

マイクロホンアレーを用いた雑音下での音声区間検出

正員 金田 豊[†]

Speech Period Detection Using a Microphone Array under Noisy Environments

Yutaka KANEDA[†], Member

あらまし 雑音下で受信した信号の音声区間を検出する技術は、音声認識の前処理などにおいて大変重要である。しかし現在のところ、この要求を十分満足する手法は見出されていない。従来の手法が一つのマイクロホンで受信した1チャンネルの信号を対象とした手法であるのに対し、本論文では、複数のマイクロホンで得られた複数の空間情報を利用した検出方法を提案する。提案手法はまず、適応形マイクロホンアレー AMNOR を用いて十数 dB の雑音抑圧を実現する。更に、この雑音抑圧能力を利用して、AMNOR 信号処理を行った信号と行わなかった信号のレベル差に基づく MLD 法と名づけた判定法により、音声区間の検出を行う。この手法は雑音の性質によらず、一つの固定されたしきい値による一定の判定規準で処理を行うもので、従来困難とされていた非定常雑音に対しても有効な点が大きな特徴となっている。妨害音声を用いて実験を行った結果、本手法は、(1)雑音区間を音声区間と判定する誤りはほぼ発生せず、AMNOR で雑音抑圧を行った結果の短時間 SN 比が 5 dB 以上となる音声区間は的確に検出する、(2)平均 SN 比が -5 dB の条件で、単語音声区間検出の正答率が 96%、また正答の場合の平均検出誤差は 20 ms である、ことが確認された。

1. まえがき

音声信号を波形として見たとき、実際に音声が存在する時間区間(音声区間)と存在しない時間区間が混在している。受信した信号からこの音声区間を検出する技術は大変有用であり、例えば音声認識の前処理として広く利用されている。また、拡声通話系における音響エコーキャンセラ⁽¹⁾の動作時期決定(同時通話の検出)においても利用価値が高い。

しかし、音声信号に雑音が重畳している場合、特に低 SN 比であったり雑音が非定常である場合に音声区間検出を行うことは容易ではない。さまざまな対策方法が提案されているが⁽²⁾、現状では要求を十分に満足する手法は見出されていない。雑音下での音声区間検出能力は、音声認識の実用化を進めていく過程で決定的な影響を与えるので、今後の重要な研究課題であると考えられている⁽³⁾。

従来の音声区間検出法の検討は、主として、一つのマイクロホンで受信した1チャンネルの信号しか手に入らないという前提のもとで行われてきた。しかし、1チャンネル信号において、音声と雑音(以下、不要な音響信号を雑音と総称する)を区別するためには信号の統

計的性質を用いるより手段がなく、その結果、所望音声と類似した雑音(例えば妨害音声)が付加されている場合には、原理的に正確な音声区間の検出は不可能であった。

これに対して本論文では、空間的に配置された複数のマイクロホン(マイクロホンアレー)で受信した複数の信号を用いた検出手法を提案する。この方法では、音の空間的情報を利用して所望音声と雑音を区別して雑音抑圧を行い、更にその結果を用いて音声区間検出を行う。この方法は、信号自体の性質を利用していないことから、雑音が所望音声と類似した非定常信号であっても効果的な音声区間の検出が可能となる。

以下本論文では、2.においてまず、空間情報を利用した雑音抑圧受信方式 AMNOR (Adaptive Microphone-array for Noise Reduction)^{(4),(5)}について簡単に説明する。そして、AMNOR を単なる受信器として利用して音声区間検出を行う場合の効果と問題点について考察する。3.においては、本論文で提案する音声区間検出法を説明し、検出判定のためのしきい値などについて論じる。4.では、提案手法の能力と有効性を検証するために行った実験結果について述べる。

[†] NTT ヒューマンインタフェース研究所, 武蔵野市
NTT Human Interface Laboratories, Musashino-shi, 180 Japan

2. AMNORの音声区間検出への適用と問題点

2.1 適応形マイクロホンアレー (AMNOR)

