

Improving Automatic Sign Language Recognition through Sign Perceptual Knowledge

Gineke A. ten Holt^{1,2}, Marcel J.T. Reinders¹, Emile A. Hendriks¹, Andrea J. van Doorn², and Huib de Ridder²

¹ Information and Communication Theory Group, Delft University of Technology, Mekelweg 4, 2628 CD, Delft, the Netherlands

² Human Information Communication Design, Delft University of Technology, Landbergstraat 15, 2628 CE, Delft, the Netherlands
[g.a.tenholt,m.j.t.reinders,
e.a.hendriks,a.j.vandoorn,h.deridder]@tudelft.nl

keywords sign language; automatic sign language recognition; sign perception;

1 Introduction

Studies in automatic sign language recognition (ASLR) have mainly focused on data-driven methods (see [1]). Little attention has been paid to the perception of sign language and how knowledge from this field can improve ASLR. In part, this is because the field of sign perceptual research itself is quite young, and knowledge is far from complete. However, recent studies ([2]) have yielded results which could significantly improve ASLR. This paper discusses two experiments that demonstrate the advantage of applying perceptual knowledge to a vision-based ASLR-system. Figure 1 illustrates the applications.

2 Application of Perceptual Knowledge

Key Frames for Information Extraction. Current vision-based ASLR systems often extract certain features for each frame in a sign movie indiscriminately. However, [2] showed that not all frames in a sign are equally informative. Especially the fragments of the sign before and after the stroke (the stroke is the central phase of the sign) can be ignored without loss of crucial information. The results of [2] also suggest that variations in timing in a sign are unimportant for its (lexical) meaning. Together, this means that a set of key frames taken from the stroke phase of a sign could contain the same amount of relevant information as all frames of a sign. This paper discusses an experiment investigating the effect of using fewer frames for feature extraction on recognition performance. The key frame selection algorithm differed for different types of signs.

Feature Value Constraints. In [2] it was investigated how variations in feature values affect the correctness of a sign. It showed that it is possible to determine constraints for feature values. Knowledge of such constraints can be used in

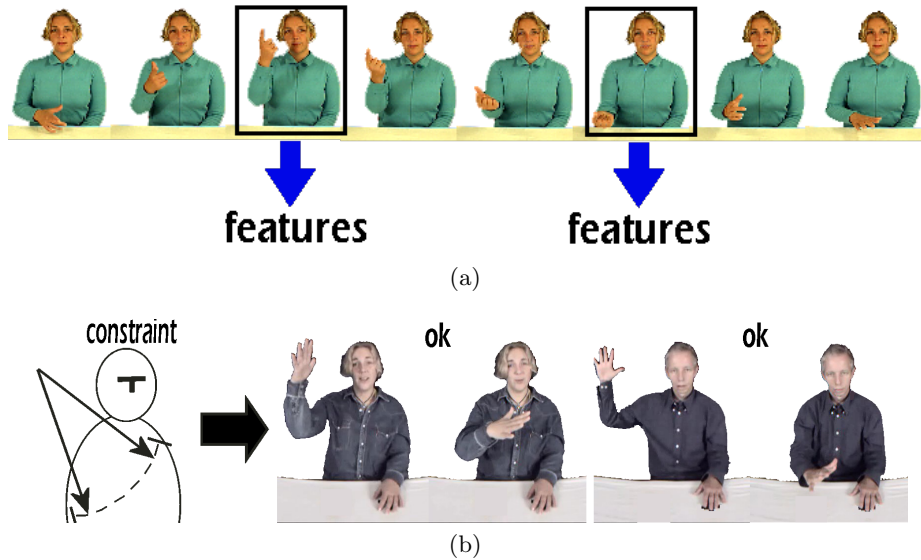


Fig. 1. Applying knowledge to ASLR. (a) Not all parts of a sign are equally informative. Select a subset of the most informative frames for feature extraction. (b) Perceptual research can give us the constraints for a feature value, e.g. the movement value: “inclined downward”. Using this knowledge, variation in signs can be handled without the need for large training sets.

ASLR to determine allowed variation. For example, say that all downward movements of which the direction does not deviate more than 30° from vertical are acceptable versions of the movement feature “straight down”. If this is known, then any movement within that range will be annotated as “straight down”. Sign variants differing in precise path direction would all receive this feature value despite their differences. This paper discusses an experiment comparing sign recognition using pre-defined constraints in the feature extraction step with regular sign recognition (which models variation statistically). The comparison is made for various training set sizes.

References

1. Ong, S.C., Ranganath, S.: Automatic sign language analysis: A survey and the future beyond lexical meaning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(6) (2005) 873–891
2. ten Holt, G.A., Arendsen, J., de Ridder, H., van Doorn, A.J., Reinders, M.J.T., Hendriks, E.A.: Sign language perception research for improving automatic sign and gesture recognition. Submitted to *SPIE Human Vision and Electronic Imaging XIV*. (2009)