Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 9960784, 10 pages
https://doi.org/10.1155/2021/9960784

Research Article

An Improved Fractional-Order Variational Optical Flow Model Combining Structure Tensor

Bin Zhu[1], **Zhaodong Wang**[2], **Lianfang Tian**[3,4], **Jinmei Guo**[1], **Lingjian Wang**[1] and **Jameel Bhutto**[3]

[1]*School of Mechanical and Electronic Engineering, Jiangxi College of Applied Technology, Ganzhou, Jiangxi, China*
[2]*College of Mechanical and Electronic Engineering, Jingdezhen University, Jingdezhen, China*
[3]*School of Automation Science and Engineering, South China University of Technology, Guangzhou, China*
[4]*Research Institute of Modern Industrial Innovation, Science and Engineering, South China University of Technology, Zhuhai 519000, China*

Correspondence should be addressed to Lianfang Tian; chlftian@scut.edu.cn

Received 28 March 2021; Revised 12 July 2021; Accepted 20 July 2021; Published 27 July 2021

Academic Editor: Erik Cuevas

Dealing with problems of illumination changes in optical flow estimation, an improved variational optical flow model is proposed in this paper. The local structure constancy constraint (LSCC) is applied in the data term of the traditional HS (Horn & Schunck) optical flow model to substitute the brightness constancy constraint. The fractional-order smoothness constraint (FSC) is applied in the smoothness term of the HS model. Then, the detailed calculation processes from the optical flow model to the optical flow value are explained. The structure tensor in LSCC is an image feature that is constant in the illumination changes scene. The fractional differential coefficient in FSC can fuse the local neighborhood optical flow vector into the optical flow vector of the target pixel, which can improve the integrity of the motion region with the same motion speed. Combining LSCC with FSC, our improved optical flow model can obtain an accurate optical flow field with clear outline in the illumination abnormity scene. The experimental results show that, compared with other optical flow models, our model is more suitable for the illumination changes scene and can be employed in outdoor motion detection projects.

1. Introduction

Motion detection in the image sequence or video is one of the basic tasks of various image processing projects, which has been widely used in the research and engineering practice of 3D reconstruction, moving object segmentation, moving object tracking, video compression, automatic driving, and so on. The variational optical flow method is one of the most commonly used methods of motion detection. The HS optical flow model [1] is the most classical variational optical flow model. It is composed of a brightness constancy constraint and global smoothness constraint. The brightness constancy constraint requires that the intensity of pixels remains constant in the process of motion, while the smoothness constraint supposes that the motion speed of all pixels in the image changes smoothly. However, both constraints are based on ideal assumptions. In practical applications, these two constraints are often violated. Illumination changes in the scene will break the brightness constant constraint and the global smoothness constraint and lead to the blur of the edge between different motion regions.

Several other constraints on the data term were proposed to improve the performance of the variational optical flow model, such as gradient constancy constraint [2], Laplacian constancy constraint [3], and Hessian constancy constraint [4], but these constraints depend heavily on the intensity difference. Illumination invariant descriptors such as binary-based [5], real value-based [6, 7], and neighborhood-based [8, 9] were applied to substitute the brightness constancy constraint. But these descriptors fail to identify the edge between different motion regions. The structure tensor [10–12], which can embody the shape feature of the motion

objective, is a local structure feature that does not change with the environmental factor. The local structure optical flow model (LSOFM) [13–15], which is established on the structure tensor feature, can improve the robustness in the illumination changes scene. But when a structure includes different motion parts, LSOFM would fail.

Because the fractional optical flow vector gradient can fuse the characteristics of the local neighborhood optical flow vector into the optical flow vector of the target pixel, the integrity of regions with the same motion speed is enhanced, so it can preserve the discontinuity of the optical flow field between different motion regions. The fractional-order variational optical flow model (FOVOFM) is proposed by Chen et al. [16], which uses the fractional-order smoothness constraint in the smoothness term of the HS model. The dual fractional order variational optical flow model (DFOVOFM) is proposed by Zhu et al. [17], which the fractional-order derivative is applied both in data term and smoothness term of the HS model. The fractional-order TV-L1 optical flow model was proposed in [18]. The fractional-order-based optical flow model neglects the structural information in image, so error would exist on the edge and some tiny motion regions.

An improved variational optical flow model is proposed in this paper. Our model is composed of LSCC and FSC. LSCC cannot only ensure the robustness of the model in scenes of illumination changes but also the contour in the image can be enhanced. FSC cannot only preserve the motion discontinuity between different motion regions but also weaken the influence of illumination changes on optical flow estimation. The model proposed in this paper combines the advantages of the two constraints and solves the illumination change problem and the motion discontinuity problem in optical flow estimation, and different motion parts in a structure can be identified.

The contribution of the paper is as follows:

(1) A novel optical flow model combining local structure tensor and fractional-order deviation is designed to improve the robustness of illumination changes

(2) An applicable model simplification method for our model is presented

2. Related Work

The research of the optical flow model never stops. An optical flow model using full four-dimensional cost volume was proposed in [19], which applied structural features and stereo matching pipelines to yield high accuracy. To solve the problems caused by untextured areas, motion boundaries, and occlusions in optical flow estimation, post-hoc confidence measures [20] are applied to assess the per-pixel reliability of the flow. Occlusions have remained one of the key challenges. In [21], a symmetric optical flow model was proposed to address occlusion problems of optical flow estimation. In [22], an interpolation correspondence method is proposed to estimate the optical flow, but these methods are usually susceptible to nonrigid motion and large displacements; to solve this problem, robust interpolation of sparse correspondences is proposed in [23].

A coarse-to-fine scheme combined with an efficient random search strategy for efficient dense correspondence is proposed to estimate the optical flow in [24]. A coarse-to-fine PatchMatch with sparse seeds to sparse matches is applied to optical flow estimation in [25]. A novel differentiable neighbor-search up-sampling (NSU) module [26] is proposed to improve the coarse-to-fine technology for avoiding the optical flow error in edges, thin bars, and holes. A novel optical flow method related to the frequency domain that uses TV-wavelet regularization is applied to accurate optical flow estimation with the frequency-domain regularization in [27].

