Motion Estimation

Presented by

Canan Bicer

Andreas Pancenko

Outline

- Motion
- Interaction Techniques
  - Sweep
  - PlayAnywhere
  - Mozzies and more
- High-Level Approaches
  - Block-matching: *full-search, 3SS*
  - Optical flow: *differential technique*
- Conclusion

Motion

- causes of image motion:
  - a moving object in the scene
  - eye movements
  - motion of the observer

- possibilities:
  - camera still, moving scene
  - moving camera, still scene
  - moving camera, moving scene

Motion Estimation

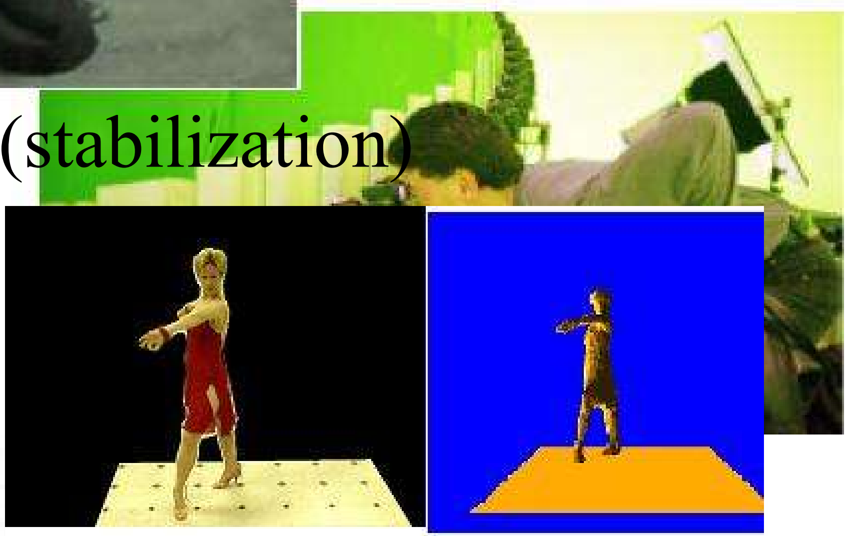
- assumptions:
  - rigid body objects move in translational movements
  - uniform illumination in time and space
  - no occlusion and uncovered background

Why estimate motion?

lots of uses:

- Track object behavior
- Motion Detection
- Video Compression
- Correct for camera jitter (stabilization)
- Align images (mosaics)
- 3D shape reconstruction
- Special effects

use optical flow to compute correspondence between different camera views



Motion Estimation

- What are the movement objects in the scene?
- What sort of motion are they undergoing?
- Where will they be in the future?

→ measure the motion

Mobile Phones

- abilities of a small computer
- mobile phones are ubiquitous in our everyday life
- part of our life

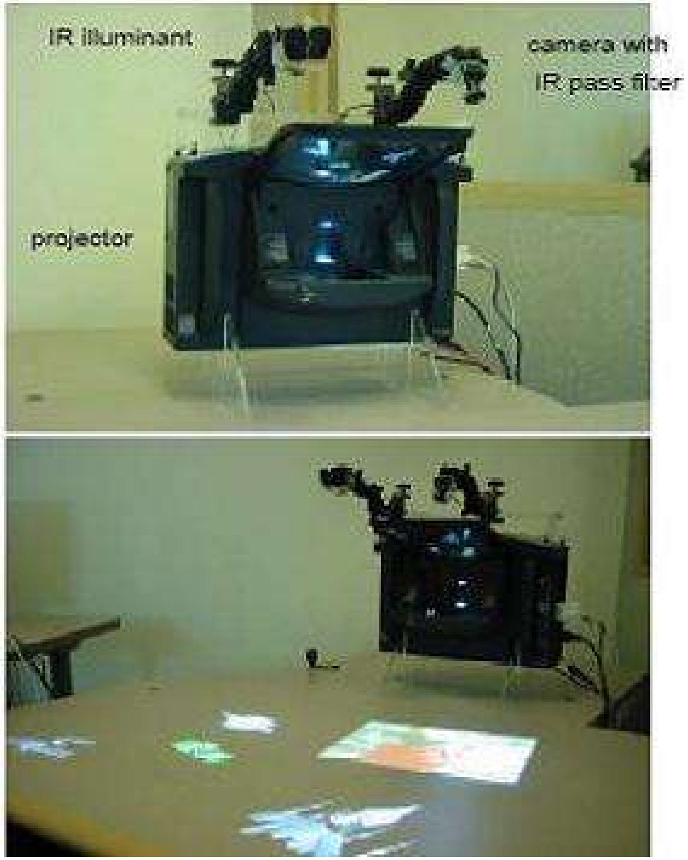
Sweep

- controlling of a cursor on a public display via mobile phone (like optical mouse)
- user is concentrated on the large display
- calculation of motion in the mobile phone
- currently still delay of 200ms

Sweep

PlayAnywhere

- without the need for specially mounted cameras.
- does not require calibration
- allows user to use hand to move projected virtual objects
- aim: no keyboard or mouse



Mozzies and *Attack of the killer virus*

- the first game based on using phone camera
- the purpose is to shoot the mosquitoes or viruses

Mozzies and *Attack of the killer virus*

- the games use real-time images from the phone camera to target the insects
- Mozzies uses the block matching

