Optical flow and diffeomorphic methods

Jan Kybic

2023

Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. IJCV 2005

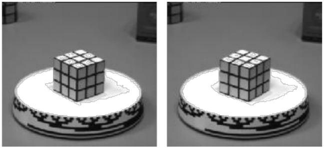
some slides from S. Kong, S. Lazebnik, K. Grauman, S. Seitz, R. Szeliski, M. Pollefeys, S. Nayar, J. Niebles, R. Krishnan

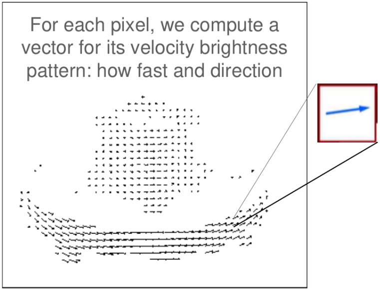
Optical flow

- Register two (or more images)
- Assume small motion
- Assume brightness constant
- Assume spatial coherence
- Sometimes allows occlusions
- Provides dense field
- Fast

Motion field

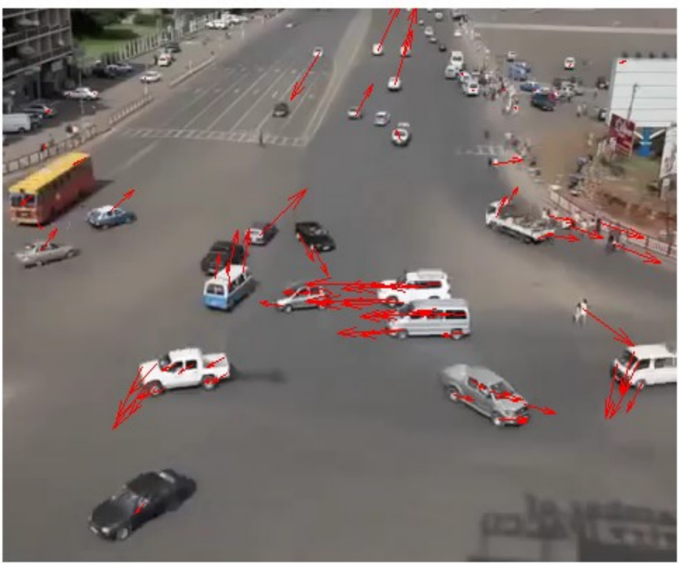
- The motion field is the projection of the 3D scene motion into the image





Input images

Output motion



Optical flow applications

- traffic monitoring
- autonomous driving
- optical mouse
- video stabilization
- motion interpolation (slow motion)
- motion magnification
- motion measurement (e.g. heart)

Brightness constancy

- Smooth image

$$f(x, y, t) := (K_\sigma * g)(x, y, t),$$

- Brightness constancy

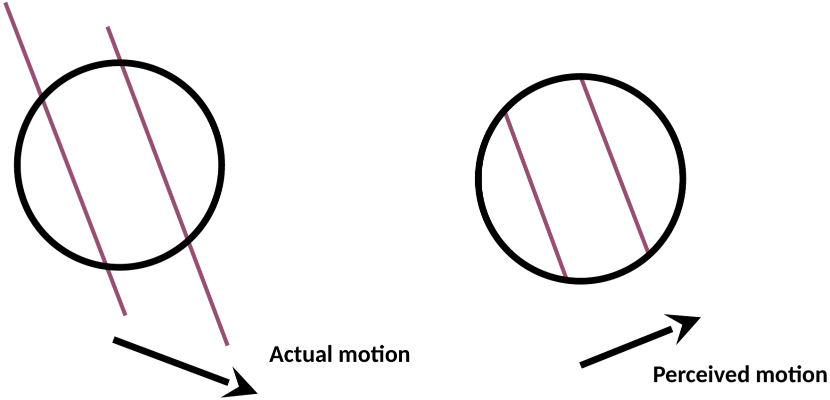
$$f(x+u,\ y+v,\ t+1) = f(x, y, t),$$

- Optic flow constraint

$$f_x u + f_y v + f_t = 0,$$

Aperture problem

Optic flow constraint is underdetermined



Only normal motion can be recevered.

