Body Part Position Estimation

Jorik Spijkerman en Sanne Bouwmeester

Introduction

Body Pose Recognition

Kinect

Initial Pose

Automatically induce body position from depth point cloud

Video Example

https://youtu.be/ZXI6gko7kG4?t=1m

Background: ICP

Fixed model of body and skeleton

Movement is concatenation of rigid transformations

ICP to estimate what these transformations are

Estimate depth from 2 images using triangulation

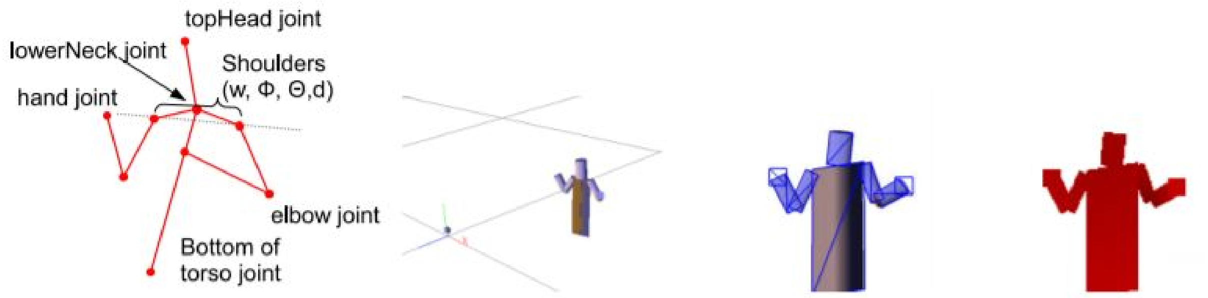
Nonlinear Body Pose Estimation from Depth Images
Daniel Grest, Jan Woetzel and Reinhard Koch

Background: Suggested Poses

Depth camera

Render depth cloud from Pose Skeleton

Compare them



(a) 3D Pose Skeleton (b) Synthesized Pose(3D) (c) Projected Pose (2D) (d) Rasterized Pose(2D)

