Title	３次元空間を考慮した多数マイクロホンによる雑音除去法に関する研究
Author(s)	森井，大貴
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1140
Rights	
Description	Supervisor:赤木 正人，情報科学研究科，修士

NOISE REDUCTION BY MULTI-MICROPHONE CONSIDERING 3-D SPACE

Masataka Morii

School of Information Science,
Japan Advanced Institute of Science and Technology

February 13, 1998

Keywords: Noise Reduction,Microphone Array,3-D Space,No-calculated Band.

1 INTRODUCTION

A large number of current speech recognition systems for clean speech show great ability in ideal enviroments which have no noise and reverberation. However,there are many reports, e.g.[4] that performance of these speech recognition systems reduces extremely in real environments because noises and reverberations distort speech.

Therefore, this paper proposes a new method of noise reduction as a front-end prosessor of speech recognition systems to make speech recognition systems robust by reducing distortions of amplitude spectrums caused by noise.

The proposed method expanded noise reduction model[1] proposed by Akagi etal, and two paired microphones are used in the horizontal and perpendicular direction. This method reduce noise coming from a point of 3-D space.

2 ALGORITHM DESCRIPTION

The new noise reduction method proposed here uses a microphone array constructed with two microphone pairs as an input device. Here, the microphone of up and down is the main-microphone pair, the microphone pair of left and right is the sub-microphone pair and the center microphone is support.(See Fig.1)

2.1 FORMULATION

Let us assume that there are one microphone pair in a noisy enviroment as shown Fig.2.

s(t) is a signal coming from the front, and n(t) is a noise coming from the direction such that the difference of arrival time between two maicrophones locating at both ends of a 3ch. microphone array is 2δ. The arrival sound at each microphone is assumed as follow,

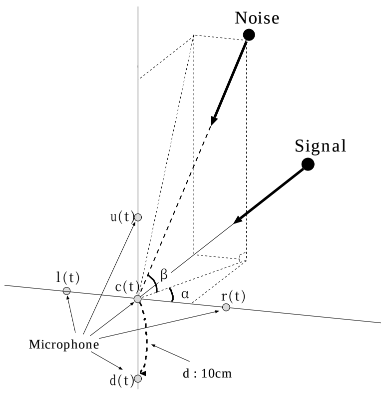

Figure 1: The structure of microphone array

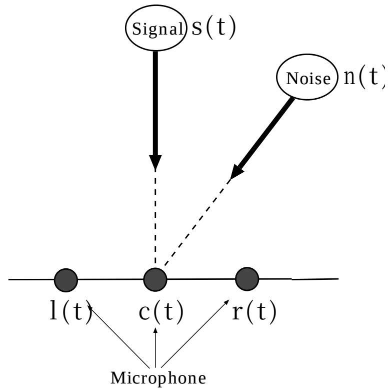

Figure 2: The relation between mic and sounds

$$left: l(t) = s(t) + n(t - \delta) \tag{1}$$

$$center: c(t) = s(t) + n(t) \tag{2}$$

$$right: r(t) = s(t) + n(t + \delta) \tag{3}$$

where l(t),c(t),r(t) are signals received by left,center and right microphones.

Then, when detection of a noise direction is 2τ, equation(1) and (3) are shifted $\pm\tau$ in time. The result $g_{lr}(t)$ is then

$$g_{lr}(t) = \frac{\{l(t+\tau) - l(t-\tau)\} - \{r(t+\tau) - r(t-\tau)\}}{4} \tag{4}$$

and $g_{lr}(t)$ is transformed to $G_{lr}(\omega)$ by the short term Fourier transformation(STFT).

$$G_{lr}(\omega) = \sin\omega\tau N(\omega)\sin\omega\delta \tag{5}$$

Estimating noise direction 2δ and assuming $\tau = \delta$, the estimated noise $\tilde{N}(\omega)$ can be represented as

$$\tilde{N}(\omega) = \begin{cases} G_{lr}(\omega)/\sin^2\omega\delta \ , & |\sin\omega\delta| > \varepsilon \\ G_{lr}(\omega) \ , & |\sin\omega\delta| \leq \varepsilon \end{cases} \tag{6}$$

where $\varepsilon \approx 0$.

The above process applies each microphone pair.

