

GEOINT 2013: **OPERATIONALIZING INTELLIGENCE** **FOR GLOBAL MISSIONS**

**Matt Turek, Sangmin Oh,
Anthony Hoogs**
Kitware, Inc.

Large-Scale Understanding of Crowd-sourced Multimedia Relationships

This work was partially supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D11PC20069. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI/NBC, or the U.S. Government.



Lots of web video...

...what does it contain?



Complex events

- Flash mobs, weddings, birthday parties, ...

Relationships between places, people, things

- Videos with similar scenes, objects, attributes, events

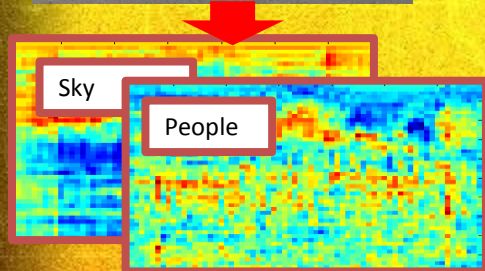
Indicators of population sentiment

- Level of violence, shouting, noise; happiness, contentment, peacefulness

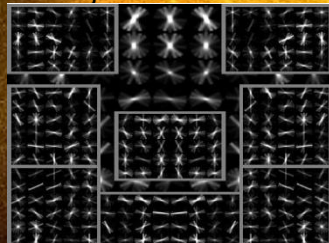
→ *Understanding Internet Video Relationships*

Video is Characterized by Features

Object Bank (~170 objects)

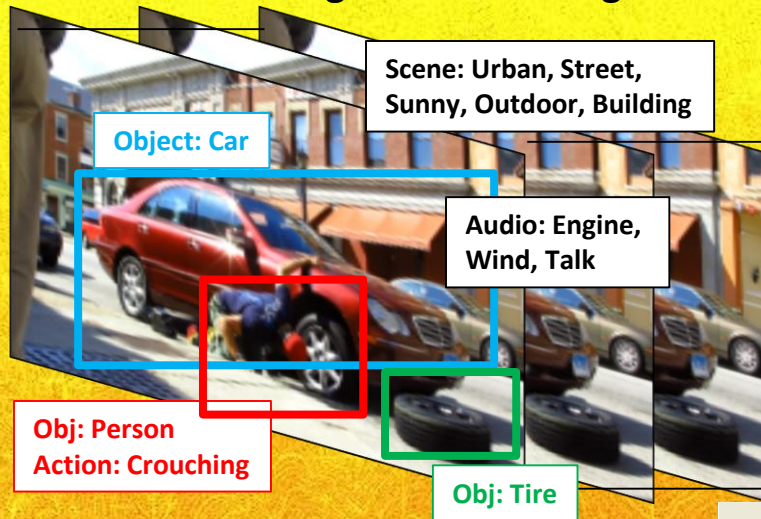


Bottom-up Discovered
Object Detections



Person with Birthday cake

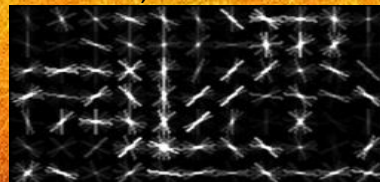
GENIE: General Engine for Indexing Events



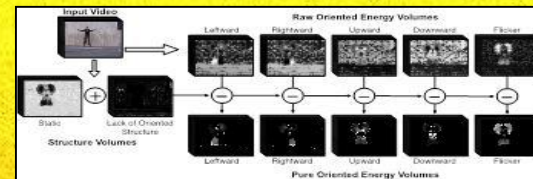
Tracking &
Videography



Histogram of Oriented
Gradients, Texture



Action Bank (~200 Actions)



Scene (~100 Attributes)

