

Research Article

Scene Flow Estimation Based on Adaptive Anisotropic Total Variation Flow-Driven Method

Xuezhi Xiang , Rongfang Zhang, Mingliang Zhai, Deguang Xiao, and Erwei Bai

School of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Xuezhi Xiang; xiangxuezhi@hrbeu.edu.cn

Received 5 December 2017; Accepted 27 February 2018; Published 20 May 2018

Academic Editor: Francesco Marotti de Sciarra

Copyright © 2018 Xuezhi Xiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Scene flow estimation based on disparity and optical flow is a challenging task. We present a novel method based on adaptive anisotropic total variation flow-driven method for scene flow estimation from a calibrated stereo image sequence. The basic idea is that diffusion of flow field in different directions has different rates, which can be used to calculate total variation and anisotropic diffusion automatically. Brightness consistency and gradient consistency constraint are employed to establish the data term, and adaptive anisotropic flow-driven penalty constraint is employed to establish the smoothness term. Similar to the optical flow estimation, there are also large displacement problems in the estimation of the scene flow, which is solved by introducing a hierarchical computing optimization. The proposed method is verified by using the synthetic dataset and the real scene image sequences. The experimental results show the effectiveness of the proposed algorithm.

1. Introduction

Scene flow was first introduced by Vedula et al. [1] in 1999. It is defined as a dense 3D motion field that describes the change of structured light in the image surfaces, which can be estimated from a calibrated stereo image sequence. We can obtain stereo image sequence by using binocular stereo vision sensor, like BumblebeeX3. As we all know, optical flow can be seen as the projection of the moving object on the retina or the camera's sensor. So scene flow can be regarded as a 3D extension of optical flow. Reliable 3D motion flow field can be used in many applications. For example, Menze and Geiger [2] focus on the autonomous driving with 3D scene flow estimation. Lenz et al. [3] used scene flow for moving object detection in urban environments. Waizenegger et al. [4] present a real-time approach for 3D structure estimation of human bodies. In general, scene flow estimation techniques can be divided into two categories. One is variational minimization method of scene flow based on the optical flow and the disparity from stereo image sequence [5]. The other is 3D point cloud parametrization method based on the 3D structure to estimate the scene flow, which allows us to directly estimate the desired unknowns [6]. Under these two frameworks, a lot of researches on optical flow estimation can be used for scene flow estimation.

In the estimation of optical flow, many regularization approaches have been derived to solve the ill-posed problem. These methods can be done by minimizing an energy function

$$\min \int_{\Omega} \Psi_{\text{data}} (I_x \cdot u + I_y \cdot v + I_t) + \Psi_{\text{smooth}} (u, v) dx dy, \quad (1)$$

where $\Psi_{\text{smooth}}(u, v)$ is a regularization term which penalizes variations u and v . Ω refers to the domain of the image plane. Ψ_{data} and Ψ_{smooth} are convex regularization functions and used to guarantee convexity and differentiability of the energy function.

Many smoothing regularizations are characterized as below. Homogenous regularization is named by Bruhn et al. [7], because the approach is a linear diffusion equation which is defined by Horn and Schunck [8]. Inspired by image filtering, Alvarez and Esclarin [9] use the squared magnitude of the image gradient to control the smoothness. Perona and Malik [10] propose an anisotropic smoothness method, which can preserve flow edges for a large image gradient and reduce the smooth strength in the consistency area. Nagel and Enkelmann [11] only smooth flow along the image edge and reduce the smooth strength in the direction perpendicular to image edge. In other words, this is the "oriented smoothness" method, which is driven by image gradient. Weickert and

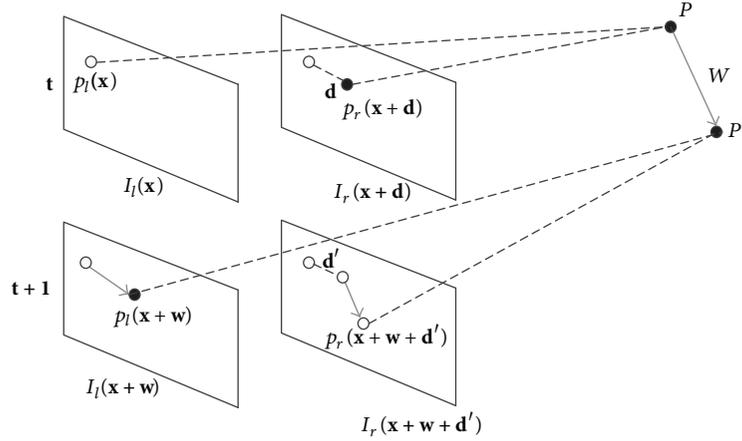


FIGURE 1: Data term in the two-view case. Motion of a single pixel point in two consecutive frames of the image sequences. Brightness and gradient can be considered equal approximately in every image.

Schnörr [12] propose a systematic classification of rotation invariant convex regularizers, which provides a unifying framework for image-driven and flow-driven methods, as well as isotropic and anisotropic regularizers.

Inspired by optical flow regularizers, we propose a scene flow estimation method, which can take the advantages of anisotropic total variation and adaptive regularizer. Different from the structure tensor, our anisotropic diffusion applies robust penalty function to the component of gradient terms in different direction. A weight function used by Blomgren and Chan [13] for image segmentation is introduced to reduce diffusion of flow field in the large gradient area.

2. Related Work

Binocular camera can be used to build the stereo vision system that is used to obtain the stereo image sequences. Li and Sclaroff [14] first estimated scene flow under binocular setting as a joint framework including optical flow estimation and stereo matching. Huguët and Devernay [5] proposed a common variational framework for scene flow estimation by coupling optical flow and stereo matching. A rotating sphere dataset with ground truth is provided publicly in this paper, which allows us to quantitatively evaluate the accuracy of the estimation. After that, KITTI benchmark was published for autonomous driving [2], and scene flow estimation can also be evaluated using this dataset. 400 dynamic scenes were annotated with detailed 3D CAD models and revealed great challenges for scene flow algorithms. Wedel et al. [15] modified Huguët and Devernay's method [5] by decoupling the stereo and motion estimation. Optical flow consistency in both views and the stereo consistency in time $t + 1$ are taken as data term, while the smoothness term of disparity change was separated from the regularization term. With GPU's development, they achieved the promising accuracy results with frame-rate at 5 fps. Basha et al. [6] proposed a 3D point cloud parametrization method which enforces multiview geometric consistency based on brightness consistency and piecewise smoothness assumption. Menze and Geiger [2] used a novel model and dataset for scene flow estimation,

which represents each element in the scene by its rigid motion parameters. And each element can be achieved by super pixel segmentation. Vogel et al. [16] proposed a piecewise rigid model, which combines rigidly moving 3D planes constraint and segmentation regularization, and their method yields better results.

