Image Sequences

- Motion pictures
- Motion field and optical flow
- Motion models
- Motion estimation

Motion pictures

Illusion of motion: display images with certain frequency (30, 48 images/sec)

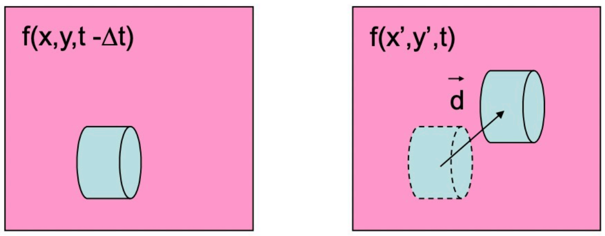
Motion estimation: determine the motion of objects in a scene

Motion compensation: predict a frame with the previous and future frames
—> use the motion estimation information to achieve compensation

Video coding basics — 两种冗余
- Spatial redundancy —> still image compression (intra-frame)
- Temporal redundancy —> differential coding of blocks with motion vectors (inter-frame)

Motion field and optical flow

Displacement: 位移（矢量）
Motion: 速度（矢量）



$$\vec{v} = \frac{\vec{d}}{\Delta t}$$

Motion field: projection of 3D motion of the objects onto the 2D image plane
三维相对速度矢量在二维图像平面上的投影 (2D motion) —> 运动场

Optical flow: apparent motion, variations in illumination influence the optical flow
illumination的改变会让optical flow estimation变差

Brightness constancy: Brightness of the point will remain the same

以Brightness constancy作为假设让optical flow estimation的性能变好， 此时两个frame相同intensity的点可建立对应关系 （光源跟着点走）

Motion models

Motion vector (block)
1 motion vector for each block of pixels (size of block: 8 x 8, 16 x 16)

Affine motion
- 9 parameters: translation, rotation, zoom in/out (3D (x, y, z)) — each 3 params
- 6 parameters: in 2D
- Complex

Translational motion
- 2D parameters: horizontal and vertical translations

$$\boxed{\vec{d} = (d_x, d_y)}$$

2-D Affine Motion Model

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} d_x \\ d_y \end{pmatrix} + \begin{pmatrix} d_{xx} & d_{xy} \\ d_{yx} & d_{yy} \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix}$$

Find the values for the parameters best fits the motion present (six parameters)

Motion estimation

Hypotheses (Assumptions)
- No occlusions [lost the corresponding point]
- Rigid objects [object shape cannot be changed]
- No illumination changes [brightness constancy]
- Locally [continuity of motion, translational motion]

Methods
- Gradient
- Block matching

Gradient method
1. Brightness consistency constraint + small motion —> spatio-temporal constraint

$$\boxed{\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0}$$

$$I(x+\Delta, y+\Delta y, t+\Delta t) = I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t + \text{higher order terms}$$

$$I(x+\Delta, y+\Delta y, t+\Delta t) \approx I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t$$

2. Aperture problem
the fact that the motion of a one-dimensional spatial structure, such as a bar or edge, cannot be determined unambiguously if it is viewed through a small aperture

Solve the aperture problem —> Lucas-Kanade method
Assumption: the neighboring points share the similar motion information = optical flow field changes smoothing in a small region

Lucas-Kanade method
Assumption: the displacement of the image contents between two nearby frames is small

$$I_x(q_1)V_x + I_y(q_1)V_y = -I_t(q_1)$$
$$I_x(q_2)V_x + I_y(q_2)V_y = -I_t(q_2)$$
$$\dots$$
$$I_x(q_n)V_x + I_y(q_n)V_y = -I_t(q_n)$$

Matrix form

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \dots & \dots \\ I_x(q_n) & I_y(q_n) \end{bmatrix} \quad v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \dots \\ -I_t(q_n) \end{bmatrix}$$

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

$$\underbrace{\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} V_x \\ V_y \end{bmatrix} = - \underbrace{\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}}_{A^T b}$$

When is this solvable?

