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

Vinay Bettadapura¹, Grant Schindler¹, Thomas Ploetz², Irfan Essa¹
¹Georgia Institute of Technology, ²Newcastle University

---

**Research Goals**
- Robust modeling and recognition of complex long term activities (ranging from several minutes to several days/weeks/months)
- Develop a purely data-driven approach that works when the structure and topology of the activities are not known a priori

**Applications**
- Activity Recognition, Anomaly Detection, Functional Categorization and Skill Assessment

**Discovering Temporal Information**

**Activity:** A sequence of discrete events. Example: A vehicle driving activity

**Observable Event:** A short observable occurrence

**Temporal Event:** Elapsed between consecutive events

**Temporal Event:** Temporal information

**Capturing Global Structure**

N-grams fail to capture long-range patterns across data. Regular-Expressions are good for such tasks.

**Encoding Local Structure**

Standard BoW ignore the order of words. To fix this, we use n-grams to encode local structure of the activities

A1:

A2:

We propose three n-gram encoding schemes:
- **Interpersed Encoding:** Modeling the time elapsed between pairs of consecutive events
- **Cumulative Encoding:** Modeling the cumulative time taken by sub-sequences of events
- **Pyramid Encoding:** Builds a pyramid of n-grams for looking at the events at multiple scales of resolution

**Results**

1. **Ocean City Surveillance**
   - 2,140 vehicle tracks with 448 parking activities
   - Binary classification: parking vs. non-parking

2. **Surgical Skill Assessment**
   - 64 videos from 16 students
   - 20 minute long videos, on avg.
   - Ground truth by expert surgeon

3. **Clustering Soccer Player Activities**
   - Full-length soccer game videos
   - Cluster players based on their roles: Defence, 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 54,777 activities
   - Ground truth: Professions of all the individuals with 23 out of the 4,623 are members of a terror group

**Software and Acknowledgements**

- **Software:** [http://www.cc.gatech.edu/cpl/projects/abow/](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.