

Human Body Posture and Pose Estimation summary



S. Kokkas, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 3.0
Date: 25/01/2021

Human body posture and pose estimation



- **Data acquisition**
 - **Sensors**
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

Sensors for Human Body Posture and Pose Estimation

- Smartphone sensors
- Wearables
- Graphene sensors
- Strain Sensors
- Night Light Camera Sensors
- Inertial Measurement Units (IMUs)
- Kinect

Human body posture and pose estimation



- **Data acquisition**
 - Sensors
 - **Datasets**
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

Datasets for Human Body Posture and Pose Estimation

- COCO
- MPII Human Pose dataset
- CMU Panoptic Dataset(OpenPose)
- DensePose
- Unite the People
- AI Challenger
- Human3.6M
- Mannequin RGB and IRS in bed-dataset
- VGG Pose

Human body posture and pose estimation



- Data acquisition
 - Sensors
 - Datasets
- **Human body models**
 - **2D Skeletons**
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

2D Skeletons

- Representation of the human body in mathematical way.
- Body parts: Edges and Joints: Nodes.
- IMU sensors.
- Real-time quaternions:

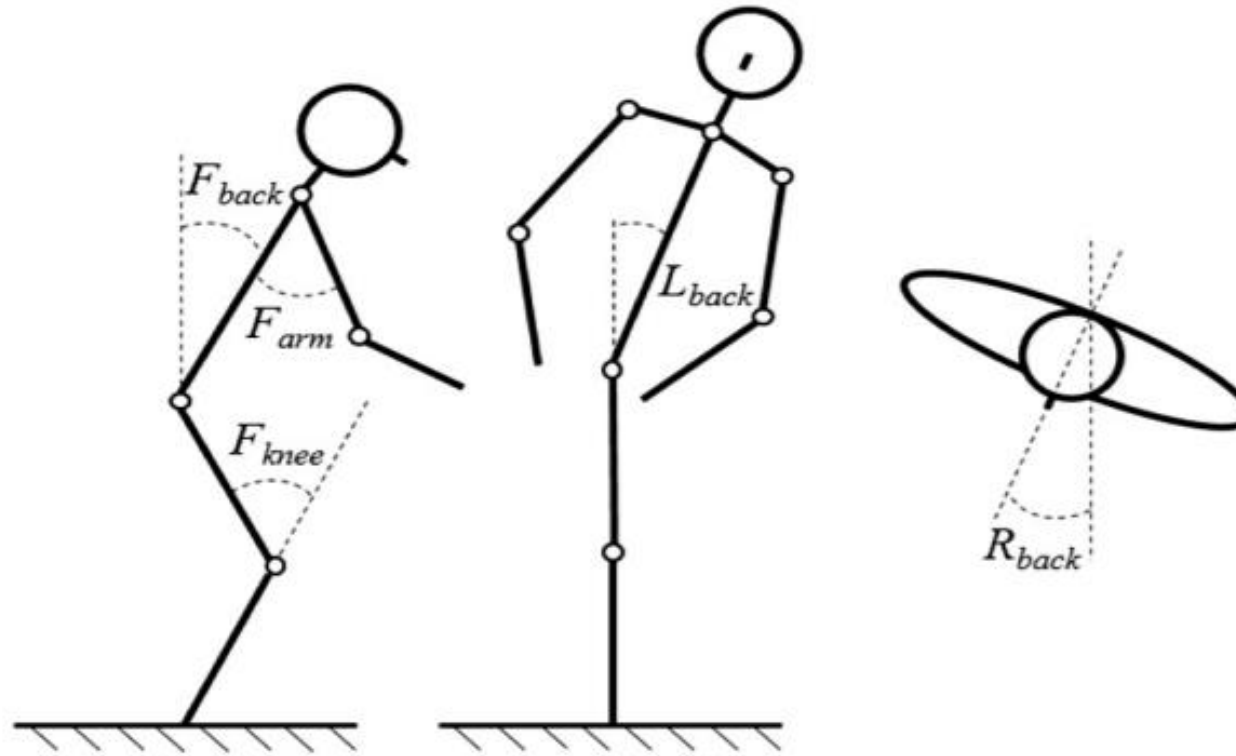
$$q = [\mathbf{w}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{z}_k]$$

- Coordinates:

$$p'_1 = [0, (-2\mathbf{w}_1 \cdot \mathbf{z}_1 + 2\mathbf{x}_1 \cdot \mathbf{y}_1)i, (\mathbf{w}_1^2 - \mathbf{x}_1^2 + \mathbf{y}_1^2 - \mathbf{z}_1^2)j, (2\mathbf{w}_1 \cdot \mathbf{x}_1 + 2\mathbf{y}_1 \cdot \mathbf{z}_1)k]$$

$$p'_2 = [0, (-2\mathbf{w}_2 \cdot \mathbf{y}_2 - 2\mathbf{x}_2 \cdot \mathbf{z}_2)i, (2\mathbf{w}_2 \cdot \mathbf{x}_2 - 2\mathbf{y}_2 \cdot \mathbf{z}_2)j, (-\mathbf{w}_2^2 + \mathbf{x}_2^2 + \mathbf{y}_2^2 - \mathbf{z}_2^2)k]$$

2D Skeletons



2D Skeleton Example.

