Georgia College of Tech Computing

# Augmenting Bag-of-Words: **Data-Driven Discovery of Temporal and Structural Information for Activity Recognition**



Vinay Bettadapura<sup>1</sup>, Grant Schindler<sup>1</sup>, Thomas Plötz<sup>2</sup>, Irfan Essa<sup>1</sup> <sup>1</sup>Georgia Institute of Technology, <sup>2</sup>Newcastle University

- Cumulative Encoding: Modeling the cumulative time
- Pyramid Encoding: Builds a pyramid of n-grams for looking at the events at multiple scales of resolution





 Binary classification: parking vs. non-parking

## 2. Surgical Skill Assessment



	Respect	Time and	Instrument	Suture	Flow of	Knowledge	Overall	Average
	for tissue	motion	handling	handling	operation	of procedure	performance	accuracy
M1: BOW baseline	66.67%	50.79%	50.79%	69.84%	49.21%	60.32%	52.38%	57.14%
M2: BOW + Time	69.84%	66.67%	65.08%	69.84%	63.49%	74.60%	68.25%	68.25%
M3: Our encoding	73.02%	74.60%	68.25%	73.02%	66.67%	80.95%	71.43%	72.56%

### **3. Clustering Soccer Player Activities**



- 2,140 vehicle tracks with 448 parking activities
- ROC Curve (Ocean-City) *k*-NN (VSM) BoW + Time Interspersed Encoding Cumulative Encodin Pyramid Encoding
  Regular Expression. With BoW With BoW + Time With our Encoding With BoW With BoW + Time With our Encoding -alse Positive Rate
- 64 videos from 16 students
- 20 minute long videos, on avg.
- Ground truth by expert surgeon <sup>1</sup>

 Initial BoWs built using HOG3D detector and HOF descriptor (with k=50), which are then augmented using our technique

- Full-length soccer game videos
- Cluster players based on their roles: Defense, Striker, Goalkeeper, Referee





### 4. Wide Area Airborne Surveillance

 Massive wide-area simulation developed by US military • Activities of 4,623 individuals for a duration of 46.5 hours • Total of **180 events** (like "eat lunch", "enter vehicle", "move", "wait", etc.) with a total of 544,777 activities • Ground truth: **Professions** of all the individuals with **23** out of the 4,623 are members of a terror group



• We can see 30.04% improvement in performance • A third of the suspect group gets correctly classified using

Augmented BoW, which previous methods had failed to capture

### Software and Acknowledgements

### Software: http://www.cc.gatech.edu/cpl/projects/abow/

• Funded by the U.S. Army Research Office (ARO) and DARPA under Contracts W911NF-11-C-0088 and W31P4Q-10-C-0262 • Parts of this work have also been funded by the RCUK

Research Hub on Social Inclusion through the Digital Economy (SiDE), and the German Research Foundation (DFG, Grant No. PL554/2-1).

We thank Kitware, Disney Research and DARPA for providing the Ocean City, Soccer and WAAS datasets, respectively.