Image and video processing

Image sequences

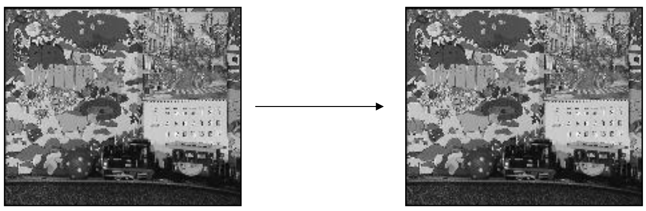
Dr. Qianni Zhang (qianni.zhang@qmul.ac.uk)

Agenda

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

Introduction

- Motion in a sequence of images
  - The illusion of motion in a sequence of pictures is given by displaying images with a certain frequency
    - e.g. 25, 30or 48 images/sec
  - The number of images shown each second depends on
    - the spatial resolution of the images (cinema, TV) as well as on
    - the amplitude of the motion



EBU723U

- 3 -

Motion estimation and compensation

- Motion estimation
  - Set of methods aiming at determining the motion of objects in a scene that describe the transformation from one 2D image to another;
  - usually from adjacent frames in a video sequence.
- Motion compensation
  - Image processing algorithms using motion information
  - to predict a frame in a video, given the previous and/or future frames by accounting for motion of the camera and/or objects in the video

EBU723U

- 4 -

Example of applications

- Motion is a rich source of information about the world:
  - Segmentation
  - surface structure from parallax
  - self-motion
  - recognition
  - understanding behavior
  - understanding scene dynamics
- Other correspondence / registration problems:
  - stereo disparity (short and wide baseline)
  - computer-assisted surgery
  - multiview alignment for mosaicing or stop-frame animation

EBU723U

- 5 -

Video coding basics

- Take advantage of:
  - Spatial redundancy
    - use methods similar to still image coding
  - Temporal redundancy
    - use inter-frame compression
      - differential coding of blocks with motion vectors

EBU723U

- 6 -

Agenda

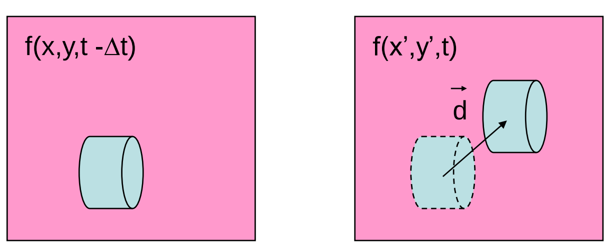
- Motion pictures
- Motion field and optical flow
- Motion models
- Motion estimation

