FLOW-DRIVEN REGULARIZATION WITH REGULARIZER VERIFICATION IN OPTICAL FLOW COMPUTATION

Tarik Arici[*]

College of Engineering and Natural Sciences,
Department of Electrical Engineering
Istanbul Sehir University, Istanbul, Turkey
tarikarici@sehir.edu.tr

Sait Celebi

Graduate School of Natural and Applied Sciences,
Istanbul Sehir University, Istanbul, Turkey
saitcelebi@std.sehir.edu.tr

ABSTRACT

Inverse problems arise in a variety of fields from computer vision to statistical physics. Optical flow is an inverse problem that uses smoothing/regularization processes to deal with the aperture problem. However, the smoothness assumption is not valid across motion boundaries. There has been extensive research on using partial differential equations (PDEs) to protect optical flow discontinuities across motion boundaries: Image-driven and/or flow-driven anisotropic regularization have been developed with significant improvement over conventional methods. However, optical flow discontinuities inferred using these models may not be correct, hence an incorrect regularization will be performed. We propose flow-driven regularization with regularizer verification using bilateral filtering. We enforce smoothness only when the local continuity of the currently estimated optical flow is supported by the brightness/intensity constancy assumption. Experiment results show that the proposed regularization performs favorably with respect to image-driven or flow-driven methods.

***Index Terms*—** Optical flow, bilateral filtering, flow discontinuity preservation.

1. INTRODUCTION

Most problems in computer vision such as motion/optical-flow estimation, stereo-disparity estimation, video synopsis, image formation modeling, texture segmentation are inverse problems, which are ill-posed [1][2][3] in the Hadamard sense: a small change on the input data may create unpredictable large fluctuations on the solution. To overcome the ill-posedness problem in inverse problems, prior information about the desired solution is incorporated via regularization. Smoothness assumption is a prior information that often helps with ill-posed problems of computer vision.

Optical flow computation algorithms exploit the smoothness information to reduce sensitivity to noise and to deal with the aperture problem. Techniques for optical flow computation can be classified into local and global methods. Lucas and Kanade [4] is a local method which assumes the optical flow is constant within some local neighborhood. On the other hand Horn and Schunck method is a global method which imposes smoothness by including a smoothness term in the global energy functional to be minimized [5]. Local methods are more robust to noise since they integrate the optical flow constraints over pixels in a local neighborhood. This helps with the aperture problem and improves computation of gradients, which are highly sensitive to noise. However, global methods yield 100% dense optical flow by minimizing an energy functional via partial differential equations (PDEs) [6]. Both local and global methods impose smoothness on the optical flow using different approaches [7].Although imposing smoothness helps with the aperture problem, it comes at the cost of smoothing motion boundaries. To prevent smoothing of flow discontinuities the regularizers in the energy functionals has to be modified. The modified regularizers lead to spatially adaptive diffusion processes. Image-driven methods hypothesize a motion boundary at intensity edges by using intensity gradients. On the other hand, flow-driven methods infer a motion boundary using optical flow gradients of the current iteration. These two assumptions fail since an intensity edge may not necessarily map to a motion boundary, and a motion boundary inferred from the current optical flow iterate's gradients may not be correct as the iterations need not converge to the global optimum, but possibly to a bad local minimum. To overcome these two failures of image-driven and flow-driven regularization approaches, we propose a flow-driven regularization in which the smoothness regularizer is verified using the brightness constancy assumption. In our approach neither the optical flow iterate gradients nor the intensity gradients are used to infer discontinuities. Optical flow smoothening is prevented if the optical flow iterate is not supported by the brightness constancy assumption. This is

[*] This research is funded by a grant from European Commission (FP7-PEOPLE-2011-CIG, No:293734) to Tarik Arici.

achieved by checking if a neighboring pixel's flow value violates the brightness constancy (BC) assumption or not. The error in the BC is inputted to the range component of a bilateral filter. As a result, bilateral filter smoothes the flow by preserving flow discontinuities.

