Université Constantine 2
جامعة قسنطينة 2

Artificial Vision

– Course 4 –

Chapter 4 : Movement and Video Analysis

Dr. Benaliouche Houda
Faculté des nouvelles technologies

Houda.benaliouche@univ-constantine2.dz

Université Constantine 2
جامعة قسنطينة 2

Artificial Vision

– Course 4 –

Chapter 4 : Movement and Video Analysis

Dr. Benaliouche Houda

Faculté des nouvelles technologies

Houda.benaliouche@univ-constantine2.dz

Etudiants concernés

Faculté/Institut	Département	Niveau	Spécialité
Nouvelles technologies	/	Master 2	Sciences de Données et Intelligence Artificielle (SDIA)

Summary

Prerequisites

- Mathematical Notions
- Algorithmic Notions

Course Objective

- A look into how machines analyze video and track movement

OUTLINE

- ✓ Definition of video analysis
- ✓ Key concept:
  - motion detection
  - object tracking
  - Motion estimation
- ✓ Techniques in movement and video analysis
  - Frame Differencing
  - Background Subtraction
  - Optical Flow
- ✓ AI approaches for movement analysis
- ✓ applications
- ✓ challenges
- ✓ conclusion

DEFINITION

Video analysis: Extracting actionable insights from video data using computational techniques.

Why movement analysis?

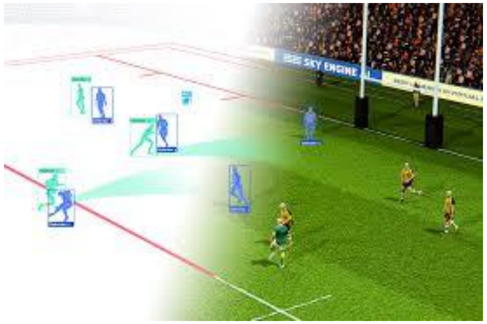
- Understand the dynamics of objects in a scene.
- Enable **automation** in systems like surveillance, sports analytics, and robotics.

Examples:

Motion detection in home security systems like Ring doorbells or Nest cameras.

Sport analytics

robotocs

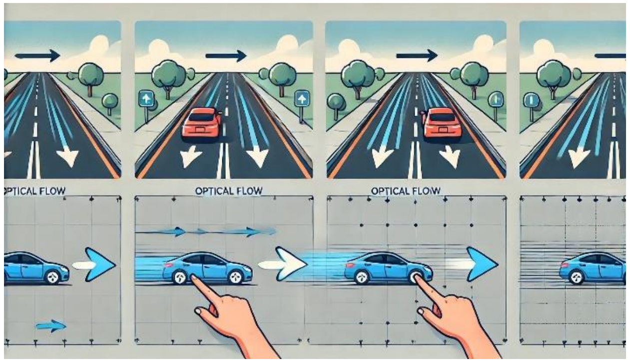


Key concept

- **Motion Detection:**
  - Recognizing areas of motion in video frames.
  - Used in alarm systems or intruder detection.

- **Object Tracking:**
  - Maintaining the identity of an object over time in a video.
  - Example: player tracking, Tracking a car through traffic,.

- **Motion Estimation:**
  - Quantifying the displacement of objects.
  - Example: Optical flow to compute pixel-wise motion.



