CS664 Computer Vision

8. Matching Binary Images

Dan Huttenlocher

Cornell University
Faculty of Computing and Information Science

Comparing Binary Feature Maps

- Binary "image" specifying feature locations
  - In x,y or x,y,scale
- Variations will cause maps not to agree precisely when images aligned
- Measures based on proximity rather than exact superposition

Binary Correlation

- Recall cross correlation

$$C[i,j] = \sum_{u=-k}^{k} \sum_{v=-k}^{k} H[u,v]F[i+u, j+v]$$

- For binary images counting number of coincident 1-valued pixels
  - Number of on pixels in AND at offset (i,j)
- SSD (sum squared difference) – XOR

$$S[i,j] = \sum_{u=-k}^{k} \sum_{v=-k}^{k} \left(H[u,v] - F[i+u, j+v]\right)^2$$

- Suffer from measuring exact agreement and not proximity

Hausdorff Distance

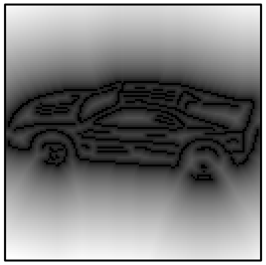
- Classical definition
  - Directed distance (not symmetric)
    - $h(A,B) = \max_{a \in A} \min_{b \in B} \|a-b\|$
  - Distance (symmetry)
    - $H(A,B) = \max(h(A,B), h(B,A))$
- Minimization term simply dist trans of B
  - $h(A,B) = \max_{a \in A} D_B(a)$
  - Maximize over selected values of dist trans
- Classical distance not robust, single "bad match" dominates value

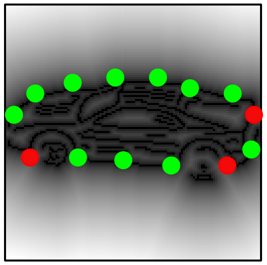
Hausdorff Matching

- Best match
  - Minimum fractional Hausdorff distance over given space of transformations
- Good matches
  - Above some fraction (rank) and/or below some distance
- Each point in (quantized) transformation space defines a distance
  - Search over transformation space
    - Efficient branch-and-bound “pruning” to skip transformations that cannot be good

Hausdorff Matching

- Partial (or fractional) Hausdorff distance to address robustness to outliers
  - Rank rather than maximum
    - $h_k(A,B) = kth_{a \in A} \min_{b \in B} \| a-b \| = kth_{a \in A} D_B(a)$
  - K-th largest value of D_B at locations given by A
  - Often specify as fraction f rather than rank
    - 0.5, median of distances; 0.75, 75th percentile





1,1,2,2,3,3,3,3,4,4,5,12,14,15

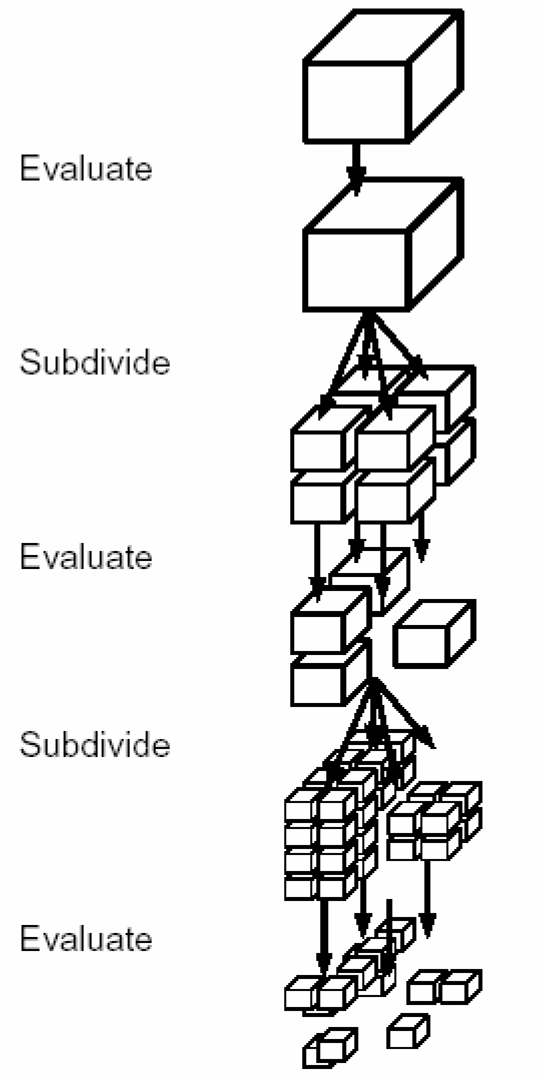
.25 .5 .75 1.0

Fast Hausdorff Search

- Branch and bound hierarchical search of transformation space
- Consider 2D transformation space of translation in x and y
  - (Fractional) Hausdorff distance cannot change faster than linearly with translation
    - Similar constraints for other transformations
  - Quad-tree decomposition, compute distance for transform at center of each cell
    - If larger than cell half-width, rule out cell
    - Otherwise subdivide cell and consider children

