Master Mathematical Analysis and Applications
Course M1 - S2

Computer vision

Motion estimation

Week 8

Mohammed Hachama
hachamam@gmail.com
http://hachama.github.io/home/
University of Khemis Miliana

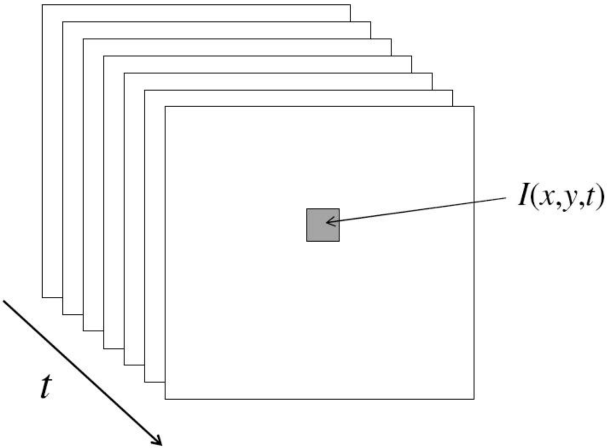
-2020-

Plan

Optical flow

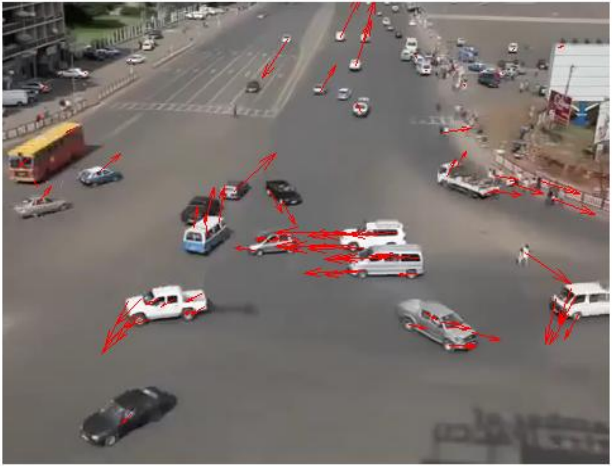
Optical flow

- A video is a sequence of frames captured over time
- Now our image data is a function of space (x, y) and time (t)



Optical flow

Motion



Optical flow

Applications

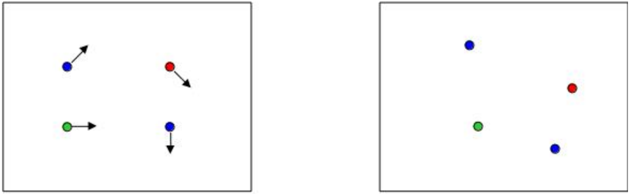
- 3D shape reconstruction
- Object segmentation
- Learning and tracking of dynamical models
- Event and activity recognition

Optical flow

- Optical flow is the apparent motion of brightness patterns in the image

  - The motion field is the projection of the 3D scene motion into the image
  - Apparent motion can be caused by lighting changes without any actual motion
- Ideally, optical flow would be the same as the motion field

Optical flow



$I(x,y,t-1)$ $\quad$ $I(x,y,t)$

- Given two subsequent frames, estimate the apparent motion field u(x,y) and v(x,y) between them
- Key assumptions
  - **Brightness constancy** : projection of the same point looks the same in every frame
  - **Small motion** : points do not move very far
  - **Spatial coherence** : points move like their neighbors

Optical flow

- Brightness Constancy Equation :
$$I(x, y, t-1) = I(x + u(x, y), y + v(x, y), t)$$
- Linearizing the right side using Taylor expansion :
$$I(x, y, t-1) \approx I(x, y, t) + I_x u(x, y) + I_y v(x, y)$$
- Hence,
$$I_x u + I_y v + I_t \approx 0$$
- One equation, two unknowns per pixel
- The component of the flow perpendicular to the gradient (i.e., parallel to the edge) is unknown
$$\nabla.(u, v) + I_t = 0$$

Optical flow

Lucas-Kanade method

Lucas-Kanade method

Solving the aperture problem

- How to get more equations for a pixel ?
- Spatial coherence constraint : pretend the pixel's neighbors have the same (u, v)
  - E.g., if we use a 5×5 window, that gives us 25 equations per pixel

$$\nabla I(x_i)[u, v] + I_t(x_i) = 0$$

$$\begin{pmatrix} I_x(x_1) & I_y(x_1) \\ \dots & \\ I_x(x_n) & I_y(x_n) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{pmatrix} I_t(x_1) \\ \dots \\ I_t(x_n) \end{pmatrix}$$