EBU723U

- 7 -

Definitions

- Displacement



- Motion

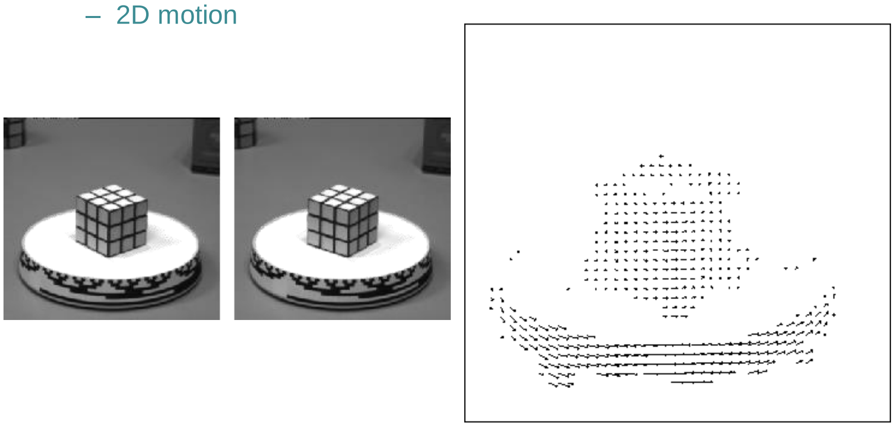
$$\vec{v} = \frac{\vec{d}}{\Delta t}$$

EBU723U

- 8 -

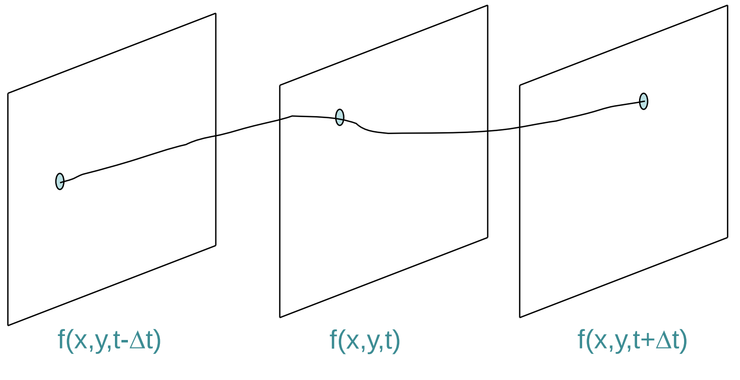
Definition: motion field

- Motion field
  - Projection of the 3D motion of the objects in a scene onto the 2D image plane
  - 2D motion

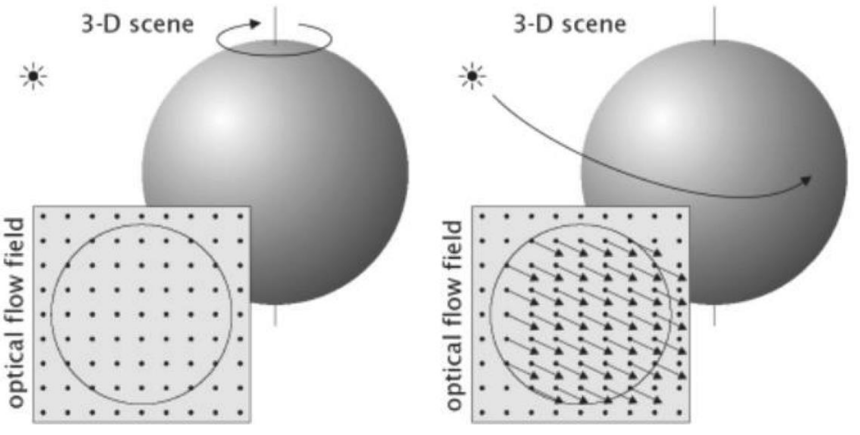


Definition: optical flow

- Optical flow
  - Apparent motion
  - Note: variations in illumination influence the optical flow!



Brightness constancy



ELEM023

- 11 -

Agenda

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

EBU723U

- 12 -

Motion models

- Ideally
  - 1 motion vector for each pixel
  - Too complex
- Semantically
  - 1 motion vector for each object
  - Second generation techniques
- In practice ...
  - 1 motion vector for each block of pixels
  - Size of the block: 8 x 8 or 16 x 16 pixels

EBU723U

- 13 -

Motion models

- Affine motion
  - 9 parameters: translation, rotation, zoom in/out
  - 6 parameters in the 2D case
  - Complex
- Translational motion
  - 2 parameters: horizontal and vertical translations

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

EBU723U

- 14 -

2-D Affine Motion Model

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

- The affine model describes the vector at each point in the image
- Need to find values for the parameters that best fit the motion present

- 15 -

Agenda

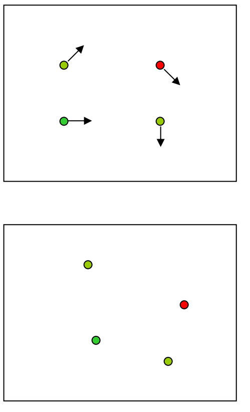
- Motion pictures
- Motion field and optical flow
- Motion models
- Motion estimation

EBU723U

- 16 -

Motion estimation

- Hypotheses
  - No occlusions
  - Rigid objects
  - No illumination changes
  - Locally
    - continuity of the motion
    - translational motion
- Methods
  - Gradient
  - Block matching
  - Pel-recursive
  - Phase correlation



H(*x*,*y*)

I(*x*,*y*)

EBU723U

- 17 -

Gradient method

- Brightness consistency constraint

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

- small motion: (Δ*x* and Δ*y* are less than 1 pixel)
  - suppose we take the Taylor series expansion of *I*:

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

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

EBU723U

- 18 -

Gradient method

- Spatio-temporal constraint

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

- This equation introduces one constraint only
  - Where the motion vector of a pixel has 2 components (parameters)
  - A second constraints is necessary to solve the system

