Pattern Recognition 44 (2011) 2057–2062

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

# Optic flow based on multi-scale anchor point movement and discontinuity-preserving regularization

P.A.G. van Dorst *, B.J. Janssen, L.M.J. Florack, B.M. ter Haar Romeny

*Eindhoven University of Technology, Den Dolech 2, Postbus 513, 5600 MB Eindhoven, The Netherlands*

ARTICLE INFO

Available online 16 September 2010

*Keywords:*
Image processing
Optic flow
Multi-scale
Scale-space
Toppoints
Discontinuity-preserving regularization

ABSTRACT

We introduce a new method to determine the flow field of an image sequence using multi-scale anchor points. These anchor points manifest themselves in the scale-space representation of an image. The novelty of our method lies largely in the fact that the relation between the scale-space anchor points and the flow field is formulated in terms of soft constraints in a variational method. This leads to an algorithm for the computation of the flow field that differs fundamentally from previously proposed ones based on hard constraints. We show a significant performance increase when our method is applied to the Yosemite image sequence, a standard and well-established benchmark sequence in optic flow research. Also, it is shown that this performance is not sensitive to slight changes in the two parameters used and that, with the same parameter values, our method yields very good results in the Rubber Whale image sequence as well.



## 1. Introduction

Optic flow describes the apparent motion in an image sequence. A variety of approaches exists to estimate this motion. Survey papers include those by Barron et al. [1] and Mitchie et al. [2].

Differential methods are based on the most widespread approach, which uses spatiotemporal derivatives to describe the local image structure. The flow field is assumed to connect points in subsequent frames of the image sequence with similar structure. For example, in one of the earliest methods, proposed by Horn and Schunck [3], this "structure" is the image intensity, which leads to the well-known optic flow constraint equation (henceforth abbreviated OFCE). An overview of current developments in differential methods can be found in Bruhn et al. [4]. A problem that is encountered by these methods is that the structure does not always remain constant over time. For example, the global image intensity may vary over time. More complex terms to describe the structure can be used to overcome this problem [5,6]. A second problem is that many possible solutions exist, since points on level-sets have the same image intensity. This requires a so-called prior, which determines a unique solution based on prior knowledge. A prior usually is a regularization term, which can for example prefer an overall smooth solution with sparse discontinuities [7–9].

Another well performing approach is that of region matching, in which the image is split up into small blocks, each of which is translated to match the image neighborhood [10]. Because of their low computational cost, these methods are widely used in applications such as temporal up-scaling of video signals and video compression.

Our method can be placed in the category of feature-tracking methods. An overview of such methods can be found in [11]. However, in contrast to most feature-tracking algorithms, the features we use do not correspond to specific points in the image sequence. Instead, we use anchor points that exist at different scales in scale-space, called toppoints (properly defined in Section 2.1). Therefore, instead of corresponding to points, the features we track actually represent entire regions in the image sequence. Using toppoints to extract the motion from an image sequence has been first proposed by Janssen et al. [12] and Florack et al. [13]. In these papers, the relation between the toppoint velocity and the flow field was implemented using a hard constraint, which means that this constraint has to be fulfilled exactly. The advantage is that their method is entirely parameter free, but the price of this is sensitivity to outliers. In the method presented in this paper, a 1-parameter soft constraint is used, yielding higher robustness against errors in the estimated toppoint velocity or deviations from the stipulated relation between toppoint velocity and the flow field.

Toppoints are found throughout the scale-space of each frame of the image sequence as isolated entities. Therefore they are truly multi-scale, in contrast to other multi-scale features which are found by applying scale-selection to points that exist at every scale, such as corners. Another use of toppoints is to reconstruct

* Corresponding author.
*E-mail addresses:* P.A.G.v.Dorst@tue.nl (P.A.G. van Dorst), B.J.Janssen@tue.nl (B.J. Janssen), L.M.J.Florack@tue.nl (L.M.J. Florack), B.M.terHaarRomeny@tue.nl (B.M. ter Haar Romeny).

