

氏名	すぎやままさひで 杉山雅英
授与学位	工学博士
学位授与年月日	昭和60年2月13日
学位授与の根拠法規	学位規則第5条第2項
最終学歴	昭和54年3月 東北大学大学院理学研究科数学専攻 前期課程修了
学位論文題目	LPC Spectral Matching Measures for Speech Recognition (音声認識のためのLPCスペクトルマッチング尺度の研究)
論文審査委員	東北大学教授 城戸 健一 東北大学教授 重井 芳治 東北大学教授 木村 正行 東北大学教授 曾根 敏夫

論文内容要旨

音声の中には音韻と呼ばれる言語的な意味をもつ音声波があり、この音声波を用いた情報の伝達は人間にとって最も自然なものと言える。その音声波に含まれる意味内容の認識が音声認識の課題である。音声認識において2つの音声の違いを定量化する手法は重要であり、音声認識のみならず音声処理の諸分野において用いられてきた。本論文の主題は音声認識を目的としたスペクトルマッチング尺度の研究である。従来から音声認識において音声の特徴抽出の方法として、Linear Predictive Coding (LPC) 法が用いられてきた。本論文においてはLPC分析法に基づくスペクトルマッチング尺度を対象としている。

1971年～1976年に音声理解プロジェクトが米国国防省のARPAの援助のもとに行なわれ、音響処理と言語処理とを融合したシステム構成の研究が進められた。同様に日本国内でも音声理解の研究がいくつかの研究グループにおいて進められた。それらの研究の結果、音声理解における言語処理の有効性が明らかにされるとともに音響処理の精密化が研究課題として提出された。

また単語音声認識の研究分野においても、従来から音声間の違いを定量化する方法として、スペクトルマッチング尺度が用いられてきた。これらのマッチング尺度は単にスペクトルの差を求めるだけであって人間の聴覚特性を考慮してはいなかった。その結果、スペクトルマッチングに基づく単語音声認識システムは、発声者の変動、雑音の混入や入力環境の変動に非常に影響を受けやすく、認識率の低下が問題になっており新しい尺度の導出が期待されていた。以上の様な観点から、著者

は音響処理の基本的な手法の1つでもあるスペクトルマッチング尺度の研究を開始した。

マッチング尺度の改良の基本的なアイディアは、より人間の聴覚の性能に近付くために人間の聴覚の特性を考慮することである。人間の聴覚の特性としてスペクトルの強度の大きい部分、すなわちホルマントと呼ばれるスペクトルのピークに敏感であること、および0～2 kHzの低域周波数に敏感であることが聴覚の研究を通して示されていた。

そこでこの2つの聴覚特性を考慮したスペクトルマッチング尺度 Peak Weighted Measure を提案し、母音識別実験、大語彙単語音声認識実験を通してその有効性を示した。また従来からの課題である入力環境の変動に対しても非常に有効であることを示した。さらに、音声情報圧縮に必要な量子化歪み最小のスペクトルの計算手順、その諸性質および近似解法を示し、圧縮したスペクトルを用いた認識法の有効性を示した。

本論文の各章の内容は次の通りである。第1章では、音声認識におけるLPCスペクトルマッチング尺度の研究状況、問題点を述べ、第2章では、準備としてLPC分析法およびLPCスペクトルマッチング尺度について述べた。入力音声のピリオドグラムを重み関数とする内積を用いて、LPC分析法は多項式に対する一般化したフーリエ展開の理論として定式化できることを示した。また従来から用いられてきたLPCスペクトルマッチング尺度は人間の聴覚特性を考慮していないことを明らかにした。第3章では、従来の尺度のもつ問題点を解決する一方法として、スペクトルのピークに重みをおいたスペクトルマッチング尺度 (Peak Weighted Measure) を提案した。母音識別実験の結果、その1つであるWLR (Weighted Likelihood Ratio) 尺度の識別誤り率は9.3%であり、従来の尺度の母音識別率に比べて2.4%少ないことを示し、ピーク重み付けが有効であることを明らかにした。第4章では、周波数軸に重みをおくスペクトルマッチング尺度の構成方法を2種類提案した。低域周波数重み付けWLR尺度の母音識別誤り率は8.2%であり、従来の尺度に比べて3.5%少ないことを示した。母音識別実験の結果から母音認識に対して周波数軸重み付け尺度の有効性を明らかにした。さらに2種類の周波数軸重み付けの方法は Peak Weighted Measure に対して認識性能の観点からは等価である事を示した。第5章では、日本の都市名641単語を用いた大語彙単語音声認識実験を行ない、提案した尺度の評価を行なった。認識実験を通してWLR尺度が従来の尺度及び Peak Weighted Measure の中で最も有効な尺度であることを示した。第6章では、ピーク重み付け尺度の中でも特に認識率の高かったWLR尺度に焦点をしばって、WLR尺度に関する低域重み付けの効果、異話者音声の認識、帯域制限音声の認識および白色雑音付加音声の認識に対する有効性を大語彙単語音声認識実験を通して示した。低域周波数に重みをおいたWLR尺度の単語認識率は97.0%であり、従来の尺度の認識率に比べて1.7%高いことを示した。また、音声の登録者と発声者とが異なる場合の低域重み付けWLR尺度の認識率は82.4%であり、従来の尺度に比べて10.6%高いことを示した。さらに音声の登録をマイクで行ない音声の入力を電話から行なうような、音声の周波数帯域が異なる場合の認識においても、低域重み付けWLR尺度の認識率は95.5%であり、従来の尺度に比べて12.9%高いことを示した。最後に登録音声・入力音声の一方に白色雑音を加えた場合の認識において、SN比が18 dBのとき低域重み付けWLR尺度の認識率は85.0%であり、従来の尺度に比べて19.7%高いことを示した。これらの結果からス