3. Proposed Method of Scene Flow Estimation

3.1. Variational Scene Flow Framework. Our goal is unify the data term and smooth term in a variational framework and obtain the scene flow and 3D structure. Our method follows the variational formulation proposed by Huguët and Devernay [5]. It has the following common form:

$$E(u, v, d_0, d_t) = \int_{\Omega} (E_{\text{data}} + \alpha E_{\text{smooth}}) dx dy. \quad (2)$$

3.2. Data Term. We first obtain a set of stereo image sequences, whose disparity is only in the horizontal direction. Two image sequences between two time steps are shown in Figure 1. According to the time and space constraint from 4 images, data term can be divided into 4 parts

$$E_{\text{data}} = \int_{\Omega} (O_l E_l + O_r E_r + O_{d_0} E_{d_0} + O_{d_t} E_{d_t}) dx dy, \quad (3)$$

where Ω usually refers to the domain of the image plane. O is an occlusion factor, which equals 0 when the pixel is occluded and equals 1 otherwise. Each part contains brightness constancy assumption and gradient constancy assumption to overcome the nonlinear illumination problem.

In other words, the brightness and the gradient of the pixel are nearly constant in the short time and in different views. According to the corresponding relationship, brightness and gradient constancy assumption are given by

$$\begin{aligned} I_l(\mathbf{x}) &= I_l(\mathbf{x} + \mathbf{w}) = I_r(\mathbf{x} + \mathbf{d}_0) + I_r(\mathbf{x} + \mathbf{w} + \mathbf{d}_t), \\ \nabla I_l(\mathbf{x}) &= \nabla I_l(\mathbf{x} + \mathbf{w}) \\ &= \nabla I_r(\mathbf{x} + \mathbf{d}_0) + \nabla I_r(\mathbf{x} + \mathbf{w} + \mathbf{d}_t), \end{aligned} \quad (4)$$

where ∇ is the vector of gradient. We introduce the following abbreviation to explain these 4 terms.

$$\begin{aligned} \Delta(I, \mathbf{x}; I', \mathbf{x} + \mathbf{dx}) &= |I'(\mathbf{x} + \mathbf{dx}) - I(\mathbf{x})|^2 \\ &+ \gamma |\nabla I'(\mathbf{x} + \mathbf{dx}) - \nabla I(\mathbf{x})|^2. \end{aligned} \quad (5)$$

The 4 components of the data term can be written as

$$\begin{aligned} E_l &= \Psi(\Delta(I_l, \mathbf{x}; I_l, \mathbf{x} + \mathbf{w})), \\ E_r &= \Psi(\Delta(I_r, \mathbf{x} + \mathbf{d}_0; I_r, \mathbf{x} + \mathbf{w} + \mathbf{d}_t)), \\ E_{d_0} &= \Psi(\Delta(I_l, \mathbf{x}; I_r, \mathbf{x} + \mathbf{d}_0)), \\ E_{d_t} &= \Psi(\Delta(I_l, \mathbf{x} + \mathbf{w}; I_r, \mathbf{x} + \mathbf{w} + \mathbf{d}_t)), \end{aligned} \quad (6)$$

where E_l and E_r are the left and the right data term energy function of optical flow, respectively. E_{d_0} is the disparity term that represents the stereo matching at time t_0 . E_{d_t} is the disparity term at time $t + 1$. Penalty function ($\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}$, $\varepsilon = 0.001$) is introduced to guarantee convexity and differentiability of the whole data term energy function and remove the outliers at the same time. Outliers are the pixels that are not consistent with the pixels affected by noise, illumination change, and occlusion. γ can be determined according to the intensity of illumination change or just rely on the experiences.

The most obvious thing should be noted here is that if there is an error in any term of u , v , d_0 , and d_t , the error would be propagated to others. We minimize the 4 data terms together to prevent the error from this situation.

3.3. Smooth Term Based on Anisotropic Total Variation. Huguet and Devernay [5] used flow-driven regularization that uses a robust and convex function to keep convexity and differentiability of the smooth term

$$\begin{aligned} E_{\text{smooth}}(u, v, d_0, d_t) &= \int_{\Omega} \Psi(|\nabla u|^2 + |\nabla v|^2) \\ &+ \rho |\nabla(d_0 - d_t)|^2 + \mu |\nabla(d_0)|^2 dx dy. \end{aligned} \quad (7)$$

Here Ψ is used to remove the outliers and preserve the discontinuities of the flow field. The magnitude of the gradient is

$$\Psi(|\nabla u|) = \sqrt{u_x^2 + u_y^2 + \varepsilon^2}. \quad (8)$$

We further derive its scene flow form by minimizing the following problems. In order to simplify the equation, we only consider the 2D case

$$\begin{aligned} F &= \int_{\Omega} \Psi(|\nabla u|^2 + |\nabla v|^2) dx dy \\ &= \int_{\Omega} \sqrt{u_x^2 + u_y^2 + v_x^2 + v_y^2 + \varepsilon^2} dx dy. \end{aligned} \quad (9)$$

It corresponds to the diffusion equation as

$$\begin{aligned} \frac{\partial F}{\partial u} &= \text{div} [\Psi'(|\nabla u|^2 + |\nabla v|^2) \nabla u] \\ &= \frac{\partial}{\partial x} \left(\frac{u_x}{\sqrt{u_x^2 + u_y^2 + v_x^2 + v_y^2 + \varepsilon^2}} \right) \\ &\quad + \frac{\partial}{\partial y} \left(\frac{u_y}{\sqrt{u_x^2 + u_y^2 + v_x^2 + v_y^2 + \varepsilon^2}} \right), \\ \frac{\partial F}{\partial v} &= \text{div} [\Psi'(|\nabla u|^2 + |\nabla v|^2) \nabla v] \\ &= \frac{\partial}{\partial x} \left(\frac{v_x}{\sqrt{u_x^2 + u_y^2 + v_x^2 + v_y^2 + \varepsilon^2}} \right) \\ &\quad + \frac{\partial}{\partial y} \left(\frac{v_y}{\sqrt{u_x^2 + u_y^2 + v_x^2 + v_y^2 + \varepsilon^2}} \right). \end{aligned} \quad (10)$$

