Vehicle Recognition in Cluttered Environments

Masters Thesis Defense

By Gerald Dalley

Signal Analysis and Machine Perception Laboratory
The Ohio State University

05 June 2002

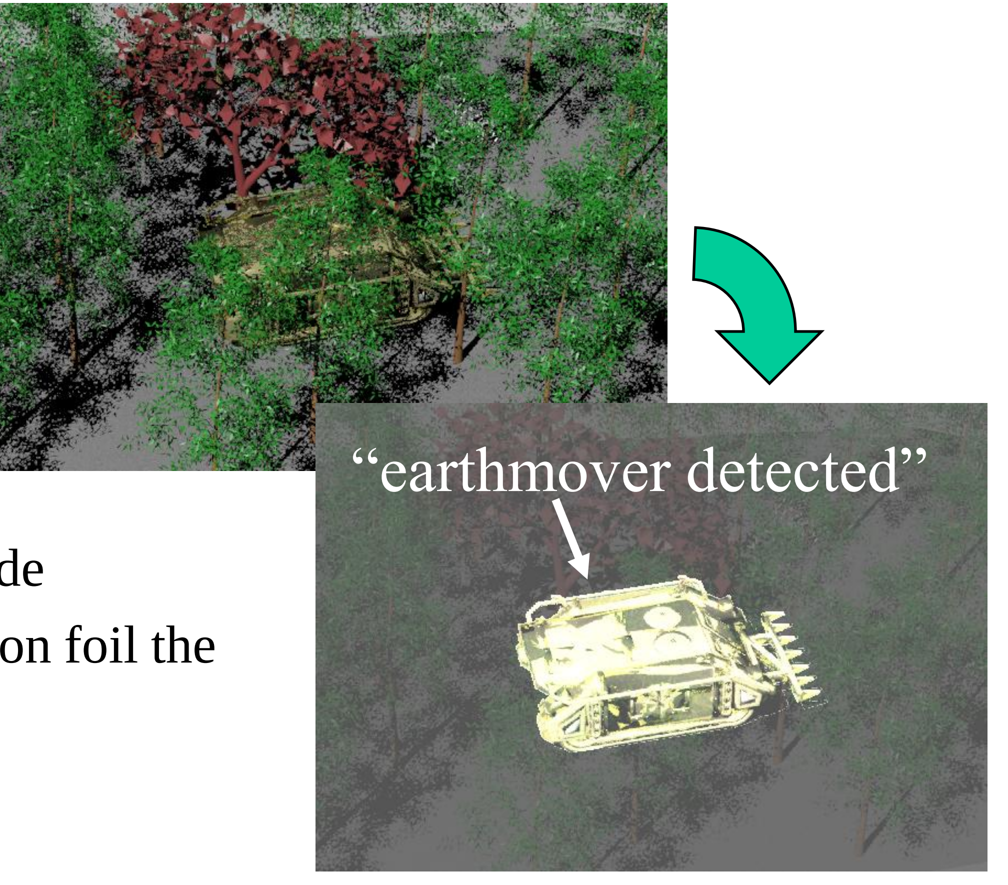
Overview

- Problem Statement and Motivation
- Recognition Steps
  - ➢ Range Image Generation
  - ➢ Local Surface Estimation and Decimation
  - ➢ Global Surface Reconstruction
  - ➢ Surface Segmentation
  - ➢ Graph Matching
- Conclusions and Future Work
- Questions

Problem Statement and Motivation

- Problem
  - Recognize vehicles
  - Military and civilian
  - Forested environment

- Motivation
  - Hostile forces tend to hide
  - Camouflage and occlusion foil the human visual system



Overview

- Objects modeled
- Clutter models
- Camera flight paths (scenes)
- Noise generation

Range Image Generation

Local Surface Fitting

Surface Reconstruction

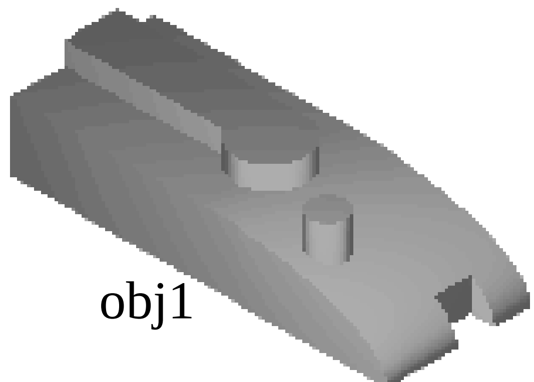
Surface Segmentation

Graph Matching

Range Image Generation:

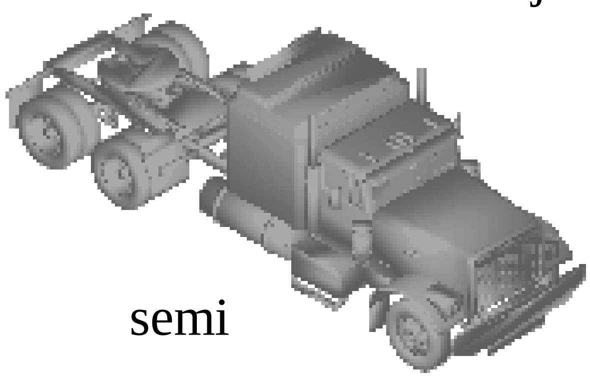
Objects Modeled

earthmover



obj1

sedan

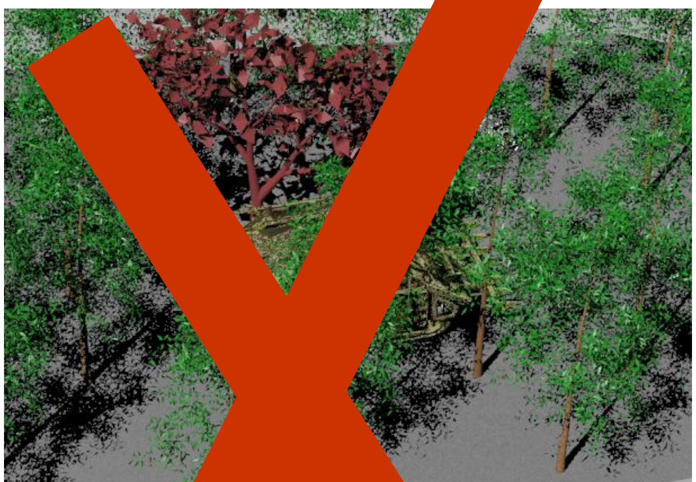


semi

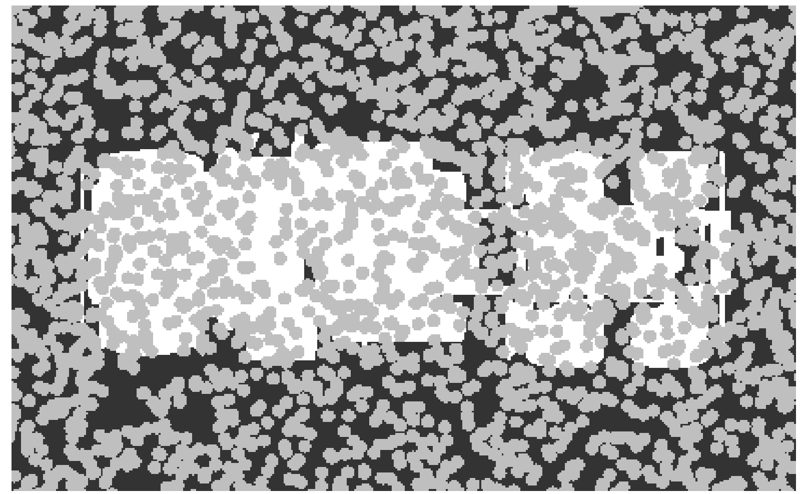
tank

Range Image Generation:

Clutter Models



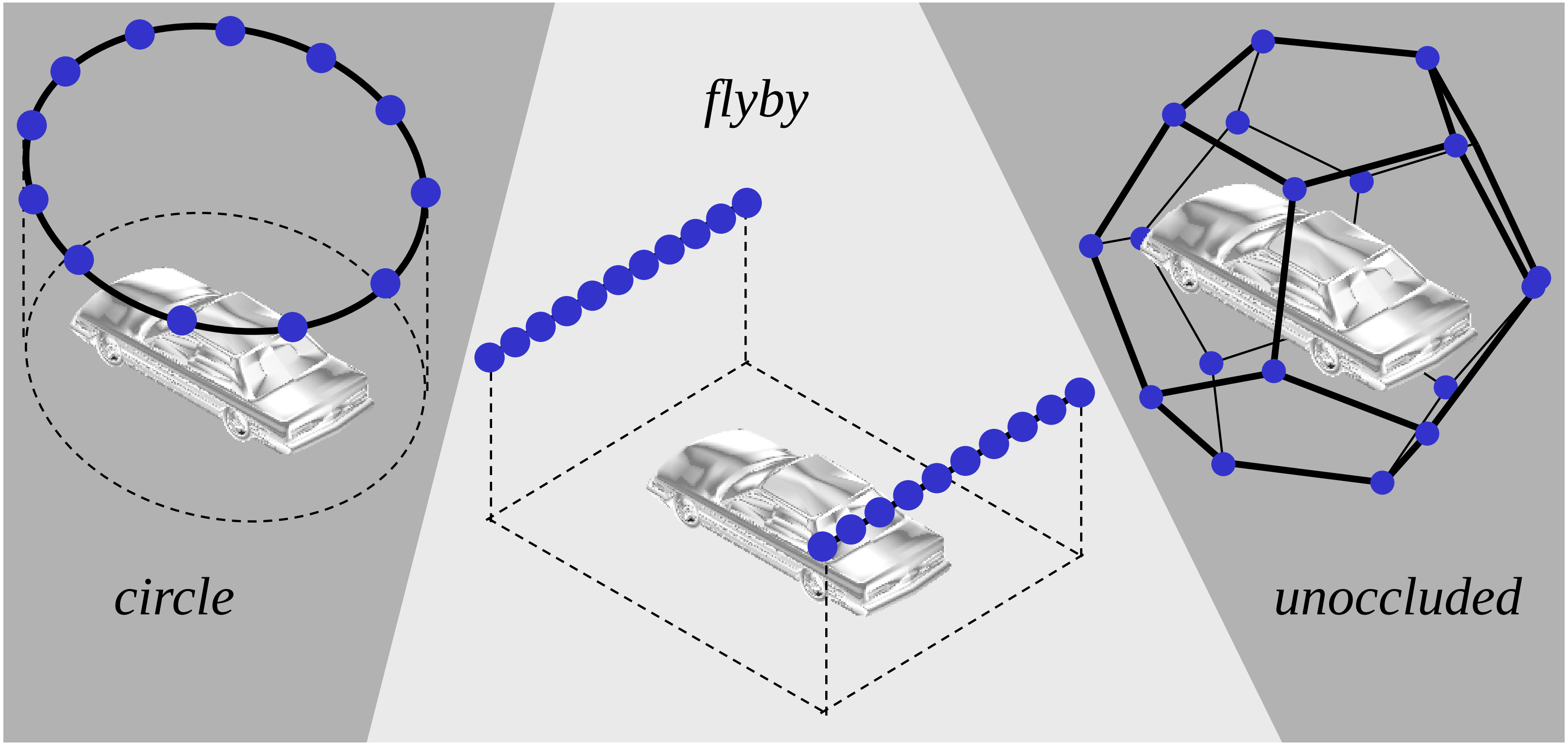
- Vis   re   c trees
  - ok goo   ..
  - oor occlusi
  - Very long runt