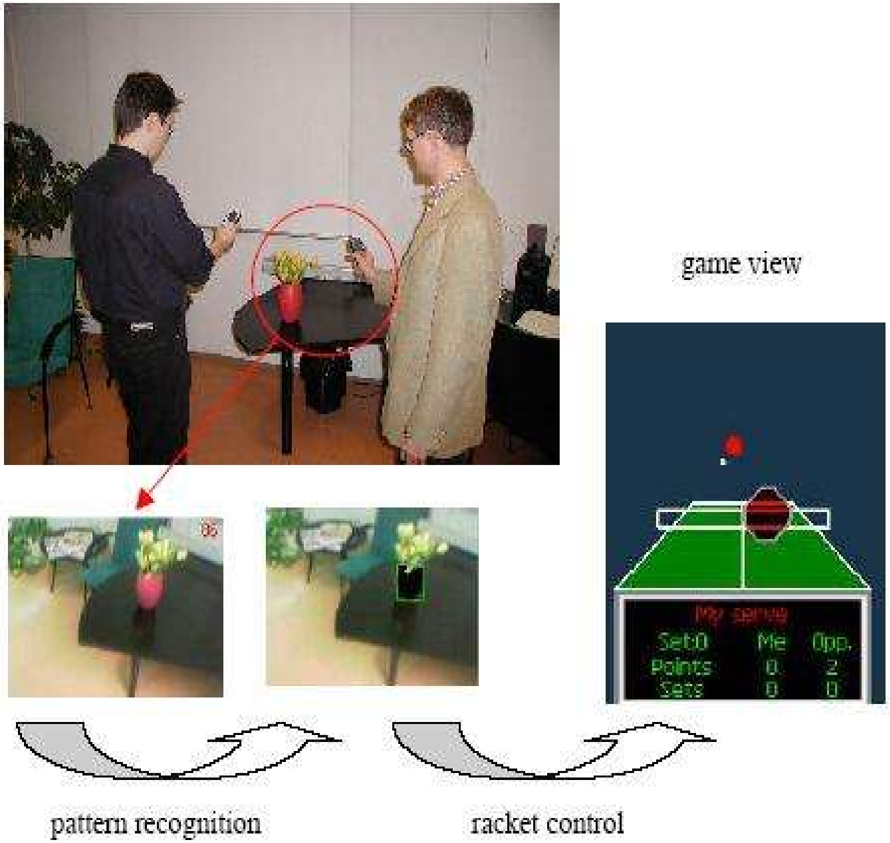
AR-Soccer

- player has to kick a virtual ball
- player`s foot is real
- the ball and the goal with goalkeeper are virtual

- the game uses the camera to detect the motion and position of the player`s foot
- the algorithm operates only in region of interest (ROI) around the interaction object (e.g. the ball)

SymBall

- virtual table tennis game
- mobile phones are used as rackets
- the users may play against each other via Bluetooth connection
- position of the virtual racket is controlled by feature detection



Disadvantages of such mobile device applications

- the user must have enough place round himself
- the display will be moved with mobile device
- communication channel can be slow

Main approaches for motion estimation

- Block Matching Algorithm
- Optical Flow

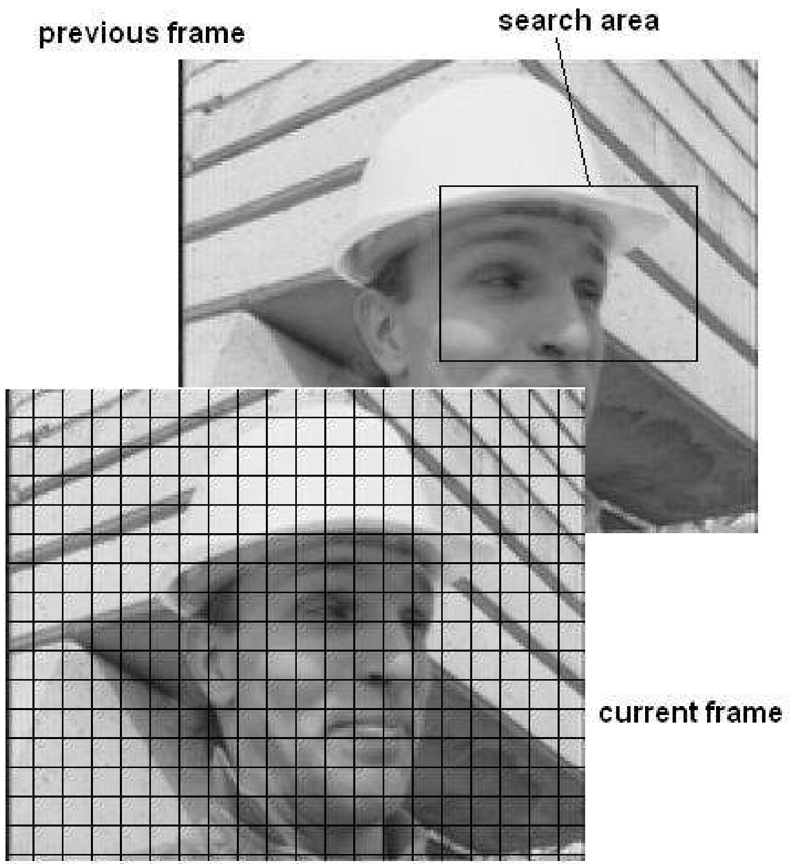
Block Matching Algorithm

- the basic principle of the BMA is the pattern comparison

- the comparison happens block by block

How Block Matching works

- the image is divided in the square blocks
- the search area will be determined
- the intensity of a block is used as a comparative criterion



How Block Matching works

- the actual block goes through the searching area
- the point with the greatest possible correspondence will be determined