With the rapid development of artificial intelligence, deep learning-based methods have been widely employed in optical flow estimation. The FlowNet [28] demonstrated that CNNs can be used to estimate the optical flow. In [29], three major improvements were designed for the FlowNet: a subnetwork was elaborated on small displacements, a stacked architecture was applied to warp of the second image with intermediate optical flow, and a schedule of the training data. A new deep network architecture called recurrent all-pairs field transforms is designed to estimate the optical flow in [30]. A deformable cost volume neural network which can estimate the multiscale optical flow in a single high resolution was proposed in [31]. PWC-Net [32], which consists of pyramidal processing, warping, and a cost volume, is designed to estimate optical flow. The network uses the current optical flow estimate to warp the CNN features of the second image and then uses the warped features and features of the first image to construct a cost volume. The drawback of deep learning-based methods is large number of ground truth are needed, and retraining the model is needed for different application scenarios.

3. HS Optical Flow Model

The HS model can be written as follows:

$$E(\mathbf{u}) = E_{\text{data}}(\mathbf{u}) + \lambda E_{\text{smooth}}(\mathbf{u}), \tag{1}$$

where $\mathbf{u} = (u, v)$ represents the optical flow vector, u is the optical flow value on x axis, v is the optical flow value on y axis, $E_{\text{data}}(\mathbf{u})$ represents the brightness constraint equation, $E_{\text{smooth}}(\mathbf{u})$ represents smoothness constraint equation, and λ is a smoothness parameter, and its value is decided by the image quality. In a blurry image with many noises disturbing, the value of λ should be big and vice versa.

The brightness constraint equation based on the bright constancy constraint, which supposes that the brightness value of a pixel is constant during the motion, is as follows:

$$E_{\text{data}}(\mathbf{u}) = \int_{\Omega} \left(I_x u + I_y v + I_t\right)^2 \mathrm{d}\mathbf{x}. \tag{2}$$

The smoothness constraint equation based on the global smoothness constraint, which supposes that the motion speed of all pixels changes smoothly, that is to say, the derivative of the speed is very small, is as follows:

$$E_{\text{smooth}}(\mathbf{u}) = \int_{\Omega} \left(|\nabla u|^2 + |\nabla v|^2\right) \mathrm{d}\mathbf{x}. \tag{3}$$

4. Local Structure Constraint Equation

The structure tensor which is composed of local geometric features of image is an invariable feature to illumination changes, image rotation, image translation, and image scaling. To increase the robustness of the optical flow model, we use the local structure constancy constraint to substitute the brightness constant constraint.

4.1. Image Structure Tensor. The local structure tensor of the two-dimensional image can be expressed as follows:

$$\widetilde{T} = (\nabla I)^T (\nabla I) = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \tag{4}$$

where I represents the brightness function, $\nabla = (\partial_x, \partial_y)$ is the tangent vector along x and y axes, and $\widetilde{T}$ is the structure tensor of a pixel. But $\widetilde{T}$ is easily be affected by noise, so it is always enhanced by Gaussian filter G_σ:

$$\mathbf{T} = G_\sigma * \widetilde{T} = G_\sigma * \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}. \tag{5}$$

4.2. Local Structure Constancy Constrain. The local structure constancy constraint (LSCC) is proposed in this paper to enhance the robustness of the optical flow model. LSCC is based on the assumption that the local structure in an image would not change during the motion:

$$\mathbf{T}(x, y, t) = \mathbf{T}(x + u, y + v, t + 1), \tag{6}$$

where $\mathbf{T}(x, y, t)$ represents the structure tensor function from a pixel (x, y) to t frame. Then, the pixel will move from (x, y, t) to $(x + u, y + v, t + 1)$, and (u, v) is the optical flow value, and $\mathbf{T}(x + u, y + v, t + 1)$ represents the structure tensor of $t + 1$ frame. Equation (6) means that the structure tensor will not change during the motion.

Using the Taylor series expansion and ignoring the minimum term, we can get

$$\mathbf{T}_x u + \mathbf{T}_y v + \mathbf{T}_t = 0, \tag{7}$$

where $\mathbf{T}_x$, $\mathbf{T}_y$, and $\mathbf{T}_t$ are the derivative of the structure tensor function on $x, y,$ and t axes, $\mathbf{T} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, $\mathbf{T}_x = \begin{bmatrix} \partial a/\partial x & \partial b/\partial x \\ \partial b/\partial x & \partial c/\partial x \end{bmatrix}$, $\mathbf{T}_y = \begin{bmatrix} \partial a/\partial y & \partial b/\partial y \\ \partial b/\partial y & \partial c/\partial y \end{bmatrix}$, $\mathbf{T}_t = \begin{bmatrix} \partial a/\partial t & \partial b/\partial t \\ \partial b/\partial t & \partial c/\partial t \end{bmatrix}$, $a = (1/l)\sum_{\mathbf{x}\in\Omega(x_0)} w I_x^2$, $b = (1/l)\sum_{\mathbf{x}\in\Omega(x_0)} w I_x I_y$, $c = (1/l)\sum_{\mathbf{x}\in\Omega(x_0)} w I_y^2$, $\Omega(x_0)$ represents the local neighborhood of the target point, l is the normalization constant, and w represents the weight function.

