DragonFly

Wearable Devices to Enable Communication via ASL

Presented at the

Association for Machine Translation in the Americas (AMTA) 2016 Conference

Austin, Texas

October 31, 2016

Overview

- Statement of User Need
- Our Technical Approach
- Project Status & Next Steps
- Wrap-Up

Statement of User Need

- There is a need for to enable signers to seamlessly communicate with non-signers.
- In day-to-day life, people do not have dedicated interpreters who are with them on a 24 X 7 basis.
- Obtaining an interpreter requires prior arrangements & advanced planning, sometimes weeks in advance.
- Ad-hoc meetings come up, and the chance to participate is lost if an interpreter can’t be found.

Current Situation

Some of the Problems…

- People you run into say, “Text ya later,” and never do…
- You feel you can’t express yourself writing in English as well as you can in ASL

Some of the Consequences…

- Inability to truly connect
- Delays in Decision Making
- Failure to move to next steps

Solving the Challenge

- Develop a tool to allow an ASL signer and a speaker of English to communicate with each other…
  - Face-to-Face
  - Naturally
  - Anytime
  - Anywhere

Our Solution

- To develop wearable devices that will facilitate interactions between signers and non-signers
  - This family of solutions is called ***DragonFly***
  - We envision ***DragonFly*** running on different types of wearables

Why?

- To help bridge the current communication barrier…
  - With ***DragonFly***, a person who uses ASL and someone who does not know ASL can express themselves completely to one another.
  - The opportunity ***DragonFly*** creates is the ability to unlock the potential of every person to fully contribute to the mission.

Technical Approach

Technical Challenges

Sign/Signer Variability

Signal Complexity

- Head/Face Gesture
- Handshape
- Hand Orientation
- Hand Placement
- Hand Motion