Criteria for similarity

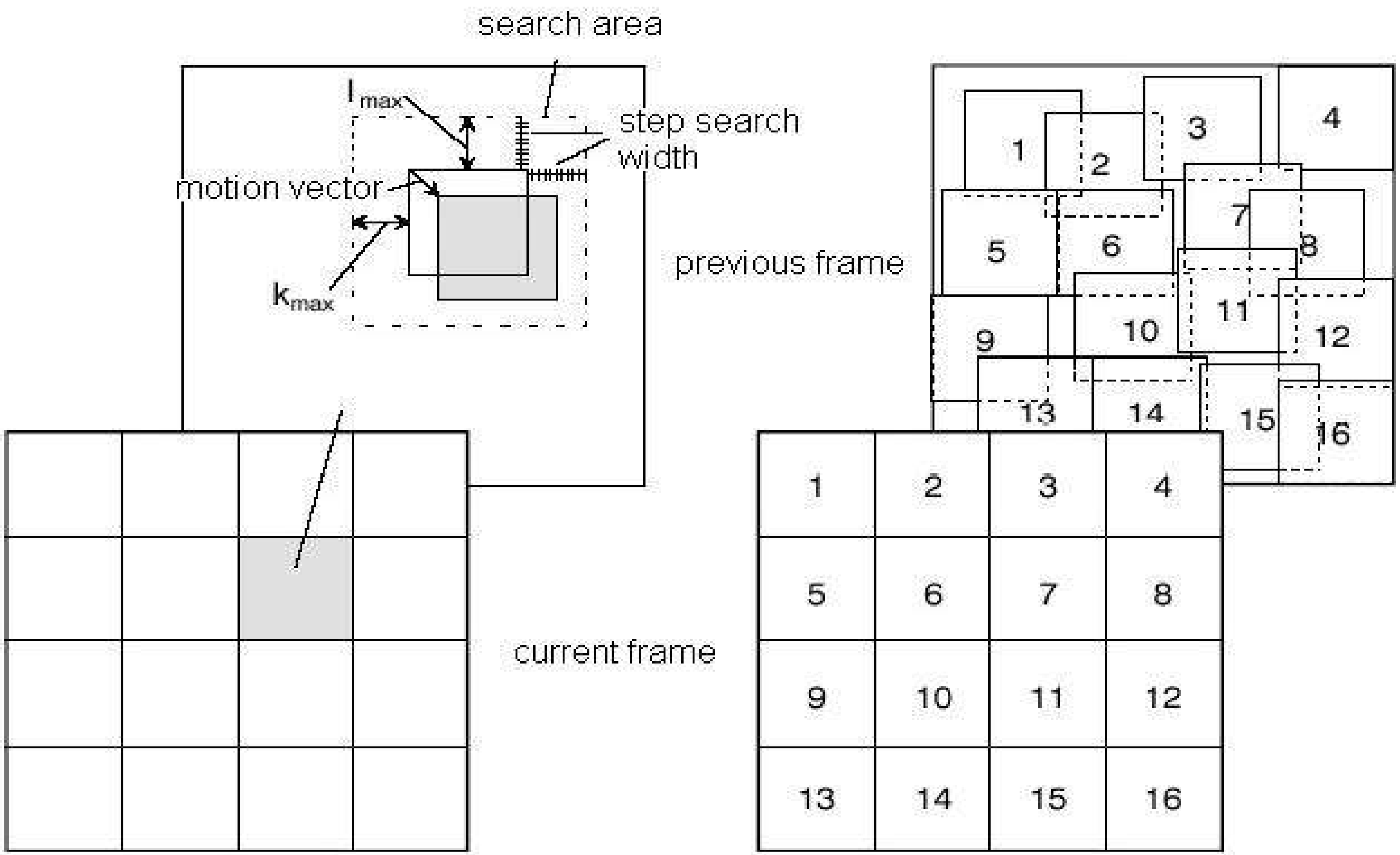
$$SAD(u,v) = \sum_{j=0}^{B-1}\sum_{i=0}^{B-1} |x_t(i,j) - x_{t-1}(i+u, j+v)|$$

$$MAD(u,v) = \frac{1}{B \cdot B}\sum_{j=0}^{B-1}\sum_{i=0}^{B-1} |x_t(i,j) - x_{t-1}(i+u, j+v)|$$

$$SSD(u,v) = \sum_{j=0}^{B-1}\sum_{i=0}^{B-1} (x_t(i,j) - x_{t-1}(i+u, j+v))^2$$

$$MSE(u,v) = \frac{1}{B \cdot B}\sum_{j=0}^{B-1}\sum_{i=0}^{B-1} (x_t(i,j) - x_{t-1}(i+u, j+v))^2$$

How Block Matching works



Full search

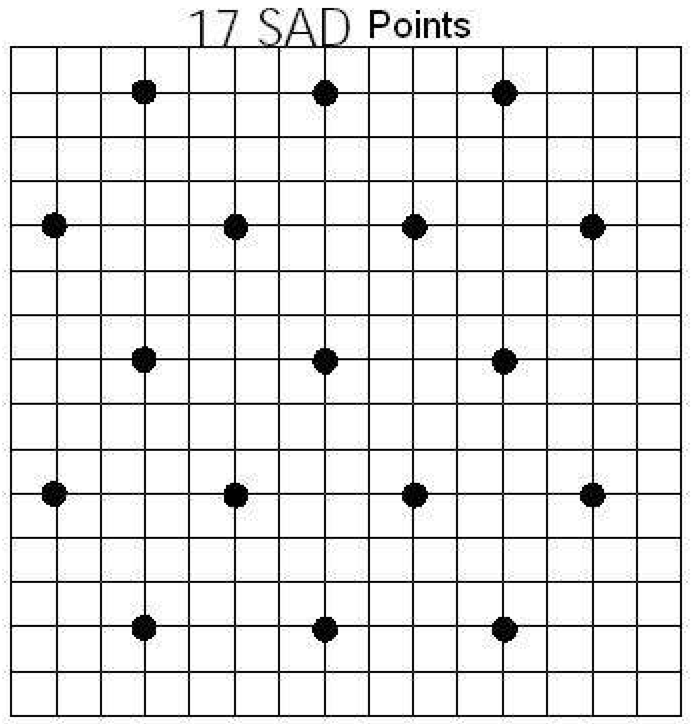
+ the simplest block matching algorithm
+ simple to implement
+ FS finds the absolute minimum always

