

# P97 教師なしアンサンブル適応法の提案と 音響モデル適応への応用

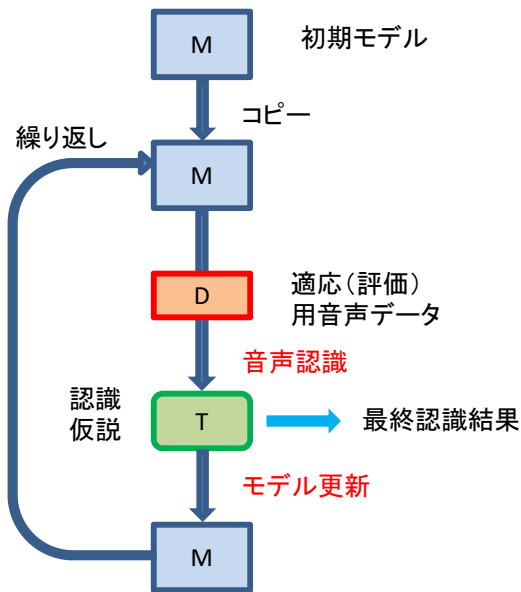
篠崎 隆宏, 久保田 雄, 古井 貞熙

東京工業大学 計算工学専攻 [www.furui.cs.titech.ac.jp](http://www.furui.cs.titech.ac.jp)



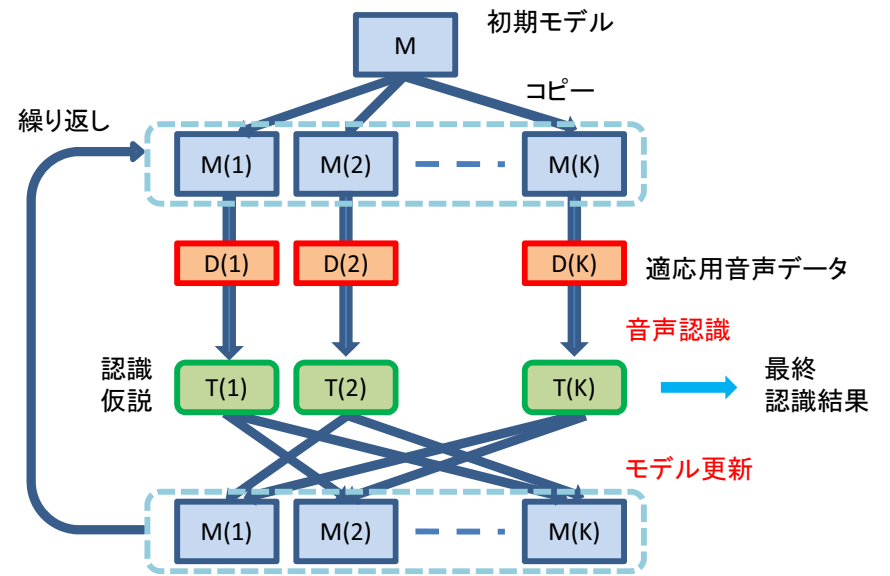
- 音声認識において、**バッチ型教師なし適応**は人手による書き起こしデータを用いずに高い認識精度を得る手法として広く用いられている。
- バッチ型教師なし適応では音声認識により認識仮説を求め、それをラベルとして用いてモデル更新を行うことを繰り返す。**認識仮説には誤りが含まれるため、適応化性能が限られる**問題がある。
- アンサンブル的手法を教師なし適応に応用することで適応化性能を向上させることを目的とし、**教師なしクロスバリデーション(CV)適応法**および**教師なし集合(Ag)適応法**を提案した。大語彙音声認識実験により顕著な適応性能改善効果があることを示す。
- 提案法は一般性があり、音声認識に限らず**幅広い応用**が考えられる。