Human Pose Estimation from a Single View Point, Real-Time Range Sensor

Matheen Siddiqui and G´erard Medioni

Data

Body position data relies on shape of body and clothes

Manipulate known data to unknown shapes and clothes

More data is useful for overfitting

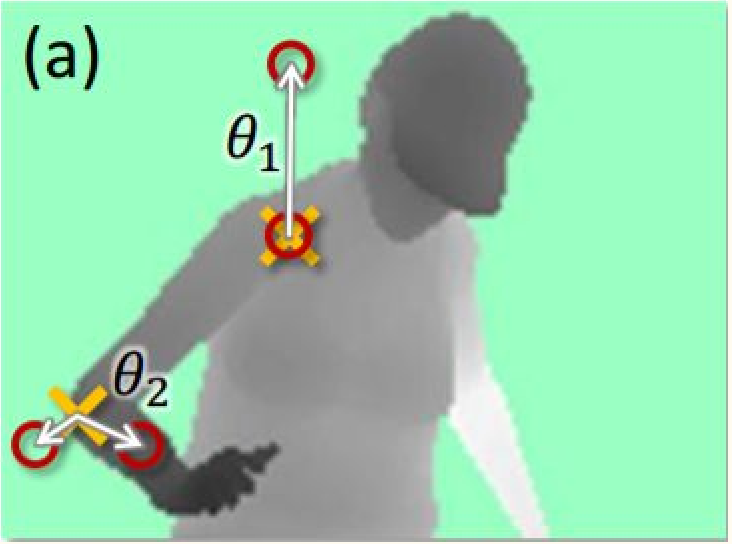
More data is useful for robustness

Add transformations and mirrored data

Classification

Depth Features

- Location
- Offset



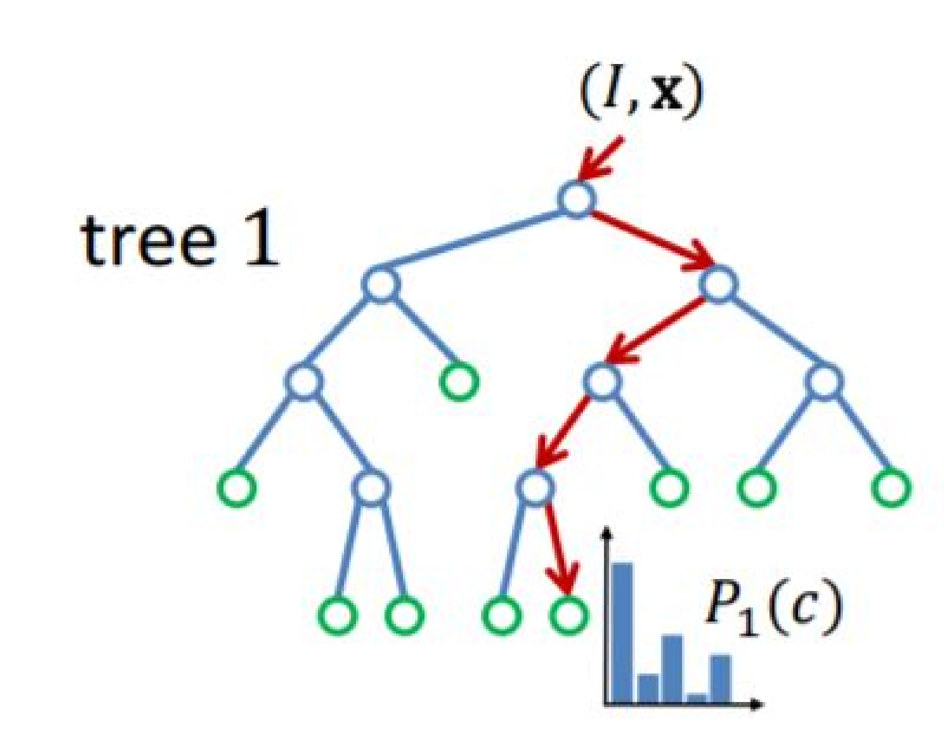
Random Forests

Goal: Classify different body parts

Split nodes
- A feature θ
- A threshold τ

Leaf node
- probability over body part classes

Several trees are averaged



Body part labeling

Local labels

Skeletal joints are used directly

Other body parts only used sparsely

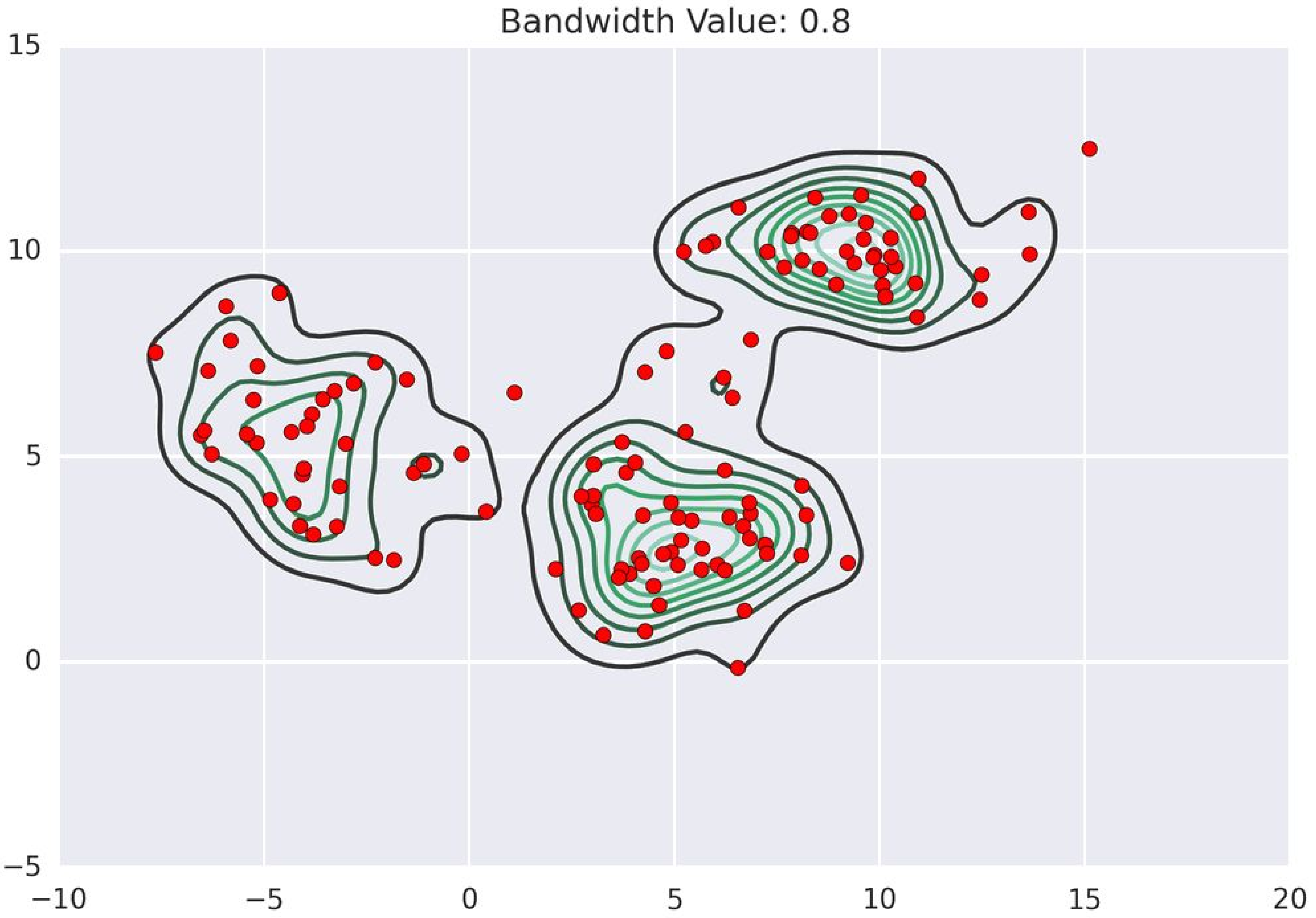