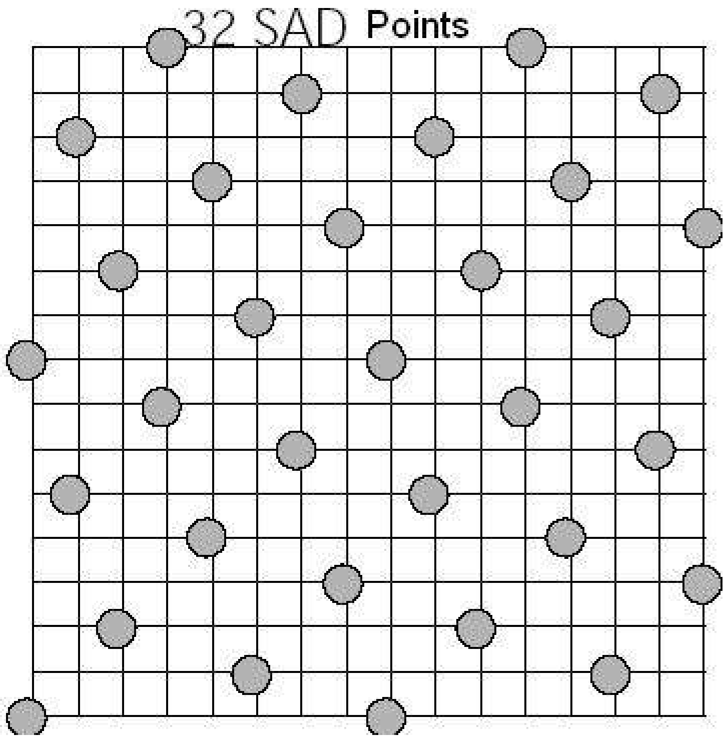
- FS is most compute-intensive solution

How to optimize the BMA

- Reducing the calculation of the similarity criterion
- Reducing the search area
- Increasing of block size

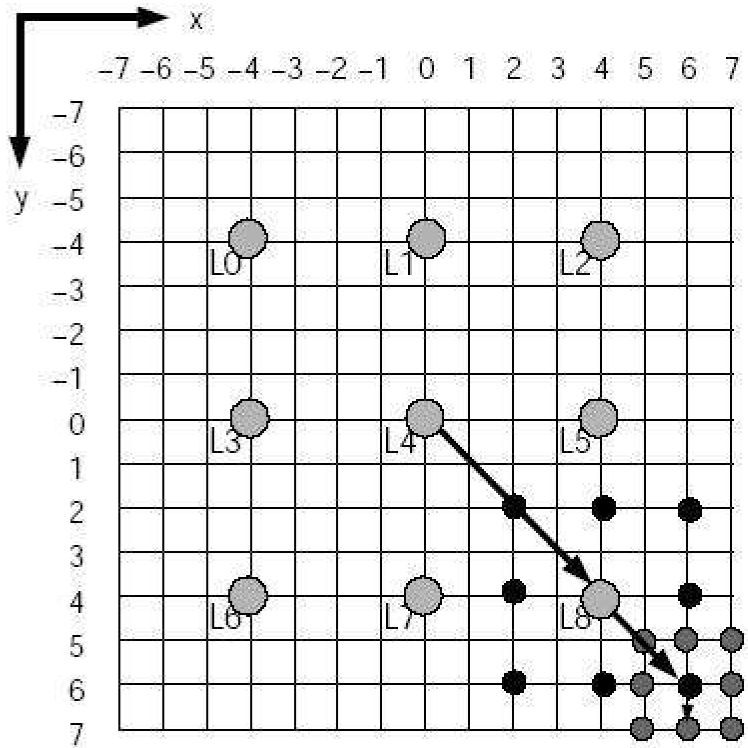
Reducing the calculation of the similarity criterion





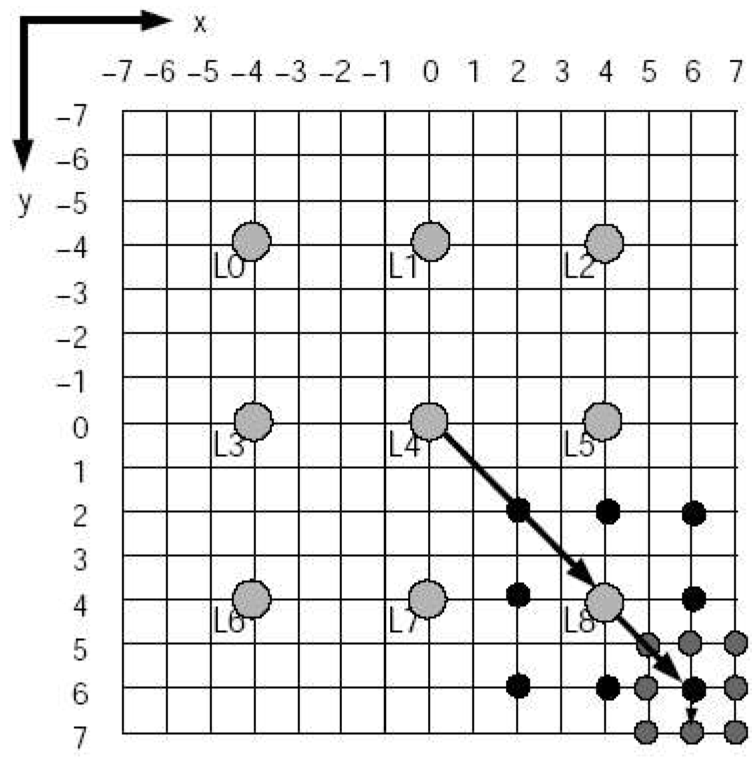
Three-Step-Search algorithm

- step search width defined
- checking of searching area in horizontal, vertical and diagonal direction
- the environment of the point with the best correspondence will be examined further



