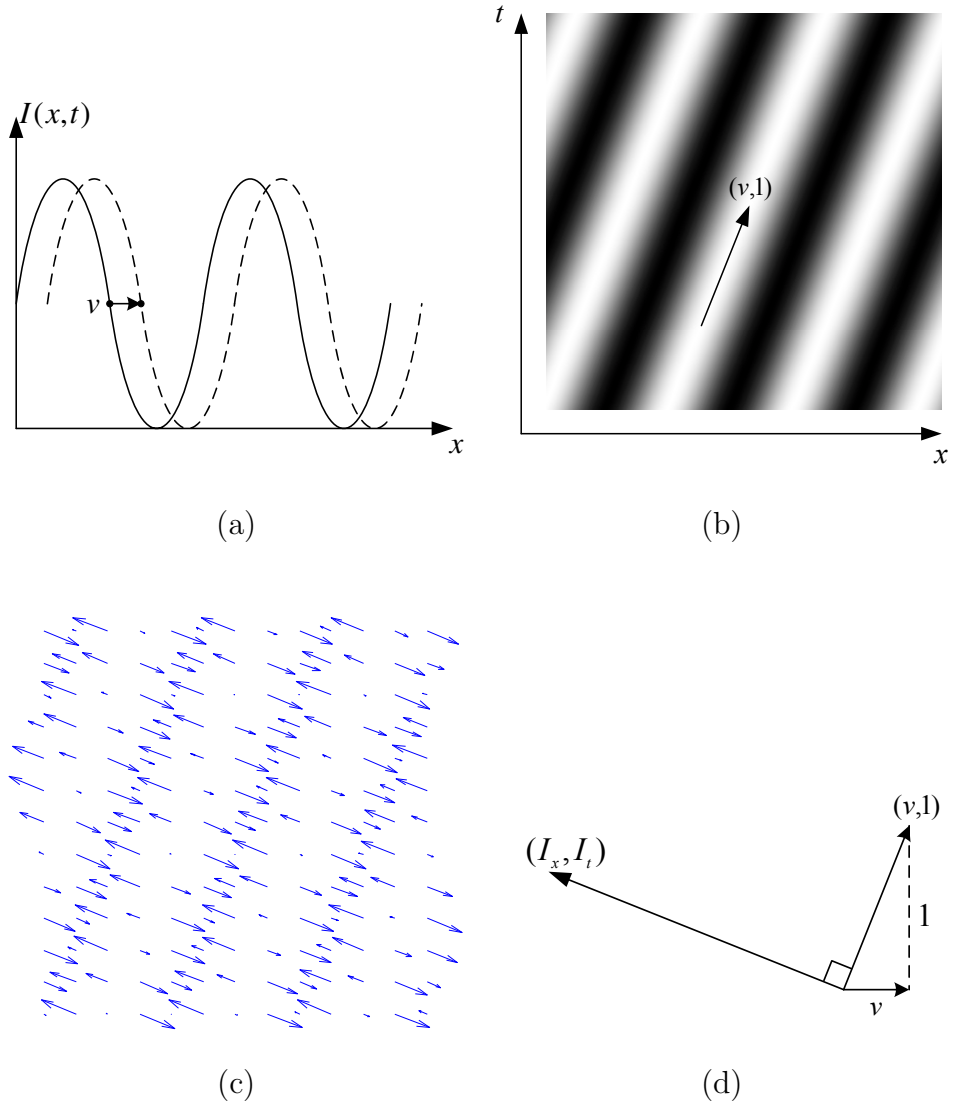


Figure 3.2. Estimate optical flow from spatio-temporal image of the 1-D shift sine pattern.

(a) A shifting 1-D sine-wave pattern. (b) The spatio-temporal image of (a). (c) The gradient of the spatio-temporal image. (d) The perpendicular relationship between the image gradient (I_x, I_t) and the direction of constant brightness $(v, 1)$.

3.2 The 3-D Gradient Constraint

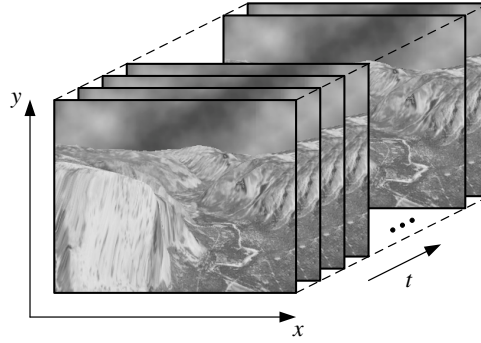


Figure 3.3. The 3-D spatio-temporal cube for 2-D video sequence.

interval of time. On this spatio-temporal image, it is clear that the image gradient (I_x, I_t) and the direction of constant brightness $(v, 1)$ are perpendicular to each other, which is shown in Figure 3.2(d).

From the above observation, it will be further demonstrated that such perpendicular relationship could further be taken advantage of to estimate the optical flow for 2-D video sequences. However, as shown in Figure 3.2(c), it should be noticed that not all pixels in the spatio-temporal image has a nonzero gradient, which means the pixels with zero or very small amplitude of image gradient is not reliable enough for calculating the optical flow for the 1-D time-variant signal.

3.2 The 3-D Gradient Constraint

In this section, the 3-D gradient constraint is derived from the perpendicular relationship between the direction of constant brightness and the 3-D gradient in the spatio-temporal cube. By locating the direction of constant brightness on the *constraint plane*, the tangent flow could be determined directly, which implies that the standard aperture problem could be solved based on the proposed constraint.

3.2 The 3-D Gradient Constraint

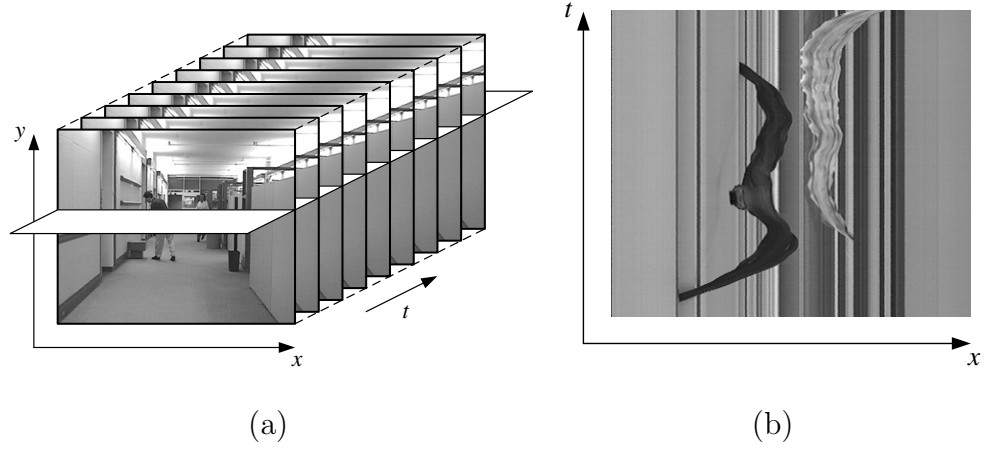


Figure 3.4. Sample the spatio-temporal cube. (a) Sample the cube with a plane parallel to the x - t plane. (b) The sample result of (a).

3.2.1 The Spatio-temporal Representation of Video

Now, consider a typical digital video containing a sequence of consecutive frames, which can be viewed as a 3-D volume data called the *spatio-temporal cube* with two spatial coordinates x and y , plus one temporal coordinate t , which is shown in Figure 3.3. Because the two point (x, y, t) and $(x + u, y + v, t + 1)$ share the same intensity under the constant intensity assumption (2.5), $(u, v, 1)$ points to the *direction of constant brightness* in the spatio-temporal cube. It's obvious that the optical flow could be estimated by finding the direction of constant brightness in the spatio-temporal cube, which is the fundamental idea of the proposed 3-D gradient constraint.

Figure 3.4 illustrates this observation by taking a image sample parallel to x - t plane in the spatio-temporal cube of the sequence *Hall and monitor*, which contains 300 frames. On the sample image, each pair of the pixels along the direction of constant brightness could be connected to construct a curve which has a constant intensity (Figure 3.4(b)). As discussed in 3.1.2, such texture could be utilized to calculate the 1-D velocity of the 1-D time-variant signal. However, only the motion parallel to the sample plane could be extracted from this sample of the spatio-

3.2 The 3-D Gradient Constraint

temporal cube, which implies that the plane on which the direction of constant brightness locates should be determined in the first place.

