Shape Context and Chamfer Matching in Cluttered Scenes

Presented by Jong Taek Lee

Index

Introduction

- How to detect a hand?

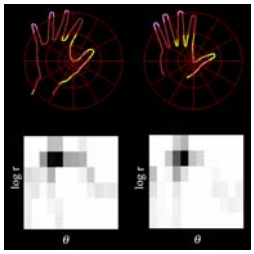
- Comparison of matching methods
  - Shape context vs. Chamfer matching
- Enhancements for shape context
  - Robustness to clutter

Previous Work

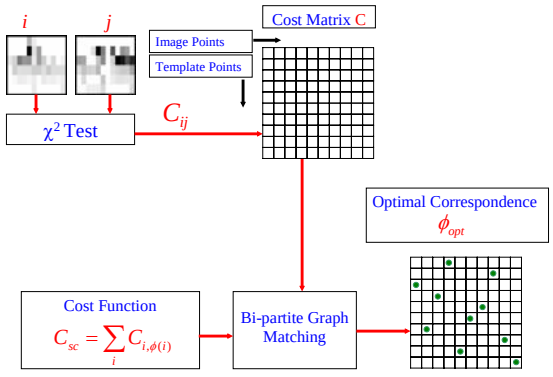
- Shape Context [Belongie *et al.*, 00]
  - Invariance to translation and scale
  - High performance in
    - Digit recognition : MNIST dataset
    - Silhouettes : MPEG-7 database
    - Common household objects: COIL-20 database
- Chamfer Matching [Barrow *et al.*, 77]
  - efficient hierarchical matching [Borgefors, 88]
  - pedestrian detection [Gavrila, 00]

Shape Context: Histogram

- Shape context of a point: log-polar histogram of the relative positions of all other points
- Similar points on shapes have similar histograms



Shape Context: Matching



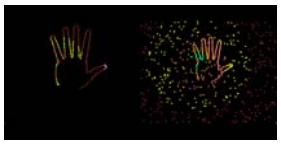
Shape-Context: Matching

Scale Invariance in Clutter ?

- Median of pairwise point distances is used as scale factor

50.5 41.6

- Clutter will affect this scale factor



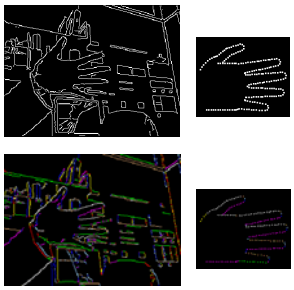
50.5 121.9

Scale Invariant in Clutter ?

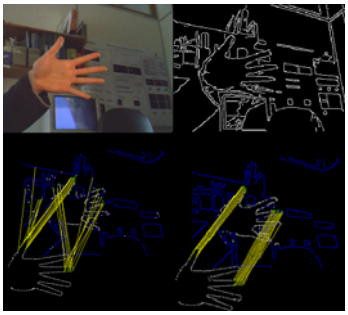
- Significant clutter
  - Unreliable scale factor
  - Incorrect correspondences
- Solution
  - Calculate shape contexts at different scales and match at different scales
  - Computationally expensive

Multiple Edge Orientations

- Edge pixels are divided into 8 groups based on orientation
- Shape contexts are calculated separately for each group
- Total matching score is obtained by adding individual χ^2 scores

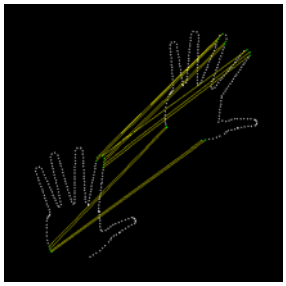


Single vs. Multiple Orientation

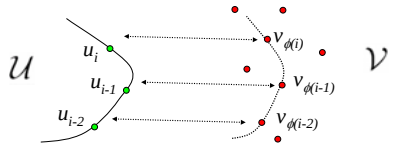


No Figural Continuity

- No continuity constraint
- Adjacent points in one shape are matched to distant points in the other



