

学位論文題名

動画像観測データの統合による物体構造推定に関する研究

学位論文内容の要旨

環境を視覚的に解析して得られた情報を利用することによって、よりヒューマンフレンドリーなシステムを実現しようという試みが行われている。中でも人物や手指などのように形状変化が可能な構造物の視覚的解析は、次世代のヒューマンインタフェースに不可欠な技術であると考えられる。人物や手指、あるいは機械を幾何学的にみれば、部品群が一定の幾何学的制約に従って運動する構造物であり、これらの解析は幾何学的モデルのもとでのモデルベース認識とみなすことができる。しかしながら、従来の研究で扱われている事例は、物体の構造モデルを事前に与えておいた場合の挙動解析に限られており、登録済みの物体しか認識ができない。また、動画像を通じて静止した対象物体の観測や単一の剛体を観測する場合などを扱い、部品間の幾何学的制約を考慮していない。

そこで本論文では、(1) 対象物体が関節構造で構成されている (2) 観測している特徴はオクルージョンの影響を受けないという2つの条件の下で、動画像の観測過程で自動的に対象物の構造モデルを生成するための手法について議論を行っている。本論文で提案する手法は、関節物体を構成する剛体と関節に関するプリミティブなモデルをあらかじめ用意し、画像列中に観測されるプリミティブな特徴の運動に対し、プリミティブなモデルを評価することで、物体のモデルを構築する。このように動画像中のプリミティブな特徴のトラッキングにより構造モデルが推定できる場合、従来のモデルベース認識のように、対象のモデルを与えることなく、対象物体の認識が行える。

本論文は、全4章から構成されており、以下では各章の概要について述べる。

第1章では、コンピュータビジョンシステムを構築するとき、従来行われてきた手法の問題点や限界について述べ、本論文で提案する手法に関連するセンサフュージョンについての説明を行う。また、本論文全体の概要と構成について記述している。

第2章では、画像中にある複数の特徴に対し、逐次的に局所的な評価を行い、ボトムアップに構造モデル構築する手法について提案している。本章では、複数の部分観測情報から(1)局所的な剛体や剛体間の関係の判定による構造モデルの構築法と、(2)観測結果に応じた動的なモデル構築についての提案を行った。

画像中の特徴量のグループ化は、階層的に、センサ情報レベル、剛体レベル、物体レベルの3層構造で物体の構造獲得を行う。センサ情報レベルでは、同じ特徴量が得られたものを統合する。剛体レベルでは、複数の部分特徴間に剛体の拘束関係が存在しているか判別し、グラフ理論を用いて剛体を構成するものをグループ化する。グラフ上では、部分特徴をノードで表し、部分特徴間に拘束関係があるときエッジで表す。各剛体は、グラフ上で完全グラフとして表せるので、特徴量について完全グラフとなるようにグループ化する。物体レベルでは、各剛

体間に関節が存在するかの判別を行う。関節の判別は、剛体間の相対運動の種類を状態で表し、それぞれの剛体間で観測された運動状態の遷移図をもとに判別する。各物体の判別にも同様に、グラフ理論を用いる。1個の物体は、グラフ上で可到達グラフで表されるので、同様に物体のグループ化を行う。各レベルでグラフを用いたグループ化は、局所的な誤判定を隣接関係を調べることで除去できる。それぞれのレベルでは、今まで行ってきた運動と異なるものが観測されたとき、無矛盾なモデルに再構築が行われる。部分情報の判別は並列に実行でき、全体の処理が複数の部分処理の集まりで表せるので、計算モデルとして黑板モデルを用いた。また、本章で提案した手法を用いて、2次元平面内を運動する多関節物体の実画像に対してエッジ情報の評価を行なった。その結果、各剛体及び回転、並進関節の獲得が逐次的に行われることが示され、本手法の有効性が示された。