Indoor/outdoor lighting
Emotion Functions
Materials Viewpoint

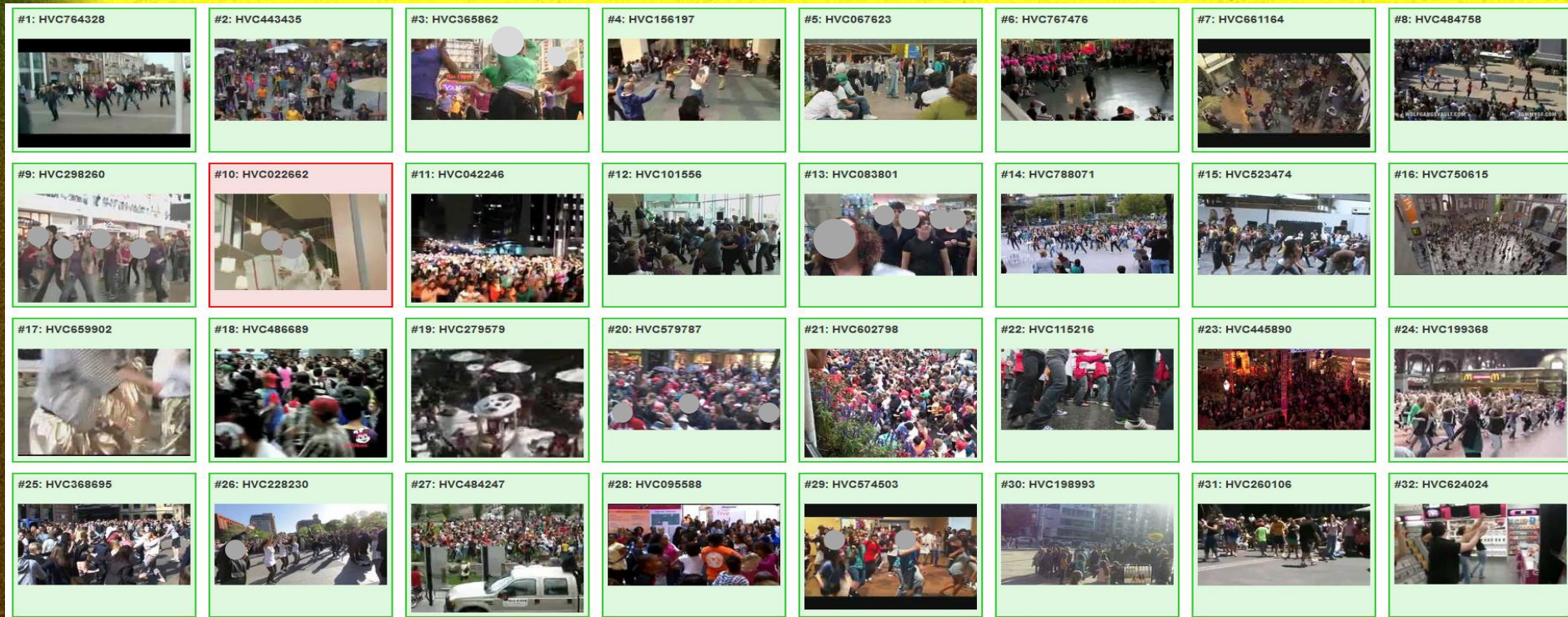
Audio (MFCCs, ASMs)



Audio Bank (~50 sound)