Key Differences Between Motion Detection and Motion Estimation

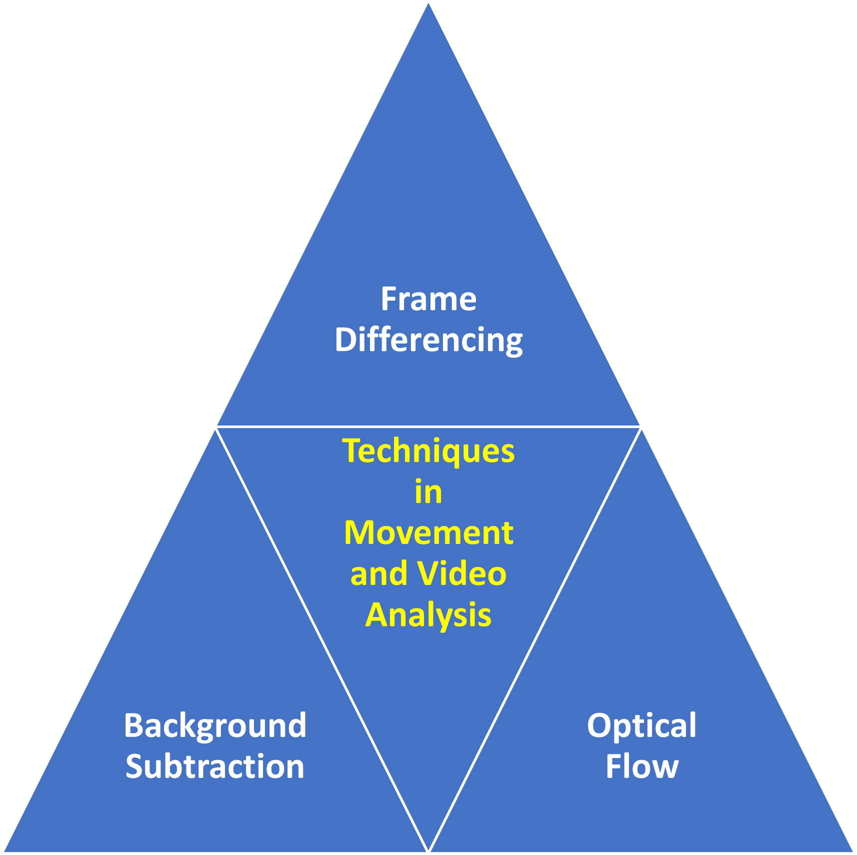
Feature	Motion Detection	Motion Estimation
Focus	Whether motion exists.	Quantifying motion (speed, direction).
Output	Binary mask or bounding boxes.	Motion vectors or optical flow fields.
Techniques	Frame differencing, background subtraction.	Optical flow, block matching.
Complexity	Low computational cost.	High computational cost.
Applications	Security, surveillance, and alarms.	Video tracking, stabilization, and robotics.

Comparison of Motion Detection, Motion Estimation, and Object Tracking

Feature	Motion Detection	Motion Estimation	Object Tracking
Focus	Whether motion exists.	Quantifying motion (speed, direction).	Maintaining identity of objects over time.
Output	Binary mask or bounding boxes.	Motion vectors or optical flow fields.	Trajectory of objects across frames.
Techniques	Frame differencing, background subtraction.	Optical flow, block matching.	Kalman filter, DeepSORT, or tracking-by-detection methods.
Complexity	Low computational cost.	High computational cost.	Varies (moderate to high, depending on occlusions and environment).
Applications	Security, surveillance, and alarms.	Video tracking, stabilization, and robotics.	Sports analytics, traffic monitoring, and video editing.

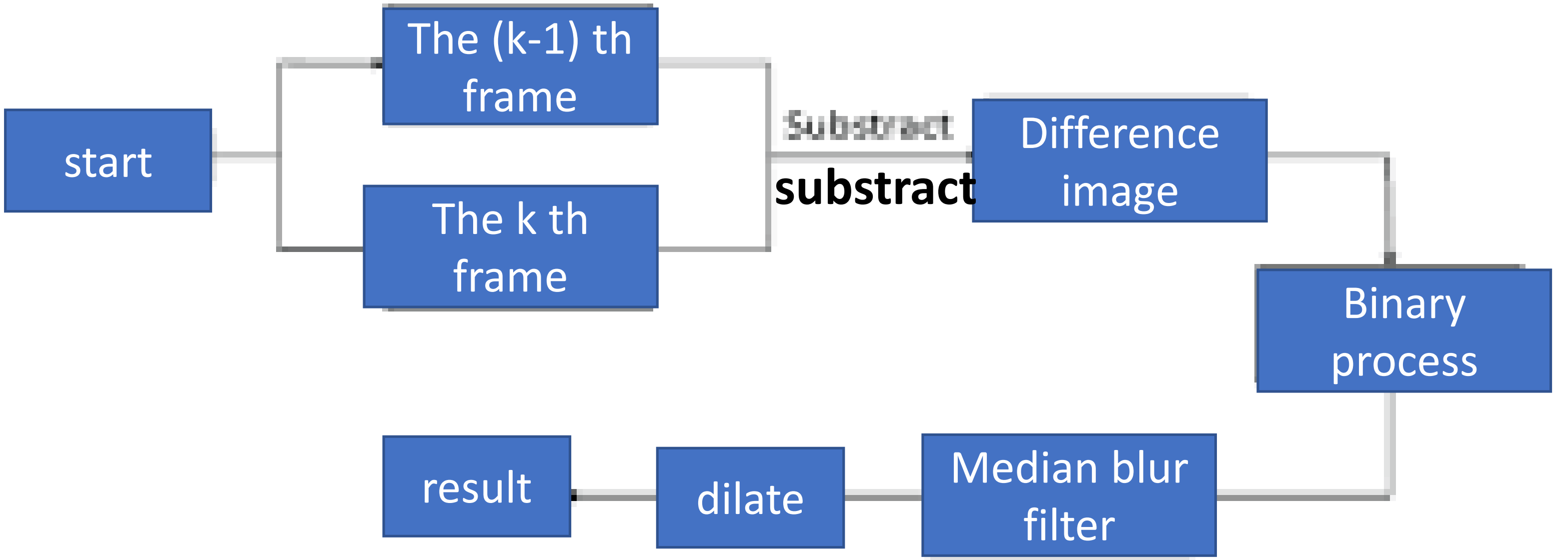
Techniques in Movement and Video Analysis