Branch and Bound Illustration

- Guaranteed (or admissible) search heuristic
  - Bound on how good answer could be in unexplored region
    - Cannot miss an answer
  - In worst case won't rule anything out
- In practice rule out vast majority of transformations
  - Can use even simpler tests than computing distance at cell center



Chamfer Distance

- Sum of closest point distances

  $$Ch(A,B) = \Sigma_{a \in A} \min_{b \in B} \| a-b \|$$

- Generally use asymmetric measure but can be symmatrized

  $$CH(A,B) = Ch(A,B) + Ch(B,A)$$

- As for Hausdorff distance minimization term is simply a distance transform
- While intuitively may appear more robust to outliers than max, still quite sensitive
  - Trimming can be useful in practice

Dilation

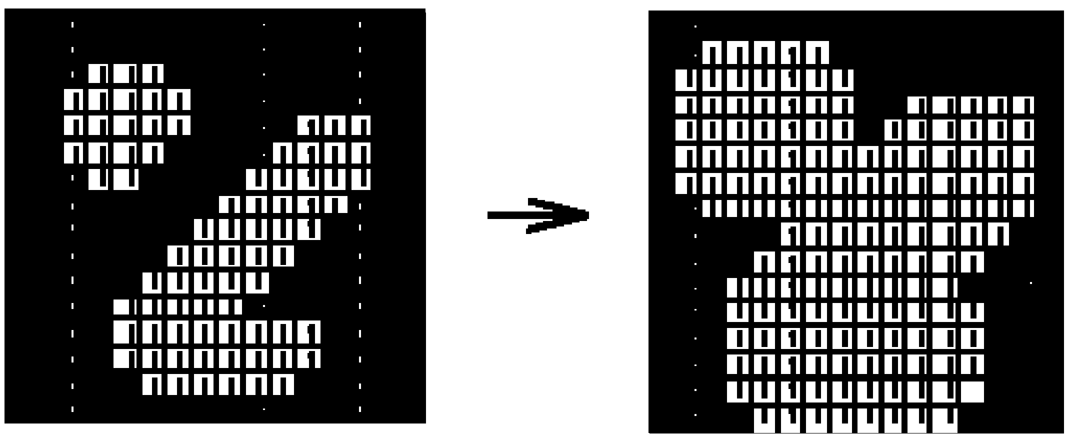
- The Minkowski sum of two point sets A,B is result of adding every point of A to every point of B
  - Note for finite sets, cardinality of result is product of set cardinalities

$$F \oplus H = \{f + g \mid f \in F, g \in G\}$$

- For binary images this is called dilation
  - As with correlation and convolution think of asymmetrically as function and kernel or mask
  - Replace each on pixel of F by mask H
    - Generally center pixel of H is on

Dilation

- Dilation by a disk of radius d corresponds to level sets of L^2 distance transform for distances $\leq d$
  - Analogously for square of radius d and L-infinity norm
  - 3x3 square example (radius 1)



Dilation and Correlation

- Correlation of F with G dilated by a disk of radius d
  - Counts number of on pixels in F at each [i,j] that are within distance d of some on pixel in G
  - Normalize the count by dividing by total number of on pixels in F
- Corresponds to the Hausdorff fraction
  - Fraction within distance d rather than distance for fraction f

