ILLUMINATION ROBUST OPTICAL FLOW ESTIMATION BY ILLUMINATION-CHROMATICITY DECOUPLING

Sungheon Park and Nojun Kwak

Graduate School of Convergence Science and Technology, Seoul National University, Korea
{sungheonpark, nojunk}@snu.ac.kr

ABSTRACT

In this paper, a novel optical flow algorithm which is robust to illumination variation is proposed. HSL color space is adopted to decouple illumination and chromaticity information. The chromaticity component is normalized by chroma and transformed to the cartesian coordinate. Then, the decoupled distance is defined using both illumination and chromaticity. Cost function for the optical flow is formulated using l_1 norm of the decoupled distance with Huber norm regularization term. The cost function is efficiently minimized by utilizing Legendre-Fenchel transform. Optical flow field is further refined via weighted median filter whose weight is also based on the decoupled distance. Experimental results show that the proposed method works robustly even in the presence of severe illumination variation.

***Index Terms*—** Optical flow, HSL color space, illumination robust

1. INTRODUCTION

Estimating motion densely in the consecutive image frames has been studied for more than 30 years. In contrast to the first algorithm of Horn and Schunck [1] which exploited l_2 distance for designing the cost, TV-L1 optical flow [2, 3] used l_1 distance which is more robust to outliers than l_2 distance. Both aforementioned methods and most optical flow algorithms are based on the intensity consistency or the gradient consistency assumption. However, the assumption does not hold when there is a severe intensity change in the consecutive image. There are a couple of attempts to overcome this problem. Mileva et al. [4] proposed the data cost using photometric invariants such as hue from HSV color space, normalized RGB, and spherical transform of RGB coordinates. However, information from intensity can be lost when considering only photometric invariants. Recently, Kumar et al. [5] achieved robustness by decoupling illumination and reflectance from the grayscale image.

In this paper, we propose a novel optical flow estimation algorithm which is robust to illumination variation. HSL color space is adopted to separate the luminance and the chromaticity component. Weighted l_1 distance of illumination and chromaticity are used as a data term of the optical flow cost to achieve the robustness on the illumination variation. Huber norm penalizer with edge-preserving weight is used as a regularization term. The optical flow is further refined using weighted median filter whose weight is also calculated from difference of colors. Based on the work of Chambolle and Pock [6], the proposed cost function can be efficiently minimized. Experiments are conducted on the MPI Sintel dataset [7] which contains severe illumination variation and on the real images with specular objects. The result shows that the proposed algorithm is not only robust to illumination changes, but it also preserves discontinuity of optical flow over boundary regions of the image.

2. ILLUMINATION-CHROMATICITY DECOUPLING USING HSL COLOR SPACE

HSL color space consists of three components, hue, saturation, and lightness. Therefore, illumination and chromaticity component can be easily decoupled by separating HSL space into L and HS. HS space is exploited in many applications to achieve the robustness to illumination change [8, 9]. We exploit chroma as a radial dimension of a color space instead of saturation which is defined as

$$C = \max(R, G, B) - \min(R, G, B). \tag{1}$$

Hue-chroma-lightness color model is a bi-conic shape. Colors near the tip of the bi-conic has smaller radius. Thus, the chroma should be normalized using the lightness component. First, lightness is shifted so that the center of the bi-cone is located at the origin. Then, the normalized chroma is calculated as

$$\hat{C} = \frac{L_{max} - |L|}{C} C_{max} \tag{2}$$

where C and L are chroma and lightness of a color respectively, and C_{max} and L_{max} are the maximum value of chroma and lightness respectively. For our implementation, value of HSL components are scaled and translated so that the lightness is between $-100 \sim 100$ and the maximum chroma is 100. As a consequence, colors on the HSL space is projected onto the chromatic plane shown in Fig 1(a). An example

(a) (b) (c)

Fig. 1. (a) Chromatic plane. Cartesian coordinate of this plane is used as a chromaticity component. (b) RGB image. (c) Visualization of chromaticity component after chroma normalization. Shading and speculation effects are removed.

of chroma normalization is shown in Fig 1(c). Shading and specular effects are removed compared to the original image. Since utilizing only hue component cannot fully represent the chromaticity of a color, Cartesian coordinate on the chromatic plane is used as a chromaticity component which will be denoted as (a, b) in this paper.