2. RELATED WORK AND BACKGROUND INFORMATION

Optical flow algorithms are based on the assumption that intensity of object texture do not change over time, which is called the optical flow constraint. Denoting intensity of a pixel at 2D coordinate (x, y) on an image t by $I(x, y, t)$, the optical flow constraint is given by

$$I(x + u_1, y + u_2, t + 1) = I(x, y, t), \qquad (1)$$

where (u_1, u_2) is the optical flow vector. The first-order Taylor expansion of $I(x + u_1, y + u_2, t + 1)$ around (x, y) yields

$$I_x u_1 + I_y u_2 + I_t = 0 \qquad (2)$$

where subscripts denote partial derivatives. This is called the optical flow constraint (OFC). For each point (x, y) there are two unknowns u_1 and u_2 but one OFC equation. To solve for a unique optical flow additional constraints are required. Since points on the same object move together, smoothness of optical flow is commonly employed via a regularizer term that includes optical flow gradients. Hence, optical flow estimation problem can be formulated as an energy minimization problem in which the energy functional is

$$E(u_1, u_2) = \int_\Omega (I_x u_1 + I_y u_2 + I_t)^2 + \lambda V(\nabla u_1 \nabla u_2)\, dxdy, \qquad (3)$$

where $\nabla = (\partial_x, \partial_y)^T$ is the spatial derivative operator, $V(\nabla u_1, \nabla u_2) = |\nabla u_1|^2 + |\nabla u_2|^2$ is the smoothness term for regularization and λ is a parameter. The optimal solution has to satisfy the Euler-Lagrange equations given by

$$\Delta u_1 - \frac{1}{\lambda} I_x (I_x u_1 + I_y u_2 + I_t) = 0, \qquad (4)$$

$$\Delta u_2 - \frac{1}{\lambda} I_y (I_x u_1 + I_y u_2 + I_t) = 0, \qquad (5)$$

where $\Delta = \partial_{xx} + \partial_{yy}$ denotes the Laplace operator. To solve for the optimal solution by a gradient descent update, a linear heat-diffusion process is obtained given as below

$$\partial_\tau u_1 = \Delta u_1 - \frac{1}{\lambda} I_x (I_x u_1 + I_y u_2 + I_t) = 0, \qquad (6)$$

$$\partial_\tau u_2 = \Delta u_2 - \frac{1}{\lambda} I_y (I_x u_1 + I_y u_2 + I_t) = 0, \qquad (7)$$

The parameter τ corresponds to the time parameter in the diffusion process. Optical flow smoothness is enforced by the underlying linear heat-diffusion process, which can be written separately as

$$\partial_\tau u_i = \Delta u_i = div(\nabla u_i), \qquad (8)$$

for $i = 1, 2$. In the above equation, the Laplacian is expressed as the divergence of the gradient. The solution of (10) for $\tau = 2\sigma^2$ is equal to filtering with a Gaussian function with scale σ. The diffusion over time (*i.e.*, smoothing) will be *isotropic*: in all directions in the same way. This will result in smoothing of flow discontinuities. Furthermore, since the diffusion rate will not change across the image, the flow will be smoothed in a *homogeneous* way throughout the image. The underlying standard heat-diffusion process therefore leads to a homogeneous isotropic smoothing of the optical flow.

Different optical flow computations methods propose to use different regularizers in the energy-functional to preserve motion boundaries. Cues inferred either from the image intensities or the optical flow discontinuities created during the diffusion process are used to modify the regularizer. Hence, optical flow regularizers can be divided into two classes: image-driven or flow-driven regularizers.