In the directions of u and v , the weights are the same and are related to the sum of the squares of the gradient. We propose a new scheme that regularization includes information about the direction of the flow and we define it as anisotropic total variation smoothness term. In the 2D case,

$$\begin{aligned} \text{TV}_{\text{anisotropic}} &= \sum_{i,j} (|\nabla u|) + \sum_{i,j} (|\nabla v|) \\ &= \sum_{i=1}^M \sum_{j=1}^N (|u_{i,j} - u_{i,j+1}| + |u_{i,j} - u_{i+1,j}|) \\ &\quad + \sum_{i=1}^M \sum_{j=1}^N (|v_{i,j} - v_{i,j+1}| + |v_{i,j} - v_{i+1,j}|). \end{aligned} \quad (11)$$

But for outliers, robustness function still should be used. Here we show that it is anisotropic. Similarly, in the 2D case, we define the following equation:

$$\begin{aligned} F &= \int_{\Omega} \Psi(|u_x|^2) + \Psi(|u_y|^2) + \Psi(|v_x|^2) \\ &\quad + \Psi(|v_y|^2) dx dy = \int_{\Omega} \sqrt{u_x^2 + \varepsilon^2} \\ &\quad + \sqrt{u_y^2 + \varepsilon^2} + \sqrt{v_x^2 + \varepsilon^2} + \sqrt{v_y^2 + \varepsilon^2} dx dy. \end{aligned} \quad (12)$$

We minimize it and obtain the following equation:

$$\begin{aligned} \frac{\partial F}{\partial u} &= \frac{\partial}{\partial x} \left(\frac{u_x}{\sqrt{u_x^2 + \varepsilon^2}} \right) + \frac{\partial}{\partial y} \left(\frac{u_y}{\sqrt{u_y^2 + \varepsilon^2}} \right), \\ \frac{\partial F}{\partial v} &= \frac{\partial}{\partial x} \left(\frac{v_x}{\sqrt{v_x^2 + \varepsilon^2}} \right) + \frac{\partial}{\partial y} \left(\frac{v_y}{\sqrt{v_y^2 + \varepsilon^2}} \right). \end{aligned} \quad (13)$$

Obviously, the weights only depend on the direction of each gradient of u_x or v_x and are not related to other direction, which can be regarded as anisotropic diffusion of flow field. In this way, we can control the rate of diffusion according to the gradient of flow field so that we can preserve the sharp edge at the position of nonflat area. And in the flat region, it has the same strength of smooth term. Then a new smooth term can be defined as

$$\begin{aligned} E_{\text{smooth}}(u, v, d_0, d_t) = & \int_{\Omega} \Psi(|u_x|^2) + \Psi(|u_y|^2) \\ & + \Psi(|v_x|^2) + \Psi(|v_y|^2) \\ & + \rho \left[\Psi(|(d_0 - d_t)_x|^2) + \Psi(|(d_0 - d_t)_y|^2) \right] \\ & + \mu \left[\Psi(|(d_0)_x|^2) + \Psi(|(d_0)_y|^2) \right] dx dy, \end{aligned} \quad (14)$$

where ρ and μ are empirical parameters.

3.4. Smooth Term Based on Adaptive Anisotropic Total Variation Flow-Driven Method. A major deficiency of the total variation that Rudin et al. [17] pointed out is staircase in the steady state solution. Marquina and Osher [18] proposed a modified method in 2000, but they just relieved this phenomenon. Different from total variation, there is no such problem with $L2$ norm. Blomgren and Chan [13] used the following norm in the image restoration:

$$R_p(u) = \int_{\Omega} |\nabla u|^{p(|\nabla u|)} dx dy, \quad (15)$$

where p function monotonously decreases and meets the following requirement:

$$p(|\nabla u|) \rightarrow \begin{cases} 2, & |\nabla u| \rightarrow 0 \\ 1, & |\nabla u| \rightarrow \infty. \end{cases} \quad (16)$$

At the edge position, we have

$$R_p(u) \approx \int_{\Omega} |\nabla u| dx dy. \quad (17)$$

In the flat region, we have

$$R_p(u) \approx \int_{\Omega} |\nabla u|^2 dx dy. \quad (18)$$

In other words, this norm can choose from $L1$ to $L2$ according to the gradient magnitude automatically. To achieve this, a rational form can be chosen as follows:

$$p(|\nabla u|) = \frac{|\nabla u| + 2}{|\nabla u| + 1}. \quad (19)$$

The Euler-Lagrange equation for (15) is

$$\begin{aligned} \frac{\partial R}{\partial u} = & \text{div} \left[(p(|\nabla u|) \cdot |\nabla u|)^{p(|\nabla u|)-1} \right. \\ & \left. + \ln p(|\nabla u|) p'(|\nabla u|) \cdot |\nabla u|^{p(|\nabla u|)} \frac{\nabla u}{|\nabla u|} \right]. \end{aligned} \quad (20)$$

Define a function to represent the equation of the bracket in (20):

$$\begin{aligned} q(|\nabla u|) = & p(|\nabla u|) \cdot |\nabla u|^{p(|\nabla u|)-1} \\ & + \ln p(|\nabla u|) p'(|\nabla u|) \cdot |\nabla u|^{p(|\nabla u|)}. \end{aligned} \quad (21)$$

Then (20) can be simplified as

$$\frac{\partial R}{\partial u} = \text{div} \left[q(|\nabla u|) \frac{\nabla u}{|\nabla u|} \right]. \quad (22)$$

Equation (23) can be derived from (21):

$$q(|\nabla u|) = \begin{cases} 2|\nabla u|, & |\nabla u| \rightarrow 0 \\ 1, & |\nabla u| \rightarrow \infty. \end{cases} \quad (23)$$

