Expresso: Typesetting Handwritten Mathematical Expressions on the Post-PC Tablet Computer

[Project Proposal]

Josef Lange
University of Puget Sound
1500 N. Warner St.
Tacoma, WA 98416
jlange@pugetsound.edu

Daniel Guilak
University of Puget Sound
3396 Wheelock Student Center
Tacoma, WA 98416-3396
dguilak@pugetsound.edu

ABSTRACT

This paper sets forth to describe our project fulfilling the capstone requirement of the Bachelor's Degree in Computer Science at the University of Puget Sound. Our project is to research, design, and implement a tablet-based software solution for the capture of handwritten mathematical expressions and compilation of its mathematical representation in an appropriate representative language (such as LaTeXor MathML).

Subjects approached in our project include image processing, artificial intelligence, application development for the Apple iOS platform (particularly in the tablet form factor), server development, as well as the adherence to a highly disciplined development cycle. We will attempt to most accurately and most briefly explain any uncommon concepts described hereafter.

1. INTRODUCTION

The utility of the touch-based computer interaction has only been recently fully realized, with the nascence of the "Post-PC" tablet in its migration away from the standard desktop usage paradigm commonly associated with computing. The modern tablet focuses on media consumption and so-called "basic" computing. In today's reality, most "Post-PC" tablets are incredibly powerful both in hardware and in software, supporting complex and challenging computations including the decoding of video, image processing, interpretation of multiple inputs, managing several network connections, and displaying high-resolution two- and three-dimensional graphics.

Our project aims to utilize this now-popular device class for the capture and conversion of handwritten mathematics to relevant language for typesetting systems. Predominantly, this functionality has been researched and implemented on desktop systems, due in part to the prior lack of power in tablet devices. With modern tablets and their ubiquitousness with the Internet the pieces line up for a cloud-based recognition system.

2. EXISTING SOLUTIONS

2.1 General Handwriting Recognition

Handwriting recognition is a complex problem, only a segment of which my project wishes to solve. Existing solutions to handwriting recognition exist, and some perform at a useable level. Frequently, these systems require a period of "training" before any recognition can be achieved. Others use artificial intelligence to conclude what a sequence of handwriting is meant to be. For full-on text recognition, this is useful and almost required for any success. Commercial products exist, including but not limited to software built into Microsoft's Windows and Apple's Mac OS X Operating Systems. Additionally, many third-party companies have produced software for a similar purpose.

2.2 Mathematical Expression Capture

General handwriting recognition solutions are frequently overpowered for the niche of mathematical expressions, in which there are fewer possible logical constructs, most of which follow a fairly standard form. Because of this, the subset of things needing recognition, and possibilities for what they could be recognized as, is significantly smaller. Academic projects, including those by Matsakis[2], Smithies et. al.[3], and Levin[1] exhibit valid solutions for the desktop model of usage, though are not implemented in a way that is ultimately accessible for the common user.

3. PROPOSED SOLUTION

This project will produce a feature-complete, usable piece of software for the Apple iOS platform. This software will have a singular function: to capture handwritten mathematical expressions via the digitizing screen, and communicate with a server which will compile the interpreted characters into the appropriate mathematical representational language (namely LaTeXor MathML).

4. MOTIVATION & IMPORTANCE

Throughout their history, touch devices have attempted to attain status as a "universal tool" – something able to function seamlessly in their users' lives. Humans communicate naturally with handwriting, and devices from Apple Newtons, to Palm Pilots, to Pocket PCs, and now to the likes of

the Apple iPad, Amazon Kindle Fire, and Microsoft Surface have tried to integrate handwriting recognition with little success since software keyboards are more familiar and efficient to most users. However, we believe that the size and shape of tablet devices are reminiscent of "slates" used in classrooms to scribble out mathematical equations.