Imposing Figural Continuity



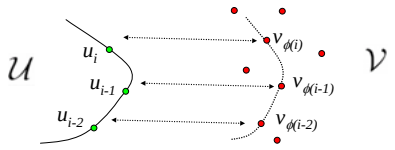
- u_i and u_{i-1} are neighboring points on the model shape $\mathcal{U}$
- ϕ is the correspondence between two shape points
- Corresponding points $v_{\phi(i)}$ and $v_{\phi(i-1)}$ need to be neighboring points on target shape $\mathcal{V}$

Imposing Figural Continuity

$$C_\phi(\mathcal{U},\mathcal{V}) = C_{sc}(\mathcal{U},\mathcal{V}) + \lambda C_{cont}(\mathcal{U},\mathcal{V}) + \mu C_{curv}(\mathcal{U},\mathcal{V})$$

$$C_{sc}(\mathcal{U},\mathcal{V}) = \sum_{i=1}^{n} C_{sc}(\mathbf{u}_i, \mathbf{v}_{\phi(i)}) \qquad C_{curv}(\mathcal{U},\mathcal{V}) = \sum_{i=2}^{n-1} \|\kappa(\mathbf{u}_i) - \kappa(\mathbf{v}_{\phi(i)})\|$$

$$C_{cont}(\mathcal{U},\mathcal{V}) = \sum_{i=2}^{n} \|\mathbf{v}_{\phi(i)} - \mathbf{v}_{\phi(i-1)}\| \qquad \kappa(\mathbf{u}_i) = \|\mathbf{u}_{i-1} - 2\mathbf{u}_i + \mathbf{u}_{i+1}\|$$

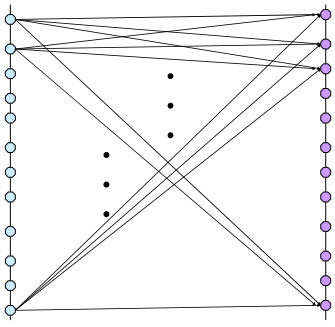


Imposing Figural Continuity

$$C_\phi(\mathcal{U},\mathcal{V}) = C_{sc}(\mathcal{U},\mathcal{V}) + \lambda C_{cont}(\mathcal{U},\mathcal{V}) + \mu C_{curv}(\mathcal{U},\mathcal{V})$$

- Minimize the cost function for ϕ
- Ordering of the model shape is known
- Use Viterbi Algorithm

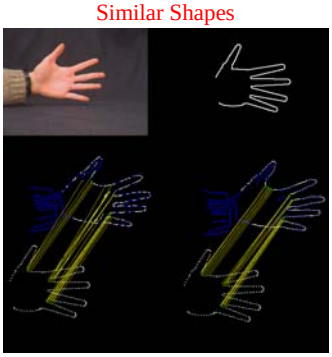
Viterbi Algorithm



- Initialization
- Propagation
  - Compute and sum up cost
  - Store a pointer
- Termination
- Optimal Path Backtracking