Three-Step-Search algorithm

- the step search width will be smaller
- TSS ends if the step search width arrive 0
- TSS is fast, but can not always find the absolute minimum



Another methods based on BMA

- Spiral Search
- Binary Search
- 2D-Logarithmic Search

etc.

media computing group
RWTH AACHEN
RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

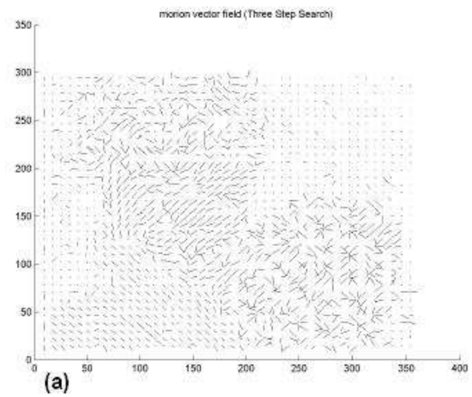
Increasing of block size

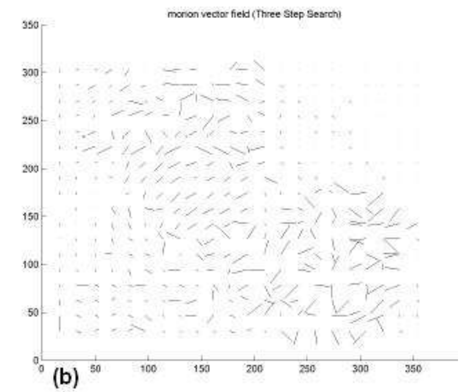
- Increasing of block size

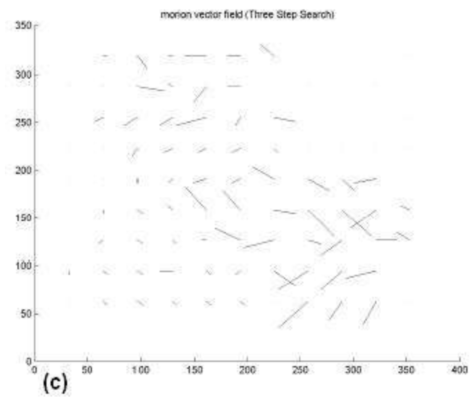
(a) original frame
(b) estimated frame (TSS 4x4)
(c) estimated frame (TSS 8x8)
(d) estimated frame (TSS 16x16)

Increasing of block size

- Vector field for increasing of block size







(a) vector field (TSS 4x4)

(b) vector field (TSS 8x8)

(c) vector field (TSS 16x16)

Results

- Slowly Motion

(a) original frame (Full Search)

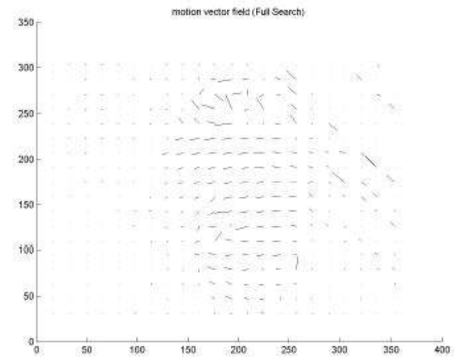
(b) estimated frame (Full Search)

(a) original frame (Three Step)

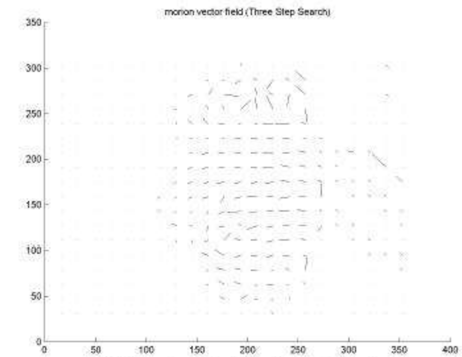
(b) estimated frame (Three Step)

Results

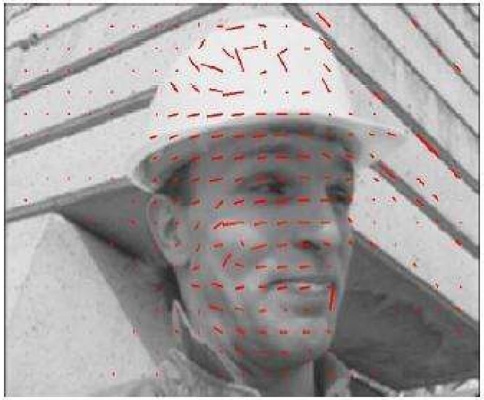
- Vector field for slowly motion



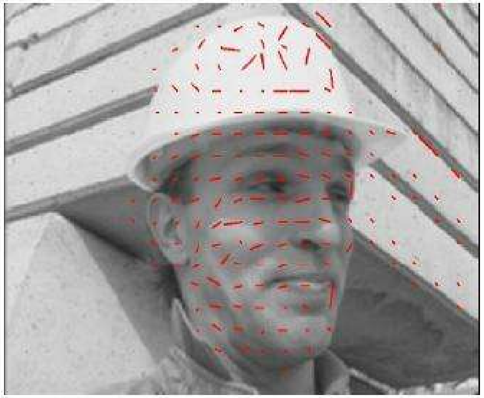
(a) vector field (Full Search)



