

# Speechrecognition

---

Robert Feldhans

9.11.2018

Seminar Robocup

Why even bother?

- faster and more general way to give robots commands
- a necessity for casual users
- user does not need additional hardware
-

**What is Speechrec? What does it consist of?**

1. Hardware
2. Localisation
3. Signal Enhancing
4. Voice Activation Detection
5. Speaker Recognition
6. Speech Recognition
7. Natural Language Processing

content...

# Hardware

---

content...

content...



# Localisation

---

content...

# Signal Enhancing

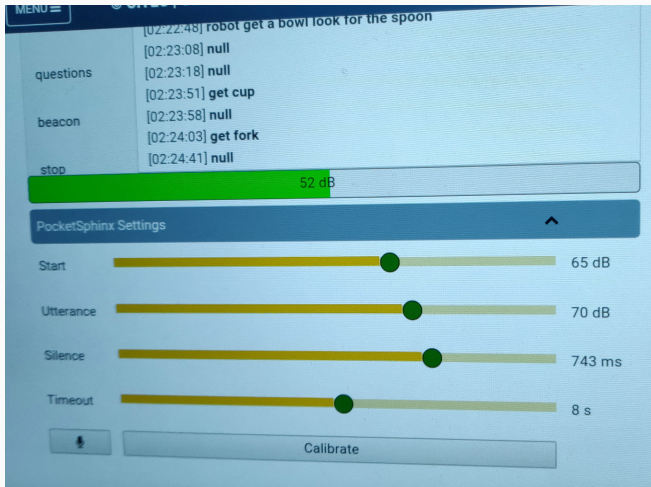
---

content...

# Voice Activation Detection

---

# What we use



**Abb. 1:** Double threshold voice activation detection

## What we use II

A voice activation detection based on audio loudness with three states:

idle Start in this state

starting switch to this state if the audio  $>$  StartDb and stay here as long as audio  $>$  UtteranceDb

ending switch to this state if the audio  $<$  UtteranceDb and stay as long as specified via Silence, then return to idle

A maximum audio length can be specified via Timeout

## Other approaches

loudness-based based on decibel calculation, it will only take into account the single most extreme value in an audio frame

energy-based in contrast to loudness-based approaches, energy calculation will take all values in an audio frame into consideration

frequency-based will calculate frequencies and search for those typically used by human speech



# Speaker Recognition

---

# Speech Recognition

---

Hidden Markov Models vs Deep Learning vs Online

content...

# Corrected Spelling vs Phoneme based recognition

content...

content...

content...

content...



# Natural Language Processing

---

- Just recognizing what was said does not solve all our problems

content...

Thanks for the Attention!

## Discussion