Then, $E_{\text{data}}(\mathbf{u})$ can be rewritten as follows:

$$E_{\text{data}}(\mathbf{u}) = \int_\Omega \left(\mathbf{T}_x u + \mathbf{T}_y v + \mathbf{T}_t\right)^2 \mathrm{d}\mathbf{x}. \tag{8}$$

5. Fractional Smoothness Constraint Equation

The fractional smoothness constraint (FSC) is applied in our model to further improve the performance of optical flow estimation. We use the fractional-order derivative to substitute the integer-order derivative in the smoothness constraint equation of the HS optical flow model:

$$\begin{aligned} \nabla u &\longrightarrow D^\alpha u, \\ \nabla v &\longrightarrow D^\alpha v, \end{aligned} \tag{9}$$

where $D^\alpha u$ represents α-order derivative of the optical flow vector u on x and y axes, $D^\alpha v$ represents α-order derivative of the optical flow vector v on x and y axes, α is real, $\alpha \in (1, 2)$, $|D^\alpha u| = \sqrt{(D_x^\alpha u)^2 + (D_y^\alpha u)^2}$, and $|D^\alpha v| = \sqrt{(D_x^\alpha v)^2 + (D_y^\alpha v)^2}$.

Then, $E_{\text{smooth}}(\mathbf{u})$ can be designed as follows:

$$E_{\text{smooth}}(\mathbf{u}) = \int_\Omega \left(\left|D^\alpha u\right|^2 + \left|D^\alpha v\right|^2\right)\mathrm{d}\mathbf{x}. \tag{10}$$

6. Improved Fractional Optical Flow Model

Combining LSCC with FSC, we construct our improved optical flow model; using the structure of the HS model, our optical flow model can be described as follows:

$$E(\mathbf{u}) = E_{\text{data}}(\mathbf{u}) + \lambda E_{\text{smooth}}(\mathbf{u}) = \int\int_\Omega \left(\left(\mathbf{T}_x u + \mathbf{T}_y v + \mathbf{T}_t\right)^2 + \lambda\left(\left|D^\alpha u\right|^2 + \left|D^\alpha v\right|\right)^2\right)\mathrm{d}x\mathrm{d}y. \tag{11}$$

Using the Euler–Lagrange equation to minimize the energy function (11), a similar procedure can be seen in [14, 15]; then, we can get

$$\begin{aligned} \mathbf{T}_{xx}(i, j)u(i, j) + \mathbf{T}_{xy}(i, j)v(i, j) + \mathbf{T}_{xt}(i, j) + \lambda\left(D_x^{\alpha *} D_x^\alpha u + D_y^{\alpha *} D_y^\alpha u\right) = 0, \\ \mathbf{T}_{yx}(i, j)u(i, j) + \mathbf{T}_{yy}(i, j)v(i, j) + \mathbf{T}_{yt}(i, j) + \lambda\left(D_x^{\alpha *} D_x^\alpha v + D_y^{\alpha *} D_y^\alpha v\right) = 0, \end{aligned} \tag{12}$$

where $\mathbf{T}_{xx}(i,j)$, $\mathbf{T}_{xy}(i,j)$, $\mathbf{T}_{xt}(i,j)$, $\mathbf{T}_{yx}(i,j)$, $\mathbf{T}_{yy}(i,j)$, and $\mathbf{T}_{yt}(i,j)$ are the second-order derivative of the structure tensor at x, y, and t axes:

$$\begin{aligned}
\mathbf{T}_{xx} &= \frac{\mathrm{d}\mathbf{T}_x}{\mathrm{d}x} = \begin{bmatrix} \frac{\partial a}{\partial x}\frac{\partial a}{\partial x} & \frac{\partial b}{\partial x}\frac{\partial b}{\partial x} \\ \frac{\partial b}{\partial x}\frac{\partial b}{\partial x} & \frac{\partial c}{\partial x}\frac{\partial c}{\partial x} \end{bmatrix}, \\
\mathbf{T}_{xy} &= \frac{\mathrm{d}\mathbf{T}_x}{\mathrm{d}y} = \begin{bmatrix} \frac{\partial a}{\partial x}\frac{\partial a}{\partial y} & \frac{\partial b}{\partial x}\frac{\partial b}{\partial y} \\ \frac{\partial b}{\partial x}\frac{\partial b}{\partial y} & \frac{\partial c}{\partial x}\frac{\partial c}{\partial y} \end{bmatrix}, \\
\mathbf{T}_{yx} &= \frac{\mathrm{d}\mathbf{T}_y}{\mathrm{d}x} = \begin{bmatrix} \frac{\partial a}{\partial y}\frac{\partial a}{\partial x} & \frac{\partial b}{\partial y}\frac{\partial b}{\partial x} \\ \frac{\partial b}{\partial y}\frac{\partial b}{\partial x} & \frac{\partial c}{\partial y}\frac{\partial c}{\partial x} \end{bmatrix}, \\
\mathbf{T}_{yy} &= \frac{\mathrm{d}\mathbf{T}_y}{\mathrm{d}y} = \begin{bmatrix} \frac{\partial a}{\partial y}\frac{\partial a}{\partial y} & \frac{\partial b}{\partial y}\frac{\partial b}{\partial y} \\ \frac{\partial b}{\partial y}\frac{\partial b}{\partial y} & \frac{\partial c}{\partial y}\frac{\partial c}{\partial y} \end{bmatrix}, \\
\mathbf{T}_{xt} &= \frac{\mathrm{d}\mathbf{T}_x}{\mathrm{d}t} = \begin{bmatrix} \frac{\partial a}{\partial x}\frac{\partial a}{\partial t} & \frac{\partial b}{\partial x}\frac{\partial b}{\partial t} \\ \frac{\partial b}{\partial x}\frac{\partial b}{\partial t} & \frac{\partial c}{\partial x}\frac{\partial c}{\partial t} \end{bmatrix}, \\
\mathbf{T}_{yt} &= \frac{\mathrm{d}\mathbf{T}_y}{\mathrm{d}t} = \begin{bmatrix} \frac{\partial a}{\partial y}\frac{\partial a}{\partial t} & \frac{\partial b}{\partial y}\frac{\partial b}{\partial t} \\ \frac{\partial b}{\partial y}\frac{\partial b}{\partial t} & \frac{\partial c}{\partial y}\frac{\partial c}{\partial t} \end{bmatrix},
\end{aligned} \tag{13}$$

where $D_x^{\alpha *} D_x^{\alpha} u$, $D_y^{\alpha *} D_y^{\alpha} u$, $D_x^{\alpha *} D_x^{\alpha} v$, and $D_y^{\alpha *} D_y^{\alpha} v$ denote the dual fractional-order derivative of optical flow value u and v on x and y axes, respectively:

