

Research Article

An Adaptive Smoothness Parameter Strategy for Variational Optical Flow Model

Hussain Zaid H. Alsharif,¹ Tong Shu ,¹ Bin Zhu ,² and Zeyad Farisi³

¹School of Business, Hunan University, Changsha, Hunan, China

²School of Mechanical and Electronic Engineering, Jiang Xi College of Applied Technology, Ganzhou, China

³Applied College, Taibah University, Medinah, Saudi Arabia

Correspondence should be addressed to Tong Shu; shutong@hnu.edu.cn

Received 12 October 2021; Revised 22 November 2021; Accepted 7 December 2021; Published 29 December 2021

Academic Editor: Jianping Gou

Copyright © 2021 Hussain Zaid H. Alsharif 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.

The smoothness parameter is used to balance the weight of the data term and the smoothness term in variational optical flow model, which plays very significant role for the optical flow estimation, but existing methods fail to obtain the optimal smoothness parameters (OSP). In order to solve this problem, an adaptive smoothness parameter strategy is proposed. First, an amalgamated simple linear iterative cluster (SLIC) and local membership function (LMF) algorithm is used to segment the entire image into several superpixel regions. Then, image quality parameters (IQP) are calculated, respectively, for each superpixel region. Finally, a neural network model is applied to compute the smoothness parameter by these image quality parameters of each superpixel region. Experiments were done in three public datasets (Middlebury, MPI_Sintel, and KITTI) and our self-constructed outdoor dataset with the proposed method and other existing classical methods; the results show that our OSP method achieves higher accuracy than other smoothness parameter selection methods in all these four datasets. Combined with the dual fractional order variational optical flow model (DFOVOFM), the proposed model shows better performance than other models in scenes with illumination inhomogeneity and abnormality. The OSP method fills the blank of the research of adaptive smoothness parameter, pushing the development of the variational optical flow models.

1. Introduction

Motion detection [1, 2] is a research hotspot of image processing, which is widely applied in motion segmentation [3], target tracking [4], and video [5] surveillance [6]. Optical flow method [7] is one of the most commonly used motion detection methods, which aims to estimate the spatial displacement of each pixel in adjacent image sequences.

Variational optical flow algorithm [8] is one of the most popular optical flow methods, which consists of three parts: data term, smoothness term, and smoothness parameter. In order to improve the performance of the variational optical flow algorithm, many improvements have been done to the data and smoothness term to solve the large displacement problem [9], enhance the robustness against noise and illumination changes [10–12], maintain the discontinuity

between different motion regions [13], highlight the contour of motion regions [14–16], and deal with occlusion problems [17, 18]. However, little attention is paid to the smoothness parameter selecting method. The smoothness parameter acts as a bridge between the data term and smoothness term that directly affects the final result of optical flow estimation. Small smoothness parameter leads to overfitting, whereas higher value causes underfitting.

The optimal smoothness parameter is determined by image quality [19]. Small value of smoothness parameter is assigned to image which has clear outline and high contrast, while large smoothness parameter is assigned to image that has blurry outline and low contrast. In optical flow estimation, many methods determine smoothness parameter through human visual perception which has no exact value, so the smoothness parameter chosen by these methods is

inaccurate. Moreover, many existing methods use the same smoothness parameter throughout the whole image, but the image quality in different region may be different and due to this, a mismatch between the smoothness parameter and image content would exist. Although some methods have been proposed for adjusting the smoothness parameter through image content, none of them can solve these problems well.

Figure 1 shows the illumination inhomogeneous image sequence with different image quality regions and the optical flow estimation results of dual fractional order optical flow model. Figures 1(a) and 1(b) are input image sequences; Figures 1(c) and 1(d) are optical flow estimation results when the smoothness parameter is $\lambda = 10$ and $\lambda = 20$, respectively. It can be seen in Figure 1(c) that the contour of the flow field on the bright part is clearer than dark part, whereas in Figure 1(d), the contour of the flow field on the dark part is clearer than bright part, so, if we set $\lambda = 10$ in bright part of the image and meanwhile $\lambda = 20$ in dark part of the image, the accuracy would be highest among these three methods.

The purpose of this paper is to solve the mismatch problem of smoothness parameter and image content in each part of an image; an adaptive smoothness parameter strategy is designed to select the optimal smoothness parameters for different regions in an image.

For this paper, the main contributions are as follows:

- (1) A new image segmentation algorithm combining SLIC with LMF is proposed to segment the whole image into several superpixel regions
- (2) We would not use the same smoothness parameter in the whole image any more; different smoothness parameters would be assigned to different superpixel region, respectively
- (3) Image quality parameters are used to calculate smoothing parameters for the first time

The rest of the paper is organized as follows. Section 2 reviews the work related to smoothness parameter selecting methods. Section 3 introduces the flowchart of the algorithm. Section 4 introduces the variational optical flow model used in this paper. Sections 5 and 6 describe the proposed image segmentation method and image quality parameters, respectively. Section 7 explains how to use image quality parameters to determine smoothness parameter for each superpixel region. The experimental results are discussed in Section 8 and conclusion is presented in Section 9.

2. Related Work

The smoothness parameter adjustment strategy for the variational optical flow model was first proposed by Nagel in 1986. He proposed oriented smoothness constrain in [20] which used the variation of gray value to adjust the value of smoothness parameter. However, this method fails in illumination insufficient scenes and illumination changes scenes. In [21], a smoothness weight selection method was

proposed; the method uses a blurring operator to calculate the weighted distance, but the spatially varying character needs to be adjusted before it can be applied to variational optical flow model. In [22], a method combining the maximum likelihood function with the cross-validation function was applied to estimate the smoothness parameter, but the method cannot be used in large displacement scenes. A data-driven method based on estimated risk in [23] was used to select the smoothness parameter, but the method is computationally expensive, especially with respect to robust data terms. A Bayesian framework is presented in [24] to estimate the variational model parameters by minimizing the objective function. In [25], an average data constant error criterion based on optimal prediction theory was proposed to dynamically select the smoothness parameter, but the method is limited to linear rigid motion. Kyong et al. applied adaptive convolution kernel function in [26] to address oversegmentation or oversmoothing problem, but when low contrast regions exist, the adaptive strategy would be invalid. In [27], weighted root mean square error criterion was used to estimate the smoothness parameter. The method estimates an initial smoothness parameter by visual perception. Then, search procedures were applied to obtain the optimal parameter around the initial smoothness parameter with the least root mean square error. However, the method has to try many times to find the optimal smoothness parameter, and the same smoothness parameter would be applied in the whole image; when uneven illumination occurs in an image, the method would not be suitable.

In recent years, many other improvements for the variational optical flow model were proposed: TV-L1-based methods [28, 29], deep learning-based methods [30–32], and filtering based methods [33–35], but none of them focus on smoothness parameter.

3. The Flowchart of the Algorithm

The flowchart of the algorithm is in Figure 2: first, the whole image was segmented into several superpixel regions by SLIC and LFM method; then, for each superpixel region, the image quality parameter is calculated; then, a neural network model is designed to use image quality parameter determining the optimal smoothness parameter; the neural network model should be trained by pairs of image quality parameters and optimal smoothness parameter firstly. At last, DFOVOFM and optimal smoothness parameter was used to estimate the optical flow for image sequences.

