

2マイクによる高指向性デジタル信号処理を使った音声認識用車載マイクシステムの開発

In-vehicle Super-directive Microphone System for Speech Recognition
Using Nonlinear Signal Processing with Two Microphones

鹿野 俊哉*
Toshiya SHIKANO

渡部 純一**
Junichi WATANABE

松尾 直司***
Naoshi MATSUO

要 旨

北米での車載音声認識用に、車内の高雑音下における認識率向上を目的とする、デジタル信号処理を用いた高指向性マイクシステムを開発した。

助手席の声などの雑音に対し指向性雑音抑制率の高い新規高指向性アルゴリズムを採用し、これをエアコン、ウインドウの開け閉めなどの車内環境に適応するために、無指向性マイク素子、アナログハイパスフィルタを組み合わせたシステムを開発した。これにより、各種雑音環境下で認識率を向上させることができた。

ABSTRACT

A super-directive microphone system using digital signal processing has been developed for in-vehicle speech recognition systems for the North American market, in order to increase the recognition accuracy of these systems in high noise environments.

The developed microphone system uses a new super-directive algorithm with a high directional noise reduction rate that functions to reduce noise, such as passenger utterances. To enable adaptation of the algorithm to different in-vehicle noise environments, such as while operating the air conditioner and opening/closing the windows, the system also incorporates an omni-directional microphone and an analog high pass filter. The use of the new system increases recognition accuracy in a variety of noise environments.

1. まえがき

Honda では、北米音声認識ナビゲーションシステムを2003年モデル ACCORD から順次適用してきた。当社がおこなったユーザアンケート調査において、音声認識システムのユーザの使用頻度は高く、操作性も高い評価が得られた。一方、当社のテストでは通常のロードノイズにおいて十分な認識率が得られているが、ユーザの感じる認識率においてはさらなる向上を求める声が多く、実際の使われ方においても改善を図る必要があった。このため、認識を阻害する要因の中で、車内において想定される特に強い雑音に注目し、デジタル信号処理を用いた高指向性マイクシステムを車載音声認識用に開発した。

* 四輪開発センター

** 富士通株式会社

*** 株式会社富士通研究所

2. 車内雑音とその対策

2.1. 車内雑音

車内の雑音として以下の音源がある。

- ・ロードノイズ
- ・窓、サンルーフを開けた状態での風切り音
- ・エアコンの騒音、およびエアコン風のマイクへの直撃
- ・助手席など同乗者の話し声

これらを物理特性で分類すると、ロードノイズ、窓、サンルーフ開け時の風切り音、エアコンの騒音などは定常雑音に類する。助手席など同乗者の話し声は非定常(突発的な)雑音に類する。また、エアコン風のマイクへの直撃は、指向性マイクの構造上の特性として、風圧による振動板の壁へ

の張り付きによる異常信号のノイズである。これらをまとめると、

- ・ 定常雑音
- ・ 非定常雑音
- ・ 風圧による異常信号

に分類され、これらのノイズに有効なマイクシステムを検討した。

2.2. 非定常雑音と高指向性デジタル信号処理

車内の非定常雑音の代表として、同乗者の話し声がある。通学の送迎やショッピングモールでの買い物、レストランでの食事、家族旅行など、複数人乗車の機会が想定されるが、こうした場合、ナビなどの情報機器の操作も頻度が高い。

また、音声認識処理は、声の特徴を元に演算されるので、雑音の中では同乗者の声が認識に強く影響する特性を持っている。ドライバの声の認識を阻害させないための同乗者の声の抑制率は 20-30dB 必要であり、ホワイトノイズなどの雑音に対する認識限界 S/N 比が 10dB 弱であるのに対し、難易度が高い。

非定常雑音抑制の手段として、複数のマイクを用いた音源の到来方向によって必要な音声を分別する、遅延和アレイ、適応型ビームフォーム等の指向性制御⁽¹⁾の方法と、複数音源間の出力の独立性を想定した独立成分分析⁽²⁾を用いて分別する方法がある。しかし、指向性制御はマイクの数

少ないと抑制率を高くできない傾向にあり、独立成分分析は演算負荷が高く車載用としては現実的でない。したがって、二つのマイクで目標とする高い抑制率が得られる、新規高指向性デジタル信号処理⁽³⁾を選択した。