3.2.2 Definition of the Constraint Plane

As mentioned previously, the gradient constraint (2.8) only restricts the optical flow solution on a *constraint line* (Figure 2.4) which is a single linear equation in the two unknown velocity components u and v . Consequently, it is always the fundamental problem for gradient-based methods that the tangent flow can not be estimated from the gradient constraint [37]. In this part, we would introduce the *constraint plane*, on which both the normal flow and the tangent flow could uniquely be determined.

The constraint plane is derived by extending the gradient constraint from the 2-D velocity space to a 3-D spatio-temporal space, which enable us to locate the direction of constant brightness in the spatio-temporal cube. By rewriting the left side of (2.8) in the inner product form, the gradient constraint can be represented as

$$(I_x, I_y, I_t) \cdot (u, v, 1) = 0, \quad (3.11)$$

where (I_x, I_y, I_t) is the local gradient in the spatio-temporal cube and $(u, v, 1)$ is the vector pointing to the direction of constant brightness. Equation (3.11) reveals that the direction of constant brightness is perpendicular to the local gradient, which is the physical meaning underlying in the gradient constraint.

Because the 3-D vector $(u, v, 1)$ is perpendicular to the vector (I_x, I_y, I_t) which has already been determined, the solution to the vector $(u, v, 1)$ is restricted on a plane S which is perpendicular to the vector (I_x, I_y, I_t) (Figure 3.5). The plane S is coined as the *constraint plane* in this paper. As $(u, v, 1)$ represents the direction of constant brightness of the current point in the spatio-temporal cube, the constraint

3.2 The 3-D Gradient Constraint

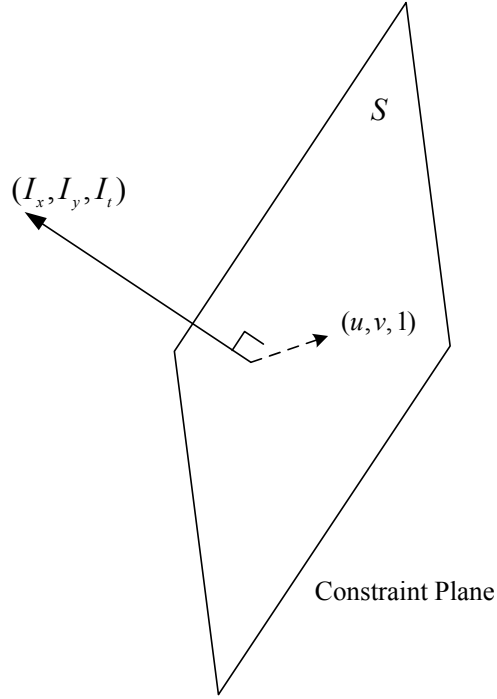


Figure 3.5. The constant plane is perpendicular to the local gradient (I_x, I_y, I_t) .

plane could be determined by the following equation,

$$I_x dx + I_y dy + I_t dt = 0, \quad (3.12)$$

where dx , dy and dt are the three axes of the coordinate, and the origin of the coordinate is the current point.

As the vector $(u, v, 1)$ is restricted on the constraint plane, the problem could be reduced to finding the 2-D direction of constant brightness on the image of the constraint plane. Once the direction of constant brightness is located on the constraint plane, the final solution to the vector $(u, v, 1)$ could be calculated, and the optical flow could therefore be obtained.

3.2.3 Construct the Constraint Plane

In order to further locate the direction of constant brightness, the image on the constraint plane has to be obtained by sampling the spatio-temporal cube. Since a

3.2 The 3-D Gradient Constraint

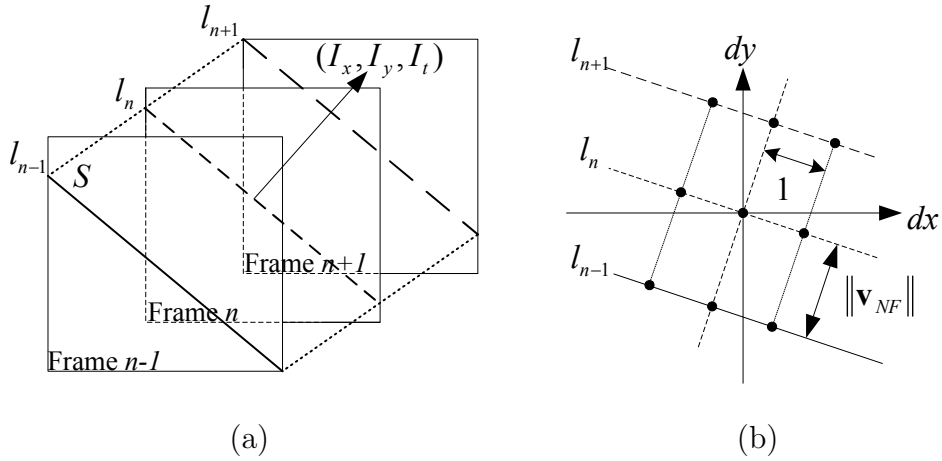


Figure 3.6. Sample the constraint plane. (a) The intersection lines on the constraint plane. (b) Sample the intersection lines.

video is a sequence of frames captured at fixed time intervals, the constraint plane is represented by the intersection lines of the constraint plane S with the consecutive video frames (Figure 3.6(a)). Here, three frames are utilized to get the direction of constant brightness, which implies that the motion is assumed to be the same across three frames.

By combining (3.12) with $dt = -1$, $dt = 0$ and $dt = 1$, respectively, the three intersection lines of the constraint plane at a point on the current frame n could be obtained, which are

$$\begin{aligned} l_{n-1} : I_x dx + I_y dy - I_t &= 0 & (\text{on frame } n-1) \\ l_n : I_x dx + I_y dy &= 0 & (\text{on frame } n) \\ l_{n+1} : I_x dx + I_y dy + I_t &= 0 & (\text{on frame } n+1). \end{aligned} \quad (3.13)$$

As shown in Figure 3.6(b), when the three intersection lines are projected onto the same frame plane, it can be calculated that the distance between these lines is $|I_t / \sqrt{I_x^2 + I_y^2}|$ which is equal to $\|\mathbf{v}_{NF}\|$. This implies that the motion of the normal flow has been compensated by positioning the intersection lines, and the tangent flow could actually be estimated from the displacement between these intersection lines, which would be further demonstrated.

3.2 The 3-D Gradient Constraint

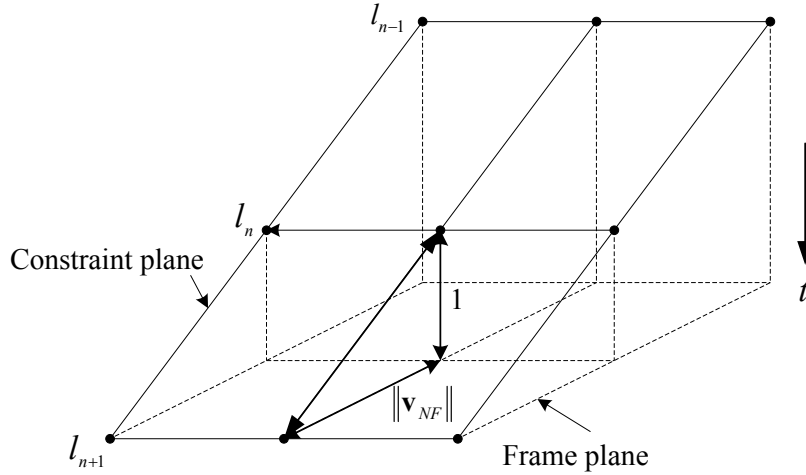


Figure 3.7. The distance between intersection lines in spatio-temporal space.

By laying a 3×3 grid on the constraint plane, the eight neighbors of the current point is obtained (Figure 3.6(b)), which are the samples of the constraint plane. The interval between the neighbor points on each intersection line is equal to 1, the same as the space between image pixels. Because these neighbors may not necessarily locate on integer points, the interpolation algorithm is needed to calculate the intensity of these points.

3.2.4 The Direction of Constant Brightness

As the image on the constraint plane is constructed by the 1-D signal of the intersection lines on consecutive frames, the direction of constant brightness can be found by utilizing the perpendicular relationship described in (3.10), in which the 1-D time-variant signal is represented as a spatio-temporal image.

