Applying Hierarchical Contextual Parsing to Recognize Isolated Typeset Math Formulas

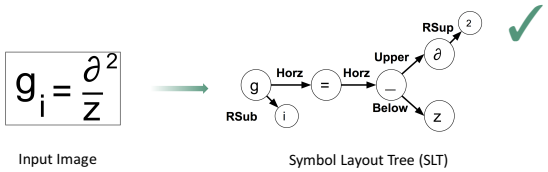
Mahshad Mahdavi, Michael Condon and Richard Zanibbi

Rochester Institute of Technology

Overview

- Recognition of scanned typeset mathematical equations can be done by extracting maximum spanning trees from line of sight graphs weighted using geometric and visual density features [5].
- Interestingly, segmentation and parsing are done without using symbol classification information, and symbol classification is done independently of expression structure recognition.
- Only two types of features are used, spatial and visual.
- Our model parsed 95.97% of expressions correctly when given symbols and 93.95% when requiring symbol segmentation from connected components.
- Overall HCP reached 90.83% expression recognition rate from connected components.

Goal:



Input Image — Symbol Layout Tree (SLT)

Features

Two types of features are used, *spatial* and *visual*. The spatial features including horizontal distance, size difference and vertical offset, minimum point distance, over- lapping area, etc. Visual density features are captured using two kinds of histograms, shape context features and 2D Grid Histograms.

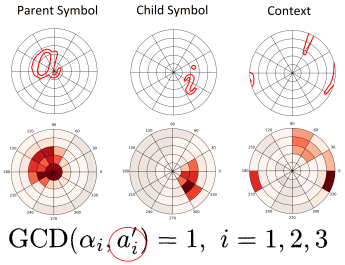


Fig.1 : **Shape Context Features (SCF) histogram** for capturing spatial density features of the relation between *a (parent)* and *i (child)* symbol.

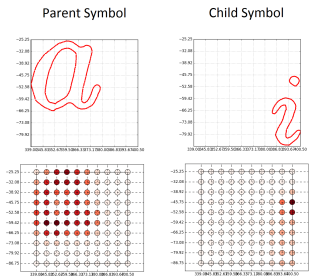
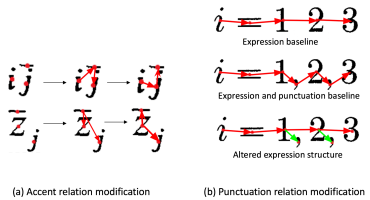


Fig.2: **2D Grid histogram** spatial densities for a relation between parent symbol *a* and child symbol *i*.

Methods

1. Data and Representation:

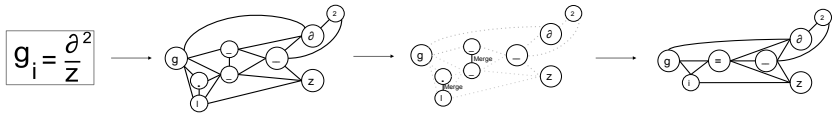
- ***Defining Punctuation Relation:*** Having the baseline punctuation separated from the main baseline
- ***Restructuring Accent Relation:*** giving accent an above relation to the symbol under it



(a) Accent relation modification (b) Punctuation relation modification

2. Segmentation with Line-Of-Sight (LOS) Graphs

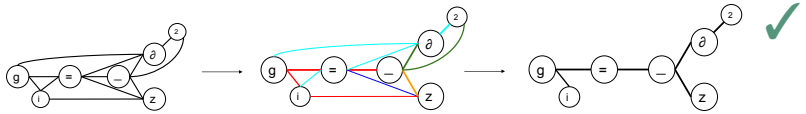
- A binary classifier is used to identify which directed LOS [1] edges should be merged. Each component pair defined by 'merge' edges is taken to be a symbol candidate.



LOS on strokes — Segmentation — LOS on symbols

3. MST-based Parsing

- In parsing, classifier is used to give each edge probability scores for each of the possible spatial relations.
- With the weighted expression graph, Edmonds' algorithm [2,3] is used to select a maximum spanning tree
- The maximum spanning tree is used as the final expression structure.



LOS on symbols — Weighted graph — MST with Edmond's

Results

- The Hierarchical Contextual Parsing (HCP) has been tested on the InftyMCCDB- 2 dataset [4] .

Table 2: **Expression Rates for Full Recognition of Typeset Math Expressions using geometric and density histograms. v1: Original Expressions, v2: Punctuation relation has been defined and Accent restructuring has been done.**

Expression Rate(%) with Segmented Symbols

Dataset	Expr Rate
Original Expressions (v1)	88.45
Modified Expressions (v2)	**90.83**

Table 1: **Detection(Det) and Detection with Classification(Det+Class) results for parsing Infty expressions using geometric and density histogram features. v1: Original Expressions, v2: Punctuation relation has been defined and Accent restructuring has been done.**

Parsing Rate(%) with Segmented Symbols

Dataset		Relationships	Expressions
Original Expressions (v1)	Det	97.82	92.15
	Det+class	97.6	91.68
Modified Expressions(v2)	Det	98.23	94.41
	Det+class	97.96	**93.95**

Parsing Rate(%) with Given Symbols

Dataset		Relationships	Expressions
Original Expressions (v1)	Det	98.81	94.35
	Det+class	98.55	93.66
Modified Expressions(v2)	Det	99.3	96.54
	Det+class	99.07	**95.97**

Conclusion

Our results provide strong support for typeset math expression recognition being done effectively by maximum spanning tree extraction with simple classifiers using spatial and visual density features.
With the transformed spatial layouts and HCP, 95.97% of expressions were parsed correctly when given symbols and 93.95% correctly parsed when requiring symbol segmentation from connected components. Overall HCP reached 90.83% expression recognition rate from connected components.

Reference

[1] Hu, Lei, and Richard Zanibbi. "Line-of-sight stroke graphs and parzen shape context features for handwritten math formula representation and symbol segmentation." *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR).* IEEE, 2016.

[2] Hu, Lei, and Richard Zanibbi. "MST-based visual parsing of online handwritten mathematical expressions." *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR).* IEEE, 2016.

[3] Edmonds, Jack. "Optimum branchings." *Journal of Research of the National Bureau of Standards B,* 71(4): 233-240, 1967.

[4] Masakazu Suzuki, Seiichi Uchida, and Akihiro Nomura. A ground-truthed mathematical character and symbol image database. In Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on, pages 675–679. IEEE, 2005.

[5] Condon, Michael Patrick Erickson. "Applying Hierarchical Contextual Parsing with Visual Density and Geometric Features to Typeset Formula Recognition." 2017.

Acknowledgment

This material is based upon work supported by Alfred P. Sloan Foundation

Alfred P. Sloan FOUNDATION