4. The Dual Fractional Order Optical Flow Model

The dual fractional order optical flow model (DFOVOFM) [36] is a fractional order version of the HS model in which the data term and smoothness term of the HS model are reconstructed by fractional order derivative. The predominance of DFOVOFM is that the model integrated the variation characteristics around the target points, which not only retains the edge and texture details, but also eliminates the influence of subtle noise, so as to improve the robustness



FIGURE 1: Result of DFOVOFM with different smoothness parameters in illumination inhomogeneous scene. (a) Input 1, (b) input 2, (c) $\lambda = 10$, and (d) $\lambda = 20$.

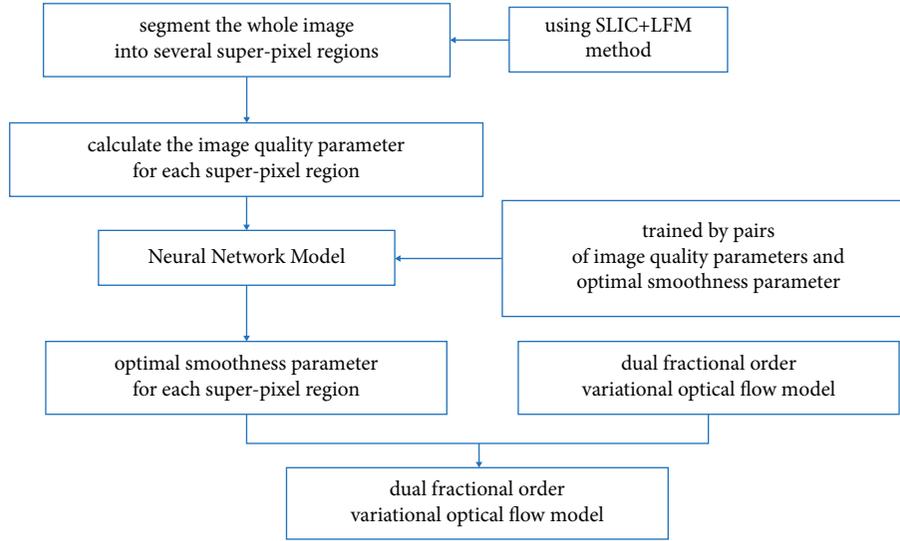


FIGURE 2: Flowchart of our algorithm.

of the model to environmental changes in the process of optical flow estimation.

The DFOVOFM is defined as

$$E(\mathbf{u}) = \int_{\Omega} \left((D_x^\alpha I \cdot u^\alpha + D_y^\alpha I \cdot v^\alpha + D_t^\alpha I)^2 + \lambda \left(|D_x^\beta u|^2 + |D_y^\beta u|^2 + |D_x^\beta v|^2 + |D_y^\beta v|^2 \right) \right) dx, \quad (1)$$

where $E(\mathbf{u})$ represents the energy function of the optical flow model, $\mathbf{u} = (u, v)$ is the optical flow vector, and u, v are the components of optical flow vector on axis x and axis y , respectively. $D_x^\alpha I$, $D_y^\alpha I$, and $D_t^\alpha I$ denote the fractional order derivatives of brightness function I on x, y, t axis. $\alpha \in R^+$ represents the order of fractional order derivative. $D_x^\beta u$, $D_y^\beta u$, $D_x^\beta v$, and $D_y^\beta v$ denote the β th order derivatives of optical flow components u, v on x, y axis, respectively, $\Omega \subset R^3$ denotes a neighborhood of point (i, j, t) , λ is the smoothness parameter, and $\mathbf{x} = (x, y)$.

5. The Superpixel Segmentation

The image quality parameters in a local region are considered to be the same if the RGB values of pixels are same or similar. Therefore, we use the RGB value and position of each pixel to divide the image into several subregions and then calculate the quality parameters of each subregion. Simple Linear Iterative Clustering (SLIC) [37] algorithm with five-dimensional space (r, g, b, x, y) is applied in this paper for image segmentation.

Firstly, the whole image is segmented into K subregions of equal size; then the coordinate and RGB value of center points are reserved. The distance between two adjacent center points is $s = \sqrt{N/K}$, where N represents the pixel number in the image.

Secondly, the Euclidean distance of RGB value d_{rgb} and position information d_{xy} between the center point and its adjacent points in $2s \times 2s$ region is calculated. A weight should be added to d_{xy} because the size of subregion may be different.

$$D_s = d_{\text{rgb}} + \frac{m}{s}d_{xy}, \quad (2)$$

where m is a compensation factor; after that, the D_s values of adjacent pixels in the $2s \times 2s$ region are ranked from small to large. The new superpixel region is composed by the first $s \times s$ pixels; then the center point of every new superpixel region is renewed.

We iterate the second procedure until the average distance between $C_n^{(k)} = (x_n^{(k)}, y_n^{(k)})$ and $C_n^{(k+1)} = (x_n^{(k+1)}, y_n^{(k+1)})$ is less than threshold Th . $C_n^{(k)}$ and $C_n^{(k+1)}$ represent the coordinates of the center point of the k th iteration and $k+1$ th iteration, $n \in [1, s \times s]$. The iteration will terminate when the following equation is true:

$$\frac{1}{s \times s} \sum_{n=1}^{s \times s} \sqrt{(x_n^{(k)} - x_n^{(k+1)})^2 + (y_n^{(k)} - y_n^{(k+1)})^2} < Th. \quad (3)$$

However, according to the conventional SLIC, even if the RGB value is very similar to the center point, the edge points of the superpixel region will be misclassified. In order to correct these erroneously segmented pixels, the local membership function [38] $m_i(x)$ is applied to the edge points:

$$m_i(x) = \frac{|I(x) - \mu_i|^2}{\sum_{i=1}^n |I(x) - \mu_i|^2}, \quad (4)$$

where $m_i(x)$ represent the proportion of pixels belongs to which superpixel region. We reclassify the pixel to the superpixel region of the maximum $m_i(x)$. n is the number of superpixel region around the pixel and $n \in [2, 3, 4]$, $I(x)$ is the brightness function of a pixel, and μ_i is the mean of brightness value of pixels in a superpixel region.

Figure 3 shows the segmentation result of Figure 1(a). Figures 3(a) and 3(b) are the results of original SLIC and improved SLIC, respectively. It can be seen that some of the edge points of superpixel regions have been classified correctly after processing by the local membership function.

6. Image Quality Parameter

Image quality parameter is the quantification of human visual perception of an image and it is one of the basic steps in image processing field. Many image processing tasks can be performed well if the image quality is evaluated accurately. Recently, many image quality parameters are proposed such as entropy [39], brightness [40], colorfulness [41], sharpness [42], and contrast [43]. Author in [44] proposed a Color Root Mean Square Enhancement (CRME) method for color image as the extended version of the gray scale Root Mean Enhancement (RME). K. Panetta in [45] proposed a Color Quality Enhancement (CQE) method which measured the color image quality by the combination of colorfulness, sharpness, and contrast.

Keeping in consideration the aforementioned image quality parameters and the characteristics of the variational optical flow model, we designed more comprehensive image quality parameters which include CRME, colorfulness, sharpness, and contrast in CQE and signal-to-noise ratio in our work.