EBU723U

- 19 -

Aperture problem

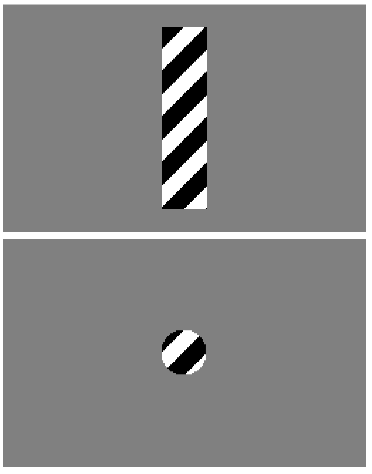
- The aperture problem
  - stems from the need to solve one equation with two unknowns, which are the two components of optical flow
  - it is not possible to estimate both components of the optical flow from the local spatial and temporal derivatives
- By applying a constraint
  - the optical flow field changes smoothly in a small neighborhood it is possible to estimate both components of the optical flow if the spatial and temporal derivatives of the image intensity are available

EBU723U

- 20 -

Aperture problem

- The aperture problem - the barberpole illusion



EBU723U

- 21 -

Solving the aperture problem

- How to get more equations for a pixel?
- By applying a constraint
  - the optical flow field changes smoothly in a small neighborhood it is possible to estimate both components of the optical flow if the spatial and temporal derivatives of the image intensity are available
- Lucas–Kanade method

EBU723U

- 22 -

Gradient method

- The Lucas–Kanade method assumes that the displacement of the image contents between two nearby instants (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

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

EBU723U

- 23 -

Gradient method

- Prob: we have more equations than unknowns

$$Av = b \longrightarrow \text{minimize } \|Av - b\|^2$$

Solution: solve least squares problem

- minimum least squares solution given by solution of:

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

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

- The summations are over all *n* pixels in the K x K window
- This technique was first proposed by Lukas & Kanade (1981)

EBU723U

- 24 -

Lucas-Kanade flow

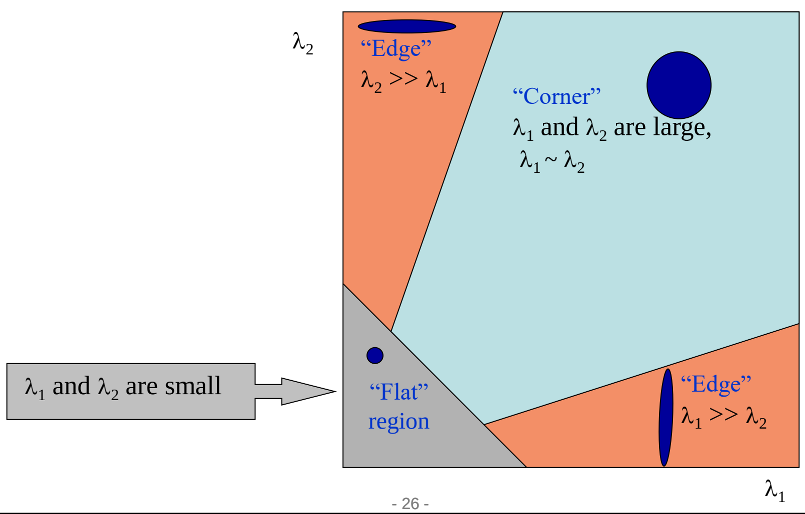
$$\underbrace{\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} V_x \\ V_y \end{bmatrix} = - \underbrace{\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}}_{A^T b}$$

- Recall the Harris corner detector: $M = A^T A$ is the *second moment matrix*
- When is this solvable?
  - **$A^T A$** 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 (λ_1 = larger eigenvalue)

- 25 -

Interpreting the eigenvalues

Classification of image points using eigenvalues of the second moment matrix:



- 26 -

Uniform region

– gradients have small magnitude
– small λ_1, small λ_2
– system is ill-conditioned

- 27 -

Edge

– gradients have one dominant direction
– large λ_1, small λ_2
– system is ill-conditioned

- 28 -

High-texture or corner region

– gradients have different directions, large magnitudes
– large λ_1, large λ_2
– system is well-conditioned

- 29 -

Errors in Lucas-Kanade

- What are the potential causes of errors in this procedure?
  - Suppose A^TA is easily invertible
  - Suppose there is not much noise in the image
- When our assumptions are violated
  - Brightness constancy is **not** satisfied
  - The motion is **not** small
  - A point does **not** move like its neighbors
    - window size is too large
    - what is the ideal window size?

