A Hierarchical Model of Shape and Appearance for Human Action Classification

Juan Carlos Niebles 1,2
Feifei Li 1,3

http://vision.cs.princeton.edu

University of Illinois at Urbana – Champaign
Universidad del Norte, Colombia
Princeton University

Highlights and Summary
- A novel model for human action categorization from video sequences.
- Our model can be characterized as a constellation of bags-of-features.
- Use of hybrid features: combines both static shape and spatio-temporal features.

Hybrid features
- Original frame
- Feature detection
- Shape context
- Feature description
- Membership assignment
- Video frame: \( w = (x, a) \)

Hierarchical model
- Mixture components
- Part layer
- Feature layer

Algorithm
- Feature extraction and description
- Learning
- Recognition

Previous Works
- Small number of features
- Large number of features
- Strong shape representation
- No geometrical or shape information

Learning
- Estimate model parameters using EM
- E-step:
- M-step:

Experimental Results
- 9 action classes, performed by 9 subjects
- Leave one out cross-validation
- Video classification performance: 72.8%

Conclusions
- The constellation of bags-of-features is able to capture semantic information of human action classes.
- Combines hybrid features: static shape features and dynamic motion features.
- Capable of classifying in both frame based and video based manner.