$$h_f(A,B) = \text{fth}_{a \in A} \min_{b \in B} \| a-b \|$$

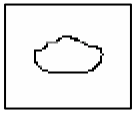
where fth quantile

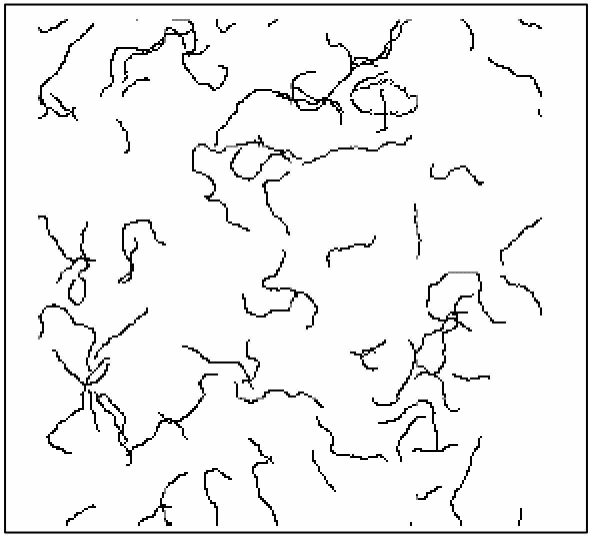
DT Based Matching Measures

- Fractional Hausdorff distance
  - Kth largest value selected from DT
- Chamfer
  - Sum of values selected from DT
    - Suffers from same robustness problems as classical Hausdorff distance
    - Max intuitively worse but sum also bad
  - Robust variants
    - Trimmed: sum the K smallest distances (same as Hausdorff but sum rather than largest of K)
    - Truncated: truncate individual distances before summing

Comparing DT Based Measures

- Monte Carlo experiments with known object location and synthetic clutter
  - Matching edge locations
- Varying percent clutter
  - Probability of edge pixel 2.5-15%
- Varying occlusion
  - Single missing interval, 10-25% of boundary
- Search over location, scale, orientation

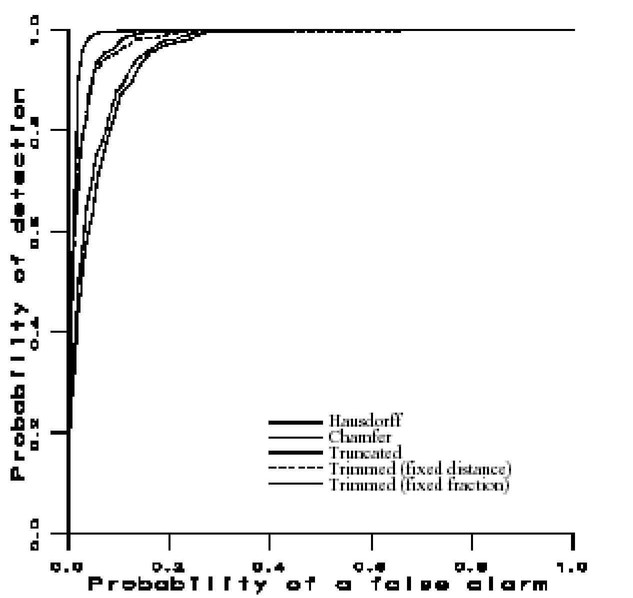


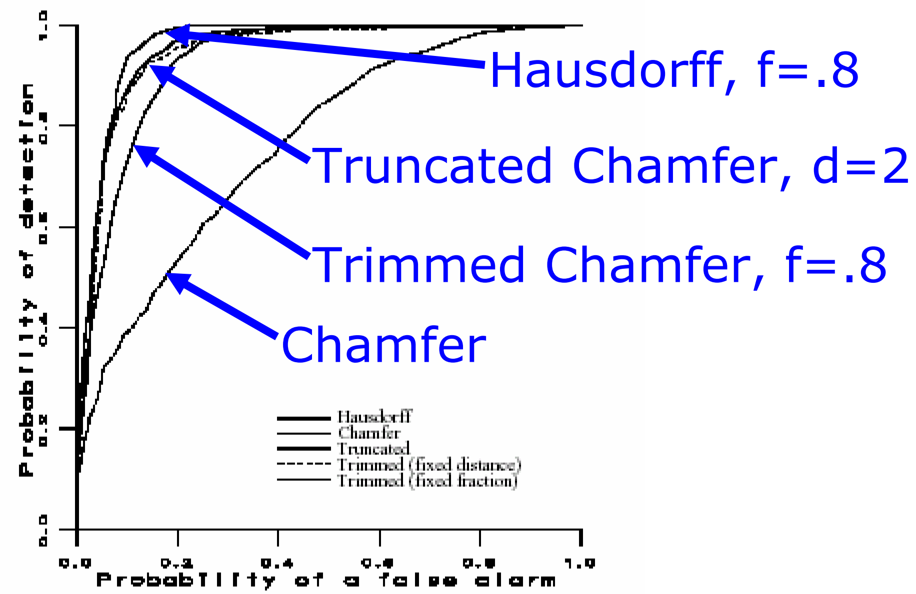


5% Clutter Image

ROC Curves

- Probability of false alarm vs. detection
  - 10% and 15% occlusion with 5% clutter
  - Chamfer is lowest, Hausdorff (f=.8) is highest
  - Chamfer truncated distance better than trimmed



