3D Modeling Using a Statistical Sensor

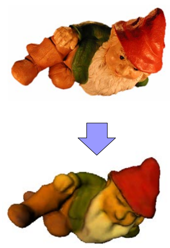
Daniel Huber and Martial Hebert

The Robotics

Introduction

Problem

- Automatically construct a digital 3D model of a real-world scene
- Input data from a 3D sensor (e.g., laser scanner) from unknown viewpoints and with unknown ordering (spatial or temporal)



Contributions

- Principled method to evaluate model hypothesis quality
- Stochastic search algorithm to find optimal model hypothesis
- New hypothesis representation allows multi-part output models when necessary

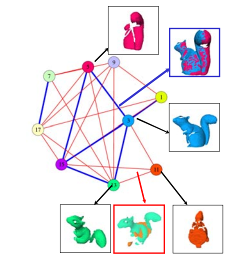
Model graphs and model hypotheses

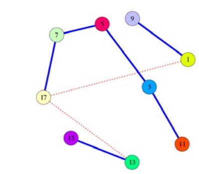
A **model graph** encodes topological relationships between overlapping views

A **model hypothesis** is a spanning tree of a model graph with an additional "visibility" attribute

= correct match = incorrect match
(hand-labeled for illustration)

= visible = hidden



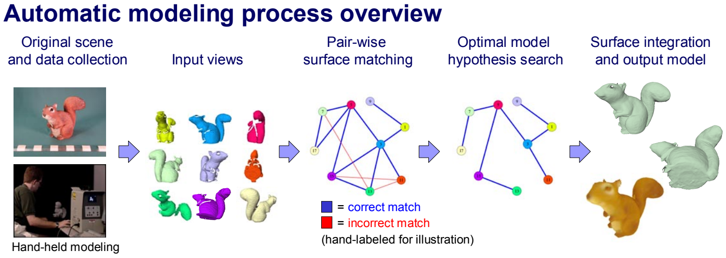


Example model hypothesis
(with three parts)

- Nodes are views
- Edges connect overlapping views (store relative pose)
- Edges arise from pair-wise surface matching
- Model graph from pair-wise surface matching is called G_{LR} (LR = local registration)

- Each hidden edge separates a model into two parts
- Visible edges connect all views within a model part
- Spanning tree simplifies search for optimal hypothesis without restricting hypothesis space

Automatic modeling process overview



Evaluating model hypothesis quality

Maximum likelihood local quality model

- Start with a simpler problem: What is the quality of registration between a pair of views?
- Want to avoid using fixed, user-defined thresholds, which are brittle
- Instead, we derive a statistical model using Bayesian decision theory

M^+ = correct match
M^- = incorrect match
x = features derived from data

Goal: estimate $P(M^+|x)$

Define local quality (Q_L) as log odds ratio

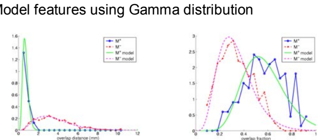
$$Q_L = \log\left(\frac{P(M^+|x)}{1 - P(M^+|x)}\right) = \log\left(\frac{P(M^+|x)}{P(M^-|x)}\right)$$

$$\frac{\Pr(M^+|x)}{\Pr(M^-|x)} = \frac{P(x|M^+)}{P(x|M^-)}\frac{P(M^+)}{P(M^-)}$$

prior probabilities

model of sensor and registration error

- Estimate $P(x|M^+)$ and $P(x|M^-)$ from labeled training data
  - Maximum likelihood parametric density estimation
  - Model features using Gamma distribution



- Estimate $P(M^+)$ and $P(M^-)$ directly from frequency in training data
- Also, define Classifier (C_L)
  - Decide M^+ if $Q_L > \lambda$, M^- otherwise
  - Used to remove worst matches from G_{LR}
- Several local quality measures have been derived using this framework, including one based on overlap distance and two based on visibility consistency

Benefits

- Combine multiple independent features in principled manner
- Features with unrelated units can be employed (e.g., color or texture similarity)

Disadvantages

- Requires labeled training data
- Separate model for each data collection scenario (e.g., buildings, terrain, small objects)

Overlap local quality measure

- This quality measure consists of two features:
  - **Overlap fraction** (F_{OV}) – Fraction of two surfaces that overlap
  - **Overlap distance** (D_{OV}) – RMS distance between surfaces in overlapping region
- For two registered surfaces S_i and S_j,

$$F_{OV} = \max\left(\frac{A(R_i)}{A(S_i)}, \frac{A(R_j)}{A(S_j)}\right)$$

A = surface area
R_i (R_j) = overlapping region of S_i (S_j)

- Estimate F_{OV} by sampling K points from S_i and from S_j, and counting the number (N_i and N_j) that overlap

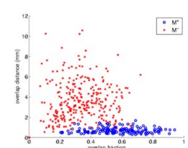
$$F_{OV} \approx \max\left(\frac{N_i}{K}, \frac{N_j}{K}\right)$$

- Similarly, estimate D_{OV} from the distances (d_{ik} and d_{jk}) between closest points in the overlapping region

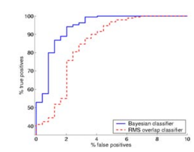
$$D_{OV} = \sqrt{\frac{\sum_{k=1}^{K}(d_{ik}^2 + d_{jk}^2)}{N_i + N_j}}$$

- Assuming independence between F_{OV} and D_{OV}, this leads to the overlap local quality measure definition

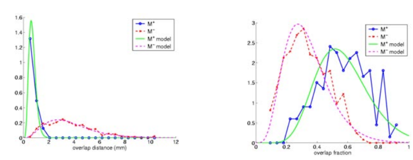
$$Q_L = \log\left(\frac{P(D_{OV}|M^+)P(F_{OV}|M^+)P(M^+)}{P(D_{OV}|M^-)P(F_{OV}|M^-)P(M^-)}\right)$$



Joint distribution of F_{OV} and D_{OV}



ROC curve comparing overlap quality measure classification performance to traditional RMS distance measure



Marginal distributions for F_{OV} and D_{OV} for a set of 443 matches (153 correct and 290 incorrect) obtained by exhaustive pair-wise surface matching on three real objects. Maximum likelihood Gamma distributions are shown overlaid.

Local quality examples

Q_L = 6.21	Q_L = 4.29	Q_L = -8.11	Q_L = -24.8
C_L = true	C_L = true	C_L = true	C_L = true

Extending to an entire model hypothesis

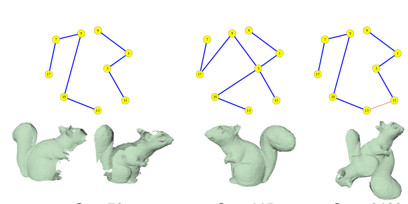
- Compute global quality (Q_G) by summing the local quality (Q_L) over all pairs of connected (not necessarily adjacent) views in G:

$$Q_G(G) = \sum_{(i,j)\in V_c} Q_L(V_i, V_j, T_{i,j})$$

V_c = set of connected views in G

- Using all connected views improves power of the quality measure by using additional constraints between non-adjacent views
- Model the effect of accumulating relative pose error, which is not encoded in the local quality model, by learning local quality models as a function of path length

Global quality examples



Q_G = 70	Q_G = 115	Q_G = -3183