

氏 名 (本 籍)	にしむらりょういち 西村 竜 一	(福 岡 県)
学 位 の 種 類	博 士 (情 報 科 学)	
学 位 記 番 号	情 博 第 77 号	
学 位 授 与 年 月 日	平 成 10 年 3 月 25 日	
学 位 授 与 の 要 件	学 位 規 則 第 4 条 第 1 項 該 当	
研 究 科 , 専 攻	東 北 大 学 大 学 院 情 報 科 学 研 究 科 (博 士 課 程) シ ス テ ム 情 報 科 学 専 攻	
学 位 論 文 題 目	ウ ェ ー ブ レ ッ ト 変 換 に 基 づ く 音 声 強 調 手 法 に 関 す る 研 究	
論 文 審 査 委 員	(主 査) 東 北 大 学 教 授 曾 根 敏 夫 東 北 大 学 教 授 根 元 義 章 東 北 大 学 教 授 牧 野 正 三 東 北 大 学 助 教 授 鈴 木 陽 一	

論 文 内 容 要 旨

第 1 章 序 論

現在、通信やその関連分野において、1チャンネルによる通信は非常に多くの割合を占めている。しかし、複数チャンネルによる信号伝送と比較して情報が少ないため、目的信号と雑音信号が混在する場合には、これらを分離するのが困難となる。特に目的信号が音声信号の場合には、発話内容が聞き取れないなどの問題が生じる。この問題の解決方法として、信号に対して音声強調処理を施すことが考えられる。1チャンネル通信の利点は、システムが簡便であることと、それ故、多くの人や場面で利用できる点である。音声モデルや学習を必要とするものでは、ある程度限定された発話者に特化した音声強調システムとなってしまったため、先の利点が損なわれてしまう。そこで、この利点を損なわない音声強調手法の実現が望まれる。このような観点に立った従来の音声強調手法では、物理的なS/Nは改善されるものの、ミュージカルノイズと呼ばれる楽音的な雑音の影響などにより、受聴環境はあまり改善されないものが多い。そこで、本論文では、信号の物理的な性能を改善するばかりでなく、受聴環境も改善するという意味において優れた音声強調手法を実現することを目的とする。

第 2 章 短区間スペクトル推定に基づく1チャンネル音声強調手法

本章では、先に述べた立場に立った従来の音声強調手法の問題点を、計算機シミュレーションを通して明らかにすることを目的とする。そのために、スペクトル減算法、ウィナー法、およびMMSE法の3つの手法を取り上げた。原音声のパワーに対する原音声と処理後の信号のパワーの比で定義されるS/D、および、処理後の信号のパワーに対する処理後の雑音信号のパワーの比で定義される(S+N)/Nのふたつの指標を用いた、その結果、1)信号解析手法として短区間フーリエ変換を用いることが、必ずしも適していないこと、2)音声に関する何らかの情報を利用することで、高い音声強調効果の実現できる可能性があること、3)雑音抑制効果を高めるためには、定常信号を仮定したアルゴリズムよりも、非定常信号であることを意識したフィルタの設計が必要であることを明らかにした。

第 3 章 ウェーブレット変換に基づくスペクトル減算法の拡張

本章では、スペクトル減算法のアルゴリズムに適した信号解析手法についての検討を行い、近年数学的な体系化が進められているウェーブレット変換を採用した。ウェーブレット変換は、被解析関数を $f(t)$ とすると、 a 、 b の2つの変数を用いて、

$$(W_\phi f)(a, b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \phi^* \left(\frac{t-b}{a} \right) dt$$

で定義される。ここで、関数 $\phi(t)$ は固定的なものではなく、幾つかの条件を満たささえすれば、どのような関数を用いても構わない。したがって、ウェーブレット変換は、信号解析のための基底系の構成に自由度がある。また、変数 a の働きにより、解析周波数に応じて時間-周波数分解能が変化するという特徴を有している。これは、人間の聴覚系における信号解析と類似な解析法であり、ウェーブレット変換による信号解析が、人間の聴覚と高い親和性を有していることを意味する。

ここで、スペクトル減算法における推定誤差の要因を明らかにするために、瞬時フレームにおける誤差についての検討を行った。その結果、ふたつの要因があることが判明した。ひとつは、音声信号と雑音信号のクロスパワースペクトルに起因するものであり、もうひとつは、瞬時フレームにおける雑音信号のパワースペクトルと平均的な雑音信号のパワースペクトルとの差に起因するものである。その上で、この誤差を軽減するには、スペクトル減算処理を行う空間において、音声のパワーの局在性を高めることが必要であると推論された。先に述べたウェーブレット変換の性質から、スペクトル減算法のアルゴリズムに対しては、ウェーブレット変換が極めて有効な信号解析手法であると考えられた。そこで、ウェーブレット変換とスペクトル減算法のアルゴリズムを組み合わせた音声強調手法である WBSS 法を新たに提案した。

