

報告番号 [※]乙 第 2033 号

主論文の要旨

題名

会話音声自動認識システムに関する研究

氏名 鹿野 清宏

主論文の要旨

報告番号	※甲第	号	氏名	鹿野清宏
<p>本論文は、著者が数年間にわたって行ってきた会話音声の自動認識システムの研究に関するものである。</p>				
<p>会話音声自動認識システムを構成するには、少なくとも次の項目について検討せねばならない。</p>				
<p>(1) 音響分析,</p>				
<p>(2) 単語認識アルゴリズム,</p>				
<p>(3) 言語情報の表現法および利用方法,</p>				
<p>(4) 認識システムの構成手法,</p>				
<p>(5) 認識システムの性能評価方法。</p>				
<p>(1)の音響分析は、音声認識技術の根幹である。本論文では、音響分析の手法として、分析精度の高いLPC (Linear Predictive Coding)分析をとりあげる。LPC分析に基づくスペクトルマッチング尺度として、COSH尺度およびWLR (Weighted Likelihood Ratio) 尺度をとりあげる。本論文で提案するWLR尺度は、スペクトル包絡のピークに重みをおいたスペクトルマッチングが行えるだけでなく、周波数軸の重みづけもできるすぐれた尺度である。また、パワーティップの新しい検出法を導入し、音韻単位のセグメンテ</p>				

主論文の要旨

報告番号	※甲第	号	氏名
------	-----	---	----

ーションの精密化を行う。(2)の単語認識アルゴリズムとして、音韻の類似度や調音結合の性質が考慮でき、かつ高速で動作するアルゴリズムを開発する。これをスタックを用いたトリーサーチのアルゴリズムで実現する。(3)の言語情報の利用に関しては、top-down的手法をとり入れ、構文情報のリスト表現を用いた新しい記述法を提案する。(4)のシステム構成法に関しては、音響処理と言語処理とを明確に分離することにより、音響処理と言語処理の普遍性を高め、対象への適応性を高める。音響処理結果のデータ形式として、音韻ラティスという新しい形式を用いる。また、会話音声認識システムを質問回答システムの中でとらえ、対象として列車の座席予約をとりあげる。(5)の性能評価に関しては、音響処理、言語処理の各部を個々に評価する。音響処理の評価を音韻認識率で行う。そのために、音韻ラティスの情報量に基づいた音韻認識率の推定方法を提案する。言語処理は、入力の音韻ラティスの音韻認識率と、処理結果の文節認識率とで評価する。また、質問回答システムとしての評価も行う。

主論文の要旨

報告番号 ※甲第 号 氏名

以下、本論文にそって各章の概要を述べる。

第1章では、会話音声認識の研究動向について述べるとともに、会話音声認識の問題点をあげ、研究の基本方針および本論文の構成について述べる。

第2章では、会話音声認識システム(Voice Q-A System)の基本構成について述べ、さらに、認識対象としてとりあげた新幹線の座席予約の概要について説明する。Voice Q-A Systemの基本構成として、次の方針をとる。

- (1) 音声による質問回答システムを目標とし、音声認識部の他に音声応答部も作成する。
- (2) 音声認識部は、役割が大幅に異なる音響処理と言語処理とに分離して作成し、各々が並列に動作できるようにする。
- (3) 音響処理と言語処理の情報の受け渡しの形式として音韻ラティスの形式をとる。

また、認識対象の困難さの評価尺度として冗長度を用いることを提案し、各種の言語情報の冗長度の大きさを評価する。

第3章では、Voice Q-A Systemの言語処理の構成法

主論文の要旨

報告番号	※甲第	号	氏名
------	-----	---	----

および Voice Q-A System I の性能評価について述べる。言語処理の構成およびアルゴリズムの特徴を以下にまとめておく。

(1) 言語処理は単語認識部、構文解析部、推論部からなる認識部と会話モデルからなる。認識部は、top-down 的な構成からなっており、left-to-right のマッチングを基本としている。

(2) 音韻ラティスと単語辞書とのマッチングのアルゴリズムとして、tree search の depth-first 法を基本とする手法を提案する。この手法では、様々な音韻の性質を音韻変形規則の形で利用でき、かつ tree search の途中過程の記憶や連想的な検索によって、実行速度も非常に高速化されている。

(3) 構文情報のリスト表現による表現方法を提案するとともに、この構文表現に基づく構文解析のアルゴリズムも明らかにする。

(4) 質問回答を通して、誤りの訂正が自然に行えるように会話モデルを作成する。

この章の後半では、実時間の 5 倍の処理時間で動作する Voice Q-A System I のオンラインシステムにつ

主論文の要旨

報告番号	※甲第	号	氏名
------	-----	---	----

いて説明し、男性8名の音声データを用いてシステムの評価を行う。音韻認識率58.5%の音韻ラティスを入力として、86.0%の文節認識率が得られる。また、質問回答実験も行われ、99.1%の予約達成率が得られている。この章の最後に、Voice Q-A System Iの問題点を明らかにし、Voice Q-A System IIの作成方針を示す。

オ4章では、音響処理の出力である音韻ラティスの自動評価方法を提案し、その有効性を確かめる。この章で提案する方法により、音韻認識のあいまいさとセグメンテーションのあいまいさを含む時系列である音韻ラティスの客観的評価方法が確立される。

