

Introduction to Open Source Robot Audition Software “HARK”

Kazuhiro Nakadai^{1,2}, Hiroshi G. Okuno³,
Toru Takahashi³, Keisuke Nakamura¹,
Takeshi Mizumoto³, Takami Yoshida²,
Takuma Otsuka³, Gökhan Ince¹

¹ Honda Research Institute Japan Co., Ltd.

² Tokyo Institute of Technology

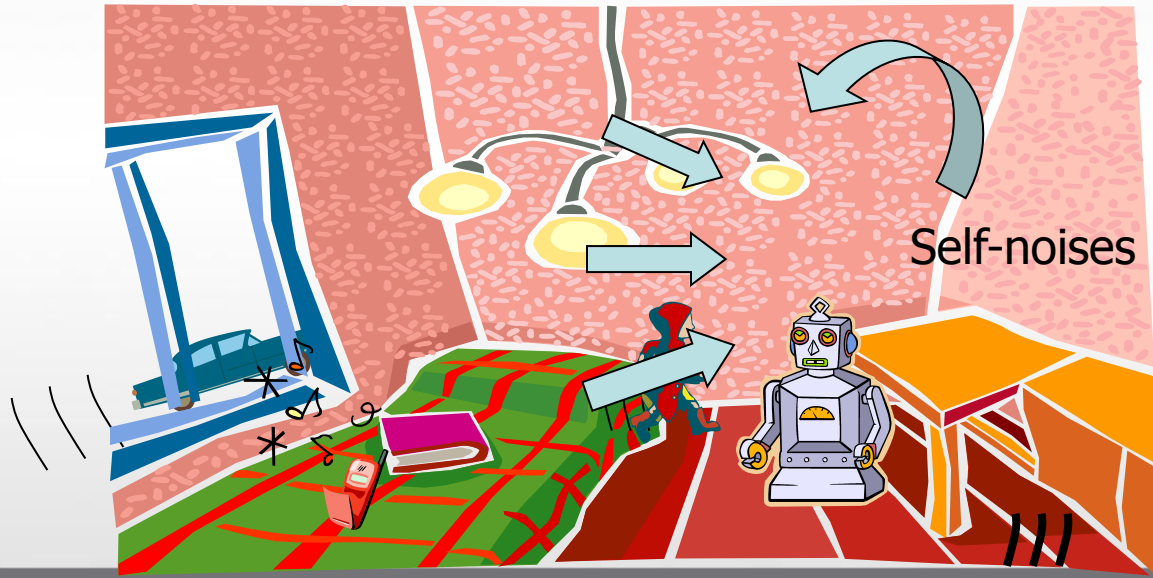
³ Kyoto University

Robot Audition [AAAI 00]

- Not a headset microphone, but *robot's own ears!*
 - Noise-robustness
 - Ego-noise (actuators, self-voice)
 - Environmental sounds
 - Simultaneous speech(barge-in)
 - Cocktail Party Robot
 - Prince Shotoku Robot



- Towards Auditory Scene Analysis



- HRI-JP Audition for Robots with Kyoto University

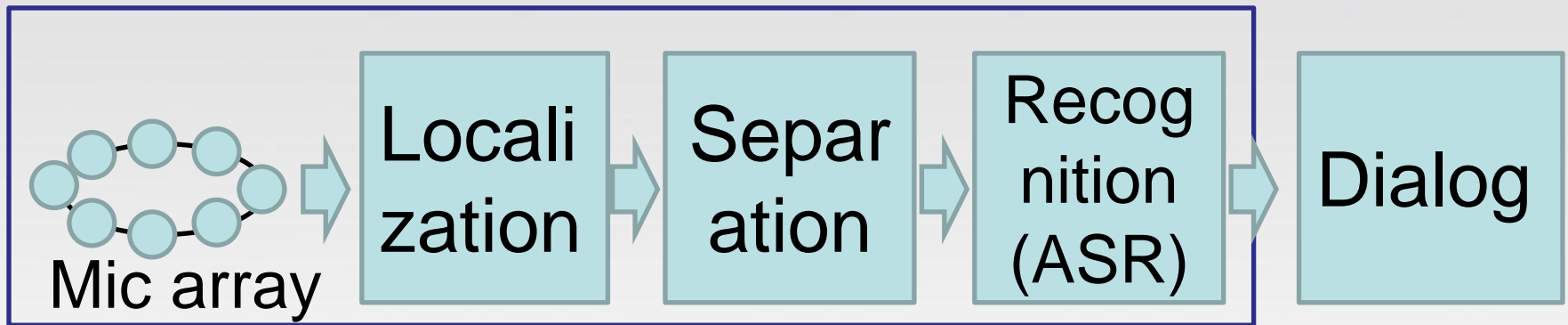


Honda Research Institute Japan **A**udition
for **R**obots with **K**yoto University

hark = listen in old English

Research purpose: Free
(Commercial: Licensing)

