

論文の内容の要旨

論文題目: 柔軟なインタフェースの実現に向けた人間行動の計測と理解

氏名: 岡 兼 司

近年, ユビキタスコンピューティング環境と呼ばれる, 人間の周辺にコンピュータが多数遍在する環境が普及しつつある. このような環境のもとでは, 遍在するコンピュータが人間の行動を自動的に計測・理解し, その結果に応じて人間にとって最適なサービスを提供するようなインタフェースが望まれる. このようなインタフェースは「柔軟なインタフェース」と呼ばれ, 近年注目を集めている. 柔軟なインタフェースの場合, 人間は周囲に潜む多数のコンピュータを明示的に操作する必要がないため, ユビキタスコンピューティング環境のもとに置かれた人間にとって, 非常に負担の少ないインタフェースであると言える.

以上に述べたような柔軟なインタフェースを実現するために最も重要な技術要素の一つとなるのが, 人間の行動を自動的に計測し理解するための技術である. そこで本研究では, 人間の各器官の動作を, 安定して自動的に計測・理解するための技術について研究を進めてきた.

ここで, 本研究の全体的な指針について述べる. 本研究の動作計測手法はマン・マシン・インタフェースに利用することを前提としており, ユーザにとってできるだけ煩わしさを感じない手法でなければならない. そのためには, 次に示す 3 個の条件を満たす必要がある. まず, ユーザに何らかのデバイスを装着する必要がないことが重要である. この問題に対して, 本研究では, コンピュータビジョンにもとづいた, 非装着・非接触型の動作計測を行うこととする. 次に, 動作計測に大幅な遅延が生じないように, 汎用コンピュータでも実時間処理が可能な程度に計算コストが低い手法を実現する. 最後に, インタフェースにも十分に適用可能な程度の精度や安定性で動作計測を行う. 本研究では, 以上の条件を満たすような手法の開発を行った.

以上の指針を踏まえた上で, まず, 人間の数多くの器官の中でも最も動作させる頻度が多い器官の一つとして, 手や指先の動作を安定して計測するための手法について研究を行った. 手指追跡手法については従来から数多く提案されてきたが, 実時間性を維持しながらも高い精度と安定性で複数の指先位置を追跡可能な手法はほとんど存在しなかった. そこで本研究では, 赤外線カメラから得られる個々の入力画像フレームの中で検出された複数の指先位置に対して, 画像フレーム間での指先同士の対応関係を考慮することにより, 複数の指先の動作を計測することを可能とした. ここでは, まず, 各画像フレームでの検出された指先位置に対して時系列フィルタを適用し, 直後のフレームでの各指先

の位置を予測する。次に、それらの予測位置と直後のフレームで実際に検出された指先位置との最適な組み合わせを決定することにより、隣り合う画像フレーム間での指先同士の対応関係を求める。本手法により、指の曲げ伸ばしによる指先の検出数が変化する場合や指先の検出に失敗した場合にも、複数の指先軌跡を実時間で安定して計測することが可能となった。指先同士の対応関係を考慮することによる指先追跡性能の向上に関しては、実験を通して証明することができた。

さらに、計測した指先動作をジェスチャとして認識するための手法についても提案した。特に、ダイレクトマニピュレーションとシンボリックジェスチャという二つの動作モードを統合的に利用するための機構についての提案を行った。従来のジェスチャ認識手法では、これら二つの動作モードを統合的に利用するという枠組み自体がほとんど存在せず、本手法のような機構の開発は斬新な試みである。本機構を実現するために、まず、計測された指先軌跡に基づいて、ダイレクトマニピュレーションとシンボリックジェスチャを自動判別する。その後、ダイレクトマニピュレーションとして判別された指先動作については指先位置情報を直接的にインタラクションに用いるとともに、シンボリックジェスチャとして判別された動作に対しては隠れマルコフモデルに基づく認識を行う。これにより、手の動作にもとづく様々な種類のインタラクションを可能とするための機構を実現することに成功した。