Lucas-Kanade

- Average MSE over a neighborhood

$$E_{LK}(u,v) := K_\rho * \left((f_x u + f_y v + f_t)^2\right).$$

- Linear system of equations at each point from $\partial_u E_{LK} = 0, \partial_v E_{LK} = 0$

$$\begin{pmatrix} K_\rho * \left(f_x^2\right) & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * \left(f_y^2\right) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * (f_x f_t) \\ -K_\rho * (f_y f_t) \end{pmatrix}$$

- Algorithm:
  - smooth image
  - calculate derivatives
  - calculate matrix at each point
  - smooth matrix coefficients in space
  - solve 2×2linear system in each point

Lucas-Kanade disadvantages

- in large homogeneous regions, matrix remains ill-conditioned
- matrix conditioning (eigenvalues) $\rightarrow$local reliability estimate
- large scale $\rho \rightarrow$ poor resolution

Horn-Schunck

- Regularization. Penalize unsmooth motion field

$$E_{HS}(u, v) = \int_\Omega ((f_x u + f_y v + f_t)^2 + \alpha(|\nabla u|^2 + |\nabla v|^2))\, dx\, dy$$

- Euler-Lagrange equations

$$0 = \Delta u - \frac{1}{\alpha}\left(f_x^2 u + f_x f_y v + f_x f_t\right),$$
$$0 = \Delta v - \frac{1}{\alpha}\left(f_x f_y u + f_y^2 v + f_y f_t\right).$$

- can be solved iteratively (computationaly complex)
- extrapolates to homogeneous locations

Combined local-global method

- Notation
$$\mathbf{w} := (u, v, 1)^\top,$$
$$|\nabla \mathbf{w}|^2 := |\nabla u|^2 + |\nabla v|^2,$$
$$\nabla_3 f := (f_x, f_y, f_t)^\top,$$
$$J_\rho(\nabla_3 f) := K_\rho * (\nabla_3 f \, \nabla_3 f^\top)$$
- Lucas-Kanade minimizes
$$E_{LK}(\mathbf{w}) = \mathbf{w}^\top J_\rho(\nabla_3 f)\, \mathbf{w},$$
- Horn-Schunck minimizes
$$E_{HS}(\mathbf{w}) = \int_\Omega (\mathbf{w}^\top J_0(\nabla_3 f)\, \mathbf{w} + \alpha |\nabla \mathbf{w}|^2)\, dx\, dy.$$

Combined local-global method (2)

- Lucas-Kanade minimizes

$$E_{LK}(\mathbf{w}) = \mathbf{w}^\top J_\rho(\nabla_3 f)\,\mathbf{w},$$

- Horn-Schunck minimizes

$$E_{HS}(\mathbf{w}) = \int_\Omega (\mathbf{w}^\top J_0(\nabla_3 f)\,\mathbf{w} + \alpha|\nabla \mathbf{w}|^2)\,dx\,dy.$$

- Combined method minimizes

$$E_{CLG}(\mathbf{w}) = \int_\Omega \left(\mathbf{w}^\top J_\rho(\nabla_3 f)\,\mathbf{w} + \alpha|\nabla \mathbf{w}|^2\right)\,dx\,dy$$

Euler-Langrange equations

$$
\begin{aligned}
0 &= \Delta u - \frac{1}{\alpha}\Big(K_\rho * \left(f_x^2\right) u + K_\rho * (f_x f_y)\, v \\
&\quad + K_\rho * (f_x f_t)\Big), \\
0 &= \Delta v - \frac{1}{\alpha}\Big(K_\rho * (f_x f_y)\, u + K_\rho * \left(f_y^2\right) v \\
&\quad + K_\rho * (f_y f_t)\Big).
\end{aligned}
$$

Spatio-temporal extension

- sequence of images, Gaussian in space+time

$$E_{CLG3}(\mathbf{w}) = \int_{\Omega\times[0,T]} (\mathbf{w}^\top J_\rho(\nabla_3 f)\,\mathbf{w} + \alpha|\nabla_3 \mathbf{w}|^2)\,dx\,dy\,dt$$