2.1. Image-Driven Regularization

These approaches assume that motion boundaries coincide with intensity edges. This assumption yields the regularizer

$$V(\nabla I, \nabla u_1, \nabla u_2) = g(|\nabla I|^2)(|\nabla u_1|^2 + |\nabla u_2|^2), \qquad (9)$$

where g is a decreasing and strictly positive function, which results in reducing the diffusion rate around intensity edges. The resulting diffusion equation is

$$\partial_\tau u_i = div(g(|\nabla I|^2)\nabla u_i). \qquad (10)$$

Depending on the intensity gradient magnitudes, the diffusion rate is changed to prevent diffusion in the vicinity of intensity edges. The diffusion rate changes across the image but diffusion does not depend on the gradient directions. Therefore, the resulting diffusion can be classified as image-driven, inhomogeneous and isotropic. An improvement over the regularization in (9) is to use an anisotropic diffusion that smoothens the flow along the intensity edges but not across [8]. This can be achieved by employing the regularizer

$$V(\nabla I, \nabla u_1, \nabla u_2) = \nabla u_1^T D(\nabla I) \nabla u_1 + \nabla u_2^T D(\nabla I) \nabla u_2, \qquad (11)$$

where $D(\nabla)$ is an underrelaxed projection matrix perpendicular to ∇I with underrelaxation parameter α and is given by

$$D(\nabla) = \frac{1}{|\nabla I|^2 + 2\alpha^2}(\nabla I^\perp \nabla I^{\perp T} + \alpha^2 I) \qquad (12)$$

From (11) and (12), one can see that the regularization term becomes smaller as the optical flow discontinuity (*i.e.*, ∇u_i) becomes perpendicular to ∇I.

Image driven techniques do not smooth motion boundaries when they coincide with intensity edges, however, they may even degrade the overall quality of the optical flow due to the fact that only a small subset of the intensity edges are actual motion boundaries.

2.2. Flow-Driven Regularization

Flow driven regularizers lead to a diffusion process with a diffusion rate that decreases when the flow gradient is large (*i.e.*, in the vicinity of a motion boundary). A typical regularizer is

$$V(\nabla u_1, \nabla u_2) = \Psi(|\nabla u_1|^2 + |\nabla u_2|^2), \tag{13}$$

where $\Psi(s^2)$ is a differentiable and increasing function. An example is

$$\Psi(s^2) = \epsilon s^2 + (1-\epsilon)\lambda^2 \sqrt{1 + \frac{s^2}{\lambda^2}}, \quad 0 < \epsilon \ll 1,\ \lambda > 0 \tag{14}$$

In the above regularizer, the rate of increase in the regularizer is decreasing with increasing $|\nabla u_1|^2 + |\nabla u_2|^2$ due to the use of square root function. This is an isotropic flow-driven regularizer, since the diffusion process is changed regardless of the flow direction. Anisotropic flow-driven regularizers use a diffusion tensor that is adaptive to the flow gradient. Further details on this can be found in [9].

Flow-driven regularization still has the same shortcoming as image-driven approaches: The inferred flow discontinuity, which is used to change the heat-diffusion process may not be correct. The diffusion process is a steepest descent method to optimize the energy-functional in (3), and generally will result in a local optimum. The flow discontinuity is inferred by using the gradient magnitude of the current flow iterate, which might be significantly different than the true flow depending on the quality of the local optimum. This may lead to a degradation of quality in the computed optical flow.

3. REGULARIZER VERIFICATION

Both image-driven and flow-driven regularizers perform poorly when the inferred motion boundary is wrong. To deal with this, we propose to verify the smoothness assumption in a locality using the brightness constancy assumption, which is used to derive the OFC. The error measure for the smoothness assumption between two neighboring pixels i and j are

$$e(i,j) = I(x^i + u_1^j, y^i + u_2^j, t+1) - I(x^i, y^i, t), \tag{15}$$

where (x^i, y^i) is the coordinate of i and (u_0^j, u_1^j) is the flow vector for j. The error measure $e(i,j)$ between i and j measures how well j's flow vector applies to i [1]. Our goal is to modify the Laplacian computation so that second derivatives are computed using one-sided directional derivatives in a way that flows across the motion boundary do not contribute to the Laplacian. Using such a Laplacian, a pixel's flow vector will be enforced to be similar to other flows that apply well to that pixel as determined by (15). To achieve this goal, we will make use of Difference of Gaussians (DOG) as an approximation to Laplacian of Gaussian (LOG) [10], and replace the Gaussian blurring with a bilateral filter. By approximating LOG with a DOG,