However, the distance between the intersection lines in the spatio-temporal space is not equal to 1. As shown in Figure 3.7, because the time interval between the three intersection lines is 1 and the distance between the projection of the intersection lines on the frame plane is $\|\mathbf{v}_{NF}\|$, the distance between the intersection lines in 3-D space is $\sqrt{1 + \|\mathbf{v}_{NF}\|^2}$. In order to utilizing the perpendicular rela-

3.2 The 3-D Gradient Constraint

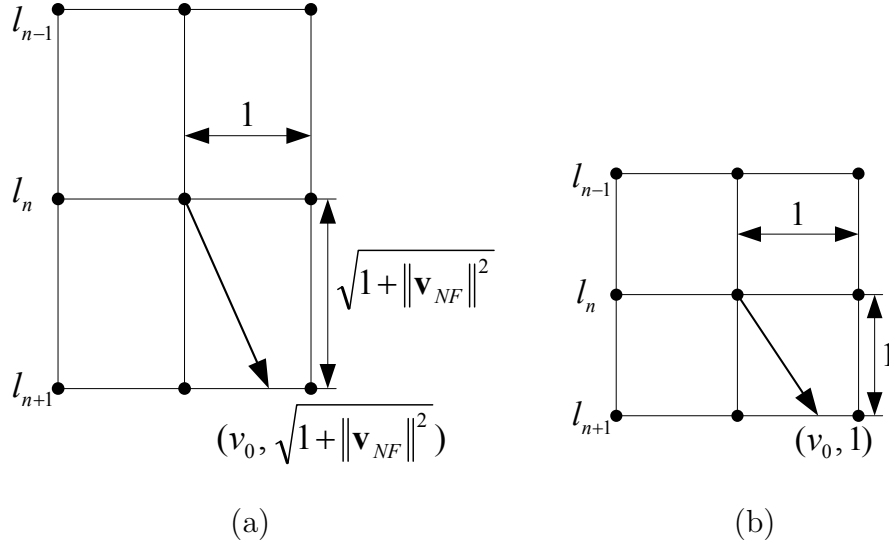


Figure 3.8. The direction of constant brightness on the constraint plane. (a) The nine points on the constraint plane. (b) Normalize the distance between the intersection lines.

tionship which has been discussed in 3.1.2, the distance between the intersection lines are normalized to 1 to construct the spatio-temporal image, which is shown in Figure 3.8(b). Based on this spatio-temporal image, the 2-D direction of constant brightness $(v_0, 1)$ can be obtained from (3.10), where v_0 is computed by (3.9) as the 1-D displacement between the intersection lines. When these three lines are stretched back to their original distance on the constraint plane (Figure 3.8(a)), the vector $(v_0, 1)$ is correspondingly stretched along the direction perpendicular to the intersection lines, and the vector is changed to $(v_0, \sqrt{1 + \|\mathbf{v}_{NF}\|^2})$, which is the 2-D direction of constant brightness on the constraint plane.

As the direction of constant brightness $(v_0, \sqrt{1 + \|\mathbf{v}_{NF}\|^2})$ is represented using the coordinate of the constraint plane, the corresponding 3-D vector in spatio-temporal space can be obtained by mapping each component of this 2-D vector to

3.2 The 3-D Gradient Constraint

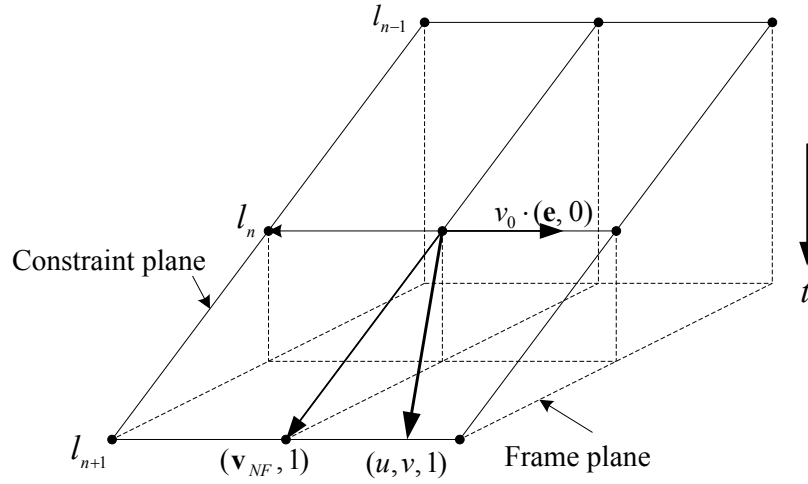


Figure 3.9. The decomposition of the direction of constant brightness.

3-D space (Figure 3.9), which is:

$$\begin{aligned} (u, v, 1) &= v_0 \cdot (\mathbf{e}, 0) + \sqrt{1 + \|\mathbf{v}_{NF}\|^2} \cdot \frac{(\mathbf{v}_{NF}, 1)}{\|(\mathbf{v}_{NF}, 1)\|} \\ &= (v_0 \cdot \mathbf{e} + \mathbf{v}_{NF}, 1), \end{aligned} \quad (3.14)$$

where \mathbf{v}_{NF} is the normal flow, \mathbf{e} is the 2-D unit vector parallel to the intersection lines on the frame plane, and v_0 is the 1-D displacement between the intersection lines (Figure 3.8(b)).

Based on (3.14), the optical flow can be represented as

$$\mathbf{v} = v_0 \cdot \mathbf{e} + \mathbf{v}_{NF}. \quad (3.15)$$

From $\mathbf{v} = \mathbf{v}_{NF} + \mathbf{v}_{TF}$, the tangent flow is

$$\mathbf{v}_{TF} = v_0 \cdot \mathbf{e}. \quad (3.16)$$

Therefore, the tangent flow \mathbf{v}_{TF} is equal to the 1-D displacement between the intersection lines which construct the constraint plane. In other words, \mathbf{v}_{TF} can be calculated by utilizing the variational information on the constraint plane.

3.3 The Standard Aperture Problem

3.3 The Standard Aperture Problem

Based on the traditional gradient constraint (2.8), the tangent flow could not be estimated directly, which is known as the *standard aperture problem* [3, 16]. As stated in 3.2, the tangent flow could be calculated based on the proposed 3-D gradient constraint, which implies that the standard aperture problem could therefore be solved.

In this section, the spatio-temporal image on the constraint plane is analyzed. Based on this analysis, an image gradient estimator is proposed to replace the regular image gradient estimator, such as the Sobel operator. The condition for solving the standard aperture problem is given in the end.

3.3.1 The Spatio-temporal Image on the Constraint Plane

As the directional derivative of a 2-D function $I(\mathbf{x})$ along the 2-D vector \mathbf{h} is given by

$$\nabla_{\mathbf{h}} I(\mathbf{x}) = \nabla I(\mathbf{x}) \cdot \mathbf{h}, \quad (3.17)$$

if the vector \mathbf{h} is perpendicular to the image gradient $\nabla I(\mathbf{x})$, the corresponding directional derivative is equal to *zero*. According to the definition in (3.13), the intersection line l_n on the constraint plane (Figure 3.6(b)) is perpendicular to the local image gradient of the current frame. Therefore, the directional derivative along the intersection line l_n on the spatio-temporal image is expected to be zero.

However, such *zero* directional derivative does not mean there is no intensity variation along this direction. Normally, as the image intensity would not only change in the direction of image gradient, except the case of barber's pole illusion which has no variation in the direction perpendicular to the image gradient, the zero directional derivative represents the local extremum of the intensity (Figure 3.10). It will be shown that the variational information at the local extremum could

3.3 The Standard Aperture Problem

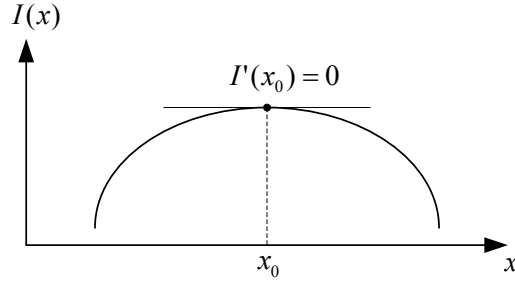


Figure 3.10. The derivative would be zero at local extremum.

still be utilized in the solving of standard aperture problem.

3.3.2 Estimate the Direction of Image Gradient for Spatio-temporal Image

Rather than a complicated computational strategy, a simple algorithm would be introduced in this part to replace the regular image gradient estimator which may fail to fully extract the variation information for spatio-temporal images.

Commonly, the 2-D image gradient (I_x, I_t) of a pixel (x, t) is calculated by the difference operators, such as the Sobel operator, which is

$$I_x(x, t) = I(x, t) * \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \quad \text{and} \quad I_t(x, t) = I(x, t) * \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (3.18)$$

Although such estimation may perform well in many other tasks, this method may result in very poor solution when the computed image gradient was used to estimate the tangent flow in the situation described in 3.3.1.

