



Design and Implementation of Indonesian Sign Language to Speech Converter

Arry Akhmad Arman, Kudrat Soemintapoera, Evita T Sekar
Electrical Engineering Department,
Institut Teknologi Bandung
Email : aa@lss.ee.itb.ac.id

Outline



- Background
- Introduction
- System Description
- Testing
- Concluding Remark

Background



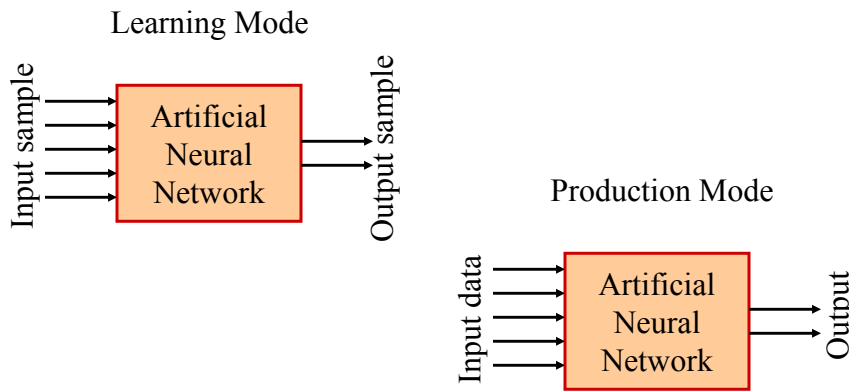
- Deaf people or mute people use "sign language" to communicate each other
- Normal people who don't understand "sign language" cannot receive information from them
- There is a need to develop a system that convert a sign language to a speech.
- With this converter, normal people can "hear" the "message" in the sign language form.

Introduction

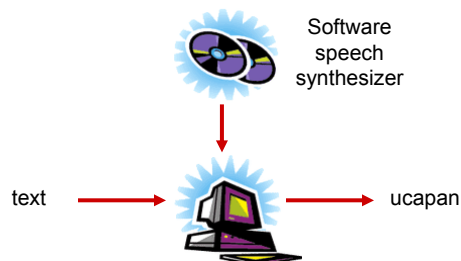


- Sign language is a language used to communicate between deaf/mute people or deaf/mute people with a normal people who understand that language.
- Sign language is a set of sign that formed by hand and also several part of human face.
- Sign language is dedicated to a country.

Introduction : ANN



Introduction : Text to Speech



- Text to Speech System (TTS) is system that convert a text in specify language to a speech in the same language.
- TTS is language dependent system.

Introduction : Text to Speech (2)



- TTS for certain language cannot be used for other language. There are a lot of modification to be done to use an existing TTS for other language.
- There are several type of TTS technologies.
- The most popular TTS technology is a diphone concatenation system. We have to build a diphone database for specify language to use this type of TTS. This type of TTS can generate the most natural and intelligible voice.
- **The first complete Indonesian Text to Speech System have been developed in ITB using diphone concatenation technology.**

Diphone Concatenation Principal

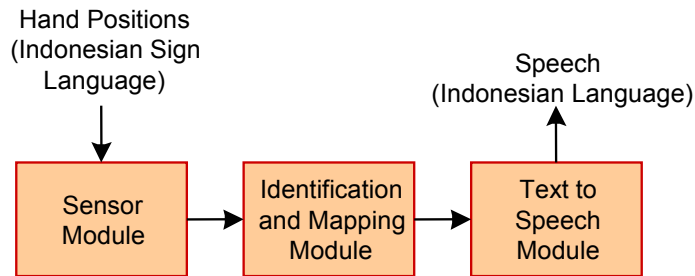


~k ko om mp pu ut te er r~

↑
diphone

Diphone database is
a complete set of diphones

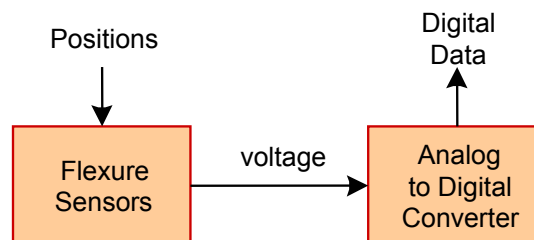
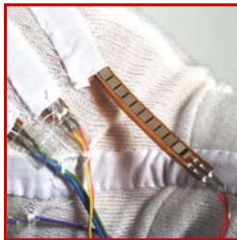
Sign Language to Speech Converter



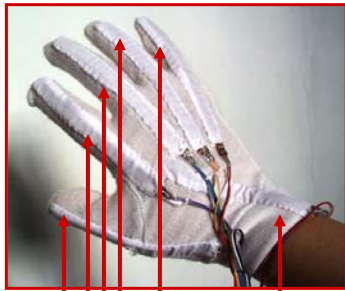
Sensor Module



- We use flexure sensors from Images Company that integrated with a glove and other handmade holder.
- This sensor is variable resistance sensor. The value of resistance will change according to the angle position of the sensor.