AMNORは指向特性を環境に応じて適応的に変化させることで効果的な雑音抑圧効果を実現する受音方式である。本節ではまず、その概要を説明する。

図1にAMNORの構成図を示す。音響信号は複数のマイクロホンより構成されるマイクロホンアレーで受音され、受音された各信号はフィルタを通した後、総和をとって出力される。このとき、この受音システムの指向特性はフィルタの特性により決定される。従って、雑音の空間的情報に基づいてフィルタ制御を行えば、周囲の雑音条件に応じた指向特性の適応制御が可能となる。例えば、雑音がある方向から到来する場合には、その方向に対して感度が低く、目的とする方向に感度の高い指向特性が実現できる。

フィルタの制御方法および実現される指向特性の詳細は文献^{(4),(5)}に述べられているが、このAMNORの雑音抑圧効果は、AMNORから見た話者と雑音源の方向が近接した(約±20°以内)場合を除けば、10 dB以上は得られることが確認されている。以下、本論文で述べる音声区間検出法は、AMNORの雑音抑圧効果を利用したものであり、話者と雑音源の方向が近接していないことを前提とする。

2.2 音声区間検出への適用と問題点

本節では、AMNORを単なる受音装置とみなし、その出力を用いて音声区間検出を行うことを考える。

AMNORによる雑音抑圧効果を表す模式図を図2(a)(b)に示す。図の横軸は時間軸を表し、縦軸は信号の各時刻におけるパワーを表している。図2(a)は定常雑音が、図2(b)は非定常雑音がそれぞれ音声に

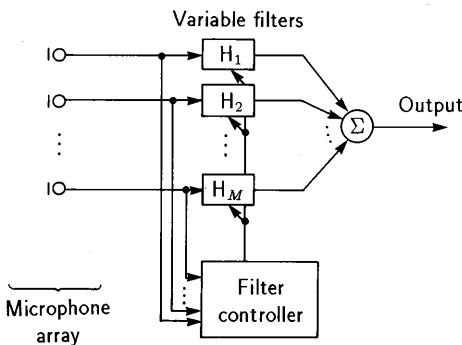


図1 AMNORの構成図
Fig. 1 Configuration of AMNOR.

重畳した場合を表している。図はAMNORによる雑音抑圧を行った信号と、行わない信号を重ね書きしたもので、破線は雑音抑圧を行わない場合の雑音レベルを、ハッチをかけた部分は抑圧された雑音のレベルを表す。また、実線白ぬきの部分は音声信号を表している。音声はある時間区間のみ存在し、AMNOR処理の有無により音声のレベルは変化しない。

図2(a)より、定常雑音の場合には雑音が抑圧されれば音声区間は明確になることが直観的にわかる。これは、定常雑音のパワーが時刻によらず一定値であるという性質によるものである。しかし、図2(b)からもわかるように、パワーの予測ができない非定常雑音の場合には、雑音が抑圧されても音声区間の検出は容易ではない。

これらのことから、AMNORを単なる受音装置とみなして、その出力を用いて音声区間検出を行うことを考えた場合、定常雑音に対しては良好な検出制度の向上が得られるが、非定常雑音に対してはあまり改善効果が得られないことがわかる。

しかし通常的环境においては、不要音声・音楽・扉の閉まる音など、非定常雑音は数多く存在するため、このことは重大な問題点である。従って、この非定常雑音下での音声区間検出問題の解決が本論文の重要な課題となる。

3. AMNORとMLD検定法を組み合わせた音声区間検出

本章では非定常雑音下での音声区間検出を可能とするために、AMNORの出力信号を含む複数チャンネルの信号を利用した手法の提案を行う。

3.1 提案する音声区間検出法

図3に提案する音声区間検出法の処理手順を示す。雑音の混入した音声はマイクロホンアレーにより受音され、AMNOR信号処理により、雑音抑圧フィルタリ

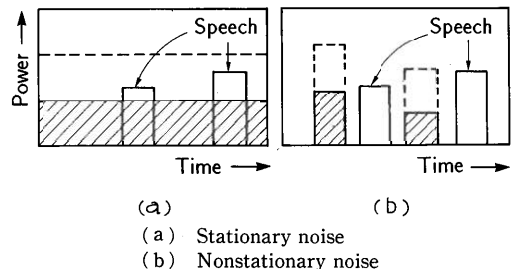


図2 AMNORによる雑音抑圧効果を表す模式図
Fig. 2 Model of AMNOR noise reduction effect.