- **Frame Differencing:**
  - Compare pixel intensity between consecutive frames to highlight moving regions.
  - Example: Detecting motion in a static camera scene.
- **Background Subtraction:**
  - Maintain a model of the static background and subtract it from each frame.
  - Application: Removing non-relevant static areas in CCTV footage.
- **Optical Flow:**
  - Horn-Schunck or Lucas-Kanade algorithms to detect per-pixel motion.
  - Used in analyzing flow direction in drone videos.



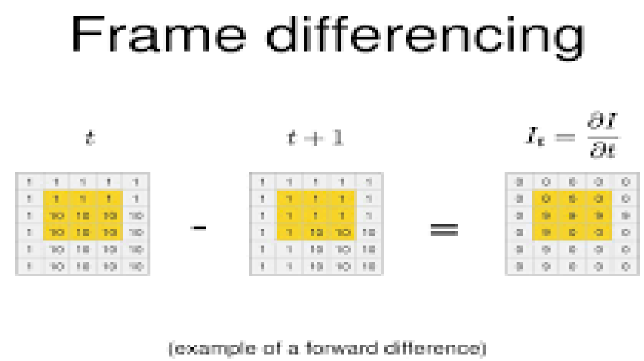
- **Frame Differencing:**

Frame Differencing is a fundamental technique in computer vision used for motion detection. It involves comparing consecutive frames (images) in a video sequence to detect changes over time. These changes often correspond to moving objects or regions.



• Frame Differencing:

How It Works:



Input Video Frames

- A video is essentially a sequence of images called frames, displayed rapidly to create the illusion of motion

Pixel-by-Pixel Comparison

- For every pixel in the current frame, subtract the corresponding pixel value in the previous frame.
  - The resulting difference is a new image called a difference frame.

Thresholding

- - Small differences (e.g., due to noise or slight illumination changes) are ignored by applying a threshold.
  - Pixels with differences above this threshold are marked as 'changed,' often represented as white in the difference frame, while unchanged pixels are black.

Result

- The white regions in the difference frame represent areas of movement

- **Frame Differencing:** **How It Works:**

Example:

Scenario:

Consider a security camera monitoring an empty room. A video sequence has the following:

- Frame 1: An empty room.
- Frame 2: A person walks into the room.

1. Original Frames:

 - Frame 1: Pixels represent the static scene.
 - Frame 2: Pixels representing the person are different from Frame 1.

2. Frame Difference:

 - Subtract Frame 1 from Frame 2. The pixels corresponding to the person will appear as white (indicating motion), while the rest will be black.

3. Thresholding:

 - Apply a threshold to ignore minor differences like lighting changes.

