

Handwritten Math Symbol Recognition

Ajinkya Dhaigude
Supriya Godge

Classification

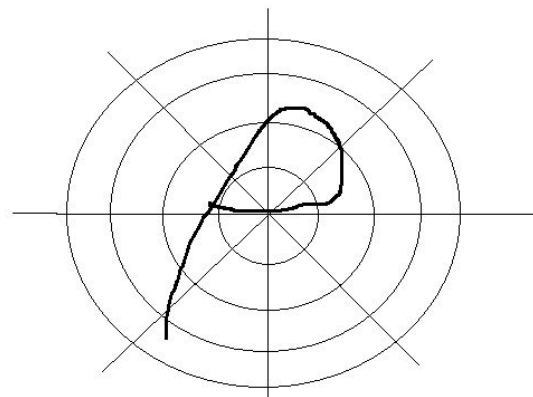
- Preprocessing
 - Scale strokes to fit within a 100x100 square while maintaining the aspect ratio
 - Generate offline image then dilate and apply Gaussian blur
 - Align images in the center of the 100x100 square



Feature extraction

- Online

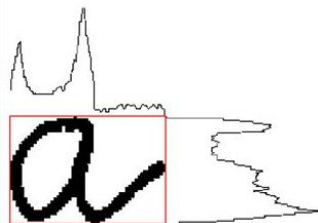
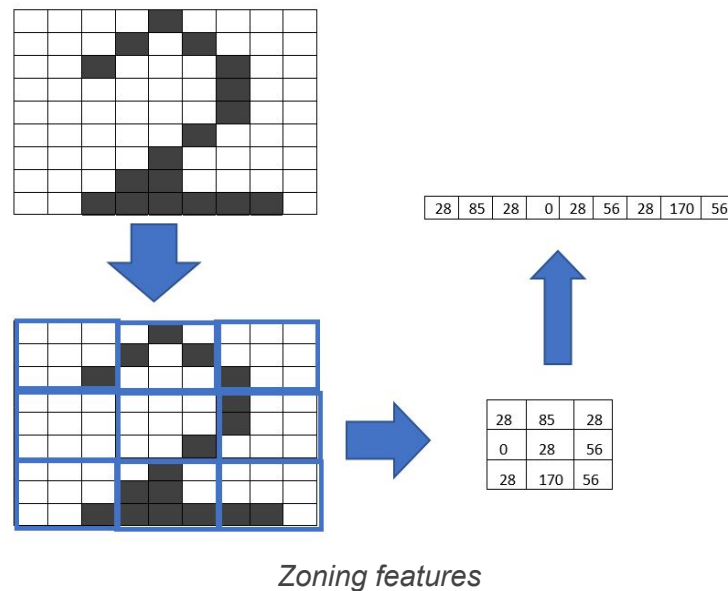
- The distance from first point of first stroke to last point of the last stroke
 - 2 features
- Added polar histogram as a new feature
 - 6 angles and 5 radii
 - 30 features



Polar histogram

Feature extraction

- Offline features
 - Zoning features
 - 81 features
 - X and Y axis projections
 - 100 features each
 - Diagonal projections
 - 398 features
- A total of 712 features



X and Y axis projections

Classifier

- Random forest
 - Number of trees: 10
 - Depth of the tree: No limit specified
- Accuracy:
 - Without junk: 84.11%
 - With junk: 77.21%



Segmentation

- Preprocessing
 - Minimize jitter by averaging neighboring coordinates
 - Scale expression to a fixed height of 100
 - Width is determined by aspect ratio
 - Change ‘,’ to *COMMA*

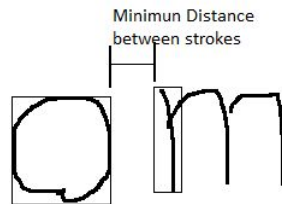
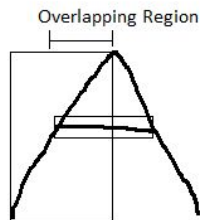


Feature extraction

- Pair of strokes are *merged* or *split*
 - Consider all pairs of adjacent strokes
 - Consider nearest neighbor if stroke is not adjacent

- Geometric features

- Provide a narrow view of only the pair
- 13 different features used
- Normalized between 0 and 1



- Shape context features

- Provide a localized view of neighboring strokes

- A total of 113 features

Hu, Lei, and Richard Zanibbi. "Line-of-sight stroke graphs and parzen shape context features for handwritten math formula representation and symbol segmentation." *Proc. ICFHR* (2016)

Classifier

- Random forest
 - Number of trees: 50
 - Depth of the trees: 50
- Result:
 - Baseline: Each stroke is a symbol

	Sophisticated					BaseLine				
	Object			Classes		Object			Classes	
	Recall	Precision	F-value	F-Value	Accuracy/Det	Recall	Precision	F-value	F-value	Accuracy/Det
After implementation on time series	89.57	84.44	86.93	72.09	82.93	63.62	44.36	52.28	44.65	85.41
After Normalization of geometric feature	91.01	87.6	89.27	79.29	87.12	NA	NA	NA	NA	NA
After merging the comma and preprocess	91.3	87.73	89.48	79.68	87.27	NA	NA	NA	NA	NA
Another training and testing file	92.21	88.59	90.37	80.14	88.69	NA	NA	NA	NA	NA
						NA	NA	NA	NA	NA
						* Baseline was only tested once.				

Errors and Challenges

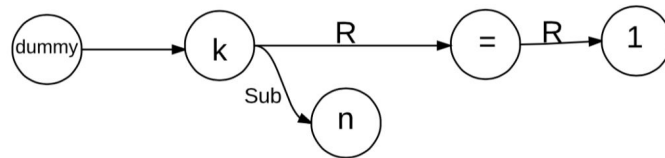
- Most common errors
 - X and $/times$, x and $)$, 2 and z
 - Some should be avoidable with a parsing context
- Challenges
 - Need to look at large data very closely
 - Need to plan out the system
 - Communication
 - Training time
- Improvements
 - Normalization of online data
 - Binary classifier for most common errors

- 2x



Parsing

- Create a Line of Sight graph of the symbols and compute the minimum spanning tree
- Improve the segmentation by using the strokes pairs generated by the Line of Sight instead of time series
- Geometric and context features



Apply Edmonds' algorithm to obtain MST.

Output: Remove dummy node for Symbol Layout Tree

Thank You

Questions ?