Detection and Tracking of Vehicles in Aerial Imagery

Jan Prokaj
University of Southern California

Mark Duchaineau
Lawrence Livermore National Lab

Abstract

In this work we consider the problem of detecting and tracking vehicles in aerial imagery. The problem is formulated as a data association problem between observations of motion and vehicle tracks. To decrease the ambiguities in association, all decisions are based on a sliding window of frames. Within a window, a Bayesian network is created, structured to represent possible associations of one observation across frames. Belief Propagation is used to do inference. Tracks are associated from one window to the next using motion and appearance similarity. The algorithm is evaluated on a real world dataset.

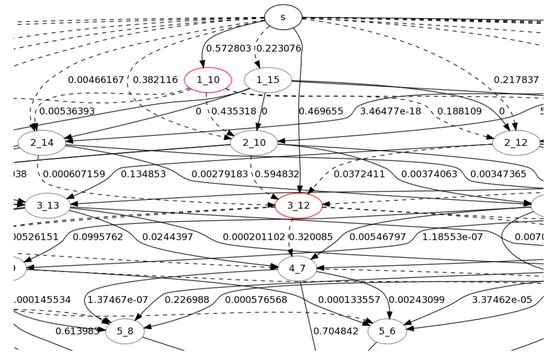
Problem

One frame of a video sequence with vehicles outlined in red. Four vehicles are enlarged and shown on the right.

- Aerial surveillance of large geographic areas with a moving camera
- Large format video (1700x1700) at 2 fps, grayscale
- Vehicles are very small size (~20x20), and some are difficult to distinguish from the background
- GOAL: Detect and track all vehicles as they move through the area under surveillance

Algorithm

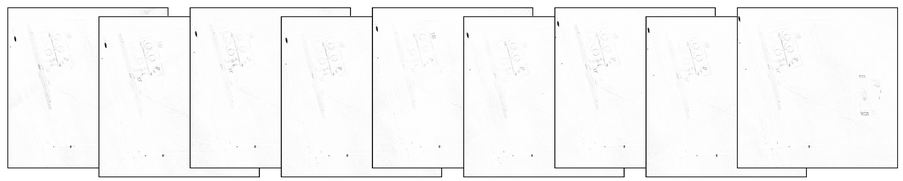
- Create a Bayesian network (tree) for each observation in the beginning of the window
  - The binary variables in the tree represent observations in subsequent frames
  - The evidence and compatibility functions of each variable are based on appearance similarity and motion likelihood
- Use Belief Propagation (max-product) to infer the most likely state of each observation
- Filter out unlikely tracks
- Associate tracks in window with existing ones



An example of a (portion of) Bayesian network representing the possible associations of an observation across multiple frames.

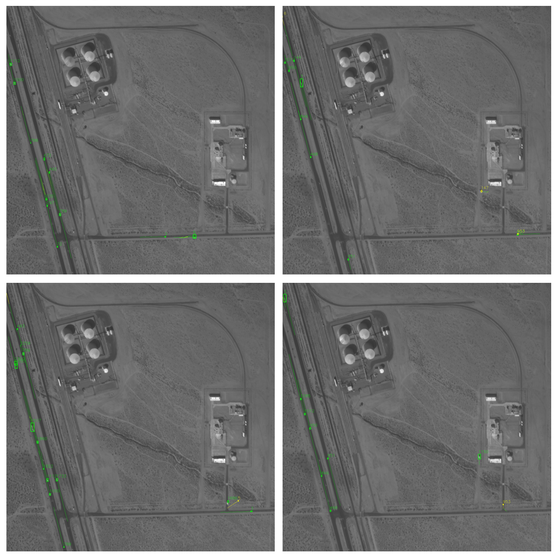
Strategy and Prior Work

- Model the background and subtract it to find moving objects
- Formulate the tracking problem as a data association problem between observations of motion and vehicle tracks
- To decrease the ambiguity present at one time instant, a sequence of observations over a window of frames is used in making an assignment decision
  - Prior work in this area centers around JPDAF or MHT algorithms, which evaluate all possible hypotheses (assignment of observations to tracks)
  - The solution space grows exponentially with longer windows and requires pruning in practice
  - Alternatives include MCMC based approaches



Background subtraction example. Darker color indicates larger background difference.

Results



Example vehicle detections and tracks.

Discussion and Conclusions

- Algorithm so far evaluated on 184 frames of one video sequence
- No missed detections
- Low number of false detections
- Almost all vehicles are tracked correctly
- Areas needing improvement:
  - Multiple tracks assigned to one vehicle
  - Efficiency

- The developed algorithm works well on the dataset evaluated, but needs more thorough testing to be trusted
- Efficiency issues can be addressed by taking advantage of the algorithm's inherent parallelism

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

IM Review & Release Number: LLNL-POST-415383