Both authors at one point or another have been frustrated with a homework assignments requiring mathematical typesetting in LaTeX. We believe it would be great for a user to jot down an equation on an iPad or other tablet device and have the LaTeX code appear on the document they are editing on their laptop or desktop computer.

This software could easily improve the classroom environment in all levels of schooling. Getting mathematical equations into type is a frustrating task for many teachers in the K-12 levels, and even frustrating for the higher-education professor and student. With this software, student and educators alike can easily convert their glyphic mathematical thoughts into usable typesetting language. This will hopefully open doors to clearer and more specific teaching and learning throughout all levels of education.

5. PROJECT ROADMAP

5.1 Research

First and foremost, additional heavy research is required. Each small piece of information leads to additional sources of information and additional projects that have attempted to achieve (with varying levels of success) the same goal. There is much to be learned from each general field from which my project draws. It will be helpful to carefully analyze previous projects of this one's nature to find commonly-used methodologies when approaching handwriting recognition. Research may also help to steer my project from any abysses of hard work. The best implementation is frequently the simplest one, and as such we aim to avoid overly-complex solutions.

Research conducted will vary from understanding general text recognition, to developing a firm knowledge of mathematical representative languages, to bezier curve approximation, to distributed artificial intelligence implementation. Not all avenues of research will produce implementation, but it is my hope that most researched topics find their way into my design and implementation. The process will include careful reading of prior works, implementing rudimentary forms of algorithms investigated, and feasibility assessments of some of the more complex and optional components.

5.2 Design

A great deal of design must be done on such a big and complex project. The first task is to intelligently segment the project into smaller, possible-to-implement pieces that can be connected to create a whole system. This will require a strong knowledge of design patterns and general software engineering concepts. It is this project's intention to be comprised of several software components that may be useful on their own, and when combined properly, efficiently and elegantly solve the intended problem.

First, we intend to build a rudimentary interface for drawing capture which will use curve approximation and line smoother to facilitate cleaner drawn glyphs, which may also increase their recognizability. The bitmap data gathered by the client will be compressed and sent to a server which will identify individual glyphs. Their relative positions and sizes will be assessed to create a best guess of what the entire mathematical expression is. This information will be stored in a data structure (most likely an expression tree of some sort) and then systematically parsed into LaTeXor MathML markup which will be sent to the laptop or desktop client where the user is writing their document.

Eventually, we would like to have the server parse an expression and send it back to the tablet for rudimentary supervised learning, to have the system benefit from global distributed input from all client software, much like Apple's Siri technology. In addition, this split between client and server allows for the easier development of clients on other tablet platforms in the future, including potentially "thin" tablets – devices with less powerful hardware specifications that primarily rely on the Internet and cloud-based computing for intensive processes.

5.3 Implementation

This project, as mentioned before, will need to be segmented, with each segment well-designed, in order for implementation to be manageable. The client will be implemented on the Apple iOS platform using the Cocoa Touch framework. Cocoa Touch is developed primarily in Objective-C, which is compatible at compilation with C and C++. The intention is to compile and distribute for the iPad device, though the iPhone and iPod touch devices are equally capable, albeit slightly more difficult to use with such a small screen and digitizer. Apple provides a Development Program for institutions of higher education, which will negate any costs of development that Apple normally charges (namely the annual $99 membership) for developers to be able to compile their code to devices for testing and usage.

The server component will be developed in Python with the Flask web framework, and deployed on the Heroku[1] cloud application platform which will allow for easy scalability in the chance that the application gains a significant amount of users.

6. REFERENCES

[1] M. Levin. *CellWriter: Grid-Entry Handwriting Recognition*. PhD thesis, University of Minnesota, Minneapolis, MN, USA, 2007.

[2] N. E. Matsakis. *Recognition of Handwritten Mathematical Expressions*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1999.

[3] S. Smithies, K. Novins, and J. Arvo. A handwriting-based equation editor. Technical report, Department of Computer Science, University of Otago, Dunedin, New Zealand, 1999.

[1] http://www.heroku.com/