Human Motion Analysis

Raquel Urtasun

TTI Chicago

Feb 22, 2010

What is this all about?

- Class webpage http://ttic.uchicago.edu/∼rurtasun/
- This class is interdisciplinary, it combines the fields of
  - computer vision
  - computer graphics
  - machine learning
  - other domains: robotics, etc

What is this all about?

- Class webpage http://ttic.uchicago.edu/∼rurtasun/
- This class is interdisciplinary, it combines the fields of
  - computer vision
  - computer graphics
  - machine learning
  - other domains: robotics, etc
- The class is composed of
  - 2h/week of theory
  - 1h/week of practice: guided exercises + final project (research).

What is this all about?

- Class webpage http://ttic.uchicago.edu/∼rurtasun/
- This class is interdisciplinary, it combines the fields of
  - computer vision
  - computer graphics
  - machine learning
  - other domains: robotics, etc
- The class is composed of
  - 2h/week of theory
  - 1h/week of practice: guided exercises + final project (research).
- I would like the class to be interactive: suggestions please ;)

What is this all about?

- Class webpage http://ttic.uchicago.edu/∼rurtasun/
- This class is interdisciplinary, it combines the fields of
  - computer vision
  - computer graphics
  - machine learning
  - other domains: robotics, etc
- The class is composed of
  - 2h/week of theory
  - 1h/week of practice: guided exercises + final project (research).
- I would like the class to be interactive: suggestions please ;)
- GOAL: Good understanding of theoretical and practical aspects for the analysis of human motion.

What is this all about?

- Class webpage http://ttic.uchicago.edu/∼rurtasun/
- This class is interdisciplinary, it combines the fields of
  - computer vision
  - computer graphics
  - machine learning
  - other domains: robotics, etc
- The class is composed of
  - 2h/week of theory
  - 1h/week of practice: guided exercises + final project (research).
- I would like the class to be interactive: suggestions please ;)
- GOAL: Good understanding of theoretical and practical aspects for the analysis of human motion.
- EXAM: 20-30 minutes presenting final project + theoretical questions.

What is this all about?

- Class webpage http://ttic.uchicago.edu/∼rurtasun/
- This class is interdisciplinary, it combines the fields of
  - computer vision
  - computer graphics
  - machine learning
  - other domains: robotics, etc
- The class is composed of
  - 2h/week of theory
  - 1h/week of practice: guided exercises + final project (research).
- I would like the class to be interactive: suggestions please ;)
- GOAL: Good understanding of theoretical and practical aspects for the analysis of human motion.
- EXAM: 20-30 minutes presenting final project + theoretical questions.

What are we going to cover in this class?

- Ways to capture human motion
- Human motion representations and direct kinematics
- Models of human pose and motion
- Human motion synthesis
- Pose estimation from images
- Human motion classification

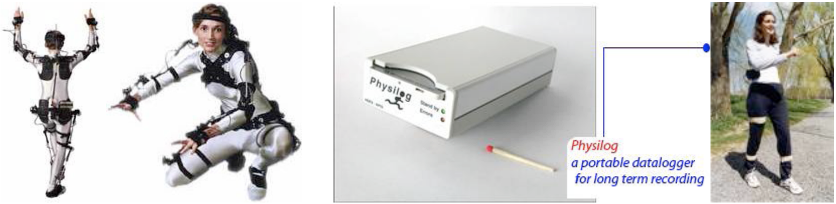
What are we going to cover in this class?

- Ways to capture human motion
  - Mechanical motion capture
  - Electro-magnetical motion capture
  - Optical motion capture
  - Video-based capture
- Human motion representations and direct kinematics
- Models of human pose and motion
- Human motion synthesis
- Pose estimation from images
- Human motion classification

Mocap?

- Just by looking at a set of dots in motion we can identify what's going on [Johansons]

Mechanical motion capture



(a) (b)

Figure: Mechanical commercial motion capture systems. (a) Gypsy system. (b) Physilog system.

