ExTOL: Automatic recognition of British Sign Language using the BSL Corpus

Kearsy Cormier

Deafness Cognition and Language Research Centre, University College London, London, UK, k.cormier@ucl.ac.uk

Neil Fox

Deafness Cognition and Language Research Centre, University College London, London, UK, neil.fox@ucl.ac.uk

Bencie Woll

Deafness Cognition and Language Research Centre, University College London, London, UK, b.woll@ucl.ac.uk

Andrew Zisserman

Department of Engineering Science, Oxford University, Oxford, UK, az@robots.ox.ac.uk

Necati Cihan Camgoz

Centre for Vision, Speech and Signal Processing, Department of Electrical and Electronic Engineering, University of Surrey, Guildford, UK, n.camgoz@surrey.ac.uk

Richard Bowden

Centre for Vision, Speech and Signal Processing, Department of Electrical and Electronic Engineering, University of Surrey, Guildford, UK, r.bowden@surrey.ac.uk

ABSTRACT

Although there have been some recent advances in sign language recognition, part of the problem is that most computer scientists in this research area do not have the required in-depth knowledge of sign language, and often have no connection with the Deaf community or sign linguists. For example, one project described as translation into sign language aimed to take subtitles and turn them into fingerspelling. This is one of many reasons why much of this technology, including sign-language gloves, simply doesn't address the challenges.

However there are benefits to achieving automatic sign language recognition. The process of annotating and analysing sign language data on video is extremely labour-intensive. Sign language recognition technology could help speed this up.

Until recently we have lacked large signed video datasets that have been precisely and consistently transcribed and translated – these are needed to train computers for automation. But sign language corpora – large datasets like the British Sign Language Corpus (Schembri et al., 2014) - bring new possibilities for this technology.

Here we describe the project "ExTOL: End to End Translation of British Sign Language" – which has one aim of building the world's first British Sign Language to English translation system and the first practically functional machine translation system for any sign language. Annotation work previously done on the BSL Corpus is providing essential data to be used by computer vision tools to assist with automatic recognition. To achieve this the computer must be able to recognise not only the shape, motion and location of the hands but also nonmanual features – including facial expression, mouth movements, and body posture of the signer. It must also understand how all of this activity in connected signing can be translated into written/spoken

language. The technology for recognising hand, face and body positions and movements is improving all the time, which will allow significant progress in speeding up automatic recognition and identification of these elements (e.g. recognising specific facial expressions or mouth movements or head movements). Full translation from BSL to English is of course more complex but the automatic recognition of basic position and movements will help pave the way towards automatic translation. In addition, a secondary aim of the project is to create automatic annotation tools to be integrated into the software annotation package ELAN. We will additionally make the software tools available as independent packages, thus potentially allowing their inclusion into other annotation software such as iLex.

In this poster we report on some initial progress on the ExTOL project. This includes (1) automatic recognition of English mouthings, which is being trained using 600+ hours of audiovisual spoken English from British television and TED videos, and developed by testing on English mouthing annotations from the BSL Corpus. It also includes (2) translation of IDGloss to Free Translation, where the aim is to produce English-like sentences given sign glosses in BSL word order. We report baseline results on a subset of the BSL Corpus, which contains 10,000+ sequences and over 5,000 unique tokens, using the state-of-the-art attention based Neural Machine Translation approaches (Camgoz et al., 2018; Vaswani et al., 2017). Although it is clear that free translation (i.e. full English translation) cannot be achieved via ID glosses alone, this baseline translation task will help contribute to the overall BSL to English translation process - at least at the level of manual signs.

ACKNOWLEDGMENTS

REFERENCES

Schembri, A., Fenlon, J., Rentelis, R., & Cormier, K. (2014). British Sign Language Corpus Project: A corpus of digital video data and annotations of British Sign Language 2008-2014 (Second Edition). London: University College London. http://www.bslcorpusproject.org.

Camgoz, N.C., Hadfield, S., Koller, O., Ney, H., & Bowden, R. (2018). Neural sign language translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7784-7793).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998-6008).