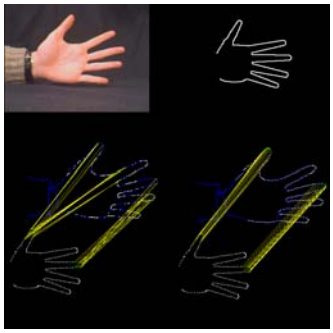
With Figural Continuity

Similar Shapes



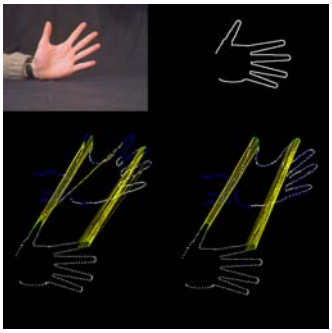
With Figural Continuity

Different Scale



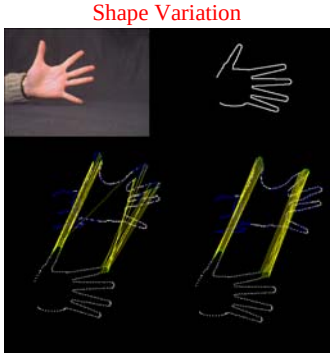
With Figural Continuity

Small Rotation



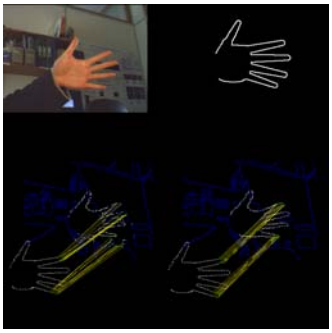
With Figural Continuity

Shape Variation



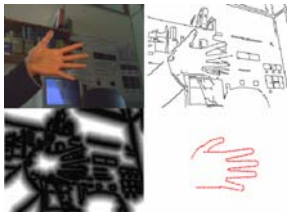
With Figural Continuity

Clutter

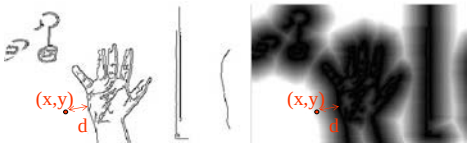


Chamfer Matching

- Matching technique cost is integral along contour
- Distance transform of the Canny edge map

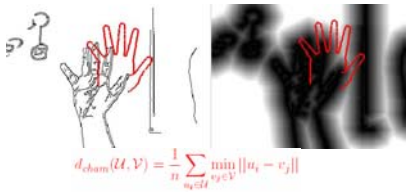


Distance Transform



- Distance image gives the distance to the nearest edge at every pixel in the image
- Calculated only once for each frame

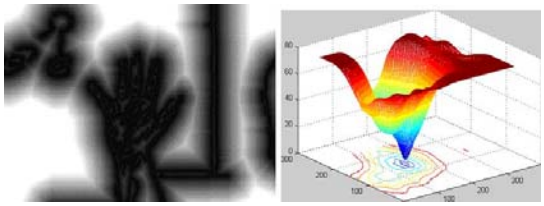
Chamfer Matching



$$d_{cham}(\mathcal{U},\mathcal{V}) = \frac{1}{n}\sum_{u_i \in \mathcal{U}} \min_{v_j \in \mathcal{V}} \|u_i - v_j\|$$

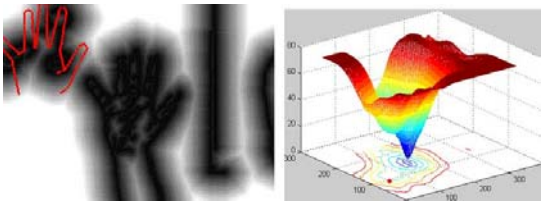
- Chamfer score is average nearest distance from template points to image points
- Nearest distances are readily obtained from the distance image
- Computationally inexpensive

Chamfer Matching

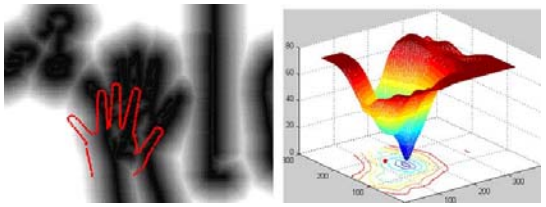


- Distance image provides a smooth cost function
- Efficient searching techniques can be used to find correct template

