Visual SLAM for Autonomous Hull Inspection

Ryan Eustice
Perceptual Robotics Laboratory (PeRL)
University of Michigan
Acknowledgements

Collaborators

- **MIT**: Franz Hover, Brendan Englot, Hordur Johannsson, Michael Kaess, John Leonard
- **Bluefin Robotics**: Jerome Vaganay
- **SeaByte**: Jose Vazquez, Scott Reed
- **FAU**: Pierre-Philippe Beaujean

Funding

- **ONR POC**: Dr. Thomas Swean
  Office of Naval Research
- **Technical POC**: Ms. Victoria Steward
  NAVEODTECHDIV
Why Underwater Inspection?

- Safety and security purposes
- No other easy way to inspect large ships

Courtesy of Seaward Marine
Autonomous Underwater Inspection

- **Challenges to localization**
  - No GPS underwater
  - Long Base Line (LBL) difficult to use in harbors
  - Metal ship hulls => Magnetic compass
  - Complex environment to navigate

- **Proposed solution**
  - In-situ navigation using SLAM:
    - Vision (Eustice, Kim)
    - Imaging sonar (Johannsson, Kaess, Leonard)
    - Path-planning for 100% coverage (Hover, Englot)

- **Benefits**
  - Ensure full coverage
  - Avoid hazards and restricted areas
  - Re-acquire targets / Change detection
  - No additional infrastructure required
Hovering Autonomous Underwater Vehicle (HAUV)

- Equipped with
  - 8 Thrusters (full 6-DOF)
  - Ring laser gyro
  - Sonars:
    - Doppler Velocity Log (DVL)
    - Multi-beam imaging sonar (DIDSON)
  - Both sonars are actuated
  - Camera + light strobe

Ship hull inspection by hull-relative navigation and control, OCEANS 2005
Planning Complex Inspection Tasks Using Redundant Roadmaps

Brendan Englot and Franz Hover
Department of Mechanical Engineering, MIT, USA
Inspecting the “Complex Areas” of a Ship

- At the stern of a ship, protruding structures make 100% sensor coverage difficult to achieve.

- These structures require the use of DIDSON in profiling mode, giving 30° range scans rather than 2D images.

- 3D mesh models are built by scanning the ship in low-resolution, long-range viewing mode (10m range).

- Model is used for planning an inspection in high-resolution, short-range viewing mode (5m range), suitable for detecting mines.

Top: Views of USCGC Seneca propeller, using DIDSON. Bottom: 3D Mesh constructed from full sonar dataset.
Our Planning Method: sample robot configurations and record their observations.

Ensure that every geometric primitive in the mesh is observed at least \( n \) times.

Use this redundant roadmap as the state space from which an inspection path is designed, using the iterative procedure detailed below:
Sampling-Based HAUUV Inspection Paths

The Redundant Roadmap method quickly computes a feasible inspection path, a suitable starting point for an improvement procedure.

Inspection Path for 100% Coverage of SS Curtiss, designed in less than two minutes by constructing a roadmap of redundancy 10.

Shortened tour after several hours of sampling-based improvement
Autonomous Ship Hull Inspection – Real-Time 6-DOF SLAM

red: camera footprint
blue: sonar footprint
Final SLAM Result (100% Camera Coverage Survey)
Sonar and Camera Derived Constraints
SLAM vs. Dead-Reckoned Navigation
Tested with a Variety of Hulls

2008 – USS Saratoga (324 m)

2010 – R/V Oceanus (54 m)

2010 – USCGC Venturous (64 m)

2011 – M/V Terry Bordelon (46 m)

2011 – USCGC Seneca (82 m)

2011 – SS Curtiss (183 m)
Intuition behind visual SLAM

- Odometry only (DVL dead-reckoned)

- Vision system

Nonsequential constraint
- camera

Sequential constraint
- odometry
Sonar Constraints – Processing

(a) Initial sonar image

(b) Smoothed

(c) Gradient

(d) Threshold

(e) Clustering

(f) Extracted Features

Initial → Smoothing → Gradient → Threshold → Clustering → Final
Sonar Constraints – Registration

• Registration – NDT (Normal Distribution Transform, Biber 2004)
  – Compact representation
  – Loose correspondences

• When to accept a registration?
  – Conservative threshold to avoid wrong matches

Frame A  B registered in A  Frame B
State Estimation

- We optimize over the full trajectory (smoothing!)