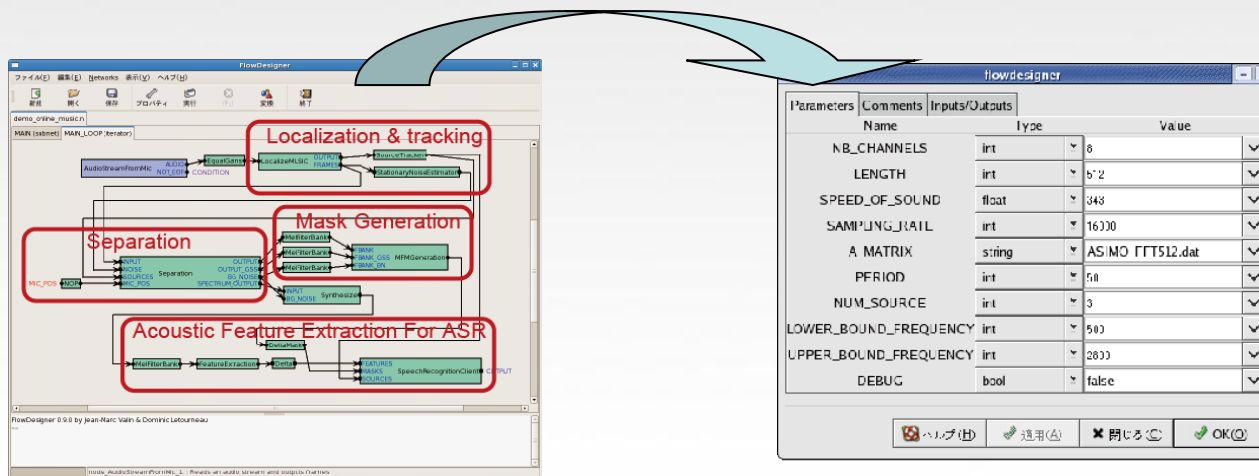
- Apr., 2008 First release
 - <http://winnie.kuis.kyoto-u.ac.jp/HARK>
 - Tutorials in Japan, Korea, France(Humanoids'09)
- Nov., 2010 Major version up to **1.0.0**
 - >50 modules
 - Linux (officially support Ubuntu 10.04 and higher)



- The following functions are provided by using a **robot-embedded microphone array** even in a **highly-noisy environment** such as simultaneous speeches
 - Sound Source Localization (SSL)
 - Sound Source Separation (SSS)
 - Automatic Speech Recognition of each separated speech

Features in HARK (1)

- **Modular architecture based on Flowdesigner [Cote 04]**
 - GUI programming environment (modules written in C++)
 - Suitable for frame-based processing like audio and vision
 - No overhead in module communication



a) Module network

b) Property setting window

Example of robot audition system with HARK

- **Support many multi-channel sound input devices**
 - ALSA based sound devices
 - TED TD-USB devices
 - SiF RASP series
 - * Can use any layout and any number of microphones

- **Advanced signal processing technologies** which take **dynamic environments** into account
 - MUSIC, GHDSS, HRLE, MFT-ASR etc.
- **Easy to install**
 - Just use conventional package management tool “apt-get” !
- **Rich documentation**
 - Manual and cookbook over 300 pages in Japanese and English
- **High interoperability with robot middleware**
 - **HARK-ROS**: seamless integration of HARK and ROS
 - **HARK-MUSIC**: music related functions like beat tracking
 - **HARK-Binaural**: binaural sound localization
 - Wrapper for OpenRTM (release is under consideration)
 - Developing Windows version of HARK (possibly in this year)