ウェーブレット変換は、基底系が直交する直交ウェーブレット変換と、直交しない非直交ウェーブレット変換に大別することができる。直交ウェーブレット変換は、高速な計算アルゴリズムが存在する反面、基底の選択性の自由度が低くなる。一方、非直交ウェーブレット変換では、多くの演算時間を必要とする代わりに、極めて柔軟に基底系を構成することができる。したがって、WBSS法において、適切に基底系を構成した非直交系のウェーブレット変換を用いることにより、物理的特性を大きく改善する音声強調が実現できるものと予想した。

計算機シミュレーションの結果を図1に示す。図から、非直交系のウェーブレット変換を用いることにより、物理的な音声強調効果が向上するこ

とが分かる。特に、音声強調が困難となる低周波数領域が優勢な雑音に対して、提案手法が有効であることが明らかとなった。また、聴取実験の結果から、直交系のウェーブレット変換を用いることで、WBSS法は聴感上の音声強調性能も改善し得る手法であることが示された。これらのことから、WBSS法はウェーブレット基底の構成を変化させることにより、用途に応じて様々な音声強調問題へ適用することが可能な手法であると言える。

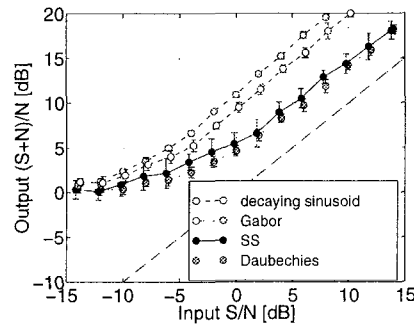


図1：付加雑音をピンクノイズとしたときの、WBSS法による $(S+N)/N$ の改善量 (decaying sinusoid: 減衰正弦波ウェーブレット [非直交], Gabor: ガボールウェーブレット [非直交], Daubechies: Daubechies ウェーブレット [直交])

第4章 カルマンフィルタとウェーブレット変換に基づく音声強調手法の提案

第2章における従来の音声強調手法の性能評価を通じて、音声に関する何らかの情報を利用することで、高い音声強調効果が実現できる可能性があることが示された。このことから、音声に関してできる限り多くの情報を用いることが、高性能な音声強調を実現するためには必要不可欠であると判断された。そこで本章において、音声信号が非定常信号であることを意識した音声強調手法の実現を試みた。たとえ雑音が付加されていても、入力信号を得た時点で、音声信号

の非定常性に関する情報の一部は得ることができる。このことから、非定常信号に対する最小平均自乗誤差フィルタとなるカルマンフィルタを音声強調へ適用することを試みた。カルマンフィルタでは、目的信号のダイナミクスに対応する状態遷移行列は、通常、既知であると仮定される。しかし、音声強調問題では、入力信号から状態遷移行列を推定する必要がある。そこで、図2に示すように、状態ベクトルや観測ベクトルを、信号のウェーブレット係数やコサイン変換係数から構成される行列で表現することを考案した。雑音が付加された音声信号を、予めウィナーフィルタやWBSS法などにより予備推定したうえで、上記の方法により状態遷移行列の推定を行い、カルマンフィルタを実現した。この音声強調アルゴリズムのブロックダイアグラムは、図3に示す通りである。

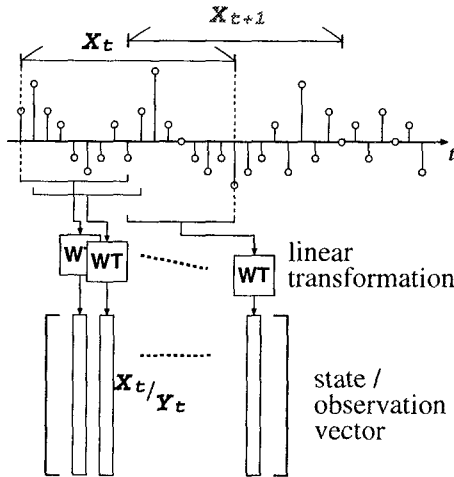


図2：状態ベクトル及び観測ベクトルの行列表現

計算機シミュレーションの結果、再生音の音声にひずみが生じ、S/Dは改善されないことが判明した。しかし、(S+N)/Nについては、WBSS法と同程度の改善を示すことが判明し、雑音を抑制することには成功したものと判断される。また、聴取実験から、聞きやすさや不快感については、従来のスペクトル減算法と比較して改善することが判明した。したがって、提案手法によって処理された音声に生じたひずみは、聴感上重大な問題を与えるものではないと判断される。更に、状態遷移行列が既知という理想的な状況においては、本手法は極めて高い音声強調効果を実現できることが判明した。このことは、状態遷移行列の推定精度を向上させることにより、本手法の性能が更に改善されることを示すものである。

