

論文の内容の要旨

論文題目 Associative-Processor-Based Intelligent Data Processing Systems
(連想プロセッサに基づく知的データ処理システム)

氏名 顧 清栄

(本文) 知恵生物と機械の最も大きな違いの一つは知恵生物が生まれつきの学習能力を持つことである。この周辺環境から過去経験のないことを学習する能力は知恵生物が生き残るために長い歳月かけて厳しい生存競争を通してやっと身に付けたものである。学習能力のなかに最も重要な役割を果たすのは特徴抽出と特徴マッチングと呼ばれる二つの能力である。特徴抽出と特徴マッチングの能力はすべての知的な活動のもとであると指摘する研究者は大勢にいる。柴田研究室は特徴抽出と特徴マッチングを主要な研究方向として研究を進んでいる。ベクトル演算に基づく特徴マッチング専用の連想プロセッサはすでに何種類も開発されていた。私の博士課程の主な研究テーマは、1. 学習データからベクトル形の特徴を抽出する汎用なアルゴリズム、2. 連想プロセッサに基づく認識ための高い計算効率のモデリング手法、という二つのことである。

ベクトル量子化の手法を用いて音声の波形をそのまま圧縮することは私の研究のスタートポイントとなった。音声波形の特徴が有効に抽出されるかどうかは一つの非常に簡単な数値 SNR(signal to noise ratio)で評価できるということは音声圧縮を選択した理由であった。ベクトル量子化は新しい手法ではないが、アナログの連想プロセッサ(当時、デジタルの連想プロセッサはまだ開発されていなかった。)とアナログメモリを圧縮システムのハードウェアの基盤となって、この制限条件の下でよい再生音質を得るには実に難しい挑戦であった。波形の特徴量或いは特徴ベクトルの抽出を最適化させるために、やれることをすべてやり尽くしたくさんのシミュレーションをした末、圧縮率は4と相当する時平均 18.3dB の再生音質が得られた。

シミュレーションの際、柴田直教授が音声波形の予測エラーから抽出した特徴ベクトルの使用頻度ヒストグラムは話者のモデルとして使えるではないかと指摘した。この指摘は直ちに一つの改善した汎用な特徴抽出アルゴリズムを導き出した。学習データなかの共通的な性質をフィルタによって取り除く、そして残った様々な個性を含む部分から典型的な特性を代表する特徴ベクトルを抽出するという手法である。この手法の有効性と簡単さは歯列並びのデータベースから典型的な歯列の並びかたを見つけないというアプリケーションによって検証された。特徴抽出時のパラメータを調整することで、ベテランの歯医者が

主観的に分類した結果と全く同じの結果を自動的に抽出することが成功した。更に、その際のパラメータを詳しく分析することによって、歯医者が患者を診断する時に無意識にある影響を受けていることが分かった。その影響のもととなる解剖学上の裏付けも明らかにした。この点は大阪大学歯学部付属病院矯正科高田健治教授に高く評価された。

同じ手法で抽出した歯列並びの異常パターンは歯医者が患者を治療する時のリファレンスモデルとして使えば、今までの定性的な治療しかできないか、歯医者の経験に任せる任意性が高い歯列の矯正治療は定量的しかも精密な治療に変わると考えられる。その上、それぞれの典型的な異常パターンの過去の矯正治療記録を分析すれば、治療前に治療効果の見込みも可能となる。矯正歯科の治療現場にとって非常に役に立つ研究成果であった。

柴田直教授が注目した特徴ベクトルの使用頻度ヒストグラムから高い計算効率を持つ認識対象のモデル作り方を開発された。通常ベクトル量子化の手法でものを認識しようとすると、認識対象ごとに認識対象のモデルとなるコードブックを作る必要がある。それと比べて、我々の手法ではすべての認識対象に対してただ一つの共通なコードブックと作っておけばよい。そして、この共通なコードブックを用いて、それぞれの認識対象の学習データに対して、ベクトル量子化の操作が行われ、共通なコードブックなかのコードベクトルの選択される回数を数える。数えた結果すなわち使用頻度ヒストグラムは認識対象のモデルとなる。従来の手法と比べれば、モデルを作るでは計算量の最も多いコードブックの計算は一回だけで済み、認識段階では従来の認識の判断標準となる SNR のような値の計算必要はなくなり、その代わりに単純な回数を数えるだけで済む。従って、モデルの作る段階でも、対象を認識する段階でも、必要な計算量は大幅に削減された。更に、もし求めた共通なコードブックには十分な汎用性があれば、新たな認識対象が追加されるたびに今のコードブックをそのまま使って、対象の使用頻度ヒストグラムを計算すればよい、コードブックを更新する必要はない。従って、このような認識システムは優れた拡張性を持っている。

二つ違うアプリケーションで、このモデリング手法の有効性を検証した。一つは text-dependent 話者認識システムであり、もう一つは言語認識システムである。

text-dependent 話者認識システムにおいては二つ少し違うモデリング手法で検証が行った。話者全部で 14 人がいる。そのうち男性は 6 名、女性は 8 名である。一つめの手法では共通なコードブックが直接に全ての話者の声を含む大きな学習データから求められ、そして、このコードブックを用いてそれぞれ話者の声だけを含む小さい学習データに対しベクトル量子化の操作が行われ、話者のモデル即ち使用頻度ヒストグラムを計算するというやり方である。二つめの手法ではまず全ての話者の声を含む大きな学習データから一つの予測機を作っておき、共通なコードブックはこの学習データの予測エラーから抽出され、その後のやり方は一番の方法と全く同じである。一番目の手法の平均認識率は 70%、二番目方法の平均認識率は 74% である。ほかの話者認識システムと比べて認識率は高くないが、このシステムの単純さと計算の簡単さを配慮すると、かなりよい認識率である。

言語認識システムにおいては上記の二番目の手法で検証が行った。言語の種類は二種類、英語と日本語である。平均認識率は 75%であり、これはほかの言語認識システムと匹敵できる認識率である。従った、我々のモデリング方法の有効性はこれら二つのアプリケーションによって検証されたと考えられる。そのうえ、このモデリング手法は特定の応用に限られなく、汎用性を持つこともこれら二つのアプリケーションによって実際に検証された。

最後結論としては私が博士課程の間に連想プロセッサに基づいて、汎用な特徴抽出アルゴリズムと高い計算効率のモデリング手法を開発し、そして実際のアプリケーションによってそれらを検証した。