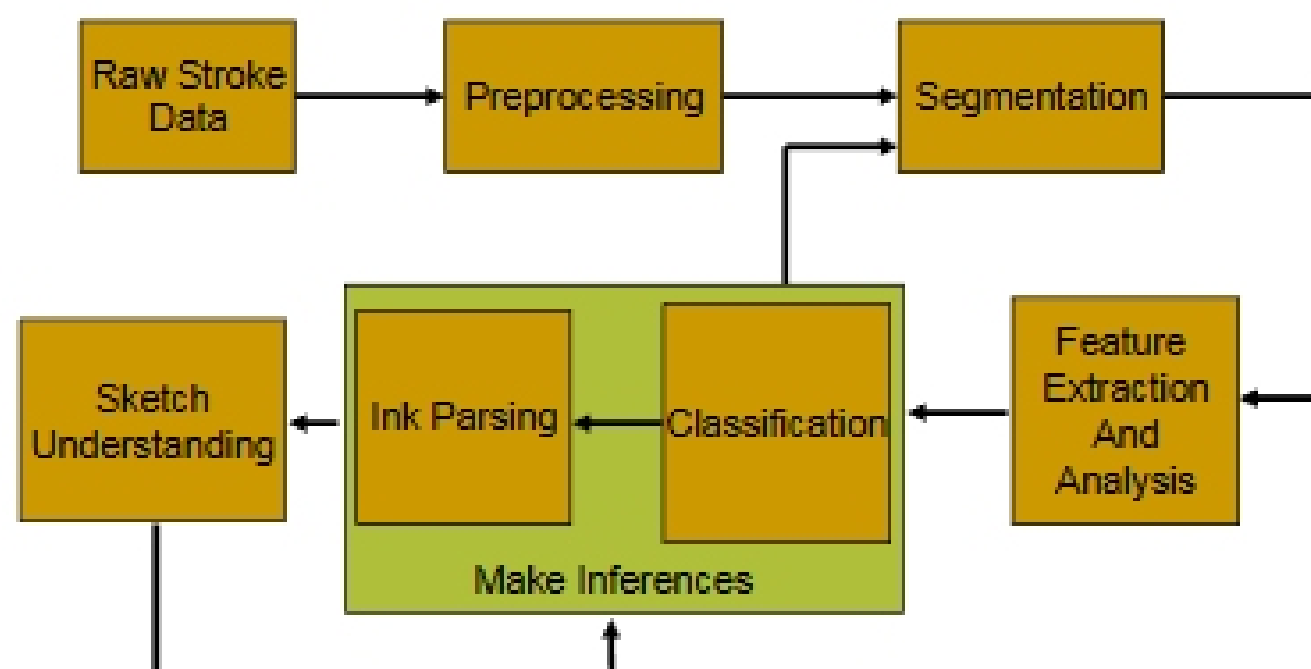


Features Extraction for Sketch-Based Recognition

Lecture #8: Feature Extraction
Joseph J. LaViola Jr.
Fall 2007

Recall Pen-Based Interface Dataflow



Feature Extraction and Analysis

- What came first, the feature or the machine learning algorithm?
- Want to distinguish sketch components from one another
- Good features are critical
- Extract important information
 - geometrical, statistical, contextual
- Examples include
 - arc length, histograms, cusps, aspect ratio
 - self-intersections, stroke area, etc...

Finding Features

- Challenging problem
 - need fast algorithms for gathering information
 - features must be good discriminators
- Often trial and error
- Can be domain specific

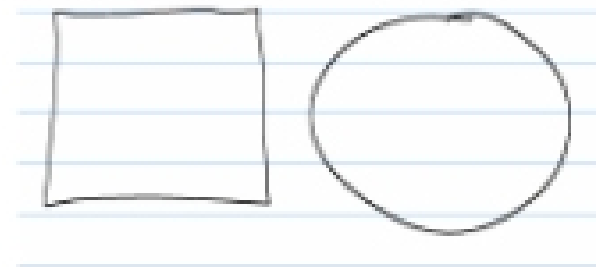
Geometric Features (1)

■ Number of strokes

- if you know how many strokes a symbol has, you can break up your recognizer into pieces (i.e., recognizer for 1 stroke symbols, recognizer for 2 stroke symbols ...)

■ Cusps

- smooth vs. jagged strokes
- distance between cusps
 - useful for when cusps are close together/far apart



Geometric Features (2)

■ Aspect ratio (width / height)

- tall vs. flat



■ Self Intersections

- loops vs. no loops
- strokes with write over
- distance between self intersections also useful
- use line segment intersection algorithm