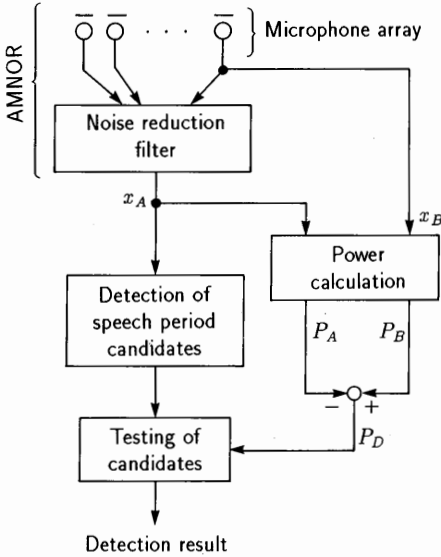


図3 提案する音声区間検出法の処理手順
Fig. 3 Proposed speech period detection method.

ングが行われる。次に、AMNORの出力信号 x_A を用いて音声区間候補の選択を行う。この候補選択の方法については、特に限定をしない。例えば、従来提案されている1チャンネル信号を用いた音声区間の検出方法を用いればよい。但し、この段階では、雑音区間が音声区間の候補として選ばれることはあるが、音声区間を棄却することはない、過剰候補選択の状態であるとする。

次に、各時刻において、 x_A およびアレーを構成する一つのマイクロホンからの出力信号 x_B のパワー P_A , P_B を計算し、それらの対数値としての差 $P_D(\text{dB}) = P_B(\text{dB}) - P_A(\text{dB})$ を求める。この P_D は、AMNOR 雑音抑圧処理を行う前の信号 x_B と行った後の信号 x_A のパワーの差を表している。

さてここで、音声と非定常雑音が図4の模式図に示すように受音されている場合を考える。図2の場合と同様に、破線は雑音抑圧を行わない場合の雑音レベルを、ハッチをかけた部分は抑圧された雑音のレベルを、また、実線白ぬきの部分は音声信号を表している。このとき、選択された候補が雑音区間であった場合には、 P_D の値は AMNOR の雑音抑圧効果に相当する値となる。

一方、AMNOR は音声レベルにほとんど変化を与えないので、 x_A , x_B に含まれる音声のレベルは等しく、候補が正しい音声区間である場合には P_D の値はほぼ零となる。

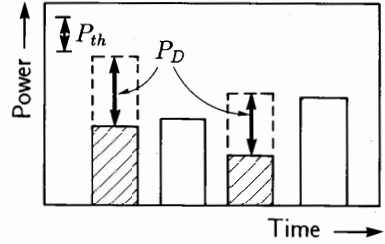


図4 MLD法による音声区間候補検定の考え方を表す模式図
Fig. 4 Concept of MLD candidate testing method.

もちろん種々の現実的条件下では、 P_D の値はこのような理想的な値をとるとは限らない。そこで、図3における候補検定部において、あるしきい値 P_{th} をもうけ、 P_D の値が P_{th} を下回った場合には候補は正しい音声区間、そうでない場合には雑音区間とみなす候補の検定を行う。

この方法は1チャンネル信号のパワーレベルではなく、複数信号間のレベル差を利用しているため、雑音レベルの大小によらず、雑音区間を音声区間と誤判定することはない。従って、雑音レベルの予測ができない非定常雑音下であっても音声区間の検出が可能となる。以降、この複数の信号チャンネル間のレベル差を用いた検定方法を、MLD (Multi-channel Level Difference) 法と呼ぶことにする。

3.2 提案手法の有効性とパラメータの定め方

前節では音声区間と非定常雑音の区間が異なっている場合について本提案手法の考え方を説明してきた。しかし、本手法の有効性としきい値 P_{th} の定め方を検討するためには、音声と非定常雑音の区間が重畳している場合について考察する必要がある。

雑音が音声区間に重畳している場合、正しい音声区間であっても P_D の値は零より大きな値となる。そこで、音声と雑音が無相関であると仮定して、雑音が重畳している音声区間の P_D の値を計算した。計算は、AMNOR の雑音に対する感度を $G (< 1)$ 、AMNOR の入力における音声と雑音のパワーをそれぞれ P_s, P_n と表して、 $P_B = P_s + P_n$ と $P_A = P_s + G \cdot P_n$ との対数軸上での差 P_D を求めた。

