Detection and Tracking of Mobile Targets in Aerial Infrared Images

Justin H. Le, Ali Pour Yazdanpanah, Emma E. Regentova

Department of Electrical and Computer Engineering,
University of Nevada, Las Vegas (UNLV),
Las Vegas, NV 89154, USA

Overview

- Detect and track moving animals or pedestrians in a dynamic environment through infrared (FLIR) video taken by unmanned aerial vehicle (UAV) or similar mobile platforms
- Collect state-of-the-art and historical methods in one application for the comparison of robustness and efficiency
- UAV observation can be a cost-efficient method of: (1) researching wildlife ecology for applications in life sciences and conservation; (2) surveillance in issues of national security and in military operations

Background

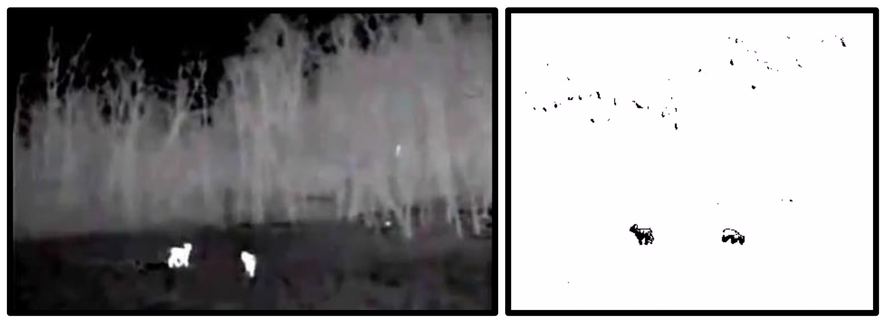
- Recording by aerial platform introduces noise (jitter) into video, rendering useless a majority of essential and well-developed image processing techniques
- FLIR images lack detail in texture, contour, and color; each detection method succeeds conditionally, varying with context
- Methods for handling above challenges have not been unified; FLIR UAV imaging remains an untapped asset in surveillance technologies

Method

- Noise due to camera motion can be removed by stabilization:
  - i. Compare adjacent video frames to form an estimate of the transformations that occur between frames (e.g., scaling, translating, rotating).
  - ii. Apply the inverse transformations to reduce motion's effects.
- After noise removal, (1) separate objects of interest from the environment (extraction), and (2) track them (optical flow):
  - i. FLIR imaging makes heat-emitting objects, such as living organisms, distinctly bright in most frames. Replace these bright objects with white blobs, blacken all other regions.
  - ii. "Optical flow" is a technique for indicating the direction of motion of objects (relative to camera). For static cameras, compute the flow of blobs. For dynamic cameras, directly compute the flow of targets in the unprocessed frames using the built-in extractor of the Lucas-Kanade technique.
- A classifier places data in a category based on a set of criteria:
  - i. Criterion: an object in a frame matches the object described by a set of still images (the model).
  - ii. Every region of a frame is compared pixel-by-pixel to the model. If a match occurs, graphically indicate it.
- Core functions are derived from OpenCV, a set of open-source software libraries that streamline the development of applications for image processing and machine learning.

Extraction and Optical Flow

Fig. 1:
Original (top) and result of foreground extraction (bottom).



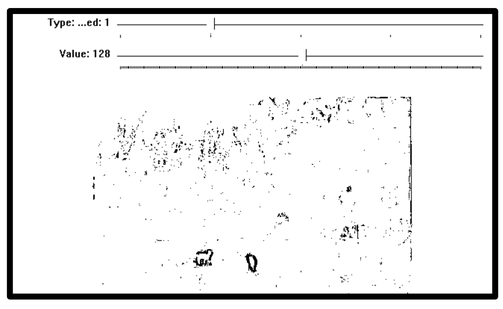
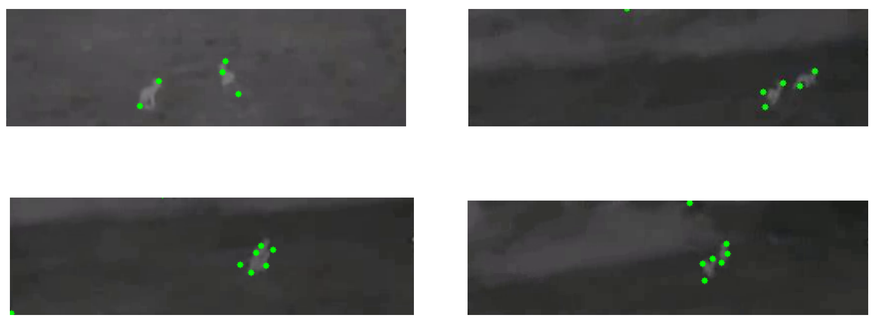


Fig. 2 (above):
Filtered result produced by thresholding based on preset values of statistical moment features.