$$\begin{aligned}
D_x^{\alpha *} D_x^{\alpha} u &= \sum_{k=-\infty}^{0} w_{|k|}^{(\alpha)} u(i-k, j) + \sum_{k=0}^{\infty} w_k^{(\alpha)} u(i-k, j) \\
&\approx \sum_{k=-L}^{-1} w_{|k|}^{(\alpha)} \nabla u(i-k, j) \\
&\quad + \sum_{k=1}^{L} w_k^{(\alpha)} \nabla u(i-k, j), \\
D_y^{\alpha *} D_y^{\alpha} u(i,j) &\approx \sum_{k=-L}^{-1} w_{|k|}^{(\alpha)} \nabla u(i, j-k) + \sum_{k=1}^{L} w_k^{(\alpha)} \nabla u(i, j-k),
\end{aligned} \tag{14}$$

where $w_0^{(\alpha)} = 1$ and $w_k^{(\alpha)} = (1 - ((\alpha+1)/k)) w_{k-1}^{(\alpha)}$, $k = 1, 2, \ldots, L$, which denote the size of the neighborhood.

Then, we can get

$$\begin{aligned}
D_x^{\alpha *} D_x^{\alpha} u + D_y^{\alpha *} D_y^{\alpha} u &\approx \sum_{(\bar{i},\bar{j}) \in \chi(i,j)} w_{k_{\bar{i}\bar{j}}}^{(\alpha)} (u(\bar{i},\bar{j}) - u(i,j)), \\
D_x^{\alpha *} D_x^{\alpha} v + D_y^{\alpha *} D_y^{\alpha} v &\approx \sum_{(\bar{i},\bar{j}) \in \chi(i,j)} w_{k_{\bar{i}\bar{j}}}^{(\alpha)} (v(\bar{i},\bar{j}) - v(i,j)),
\end{aligned} \tag{15}$$

where $\chi(i,j)$ represents a neighborhood of point (i,j), $k_{\bar{i}\bar{j}}$ is equal to $\max(|\bar{i} - i|, |\bar{j} - j|)$, $u(\bar{i},\bar{j})$ and $v(\bar{i},\bar{j})$ are the mean of the optical flow vector in the neighborhood, and $u(i,j)$ and $v(i,j)$ are the coordination of the optical flow vector on x and y axes of each pixel in the neighborhood.

Then, the optical flow estimation formula is as follows:

$$\begin{aligned}
\mathbf{T}_{xx}(i,j)u(i,j) + \mathbf{T}_{xy}(i,j)v(i,j) + \mathbf{T}_{xt}(i,j) + \lambda \left(\sum_{(\bar{i},\bar{j}) \in \chi(i,j)} w_{k_{\bar{i}\bar{j}}}^{(\alpha)} (u(\bar{i},\bar{j}) - u(i,j)) \right) = 0, \\
\mathbf{T}_{yx}(i,j)u(i,j) + \mathbf{T}_{yy}(i,j)v(i,j) + \mathbf{T}_{yt}(i,j) + \lambda \left(\sum_{(\bar{i},\bar{j}) \in \chi(i,j)} w_{k_{\bar{i}\bar{j}}}^{(\alpha)} (v(\bar{i},\bar{j}) - v(i,j)) \right) = 0,
\end{aligned} \tag{16}$$

where λ is the smoothness parameter, according to Huerst and Xie [33]. The value of parameter λ is determined by the image quality; when the image is dark, unidentified, blurry, or noisy, λ should be large. Otherwise, λ should be small. In our algorithm, we choose $\lambda \in [5.0, 10.0]$ in relatively clear images, and $\lambda \in [10.0, 15.0]$ in relatively blurry images. α is

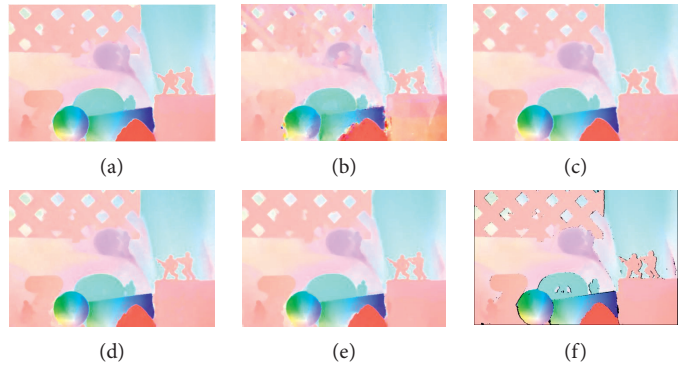


Figure 1: Optical flow estimation results of the army sequence in the Middlebury dataset with different algorithms. (a) HS. (b) FOVOFM. (c) DFOVOFM. (d) LSOFM. (e) Ours. (f) GT.

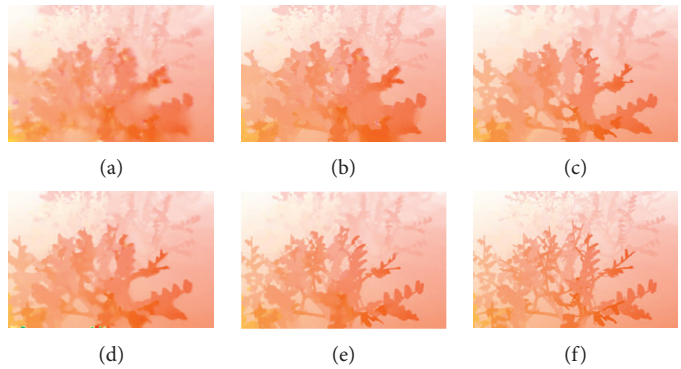


Figure 2: Optical flow estimation results of the grove sequence in the Middlebury dataset with different algorithms. (a) HS. (b) FOVOFM. (c) DFOVOFM. (d) LSOFM. (e) Ours. (f) GT.

the order of fractional derivation in the smoothness term, according to Chen et al. [15], and $\alpha = 1.3$ or 1.4. The accuracy in the case of $\alpha = 1.3$ is relatively higher than in the case of $\alpha = 1.4$, while the latter case needs less iteration times.