以下に新規高指向性アルゴリズムの原理を Fig. 1 に示す。

Fig. 1(a) のように、ドライバ発声の音波が二つのマイクに到達する際、到来方向によって一定の時間差を生ずる。この時間差を周波数軸変換してグラフにすると、Fig. 1(b) のように、ドライバの発声は右上がりの直線近傍にプロットされ、助手席側の声と明確に分離される。

新規高指向性アルゴリズムの効果を Fig. 2 に示す。ドライバ、助手席同乗者の発声を交互におこなった時の音圧レベルの時系列を Fig. 2(a) に示す。これを新規高指向性アルゴリズムで処理すると、Fig. 2(b) のように助手席同乗者の発声だけ選択的に強く抑制され、ドライバの発声だけ切り出すことができる。指向性制御ではあるが、人の聴覚に備わるカクテルパーティ効果⁽⁴⁾のように、あたかも雑踏の中からドライバの発声を聞き分けることを可能にする。

2.3. 風圧による異常信号と無指向性マイク素子

一般的に、車載の音声認識用、ハンズフリー電話用のマイク素子は、ドライバの声を選択的に拾いたいため指向性マイク素子が使われている。しかし、指向性マイク素子⁽⁵⁾、⁽⁶⁾は、Fig. 3(a) のようにエアコンの風をマイクに向けると音声信号に異常な信号を出力する。これは、Fig. 4(a) のように指向性を出すため振動板前面の音圧とマイク背面に回りこんだ振動板背面音圧とが干渉する構造であることに起因する。すなわち、指向性マイク素子は振動板背圧が開放さ

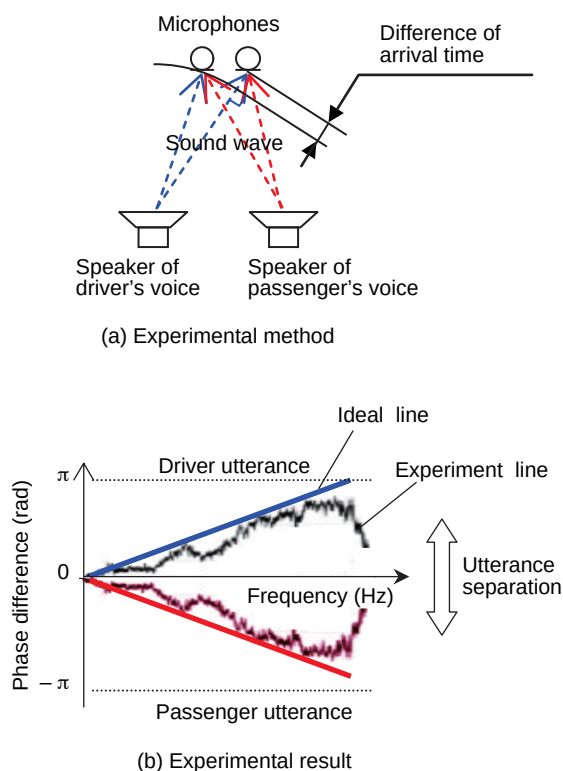


Fig. 1 Principle of channel separation between driver utterance and passenger utterance

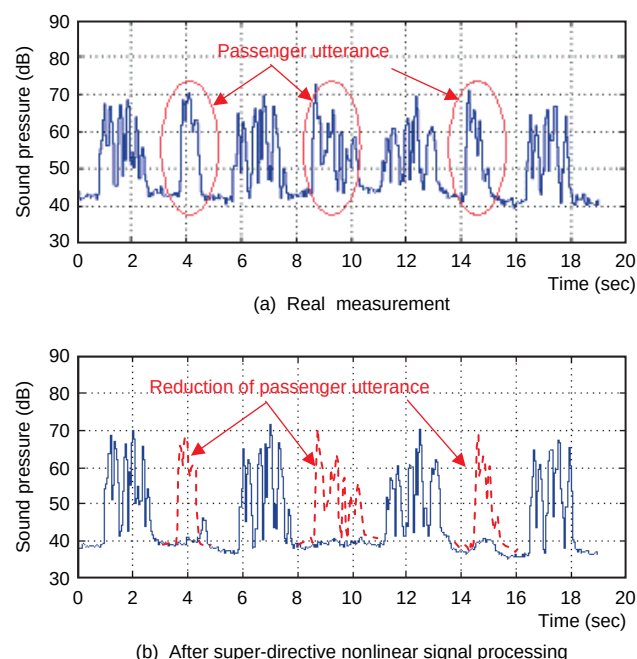


Fig. 2 Effect of channel separation

れており、振動板前面への風圧や、マイク素子にかかる振動などで音圧以外でも容易に振動板が変位する。このため振動板が前後の壁に張り付き、異常音を発する。これに対し Fig. 4(b) に示す無指向性マイク素子⁽⁷⁾は、振動板の大きな変位を背圧で押さえ込むため風圧等で容易に壁に張り付くことはなく、Fig. 3(b) に示すように信号出力に異常は現れない。

