

学位論文題名

ジェスチャ推定のための画像処理方式に関する研究

学位論文内容の要旨

マルチメディアネットワークはインターネットの普及やFiber to The Homeの実現により、将来の社会基盤として成長する可能性を秘めている。このマルチメディアネットワークを根付かせるためには、人間とコンピュータとのコミュニケーション手段、いわゆるマン・マシン・インタフェースの技術が重要となるであろう。近年、マン・マシン・インタフェースにおける研究分野では、コミュニケーションの道具として、従来のキーボードやマウスといった機器の代わりに身振り手振り(ジェスチャ)を利用するという動きがある。これは、ジェスチャが我々の生活において最も手軽な意思の伝達手段の一つであり、音声を補うものとして重要な位置を占めているためである。このジェスチャをインタフェースとして利用するためには、次のような点を考慮する必要がある。(1)人が長時間利用しても負担とならないこと。(2)ジェスチャ推定の処理が高速にできること。(3)一般家庭内でも利用が可能なこと。(4)不特定の人が利用できること。

従来のジェスチャ推定手法は、データグローブのようなセンサを身体に装着する方法とビデオカメラなどで身体を観測する方法とに大別できる。前者の方法では(1)の点に関して大きな問題を持ち、後者は(2)から(4)の点において問題を抱えている。しかし日常生活で用いるインタフェースを実現するのならば、(1)の問題を解決することが最優先となるであろう。そこで本論文では、ビデオカメラを利用したジェスチャ推定法の確立を目的とし、上述した4点を解決し、推定に必要な種々のパラメータ情報を獲得するための画像処理方式の提案を行う。

本論文では、ジェスチャパラメータの獲得に対して2つの観点からアプローチしている。1つは、画像内の対象物体(手など)の形状を直接解析し、形状パラメータを求める方式で、もう1つは、形状の直接の解析は行わずに、彩色したグローブの画像内の色の情報から間接的に求めるという方式である。前者のアプローチに対しては動画像から手指の形状を正確に抽出し、それを追跡する手法の提案を行い、後者のアプローチを実現する手法として手の形状を効率よく推定できるようにグローブを作成し、画像内で検出された色の情報から安定に手の形状を推定するアルゴリズムを開発している。また、本研究で対象としている環境は我々の日常生活における環境とし、話者がその環境内を自由に移動できるようにカメラは固定されていないものとしていることから、両アプローチに共通した課題として、画像内における話者の位置決定を行うことが必要となり、この方法についても論じている。

本論文は、全体が以下に述べる全5章により構成されている。

第1章は序論であり、本論文における研究が行われるに至った背景と目的、及び本研究に関連のある手法について述べ、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、話者のカラー画像を1枚例示することによって画像の適合性エネルギー関数を適応的に定義したアクティブネットを用いて話者や手の部分を安定に抽出する方法の提案を行っている。提案手法では、話者の画像と入力された画像の3次元カラーヒストグラムから各色における画素数の比を求めるRatio Histogramという評価関数を利用し、その評価関数が高い値を返す色、すなわち話者の一部の色を指標とするエネルギー関数を定義し、エネ

ギー最小化によって画像内から対象物体の領域を抽出するアクティブネットを利用することで安定な話者の位置決定を行う。このようにすることで、我々の日常生活の環境のように様々なものがある場合でも話者だけを選別して特定することが可能になる。複雑な背景を持つ環境下で話者を撮影し、検出する実験を行った結果、また、入力画像内の背景領域に対象物体の一部と同じ色を含む環境下での実験結果から、例示画像に応じてエネルギー関数が定義されることにより安定して対象物体の抽出が可能であることを明らかにしている。

第3章では、アクティブネットのリンクを切断させることによって対象物体形状を表現しやすい構造に再構成し、手指のように複雑な形状の物体を正確に抽出する方法の提案を行っている。また、アクティブネットと入力画像の濃淡値との関係にぬれのアナロジーを適用し、その関係から得られる外部強制力を定義することによってアクティブネットに膨張する性質を与え、手指の追跡を行う方法についても述べている。提案した手法では、画像の適合性エネルギーを再定義することにより局所的最小の問題を回避するとともに、2本の指の間のような不連続領域の検出を行っている。次に、不連続領域におけるリンクを切断することによりアクティブネット構造の再構成を行っている。このようにすることで、アクティブネットの柔軟性が増し、手指のように凹凸の激しい形状をもつ物体を正確に抽出することができるようになる。更に、リンクの切断を繰り返すことによってアクティブネットは分裂し、左右の手を同時に抽出することも可能になる。手指動作の追跡では、対象物体の内部にある最外郭格子点に外側へ移動する力を加え、アクティブネットを対象物体のある方向へと移動させることを行っている。