Sensor Unit



fingers

wrist

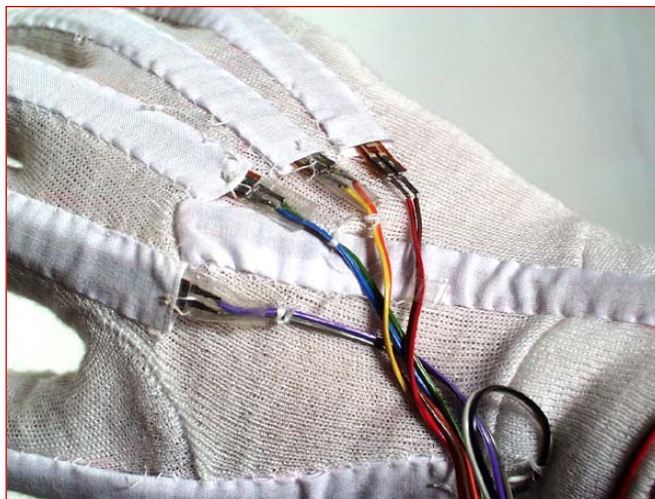


elbow

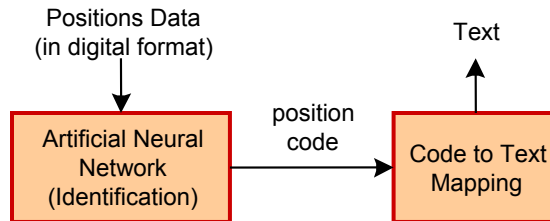


Shoulder joint

Sensor Unit



Identification Modul

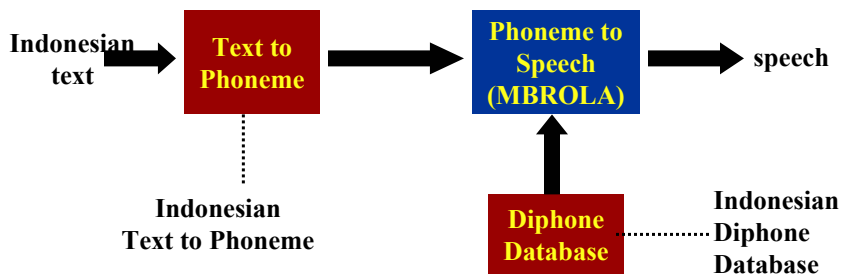


- In fact, the same input (i.e. finger) position for the sensor cannot guarantee to desire the same exact value. User cannot exactly form the same angle to represent every sign in sign language. So, there is varied result from the sensors.
- To handle that condition, we choose Artificial Neural Network.

Text to Speech Module



- Text to Speech module implemented using existing Indonesian TTS that already developed in previous research activity in our laboratory.



- Processed in Belgium
- Published in Internet for non commercial use

Testing : Pre-Test



- Sensor system is one of the most important unit in this system. The final result of the system is very determined by the sensor accuracy.
- From the pre-test phase, we found that the result from sensor module is not consistence.
- The inconsistency came from the sensor in the elbow and shoulder joint position.
- The good result from a sensor at fingers and wrist position is good.
- In fact, we use only one sensor for each position. Finger and wrist movement or position need only one degree of freedom, but the elbow and shoulder joint need than one more degree of freedom.

Testing & Result : Final Test 1



- Testing procedure for finger and wrist position :
 - Choose a sample word that **formed only by fingers and wrist sensors** (22 words)
 - Train ANN with sample words. Each word repeated 10 times.
 - Apply the network weights to the system.
 - Try every trained sign as an input to the system and let the system identify the word. Repeat this test 10 time for each sign.
- Result : 83.18% of trained signs (words) can be identified by the system

Testing : Final Test 2



- Testing procedure for all sensor position :
 - Choose a sample word that **formed by all sensors** (72 words)
 - Train ANN with sample words. Each word repeated 10 times.
 - Apply the network weights to the system.
 - Try every trained sign as an input to the system and let the system identify the word. Repeat this test 10 time for each sign.

- Result : 49.58% of trained signs (words) can be identified by the system

Concluding Remark



- For this application, the movement/position that only consist of one degree of freedom, one sensor its enough to identify the status.
- Sign language to speech converter with a sign formed only by fingers and wrist is possible to implement with good result.
- To detect a node status that consist of more that one degree of freedom, one sensor it not possible to use. So, to build a sign language to speech converter that include the signs that formed by elbow and shoulder joint, it is need to develop more sophisticated sensory system.