[1]

1. B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. 1981

Lucas-Kanade method

- Linear least squares problem (over-constrained) :

$$\overset{n\times 2}{A}\ \overset{2\times 1}{d} = \overset{n\times 1}{b}$$

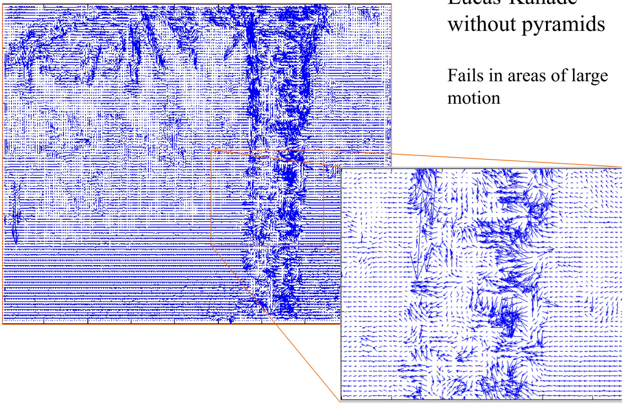
- Solution given by

$$(A^T A)d = A^T b$$

$$\begin{pmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}$$

  - $A^T A$ should be invertible
  - $A^T A$ should not be too small due to noise : eigenvalues λ_1 and λ_2 of $A^T A$ should not be too small
  - $A^T A$ should be well-conditioned : λ_1/λ_2 should not be too large ($\lambda_2 =$ larger eigenvalue).

