21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Exploiting Ramp Structures For Improving Optical Flow Estimation

Abhishek Singh and Narendra Ahuja
University of Illinois at Urbana-Champaign
abhishek_singh@ieee.org, n-ahuja@illinois.edu

Abstract

The underlying principle behind most optical flow algorithms is that the brightness of a pixel remains the same as it flows from one frame to the next. The first order Taylor approximation used in formulating this brightness constancy principle may not be accurate when intensity profiles change non-linearly. In this paper, we propose a method of alleviating the effect of this approximation. Instead of computing image gradients using conventional horizontal and vertical filters of fixed coefficients and sizes, we propose to obtain the gradient information by an explicit examination of ramp profiles at a given location, in all directions. The gradient information obtained using the proposed analysis is more robust under non-linear changes in intensity profiles. Our results demonstrate that by incorporating the ramp structure information as proposed, we are able to improve existing optical flow algorithms.

1. Introduction

New databases and objective evaluation methodologies such as the Middlebury benchmark have spurred recent research interest in optical flow estimation. This is evident by the number of recent additions to the Middlebury rankings [3].

Although several methods for optical flow estimation have been proposed, the central idea common to most of them remains the same. The optical flow problem is posed as the optimization of an energy function that is the weighted sum of two terms - a data term, and a regularization or smoothness term.

The most commonly used data term used in the flow literature is based on the *brightness constancy* principle, which assumes that when a pixel flows from one image to the other, its intensity or color does not change. This principle is formulated by using a first order Taylor series approximation of the image at the particular location. Such a formulation, however, is only accurate if the image has a linearly varying intensity profile near that location. In areas where intensities vary non-linearly, the first order approximation is no longer accurate.

In this paper, we propose a method to alleviate this problem. Instead of using conventional first order gradient operators to formulate the brightness constancy principle, we propose to do a more detailed analysis of the local image structure. Instead of using predefined, fixed-size filters to compute the gradient, we obtain the gradient information by using the *ramp transform* of the image, which was proposed earlier by our group [1]. The ramp transform provides gradient information by an explicit examination of ramp profiles at a given location, in all directions. This reflects the local image structure better than conventional gradient filters. As we show in Section 3, exploiting this additional ramp structure information can help alleviate the inaccuracy of the first order approximation of the brightness constancy principle, in cases where there are non-linear changes in intensity profile.

Our idea is applicable to all existing optical flow methods that use the gradient based formulation of the brightness constancy principle, and can be used to improve over their existing results. We have applied the proposed idea to the classical framework of Black and Anandan [5], as well as a more recent algorithm of Sun et al [6], and have observed consistent improvement in all cases.

2 The Brightness Constancy Principle

The brightness constancy principle states that pixel intensities are translated from one frame to the next.

$$I(\mathbf{x}, t) = I(\mathbf{x} + \mathbf{u}, t + 1). \quad (1)$$

Consider first, the simple case of estimating flow in the 1-D case. Let $f_1(x)$ and $f_2(x)$ be 1-D signals at two time instants ($f_2(x) = f_1(x - u)$). Using a first order Taylor approximation of $f_1(x - u)$ about x, it is easy to

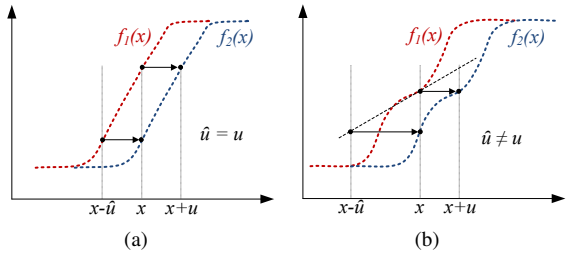

Figure 1. *(a) Computing flow in a 1-D signal, with a linear profile. First order approximation of the brightness constancy principle holds, and the flow can be computed exactly. (b) Estimating flow in case of a non-linear intensity profile. The first order approximation using the conventional gradient can yield incorrect estimates of the flow.*

