HOMOGRAPHY-BASED GROUND PLANE DETECTION FOR MOBILE ROBOT NAVIGATION USING A MODIFIED EM ALGORITHM

Daniel Conrad
and
G. N. DeSouza
Outline

- Introduction
- Homography Relation
- EM Algorithm and Modification
- Navigation Algorithm
- Experimental Results
- Conclusions and Current Work
Our goal is to perform plane detection (ground plane and walls) for mobile robot navigation.

In the simple case of ground plane detection, we need to classify/cluster pixels into two groups:
- Ground pixels
- Non-Ground pixels

We use the homography of the ground plane along with a Modified EM algorithm to achieve this clustering.
Plane in a scene induces a Homography between images where \( sx' = Hx \).

H can be decomposed into:

\[
H = \hat{A}(R + \frac{t}{d}n^T)A^{-1}
\]
Expectation Maximization Algorithm

- Common algorithm in many learning problems
- Iterative algorithm divided into two steps:
  - E-Step computes the expected value of: \( P(C_j | x_i; \theta_t) \)
  - M-Step refines it using the MLE (\( \theta_t \)) obtained from:

\[
\theta_{t+1} = \arg\max_{\theta} \sum_i \sum_j P(C_j | x_i; \theta_t) \ln P(x_i, C_j | \theta_t)
\]
Common algorithm in many learning problems
Iterative algorithm divided into two steps:
- E-Step computes the expected value of: $P(C_j|x_i; \theta_t)$
- M-Step refines it using the MLE ($\theta_t$) obtained from:

$$\theta_{t+1} = \arg\max_\theta \sum_i \sum_j P(C_j|x_i; \theta_t) \ln (P(x_i|C_j; \theta_t)P(C_j))$$

$P(C_j)$ is updated every iteration using:

$$P(C_j) = \frac{1}{N} \sum_{j=1}^{N} P(C_j|x_i; \theta_t)$$
Another way of defining the problem is:

- E-Step computes the expected value of: \( P(C_j | x_i; \theta_t) \)
- M-Step refines it using the MLE (\( \theta_t \)) obtained from:

\[
\sum_i \sum_j P(C_j | x_i; \theta_t) \frac{d}{d\theta_j} \ln P(x_i | C_j; \theta_t) = 0
\]

- \( P(C_j) \) is updated every iteration using:

\[
P(C_j) = \frac{1}{N} \sum_{j=1}^{N} P(C_j | x_i; \theta_t)
\]
Applying to Ground Detection

- \( x_i \) are the point correspondences using SIFT
- \( \Theta_i \) are the parameters in:

\[
H = \hat{A}(R + \frac{t}{d} n^T)A^{-1}
\]

- The likelihood \( P(x_i|C_j; \theta_t) \) can be defined as:

\[
P(x_i|C_{ground}; \theta_t) = \frac{\exp\left(-\frac{err_i^2}{2\sigma^2}\right)}{\sum_i \exp\left(-\frac{err_i^2}{2\sigma^2}\right)}
\]

- where:

\[
err_i = \|\hat{p}_i - \frac{H_{ground}p_i}{s}\|
\]
Applying to Ground Detection

- Choice of $\sigma$ (empirically)
Recall (M-Step):

\[ \theta_{t+1} = \arg \max_\theta \sum_i \sum_j P(C_j | x_i; \theta_t) \ln (P(x_i | C_j; \theta_t) P(C_j)) \]

Problem: To update \( \theta_t \) we have to take the derivative of \( P(x_i | C_{ground}; \theta_t) \)

Solution: Use an optimization to find a “optimal” solution to the M-Step
Modified EM

- Choice of optimization algorithm

![Graph comparing Simplex vs Levenberg Marquardt](image-url)
Navigation Algorithm

- Initialization
  - Only needs to be performed once
- Feature Detection and Matching
  - Collect two images
  - Run SIFT to determine pixel correspondences (x)
- Pixel Classification
  - Classify pixels using Modified EM algorithm
- Robot Motion Control
  - Calculate new heading for mobile robot
Navigation algorithm tested on an HP Pavilion running Intel® Core™2 Duo CPU @ 2.0GHz

Used two P3DX mobile robots from Mobile Robots Inc.
- One robot “wandered“ down the hallway
- Second robot tracked the “wandering“ robot using the navigation above
Results

- 1000 frame collected
  - 96,075 correspondences to classify
  - Ground truth collected for random 200 frames

<table>
<thead>
<tr>
<th></th>
<th>Total number of pixels Classified</th>
<th>Correct Classification Percentage</th>
<th>Incorrect classification from SIFT</th>
<th>Incorrect classification from MEM alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground</td>
<td>88,145</td>
<td>99.62%</td>
<td>270 (0.3%)</td>
<td>71 (0.08%)</td>
</tr>
<tr>
<td>Non-ground</td>
<td>7,930</td>
<td>99.4%</td>
<td>24 (0.3%)</td>
<td>22 (0.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>96,075</td>
<td>99.6%</td>
<td>294 (0.3%)</td>
<td>93 (0.1%)</td>
</tr>
</tbody>
</table>

- Total Correct Classification Percentage: 99.6%
Example Results
Conclusions and Current Work

- Introduced a Modified EM algorithm for Ground Plane Detection
- An improved version to detect multiple planes
- Applied to outdoor scenes