$$G_{\tau_1}(|j-i|) * \Delta u_i \approx G_{\tau_2}(|j-i|) * u_i - G_{\tau_1}(|j-i|) * u_i, \tag{16}$$

which is a good approximation when $\tau_2 = 1.6^2 \tau_1$ [10]. $G_\tau(|j-i|)$ denotes a Gaussian with scale parameter τ defining the neighborhood size, and $|j-i|$ is the Euclidean distance between pixels i and j. Since Gaussian is a self-similar function, (16) can be rewritten as

$$\Delta u_i \approx G_{\tau_2 - \tau_1}|j-i| * u_i - u_i, \tag{17}$$

and $\tau_2 - \tau_1$ becomes the time length of diffusion iterations. To achieve an adaptive smoothing as described above, we replace the Gaussian filter by a bilateral filter. The bilateral filter, denoted by $BF[.]$ is defined by

$$BF[u_i] = \frac{1}{W_i} \sum G_{\sigma_s}(|j-i|) G_{\sigma_r}(e(i,j)), \tag{18}$$

where G_{σ_s} is a spatial component and G_{σ_r} is a range component that penalizes flows that have large $e(i,j)$ [11]. Using a 3×3 box function for the spatial component as the standard discrete Laplace operator does, and substituting (18) for the Gaussian in (17)

$$\Delta u_i \approx \frac{1}{W_i} \sum_{j \in S_{3\times 3}} G_{\sigma_r}(e(i,j)) u_j - u_i. \tag{19}$$

Inspecting (19), we can see that the Laplacian is computed using neighboring flows that apply well to a pixel i, which is determined by $e(i,j)$. Hence, the smoothness assumption is verified using the error measure via brightness constancy requirement. Each diffusion iteration by using (19) smoothes a flow vector u_i towards an *average* flow vector $\frac{1}{W_i} \sum_{j \in S_{3\times 3}} G_{\sigma_r}(e(i,j)) u_j$ that takes into account motion boundaries via bilateral filtering.

4. RESULTS

The performance is analyzed using the Middlebury Training set [12]. The performance measures are average angular error (AAE) and average end-point error (EPE). We used the Horn and Schunck implementation of [13], which is an improved version of the classic Horn and Schunck method. It uses a multi-resolution implementation with median filtering of intermediate/coarse optical flow estimates and uses the Rudin-Osher-Fatemi (ROF) structure-texture decomposition method

[1] Note that $e(i,j)$ is not necessarily equal to $e(j,i)$.

	Venus	Dimetrodon	Hydrangea	RubberWhale	Grove2	Grove3	Urban2	Urban3
Horn and Schunck	5.533	4.675	2.218	3.798	2.853	6.809	4.060	7.520
Proposed method	5.262	4.647	2.181	3.699	2.796	6.605	3.900	6.931
Anisotropic image-driven	5.342	4.685	2.209	3.785	2.855	6.818	4.067	7.468
Anisotropic flow-driven	5.613	4.694	2.309	3.831	3.094	6.610	5.739	7.279

Table 1. Average angular error (AAE) on Middlebury training set

	Venus	Dimetrodon	Hydrangea	RubberWhale	Grove2	Grove3	Urban2	Urban3
Horn and Schunck	0.337	0.224	0.190	0.118	0.204	0.690	0.459	0.856
Proposed method	0.327	0.223	0.182	0.115	0.199	0.666	0.436	0.836
Anisotropic image-driven	0.328	0.224	0.189	0.118	0.204	0.692	0.462	0.858
Anisotropic flow-driven	0.381	0.224	0.204	0.121	0.224	0.680	1.138	0.913

Table 2. Average end-point error (EPE) on the Middlebury training set

to pre-pocess input sequences, which gives robustness against lightning and shadow changes. The proposed method outperforms the Horn and Schunck and anisotropic image-driven and flow-driven regularization methods in all the sequences. Surprisingly, image-driven and flow-driven methods perform worse than Horn and Schunck, which we believe is because of the optimized Horn and Schunck implementation used and the failure of both methods to correctly detect motion boundaries.

