**2017 RIPS-LAPD Project:**

**Conversational Turn-Taking in Police Body-Worn Video**

Industry Sponsor: Deputy Chief Sean Malinowski (LAPD Chief of Staff); Sgt. Dan Gomez, Mr. Arnold Suzukamo (LAPD-IT Bureau).

Academic Mentor: Jason Xu (UCLA)

Academic Supervisors: Jeff Brantingham, UCLA Anthropology; Dr. Craig Uchida, Justice & Security Strategies

**Introduction**

Body-worn video (BWV) or on-body cameras provide a novel means to collect very fine-information about police-public interactions. The general use model requires officers to initiate recording of video whenever there is an encounter with a member of the public. During such interactions, BWV is recorded in real-time. Recording is terminated at the officer’s discretion. BWV is not streamed or reviewed in real-time, but rather is uploaded to a secure cloud storage system at the end of an officer’s shift.

BWV is designed to provide another line of evidence for the actions of individuals and the outcomes of interactions between police and members of the public. BWV is therefore evidence relevant to legal proceedings like any other form of evidence collected by police. In a limited number studies, BWV has been shown to reduce the likelihood that situations escalate to a point requiring use of force.

There are considerable challenges facing wide-spread use of BWV. Even small scale deployments are expected to lead to massive volumes of video data that will quickly outstrip the ability of law enforcement agencies to analyze. The resulting fallback position will be to review BWV footage only when it corresponds to adverse outcomes (e.g., use of force). Most video will go unused. Many of the potential benefits of BWV may therefore go unrealized.

**The 2017 LAPD-RIPS Project**

The 2017 RIPS-LAPD team will work to develop methods for the automatic discrimination and labeling of audio-video segments into the following categories: (1) the focal police officer speaking; (2) other actors speaking; and (3) overlapping speech involving the focal officer and others. The focal police officer is defined as the officer wearing the camera. The goal is not speech content recognition, or transcription. Rather we wish to identify when police officers exclusively are speaking relative to one or more other actors in a video scene and when the officer and others are trying to override one another with speech. Measures of conversational turn taking may then be computed. Conversational turn taking may provide evidence of when interactions are escalating or de-escalating without specific knowledge of the content of speech.
Understanding when interactions escalate and de-escalate can be of tremendous value in helping to minimize the risk of adverse outcomes in police-public interactions.

The project will rely on a range of data types BWV metadata (e.g., time stamps), BWV audio, and the video images themselves. Computations may be done in Matlab, Mathematica, C, C++, R, Java, or other appropriate computational language.

**Key Milestones:**

1. Statistical assessment of LAPD BWV and other associated data.
2. Develop speech segmentation methods.
5. Present to LAPD.

**References**