Sensor Variability

Session Variability

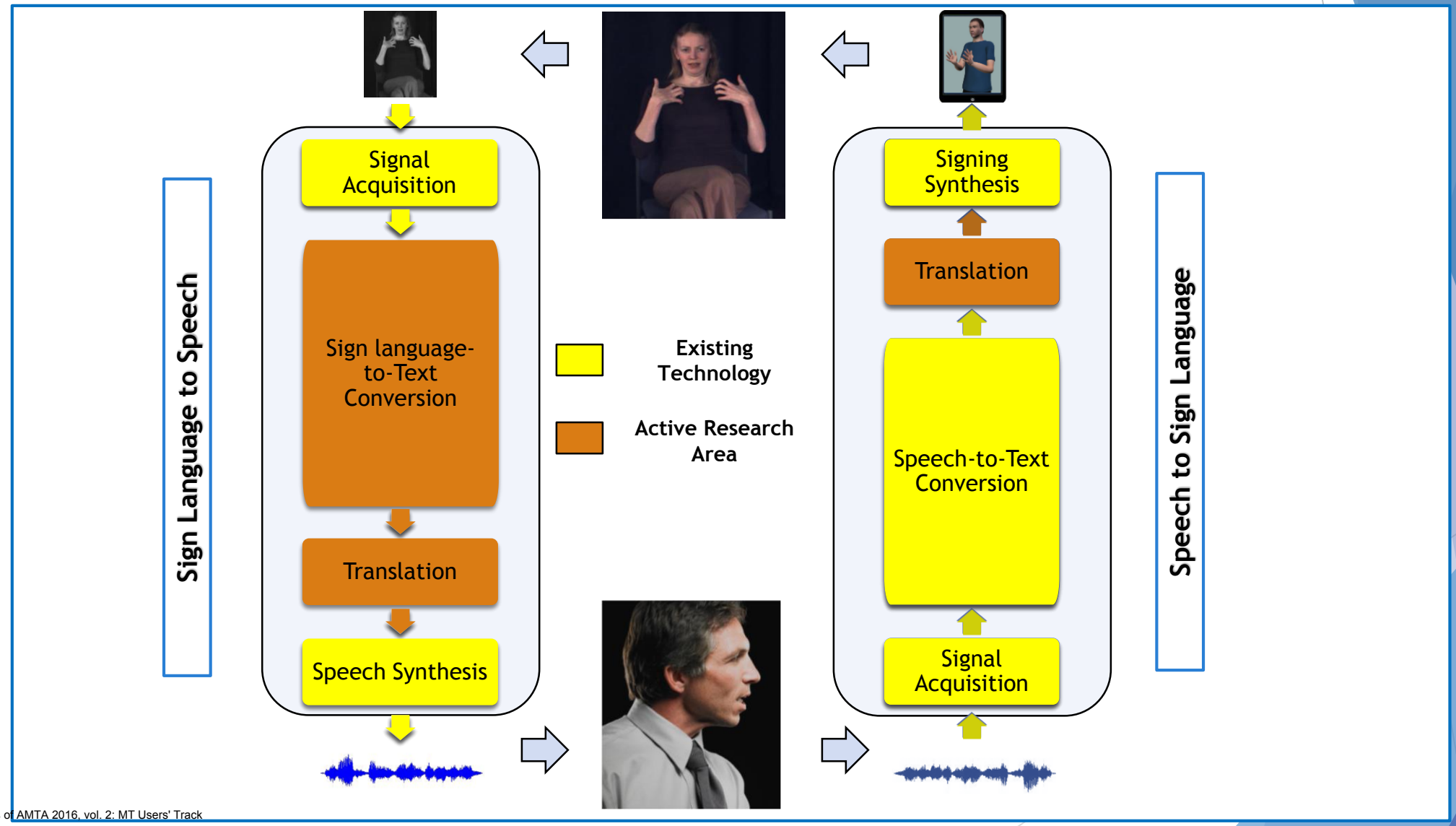
e.g. observation angle

Data Availability

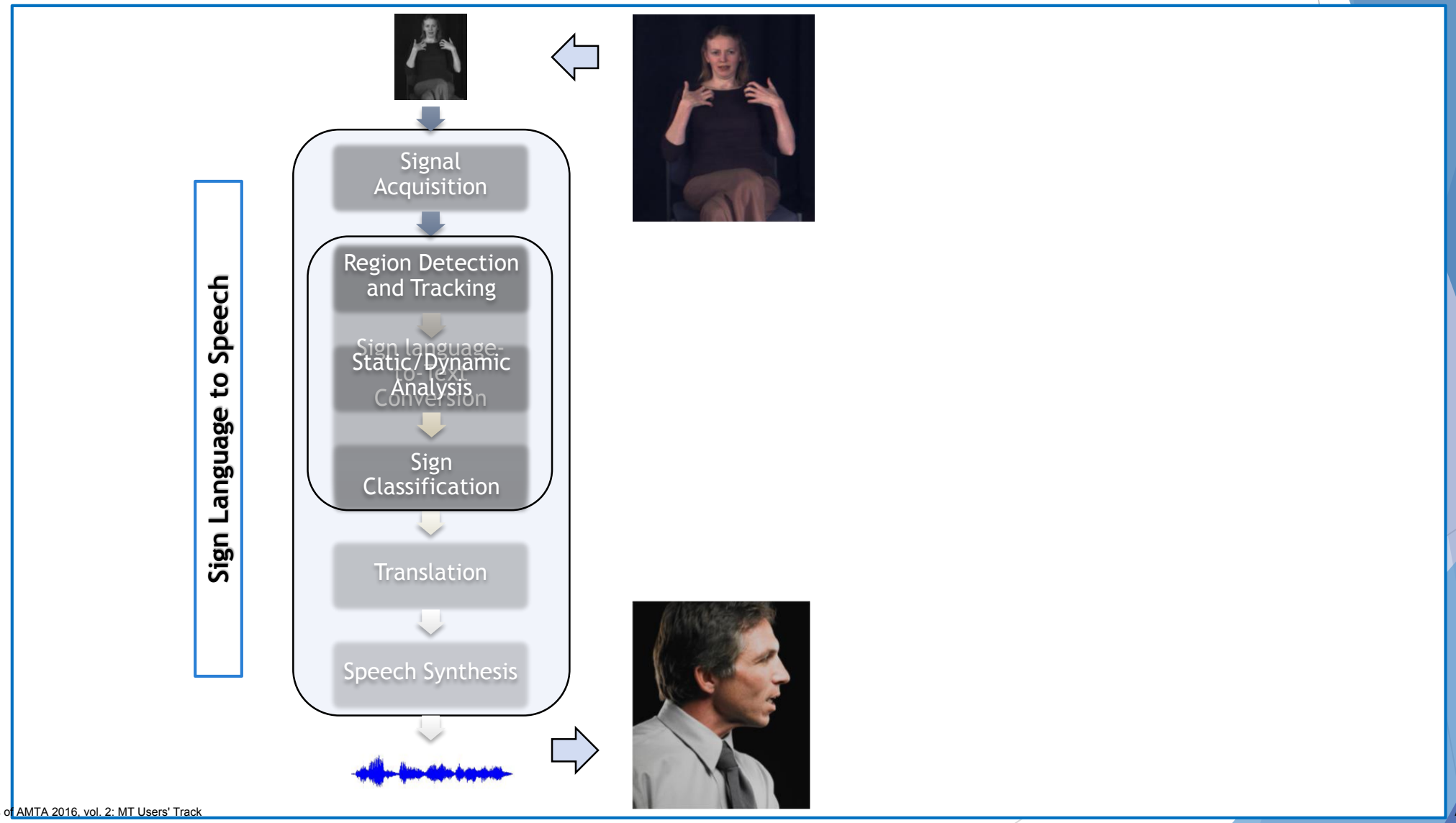
- Limited availability of well annotated ASL<->English content for development and evaluation (e.g. ASLLVD)
- Technical challenges remain for exploiting loosely annotated content (e.g. ASL w/ closed captioning)