(b) vector field (Three Step)



(a) vector field (Full Search)



(b) vector field (Three Step)

Results

- Fast motion

(a) original frame (Full Search)

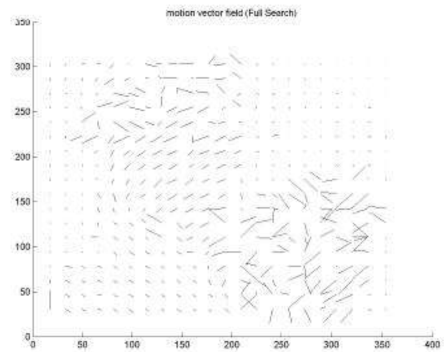
(b) estimated frame (Full Search)

(a) original frame (Three Step)

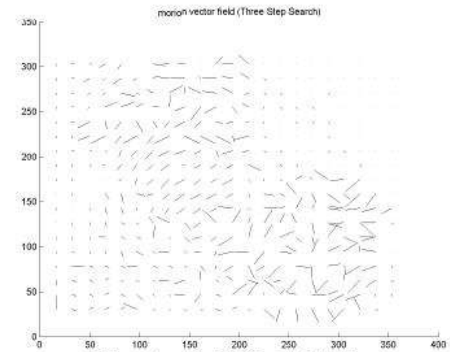
(b) estimated frame (Three Step)

Results

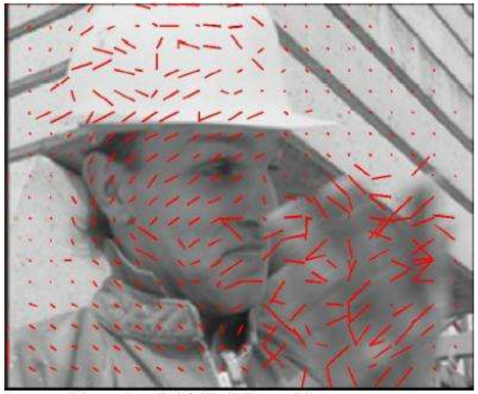
- Fast motion



(a) vector field (Full Search)



(b) vector field (Zhree Step)


(a) vector field (Full Search)


(b) vector field (Three Step)

What is optical flow?

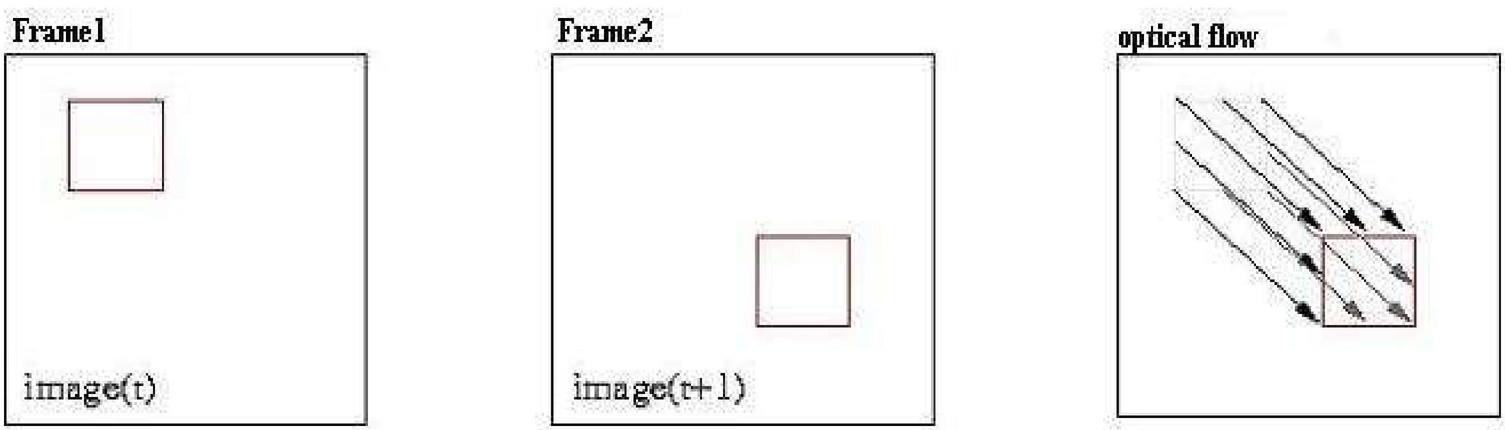
- *Optical flow:* representation of the projections of 3D motion on a sequence of 2D images
- displacement field for each of the pixels in an image sequence
- demonstrates visual variation of brightness pattern in sequenced images

media computing group

RWTH AACHEN RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

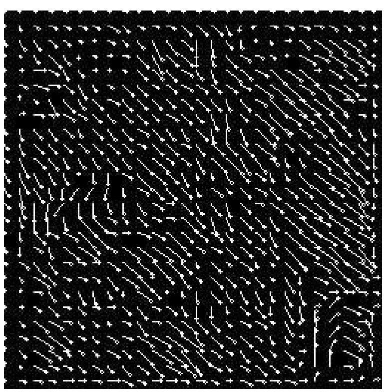
What is optical flow?

