University Defence Research Collaboration (UDRC)
Contextual Image Processing

Iain Rodger & Neil M. Robertson – Heriot Watt University; Barry Connor – Thales Optronics
ir45@hw.ac.uk

Problem Overview

Key Info:

- The human visual system can quickly accomplish complete scene understanding, but is limited by information processing. To counteract this and improve situational awareness, sensor suites are constructed to collect and process environment specific information.
- Target detection systems are critical in situational awareness activities for both military and surveillance applications. Segmentation, detection and tracking stages are common tasks within these .
- Such systems are, however, hindered by complex, cluttered scenes as well as object occlusions and dynamic lighting.
- Total scene understanding, akin to the human visual system, should help to overcome these issues. Utilising **context** and **segmentation** is one way to achieve this.

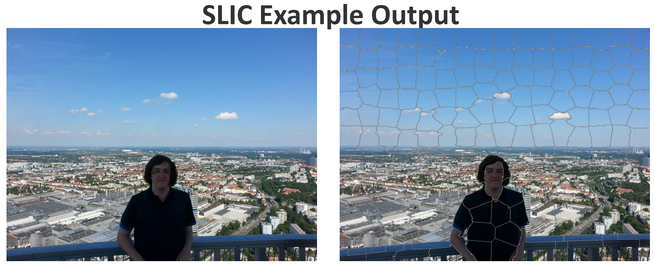
Background

Segmentation Definition: Process of separating or parsing an image into set of regions that do not overlap, but form an entire image upon unification. These regions will be similar in nature, such as spatial or colour proximity.

Segmentation Algorithm: The method used here is a super-pixel based algorithm. The Simple Linear Iterative Clustering (SLIC) technique from Achanta et al. It is a *subjective* approach to segmenting an image.

SLIC Mechanics: Clusters pixels in a 5 –dimensional colour and image plane space, using a combined distance measure.

SLIC Example Output



Region Extraction Method

1) Extract label map from SLIC output image.

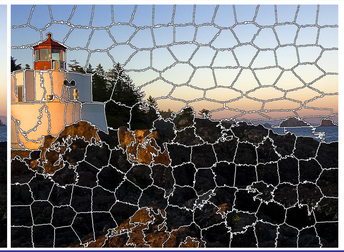
2) Identify pixels within each superpixel using label map.

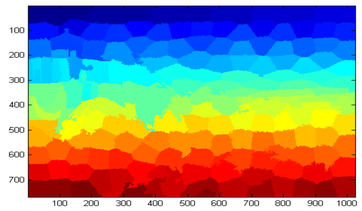
3) Use histogram statistics to compute similarities of each superpixel. Similarity metric is the **Bhattacharya Distance**: $\boldsymbol{d(y)} = \sqrt{1 - \rho[\widehat{\boldsymbol{p}}(\boldsymbol{y}), \widehat{\boldsymbol{q}}]}$

4) Grow regions of most similar superpixels via global threshold.

5) Use adjacency of superpixels to reassign / re-label spurious regions.

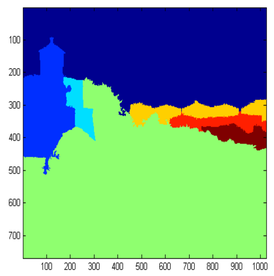
6) Display outputs.





Initial Results

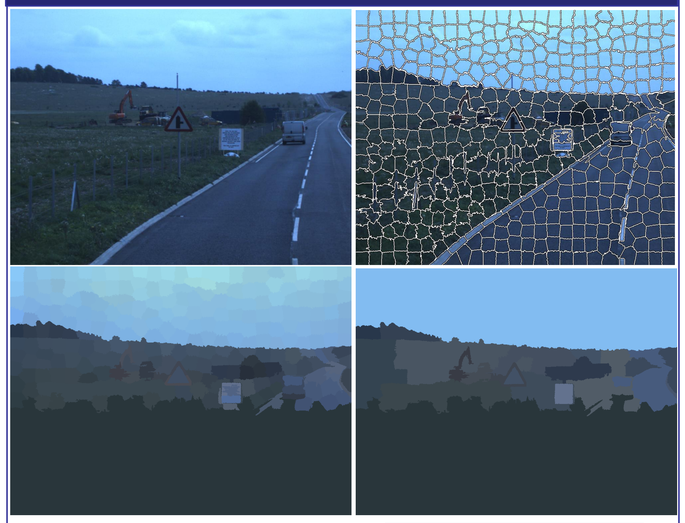
Extracted Regions from Lighthouse image above. LHS before step 5 / RHS after step 5.



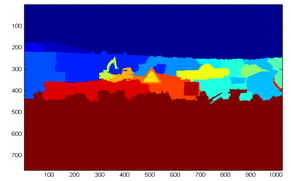
Label map from final region extraction.

Thermal image region extraction. Final regions obtained before and after adjacency pass.

Initial Results



Input image from Hydravision project (RGB camera). Extracted regions are shown above with before/after adjacency pass results shown. Label map (shown right) highlights the final regions.



Next Steps and Future Work

This work is in its infancy and there are some obvious immediate goals, in addition to longer term objectives. These are:

- Perform thorough evaluation to quantify region segments accuracy.
- Investigate more possibilities for using the Bhattacharya distance metric.
- Test algorithm on data representative of surveillance problems.
- Explore parameter sensitivity and effects of varying them.
- Investigate local adaptive threshold methods.
- Introduce detection/tracking information and construct contextual framework.
- Ultimately, a foreground/background extractor that is capable of total scene understanding is desired.

HERIOT WATT UNIVERSITY
THALES
EPSRC
Engineering and Physical Sciences Research Council