- 500 Discs
- Radius of 100mm
- Volume of 12.2 x 12.2 x 2 meters

Range Image Generation:

Camera Flight Paths (Scenes)



Range Image Generation:

Noise Generation

- Isotropic additive Gaussian noise
- Standard deviations of:
  - 0mm
  - 2mm
  - 4mm
  - 8mm
  - 16mm
  - 32mm

Local Surface Estimation and Decimation:

Overview

- **Assumption:** Vehicles are composed primarily of large, low-order, low-curvature surfaces.
- **Constraint:** 10 tank views → more than 220,000 range points (too many)

- Point Selection (Decimation)
- Principle Component Analysis
- Biquadratic Surface Fits

Range Image Generation

Local Surface Estimation

Surface Reconstruction

Surface Segmentation

Graph Matching

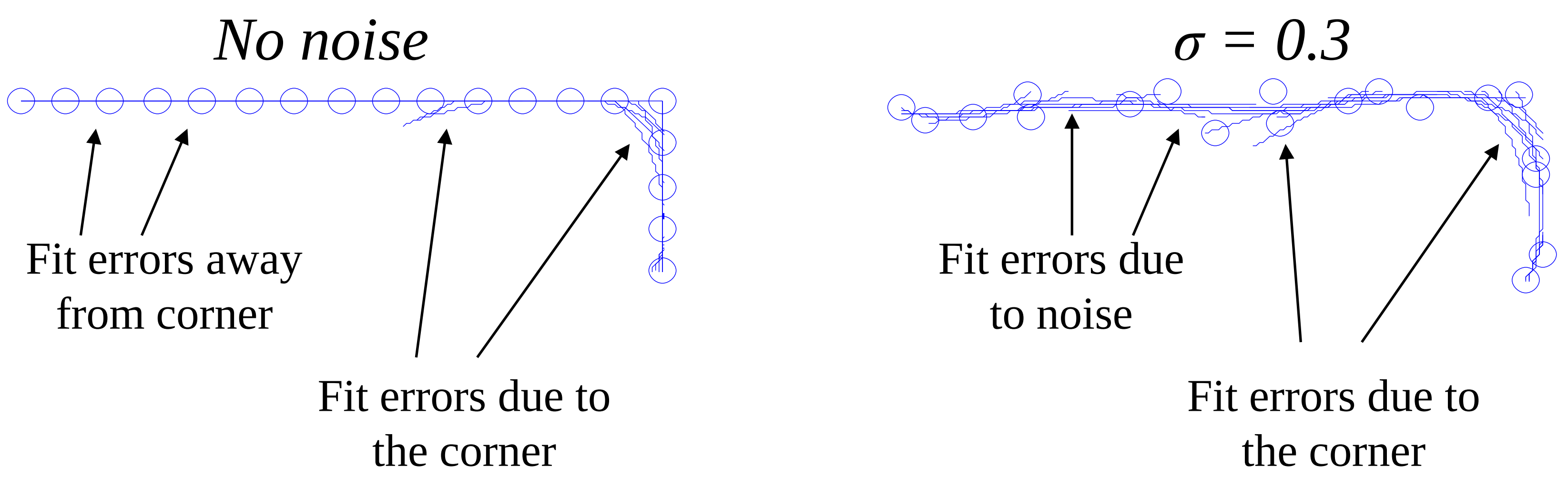
Local Surface Estimation and Decimation:

Point Selection

Method 1:

1. Randomly select 1% (for example) of the original points
2. Make local surface estimates based on selected points

Problems:



Point Selection (cont'd.)

Method 2: In the region of interest…

- Collect range image points into cubic voxel bins (*128x128x128mm*)
- Discard bins that have:
  - Too few points
  - Points that do not represent biquadratic surfaces well
- Retain only the centroids of the bins and their surface fits

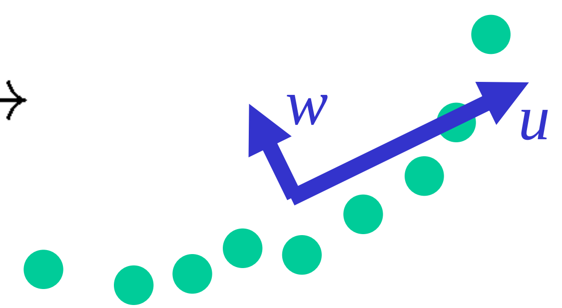
Local Surface Estimation and Decimation:

Principle Component Analysis

- Define a local neighborhood $N_i = \{p_{i_j}\}$ about point p_i
- Find the centroid of the neighborhood, $\widehat{p}_i$
- Find the 3x3 covariance matrix:

$$\mathbf{S} = \sum_{j=1}^{|N_i|} \left[p_{i_j} - \widehat{p}_i\right] \left[p_{i_j} - \widehat{p}_i\right]^T$$

- Smallest eigenvector of $\mathbf{S} \rightarrow$ normal estimate.

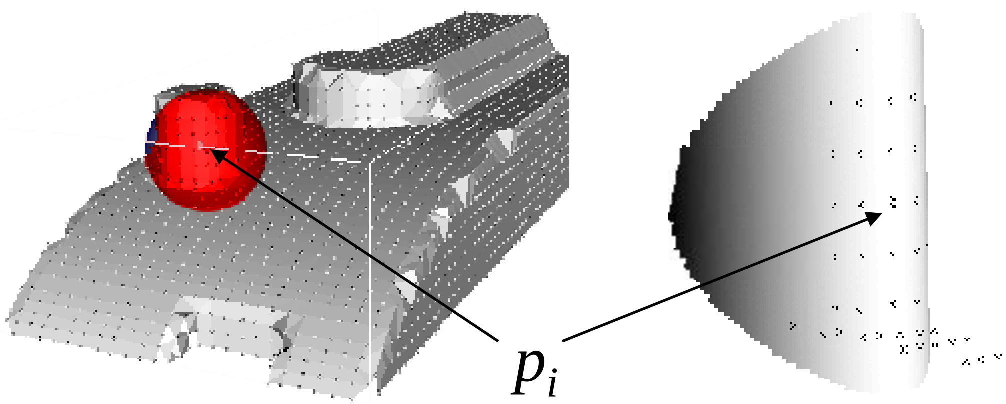


Local Surface Estimation and Decimation:

Biquadratic Surface Fits

$$w = f(u,v) = a_0u^2 + a_1uv + a_2v^2 + a_3u + a_4v + a_5$$

$$\begin{bmatrix} u_0^2 & u_0v_0 & v_0^2 & u_0 & v_0 & 1 \\ u_1^2 & u_1v_1 & v_1^2 & u_1 & v_1 & 1 \\ u_2^2 & u_2v_2 & v_2^2 & u_2 & v_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \approx \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \end{bmatrix}$$



Global Surface Reconstruction:

Overview

- Motivations
- Post-Processing

Range Image Generation

Local Surface Fitting

Surface Reconstruction

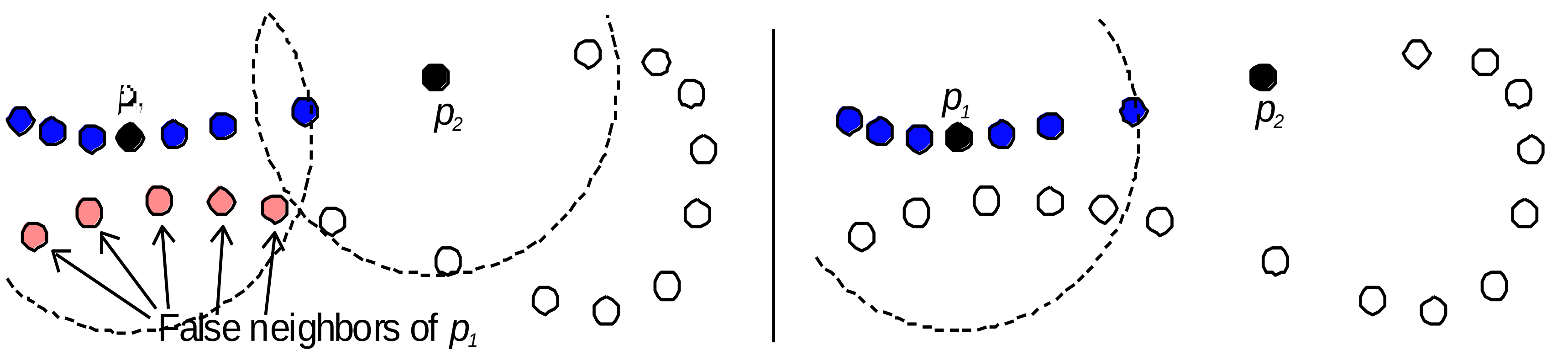
Surface Segmentation

Graph Matching

Global Surface Reconstruction:

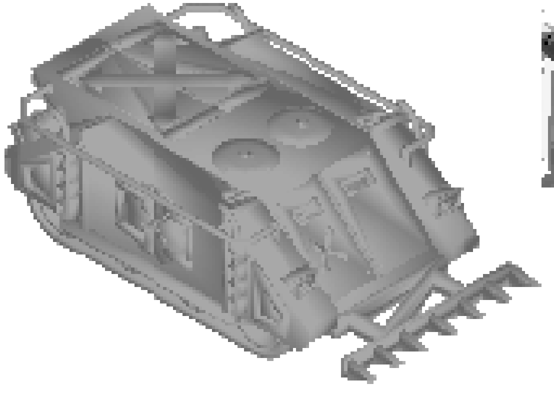
Motivations

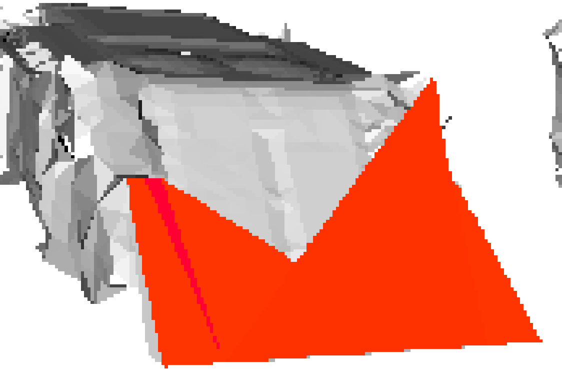
- Easy, unambiguous nearest-neighbor identification
- Fast searches over small cardinality
- Makes rendering easier
- Avoids incorrect groupings of nearby surfaces

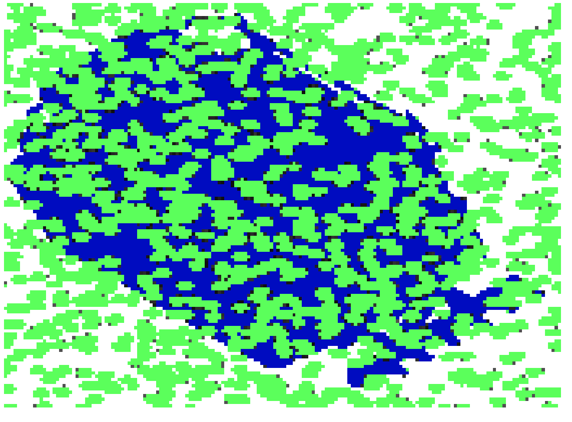


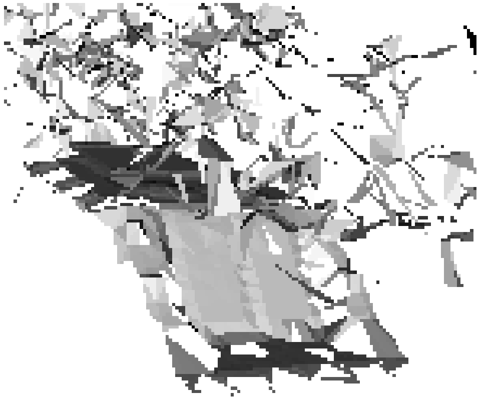
Global Surface Reconstruction:

Post-Processing









Surface Segmentation:

Overview

- **Motivation:** Correspondence is hard
- Some Techniques Not Used
- Spectral Clustering
  - ➢ An overview
  - ➢ Normalized cuts
  - ➢ Our affinity measure
- Results

Range Image Generation

Local Surface Fitting

Surface Reconstruction

Surface Segmentation

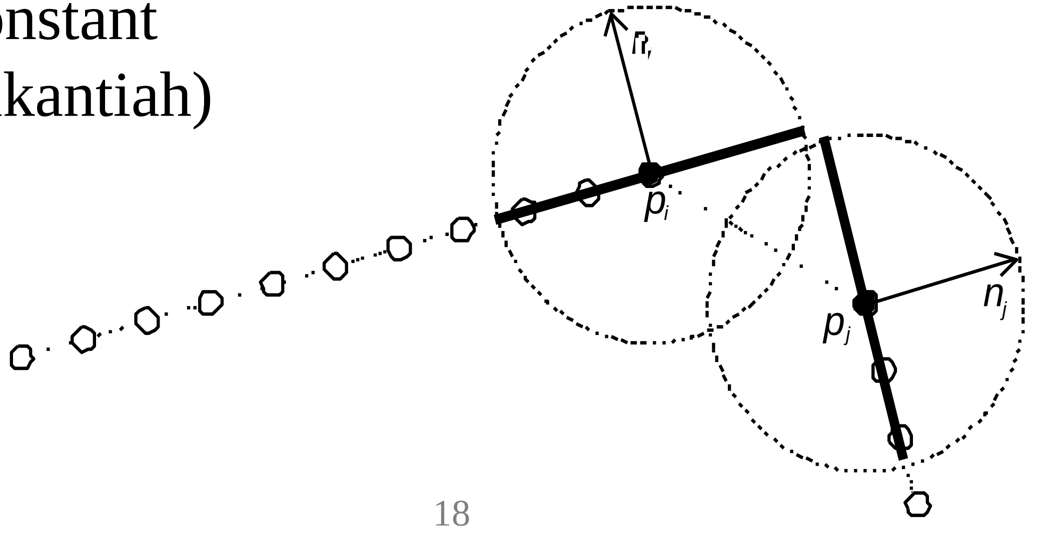
Graph Matching

Surface Segmentation:

Some Techniques Not Used

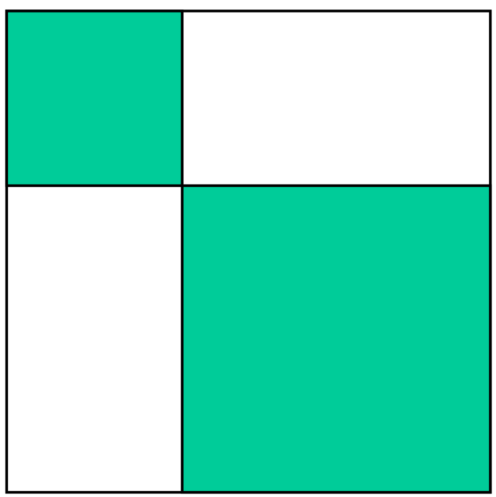
Robust Sequential Estimators (Mirza)

Regions of Constant Curvature (Srikantiah)

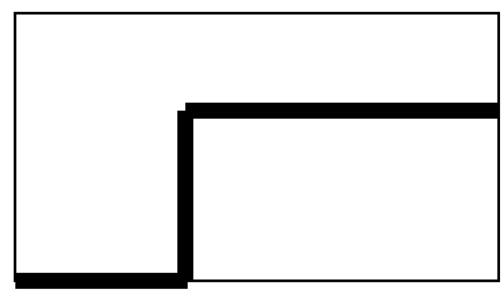


Overview of Spectral Clustering

- Two surface points have an affinity…



A_{ij}



y_1, where

$Ay_i = \lambda_i y_i$

Surface Segmentation:

Normalized Cuts

- What is being minimized:
  - $cut\,(B,C) = \sum_{b \in B, c \in C} \mathbf{A}_{b,c}$
  - $assoc\,(B,V) = \sum_{b \in B, v \in V} \mathbf{A}_{b,v}$
  - $Ncut\,(B,C) = \frac{cut(B,C)}{assoc(B,V)} + \frac{cut(B,C)}{assoc(C,V)}$