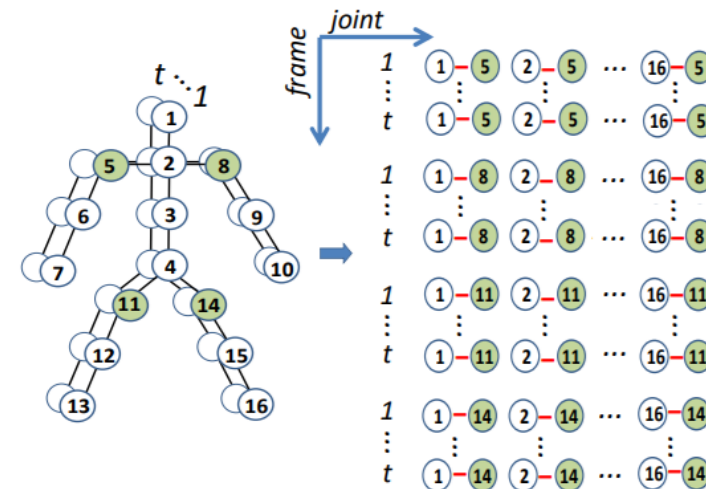
Human body posture and pose estimation



- Data acquisition
 - Sensors
 - Datasets
- **Human body models**
 - 2D Skeletons
 - **3D Skeletons**
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

3D Skeletons

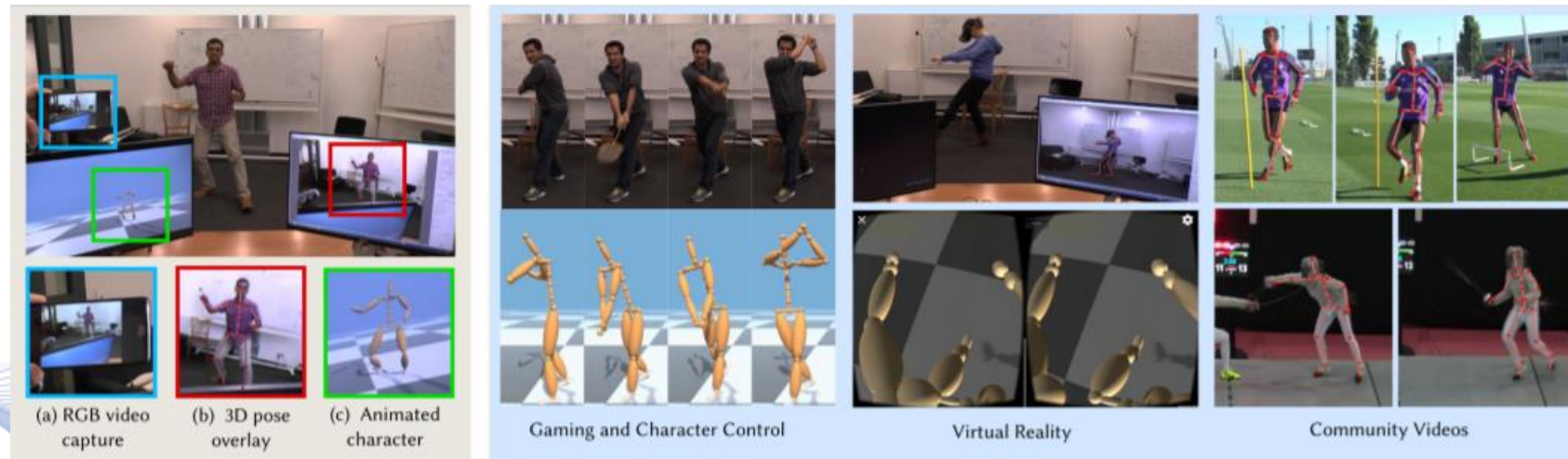
- RGB Video.
- Time Dimension.
- Skeleton sequence: Chain of frames of concatenating joints.
- Dimensions: $(m - 1) \times n$



Clip Generation of a skeleton sequence. [QIU2017]

3D Skeletons

- By evaluating the E_{total} , we can project the joints and create the 3D Skeleton.



VNect 3D Skeleton Estimation Results. [DUS2017]

Human body posture and pose estimation



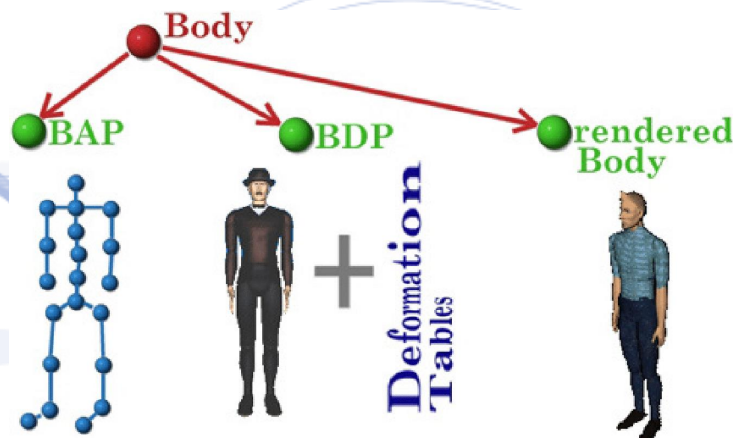
- Data acquisition
 - Sensors
 - Datasets
- **Human body models**
 - 2D Skeletons
 - 3D Skeletons
 - **MPEG-4**
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

MPEG-4 Descriptors

- Video stream compression and processing.
- Fields: 1) Digital Television, 2) Interactive graphics and 3) Interactive Multimedia.
- Applications: Body and Face.
- Body: ***Body Animation Parameters(BAPs)*** and ***Body Definition Parameters(BDPs)***.
- Face: ***Face Animation Parameters(FAPs)*** and ***Face Definition Parameters(FDPs)***.

MPEG-4 Descriptors

- Body:
 - Visualization as nodes.
 - **BodyNode** consists of three nodes.
 - First sub-node: BAPs (296 parameters) that describe the topology of the body.
 - Second sub-node: BDPs that associate joints with each other.



