Visual Homing with a Pan-Tilt based Stereo Camera

Paramesh Nirmal, Student Member SPIE
Fordham University Robotics and Computer Vision Laboratory
Bronx, NY
www.cis.fordham.edu/rcvlab
Visual Homing

- The ability of an agent to return to a previously visited location using vision-based techniques
- Agent has goal image stored in memory
Applications

• Navigation for autonomous robots

• Minimal Requirements
  • Terrestrial (2D) – Moller 2001, Vardy 2005, Churchill 2008
  • Aerial (3D) – Steinmann 2011

Homesick 😞
Holistic Approach

- Operates on whole image
- Warping method – Franz et al, Moller
- Warps current image by until it looks like goal image
Feature-based Approach

• Look for correspondence between goal image and current image
• Homing in Scale Space (HiSS) – Churchill, Vardy 2008
• Uses omni-directional camera for full field view
• Look for match in SIFT features
• No depth information, only scale of feature (whether it has expanded or contracted)
Feature-based Approach

(a) Contracted features

(b) Expanded features
Sensor Fusion

- Sensor Fusion for visual homing
- Laser Camera Fusion – Choi et al. (IROS ‘12)
- Make use of visual info as well as depth info
Stereovision

- Passive visual sensor
- Disparity to determine depth
- Biologically inspired – human beings also use other depth perception cues
Pan-Tilt

- Stereo-view has a limited field
- Pan-tilt unit allows robot to obtain a wide-field view
- Construct a panorama by concatenating images
- Panoramic depth image
Panoramic images

• Imperfect “panoramas”
• Increases chances of SIFT match
• Reduces computational complexity
• No disadvantage of unstitched panoramas
Homing with Stereo Vision (HSV)

- \( f = \{ f^u, f^v, f^\theta, f^\sigma, f^d \} \) – original SIFT vector, used in Homing in Scale Space
- \( f = \{ f^x, f^y, f^z, f^\alpha, f^d \} \) – redefined w/ stereo-data
  - \( \alpha = \text{atan2}(f^x, f^z) \)

SIFT tools obtained from David Lowe: http://www.cs.ubc.ca/~lowe/keypoints/
Homing with Stereo Vision (HSV)

HSV ($G$ – home image, home stereo-data)

1. Capture current image
2. Perform SIFT feature match $M$
3. For each matched feature $m$ in $M$:
   - $\theta m = G^{i,\alpha} - C^{j,\alpha}$  \hspace{1cm} $m = (i, j)$
   - $d m = G^{i,z} - C^{j,z}$  \hspace{1cm} $m = (i, j)$
4. Move by $h = [\Theta, D]$
5. Repeat 1-4 until $\Theta$ and $D$ are small enough

$$\Theta = \frac{1}{n} \sum_{m \in M} (\theta m)$$
$$D = \frac{1}{n} \sum_{m \in M} (d m)$$

($n$ matched features)
Feature Match
Feature Match
Feature Match
Performance

- Converges to goal location
Results

- HSV vs HiSS (Homing in Scale Space)
  - Reduced angular error (59% reduction)

Angular Error (HSV vs HiSS)
Results

- HSV vs HiSS (Homing in Scale Space)
  - Reduced number of steps (37% reduction)

# Steps (HSV vs HiSS)

![Bar chart showing the number of steps for different position differences (m) for HSV and HiSS, with a total average of 8 steps. HSV has fewer steps compared to HiSS for all position differences.]
Results

- HSV vs HiSS (Homing in Scale Space)
  - Increased return rate (61% increase)

Return Rate (HSV vs HiSS)

- Position difference (m)
- Return Rate HSV
- Return Rate HiSS

Results

- HSV vs HiSS (Homing in Scale Space)
  - Increased return rate (61% increase)
Distance error

• Limited accuracy
• Depends on image resolution

• 20cm apart
• Unnoticeable difference
Conclusions

• Algorithm is egocentric, no topological info
• Assumptions:
  • No obstructions
• Independent of dynamics in the environment
• HSV is enhanced by superior knowledge of the environment
• Theoretically can home in a single step
• Completely dependent on SIFT match
• Pan-tilt takes time
Acknowledgements

• Defense Threat Reduction Agency (DTRA) award #HDTRA1-11-1-0038 to GeorgiaTech and Fordham University supports my graduate study at Fordham University

• Special thanks to Dr. Damian Lyons for his valuable mentorship and support