Engine Human Chat
Explosion Animal W

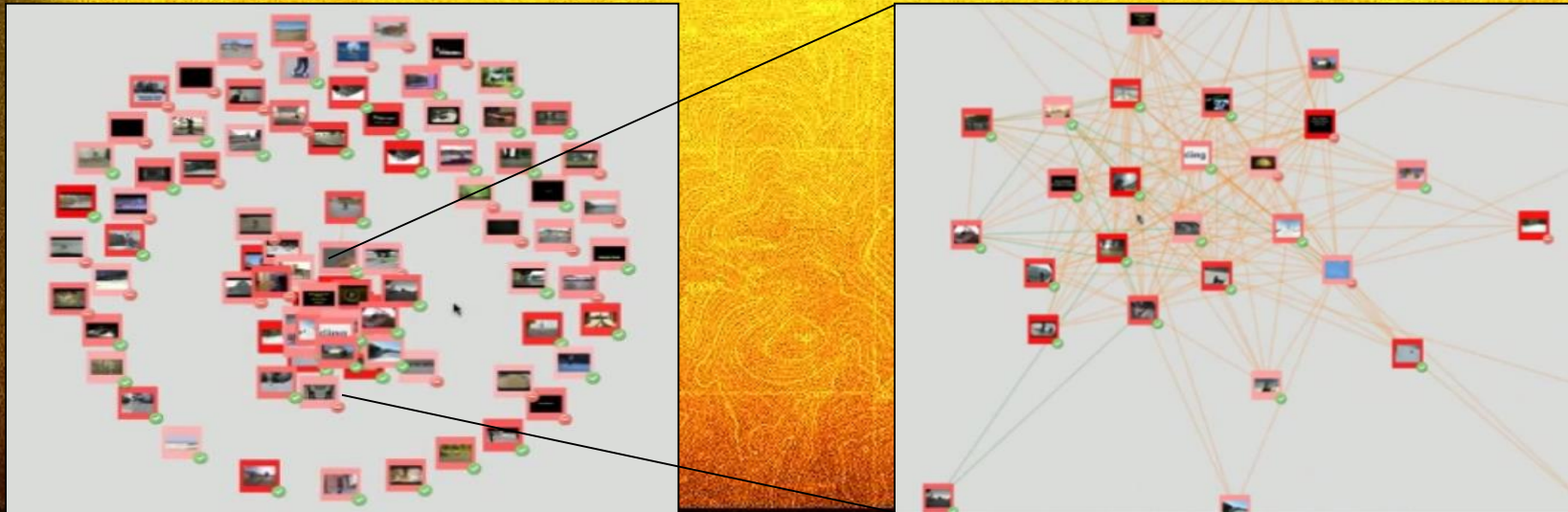
GENIE Top 32 for “Flash Mob”



Precision @ 32 = 96% as shown; average precision = 74.3%; archive contains 26K videos including ~100 true positives

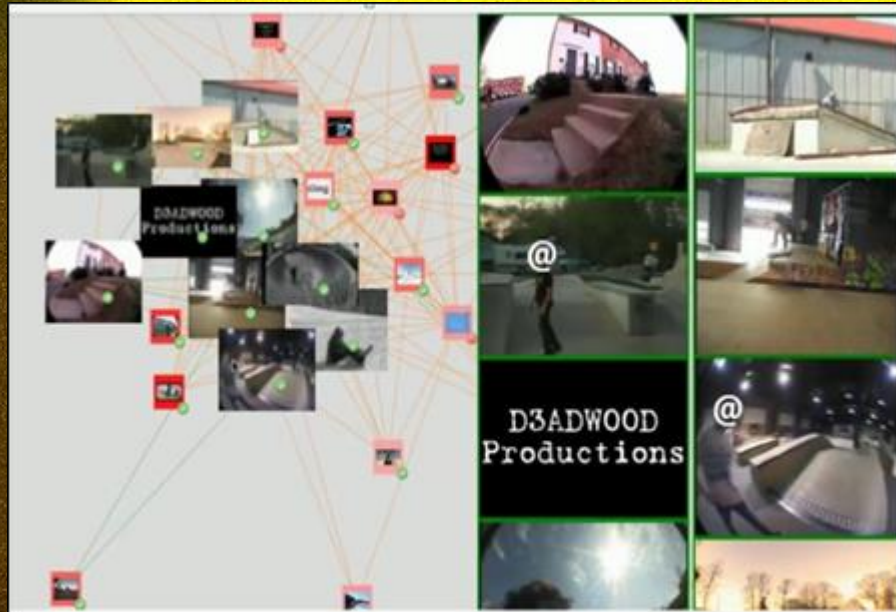
GENIE Graph-based Analysis

- Quickly triage retrieval results by their similarity and variation
- Nodes: images, videos, other multimedia items
- Edges: exists if two items are similar
 - Different colors show different similarity measures (e.g., features)
- Similar items are close, dissimilar items are far