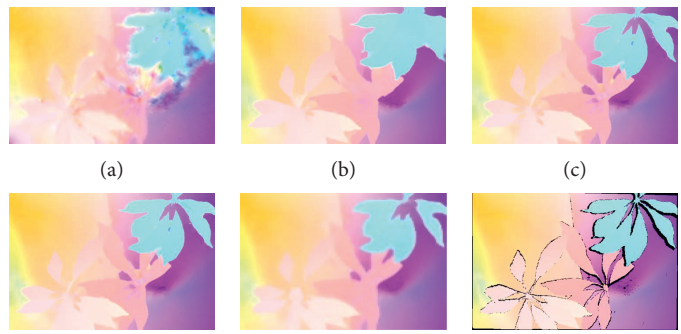
We can obtain the accurate optical flow value by repeated iteration utilizing the Gauss–Seidel method. The iteration will stop when the accuracy reaches the requirement (the difference of the optical flow value between two adjacent iterations is less than the threshold).

7. Experimental Result and Analysis

To testify the superiority and practicability of our optical flow model, Middlebury dataset [34], MPI_Sintel dataset [35], KITTI dataset [36], and an outdoor real scene video are used in experiments. Vertical comparisons are employed to explain the model evolution process; these models include HS model [1], LSCOFM [15], FOVOFM [16], and DFOVOFM [17]. Three widely used high performance models are employed for horizontal comparison: HAST [37], MDP_Flow [38], and PH_Flow [39]. All experiments are run in a PC equipped with Intel Xeon(R) X5600, 3.4 GHZ, and 4 GB memory. The operating system is 64 byte Windows 7, and the software platform is Matlab 10.1.

Figures 1–3 show the color code map by optical flow estimation of different algorithms in the Middlebury database. They are army sequence, grove sequence, and scheflera sequence, respectively. It can be seen in these figures that, from left to right, we can get more and more clear outlines, which means the optical flow results are getting more and more accurate. But, compared to the GT, some highly-textured thin outlines disappear or become adjoined in the results of our model, such as soldier's gun in army, the end of branches in Figure 2(e), and the rhizome parts of leaves in Figure 3(e). That is because using the fractional order derivative, the thin parts would be assimilated by the optical flow field of surrounding pixels. Table 1 shows the AEE/AAE of different algorithms in the Middlebury dataset. We can see that, from top to down, the error rate is becoming more and more small, and the error rate of our model is the lowest.

Figures 4 and 5 show the color code map by optical flow estimation of different algorithms in the MPI_Sintel dataset. They are Cave3 sequence and Shaman1 sequence, respectively. The results presented show that our method can acquire clear and integral motion outlines under the low contrast scene (Cave3) and illumination insufficient scene (Shaman1). HAST can embody motion details such as the end of the cone in Cave3 and the boy's hair in Shaman1, but fails in large nontextured regions such as the tail in Cave3



(a) (b) (c) (d) (e) (f)

Figure 3: Optical flow estimation results of the scheflera sequence in Middlebury dataset with different algorithms. (a) HS. (b) FOVOFM. (c) DFOVOFM. (d) LSOFM. (e) Ours. (f) GT.

Table 1: The AEE/AAE of different algorithms for the Middlebury dataset.

	Algorithm	Army	Mequn	Scheflera	Woden	Grove	Urban	Yosemite	Teddy
AEE	HS	0.22	0.61	1.01	0.78	1.26	1.43	0.16	1.51
	FOVOFM	0.11	0.21	0.32	0.19	0.89	1.25	0.14	0.64
	DFOVOFM	0.10	0.22	0.30	0.16	0.67	0.47	0.17	0.51
	LSOFM	0.11	0.21	0.26	0.17	0.69	0.46	0.17	0.45
	Ours	0.08	0.17	0.18	0.16	048	0.31	0.18	0.38
AAE	HS	8.01	9.13	14.2	12.4	4.64	8.21	4.01	9.16
	FOVOFM	3.82	3.61	5.66	3.93	3.24	4.12	2.73	4.48
	DFOVOFM	3.18	3.21	5.20	2.93	2.94	2.59	2.67	2.33
	LSOFM	2.85	2.24	3.20	3.02	2.54	2.57	2.73	1.73
	Ours	2.64	1.98	2.02	2.83	2.34	2.43	2.23	1.32

(a) (b) (c) (d) (e)

Figure 4: Optical flow results of Cave3 sequence in the MPI_Sintel dataset with different algorithms. (a) HAST. (b) MDP_Flow. (c) PH_Flow. (d) Ours. (e) GT.

and the gusset parts in Shaman1. Similar situations are shown in PH_Flow and MDP_Flow. The deficiencies of our model also exist in the thin detail parts: the end of the cone in Cave3 and details of boy's hair are missing in Shaman1. Table 2 shows the AEE/AAE of different algorithms in the MPI_Sintel dataset. We can see in Table 2 that, in sequences with existing large displacement and occlusion problems, the accuracy of our method is lower than its components. When comes to sequences with illumination changes and illumination insufficient, the accuracy of our method is higher than other algorithms.