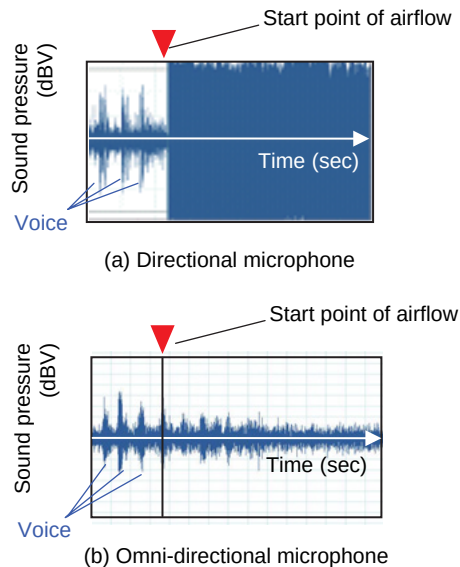


Fig. 3 Comparison of acoustic waves

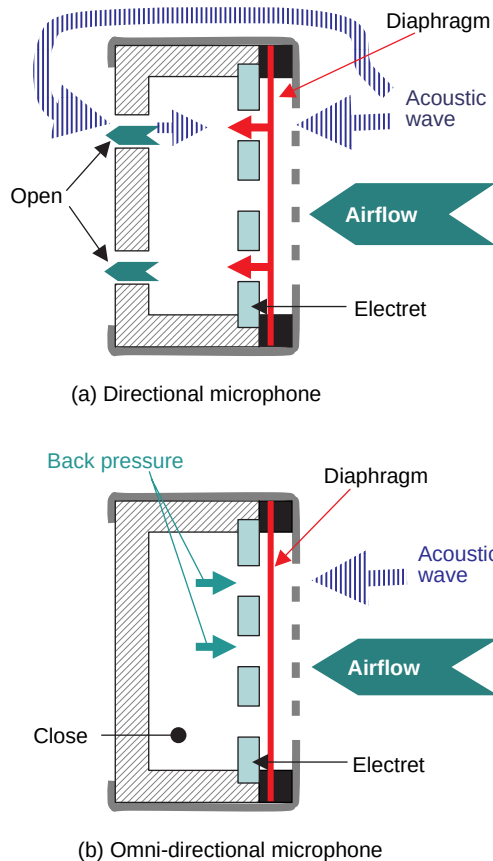


Fig. 4 Comparison of microphone structures

車内でマイクの設置されているルーフ前部は、Fig. 5 のようにダッシュボードのベント口より吹き付けるエアコン風を受けるレイアウトとなっている。したがって、風圧に影響されない振動板背面密封型(厳密には大気圧をキャンセルする差圧を抜く穴がある)の無指向性マイク素子が耐風圧には優れている。2個のマイク素子によるデジタル信号処理でドライバへの指向性を得て、指向性マイク素子以上のドライバ音声の品質が確保されることにより、風圧に強い無指向性マイク素子を初めて採用することができた。



