GASLA: Enhancing the Applicability of Sign Language Translation

Jiao Li, Yang Liu, Weitao Xu, Zhenjiang Li
City University of Hong Kong
Background

A communication bridge needs to be built!
Sign Language Translation

Data stream → Data analysis → Machine learning → Translation results → How are you?
Existing works

[IMWUT’18]
Wireless based
not ubiquitous
multipath effect

[SenSys’17]
Camera based
privacy leakage
ambient light

[MobiCom’19]
Wearable based
ubiquitous
privacy-protected
Applicability issues

- Expensive setup overhead
- Low quality data
- Unreliable system performance
- High maintenance cost
- Adding more sentence

Diagram:

- ASL library
  - Words
  - Sentences
- Words library
- Sentence library
  - Nice to meet-you
  - nice, meet-you

Word data collection
Sentence data collection
Expensive setup overhead
Low quality data
Unreliable system performance
Our idea

- Sentences can be crafted by words!

(a) Nice

(b) Meet-you

(c) Nice meet-you

- Connect two words directly?

The answer is NO!
Observation

- Two traces share *similarity*
- *Separation point*
Our system----GASLA

GASLA

Sentence Library Construction
- Preprocessing
- Separation point localization
- Concatenation

Word-level Library

Word Data Collection

ASL Translation System
- Spectrum Generation
- Translator

collected sentence samples

sentence

word samples

Training data flow

Inference data flow
Trajectory selection

- Moving trajectory? 🚫
- Orientation? ✔

![3D plots showing trajectory selection criteria](image)
Sentence Library Construction

- Step 1: Preprocessing
  - a) signal cropping
  - b) resampling
Step 2: Separation point localization
Sentence Library Construction

• Step 2: Separation point localization

\[ \min_{r_y \in \text{space}} d(\text{ori}_x', r_y \cdot \text{ori}_y) \]

Rotation matrix

first trace

second trace

cone-like search space

• Step 3: Concatenation
ASL Translation

generated

collected

Frequency (Hz)

Time (s)

(a) X

(b) Y

(c) Z

(d) X

(e) Y

(f) Z

generated
collected

CityU
ASL Translation

• Our target: provide an independent component

• Network input
• Neural network structure
Experiment setup

- **Hardware**
  LG Watch, SAMSUNG Galaxy S7, Desktop (Intel i7-8700K CPU and Nvidia 2080Ti GPU)

- **Data collection**
  69 words and 41 sentences
  6 volunteer

- **Metric**
  \[\text{Accuracy} = 1 - \text{word error rate}\]

- **Methods**
  SignSpeaker(SP): the state-of-the-art
  GASLA(GA): our design
Evaluation results

- Overall performance

- Quality of generated sensory data
Conclusion

• We identify an applicability issue commonly in prior ASL systems.

• We propose effective techniques to address these issues and provide a clear interface to existing ASL systems.

• We develop a prototype of GASLA and conduct extensive experiments.