従来型バッチ型適応では、初期モデルを用いた認識対象音声の認識と、認識結果を用いたモデル更新を繰り返す。モデル更新には種々の手法が用いられる。



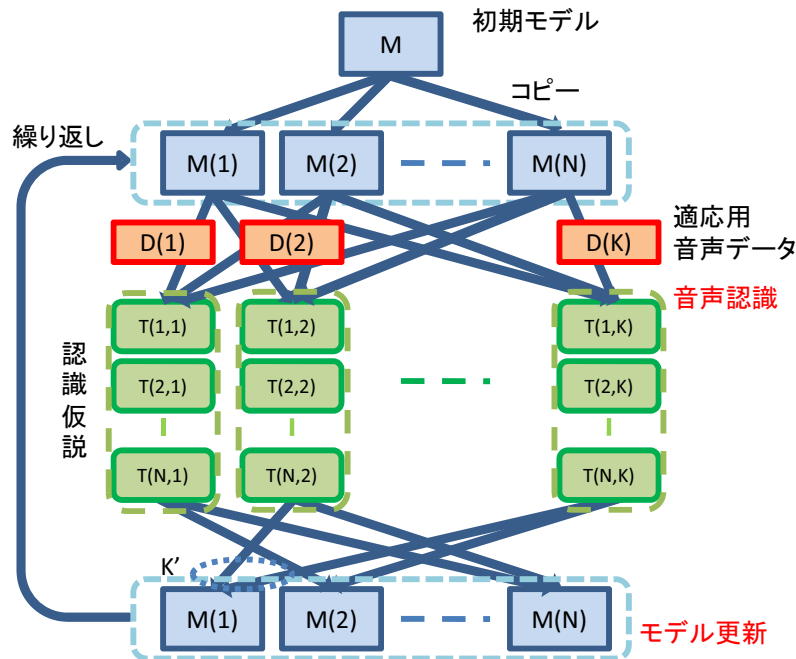
従来型教師なしバッチ適応法

教師なしCV適応では、モデル適応に用いる認識対象音声データをK個の排他的なサブセットに分割する。認識ステップとモデル更新ステップで使用するデータをCV法により分離することで繰り返しループ中での認識誤りの影響を低減し、適応化性能の向上を図る。



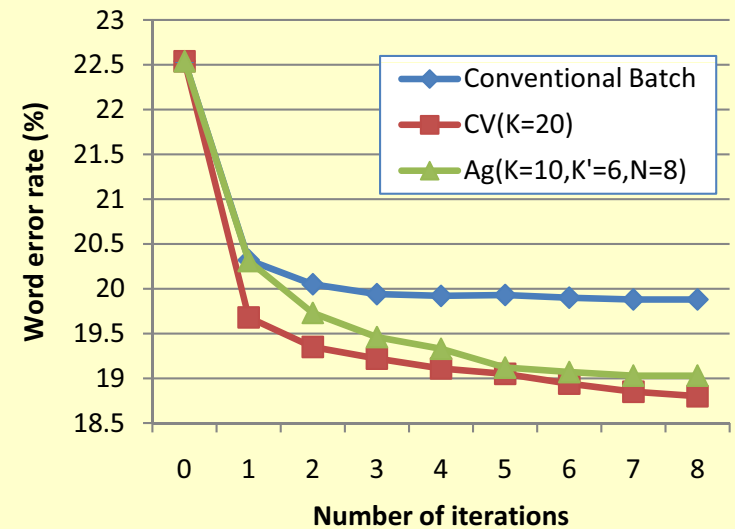
教師なしクロスバリデーション(CV)適応法

教師なし集合適応では、K個に分割した認識対象音声のサブセットそれぞれをN個のモデルを用いて認識する。K個中ランダムに選んだK'個のサブセットに対応するNK'個の認識結果を用いてN個のモデルそれぞれを更新する。N個のモデルを用いた認識結果を同時にモデル更新に用いることで、個々のモデルによる認識誤りの影響を低減し、適応化性能の向上を図る。



教師なし集合(Ag)適応法

### 実験結果 (音響モデルの教師なし話者適応)



#### 関連文献

- T. Shinozaki, Y. Kubota, and S. Furui, "Unsupervised cross-validation adaptation algorithms for improved adaptation performance," ICASSP, pp. 4377--4380 (2009).
- 篠崎 隆宏, 「音響モデルのアンサンブル学習」, 2009-SLP-77, No. 11 (2009-7).