A non-regular optical flow for dynamic textures

Sándor Fazekas and Dmitry Chetverikov

Computer and Automation Research Institute of the
Hungarian Academy of Sciences (MTA SZTAKI, Budapest)
fazekas@phy.bme.hu, csetverikov@sztaki.hu

Abstract. We propose a non-regular optical flow based on *brightness conservation assumption* allowing brightness to propagate to its neighborhood, and thus to describe brightness changes which cannot be described based on *brightness constancy assumption* – the starting point of regular optical flow calculation methods. Brightness constancy holds well for motion of rigid objects with Lambertian surfaces, but it is less appropriate for complex physical processes typical for dynamic textures. Our experimental results demonstrate that the proposed non-regular optical flow is more suitable in this latter case.

1 Introduction

Dynamic textures form basic building blocks of motion detection and recognition systems as well as data mining tools of multimedia databases. A dynamic texture is a spatially repetitive, time-varying visual pattern possessing certain temporal stationarity [20, 23, 10]. Typical dynamic textures are recordings of physical processes such as surface waves, fire and smoke, flag blown by the wind or the collective motion of different distinct elements such as a walking crowd or a flock of birds. The study of dynamic textures aims to extend texture analysis from spatial domain to the temporal domain, treating motion as integral part of the visual pattern. Detecting, segmenting, recognizing, and classifying dynamic textures can rely on visual aspects such as geometry, motion, or both.

Dynamic textures are common in natural scenes, however, in many cases only parts of the scene form a dynamic texture. Furthermore, their spatial location and extent can vary and they can be also partially transparent. All these make it difficult to separate them from a textured background. Due to these problems the geometric information can be misleading. To be able to handle dynamic textures, differences in dynamics must be analyzed. Many current methods [8] are based on optical flow calculation estimating the velocity field which describes the motion of image regions.

The widely accepted definition of dynamic textures based on the concept of spatiotemporal self-similarity covers a broad range of motion patterns. Dynamic textures can be grouped in two main categories: *weak* and *strong* dynamic textures, according to general characteristics of the underlying dynamics.

A *weak dynamic texture* such as a simple moving texture (or a combination of distinct moving textures) does not posses any intrinsic (or hidden) dynamics.

Following its parts in appropriate local moving coordinate systems, i.e. "going with the flow", a weak dynamic texture becomes static. Such local coordinate systems are computed using standard optical flow algorithms relying on the *brightness constancy assumption* [17, 19].

Characteristics of *strong dynamic textures* possessing an intrinsic dynamics cannot be fully captured in the above approach. Self-occlusion, material diffusion, and other physical processes not obeying the brightness constancy assumption make regular optical flows less applicable in this case.

Here, we present an alternative to the brightness constancy assumption, which we call the *brightness conservation assumption*. Under this assumption the brightness of an image point (in one frame) can propagate to its neighborhood (in the next frame). While a static or weak dynamic texture obeys the brightness constancy assumption – as we are going to show – a strong dynamic texture is better modeled by the brightness conservation assumption. Based on this assumption we propose an algorithm for optical flow calculation.

We test our method on videos taken from the DynTex database [21]. The experimental results demonstrate the better applicability of the proposed non-regular optical flow to the study of dynamic textures.

2 Background

2.1 Optical flow

The optical flow concept arose from the studies of the human visual system [14] and aims to model low-level motion perception. Aside from algorithmic complexity and numerical stability problems, a major difficulty in calculating an optical flow is caused by the *aperture problem* [24]. This affects both computer vision algorithms [16] and human motion perception [15]. Due to the aperture problem, only the velocity component orthogonal to edges and contours – the so called *normal flow* – can be defined without ambiguity, unless the motion estimation is extended to larger regions [17, 19].

The standard way to calculate an optical flow is based on the *brightness constancy assumption*

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

where I is the image function (i.e. $I(x, y, t)$ gives the brightness in point (x, y) at time t) and (u, v) is the flow. The above equation states that the brightness of a small image region is constant (from frame-to-frame) along the flow.