ベクトルのピーク重み付け及び低域周波数重み付け尺度の有効性が明らかになった。第7章では、音声の圧縮法であるベクトル量子化法の中で重要な概念である量子化歪み最小のスペクトル(centroid)の導出方法を確立し、歪み最小スペクトルを用いた音声認識システムの有効性を示した。WLR尺度に対するcentroidはスペクトルの相関係数の相加平均で近似できることを示した。また641単語音声認識に対しては5000個程度の学習用音声サンプルから256個の代表的な音声スペクトルパターンに圧縮しても認識率の劣化は小さいことを明らかにした。第8章は、本論文の結論である。

本研究を通して得られた結果を個条書きにしておく。

- (1) スペクトルのピークに対する重み付けマッチング尺度は母音の認識及び単語音声の認識において有効である。また、低域周波数での重み付けマッチング尺度も同様に有効である。特にWLR尺度は認識性能が高く計算量も比較的少なく解析しやすい尺度であり有効な尺度である。
- (2) WLR尺度は特定話者音声認識に対して有効であるばかりでなく、不特定話者の音声認識や電話入力音声・雑音下音声のような歪んだ音声の認識に対しても有効である。このことはWLR尺度を用いることにより堅牢なシステムを構成できる可能性を示している。
- (3) WLR尺度に対する歪み最小スペクトルは相関係数の相加平均で近似できる。また641単語音声の認識のためには、5000サンプルから圧縮した256個のスペクトルパターンで十分であり、256個を用いたときの認識性能の劣化は小さい。

本研究の結果は、スペクトルマッチング尺度の問題点を解決し音声理解における音響処理の精密化に適用されるだけでなく、単語レベルでの音声認識の性能向上に適用することができた。また聴覚の特性を考慮することの重要性が改めて明らかになった。

審査結果の要旨

線形予測分析は、音声波の大部分を占める有声音の生成モデル同定に基づいているため、音声認識のための周波数スペクトル推定的手段として有用であると考えられ、音声認識の研究においては、線形予測分析によって推定されたスペクトル(LPCスペクトル)がよく用いられている。しかし、不特定の話者を対象にするときは、LPCスペクトルマッチングによる音素認識率は、期待された程に高くなり得なかった。著者は、その原因が、スペクトルのあらゆる部分が平等にマッチングに寄与することにあると考え、スペクトルの重要な部分に重点をおいたマッチング尺度の研究を行った。本論文は、この研究成果をまとめたもので、8章よりなる。

第1章は序論である。

第2章は本研究における分析とマッチングの方法の説明である。

第3章では、ピークに重みをおいたスペクトルマッチング尺度を提案し、その方法を詳細に述べると共に、母音認識実験によってその有用なことを明らかにしている。これは、優れた着想に基づく提案である。

第4章では、母音認識においては低周波の範囲のマッチングを重視すべきであるという考えから、周波数軸に重みをおいたスペクトルマッチング尺度を提案し、実験において、よい結果を得ている。

第5章では、前2章の研究に基づいて開発したLPCスペクトルマッチング尺度を用い、大語彙単語音声認識実験によって、尺度の評価を行っている。

第6章では、ピークに重みをつけた尺度の中でも特に高い認識率を得ることができたWLR(Weighted Likelihood Ratio)尺度を取り上げ、低域重み付けと併用して、異話者音声の認識、帯域制限音声の認識および白色雑音付加音声の認識に対して有効なことを示している。

第7章では、音声の情報圧縮に使われるベクトル量子化法の中で重要な概念である量子化歪み最小のスペクトルの導出法を確立し、音声認識系に適用している。

第8章は結論である。

以上要するに本論文は、音声認識におけるLPCスペクトルマッチング尺度について研究し、ピークと低周波部に重みをおいたスペクトルマッチング尺度を提案し、その有効なことを実験によって明らかにすると共に音声認識に重要な概念を与えたもので、情報工学の進歩に寄与するところが少なくない。

よって、本論文は工学博士の学位論文として合格と認める。