31 body part tags

Joint Position estimation

Input: Body part clusters with body part class labels

Pooling per body part, in particular joints

Mean shift


Bandwidth Value: 0.8
15
10
5
0
−5
−10
−5
0
5
10
15
20

Joint Position estimation

Input: Body part clusters with body part class labels

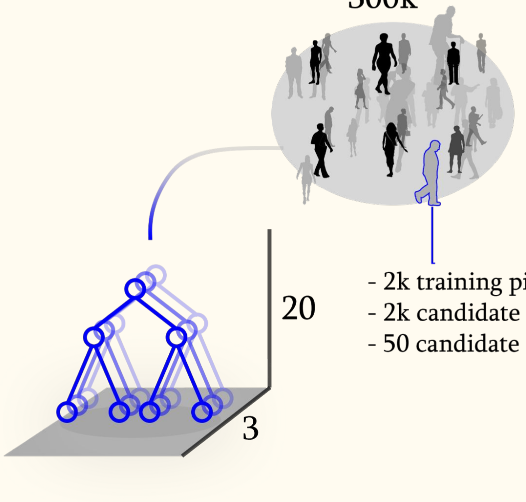
Pooling per body part, in particular joints

Mean shift

- Weighted Gaussian Kernel to compute mean
- Several joint proposals, weighted with confidence estimate
- Account for body depth

Experiments

Training



- 2k training pixels
- 2k candidate features (θ)
- 50 candidate thresholds (τ)

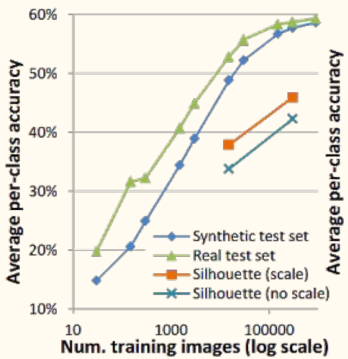
Testing

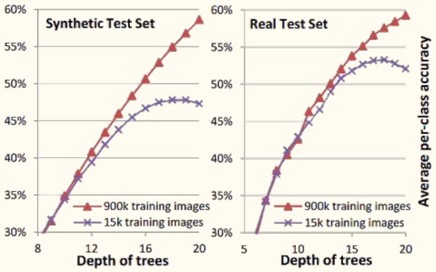
- Synthetic depth images (5k)
- Real depth images (8.8k)