Take a 3×3 sample (Figure 3.11(a)) from the spatio-temporal image as an example, in which the signal is shifting 1 pixel to the left by each frame and the pixel in the center is a local minimum along the intersection line l_n which is the three pixels in the middle row. If we use (3.18) to calculate the image gradient of

3.3 The Standard Aperture Problem

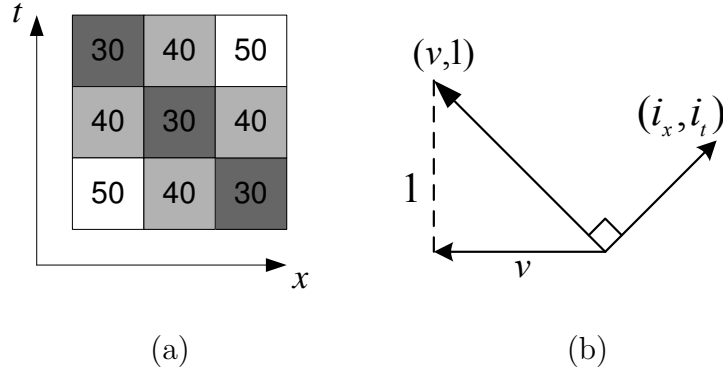


Figure 3.11. Estimate local gradient of a pixel by its 8-adjacent. (a) The 3×3 sample from the spatio-temporal image. (b) The direction of maximum intensity change and the corresponding velocity.

the pixel in the center, the gradient will be $(I_x, I_t) = (0, 0)$, which implies that there is no variational information to estimate the 1-D motion. However, if we utilized the absolute difference between the current pixel and its neighbors, the 1-D motion could still be extracted.

Because the gradient of a 2-D function is defined as a vector pointing to the direction of the steepest slope, which means the image gradient could be determined by finding the direction with the maximum intensity change. In addition, according to (3.9), the 1-D velocity is computed from the quotient of the two image gradient components

$$v = -\frac{I_t}{I_x} = -\frac{i_t}{i_x}, \quad (3.19)$$

where (i_x, i_t) represents the direction of maximum intensity change (Figure 3.11(b)). Therefore, what we need is only the direction of image gradient rather than the exact image gradient, and the direction of image gradient at a pixel in the spatio-temporal image of the constraint plane could be computed by detecting the *direction of maximum intensity change*:

Step 1 Take the absolute difference between the outlying pixels and the pixel in the center along the four directions g_i ($i = 1, 2, 3, 4$) shown in Figure 3.12 as

3.3 The Standard Aperture Problem

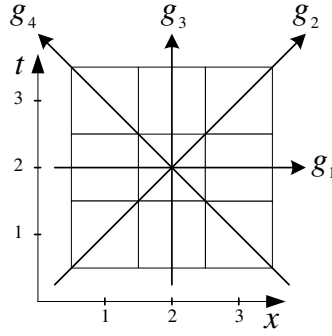


Figure 3.12. Estimate the direction of image gradient based on the absolute intensity difference along the 4 directions.

the weights of the corresponding directions, which are

$$\begin{aligned}
 d_1 &= |I(2, 3) - I(2, 2)| + |I(2, 1) - I(2, 2)| \\
 d_2 &= [|I(1, 1) - I(2, 2)| + |I(3, 3) - I(2, 2)|] / \sqrt{2} \\
 d_3 &= |I(1, 2) - I(2, 2)| + |I(3, 2) - I(2, 2)| \\
 d_4 &= [|I(1, 3) - I(2, 2)| + |I(3, 1) - I(2, 2)|] / \sqrt{2}.
 \end{aligned} \tag{3.20}$$

Step 2 Compare the four weights and find the direction with the maximum absolute difference, which is labeled as g_{max} .

Step 3 Estimate the direction of maximum intensity change by taking the weighted vector sum of g_{max} and its two neighbors, which is

$$(i_x, i_t) = \begin{cases} g_1 \cdot d_1 + g_2 \cdot d_2 - g_4 \cdot d_4 & \text{if } g_{max} = g_1 \\ g_1 \cdot d_1 + g_2 \cdot d_2 + g_3 \cdot d_3 & \text{if } g_{max} = g_2 \\ g_2 \cdot d_2 + g_3 \cdot d_3 + g_4 \cdot d_4 & \text{if } g_{max} = g_3 \\ -g_1 \cdot d_1 + g_3 \cdot d_3 + g_4 \cdot d_4 & \text{if } g_{max} = g_4 \end{cases}$$

where $g_1 = (1, 0)$, $g_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $g_3 = (0, 1)$ and $g_4 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.

To illustrate the way that this scheme works, it is applied to the example shown in Figure 3.11(a). First, the four weights are calculated as $d_1 = 20$, $d_2 = 40$, $d_3 = 20$

3.3 The Standard Aperture Problem

and $d_4 = 0$. Then, the maximum of the four weights is d_2 . Finally, the direction of maximum intensity change could be determined as $(i_x, i_t) = (20 + 20\sqrt{2})(1, 1)$ which is perpendicular to the direction of constant brightness $(-1, 1)$ (Figure 3.11(b)).

As the image gradient is calculated by examining the absolute intensity difference along all four possible directions at a pixel position, the variational information is fully extracted in the local area. Besides, this approach also accomplished to extract the variational information for the spatio-temporal image where local extremum would appear as illustrated in the Figure 3.11(a).

3.3.3 Condition for Solving The Standard Aperture Problem

In this paper, as the *direction of maximum intensity change* (i_x, i_t) are utilized to compute the tangent flow using (3.19), the *solvability condition* of the standard aperture problem is

$$i_x \neq 0 \tag{3.21}$$

on the *constraint plane*. In addition, $i_x = 0$ on the constraint plane implies that there is no brightness variation along the direction perpendicular to the image gradient at the current point, which appears like the barber's pole illusion in the local area. In that case, the tangent flow at this point cannot be obtained, thus the standard aperture problem could not be solved, and only normal flow could be obtained in the local area. Besides, because the 3-D gradient constraint is extended from the gradient constraint, which is valid in the condition that the motion between consecutive frames is very small, the case of large displacement could not be handled by the proposed constraint. As stated in Chapter 2, the multi-resolution technique and the coarse-to-fine strategy [25, 26] should be incorporated in such cases.

As the standard aperture problem could be solved, the estimation of tangent

3.4 The Optical Flow Estimation Scheme

flow could stop relying on those smoothness assumptions, which provides a potential to get a more precise estimation of optical flow at the edge of motion objects as well as the objects with non-rigid deformation.

3.4 The Optical Flow Estimation Scheme

From (3.16) the tangent flow can be directly estimated from the spatio-temporal image of the constraint plane. Therefore, the full flow could be obtained by respectively calculating the normal flow and the tangent flow, and taking the vector sum of the two optical flow components. The optical flow estimation procedure for a pixel on frame n of a pre-smoothed sequence is:

Step 1 Estimate the local gradient (I_x, I_y, I_t) and compute the normal flow \mathbf{v}_{NF} by (2.11).

Step 2 Locate the three intersection lines of constraint plane by (3.13) and get the intensity of the nine points on the intersection lines using image interpolation.

Step 3 Estimate the direction of image gradient on the constraint plane using the approach in 3.3.2 and compute the tangent flow \mathbf{v}_{TF} by (3.19).

Step 4 Get the optical flow \mathbf{v} by taking the vector sum of \mathbf{v}_{NF} and \mathbf{v}_{TF} , which is

$$\mathbf{v} = \mathbf{v}_{NF} + \mathbf{v}_{TF}.$$

From the steps listed above, both the two components of the optical flow, the normal flow and the tangent flow, are calculated by closed-form equations. And it's not so time-consuming to estimate the image gradient and implement the interpolation. Therefore, a real-time performance could be expected in estimating the optical flow of high definition video sequences.

Chapter 4

Experimental Results

4.1 Implementation

As described in 3.4, because the tangent flow could be estimated directly based on the 3-D gradient constraint, the proposed optical flow estimation algorithm is implemented by respectively calculating the normal flow and tangent flow and taking their vector sum. First, the image of the current frame I_n and its two neighbor frames I_{n-1} , I_{n+1} are pre-smoothed by a Gaussian filter with a standard deviation of 1.5 pixel in space and 1.5 frame in time. Then, the local gradient (I_x, I_y) in space and I_t in time is calculated by utilizing the two point central-difference kernel $(-0.5 \ 0 \ 0.5)$ along the three corresponding directions in the spatio-temporal space. The normal flow \mathbf{v}_{NF} is computed by (2.11), and the 3×3 image of the constraint plane is obtained by using the bi-cubic interpolation on the corresponding frames because the these points may not locate on integer indexed positions. Next, the direction of image gradient in the spatio-temporal image of the constraint plane is calculated using the approach introduced in 3.3.2. The tangent flow \mathbf{v}_{TF} is estimated by using (3.19). Finally, by taking the vector sum of the normal flow and the tangent flow, the estimation of the full flow \mathbf{v} is obtained.