Fig. 5 Direct airflow from air conditioner

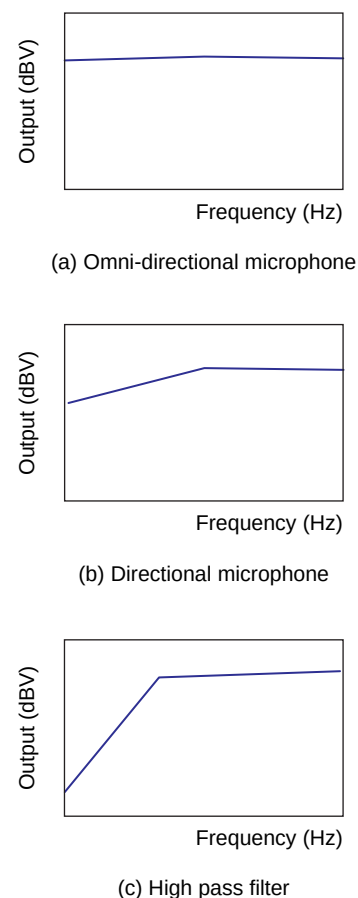


Fig. 6 Frequency response of microphone and high pass filter

2.4. 定常雑音とアナログハイパスフィルタ

定常雑音抑制の信号処理は、スペクトラルサブスクリプションなどさまざまな手法が考案されているが、雑音抑制率を高めると原音に対して処理後の音のひずみが大きくなるという傾向がある。音のひずみが大きくなると、後段の音声認識演算の中の音素を選択するための音響モデルとアンマッチが生じ、認識率が低下する懸念がある。

また、無指向性マイク素子の周波数特性 Fig. 6(a) は、指向性マイク素子の低域減衰特性 Fig. 6(b) と異なり、低域に集中する車内特有の定常雑音を減衰することができない。

そこで、強い抑制のデジタル信号処理を避け、かつ無指向性マイク素子の課題を解決する定常雑音抑制の手段として、アナログデジタル変換処理前に、Fig. 6(c) のように指向性マイク素子の低域減衰特性よりも急峻なアナログハイパスフィルタを設けた。ハイパスフィルタのクロスオーバー周波数は、使用する音声認識アプリケーションに必要な周波数以下に設定した。

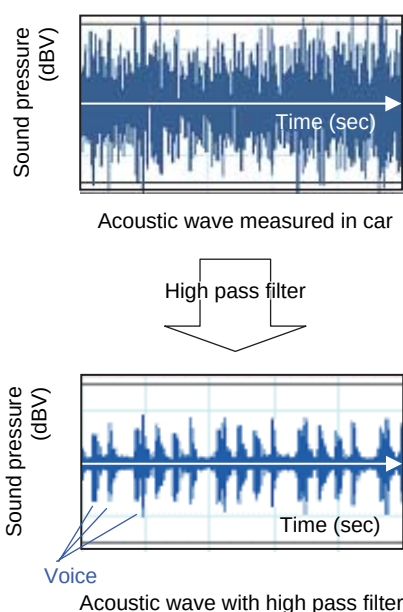


Fig. 7 Effect of high pass filter with omni-directional microphone

音声認識演算時も数十の周波数バンドでフィルタリングするが、デジタルに変換される前であるため、デジタル信号の情報量を高め、処理後の信号の品質を向上することができる。効果は Fig. 7 に示すように、低周波の雑音成分が抑制されて発声信号の S/N 比を改善する。

3. 車内雑音抑制システム

2.2. 節から 2.4. 節の、各項目の車内雑音環境に対応するため、新規高指向性アルゴリズムに、無指向性マイク素子の採用と、アナログハイパスフィルタを組み合わせ、高い雑音抑制性能を得る方法を開発した。開発した雑音抑制マイクシステムの概観を Fig. 8 に示す。

処理の流れを Fig. 9 に示す。ドライバの発声は二つのマイク素子で電気信号に変換され、二つの信号それぞれがハイパスフィルタを通り、その後アナログ・デジタル変換される。この二つの信号をデジタル信号処理に入力して高指向性の雑音抑制処理を施す。以上が Microphone unit の中でおこなわれる処理である。これから出力した信号を、カーナビゲーションユニットの中の Speech recognition engine に渡す流れになっている。こうすることにより、現行のカーナビゲーションユニットを変更せずに組み合わせて使用することができた。