MPEG-4 Body. [THAL2004]

MPEG-4 Descriptors



Facial Feature extraction using FAPs from MPEG-4.[MIN2004]

Human body posture and pose estimation



- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- **Human body posture estimation**
 - **2D human body posture estimation**
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

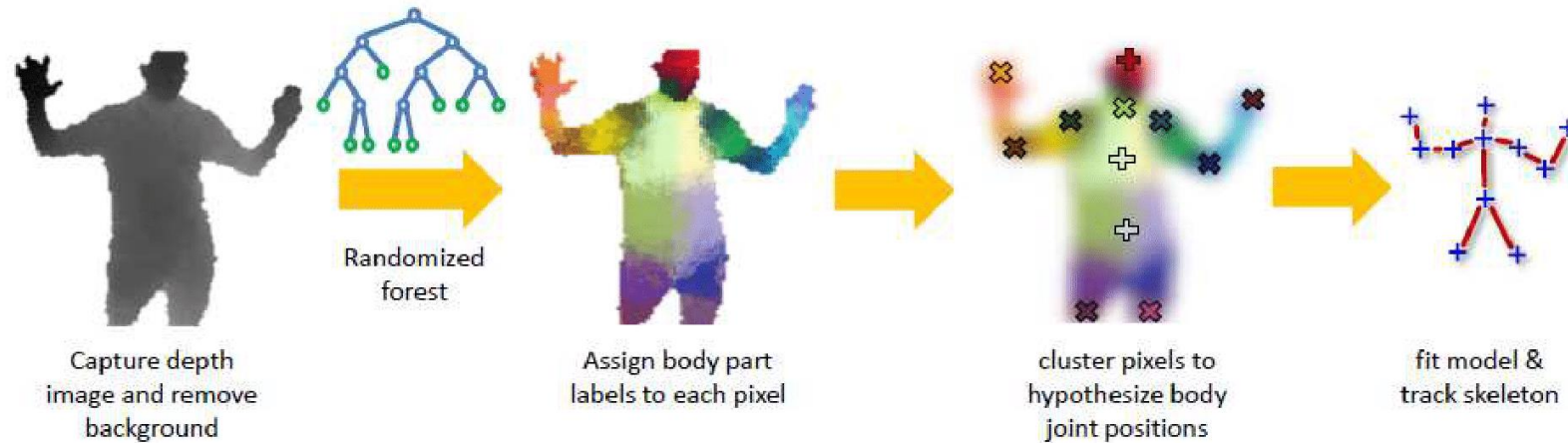
2D Human Body Posture Estimation



- **Posture** means an intentionally or habitually assumed position of the joints.
- Human key posture points include rich 2D appearance, angular point, and multi-point autocorrelation.
- In order to achieve the 2D Posture Model, first we need to detect the human body.

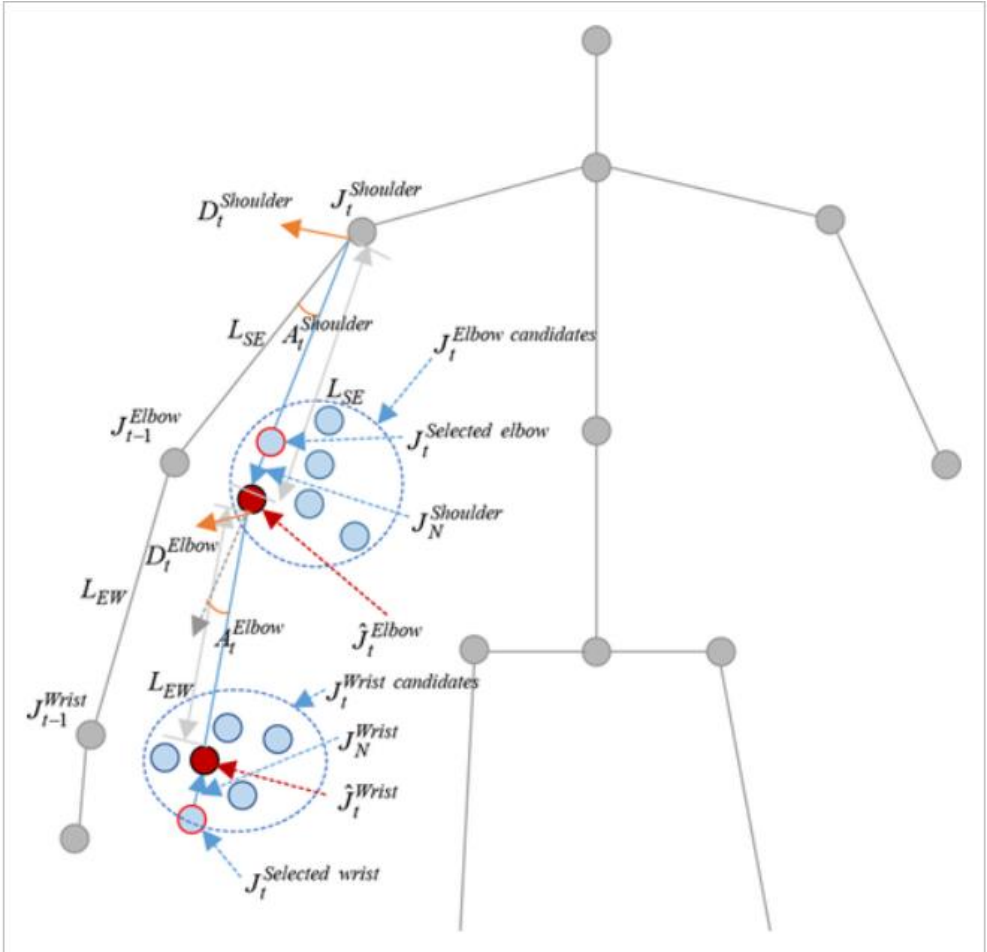
2D Human Body Posture Estimation

- Kinect sensors to get a skeleton sequence.



