Introduction to motion estimation
Definitions

- **Motion estimation** is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence.

- 3D scene is a projection onto a 2D scene.
Definitions

- Successive video frames may contain the same objects (still or moving).
- Motion estimation examines the movement of objects in an image sequence to try to obtain vectors representing the estimated motion.
Definitions (II)

- The motion vectors may relate to the whole image (global motion estimation) or specific parts, such as rectangular blocks, arbitrary shaped patches or even per pixel.
- In motion estimation an exact 1:1 correspondence of pixel positions is not a requirement.
Definitions (III)

- **Motion compensation**: describes a picture in terms of the transformation of a reference picture to the current picture. The reference picture may be previous in time or even from the future.
- It is the action to apply the motion vectors to an image to synthesize the transformation to the next image.
Definitions (IV)

- **Optical flow or optic flow**: is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (an eye or a camera) and the scene.
**Motion Estimation**

- **Direct Methods:**
  - **Block-matching algorithm**
  - Phase correlation and frequency domain methods
  - Pixel recursive algorithms
  - MAP/MRF type "Bayesian" estimators
  - Optical flow
Motion Estimation (II)

- Indirect Methods:
  - *Indirect methods* use features, such as Harris corners, and match corresponding features between frames, usually with a statistical function applied over a local or global area. The purpose of the statistical function is to remove matches that do not correspond to the actual motion.
  - Statistical functions that have been successfully used include RANSAC
Block-matching algorithm

- BMA: is a way of locating matching blocks in a sequence of digital video frames for the purposes of motion estimation.
- Goal: it is to find a matching block from a frame ’i’ in some other frame ’j’, which may appear before or after ’i’. This can be used to discover temporal redundancy in the video sequence, increasing the effectiveness of interframe video compression.
Block-matching algorithm

**Activity**: Find at least 7 BMA and make a short description about 3 of them. (30 minutes)

**Link**: http://www.mathworks.com/matlabcentral/fileexchange/8761
Optical Flow

• **Reference:**

• **Activity:**
  
  ◦ Briefly define what is optical flow motion estimation method
  
  ◦ Describe, without mathematical formula if it is possible, three optical flow techniques (30 minutes)
Pyramid methods in image processing

- Reference:

- Pyramids are an example of a multi-resolution representation of the image.
- Pyramids separate information into frequency bands.
- In the case of images, we can represent high frequency information (textures, etc.) in a finely sampled grid.
- Coarse information can be represented in a coarser grid (lower sampling rate acceptable).
- Thus, coarse features can be detected in the coarse grid using a small template size.
- This is often referred to as a multi-resolution or multi-scale resolution.
- The goal is to reduce the time processing for looking template in an image.
Pyramid methods (II)

- Pyramid correspond to a mixture of the spatial and frequency domains
- Position information at each level is known to the accuracy of that grid resolution
- Contributing frequencies are bandwidth limited at each grid resolution
Pyramid methods (III)

Original Image (Level 0)

The Number of levels is a function of the application (3-5 is typical)

Decimation (Level 1)

Decimation (Level 2)

Decimation (Level 3)
Pyramid methods (IV)

- For each decimation, we convolve with the same size Gaussian kernel, (low pass filter matrix)
- While the width of this filter stays fixed, its effective width actually doubles at each pyramid level

![Gaussian kernel matrix](image)

![Gaussian kernel curves](image)

- Level 1
- Level 2
- Level 3
Pyramid methods (V)

- As a consequence, only lower frequencies are passed through (low-pass filtering)
Pyramid methods

- Pyramid Level at same resolution
Pyramid methods

- Pyramids correspond to a decomposition of an image into spatial/frequency bands
- Higher frequency require a larger image (sample) size to be represented, while lower frequency can be accommodate in a coarse image
- This “coarse to fine” approach can provide tremendous increases in computational efficiency
- Laplacian pyramids correspond to a band pass while Gaussian a low pass representation
- Pyramids are used in many applications beyond target tracking and image fusion
Hierarchical estimation techniques
1. Motion Estimation: Introduction

- Temporal variations in a sequence are mainly due to movement.
- The sources of the movement may include:
  - The objects in the image
  - Camera Movement:
    - Rotation
    - Translating
    - Approaches
1.2 What is the estimation?

- Identification of movement between two frames $\rightarrow$ motion vectors
- The motion estimation is useful for:
  - Video Compression:
    - Eliminates the temporal redundancy $\rightarrow$ less processing time
  - Video Processing
    - Noise suppression
2. Movement models

- Full image: no Division.

- Pixel to pixel: movement = offset + speed + acceleration
2. Movement models

- Blocks: Division in rectangular blocks

- Regions: general Division, grouping pixel to pixel of the same surface
3. Hierarchical estimation techniques

- Two different parts:
  - Calculation of the resolutions of images
  - Motion Estimation.
3.1 Calculation of the resolutions of images

- The method does not determine a fixed number of levels, it will depend on:
  - The image size.
  - The image quality.
  - The resulting compression quality.
  - Resulting compression level.
  - Application to work.

- 1st: Resolution pyramidal representation.
3.2 Motion Estimation

- Determine movement between a frame and its precedent vectors
  - The estimate is made possible by the Division of each frame into various smaller blocks
  - Previous Frame: divided into different regions search
  - Each region search: divided into different blocks
  - Each region: divided into different areas
- These hierarchies range from the highest to the lowest
- Need to calculate these motion vectors
3.2.1 A motion vector calculation

- Example:
  1. We start at level 2 (lower resolution image) and look for the most similar block to the large one in the following image with a size 7 search region. The search can be done in $N$ steps. Suppose the optimal shift is 6 in the direction of the X and 5 in the of the Y.
3.2.1 A motion vector calculation

2. For the ‘t’ image we want to find the small square. The t+1 image focus on the position of the big square that we have found the small square and allow it to move up to 3 units in each direction.