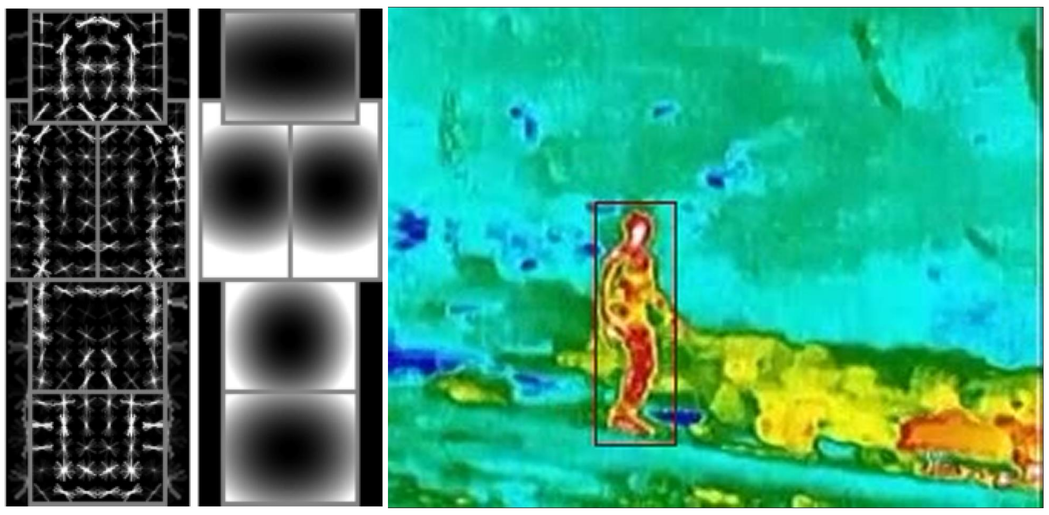
Fig. 3:
From left to right, two animals in a field are detected, highlighted with dots, and tracked as they move.

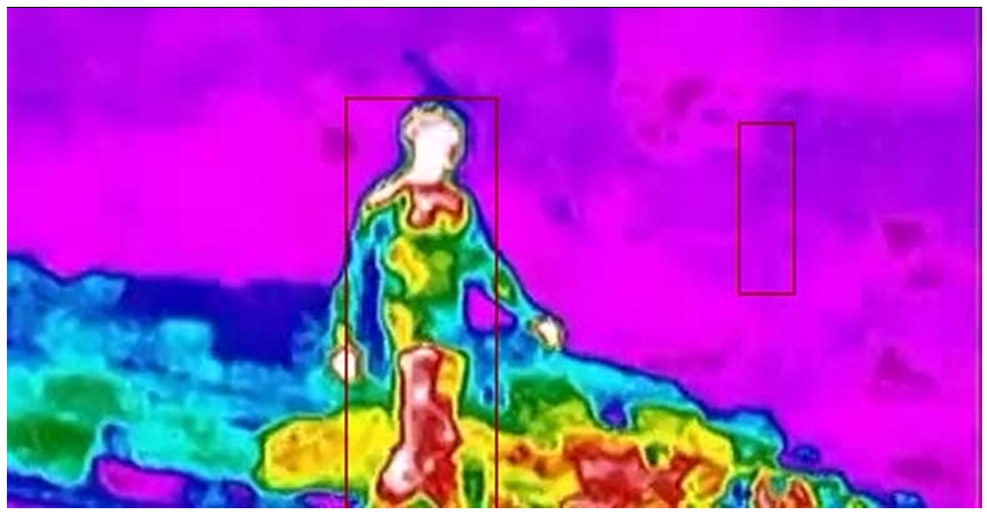


Classification

Fig. 4:
Objects that resemble the specified model (left) are bound by colored boxes.

Fig. 5 (right):
Training the classifier with non-IR images causes noise to be classified as a match.





Conclusions

- Preliminary experiments show that state-of-the-art detection/tracking methods developed in computer vision can be combined to build an efficient framework for human and wildlife detection in video sequences. The challenges include factors such as scale, noise, and occlusion.
- Further developments include vast training of the system for differentiation among animals of the same scale and quantification of targets in real-time.

Acknowledgement

This material is based upon work supported by NASA EPSCoR under Cooperative Agreement No. NNX10AR89A

Fig. 4: Felzenszwalb, et al, "Object Detection with Discriminatively Trained Part Based Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009.