4.2 Results

In this section, the performance of the proposed optical flow estimation method will be explored, using the synthetic sequences *Sine*, *Translating Tree*, *Diverging Tree* and *Yosemite*, as well as the real-world sequence *Taxi*. All these sequences are available in `ftp://ftp.csd.uwo.ca`. To enhance the precision of the proposed approach, the median filter is applied to the obtained optical flow field to get rid of the vectors with large estimation error. Here, the u component and the v component of the optical flow field are filtered by the median filter separately. In order to demonstrate the full potential of the proposed method, we do not utilize any further refinement or robust scheme.

Based on the 3-D gradient constraint, the result utilizing Sobel operator and the *direction of maximum intensity change* are shown respectively to demonstrate the novel effect of using the proposed gradient estimator on the constraint plane. In addition, the precision of the two methods refined by the median filter are compared in terms of AAE (average angular error) and STD (standard deviation of angular error) (shown in Table 4.1). The estimation error of the flow field using the Sobel operator and the proposed gradient estimator is also compared by demonstrating the proportion of pixels with the angular error less than 1° , 2° , 3° , 5° and 10° (Table 4.2), respectively. In the end, the execution time of the proposed method is shown for video sequences with different resolutions.

Figure 4.1 shows the optical flow results of sequence *Sine*. The 2-D sine pattern (Figure 4.1(a)) is moving (1, 1) by each frame to top right direction. The optical flow result of the algorithm utilizing the Sobel operator (Figure 4.1(c)) and the algorithm using the proposed gradient estimator (Figure 4.1(e)) get similar precision. By further filtering the result in Figure 4.1(c) and Figure 4.1(e) using the median filter, respectively, both the flow fields are fairly precise as the ground truth.

4.2 Results

Table 4.1

PRECISION COMPARISON IN AAE AND STD.

Sequence	Sobel operator		Proposed	
	AAE($^{\circ}$)	STD($^{\circ}$)	AAE($^{\circ}$)	STD($^{\circ}$)
Sine	0.69	0.21	1.02	0.09
Translating Tree	26.44	13.97	3.87	2.72
Diverging Tree	12.42	7.84	3.19	1.77
Yosemite	20.48	15.61	10.69	10.74

Table 4.2

PRECISION COMPARISON IN ERROR DISTRIBUTION (%).

Sequence	Sobel operator					Proposed				
	$< 1^{\circ}$	$< 2^{\circ}$	$< 3^{\circ}$	$< 5^{\circ}$	$< 10^{\circ}$	$< 1^{\circ}$	$< 2^{\circ}$	$< 3^{\circ}$	$< 5^{\circ}$	$< 10^{\circ}$
Sine	93.3	99.8	99.9	100	100	51.0	100	100	100	100
Translating Tree	0.00	0.00	0.00	0.028	5.82	8.32	20.5	39.1	76.4	97.1
Diverging Tree	0.52	1.91	4.02	11.35	48.65	5.21	26.6	50.0	89.4	99.2
Yosemite	1.05	3.26	6.78	16.3	34.3	4.67	12.0	20.2	39.6	69.8

As shown in Figure 4.2(a) and Figure 4.3(a), the texture of the natural image in the sequences *Translating tree* and *Diverging tree* is much more complicated than the simple sine pattern in the sequence *Sine*. In such circumstances, a large number of motion vectors are wrongly estimated in the result using Sobel operator (Figure 4.2(c) and Figure 4.3(c)), while the result using the direction of maximum intensity change estimator (Figure 4.2(e) and Figure 4.3(e)) is much more reliable. Besides, the proposed gradient estimator also yields better objective precision in the comparison of AAE and STD (Table 4.1), as well as higher proportion of pixels with smaller angular error (Table 4.2).

As a more complicated scene, the *Yosemite* is involved with multiple motion of

4.2 Results

Table 4.3

PRECISION COMPARISON OF THE SEQUENCE *Yosemite*.

Technique	AAE(°)	STD(°)
Negel [19]	10.22	16.51
Horn and Schunck [19]	9.78	16.19
Uras <i>et al.</i> [19]	8.94	15.61
Our method	10.69	10.74
Alvarez <i>et al.</i> [26]	5.53	7.40
Brox <i>et al.</i> [10]	2.46	7.31
Xiao <i>et al.</i> [17]	2.57	6.07

different objects located in different depths, which is shown in Figure 4.4(a). Due to a large degree of illumination change in the area of clouds, accurate optical flow estimation is difficult to achieve. However, after the refinement of the median filter (Figure 4.4(f)), the motion boundaries are well preserved and the flow field in the area of mountain is estimated accurately. As this sequence is widely utilized in the evaluation of optical flow techniques, a comparison of the estimation precision between the proposed method and some existing methods are shown in Table 4.3, in which AAE and STD are used in the comparison.

For a real-world video sequence *Taxi* which contains multiple objects under different motion and object occlusion, the result is shown in Figure 4.5. The proposed method provides not only accurate moving direction but also sharp motion boundaries. Further note that the speed of the three moving cars is also reflected correctly: the one in the middle is turning slowly, and the other two are moving horizontally with higher speed.

To demonstrate the potential of utilizing the proposed algorithm in time-critical applications, the execution time is shown in Table 4.4. Such short execution time reveals the potential of achieving real-time performance of optical flow estimation.

4.2 Results

Table 4.4

EXECUTION TIME.

Sequence	Image size	Time (sec)
<i>Sine</i>	100×100	0.38
<i>Taxi</i>	256×190	0.84
<i>Yosemite</i>	316×252	1.50

Our proposed algorithm is programmed in MATLAB language and run on 2.13 GHz Intel Core 2 Duo CPU processor. For an image sequence with frame size 316×252 , the resulted computation time is 1.5 sec/frame, which is much faster than most algorithms even programmed in C/C++ language, for example, 4 sec/frame in [17], 23 sec/frame in [10]. Furthermore, our algorithm yields a close performance, compared to the method utilizing multigrid algorithms [38] implemented in C/C++, which is 0.15 sec/frame for a 160×120 sequence.

Based on these results, it is clear demonstrated that the intensity variation on the image of constraint plane is reliable to estimate the *tangent flow* with the proposed direction of maximum intensity change estimator. Although our approach is not the best in terms of estimation *accuracy*, however, it provides the fastest optical flow estimation, since only the closed-form equations are exploited without incorporating any smoothness constraint or additional processing techniques, while yielding a fairly precise optical flow field.

