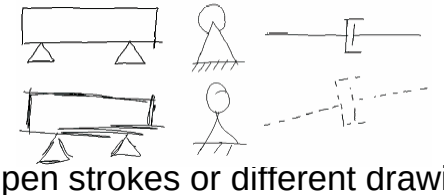
An Image-Based Trainable Symbol Recognizer for Sketch-Based Interfaces

Levent Burak Kara, Thomas F. Stahovich (2004)

Nazim Ashraf

Symbol Recognition

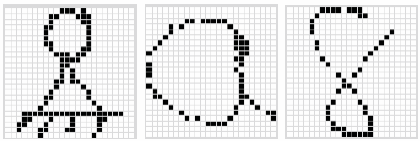
- Image-based recognition approach
- Advantages:
  - Segmentation is eliminated
  - Can recognize ‘sketchy’ symbols:



  - Multiple pen strokes or different drawing orders does not matter

Preprocessing

- Input Symbols are quantized into 48x48 bitmap images:



- Quantization
  - Reduces amount of data
  - Preserves the main characteristics

Template Matching

- Task:
  - Match unknown template against templates in the database
- Similarity Measures
  - Hausdorff Distance
  - Modified Hausdorff Distance
  - Tanimoto Coefficient
  - Yule Coefficient

Hausdorff Distance

- Hausdorff distance: *maximum distance of a set to the nearest point in the other set*
- Given sets A and B,
  - $H(A,B) = max(h(A,B), h(B,A))$
    where $h(A,B) = \max_{a \in A}(\min_{b \in B} \|a - b\|)$

  If *h(A,B)* = *d*, then every point in set A is at least *d* distance away from some point in set B

Hausdorff Distance

- Very sensitive to outliers
- Partial Hausdorff Distance:
  - Rank points in A according to their distances to points in B in descending order
  - Select the k^{th} ranked point

$$h^k(A,B) = \underset{a \in A}{k^{th}} \min_{b \in B} \|a - b\|$$

Modified Hausdorff Distance

- Replaces the *max* operator by average

$$h_{mod}(A,B) = \frac{1}{N_a} \sum_{a \in A} \min_{b \in B} \|a - b\|$$

  - N_a: *number of points*

$$MHD(A,B) = max(h_{mod}(A,B), h_{mod}(B,A))$$

Tanimoto Similarity Coefficient

- Given binary images, A and B:

$$T_{sc}(A,B) = \alpha \cdot T(A,B) + (1-\alpha) \cdot T^{C}(A,B)$$

  where

  - T(A, B): Tanimoto coefficient
  - T^C(A, B): Tanimoto coefficient complement

Tanimoto Similarity Coefficient

- Tanimoto coefficient: Measure of matching black pixels

$$T(A,B) = \frac{n_{ab}}{n_a + n_b - n_{ab}}$$

where

n_a: Black pixels in A

n_b: Black pixels in B

n_{ab}: Overlapping black pixels

Tanimoto Similarity Coefficient

- Tanimoto coefficient complement: Measure of matching white pixels

$$T_{sc}(A,B) = \alpha \cdot T(A,B) + (1-\alpha) \cdot T^C(A,B)$$

- α is the weighting factor. Used range: [0.5, 0.75]
- Black pixels more important

Yule Coefficient

- Defined as:

$$Y(A,B) = \frac{n_{ab} \cdot n_{00} - (n_a - n_{ab}) \cdot (n_b - n_{ab})}{n_{ab} \cdot n_{00} + (n_a - n_{ab}) \cdot (n_b - n_{ab})}$$

Combining Classifiers

- Outputs different:
  - First two classifiers measure *dissimilarity*
  - Last two measure *similarity*
- Output range different:
  - Normalize the values to the range 0 to 1