Our Affinity Measure

$$v_{i,j} = p_i - p_j$$

$$E \geq 1$$

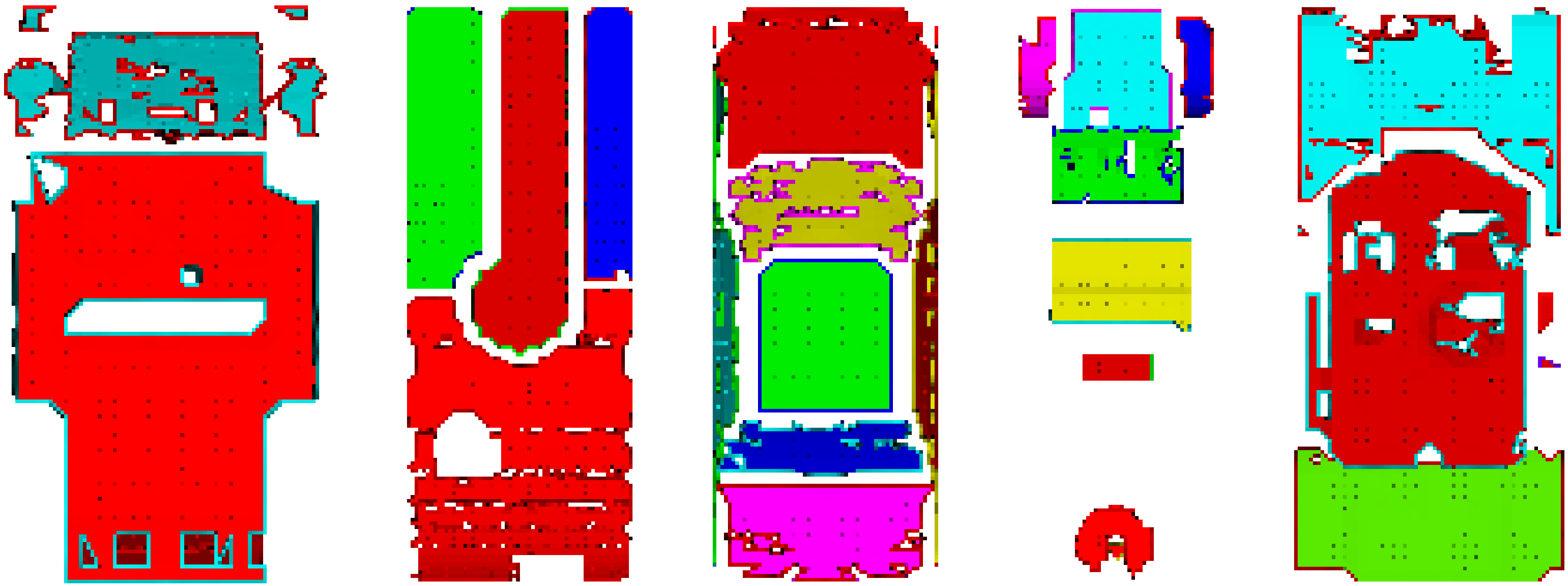
$$d\left(p_i, p_j\right) = \sqrt{E\left(v_{i,j} \circ n_i\right)^2 + \frac{||v_{i,j} - (v_{i,j} \circ n_i)n_i||^2}{E}}$$

$$\mathbf{a}_{i,j} = exp\left(-\frac{\cos^{-1}(n_i \circ n_j)}{2\left(\frac{10^o \pi}{180^o}\right)^2}\right) exp\left(-\frac{d(p_i, p_j)}{2r^2}\right)$$

$$\mathbf{A}_{i,j} = min\left(\mathbf{a}_{i,j}, \mathbf{a}_{j,i}\right)$$

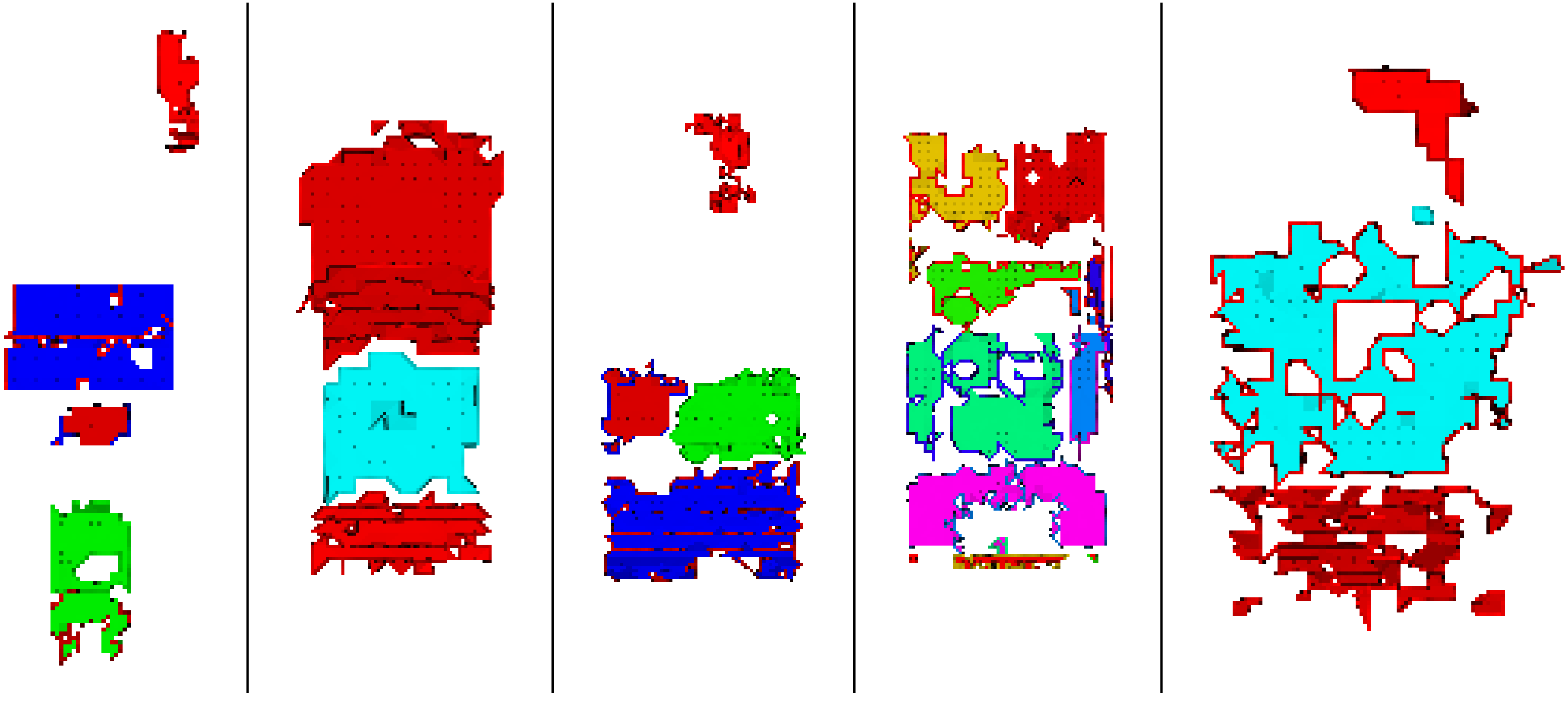
Surface Segmentation:

Unoccluded Results



Surface Segmentation:

Which Objects Are These?



Graph Matching:

Overview

- Match tree example
- Error measures
- Entropy
- Results
- What caused problems?

Range Image Generation

Local Surface Fitting

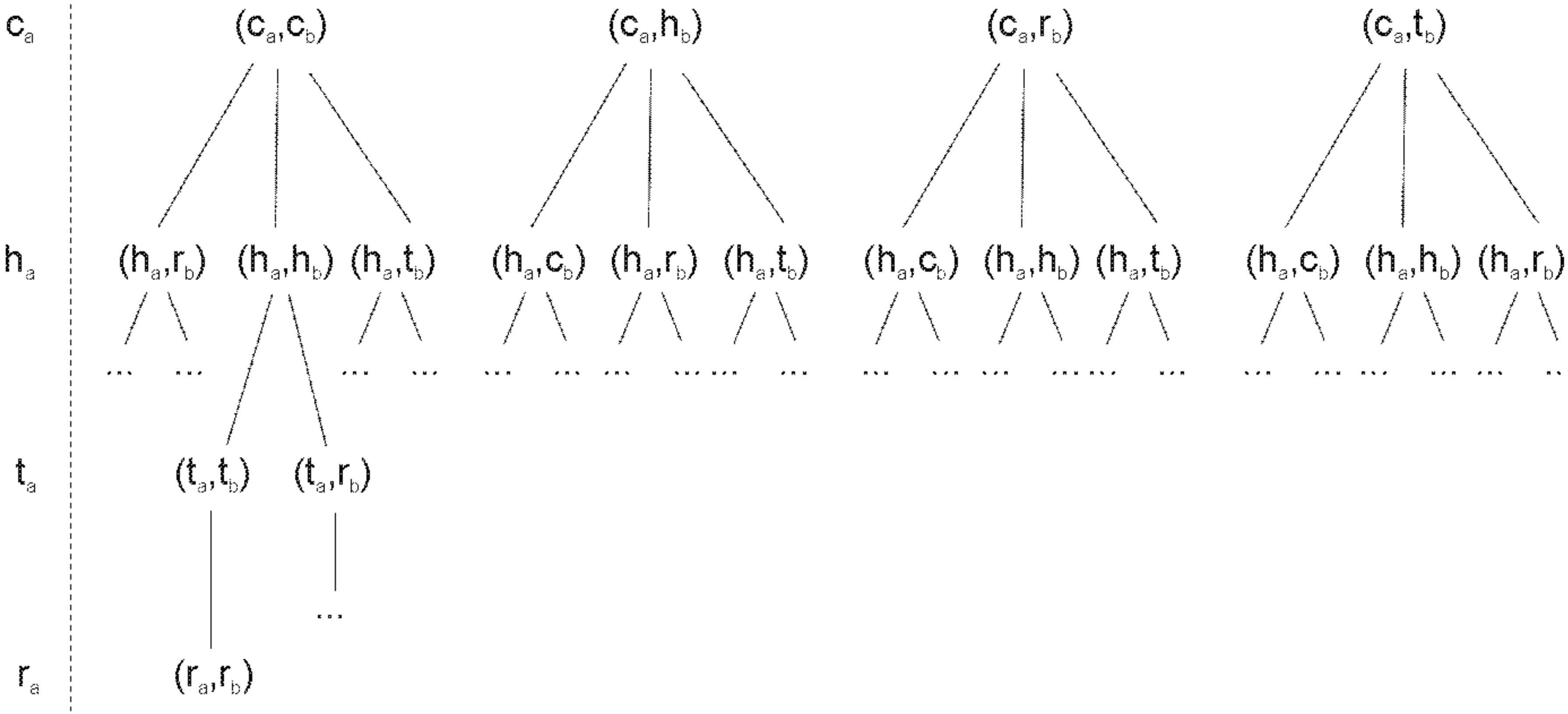
Surface Reconstruction

Surface Segmentation

Graph Matching

Graph Matching:

Match Tree Example



Graph Matching:

Error Measures

- Unary Error
  - Area, Elongation, Thickness
- Orientation Error
  - How poorly pairs of normals match up
- Centroid Distance Error
  - How poorly pairs of centroids match up
- Cumulative Area Error
  - What percentage of the model area is not matched up

Graph Matching:

Entropy

$$S\left(M\right) = 1 - E\left(M\right)$$

$$s\left(\{M_1, \ldots, M_N\}\right) = \sum_{k=1}^{N} S\left(M_k\right)$$

$$entropy\left(\{M_1, \ldots, M_N\}\right) - \sum_{k=1}^{N} \frac{S(M_k)}{s(M_k)} \, log_2 \left(\frac{S(M_k)}{s(M_k)}\right)$$

Graph Matching:

Results: earthmover

Scene	Noise Level	Score	Match Rank	Entropy
circle	0	0.824	1	2.16
circle	2	0.764	1	2.16
circle	4	0.819	1	2.19
circle	8	0.865	1	2.16
circle	16	0.797	1	2.15
circle	32	0.777	1	2.12
flyby	0	0.775	1	2.17
flyby	2	0.757	1	2.17
flyby	4	0.771	1	2.17
flyby	8	0.747	1	2.17
flyby	16	0.707	1	2.16
flyby	32	0.482	1	1.94
unoccluded	0	1.000	1	2.03
unoccluded	2	0.921	1	2.07
unoccluded	4	0.905	1	2.08
unoccluded	8	0.827	1	2.11
unoccluded	16	0.825	1	2.11
unoccluded	32	0.824	1	2.15

Recognition rate: 100%

Graph Matching:

Results: obj1

Scene	Noise Level	Score	Match Rank	Entropy
circle	0	0.756	2	1.86
circle	2	0.811	1	1.95
circle	4	0.816	1	1.96
circle	8	0.662	1	2.27
circle	16	0.468	4	2.25
circle	32	0.241	5	2.26
flyby	0	0.796	1	1.94
flyby	2	0.561	1	2.28
flyby	4	0.730	1	2.23
flyby	8	0.272	5	2.26
flyby	16	0.281	5	2.29
flyby	32	0.000	5	1.96
unoccluded	0	1.000	1	1.89
unoccluded	2	0.810	1	2.20
unoccluded	4	0.726	1	2.23
unoccluded	8	0.794	1	2.21
unoccluded	16	0.839	1	2.07
unoccluded	32	0.179	4	1.93
Recognition rate: 64%				

Graph Matching:

Results: sedan

Scene	Noise Level	Score	Match Rank	Entropy
circle	0	0.641	1	1.64
circle	2	0.723	1	1.65
circle	4	0.767	1	1.63
circle	8	0.589	1	1.69
circle	16	0.507	1	1.72
circle	32	0.486	1	1.71
flyby	0	0.683	1	1.61
flyby	2	0.576	1	1.68
flyby	4	0.662	1	1.62
flyby	8	0.602	1	1.68
flyby	16	0.447	1	1.72
flyby	32	0.291	1	1.71
unoccluded	0	1.000	1	1.46
unoccluded	2	0.937	1	1.50
unoccluded	4	0.894	1	1.52
unoccluded	8	0.918	1	1.51
unoccluded	16	0.666	1	1.67
unoccluded	32	0.451	1	1.70

Recognition rate: 100%

Graph Matching:

Results: semi

Scene	Noise Level	Score	Match Rank	Entropy
circle	0	0.767	1	2.08
circle	2	0.733	1	2.16
circle	4	0.722	1	2.13
circle	8	0.708	1	2.16
circle	16	0.703	1	2.17
circle	32	0.374	2	2.18
flyby	0	0.611	1	2.12
flyby	2	0.580	1	2.14
flyby	4	0.598	1	1.89
flyby	8	0.599	1	2.06
flyby	16	0.545	1	2.08
flyby	32	0.256	3	2.16
unoccluded	0	1.000	1	1.95
unoccluded	2	0.770	1	2.09
unoccluded	4	0.737	1	2.11
unoccluded	8	0.736	1	2.11
unoccluded	16	0.677	1	2.01
unoccluded	32	0.269	2	2.16
Recognition rate: 83%				

Graph Matching:

Results: tank

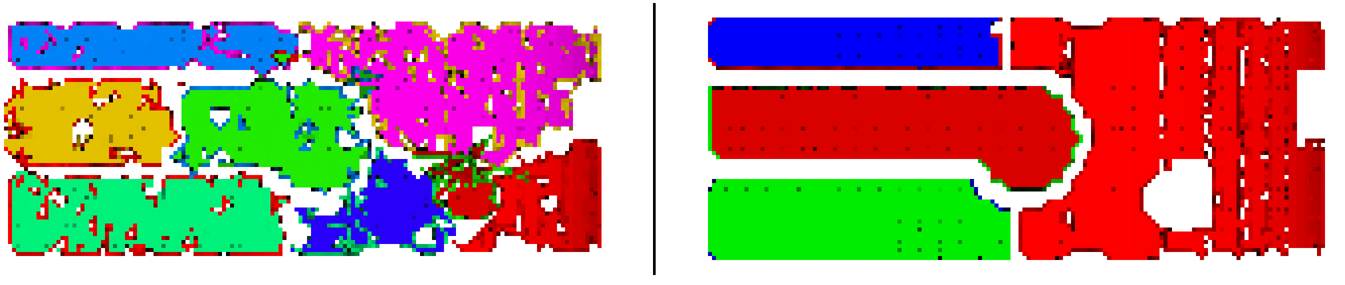
Scene	Noise Level	Score	Match Rank	Entropy
circle	0	0.603	2	2.26
circle	2	0.628	1	1.93
circle	4	0.682	1	1.92
circle	8	0.636	1	1.93
circle	16	0.588	1	1.95
circle	32	0.511	2	2.26
flyby	0	0.444	1	2.30
flyby	2	0.187	5	2.29
flyby	4	0.430	2	2.29
flyby	8	0.234	5	2.30
flyby	16	0.281	3	2.28
flyby	32	0.119	4	2.11
unoccluded	0	1.000	1	2.06
unoccluded	2	0.826	1	2.08
unoccluded	4	0.831	1	2.06
unoccluded	8	0.473	2	2.11
unoccluded	16	0.697	1	2.26
unoccluded	32	0.237	4	2.13

Recognition rate: 50%

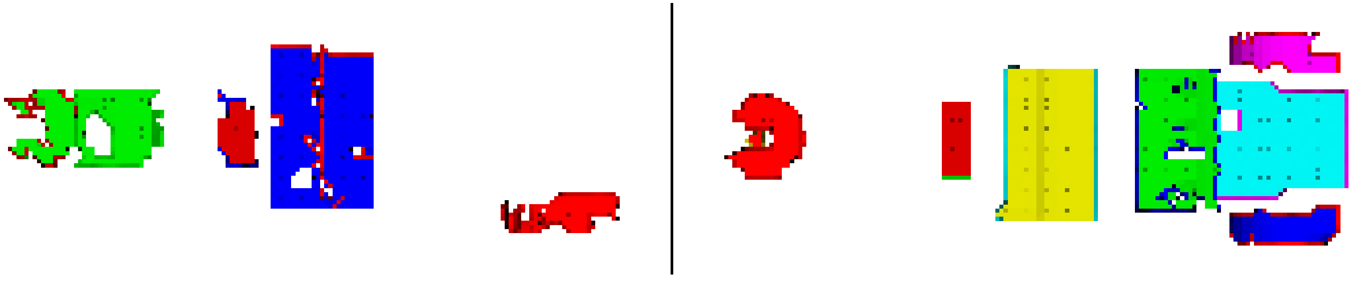
Graph Matching:

What Caused Problems?

- 19 total incorrect recognition results
  - 12: over-segmentation



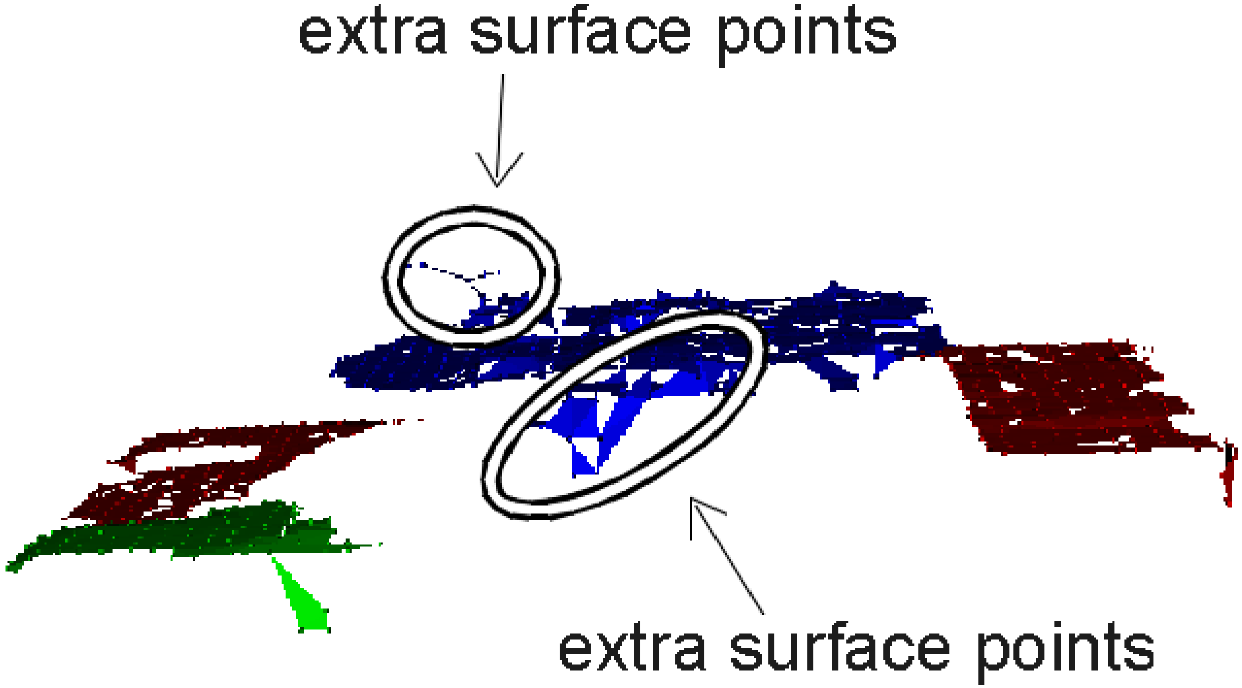
  - 10: area errors (including non-existent segments)



Graph Matching:

What Caused Problems? (cont'd.)