In the smooth region of flow, the regularization is similar to linear smoothing, and total variation model will be used closing to the flow edge. Combining the robust function $\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}$, our scene flow smooth term can be written as

$$\begin{aligned} E_{\text{smooth}}(u, v, d_0, d_t) = & \int_{\Omega} \Psi(|u_x|^2)^{p(|u_x|)} \\ & + \Psi(|u_y|^2)^{p(|u_y|)} + \Psi(|v_x|^2)^{p(|v_x|)} \\ & + \Psi(|v_y|^2)^{p(|v_y|)} + \rho \left[\Psi(|(d_0 - d_t)_x|^2)^{p(|(d_0 - d_t)_x|)} \right. \\ & \left. + \Psi(|(d_0 - d_t)_y|^2)^{p(|(d_0 - d_t)_y|)} \right] \\ & + \mu \left[\Psi(|(d_0)_x|^2)^{p(|(d_0)_x|)} \right. \\ & \left. + \Psi(|(d_0)_y|^2)^{p(|(d_0)_y|)} \right] dx dy. \end{aligned} \quad (24)$$

3.5. Optimization

3.5.1. Euler-Lagrange Equations. Before the introduction of minimizing the energy function, we first introduce the following abbreviations for the left image. The similar abbreviation is also applicable to the right image.

$$\begin{aligned} I_{Ix} & := \partial_x I_l(\mathbf{x} + \mathbf{w}), \\ I_{Ixz} & := \partial_x I_l(\mathbf{x} + \mathbf{w}) - \partial_x I_l(\mathbf{x}), \\ I_{Iy} & := \partial_y I_l(\mathbf{x} + \mathbf{w}), \end{aligned}$$

$$\begin{aligned}
 I_{lyz} &:= \partial_y I_l(\mathbf{x} + \mathbf{w}) - \partial_y I_l(\mathbf{x}), \\
 I_{lz} &:= I_l(\mathbf{x} + \mathbf{w}) - I_l(\mathbf{x}), \\
 I_{lyy} &:= \partial_{yy}^2 I_l(\mathbf{x} + \mathbf{w}), \\
 I_{lxx} &:= \partial_{xx}^2 I_l(\mathbf{x} + \mathbf{w}), \\
 I_{lxy} &:= \partial_{xy}^2 I_l(\mathbf{x} + \mathbf{w}), \\
 I_l^{t+1} &:= I_l(\mathbf{x} + \mathbf{w}).
 \end{aligned} \tag{25}$$

We also use abbreviations to replace the partial derivatives of the data term.

$$\begin{aligned}
 \Psi'_{fl} &:= \partial_x \Psi \left(|I_l(\mathbf{x} + \mathbf{w}) - I_l(\mathbf{x})|^2 \right. \\
 &\quad \left. + \gamma |\nabla I_l(\mathbf{x} + \mathbf{w}) - \nabla I_l(\mathbf{x})|^2 \right), \\
 \Psi'_{fr} &:= \partial_x \Psi \left(|I_r(\mathbf{x} + \mathbf{w} + \mathbf{d}_t) - I_r(\mathbf{x} + \mathbf{d}_0)|^2 \right. \\
 &\quad \left. + \gamma |\nabla I_r(\mathbf{x} + \mathbf{w} + \mathbf{d}_t) - \nabla I_r(\mathbf{x} + \mathbf{d}_0)|^2 \right), \\
 \Psi'_{st} &:= \partial_x \Psi \left(|I_r(\mathbf{x} + \mathbf{w} + \mathbf{d}_t) - I_l(\mathbf{x} + \mathbf{w})|^2 \right. \\
 &\quad \left. + \gamma |\nabla I_r(\mathbf{x} + \mathbf{w} + \mathbf{d}_t) - \nabla I_l(\mathbf{x} + \mathbf{w})|^2 \right).
 \end{aligned} \tag{26}$$

It should be noted that the regularization term we used is anisotropic, so the Euler-Lagrange equations are not the same as (21). Considering anisotropic total variation of individual variables

$$F = \Psi \left(|u_x|^2 \right)^{p(|u_x|)} + \Psi \left(|u_y|^2 \right)^{p(|u_y|)}, \tag{27}$$

we can get the following Euler equations:

$$\begin{aligned}
 \frac{\partial F}{\partial u_x} &= \left[2|u_x| \cdot p(|u_x|) \cdot \psi'(|u_x|^2) \cdot \Psi(|u_x|^2)^{p(|u_x|)-1} \right. \\
 &\quad \left. + \ln \Psi(|u_x|^2) \cdot p'(|u_x|) \cdot \Psi(|u_x|^2)^{p(|u_x|)} \right] \frac{u_x}{|u_x|}, \\
 \frac{\partial F}{\partial u_y} &= \left[2|u_y| \cdot p(|u_y|) \cdot \Psi'(|u_y|^2) \right. \\
 &\quad \cdot \Psi(|u_y|^2)^{p(|u_y|)-1} + \ln \Psi(|u_y|^2) \cdot p'(|u_y|) \\
 &\quad \left. \cdot \Psi(|u_y|^2)^{p(|u_y|)} \right] \frac{u_y}{|u_y|}.
 \end{aligned} \tag{28}$$

We define the q_1 function as follows:

$$\begin{aligned}
 q_1(s) &= 2s \cdot p(s) \cdot \Psi'(s^2) \cdot \Psi(s^2)^{p(s)-1} + \ln \Psi(s^2) \\
 &\quad \cdot p'(s) \cdot \Psi(s^2)^{p(s)},
 \end{aligned} \tag{29}$$

and then (28) can be simplified as

$$\begin{aligned}
 \frac{\partial F}{\partial u_x} &= q_1(|u_x|) \frac{u_x}{|u_x|}, \\
 \frac{\partial F}{\partial u_y} &= q_1(|u_y|) \frac{u_y}{|u_y|}.
 \end{aligned} \tag{30}$$