- a velocity vector is found for each pixel which says:
  - the velocity of a pixel
  - the direction of its movement

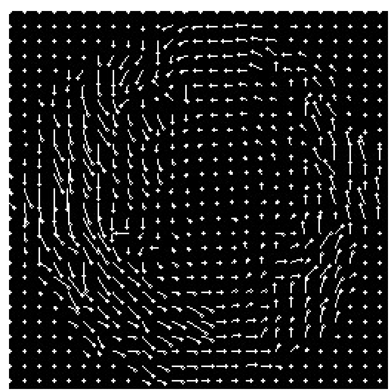


More optical flow examples

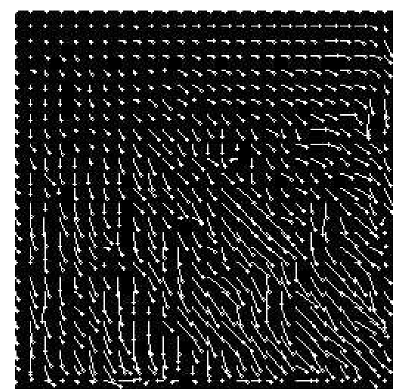
Translation



Rotation



Scaling

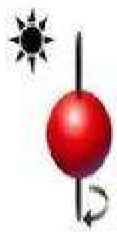


Motion field and optical flow

- Motion field
  - 2D projection of velocity of the image points
  - not directly measurable from the image
- Optical flow
  - best case: will be the same as the motion field, but not always the case
  - measurable from the image

Motion field != Optical flow

- a rotating ball

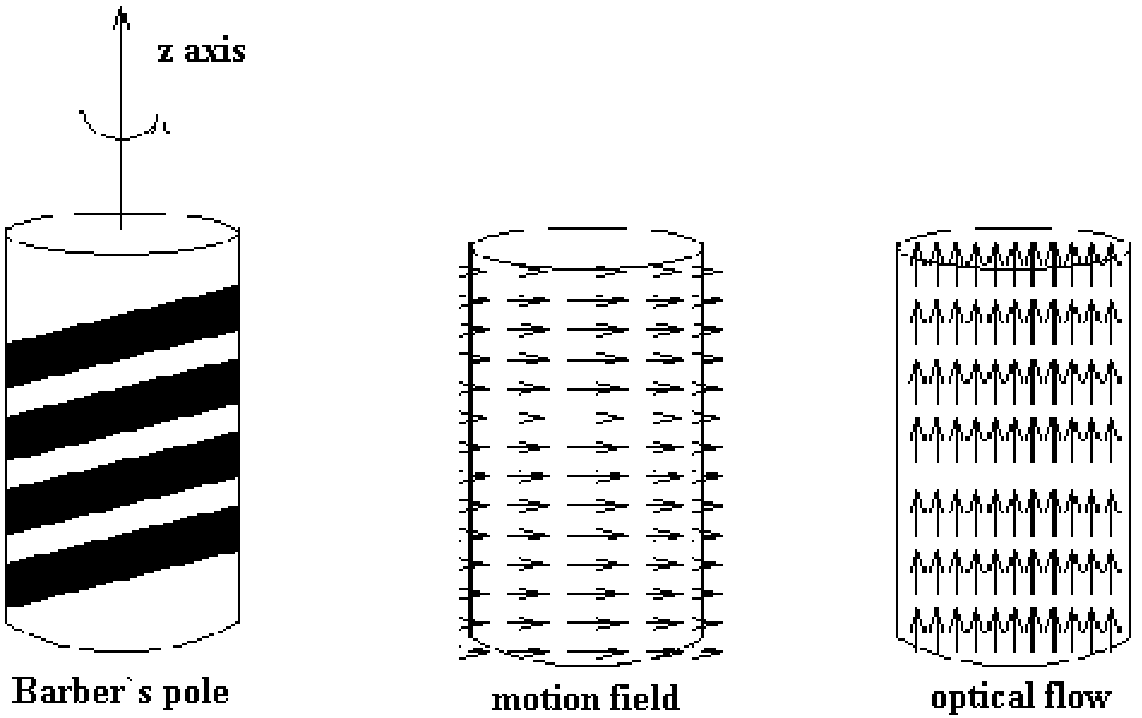


  - no perceptive variation is recognizable
  - a rotation takes place → *motion field*

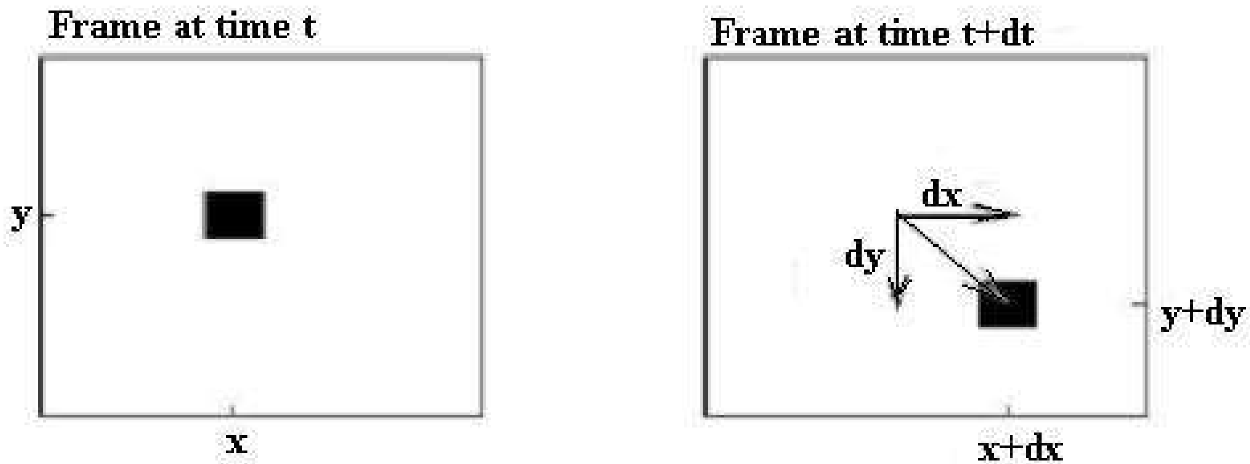
- a moving light

  - ball is illuminated by a moved source of light
  - scene objects do not move
  - image changes → *optical flow*

