

研究種目：若手研究（B）

研究期間：2007～2009

課題番号：19700165

研究課題名（和文） 統計的アプローチによる画像変動に頑健なモデル構造の探究

研究課題名（英文） An exploration of robust model structures for image variations based on statistical approaches

研究代表者

南角 吉彦 (NANKAKU YOSHIHIKO)

名古屋工業大学・大学院工学研究科・助教

研究者番号：80397497

研究成果の概要（和文）：

本研究の目的は、画像認識における様々な変動に対応可能な汎用的な画像モデルを提案することである。この問題に対して、これまで画像の位置ずれや大きさの変動に対応可能な確率モデルとして、分離型格子HMMが提案されてきた。本研究では、分離型格子HMMに確率的主成分分析や因子分析といった特徴抽出を含むモデル構造を組み込むことにより、認識率の改善を行った。また、より複雑な変動に対応するため、分離型格子HMMを拡張し、画像の回転を表現可能なモデルや状態継続長を組み込んだモデルを提案した。

研究成果の概要（英文）：

The objective of this research is to find general model structures which are robust to image variations for image recognition. Separable lattice HMMs (SL-HMM) have been proposed as statistical models which can deal with location and size variations. This research proposed a new model structure which integrates linear feature extraction based on probabilistic principal component analysis and factor analysis into SL-HMMs. Furthermore, SL-HMMs were extended to models which can represent rotation variations and which include explicit state duration models.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,500,000	0	1,500,000
2008年度	800,000	240,000	1,040,000
2009年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,200,000	510,000	3,710,000

研究分野：音声・画像情報処理

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：パターン認識，画像認識，統計モデル，隠れマルコフモデル

1. 研究開始当初の背景

近年のコンピュータの普及により、人間と同様の手段でコミュニケーションのでき

る視覚・聴覚情報を用いたインターフェイスの開発が期待されている。また、音声認識や画像認識は、その重要な要素技術であり、盛んに研究が行われている。しかし、音声認識では、隠れマルコフモデル(HMM)に基づく手法が確立されつつあるのに対し、画像認識では、認識対象が多さとデータの複雑さから様々な手法が乱立している状態である。そのため、様々な画像に適用可能な画像認識のためのスタンダードな統計モデルの構築が期待される。

2. 研究の目的

本研究では、顔画像認識、文字認識、ジェスチャ・手話認識、リップリーディングなどの様々な画像に適用可能な統計モデルを考え、音声認識におけるHMMのような、画像認識のための汎用的な統計モデルの構築を目指す。これまでの画像認識の研究は、大きく以下の2つに分類することができる。

手法1. 特定のタスクに対して、人間の経験的な知識を積極的に利用する方法

手法2. 画像認識を多次元特徴空間のパターン識別として、客観的に捕らえる方法

手法1は、実用を重視した手法であり、少量の学習データでも高い認識率が得られる可能性があるが、タスクに応じて人間が思考錯誤を繰り返す必要がある。一方、汎用的な画像認識の枠組を考えれば、手法2を選択する必要があるが、画像認識の問題をパターン識別の問題として取り扱うまでには、画像の切り出し、正規化、特徴抽出などの前処理が必要となる。しかし、手法2の多くの研究では、これらの前処理は研究の対象とせず、人手による画像の切り出しやヒューリスティックな正規化手法を用いることが多かった。本研究では、これらの前処理を含むモデル構造を考え、モデルの学習と同時に最適化する枠組みを構築する。認識対象が含まれた原画像を直接モデル化することにより、煩雑な前処理を自動化することができる。また、手法2では、学習データが画像であることを考慮してい

ないものが多いが、提案するモデルでは、多くの画像に共通するデータ構造(例えば、観測が2次元に配置されているなど)を考慮し、認識対象に対して汎用性を確保しながら、多くの画像データに共通の構造を組み込んだ統計モデルの構築を目指して研究を行った。

3. 研究の方法

これまでの研究で、画像の性質を考慮したモデル構造として、分離型格子HMM(分離型格子HMM)を提案した。このモデルは、隠れマルコフモデルの多次元への拡張であり、対象物の位置や大きさの変動をマルコフ連鎖を持つ確率変数として表現したモデルである。本研究では、このモデルを拡張し、画像の回転などのより複雑な画像変形を考慮したモデル構造について探求する。さらに、本研究では、確率的主成分分析や因子分析、線形判別分析などの特徴抽出を含むモデル構造を分離型格子HMMに組み込むことを考える。構築された画像モデルでは、画像変動に頑健な特徴量を自動抽出することができ、認識率の改善が期待できる。