Assuming that I depends smoothly on x, y, and t, the first-order Taylor approximation of the above brightness constancy assumption leads to the *optical flow constraint* equation

$$I_t + uI_x + vI_y = 0, \tag{2}$$

where I_x, I_y, and I_t are the partial derivatives of I. If (u, v) satisfies the optical flow constraint, $(u + rI_y, v - rI_x)$ does also satisfy it for any random function

r. In order to overcome this strong ambiguity – which is a manifestation of the aperture problem – additional *smoothness constraints* are needed.

Horn and Schunck [17] formulate a variational problem with the Lagrangian

$$L_{HS}(u,v) = (I_t + uI_x + vI_y)^2 + \alpha(u_x^2 + u_y^2 + v_x^2 + v_y^2), \qquad (3)$$

where the first term corresponds to the optical flow constraint and the second term – controlled by the parameter α – is a smoothness constraint with u_x, u_y, v_x, and v_y being the flow derivatives. The optical flow (u,v) is calculated based on calculus of variations [16]. The accuracy of the method can be enhanced by using a coarse-to-fine scheme [2, 5].

Many similar optical flow calculation methods based on calculus of variations exist (e.g. [6, 1]). There are also other techniques such as structure tensor methods (see the Lucas-Kanade method [19] and its variants), frequency and phase-based methods, and methods based on region-matching [3]. All these, in explicit or implicit form, rely on the brightness constancy assumption.

2.2 Dynamic textures

Several examples can be taken from biological systems to emphasize the importance of motion in visual sensing. Studies of visual perception [7] revealed that humans use motion directly in recognizing aspects of their environment. Insects are essentially blind to anything that is standing still and the camouflage strategies of some animals are effective only as long as they are not moving. The observation that motion adds relevant information to visual patterns, explains the current effort in computer vision research to extend the already classical field of texture analysis from the spatial domain to the temporal domain.

Currently the most popular methods used to analyze dynamic textures are based on optical flow calculation (for a recent review see [8]). An alternate approach computes geometric properties in the spatiotemporal domain [26], while other methods are based on spatiotemporal filtering [25] and spatiotemporal transforms [22]. There are also model-based methods [23, 11] which use estimated model parameters as features.

The methods based on optical flow, in a computationally efficient and natural way, characterize the local changes in dynamic textures with features extracted from an estimated velocity field describing the motion of small image regions. In this approach a dynamic texture can be viewed as a sequence of instantaneous motion patterns. When necessary, geometric and color information can also be added to form a complete set of features for both motion and appearance based detection, segmentation, and recognition.

Most of the work done so far in studying dynamic textures used the normal flow, partly as an influence of the successful pioneering work of Nelson and Polana [20] and partly because the calculation of the normal flow is easy and fast. However, the close relation of normal flow and spatial gradients – and hence contours and shapes – implies that the normal flow correlates with appearance features and thus it does not characterize the "pure dynamics". Even though this was recognized already at the early state of these studies [20], no solution was proposed. Later, to overcome this problem, Fablet and Bouthemy [12] used only the magnitude of the normal flow and recently Lu et al. [18] as well as Fazekas and Chetverikov [13] stressed the necessity to apply a complete flow calculation in extracting characteristics of dynamic textures.

3 Brightness Conservation Assumption

The brightness constancy assumption, as defined in Eq. 1, states that for two consecutive frames in a video sequence the brightness of a point (x, y) in a frame is identical to the brightness of a point $(x+u, y+v)$ in the next frame. In other words, it states that by warping the image space according to a displacement field (u, v) the images can be brought into point-by-point equality.

There are situations – typical for dynamic textures – in which the above assumption is not applicable. Such situations appear in case of occlusions or when the captured scene includes glinting surfaces such as a water surface or physical processes such as fire and smoke. In these situations, equality of consecutive frames cannot be achieved by warping the image space because there are changes in brightness which violate the brightness constancy assumption.

Below we propose a solution for this problem by defining an optical flow which can model not only space warps but also changes in brightness. A straightforward approach is to consider that the flow "carries" the brightness as a "physical quantity" and the changes in brightness are encoded in the divergence of the flow. A similar technique was developed to detect motion of gaseous materials (e.g. smoke or vapor) by Béréziat et al. [4] and Cuzol et al. [9].