計算結果を図5に、受信SN比(AMNORの入力におけるSN比) P_s/P_n をパラメータとして表示した。この図より、 P_D の値は、受信SN比が0dBよりある程度大きい場合には零に近い値となり、SN比が0dBより小さくなるにつれて(図の左方向) P_D は大きくなる。そして、受信SN比が $G(\text{dB})$ 以下では、 P_D の値は音声

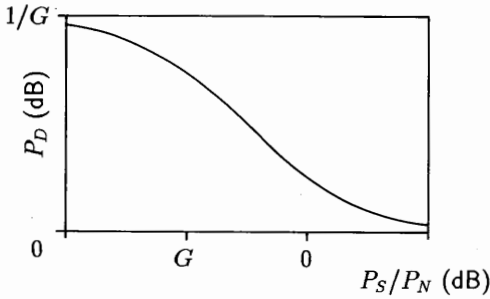


図 5 雑音の重畳した音声区間におけるチャンネル間レベル差 P_D

Fig. 5 Channel level difference P_D at the period where speech and noise overlap.

区間であっても雑音区間における値 $1/G$ と同程度となることがわかる。なお、対数をとったときに正值となる $1/G$ を G_R と表して、これを以降、雑音抑圧量と呼ぶことにする。

以上の結果を図 6(a)(b) に模式図として示した。図 6(a) は AMNOR の出力においても SN 比が悪い場合を示している。このとき、音声区間においても雑音成分の比率が大きく、 P_D の値は雑音区間における P_D の値とほぼ等しくなってしまう。そしてその結果、音声区間を雑音区間と判定するという誤りが生じる。このような状況は本手法の適用限界を示すものである。

図 6(b) は AMNOR 出力の SN 比が比較的良好な場合である。この場合には、音声区間における雑音成分の比率はそれほど大きくないため、雑音が抑圧されても P_D の値は雑音区間における P_D の値ほど大きくはならない。従って、図 6(b) に示すように P_{th} の値を雑音区間における P_D の値 G_R に近い値に定めておけば、雑音と音声为重畳していても音声区間の検出は可能である。

AMNOR は指向性受音機器であるので、その雑音抑圧量 G_R は、話者と雑音源の方向、室内の反響条件などによって異なってくる。しかし、話者と雑音源の方向が $\pm 20^\circ$ 以内の場合を除けば、悪くとも 10 dB 程度の雑音抑圧効果は期待できるので、 P_{th} の値としては 10 dB 弱と定めておくのが妥当であると考えられる。

なお、 G_R および P_{th} の値と、誤判定を生じる限界 SN 比との定量的関係は付録に示した。

3.3 提案手法に対する考察

3.3.1 MLD 法と AMNOR の組合せについて

MLD 法は必ずしも AMNOR と組み合わせる必要がなく、他の指向性受音機器、例えば単一指向性マイ

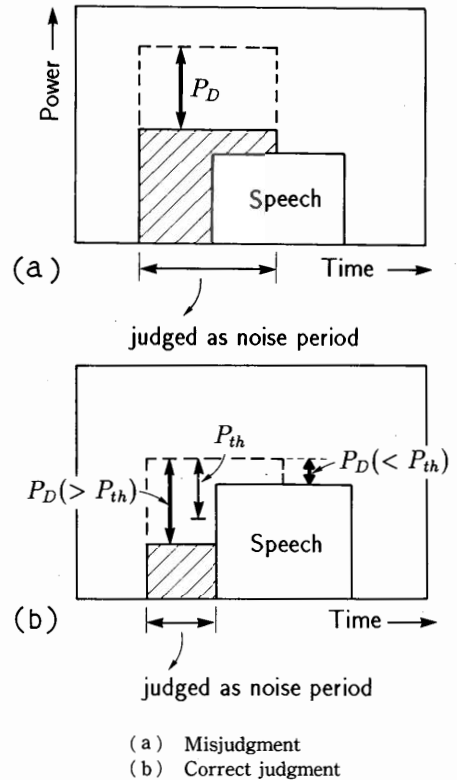


図 6 非定常雑音の重畳した音声区間に対する本手法の有効性を表す構成図

Fig. 6 Judgment of the proposed method at the period where speech and noise overlap.