4. 研究成果

本研究では、これまでに提案されてきた分離型2次元HMMに、確率的な主成分分析や因子分析に基づいた特徴抽出手法を組み込んだ新しいモデルを定義した。本研究では、このモデルを可変固有顔モデルと呼ぶ。確率的な主成分分析や因子分析は、画像認識の代表的な手法である固有顔法や部分空間法に対応する確率的な手法であり、これらの手法をモデル構造として組み込むことにより、大幅な認識率の改善が期待できる。提案モデルでは、様々な画像変動のうち、位置や大きさの変動のような幾何学的な変動は分離型2次元HMMの構造で、照明条件などの違いによる画像変動には、確率的な主成分分析や因子分析の構造で吸収することにより、画像変動にロバストな認識が実現可能となる。

可変固有顔モデルの学習には、EMアルゴリズムと呼ばれる学習アルゴリズムを用い

る。このアルゴリズムでは、画像変動を表す隠れ変数の期待値計算とモデルパラメータの推定を交互に繰り返す。しかし、理論的には導出されたEMアルゴリズムで提案モデルの学習は可能であるものの、実際には隠れ変数の組み合わせが膨大になるため、計算量的に実現は困難である。そこで本研究では、変分近似に基づいた学習アルゴリズムを導出した。このアルゴリズムでは、期待値計算における事後確率に独立性を仮定し、事後確率分布の要素分布を個別に更新することにより計算量の削減を行っている。

提案した可変固有顔モデルの性能を評価するために、顔画像認識実験を行った。従来の部分空間法に対応する確率的成分分析や因子分析に対して、提案手法では大幅な認識率の改善が得られた。特に意図的に位置ずれや大きさの変動を加えたデータに対しては、従来手法はこれらの画像変動に対応できないため、大幅な認識率の低下が見られたが、提案手法では分離型2次元格子HMMの構造により、変動を吸収し認識率低下を防ぐことができた。

分離型格子HMMは、画像の位置ずれや大きさの変動に対応可能であったが、より複雑な変動に対応するために、分離型格子HMMを拡張し、回転変動も表現可能なモデル構造を提案した。このモデルでは、分離型格子HMMの縦・横の状態遷移に加えて、傾きを表現するシフト状態を導入することにより、回転を含む複雑な画像変形を表現している。また、このモデルのための学習アルゴリズムを導出し、顔画像認識に適用した。評価実験により、提案モデルが回転変動に頑健であることを示した。

従来の分離型格子HMMは、基本的にはHMMの多次元への拡張であるため、HMMが持つ問題点を継承している。その問題の一つとして、1次マルコフ性を仮定したことによる状態遷移をモデル化する際の脆弱性がある。すなわち、状態の継続確率が指数分布に限定されるという問題があった。この問題に対して、本研究では、分離型格子HMMに明示的な状態継続長分布を組み込んだモデル

構造を提案した。これにより、画像変動をより正確にモデル化できるようになった。

本研究で提案したモデルの学習には、変分近似に基づいたEMアルゴリズムを用いている。しかし、EMアルゴリズムには初期値依存性の問題が存在する。特に、変分近似を用いた場合、事後分布の要素分布を個別に更新するため初期値依存性の問題は、より深刻なものとなる。この問題を解決するために、本研究では、確定的アニーリングに基づくEMアルゴリズム(DAEMアルゴリズム)を適用した。それぞれのモデル構造に対して、アニーリングに基づく学習アルゴリズムを導出し、認識実験による評価を行ったところ、認識性能の大幅な改善が得られることを確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計5件)

- ① Yoshiaki Takahashi, Akira Tamamori, Yoshihiko Nankaku, Keiichi Tokuda, "Face recognition based on separable lattice 2-D HMM with state duration modeling," 2010 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), 査読有, pp.2162-2165, Dallas, Texas, U.S.A., March 14-19, 2010.
- ② Akira Tamamori, Yoshihiko Nankaku, Keiichi Tokuda, "An extension of separable lattice 2-D HMMs for rotational data variations," 2010 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), 査読有, pp.2206-2209, Dallas, Texas, U.S.A., March 14-19, 2010.
- ③ Yoshihiko Nankaku, and Keiichi Tokuda, "Face recognition using hidden Markov eigenface models," 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007), 査読有, Vol.

2, pp.469-472, Hawaii, USA, April 15-20, 2007.

- ④ 高橋 良彰, 玉森 聡, 南角 吉彦, 徳田 恵一, "状態継続長を考慮した分離型 2 次元格子 HMM による顔画像認識," 電子情報通信学会技術研究報告, 査読無, vol.108, no.484, PRMU2008-262, pp.153-158, 東北工業大学, March 2009.
- ⑤ 玉森 聡, 南角 吉彦, 徳田 恵一, "回転変動を考慮した分離型 2 次元 HMM による顔画像認識," 電子情報通信学会技術研究報告, 査読無, vol.108, no.484, PRMU2008-263, pp.159-164, 東北工業大学, March 2009.

6. 研究組織

(1) 研究代表者

南角 吉彦 (NANKAKU YOSHIHIKO)
名古屋工業大学・大学院工学研究科・助教
研究者番号：80397497

(2) 研究分担者

なし

(3) 連携研究者

なし