Kinect Skeleton Sequence. [SMIS2011]

2D Human Body Posture Estimation



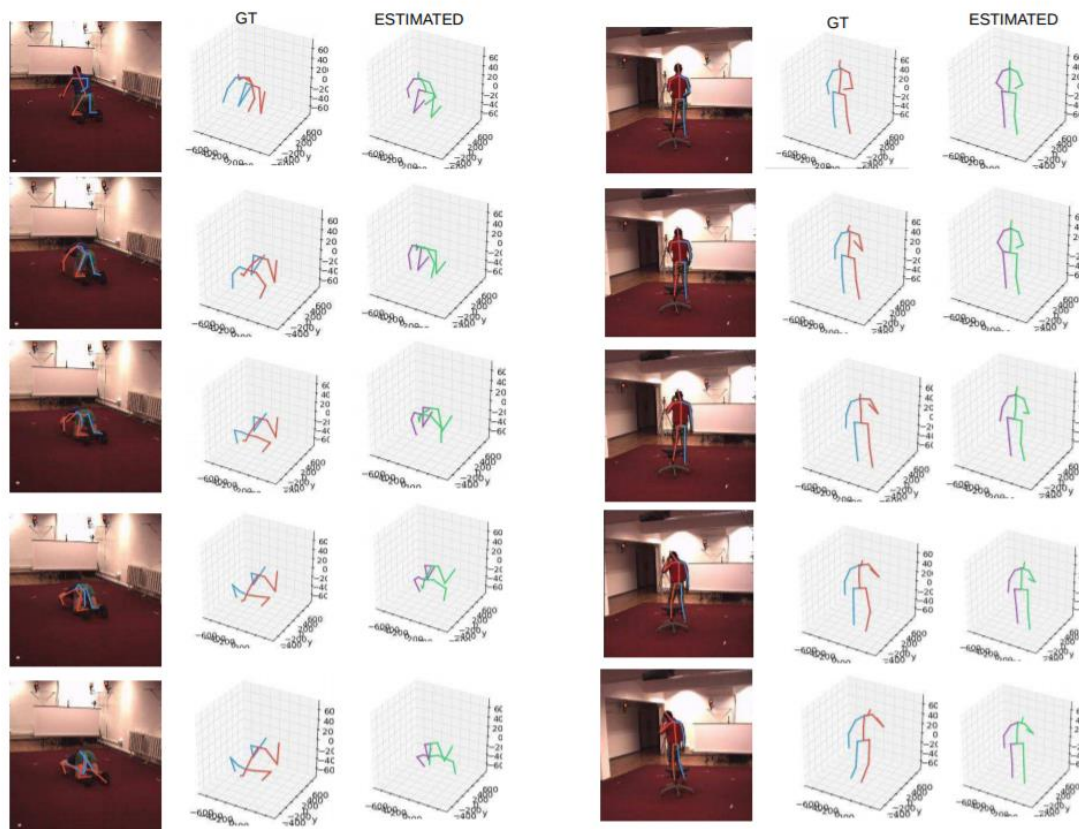
Example of 2D Human Body Posture Estimation.
[KIM2017]

Human body posture and pose estimation



- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- **Human body posture estimation**
 - 2D human body posture estimation
 - **3D human body posture estimation**
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

3D Human Body Posture Estimation



Video frames, ground truth and prediction. [RAY2018]

Human body posture and pose estimation

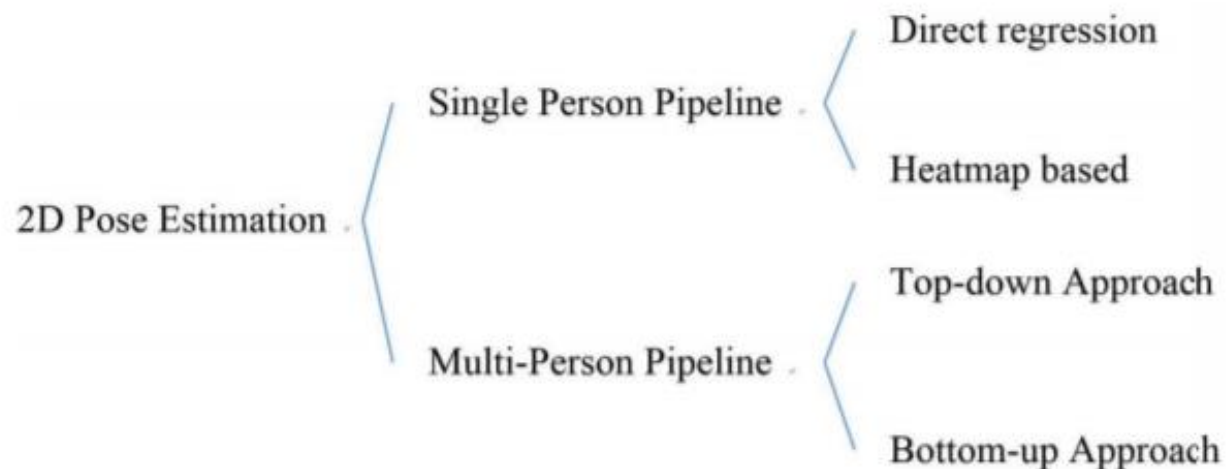


- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- **Human body pose estimation**
 - **2D Human body pose estimation**
 - 3D Human body pose estimation
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

Human Body Pose Estimation

- **Pose** is the combination of *position* and *orientation* of an object vs the camera.
- It can be considered as a regression problem
- Deep Learning Methods.

2D Human Body Pose Estimation



Sections of 2D Human Body Pose Estimation. [DAN2019]

2D Human Body Pose Estimation



- ***Single Person Pipeline:***
 - Localization of anatomical keypoints for a single person.
 - Spatial dependencies between keypoints.
 - Convolutional Neural Network.
 - Subpart of Multi-Person Pipeline.