クロホンの出力と無指向性マイクロホンの出力を用いても原理的には適用可能である。しかし、単一指向性マイクロホンは、正面方向から $\pm 90^\circ$ 以上の範囲から到来する雑音には大きな雑音抑圧効果をもつが、 $\pm 90^\circ$ 以内の範囲から到来する場合には、数 dB 程度の雑音抑圧能力しか有しない。そしてその場合には音声区間と雑音区間の P_D の違いも数 dB 程度となるため、種々の要因による P_D の変動を吸収できず、検出誤差が生じやすくなる。これに対して AMNOR は、雑音の到来方向に応じた指向特性を形成するため、広範囲の雑音到来方向にわたり、10 dB 以上の雑音抑圧効果を実現し、音声区間と雑音区間における P_D の違いも大きい。このように、予想される音響条件の多くにおいて、ある程度以上の雑音抑圧効果を実現することは MLD 法の適用にあたって重要であり、その意味で MLD 法と AMNOR との組合せは効果的である。

3.3.2 音声区間候補選択の必要性

以上述べてきた説明だけからは、 P_D の値のみで音声

区間の検出は可能で、図3に示す音声区間候補の選択は不要であるとも考えられる。しかし例えば、図4に見られるような音声も音響的雑音も存在しない時間区間を考える。そのとき、信号のパワーは増幅器などから発生する電氣的雑音のパワーが支配的であるとすると、その区間が音声区間でなくとも P_d の値は零に近い値をとり得る。そして、 P_d の値だけを見る場合、この区間を音声区間と判定してしまう。従って、簡単な音声区間候補選択手法であっても、明らかに音声区間とはみなせない区間をあらかじめ除外しておくことは有効である。

4. 実験結果

本章では、提案手法における候補検定の効果を調べ、単語音声区間検出に対する提案手法の有効性を確認するために行った実験結果を述べる。

4.1 実験条件

実験は容積が約 100 m^3 、残響時間が 0.4 秒の実験室で行った。非定常雑音としては妨害音声（ラジオのニュース）を発生させ、所望音声としては、単語音声（都市名）を 100 サンプル受信した。音声区間の正解を得る必要があるため、妨害音声と所望音声は別個に受信し、計算機上で加算した。AMNOR は単一指向性マイクロホン6個よりなるアレーを用いて構成した。マイクロホン配置、および目的音源、雑音源配置を図7に示す。各指向性マイクロホンの正面方向は図の下方向を向くように配置した。この条件でのAMNORの雑音抑圧量は約 14 dB であった。信号は 8 kHz でサンプリングし、信号パワーの算出は窓長 30 ms で 10 ms ごとに行なった。

音声区間候補の選択法は、以下に示す時間方向のレベル差に基づく方法（以下TLD法と呼ぶ）を用いた。

(1) 図8に示すように、各発声を一定の長さで取り込み、その中での最大パワーと最小パワーの対数値としての差 $P_M(\text{dB})$ を求める。

(2) ある区間のパワーレベルが、最小パワーレベルより $P_M \times 0.5$ 以上大きくなった区間を音声区間の候補とする。

本実験においては、このTLD法を雑音がないときの音声に対して適用して得られた区間を、音声区間の正解とみなした。無雑音時の P_M の値は約 30 dB であったので、このことは、音声の最大パワーから 15 dB 以内の区間を音声区間の正解とみなしたことになる。この正解区間は音声区間のうちパワーの大きな部分

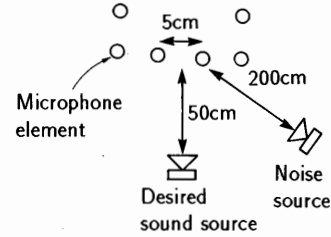


図7 マイクロホンおよび音源配置
Fig.7 Microphones and sound sources arrangement.

