

研究概要報告書

資料 - 1 2

(/)

研究題名	ニューラル・ネットに基づく音声情報圧縮	報告書作成者	森島繁生
研究従事者	森島繁生		
研究目的	<p>バックプロパゲーションを学習規則として用いた多層パーセプトロンは、パターン認識や信号処理など多くの分野で注目され、活発に研究が行われている。同一層内にリンクが存在しない多層パーセプトロンにおいて、中間層のユニット数を入出力層のそれよりも減少させることにより、中間層において入力信号の情報圧縮が可能となる。このような3層砂時計構造を有するニューラル・ネットが学習した内部構造について検討を加える。また、中間層出力の量子化による音声符号化器を実現し性能評価を試みる。次に、5層砂時計構造を有するネットワークについて同様の実験的考察を行い、その特性解析と符号化器実現の可能性について検討することを目的とする。</p>		
研究内容	<p>ニューラルネットが信号の冗長度削減のための最適な圧縮アルゴリズムを自動的に習得することを期待し、入力データおよび学習教師データとして音声信号サンプルをそのまま用いることとする。この時、入出力層のユニット数と中間層のユニット数を任意に設定して、フレーム周期と圧縮率をコントロールする。また、中間層出力値を量子化することによって音声符号化が可能となる。学習方式はバックプロパゲーションを用いており、非線形関数としてシグモイドを用いている。3名の話者のデータを用いて実験を行った結果、恒等写像を実現する3層ニューラル・ネットは大部分シグモイドの線形部分のみを利用してマッピングを実現していることが解った。2次元空間上のランダムに8つ選択した点を用いて、図1に示すネットワークで実験を行った。この結果、図2に示すような再現線が引かれ、3層ネットは線形写像を学習していることが解った。●は学習点、○は出力層に出現した再現点を示す。次に図3のような5層ネットワークを用いて同様の実験を行った。この学習によって実現された恒等写像によって生じる再現線の様子を図4に示す。この場合、学習点上をきれいに通る曲線が引かれ、学習点が誤差なく出力層に現われていることが解る。したがって、5層の場合には非線形特性を生かした有効な情報圧縮の可能性はある。3層ネットによって実現される音声符号化器は学習点からの最小2乗誤差を実現する線形空間上への写像に過ぎず、従来の圧縮手法の性能を上回る可能性は薄い。5層の場合には非線形特性を生かした有効な音声符号化器の実現可能性を確認した。</p>		

様式 - 9

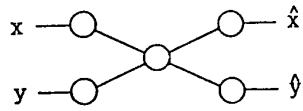


図1 実験用3層ネットワーク

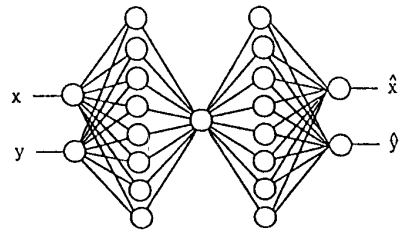


図3 実験用5層ネットワーク

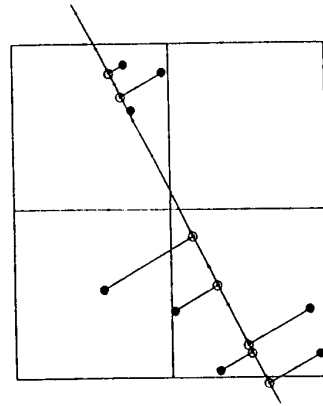


図2 3層ネットによる再現線の様子

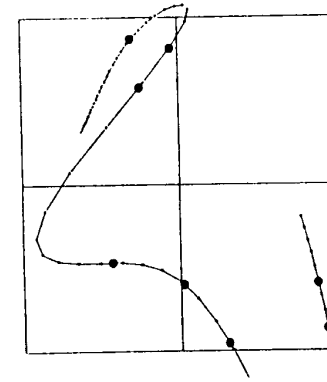


図4 5層ネットによる再現線の様子