6.1. CRME. The Color Root Mean Enhancement (CRME) is designed to not only measure the contrast for each color plan, but also measure the difference between the color cube center and the surrounding.

$$CRME = \frac{1000}{k_1 k_2} \sqrt{\sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \left| \frac{\log |I_{i,j} - \sum_{c=1}^3 \zeta_c I_{c1} + I_{c2} + \dots + I_{cn}/n|}{\log |I_{i,j} + \sum_{c=1}^3 \zeta_c I_{c1} + I_{c2} + \dots + I_{cn}/n|} \right|}. \quad (5)$$

To calculate CRME, the image is segmented into $k_1 * k_2$ blocks, $I_{i,j}$ is the center of the block i, j , and $((I_1 + I_2 + \dots + I_n)/n)$ is the mean of the intensity in block i, j . ζ_c is the weight of the color plane, according to NTSC (National Television System Committee) standard, and $\zeta_r = 0.299$, $\zeta_g = 0.587$, and $\zeta_b = 0.144$.

6.2. Contrast. Contrast refers to the measurement of different brightness levels between the brightest white and the darkest black in an image. The larger the difference between the brightest and the darkest, the greater the contrast.

$$\text{contrast} = \sum_{c=1}^3 \zeta_c \frac{1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log \left(\frac{I_{\max,k,l} + I_{\min,k,l}}{I_{\max,k,l} - I_{\min,k,l}} \right). \quad (6)$$

To calculate contrast, the image is segmented into $k_1 * k_2$ blocks, $I_{\max,k,l}$ and $I_{\min,k,l}$ are the max and min intensity value in block k, l , and $\zeta_r = 0.299$, $\zeta_g = 0.587$, and $\zeta_b = 0.144$.

6.3. Sharpness. Sharpness is responsible for the presentation of the details, edges, and textures in the image. For example, in the case of high sharpness, not only are the wrinkles and



FIGURE 3: Superpixel segmentation results of (a) original SLIC and (b) improved SLIC method.

spots of the face on the picture clearer, but also the bulge or depression of the facial muscles can be identified.

$$\text{sharpness} = \sum_{c=1}^3 \zeta_c \frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log\left(\frac{I_{\max,k,l}}{I_{\min,k,l}}\right). \quad (7)$$

To calculate contrast, the image is segmented into $k_1 * k_2$ blocks, $I_{\max,k,l}$ and $I_{\min,k,l}$ are the max and min intensity value in block k, l , and $\zeta_r = 0.299$, $\zeta_g = 0.587$, and $\zeta_b = 0.144$.

6.4. Colorfulness. Color is the most sensitive formal element that can cause our common aesthetic pleasure. Color is one of the most expressive elements, because its nature directly affects people's feelings.

$$\text{colorfulness} = 2 * \log\left(\frac{\sigma_\alpha^2}{\mu_\alpha^{0.2}}\right) * \log\left(\frac{\sigma_\beta^2}{\mu_\beta^{0.2}}\right), \quad (8)$$

where $\alpha = R - G$, $\beta = 0.5(R + G) - B$, N is the pixel number of a superpixel region, $p \in [1, N]$, $\mu_\alpha = 1/N \sum_{p=1}^N \alpha_p$, $\sigma_\alpha^2 = 1/N \sum_{p=1}^N (\alpha_p^2 - \mu_\alpha^2)$, $\mu_\beta = 1/N \sum_{p=1}^N \beta_p$, $\sigma_\beta^2 = 1/N \sum_{p=1}^N (\beta_p^2 - \mu_\beta^2)$.

6.5. Signal-to-Noise Ratio (SNR). SNR is a physical quantity that represents the degree of local image region affected by noise and distortion. SNR is defined as

$$\text{SNR} = 20 \log_{10}\left(\frac{\mu}{\delta}\right), \quad (9)$$

where μ indicates the mean of $\|\nabla_2 I\|_{(i,j)}$ and δ is the corresponding standard deviation.

$$\|\nabla_2 I\|_{(i,j)} = \sqrt{I_x^2 + I_y^2}, \quad (10)$$

where I_x and I_y are the gradient of brightness function $I(x, y)$ for x - and y - axis, respectively. In general, the higher value of SNR results in better image quality.

7. The Calculation of Smoothness Parameter

A neural network model [46–48] is designed to obtain smoothness parameter using image quality parameters.

7.1. The Neural Network Model. The input of the model is composed of five nodes which represent five image quality parameters and the output of the model has only one node

(the corresponding smoothness parameter). After numerous attempts, we finally determined the number of hidden layers and the number of nodes in each layer: five hidden layers: the first layer has 10 nodes, the second layer has 20 nodes, the third layer has 20 nodes, the fourth layer has 10 nodes, and the fifth layer has 5 nodes. The sigmoid function was chosen as the active function. Softmax was chosen as the loss function.

We initialize the weight parameters using the Gaussian distribution with mean value of 0 and variance of 0.01, and then the weight parameters are adjusted by gradient descent method and backpropagation function. The self-built database is used as the training sample, the batch normalization training model is used, and the random deactivation is used to avoid the overfitting of the model. The momentum value is selected as 0.99, the weight penalty item is set as 0.01, and the learning rate is 0.02.

7.2. Self-Construction Dataset. Twelve classes of database are used to construct the training set that are Middlebury database, the MPI_Sintel_final database, MPI_Sintel_clean database, KITTI database, and image sequences of eight outdoor videos with different conditions. It is well-known that higher number of training samples results in higher accuracy of the model. In our model, by increasing the number of samples, the accuracy of model also increases smoothly until it reaches 98.5%, but the accuracy increases slightly when the number of samples exceeds 300. Due to this, we use 300 samples for each class. For each sample, i.e., image sequence, the image quality parameters are calculated and the optimal smoothness parameter is obtained by [27]; then they are used to train the neural network model. Table 1 shows some of the training data.

8. Experimental Result and Analysis

Three public databases: Middlebury [49], MPI_Sintel [50], KITTI [51] and several outdoor videos are used in these experiments. Our proposed algorithm is compared with DFOVOFM [36], TV_L1 [52], HAST [53], MDP_Flow [54], PH_Flow [55], and WRMS based smoothness parameter selection method [27]. All the simulation results are performed on MATLAB 10.0 with system configuration of Windows 7, Intel 3.3 GHZ, and 16 GB RAM.

The intention of these experiments is to show the superiority of our OSP method.

TABLE 1: Image quality parameters and OSP of optical flow dataset.

Dataset	Seq	CRME	Sharp	Contrast	Color	sn	osp
Mid	Army	75	76	85	86	83	3.4
	Backyard	86	83	83	72	76	8.3
MPI_f	Market4	77	68	76	66	67	11.2
	Cave3	43	32	58	63	65	10.5
MPI_c	Market4	95	93	92	93	92	2.8
	Cave3	64	53	94	91	93	3.2
KITTI	00015	68	64	67	67	73	9.6
	00022	63	53	69	58	65	11.3
Outdoor video	Video1	76	67	83	72	78	8.7
	Video2	62	74	85	63	50	10.4
	Video3	63	46	76	51	67	15.7
	Video4	65	41	79	37	23	21.5
	Video5	57	36	53	35	72	21.3
	Video6	47	32	56	27	53	24.5
	Video7	37	24	36	21	41	25.6
	Video8	38	23	37	16	14	27.3

In Figure 4 and Table 2, we applied DFOVOFM with different fixed smoothness parameter and our OSP method to show the relationship between the image content and optimal smoothness parameter, and our OSP method has the highest accuracy. In Figure 5 and Table 3, we compared DFOVOFM + OSP with other classical optical flow models; we can see our model is excellent in illumination uneven scenes, illumination insufficient scenes, and illumination changes scenes.

In Figure 6 and Table 4, we compared DFOVOFM + OSP with DFOVOFM + WRMS, TV-L1 + OSP, and TV-L1 + WRMS; we can see DFOVOFM + OSP is the best among these four algorithms.

