

Continuous Sign Language recognition for the design of a gestural server

Yanis Ouakrim

Gipsa-Lab, LISN

October 2021

Table of contents

- 1 What did I do so far?
- 2 What did I understand of the subject so far?
- 3 Listed objectives of the thesis
- 4 Data
- 5 Mediapipe experiment
- 6 Perspectives

What did I do so far?

- Signing papers
- Getting to know people involved in the project from the various teams
- Reading Valentin Belissen's work and Koller's papers
- Testing Mediapipe holistic on video clips

What did I understand so far?

Goal

To investigate deep learning methods for natural sign language recognition, which would better take co-articulation, spatialization and illustrative structures into account.

Continuous Sign Language Recognition for the Design of a Gestural Server

Definition

Continuous Sign Language Recognition: Obtaining glosses from video input containing a continuous flow of signs.

Listed objectives of the thesis

- **Automatic sign spotting** : this approach will allow querying a sign language database directly using a short SL video, without any textual entry ; this detection will automatically annotate videos to enrich the training databases
- **Automatic conversion of sign language sequences into written text** : A special focus will be put on sequence-to-sequence models [8]. The problem of adapting a pre-trained model to a new signer or a new recording set-up (e.g. video device, recording environment, etc.) will also be addressed.

Listed objectives of the thesis

- Building realistic datasets of sign language (in collaboration with industrial partners)
- Image processing/feature extraction/representation learning
- Deep learning pipeline for decoding sign language.

Data challenges

Need for a dataset that fits the need of a Gestural Server:

- What type of gestures? Lexical ? Partially Lexical? Non Lexical Signs?
- What type of discourse?
- What level of spontaneity?
- What type of annotation?

Existing datasets

Name	PHOENIX	DictaSign	SIGNUM	GSL 20	CUNY
Year	2014	2020	2009	2020	
SL	DGS	LSF	DGS	GSL	AS
Tot. Time [h]	12.5	8.0	33.5	9.59	
Videos	386	207	0	310	
#Signers	9	18	25	7	
Vocabulary Size	2048	2252	455	310	
Annotation type	Glosses + Text	Glosses	Glosses + Text	Glosses	

Table: Continuous Sign Language Corporuses

Testing Mediapipe Holistic



Figure: Predicted body, face and hands poses

Testing Mediapipe Holistic

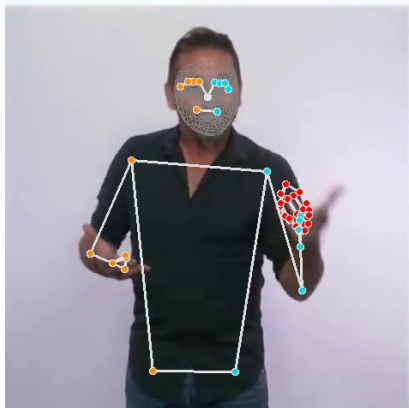


Figure: Predicted body, face and hands poses

What did I do so far?

What did I understand of the subject so far?

Listed objectives of the thesis

Data

Mediapipe experiment

Perspectives

Perspectives

- SOTA of existing datasets in Continuous Sign Language
- Reading papers on SLT
- Completing registration to doctoral school
- Onboarding at LISN