

# ニューラルネットワークを用いた音声認識技術の研究\*

鷗野 俊寿\*<sup>1</sup>

Study on Speech Recognition Technology using Neural Network

Toshihisa HIBARINO

電子機器を開発する際に、内部回路を自由に書き換えることが可能なFPGA（フィールド・プログラマブル・ゲート・アレイ）を利用すれば機器の高速化、短納期化に寄与することができる。本研究では、FPGAを利用した音声認識機器の開発を行うため、ニューラルネットワークによる認識アルゴリズムの研究を行った。

キーワード：FPGA、HDL、ニューラルネットワーク、音声認識

## 1 はじめに

近年、電子機器の高性能化、低価格化にともなって、パーソナルコンピュータや携帯電話などの情報機器が広く普及しつつある。情報機器が一般家庭へ普及するにつれてユーザー親和性に優れた音声によるユーザーインターフェースが注目され、パーソナルコンピュータ上で動作する音声認識ソフトや音声認識機能をもった携帯機器などの開発が求められている。

そこで本研究ではFPGAを利用した音声認識機器の開発を行うため、ニューラルネットワークによる認識アルゴリズムの研究を行ったので報告する。

## 2 方法

マイクロフォンから入力された音声信号はA/D変換器でデジタル信号に変換され、FFT(Fast Fourier Transform)で周波数成分に分解される。その後、ベクトル量子化された音声信号が認識処理部に渡される。

音声認識の手法としては、DP(Dynamic Programming)マッチングによる方法、HMM(隠れマルコフモデル：Hidden Markov Model)による方法がよく知られているが、本研究ではニュー

ラルネットワークモデルの1つである自己組織化ネットワークを用いて音素レベルでの認識を行う。このネットワークモデルは教師信号なしの競合学習を行いパターンを分類する能力に優れているため、未学習の多様な音声信号に対しても柔軟かつ効率よく認識することが期待できる。

音声認識機器の開発にはFPGAを用いる。このICは内部回路を自由に書き換えることが可能なため、複雑な信号処理を高速に、コンパクトな回路で実現することができる。FPGAの開発はHDL(ハードウェア記述言語：Hardware Description Language)での言語設計を行う。

また、FPGAのリコンフィギャブルな特性を利用することで、音声認識の学習の過程でニューロンの数やネットワークの形態を動的に変化させ、認識に最適な回路を構築するようにする(図1)。

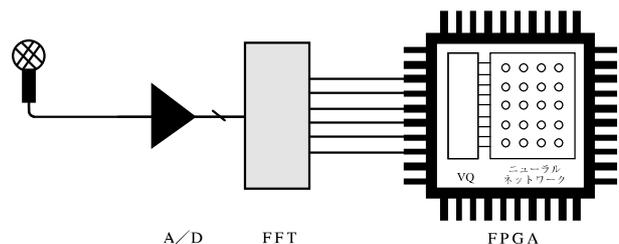


図1 音声認識機器の模式図

\* FPGAを利用した音声認識技術に関する研究  
(第1報)

\* 1 機械電子・デザイン部

### 3 結果及び考察

FFT変換された音声信号を、自己組織化ニューラルネットワークモデルを用いて音素単位のグループに分類するプログラムの開発を行った。

音声信号を10msごとにFFT変換し、ベクトル量子化を行う。ニューラルネットワークへ入力された信号は教師なし学習を行い、自己組織化Featureマップ上で各音素ごとの領域に分類される(図2)。

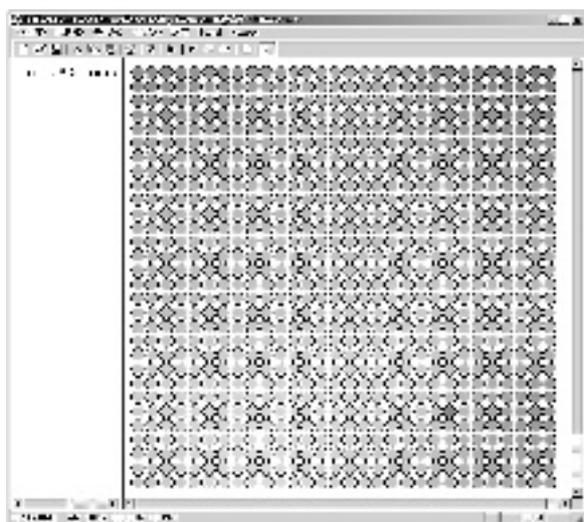


図2 音声信号の自己組織化Featureマップ

プログラムの動作確認を行った結果、大まかな分類は可能だが、一部の音素(特に破裂音)については誤認識が多いことが分かった。

### 4 まとめ

FPGAを利用した音声認識機器の開発を行うため、ニューラルネットワークによる認識アルゴリズムの研究を行った。より認識率を上げていくために、今後は、音素の時系列の変化を認識アルゴリズムに取り組んでいく予定である。

### 5 参考文献

- 1) 長谷川裕恭: VHDLによるハードウェア設計入門, CQ出版社(1995年)
- 2) R・ピール, T・ジャクソン: ニューラルネットワーク情報処理, 海文堂出版(1993年)