Electro-magnetical motion capture

(a) (b) (c) (d)

Figure: Electro-magnetical motion capture systems. (a) Liberty mocap system from Polhemus. (b) Cabled Flock of Birds system from Ascension. (c) Wireless electro-magnetical mocap system Motion Star from Ascension. (d) The motion capture for Lara Craft movie was performed by Motion Star.

Optical motion capture

Figure: Optical motion capture system.

Optical motion capture

Figure: Optical motion capture system.

Optical motion capture

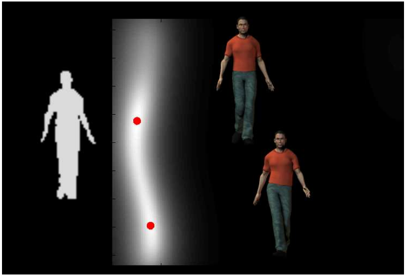
Figure: Optical motion capture system.

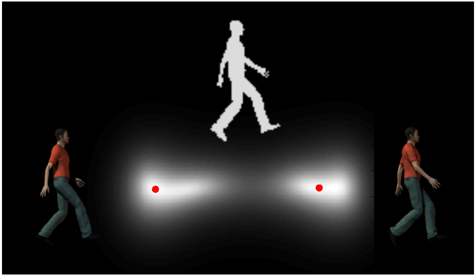
Video-based motion capture

- The previous systems are very expensive and/or invasive.
- We would like to capture motion from video.
- The holy grail: use a single camera.

Why is it difficult?

- Poor imaging: motion blurred, occlusions, etc.
- The mapping is generally multimodal: an image observation can represent more than one pose.



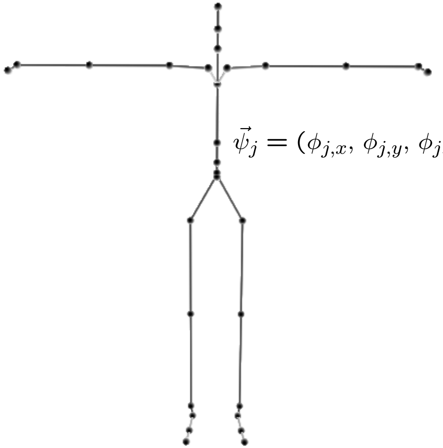


What are we going to cover in this class?

- Ways to capture human motion
- Human motion representations and direct kinematics
  - Kinematic tree
  - Representations: 3D locations, joint angles, axis angles, quaternions
- Models of human pose and motion
- Human motion synthesis
- Pose estimation from images
- Human motion classification

Human Motion Representations

- The human body can be approximated as a kinematic tree



$$\vec{\psi}_j = (\phi_{j,x}, \phi_{j,y}, \phi_{j,z})$$

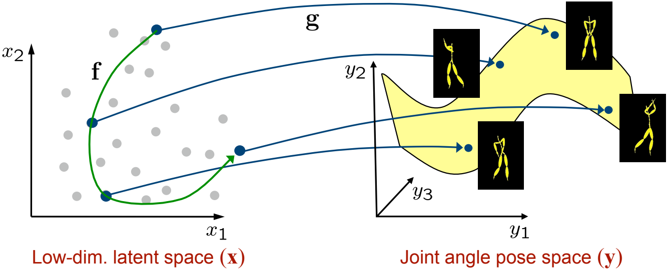
- Of course in reality it is much more complicated!

What are we going to cover in this class?

- Ways to capture human motion
- Human motion representations and direct kinematics
- Models of human pose and motion
  - Latent variable models: PCA, FA, GPLVM, etc.
  - Dynamical systems: LDS, etc.
- Human motion synthesis
- Pose estimation from images
- Human motion classification

Models of human pose and motion

- Latent variable models: PCA, FA, GPLVM, etc.
- Dynamical systems: LDS, etc.



- This will allow for synthesis (with and without constraints) and video-based motion capture.

What are we going to cover in this class?