4.2 Results

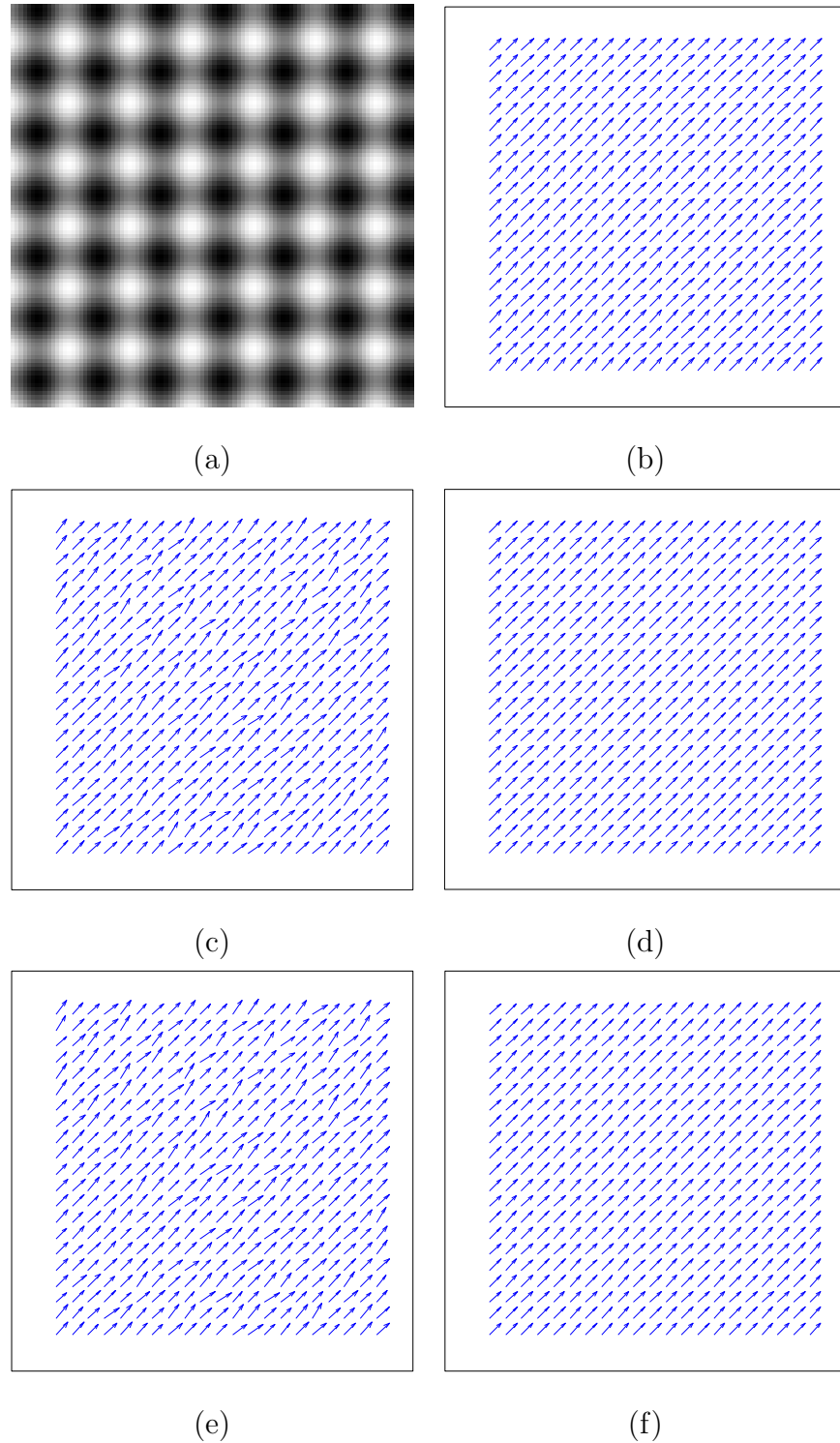


Figure 4.1. Estimated optical flow field for sequence *Sine*. (a) The 9th frame of *Sine*. (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.

4.2 Results

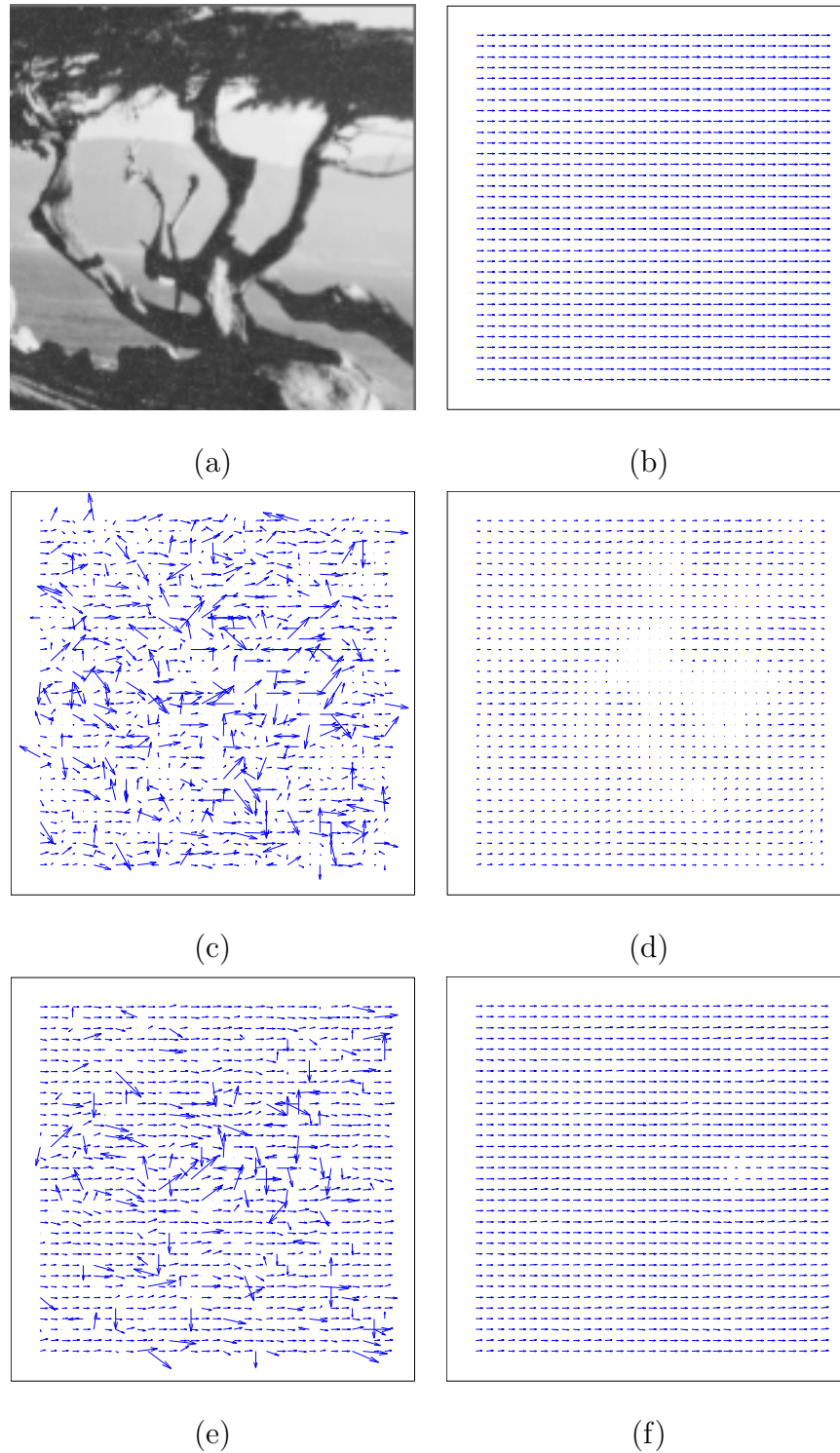


Figure 4.2. Estimated optical flow field for sequence *Translating Tree*. (a) The 9th frame of *Translating Tree*. (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.