In Figures 7 and 8, we demonstrate that, compared to other algorithms, the DFOVOFM + OSP is excellent in outdoor environment.

Figure 4 shows the optical flow estimation results for DFOVOFM with different smoothness parameters. From up to down in Figure 4, there are Army and Grove in Middlebury, Market_4 and Ambush_3 in MPI_Sintel, and two outdoor videos V1 and V2mmc5. First three columns from left to right represent different smoothness parameters having value λ equal to 3, 10, and 20, whereas fourth, fifth, and sixth column represent the WRMS based smoothness parameter selection method, our proposed method, and Ground Truth. It can be seen from Figure 4 that performance of each database changes by varying the value of λ . We get better results in Middlebury database when $\lambda = 3$. However, MPI_Sintel final database performs better when $\lambda = 10$, while results in outdoor videos are better when $\lambda = 20$; we can see that different image quality would have different optimal smoothness parameter. Apparently, it can be clearly seen that our proposed method achieves better results than existing methods for all databases. Our method shows clearer contour than other methods such as gun in the soldier's hand in sequence Army, the end of the branches in Grove, Monster's tail in Market_4, and the contour giant head and waving stick in Ambush_1. In the case of the outdoor videos V1, V2, the result of our method has most complete outline and least error regions caused by

illumination changes. Table 2 shows the error rate (AEE/AE, Average Endpoint Error/Average Angular Error) for the DFOVOFM with different smoothness parameters in Middlebury database. The red with bold font letter indicates the minimal error rate for fixed λ while black with bold font letter shows the error rate of our method. It can be seen in Table 2 that the value of λ affects the accuracy of the optical flow estimation and the error rate of our method is the lowest.

Figure 5 shows the optical flow estimation results for different algorithms in MPI_Sintel database. Each row from top to bottom represents different image sequences that are Market_1, Cave_3, Shaman_1, Ambush_1, and Market_3, whereas each column from left to right shows result of different models that are HAST, MDP_Flow, PH_Flow, our DFOVOFM + OSP, and Ground Truth. It can be seen in Figure 5, compared to other algorithms, the predominance of our method is the robustness in the illumination changes and insufficient illumination scenes, such as Shaman_1; we can see intact contour and bright color for our optical flow estimation result but in other algorithms, the gusset parts are missing. Meanwhile, illumination changes happened on the right side of the running girl in Market_1 which caused big optical flow estimation error in the results of HAST, MDP_Flow, and PH_Flow and our method is the most accurate one. However, our algorithm has a drawback that it does not show the movement of tinny details and some texture details are missing that can be seen in the end of the cone in Cave_3 and the beard of the black man in Ambush_1.

Table 3 shows the average AEE/AE of different algorithms in MPI_Sintel database, where s0-10, s10-40, and s40+ represent the largest displacement that are less than 10 pixels, between 10 and 40 pixels, and exceeding 40 pixels, respectively. We can see in Table 3 that the accuracy of our method is higher than HAST while little lower than MDP_Flow and PH_Flow, because some sequences in MPI_Sintel include large displacement and occlusion problems due to which the accuracy of our method is slightly less than its components.

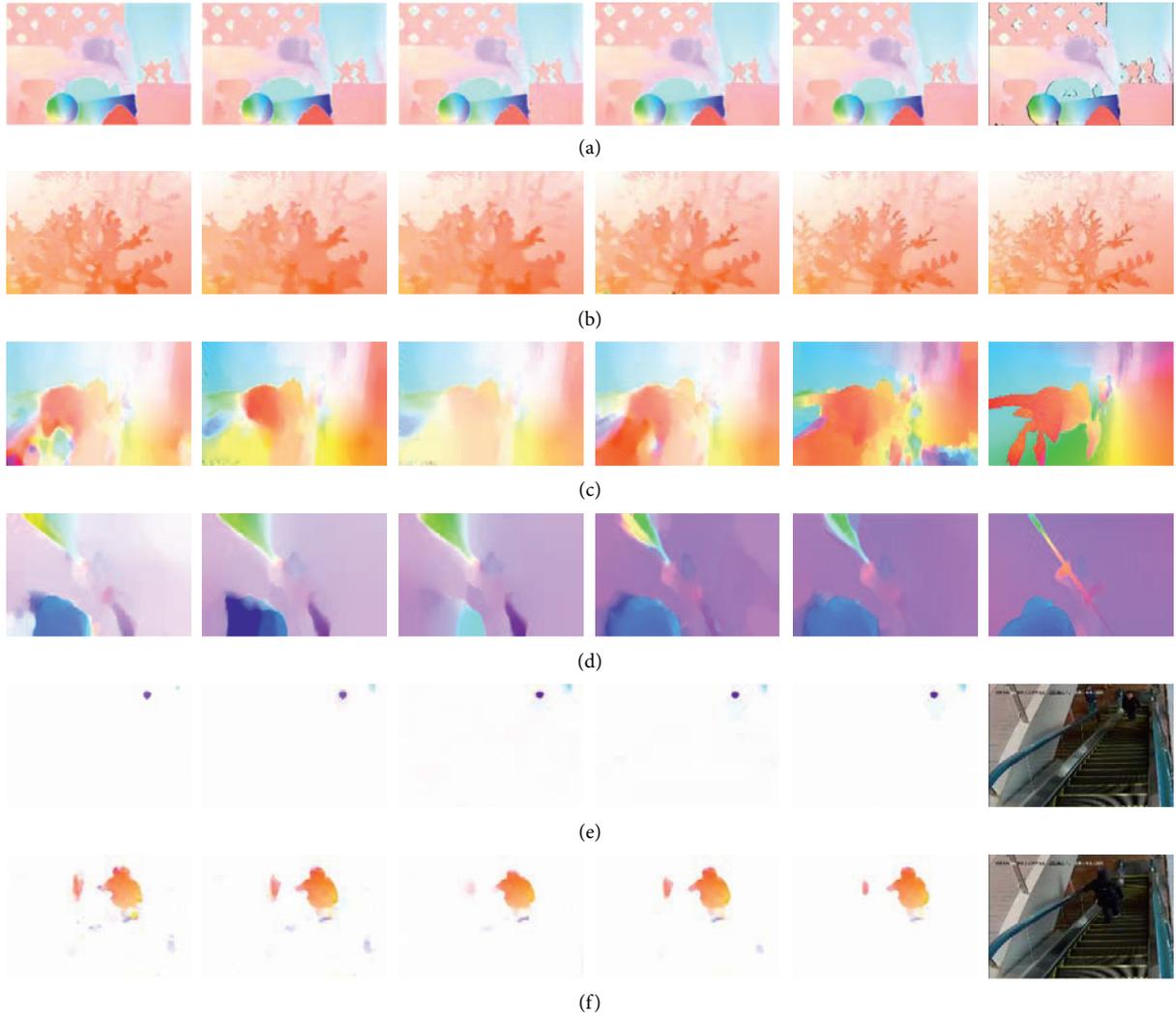


FIGURE 4: Visual comparison of different image sequences from different database with different smoothness parameter setting. (a) Army in Middlebury. (b) Grove in Middlebury. (c) Market_4 in MPI_Sintel. (d) Ambush_3 in MPI_Sintel. (e) Image sequence in outdoor video V1. (f) Image sequence in outdoor video V2.

TABLE 2: The AEE/AAE of DFOVOFM with different smoothness parameters in Middlebury.