- 30 -

Iterative Refinement

Iterative Lukas-Kanade Algorithm

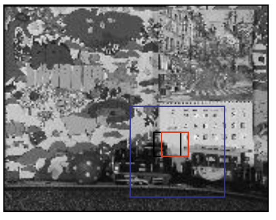
1. Estimate velocity at each pixel by solving Lucas-Kanade equations
2. Warp H towards I using the estimated flow field
   - use image warping techniques
3. Repeat until convergence

Gradient methods

- Advantages
  - Dense motion field
  - Appropriate for image sequence analysis
- Disadvantages
  - The additional constraint increases the energy of the error
  - The dense motion field is represented with a large amount of data

Block matching

- Image in the sequence
  - divided into small blocks
  - for each block
    - a similar block is searched for in the previous image
    - the relationship between the 2 blocks generates the motion vector for the block under analysis

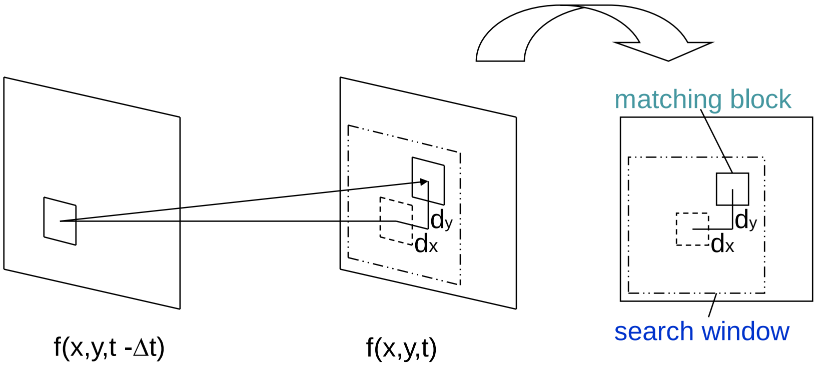


f(x,y,t -Δt) f(x,y,t)

EBU723U

- 33 -

Block matching



EBU723U

- 34 -

Block matching

- Similarity measure
  - Cross Correlation Function
  - Pel Difference Classification
  - Sum of Absolute Difference / Error

$$SAD = \sum_{B} \left\| f(x,y,t) - f(x-d_x, y-d_y, t-\Delta t) \right\|$$

  - Mean Squared Difference / Error

$$MSE = \frac{1}{B}\sum_{B}\left(f(x,y,t) - f(x-d_x, y-d_y, t-\Delta t)\right)^2$$

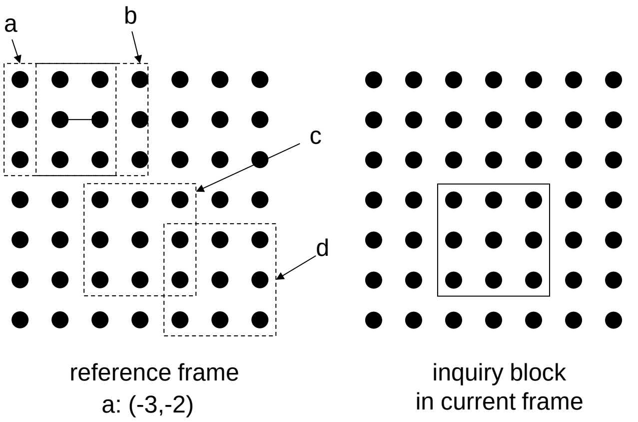
  - Integral Projection

EBU723U

- 35 -

Block matching

- Search algorithms
  - full-search
  - logarithmic
  - n-step
  - conjugate
  - pyramidal
  - multigrid



a: (-3,-2)
b: (-3,-1)
c: (0,0)
d: (1,2)