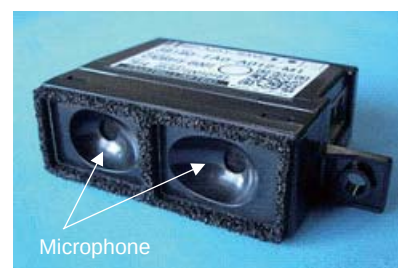


Fig. 8 Digital signal processing microphone unit

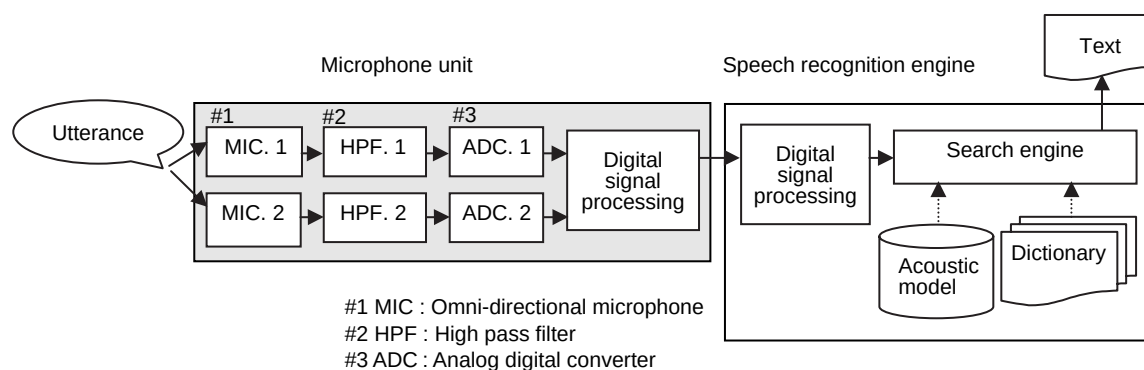


Fig. 9 Block diagram of voice recognition system

4. 結果

4.1. 非定常雑音の抑制効果

実車走行にて、ドライバのみの発話とドライバおよび助手席の同時発話について、新旧マイクの音声認識比較テストをおこなった。助手席の声は、音圧はドライバと同等で、発話区間がドライバの発話と重複し、かつドライバより長くなるようにした。

結果を Fig. 10 に示す。助手席の発話があっても高い認識率を確保することができた。

4.2. エアコン風当たり異常信号の抑制効果

アイドリングにて、エアコンを止めた状態、風量最大で風向き肩方向にした状態、同じく風向きを頭上方向(マイク方向)にした状態について、新旧マイクの音声認識比較テストをおこなった。結果を Fig. 11 に示す。

背面密封型である無指向性マイク素子の採用により、マイク素子にエアコン風が直撃しても高い認識率を確保することができた。また、エアコン風量最大時は、騒音自体も大きい。これは、エアコン噴出し口等での乱流音が車室容積で決まる共鳴音を発生させるため、後述のウインドスロップ現象⁽⁸⁾の一種である。騒音自体が大きいにもかかわらず高い認識率が得られたことは、2.4. 節で述べたアナログハイパスフィルタも有効に働いていることを示している。

4.3. 定常雑音の抑制効果

窓閉め、窓開けの各状態で音声認識テストをおこなった。窓開け走行テストでは、車内騒音が70dBA以上となり、通常の音圧での発話では声が雑音に埋もれS/N比が取れないため、発話の音圧をその時の S/N 比が数 dB となるよう

音圧を上げてテストした。

窓開け時は騒音が大きく会話が困難な状態であったが、Fig. 12 のように窓閉め時とほぼ変わらない認識率を得た。窓開けや、サンルーフ開け時の風切音は、ウインドスロップ現象による音であり、乱流音と車室容積に対するヘルムホルツ共鳴との連成によって発生する。したがって、音の中心周波数は車速によらず一定であり、車載マイクでは共鳴周波数の倍音を含め、約 300Hz 以下で影響を受ける。この共鳴音による雑音は急峻なアナログハイパスフィルタによって抑制されたため、音声認識率の低下を防ぐことができた。