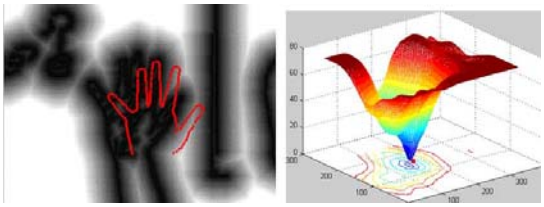
Chamfer Matching



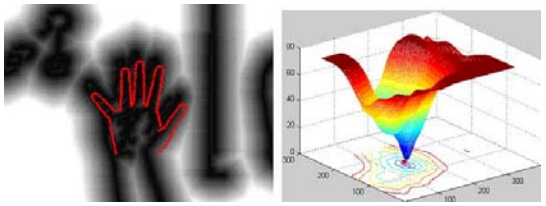
Chamfer Matching



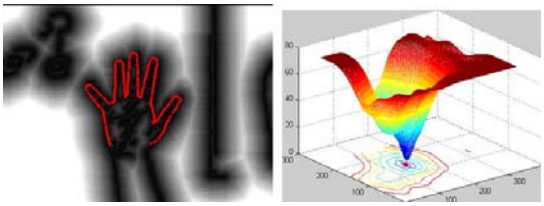
Chamfer Matching



Chamfer Matching

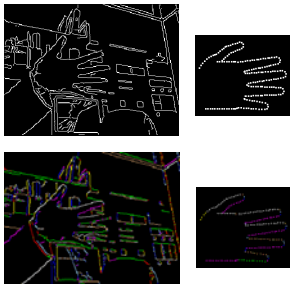


Chamfer Matching



Multiple Edge Orientations

- Similar to Gavrila, edge pixels are divided into 8 groups based on orientation
- Distance transforms are calculated separately for each group
- Total matching score is obtained by adding individual chamfer scores



Applications: Hand Detection

- Initializing a hand model for tracking
  - Locate the hand in the image
  - Adapt model parameters
  - No skin color information used
  - Hand is open and roughly front-parallel

Results: Hand Detection

Original Shape Context | Shape Context with Continuity Constraint | Chamfer Matching

Results: Hand Detection

Original Shape Context | Shape Context with Continuity Constraint | Chamfer Matching

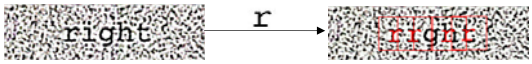
Applications: CAPTCHA

- Completely Automated Public Turing test to tell Computers and Humans Apart [Blum *et al.*, 02]
- Used in e-mail sign up for Yahoo accounts
- Word recognition with shape variation and added noise

Examples:

EZ-Gimpy results

Chamfer cost for each letter template



Word matching cost: average chamfer cost
+ variance of distances

Top 3 matches (dictionary 561 words)

right 25.34 fight 27.88 night 28.42

89.5% correct matches using 1 template per letter
93.2% correct matches using 2 templates per letter

Shape context 82.7% [Mori & Malik, 03]

Discussion

- The original shape context matching
  - Not invariant in clutter
  - Iterative matching is used in the original shape context paper
  - Correct point correspondence in the initial matching is quite small in substantial clutter
  - Iterative matching will not improve the performance

Discussion

- Shape Context with Continuity Constraint
  - Includes contour continuity & curvature
  - Robust to substantial amount of clutter
  - Much better correspondences and model alignment just from initial matching
  - No need for iteration
  - More robust to small variations in scale, rotation and shape.

Discussion

- Chamfer Matching
  - Variant to scale and rotation
  - More sensitive to small shape changes than shape context
  - Need large number of template shapes

  But
  - Robust to clutter
  - Computationally cheap compared to shape context

Conclusion

Use shape context when
- There is not much clutter
- There are unknown shape variations from the templates (e.g. two different types of fish)
- Speed is not the priority

Conclusion

Chamfer matching is better when
- There is substantial clutter
- All expected shape variations are well-represented by the shape templates
- Robustness and speed are more important

References

- A. Thayananthan, B. Stenger, P. H. S. Torr, and R. Cipolla. Shape Context and Chamfer Matching in Cluttered Scenes. CVPR 2003.
- D. Gavrila. Pedestrian Detection from a Moving Vehicle. ECCV 2000
- S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. PAMI 2002.

The original version of this presentation file is from A. Thayananthan, B. Stenger, P. H. S. Torr, and R. Cipolla

Thank You!