Multi-Stage Localization Given Topological Map for Autonomous Robots

Mohammed Abdel-Megeed Salem
Faculty of Computer and Information Science, Ain Shams University in Cairo
Salem@cis.asu.edu.eg
Outline

• Motivation
• Introduction
• Robot Localization and Place Recognition
  – *Global Features Extraction and Classification*
  – *Local Features Extraction and Classification*
• Results
• Conclusion
Motivation

- Vision-based place recognition for autonomous robots to navigate in a human-inhabited environment, given a topological map.
Multistage Vision-based Localization

Sensing

Image Stream

Global Features Extraction and Classification

Ambiguous Classification?

Place Recognition

No
Multistage Vision-based Localization

Sensing

Image Stream

Global Features Extraction and Classification

Ambiguous Classification?

Yes

Local Features Extraction and Classification

Place Recognition

No

Place Recognition

Mohammed A-Megeed Salem

ICCES 2012

Multi-Stage Localization Given Topological Map for Autonomous Robots
Multistage Vision-based Localization

Training
- Sensing
  - Image Stream
- Global Features Extraction
  - G. F. Vector
- Ambiguous Place Det. (EM)
- Local Features Extraction
  - L. F. Vector
- Training (SVM)
  - Trained Data
  - Amb. Places Knowledgebase
  - L. F. Knowledgebase

Testing
- Sensing
- Global Features Extraction
- Classification (SVM)
  - Is Amb. Place?
    - No
      - Class Label
    - Yes
      - Matching
      - Local Features Extraction
      - Class Label
Global Features Extraction

• Global features are recommended as long as the overall composition of the image is desired to be represented, rather than the foreground object.
• Typical examples are color histogram and vectors of principle component.
Global Features Extraction

• The global feature vector consists of statistical, spectral and shape features.
  – The image is divided into 9 disjoint regions
  – The color model is transformed to the HSV
  – Mean, and standard deviation of each channel for each block were computed
  – Statistical Features: 54 for each frame
Global Features Extraction

• The global feature vector consists of statistical, spectral and shape features.
  – The image is divided into 9 disjoint regions
  – The 2D Fourier Transform is computed
  – The highest 10 frequency magnitudes and the corresponding phases are recorded
  – Spectral Features: 360 for each frame
Global Features Extraction

• The global feature vector consists of statistical, spectral and **shape** features.
  – The Hough Line Transform is used
  – The range of 360 degrees is divided into 6 bins.
  – The average length of lines in the bin and
  – The ratio of the number of lines in the bin and the total number of lines.
  – Shape Features: 12 for each frame.
Local Features Extraction

• Scale Invariant Feature Transform (SIFT)
  – Invariant to rotation, viewpoint change and scaling.
  – Robust against changes in illumination, and image noise.

• Good results were obtained with $2 \times 2$ array of histograms with 8 orientation bins.

• Local Features: $2 \times 2 \times 8 = 32$ elements for each interest point.
Results

• Datasets of the international competition ImageCLEF
• two subsets represent different floors in an indoor office environment.
• First subset contains 4782 images
• Second subset contains 4138 images
## Results

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame Size</td>
<td>200x200 Pixels</td>
</tr>
<tr>
<td>Training Set Size</td>
<td>75%</td>
</tr>
<tr>
<td>Testing Set Size</td>
<td>25%</td>
</tr>
<tr>
<td>Acquisition Rate</td>
<td>30 frame per second</td>
</tr>
</tbody>
</table>
## Results

### Training Performance Measures

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Features Average Extraction Time</td>
<td>200ms</td>
</tr>
<tr>
<td>Local Features Average Extraction Time</td>
<td>420ms</td>
</tr>
<tr>
<td>Global Features Vector Length</td>
<td>426</td>
</tr>
<tr>
<td>Avg. Local Feature Vector Length</td>
<td>120 Interest Point</td>
</tr>
<tr>
<td>SVM Training Time</td>
<td>10890 ms</td>
</tr>
<tr>
<td>Matching Training Time</td>
<td>0 ms</td>
</tr>
<tr>
<td>Total Training Time</td>
<td>1177001 ms ð19 min</td>
</tr>
</tbody>
</table>
## Results

### Testing Performance Measures

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Key Frame Rate</td>
<td>1.7 fps</td>
</tr>
<tr>
<td>SVM Classification Time</td>
<td>0 ms</td>
</tr>
<tr>
<td>Point Matching Class. Time</td>
<td>5 ms</td>
</tr>
<tr>
<td>Rec. Rate for not ambiguous Places</td>
<td>5 fps</td>
</tr>
<tr>
<td>Rec. Rate for ambiguous Places</td>
<td>2.4 fps</td>
</tr>
<tr>
<td>Error Rate</td>
<td>11%</td>
</tr>
</tbody>
</table>
Conclusion

• We have proposed a novel approach of localization for autonomous robots using a multistage feature extraction and classification.

• The first classification stage is based mainly on global features.

• In certain cases the system adds a stage of classification based on local features.
Conclusion

- Ambiguous detection allows the system to be robust under different and imaging conditions and providing low misclassification rate.
- The system is tested using ImageCLEF dataset, and data set for in house environment.
Multi-Stage Localization Given Topological Map for Autonomous Robots

The International Conference on Computer Engineering and Systems November 27-29, November, 2012, Cairo, Egypt

Dr. Mohammed Abdel-Megeed M. Salem

Faculty of Computer and Information Sciences,
Ain Shams University,
Abbassia, Cairo, Egypt
Tel.: +2 011 1727 1050
Email: salem@cis.asu.edu.eg