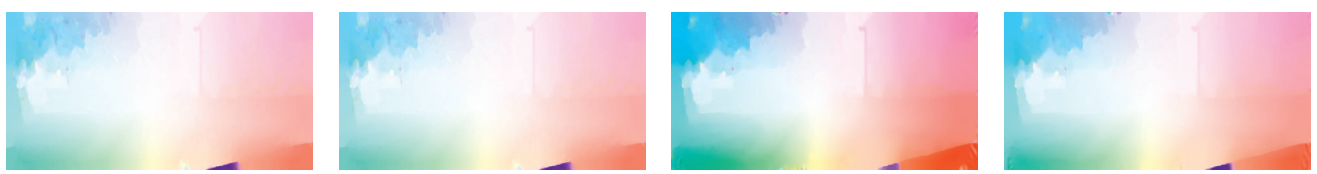
The KITTI dataset is a challenging real-world dataset in which most sequences are affected by illumination abnormity, such as illumination changes, illumination uneven, illumination insufficient, and so on. The experimental results are shown in Figures 6–8. They are sequence 00004, 00009, and 00011, respectively. We can see that our method obtains clear motion outlines in all image sequences. Despite the illumination uneven conditions, the tree outline can be seen clearly in 00004 and the car outline can be seen clearly in 00009. In sequence 00011 with the illumination normal scene, differences between different algorithms are small. Table 3 shows the AEE/AAE of different algorithms in the KITTI dataset. We can see that, in illumination abnormity scenes, the error rate of our model is the lowest, and the average accuracy of our model for the whole sequences in KITTI is the highest.

(a) (b) (c) (d) (e)

Figure 5: Optical flow results of Shama1 sequence in the MPI_Sintel dataset with different algorithms. (a) HAST. (b) MDP_Flow. (c) PH_Flow. (d) Ours. (e) GT.

Table 2: The AEE/AAE of different algorithms in the MPI_SINTEL dataset.

	Algorithm	Market1	Wall	Cave3	Shaman	Market3	Market4	Ambush1	Ambush3
AEE	HAST	3.64	8.39	9.82	2.41	2.20	40.12	43.87	14.12
	PH_Flow	3.64	6.45	8.05	1.41	2.06	28.11	38.69	12.98
	MDP_Flow	0.10	0.22	0.30	0.16	0.67	0.47	0.17	0.51
	LSOFM	0.11	0.21	0.26	0.17	0.69	0.46	0.17	0.45
	Ours	0.08	0.17	0.18	0.16	048	0.31	0.18	0.38
AAE	HAST	8.01	9.13	14.2	12.4	4.64	8.21	4.01	9.16
	PH_Flow	3.82	3.61	5.66	3.93	3.24	4.12	2.73	4.48
	MDP_Flow	3.18	3.21	5.20	2.93	2.94	2.59	2.67	2.33
	LSOFM	2.85	2.24	3.20	3.02	2.54	2.57	2.73	1.73
	Ours	2.64	1.98	2.02	2.83	2.34	2.43	2.23	1.32

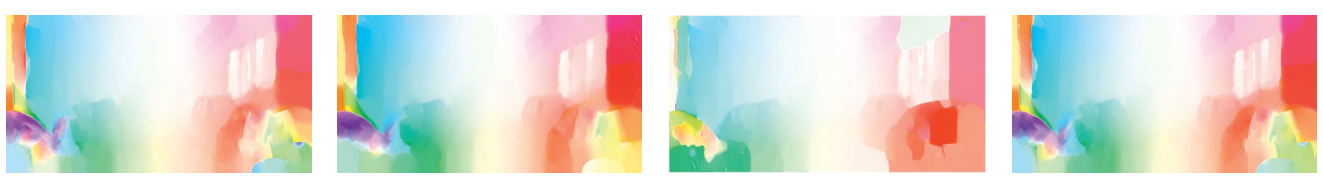


(a) (b) (c) (d)

Figure 6: Visual comparison of 00004 sequence in the KITTI dataset with different methods. (a) HAST. (b) MDP_Flow. (c) PH_Flow. (d) Ours.

Table 3: The AEE of TV_L1 and DFOVOF combined with WRMS-based method or our adaptive method in KITTI.

Algorithm	AVE	000004	000009	000011	000014	000015
HAST	8.85/15.23	2.16/5.53	18.26/25.92	8.05/14.97	5.82/8.03	9.93/18.74
MDP_Flow	8.25/14.73	1.87/5.13	16.91/24.25	7.46/13.89	5.03/7.29	8.26/16.53
PH_Flow	8.57/14.92	1.94/5.26	17.68/25.64	7.73/14.54	5.47/7.83	9.75/17.42
Ours	8.15/13.65	1.73/5.15	16.65/23.64	7.76/14.68	4.79/7.17	8.18/15.72

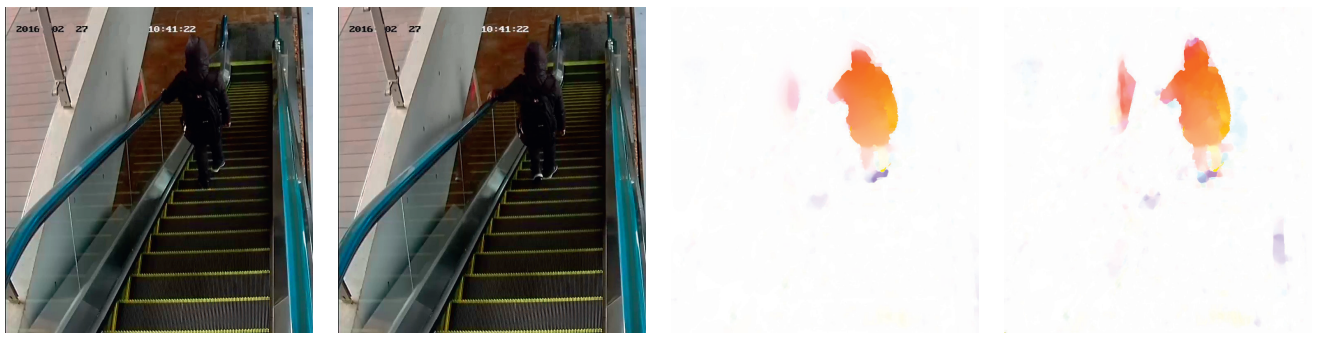


(a) (b) (c) (d)

Figure 7: Visual comparison of 00009 sequence in the KITTI dataset with different methods. (a) HAST. (b) MDP_Flow. (c) PH_Flow. (d) Ours.

(a) (b) (c) (d)

Figure 8: Visual comparison of 00011 sequence in the KITTI dataset with different methods. (a) HAST. (b) MDP_Flow. (c) PH_Flow. (d) Ours.

