

氏 名	中島 佑樹
学位の種類	博 士 (工学)
学位記番号	工博甲第 4 5 8 号
学位授与の日付	平成 3 0 年 9 月 2 1 日
学位授与の条件	学位規則第 4 条第 1 項該当
学位論文題目	特徴量変換に基づく物体識別に関する研究
論文審査委員	主 査 准教授 タン ジュークイ 教 授 森江 隆 " 大屋 勝敬 " 金 亨燮

## 学 位 論 文 内 容 の 要 旨

近年、デジタルカメラと計算機は高性能化、低価格化が進んでおり、膨大な画像を用いた画像処理に関する研究開発は世界中で盛んに行われている。その中で、最も注目されている画像処理技術の一つに画像識別がある。これは画像に付与されたクラスを識別する技術である。従来、人が目で見えて観測し、観測された物体が何であることを識別することの代替となる技術であるため、画像処理の中でも応用の幅が広い。画像識別ではカメラから得られた画像の RGB 値から特徴量に変換後、機械学習により識別する。特徴量の計算式の設計は手動で設計する手法と自動で特徴量の計算式を最適化する手法がある。本研究では手動で設計する手法から始めて、特徴量の計算式を最適化する手法までを提案している。

第 1 章では序論を述べる。人が目で見えて識別することを機械的に代替する試みの応用先は広い。特に近年の自動運転車や ITS(Intelligent Transport System)における画像識別の需要は大きい。さらに計算機性能向上や開発環境の整備といった技術的な背景を受けて画像識別の研究が盛んに行われている。

従来の研究では画像識別に有効な特徴量を研究者が手動で設計することが主流であった。2012 年の ILSVRC(ImageNet Large Scale Visual Recognition Competition)において AlexNet が提案されて以降は、現在に至るまで特徴量の設計をニューラルネットワークにより実現する手法が提案され、主流となりつつある。しかし、この手法には二つの問題がある。一つはニューラルネットワークでは微分可能な関数しか扱えないことである。微分可能でない関数による特徴量の記述による識別精度向上の報告があることを考えれば、これは識別精度向上を阻害している可能性がある。二つ目はセキュリティの問題である。ニューラルネットワークを用いた画像識別に対して、故意に別の画像に分類する手法が確立されている。特に One Pixel Attack と呼ばれる攻撃手法では 1

画素のみ値を変えるだけで別の画像であると誤認識させることに成功している。

本研究では、画像識別の特徴量を手動で設計する手法の提案から始まり、この問題に対してニューラルネットワークを用いない特徴量の計算式を最適化する手法を提案している。提案法では特徴量の計算式をグラフ理論のネットワークにより表現し、これを、組合せ最適化問題を解くことにより最適化する。

第2章では、特徴量を手動で設計する HOG(Histograms Oriented Gradients)拡張法を提案する。HOG は 2005 年に Dalal らによって提案された特徴量である。この手法は複数の拡張が提案されているが、それらを統合しつつ、別のパラメータに対しても拡張を行うことにより識別精度が向上できることを示す。

第3章では、自動での特徴量計算式の最適化手法を提案する。まず、特徴量の計算式をグラフ理論のネットワークで表現する。これを FTN(Feature Transform Network)と呼ぶ。FTN に対して評価関数と制約式を与えた組合せ最適化問題 FTOP(Feature Transform Optimization Problem)を定義する。次に、FTOP の求解法を提案する。FTOP の評価関数と画像識別の汎化性能に対する相関を計測することにより、FTOP を求解することが物体識別に有効であることを確認した。その結果相関係数 0.8973 が得られたため、有効であると結論付けた。

第4章では、FTOP を画像識別による歩行者識別に適用する手法を提案する。そのために FTN のノードに用いるべき関数や、識別に用いる機械学習と評価関数、およびそれらの連携方法を述べる。画像識別による歩行者識別に FTOP が有効であることを実験的に示すために従来法との比較実験を行った。手動で設計された sp-Cov とニューラルネットワークの手法である GoogLeNet との比較を実施した結果、提案法が最も良い結果を得た。FTOP の求解では評価値の計算に 97%以上の時間を要し、求解のボトルネックになっていることを指摘している。これに対し、アルゴリズム上の解決策として評価値計算の打切りが考えられることや、ハードウェア上の解決策として GPU(Graphic Processing Unit)や FPGA(Field Programmable Grid Array)の利用が考えられることを述べる。

第5章では、上記のまとめとして結論を延べると共に、FTOP 自体は画像識別に限定されず、分類問題であれば適用できる手法であることに触れ、FTOP の応用の広がりの可能性について述べている。

本研究の独創点は以下の通りである。まず、提案法はニューラルネットワークを用いない特徴量の計算式の最適化手法であることである。これにより微分可能でない関数を特徴量の計算式に含めながら最適化が可能である。また、以下の三点の機能を有している。(1)機械学習の計算環境によって特徴量ベクトルの次元数の上限が決まる。これには特徴量ベクトルの次元数を指定することにより対応できる。(2)画像識別が用いられるアプリケーションによって特徴量の計算時間の要件が決まる。これにはノードに使用する関数の処理時間をコストとして、FTN 全体のコストに対する上限を指定すること