4.2 Results

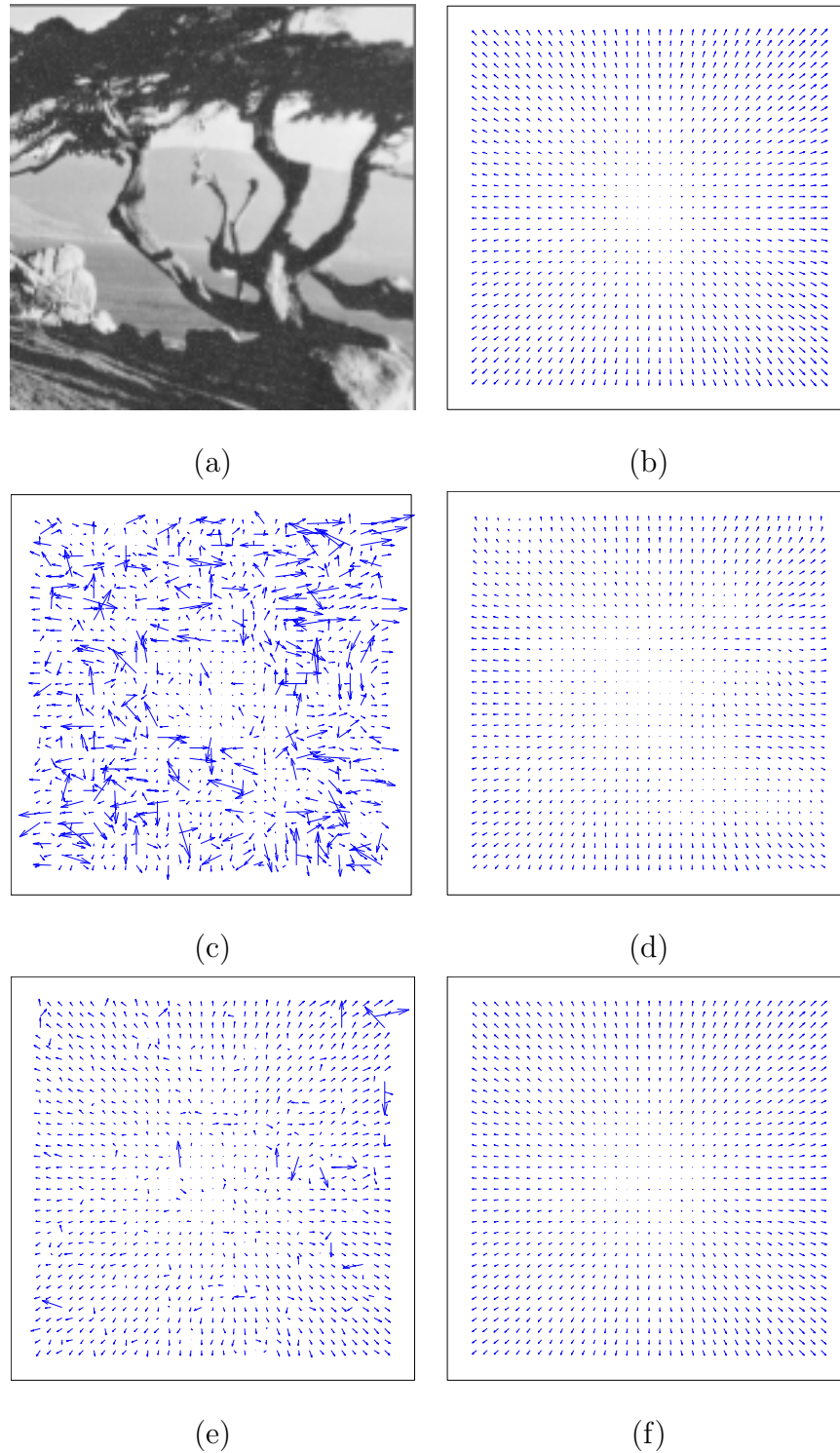


Figure 4.3. Estimated optical flow field for sequence *Diverging Tree*. (a) The 9th frame of *Diverging Tree*. (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.

4.2 Results

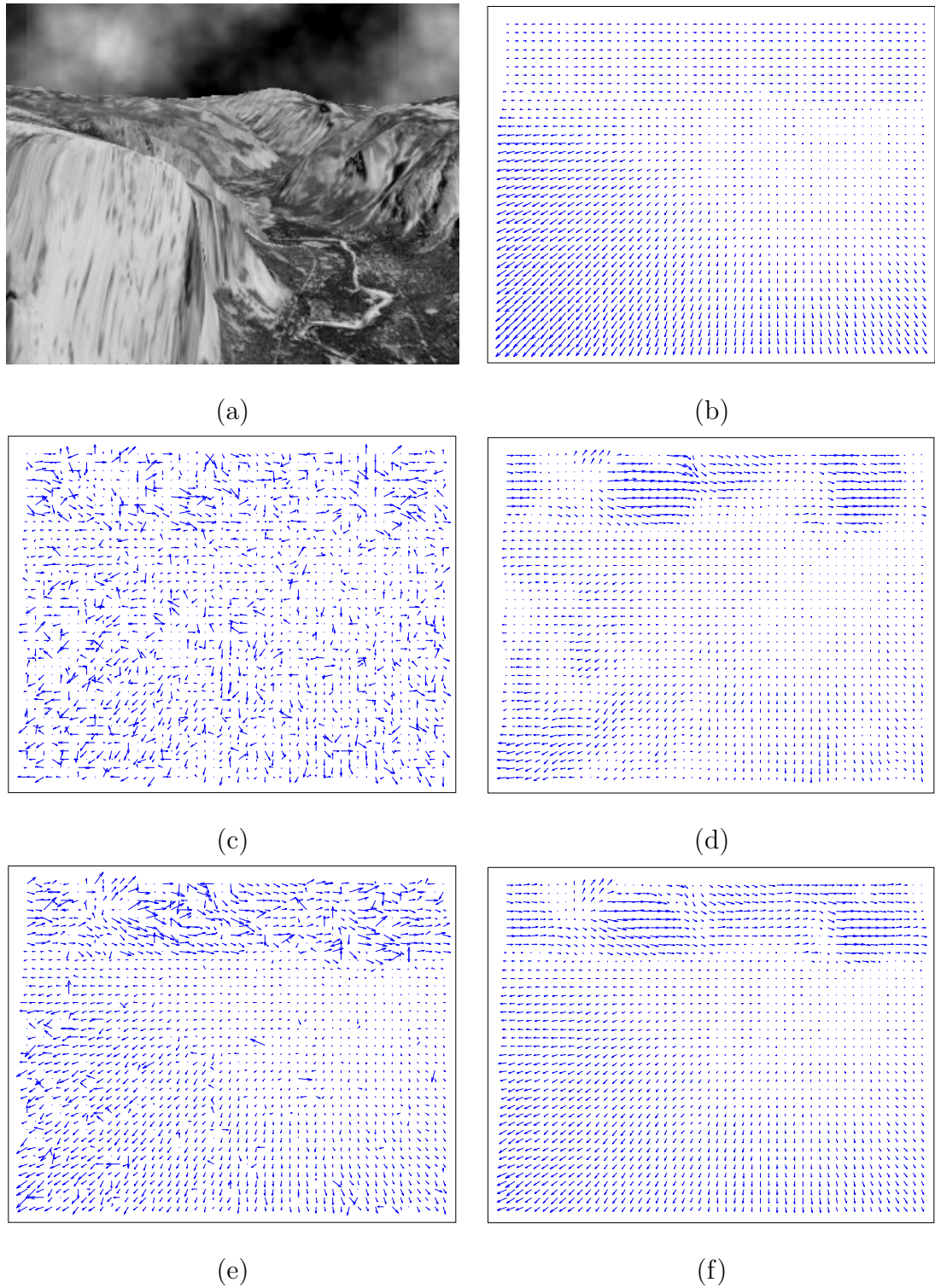


Figure 4.4. Estimated optical flow field for sequence *Yosemite*. (a) The 9th frame of *Yosemite*. (b) Ground truth. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter.. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.

4.2 Results

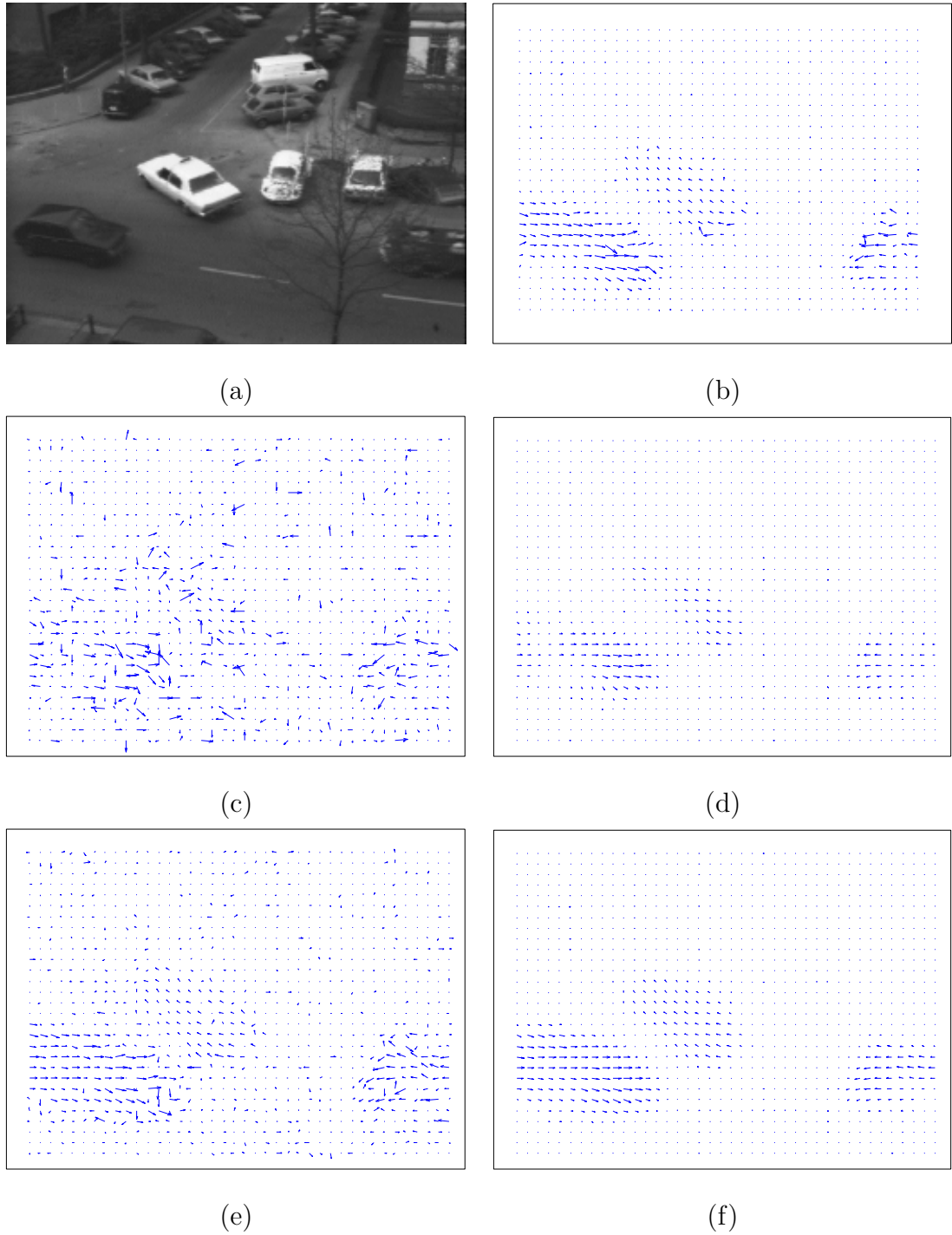


Figure 4.5. Estimated optical flow field for sequence *Taxi*. (a) The 9th frame of *Taxi*. (b) Horn and Schunck 100 iterations. (c) Proposed method using Sobel operator. (d) Refined result of (c) using median filter. (e) Proposed method using direction of maximum intensity change. (f) Refined result of (e) using median filter.