	Algorithm	Army	Mequn	Scheflera	Woden	Grove	Urban	Yosemite	Teddy
AEE	$\lambda = 3$	0.08	0.19	0.27	0.16	0.52	0.40	0.16	0.40
	$\lambda = 5$	0.10	0.17	0.25	0.18	0.61	0.36	0.20	0.37
	$\lambda = 8$	0.10	0.19	0.23	0.20	0.64	0.30	0.22	0.43
	$\lambda = 10$	0.11	0.21	0.24	0.25	0.69	0.34	0.30	0.47
	$\lambda = 15$	0.13	0.23	0.28	0.26	0.73	0.42	0.35	0.53
	$\lambda = 20$	0.13	0.25	0.30	0.32	0.78	0.51	0.40	0.58
	OSP	0.08	0.16	0.19	0.15	0.45	0.29	0.15	0.35
AAE	$\lambda = 3$	2.67	2.13	2.28	2.75	2.34	2.68	2.25	1.46
	$\lambda = 5$	2.69	2.03	2.16	2.93	2.39	2.56	2.36	1.33
	$\lambda = 8$	2.73	2.15	2.04	3.14	2.43	2.45	2.48	1.52
	$\lambda = 10$	2.85	2.24	2.20	3.22	2.54	2.57	2.73	1.73
	$\lambda = 15$	2.92	2.36	2.34	3.45	2.65	2.72	2.92	1.95
	$\lambda = 20$	3.04	2.53	2.46	3.62	2.77	2.86	3.05	2.16
	OSP	2.63	1.98	1.96	2.71	2.26	2.42	2.13	1.25

OSP is the result of our own algorithm, and other bold values are the lowest one (using the fixed smoothness parameter) for each sequence.

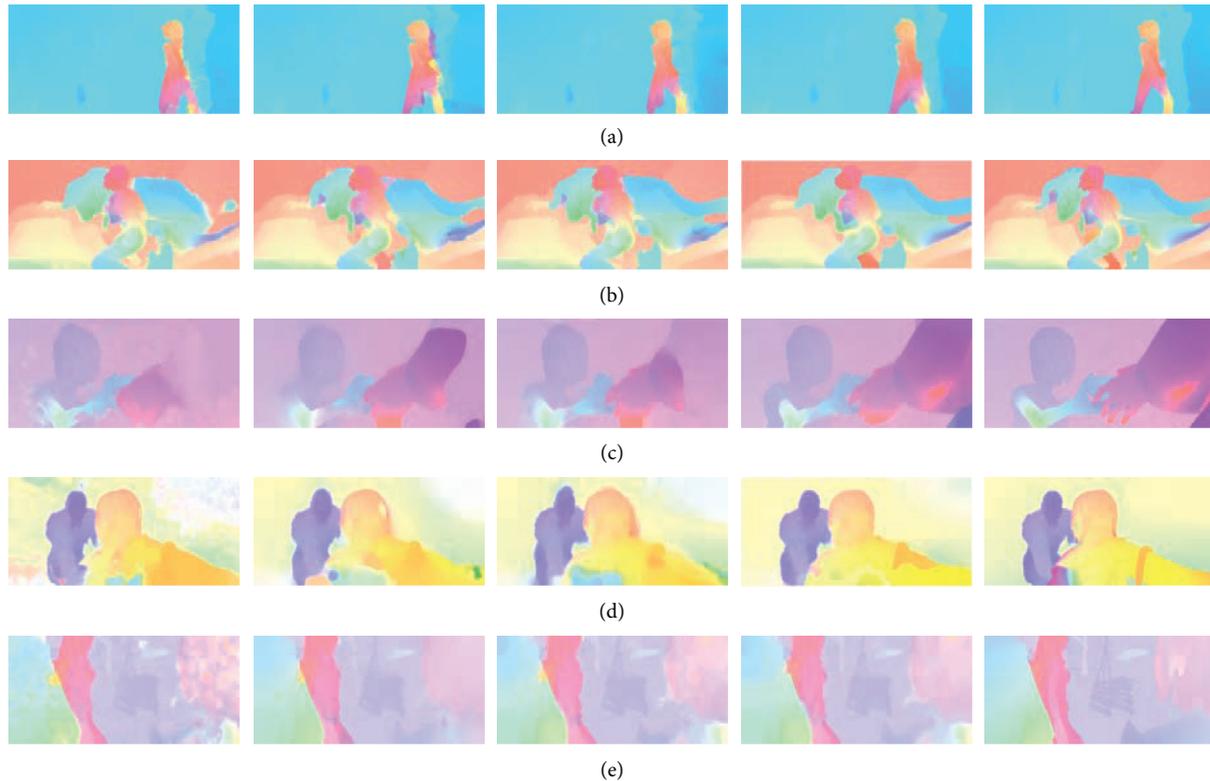


FIGURE 5: Visual comparison of different image sequences from MPI_Sintel database of different algorithm. (a) Market_1. (b) Cave_3. (c) Shaman_1. (d) Ambush_1. (e) Market_3.

TABLE 3: The AEE/AE of DFOVOFM with different smoothness parameters in MPI_SINTEL.

	Algorithm	AEE	s0-10	s10-40	s40+
Clean	HAST	6.802	0.683	3.022	46.33
	MDP_Flow	5.837	0.640	2.603	39.46
	PH_Flow	4.388	0.712	2.855	40.23
	DFOVOFM + OSP	5.236	0.634	2.935	42.87
Final	HAST	9.185	1.614	5.856	54.78
	MDP_Flow	8.445	1.420	5.449	50.51
	PH_Flow	7.432	1.360	5.428	46.03
	DFOVOFM + OSP	7.94	1.337	5.473	48.15

Bold values are the result of our OSP method.

Figure 6 shows the optical flow estimation results for different algorithms in KITTI database. Each row from top to bottom represents different image sequences that are 00004, 00009, 00011, 00014, and 00015 which include illumination inhomogeneous scenes, whereas each column from left to right shows result of different models that are TV_L1 + WRMS, DFOVOFM + WRMS, TV_L1 + OSP, and DFOVOFM + OSP. We can see that results of both TV_L1 + OSP and DFOVOFM + OSP are better than DFOVOFM + WRMS and TV_L1 + WRMS in all five image sequences. Table 4 shows the AEE of the four optical flow estimation methods for 00004, 00009, 00011, 00014, and 00015 image sequences in KITTI database; AVE means the average of AEE of 50 image sequences. It can be justified that our OSP methods have higher accuracy than WRMS methods.