Considering an arbitrary region Ω of an image I, brightness conservation can be defined as the equilibrium of the total brightness change on Ω and the brightness carried in and out through its boundary. With notations usual in physics, this is

$$\int_{\Omega} \partial_t I \, \mathrm{d}A + \int_{\partial\Omega} I \, \boldsymbol{f} \cdot \boldsymbol{n} \, \mathrm{d}L = 0, \tag{4}$$

where $\partial_t I$ is the time derivative of I, $\boldsymbol{f} = (u, v)$ is the flow, $\partial\Omega$ denotes the boundary of Ω, and $\boldsymbol{n}$ is the external normal of $\partial\Omega$.

Through mathematical transformations and the divergence theorem one can derive the *continuity equation*

$$\partial_t I + \nabla I \cdot \boldsymbol{f} + I \, \mathrm{div}(\boldsymbol{f}) = 0. \tag{5}$$

Using the notations of Eq. 2, this is equivalent with

$$I_t + uI_x + vI_y + I(u_x + v_y) = 0, \tag{6}$$

where u_x and v_y are partial derivatives of the flow.

The above equation is the first order Taylor approximation of

$$I(x+u, y+v, t+1) = I(x, y, t)(1 - u_x - v_y), \tag{7}$$

which we call the *brightness conservation assumption*. This should be compared to the brightness constancy assumption defined in Eq. 1.

A flow (u, v) satisfying Eq. 7 does not only describe a warp of the image space but can also balance brightness changes through $(u_x + v_y)$. Consequently, such a flow can capture more dynamic information than a regular optical flow, and thus it is more suitable for the study of dynamic textures.

4 Optical flow based on brightness conservation

Similar to the derivation of the Horn-Schunck Lagrangian (see Eq. 3), we formulate a Lagrangian based on the brightness conservation assumption and the corresponding continuity equation defined above. Using this in a variational scheme, a non-regular optical flow can be calculated.

If (u, v) satisfies Eq. 6, then the flow $(u + rI_y, v - rI_x)$ does also satisfy it for any $r = \xi(I)$, where ξ is an arbitrary function. This ambiguity resembles the aperture problem and can be handled by imposing smoothness and consistency constraints. In this way we obtain the Lagrangian

$$L(u,v) = (I_t + uI_x + vI_y + I(u_x + v_y))^2 + \alpha(u_x^2 + u_y^2 + v_x^2 + v_y^2) + \beta(u^2 + v^2). \quad (8)$$

Here the first term assures brightness conservation, the second term imposes a certain level of smoothness, while the third term makes sure that (u, v) is consistent with the assumed first order Taylor approximation.

The optical flow is calculated by minimizing

$$F(u,v) = \int_I L(u,v)\,\mathrm{d}x\mathrm{d}y \quad (9)$$

with calculus of variations [16]. This is done by solving numerically the corresponding Euler-Lagrange equations.

Introducing the notation $R(u, v) = I_t + uI_x + vI_y + I(u_x + v_y)$, the Euler-Lagrange equations for u and v are

$$\beta u - IR_x(u,v) - \alpha(u_{xx} + u_{yy}) = 0, \quad (10)$$

$$\beta v - IR_y(u,v) - \alpha(v_{xx} + v_{yy}) = 0. \quad (11)$$

Here $R_x(u, v)$ and $R_y(u, v)$ denote the partial derivatives of $R(u, v)$. The above equations can be discretized by assuming central derivatives of u and v.

Similar to the method described in [17], Eq. 10 and 11 are solved by solving iteratively the matrix equation

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} D \\ E \end{pmatrix}, \quad (12)$$

where

$$A = I(-I_{xx} + 2I) + 4\alpha + \beta, \quad (13)$$

$$C = -II_{xy}, \quad (14)$$

$$D = I(I_{tx} + I_x(2u_x + v_y) + I_y v_x + I(\bar{u}^x + v_{xy})) + \alpha\bar{u}. \quad (15)$$

Here, we used the notation $\bar{u} = \bar{u}^x + \bar{u}^y$ and

$$\bar{u}^x(x, y) = u(x-1, y) + u(x+1, y), \quad (16)$$

$$\bar{u}^y(x, y) = u(x, y-1) + u(x, y+1). \quad (17)$$

The definition of $\bar{v}^x$ and $\bar{v}^y$ is analogous. The terms B and E can be derived from A and D using the simultaneous substitutions $u \leftrightarrow v$ and $x \leftrightarrow y$.