2D Human Body Pose Estimation

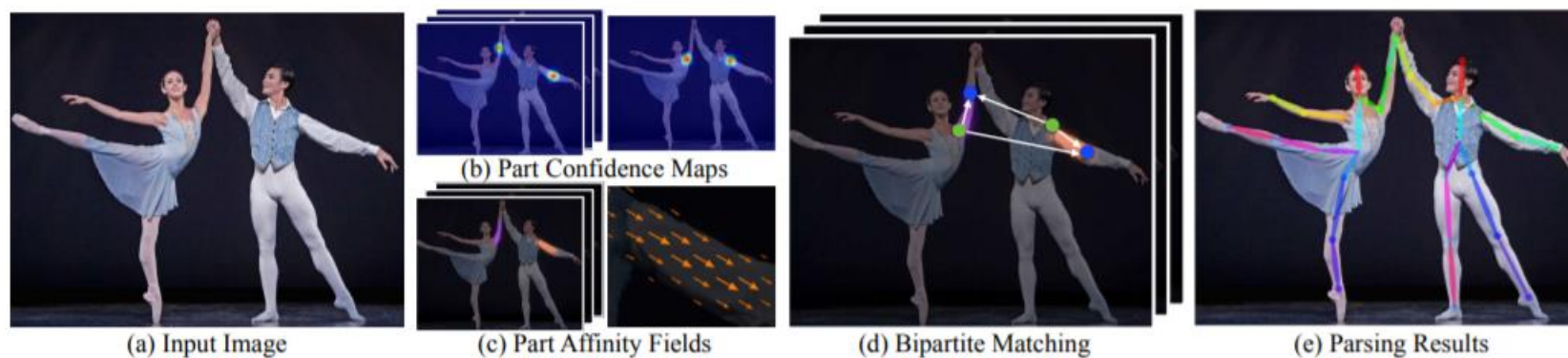


- ***Multi-Person Pipeline:***

- Top-Down : Each human is detected and each human's keypoints are selected.
- Bottom-Up: All keypoints are identified, reference to possible humans in the image.
- Most famous approach is Bottom-Up *Part Affinity Fields* (PAFs).

2D Human Body Pose Estimation

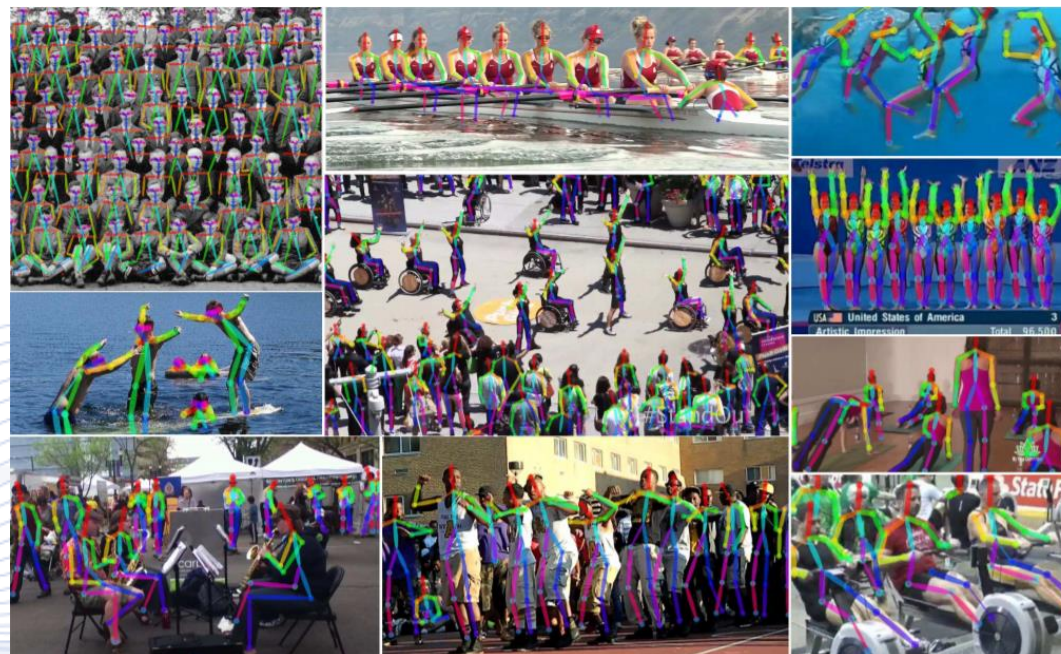
- ***Multi-Person Pipeline:***



Overall Pipeline of PAFs.[CAO2017]

2D Human Body Pose Estimation

- **Multi-Person Pipeline:**



PAFs Results in Multi-Person Pose Estimation.[CAO2017]

Human body posture and pose estimation



- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- **Human body pose estimation**
 - 2D Human body pose estimation
 - **3D Human body pose estimation**
- Face Pose Estimation
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