Figures 7 and 8 show the experimental results of the outdoor videos V3 (illumination changes and inhomogeneous regions) and V4 (illumination changes and insufficient regions). First two columns of these figures are input image sequences while remaining columns from left to right show the results of HAST, MDP_Flow, PH_Flow, TV_L1 + WRMS, DFOVOFM + WRMS, TV_L1 + OSP, DFOVOFM + OSP, and manual segmentation result, respectively. Results obtained by HAST technique as depicted in Figures 7(c) and 8(c) have relatively lesser error regions, but the object outline is not intact and several parts of human body are missing. For MDP_Flow method, we can get better intact human body in Figures 7(d) and 8(d), but the results contain more error regions meantime. We can see that DFOVOFM model achieves better results than other models while DFOVOFM + OSP produces even more better

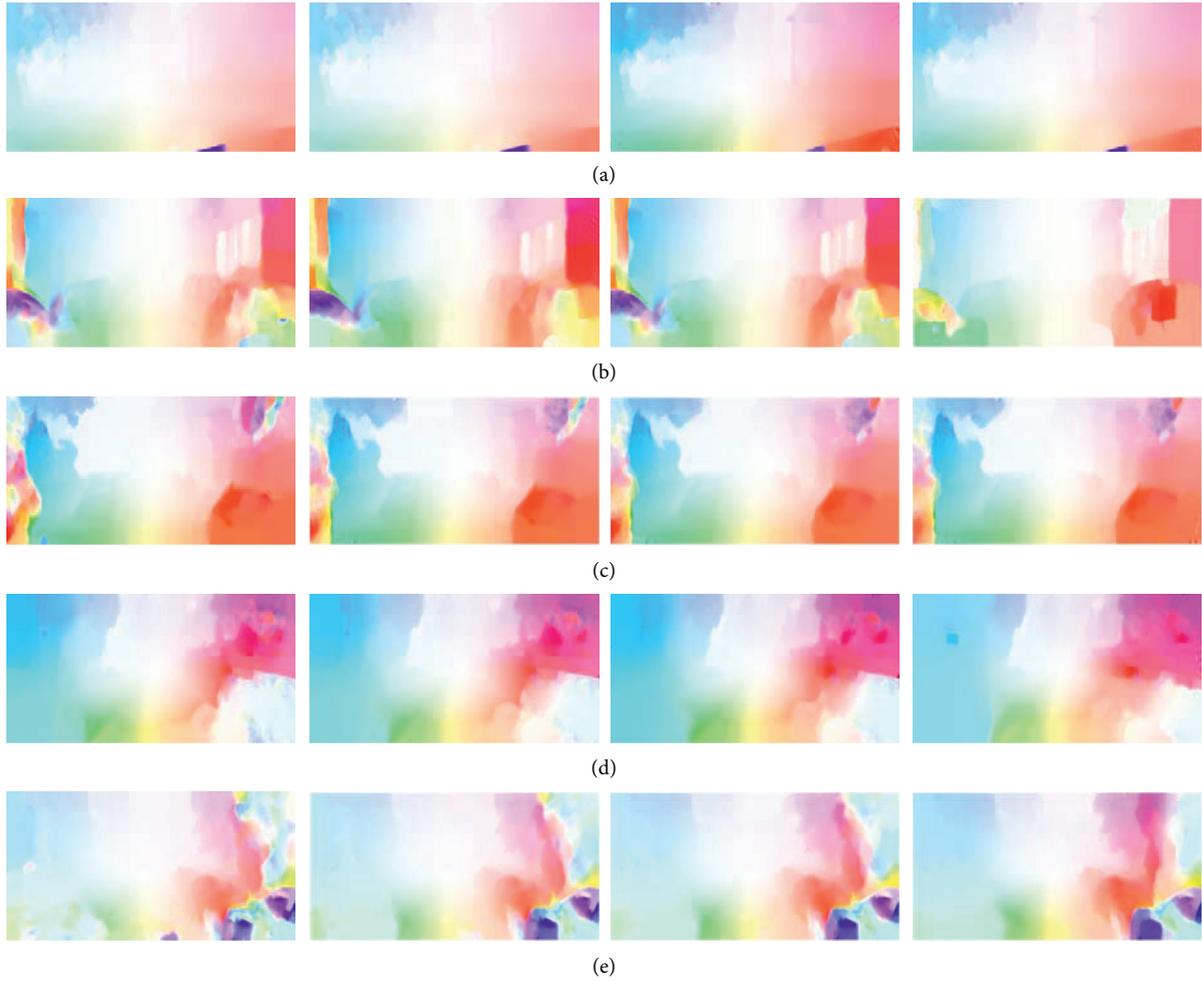


FIGURE 6: Visual comparison of different image sequences with different algorithm. (a) 00004 sequence. (b) 00009 sequence. (c) 00011 sequence. (d) 00014 sequence. (e) 00015 sequence.

TABLE 4: AEE/AAE of DFOVOFM with different smoothness parameters in KITTI.

Algorithm	AVE	000004	000009	000011	000014	000015
TV-L1 + WRMS	8.85/15.23	2.16/5.53	18.26/25.92	8.05/14.97	5.82/8.03	9.93/18.74
TV-L1 + OSP	8.25/14.73	1.87/5.13	16.91/24.25	7.46/13.89	5.03/7.29	8.26/16.53
DFOVOFM + WRMS	8.57/14.92	1.94/5.26	17.68/25.64	7.73/14.54	5.47/7.83	9.75/17.42
DFOVOFM + OSP	8.13/13.85	1.83/5.04	16.45/23.34	7.24/13.46	4.86/7.03	8.05/15.46

Bold values are the result of our OSP method.

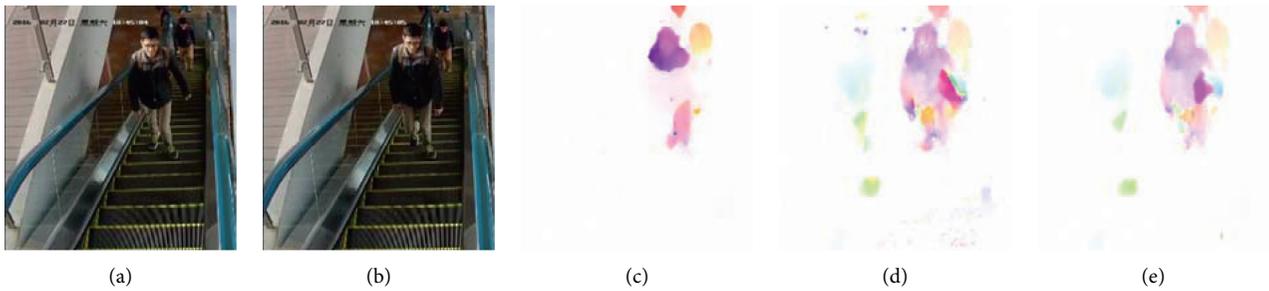


FIGURE 7: Continued.

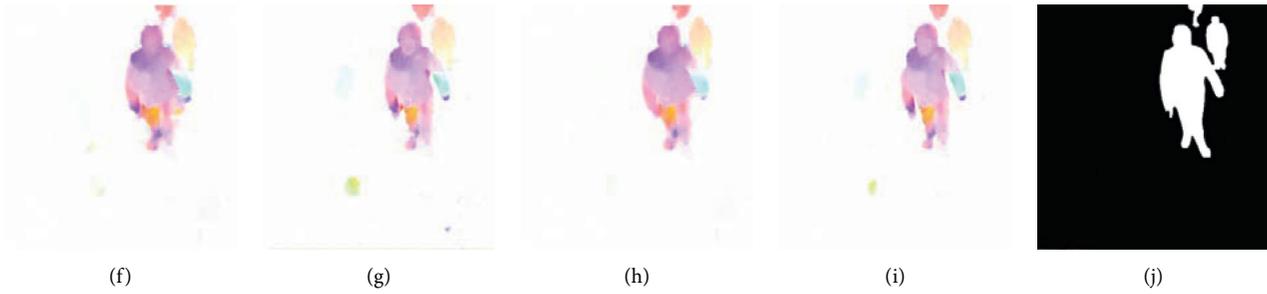


FIGURE 7: Visual comparison of an image sequence from outdoor video V3mmc5 of different algorithms. (a) Input 1. (b) Input 2. (c) HAST. (d) MDP_Flow. (e) PH_Flow. (f) TV_L1 + WRMS. (g) DFOVOFM + WRMS. (h) TV_L1 + OSP. (i) DFOVOFM + OSP. (j) GT.

