Case ID:M20-304P

Published: 5/7/2021

Inventors

Sandeep Gupta

Ayan Banerjee

Contact

Shen Yan
shen.yan@skysonginnovations.com

Machine Recognition of New Gestures After Training with Sign Language

Background

In American Sign Language (ASL), all of the nearly 10,000 gestures for English words are composed using a set of over 80 handshapes, six locations and around 20 unique movements. Each handshape, movement, and location has a semantic relation with the English word and can be considered as concepts. Each gesture can be expressed using a unique ordering of start handshape and start location, a movement type, end handshape and end location, which is the canonical form for that gesture. If a machine learns these unique concepts, then by combining them following a language, there is potential to recognize gestures that it has never seen before. This concept of recognizing previously unseen classes without access to training data is known as zero-shot learning. Notable applications include ASL learning, training personnel in various domains such as construction and military, and validating the quality of unsupervised physiotherapeutic exercises.

Invention Description

Researchers at Arizona State University have developed the concept of using a canonical form as the intermediate modular representation required for zero-shot learning of gesture-based languages. The fundamental difference is in the definition of a concept that enables soft matching and the usage of canonical forms that convert an example into concepts arranged in spatiotemporal order. The present system and method apply this embedding strategy for zero-shot learning of ASL gestures.

In experiments, two datasets were utilized: (1) IMPACT Lab training dataset, consisting of 23 ASL gestures each executed three times from 130 first-time ASL learners, and (2) the ASLTEXT dataset, consisting of 190 gestures each executed six times on an average. The developed system was able to recognize 19 arbitrarily chosen—and previously unseen in the IMPACT dataset—gestures from seven individuals who were not a part of 130 learners. From the ASLTEXT dataset, 34 unseen gestures were recognized without any retraining. Normalized accuracy on the ASLTEXT dataset is 66% which is 13.6 % higher than the state-of-art technique.

Potential Applications

- Sign language gesture transcription

- Video-based gesture recognition

- Education for hearing-impaired individuals

- Tracking of physiotherapeutic movements and training progress

Related Publication (PDF): Concept Embedding through Canonical Forms: A Case Study on Zero-Shot ASL Recognition

Research Homepage of Professors Sandeep Gupta and Ayan Banerjee