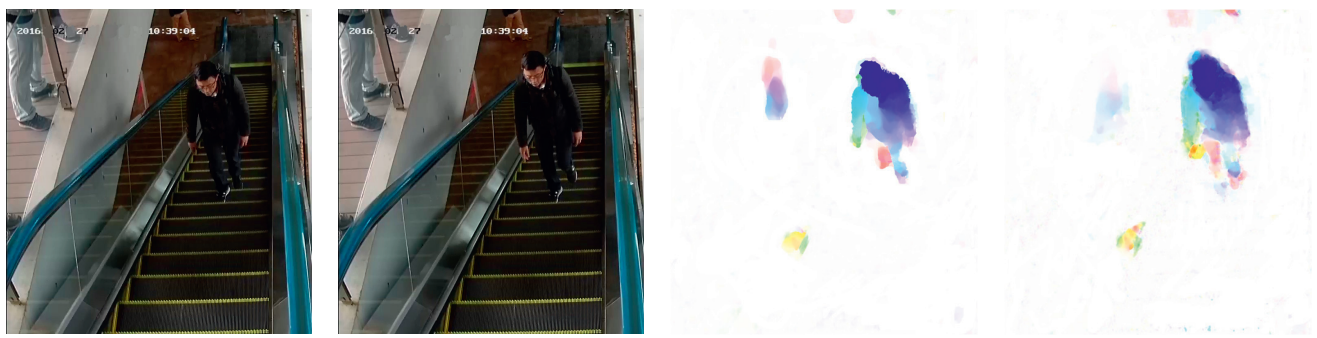


(a) (b) (c) (d)

Figure 9: Inputs and visual comparison of an image sequence from the outdoor video of different algorithms. (a) Input1. (b) Input2. (c) HAST. (d) MDP_Flow.

(a) (b) (c) (d)

Figure 10: Visual comparison of an image sequence from the outdoor video of different algorithms. (a) PH_Flow. (b) DFOVOFM. (c) LSOFM. (d) Ours.



(a) (b) (c) (d)

Figure 11: Inputs and visual comparison of an image sequence from the outdoor video of different algorithms. (a) Input1. (b) Input2. (c) HAST. (d) MDP_Flow.

Figures 9–12 show the color code map of optical flow estimation by different algorithms in an outdoor video. We can see that compared to LSOFM, DFOVOFM has less noise affected regions, but the outline is not so clear. Our model achieves better results with the most intact object outline and least noise affected regions than other models. In Figure 10(d), different motion parts in the right sole can be seen, the optical flow field of the first half foot is pink, and the second half is yellow. The motion of one finger in the right hand in Figures 11 and 12 is different from the other part on that hand, and we can see different color regions in Figure 12(d).

(a) (b) (c) (d)

Figure 12: Visual comparison of an image sequence from the outdoor video of different algorithms. (a) PH_Flow. (b) DFOVOFM. (c) LSOFM. (d) Ours.

8. Conclusion

An improved variational optical flow model is proposed in this paper to solve problems of illumination changes in optical flow estimation. The model applies LSCC and FSC in the data term and smoothness term of the HS model. Then, the detailed calculation processes from the optical flow model to optical flow value are explained. Experimental results in three public datasets and an outdoor video show that our model is superior to other algorithms in illumination abnormity scenes. The disadvantage of our model is that it cannot embody the motion of thin parts. Experimental results testify the performance of our model. Finally, we have conclusions that our model is more suitable for the illumination changes scene and can be employed in outdoor motion detection projects.

Data Availability

Middlebury dataset, MPI_Sintel dataset, and KITTI dataset can be found in references [17–19]. Our own outdoor video can be easily obtained by a normal RGB camera.

Additional Points

The model proposed in this paper has some limitations: When the difference between the RGB value of background and foreground all less than "1," the model would fail; because of the large amount of computation, the running speed of the algorithm is 30 seconds per frame, so it cannot be used for real-time application; when the average brightness is less than 5, the optical flow estimation result of the algorithm would not be accurate.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Department of Natural Resources of Guangdong Province-Offshore Wind Power Project, 2020, Key-Area Research and Development Program of Guangdong Province, 2019B020214001, Key-Area Research and Development Program of Guangdong Province, 2018B010109 001, Guangzhou Industrial Technology Major Research Plan, 20190101121006001, and Fundamental Research Funds for the Central Universities, 2018KZ05.

References

[1] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1, pp. 185–203, 1981.

[2] Z. Tu, W. Xie, J. Cao et al., "Variational method for joint optical flow estimation and edge-aware image restoration," *Pattern Recognition*, vol. 65, no. 3, pp. 11–25, 2016.

[3] N. Monzón, A. Salgado, and J. Sánchez, "Regularization strategies for discontinuity-preserving optical flow methods," *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1580–1591, 2016.

[4] C. Zhang, Z. Chen, M. Wang, M. Li, and S. Jiang, "Robust non-local TV- L1 optical flow estimation with occlusion detection," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 4055–4067, 2017.

[5] M. A. Mohamed, H. A. Rashwan, B. Mertsching, M. A. Garcia, and D. Puig, "Illumination-robust optical flow using a local directional pattern," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 9, pp. 1499–1508, 2014.

[6] M. Drulea and S. Nedevschi, "Motion estimation using the correlation transform," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 22, no. 10, pp. 3260–3270, 2013.

[7] R. W. Liu, L. Shi, W. Huang, J. Xu, S. C. Yu, and D. Wang, "Generalized total variation-based MRI rician denoising model with spatially adaptive regularization parameters," *Magnetic Resonance Imaging*, vol. 32, no. 6, pp. 702–720, 2014.

[8] S. Ali, C. Daul, E. Galbrun, and W. Blondel, "Illumination invariant optical flow using neighborhood descriptors," *Computer Vision and Image Understanding*, vol. 145, no. 12, pp. 95–110, 2016.