ASL: American Sign Language
ASLLVD: American Sign Language Video Dataset

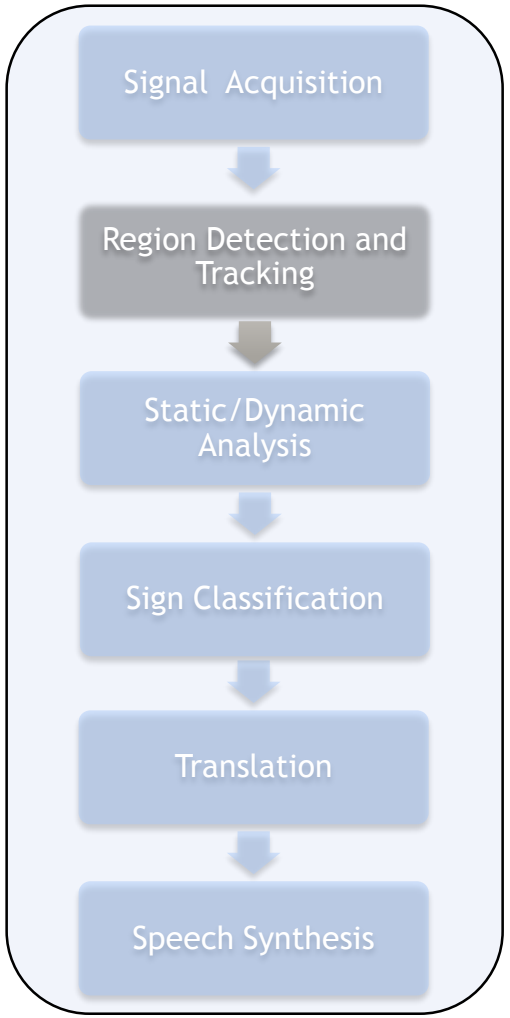
System Overview



Sign Language to Speech

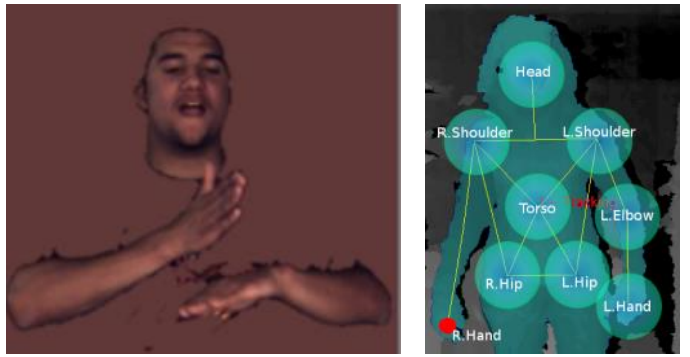


Sign Language to Speech

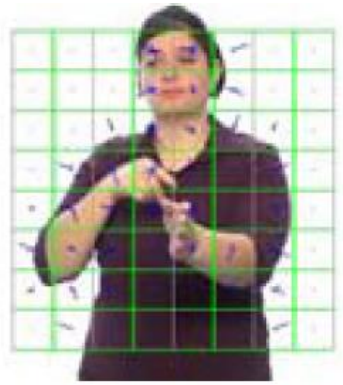


- **Signer Isolation**

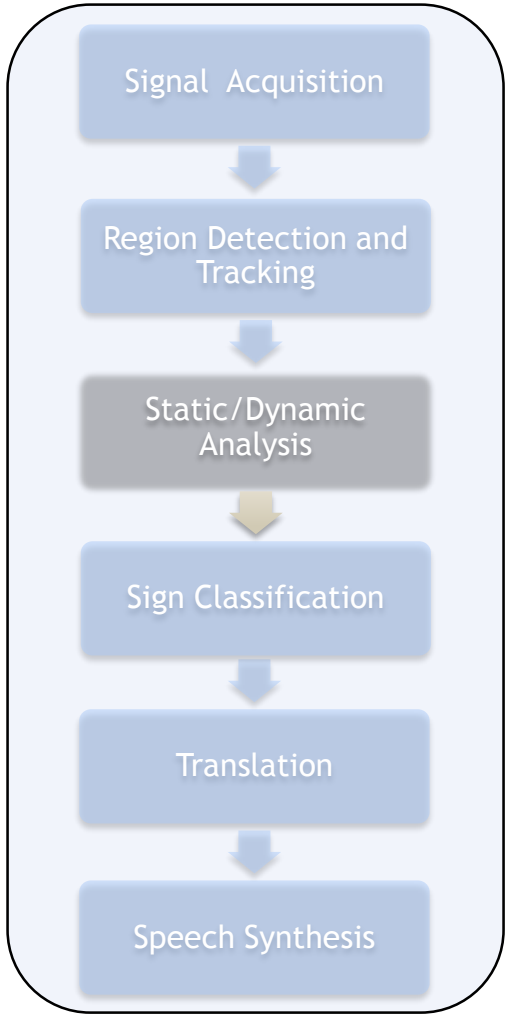
- **Feature Localization**
  - **Face and Hands**
  - **Whole Body**