により対応できる．(3)初めから最適化することも可能であるが，最適化処理時間の短縮のために従来研究のノウハウを用いることができるように，関数と関数の接続関係を制限する機能がある．

## 学 位 論 文 審 査 の 結 果 の 要 旨

ディジタルカメラと計算機の高性能化，低価格化は現代の科学技術の大きな成果であり，産業界やサービス業界，また高齢層から若年層社会に至るまで多大な利便性をもたらしている．一方インターネットの普及により，インターネット上にある膨大かつ多様なデータを活用して新たな技術革新が起きようとしている．近年，コンピュータビジョン分野においても，ビッグデータを用いた画像識別や解析の研究開発が国内外で盛んに行われるようになってきている．

画像内の物体を識別し解析するには，その物体の識別に有効と思われる特徴量を考案・設計し，その特徴量に基づく識別器を構築しなければならない．従来の物体識別技術では，特徴量は人が発見的に設計していたが，2012年以降は特徴量の設計をニューラルネットワーク(NN)により自動化する技術が広がっている．しかし NN では，中身がブラックボックスである上に，微分可能な関数しか扱えないという問題や，物体識別の結果を故意に操作して別のクラスに分類する手法が報告されているというセキュリティの問題がある．そこで本研究は，特徴量の発見的設計法を人検出を例として対比的に示すとともに，NN を用いずに，特徴量変換の計算式を最適化することによって特徴量の設計を自動化し，より高性能な識別器を構築する手法を提案している．

本論文では，著者はまず，研究の背景について述べている．機械による画像識別技術の重要性・応用分野の広さについて述べ，近年の技術の推移と，主要な技術となりつつある NN の問題点を指摘している．

次に著者は，特徴量の発見的設計法として，一般物体検出に有効な特徴量である HOG(Histogram of Oriented Gradients) 特徴量に対し，勾配方向のビン数の異なる複数のヒストグラムを用いるという新たな HOG 拡張法を提案し，その有効性を示すために，標準データセットの一つである INRIA Person Dataset を用いて実験を行い，提案法の方が従来法より識別精度が向上することを示している．

次に著者は，特徴量を自動取得する最適化法を提案している．近年物体識別によく用いられる AlexNet, GoogLeNet や YOLO(You Only Look One) は，NN と勾配法を用いて畳み込みフィルタの重みを最適化する手法であるが，これに対して提案法は，NN と勾配法を用いず，特徴量変換の計算式を最適化する手法である．提案法は，まず特徴量変換の計算式をグラフ理論のネットワーク構造を用いて表現する．これを FTN(Feature Transform Network)と呼ぶ．FTN に対して評価関数と制約式を与えた組合せ最適化問題 FTOP(Feature Transform Optimization Problem)を定義する．

FTOP の求解のために、機械学習で得られた結果を反復評価しながらグラフの近傍操作を繰り返し、最適解を取得する。

実験ではまず、FTOP の評価関数と画像識別の汎化性能との間には相関があるという仮説に対して検証実験を行い、その結果相関係数約 0.90 が得られ、この仮説が有効であることを確認している。次に提案法の有効性を示すために歩行者識別問題を選び、手動で設計された sp-Cov(Spatial Covariance)および NN の手法である GoogLeNet との比較識別実験を行い、提案法は従来法より次元数が最も小さく、識別性能が最も高いことを示している。

最後に著者は本研究をまとめ、今後の課題について言及している。

以上のように本論文は、NN を用いず、特徴量変換の計算式を最適化することによって特徴量の設計を自動化し、高性能な識別器を構築する手法を提案し、それを歩行者識別に適用して有効性を示している。提案法は NN の問題点を解決するとともに、従来の発見的設計法も改善している。また、NN におけるブラックボックス問題に対して、提案法では明確にパラメータが選択され、最適化過程においても可視化が可能である。よって本論文の成果は、計測工学、画像計測、特に物体検出・識別分野への貢献が大きいものと考えられる。

なお、本研究に関して、審査委員および公聴会における出席者から、HOG 特徴量の拡張法において、異なる複数の特徴量の増加による性能向上の可能性、特徴量の数およびその組み合わせ数の上限、CNN Inception の詳細、最適解において AUC(Area Under the Curve)と評価値との相関関係の関連、Adversarial example と呼ばれる学習済の NN を欺くような問題が提案法で発生する可能性等、種々の質問がなされたが、いずれも著者からの適切な説明によって質問者の理解が得られた。

以上により、論文審査及び最終試験の結果に基づき、本学位論文審査委員会において慎重に審査した結果、本論文が、博士（工学）の学位に十分値するものであると判断した。