Segmentation of Continuous Sign Language

A THESIS PRESENTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF

DOCTOR OF PHILOSOPHY

IN ENGINEERING

MASSEY UNIVERSITY, PALMERSTON NORTH,

NEW ZEALAND

SHUJJAT KHAN

2014

ABSTRACT

Sign language is a natural language of deaf people comprising of hand gestures, facial expressions and body postures. It has all the constituents that are normally attributed to a natural language, such as variations, lexical/semantic processes, coarticulations, regional dialects, and all the linguistic features required for a successful communication. However, sign language is an alien language for a vast majority of the hearing community so there is a large communication barrier between both the sides. To bridge this gap, sign language interpreting services are provided at various public places like courts, hospitals and airports. Apart from the special needs, the digital divide is also growing for the deaf people because most of the existing voice-based technologies and services are completely useless for the deaf. Many attempts have been made to develop an automatic sign language interpreter that can understand a sign discourse and translate it into speech and vice-versa. Unfortunately, existing solutions are designed with tight constraints so they are only suitable for use in a controlled environment (like laboratories). These conditions include specialized lighting, fixed background and many restrictions on the signing style like slow gestures, exaggerated or artificial pause between the signs and wearing special gloves. In order to develop a useful translator these challenges must be addressed so that it could be installed at any public place.

In this research, we have investigated the main challenges of a practical sign language interpreting system and their existing solutions. We have also proposed new solutions (like robust articulator detection, sign segmentation, and availability of reliable scientific data) and compared them with the existing ones. Our analysis suggests that the major challenge with existing solutions is that they are not equipped to address the varying needs of the operational environments. Therefore, we designed the algorithms in a way that they stay functional in dynamic environments. In the experiments, our proposed articulator segmentation technique and boundary detection method have outperformed all the existing static approaches when tested in a practical situation. Through these findings, we do not attempt to claim a superior performance of our algorithms in terms of the quantitative results; however, system testing in practical places (offices) asserts that our solutions can give consistent results in dynamic environments in comparison to the existing solutions.

Temporal segmentation of continuous sign language is a new area which is mainly addressed by this thesis. Based on the conceptual underpinnings of this field, a novel tool called DAD signature has been proposed and tested on real sign language data. This

segmentation tool has been proven useful for sign boundary detection using the segmentation features (pauses, repetitions and directional variations) embedded in a sign stream. The DAD signature deciphers these features and provides reliable word boundaries of sentences recorded in a practical environment. Unlike the existing boundary detectors, the DAD approach does not rely on the artificial constraints (like slow signing, external trigger or exaggerated prosody) that restrict the usability of an interpreting system. This makes DAD viable for practical sign language interpreting solutions.

As signified in this dissertation, the development of the much awaited useful sign language interpreter is achievable now. We have established that by making use of our proposed techniques, the strict design constraints of the existing interpreters can be mitigated without affecting the overall system performance in a public place. In a nutshell, our research is a step forward towards the possibility of turning the idea of a practical automatic interpreter into a reality.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisors A/Prof. Gourab Sen Gupta and A/Prof. Donald Bailey for their continuous support of my PhD research. I would like to thank them for their patience, motivation, enthusiasm, and their amazing sense of humour which is always required for an ideal learning environment.

Higher education commission Pakistan deserves a profound acknowledgement for providing me this opportunity and continuous support.

I would like to thank Annette Scott and everybody from the Deaf Aotearoa, Palmerston North for their active participation in my research.

Last, but not the least, I would like to thank my family: my parents and my sisters for supporting me throughout my life.

GLOSSARY

Inflect	To modify a basic word
Prosody	Rhythm, style
Iconicity	Conceived similarity between the form of a sign and its meaning[1]
Articulator	Communicating organ
Coarticulation	A linguistic process in which morphology of a sign is affected by the neighbouring one
Lexicon	Sign, word, or gesture are used interchangeably
Gloss	Isolated sign/gesture
DAD	Delayed Absolute Difference
Dademe	A de-facto terminology defining the sub units of a sign in terms of DAD's segmentation features
Semiotic	Study of signs and sign processes

[1] http://en.wikipedia.org/wiki/Iconicity

TABLE OF CONTENTS