Motion field != Optical flow

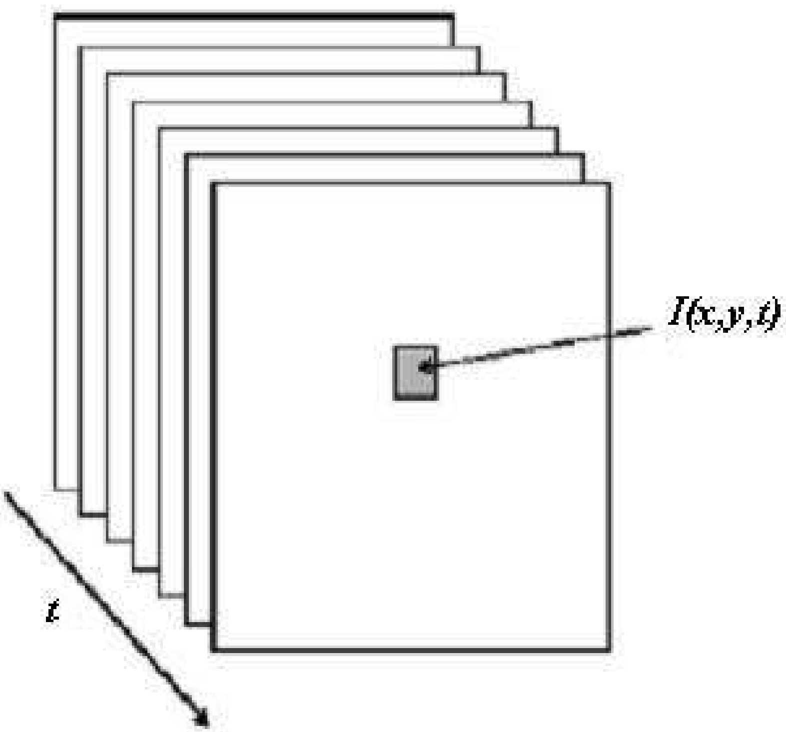


OF – mathematical:



- *Left:* image at the time *t* with the point at the pixel *(x,y)*
- *Right:* image at the point *t+dt* with the velocity vector $(u,v) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right)$

OF-mathematical



- adding a third dimension
- intensity value *I(x,y,t)* of pixel *(x,y)* and time *t*

optical flow is ambiguous

- assumption:
  - for sufficient small time steps a point possesses at the time $t+dt$ the same intensity as at the time t

→ **Brightness constancy constraint:**

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

1st constraint

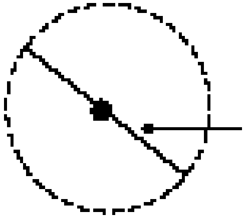
- using Taylor expansion and another transformation you get the **optical flow constraint equation (OFCE):**

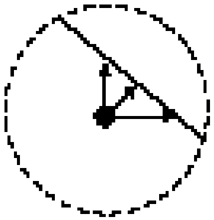
$$\rightarrow \boxed{I_x u + I_y v + I_t = 0}$$

but...

- no solution (1 equation, 2 unknowns)
- local information not completely enough to compute the optical flow *(u,v)* of one point
- can not determine the component of flow orthogonal to it

→ **Aperture problem**





2nd constraint

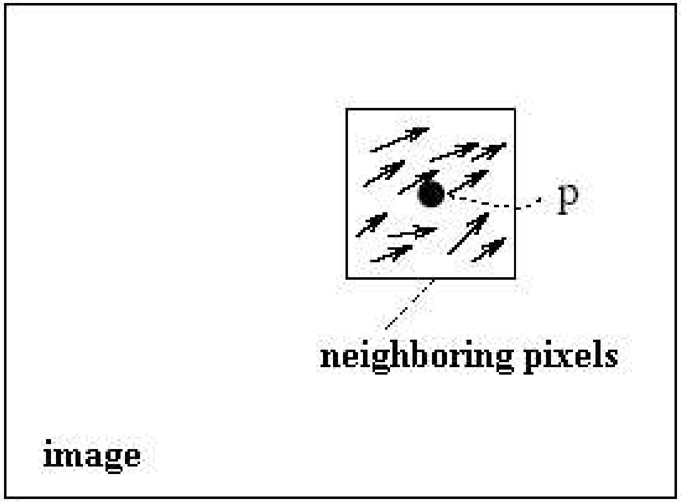
- a further constraint is necessary
- constraint that the flow field is smooth
- environment of a regarded point is differentiable

→ **Smoothness constraint**

- well-known approach: **Differential techniques** (Horn & Schunck)

Differential Techniques (Horn&Schunck)

- based on spatial and temporal variations of the image at all pixels
- neighboring pixels in the image should have similar optical flow



Differential Techniques (Horn&Schunck)

- mathematically:

  sum of the amount of the gradients have to become minimal

$$\rightarrow \quad E_r(u,v) := |\nabla \mathbf{u}|^2 + |\nabla \mathbf{v}|^2$$

- evolving function have to be solved iteratively
- iteration stops if *t* and *t+1* are very similar

Differential Techniques (Horn&Schunck)

- assumptions:
  - brightness constancy
  - neighboring velocities are similar
- features:

  + image first derivatives only

  + incorporate global information

  - iterative

  - smoothness is violated across motion boundaries

	present motion	method	accuracy & efficiency	problems
BM	block-based	exhaustive search	good trade-off between AE	finding of right block-size, search-area
OF	pixel-based	gradient descent	most accurate, not well-suited to real-time processing	only for small motion, most costly to estimate

Thank you for listening!

Merry X-mas
and
Happy New Year!