



Embodied Audition for Robots

Introduction & Overview

Heinrich Löllmann, Christine Evers, and Radu Horaud

EUSIPCO Tutorial

August 31, 2015

Embodied Audition for Robots

- ▶ **All slides of the Tutorial are available at**
<http://robot-ears.eu/eusipco-tutorial-on-embodied-audition-for-robots/>

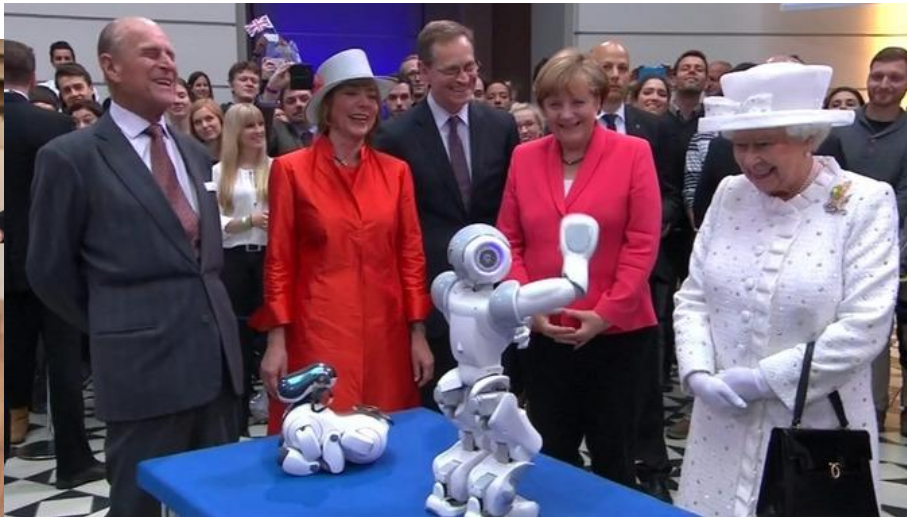


Motivation

- ▶ **Humanoid robots, which can interact with persons like human beings, attract and fascinate people likewise**
- ▶ **Potential applications are**
 - support of elderly people
 - reception of customers, e.g., in hotels or shops
 - healthcare assistant in hospitals
 - greeting the Queen



Source: Aldebaran Robotics



Source: RTV



Source: SoftBank/Aldebaran Robotics

Motivation

- ▶ **Speech is the most natural and effective way for humans to communicate with a humanoid robot (as with humans)**
- ▶ **For a natural human-robot interaction (HRI), the robot should**
 - look to the person it is talking to
 - source detection, localization and tracking
 - understand the content of the conversation
 - automatic speech recognition (ASR)
 - respond in a reasonable manner
 - speech-dialog systems
 - show human-like gestures
 - robot gesturing



Source: Polly Braden

- ➔ **What are the specific challenges for *robot audition*?**
- ➔ **Which concepts are promising to tackle these challenges?**

Challenges in Robot Audition

- ▶ **Distorted recordings**
 - background and interfering noise
- ▶ **Room reverberation**
 - no close distance between robot and person
- ▶ **Robot ego-noise**
 - caused by motors (actuators) and cooling fan
- ▶ **Acoustic feedback for a speaking robot**
 - loudspeakers with nonlinear transmission characteristic
- ▶ **Robot (head) and/or person are moving**
 - speaker localization and tracking
- ▶ **Microphone and/or camera signals possibly unsynchronized**



Focus

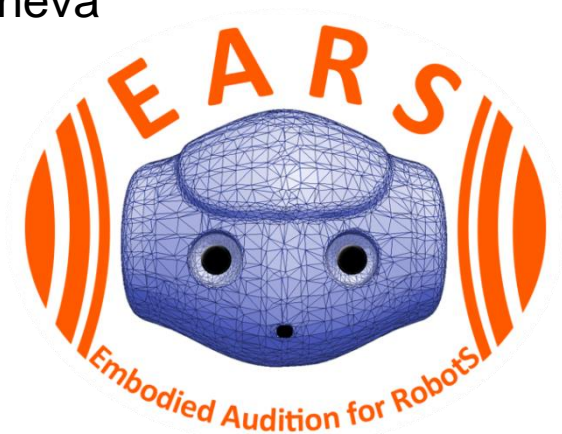
▶ ***This tutorial provides an overview and details the challenges, concepts, and state of the art developments in embodied audition for robots.***

▶ **Main topics**

- algorithms for ego-noise reduction and signal extraction
- beamforming and microphone array design
- speaker localization and tracking
- acoustic scene mapping
- audio and video modalities for speaker detection and tracking

The EARS Project

- ▶ **Tutorial presents many results developed with the EU-funded project *Embodied Audition for Robots (EARS)***
 - **Main focus:** Developing the fundamentals for a natural dialogue between humans and robots in *adverse acoustical environments*.
 - **Runtime:** 3 years (Jan. 2014 – Dec. 2016)
 - **6 Partners:**
 - FAU Erlangen-Nuremberg (Project Coordinator)
 - Imperial College London
 - Ben-Gurion University of the Negev, Beer Sheva
 - Humboldt University Berlin
 - INRIA Grenoble
 - Aldebaran Robotics SA, Paris



Program Overview

- 9:00 **Embodied Audition for Robots: Introduction & Overview**
H. Löllmann, C. Evers, R. Horaud
- 9:15 **Audio-Vision Fusion – Challenges and Concepts**
R. Horaud
- 9:40 **Acoustic Signal Enhancement**
H. Löllmann
- 10:05 **Acoustic Source Localisation for Environment Awareness**
C. Evers
- 10:30 **Coffee break**
- 10:45 **Microphone Array Design and Beamforming**
H. Löllmann
- 11:10 **Acoustic Scene Mapping**
C. Evers
- 11:35 **Fusion of Audio and Video Modalities**
R. Horaud
- 12:05 **Summary & Outlook**
H. Löllmann, C. Evers, R. Horaud
- 12:15 **Closing**

