Roke

Part of the
Chemring Group

UDRC WAMI Challenge

Richard Lowe

richard.lowe@roke.co.uk

Contents

Roke

Part of the
Chemring Group

Problem

Problem

Detecting vehicles and pedestrians (targets) in urban areas using an airborne camera, from anything other than a vertical viewpoint, is difficult as much of the road network is occluded by:

- Buildings;
- Trees;
- Shadows.

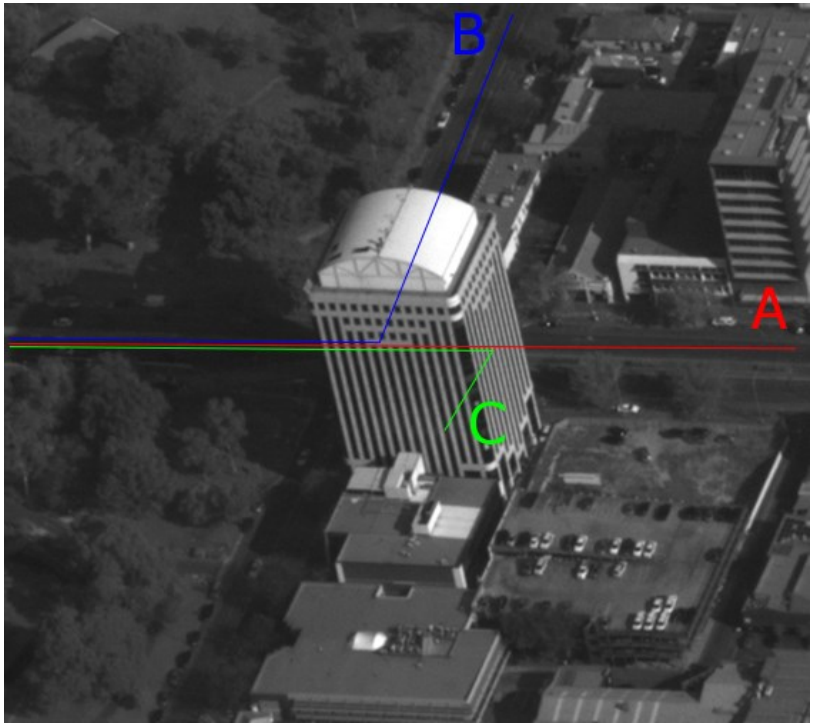
Problem

Determining tracks of a target's movement through such an area is then also difficult, as there is a need to distinguish between:

a) temporary occlusion;

b) change in velocity;

c) permanent occlusion where targets are no longer visible in the state space (having entered a building, for example).

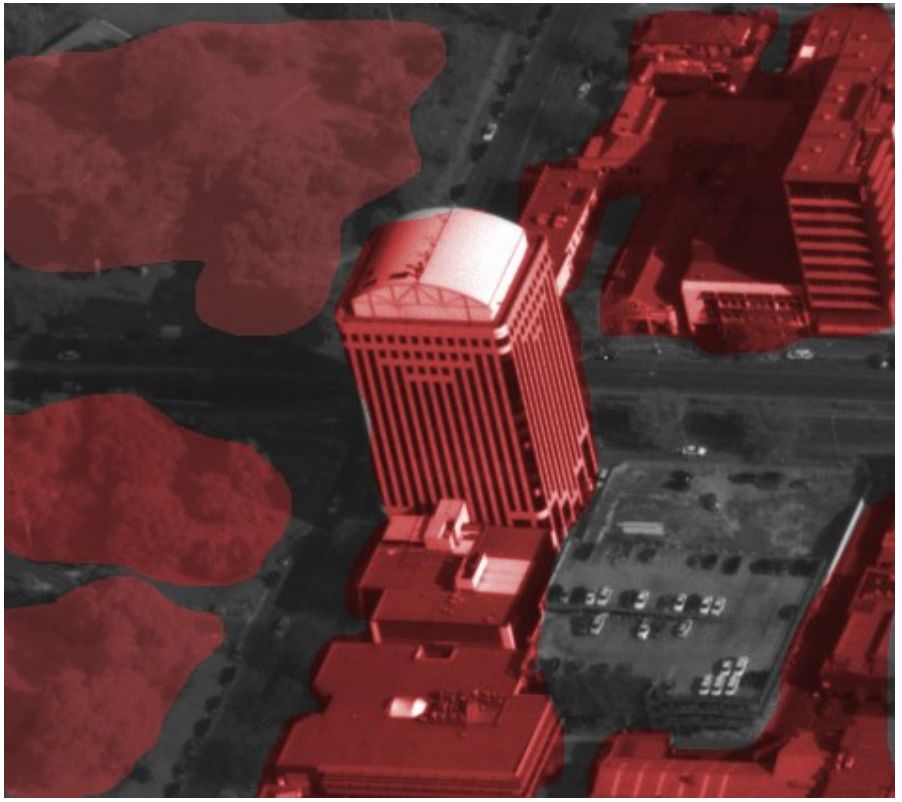


Problem

- Some Bayesian trackers contain models of the detection likelihood of a target;
- They require a high level of uncertainty in that model if there is a variable amount of occlusion;
- This high level of uncertainty in the detection likelihood propagates a larger proportion of candidate tracks;
- This is undesirable in regions with low levels of occlusion such as extra-urban areas.