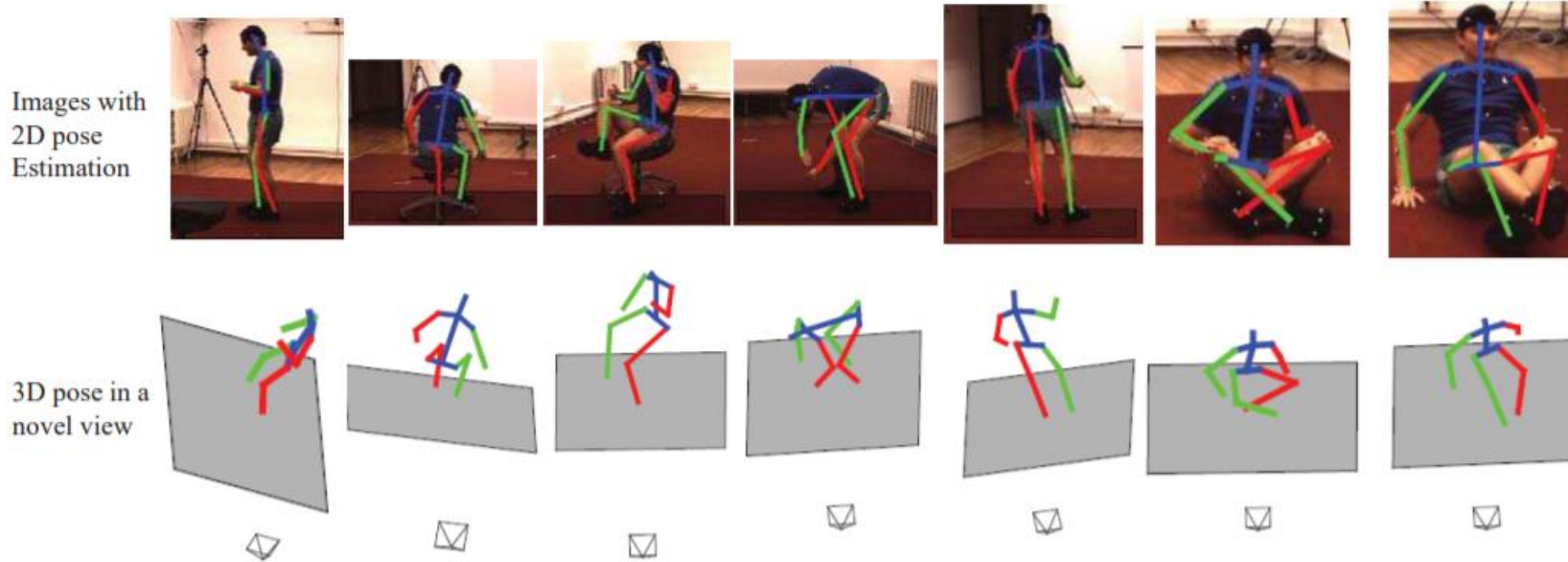
3D Human Body Pose Estimation



- We can divide ***3D Human Body Pose Estimation*** in two categories:
 - Probabilities.
 - Direct Estimation.

3D Human Body Pose Estimation

- **Probabilities:**

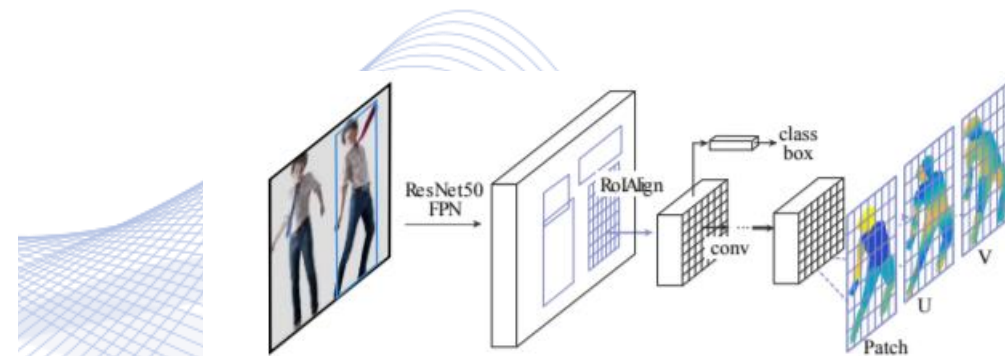
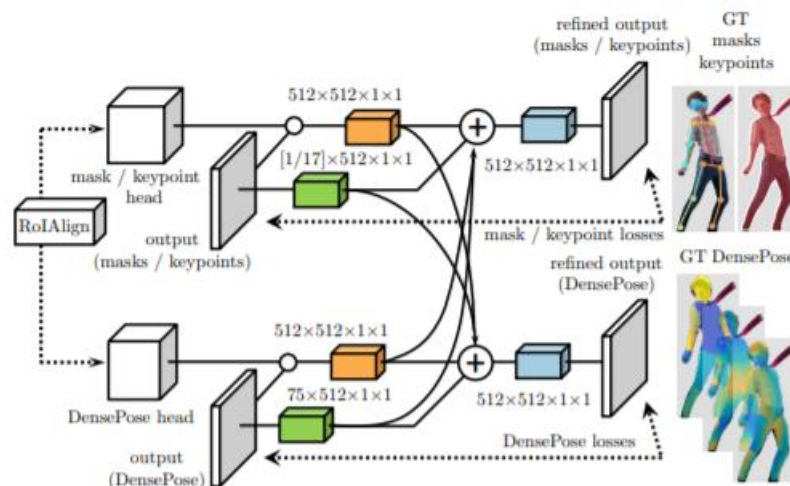


Probabilistic Method Results.[CHEN2017]

