Applying Pattern-based Classification to Sequences of Gestures

Suzanne Aussems[1] Mingyuan Chu[2] Sotaro Kita[1] Menno van Zaanen[3]

s.aussems@warwick.ac.uk mingyuan.chu@mpi.nl s.kita@warwick.ac.uk m.m.vanzaanen@uvt.nl

[1] University of Warwick [2] Max Planck Institute for Psycholinguistics [3] Tilburg University

Introduction

- Previous research related type, frequency, and salience of isolated gestures to personality traits, cognitive skills, and empathy levels in adults (Hostetter & Alibali, 2006; Chu & Kita, 2011; Hostetter & Potthoff, 2012; Chu, Meyer, Foulkes, & Kita, 2014)
- More information might be hidden in *gesture sequences* than in single gestures
- Empathic people might structure their gestures differently from less empathic people, because gestural communication is shaped in part by speakers' desire to communicate information clearly to their listeners (Hostetter, Alibali, & Schrager, 2011)
- We present a pattern-based classification approach to predict empathy levels in adults based on *sequences of gestures* they produced

Approach

- Pattern-based sequence classification (PBSC) is an approach that aims to identify discriminative patterns in longer sequences of symbols (van Zaanen & Gaustad, 2010)
- PBSC has been used to identify patterns in written language (van Zaanen & van de Loo, 2012) and musical notations (van Zaanen, Gaustad, & Feijen, 2011)
- Subsequences (n-grams), representing consecutive gestures, are used as patterns

Training	Classification
1. Extract all potential patterns from dataset	1. Apply patterns to unseen sequence
2. Identify patterns that help best to classify	2. Select class that matches best

Weights

- Patterns receive weight per class
  - High weight means high classification strength
  - *tf*idf* is measure taken from Information Retrieval
- Term Frequency: *tf* of pattern i in class j measures regularity
- Inverse Document Frequency: *idf* of pattern i measures discriminative power
- *tf*idf* is the combination of *tf* and *idf*

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

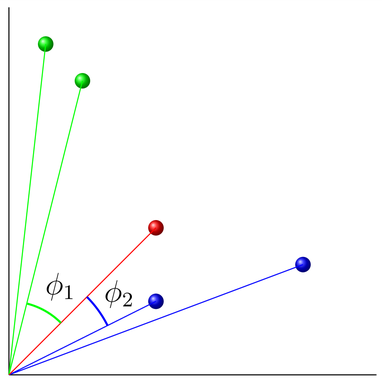
$$idf_i = \log \frac{|D|}{|\{d : t_i \in d\}|}$$

$$tf{*}idf_{i,j} = tf_{i,j} \times idf_i$$

Vector space

- Patterns can be seen as vectors
- Vector space has classes as dimensions
- Sequences of gestures are represented as vectors
  - Sum of matching patterns
- Classification according to k-NN ($k = 1$)
  - Cosine distance metric

