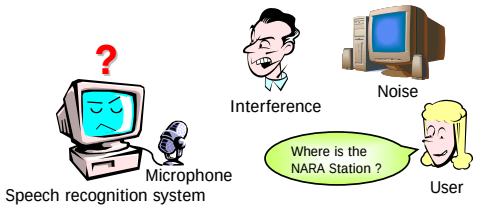
Fast-Convergence Blind Source Separation Combining ICA and Beamforming

December 22, 2004

Tsuyoki NISHIKAWA

Speech and Acoustics Laboratory

Hands-Free Speech Recognition System



Speech recognition performance significantly degrades.

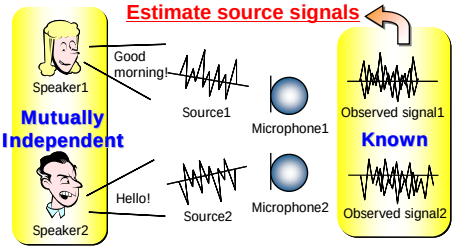
Conventional Approach

- **Goal**
  - **Realization of a high quality hands-free speech recognition system**

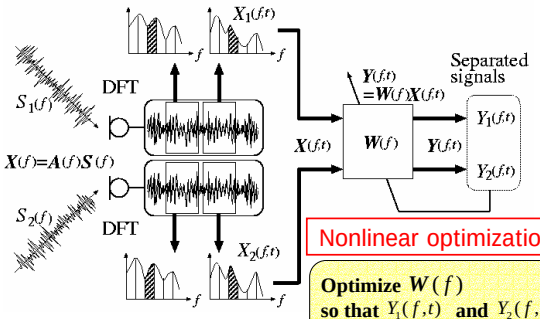
- Microphone array
  - Receiver which consists of multiple elements
  - To enhance target speech or reduce interference
- Problem of microphone array processing
  - A priori information is required.
    - Directions of arrival of the sound sources
    - Breaks of target speech for filter adaptation

Blind Source Separation (BSS) Based on Independent Component Analysis (ICA)

- BSS is the approach taken to estimate source signals only from mixed signals.
- ICA (Comon, 1994) using independence between the source signals is mainly used.



Frequency-Domain ICA



Nonlinear optimization

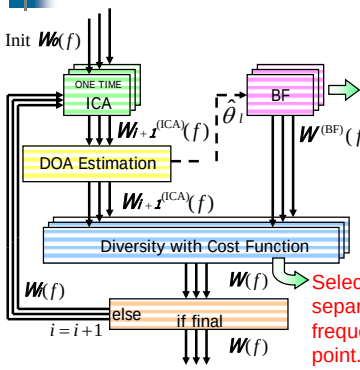
Optimize $W(f)$ so that $Y_1(f,t)$ and $Y_2(f,t)$ are mutually independent.

Problems and Our Approach

- Conventional ICA-based BSS must perform a nonlinear optimization.
  - ⇒ Low-convergence problem
  - ⇒ Trap on Local-minizer

- To combine ICA and beamforming (BF)
- Temporal alternation between ICA and BF through optimization can realize fast- and high-convergence.

Proposed Algorithm Overview
Combination of ICA and BF



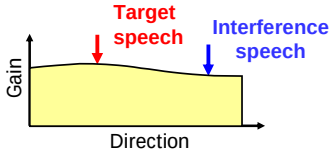
Null Beamforming to reduce interference from direct-path

Select the most suitable separation matrix in each frequency and each iteration point.

Direction of Arrival (DOA) Estimation

(1) To find nulls in the directivity pattern every frequency bin.

(2) To determine DOAs by the inter-frequency-averaged null direction.