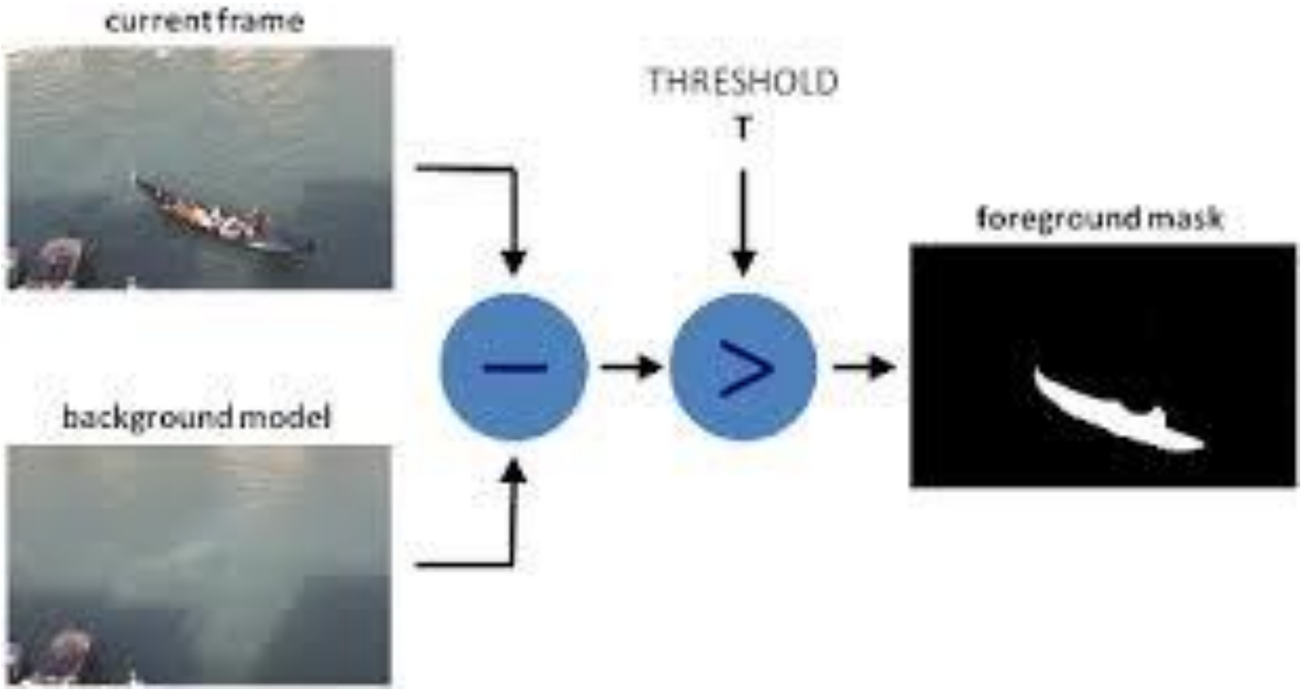
- **Frame Differencing:**

Simplified Definition = spot the difference game !

Frame Differencing is like playing a 'spot the difference' game between two pictures taken one after the other. Imagine taking a photo of an empty room and then another when someone walks in. You compare the two photos, and wherever you see changes, that's where the person moved. It helps cameras notice movement!

- **Background Subtraction:**

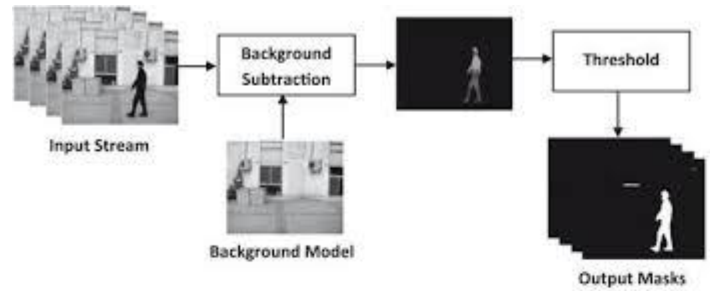
Background Subtraction is a widely used technique in computer vision to detect moving objects in video streams. It involves separating the foreground (moving objects) from the static background.



Techniques in Movement and Video Analysis

- **Background Subtraction :**

How It Works:



Background Model Creation

- The algorithm creates a model of the background by analyzing multiple frames in the video. This model represents the static or slowly changing parts of the scene

Frame Comparison

- Each incoming video frame is compared to the background model.
  - Differences between the current frame and the background are identified as foreground .

Foreground Extraction

- Pixels with significant differences are marked as part of the moving objects (foreground).

Refinement

- Noise and small changes (e.g., tree branches swaying) are often reduced using filters and morphological operations.

- **Background Subtraction::**

Example:

Scenario:

Imagine a surveillance camera monitoring a parking lot. The background (empty parking lot) is static, while cars moving through the area are the foreground.