- **A^TA** should be invertible
- **A^TA** should not be too small due to noise
  - eigenvalues λ_1 and λ_2 of **A^TA** should not be too small
- **A^TA** should be well-conditioned
  - λ_1/λ_2 should not be too large (λ_1 = larger eigenvalue)

• Flat region (small λ_1, λ_2) is difficult to be computed by Lucas-Kenade flow
• Edge (large λ_1, small λ_2) is difficult to be computed by Lucas-Kenade flow
—> ill conditioned
• Corner —> well conditioned, can be solved

Assumptions of Lucas-Kenade Algorithm
1. Brightness constancy
2. The motion is not small
3. A point does not move like its neighbor

Gradient methods (优劣)
Advantages:
• Obtain dense motion field
• Appropriate for image sequence analysis
Disadvantages:
• Additional constraint —> increase the energy of error
• Dense motion field —> a large amount of data

Gradient method (summary)
1. spatio-temporal constraint (brightness constancy + small motion)
2. Aperture problem

Assumption for gradient method
1. Brightness constancy
2. The motion is not small
3. A point does not move like its neighbor

Advantages & Disadvantage
Advantage: dense motion field, appropriate for image sequence analysis
Disadvantage: dense motion field —> large amount of data, assumption constraint

Lucas-Kanade method

$$\begin{aligned}
I_x(q_1)V_x + I_y(q_1)V_y &= -I_t(q_1) \\
I_x(q_2)V_x + I_y(q_2)V_y &= -I_t(q_2) \\
&\cdots \\
I_x(q_n)V_x + I_y(q_n)V_y &= -I_t(q_n)
\end{aligned}$$

Matrix form

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \cdots & \cdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix} \qquad v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \qquad b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \cdots \\ -I_t(q_n) \end{bmatrix}$$

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

$$\underbrace{\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} V_x \\ V_y \end{bmatrix} = -\underbrace{\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}}_{A^T b}$$

Solvable
1. Edge (large λ_1, small λ_2), flat (small λ_1, λ_2) region —> difficult to solve ()
2. Corner —> easy to solve

Practice: gradient method (Lucas-Kenade constraints)

a) This question is about **Motion estimation**.

[12 marks]

i) Given an image I at pixel (x, y) and time t, $I(x, y, t)$, derive the spatio-temporal constraint from the brightness constraint in gradient-based motion estimation.

Assume : small motion

(3 marks)

$$I(x+\Delta x, y+\Delta y, t+\Delta t) = I(x,y,t) + I_x \cdot \Delta x + I_y \cdot \Delta y + I_t \cdot \Delta t + \text{higher order terms}$$

$$I(x+\Delta x, y+\Delta y, t+\Delta t) \approx I(x,y,t) + I_x \Delta x + I_y \cdot \Delta y + I_t \cdot \Delta t \quad \text{(ignore)}$$

According to the brightness constancy:

$$I(x+\Delta x, y+\Delta y, t+\Delta t) = I(x,y,t)$$

$$I_x \Delta x + I_y \Delta y + I_t \Delta t = 0$$

$$I_x U_x + I_y U_y + I_t = 0$$

ii) What is the aperture problem in motion estimation?

(3 marks)

Aperture problem is the motion cannot be determined if it is viewed through a small aperture. It indicates that if there is an object with strips in a certain motion, if we view it through different aperture kinds, the motion seems different. In motion estimation, we solve aperture problem to obtain constraint equations to solve parameters. We assume neighboring pixels have similar motion.

本题设x，y坐标系

iii) We have two sequential images at time t and $t+1$. Using Lucas-Kanade algorithm with 3×3 neighborhood (in red box), estimate the motion vector of the gray pixel, $v(2,2) = [v_x(2,2), v_y(2,2)]$, at time t. Round the final answer to two decimal places.

(6 marks)

t

6	8	8	8	8
8	8	8	8	8
7	7	7	8	8
0	0	7	8	8
1	1	7	8	7

$t+1$ (→ y, ↓ x)

6	8	8	8	8
6	7	7	7	8
0	0	0	7	8
0	1	1	7	8
0	1	1	7	7

① Calculate I_x, I_y, I_t

I_x, kernel $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$: $\begin{bmatrix} -1 & -1 & 0 \\ -7 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

I_y, kernel $[-1 \quad 1]$: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & 1 & 0 \end{bmatrix}$

I_t, $\begin{bmatrix} -1 & -1 & -1 \\ -7 & -7 & -1 \\ 1 & -6 & -1 \end{bmatrix}$

② $A \cdot v = b$

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ \vdots & \vdots \\ I_x(q_9) & I_y(q_9) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & 0 \\ 0 & 0 \\ -7 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 7 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$v = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

$$b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \vdots \\ -I_t(q_9) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 7 \\ 7 \\ 1 \\ -1 \\ 6 \\ 1 \end{bmatrix}$$