3D Human Body Pose Estimation

- **Direct Estimation:**

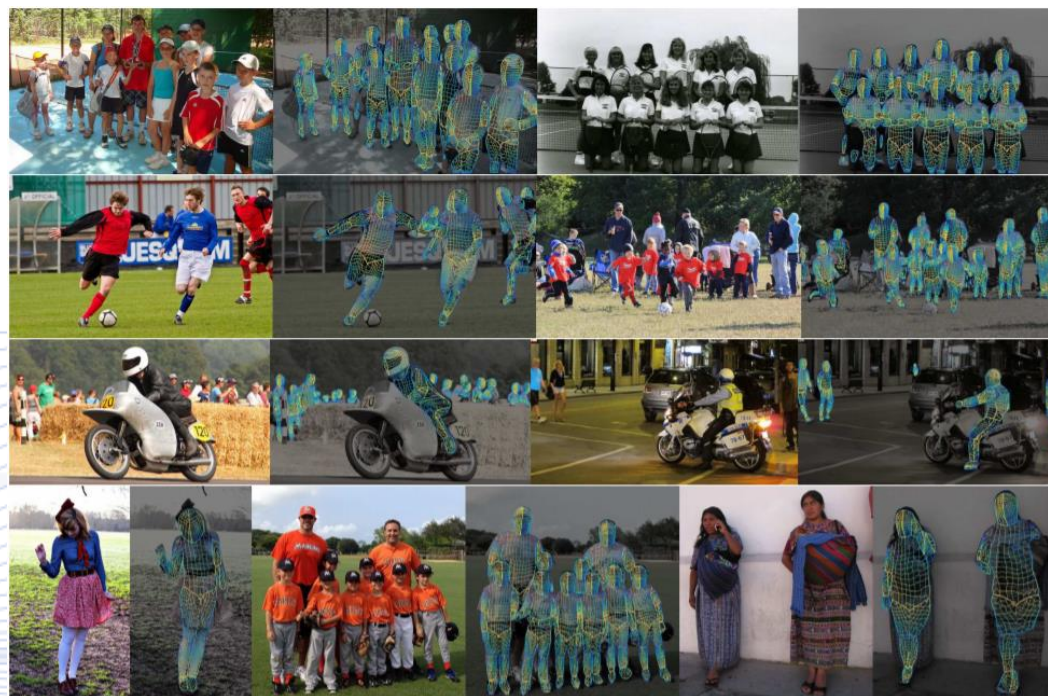
- This is the preprocessing. The data are used as input in *ResNet50* and a *Cross-Cascading Architecture*.



Cross-Cascade and General Architecture of DensePose.[GUL2018]

3D Human Body Pose Estimation

- ***Direct Estimation:***



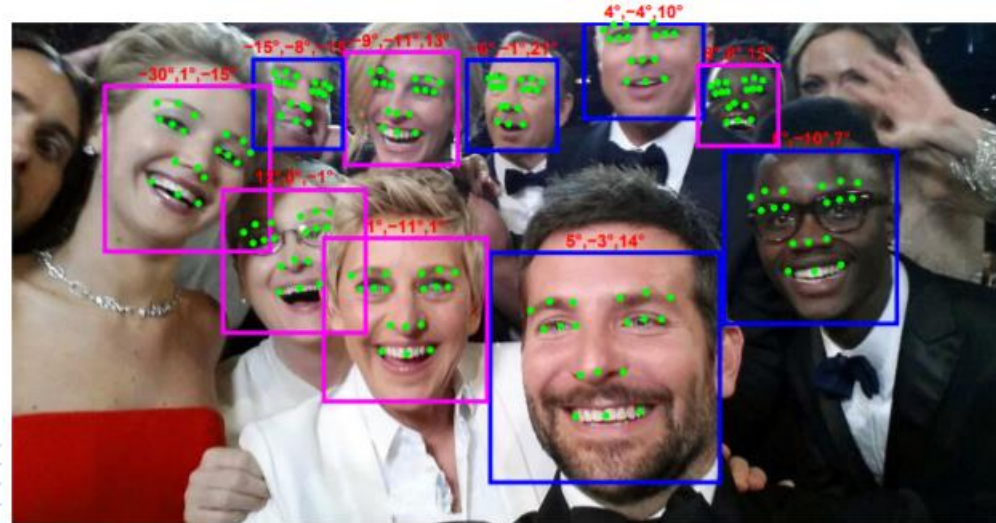
Example Results of DensePose.[GUL2018]

Human body posture and pose estimation



- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- **Face Pose Estimation**
- Software tools
 - Convolutional Neural Networks
 - Deep Convolutional Neural Networks

Face Pose Estimation



Hyperface detecting faces of Hollywood stars.[RAN2019]