- **Motion Tracking**



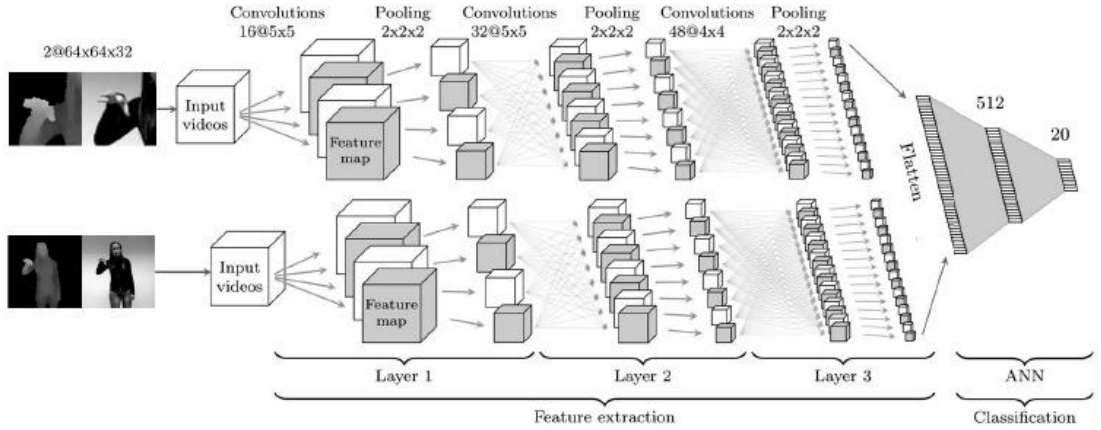
Sign Language to Speech



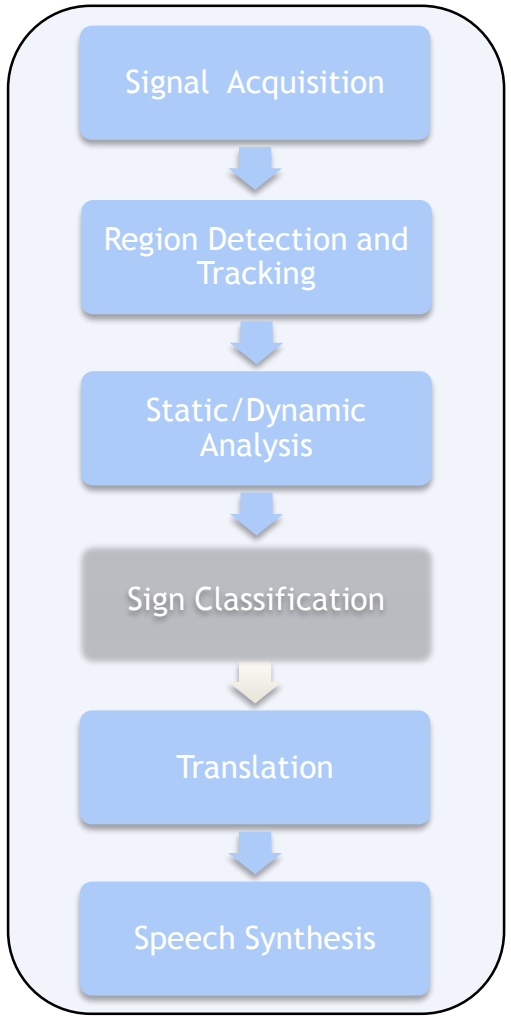
- **Classify Individual Hand Gestures**

- **Evaluate Facial Cues**

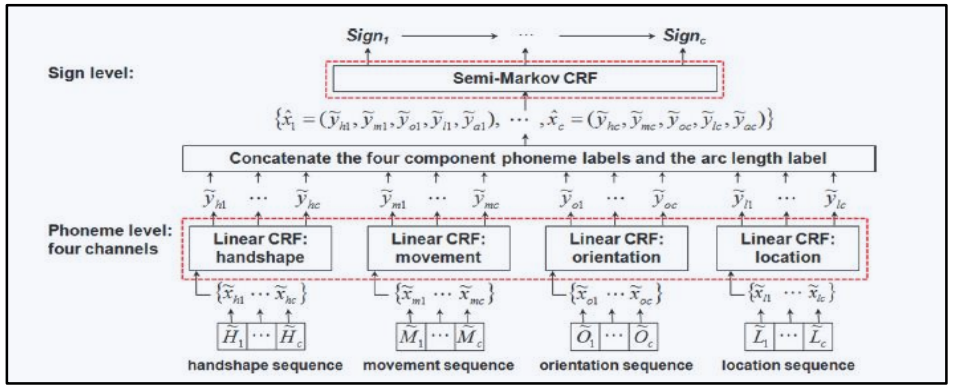
- **Performed on single frame or short series of video**
- **Typically adopt image machine learning methods.**