Since $\tilde{N}(\omega)$ becomes infinite when $\omega\delta = n\pi$, inaccurate frequency band exist as

$$f = n/2\delta, \quad n = 1, 2, \cdots \ [Hz] \tag{7}$$

Therefore, compensating in accurate frequency bands by other microphone pairs, $\tilde{N}(\omega)$ can be estimated as follow;

$$\tilde{N}(\omega) = \begin{cases} G_{du}(\omega)/\sin^2\omega\delta_{du}, & |\sin\omega\delta_{du}| > \varepsilon_1 \\ G_{lr}(\omega)/\sin^2\omega\delta_{lr}, & |\sin\omega\delta_{du}| \leq \varepsilon_1, \\ & and\, |\sin\omega\delta_{lr}| > \varepsilon_2 \\ G_{du}(\omega), & |\sin\omega\delta_{lr}| \leq \varepsilon_2 \end{cases} \tag{8}$$

where ε_1, ε_2 is a certain small value.

Then, the objective signal is given by subtracting the estimated noise from the signal received by the main-microphone pair.

Noise direction estimation is done by crosscorrelation.

3 EXPERIMENTS AND RESULTS

Two experiments were conducted for evaluation. One is simulation experiments using operating on synthetic signals, and the other is real environment experiments using real signals that presented by speaker in asoundproof room(reverberation time:about 50 ms).

3.1 DATA AND CONDITION

Let us assume that speech (ATR,mht14348/bunri/) come from the front and noise (Sweep Tone:start frequency 1 kHz, end frequency 6 kHz, time 48000 point, Band Noise:center frequency 3.4kHz, bandwidth 3kHz, duration time 48000point) come from direction shown in Table.1. Sampling frequency is 48 kHz.

Table 1: Arrival direction of noise

Simulation Experiments	Sweep Tone : Azimuth (right 30°), Angle of elevation (up 30°)	Data 1
	Sweep Tone : Azimuth (right 60°), Angle of elevation (up 30°)	Data 2
	Band Noise : Azimuth (right 30°),Angle of elevation (up 30°)	Data 3
Real Environment Experiments	Sweep Tone : Azimuth (right 45°), Angle of elevation (up 20°)	Data 4
	Band Noise : Azimuth (right 45°), Angle of elevation (up 20°)	Data 5

Data2 is adopted for performance test of this method when main-microphone and sub-microphone pairs have equal arrival time difference of noise .

Parameters are set up shown in Table.2.

3.2 ARRIVAL TIME DIFFERENCE OF NOISE

The arrival time difference of noise is calculated as

$$\begin{aligned} main-microphone\ pair\ &:\ \delta_{ud}=(2d/c)\cdot\sin\beta \\ sub-microphone\ pair\ &:\ \delta_{lr}=(2d/c)\cdot\cos\alpha\cdot\cos\beta \end{aligned} \tag{9}$$

α:Azimuth, β:Angle of elevation, c:Speed of sound, $2d$:Space of microphone pair (See Fig.1)

The arrival time difference and inaccurate frequency band are shown in Table.3.

Table 2: Parameter Establishment

Parameter	Establishment Volue (Simulation)	Establishment Volue (Real Environment)
Sampling frequency	48kHz	the same as left
Frame Length	2048 point（Sweep Tone） 1024 point（Band Noise）	4096 point
Frame Period	1024 point（Sweep Tone） 512 point（Band Noise）	2048 point
Window Types	Hamming（Direction Estimation） Triangle（Noise Reduction）	the same as left
ε_1	0.7（Sweep Tone） 0.05（Band Noise）	0.9
ε_2	0.5（Sweep Tone） 0.05（Band Noise）	0.7

Table 3: Arrival time difference of noise,inaccurate frequency band

Arrival direction of noise	Arrival time difference （main-pair,sub-pair）	Inaccurate frequency band （main-pair,sub-pair）
$\alpha : 30^\circ$ to the right,$\beta : 30^\circ$ to the up	7point、11point	3.4*n[kHz]、2.2*n[kHz]
$\alpha : 60^\circ$ to the right,$\beta : 30^\circ$ to the up	7point、7point	3.4*n[kHz]、3.4*n[kHz]
$\alpha : 45^\circ$ to the right,$\beta : 20^\circ$ to the up	5point、9point	4.8*n[kHz]、2.6*n[kHz]

3.3 RESULTS

For evaluation, the following signal-to-noise ration(SNR) is adopted.

$$SNR = 10log_{10}\frac{\sum_n s^2(t_n)}{\sum_n \{s(t_n) - \tilde{s}(t_n)\}^2} \qquad (dB) \tag{10}$$

where $s(t_n)$ is an original sound wave and $\tilde{s}(t_n)$ is a noise-reduced sound wave.