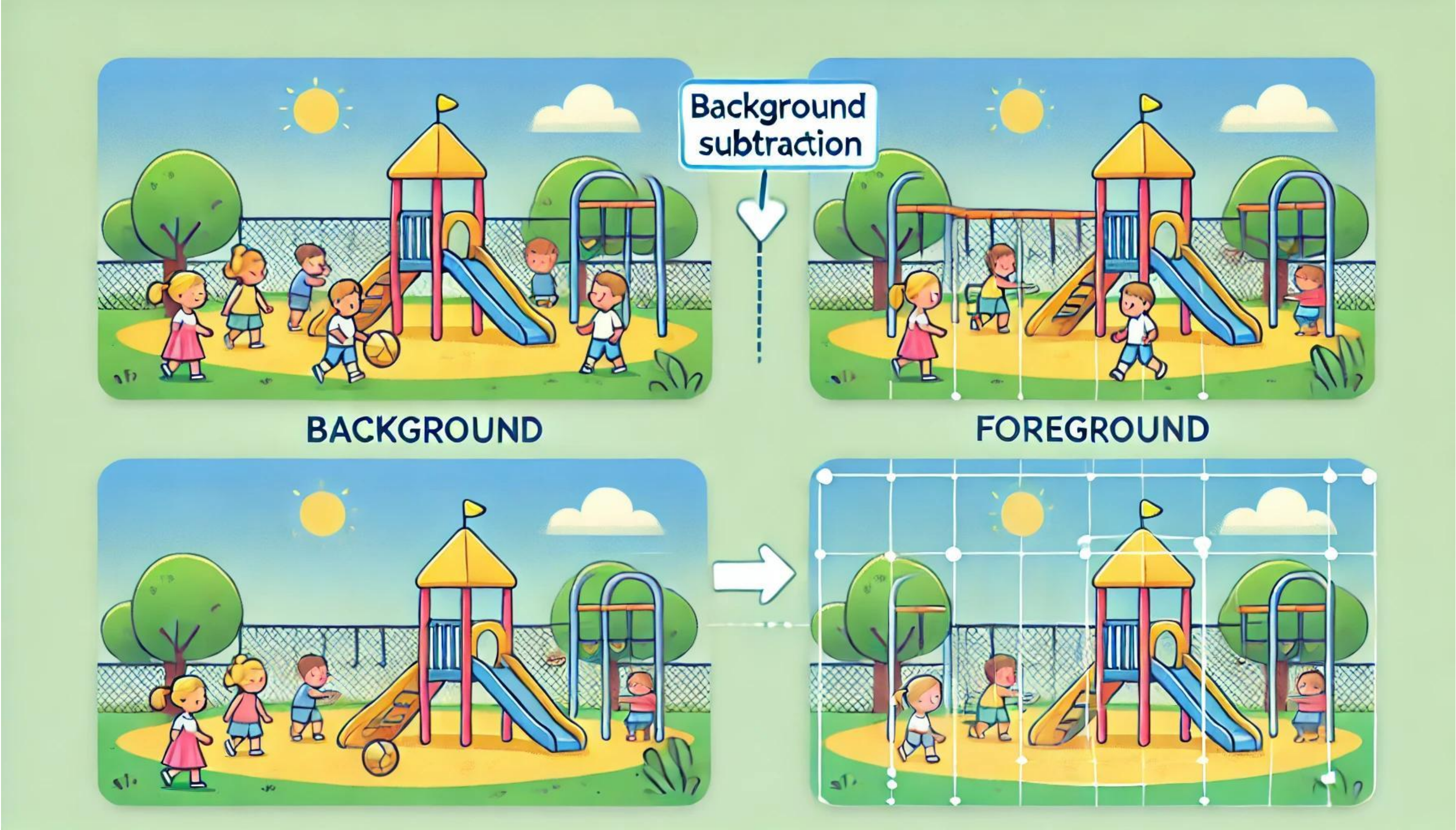
Steps:

1. Background Model: The camera learns the empty parking lot as the background.
2. New Frame: A car enters the frame.
3. Subtraction: The car’s pixels differ from the background model.
4. Foreground Detection: The car is highlighted as a moving object.

- **Background Subtraction::**

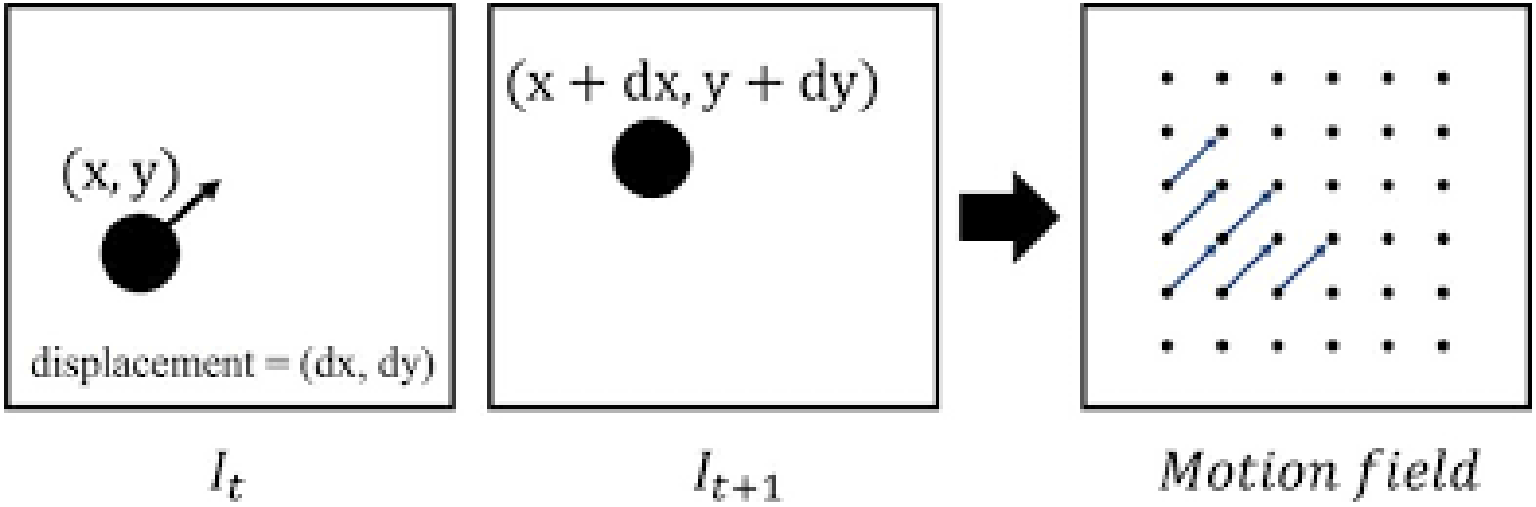
  Simplified Definition = watching playground

