ISOMER: Informative Segment Observations for Multimedia Event Recounting

Chen Sun1, Brian Burns2, Ram Nevatia2, Cees Snoek3, Bob Bolles2, Greg Myers2, Wen Wang2 and Eric Yeh2
1 University of Southern California 2 SRI International 3 University of Amsterdam

Introduction

- **Goal**
  - Multimedia Event Detection (MED): Video level event classification
  - Multimedia Event Recounting (MER): Key evidence recounting for a video event

- **Challenges**
  - Large video dataset (~100,000 test videos)
  - Huge intra-event variations
  - Irrelevant segments within video events

- **Our Approach**
  - Utilization of multimodal concept detectors
  - Event-specific key evidence localization and weighting
  - Concept selection and fusion for MER generation

Event Classification (MED) and Recounting (MER)

- **Event Classification**
  - Train SVM classifiers for every mode

- **Visual & Motion Observations**
  - Align visual concept responses (CR) with motion CR via max-pooling
  - Learn evidence templates from video-level CR
  - Rank video segments and concepts with evidence templates
  \[ P_k(c) = \frac{1}{1 + \exp(-w_k X_c)} \]
  - Map irrelevant concept names to more abstract terms
  - Maintain a white list of event related concept names

- **ASR Observations**
  - Rank text by learning sparse linear SVMs with log-mapped n-gram counts
  - Post-process with WordNet and topic modeling techniques

- **OCR Observations**
  - Log-linear classifier over word frequencies

- **Fusion of Observations**
  - Use thresholds and set a observation count limit

Dataset

- **TRECVID MER 2013 Progress Set**
- **20 events**

Evaluation results

- **MER applied only to videos with high confidence MED scores**
- **Human evaluators at NIST**
- **BEST MER text description quality**
  - 2.5 out of 4.0
- **Able to locate a small proportion of informative video segments**
  - 40% of video length on average
  - 64% accuracy (percentage of correctly assigned event labels by evaluators)

Conclusion

- **Concept detectors are informative but noisy**
- **Event prior from MED is important for MER**
- **Rank noisy concept detections**

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA) via the Department of Interior National Business Center (DOI/NBC), contract number D11PC0067. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.