- Ways to capture human motion
- Human motion representations and direct kinematics
- Models of human pose and motion
- Human motion synthesis
  - Inverse kinematics
  - ML approaches: NN (motion graphs), LVMs, etc.
  - Space-time constraints: physics.
- Pose estimation from images
- Human motion classification

Human motion synthesis: NN

- Using simple NN techniques

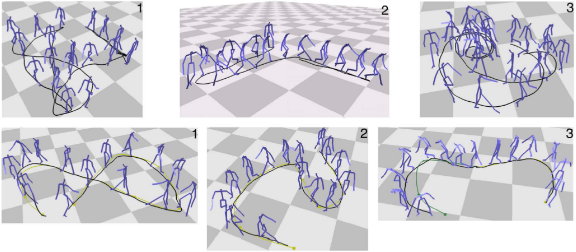


Figure: Motion graphs: [L. Kovar, M. Gleicher and F. Pighin, Siggraph 2002]

Human motion synthesis: NN

- Using simple NN techniques

Figure: Motion graphs: [L. Kovar, M. Gleicher and F. Pighin, Siggraph 2002]

Human motion synthesis: NN

- Using simple NN techniques

Figure: Motion graphs: [L. Kovar, M. Gleicher and F. Pighin, Siggraph 2002]

Human motion synthesis: LVMs

- Sophisticated latent variable models allow for less training data.

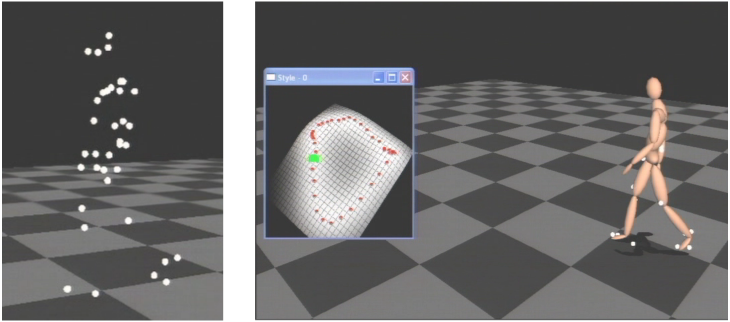


Figure: Syle-IK: [K. Grochow, S. Martin, A. Hertzmann and Z. Popovic, Siggraph 2004]

Human motion synthesis: LVMs

- Sophisticated latent variable models allow for less training data.
- Replay the same motion

Figure: Syle-IK: [K. Grochow et al., Siggraph 2004]

Human motion synthesis: LVMs

- Sophisticated latent variable models allow for less training data.
- Keyframing of joint trajectories.

Figure: Syle-IK: [K. Grochow et al., Siggraph 2004]

Human motion synthesis: LVMs

- Sophisticated latent variable models allow for less training data.
- Deal with missing data

Figure: Syle-IK: [K. Grochow et al., Siggraph 2004]

Human motion synthesis: LVMs

- Include dynamics.
- Good generalization properties

Figure: [R. Urtasun et al., ICML 2008]

Human motion synthesis: LVMs

- Sophisticated latent variable models allow for less training data.
- Animate from images

Figure: Syle-IK: [K. Grochow et al., Siggraph 2004]

- We would like less interaction to compute a pose from an image

Human motion synthesis: Physics

- Force the animation to satisfy the law of physics

Figure: [K. Liu and Z. Popovic, Siggraph 2005]

Human motion synthesis: Physics

- Force the animation to satisfy the law of physics

Figure: [Y. Abe, M. daSilva and J. Popovic, Siggraph 2005]

What are we going to cover in this class?