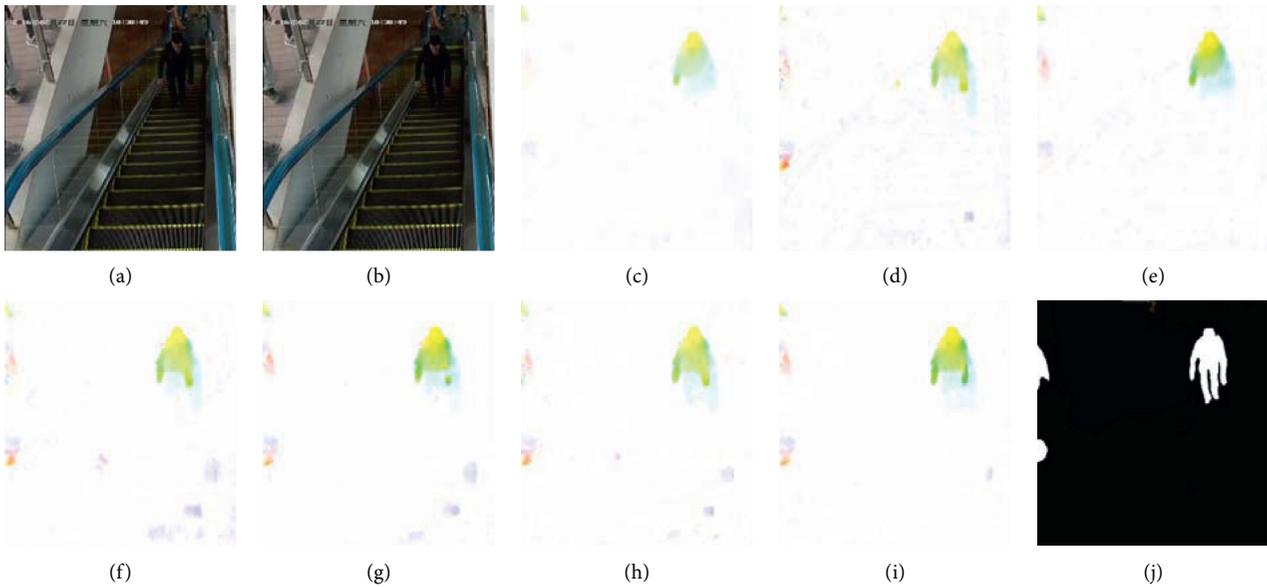


FIGURE 8: Visual comparison of an image sequence from outdoor video V4mmc5 of different algorithm. (a) Input 1. (b) Input 2. (c) HAST. (d) MDP_Flow. (e) PH_Flow. (f) TV_L1 + WRMS. (g) DFOVOFM + WRMS. (h) TV_L1 + OSP. (i) DFOVOFM + OSP. (j) GT.

results with most intact object outline and least error regions. Comparing WRMS methods in Figures 7(f), 7(g), 8(f), and 8(g) with our OSP methods in Figures 7(h), 7(i), 8(h), and 8(i), we can see our OSP method achieves better optical flow estimation.

9. Conclusion

An adaptive smoothness parameter method for variational optical flow model is proposed in this work to solve a low matching rate problem of smoothness parameter and the image content. The main idea of this method is to assign different smoothness parameters for subregions with different image quality parameters. These subregions can be obtained by combining SLIC with LMF while smoothness parameters are calculated by image quality parameters (CRME, contrast, sharpness, colorfulness, and signal to noise ratio) and self-designed neural network model. Simulation results validate that, in comparison to the WRMS method, our method achieves better results especially in

scenes with different image quality regions. Furthermore, combining with DFOVOFM, compared to other optical flow estimation methods, the method shows better performance in illumination abnormal scenes such as illumination changes, illumination insufficiency, and illumination inhomogeneity.

Data Availability

The images used in this study are publicly available.

Conflicts of Interest

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

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant nos. 71771080, 71790593, and 71988101).

References

- [1] M. Nazhamaiti, X. u. Han, L. u. Zheyu, and Q. Fei, "Near-sensor motion detection with energy harvesting image sensor for always-on visual perception," *IEEE Transactions on Circuits and Systems*, vol. 68, no. 9, pp. 3078–3082, 2021.
- [2] F. Guo, W. Wang, Z. Shen, J. Shen, L. Shao, and D. Tao, "Motion-Aware rapid video saliency detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 12, pp. 4887–4898, 2020.
- [3] H. Wang, J. Peng, X. Zheng, and S. Yue, "A robust visual system for small target motion detection against cluttered moving backgrounds," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 3, pp. 839–853, 2020.
- [4] S. Muthu, R. Tennakoon, T. Rathnayake, R. Hoseinnezhad, D. Suter, and A. Bab-Hadiashar, "Motion segmentation of RGB-D sequences: combining semantic and motion information using statistical inference," *IEEE Transactions on Image Processing*, vol. 29, no. 12, pp. 5557–5570, 2020.
- [5] P. Hu, G. Wang, X. Kong, J. Kuen, and Y. Tan, "Motion-guided cascaded refinement network for video object segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 1957–1967, 2020.
- [6] T. Zhuo, Z. Cheng, P. Zhang, Y. Wong, and M. Kankanhalli, "Unsupervised online video object segmentation with motion property understanding," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 237–249, 2020.
- [7] A. Agarwal, S. Gupta, and D. ., K. Singh, "Review of optical flow technique for moving object detection," in *Proceedings of the IEEE International Conference on Contemporary Computing and Informatics*, pp. 409–413, IEEE, Noida, India, October 2016.
- [8] B. Wang, Z. Cai, and L. Shen, "An analysis of physics-based optical flow," *Journal of Computational and Applied Mathematics*, vol. 276, no. 32, pp. 62–80, 2015.
- [9] Z. Chen, H. Jin, and Z. Lin, "Large displacement optical flow from nearest neighbour fields," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 2443–2450, IEEE, Portland, December 2013.
- [10] H. ., A. Rashwan, D. Puig, and M. Garcia, "Improving the robustness of variational optical flow through tensor voting," *Computer Vision and Image Understanding*, vol. 116, no. 9, pp. 953–966, 2012.
- [11] M. Mohamed, H. Rashwan, B. Mertsching, and P. Domenec, "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.
- [12] H. Li, J. Xu, and S. Hou, "Optical flow enhancement and effect research in action recognition," in *Proceedings of the 2021 IEEE 13th International Conference on Computer Research and Development (ICCRD)*, pp. 27–31, Beijing, China, January 2021.
- [13] 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.
- [14] Z. ., G. Tu, W. Xie, J. Cao, C. v. Gemeren, and R. Poppea, "Variational method for joint optical flow estimation and edge-aware image restoration," *Pattern Recognition*, vol. 65, pp. 11–25, 2016.
- [15] X. Xing, Y. Yongjie, and X. Huang, "Real-time object tracking based on optical flow," in *Proceedings of the 2021 International Conference on Computer, Control and Robotics (ICCCR)*, pp. 315–318, Shanghai, China, January 2021.
- [16] X. Zhang, H. Hu, and F. Jiang, "Pedestrian detection based on hierarchical co-occurrence model for occlusion handling," *Neurocomputing*, vol. 168, pp. 861–870, 2015.
- [17] C. Zhang, Z. Zhou, Z. Chen, W. Hu, M. Li, and S. Jiang, "Self-attention-based multiscale feature learning optical flow with occlusion feature map prediction," *IEEE Transactions on Multimedia*, vol. 35, no. 23, pp. 3096–4008, 2021.
- [18] 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.
- [19] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and the noise variance in image restoration and their relation," *IEEE Transactions on Image Processing*, vol. 1, no. 3, pp. 322–336, 1992.
- [20] V. Solo, "A sure-fired way to choose smoothing parameters in ill-conditioned inverse problems," in *Proceedings of the 3rd IEEE International Conference on Image Processing*, pp. 89–92, IEEE, Lausanne, Switzerland, September 1996.
- [21] Xi Yang and R. Klette, "Evaluation of motion analysis on synthetic and real-world image sequences," in *Proceedings of the 25th International Conference of Image and Vision Computing New Zealand*, pp. 1–9, New Zealand, November 2010.
- [22] J. Porée, M. Baudet, F. Tournoux, G. Cloutier, and D. Garcia, "A dual tissue-doppler optical-flow method for speckle tracking echocardiography at high frame rate," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2022–2032, 2018.
- [23] L. Ng and V. Solo, "A data-driven method for choosing smoothing parameters in optical flow problems," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 360–363, IEEE, Santa Barbara, USA, March 1997.
- [24] K. Kai and R. Mester, "Bayesian model selection for optical flow estimation," in *Proceedings of the 29th DAGM Conference on Pattern Recognition*, pp. 142–151, Springer, Berlin, Heidelberg, September 2007.
- [25] H. Zimmer, A. Bruhn, and J. Weickert, "Optic flow in harmony," *International Journal of Computer Vision*, vol. 93, no. 3, pp. 368–388, 2011.
- [26] J. ., L. Kyong, D. Kwon, and I. ., D. Yun, "Optical flow estimation with adaptive convolution kernel prior on discrete framework," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2504–2512, IEEE, San Francisco, USA, June 2010.
- [27] 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. 223–232, 2012.
- [28] C. Zhang, L. Ge, Z. Chen, M. Li, W. Liu, and H. Chen, "Refined TV-L1 optical flow estimation using joint filtering," *IEEE Transactions on Multimedia*, vol. 22, no. 2, pp. 349–364, 2020.
- [29] P. Kumar, "A duality based approach for fractional order Tv-model in optical flow estimation," in *Proceedings of the 5th International Conference on Computing, Communication and Security (ICCCS)*, pp. 1–5, Patna, India, October 2020.
- [30] T.-W. Hui, X. Tang, and C. C. Loy, "A lightweight optical flow CNN-revisiting data fidelity and regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 8, pp. 2555–2569, 2021.