複数の異なる手の形状の画像や手を振る動作を含む動画像列に適用した実験は、本手法を用いることにより手指形状の抽出や追跡が可能であることを示している。また、指を曲げる動作の動画像に適用した結果では、提案するアクティブネットが手の形状推定のために従来から利用されている手の3次元モデルと同等の振る舞いをすることも示している。

第4章は、色の異なる複数のパッチを付けたグローブを利用し、画像内で見えているパッチの色の組み合わせから間接的に手の形状と動作を推定するためのジェスチャパラメータを求める方法の提案を行っている。本手法では、人間の手の幾何学的な拘束（手のひらと手の甲は同時に見えないなど）を考慮することでグローブを作成し、画像から抽出されたグローブ表面のパッチのカラーヒストグラムの平均値をパラメータとして検出している。また、手の動作に関するパラメータには抽出されたパッチの画像上での重心の軌跡やカラーヒストグラムの平均値の変化を利用している。試作したグローブを装着させた手の画像からジェスチャパラメータを検出し、評価を行った結果は、安定してパラメータが求められることを示している。また、それらのパラメータを利用して手の形状を推定した結果により色の組合せから間接的に手の形状を推定できることが確認された。

第5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ている。

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

ジェスチャ推定のための画像処理方式に関する研究

本論文は、計算機視覚を用いて人物のジェスチャを解析するための手法について論じたもので、弾性体モデルの一種であるアクティブネット、および色彩分布のモデルであるカラーレシオヒストグラムを応用したいくつかの技法を提案しており、これにより日常的な情景下でのジェスチャ推定が可能であることを示している。

本論文は以下に示す5章から構成されている。

第1章は序論で、本論文の背景と目的、及び本研究に関連のある手法について述べ、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、話者の例示カラー画像により画像の適合性エネルギー関数を適応的に定義したアクティブネットを用いて、話者の全身、あるいは手の部分を安定に抽出する方法の提案を行っている。提案手法では、話者の画像と入力された画像のカラーヒストグラムからレシオヒストグラムを算出し、これをエネルギー関数とするアクティブネットのエネルギー最小化によって画像内の話者の位置決定を行っている。この手法を用いて、複雑な背景を持つ環境下での話者検出実験をおこない、日常的環境のもとでも話者だけを選別して特定することが可能であることを示している。

第3章では、アクティブネットのリンクを切断させることによって対象物体形状を表現しやすい構造に再構成し、手指のように複雑な形状を正確に抽出する手法の提案を行っている。また、アクティブネットと入力画像の濃淡値との関係にぬれのアナロジーを適用し、その関係から得られる外部強制力を定義することによってアクティブネットに膨張する性質を与え、手指の追跡を行う方法についても論じている。提案法では、画像の適合性エネルギーを再定義することにより局所的最小の問題を回避するとともに、指間のような不連続領域の検出を行っている。また、不連続領域におけるリンクを切断することによりアクティブネット構造の再構成を行っている。更に、手指動作を動画像中で追跡するために、対象物体の内部の最外郭格子点に外側への移動力を加え、アクティブネットが対象物体方向へと吸引される性質も付与している。これらの結果、アクティブネットの柔軟性が増し、手指のように凹凸の激しい形状をもつ物体を正確に追跡することが可能であることを実験により明らかにしている。

第4章は、色の異なる複数のパッチを付けたグローブを利用し、画像内で見えているパッチの色の組み合わせから間接的に手の形状と動作を推定するためのジェスチャパラメータを求める方法の提案を行っている。本手法では、人間の手の幾何学的な拘束（手のひらと手の甲は同時に見え

ないなど)を考慮することでグローブを作成し、画像から抽出されたグローブ表面のバッチのカラーヒストグラムを平均値をパラメータとして検出している。また、手の動作に関するパラメータには抽出されたバッチの画像上での重心の軌跡やカラーヒストグラムの平均値の変化を利用している。試作したグローブを装着させた手の画像からジェスチャパラメータを検出し、評価を行った結果は、安定してパラメータが求められることを示している。また、それらのパラメータを利用して手の形状を推定した結果により色の組合せから間接的に手の形状を推定できることを明らかにしている。

第5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論じている。

これを要するに、本論文は、人物のジェスチャを画像処理によって推定するための位置同定法、および形状解析法について論じたもので、色彩と弾性体に基づく解析モデルの提案を行っており、その結果得られた数々の新知見は情報工学、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、博士(工学)の学位を授与される資格があるものと認める。