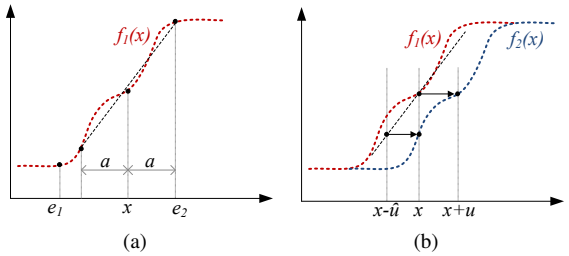

Figure 2. *(a) Slope of the dotted black line is the ramp magnitude of the signal at x, as defined by Eqn. 4. (b) Estimating the flow using the ramp magnitude in place of the true derivative of the signal at x. The estimated flow $\hat{u}$ is more accurate than in Fig. 1(b), since the ramp magnitude is more robust to small non-linearities in the signal profile..*

derive an estimation of the flow magnitude as,

$$\hat{u} = \frac{f_1(x) - f_2(x)}{f_1'(x)}. \tag{2}$$

This can be generalized to the 2-D case to yield the well known gradient constraint equation,

$$\nabla I(\mathbf{x}, t).\mathbf{u} + \frac{\partial I(\mathbf{x}, t)}{\partial t} = 0. \tag{3}$$

Eqn. 2, however, is accurate only if the signals have a linear intensity profile, such as in Fig. 1(a). On the other hand, if the signal has non-linear variation, Eqn. 2 would fail to give an accurate estimation of the flow, as can be seen from the example in Fig. 1(b). This is because the conventional way of computing gradients can become very sensitive to local non-linearities that may be present in the signal profile.

We now describe how the use of the ramp transform can yield a more accurate description of the gradient in areas of such non-linear intensity profiles.

3 Proposed Approach

The ramp transform was proposed in [1], for detection of ramps in images, by analyzing multiple intensity profiles passing through a given pixel in all different directions. It has been successfully used to design a low level, multiscale, hierarchical segmentation algorithm [1], which has been used for higher level visual recognition tasks [2], [7] as well.

A ramp, in the 1-D case, is characterized by its strictly increasing (or decreasing) profile, between the end points e_1 and e_2 as shown in Fig. 2(a). We define the ramp magnitude at a point x to be,

$$R(x) = \frac{|f(x+a) - f(x-a)|}{2a}, \tag{4}$$

where $a = min\{|x - e_1|, |x - e_2|\}$.

$R(x)$ is related to the gradient magnitude of f at x. Conventional gradient magnitudes are computed using fixed size filters (with fixed coefficients). However, in the case of $R(x)$, the 'filter size', a is a function of the ramp end points e_1 and e_2. Therefore, the proposed measure is more adaptive to the structure of the ramp. It is computed using information from the entire ramp profile, and is therefore less sensitive to the local non-linear behavior of the intensity profile within a ramp.

Fig. 2 shows how the ramp magnitude $R(x)$ can be used to yield a potentially better estimate of the flow in the 1-D case.

In the 1-D case, it is easy to compute the end points of the ramp by simply looking at regions where the signal has a strictly increasing or decreasing profile. However, in the 2-D case, infinitely many intensity profiles pass through a given point. To compute the ramp magnitude $R(\mathbf{x})$ for the 2-D case, we follow the approach of [1], as follows:

Let $r(\mathbf{x}, \theta)$ be the ramp magnitude measured at the point x in the image, by analyzing the 1-D ramp along a direction making an angle θ with the horizontal. We compute this ramp magnitude along several different directions θ, between 0 and π.

The ramp magnitude $R(\mathbf{x})$ at location $\mathbf{x}$ is now defined to be,

$$R(\mathbf{x}) = \max_{\theta} r(\mathbf{x}, \theta). \tag{5}$$

The ramp magnitude at $\mathbf{x}$ in the 2-D space, therefore, corresponds to the ramp magnitude along the direction of steepest ascent (or descent).

Using the ramp magnitude and the corresponding direction $\theta^* = \arg\max_\theta r(\mathbf{x}, \theta)$, we can define a ramp based image gradient measure as,

$$\nabla_R I(\mathbf{x}, t) = [R(\mathbf{x}) \cos\theta^*, R(\mathbf{x}) \sin\theta^*]^T \tag{6}$$

