Real-Time Video Analytics for Situation Awareness

Presenters: Salvador Medina, Junwei Liang
PI: Alexander Hauptmann
Carnegie Mellon University
This presentation was produced by guest speaker(s) and presented at the National Institute of Standards and Technology’s 2019 Public Safety Broadband Stakeholder Meeting. The contents of this presentation do not necessarily reflect the views or policies of the National Institute of Standards and Technology or the U.S. Government.

Posted with permission
Introduction

• Ubiquitous camera phones capture public safety events and media is rapidly shared

• Goal: develop **video analytics tools** based on AI
  – Harness the information of event-based video recordings
  – Make video analysis for public safety less labor-intensive and more manageable at a large scale
  – Enhance the decision making capability of analysts
Overview

Event Begins

Information Flow

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-Localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization

Command Center
Social Media Event Detection & Analysis

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization

Command Center

Event Begins

Information Flow
As a major public safety event unfolds, social media is an important source of information.

**Problem:**
- Many social media entries are unrelated or uninformative.

**Solution:**
- Filter “useful” entries with text, image, and/or video.
- Automatically identify useful social media posts in real-time.

![Image of a diagram showing event detection and analysis process.](image-url)
That’s some of the damage caused by high winds and heavy rain in Goldsboro. Lights and some traffic signals are out, neighborhood streets are flooded. Avoid downed power lines! #abc11 https://t.co/D54qJlQmnl
AnonUser @anonuser Oct 3, 2017 10:51:19 PM GMT-07:00 DST
Look at the window the shooting is from. Real or fake?
Crowd Counting

- **Goal**
  - Count number of persons in real-time in a given scene
- **Possible applications:**
  - Occupancy monitoring for safety
  - Situation assessment
  - Crowd management
  - Response coordination
Crowd Counting
Traffic Accident Prediction

Event Begins

Information Flow

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization

Command Center
Object Detection & Tracking in Surveillance Videos

- Our tools are currently among the best in these scenarios

* Source Code and Model: https://github.com/JunweiLiang/Object_Detection_Tracking
Traffic Accident Prediction

• Predict and detect crashes from surveillance cameras
  – Fast notifications to first responders
  – **Ready to use on any camera stream**
• Proactive safety check
  – Speed and distance check of normal traffic flow
  – Provide insight about high-risk intersections
Traffic Accident Detection - Video Example

Traffic Danger Recognition With Surveillance Cameras Without Training Data
Person Identification

Goal: Locate person-of-interest in surveillance videos solely based on appearance
Person Identification
Gait Recognition for Person identification

with bag, 54°, id:1

with coat, 90°, id:1

normal, 144°, id:1

PTZ Camera Automated Capture

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization
Last Year’s Auto Person Capture

- Used Pan-Tilt-Zoom Camera
- PTZ camera system can take high-resolution pictures of people that are far away
Last Year’s Auto Person Capture

- Problem: PTZ camera movement has a delay
Last Year’s Auto Person Capture

• Pan-Tilt-Zoom Camera
  – We can see that the PTZ movement takes time
  – Therefore we need a system that:
    • predicts people’s trajectories
    • filters out targets if the system can predict activities
Can you predict what’s next?

Where are they going to go?

What are they going to do?
How can a system predict a person’s future?
How can a system predict a person’s future?

- People navigate in a scene with a specific purpose in mind
- People’s purpose can be inferred from:
  - their appearance
  - body language
  - nearby environment
Our Model - NEXT

- We designed an AI model with a *Person Behavior Module* and *Person Interaction Module* which considers:
  - Target person
  - Target person interactions with the scene and other objects.

* Source Code and Model: https://github.com/google/next-prediction
Demonstration - Single Person Prediction

Yellow path: observable trajectory; Heatmap: trajectory prediction; Text: activity prediction
Automatic Person Picture Capturing Using Pan-Tilt-Zoom Cameras
Automatic Person Picture Capturing Using Pan-Tilt-Zoom Cameras

Activity of Interest: Opening Trunk of a Car
Video Synchronization

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization
Video Synchronization

- Videos from social media usually do not have any meta-data like GPS and time.
- It is necessary to put all relevant videos in a global timeline first to understand an event.
Video Synchronization

• This year we added frame-accurate synchronization tool for video alignment
## Camera Geo-Localization

### Event Begins

### Information Flow

#### Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

#### Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

#### Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization

#### Command Center
Camera Geo-Localization

– Usually social media videos have no GPS metadata
– To locate a video we match Google Street View images to video frames to infer its position
3D Scene Reconstruction

**Real-Time Event Analysis**
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

**Event Reconstruction**
- Camera Geo-localization
- 3D Scene Reconstruction

**Shooter Event Analysis**
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization

Command Center

Event Begins

Information Flow
3D Reconstruction with Camera Localization

• Synchronizing multiple videos at a large scale makes it hard to understand the situation
• Our goal is:
  – To reconstruct the scene in 3D to project videos into augmented reality to understand the event
  – Pinpoint each camera location
Demonstration – Boston Dataset