3. Suppose the closest block is obtained moving X to left three units and three up in Y. The X 6-3 = 3 and in the Y, 5 + 3 = 8.
3.2.2 Spread of motion vector

- As the motion vectors are calculated, they are extrapolated to a resolution image better.
- This technique makes an initial motion estimation for low resolution images.
- Initial vectors are propagated to higher levels and finally carried out a refinement of the propagated vectors.
- These hierarchies range from the highest to the lowest.
4. Advantages

- The advantages of hierarchical method:
  - Low computational complexity
  - High efficiency and flexibility: especificables parameter range (block size, number of levels, scaling factor...)
  - Reduces processing time
  - High spatial accuracy and quality of results
5. Applications

- Video codecs
- Cartography: movements of Earth, progress on the coast, channel of the rivers, etc.
- Create panoramic views
- Medicine: clinical disease progression
- Traffic management
7. Conclusions

- One of the simplest methods algorithmically.
- Good relationship quality-time
- Very good choice for videos with simple movements and still camera
- This is the most used video compression method (MPG2).
- Greater movement of camera and object, need more levels to maintain quality.
Activity I

- Describe applications (3) based on image processing focus on (15min)
  - Intelligent Transportation Systems
  - Logistic
  - Traffic control/management
  - Intelligent Vehicles
  - Ambient Intelligent Systems
Image compression
JPEG (Joint Photographic Experts Group) (I)

- Algorithm designed to compress images in 24 bits or grayscale, with loss
- File Format is defined
- It only works with still images
- Sometimes is called JPG
- To decompress the image does not get exactly the same image before compression.
**JPEG**(II)

- The compression rate can be adjusted
- Compression too high -> loss of quality -> very small files
- Compression too low -> quality similar to original -> very large files
- The quality loss accumulates
JPEG allows four modes of operation: hierarchical, sequential (baseline), progressive and lossless.
JPEG\textsuperscript{(IV)}

- Baseline: the simplest, the image is shifted from left to right and top to bottom. With losses.
- Hierarchical coding process is applied to different ranges of resolutions that can be decompressed separately.
- Progressive allows reconstruction of the image in several passes. Loss.
- No losses. Reversible process.
JPEG - Activity I (15min)

- Main Characteristics
- What is the best ratio compression?
- Why use JPEG?
- Compare JPEG and GIF (losses, simple draws, interlacing, transparent, etc)
- How is the JPEG file structure?
MPEG II

- ITU-H.262 and ISO/IEC-13818
- Applications:
  - Digital recording
  - TV recording.
  - Transmission and communications
- Flexible, and it contains a set of rules that apply depending on application
MPEG II. Involved signals

- MPEG-2 encoding generates a serial digital signal, its meaning is not obvious. The decoder can recover the audio and video signals, and auxiliary data too.

- The original signals of a MPEG-2 are baseband, both video and audio. Analog Video (YPbPr) or digital (YCbCr), standard definition or high definition.

- Analog baseband signals need A/D converters in the encoder and expensive equipment.
MPEG II. Applications

I - BROADCAST APPLICATIONS: MULTIC'HANNEL RELEASE

- Are standardized based on MPEG-2: direct broadcast satellite, cable and terrestrial TV.
- Initially expensive decoders.
- Standardized Applications within the DVB (Digital Video Broadcasting)
- Applications MVDS (Multipoint Video Distribution System)
2- VIDEO ON DEMAND AND INTERACTIVE VIDEO STORE

• Servers VOD (Video on demand) must store large amounts of digital video, so you can have quick access for onward transmission

• Using MPEG-2 involves large savings but random access is more complicated
MPEG II. Applications

3- TRANSPORT OVER TELECOMMUNICATIONS DIGITAL NETWORK

- Compressed video with MPEG-2 can be distributed on many different digital tx systems.
- In this way, telecom operators enter into the broadcast business.
- There are DVB standards (ETSI) for interconnection of MPEG-2 signals to PDH and SDH circuits.
3.1-TRANSPORT OVER PACKET SWITCHING NETWORKS (WITH POTENTIAL LOSS OF PACKAGES)

- ATM allows reliable transmission of compressed video using AAL-1 layer.
- There are other packet switched networks without QoS, the compressed data divided into two trains of data: high priority (useful imaging) and low priority (along with the former giving an improved image)
4 - SMALL BROADCAST APPLICATIONS

- Release to business and education.
- The number of decoder per encoder is much smaller, more cost-sensitive.
MPEGII. Applications (VI)

- 5 – ENG DISTRIBUTION (Electronic News Gathering)
- Terrestrial radio links and satellite links (FM)
- Uncompressed digital modulation requires a higher bandwidth than is available for these channels
- The compression allows transmission of video signals with excellent quality in these channels.
MPEG II. Applications (VII)

- 6 - COMPUTER MULTIMEDIA AND VIDEO
- The highest quality MPEG-2 requires higher bit rate and more performance in processing and decoding
- Today's computers already have it and has several decoder cards.
- MPEG-2 in this environment has problems of cost and physical size
MPEG II. Applications (VIII)

7- HDTV

- Thanks to the introduction of flat-screen HDTV
- In Europe Eureka-256 satellite HDTV does not use MPEG-2, but it is very similar
- High cost of high-definition set-top boxes, monitors, and the availability of programming. Sure?
MPEG II – Activity II (30min)

- Look for information about the block diagram for video compression (.avi). Describe it.
- What is MJPEG?