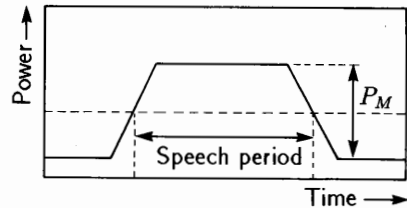


図8 時間方向のレベル差に基づく音声区間検出法 (TLD法)

Fig.8 Speech period detection method based on the short-time power difference in one-channel signal (TLD method).

(音声区間主要部)であり、レベルの低い子音・母音部分は含んでいないため、真の音声区間とは異なったものである。しかし、実際の音声認識システムなどでは、この音声区間主要部さえ正確に抽出されれば、端点フリーのマッチング方法などによって、低レベル音声区間の欠落に対処することもできる⁽⁶⁾。従って、音声区間主要部を雑音下でも誤りなく検出することは重要な技術と考えられる。

また、本方法のパラメータであるチャンネル間レベル差のしきい値 P_M は全実験を通じて 8 dB と設定した。

4.2 MLD法による検定効果

MLD法による検定の効果を確認するために、検定の有無による誤判定を調べた。誤判定は雑音区間を音声区間と誤判定した場合と、音声区間を雑音区間と誤判定した区間に分類し、全区間長に占める誤判定となった区間長の比率を求めた。平均受信 SN 比は $-15, -10, -5, 0, +5 \text{ dB}$ の5条件で実験を行った。なお、SN 比は音声の母音区間の平均パワーを用いて定めた。

実験結果を表1に示す。この表の数値の絶対的な値は、非定常雑音が多長に占める割合などに依存するためその意味付けは難しいが、数値の相対的な関係から以下のことがわかる。

(1) 検定を行わない場合に発生する雑音区間を音

表1 MLD法による検定の効果

平均受音 SN 比	雑音区間を音声区間と誤判定		音声区間を雑音区間と誤判定	
	検定なし	検定あり	検定なし	検定あり
-15 dB	19%	1.5%	8%	18%
-10 dB	18%	0.6%	4%	10%
-5 dB	16%	0.4%	1%	5%
0 dB	7.2%	0.1%	0.3%	1.4%
5 dB	0.2%	0.1%	0.1%	0.2%

声区間と判定する誤りは、検定を行えばほとんど生じない。

(2) 音声区間を雑音区間と判定する誤りは、SN比が低い場合、この検定を行ったほうが生じやすい。

次に、音声区間を雑音区間と判定する誤りを生じる区間を個別に調査してみた。その結果、この誤りは、非定常雑音のパワーが特に大きく、AMNORで雑音が低減された信号の短時間SN比が5dB以下のときに発生していることがわかった。このことを言い換えると、短時間SN比が5dB以上であれば、音声区間の検出が適切に行われていることが確認された。

検定を行うことで不利になる、AMNOR出力のSN比が5dB以下の場合というのは、仮にその区間が音声区間であることがわかったとしても、低SN比であるため、それ以降の処理において利用することが難しい区間である。そのように考えた場合、雑音区間を音声区間と誤判定することはなく、SN比の良い音声区間の適切な検出を可能とするMLD法による候補検定は有効な手段であると考えられる。

4.3 単語区間検出実験例

単語区間検出の実験結果の一例を図9に示す。この実験例においては「名古屋」「東京」「大阪」という単語を続けて発声した。図(a)は妨害音声が無い状態での信号のパワーを表し、TLD法を用いて検出した単語区間の「正解」を図の上に矢印で示している。なお、検出された二つの音声区間に挟まれる200ms以内の非音声区間は、単語区間の一部であるとみなして表示した。

図(b)は平均SN比が-5dBの妨害雑音下で単一指向性マイクロホンで受音した信号を示す。この信号にTLD法を用いた場合には、ほぼ全区間が単語区間と判定されてしまい、単語区間検出は不可能である。