- Ways to capture human motion
- Human motion representations and direct kinematics
- Models of human pose and motion
- Human motion synthesis
- Pose estimation from images
    - Inverse kinematics
    - Discriminative models: NN, regression, structured-output.
    - Generative models: Kalmann filters, particle filter, etc.
    - Likelihood models
    - Priors: shape models, motion models, joint limits, physics.
- Human motion classification

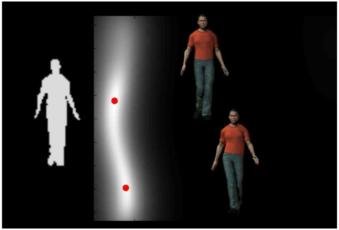
Discriminative approaches to tracking

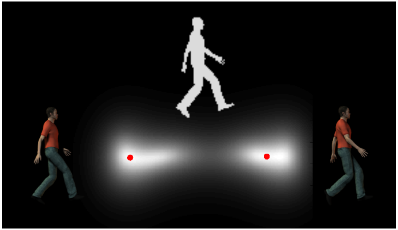
- Directly model the mapping from image observations (features) to pose


- The main challenges are:
  - Poor imaging: motion blurred, occlusions, etc.
  - Need of large number of examples to represent all possible poses.

Challenges for ML

- This problem cannot be solved directly as a regression task, since the mapping is multimodal: an image observation can represent more than one pose.





- It is typically addressed as a mixture of experts.

Humaneva

- Good results in control environments

Figure: [R. Urtasun and T. Darrell, CVPR 2008]

Generative model for video-based motion capture

- Generate a pose, and evaluate a likelihood function
- Large collection of image likelihoods
- Typically used with particle filters or minimization schemes
- Easy to employ priors
  - Shape models
  - Joint limits
  - Motion models
  - Physics

Shape priors: Implicit surfaces

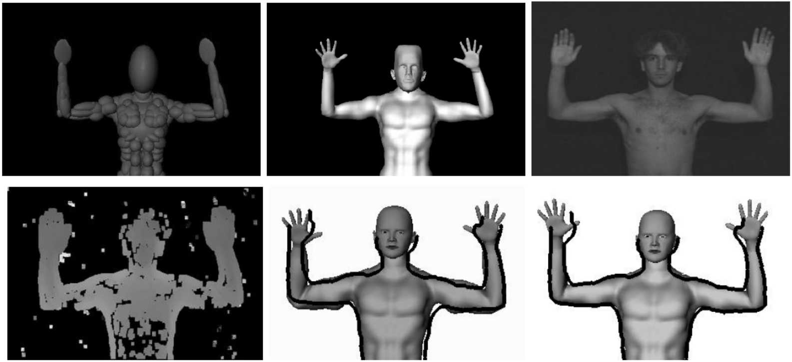


Figure: [Plaenkers et al. ECCV 2002]

Shape priors: scan

Figure: Scape model: [D. Anguelov et al. SIGGRAPH 2005]

Shape priors: scan

Figure: Scape model: [D. Anguelov et al. SIGGRAPH 2005]

Joint limit priors

- Simple joint limits.

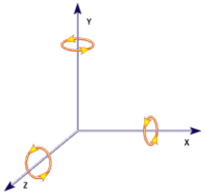


Figure: Min-max joint limits

Joint limit priors

- Simple joint limits.
- Sophisticated joint limits

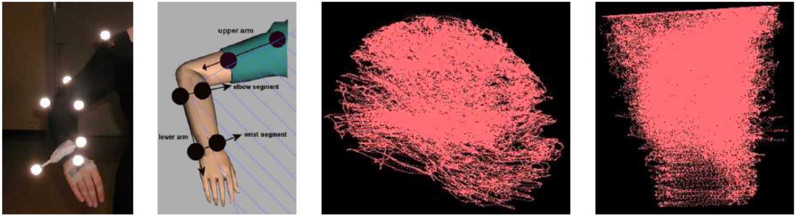


Figure: Learning joint limits from data, [L. Herda, R. Urtasun and P. Fua, CVIU 2005]

Joint limits [L. Herda, R. Urtasun and P. Fua, CVIU 2005]