Problem

- Terrain models can provide some information about an urban area by providing building heights, but they cannot cope with real-time changes in the environment (for example due to building work);
- Knowledge of occlusion using real-time information would aid our filter design, as we can:
  - reduce the detection likelihood in occluded areas;
  - propagate candidate targets;
  - improve re-identification quality as the search space for detection is reduced.



Roke

Part of the
Chemring Group

Challenge

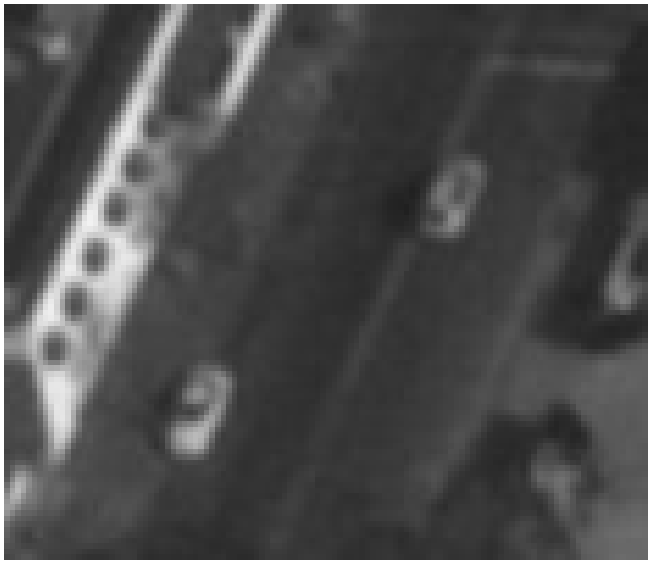
Challenge

Determine and label parts of each image that are occluded due to buildings.

- The dataset is a set of frames recorded over Adelaide's business district;
- Given a sequence of images and camera metadata, produce a binary mask for each image where pixels are given a value of occluded (0) or visible (1);
- The binary mask should treat the building footprint as equivalent to occluded ground;
- Novel approaches that can detect occlusions due to trees or areas in deep shadow are also welcomed, however such features won't form part of the scoring metric.

Dataset: Target Level

Resolution of a single target:



Dataset: Street Level

Resolution of a building:

Dataset: Single Camera

Resolution of a single camera:

Dataset: Whole Scene

Resolution of all six cameras:

Dataset format

Series of 48 RAW frames each with 6 cameras:

- Labelling format: date–frame_number–camera_number-VIS;
  - Example: 20140521042753-03032396-05-VIS
- Each camera image is 3248x4872 pixels;
- The framerate is approximately 3Hz;
- There is a small percentage of overlap between each camera;
- The pointing angle of the camera is approximately 45° from horizontal, but can be calculated with the accompanying metadata.

Dataset format

- Image rotation is camera dependent:
  - Even numbers are rotated clockwise and flipped
  - Odd numbers are rotated anticlockwise and flipped
- Each camera is laid out in the frame as follows:

3	**1**	**5**
2	0	4

POS Files

- Each frame comes with a POS file to specify the camera metadata:

	Camera								IMU Velocity					
Sensor name	**Yaw**	**Pitch**	**Roll**	**Latitude**	**Longitude**	**Altitude (ft)**	**IMU time (s)**	**GPS week**	**North**	**East**	**Up**	**IMU status**	**Offset from UTC**	**DST Flag**
$LAIR2	109.857426	1.803828	29.014919	39.753101	-84.130597	16843.046851	332717.087305	1554	-33.635874	83.931935	0.340803	3	-4	1

Reading the RAW files

```
#!/usr/bin/python

import numpy as np


file_path = /path/to/file

Even=True

height = 3248

width = 4872

img = np.fromfile(file_path,dtype="B")
img.shape = (height, width)
img = np.rot90(img)


if even:
    img = np.flipud(img)
else:
    img = np.fliplr(img)
```

Scoring

- Each image will be compared with our ground-truth data to determine how many buildings are causing occlusion;
- A score will be calculated based on the similarity to our ground-truth using the Jaccard index:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \qquad 0 \leq J(A,B) \leq 1$$

Submission

- Submission date will be in October;
- submission portal details will be distributed at a later date;
- data can be submitted as a zipped set of binary masks in png format, labelled using the same format as the input:
  - 20140521042753-03032396-00-MASK
  - (Date – frame – camera – MASK)