ROBUST PEOPLE DETECTION BASED ON APPEARANCE AND SHAPE

Pishchulin L., Jain A., Wojek C., Andriluka M., Thormählen T., Schiele B. - MPI Informatics, Saarbrücken
{leonid, ajain, cwojek, andriluk, thormae, schiele}@mpi-inf.mpg.de

max planck institut informatik

Abstract

In this work, we investigate how 3D shape models from computer graphics can be leveraged to ease training data generation. In particular we employ a rendering-based reshaping method to generate thousands of synthetic samples from only a few persons and views. Experiments on a challenging multi-view dataset indicate that the data from just eleven persons suffices to achieve good performance, while combination of our synthetic data with real data outperforms even the state of the art.

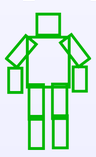
Motivation

- Collection of large representative dataset is hard
- Manual data annotation is tedious and noisy
- Synthetic data produced by current methods (e.g. [5]) lacks realism and models to control its variability

People Detection Model

Pictorial structures model [2]

- Flexible configuration of body parts with pose prior
- AdaBoost part detectors learned from dense shape context descriptor
- Inference by sum-product belief propagation



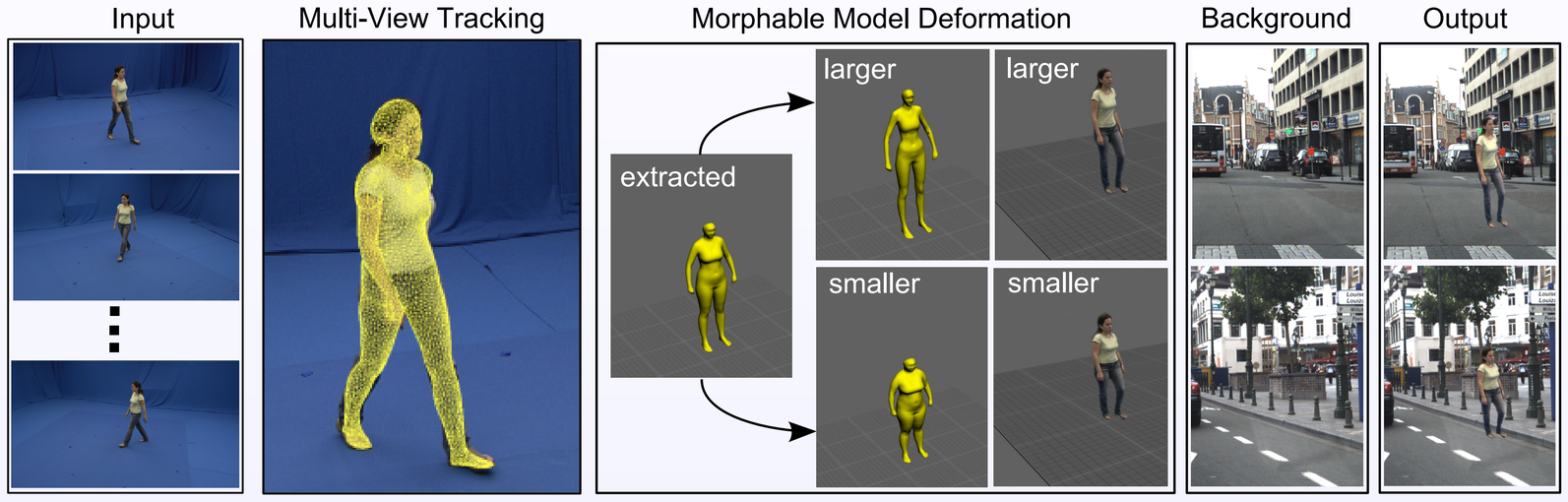
3D Body Model

Statistical 3D human shape model [3]

- Learn shape from 3D body scans, represent via PCA
- Use kinematic skeleton for motion modeling

Proposed Approach (CVPR'11)

1. Generate realistic synthetic data by MovieReshape [4]
   - Employ statistical 3D model of human shape
   - Fit the model to multi-view sequences
   - Sample 3D model height parameters to drive 2D image deformations
2. Combine reshaped humans with backgrounds



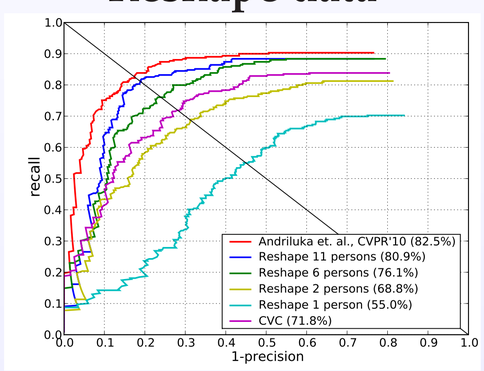
⇒ **Realistic distributions of human appearance and shape**

Results

- Samples from training datasets

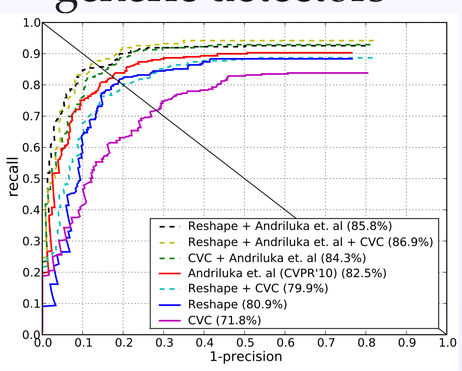
- Evaluation on challenging multi-view data [1]

Results using Reshape data

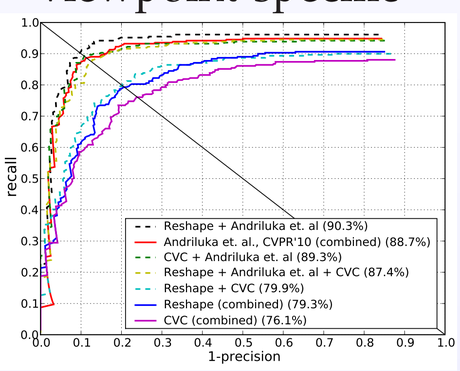


Combining different datasets

generic detectors



viewpoint-specific



References

[1] Andriluka M., Roth S., Schiele B., 'Monocular 3d pose estimation and tracking by detection', in *CVPR*, 2010

[2] Felzenszwalb P .F., Huttenlocher D. P., 'Pictorial structures for object recognition', in *IJCV*, 2005

[3] Hasler N., Stoll C., Sunkel M., Rosenhahn B., Seidel H.-P., 'A statistical model of human pose and body shape', in *Eurographics*, 2009

[4] Jain A., Thormählen T., Seidel H.-P., Theobalt C., 'Moviereshape: Tracking and reshaping of humans in videos', in *SIGGRAPH Asia*, 2010

[5] Marin J., Vazquez D., Geronimo D., Lopez A., 'Learning appearance in virtual scenarios for pedestrian detection', in *CVPR*, 2010