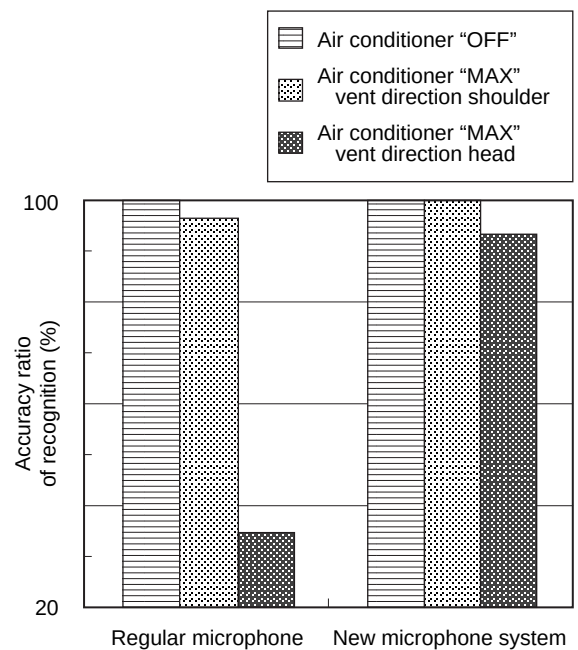


Fig. 11 Recognition accuracy with airflow from air conditioner

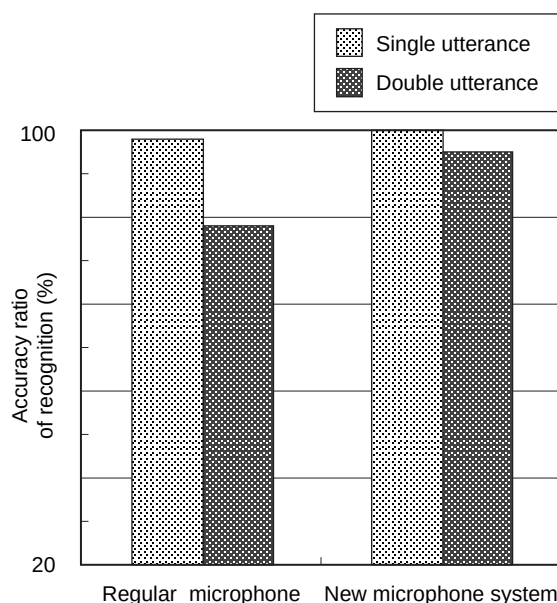


Fig. 10 Recognition accuracy for single utterance and double utterance

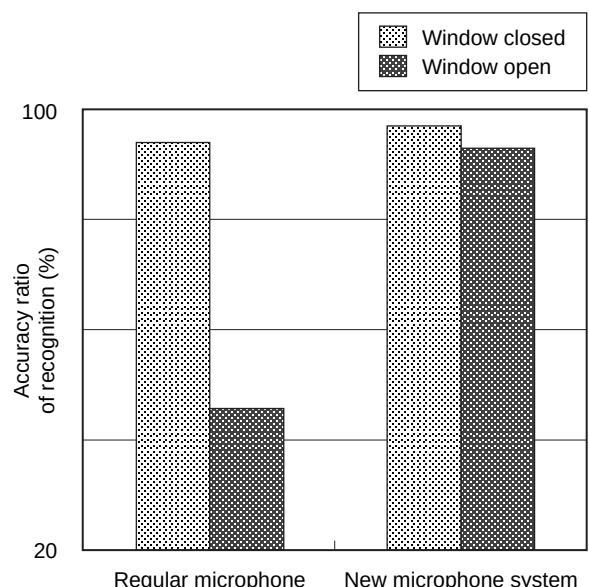


Fig. 12 Recognition accuracy with window closed and window open

5. まとめ

車の実際の使われ方における音声認識率向上を図るため、指向性雑音抑制率の高い新規高指向性アルゴリズム、無指向性マイク素子、アナログハイパスフィルタを組み合わせたマイクシステムを開発した。これらの組み合わせが互いに成立性を補完しあうことにより、通常のロードノイズだけでなく、ウィンドスロップ現象を伴うような窓、サンルーフ開け時の風切り音や、エアコンの騒音、およびエアコン風のマイクへの直撃、助手席など同乗者の話し声に対しても、高い音声認識率を得ることができた。

参考文献

- (1) 金田豊：マイクロホンアレーによる指向性制御，日本音響学会誌 Vol.51, No.5, p.390-394(1995)
- (2) Hyvärinen, A., Karhunen, J., Oja, E.: 独立成分分析，東京電機大学出版局，532p.(2005)
- (3) 古井貞熙：音響・音声工学 p.28(1992)
- (4) 特開 2007-318528
- (5) 中村仁一郎：単一指向性コンデンサ型マイクロホン・放送技術者のためのマイクロホン講座，放送技術，昭和 57 年，5 月号，p.464-468(1982)
- (6) 中村仁一郎：エレクトレット・放送技術者のためのマイクロホン講座，放送技術，昭和57年，11月号，p.1101-1108(1982)
- (7) 中村仁一郎：無指向性コンデンサ型マイクロホン・放送技術者のためのマイクロホン講座，放送技術，昭和 57 年，4 月号，p.364-372(1982)
- (8) 炭谷圭二，前田和弘，一之瀬健一：自動車と流体力学・車体周り流れと空力特性，ながれ，Vol.23, No.6, p.445-454(2004)

■ 著 者 ■



鹿野 俊哉



渡部 純一



松尾 直司