  Background Subtraction is like watching a playground and remembering how it looks when no one is there (the background). When kids come in and play, they are easy to spot because they don't match the empty playground. This method helps cameras figure out what's moving and what's not.



- Optical Flow:

Optical Flow is a technique in computer vision that estimates the motion of objects, surfaces, or edges between consecutive frames in a video. It measures how the brightness of pixels changes over time, giving us an idea of how objects are moving in a scene.

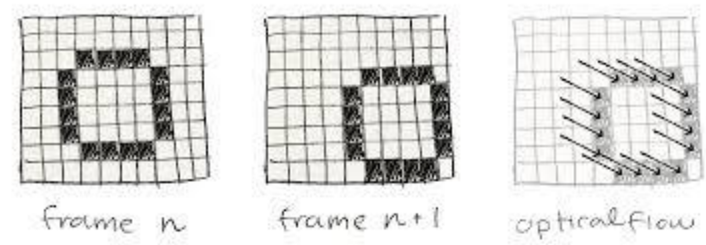


- Optical Flow :

How It Works:

Motion estimation

Optical flow calculates the motion of pixels based on changes in intensity (brightness) between two frames.

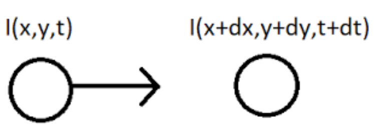


Assumptions

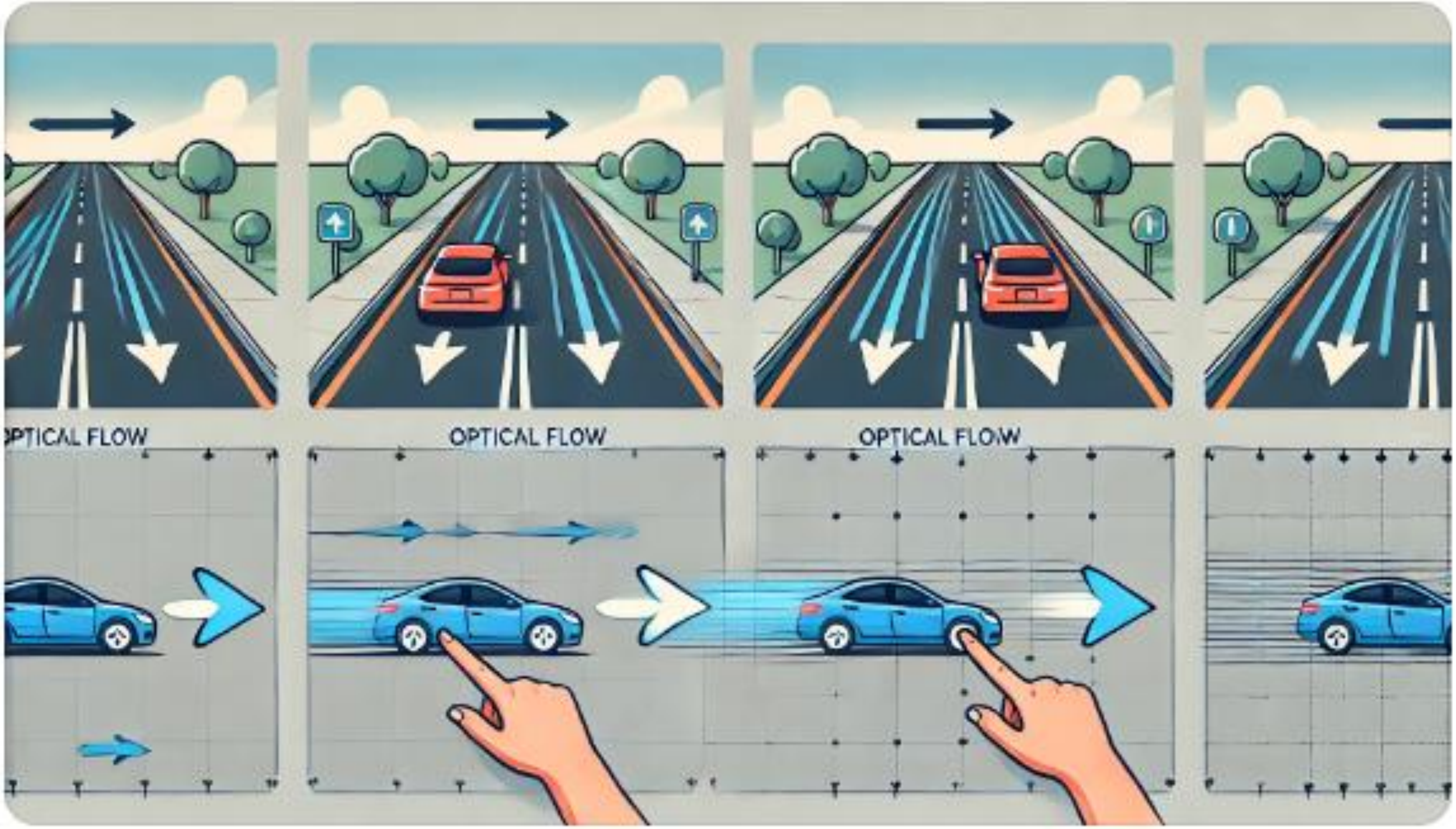
- Brightness Constancy: the brightness of an object remains constant as it moves
- Temporal Smoothness: the motion changes gradually over time,