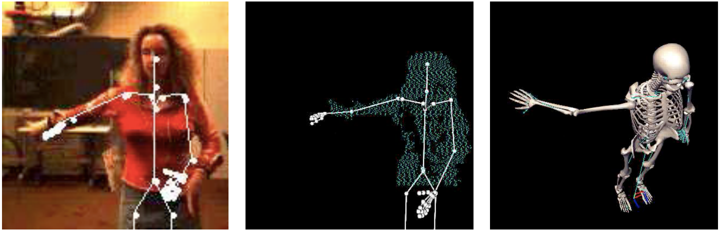


Figure: Tracking without joint limits

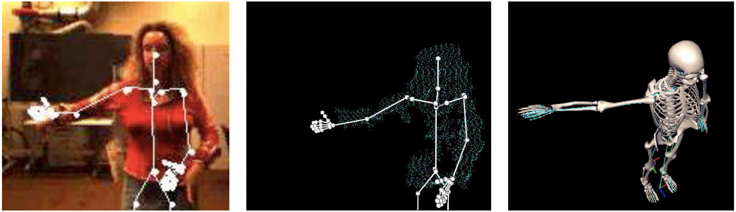


Figure: Tracking with joint limits

Motion priors for 3D people tracking

- Learn off-line prior models from Mocap.
- Use then online to constrain the tracking.

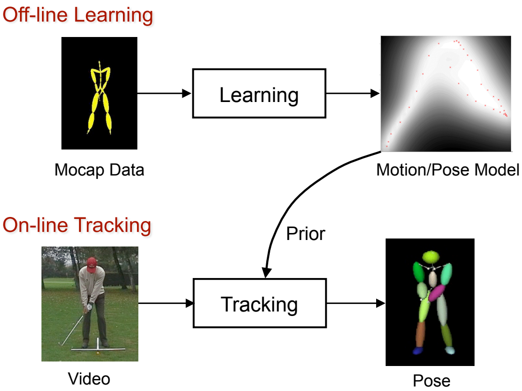
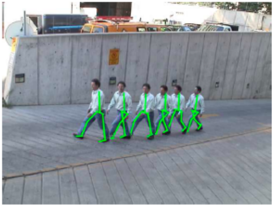
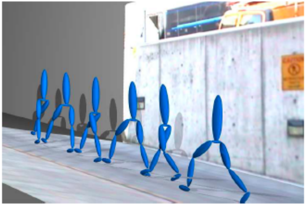
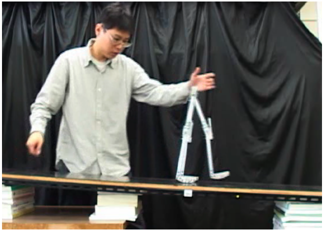


Figure: [R. Urtasun, D. Fleet and P. Fua, CVPR 2006]

Occlusion handling and training from a single example

Physical priors [Brubacker et al. CVPR 2008]

- Very new domain for people tracking.
- Very few approaches, mostly based on he passive walker

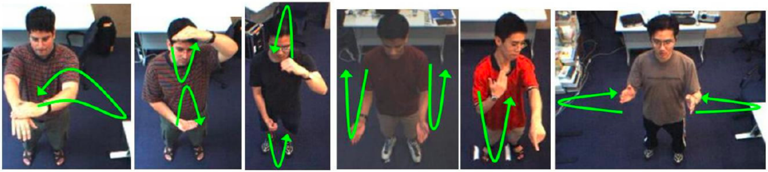





What are we going to cover in this class?

- Ways to capture human motion
- Human motion representations and direct kinematics
- Models of human pose and motion
- Human motion synthesis
- Pose estimation from images
- Human motion classification
  - Gesture recognition: Discriminative LVMs, Structured-output methods
  - Activity recognition
  - Gait analysis

Gesture recognition

- Traditional gestures



- Gestures for communication: Sign language

Activity recognition

- Simple activities: walking, running, jogging.
- Sport events: hierarchical roles

- Activity recognition on the wild: YouTube.

Gait analysis

- From different sources
  - mocap
  - video
- Different goals:
  - Surveillance
  - Orthopedics
  - Style

What to expect?

- We've seen a lot of cool animations... ;)
- ... expect also to see complex math :(
- Good understanding of the problems and the fundamental techniques ...
- ... from a practical and theoretical point of view
- Guided exercises + Research project

Other announcements

- I'm looking for PhD students!
- I'm happy to discuss research, come and see me!
- See you next week