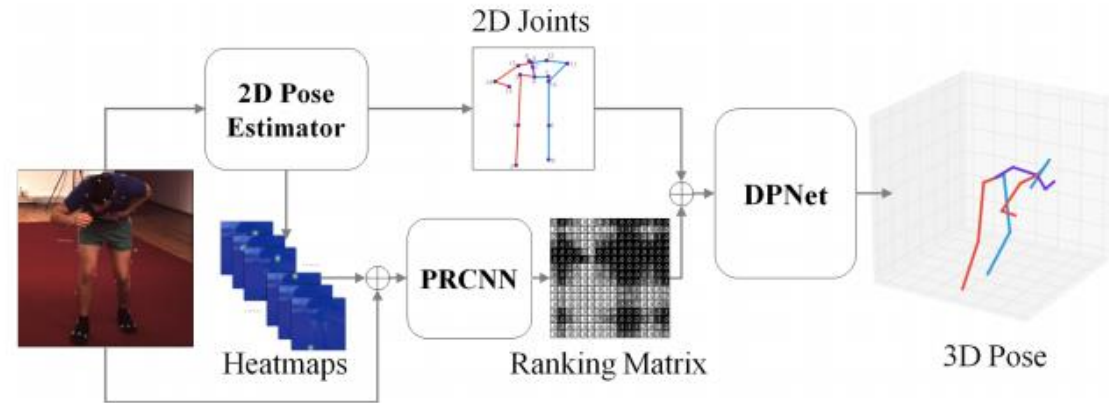
Human body posture and pose estimation



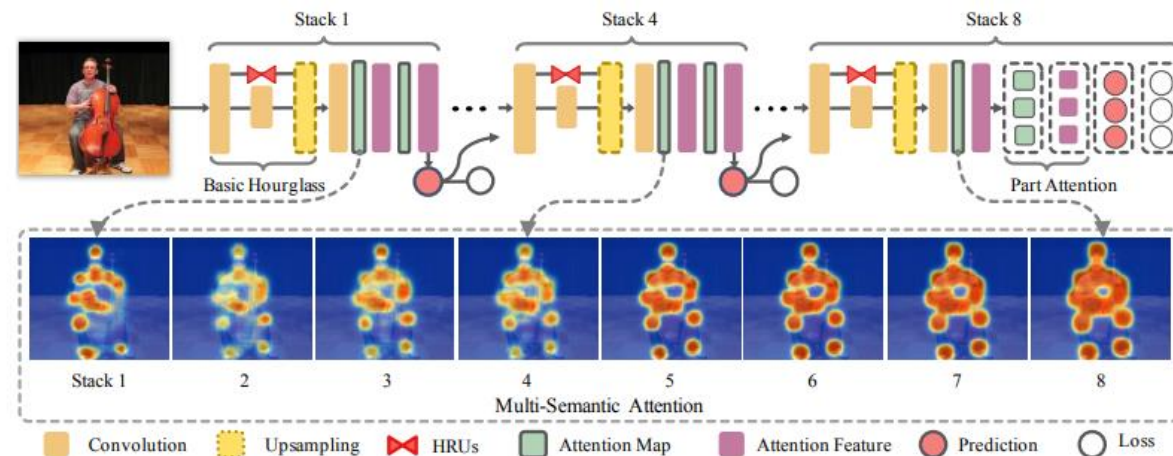
- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- **Software tools**
 - **Convolutional Neural Networks**
 - Deep Convolutional Neural Networks

Convolutional Neural Networks

- DRPose3D.
- Neural Body Fitting (NBF).
- In the Wild.
- Single-shot Multi-Person.
- BodyNet.
- ConvNet.
- Multi-Context Attention.



DRPose3D Architecture.[MIN2018]



Multi-Context Attention Architecture.[CHU2017]

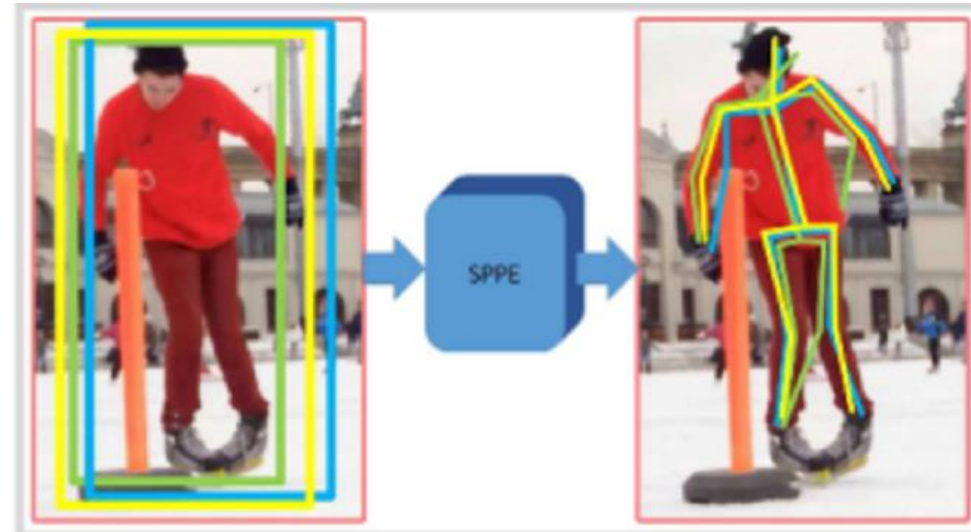
Human body posture and pose estimation



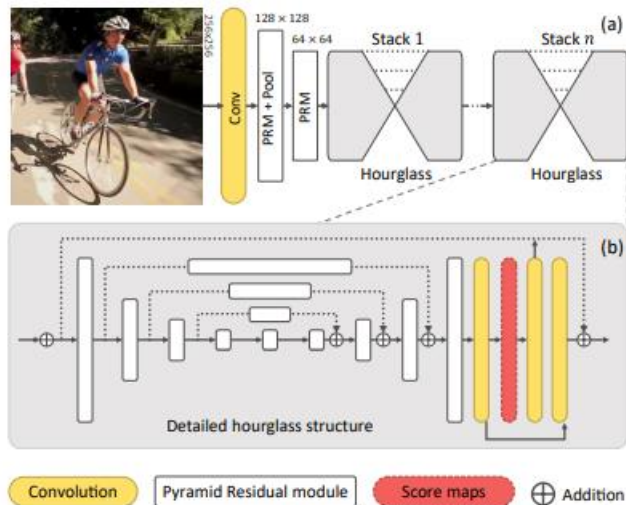
- Data acquisition
 - Sensors
 - Datasets
- Human body models
 - 2D Skeletons
 - 3D Skeletons
 - MPEG-4
- Human body posture estimation
 - 2D human body posture estimation
 - 3D human body posture estimation
- Human body pose estimation
 - 2D Human body pose estimation
 - 3D Human body pose estimation
- Face Pose Estimation
- **Software tools**
 - Convolutional Neural Networks
 - **Deep Convolutional Neural Networks**

Deep CNNs

- RMPE (AlphaPose).
- High Resolution Net (HRNet).
- Feature Pyramids.



Annotations transformed through SPPE to skeleton. [FAN2017]



Proposed Architecture of Feature Pyramids. [YAN2017]

Bibliography

- [PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.
- [PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).
- [PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.
- [NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.
- [PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

Q & A

Thank you very much for your attention!

**More material in
<http://icarus.csd.auth.gr/cvml-web-lecture-series/>**

**Contact: Prof. I. Pitas
pitass@csd.auth.gr**