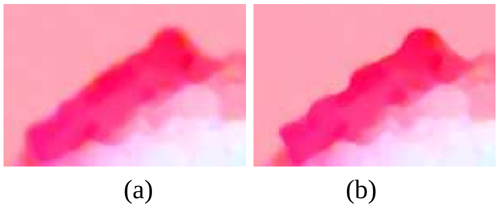
(a) (b)

Fig. 1. A cropped optical flow from the Hydrangea sequence computed by (a) Horn-Schunck, and (b) our proposed method. The motion boundaries using the proposed method are preserved. The color encoding scheme in [12] is used.

5. CONCLUSION AND FUTURE WORK

We have shown that regularizer verification by expilicitly checking for brightness/texture constancy for inferring motion boundaries outperforms image-driven and flow-driven techniques as well as the Horn-Schunck method. Interestingly, energy of the flow computed using our method is larger than that of other methods. This shows that our method minimizes some other energy function. The first author has previously shown that modifying the regularizer used in an ill-defined problem is related to a constrained minimization problem using a primal-dual optimization scheme [14]. As future work, we will explore this relation to design more powerful regularizers tailored for optical flow computation.

6. REFERENCES

[1] Alex Rav-Acha, Pushmeet Kohli, Carsten Rother, and Andrew Fitzgibbon, “Unwrap mosaics: A new representation for video editing,” *ACM Transactions on Graphics (SIGGRAPH 2008)*, August 2008.

[2] A. Rav-Acha, Y. Pritch, and S. Peleg, “Making a long video short: Dynamic video synopsis,” *CVPR*, June 2006.

[3] T. Hofmann, J. Puzicha, and J.M. Buhmann, “Unsupervised segmentation of textured images by pairwise data clustering,” 1996, pp. III: 137–140.

[4] Bruce D. Lucas and Takeo Kanade, “An iterative image registration technique with an application to stereo vision,” 1981, pp. 674–679.

[5] Berthold K. P. Horn and Brian G. Schunck, “Determining optical flow,” *Artifical Intelligence*, vol. 17, pp. 185–203, 1981.

[6] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, “Performance of optical flow techniques,” *International Journal of Computer Vision*, vol. 12, pp. 43–77, 1994.

[7] Andrs Bruhn, Joachim Weickert, and Christoph Schnrr, “Lucas/kanade meets horn/schunck: Combining local and global optic flow methods,” *International Journal of Computer Vision*, vol. 61, pp. 211–231, 2005.

[8] Luis Alvarez, Joachim Weickert, and Javier Sanchez, “Reliable estimation of dense optical flow fields with large displacements,” *International Journal of Computer Vision*, vol. 39, pp. 41–56, 2000.

[9] Joachim Weickert and Christoph Schnörr, “A theoretical framework for convex regularizers in pde-based computation of image motion,” *International Journal of Computer Vision*, vol. 45, no. 3, pp. 245–264, 2001.

[10] Richard Young, “The gaussian derivative model for spatial vision: I. retinal mechanisms,” *‘Spatial Vision 2*, pp. 273–293, December 1987.

[11] S. Paris and F. Durand, “A Fast Approximation of the Bilateral Filter Using a Signal Processing Approach,” *Int J Comput Vis*, vol. 81, no. 1, pp. 24–52, Jan. 2009.

[12] Simon Baker, Daniel Scharstein, J. P. Lewis, Stefan Roth, Michael J. Black, and Richard Szeliski, “A database and evaluation methodology for optical flow,” in *In Proceedings of the IEEE International Conference on Computer Vision*, 2007.

[13] Deqing Sun, Stefan Roth, and Michael J. Black, “Secrets of optical flow estimation and their principles,” in *CVPR*, 2010, pp. 2432–2439.

[14] Tarik Arici and Vural Aksakalli, “Energy-minimization based motion estimation using adaptive smoothness priors,” in *‘Proc. International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*.