Because, in the above equations, a first order Taylor approximation was assumed, u and v need to be small. This is achieved by using a coarse-to-fine scheme [2, 5]: The images are warped according to the flow calculated at a coarse scale and small corrections are added repeatedly at finer scales.

5 Experimental results

We demonstrate the accuracy of the proposed non-regular flow on two video sequences taken from the DynTex database [21] showing a flapping flag (`6483c10`) and a small waterfall (`6481h20`). We compare the non-regular flow with the regular Horn-Schunck flow. The input images were pre-blurred with a Gaussian filter having $\sigma = 0.4$ and processed channel-by-channel in the YCrCb color space (scaled to $[0, 1]$). Coarse-to-fine iterations were run over a Gaussian pyramid having 20 levels and scale factor 0.9. On each level 20 iterations were executed. Both the non-regular flow and the Horn-Schunck flow was calculated with $\alpha = 0.01$. The consistency parameter of the non-regular flow was $\beta = 0.1$.

The results are presented in Fig. 1. The flows are calculated between the first two frames of the video sequences. The second frame was warped backward with

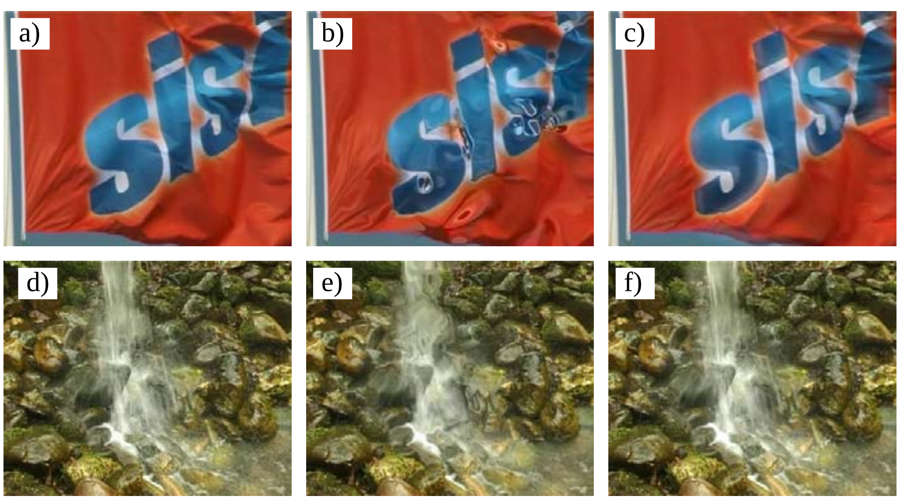

Fig. 1. Experimental results: (a, d) Original frames; (b, e) Motion compensation with Horn-Schunck flow; (c, f) Motion compensation with non-regular flow.

the calculated flow in order to be compared with the first frame. The non-regular flow could reproduce the original image almost perfectly, while the Horn-Schunck flow produced visible errors, which do not originate from numerical instabilities but they are the consequence of the fact that a regular flow cannot equilibrate certain changes in brightness.

6 Conclusions and discussion

An alternative to the *brightness constancy assumption*, namely the *brightness conservation assumption* was presented and based on this an algorithm for calculating a non-regular optical flow was proposed. The complexity of this algorithm is the same as the complexity of the Horn-Schunck method [17], however, it requires more floating-point calculations per iterations, and thus it possesses a larger pre-factor.

We have demonstrated that a strong dynamic texture is better modeled with the brightness conservation assumption. Brightness conservation is less constrained than brightness constancy, consequently a non-regular optical flow is expected to be always more accurate in capturing dynamic information than a regular flow, and thus to give an effective and theoretically well-founded tool for studying dynamic textures.

The information encoded in the presented non-regular flow can be used to extract features of dynamic textures (similar to [13]), which should give a more precise characterization of the dynamics than features extracted from a regular flow. Extensions of this work should investigate this quantitatively.

