Tools for Video Analysis

Alex Hauptmann
Carnegie Mellon University
alex@cs.cmu.edu
Vision

Ubiquitous camera phones allow dramatic events to be captured on video and rapidly shared via social media. Our goal is to create computer vision and machine learning-based tools to rapidly process and analyze this video. The tools are designed to make public safety analyst workloads less stressful and more manageable. They seek to augment the expertise of the analyst, not to replace decision-making. Our work ranges from satellite imagery analysis in conflict areas to urban event reconstruction based on user uploaded images and video. We also develop tools for large scale mining of internet video. Example tools are:

- synchronized event video visualization,
- video geo-localization,
- audio synchronization,
- group counting,
- gunshot detection,
- event analysis in online video
Long-term Surveillance Video Analysis

Large surveillance video set

Summarization-by-tracking

Visual diary

Pose Estimation

Detect “taking a bite”
We collected a urban event dataset containing social multimedia data of the Boston Marathon 2013 finish line bombing event, including Google StreetView imagery, map meta data and digital elevation map data. Test video clips are labeled with geolocation and aligned to a global timeline.

- 1,066 relevant video clips
- 22,912 geo-tagged related images
- 1,760 satellite images
- 5h 29m 41s video length
- Reference labels:
  - 109 event video clips with geo-location
    - precision within 10 meters
  - 44 event videos aligned to global timeline
    - precision within 1s
Boston Marathon Dataset

- Diverse on data types
- Diverse on timeline
  - pre-explosion
  - 1st explosion
  - 2nd explosion
  - post-explosion
- Diverse on camera positions
  - 45 degree view
  - bird's eye view
  - street view
Ground Camera Location Reconstruction

Blue crosses: matched regions
Star: camera location
Video Localization

- Uses reliably geotagged images from services such as Google StreetView and Flickr to suggest potential locations for scenes from metadata-free videos.
- Analysts can confirm or reject these suggestions through a custom-built viewer interface based on visual cues.
Synchronized video player

This tool helps visualize, geo-locate, and simultaneously play and adjust the offset of multiple synchronized videos. Features:
• Play multiple videos synchronously.
• Location displayed if specified for each video.
• Allows refinement of time offset or location information for each video.
This tool creates a unique sound print for a video clip using an algorithm that recognizes a standardized vocabulary of “features” (such as screaming and explosions). It compares the sequence of these features in each clip to all others and looks for reasonable matches. These matches can be further confirmed by a human.
Person Detection and Counting

This tool detects persons and counts persons in each frame of a video. A history of the number of people detected in each frame is shown.
Gunshot Detection

This tool detects gunshots in the audio track of a video, and lets the analyst know at what time gunshots likely occurred. It also provides an estimate of the number of gunshot events taking place in the video.