According to the variational principle, we minimize the energy function and achieve

$$\nabla E(u, v, d_0, d_t) = 0. \tag{31}$$

Equation (31) corresponds to the 4 Euler equations. We calculate the partial derivation of E with respect to u and let it be equal to 0 to get the following equation:

$$\begin{aligned}
 \beta_l \Psi'_{fl} &\cdot \left(I_{lx} I_{lz} + \gamma (I_{lxx} I_{lzx} + I_{lxy} I_{lyz}) \right) + \beta_r \Psi'_{fr} \\
 &\cdot \left(I_{rx} I_{rz} + \gamma (I_{rxx} I_{rxz} + I_{rxy} I_{ryz}) \right) + \beta_{st} \Psi'_{st} \\
 &\cdot \left((I_r^{t+1} - I_l^{t+1}) (I_{rx} - I_{lx}) \right. \\
 &\quad \left. + \gamma \left((I_{rx} - I_{lx}) (I_{rxx} - I_{lxx}) \right. \right. \\
 &\quad \left. \left. + (I_{ry} - I_{ly}) (I_{rxy} - I_{lxy}) \right) \right) \\
 &- \alpha \left(\frac{\partial}{\partial x} \left[q_1(|u_x|) \frac{u_x}{|u_x|} \right] \right. \\
 &\quad \left. + \frac{\partial}{\partial y} \left[q_1(|u_y|) \frac{u_y}{|u_y|} \right] \right) = 0.
 \end{aligned} \tag{32}$$

The partial derivatives of E with respect to v , $d_0 - d_t$, and d_0 produce similar equations as given below:

$$\begin{aligned}
 \beta_l \Psi'_{fl} &\cdot \left(I_{ly} I_{lz} + \gamma (I_{lxy} I_{lzx} + I_{lyy} I_{lyz}) \right) + \beta_r \Psi'_{fr} \\
 &\cdot \left(I_{ry} I_{rz} + \gamma (I_{rxy} I_{rxz} + I_{ryy} I_{ryz}) \right) + \beta_{st} \Psi'_{st} \\
 &\cdot \left((I_r^{t+1} - I_l^{t+1}) (I_{ry} - I_{ly}) + \gamma \left((I_{rx} - I_{lx}) \right. \right. \\
 &\quad \left. \left. \cdot (I_{rxy} - I_{lxy}) + (I_{ry} - I_{ly}) (I_{ryy} - I_{lyy}) \right) \right) \\
 &- \alpha \left(\frac{\partial}{\partial x} \left[q_1(|v_x|) \frac{v_x}{|v_x|} \right] + \frac{\partial}{\partial y} \left[q_1(|v_y|) \frac{v_y}{|v_y|} \right] \right) \\
 &= 0, \\
 \beta_r \Psi'_{fr} &\cdot \left(I_{rx} I_{rz} + \gamma (I_{rxx} I_{rxz} + I_{rxy} I_{ryz}) \right) + \beta_{st} \Psi'_{st} \\
 &\cdot \left(I_{rx} (I_r^t - I_l^t) + \gamma (I_{rxx} (I_{rx} - I_{lx}) \right. \\
 &\quad \left. + I_{rxy} (I_{ry} - I_{ly})) \right) \\
 &- \alpha p \frac{\partial}{\partial x} \left[q_1(|(d_t - d_0)_x|) \frac{(d_t - d_0)_x}{|(d_t - d_0)_x|} \right]
 \end{aligned}$$

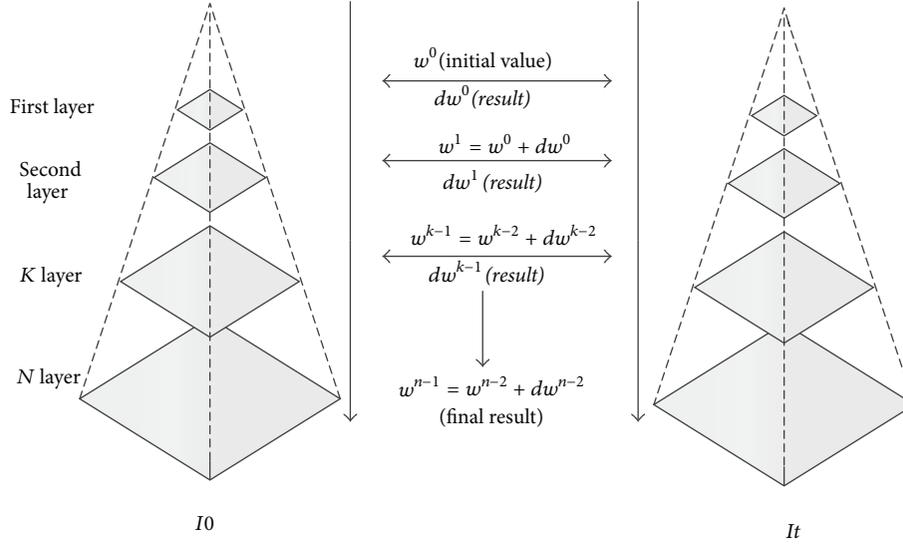


FIGURE 2: Multiresolution strategy for scene flow computation. The left side pyramid is the image at time t . The right side pyramid is the image at time $t + 1$.

$$\begin{aligned}
 & + \alpha \rho \frac{\partial}{\partial y} \left[q_1 \left(|(d_t - d_0)_y| \right) \frac{(d_t - d_0)_y}{|(d_t - d_0)_y|} \right] = 0, \\
 & \beta_r \Psi'_{fr} \cdot (I_{rx}^t I_{rz} + \gamma (I_{rxx}^t I_{rxz} + I_{rxy}^t I_{ryz})) + \beta_{st} \Psi'_{st} \\
 & \cdot (I_{rx}^t I_{rz} + \gamma (I_{rxx}^t (I_{rx}^t - I_{lx}^t) + I_{rxy}^t (I_{ry}^t - I_{ly}^t))) \\
 & - \alpha (\rho + \mu) \left(\frac{\partial}{\partial x} \left[q_1 (|(d_0)_x|) \frac{(d_0)_x}{|(d_0)_x|} \right] \right. \\
 & \left. + \frac{\partial}{\partial y} \left[q_1 (|(d_0)_y|) \frac{(d_0)_y}{|(d_0)_y|} \right] \right) = 0.
 \end{aligned} \tag{33}$$

3.5.2. Multiresolution Strategy. In order to solve the problem of large displacements in the scene and avoid local minima, we introduce the multiresolution method. Referring to Brox et al. [19], the sampling factor is set to 0.9 to avoid a larger span of adjacent layers and local minima. At the same time, we also have a serious problem in the scene flow estimation. If we start to calculate from low resolution layer, there will be a false disparity which is passed to the high resolution layer. For this reason, we start to calculate scene flow from middle resolution layer. A schematic diagram that uses the pyramid hierarchy is given in Figure 2.