Similar to the approach of [5], we decouple the illumination component and the color component. For two colors which have (a_1, b_1), (a_2, b_2) as a chromaticity component and L_1, L_2 as a lightness component respectively, we define the decoupled difference between the two colors as follows.

$$\lambda||L_1 - L_2||_k + ||(a_1, b_1) - (a_2, b_2)||_k \tag{3}$$

where $||\mathbf{x}||_k$ denotes l_k norm of a vector $\mathbf{x}$ and λ is a constant. Note that when $k = 1$, (3) is equal to

$$||(\lambda L_1, a_1, b_1) - (\lambda L_2, a_2, b_2)||_1. \tag{4}$$

To make the framework robust to illumination, λ will have a small value less than 1. This weighted distance will be exploited as a core of our optical flow framework which will be explained in the following section.

3. ILLUMINATION ROBUST OPTICAL FLOW

Armed with the decoupled color distance described in Section 2, a novel illumination-robust optical flow framework is proposed in this section. We will describe how the cost of the data term and regularization cost is designed and how it is optimized.

3.1. Cost design

Let $\Omega \subset \mathbb{R}^2$ denote image domain and $\mathbf{u} = (u_1, u_2)^T$ denote the optical flow at $\mathbf{x} \in \Omega$. Then, we propose the cost function of the optical flow to be minimized as

$$\alpha \int_\Omega ||I_1(\mathbf{x} + \mathbf{u}) - I_0(\mathbf{x})||_1 \mathrm{d}\mathbf{x} + \int_\Omega (g_x(\mathbf{x})|\nabla u_1|_\epsilon + g_y(\mathbf{x})|\nabla u_2|_\epsilon)\mathrm{d}\mathbf{x} \tag{5}$$

where $I(\mathbf{x})\colon \Omega \Rightarrow \mathbb{R}^3$ is a function that returns $(\lambda L, a, b)$ of the image at position $\mathbf{x}$ and $\nabla v = \sqrt{(\frac{\partial v}{\partial x})^2 + (\frac{\partial v}{\partial y})^2}$. l_1 norm is used for the data term to achieve robustness to outliers. To avoid staircasing effect in TV-L1 regularization [10], Huber penalty function is used as a regularization term which is defined as

$$|s|_\epsilon = \begin{cases} \frac{s^2}{2\epsilon} & \text{if} |s| \leq \epsilon \\ s - \frac{\epsilon}{2} & \text{otherwise} \end{cases} \tag{6}$$

where ϵ is a small positive constant. $g(\mathbf{x})$ is the weight function which preserves discontinuity on the boundary region following the approaches in [11, 12]. Unlike conventional edge preserving functions, the decoupled distance is used to define g_x, g_y as follows.

$$g_x(\mathbf{x}) = \exp(-h(\mathbf{x})\frac{||\frac{\partial(a,b)}{\partial x}||^2 + \lambda||\frac{\partial L}{\partial x}||^2}{c_g}). \tag{7}$$

The regularizer will have a small weight where there is a huge difference of colors or illumination. Since the parameter λ has a value smaller than 1, the weight will be more sensitive to the difference of color rather than the illumination variation. $h(\mathbf{x})$ is the function to deal with very small chromaticity, which is defined as

$$h(\mathbf{x}) = 1 - \exp(-\frac{(L_{max} - |L(\mathbf{x})|)^2}{c_h}). \tag{8}$$

$L(\mathbf{x})$ is the lightness at $\mathbf{x}$, and c_h is a constant that controls the shape of exponential curve. As very dark or very bright regions have unreliable a, b values, those regions will be regularized even if the difference of color is large. g_y is analogous to g_x except that the derivative is in y-direction.