- 4: mis-aligned segmentation



Conclusions

- System features
  - Modular design
  - Handles pessimistic levels of clutter
  - 100% recognition on *earthmover* and *sedan*
- Reliable segmentation is important when doing graph matching

Future Work

- Articulation
- Larger modelbase
- Iterative recognition
- Alternative segmentation methods
  - Other affinity matrix normalizations
  - Tensor voting
  - Enhanced version of Srikantiah's algorithm
- Verification
- Alternative recognizers (*e.g.* SAI)
- E3D! (hopefully, for the remaining SAMPL crowd)

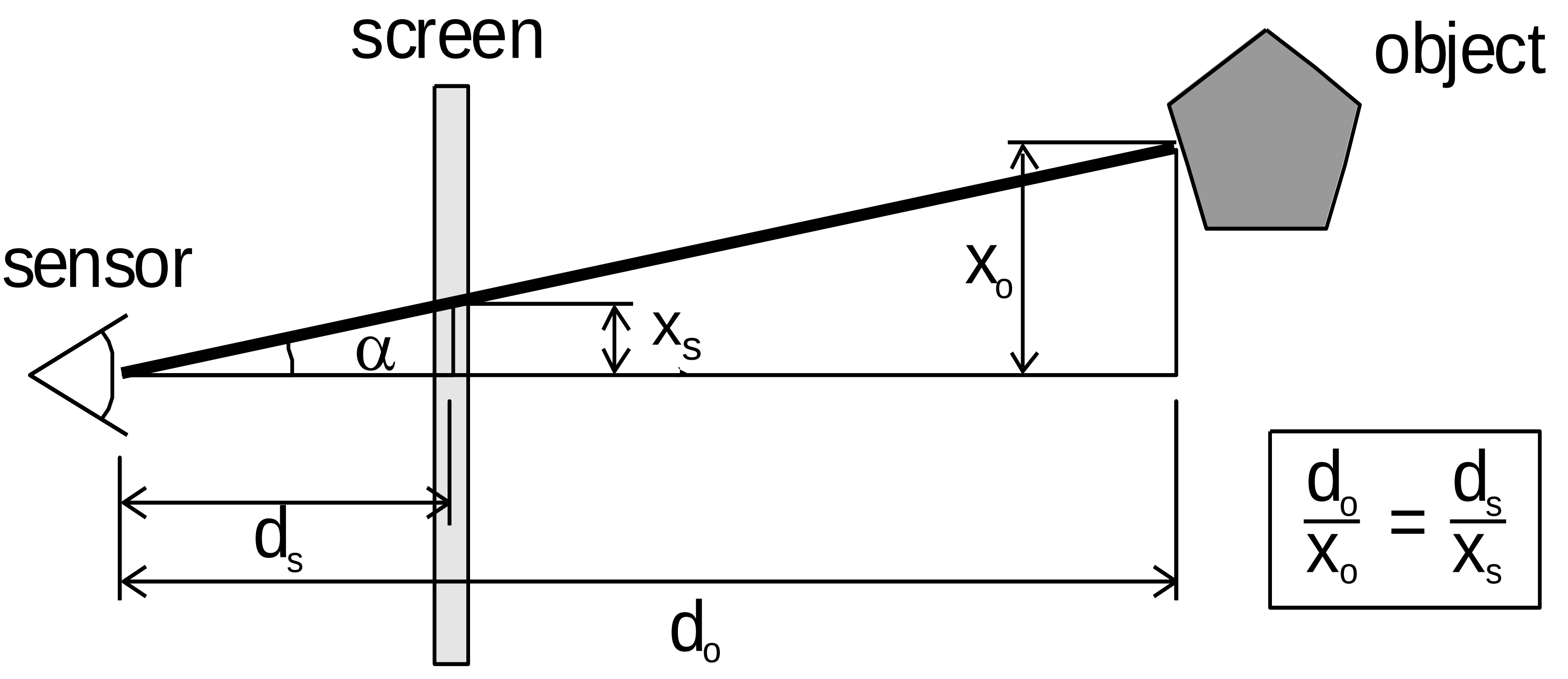
Your Questions…

EXTRA

SLIDES

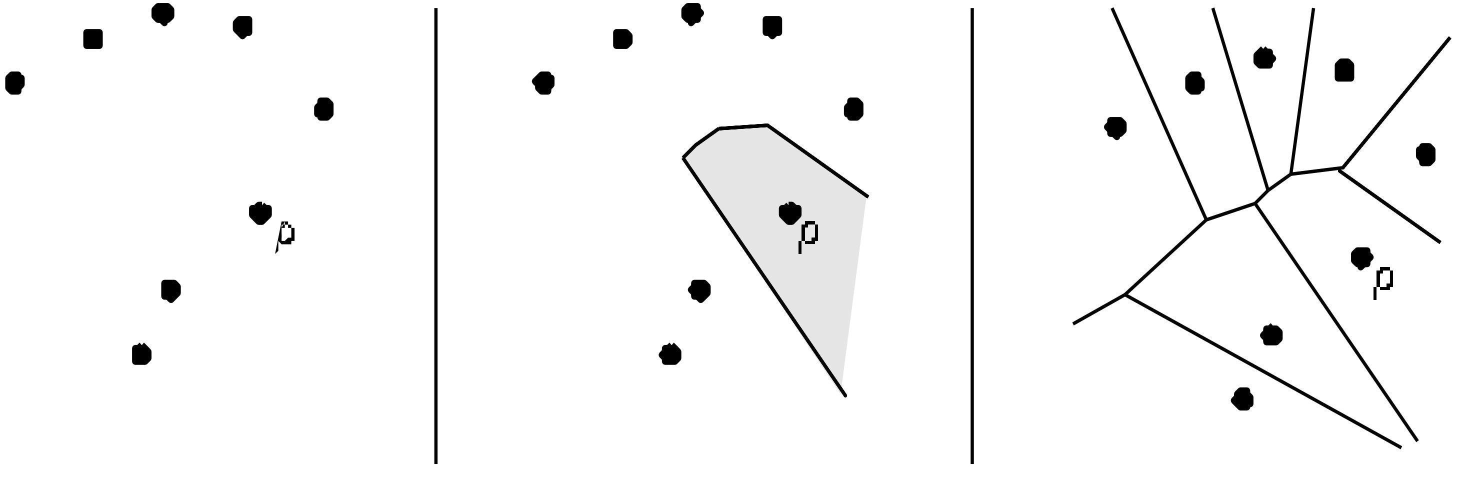
Range Image Generation:

Ray Tracing



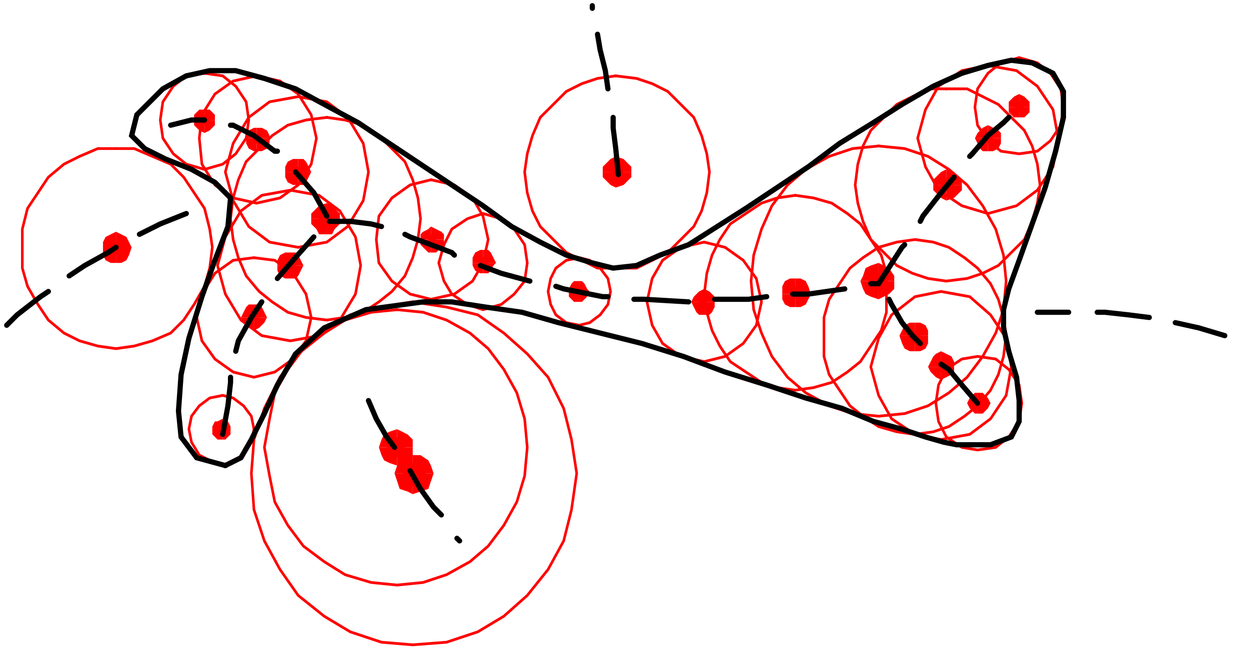
Voronoi Diagrams

- Voronoi cell = locus of points closer to a given sample point than any other point