In each level of the pyramid, we use the warping method. For the two frames of the K level, I_0^k is obtained according to a certain sampling standard. I_t^k can be obtained by motion compensation based on the image I_t^{k-1} in upper level. We only need to minimize the residual of each layer of the scene flow and the error generated by the large displacement can be avoided. The evaluation of scene flow requires minimizing partial differential equations (32)-(33). This is

done by using Euler-Lagrange equations. We use a nested fixed point iteration scheme suggested by Brox et al. [19] and linearize equations in each layer of pyramid. The inner loop is used to linearize brightness constancy assumption and outer loop linearizes the robust penalty function Ψ . We use SOR (Super Relaxation Iteration) as an iterative method. The full algorithm for scene flow estimation is given in the algorithm scheme (see Algorithm 1).

4. Experiments and Results

In our method, scene flow can be represented by optical flow (u, v) and disparity (d_0, d_t) . We used data sets from Middlebury website to evaluate the accuracy of scene flow estimation. The datasets are captured from 8 calibrated cameras in each scene. The images from cameras 2 and 6 are taken as the stereo pairs at time t , while the images from cameras 4 and 8 are taken as the corresponding stereo pairs at time $t + 1$. We evaluate the accuracy by calculating the RMS error in terms of optical flow, disparity map at time t , and disparity map at $t + 1$ as well as the average angular error (AAE).

We compared our results with Huguet and Devernay [5] and Basha et al. [6] using Venus, Cones, and Teddy from Middlebury datasets as shown in Figures 3, 4, and 5. Table 1 summarizes the RMS errors for optical flow, disparity at time t , and disparity at time $t + 1$ to measure the deviation with the ground truth. AAE errors are computed to measure the deviation with the standard flow field.

RMS errors of optical flow (O.F.), disparity at time t (Disp. t), disparity at time $t + 1$ (Dip. $t + 1$), and AAE errors of optical flow are computed to evaluate the accuracy of our method.

Table 1 shows that the results of our method are more accurate than those of Huguet and Devernay [5] for RMS of optical flow and disparity. Comparing with Basha et al. [6], AAE errors of our method are more accurate. For the RMS errors in disparity t , our method is better than Huguet

```

Ensure: The images at time  $t$  and  $t + 1$  are in the same level
Require:  $a > b \geq c \geq 0$ 
for level =  $c$  to  $b$  do
  Left flow  $(u_l, v_l)$  adaptive anisotropic flow-driven
  Right flow  $(u_r, v_r)$  adaptive anisotropic flow-driven
end for
stereo  $(d_0, d_t)$ 
for level =  $b$  to  $a$  do
  scene flow  $(u_t, v_t, d_0, d_t)$  adaptive anisotropic total variation flow-driven
end for

```

ALGORITHM 1: Scene flow estimation.

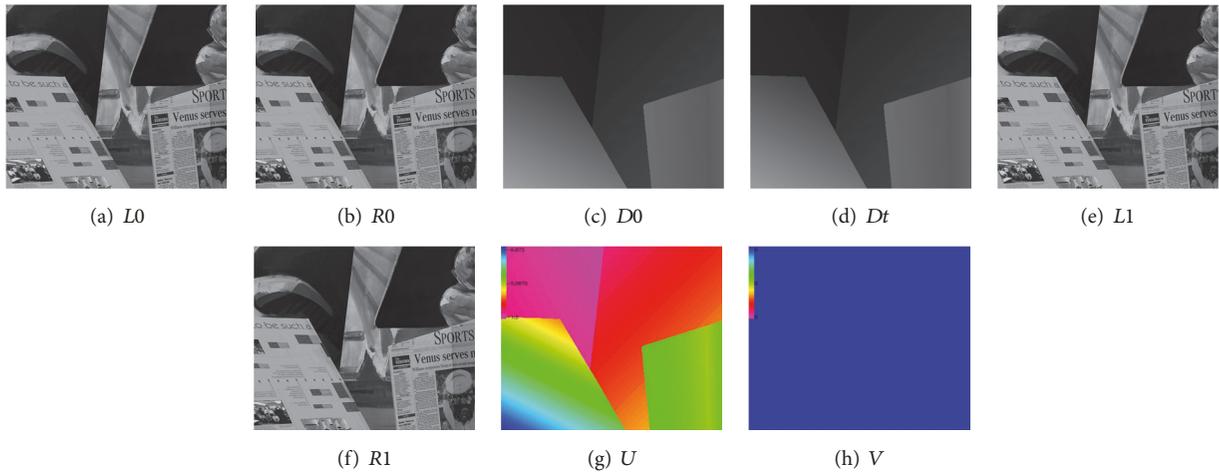


FIGURE 3: Venus of Middlebury stereo datasets. (a) and (b) are left and right image in time t . (c) and (d) are ground truth of disparity maps. (e) and (f) are left and right image in time $t + 1$. (g) is ground truth of optical flow in horizontal direction. (h) is ground truth of optical flow in vertical direction.

TABLE 1: Comparison of scene flow on Middlebury data set.

	RMS O.F.	RMS Disp. t	RMS Disp. $t + 1$	AAE
Cones				
Huguet and Devernay	1.10	2.11	5.24	0.69
Basha et al.	0.58	2.48	2.49	0.39
Ours	0.79	1.22	2.71	0.56
Teddy				
Huguet and Devernay	1.25	2.27	6.93	0.51
Basha et al.	0.57	2.83	2.86	1.01
Ours	0.59	1.20	4.06	0.34
Venus				
Huguet and Devernay	0.31	0.97	1.48	0.98
Basha et al.	0.16	1.06	1.06	1.58
Ours	0.30	1.00	2.20	1.15

and Devernay and Basha et al.'s method. This benefits from proposed smooth term based on adaptive anisotropic total variational method.

Figure 6 shows the performance on the Venus dataset quantitatively. In terms of the component U , our method

keeps more smooth and sharp boundary when comparing with Huguet and Devernay [5].

Cones, Teddy, and Venus datasets have complex edges in the scenes. Traditional scene flow methods are suffering from the isotropic diffusion and receive lower flow accuracy. The proposed adaptive anisotropic diffusion smooth term can smooth scene flow along the direction of motion edges and choose $L1$ or $L2$ regularization automatically, which can receive high flow accuracy at the position of edges and preserve motion edges.

In order to compare our method with other scene flow methods, we used a rotating sphere dataset which includes relatively complex motion and consists of the hemispheres in opposite directions.