- Euler-Lagrange equations

$$\Delta_3 u - \frac{1}{\alpha}(J_{11}u + J_{12}v + J_{13}) = 0,$$
$$\Delta_3 v - \frac{1}{\alpha}(J_{12}u + J_{22}v + J_{23}) = 0.$$

$$\Delta_3 := \partial_{xx} + \partial_{yy} + \partial_{tt}.$$

- more computationally complex - the whole sequence is processed together

Robust (nonquadratic) retularization

- Quadratic penalty

$$E_{CLG3}(\mathbf{w}) = \int_{\Omega\times[0,T]} (\mathbf{w}^\top J_\rho(\nabla_3 f)\,\mathbf{w} + \alpha|\nabla_3\mathbf{w}|^2)\,dx\,dy\,dt$$

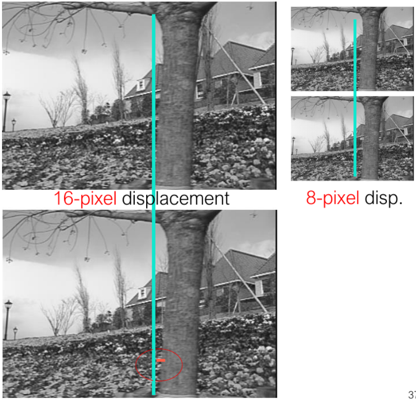
- Nonquadratic penalty

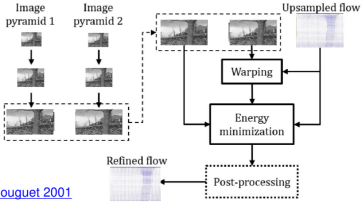
$$E_{CLG3-N}(\mathbf{w}) = \int_{\Omega\times[0,T]} (\psi_1(\mathbf{w}^\top J_\rho(\nabla_3 f)\,\mathbf{w}) + \alpha\,\psi_2(|\nabla_3\mathbf{w}|^2))\,dx\,dy\,dt$$

$$\psi_i(s^2) = 2\beta_i^2\sqrt{1+\frac{s^2}{\beta_i^2}},$$

- Euler-Lagrange equations nonlinear

Coarse-to-fine flow estimation





37 Bouguet 2001

Multiresolution and combined optical flow

- Image pyramid
- Flow at coarse resolution
  - **not** used as initialization
  - used to warp the original sequence
  - needs to be upsampled and scaled
- Final flow is a **sum** of motions at all scales

$$
\begin{aligned}
&E^{m}_{CLG3-N}(\delta\mathbf{w}^m) \\
&\quad = \int_{\Omega\times[0,T]} (\psi_1(\delta\mathbf{w}^{m\top} J_\rho(\nabla_3 f(\mathbf{x}+\mathbf{w}^m))\,\delta\mathbf{w}^m) \\
&\qquad + \alpha\,\psi_2(|\nabla_3(\mathbf{w}^m + \delta\mathbf{w}^m)|^2))\,\mathbf{dx}
\end{aligned}
$$

Implementation

- Flow field discretized, grid size *h*
- Spatial derivatives - 6th order finite differences
- Discretized Euler-Lagrange equations

$$0 = \sum_{j \in \mathcal{N}(i)} \frac{u_j - u_i}{h^2} - \frac{1}{\alpha} (J_{11i}\, u_i + J_{12i}\, v_i + J_{13i}) ,$$

$$0 = \sum_{j \in \mathcal{N}(i)} \frac{v_j - v_i}{h^2} - \frac{1}{\alpha} (J_{21i}\, u_i + J_{22i}\, v_i + J_{23i}) \tag{32}$$

  J_{nmi} are components of $J_\rho(\nabla f)$ at pixel *i*
- Sparse linear system of equations

