Semester project

Espen Johansen Velsvik

February 24, 2016

1 Motivation

When an observer of an object moves relative to the object, there is an apparent relative motion in the image plane of the observer. The problem of determining this relative motion from a sequence of images is called the Optical Flow problem. The analysis is not so much dependent on prior knowledge of the scene, but on the image sequence itself. This independency makes it applicable in many different fields. More concisely, one wants to find a flow vector components $u, v \in \mathbb{R}$ by looking at the change in brightness $f(x,t)$ at a specific pixel from one frame to another. This is problematic because we can not independently determine a vector of 2 components using one constraint coming from the change in image brightness at a point x. Thus we need to impose other constraints to make the problem solvable.

2 The Brightness constancy assumption

The starting point for the variational approach to optical flow is the so called brightness constancy assumption of Horn and Schunck [1]. Let $f(x,t)$ be the grayscale value of some image sequence. To constrain our problem we make an assumption regarding invariance in the brightness: a point moving with velocity $\frac{dx}{dt} = u(x,t)$ along the trajectory $x(t)$ over time t does not change its appearance. This assumption is called the brightness constancy assumption, and it means that if the scene has the same lighting, then movement of an object along a trajectory does not change its brightness. In mathematical notation this is (under perfect conditions) equivalent to the following:

$$\frac{d}{dt} f(x(t), t) = 0.$$

By using the chain rule for differentiation, and defining $\frac{dx}{dt} = u$ and $\frac{dy}{dt} = v$, one gets

$$\frac{\partial f}{\partial x} u + \frac{\partial f}{\partial y} v + \frac{\partial f}{\partial t} = 0,$$

or equivalently $\nabla f \cdot (u, v) = \frac{\partial f}{\partial t}$. So from this equation we can only determine the components of the movement in the direction of the gradient, or what is known as the normal flow. This is called the Aperture problem, and to be able to compute the components of the flow vector one needs another constraint. Following the notation and terminology of [3], the constraint coming from the brightness constancy assumption is called the data term $M(u, v)$

3   The Smoothness constraint

The second constraint is referred to as the smoothness constraint. It says that points can not move independently in the brightness pattern. There has to be some smoothness in the flow vector for points belonging to the same object. In other words, points on the same object moves with the same velocity. A natural way of obtaining a smoother solution would be to minimize some term depending on the sizes of the gradients in some direction. For now we let this smoothness term be $V(\nabla u, \nabla v)$.

4   The variational formulation

To combine the two constraints into one term we form a global energy function consisting of the data term and the smoothness term:

$$E(u,v) = \frac{1}{1} \int_\Omega (M(u,v) + \frac{1}{\sigma^2} V(\nabla u, \nabla v)) dx dy, \tag{1}$$

where $\sigma > 0$ is a regularization parameter. The problem is now to find the minima of the energy functional $E(u,v)$. From calculus of variations we have that if $y*$ minimizes a functional F

$$\begin{aligned} \frac{\partial M}{\partial u} - \frac{1}{\sigma^2}(\frac{\partial}{\partial x}\frac{\partial V}{\partial u_x} + \frac{\partial}{\partial y}\frac{\partial V}{\partial u_y}) = 0 \\ \frac{\partial M}{\partial v} - \frac{1}{\sigma^2}(\frac{\partial}{\partial x}\frac{\partial V}{\partial v_x} + \frac{\partial}{\partial y}\frac{\partial V}{\partial v_y}) = 0, \end{aligned} \tag{2}$$

where the notation $q_{x_i} = \frac{\partial q}{\partial x_i}$ is used.

$$\begin{aligned} \frac{\partial M}{\partial u} - \frac{1}{\sigma^2} div(\mathbf{D}\nabla u) \\ \frac{\partial M}{\partial v} - \frac{1}{\sigma^2} div(\mathbf{D}\nabla v), \end{aligned} \tag{3}$$

where $\mathbf{D}$ is a diffusion matrix steering the direction of the diffusion.

5   The approach by Horn and Schunck

The research of Horn and Shunck [1] has formed the basis of further research in the field of optical flow. They proposed the following quadratic penalised data term

$$M(\mathbf{w}) = (\nabla f \mathbf{w} + \frac{\partial f}{\partial t})^2 = (\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})^2, \tag{4}$$

where $\mathbf{w} = (u,v)$. Most of the methods in variational optical flow use this data term. The contribution to (2) are the following two terms

$$\begin{aligned} \frac{\partial M}{\partial u} = 2(\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})\frac{\partial f}{\partial x} \\ \frac{\partial M}{\partial v} = 2(\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})\frac{\partial f}{\partial y}. \end{aligned} \tag{5}$$

The smoothness term used by Horn and Schunck is the following

$$V(\nabla u, \nabla v) = |\nabla u|^2 + |\nabla v|^2.$$

This is a homogeneous regularizer which means that it performs an equal amount of diffusion in all directions. In the framework of (3), this is equivalent to the diffusion matrix $\mathbf{D}$ being the identity matrix. Using this function as a flow regularizer gives

$$\frac{\partial V}{\partial q_{x_i}^i} = 2q_{x_i}^i,$$

for $q^i = u, v$ and $x_i = x, y$. Dividing by the factor of 2 in all terms results in (2) taking the form

$$
\begin{aligned}
(\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})\frac{\partial f}{\partial x} - \frac{1}{\sigma^2}(\frac{\partial}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial}{\partial y}\frac{\partial u}{\partial y}) = 0 \\
(\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})\frac{\partial f}{\partial y} - \frac{1}{\sigma^2}(\frac{\partial}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial}{\partial y}\frac{\partial v}{\partial y}) = 0
\end{aligned}
\tag{6}
$$

or equivalently

$$
\begin{aligned}
(\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})\frac{\partial f}{\partial x} - \frac{1}{\sigma^2}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})u = 0 \\
(\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t})\frac{\partial f}{\partial y} - \frac{1}{\sigma^2}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})v = 0
\end{aligned}
$$