3.2. Cost optimization

To minimize (5), we first separate the variable $\mathbf{u}$ into $\mathbf{u}$ and $\mathbf{v}$ with the weight parameter θ. Then, the cost to be minimized becomes

$$\alpha \int_\Omega ||I_1(\mathbf{x} + \mathbf{u}) - I_0(\mathbf{x})||_1 \mathrm{d}\mathbf{x} + \int_\Omega \frac{1}{2\theta}(\mathbf{u} - \mathbf{v})^2 \mathrm{d}\mathbf{x} + \int_\Omega (g_x(\mathbf{x})|\nabla v_1|_\epsilon + g_y(\mathbf{x})|\nabla v_2|_\epsilon)\mathrm{d}\mathbf{x}. \tag{9}$$

Optimization of $\mathbf{u}$ and $\mathbf{v}$ is based on the primal-dual algorithm using Legendre-Fenchel transform. Chambolle and Pock [6] proposed fast iterative algorithm for solving the primal-dual problem, which can be directly applied to solve (9). First, $\mathbf{u}$ is updated to minimize the cost with fixed $\mathbf{v}$. Following the conventional optical flow formulation, $I(x)$ is linearized at the current estimate $\mathbf{u_0}$. Then, the problem boils down to solving the following point-wise minimization at every point in Ω.

$$\min_{\mathbf{u}} \quad \alpha||\mathbf{A}\mathbf{u} + \mathbf{b}||_1 + \frac{1}{2\theta}(\mathbf{u} - \mathbf{v})^2 \tag{10}$$

where $\mathbf{A}$ is the 3×2 Jacobian of I_1 at $\mathbf{x} + \mathbf{u_0}$, and $\mathbf{b} = I_1(\mathbf{x} + \mathbf{u_0}) - \mathbf{A}\mathbf{u_0} - I_0(\mathbf{x})$. Similar to [12, 13], (10) can be minimized using Algorithm 2 of [6]. $\mathbf{u}$ can be calculated by the iteration of following primal-dual update.

$$\begin{aligned}
&\mathbf{y}^{(n+1)} = \Pi(\mathbf{y}^{(n)} + \alpha\sigma_n(\mathbf{A}\bar{\mathbf{u}}^{(n)} + \mathbf{b}) \\
&\mathbf{u}^{(n+1)} = \frac{\theta}{\theta+\tau_n}(\frac{\tau_n \mathbf{v}}{\theta} + \mathbf{u}^{(n)} - \tau_n \mathbf{A}^T \mathbf{y}^{(n+1)}) \\
&\rho_n = \frac{1}{\sqrt{1+\frac{2\tau_n}{\theta}}},\ \tau_{(n+1)} = \rho_n \tau_n,\ \sigma(n+1) = \frac{\sigma_n}{\rho_n} \\
&\bar{\mathbf{u}}^{(n+1)} = \mathbf{u}^{(n+1)} + \rho_n(\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)})
\end{aligned} \tag{11}$$

with $\mathbf{u}^{(0)}$ initialized from the previous iteration and $\bar{\mathbf{u}}^{(0)} = \mathbf{u}^{(0)}$, $\mathbf{y}^{(0)} = \mathbf{0}$, $\tau_0 = \sigma_0 = \frac{1}{\alpha B_A}$ where B_A is the maximum norm of $\mathbf{A}$. $\mathbf{y}$ is the dual variable which is a 3-dimensional vector. Π is the function that projects a vector inside the unit ball which is defined as

$$\Pi(\mathbf{s}) = \frac{\mathbf{s}}{\max(1, ||\mathbf{s}||_2)}. \tag{12}$$

Second, $\mathbf{v}$ is updated with fixed $\mathbf{u}$. The minimization problem has the following form:

$$\min_{\mathbf{v}} \int_\Omega (\frac{1}{2\theta}(\mathbf{u}-\mathbf{v})^2 + g_x(\mathbf{x})|\nabla v_1|_\epsilon + g_y(\mathbf{x})|\nabla v_2|_\epsilon)\mathrm{d}\mathbf{x}. \tag{13}$$