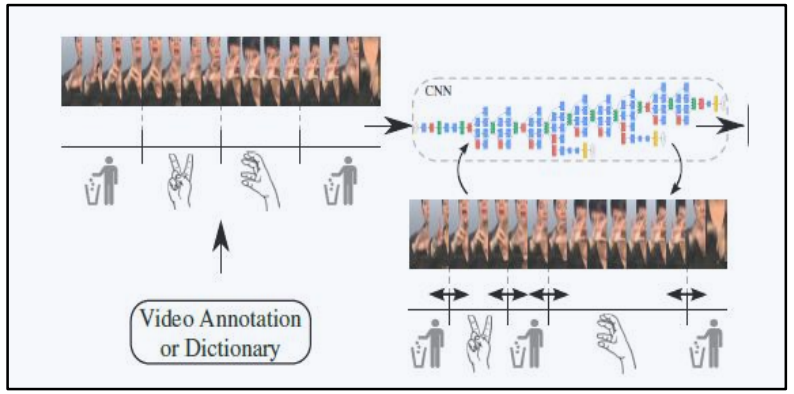


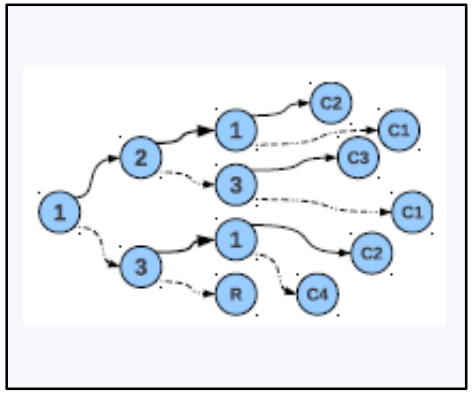
Sign Language to Speech



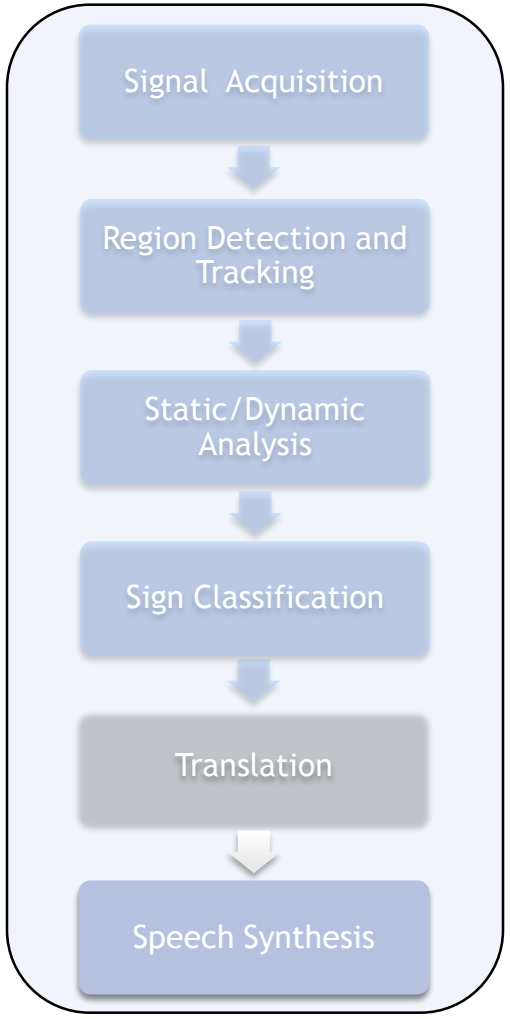
- Given gesture sequences, we want to identify discrete signs
- Guided by a lexicon (e.g. signing dictionary)
- Typically involves dynamic machine learning methods
  - Markov Modeling
  - Neural Networks
  - Sequential Pattern Boosting







Sign Language to Speech



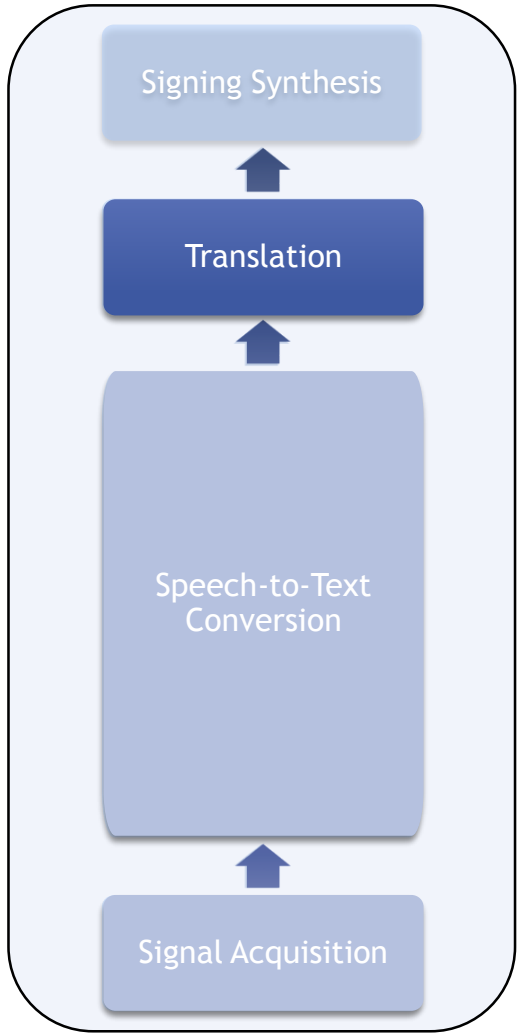
Translation:

- **Convert recognized signing sequences into an English sentences.**
- **Example of Machine Translation (MT) Problem**
- **ASL is not structured like English and is more like Japanese, in that it is a Topic-Comment language. It must be ordered correctly before it is converted to speech, so that it conforms to English syntax and is readily understood by an English speaker.**
- **Requires sizable database of parallel ASL-English data**
  e.g. Television Corpus of Closed-Caption + ASL

Speech to Sign Language

Translation

- **Inverse operation of the MT problem**
- **English sentences converted into sequences of manual and non-manual gestures**