Figure 7 shows the left and right views of the sphere. The 3D motion u, v and disparity are also shown for comparison. Figure 8 shows results computed by our method. Table 2 summarizes the RMS errors of sphere sequence.

Table 2 shows that RMS error of our method in optical flow is the same as that of Vogel et al. [16] and better than others. For the RMS error in disparity at time t , our method produces a preferable result.

Rotating sphere dataset has opposite directions motion, which is a very challenging task. The proposed method can

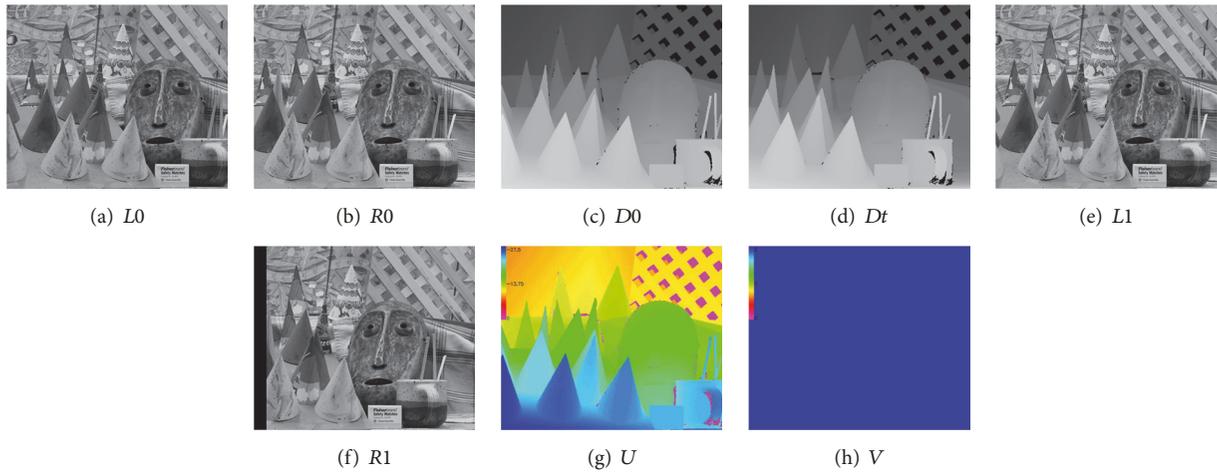


FIGURE 4: Cones of Middlebury stereo datasets. (a) and (b) are left and right image in time t . (c) and (d) are ground truth disparity maps. (e) and (f) are left and right image in time $t + 1$. (g) is ground truth of optical flow in horizontal direction. (h) is ground truth of optical flow in vertical direction.

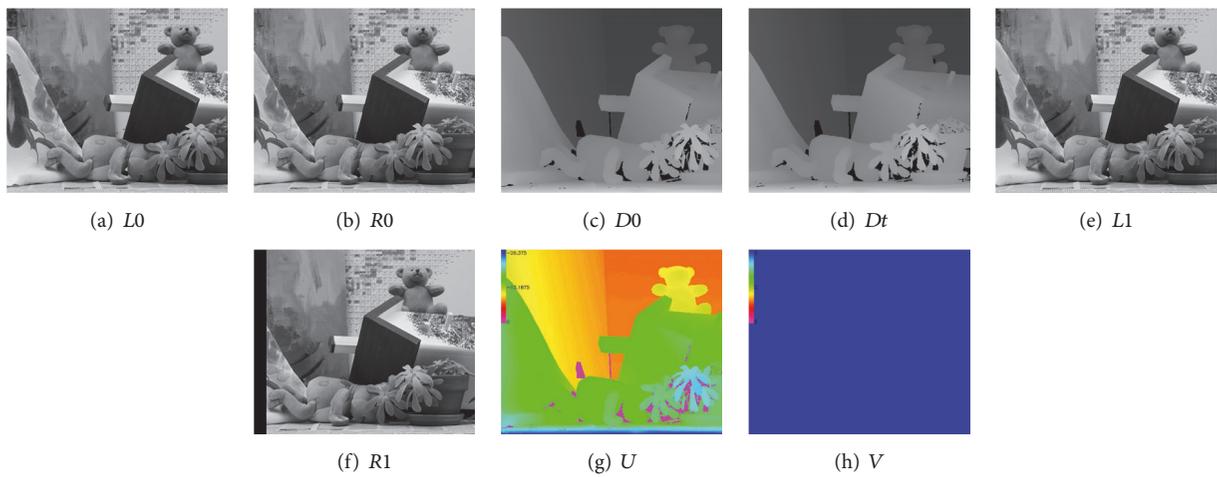


FIGURE 5: Teddy of Middlebury stereo datasets. (a) and (b) are left and right image in time t . (c) and (d) are ground truth disparity maps. (e) and (f) are left and right image in time $t + 1$. (g) is ground truth of optical flow in horizontal direction. (h) is ground truth of optical flow in vertical direction.

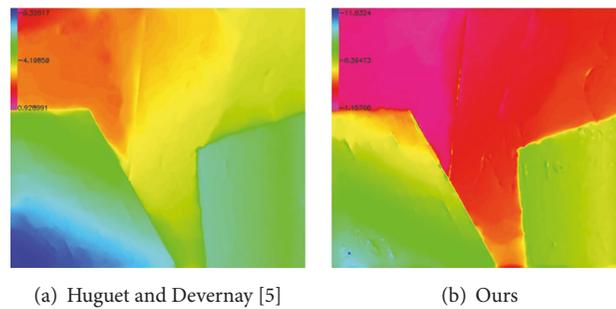


FIGURE 6: Comparison of our U component and Huguet and Devernay's [5]. (a) is the U component of Huguet and Devernay. (b) is our U component.

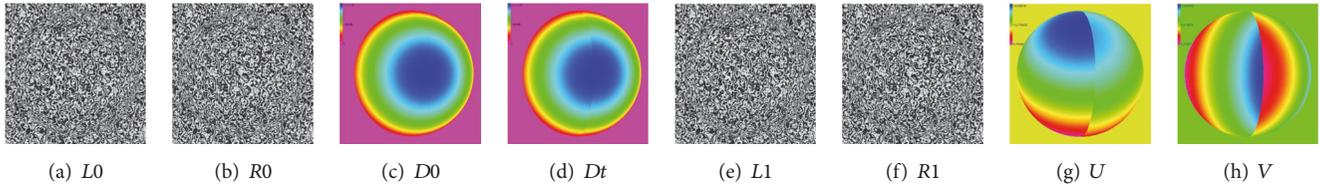


FIGURE 7: Rotating sphere which consists of the hemispheres in opposite directions. (a) and (b) are left and right images in time. (e) and (f) are left and right images in time. (c), (d), (g), and (h) are ground truth of d_0 , d_t , u , and v , respectively.

