

TR-H-029

DTWに基づく音声認識アルゴリズムと
統計的参照パターン設計法

小森 隆

1993. 9. 24

ATR 人間情報通信研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-chō, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011

Facsimile: +81-7749-5-1008

概要

DTWとは、本来、プロトタイプを用いる距離分類器のための非線形時間伸縮(正規化)法である。特に、連続音声認識のために様々な実装法が提案されており、計算量や記憶容量の観点からその優劣が競われてきた。最近、音声認識のための基盤技術としての座を隠れマルコフモデル(hidden Markov model; HMM)に明けわたした感もないではないが、最適時間伸縮関数の探索というDTWの理念は、HMMにおいても脈々と生きている。従って、より本格的な音声認識システムの設計のために不可欠なDTW探索の問題を理解することの意義は、極めて大きいと思われる。

本技術覚書は、ATR 音声翻訳通信研究所およびATR 人間情報通信研究所の共催による音声信号処理に関する一連の講演「ATR Regular Talk on Speech Processing」の一環として、平成5年8月4日、9日の2回にわたり本稿と同じ題名で行われた講演の配布資料およびスライドに若干の修正を加えてまとめたものである。本稿が、これから音声認識に取り組もうと思っている技術者にとって有益なものになれば幸いである。