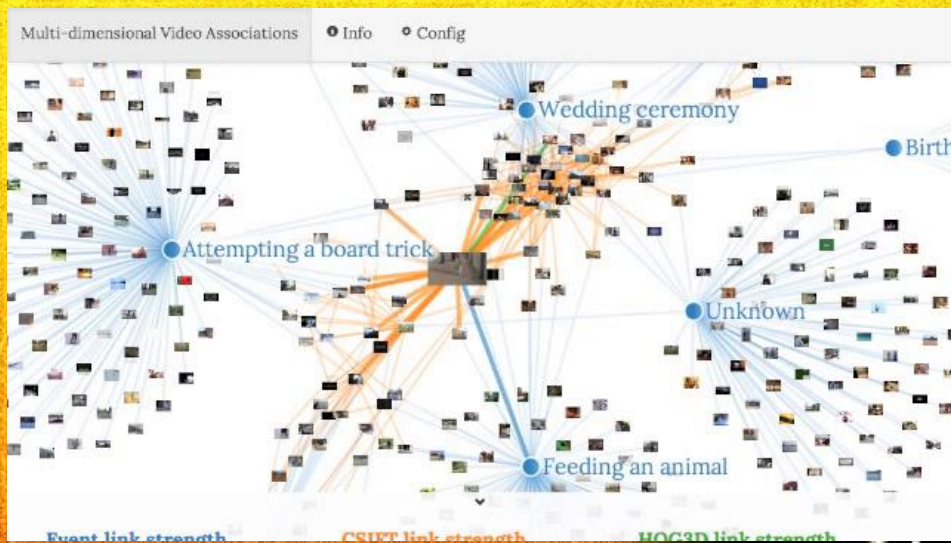
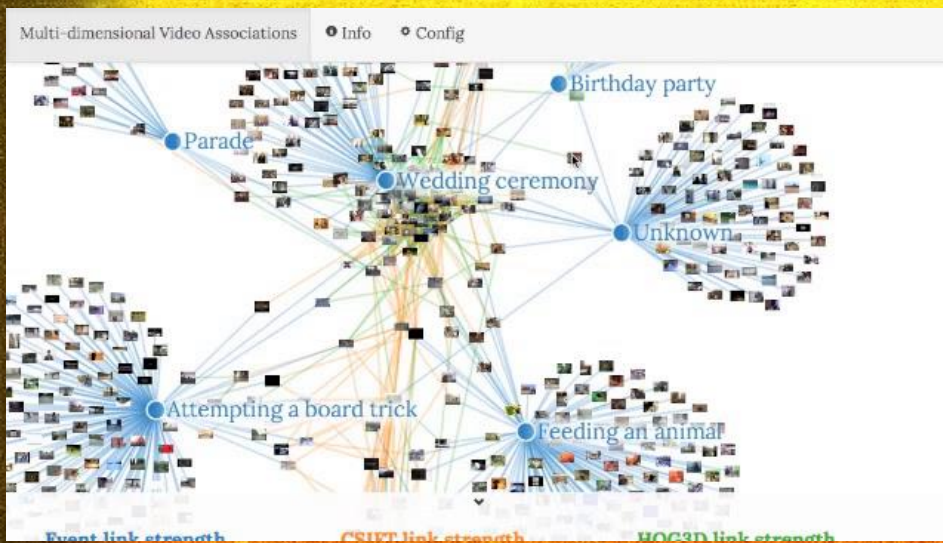
GENIE Graph-based Analysis

- Iterative Query Refinement (IQR) process learns from quick user feedback to rapidly improve results



Graph-based Analysis across Events

- Videos are linked by event type and visual features
- Analyst can adjust weights of link types to examine relationships
- Videos located between event types indicate multiple aspects of visual characteristics.



Takeaways

- Large-scale internet video understanding is ready for transition to focused intelligence operations
- Find complex events, objects, links
- Visualize and browse large internet video collections and query results