- [31] A. Li, B. Zheng, L. Li, and C. Zhang, "Optical flow estimation and denoising of video images based on deep learning models," *IEEE Access*, vol. 8, no. 9, pp. 144122–144135, 2020.
- [32] M. Zhai, X. Xiang, R. Zhang, N. Lv, and A. El Saddik, "Learning optical flow using deep dilated residual networks," *IEEE Access*, vol. 7, no. 5, pp. 22566–22578, 2019.
- [33] J. Chen, Z. Cai, J. Lai, and X. Xie, "A filtering-based framework for optical flow estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 5, pp. 1350–1364, 2019.
- [34] X. Wang, Z. He, R. Sun, L. You, J. Hu, and J. Zhang, "A crowd behavior identification method combining the streakline with the high-accurate variational optical flow model," *IEEE Access*, vol. 7, no. 10, pp. 114572–114581, 2019.
- [35] L. Mei, J. Lai, X. Xie, J. Zhu, and J. Chen, "Illumination-invariance optical flow estimation using weighted regularization transform," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 2, pp. 495–508, 2020.
- [36] B. Zhu, L. Tian, and Q. Du, "An adaptive dual fractional order variational optical flow model for motion estimation," *IET Computer Vision*, vol. 13, no. 3, pp. 277–284, 2018.
- [37] R. Achanta, A. Shaji, and K. Smith, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [38] R. , R. Gharieb, G. Gendy, and A. Abdelfattah, "Adaptive local data and membership based KL divergence incorporating C-means algorithm for fuzzy image segmentation," *Applied Soft Computing*, vol. 59, no. 2017, pp. 143–152, 2017.
- [39] X. Li, B. Liu, and H. Huang, "-reference image quality assessment based on spatial and spectral entropies," *Signal Processing: Image Communication*, vol. 29, no. 8, pp. 856–863, 2014.
- [40] F. Li, F. Shao, and Q. Jiang, "Local and global sparse representation for no-reference quality assessment of stereoscopic images," *Information Sciences*, vol. 422, pp. 110–121, 2018.
- [41] W. Yan, N. Li, and Z. Li, "An imaging-inspired no-reference underwater color image quality assessment metric," *Computers & Electrical Engineering*, vol. 70, no. 6, pp. 904–913, 2018.
- [42] Y. Zhang, J. Wu, and X. Xie, "Blind image quality assessment with improved natural scene statistics model," *Digital Signal Processing*, vol. 57, no. 13, pp. 56–65, 2016.
- [43] M. Oszust, "No-reference image quality assessment with local features and high-order derivatives," *Journal of Visual Communication and Image Representation*, vol. 56, no. 8, pp. 15–26, 2018.
- [44] Z. Zeng, W. Yang, and W. Sun, "No-reference image quality assessment for photographic images based on robust statistics," *Neurocomputing*, vol. 313, no. 3, pp. 111–118, 2018.
- [45] K. Panetta, C. Gao, and S. Agaian, "No reference color image contrast and quality measures," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 3, pp. 643–651, 2013.
- [46] M. Rincon, "Numerical analysis for a nonlinear model of elastic strings with moving ends," *Applied Numerical Mathematics*, vol. 135, no. 3, pp. 146–164, 2019.
- [47] Y. Qi, J. Lu, and R. Scardovelli, "Computing curvature for volume of fluid methods using machine learning," *Journal of Computational Physics*, vol. 377, no. 63, pp. 155–161, 2019.
- [48] R. Wason, "Deep learning: evolution and expansion," *Cognitive Systems Research*, vol. 52, pp. 701–708, 2018.
- [49] S. Baker, D. Scharstein, and J. , P. Lewis, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011.
- [50] D. Butler, J. Wulff, and G. Stanley, "A naturalistic open source movie for optical flow evaluation," in *Proceedings of the 12th European Conference on Computer Vision*, pp. 611–625, Springer, Berlin, Heidelberg, October 2012.
- [51] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proceedings of the IEEE Conference Computer Vision and Pattern Recognition*, pp. 3354–3361, IEEE, RI, USA, October 2012.
- [52] C. Zach, T. Pock, and H. Bischof, "A duality based approach for realtime TV-L1 optical flow," *Dagm*, vol. 4713, pp. 214–223, 2007.
- [53] Y. Hu, R. Song, and Y. Li, "Highly accurate optical flow estimation on superpixel tree," *Image and Vision Computing*, vol. 52, no. 3, pp. 167–177, 2016.
- [54] X. Li, J. , Y. Jia, and M. Yasuyuki, "Motion detail preserving optical flow estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1744–1757, 2012.
- [55] J. Yang and H. Li, "Dense, accurate optical flow estimation with piecewise parametric model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1019–1027, Boston, USA, October 2015.