Combination Rule

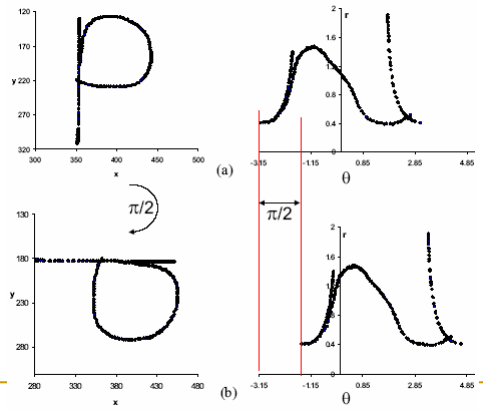
- Combination rule:
  - Compute normalized distance as sum of normalized distances
- Unknown pattern assigned to class having minimum combined normalized distance

Handling Rotations

- Template Matching sensitive to rotations
- For rotation invariance:
  - First rotate the pattern into the same orientation
- Method:
  - Incrementally rotate until best alignment
  - Very expensive

Polar Coordinates

- Rotations in Cartesian coordinates ~ Translations in Polar coordinates



Polar Transform

- Polar coordinates

$$r = \sqrt{(x - x_o)^2 + (y - y_o)^2}$$

$$\theta = tan^{-1}(\frac{y - y_o}{x - x_o})$$

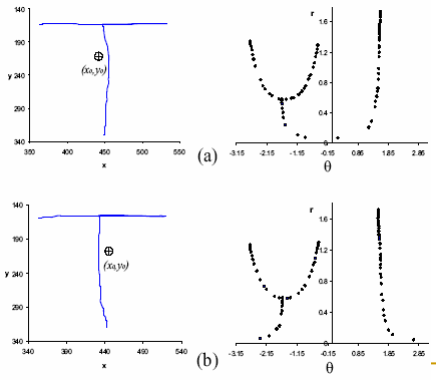
- Symbol drawn is transformed into polar coordinates

Calculating Rotation

- Slide-and-compare:
  - Incrementally displace along θ axis
  - Determine matching
  - Best match indicated amount of rotation needed
- Polar images are also 2D binary patterns

Issues with Polar Transform

- For small values of r, the θ coordinate is sensitive to the centroid



Solution

- Modify the MHD:

$$h_{mod\_weighted}(A,B) = \frac{1}{N_a}\sum_{a\in A} w(a_r)\cdot \min_{b\in B}\|a-b\|$$

$$w(r) = r^{0.10}$$

w(r) ~ 1 for large *r*

w(r) ~ 0 for small *r*

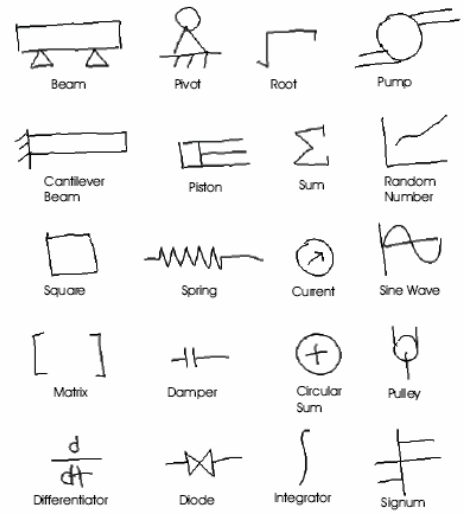
- Assign small values to pixels near centroid
- MHD now governed by pixels farther from the centroid

Recognition Process

- Polar Transform as Pre-Recognizer
  - Produces false positives
  - Rarely results in false negative matches
- Keep top 10%; discard else

User Studies

- Symbols used



User Studies

- Users asked to draw 20 symbols
- 5 users
  - Provided 3 sets of these symbols
- 4 different types of tests carried out
  - Test 1: Single definition, user-dependent
  - Test 2: Two definition sets, user-dependent
  - Test 3: Twelve definition sets, user-independent
  - Test 4: Fourteen definition sets, partially user-independent

Results

	Top 1 (%)	Top 2 (%)	Recog. Time (ms)
Test1	90.7	96.3	332
Test2	95.7	98.3	354
Test3	94.7	97.3	623
Test4	98.0	99.0	674

Column 1: Class ranked highest is indeed the correct class

Column 2: Correct class is either the highest or second highest ranked class

Questions