Results of SNR are shown in Table.4 and 5.

Results of power spectrum of Data 1 and 2 are shown in Fig.3 .

Table 4: SNR（simulation experiments）

	Data 1	Data 2	Data 3
Noise-added speech wave	-10.45 (dB)	-10.45 (dB)	-3.61 (dB)
Noise-reduced speech wave of main-pair	-2.77 (dB)	-2.77 (dB)	11.27 (dB)
Noise-reduced speech wave of sub-pair	-4.97 (dB)	-2.77 (dB)	9.54 (dB)
Noise-reduced speech wave of this method	8.61 (dB)	-2.15 (dB)	18.84 (dB)

Table 5: SNR（Real environment experiments）

	Data 4	Data 5
Noise-added speech wave	-1.83 (dB)	-1.86 (dB)
Noise-reduced speech wave of main-pair	3.07 (dB)	2.05 (dB)
Noise-reduced speech wave of sub-pair	3.54 (dB)	3.68 (dB)
Noise-reduced speech wave of this method	4.64 (dB)	3.68 (dB)

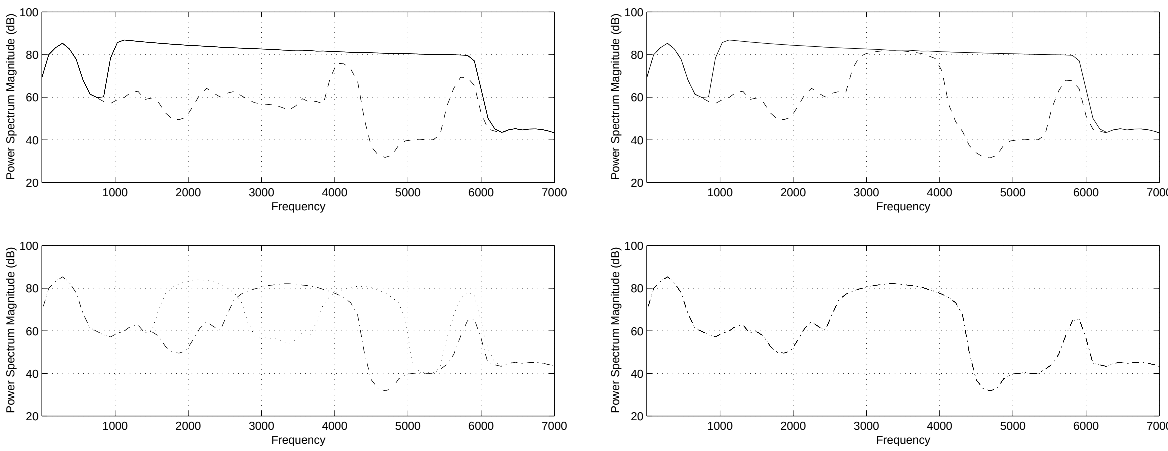

Figure 3: Power Spectrum（Fig.left：Result of Data1, Fig.right：Result of Data2）Solid line：Noise-added speech wave, Broken line：Noise-reduced speech wave of this method, Chain：Noise-reduced speech wave of main-pair, Dotted line：Noise-reduced speech wave of sub-pair

4 CONCLUSION

This paper proposed a method that reduces noise coming from a point of 3-D space with main and sub-microphone pairs.

SNRs can increase by 10 to 20 dB in simulations.

References

[1] Akagi.M, Mizuma.M:" Noise Reduction by Paired Microphones, "IEICE. SP,(1997.7)

[2] L.J.Griffiths and C.W.Jim:" An alternative approach to linearly constrai ned adaptive beamforming, "IEEE Trans. AP-30,1,(1982)

[3] Kaneda.Y. and Ohga.J: " Adaptive microphonearray system for noise reduction, "IEEE trans.ASSP,34,6,1391-1400,(1986)

[4] M.Tobita,N.Sugamura:" Effects of enviromental conditions on speech recognition, " J.Acoust.Soc.Jpn(J)., vol.51, no.4, pp.331-335,(1995)