$N$-Protractor: A Fast and Accurate Multistroke Recognizer

Lisa Anthony  
Information Systems  
UMBC  
lanthony@umbc.edu

Jacob O. Wobbrock  
Information School | DUB Group  
University of Washington  
wobbrock@uw.edu

Graphics Interface 2012
Pen and Finger Gesture Input

My Word Coach (DS)

Obenkyo (Android)

Mr. Spiff’s Revenge (PC)
Pen and Finger Gesture Input

Gesture recognition for new platforms requires rapid prototyping but existing approaches are complex.

$1 is a simple, accurate recognizer to address this challenge for unistrokes.

– “one dollar” = cheap, fast, easy.
– [Wobbrock et al, UIST 2007]
$N$ Multistroke Gesture Recognizer

$N$ extends $1$ from unistroke to multistroke
- Include pen-up periods
- Recognize based on the entire path the arm takes in the air above the surface

Template matcher: point per point correspondence (like $1$)

[Anthony & Wobbrock, GI 2010]
Open Source and Easy to Adapt

Open-source (C#, Java, more…)
  – Rapid pick-up by community
  – AlphaCount iPhone app
  – iOS port of the recognizer

~100 lines of pseudo-code

Supports independent app development
Space and Time Constraints
$N$ represents and tests all possible *permutations* of a given gesture (e.g., 8 ways to make an “x”)

- Easier for the user not to define them all
- Very costly in space and time for the system

**The 8 ways to write an “x” gesture**

**$N$’s internal representation of an “x”**

**Multi-stroke representation of an “x”**
$N$ uses **GSS** to find the best alignment between two gestures:
- Rotation-invariant, tolerant to small variations
- All pairwise comparisons
- Iterative, time-consuming

**Approximation of Golden Section Search**
Search through a space
Protractor is a closed-form method to find best alignment for $1$

- Approx. 80% time savings (for 16 gestures when 9 templates loaded)
- [Li, CHI 2010]
- http://yangl.org/protractor/protractor.pdf

Protractor’s closed-form alignment [Li, CHI 2010]
Evaluating $N$-Protractor

Added Protractor matching method to $N$ to replace GSS
  - Protractor is essentially $1$ with a different matching method

Cut out one source of time complexity in $N$ (iterative alignment search)
  - Permutations cost remains

$N$-GSS: $[(x_1, y_1), (x_2, y_2), (x_n, y_n)]$
$N$-Protractor: $[x_1, y_1, x_2, y_2, \ldots x_n, y_n]$

**Internal representations**
Gesture Corpora Tested

**Algebra Corpus**
- 40 middle and high school students
- Unconstrained math symbols
- [Anthony et al, AIEd 2007]

**Mixed Multistroke Corpus**
- 10 able-bodied adults
- Gesture and UI oriented symbols (1-3 strokes)

**$1 Unistroke Corpus**
- 10 able-bodied adults
- Gesture and UI oriented symbols (1 stroke only)
- [Wobbrock et al, UIST 2007]

MMG Corpus
Recognition Accuracy

- GSS and Protractor methods minimally different recognition accuracy
- Algebra dataset still most challenging for $N$-family
- Unistrokes easiest

Recognition accuracy with Protractor vs. GSS for different corpora

Anthony & Wobbrock. $N$-Protractor: A Fast and Accurate Multistroke Recognizer, GI 2012
Recognition Speed

- GSS consistently slower than Protractor on all datasets
- MMG: Protractor is more than 10x faster than GSS
- Benefits of Protractor magnified for multistrokes
- (More data was available for Algebra dataset)

Time per recognition with Protractor vs. GSS for different corpora
Impact of Input Method and Speed

- Just for $N$-Protractor
- Gestures entered via finger more accurate than stylus
- No difference due to input speed

Accuracies with Finger vs Stylus Input
Most Confusable Gestures

<table>
<thead>
<tr>
<th>Tested gesture</th>
<th>Confused gesture</th>
<th>No. of confusions</th>
<th>% tests confused</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclamation point</td>
<td>Half note</td>
<td>2488</td>
<td>4.6%</td>
</tr>
<tr>
<td>Line</td>
<td>[no result]</td>
<td>1466</td>
<td>2.7%</td>
</tr>
<tr>
<td>Exclamation point</td>
<td>N</td>
<td>1162</td>
<td>2.2%</td>
</tr>
<tr>
<td>H</td>
<td>N</td>
<td>822</td>
<td>1.5%</td>
</tr>
<tr>
<td>Six point star</td>
<td>Null</td>
<td>777</td>
<td>1.4%</td>
</tr>
<tr>
<td>Five point star</td>
<td>[no result]</td>
<td>749</td>
<td>1.4%</td>
</tr>
<tr>
<td>Exclamation point</td>
<td>Arrow</td>
<td>737</td>
<td>1.4%</td>
</tr>
<tr>
<td>Exclamation point</td>
<td>T</td>
<td>703</td>
<td>1.3%</td>
</tr>
<tr>
<td>P</td>
<td>D</td>
<td>608</td>
<td>1.1%</td>
</tr>
<tr>
<td>Half note</td>
<td>[no result]</td>
<td>561</td>
<td>1.0%</td>
</tr>
<tr>
<td>N</td>
<td>P</td>
<td>553</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Most confused gesture pairs in MMG corpus

- No symbol mis-recognized more than 4.6% of the time
- Worst recognized symbol was **exclamation point**
- Confusable pairs indicate small strokes and 1D gestures most difficult
- No reciprocal confusability relationships
Summary

✔ $N$-Protractor highly **accurate**.
✔ $N$-Protractor **>10x faster** than $N$-GSS on MMG.
✔ $N$-Protractor more accurate for **finger** than stylus gestures.
✔ No impact of **input speed**.
✔ $N$-Protractor has trouble with gestures with very small component strokes (e.g., **exclamation points**) and thin gestures (e.g., **lines**).
Next Steps

Working on optimizations to reduce number of comparisons and permutations stored.

Coming soon: $P – a simple, fast and accurate point-cloud approach to gesture recognition!
Questions?

Contact:
- Lisa Anthony, PhD
  - lanthony@umbc.edu
- Jacob O. Wobbrock, PhD
  - wobbrock@uw.edu

Pseudo-code:

References