図(c)は同じ条件下でAMNORを用いて受音した信号と、TLD法を用いて単語区間の検出を行った結果

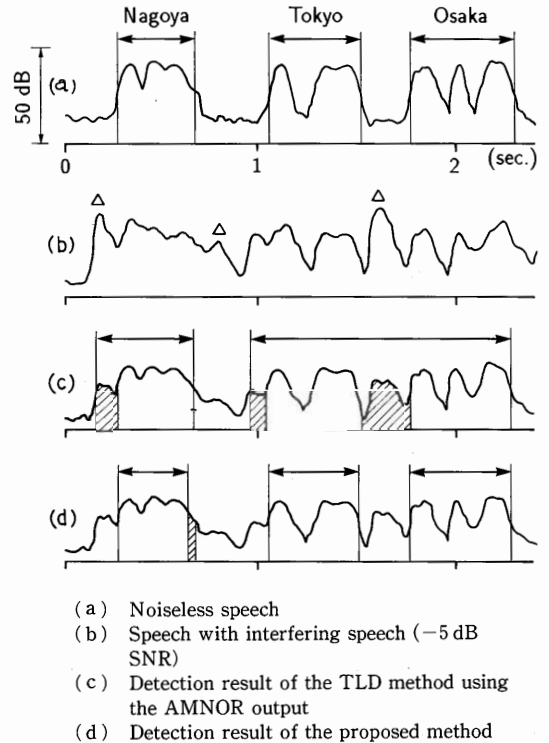


図9 単語区間検出実験結果の一例
Fig.9 Example of the word period detection experiment.

を示す。図(b)に比べてSN比が改善されており、所望の音声の形がある程度わかるようにはなっている。

また、図(b)に見られる非定常雑音のパワーの鋭いピーク(代表的なものを△印で示した)が図(c)では大きく抑圧されていることがわかる。この理由は、鋭いピークの部分は主として雑音の直接音および初期反射音などの方向性の明瞭な成分のパワーを表しており、AMNORはそのような方向性の雑音を、多重反射された雑音に比べて、より効果的に抑圧するからである。実際、この実験における平均的雑音抑圧効果は約14dBであったが、雑音パワーのピーク部分は20dB近く抑圧されている。このようにAMNORは非定常雑音パワーの時間的平滑化の効果ももつため、時間方向のレベル差を音声区間検出のための手がかりとする場合には有利な効果を与える。

しかしこの例では、妨害音声のパワーの大きな部分が依然として残されているため、時間方向のレベル差のみを用いるTLD法で判定した結果は図(c)矢印のようになり、ハッチで示した部分が誤った検出結果になってしまう。

図(d)はMLD法による検定を行う本提案手法の結果を示す。図(c)において誤って音声区間として選ばれた雑音区間は正しく棄却されており、本手法の有効性がわかる。また、図(d)にハッチで示した区間は誤って雑音と判定された音声区間を表している。この例で見られるように、検定の結果生じる、音声区間を雑音区間と判定する誤りの多くは単語の語尾の部分に生じる。その大きな理由は、残響のある室内で受音される所望音声の語尾の部分は、直接音成分より不特定な方向から到来する多重反射音(残響音)の成分の比率が大きく、AMNORはそのような目的音声の反射音を低減する効果をもつため、 P_0 の値が低減することによるものである。

4.4 単語区間検出実験結果

最後に、以上述べてきた本手法の有効性を定量的に求めた。単語区間検出結果の評価方法としては、単語の始端、終端における検出誤差が50ms以内であったものを正解とし、その正答率を求めた。受音SN比は-5dBと設定した。まず、単一指向性マイクロホンの出力にTLD法を適用するという従来法では、図9(b)からもわかるように正確な検出はほぼ不可能であった。次に、AMNORの出力にTLD法のみを適用した場合の正答率は43%であった。これに対して、AMNORの出力にTLD法を適用して候補選択を行い、MLD法により検定を行う本提案手法では、96%の正答率を得ることができた。更に、正解であった場合について、始端と終端における平均検出誤差を求めた結果、その誤差は約20msであった。このように、適応形マイクロホンアレーAMNORにMLD検定法を組み合わせた音声区間検出法の有効性が確認された。

5. む す び

本論文では、適応形雑音抑圧マイクロホンアレーAMNORとMLD法と名づけた検定法を組み合わせて、従来困難であった非定常雑音下でも有効に動作可能な音声区間検出法の提案を行った。

