学位報告

報告番号	※甲　　　第　　　号

主 論 文 の 要 旨

論文題目　Automatic construction of image feature extraction programs by using linear genetic programming

線形遺伝的プログラミングを用いる画像特徴抽出プログラムの自動合成

氏　　名　WATCHAREERUETAI　Ukrit

論 文 内 容 の 要 旨

本論文では，物体認識に対して，特徴抽出プログラムの自動合成のための線形遺伝的プログラミング（GP）に基づくシステムを記述する．我々は，そのようなシステムの効率性と有効性を改良する様々な方法を提案する．

研究の背景

本論文は，物体認識に対する画像特徴抽出のプログラミングの困難と煩雑さに関する問題を考慮する．通常、特徴抽出プログラムは専門家がデザインする．専門家は，解決する問題に対する有効な特徴を選択する．そして，そのような特徴を抽出できるプログラムを作成し，テストし，性能を評価する．もし，作成されたプログラムの性能が，十分良くないなら，ほかの特徴を選択し，プログラムを修正しなければならない．これは，大変な作業である．一般的に，専門家が自分の経験及び知識によって、プログラムをデザインする．さらに、画像から抽出できる特徴の数は非常に多く、専門家が全てを検討できない．したがって，有効な特徴が見つからないこともある。その上，問題領域の知識に基づくデザインされるので、特定の問題しか解決できず、他の問題に適用できない。したがって、他の問題を解決するために再デザインしなければならない。

代替手段として，専門家は、イメージから候補特徴のセットを抽出するプログラムをデザインし，forward selection や backward elimination などの特徴選択（特徴削減）テクニックを用い，候補特徴のセットから有効な特徴を選択する．また，主成分分析などの次元圧縮方法も利用でき，元々の特徴を新たな特徴空間に変換し，重要な特徴のみを選択する．このアプローチは，最終の特徴セットを自動的に選択するが，専門家は，自分で候補特徴セットを事前に選択し，その候補特徴セットを抽出するプログラムをデザインしなければならない．

特徴抽出プログラムのデザインの困難さに対処するため，さまざまな研究者が，特徴抽出プログラム（または，物体認識プログラム）を自動的に合成するシステムを研究してい

る．具体的に，それらの研究者はプログラムデザインの問題を最適化問題に変換し，計算機で解決する．このシステムは，ユーザ（専門家ではなくても良い）が，訓練画像と真値画像を入力し，評価関数のみを与える。システムが，ランダムに，原始オペレータで合成されるプログラムの集団（初期集団）を生成し，その後， GP などの進化アルゴリズム（EA）を用い，合成されるプログラムの性能を最適化する．このシステムは，同時に，特徴抽出と特徴選択を行う．特徴抽出の自動合成システムの利点は，次のとおりである記述する．1）これは，完全自動的なシステムである．2）専門家は不要である．3）問題領域の知識は不要である．4）合成されるプログラムは，考慮されるだけではなく，実際にテストし，評価される．5）このシステムは，専門家が考慮していない非通常のプログラムを見つけるかもしれない．

しかしながら，このようなアプローチの重要な問題は，通常の EA では，多くのプログラムを実行し，評価するので，最適化時間が長い．したがって，この問題を解決するために，効率で有効な EA が望ましい．さらに，ほとんどの従来手法は，一つの適合度関数で最適化する単目的アプローチであるが，実際には，我々は様々な目的を同時に達成するプログラムを希望することもよくあるので，効率で有効な多目的アプローチが必要である．その上，自動的に合成されるプログラムの性能は，人間がデザインするプログラムと同等の性能であるのかという質問を生じる．これは，重要な質問であるが，人間がデザインするプログラムと比較する研究は今まで極めて少ない

研究の目的

本論文の主要目的は，問題領域の知識なしで，物体認識のための特徴抽出プログラムを自動的に合成できる進化システムを開発することである．もっと具体的な目的は，前に記述された問題を解決すること，つまり，1）その進化システムのための効率で有効な EA を開発すること，2）進化システムに有効な多目的最適化方法を考案すること，3）進化システムで合成されるプログラムの性能を評価すること．