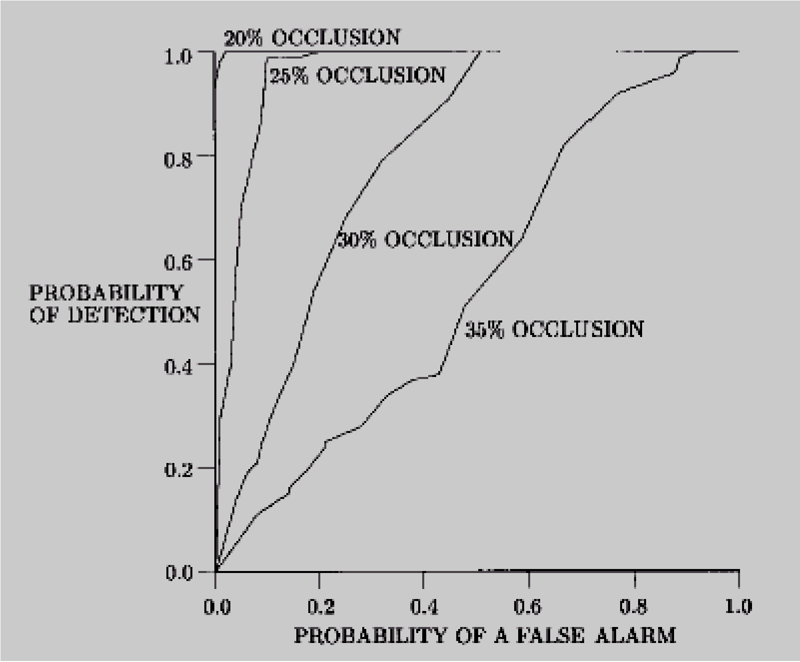


Edge Orientation Information

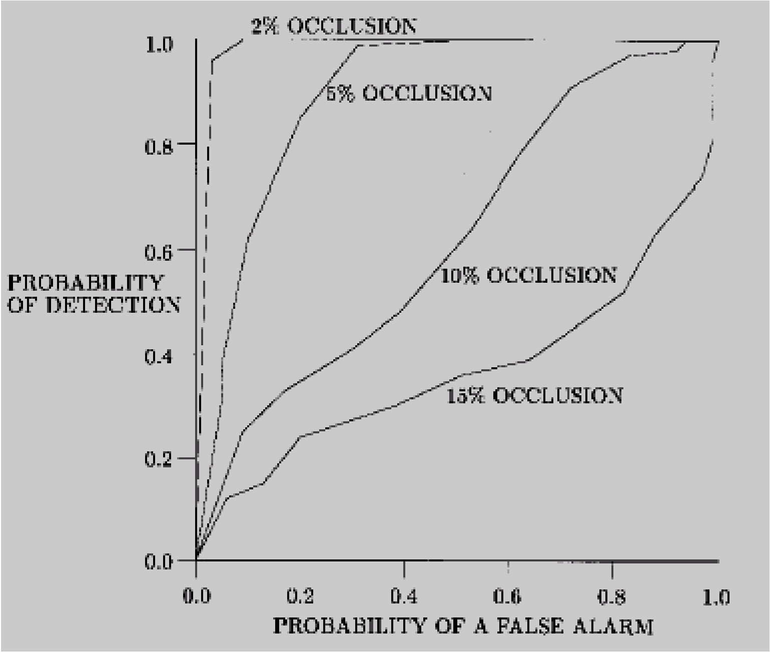
- Match edge orientation as well as location
  - Edge normals or gradient direction
- Increases detection performance and speeds up matching
  - Better able to discriminate object from clutter
  - Better able to eliminate cells in branch and bound search
- Distance in 3D feature space $[p_x, p_y, \alpha p_o]$
  - α weights orientation versus location
  - $\mathrm{kth}_{a \in A} \min_{b \in B} \| a-b \| = \mathrm{kth}_{a \in A} D_B(a)$

ROC's for Oriented Edge Pixels

- Vast improvement for moderate clutter
  - Images with 5% randomly generated contours
  - Good for 20-25% occlusion rather than 2-5%



Oriented Edges



Location Only

Summary of DT Based Matching

- Fast compared to explicitly considering pairs of model and data features
  - Hierarchical search over transformation space
- Important to use robust distance
  - Straight Chamfer very sensitive to outliers
    - Truncated DT can be computed fast
- No reason to use approximate DT
  - Fast exact method for L_2^2 or truncated L_2^2
- For edge features use orientation too
  - Comparing normals or using multiple edge maps