Chapter 5

Conclusions

In this thesis, the 3-D gradient constraint is proposed to estimate the optical flow without imposing any smoothness assumption of the motion field, which is the fundamental contribution to optical flow estimation. By extending the constraint *line* in the velocity space to the constraint *plane* in the spatio-temporal space, this new constraint enable us to estimate the tangent flow directly. This implies that the *standard aperture problem* could be solved, and the solvability condition for this problem is derived. Besides, as the local extremum would appear in the spatio-temporal image of the constraint plane, a gradient estimator which is named as the *direction of maximum intensity change* is introduced to fully extract the variational information on the constraint plane for tangent flow calculation.

Simulation results obtained from the synthetic and real-world sequences have shown that the intensity variation on the direction perpendicular to the image gradient indeed provides a reliable information for conducting tangent flow estimation. In addition, because the optical flow is calculated by utilizing a set of closed-form equations without incorporating any smoothness assumptions or calling for iterative numerical computations, the fairly low computational complexity of the proposed approach yields a very high potential for time-critical applications.

Bibliography

- [1] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.
- [2] S. Zhu and K. Ma, “A new diamond search algorithm for fast block-matching motion estimation,” *IEEE Transactions on Image Processing*, vol. 9, no. 2, p. 287, 2000.
- [3] B. K. P. Horn, *Robot Vision*. MIT Press, 1986, ch. 12 Motion Field & Optical Flow, pp. 278–298.
- [4] B. D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, April 1981, pp. 674–679.
- [5] P. Anandan, “A computational framework and an algorithm for the measurement of visual motion,” *International Journal of Computer Vision*, vol. 2, no. 3, pp. 283–310, 1989.
- [6] D. J. Fleet and A. D. Jepson, “Computation of component image velocity from local phase information,” *International Journal of Computer Vision*, vol. 5, no. 1, pp. 77–104, 1990.

-
- [7] A. Jepson and M. J. Black, “Mixture models for optical flow computation,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 1993, pp. 760–761.
 - [8] H. H. Nagel and W. Enkelmann, “An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 5, pp. 565–593, 1986.
 - [9] J. Bigün, G. H. Granlund, and J. Wiklund, “Multidimensional orientation estimation with applications to texture analysis and optical flow,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 775–790, 1991.
 - [10] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping,” in *Computer Vision - ECCV 2004*. Springer, 2004, pp. 25–36.
 - [11] A. Bruhn, J. Weickert, and C. Schnörr, “Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods,” *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.
 - [12] E. P. Simoncelli, E. H. Adelson, and D. J. Heeger, “Probability distributions of optical flow,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 1991, pp. 310–315.
 - [13] S. Roth and M. J. Black, “On the spatial statistics of optical flow,” *International Journal of Computer Vision*, vol. 74, no. 1, pp. 33–50, 2007.
 - [14] M. J. Black and A. D. Jepson, “Estimating optical flow in segmented images using variable-order parametric models with local deformations,” *IEEE Trans-*

- actions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 972–986, 1996.
- [15] H. Liu, R. Chellappa, and A. Rosenfeld, “Accurate dense optical flow estimation using adaptive structure tensors and a parametric model,” *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1170–1180, 2003.
- [16] M. J. Black and P. Anandan, “The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields,” *Computer Vision and Image Understanding*, vol. 63, pp. 75–104, 1996.
- [17] J. Xiao, H. Cheng, H. Sawhney, C. Rao, and M. Isnardi, “Bilateral filtering-based optical flow estimation with occlusion detection,” in *Computer Vision - ECCV 2006*, vol. LNCS 3951. Springer, 2006, pp. 211–224.
- [18] E. C. Hildreth, *Measurement of visual motion*. MIT Press, Cambridge, MA, 1984.
- [19] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, “Performance of optical flow techniques,” *International Journal of Computer Vision*, vol. 12, pp. 43–77, 1994.
- [20] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills, “Recovering motion fields: An evaluation of eight optical flow algorithms,” in *British Machine Vision Conference*, 1998, pp. 195–204.
- [21] B. McCane, K. Novins, D. Crannitch, and B. Galvin, “On benchmarking optical flow,” *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 126–143, 2001.

-
- [22] E. Memin and P. Perez, “A multigrid approach for hierarchical motion estimation,” in *Sixth International Conference on Computer Vision*, 1998, pp. 933–938.
 - [23] A. Bruhn, J. Weickert, C. Feddern, T. Kohlberger, and C. Schnorr, “Variational optical flow computation in real time,” *IEEE Transactions on Image Processing*, vol. 14, no. 5, pp. 608–615, 2005.
 - [24] J. Weickert and C. Schnörr, “Variational optic flow computation with a spatio-temporal smoothness constraint,” *Journal of Mathematical Imaging and Vision*, vol. 14, no. 3, pp. 245–255, 2001.
 - [25] E. Memin and P. Perez, “Dense estimation and object-based segmentation of the optical flow with robust techniques,” *IEEE Transactions on Image Processing*, vol. 7, no. 5, pp. 703–719, 1998.
 - [26] L. Alvarez, J. Weickert, and J. Sánchez, “Reliable estimation of dense optical flow fields with large displacements,” *International Journal of Computer Vision*, vol. 39, no. 1, pp. 41–56, 2000.
 - [27] L. Alvarez, R. Deriche, T. Papadopoulos, and J. Sánchez, “Symmetrical dense optical flow estimation with occlusions detection,” in *Proceedings of the 7th European Conference on Computer Vision-Part I*. Springer-Verlag London, UK, 2002, pp. 721–735.
 - [28] H. Haußecker and H. Spies, *Handbook of Computer Vision and Applications*. Academic Press, 1999, vol. 2, ch. 13 Motion, pp. 310–396.
 - [29] C. Fermüller, D. Shulman, and Y. Aloimonos, “The statistics of optical flow,” *Computer Vision and Image Understanding*, vol. 82, no. 1, pp. 1–32, 2001.

-
- [30] H. H. Nagel, "Displacement vectors derived from second-order intensity variations in image sequences," *Computer Vision, Graphics and Image Processing*, vol. 21, no. 1, pp. 85–117, 1983.
 - [31] S. Uras, F. Girosi, A. Verri, and V. Torre, "A computational approach to motion perception," *Biological Cybernetics*, vol. 60, no. 2, pp. 79–87, Dec 1988.
 - [32] A. Singh, "An estimation-theoretic framework for image-flow computation," in *Third International Conference on Computer Vision*, Dec 1990, pp. 168–177.
 - [33] D. J. Fleet and Y. Weiss, *Handbook of Mathematical Models in Computer Vision*. Springer, 2005, ch. 15 Optical flow estimation, pp. 239–258.
 - [34] D. J. Fleet and A. D. Jepson, "Stability of phase information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 12, pp. 1253–1268, 1993.
 - [35] J. Magarey and N. Kingsbury, "Motion estimation using a complex-valued wavelet transform," *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 1069–1084, 1998.
 - [36] E. Bayro-Corrochano, "The theory and use of the quaternion wavelet transform," *Journal of Mathematical Imaging and Vision*, vol. 24, no. 1, pp. 19–35, 2006.
 - [37] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade, "Three-dimensional scene flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 475–480, 2005.
 - [38] A. Bruhn and J. Weickert, "Towards ultimate motion estimation: Combining highest accuracy with real-time performance," in *Tenth IEEE International Conference on Computer Vision*, vol. 1, 2005, pp. 749–755.