Object tracking in a monocular system with ego-motion

Pratik Sajnani (ps8369@rit.edu)
Advisor : Dr. Zack Butler, Rochester Institute of Technology

ABSTRACT
- Tracking points across multiple images is a fundamental operation in many computer vision applications
- For tracking algorithms there is a tradeoff between robustness and accuracy
- Tracking obstacles in a moving robot’s path is a real-time application that requires generation of useful data efficiently
- Challenges include lighting changes, spurious motion, out-of-frame errors, camera motion

FEATURE DETECTION
- Corner detection is used to detect areas of images rich in features
- Good features have big eigenvalues which implies
  - Texture
  - Color
- Features are used to identify objects that can be tracked irrespective of their importance

TRACKING METHODS
- **Lucas-Kanade method** is an unsupervised learning method to detect and track points in an image
  - Features : corner detector output
  - Uses down-sampled versions of the images as starting point
  - Every subsequent pair of frames is processed
- **Kernelized Correlation Filters (KCF)** is a state-of-art high-speed tracking algorithm
  - Features : Histogram-of-Gradients from within image patch
  - Requires an initial bounding box specification for supervised learning
  - Reliable in the presence of occlusion and dynamic lighting changes

RESULTS
- Lucas-Kanade was measured using number of points detected on target
- KCF was measured using manually generated test data
- Out-of-frame errors were not a part of the measurement

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucas-Kanade</td>
<td>78.3%</td>
<td>98.2%</td>
</tr>
<tr>
<td>KCF</td>
<td>93.8%</td>
<td>99.1%</td>
</tr>
</tbody>
</table>

CONCLUSIONS
- Tracking using an unsupervised learning method requires a high degree of statistical regularity in the scene
- Bounding boxes are a good pre-processing step to eliminate ego-motion problem
- Out-of-frame error handling requires a hard-reset functionality in the visual pipeline

REFERENCES