The components of a vector $\mathbf{v}$ can be updated individually by applying Algorithm 3 of [6]. The iterative update scheme on the element v_1 becomes

$$\begin{aligned}
&\mathbf{p}^{(n+1)} = g_x(\mathbf{x})\Pi(\frac{\mathbf{p}^{(n)} + \sigma|\nabla \bar{v_1}^{(n)}|}{g_x(\mathbf{x}) + \sigma\epsilon}) \\
&v_1^{(n+1)} = \frac{\theta}{\theta+\tau}(\frac{\tau}{\theta} + v_1^{(n)} - \tau \mathrm{div}(\mathbf{p}^{(n+1)})) \\
&\bar{v_1}^{(n+1)} = 2v_1^{(n+1)} - v_1^{(n)}
\end{aligned} \tag{14}$$

where $v_1^{(0)}$ is initialized from the previous iteration and the dual variable $\mathbf{p}$ is a 2-dimensional vector initialized to $\mathbf{p}^{(0)} = \mathbf{0}$, $\bar{v_1}^{(0)} = v_1^{(0)}$, $\tau = \sqrt{\frac{\epsilon\theta}{8}}$, and $\sigma = \sqrt{\frac{1}{8\epsilon\theta}}$. v_2 can be optimized using the same procedure as (14) where g_x is replaced with g_y.

3.3. Weighted median filtering

In coarse to fine optical flow framework, median filter is often used to remove outliers of flow vectors. Sun et al. [14, 15] revealed the underlying principle of the median filter, and they proposed the weighted median filter to effectively filter outliers near the boundary of objects. Adopting this concept, we design the weighted median filter whose weight is determined by the difference of colors:

$$w_{i,j,i',j'} = h(i,j)\exp(-\frac{||(a,b)-(a',b')||_2^2 + \lambda||L-L'||_2^2}{c_m}) \tag{15}$$

where $I_0(i,j) = (L,a,b)$ and $I_0(i',j') = (L',a',b')$. As in [14], we follow the solution of [16] to solve the weighted median. Filter size is varied from 3×3 to 9×9 in accordance with the size of image.

3.4. Implementation details

The optimization scheme explained in Section 3.2 will be incorporated to coarse to fine framework with additional warping steps. The whole procedure is explained in Algorithm 1. To ensure convergence, 30 inner iteration and 10 outer iterations are performed with 5 warping steps for each scale. Scale factor for image pyramid is set to $\eta = 0.75$. For our experiment, parameters are set to $\lambda = 0.2, \alpha = 10^{-5}, c_m = 10^2, c_g = 10, c_h = 10$, and $\epsilon = 10^{-1}$. The parameter θ is initialized to 0.3 and multiply by $\mu = 0.9$ for every outer loop in order to force convergence.

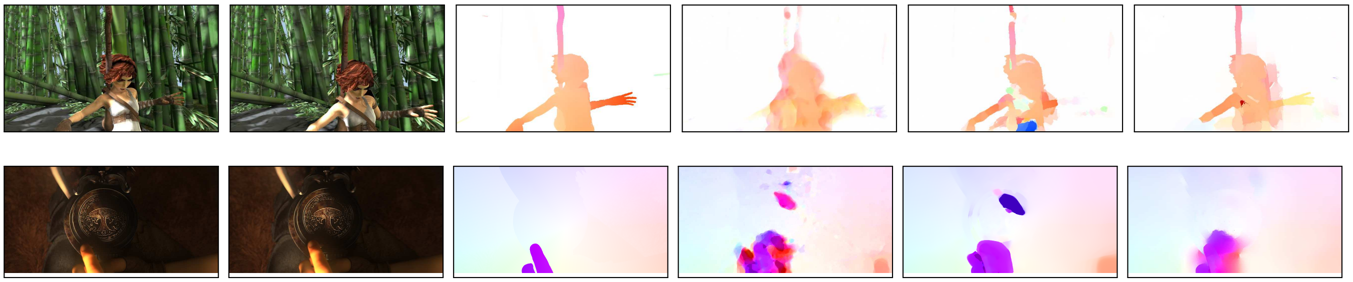
Algorithm 1 Illumination robust optical flow framework

