Recognizing Humans in Motion: Trajectory-based Aerial Video Analysis

Yumi Iwashita[1]
yumi@ieee.org

M. S. Ryoo[1]
mryoo@jpl.nasa.gov

Thomas J. Fuchs
Thomas.Fuchs@jpl.nasa.gov

Curtis Padgett
curtis.w.padgett@jpl.nasa.gov

Jet Propulsion Laboratory,
California Institute of Technology
Pasadena, CA, USA

Aerial surveillance images cover a wide area at low resolution [1, 2]. In order to detect objects (e.g., pedestrians) from such videos, conventional methods either utilize appearance information from raw videos or extract blob information from background subtraction results. However, people seen in low resolution images have less appearance information, and hence are very difficulty to classify based on their appearance or blob size. In addition, due to heavy camera movements caused by aerial vehicle ego-motion and wind, the system is expected to generate many noisy false detections including parallax. Figure 1 shows an example aerial image and a magnified area of the image. As we are able to observe in Figure 1 (b), appearance of people (yellow dotted circles) and rocks (red cricles) show very similar visual characteristics.

This paper proposes a novel method to detect/recognize people by classifying object candidates in low resolution aerial images. The idea is to detect and classify objects from aerial videos based on their motion: we analyze a trajectory of each object candidate, deciding whether it is a person-of-interest or simple noise based on how it moved. After objects are tracked by a Kalman filter-based tracking, we represent their motion as multi-scale histograms of 'orientation changes', which efficiently captures movements displayed by objects. Finally, random forest classifiers are applied to our new representation to make the decision.

Our histogram is based on an idea that tracks originated by humans participating in activities will contain movements completely different from tracks generated by noisy false object detections. The histogram is calculated using relative orientation changes between every pair of vectors, which are defined using three consecutive observed points with a constant frame difference as shown in Figure 2. Figure 3 shows examples of calculated histograms for a person, a car, and a false detection casued stabilization errors. In addition, we propose a multi-magnitude version of our histogram of orientation changes, which is descriged to capture movement magnitudes as well as orientation changes of the objects.

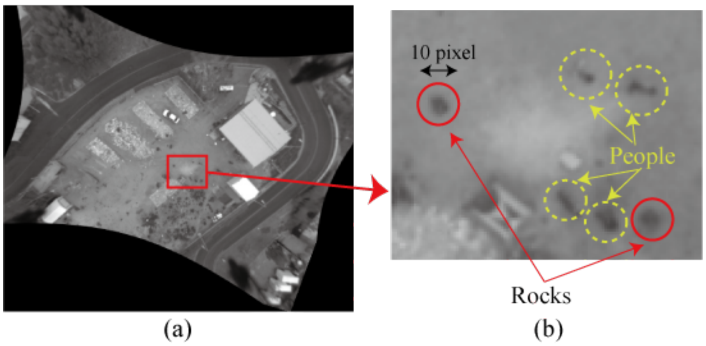

Figure 1: (a) An example of rectified and stabilized images, (b) an enlarged area including rocks and people whose appearances are similar.

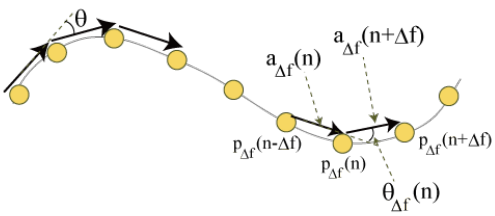

Figure 2: An example of orientation changes.

[1]These authors contributed equally to the paper.

The experimental results illustrate that our approach recognizes objects-of-interest (i.e., humans) even when there exist a large number of false detection/tracking. Figures 4 (a) $\sim$ (c) show example tracking results, classification results of our approach, and the ground truths. We are able to observe that our approach is able to successfully recognize humans from low resolution aerial videos with heavy ego-motion, by discarding many false detections while maintaining true detections based on our histogram representation.

The contributions of this paper are (i) the introduction of a new concept that objects can be better recognized using their motion information particularly in aerial video and (ii) our novel feature representation to capture object motion effectively and efficiently.

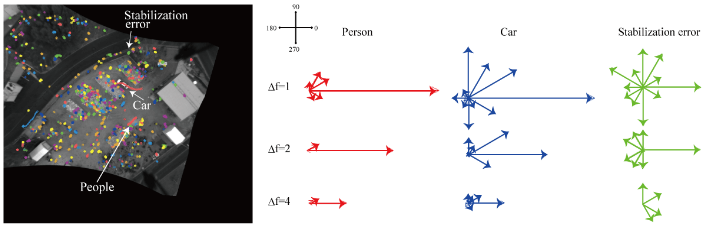

Figure 3: Example histograms describing trajectory orientation changes.

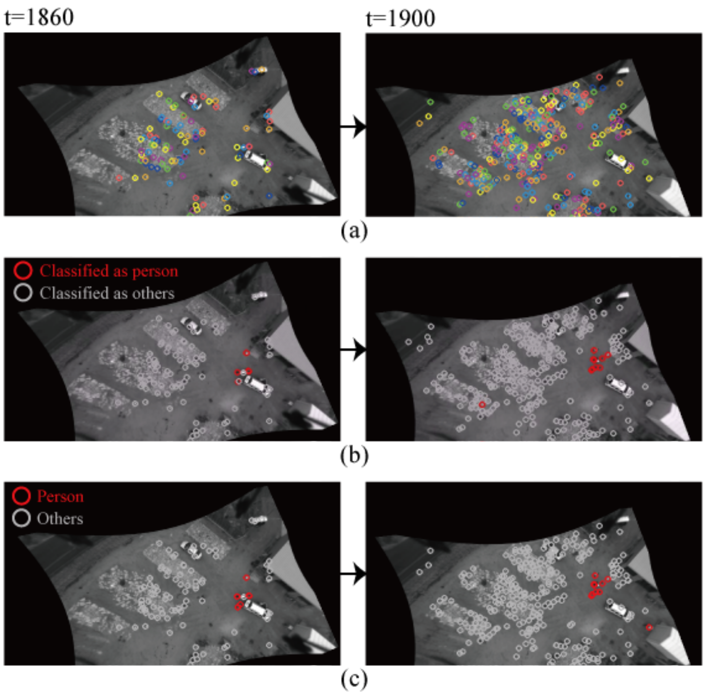

Figure 4: (a) Example results of object tracking, (b) classification results by the proposed method, and (c) ground truth. The video contains the activity of 'a group of person unloading an object from the vehicle', and we are able to observe that our approach correctly recognizes human tracks as opposed to other noise.

[1] V. Reilly, H. Idrees & M. Shah, *Detection and tracking of large number of targets in wide area surveillance*, ECCV 2010.

[2] T. Pollard & M. Antone, *Detecting and tracking all moving objects in wide-area aerial video*, CVPRW 2012.