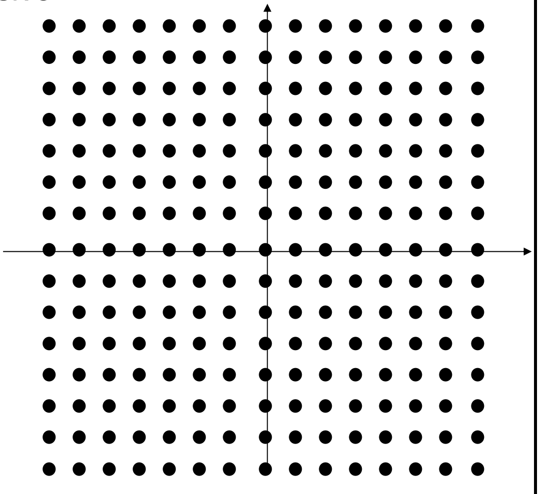
EBU723U

- 36 -

Full search

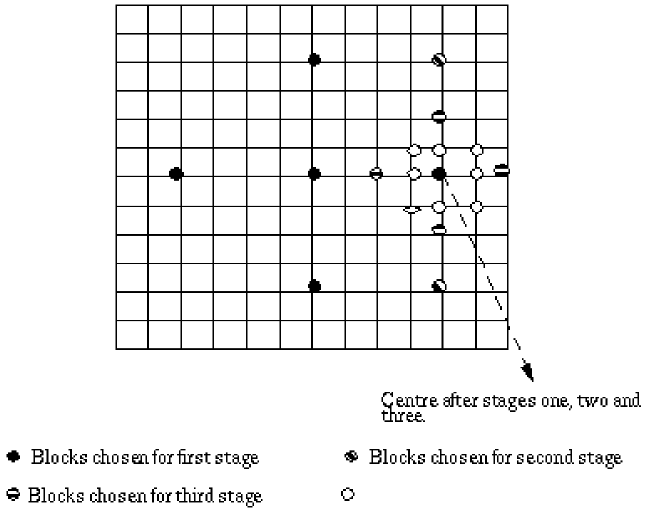
- Computationally expensive
- Optimal solution

An example of range ± 7



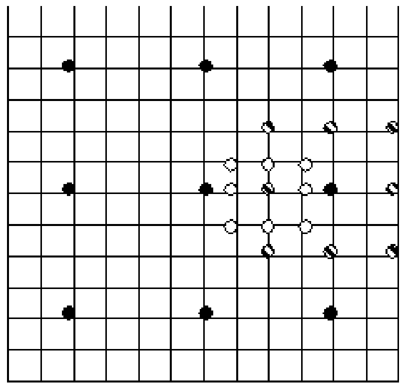
It searches $(2x7+1)^2$=225 points in total

Logarithmic search



3-step search

- Repeat algorithm 3 times
- Examines 25 points
- Assumes a uniform distribution of MV's

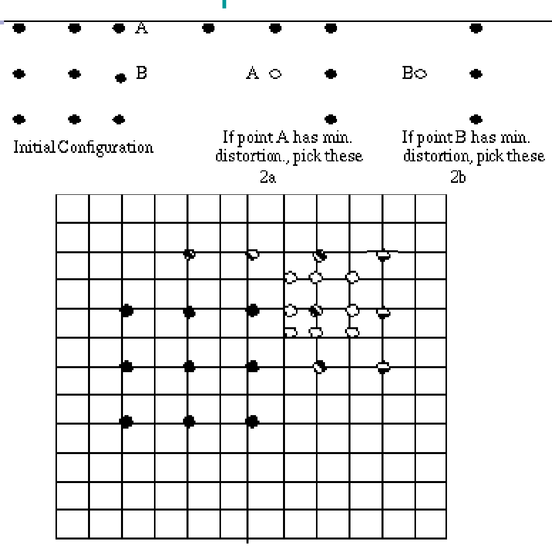


● Blocks chosen for the first stage
◐ Blocks chosen for the second stage
○ Blocks chosen for the third stage

EBU723U

- 39 -

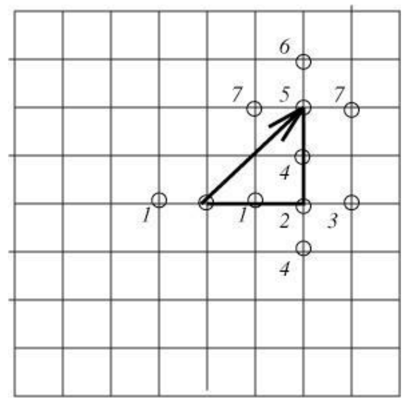
4-step search



● Initial Set of points
◐ Points for the second stage
⊖ Points for the third stage
○ Final set of points