$$\cos(\phi) = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$



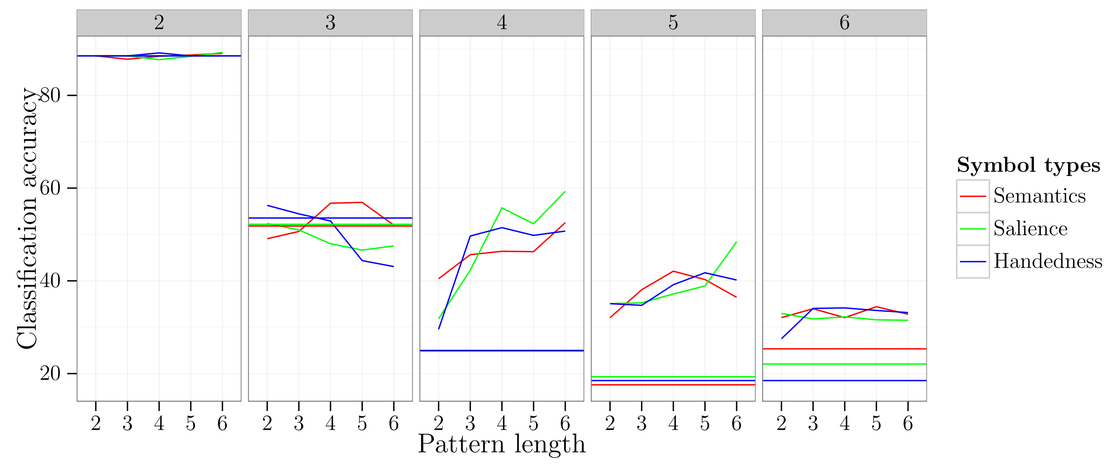
Dataset

- 122 English native speakers (19.41 ($\pm$4.85) year old)
- 11,032 annotated speech-accompanying gestures
- Three symbolic data representations (symbol types)
  - Semantics (type of gesture): 7 unique symbols
  - Salience (size of gesture): 4 unique symbols
  - Handedness (which hand(s)): 3 unique symbols
- Empathy Quotient (Baron-Cohen & Wheelwright, 2004)
  - Scores (0–80) linearly subdivided into 2–6 classes

Method

- Three distinct datasets (one per symbol type)
  - relate symbolic gesture sequences to empathy levels
- Compare performance of PBSC system using longer patterns ($n = 2$, 3, 4, 5, or 6) with the performance of the system single gestures ($n = 1$)
- Measure accuracy of the system through 10-fold cross-validation

Results



- Straight horizontal lines represent pattern length (n) = 1 (baseline of single gestures)
- Panels represent different number of classes (i.e., 2 means low versus high empathy)
- Increasing number of classes: *idf* has more effect, but there is less training data available
- Significant effects with 4 or 5 empathy classes: longer patterns outperform short patterns

Conclusion

- The pattern-based sequence classifier (PBSC) works well with gesture data
- Gesture sequences predict empathy levels in adults better than single gestures
- Typically, the best system performance is obtained with 4 or 5 empathy classes
  - Short patterns have low discriminative power, but occur frequently
  - Long patterns have high discriminative power, but do not occur frequently enough
- The system's performance depends on a complex interaction between the number of unique symbols, the number of classes, and the amount of training data
- PBSC is potentially widely applicable (e.g., visual, auditory, and motor sensory domains)

Acknowledgments

This research was supported by an Economic and Social Research Council Grant RES-062-23-2002 granted to Sotaro Kita and Antje Meyer. We thank Antje Meyer for allowing the use of the dataset. Our gratitude goes to Farzana Bhaiyat, Christina Chelioti, Dayal Dhiman, Lucy Foulkes, Rachel Furness, Alicia Griffiths, Beatrice Hannah, Sagar Jilka, Johnny King Lau, Valentina Lee, Zeshu Shao, Callie Steadman, and Laura Torney for their help with data collection and coding. We thank Birmingham City University, Bishop Vesey's Grammar School, City College Birmingham, CTC Kingshurst Academy, and University College Birmingham for their participation in our research.

References

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163-175.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, *140*(1), 102-116.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology*, *143*(2), 694-709.

Hostetter, A. B., & Alibali, M. W. (2006). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture*, *7*(1), 73-95.

Hostetter, A. B., Alibali, M. W., & Schrager, S. M. (2011). If you don't already know, I'm certainly not going to show you! Motivation to communicate affects gesture production. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture.* Amsterdam, The Netherlands: Benjamins.

Hostetter, A. B., & Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture*, *12*(1), 62-83.

van Zaanen, M., & Gaustad, T. (2010). Grammatical Inference as Class Discrimination. In J. Sempere & P. Garcia (Eds.), *Grammatical Inference: Theoretical Results and Applications.* Berlin/Heidelberg: Springer.

van Zaanen, M., Gaustad, T., & Feijen, J. (2011). Influence of Size on Pattern-based Sequence Classification. In P. van der Putten, C. Veenman, J. Vanschoren, M. Israel, & H. Blockeel (Eds.), *Proceedings of the 20th Dutch Conference on Machine Learning* (p. 53-60). The Hague, The Netherlands.

van Zaanen, M., & van de Loo, J. (2012). Learning Interpretations Using Sequence Classification. In J. Heinz, C. de la Higuera, & T. Oates (Eds.), *Proceedings of the Eleventh International Conference on Grammatical Inference* (p. 220-223). Washington, DC.