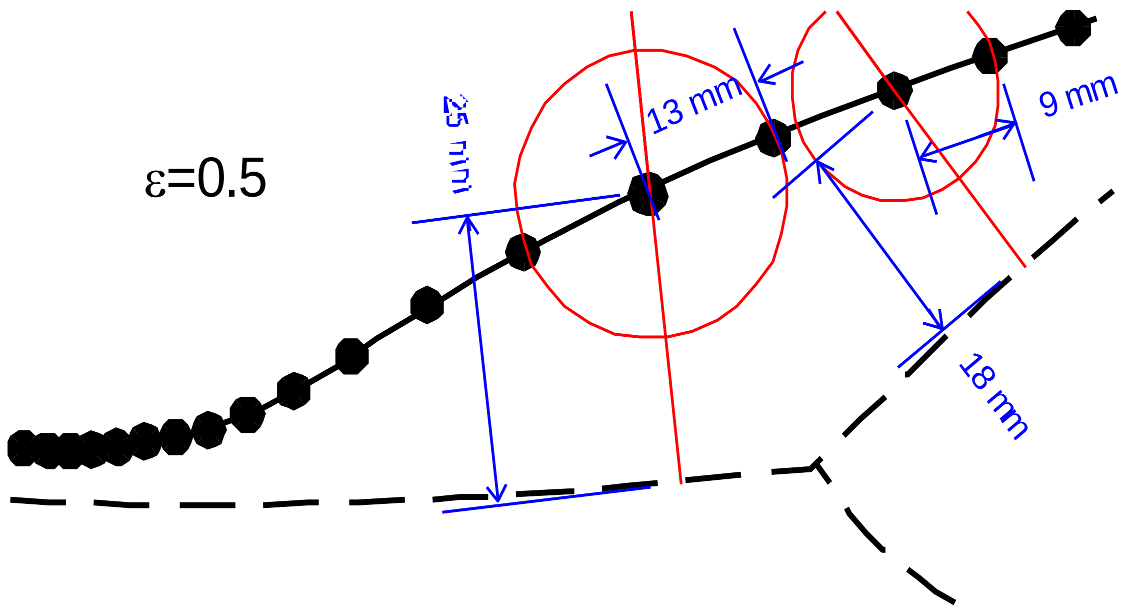
Medial Axis

- Medial axis = locus of points equidistant from at least two surface points (considering the *original* surface)

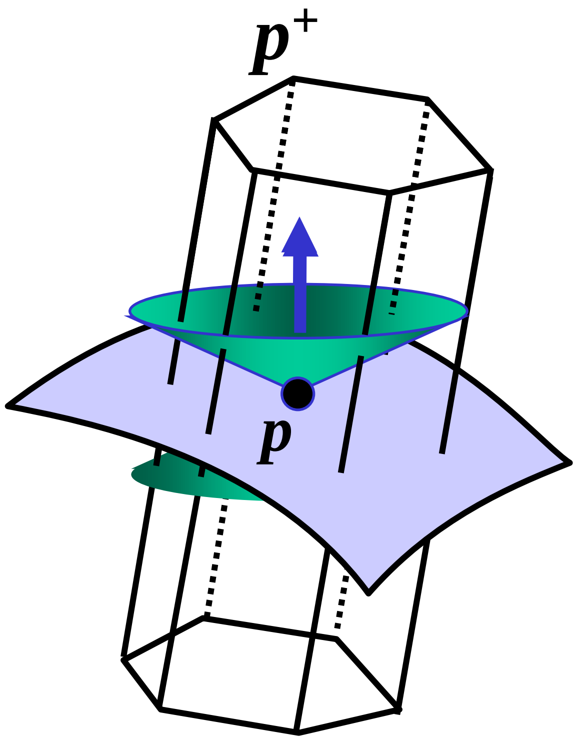


ε-sampling

- ε-sampling = Samples are at most ε times the distance to the medial axis



Cocone



Voronoi cell of ***p***

- $\boldsymbol{p}^+ \equiv$ pole of ***p*** = point in the Voronoi cell farthest from ***p***
- ε < *0.06* →
  - the vector from ***p*** to $\boldsymbol{p}^+$ is within π/8 of the true surface normal
  - The surface is nearly flat within the cell

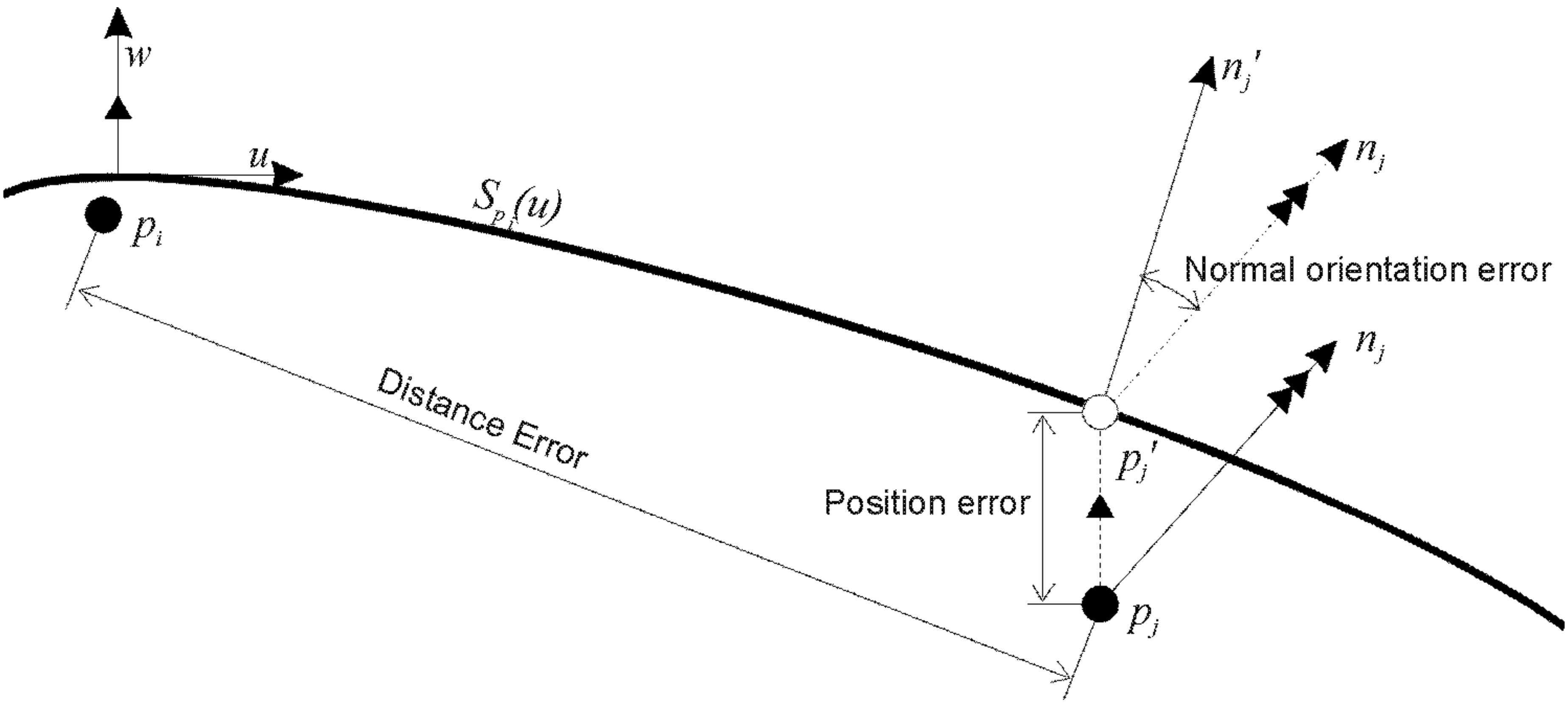
Surface Segmentation:

Normalized Cuts

- What is being minimized:
  - $cut(B,C) = \sum_{b \in B, c \in C} \mathbf{A}_{b,c}$
  - $assoc(B,V) = \sum_{b \in B, v \in V} \mathbf{A}_{b,v}$
  - $Ncut(B,C) = \frac{cut(B,C)}{assoc(B,V)} + \frac{cut(B,C)}{assoc(C,V)}$
- Matrix calculations
  - $\mathbf{D}_{i,i} = \sum_j \mathbf{A}_{i,u}$
  - $(\mathbf{D} - \mathbf{A}) y_i = \lambda_i \mathbf{D} y_i$