0031-3203/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2010.09.003

an image from the values of derivatives taken at toppoint positions, cf. the papers by Lillholm [14], Nielsen [15] and Janssen [16]. In these papers it is shown that features at the toppoint positions can be used to efficiently represent the information contained in an image. An important property is that the amount of toppoints found in a certain area of the image is proportional to the amount of information in that area.

## 2. Theory

In this chapter we will first explain how the scale-space representation of an image is defined and what toppoints are. Also important properties of toppoints are mentioned and we try to give toppoints a more intuitive meaning with some visualizations. Next we explain how to calculate toppoint velocities, and the method used to obtain the actual flow field from the toppoint velocities.

### 2.1. Scale-space and toppoints

The scale-space representation $f_s(x,y)=f(x,y;s)$, where $f\in\mathbb{C}^\infty(\mathbb{R}^2\times\mathbb{R}^+)\cap\mathbb{L}_2(\mathbb{R}^2\times\mathbb{R}^+)$, of a static scalar image $f_0\in\mathbb{L}_2(\mathbb{R}^2)$ is defined by the convolution of the image with a Gaussian kernel $\phi_s(x,y)=\phi(x,y;s)$, where $\phi\in\mathbb{C}^\infty(\mathbb{R}^2\times\mathbb{R}^+)\cap\mathbb{L}_2(\mathbb{R}^2\times\mathbb{R}^+)$ and $s\in\mathbb{R}^+$ denotes the scale (for tutorial books on scale-space see ter Haar Romeny [17], Florack [18] and Lindeberg [19]):

$$f:\mathbb{R}^2\times\mathbb{R}^+\to\mathbb{R}:(x,y;s)\mapsto f(x,y;s)\overset{\text{def}}{=}(f_0*\phi_s)(x,y),$$

$$\phi_s(x,y)=\phi(x,y;s)\overset{\text{def}}{=}\frac{1}{4\pi s}\exp\left(-\frac{x^2+y^2}{4s}\right). \tag{1}$$

This results in a 3D function, where a slice of constant scale represents a blurred version of the original image.

The scale-space of an image fulfills the heat equation, since Green's function of the Laplacian operator is a Gaussian kernel:

$$\partial_s f(x,y;s)=\Delta_{(x,y)}f(x,y;s),\quad \partial_s f(x,y;0)=f_0(x,y). \tag{2}$$

The Laplacian in the spatial directions $x$ and $y$ is denoted by $\Delta_{(x,y)}$.

A singular point in scale-space, also called a toppoint, occurs when the following conditions are fulfilled (see Gilmore et al. [20] and Damon [21]):

$$\begin{bmatrix}\nabla_{(x,y)}f\\ \det\mathbf{H}\end{bmatrix}=\begin{bmatrix}f_x\\ f_y\\ f_{xx}f_{yy}-f_{xy}^2\end{bmatrix}=\mathbf{0},$$

where

$$\mathbf{H}\overset{\text{def}}{=}\begin{bmatrix}f_{xx} & f_{xy}\\ f_{xy} & f_{yy}\end{bmatrix}. \tag{3}$$

The gradient operator with respect to $x$ and $y$ is denoted by $\nabla_{(x,y)}$ and partial derivatives of $f$ are indicated by self-explanatory subscripts. The condition states that the gradient is zero at toppoints, which in general occurs at extrema and saddle points in 2D images. These extrema and saddle points exist at every scale, and form the so-called critical paths through scale-space. When two critical paths, corresponding to a saddle point and an extremum, collide as scale increases, an annihilation takes place. A pair of two critical paths can also be created when moving up in scale, which is called a creation. The points in scale-space where these events take place are called toppoints. As a consequence, toppoints are locations in scale-space where a topological change occurs. Fig. 1 shows how two Gaussian blobs merge when scale increases, causing the maximum of the smallest blob to annihilate with the saddle point between the two blobs, creating a toppoint at the scale where this occurs.