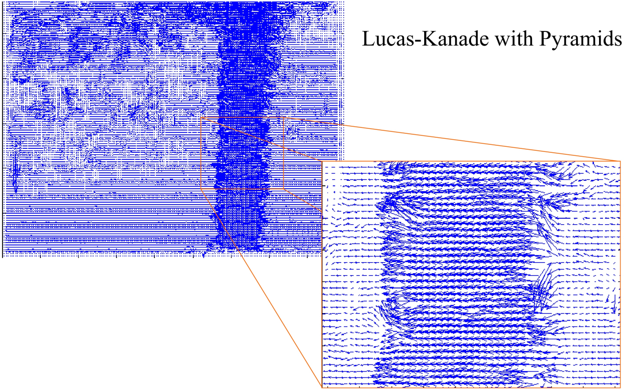
Lucas-Kanade method



Lucas-Kanade without pyramids

Fails in areas of large motion

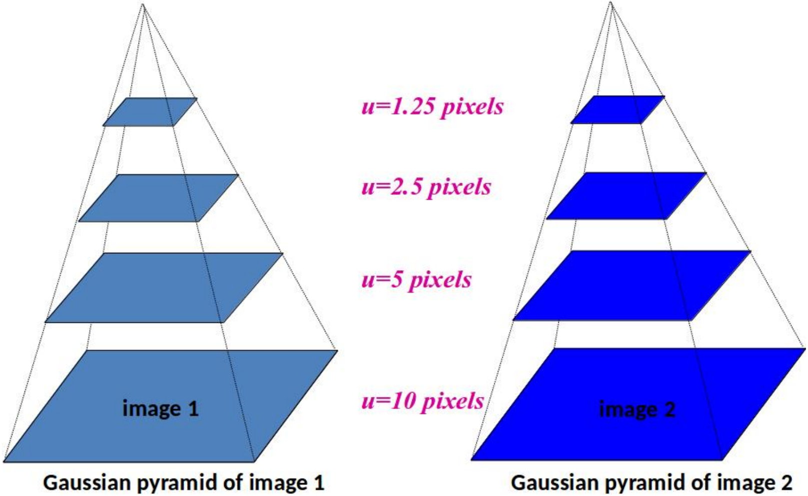
Lucas-Kanade method



Lucas-Kanade with Pyramids

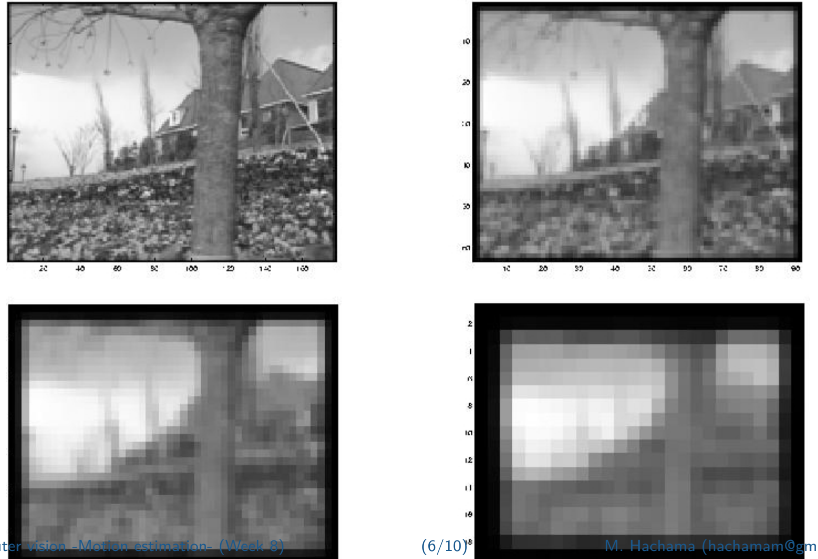
Lucas-Kanade method

Pyramids for large motion



Lucas-Kanade method

Pyramids for large motion



Horn and Schunck

Horn and Schunck

The flow is formulated as a global energy function which is should be minimized :

$$\int (I_x u \ + \ I_y v \ + \ I_t)^2 + \lambda(\|\nabla u\|^2 + \|\nabla v\|^2) \quad dx\, dy$$

$$\int (I_x u \ + \ I_y v \ + \ I_t)^2 + \lambda(u_x^2 + u_y^2 + v_x^2 + v_y^2) \quad dx\, dy$$

$$(I_x u \ + \ I_y v \ + \ I_t)\, I_x \ + \ \lambda\,(\Delta u) \ = \ 0$$
$$(I_x u \ + \ I_y v \ + \ I_t)\, I_y \ + \ \lambda\,(\Delta v) \ = \ 0$$

Global models

Global (parametric) models

Translational alignment

- Data : Template image I_0, Reference image I_1.
- Sum of squared differences (SSD) function

$$E_{SSD}(\mathbf{u}) = \sum_i [I_1(x_i + u) - I_0(x_i)]^2 = \sum_i e_i^2,$$

where $\mathbf{u} = (u, v)$ is the displacement

- Robust error metrics

$$E_R(\mathbf{u}) = \sum_i \rho([I_1(x_i + u) - I_0(x_i)) = \sum_i \rho(e_i),$$

Global (parametric) models

Translational alignment

- Bias and gain (exposure differences)
- Linear (affine) intensity variation model

$$I_1(x + u) = (1 + \alpha)I_0(x) + \beta,$$

where β is the bias and α is the gain

- Least squares formulation

$$E_{SSD}(\mathbf{u}) = \sum_i [I_1(x_i + u) - (1 + \alpha)I_0(x_i) - \beta]^2.$$

Global (parametric) models

Translational alignment

- Correlation : maximize the product (or cross-correlation) of the two aligned images

$$E_{CC}(\mathbf{u}) = \sum_i I_0(\mathbf{x}_i) I_1(\mathbf{x}_i + u)$$

- Normalized cross-correlation is more commonly used

$$E_{NCC}(\mathbf{u}) = \frac{\sum_i [I_0(\mathbf{x}_i) - \bar{I}_0)]\,[I_1(\mathbf{x}_i + \mathbf{u}) - \bar{I}_1]}{\sqrt{\sum_i [I_0(\mathbf{x}_i) - \bar{I}_0)]^2}\,\sqrt{[I_1(\mathbf{x}_i + \mathbf{u}) - \bar{I}_1]^2}}$$

where

$$\bar{I}_0 = \frac{1}{N}\sum_i I_0(x_i), \quad \bar{I}_1 = \frac{1}{N}\sum_i I_1(x_i + u)$$

Global (parametric) models

Translational alignment : Optimization

- Full search over some range of shifts, using either integer or sub-pixel steps (use of a hierarchical approach).
- Gradient descent : Incremental refinement.

Global (parametric) models

Translational alignment : Optimization

- Fourier-based alignment : (when the search range corresponds to a significant fraction of the larger image, as is the case in image stitching)

$$\mathcal{F}\{I_1(x+u)\} = \mathcal{F}\{I_1(x)\}e^{-ju.\omega} = \mathcal{I}_1(\omega)e^{-ju.\omega}$$

$$F\{E_{CC}(u) = \mathcal{F}\left\{\sum_i I_0(x_i)I_1(x_i+u)\right\} = \mathcal{I}_0(\omega)\mathcal{I}_1^*(\omega),$$

where $\mathcal{I}_1^*(\omega)$ is the complex conjugate of $\mathcal{I}_1(\omega)$. Fourier-based convolution is often used to accelerate the computation of image correlations.