Camera or Sonar

Pose graph including loops

- Incremental smoothing and mapping (iSAM) for efficient solution and access to covariances
  (open source at http://people.csail.mit.edu/kaess/isam)

"iSAM: Incremental Smoothing and Mapping" by M. Kaess, A. Ranganathan, and F. Dellaert,
IEEE Trans. on Robotics, TRO, vol. 24, no. 6, Dec. 2008, pp. 1365-1378,
Visual SLAM on a Clean Hull

- R/V Oceanus, Woods Hole Oceanographic Institution
- Jan 2010
Visual SLAM Results for R/V Oceanus

- **Search survey**
  - (4) legs, each 20 m in length
  - 425 images @ 1 fps
  - 185 cross-track matches

- Uncertainty increases monotonically (DR)
- Uncertainty is bounded (VAN)
Example Along-Track Registration (Strong Prior)
Example Cross-Track Registration (Strong Prior)
Example Cross-Track Registration (Weak Prior)
Image Saliency for Active SLAM

- Not all images are equal for SLAM
- Adapt vehicle trajectory based upon localization error and image saliency
Visual Saliency using Bag-of-words

Ayoung Kim
Bag of Words Visual Saliency for Active SLAM
Saliency Definition
Saliency Definition

- Local saliency
  - Single image
  - Color/Grayscale
  - Texture richness / Registrability

- Global saliency
  - Image stream
  - Color/Grayscale
  - Rarity
Bag of Words (BoW) Vocabulary Representation

- Vocabulary should be representative
  - Build vocab. online.
  - Start from zero vocab.


- Independence between input images
  - Update vocab. when robot has enough spatial separation

- Related Work

Saliency Definition

- Local Saliency
  - Entropy
  - Histogram of BoW
    \[ e = - \sum_{i=1}^{\mid w \mid} p(w_i) \log_2 p(w_i) \]
    \( (\mid w \mid = \text{size of vocabulary}) \)
  - Normalization is needed
    \[ S_L = \frac{\sum_i p(w_i) \log_2 p(w_i)}{\log_2 \mid w \mid} \]

- Global Saliency
  - Information
  - Statistics of BoW (idf)
    \[ s_i(t) = \sum_{w=1}^{n_d} \log_2 \frac{N(t)}{n_w(t)} \]
    \( (n_d = \text{Total number of images}) \)
  - Normalization is needed
    \[ S_{G,i}(t) = \frac{\sum_{w=1}^{n_d} \log_2 N(t)/n_w(t)}{\max_j s_j(t)} \]
    \( (N = \text{Total number of images}) \)
Saliency Evaluation
Evaluation - Local Saliency BoW Measure

- Normalized entropy score over image vocab distribution
- Compares well to a Hue derived entropy measure, but has the advantage of working equally well with grayscale

Johnson-Roberson, M. Large-Scale Multi-sensor 3D Reconstructions and Visualizations of Unstructured Underwater Environments The University of Sydney, 2010

Evaluation - Global Saliency BoW Measure

- Normalized cumulative inverse document frequency (idf)
- Discriminates based upon temporal occurrence

More salient

(a) $S_G=0.78 / S_L=0.81$
(b) $S_G=0.74 / S_L=0.77$
(c) $S_G=0.72 / S_L=0.78$
(d) $S_G=0.47 / S_L=0.74$
(e) $S_G=0.46 / S_L=0.65$

Less salient

(f) $S_G=0.66 / S_L=0.73$
(g) $S_G=0.65 / S_L=0.76$
(h) $S_G=0.61 / S_L=0.73$
(i) $S_G=0.49 / S_L=0.61$
(j) $S_G=0.47 / S_L=0.68$
Evaluation - Global Saliency

- Normalized cumulative inverse document frequency ($\text{idf}$)
- Discriminates based upon temporal occurrence

More salient

(a) $S_G=0.78 / S_L=0.81$
(b) $S_G=0.74 / S_L=0.77$
(c) $S_G=0.72 / S_L=0.75$

(f) $S_G=0.66 / S_L=0.73$
(g) $S_G=0.65 / S_L=0.76$
(h)
Hull Saliency Maps

(c) Local / global saliency map on USCGC Venturous.

(d) Local saliency map on SS Curtiss.
Saliency Incorporated
Information Gain
Geometric Information Gain

- Geometrical information gain (Ila, 2010)
  - EIF (Extended Information Filter) \( \eta = \Lambda \mu \) and \( \Lambda = \Sigma^{-1} \)
  - Mutual information defined from entropy
    \[
    I_g = \sum_{x \in X, z_i \in Z} p(x, z_i) \log \frac{p(x, z_i)}{p(x)p(z_i)}
    \]
    \[
    = H(X) + H(Z) - H(X, Z)
    \]
    \[
    = H(X) - H(X|Z),
    \]
    \[
    = \frac{1}{2} \ln \frac{|\Lambda + \Delta \Lambda|}{|\Lambda|} \quad \Delta \Lambda = H^T \Sigma_y^{-1} H
    \]
  - Expected information gain by making a measurement

Information-Based Compact Pose SLAM, IEEE Trans. Robotics, 2010
V. Ila, J. Porta, J. Andrade-Cetto
Information Gain *with* Local Saliency

- **Local saliency**
  - Build map efficiently.
  - Remove unlike candidates.

43.7% removed by imposing 0.6 local saliency thresh.
Future Work

- Incorporating visual saliency with control for active SLAM
- Multi-modality registration in complex areas
- Re-localization and managing large maps
- Robust handling of correspondence errors