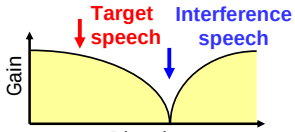


Directivity pattern of separation filter estimated by ICA

Null Beamforming

(1) Null beamforming is approximately optimal for signal separation when the effect of the reverberation is negligible

(2) However, null beamforming still has the advantage that there is no difficulty with respect to the slow convergence of optimization



Directivity pattern of separation filter by BF

Cost Function for Diversity

Diversity Criterion between ICA and BF

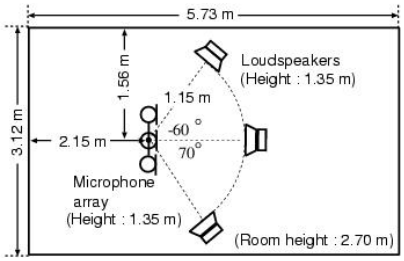
$$W_i(f) = \begin{cases} W^{(ICA)}(f) & \left(J^{(ICA)}(f) \le J^{(BF)}(f)\right) \\ W^{(BF)}(f) & \left(J^{(ICA)}(f) > J^{(BF)}(f)\right) \end{cases}$$

$$J^{(ICA)}(f) = \frac{\left|\left\langle Y_1^{(ICA)}(f,t) \cdot Y_2^{(ICA)}(f,t)^* \right\rangle_t\right|}{\left\langle \left|Y_1^{(ICA)}(f,t)\right|^2 \right\rangle_t^{\frac{1}{2}} \cdot \left\langle \left|Y_2^{(ICA)}(f,t)\right|^2 \right\rangle_t^{\frac{1}{2}}}$$

$$J^{(BF)}(f) = \frac{\left|\left\langle Y_1^{(BF)}(f,t) \cdot Y_2^{(BF)}(f,t)^* \right\rangle_t\right|}{\left\langle \left|Y_1^{(BF)}(f,t)\right|^2 \right\rangle_t^{\frac{1}{2}} \cdot \left\langle \left|Y_2^{(BF)}(f,t)\right|^2 \right\rangle_t^{\frac{1}{2}}}$$

$J(f) \to 0$
⇓
Good separation

Experimental Condition

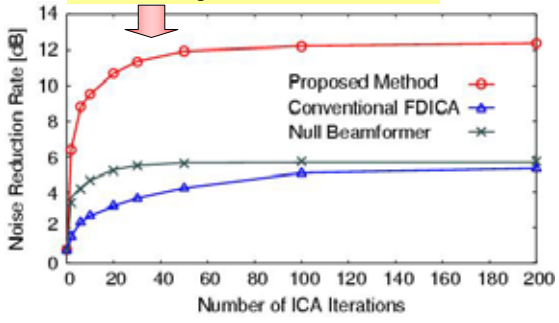


Reverberation time: 300 ms

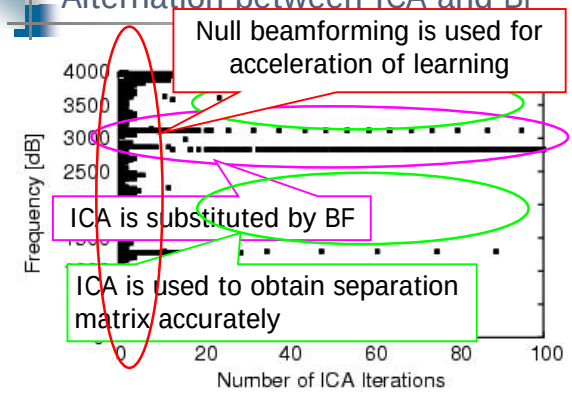
- Sound source: 12 combinations (2-male and 2-female speech)
- Evaluation: Noise Reduction Rate (Output SNR – Input SNR [dB])

Separation Results

Fast convergence can be achieved



Alternation between ICA and BF



Demonstration of BSS

- Blind separation of two speech signals
- Hardware
  - Laptop PC: Sony VAIO type U
    - CPU: Pentium M 1 GHz
    - Memory: 512 MB
  - Microphone: Panasonic electret condenser microphone (Ominidirectional)
- Software
  - Proposed fast-convergence BSS

Conclusion

- NRR improvement by proposed method
- Fast convergence
- Automatic diversity between ICA and BF
  - BF is used for acceleration of learning at early times in the iteration because BF is rough approximation of unmixing matrix.
  - ICA is used after the early part because ICA can update the unmixing matrix more accurately.
  - ICA is substituted by BF through whole iteration points at low-convergence frequency region.