

	タマモリ アキラ
氏 名	玉森 聡
学位の種類	博士（工学）
学位記番号	論博第968号
学位授与の日付	平成26年12月17日
学位授与の条件	学位規則第4条第1項該当 課程博士
学位論文題目	STATISTICAL MODELS INCLUDING NORMALIZATION PROCESSES FOR IMAGE RECOGNITION (画像認識のための正規化プロセスを含んだ統計モデル)
論文審査委員	主 査 准教授 李 晃 伸 教授 北 村 正 教授 岩 田 彰 教授 本 谷 秀 堅

## 論文内容の要旨

For many years, many researchers of pattern recognition have developed the field of image recognition as the main focus of pattern recognition and various techniques have been proposed. Especially, statistical approaches based on Principal Component Analysis (PCA) such as eigenface methods and subspace methods show good recognition performance in many applications. However, if images contain geometric variations such as size, location and rotation, the recognition performance is significantly degraded. Therefore, normalization processes for such geometric variations are required prior to applying these methods.

In many image recognition systems, the normalization process is included in the pre-process part of the classification, and heuristic normalization techniques are used. However, it is necessary to develop the normalization technique for each task, because such heuristic techniques usually use task dependent information. Furthermore, in image recognition, the final objective is not to accurately normalize images for human perception but to achieve a better recognition performance. Therefore, it is natural to use the same criterion for both training classifiers and normalization. This means that the normalization process should be integrated into classifiers.

HMM based techniques for image recognition have been proposed to reduce the influence of geometric variations. Geometric matching between input images and model parameters is represented by discrete hidden variables, and the normalization process is included in calculating probabilities. However, the extension of HMMs to multi-dimensions generally leads to an exponential increase in the amount of computation for its training algorithm.

## 論文審査結果の要旨

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。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。また近年のインターネットの普及やハードウェアの進歩により、様々な研究分野において、大量のデータに基づいた統計的な手法が主流になってきており、音声認識では、隠れマルコフモデル (Hidden Markov Model; HMM) に基づく手法や重み付き有限状態トランスデューサに基づき手法が確立されつつある。しかし、画像認識の分野では、認識対象が多さとデータの複雑さから多種多様な手法が提案されている状況である。

従来の画像認識の研究は、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、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、と大きく2つに分類することができる。前者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。また後者は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。しかし多くの研究では、これら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。また、後者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。

これまでに、照明条件や幾何学的変動といった画像の性質を考慮した統計モデルとして、分離型格子2次元HMM (Separable Lattice 2-D HMM) が提案されている。このモデルは、隠れマルコフモデルの2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として表現している。そして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可能である。しかしながら、分離型格子2次元HMMには後述するいくつかの問題点が存在することが知られている。そこで本論文ではこれらの問題点が解消されるよう分離型格子2次元HMMを拡張することで、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・タスクに適用可能な、画像認識ための高性能かつ汎用的な統計モデルを提案している。

まず、分離型格子2次元HMMでは対応が困難であった回転変動を吸収するためのモデルが提案されている。ある特定方向での観測ラインに関する状態アラインメントのシフトを表現する HMM状態系列を新たに導入する。これにより、位置・大きさの変動だけでなく回転変動にも対応可能となることが期待される。位置や大きさの変動だけでなく、回転変動を含む画像認識実験の結果、提案モデルは分離型格子2次元HMMと比較して良好な認識性能を持つことが示された。さらに、状態アラインメントを可視化することで、提案モデルは認識対象の位置や大きさだけでなく、回転変動を正しく正規化できることが確認された。

続いて本論文では、静的及び動的特徴を含む特徴ベクトルを状態出力ベクトルとする分離型格子2次元HMMに対して、静的・動的特徴間の関係を明示的に導入し、分離型格子2次元HMMを再定式化したモデルが提案されている。分離型格子2次元HMMは従来、出力確率の条件付き独立性の仮定および各状態内で統計量が区分的定常というモデル化能力に関する脆弱性を持っており、提案モデルはこれらを克服するモデルとなっている。提案モデルのモデルパラメータの数は分離型格子2次元HMMのパラメータ数から増加することではなく、隣接する観測間の相関を効率よく捉えることが可能となる。画像認識実験の結果より、提案モデルは分離型格子2次元HMMよりも良好な認識性能を持つことが示された。

以上のように、本論文では汎用的な画像認識システムの構築のための、より高精度な統計モデルを提案し、その有効性を示した。また、本研究の内容は国内外の論文誌・国際学会にて公表されている。よって本研究は、情報工学の分野において寄与するところが多大であり、博士論文として充分価値あるものと認める。