以上の手指動作の計測・理解についての研究により、手と指先動作を実時間で計測することが可能となった。その一方で、本研究における新たな問題も明らかになってきた。それは、計測された手指動作が意識的に行われたものであるかどうかを判断するのが困難であるという問題である。この問題が原因で、ユーザが意図しない偶発的な手動作についてもジェスチャとして認識してしまうという状況が頻繁に発生し、手動作の安定した認識・理解に向けての大きな障害となることがある。これは、手のジェスチャを認識するために手の動作から得られる情報だけしか利用しておらず、ユーザの意識的な動作を検出するために十分な情報が得られていないためである。

この問題に対して、本研究では、ユーザの意図と密接な関係を持つ情報として、ユーザの視線方向や顔の向きに関する情報についても同時に計測することとした。このうち、視線方向に関しては、特殊な装着型デバイスや目の周辺の高解像度画像が必要とされ、一般的なインタフェースへの応用は困難である。そこで本研究では、ユーザの顔の向きの計測を目的とし、頭部姿勢を 3 次元的に実時間で推定するための手法を提案した。

本研究の頭部姿勢推定手法では、入力画像への雑音や部分的な遮蔽への頑健性の観点から、時系列フィルタの一つであるパーティクルフィルタを利用する。その上で、本研究では、インタフェースへの適用を前提として、次の 2 点を同時に実現することが特に重要と考える。それは、ユーザ頭部の突発的な動作に対しても安定して追従することが可能であるという点と、ユーザが空間中のある一点に注目して頭部が静止している際には高精度に頭部姿勢を推定することが可能であるという点である。しかしながら、これらを同時に実現可能な手法は従来手法にはほとんど存在しなかった。そこで本研究では、パーティクルフィルタにおける仮説の拡散要素をユーザ頭部の速度に応じて適応的に制御するという手法を導入することにより、この 2 点を同時に実現することに成功した。このような適応的拡散制御による姿

姿勢推定性能の向上に関しては、実画像を用いた実験を通して証明した。

また、インタフェースに応用する上では、任意のユーザが利用可能なシステムであることも求められる。しかしながら、従来手法ではシステムを使用するユーザを限定し、そのユーザ専用のモデルを追跡開始前にあらかじめ構築しておくという方法を採用している手法が多く見られる。これに対し、本システムでは、ユーザが画像領域内に入ってきた瞬間にそのユーザ専用の 3 次元頭部モデルを自動的に構築ための機構を実現した。ここでは、2 台のカメラにより撮影された入力画像から、顔器官検出技術やステレオマッチング技術などを用いて、ユーザ頭部の複数の特徴点の 3 次元位置を計算する。ここで計算された特徴点位置を用いて 3 次元頭部剛体モデルを構築し、そのモデルを使用して頭部姿勢推定を開始する。これにより、任意のユーザに対して特別な事前準備なしに頭部姿勢を計測することが可能となった。

さらに、頭部姿勢の実時間推定と並行してユーザの顔形状変化を分析することにより、頭部変形モデルの自動構築も実現した。このように頭部変形モデルの構築と頭部姿勢推定の並行実行を試みた手法は過去に提案されておらず、本手法は非常に斬新なアプローチであると言える。本手法では、頭部剛体モデルを用いた姿勢推定の結果を手がかりとして頭部の各特徴点の正確な動きを実時間で計測する。こうして計測した特徴点の動きを分析することにより、姿勢推定を継続しながらも、新たな頭部変形モデルを構築するという手法を導入した。このように構築した頭部変形モデルを用いて頭部姿勢を推定することにより、表情変化などを原因とする顔の変形が発生した場合にも、安定して頭部姿勢を推定することが可能となった。この推定性能の向上に関しても、実画像を用いた実験により証明できた。

以上の研究により、インタフェースへの応用を前提とした上で、手指動作と頭部姿勢を計測することが可能となった。そこで本論文では、これらの動作計測手法を実際のアプリケーションとして応用した例について幾つか紹介する。そして、それら応用例の有効性を考察するとともに、本研究の計測手法の有効性についても考察を行った。