5.1 Discretizing the Horn and Schunck method

Let now our image be of dimension m-by-n, and let f^i be the image at $t = i$. Also, we let $(x^i)_{i \in [mn]}$ define a regular 2-dimensional grid in Ω. The corresponding vector representation of the image f is denoted as $\tilde{\mathbf{f}}(x^i) \in \mathbb{R}^{mn}$. Continuing this notation, the discrete flow values $\tilde{\mathbf{w}}$ is represented as the following vector in $\mathbb{R}^{2mn}$:

$$\tilde{\mathbf{w}} = \begin{bmatrix} u(x^i)_{i \in [mn]} \\ v(x^i)_{i \in [mn]} \end{bmatrix}.$$

For the discretization of the image gradient in (6) the sobel filters

$$
G_x = \begin{bmatrix} \text{-1} & 0 & 1 \\ \text{-2} & 0 & 2 \\ \text{-1} & 0 & 1 \end{bmatrix}
\qquad \text{and} \qquad
G_y = \begin{bmatrix} \text{-1} & \text{-2} & \text{-1} \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}
$$

were convolved with the image f^1 to produce the vectors $\tilde{\mathbf{d}}_x(x^i)$ and $\tilde{\mathbf{d}}_y(x^i)$ in $\mathbb{R}^{mn}$, approximating the image gradients of $\tilde{\mathbf{f}}(x^i)$ in x- and y-direction respectively. The time derivative $\frac{\partial f}{\partial t}$ is discretized using forward difference as shown below:

$$\mathbf{c} = \tilde{\mathbf{f}}^2(x^i) - \tilde{\mathbf{f}}^1(x^i),$$

where $\mathbf{c}(x^i)$ is a vector in $\mathbb{R}^{mn}$. When choosing the derivative approximations for the flow vector in (6), the first derivative was approximated using backward difference, and the second was approximated using forward difference. The resulting system to be solved is shown below:

$$(D^T D + \frac{1}{\sigma^2} L^T L)\tilde{\mathbf{w}} = -D^T \mathbf{c}. \tag{7}$$

D is the following block matrix:

$$D = \left[\begin{array}{c|c} D_x & D_y \end{array} \right],$$

where D_x and D_y are diagonal matrices in $\mathbb{R}^{mn \times mn}$ with the elements of $\tilde{\mathbf{d}}_x(x^i)$ and $\tilde{\mathbf{d}}_y(x^i)$ along its diagonals respectively. L is the following block matrix:

$$L = \left[\begin{array}{cc|cc} L_x & 0_{mn,mn} \\ L_y & 0_{mn,mn} \\ 0_{mn,mn} & L_x \\ 0_{mn,mn} & L_y \end{array} \right],$$

$Lx \in \mathbb{R}^{mn \times mn}$ being the following mn-by-mn block diagonal matrix:

$$L_x = \left[\begin{array}{c|c|c|c} I_{m,m} & -I_{m,m} & 0 & \cdots \\ \hline 0 & \ddots & \ddots & 0 \\ \hline 0 & 0 & I_{m,m} & -I_{m,m} \end{array} \right],$$

and $L_y \in \mathbb{R}^{mn \times mn}$ the following band matrix:

$$L_y = \left[\begin{array}{cccc} 1 & -1 & & 0 \\ 0 & \ddots & \ddots & \\ & & 1 & -1 \end{array} \right].$$

The matrix multiplication with L gives a forward difference approximation, and $-L^T$ gives a backward difference approximation. The matrix product $-L^T L$ is then the discretization of the second derivative in (6).

6   Anisotropic image-driven regularization

The homogeneous regularizer of Horn and Schunck smoothes the flow field in all directions. Since we are looking for a flow field describing relative movement of objects, the boundary between objects will not have a smooth flow field unless they are moving at the same relative velocity. The anisotropic regularizer of Nagel and Enkelmann [2] performs smoothing along the image gradients and prevents smoothing across image edges. This is done by introducing a regularised projection matrix P, defined as

$$P(\nabla f) = \frac{1}{|\nabla f|^2 + 2\kappa^2}(\nabla^\perp f(\nabla^\perp)^T + \kappa^2 I),$$

where $\nabla^\perp f = \left[-\frac{\partial f}{\partial y}, \frac{\partial f}{\partial x} \right]^T$, and $\kappa > 0$ is a regularization parameter. The smoothness term of Nagel and Enkelmann can now be written as the following

$$V(\nabla u, \nabla v) = \nabla^T u P(\nabla f) \nabla u + \nabla^T v P(\nabla f) \nabla v$$

4

References

[1] Horn, B. and Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 17, 185-203.

[2] Nagel, H. and Enkelmann, W. (1986). An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 565-593

[3] Zimmer, H., Bruhn, A., Weickert, J. (2011). Optic Flow in Harmony *International Journal of Computer Vision*, 93, 368-388