EBU723U

Conjugate search

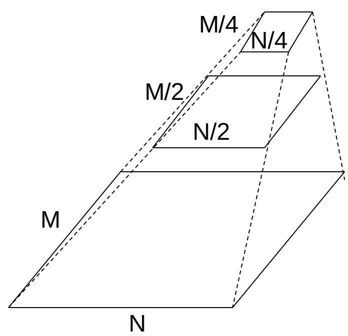


EBU723U

- 41 -

Pyramidal search

Creation of pyramid:
- mean intensity
- subsampling

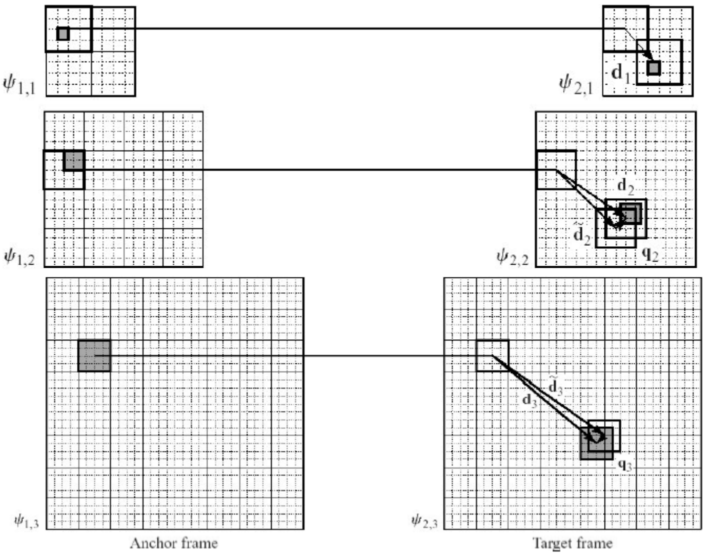


Multi-resolution representation by pyramid

EBU723U

- 42 -

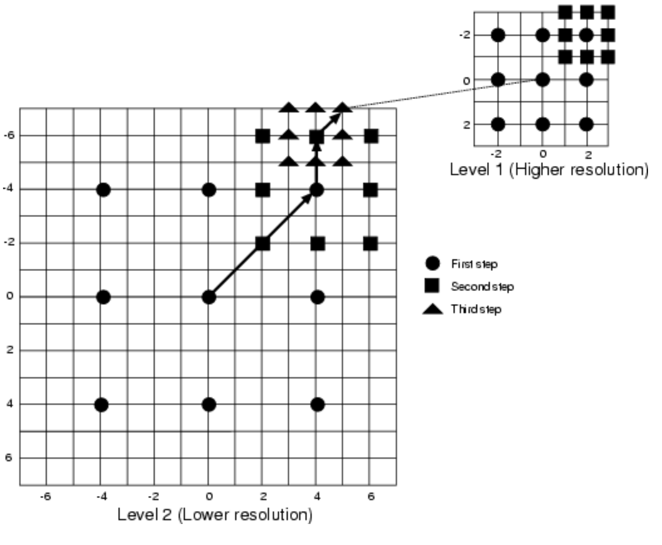
Pyramidal search



EBU723U

- 43 -

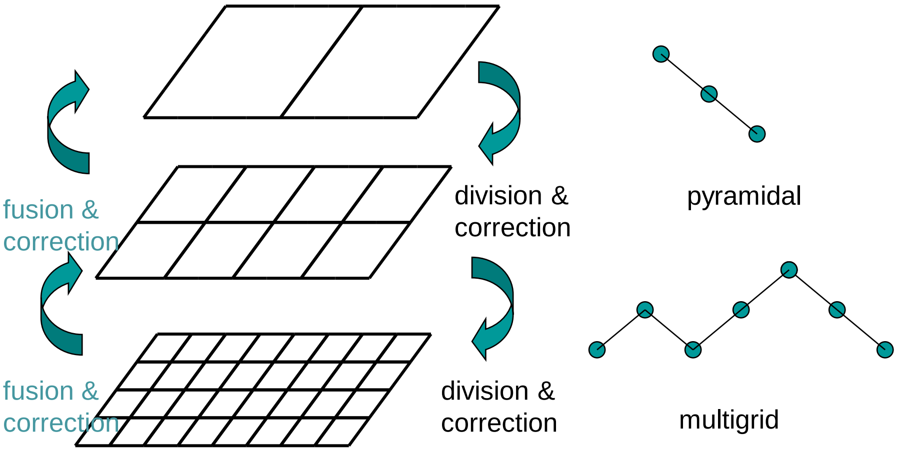
Fast BMA : Hierarchical Search



ELEM023

- 44 -

Multigrid search



EBU723U

- 45 -

Block matching methods

- Advantages
  - Directly minimise the prediction error
  - Regular structure (especially using full-search)
- Disadvantages
  - Complexity
  - Large prediction error at the edges of moving objects

EBU723U

- 46 -

What did we learn?

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

EBU723U

- 47 -