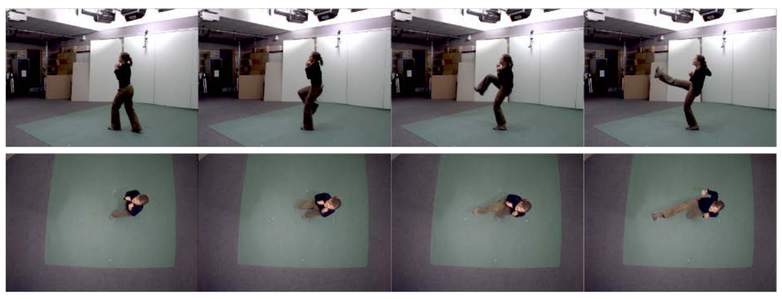
Unlabelled 3D Motion Examples Improve Cross-View Action Recognition

Ankur Gupta, Alireza Shafaei, James J. Little, Robert J. Woodham
University of British Columbia

BMVC 2014 NOTTINGHAM

Cross-view action recognition problem



Given:
- Correspondence labels
- Or partial annotations in the test view.

What if no supervision is available?

- Previous methods depend on multi-view, annotated video data to learn feature transformations.
- Or use 3D or 2D pose estimation which is often unreliable.
- Our solution: learn human motion specific geometric transformations of features using unlabelled mocap.

Learning codeword transformation

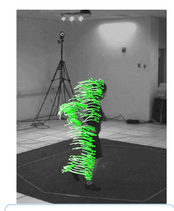
We assign each trajectory feature to its closest codeword. Given the example pairs of corresponding codewords for a viewpoint change Δ:

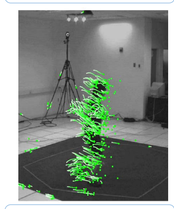
- We generate a matrix N, where each entry is the probability of transformation from one codeword to the other.
- We simply count the co-occurrences of the observed source-target codewords to estimate these probabilities.

The transformation function $f_\Delta(\boldsymbol{x}) = N\boldsymbol{x}$ maps a BoW based action descriptor $\boldsymbol{x}$ from the source view to the target view.

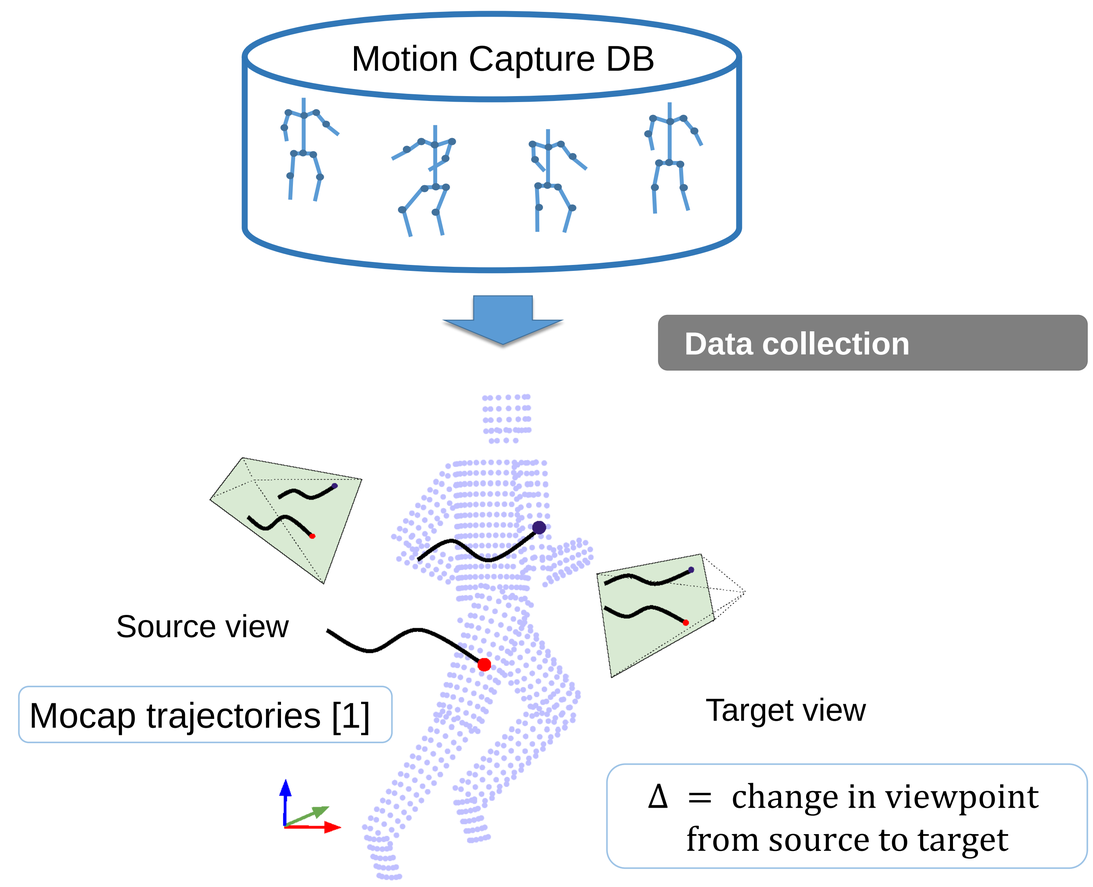
Summary

- We generate mocap trajectory features and their corresponding transformed versions using unlabelled motion capture data.
- We learn how the codewords for these features transform as a function of the change in viewpoint.
- Assuming similarity in shape between mocap trajectories and dense trajectories, we can apply the transformation function to the BoW of dense trajectories to "hallucinate" multi-view examples.


Mocap trajectories


Dense trajectories

Our Approach



Results

Method	Average accuracy
Ours	**71.7%**
nCTE [1]	67.4%
w/o Augmentation	62.1%
Hankelets [3]	56.4%

Selected References

[1] Ankur Gupta, Julieta Martinez, James J. Little, and Robert J. Woodham. 3D Pose from Motion for Cross-view Action Recognition via Non-linear Circulant Temporal Encoding. In CVPR, 2014.
[2] Heng Wang, Alexander Klaser, Cordelia Schmid, and Cheng-Lin Liu. Action recognition by dense trajectories. In CVPR, 2011.
[3] Binlong Li, Octavia I. Camps, and Mario Sznaier. Cross-view Activity Recognition using Hankelets. In CVPR, 2012.

NSERC CRSNG

Project page & code: http://cs.ubc.ca/research/motion-view-translation/

WestGrid compute • calcul CANADA