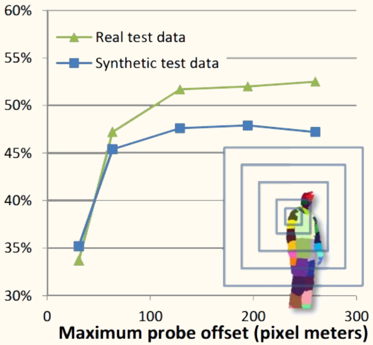
Results

Results

Classification

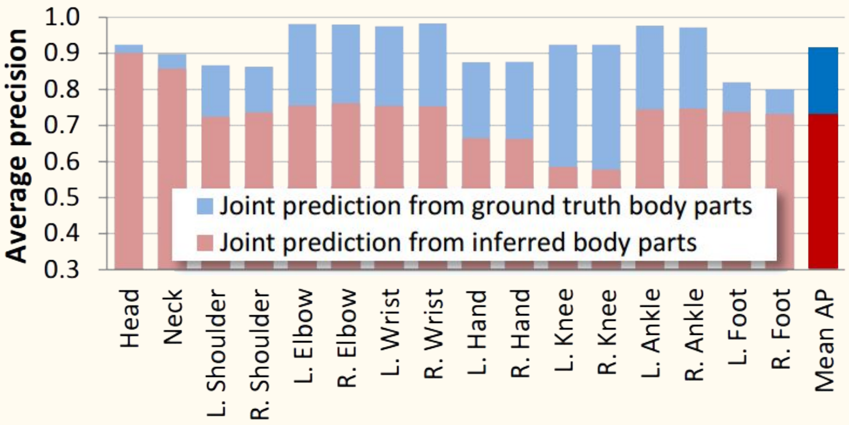






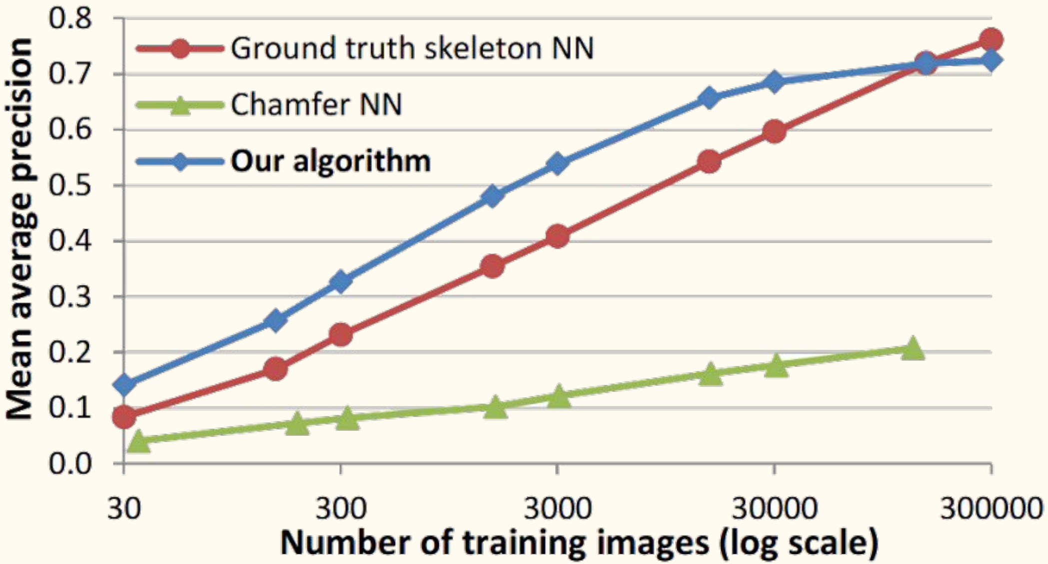
Joint precision (0.731 mAP)