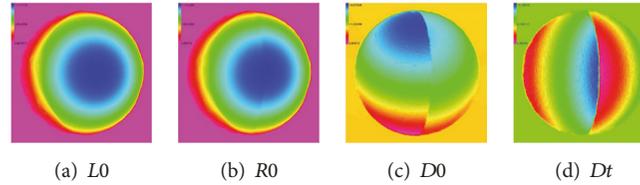


FIGURE 8: Rotating sphere which consists of the hemispheres in opposite directions. (a) and (b) are d_0 , d_t computed by our method. (c) and (d) are u , v computed by our method.

TABLE 2: 2D errors for rotating sphere sequence.

	Huguet and Devernay [5]	Wedel et al. [15]	Vogel et al. [16]	Ours
RMS O.F.	0.69	0.77	0.63	0.63
RMS Disp.	3.80	10.90	2.84	3.67

The RMS errors of Huguet and Devernay [5], Wedel et al. [15], Vogel et al. [16], and our method are compared using optical flow (O.F.) and disparity map (Disp.).

smooth flow field along the motion edge and can also reduce the rate of diffusion at the position of motion edges. In this way, our method can preserve motion edges better.

5. Conclusion

In this article, we proposed an adaptive anisotropic total variation flow-driven regularization method in the variational framework for scene flow estimation. According to the corresponding relationship of the stereo image sequence, data term is divided into 4 parts and meets the brightness and gradient constancy assumption. Regularization term has the advantages of anisotropic smoothing, which applies different weights in the different directions. $L1$ and $L2$ norm can be chosen automatically according to the gradient of the flow field to preserve the boundary of scene flow and reduce the staircase effect at the same time. By minimizing the energy function, we obtain dense optical flow and disparity map. Pyramid optimization method and warping technology are also referenced to deal with large displacement problems.

In the section of experiments, RMS errors of our method for Middlebury datasets Cones and Teddy are better than those of Huguet and Devernay [5] and close to those of Basha et al. [6]. For the rotation sphere dataset, our method's result is better than those of Huguet and Devernay [5] and Wedel et al. [15] and has the same accuracy as that of Vogel et al.

[16]. The experimental results show the effectiveness of our method.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by projects of National Natural Science Foundation of China (61401113) and Natural Science Foundation of Heilongjiang Province of China (LC201426).

References

- [1] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade, "Three-dimensional scene flow," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, vol. 2, pp. 722–729, Corfu, Greece, September 1999.
- [2] M. Menze and A. Geiger, "Object scene flow for autonomous vehicles," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*, pp. 3061–3070, Boston, Mass, USA, June 2015.
- [3] P. Lenz, J. Ziegler, A. Geiger, and M. Roser, "Sparse scene flow segmentation for moving object detection in urban environments," in *Proceedings of the 2011 IEEE Intelligent Vehicles Symposium, IV '11*, pp. 926–932, Baden-Baden, Germany, June 2011.
- [4] W. Waizenegger, I. Feldmann, O. Schreer, and P. Eisert, "Scene flow constrained multi-prior patch-sweeping for real-time upper body 3D reconstruction," in *Proceedings of the 20th IEEE International Conference on Image Processing, ICIP '13*, pp. 2086–2090, Melbourne, Australia, September 2013.
- [5] F. Huguet and F. Devernay, "A variational method for scene flow estimation from stereo sequences," in *Proceedings of the IEEE*

11th International Conference on Computer Vision, ICCV, Rio de Janeiro, Brazil, October 2007.

- [6] T. Basha, Y. Moses, and N. Kiryati, "Multi-view scene flow estimation: a view centered variational approach," *International Journal of Computer Vision*, vol. 101, no. 1, pp. 6–21, 2013.
- [7] A. Bruhn, J. Weickert, T. Kohlberger, and C. Schnörr, "A multigrid platform for real-time motion computation with discontinuity-preserving variational methods," *International Journal of Computer Vision*, vol. 70, no. 3, pp. 257–277, 2006.
- [8] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–203, 1981.
- [9] L. Alvarez and J. Esclarin, "A PDE model for computing the optical flow," in *Proceedings of the CEDYA XVI*, pp. 1349–1356, 1999.
- [10] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [11] 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.
- [12] J. Weickert and C. Schnörr, "A theoretical framework for convex regularizers in PDE-based computation of image motion," *International Journal of Computer Vision*, vol. 45, no. 3, pp. 245–264, 2001.
- [13] P. Blomgren and T. F. Chan, "Color TV: Total variation methods for restoration of vector-valued images," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 304–309, 1998.
- [14] R. Li and R. Sclaroff, "Multi-scale 3D scene flow from binocular stereo sequences," in *Proceedings of the IEEE Workshop on Motion and Video Computing, MOTION 2005*, pp. 147–153, USA, January 2005.
- [15] A. Wedel, C. Rabe, T. Vaudrey, T. Brox, U. Franke, and D. Cremers, "Efficient dense scene flow from sparse or dense stereo data," in *Computer Vision – ECCV 2008*, vol. 5302 of *Lecture Notes in Computer Science*, pp. 739–751, 2008.
- [16] C. Vogel, K. Schindler, and S. Roth, "Piecewise rigid scene flow," in *Proceedings of the 14th IEEE International Conference on Computer Vision, ICCV '13*, pp. 1377–1384, Sydney, NSW, Australia, December 2013.
- [17] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [18] A. Marquina and S. Osher, "A new time dependent model based on level set motion for nonlinear deblurring and noise removal," in *Proceedings of the International Conference on Scale-Space Theories in Computer Vision*, pp. 429–434, 1999.
- [19] 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*, vol. 3024 of *Lecture Notes in Computer Science*, pp. 25–36, Springer, Berlin, Germany, 2004.



Hindawi

Submit your manuscripts at
www.hindawi.com