Flow vectors

- Motion is represented using arrows called flow vectors, which indicates the direction and the speed of movement of each pixel



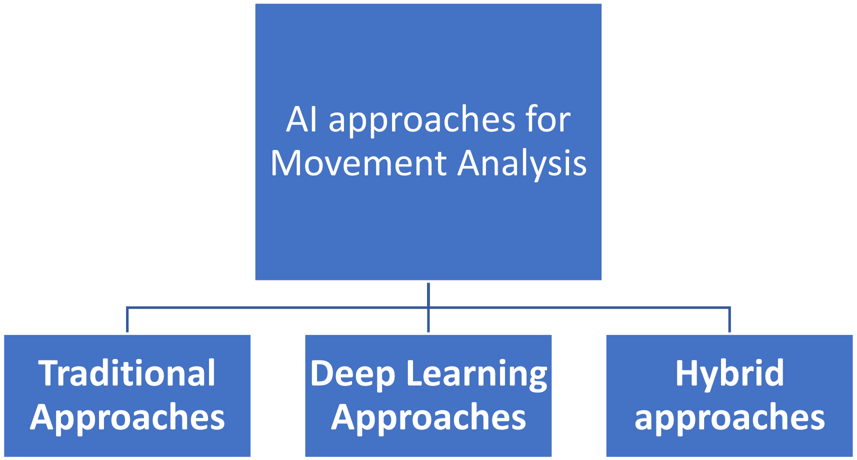
Optical flow



- Optical Flow :Here is an illustration representing Optical Flow, showcasing the motion vectors between two consecutive frames

AI approaches for Movement Analysis

- ✓ **Traditional Approaches:**
  - ✓ Features like Histogram of Oriented Gradients (HOG) combined with SVMs.
  - ✓ Example: Tracking pedestrians using handcrafted features.
- ✓ **Deep Learning Approaches:**
  - ✓ CNNs: Capture spatial features for object detection.
  - ✓ RNNs/LSTMs: Process temporal sequences for activity recognition.
  - ✓ Example: OpenPose for pose estimation in sports.
- ✓ **Hybrid approaches:**
  - ✓ Combination of traditional and deep learning approaches
- learning-based methods have replaced older techniques due to higher accuracy and robustness.



AI approaches for Movement Analysis

AI Approaches for Movement Analysis

Below is a graphical representation of AI approaches used in movement analysis, highlighting their algorithms, techniques, advantages, and weaknesses.

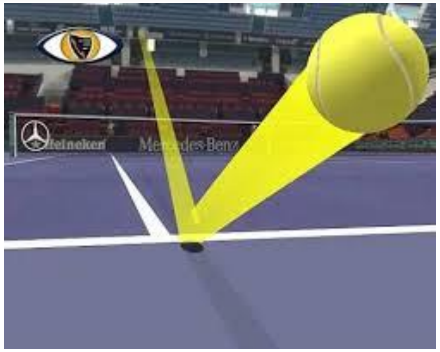
Approaches Table Graphical Representation

AI Approach	Algorithms and Techniques	Advantages	Weaknesses
Traditional Machine Learning	HOG + SVM, KNN, Decision Trees	Uses hand-crafted features; relatively simple.	Less robust to environmental changes and noise.
Deep Learning	CNNs, RNNs, LSTMs, YOLO, OpenPose	Automated feature extraction, high accuracy.	Requires large datasets and high computational power.
Hybrid Approaches	Combination of traditional and deep learning methods.	Balances simplicity and accuracy.	Complex to design and implement.

