Robust Visual Robot Localization Across Large Perceptual Changes

Tayyab Naseer

University of Freiburg, Germany
Autonomous Intelligent Systems Group at Univ. Freiburg

http://ais.informatik.uni-freiburg.de/
http://www2.informatik.uni-freiburg.de/~naseer/
GPS based Localization

If GPS is inaccurate?
Motivation

- Efficient visual localization and consistent mapping across seasons
Visual Localization under Similar Perceptual Conditions

Image at time $t_0$    Image at time $t_k$
Does it Work in this Scenario?
Does it Work in this Scenario?

Feature based approaches fail in this case
Robust Image Matching

- Robust and dense description of images
- Compute HoG descriptors on grids

Summer Image

Winter Image

Robust Visual Robot Localization Across Seasons using Network Flows
Tayyab Naseer, Luciano Spinello, Wolfram Burgard, Cyrill Stachniss
Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), Quebec, Canada, 2014
Idea: Why HOG is Better?

- Represents gradient orientation and magnitude information for each cell
- **Spatial** feature description

\[
c_{ij} = \frac{h_{d_j} \cdot h_{q_i}}{||h_{d_j}|| ||h_{q_i}||}
\]
Similarity Matrix Analysis

Is global best match the optimal choice?
Flow Network

\[G = (X, E)\]
\[s \in X, t \in X, (u, v) \in E, F \in \mathbb{N}\]
\[c(u, v) > 0, f(u, v) \geq 0\]
Minimum Cost Flow Problem

Given an amount of flow, the cheapest way of sending flow through a flow network:

$$\min \sum_{(u,v) \in \mathcal{E}} w(u, v) \cdot f(u, v)$$
Flow Network

Nodes

- White matching node models a successful match of an image pair
- Red hidden node describes that the pair of images could not be matched
Flow Network: Motion Models

- STOP
- Linear
- Double Speed

Diagram showing the progression of states over time.
Flow Network Graph
Edge Costs

\[
\begin{align*}
\text{edge cost to node} & : i,j \\
\mathbf{w}_{ij} & = C^{-1}_{ij} = \left( \frac{\mathbf{h}_{d_j} \cdot \mathbf{h}_{q_i}}{\|\mathbf{h}_{d_j}\| \|\mathbf{h}_{q_i}\|} \right)^{-1}
\end{align*}
\]
Flow Network
What does it Achieve?
Experiments

Datasets recorded in Freiburg city, Germany

Summer 6915 images
Winter 30790 images
Different Velocities

- Our method
- SeqSLAM
- Ground truth

Precision vs. Recall

- Our method
- SeqSLAM
- HOG-bm
Partial Trajectories

![Graph with two subplots]

- **Left subplot**: Shows two trajectories labeled as Query (Summer) and Database (Winter). The trajectories are represented by red dots for Our method and gray circles for Ground Truth.

- **Right subplot**: Displays a precision-recall curve with three lines:
  - Red line: Our method
  - Blue dashed line: SeqSLAM
  - Green line: HOG–bm

The x-axis represents Recall, ranging from 0 to 1, and the y-axis represents Precision, also ranging from 0 to 1.
Video
Efficient Localization using Rough GPS Prior

- Compute a small fraction of the computationally expensive image comparisons.
- Eliminating the need to setup and solve a network flow problem.
Efficient Localization using Rough GPS Prior

- Number of descriptor comparisons needed to build the cost matrix.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q$</td>
<td>79</td>
<td>676</td>
<td>1,213</td>
<td>1,266</td>
<td>1,428</td>
</tr>
<tr>
<td>$D$</td>
<td>943</td>
<td>361</td>
<td>596</td>
<td>3,601</td>
<td>1,476</td>
</tr>
<tr>
<td>[15]</td>
<td>74,498 s/0.7 s</td>
<td>244,037 s/2 s</td>
<td>722,948 s/6 s</td>
<td>4,558,866 s/55 s</td>
<td>2,107,728 s/19 s</td>
</tr>
<tr>
<td>GPS</td>
<td>74,498 s/0.7 s</td>
<td>134,791 s/1 s</td>
<td>298,432 s/2 s</td>
<td>2,620,748 s/23 s</td>
<td>1,102,689 s/10 s</td>
</tr>
<tr>
<td>500m</td>
<td>50,643 s/0.2 s</td>
<td>38,334 s/0.23 s</td>
<td>72,788 s/0.5 s</td>
<td>621,369 s/4 s</td>
<td>106,672 s/0.6 s</td>
</tr>
<tr>
<td>GPS</td>
<td>38,313 s/0.1 s</td>
<td>26,841 s/0.1 s</td>
<td>45,457 s/0.2 s</td>
<td>288,312 s/1 s</td>
<td>76,171 s/0.25 s</td>
</tr>
</tbody>
</table>

Olga Vysotska, Tayyab Naseer, Luciano Spinello, Wolfram Burgard, Cyrill Stachniss
Efficient Localization using Rough GPS Prior

Olga Vysotska, Tayyab Naseer, Luciano Spinello, Wolfram Burgard, Cyrill Stachniss
Robust Visual SLAM Across Seasons

- Approach to visual SLAM that deals with large perceptual changes
- Leverage sequential information for robust loop closure detection
- Robust global image description using convolutional neural networks (CNNs) to achieve more distinctive image matching.

Robust Visual SLAM Across Seasons
Tayyab Naseer, Michael Ruhnke, Luciano Spinello, Cyrill Stachniss, Wolfram Burgard
Quantitative Evaluation of Deep Features

Robust Visual SLAM Across Seasons
Tayyab Naseer, Michael Ruhnke, Luciano Spinello, Cyrill Stachniss, Wolfram Burgard
Metric Evaluation

- 8 km long trajectory, Pittsburgh
- 1678 images at 3Hz

Data Associations

Optimized Trajectory

Odometry
### Metric Evaluation

- **1.6 km long trajectory, Freiburg**
- **676 images at 1Hz**

**Our Approach**

- **FABMAP**
- **SeqSLAM**
Vision-based Markov Localization

- Discrete Bayes filter to estimate the state.
- Sensor model based on whole image descriptors.
- Compute probability distribution over the whole state space.
- Handle more complex trajectories
- Run online on a robot
Vision-based Markov Localization

- State at time $t$ is represented by a distribution of a random variable $x_t$, conditioned on the sensor data history
  \[ \text{Bel}(x_t) = p(x_t | z_1, \ldots, z_t) \]

- Belief can be computed recursively and efficiently without loss of information
  \[ \text{Bel}(x_t) = \alpha_t p(z_t | x_t) \sum_{x_{t-1}} P(x_t | x_{t-1}) \text{Bel}(x_{t-1}) \]
Evaluation

Vision-Based Markov Localization Across Large Perceptual Changes

Tayyab Naseer, Benjamin Suger, Michael Ruhnke, Wolfram Burgard

Conclusion

- Across season place recognition
- Exploit dense image description
- Consistent trajectories across large perceptual changes
- Efficient solution as network flow problem
- Use Rough GPS Prior for efficient computations
- Markov Localization for online mobile operation
Thank you for your attention!