第5章 結論

本論文では、物理的な性能改善ばかりでなく、聴感上での聞きやすさも改善する音声強調手法の実現を目的として、大きく分けてふたつの手法を提案した。

ひとつは、ウェーブレット変換とスペクトル減算法のアルゴリズムを組み合わせることによって、高い音声強調性能を実現する手法である。もうひとつの手法は、音声の非定常性に着目したカルマンフィルタを用いる手法である。同じ条件下においては、どちらの手法も従来の手法と比較して、高い、あるいは、少なくとも同程度の音声強調効果が実現できることが示された。また、カルマンフィルタを用いる手法は、状態遷移行列の推定精度を向上させることにより、音声強調性能が更に向上することを示した。

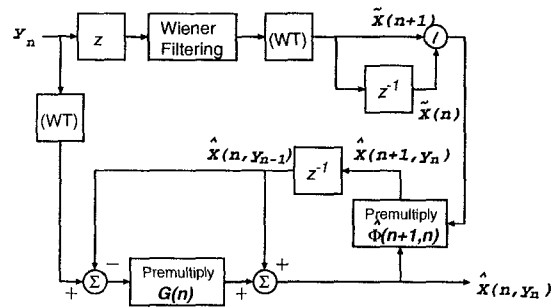


図3：カルマンフィルタを用いた音声強調手法のブロックダイアグラム

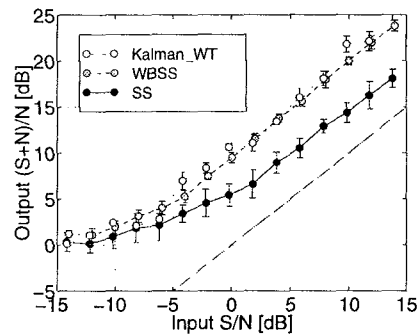


図4：付加雑音をピンクノイズとしたときの、カルマンフィルタを用いた音声強調手法による(S+N)/N (Kalman-WT：提案手法, WBSS：減衰正弦波ウェーブレットを用いた場合のWBSS法)

審査結果の要旨

音声信号と雑音信号が混在している場合、1チャンネルの通信路では、限られた情報しか得られないため、音声強調処理を行って雑音を抑制することには困難が伴う。また、従来の音声強調手法では、SN比はある程度改善されるものの、実際に受聴したときの主観的評価はあまり改善されない。著者は、この点に着目し、物理的にも主観的にもより効果的な音声強調手法を実現するための研究を行ってきた。本論文は、その研究の成果をまとめたもので、全編5章よりなる。

第1章は、序論である。

第2章では、従来の音声強調手法について、物理的な観点から性能を検討し、音声強調に用いる解析手法としては、短区間フーリエ変換よりも時間周波数平面上で音声信号を効果的に局在化させることが可能で、聴覚の周波数分析機構とも親和性の高いウェーブレット変換の方が有望であることを述べている。また、雑音抑制効果を高めるには、音声信号が非定常信号であることを考慮したフィルタの設計が重要であることを指摘している。

第3章では、広く用いられているスペクトル減算法にウェーブレット変換を導入すれば、スペクトル減算処理を行う空間での音声パワーの局在性が高まることから、効果的な音声強調が達成できるものと推論し、信号を時間周波数平面上で減算する新しい音声強調手法を提案している。また、この場合、基底系として非直交系ウェーブレットを用いることにより、従来の手法の中で最も良い性能を示すウィナー法に比べ、 $(S+N)/N$ を最大で5dB程度改善できること、直交系ウェーブレットを用いた場合には、良好な主観的評価が得られることを示した。これらは、有用性の高い音声強調アルゴリズムの提案として、高く評価できる。

第4章では、音声为非定常信号であることに着目し、音声強調問題にカルマンフィルタを用いることを試みている。すなわち、状態遷移行列を推定するため、音声の特徴を考慮して、1サンプルごとに計算したウェーブレット変換の結果から作られる行列で状態遷移行列を表現することを提案している。また、予備的に音声強調した入力信号から状態遷移行列を求めた場合には、ここで提案した手法が第3章で提案した手法と同等の性能をもつことを明らかにし、さらに、音声の状態遷移行列の推定精度を高めることにより、物理的にも主観的にも極めて良好な音声強調が達成できることを示している。この提案は、より高度な音声強調手法の実現への足掛りとなるものとして評価できる。

第5章は、結論である。

以上要するに、本論文は、1チャンネルの音声通信路において、物理的なSN比と聴感上の評価の双方を改善できる音声強調手法について、新しい有用な手法を提案したもので、システム情報科学ならびに音響通信工学の発展に寄与するところが少なくない。

よって、本論文は博士（情報科学）の学位論文として合格と認める。