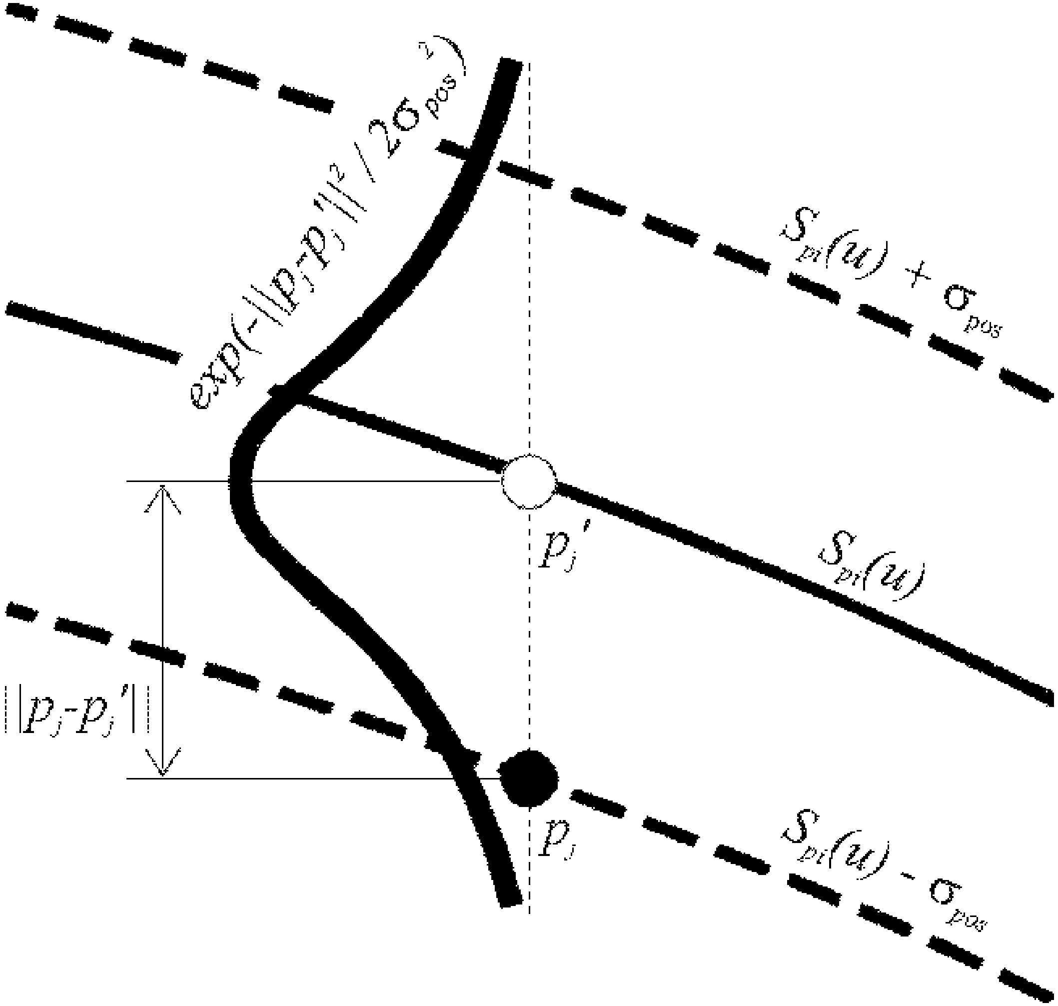
Surface Segmentation:

Probabilistic Affinity Framework



Surface Segmentation:

Probabilistic Position Affinity



Surface Segmentation:

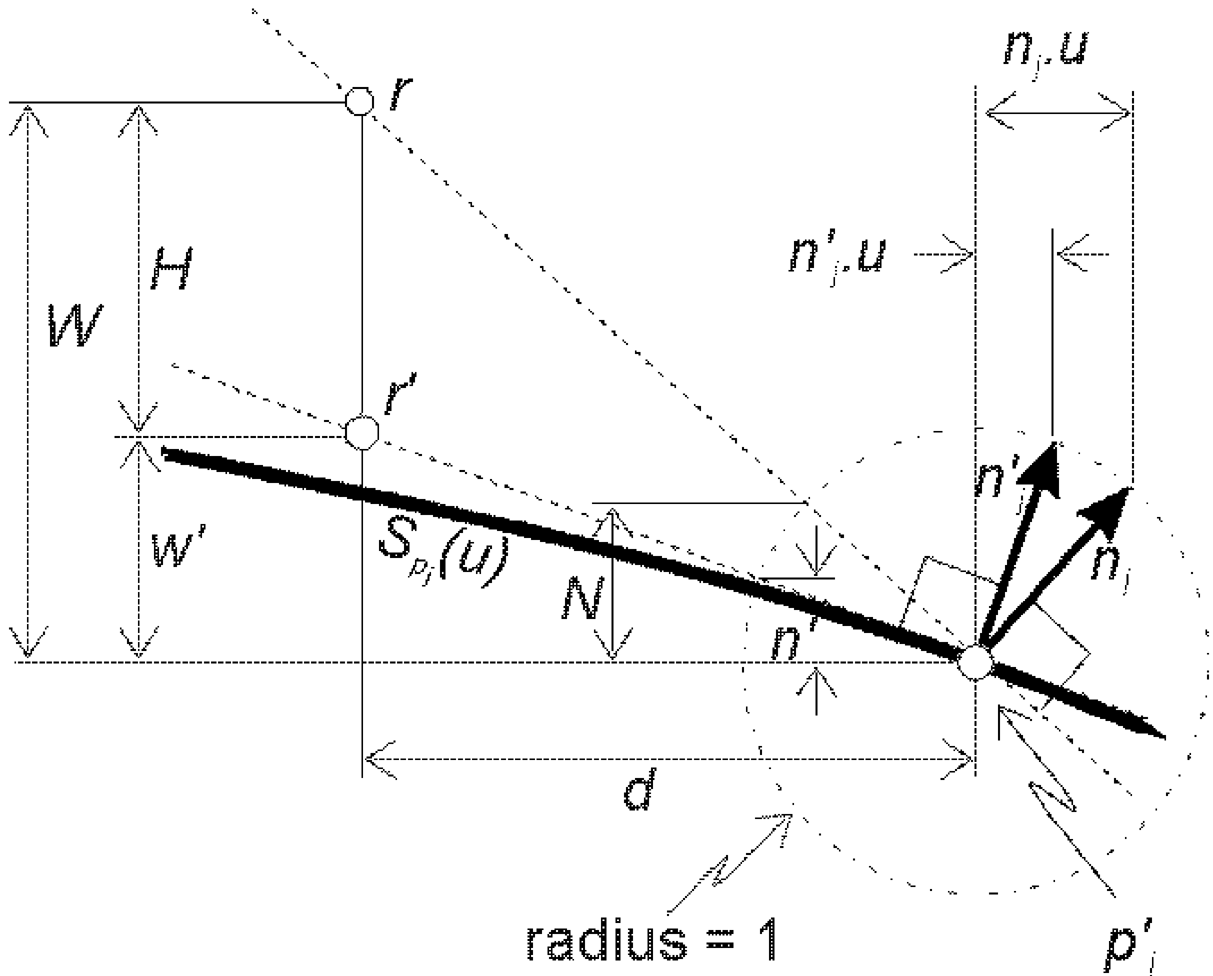
Probabilistic Position Affinity

$$P[p_j \in S_{p_i}(u,v)] \propto A_{pos}\,(p_i, p_j)$$

$$= \exp\left(-\frac{||p_j - p'_j||^2}{2\sigma_{pos}{}^2}\right)$$

Surface Segmentation:

Probabilistic Normal Affinity



Surface Segmentation:

Probabilistic Normal Affinity

$$A_{norm,2D}\left(p_i,p_j\right)=\frac{exp\left(-\left(\frac{nd}{\sqrt{1-n^2}}-\frac{n'd}{\sqrt{1-n'^2}}\right)^2\Big/4\sigma_{pos}{}^2\right)}{\left|\left(1\Big/\sqrt{d^2+\frac{(nd)^2}{1-n^2}}\right)-\frac{(nd)^2}{1-n^2}\left(d^2+\frac{(nd)^2}{1-n^2}\right)^{-3/2}\right|}$$

Graph Matching:

Error Measures

$$E_U(s_i, m_i) = 1 - exp\left(-\,{d_u}^2(s_i, m_i)\right) \qquad (1)$$

$$C_U(\{(s_1, m_1), ..., (s_x, m_x)\}) = \sum_{i=1}^{x} E_U(s_i, m_i) \qquad (2)$$

$$d_u(s_i, m_i) = \sqrt{{d_a}^2(s_i, m_i) + {d_t}^2(s_i, m_i) + {d_e}^2(s_i, m_i)} \qquad (3)$$

$$d_a(s_i, m_i) = -\,log_{(1.0+f_a)}\left(\frac{min(a_{s_i}, a_{m_i})}{max(a_{s_i}, a_{m_i})}\right) \qquad (4)$$

Graph Matching:

Error Measures (cont'd.)

$$d_t\left(s_i, m_i\right) = w_t|t_{s_i} - t_{m_i}| \quad (1)$$

$$d_e\left(s_i, m_i\right) = -w_e\, ln\left(1 - |e_{s_i} - e_{m_i}|\right) \quad (2)$$

$$E_O\left(s_i, s_j, m_i, m_j\right) = \frac{2}{\pi}\left|\left(cos^{-1}\left(n_{s_i} \circ n_{m_i}\right)\right) - \left(cos^{-1}\left(n_{s_j} \circ n_{m_j}\right)\right)\right| \quad (3)$$

$$C_O\left(\{(s_1, m_1), ..., (s_x, m_x)\}\right) = \sum_{i=1}^{x}\sum_{j=1}^{x} E_O\left(s_i, s_j, m_i, m_j\right) \quad (4)$$

Graph Matching:

Error Measures (cont'd.)

$$d_R(a,b) = \begin{cases} 0 & \text{if } a = 0 \text{ and } b = 0 \\ |a-b| / \, max(a,b) & \text{otherwise} \end{cases} \quad (1)$$

$$C_D(\{(s_1,m_1),...,(s_x,m_x)\}) = \sum_{i=1}^{x}\sum_{j=1}^{x} d_R\left(||c_{s_i} - c_{s_j}||, ||c_{m_i} - c_{m_j}||\right) \quad (2)$$

$$C_A(\{(s_1,m_1),...,(s_x,m_x)\}) = min\left(1, \left|1 - \sum_{i=1}^{x} a_{s_i} \Big/ \sum_{i=1}^{X} a_{m_i}\right|\right) \quad (3)$$

$$E(M) = \frac{1}{x} C_U(M)^{e_u} C_O(M)^{e_o} C_D(M)^{e_d} C_A(M)^{e_a} \quad (4)$$