Applications

•**Sports Analytics:** Track players for performance evaluation or strategy analysis.

- Example: Hawk-Eye in tennis for tracking ball movement.



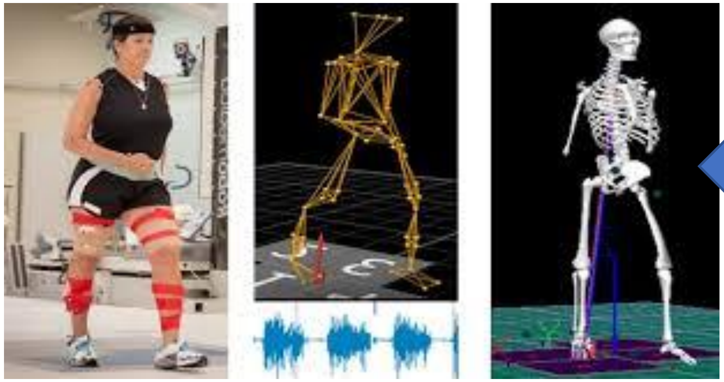
•**Surveillance Systems:** Identify unusual patterns like loitering or theft.

- Example: Smart surveillance systems in airports.

•**Autonomous Vehicles:** Detect and track pedestrians, vehicles, and traffic signs.

- Example: Tesla’s autopilot system.

.



•**Healthcare:** Analyze gait or motion for rehabilitation therapy.

- Example: Using Kinect for gait analysis in stroke patients.

.

Challenges in Video Analysis

•**Occlusion:** Objects getting obscured (e.g., a car passing behind a tree).
•**Complex Motion:** Fast, erratic movements, e.g., in sports or drones.
•**Lighting Variations:** Poor or changing lighting conditions.
•**Real-Time Requirements:** Processing high-resolution videos quickly.

Example:

•Tracking players in a crowded soccer field under varying light conditions.

Visuals:

•Examples of occluded or blurred frames.

Tools and Frameworks

1. **OpenCV:** Library for motion tracking, background subtraction, and object detection.
2. **TensorFlow/PyTorch:** Building custom deep learning models for video analysis.
3. **MediaPipe:** Pre-built solutions for pose estimation, hand tracking, etc.
4. **NVIDIA DeepStream:** Optimized pipelines for real-time video analytics.

Case Studies

1.Traffic Monitoring System:

1. Goal: Identify speeding vehicles.
2. Technique: Combine motion tracking with speed calculation from video frames.

2.Pose Estimation in Fitness:

1. Goal: Detect incorrect postures during exercise.
2. Tools: OpenPose, TensorFlow.

Visuals:

•Screenshots of system outputs for traffic monitoring and pose estimation.

Conclusion

- Movement and video analysis is central to many cutting-edge technologies.
- Challenges like occlusion and real-time processing require innovative solutions.
- Future directions: Integration with edge computing and 3D video analysis.

References

1. Dana H. Ballard & Christopher M. Brown. Computer Vision Prentice Hall, Inc, 1982
2. Robert M. Haralick & Linda G. Shapiro. Computer and Robot Vision, Vol-I, Addison-Wesley Publishing Company, 1992
3. Robert M. Haralick & Linda G. Shapiro. Computer and Robot Vision, Vol-II, Addison-Wesley Publishing Company, Inc, 1993
4. Linda Shapiro & Azriel Rosen eld. Computer Vision and Image Processing, Academic Press, Inc, 1992
5. Berthold Klaus Paul Horn. Robot Vision , MIT Press McGraw-Hill Book Company, 1986
6. Robert J. Schalko. Digital Image Processing and Computer Vision, John Wiley & Sons Inc, 1989
7. George Stockman and Linda Shapiro. Three Dimensional Computer Vision. Prentice Hall 2000.
8. David Marr. Vision, W. H Freeman and Company, NY, 1982
9. Rafael C. Gonzalez and Paul Wintz. Digital Image Processing, Third edition, Addison Wesley, MA. (Now with Prentice Hall, eective 1999).
10. Ernest Hall. Computer Image Processing and Recognition, second edition, Academic press 1982.
11. Azriel Rosenfeld and Avinash C. Kak. Digital Picture Processing, Vol. 1 & Vol. 2, Academic Press, 1982.
12. Robert J. Schalko. Digital Image Processing and Computer Vision: An introduction to theory and implementations, John Wiley & Sons, New York, 1989.
13. William K. Pratt. Digital Image Processing, John Wiley & Sons, 1993.