

# 雑音下音声認識のための重み付け分散拡大に基づく 単語 HMM の耐雑音性の改善

漢野救泰\* 船田哲男\*\*

## 研究の背景

工場のような実環境下では、非定常な高騒音の発生など周囲の状況の変化により、雑音のスペクトルや SN 比が急激に変動することがある。認識時の雑音の種類や SN 比が学習時や適応時と異なる場合、音声認識性能は著しく低下する。このため、実環境下で音声認識を可能にするためには、未知の非定常雑音が混入した場合などの雑音条件変動に対して、HMM(隠れマルコフモデル)を頑健にする必要がある。

## 研究内容

本研究では、HMM の耐雑音性向上のため、HMM の各状態あるいは各分布のパワーによる重み付け分散拡大を行い、雑音の影響を受けやすい状態からの出力確率を、観測ベクトルの違いにより大きく変動しないように制御する手法を提案した。

本手法による認識性能改善効果は、語彙数 50 の単語音声と 2 種類の工場の雑音を使用した不特定話者単語認識実験で評価された。HMM としては、クリーン音声 HMM と 2 種類の雑音付加音声 HMM を使用した。15dB の SN 比の雑音付加音声で学習した HMM(15dB-HMM)と 5 種類の SN 比の雑音付加音声で学習した HMM(5SNR-HMM)である。そして、定常時が比較的静かな環境であればクリーン音声 HMM, 騒音環境では雑音付加音声 HMM を使用し、その雑音条件での最高の認識性能を維持するとともに、異なる雑音条件となった場合でも従来の性能低下を改善する方法について検討した。図は、15dB-HMM を使用して、学習時とは異なる雑音を付加し、広範囲の SN 比での認識性能(単語認識率)を、分散拡大した特徴パラメータごとに示している。all(1)は通常の HMM を表し、(3,4,5,6)D はパワーの低い分布に対して拡大率を大きくしたことを表している。また、[ $\cdot$ ]は分散拡大したパラメータを表している。図で、pow は正規化対数パワー、cep はケプストラムであり、 $\Delta$  pow,  $\Delta$  cep はそれらの動的パラメータである。

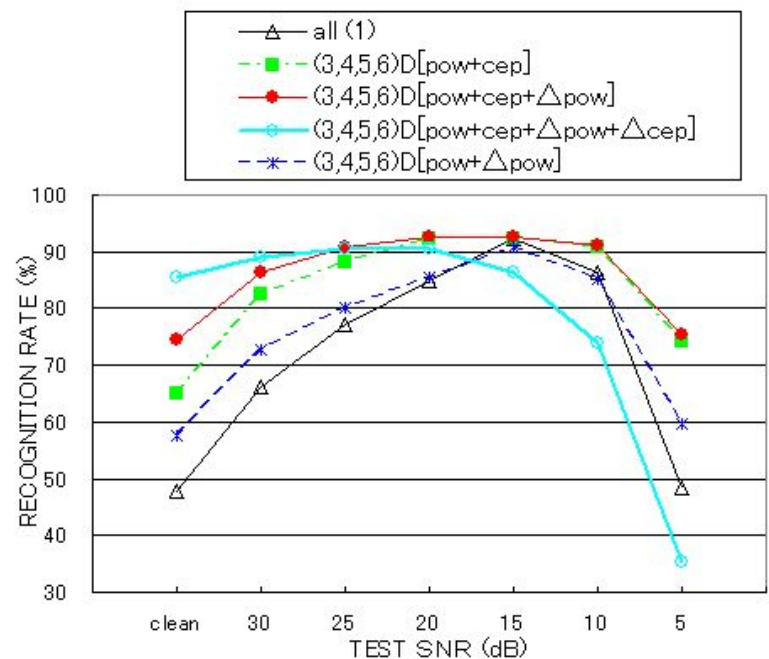


図 15dB-HMM の各種 SN 比における認識性能

## 研究成果

評価実験の結果、いずれの HMM においても、以下の確認ができた。

- (1) 雑音条件の変動に対して広範囲の SN 比における認識率が向上し、単語 HMM の耐雑音性が改善できた。
- (2) 重み付け分散拡大は拡大率一定よりも認識性能を顕著に改善できることが明らかになった。
- (3) 改善に最適な分散拡大条件(特徴パラメータ, 拡大率)が得られた。

## 論文投稿

電子情報通信学会論文誌 D-II Vol. J87-D-II, No. 5, 2004. p.1052-1061.

\*電子情報部 \*\*金沢大学工学部