[9] W. Lu, J. Duan, Z. Qiu, Z. Pan, R. W. Liu, and L. Bai, "Implementation of high-order variational models made easy for image processing," *Mathematical Methods in the Applied Sciences*, vol. 39, no. 14, pp. 4208–4233, 2016.

[10] M. Andersson and H. Knutsson, "Transformation of local spatio-temporal structure tensor fields," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pp. 25–31, IEEE, Hong Kong, China, April 2003.

[11] W. Förstner and E. Gülch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *International Commission Conference on Fast Processing of Photogrammetric Data*, pp. 281–305, Springer, Interlaken, Switzerland, June 1987.

[12] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 357–365, IEEE, Seattle, WA, USA, June 1994.

[13] Z. Q. Lu, X. Xie, and J. H. Pei, "A robust optical flow computation," *Journal of Electronics*, vol. 5, no. 3, pp. 635–641, 2007.

[14] A. Bruhn, J. Weickert, and C. Schnoerr, "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.

[15] Z. Chen, C. Zhang, X. Zhang, and M. Li, "Coarse-to-fine optical flow estimation with image structure tensor," in *Proceedings of the International Congress on Image and Signal Processing*, pp. 478–485, IEEE, Hangzhou, China, December 2013.

[16] D. Chen, H. Sheng, Y. Chen, and D Xue, "Fractional-order variational optical flow model for motion estimation," *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, vol. 371, Article ID 20120148, 2013.

[17] B. Zhu, L. F. Tian, Q. L. Du et al., "An Improved Fractional order optical flow model for motion estimation," *Mathematical Problems in Engineering*, vol. 2018, Article ID 6278719, 6 pages, 2018.

[18] S. G. Bardeji, I. N. Figueiredo, and E. Sousa, "Optical flow with fractional order regularization: variational model and solution method," in *Proceedings of the 5th International Workshop on Analysis and Numerical Approximation of Singular Problems*, pp. 188–200, Lagos, Portugal, April 2017.

[19] J. Xu, R. Ranftl, and V. Koltun, "Accurate optical flow via direct cost volume processing," in *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition*, pp. 5807–5815, Honolulu, HI, USA, June 2017.

[20] AS. Wannenwetsch, M. Keuper, and S. Roth, "Probflow: joint optical flow and uncertainty estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1182–1191, Venice, Italy, October 2017.

[21] J. Hur and S. Roth, "Mirrorflow: exploiting symmetries in joint optical flow and occlusion estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 312–321, Venice, Italy, October 2017.

[22] Y. Hu, Y. Li, and R. Song, "Robust interpolation of correspondences for large displacement optical flow," in *Proceedings of the IEEE Conference Computer Vision Pattern Recognition*, pp. 481–489, Honolulu, HI, USA, July 2017.

[23] S. Shi, D. Zhang, C. Zhang, Z. Chen, C. Feng, and B. Fan, "Large displacement optical flow estimation based on robust interpolation of sparse correspondences," *IEEE Access*, vol. 8, pp. 227360–227372, 2020.

[24] Y. Li, Y. Hu, R. Song, P. Rao, and Y. Wang, "Coarse-to-fine PatchMatch for dense correspondence," *IEEE Transaction on Circuits System Video Technology*, vol. 28, no. 9, pp. 2233–2245, 2018.

[25] J. Chen, Z. Cai, J. Lai, and X. Xie, "Efficient segmentation-based PatchMatch for large displacement optical flow estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 12, pp. 3595–3607, 2019.

[26] Y. Deng, J. Xiao, S. Z. Zhou, and J. Feng, "Detail preserving coarse-to-fine matching for stereo matching and optical flow," *IEEE Transactions on Image Processing*, vol. 30, pp. 5835–5847, 2021.

[27] J. Chen, J. Lai, Z. Cai, X. Xie, and Z. Pan, "Optical flow estimation based on the frequency-domain regularization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 1, pp. 217–230, 2021.

[28] A. Dosovitskiy, P. Fischer, E. Ilg et al., "Flownet: learning optical flow with convolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, December 2015.

[29] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "Flownet 2.0: evolution of optical flow estimation with deep networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2462–2470, Honolulu, HI, USA, July 2017.

[30] Z. Teed and J. Deng, "RAFT: recurrent all-pairs field transforms for optical flow," 2020, https://arxiv.org/abs/2003.12039v3.

[31] Y. Lu, J. Valmadre, H. Wang, J. Kannala, M. Harandi, and P. Torr, "Devon: deformable volume network for learning optical flow," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pp. 2705–2713, Snowmass, CO, USA, March 2020.

[32] D. Sun, X. Yang, M. Y. Liu, and J. Kautz, "PWC-net: cnns for optical flow using pyramid, warping, and cost volume," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8934–8943, Salt Lake City, UT, USA, June 2018.

[33] W. Huerst and W. Xie, "Weighted root mean square approach to select the optimal smoothness parameter of the variational optical flow algorithms," *Optical Engineering*, vol. 51, no. 3, pp. 720–732, 2012.

[34] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011.

[35] D. J. Butler, J. Wulff, G. B. Stanley et al., "A naturalistic open source movie for optical flow evaluation," in *Proceedings of the 12th European Conference on Computer Vision*, pp. 611–625, Florence, Italy, October 2012.

[36] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proceedings of theIEEE Conference Computer Vision Pattern recognition*, pp. 3354–3361, Providence, RI, USA, June 2012.

[37] Y. Hu, R. Song, Y. Li, P. Rao, and Y. Wang, "Highly accurate optical flow estimation on superpixel tree," *Image and Vision Computing*, vol. 52, no. 3, pp. 167–177, 2016.

[38] X. Li, J. Jiaya, and Y. Matsushita, "Motion detail preserving optical flow estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1744–1757, 2012.

[39] J. Yang and H. Li, "Dense, accurate optical flow estimation with piecewise parametric model," in *Proceedings of theIEEE Conference on Computer Vision and Pattern Recognition*, pp. 1019–1027, Boston, MA, USA, June 2015.