Input: Two images Im_0 and Im_1
Output: Optical flow $\mathbf{u}$ from Im_0 to Im_1

```
Initialize image pyramids, u ← 0, v ← 0
for s = 1 to number of image levels do
  for i = 1 to number of warps do
    u_0 ← u.
    for j = 1 to number of outer loop do
      for k = 1 to number of inner loop do
        Update u via iterating (11).
      end for
      for k = 1 to number of inner loop do
        Update v via iterating (14).
      end for
      θ ← μθ
    end for
    Weighted median filtering on u and v
  end for
  Upsample u, v for next pyramid level
end for
```

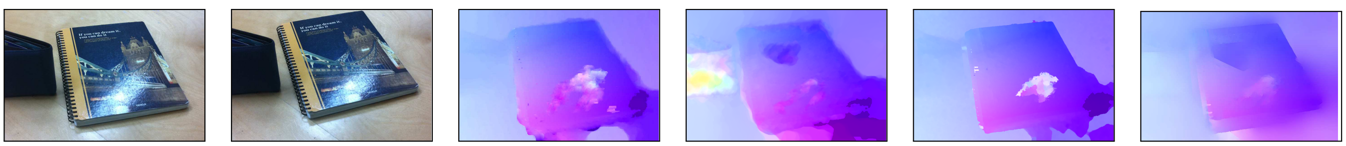
4. EXPERIMENTAL RESULTS

Experiments are performed on the publicly available MPI Sintel dataset [7]. The dataset consists of the training set with ground truth optical flow and the test set. Two sequences (*bamboo_2* and *shaman_3*) of the training set which contains severe illumination variation is selected to measure the performance of our method. Each scene consists of 50 consecutive image frames with two different versions, clean and final. We compared our method with three optical flow algorithms. The first method is TV-L1 optical flow algorithm from [6] which use grayscale image as input (TV-L1 gray). The method involves additional illumination variation handling in its cost function. Second, optical flow that exploits spherical RGB



(a) frame 1 (b) frame 2 (c) Ground truth (d) Spherical RGB (e) MDP-flow (f) Our method

Fig. 2. First row: Qualitative result on *bamboo_2* sequence. Optical flow of the specular region on the clothes and the shadow on the head of Sintel is correctly estimated in our method. Also, the discontinuity of the flow field over the boundary region can be observed. EPE for the frame is (d) 3.336, (e) 3.257, (f) 2.585. Second row: Qualitative result on *shaman_3* sequence. Our method shows robustness to shade variation. EPE for the frame is (d) 1.341, (e) 1.191, (f) 0.950.



(a) frame 1 (b) frame 2 (c) TV-L1 gray (d) Spherical RGB (e) MDP-flow (f) Our method

Fig. 3. Illumination variation on real images. Specular regions are moved in opposite direction to the object. Our method successfully smoothes out the effect of the speculation in the flow field.

		TV-L1 gray	Spherical RGB	MDP flow	Our method
Clean	*bamboo_2*	1.295	1.134	0.883	1.018
	shaman_3	0.266	0.246	0.231	0.273
Final	*bamboo_2*	1.364	1.296	1.006	1.170
	shaman_3	0.581	0.819	0.564	0.743
Overall		0.876	0.874	0.671	0.801

Table 1. EPE on the clean and final dataset of *bamboo_2* and *shaman_3* sequence from the MPI Sintel dataset.

coordinate is implemented (Spherical RGB). In [4], spherical RGB coordinate showed superior result to HSV or normalized RGB in terms of robustness to illumination changes. Unlike original formulation of [4], l_1 distance of data term and Huber norm regularization are used for implementation. Number of iterations for TV-L1 gray and Spherical RGB method is the same as our method. Lastly, MDP-flow [17] is compared as a state-of-the-art optical flow method.

The performance is measured in terms of average endpoint error (EPE). The result is shown in Table 1. Our method gives superior result to TV-L1 gray and Spherical RGB methods. Our method is inferior to the MDP-flow mainly due to the lack of occlusion or large displacement handling. Though the difference of EPE is not significant, our method performs superior to all the other methods in the presence of a severe illumination change. The examples of the results are illustrated in Fig. 2. Color coding scheme in [18] is used to visualize optical flow fields. As one can see in the first row of Fig. 2, our method correctly estimates the optical flow at the specular region of Sintel's clothes and at the shaded region of Sintel's head while MDP-flow shows discontinuity over those regions. Moreover, it can be observed at the hair of Sintel that our method preserves discontinuity on the boundary regions unlike Spherical RGB method. In the second row of 2, it is observed that flow of moving shade is captured in the other methods while our method almost completely smoothes out the effect of the moving shade.