貢献

本論文の主要な貢献は下のように記述される．

1) 我々は，線形 GP のための正準変換を提案する．提案される変換は，冗長性がある元々の線形 GP の表現を，冗長性がなく，より単純なフォームに変換する．このフォームを，正準形と呼ぶ．正準変換により，我々は，二つの表現が同じプログラムを表すか否かを判定できる．提案される変換は，プログラムの実行と比べて，非常に速い．
2) 我々は，正準変換の利用により，線形 GP に基づく特徴抽出プログラムの自動合成システムの効率性と有効性を改良する方法を提案する．三つのテクニックを記述する．第一は，fitness retrieval，第二は，intermediate-result caching，第三は，prohibition of redundant individuals と呼ぶ．fitness retrieval テクニックは，進化システムを，以前に見つかったプログラムを表す個体（冗長な個体と呼ぶ）の実行を避けるようにする．このテクニックによって，80％の計算時間を減少することができる．intermediate-result caching テクニックは，最近実行されたプログラムの部分の実行を避けることができる．したがって，fitness retrieval を用いる線形 GP システムと比べて， 25％の計算時間を減少できる．prohibition of redundant individuals テクニック GP は，生成された子孫が冗長な個体であるか否かを判定し，もし，そうならば，冗長な個体にならないように，突然変異オペレータをその個体に繰り返して適用する．

これは，見つかっていない新しいプログラムを生成することを，GPに強制し，探索性能を改良できる．

3) 我々は，GPの学習を加速するために，hierarchical structure GP (HSGP)と言う新しいGP学習方法を提案する．HSGPは，2分木など木構造の設定に位置する多数の学習ノード（LN）により，進化過程を行う．進化過程は，一番下のLNから始まり，一番上のLNまで行う．それぞれのLNは，従来のGPを実行し，進化された集団を，上の段階のLNに送り，さらに進化させる．下の段階のLNは，より小さい学習サブセットで学習するが，一番上のLNは，全ての学習セットで学習する．このようにして，我々は，無駄の多い実行を減少でき，60-70%の計算時間を減らせる．
4) 我々は，新しい多目的GP (MOGP) を提案する．提案手法は，進化過程での冗長性を有効に制御し，GPの性能を改良できる．提案手法は，有名な進化的多目的最適化（EMO）の一つ，non-dominated sorting genetic algorithm 2 (NSGA-II)，を修正したものである．主要な変更は，冗長性を制限する三つのメカニズムの導入である．冗長性を制限する三つのメカニズムは，集団の多様性だけではなく，収束率も改善できる．提案するMOGPは，従来のEMOの二つ，NSGA-IIとstrength-Pareto evolutionary algorithm 2 (SPEA2)と比べて，有意により良い性能を与える．

論文の構成

本論文の構成は次のように，記述する．

第1章は，本研究の背景，目的，関連研究を記述し，そして，本論文の特徴と貢献を説明する．

第2章は，EA，GP及び本研究で用いる線形GPを紹介する．

第3章は，線形GPに基づく特徴抽出プログラムの自動合成システムを具体的に説明する．

第4章は，線形GPの表現の冗長性の原因を議論し，その表現を冗長性がない正準形に変換する方法を述べる．また，正準変換を利用し，線形GPの効率性と有効性を改良する方法も述べる．

第5章は，HSGPと言うGPを加速する新たな方法の主要なアイディアと詳細を述べる．これは，HSGPが木構造の設定に位置する多数のLNを利用し，進化過程を行う方法を述べる．

第6章は，多目的アプローチを考慮し，特徴抽出プログラムの自動合成システムのための有効的な新しいMOGPを示す． semi-elitist truncation, phenotypic-uniform samplingと prohibition of redundant individualsという冗長性を制限する三つのメカニズムを述べる．

第7章は，このシステムで合成された特徴抽出プログラムが，人間がデザインしたプログラムと比べて，性能が同等あるいはそれ以上であるかという質問に答える．比較は，雑草検出問題と虹彩領域分割問題の実際問題の二つで行う．

第8章は，本論文のまとめと，将来研究のアイディアを記述する．