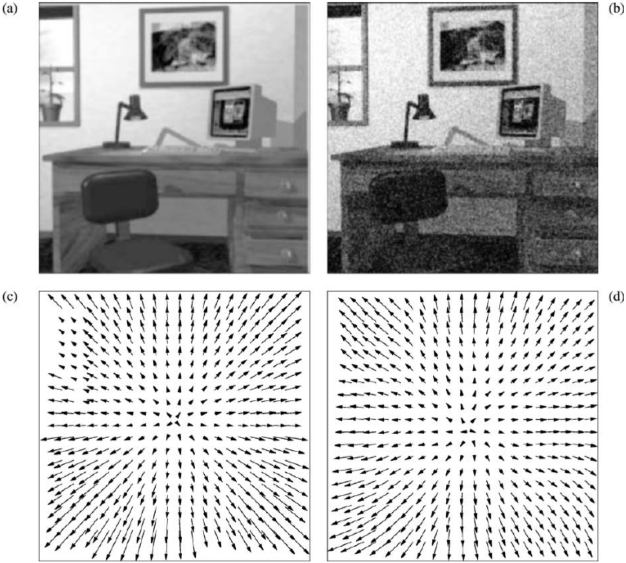
Successive overrelaxation (SOR)

$$u_i^{k+1} = (1-\omega)\, u_i^k \; + \; \omega \frac{\sum_{j\in\mathcal{N}^-(i)} u_j^{k+1} + \sum_{j\in\mathcal{N}^+(i)} u_j^k - \frac{h^2}{\alpha}\left(J_{12i}\, v_i^k + J_{13i}\right)}{|\mathcal{N}(i)| + \frac{h^2}{\alpha} J_{11i}},$$

$$v_i^{k+1} = (1-\omega)\, v_i^k \; + \; \omega \frac{\sum_{j\in\mathcal{N}^-(i)} v_j^{k+1} + \sum_{j\in\mathcal{N}^+(i)} v_j^k - \frac{h^2}{\alpha}\left(J_{21i}\, u_i^{k+1} + J_{23i}\right)}{|\mathcal{N}(i)| + \frac{h^2}{\alpha} J_{22i}} \qquad (34)$$

relaxation parameter $\omega \in (0,2)$
components updated sequentially - only storage Nrequired
4ms/iterations on 316×252images, 1000 iterations
multigrid techniques may achieve real time

Example results



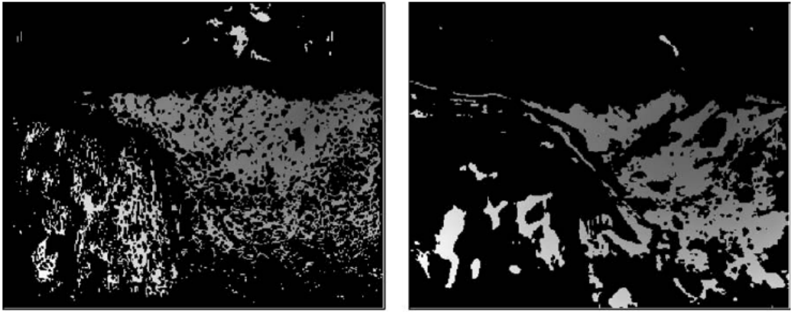
Comparison with other methods

Technique	Multiscale	Spatiotemporal information	Spatiotemporal constraint	AAE
Horn/Schunck, original (Barron et al., 1994)	–	√	–	31.69°
Singh, step 1 (Barron et al., 1994)	–	–	–	15.28°
Anandan (Barron et al., 1994)	–	–	–	13.36°
Singh, step 2 (Barron et al., 1994)	–	–	–	10.44°
Nagel (Barron et al., 1994)	–	√	–	10.22°
Horn/Schunck, modified (Barron et al., 1994)	–	√	–	9.78°
Uras et al., unthresholded (Barron et al., 1994)	–	√	–	8.94°
2-D CLG linear	–	–	–	7.09°
3-D CLG linear	–	√	√	6.24°
2-D CLG nonlinear	–	–	–	6.03°
Alvarez et al. (2000)	√	–	–	5.53°
Mémin and Pérez (1998)	√	–	–	5.38°
3-D CLG nonlinear	–	√	√	5.18°
2-D CLG nonlinear multires	√	–	–	4.86°
Mémin and Pérez (1998)	√	–	–	4.69°
3-D CLG nonlinear multires	√	√	√	4.17°

Confidence measure

- Pixel contributions to energy functional E_i
- Take $p$percent of pixels with the lower E_i

Confidence measure examples



(a) (b)

(a) confidence measure, (b) lowest error