Referee for Rock-Paper-Scissors Sound Game



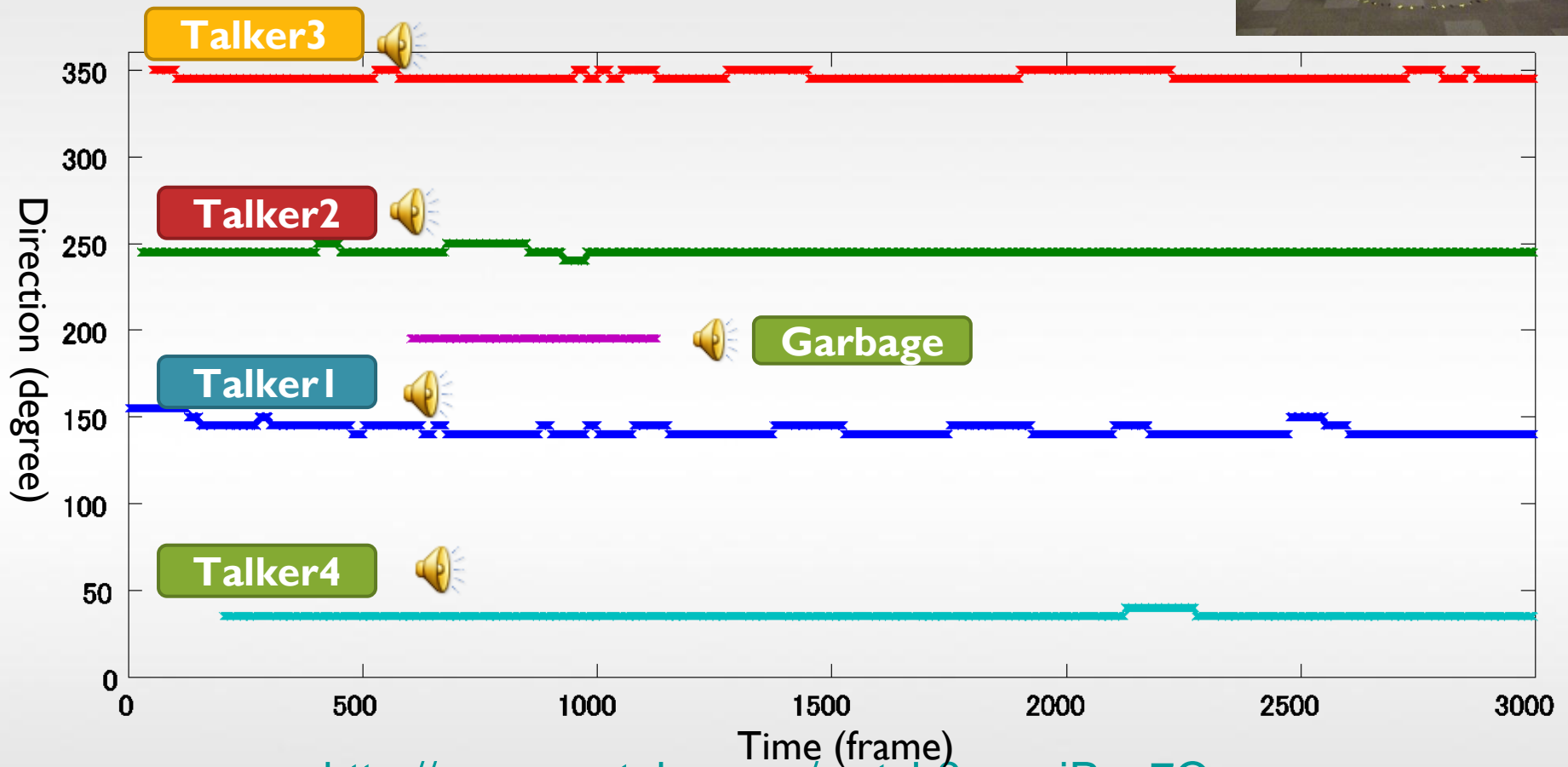
Four Simultaneous Speech Recognition (Meal Order Taking Task)

Experiment with Texai

- Reverberant conference room (RT > 1s), around 20m x 10m.



Recorded



<http://www.youtube.com/watch?v=xpjPun7Owxg>

Sound archive and reconstruction

Reconstruction using sound location and recognition result



Scene

Reconstruction of sound with specific directions interactively



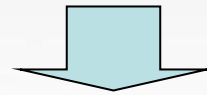
Sound Lifelog : Visualization for Sound Archives





- Sound source localization with Generalized EigenValue Decomposition (GEVD)
- Sound source identification with Hierarchical GMM

- Introduced open source robot audition software HARK
 - Can build a highly noise-robust real-time system using microphone array processing
 - GUI-programming and customization
 - Rich documentation



- Contribution to robotics and other research fields
- Just download and use it.

“Using is believing !”

- Special thanks to
 - HARK team (Okuno Lab., Kyoto Univ. and HRI-JP)
 - Dr. Shunichi Yamamoto, Honda R&D
 - Dr. Jean-Marc Valin, CSIRO
- For more information on “Robot Audition”,
<http://winnie.kuis.kyoto-u.ac.jp/HARK/>
<http://winnie.kuis.kyoto-u.ac.jp/SIG/>