- Synthetic test set, worse than real set

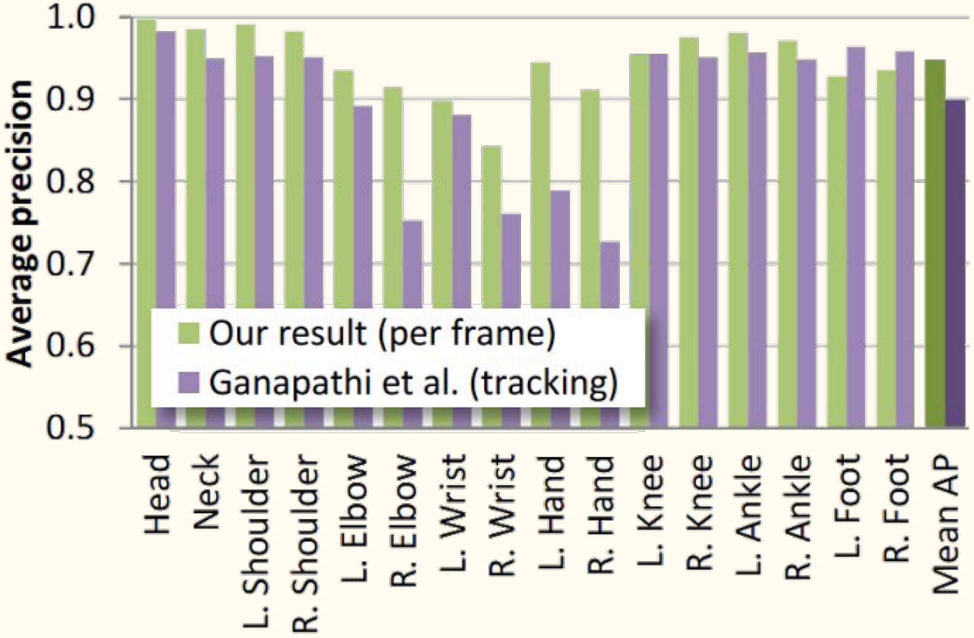


Joint prediction accuracy

In comparison(1)

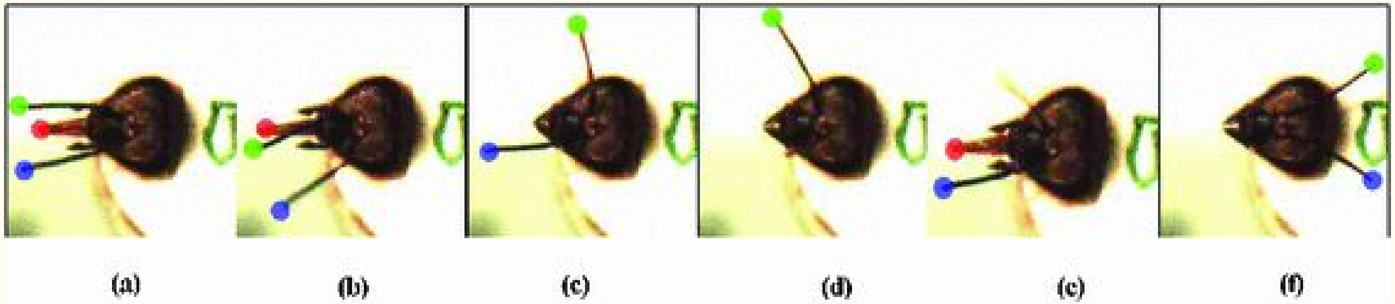


In comparison(2)



Discussion

- How is 3 trees enough?
- Deeper trees?
- Good generalization



Shen M., Duan L., Deussen O. (2016) Single-Image Insect Pose Estimation by Graph Based Geometric Models and Random Forests. In: Hua G., Jégou H. (eds) Computer Vision – ECCV 2016 Workshops. ECCV 2016. Lecture Notes in Computer Science, vol 9913. Springer, Cham

Conclusion

- Fast
- Accurate

- Serves as a great fundament to other research

Questions?