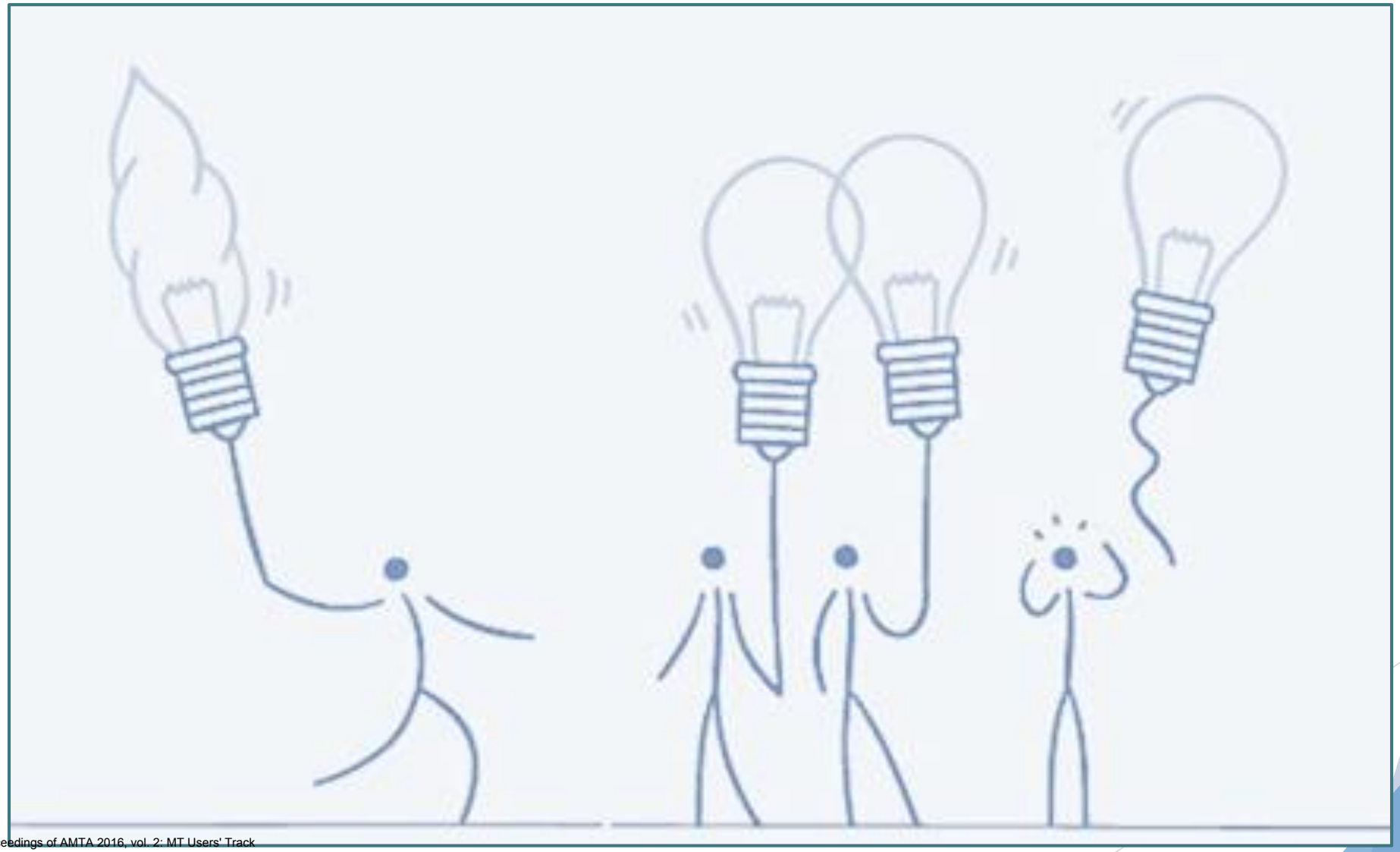
Next Steps


Sign Language to Speech
Signal Acquisition
Sign language-to-Text Conversion
Isolated Signs
Translation
Speech Synthesis
Signing Synthesis
Translation
Speech-to-Text Conversion
Speech to Sign Language
Goals:
• Proof-of-Concept System
• Recognize a set of ~50 isolated signs
• Laptop-based, not necessarily real time
• Obtain feedback from users
ASL Video
English Text: Ask me

Project Status

- We have initiated an extensive literature search to leverage best practices and relevant research done to date.
- We are compiling data, annotation, algorithmic, and system requirements
  - Identified and aggregating annotated ASL datasets
    - ASLLVD RVL-SLLL Gallaudet
  - Identified relevant CNN models for feature extraction
    - VGG and DeepHand models
- We have begun work on a prototype for ASL recognition capabilities.

Questions?



Thank You!

Contact Info:
po17b@icloud.com
nmalyska@ll.mit.edu