③ $(A^T A) \cdot v = A^T \cdot b$

$$A^T \cdot A = \begin{bmatrix} 52 & 7 \\ 7 & 51 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} -1.04 \\ 0.26 \end{bmatrix}$$

$$A^T \cdot b = \begin{bmatrix} -52 \\ 6 \end{bmatrix} \qquad v(2,2) = [-1.04 \quad 0.26]$$

Block matching method
基本思路: divide images into small blocks, for each block, a similar block is searched for in the previous image. The relationship between the 2 blocks generate the motion vector for the block under analysis.

- Similarity measure
- Search Algorithms

Similarity measure:
SAD, MSE has some problems, in this case:

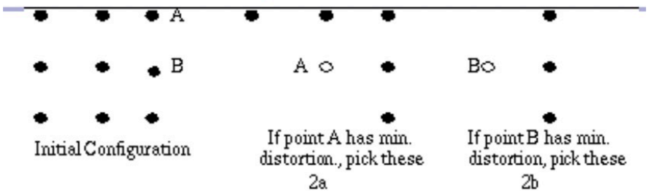
Cross-correlation can be used in this scenario

Search Algorithms
Range [-7, +7], it searches ()
Optimal solution, computational $(2 \cdot 7 + 1)^2 = 225$

Logarithmic search — remove redundancy

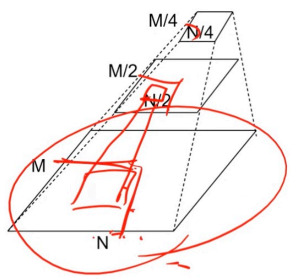
3-step search
Repeat algorithm 3 times (examine: 25 points)

4-step search — the motion vector tendency of the direction



Conjugate search

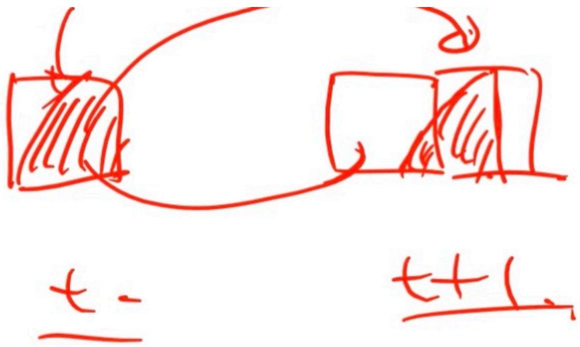
Pyramidal search
Don't need to search whole area, search the area with low-resolution, map it to the higher resolution then search again and map again



Multi-resolution representation by pyramid

Block matching methods (优劣)
Advantages:
- minimize the prediction error
- Rgular structure

Disadvantages:
- Complexity
- Large prediction error at the edge of moving objects (the motion within the box should be the same)



Block matching method (summary)

基本思路: divide images into small blocks, for each block, find the similar block in the search window of the previous images using a specific search algorithm. Use the suitable similarity measure to choose the similar search window. The relationship between 2 blocks generate the motion vector for the block.

Similarity measure
Search algorithm: full search, logarithmic, 3-step, 4-step, conjugate, pyramidal

Advantages & disadvantages
Advantages:
- minimize the prediction error
- Regular structure (especially using full-search)

Disadvantages:
- Complexity
- Large errors at the edges of moving objects

a) This question is about **Motion estimation**.

[6 marks]

i) Explain how block matching works for motion estimation.

(4 marks)

ii) What advantages and disadvantages does block matching have?

(2 marks)

(1) The block matching is one of the motion estimation method. The images in a sequence can be divided into small blocks, for each block, find the similar block in the search window of the previous image. The relationship between the two blocks give the motion vector. The similarity can be measured using various methods, such as SAD, MSE, cross-correlation and etc. There are multiple kinds of search algorithm can be used in the block matching method, such as full search (regular structure), logarithmic search, 3-step/4-step search, conjugate search and pyramidal search.

(2) Advantages: It minimizes the prediction error, it has regular structure (full-search). Disadvantages: High complexity, the prediction error at the edges of moving objects is high.