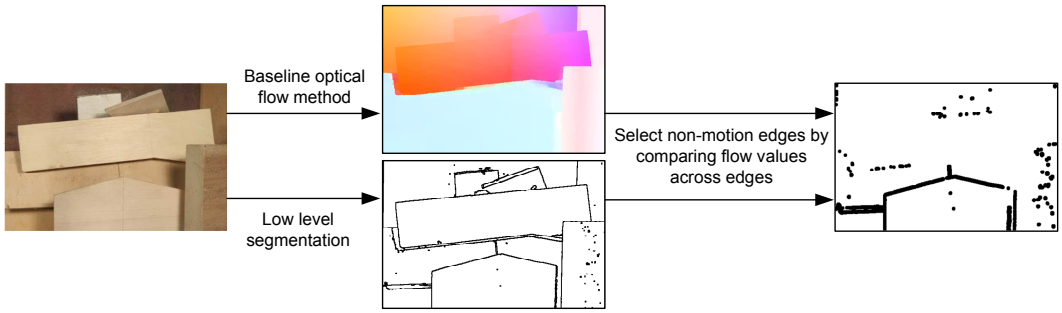

Figure 3. *Steps involved in detecting the non-motion edges of the image. Out of all the edges produced by the low level segmentation algorithm, the non-motion edges are detected by looking for similar flow values across the edges. A dilation operation is performed on these edges to yield the final non-motion edge mask.*

Conventional horizontal and vertical gradient filters face problems while estimating gradients at sharp corners. Since we explicitly scan several values of the direction θ for searching for the largest ramp magnitude, we avoid such problems.

We propose to use the quantity in Eqn. 6 as the measure of the image gradient in the gradient constraint equation of Eqn. 3 in order to compute optical flow.

The brightness constancy relation is not valid along motion discontinuities. We therefore choose to use the ramp gradient of Eqn. 6 only in the ramps that are along the non-motion boundaries of the image. The non-motion boundaries are determined as follows: We first use a baseline optical flow method to obtain the flow field between the image pair. We then use the low level segmentation algorithm of [1] on the first image to obtain *all* the edges present in the image. Some of these edges (ramps) correspond to motion boundaries, whereas others do not. To determine the edges corresponding to non-motion boundaries, we compare the flow values across the edges. If the difference of the flow values on either side of the edge is smaller than a threshold, the edge is labeled as a non-motion edge. The non-motion edge map thus obtained is then dilated using a 5×5 mask. This whole process is summarized in Fig. 3.

4 Experiments and Results

In all our evaluations, we first run a baseline flow algorithm and obtain a flow field, which is used for obtaining non-motion edges of the image as described earlier. We then re-run the algorithm with the same parameterization (as suggested by authors of the algorithms), but now using the proposed ramp gradients in place of conventional gradients, along the non-motion edges. We compare the flow fields obtained after the two runs, to see the improvement the proposed technique can bring over the baseline algorithm.

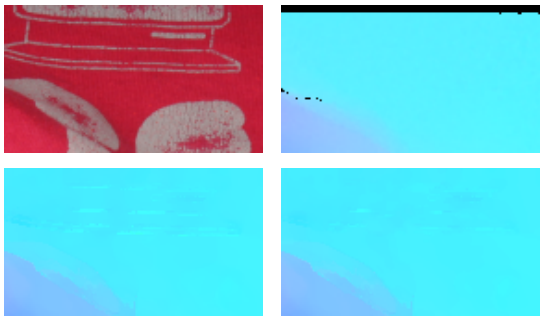

Figure 4. *Top Left: Patch from the 'Dimetrodon' sequence of the Middlebury set. Top Right: Ground truth flow field. Bottom Left: Result obtained using the Classic+NL-fast algorithm [6]. Bottom Right: Result obtained after incorporating ramp information.*

We test the proposed idea on multiple baseline algorithms. We first test on the *Classic++* algorithm, which is essentially the classical formulation of Black and Anandan [5], but implemented using modern numerical methods, as described in [6]. The second baseline used is the *Classic+NL* algorithm of Sun et al [6], which essentially performs medial filtering of intermediate flow fields, by adding an extra non-local term in the classical cost function. This method is among the best performing algorithms in the current list of published methods in the Middlebury benchmark.

Table 1 shows the results of applying the proposed technique on the *Classic++* algorithm, on the Middlebury training set (with publicly available groundtruth). We report performance in terms of Average Angular Error [4]. We see improvements in most cases.