第3章では、第2章で提案した手法の問題点である、推定の安定性、3次元構造の獲得の問題を解決するために、ロバストな推定を行う因子分解法を拡張した手法を提案した。本章で提案する手法は、3次元物体の特徴点の運動を正射影変換で投影した2次元の投影像に対し、因子分解法の枠組みを用いて、多関節構造の推定を行う。従来の因子分解法は、3次元空間を運動する単一の剛体について適用し、ロバストな推定を行うものであるが、因子分解法を多関節物体に適用するために、本章では、(1) 観測ノイズと特異値の関係を明らかにし、(2) 各剛体の分離を行い、(3) 剛体間に回転関節が存在するかの判定を行うための手法を新たに提案した。多関節物体に因子分解法を適用するために、3段階にわけて推定を行う。第一の段階では、各特徴点の観測情報を集めた観測行列を特異値分解し、観測ノイズが支配的な特異値と、運動情報が支配的な特異値の分離を行い、対象物体の運動の次元を求める。本論文では観測ノイズの分散をもとに、観測ノイズが支配的であるか、運動情報が支配的であるかを判別する閾値の決定法の提案を行った。第二の段階では、特異値分解した結果の中で、観測情報が支配的な項について、剛体の満たす拘束関係をもとに、それぞれの剛体に分離し、各剛体の形状を表す行列と運動を表す行列を求めた。第三の段階では、回転関節の判別を行うために、2剛体の運動を表す行列に対し、回転関節の拘束関係式に当てはめて求めた行列を特異値分解し、2剛体間の運動の次数から、拘束関係の有無を判定する。次数の判定は、第一段階と同様に行う。運動の次数は、2剛体に拘束関係がある場合、独立して運動している場合より次数が小さくなり、関節の情報は、ノイズが支配的と観測された特異値に対応するベクトルに現れる。本章で提案した手法をもとに、回転関節を持った物体のシミュレーション画像と実画像の2通りの評価実験を行った。その結果、ノイズの評価、各剛体の切りわけ、各剛体の形状の推定、回転関節の評価が正しく行われ、本手法の有効性が示された。

第4章では、本論文における研究の総括を行うとともに、残された課題について述べる。

学位論文審査の要旨

主 査 教 授 青 木 由 直
副 査 教 授 栃 内 香 次
副 査 教 授 北 島 秀 夫
副 査 助 教 授 川 嶋 稔 夫

学 位 論 文 題 名

動画像観測データの統合による物体構造推定に関する研究

環境を視覚的に解析して得られた情報を利用することによって、よりヒューマンフレンドリなシステムを実現しようという試みが行われている。なかでも人物や手指などのように形状変化が可能な構造物の視覚的解析は、次世代のヒューマンインタフェースに不可欠な技術であると考えられる。人物や手指は、個々の部品群が一定の幾何学的制約に従って運動する物体であり、これらの運動解析は重要な問題である。従来行われている人物や手指などの多関節物体の解析法の多くは、モデルベースト解析によるものであるが、対象物体の正確なモデルをあらかじめシステムに与えておく必要があった。一方、観測情報から直接、対象物体の形状モデルを求める手法も提案されているが、単一の剛体の運動解析に限られており、複雑な幾何学的制約の下で運動する構造物の推定が課題であった。

本論文で著者は、多関節物体を構成する各部品の形状推定と、幾何学的制約の推定を行う二つの手法を提案している。本論文で提案された手法では、対象物体を構成する剛体の性質に関するモデルと剛体間の幾何学的拘束関係に関するモデルが予め与えられ、観測している特徴はオクルージョンの影響を受けないことが仮定されている。このような仮定の下で、(1) 動画像から得られた幾何学的特徴群を制約にしたがってボトムアップ的に統合してモデル構築を行う方法、及び(2) 動画像中の特徴点群の追跡データを一括して因子分解することでモデル構築を行う方法について検討し、実験を通じてその有効性を実証している。

本論文は以下に示す4章から構成されている。

第1章では、序論で、物体構造の推定の推定に関連するコンピュータビジョン分野のこれまでの研究成果を概括している。また、本論文で提案されている手法の基礎となるセンサフュージョンの手法についても述べている。

第2章では、動画像から得られる様々な幾何学的特徴群に対し、局所的な関係評価を行い、評価結果を組み合わせることで、ボトムアップにモデル構築を行う手法を提案してい

る。構築されたモデルは観測情報をもとに逐次的に評価され、必要に応じてモデルの再構築が行われる。逐次的解析法を用いることにより、対象物体の解析を行う有限長の観測情報列を選択する必要がなく、それまでの観測に対して最も適切なモデルが構築される。本章で提案している手法を2次元平面内を運動する多関節物体の解析に適用した結果、各剛体及び関節構造のモデルが徐々に精緻化していくことが示され、手法の有効性が示されている。

第3章では、動画像から得られる特徴点の運動に対し、因子分解法を拡張した手法を用い、一括処理することで3次元構造を獲得する手法が提案されている。因子分解法を用いることで、観測情報の誤差に対してロバストな推定が実現できることが示されている。またアルゴリズム中で用いられるいくつかの閾値と観測雑音の関係についての数値的な解析が行われており、自動的な閾値決定法が提案されている。本章で提案した手法を、3次元運動物体のシミュレーションデータ、及び実画像データについて適用した結果、対象物の3次元構造がロバストに推定できることを実験を通じて明らかにしている。

第4章は結論で、提案した2手法の特徴について総括するとともに、各手法を実用的なシステムへ組み込む際の手法について議論している。

これを要するに、著者は、動画像による物体構造推定法について、いくつかの新知見を得たものであり、情報メディア工学に貢献するところ大なるものがある。

よって著者は、北海道大学博士(工学)の学位を授与される資格あるものと認める。