本手法は、まずAMNORにより雑音下での音声を受音する。このとき、マイクロホンからみて雑音源と話者の方向は近接(±20°程度)しておらず、AMNORは良好な(10dB以上)雑音抑圧動作を行うものと仮定する。次に、AMNORの出力に対し、従来法を用いて音声区間の候補を選定する。そして、AMNORの入・出力信号間(雑音抑圧前後)のレベル差と一定のしきい値との大小関係によって、その候補の検定を行う

(MLD法)。候補区間が雑音区間であるときには、AMNORの入・出力信号間にレベル差が生じることから、この方法は、従来法では多く発生した高レベルの非定常雑音を音声と判定する誤りを生じることではなく、ある程度以上のSN比をもつ音声区間を的確に検出できる。

AMNORの雑音抑圧量が通常の室内での平均的な値である14dBの場合の実験結果より、以下の結果を得た。

(1) 本手法は、AMNOR出力の短時間SN比が5dB以上の音声区間をほぼ適切に検出することができた。

(2) 指向性マイクロホンの出力信号における平均的SN比が-5dBとなる妨害音声下で単語音声区間の検出実験を行った。音声の最大パワーから15dB以内の区間を音声区間の「正解」とみなしたとき、検出正答率は96%、また、正答となった場合の平均検出誤差は20msが得られ、本音声区間検出法の有効性が示された。

これらの結果より、非定常雑音下での音声区間検出を可能とする方法の提案、という本研究の目標は一応達成できたと考える。しかし現時点では、動き回る雑音源に対するAMNORの性能は十分には保証されておらず、この方法をより一般的な雑音環境で有効とするためには、移動音源に対するAMNORの性能向上が今後の課題となる。また、本手法を音声認識に適用した場合の性能評価も今後の重要な課題である。

謝辞 本研究の遂行にあたり、御討論・御指導頂きました当所古井貞熙博士に感謝致します。

文 献

- (1) 信学技報, EA89-1~9 (1989-04)など。
- (2) H. Kobatake, K. Tawa and A. Ishida: "Speech/non-speech discrimination for speech recognition system under real life noise environments", Proc. of ICAS-SP'89, S7. 5, pp. 365-368 (May 1989).
- (3) 坪井宏之, 竹林洋一: "音声認識技術とその実用化動向", 電学論(C), 108, 10, pp. 755-760 (昭63-10).
- (4) Y. Kaneda and J. Ohga: "Adaptive microphone-array system for noise reduction", IEEE Trans. Acoust., Speech & Signal Process., ASSP-34, 6, pp. 1391-1400 (Dec. 1986).
- (5) 金田 豊: "適応形雑音抑圧マイクロホンアレー(AMNOR)の指向特性", 音響誌, 44, 1, pp. 23-30 (昭63-01).
- (6) 古井貞熙: "ディジタル音声処理", 東海大学出版会(昭60).

付 録

本手法の適用限界 SN 比

本文 3.2 より、雑音に対する AMNOR の感度を G 、AMNOR 入力における SN 比を R_1 とすると、

$$P_D = \frac{R_1 + 1}{R_1 + G} \quad (\text{A} \cdot 1)$$

と表すことができる。音声区間を雑音区間として棄却する誤りは、 $P_D \geq P_{th}$ の場合に生じるので、限界 SN 比（その値より小さくなると誤りを生じるという SN 比） R_{1m} は、 $P_D = P_{th}$ を解いて、

$$R_{1m} = \frac{1 - G \cdot P_{th}}{(P_{th} - 1)} \quad (\text{A} \cdot 2)$$

と求められる。また、AMNOR 出力における限界 SN 比 R_{2m} は、

$$R_{2m} = R_{1m}/G \quad (\text{A} \cdot 3)$$

と表される。

本論文での実験では、 $P_{th} = 8 \text{ dB}$ 、 $G_R = 1/G = 14 \text{ dB}$ であるので、式(A・3)より $R_{2m} = 5.5 \text{ dB}$ となり、実験結果をほぼ裏づけるものとなっている。

(平成元年 10 月 13 日受付, 2 年 3 月 15 日再受付)



金田 豊

昭 50 名大・工・電気卒、昭 52 同大大学院修士課程了、同年 NTT 武蔵野研究所勤務、現在に至る。空間で得られる複数の情報を利用した音響信号処理の研究に従事。