Lastly, we experimented on the real image that contains specular objects. The result is shown in Fig. 3. Speculated region is moved in opposite direction to the direction of moving objects. It can be observed that our method correctly estimate the direction of the flow at the specular region.

5. CONCLUSION

In this paper, an optical flow algorithm that contains robustness to illumination variation is proposed. The proposed method easily decouple illumination and color information using HSL color space. By exploiting the proposed decoupled distance with chroma normalization step, both illumination information and chromaticity information can be used for optical flow estimation while maintaining robustness to illumination variation. The decoupled color distance is used not only for the data term of the cost function but for the weights of regularization term and median filter. The experimental results validate robustness of the proposed optical flow algorithm in the presence of severe speculation or shade variation. Adopting different color difference measure that is similar to the human perception can be taken into consideration as a future work.

6. REFERENCES

[1] B. K. Horn and B. G. Schunck, "Determining optical flow," *ARTIFICAL INTELLIGENCE*, vol. 17, pp. 185–203, 1981.

[2] C. Zach, T. Pock, and H. Bischof, "A duality based approach for realtime tv-l1 optical flow," in *DAGM symposium on Pattern Recognition*, vol. 4713 of *Lecture Notes in Computer Science*, pp. 214–223. 2007.

[3] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers, "An improved algorithm for tv-l 1 optical flow," in *Statistical and Geometrical Approaches to Visual Motion Analysis*, pp. 23–45. Springer, 2009.

[4] Y. Mileva, A. Bruhn, and J. Weickert, "Illumination-robust variational optical flow with photometric invariants," in *DAGM symposium on Pattern Recognition*, pp. 152–162. Springer, 2007.

[5] A. Kumar, F. Tung, A. Wong, and D.A. Clausi, "A decoupled approach to illumination-robust optical flow estimation," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 4136–4147, Oct 2013.

[6] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[7] D. Butler, J. Wulff, G. Stanley, and M.J. Black, "A naturalistic open source movie for optical flow evaluation," in *European Conference on Computer Vision*, vol. 7577 of *Lecture Notes in Computer Science*, pp. 611–625. 2012.

[8] H. Fleyeh, "Color detection and segmentation for road and traffic signs," in *IEEE Conference on Cybernetics and Intelligent Systems*, 2004, vol. 2, pp. 809–814.

[9] D. Grest, J. Frahm, and R. Koch, "A color similarity measure for robust shadow removal in real time," in *Vision, Modeling and Visualization*, 2003, pp. 253–260.

[10] M. Werlberger, W. Trobin, T. Pock, A. Wedel, D. Cremers, and H. Bischof, "Anisotropic huber-l1 optical flow," in *British Machine Vision Conference*, 2009, vol. 1, p. 3.

[11] R. A. Newcombe, S. J. Lovegrove, and A. J. Davison, "Dtam: Dense tracking and mapping in real-time," in *IEEE International Conference on Computer Vision*. IEEE, 2011, pp. 2320–2327.

[12] R. Garg, A. Roussos, and L. Agapito, "A variational approach to video registration with subspace constraints," *International journal of computer vision*, vol. 104, no. 3, pp. 286–314, 2013.

[13] L.L. Rakêt, L. Roholm, M. Nielsen, and F. Lauze, "Tv-l 1 optical flow for vector valued images," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer, 2011, pp. 329–343.

[14] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 2432–2439.

[15] D. Sun, S. Roth, and M. J. Black, "A quantitative analysis of current practices in optical flow estimation and the principles behind them," *International Journal of Computer Vision*, vol. 106, no. 2, pp. 115–137, 2014.

[16] Y. Li and S. Osher, "A new median formula with applications to pde based denoising," *Commun. Math. Sci.*, vol. 7, no. 3, pp. 741–753, 09 2009.

[17] L. Xu, J. Jia, 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.

[18] 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.