Table 2 shows the results of applying the proposed method on the *Classic+NL-fast* algorithm, which is

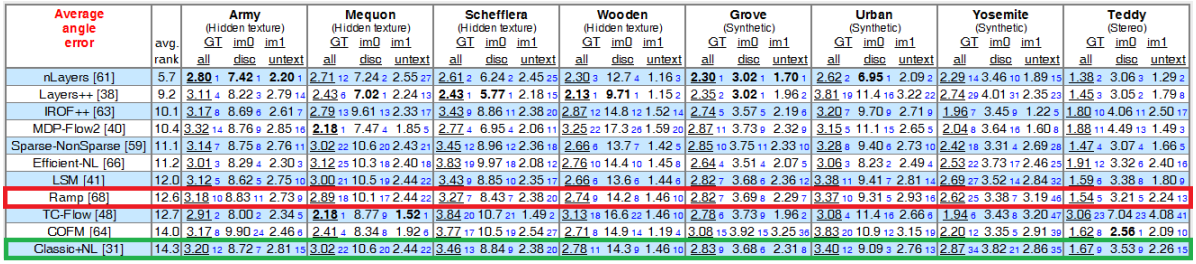

Figure 5. *Results of using the proposed approach with the Classic+NL algorithm [6], on the Middlebury test images. Our results are labeled as 'Ramp[68]' and are shown in the red box. We are able to consistently improve over the Classic+NL algorithm in the green box. The results can be viewed online at http://vision.middlebury.edu/flow/eval/results/results-a1.php*

Table 1. *Average Angular Error of Classic++ algorithm [6] and our approach, on seven Middlebury image pairs.*

Image Pair →	*Dimetrodon*	*Hydrangea*	*RubberWhale*	*Grove2*	*Grove3*	*Urban2*	*Urban3*
Classic++	2.532	1.786	2.646	2.068	6.055	2.541	4.598
Ours	2.527	1.784	2.630	2.066	6.032	2.528	4.594

Table 2. *Average Angular Error of Classic+NL-fast algorithm [6] and our approach, on seven Middlebury image pairs.*

Image Pair →	*Dimetrodon*	*Hydrangea*	*RubberWhale*	*Grove2*	*Grove3*	*Urban2*	*Urban3*
Classic+NL-fast	2.280	1.824	2.401	1.410	4.927	2.034	3.160
Ours	2.267	1.814	2.388	1.394	4.922	2.018	2.983

a fast implementation of the *Classic+NL* algorithm. Again, we see an improvement in all cases. Fig. 4 shows a qualitative comparison on an image patch from the 'Dimetrodon' pair of the Middlebury set.

We also test our approach on the Middlebury *test* images (with hidden ground truths), with the *Classic+NL* algorithm. The results from our algorithm are publicly available on the Middlebury evaluation page. Fig. 5 shows a screenshot of the evaluation. The proposed method (called 'Ramp[68]' in the figure) utilizing the ramp information improves over the performance of the *Classic+NL* baseline.

5 Conclusion

We have presented a way of alleviating the first order approximation problem in the formulation of brightness constancy principle, in case of non-linear variation of image intensities. We have shown that the performance of existing algorithms can be improved by incorporating the proposed ramp information for estimating the gradient. Although the gain in performance is small, the improvement is consistently seen across different algorithms and image pairs. We see this as an important cue for further investigation and research. It is reasonable to expect gains in performance if the proposed idea is incorporated in other flow algorithms as well, which use the gradient constraint equation.

6 Acknowledgments

The authors thank Deqing Sun for his optical flow code [6] (available online), and Emre Akbas for providing the segmentation algorithm [1]. This work was supported by ONR grant N00014-12-1-0259. Abhishek Singh was also supported by the Joan and Lalit Bahl Fellowship from the Dept. of ECE, UIUC.

References

[1] E. Akbas and N. Ahuja. From ramp discontinuities to segmentation tree. In *Proceedings of ACCV*, 2009.
[2] E. Akbas and N. Ahuja. Low-level image segmentation based scene classification. In *IEEE ICPR*, 2010.
[3] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *IJCV*, 92(1):1–31, 2011.
[4] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, Feb. 1994.
[5] M. J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *CVIU*, 63(1):75 – 104, 1996.
[6] D. Sun, S. Roth, and M. Black. Secrets of optical flow estimation and their principles. In *IEEE CVPR*, 2010.
[7] S. Todorovic and N. Ahuja. Region-based hierarchical image matching. *IJCV*, 78:47–66, June 2008.