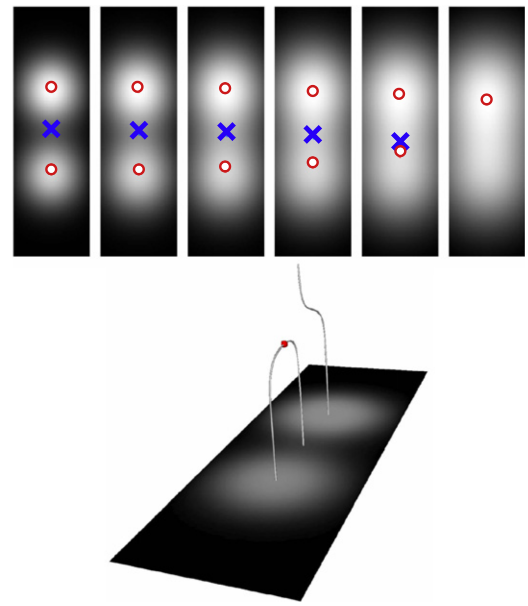

**Fig. 1.** (Top) A series of scale-space slices of an image of two Gaussian blobs of different sizes, where scale increases to the right. Red circles denote maxima and blue crosses denote saddle points. A toppoint is located between the 5th and 6th slice, where a maximum and a saddle point annihilate. (Bottom) The critical paths of the scale-space of the same image, where a toppoint is indicated by a red dot. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

A well-posed formulation of spatial derivatives of an image in scale-space is given by partially integrating the convolution product of a derivative of the image $f_0$ with a Gaussian filter $\phi_s$, see Eq. (1), using the property that $\phi_s$ is a Schwartz function:

$$(\partial_x^n\partial_y^m f_0*\phi_s)(x,y)=(f_0*\partial_x^n\partial_y^m\phi_s)(x,y). \tag{4}$$

In fact, because $f_0$ is often not $(m+n)$ times differentiable, we define the scale-space of an image derivative by the right hand side of Eq. (4). This results in a lower-bound on the scale at which derivatives can be calculated numerically, which increases with derivative order. Derivatives with respect to scale can be calculated using only spatial derivatives by means of Eq. (2).

### 2.2. Toppoint velocity

If we consider a sequence of successive images, or a movie, in which objects move, the toppoints will move as well. The movement of toppoints in spatial and scale direction is defined as: $(\dot{x},\dot{y},\dot{s})\in\mathbb{R}^3$. Note that e.g. $\dot{x}(t)=\partial_t x(t)$ represents the time derivative of the $x(t)$ position of the toppoint. An expression for this toppoint movement can be obtained by implicitly differentiating the definition of toppoints as stated in Eq. (3) with respect to the time parameter $t$:

$$\frac{d}{dt}\begin{bmatrix}\nabla_{(x,y)}f\\ \det\mathbf{H}\end{bmatrix}=\begin{bmatrix}f_{xt}+\dot{x}f_{xx}+\dot{y}f_{xy}+\dot{s}f_{xs}\\ f_{yt}+\dot{x}f_{xy}+\dot{y}f_{yy}+\dot{s}f_{ys}\\ \partial_t\det\mathbf{H}+\dot{x}\partial_x\det\mathbf{H}+\dot{y}\,\partial_y\det\mathbf{H}+\dot{s}\partial_s\det\mathbf{H}\end{bmatrix}=\mathbf{0}$$

$$\Rightarrow\begin{bmatrix}f_{xx} & f_{xy} & f_{xs}\\ f_{xy} & f_{yy} & f_{ys}\\ \partial_x\det\mathbf{H} & \partial_y\det\mathbf{H} & \partial_s\det\mathbf{H}\end{bmatrix}\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{s}\end{bmatrix}=-\begin{bmatrix}f_{xt}\\ f_{yt}\\ \partial_t\det\mathbf{H}\end{bmatrix}. \tag{5}$$

If the matrix is invertible, Eq. (5) supplies us with a scheme to calculate the movement of toppoints in an image sequence. The