

専攻	システム情報	学籍番号	893409	指導教官氏名	中川 聖一
申請者氏名	甲斐 充彦				湯浅 太一
					井上 克巳

論文要旨

論文題目	自然発話のための音声認識システムに関する研究
------	------------------------

(要旨 和文 1,200 字程度)

(1)

音声認識の研究では、近年の計算機の処理能力の向上と、大量のデータベースを用いた統計的な手法の導入により大きな進展を見せた。これらの多くは、従来の制約であった孤立単語認識から連続音声認識へ、特定話者から不特定話者の認識へ、小・中語彙から大語彙へ、と目標を高くすることを可能にした。しかし、従来の音声認識技術はていねいに朗読された音声を対象とし、いわゆる書き言葉用の文法に従う発話を仮定してきたため、そのような制約から外れた発話に対しては十分な認識性能を保持することができなかった。また、自然な発話では、発音があいまいな冗長語や、言い直し、言い淀み、助詞落ちなどの現象が見られるため、従来のように音素や単語レベルの知識と、より上位の構文、意味的な知識を分割したアーキテクチャでは、認識性能の限界が予想された。

本論文は、こうした背景で、まず文脈自由文法による構文知識を音声照合の処理に採り入れ、統合的な音声の仮説照合と探索を行なうアルゴリズムを提案した。一つは、従来の単語レベルの照合を行なうワードスポッティング法の利用による拡張連続 DP 法の原理に基づいて実現し、もう一つはパターンマッチング問題としての連続音声認識において最適な探索を行なう One Pass DP 法の原理に基づいて実現した。前者は音声照合の計算量が語彙数のオーダである効率的な手法で、後者はビームサーチ法を用いて構文制約を動的に展開し、照合仮説の枝刈りを行なうことで計算量の増加を抑えた。実験的な検討に基づいて、それぞれのアルゴリズムの有効性を示した。

また、本研究では未知語・不要語を扱うための手法を検討した。一般に

連続音声認識ではサブワード単位の音響モデルを用いることが多いため、未知語も何らかのサブワードの系列で表される。そこで、未知語としてのモデルを、任意のサブワードモデルの接続に対応させることで、既知語と未知語のそれぞれの尤度の比に基づく未知語の検出が考えられる。初めに、未知語、冗長語をそのような方法で処理することを試みた。このような方法は登録語以外の発話一般に対して適用できるので、発話のリジェクションとしての有効性も考えられる。このような手法の有効性を客観的に知るため、孤立単語認識のシミュレーション実験によって単語認識性能と未知語検出性能との関係を求め、実音声による未知語検出の実験においても同様な傾向があることを示した。

音声認識処理において間投詞、言い直しなどを扱うためには、一般に不要語としての処理が必要であり、倒置、助詞落ち、非文法的な発話などに関しては言語解析に関する対処が必要である。これまで、自然な発話においてどのような認識手法が有効であるかの明確な比較はなされていない。そこで本研究では、不要語としての照合方式や言語解析法が異なる複数のシステムを実現し、自然な発話に対する認識実験によって比較・評価した。そして未知語検出などのために検討したサブワードモデルに基づく未知語処理法が自然な発話の認識において有効に働くことを明らかにし、さらに構文・意味的レベルの制約が統合された認識手法の有効性を示した。