オ5章では、音響処理の精密化の一環として、LPCスペクトルマッチング尺度についての検討を行う。5章の前半では、従来からのLPCスペクトルマッチング尺度をとりあげ、母音標準パターンとのマッチングによる母音識別実験を通しての評価を行う。その結果、次の結論が得られている。
(1) 尺度値が最小になるようにパワー比を設定した

主論文の要旨

報告番号 ※甲第 号 氏名

条件でのスペクトル・マッチングが適当である。

(2) 残差パワーを正規化した条件,あるいは尺度値最小の条件でのCOSH尺度が母音の識別にはすぐれている。5章の後半では,新しいLPCスペクトル・マッチング尺度として,LPC Peak Weighted Measureを提案し,理論的解析および母音識別実験による評価を行う。その結果,次の結論が得られる。

(1) LPC Peak Weighted Measureは,LPCスペクトル包絡のピーク付近に重みをおいたマッチングが可能であり,ホルマント・マッチングとスペクトル・マッチングの両者の長所をあわせもつ尺度である。

(2) LPC Peak Weighted Measureは,周波数軸の重みづけも可能である。

(3) LPC Peak Weighted Measureのうち,WLR (Weighted Likelihood Ratio)は,もっともすぐれた尺度である。5章の最後では,男性4名の発声した641都市名の単語音声データを対象として,LPCスペクトル・マッチング尺度による大語の単語音声認識実験を行う。その結果,WLRは96.7%の単語認識率を達成し,従来からの尺度よりもすぐれていることがわかる。

主論文の要旨

報告番号	※甲第	号	氏名
------	-----	---	----

第6章では、音響処理の改善を目的として、Voice Q-A System IIを作成する。とくに、次の点の改善を行う。

- (1) 再帰的な構成からなる音響処理、
- (2) 音韻単位のセグメンテーションの精密化、
- (3) LPCスペクトルマッチング尺度による母音認識の精密化、
- (4) 母音標準パターンの学習法の簡易化、
- (5) 発声者によらない子音認識法。

まず、LPCスペクトルマッチング尺度として、CosH尺度をとり上げ、Voice Q-A System II-1を作成する。男性9名の音声データで認識実験を行い、66.3%の音韻認識率、96.5%の文節認識率が得られ、Voice Q-A System Iに比べ、音韻認識率で8%、文節認識率で10%の向上が達成される。次に、マッチング尺度として、Peak Weighted MeasureのWLRをとりあげ、Voice Q-A System II-2を作成し、性能の向上をはかる。System II-1に比べ母音認識率で0.8%の向上が得られ、66.8%の音韻認識率、96.9%の文節認識率が達成される。これにより、すべての発声者で

主論文の要旨

報告番号	※甲第	号	氏名
------	-----	---	----

音韻認識率 63% , 文節認識率 95% を達成するという当初の目標がほぼ達成される。

最後に、本論文の研究成果で、音声認識の研究にとって大きく貢献すると思われる事項を列挙する。

- (1) 質問回答システムの観点から会話音声認識システムを作成し、高い認識性能を達成した。このことにより、会話音声認識システムの実現の可能性が示された。本論文の Voice Q-A System II は、今後の会話音声認識システムの研究の基礎となりうる。
- (2) 言語処理での具体的アルゴリズムを提案し、その有効性が確かめられた。とくに、音韻うティスを入力とした場合の tree search による単語認識アルゴリズム、リスト表現による構文解析アルゴリズムなどが提案された。これらの手法は、会話音声認識システムの言語処理での基本的な手法となりつつある。また、音響処理と言語処理との間の情報の受け渡しとして、音韻うティスの形式を提案した。音韻うティスの形式も様々な研究機関で用いられ始めており、その有効性が確かめられている。

主論文の要旨

報告番号	※甲第	号	氏名
<p>(3) Voice Q-A System II の音響処理を音韻単位のセグメンテーションに重点をおいて、再帰的な構成で作成した。この構成法は、今後の音響処理の基本となり得る。また、この音響処理は、LPC分析に基づいて作成されている。諸研究機関でもLPC分析に基づいた音声認識の研究を始めており、それらの研究に対して、先導的な役割を果たすと考えられる。</p>			
<p>(4) LPCスペクトルマッチング尺度について、母音識別実験の観点から検討した。その結果、従来からの尺度のうちでは、残差パワーを正規化した条件でのLPCケプストラム距離とCOSH尺度がすぐれていることが明らかになった。さらに、ホルマントマッチングとスペクトルマッチングの長所を兼ね備えたPeak Weighted Measureを提案し、理論的解析および母音識別実験でその有効性を確かめた。Peak Weighted Measureのうち、とくに、WLR (Weighted Likelihood Ratio) 尺度はすぐれていることが明らかになった。このPeak Weighted Measureの提案により、スペクトルマッチングの新しい分野が開けた。</p>			

主論文の要旨

報告番号

※甲第

号

氏名

(5) 音韻タイプの自動評価システムを作成し、音響処理の客観的な評価方法を確立した。この方法は、会話音声認識システムの研究を進める上で、今後とも役立つと考えられる。

以上に述べたように、本論文の研究結果は、音声認識の基礎的な技術レベルの向上に寄与するところが極めて大きい。この研究成果をもとに、近一年中に実用的な会話音声認識システムが作成できることを確信する。