• Sparse 3D reconstruction for Boylston Street in Boston
Demonstration – Boston Dataset

• Dense 3D Reconstruction of Boylston Street from Google Street View Data
Demonstration – Boston Dataset

• Camera Localization in reconstructed 3D scene

* Source Code and Model: https://github.com/JunweiLiang/VERA_3D_Reconstruction
Gunshot Detection

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization
Gunshot Detection

• Our tools can detect video segments that contain one or more gunshots.
  – This significantly reduces the inspection time when dealing with many or very long videos.
Gunshot Counting & Gun Type Classification

• After gunshot segments are identified, our system can:
  – Count gunshots
  – Classify gun types
Gunshot Counting & Gun Type Classification

- Web tool interface
Gunshot Counting & Gun Type Classification
Shooter Geo-Localization

Real-Time Event Analysis
- Social Media Event Detection & Analysis
- Crowd Counting
- Traffic Accident Prediction
- Person Identification
- PTZ Camera Automated Capture

Event Reconstruction
- Camera Geo-localization
- 3D Scene Reconstruction

Shooter Event Analysis
- Gunshot Detection & Gun Type Classification
- Shooter Geo-Localization

Event Begins

Information Flow

Command Center
Shooter Geo-Localization

• Our system can determine the shooter’s location from social-media videos where a gunshot was detected

• This localization is based on:
  1. Gunshot Detection
  2. Video Synchronization
  3. Camera Geo-Localization
Shooter Geo-Localization

• The system can estimate:
  • The shooter distance from the camera
  • The direction the shooter is shooting from given recordings from two cameras

Assuming:
  • Videos are synchronized
  • Cameras are geo-localized
Estimating Shooter Distance

• The system can estimate the distance of the shooter to the camera if:
  1. The bullet is super-sonic
  2. Bullet shockwave sound ("crack") is recorded
  3. Muzzle blast sound is recorded
     • Muzzle blast sound: the sound when the bullet leaves the barrel

See technical report at https://vera.cs.cmu.edu for more details
Estimating Shooter Distance

Bullet shockwave facts:

1. Will arrive before the muzzle blast sound
2. Have a sharper angle if the bullet is faster

Note: cannot be heard behind the rifle

See technical report at https://vera.cs.cmu.edu for more details
Estimating Shooter Distance

- Mark the **shockwave** sound and **muzzle blast** sound on the video segment that **gunshot is detected**
Estimating Shooter Distance

- Bullet shockwave sound
Estimating Shooter Distance

- Muzzle blast sound
Estimating Shooter Distance

• Distance can be computed based on:
  – Time difference between the bullet shockwave and muzzle blast
  – Speed of the bullet
  – Speed of sound

• Since the bullet and speed of sound are estimated, there is a range of possible distances

• A probable minimum and maximum distances around the camera form a donut shape of possible locations
Estimating Shooter Distance

The *donut* is placed on the map around the camera.

- Inner circle: Possible shooter locations if bullet speed is faster.
- Outer circle: Possible shooter locations if bullet speed is slower.

Mandalay Bay Hotel

See technical report at https://vera.cs.cmu.edu for more details.
Estimating Shooter Distance

The *donut* is placed on the map around the camera

Mandalay Bay Hotel

45 degree view from another angle

Camera location
Estimating Shooter Direction

• To estimate the direction of the shooter, we need:
  – A pair of synchronized videos both recording the muzzle blast sound of the gunshot
  – **Accurate camera locations**
Estimating Shooter Direction

• We make use of the \textit{time difference} of the arrival of the muzzle blast sound at the two cameras
  – Three bowl-shaped lines (hyperbolas) can be drawn to determine the likely location of the shooter
Estimating Shooter Direction

The shooter is located within the three lines, where the middle green one is most likely.
Estimating Shooter Direction

The shooter is located within the three lines, where the middle green one is most likely 45 degree view from another angle.

Mandalay Bay Hotel

Camera locations
Demonstration of the web interface
Summary

• **Real-Time Event Analysis**
  – Social Media Event Detection and Analysis
  – Crowd Counting
    • Traffic Accident Prediction
    • Person Re-identification using Gait Recognition
    • Automatic Person Picture Capturing Using Pan-Tilt-Zoom Cameras ([https://github.com/google/next-prediction](https://github.com/google/next-prediction))

• **Event Reconstruction**
  – Video Synchronization ([https://vera.cs.cmu.edu/](https://vera.cs.cmu.edu/))
  – Camera Geo-Localization
  – 3D Reconstruction ([https://vera.cs.cmu.edu/VERA_3D_Reconstruction](https://vera.cs.cmu.edu/VERA_3D_Reconstruction))

• **Shooter Event Analysis**
  – Gunshot Detection & Gun Type Classification ([https://vera.cs.cmu.edu/](https://vera.cs.cmu.edu/))
  – Shooter Geo-Localization ([https://vera.cs.cmu.edu/](https://vera.cs.cmu.edu/))
Thank you!

Any questions?
Come back for the

Next Session

1:50 PM