Acknowledgments: This work was supported by MUSCLE (FP6-507752).

References

1. T. Amiaz and N. Kiryati. Piecewise-smooth dense optical flow via level sets. *Int. J. Comp. Vision*, 68(2):111–124, 2006.
2. P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *Int. J. Comp. Vision*, 2(3):283–310, January 1989.
3. J. L. Barron, D. J. Fleet, and S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12:43–77, 1994.
4. D. Béréziat, I. Herlin, and L. Younes. A generalized optical flow constraint and its physical interpretation. In *Proc. Conf. Comp. Vision Pattern Rec.*, pages 487–492, 2000.
5. Michael J. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Comp. Vision and Image Understanding*, 63(1):75–104, 1996.
6. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and J. Matas, editors, *European Conference on Computer Vision*, volume 3024 of *Lecture Notes in Computer Science*, pages 25–36. Springer - Berlin, 2004.

7. Vicki Bruce, Patrick R. Green, and Mark Georgeson. *Visual Perception*. Psychology Press (UK), 1996.
8. D. Chetverikov and R. Péteri. A brief survey of dynamic texture description and recognition. In *4th Int. Conf. on Computer Recognition Systems*, pages 17–26, 2005.
9. A. Cuzol and E. Mémin. Vortex and source particles for fluid motion estimation. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Lecture Notes in Computer Science*, volume 3459, pages 254–266. 2005.
10. G. Doretto, A. Chiuso, S. Soatto, and Y. N. Wu. Dynamic textures. *Int. J. Comp. Vision*, 51:91–109, 2003.
11. G. Doretto, E. Jones, and S. Soatto. Spatially homogeneous dynamic textures. In *Proc. European Conf. on Computer Vision*, volume 2, pages 591–602, 2004.
12. R. Fablet and P. Bouthemy. Motion recognition using nonparametric image motion models estimated from temporal and multiscale co-occurrence statistics. *IEEE Trans. Pattern Analysis and Machine Intell.*, 25:1619–1624, 2003.
13. S. Fazekas and D. Chetverikov. Normal versus complete flow in dynamic texture recognition: A comparative study. In *Int. Workshop on Texture Analysis and Synthesis*, pages 37–42, 2005.
14. J. Gibson. *The Senses Considered as Perceptual Systems*. Houghton Mifflin - Boston, 1966.
15. E. C. Hildreth. The analysis of visual motion: From computational theory to neural mechanisms. *Annual Review of Neuroscience*, 10:477–533, 1987.
16. B. K. P. Horn. *Robot Vision*. McGraw-Hill - New York, 1986.
17. B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203, August 1981.
18. Z. Lu, W. Xie, J. Pei, and J. Huang. Dynamic texture recognition by spatiotemporal multiresolution histograms. In *Proc. of the IEEE Workshop on Motion and Video Computing (WACV/MOTION)*, pages 241–246, 2005.
19. B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. of the 7th International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.
20. R. C. Nelson and R. Polana. Qualitative recognition of motion using temporal texture. *CVGIP: Image Understanding*, 56:78–89, 1992.
21. R. Péteri, M. Huskies, and S. Fazekas. DynTex: A comprehensive database of Dynamic Textures. `www.cwi.nl/projects/dyntex/`, 2006.
22. J.R. Smith, C.-Y. Lin, and M. Naphade. Video texture indexing using spatiotemporal wavelets. In *Proc. Int. Conf. on Image Processing*, volume 2, pages 437–440, 2002.
23. M. Szummer and R.W. Picard. Temporal texture modeling. In *Proc. Int. Conf. Image Processing*, volume 3, pages 823–826, 1996.
24. D. Todorovic. A gem from the past: Pleikart Stumpf's anticipation of the aperture problem, Reichardt detectors, and perceived motion loss at equiluminance. *Perception*, 25:1235–1242, 1996.
25. R. P. Wildes and J. R. Bergen. Qualitative spatiotemporal analysis using an oriented energy representation. In *Proc. European Conf. on Computer Vision*, pages 768–784, 2000.
26. J. Zhong and S. Scarlaroff. Temporal texture recognition model using 3D features. Technical report, MIT Media Lab Perceptual Computing, 2002.