

センサやコントローラなしにジェスチャーで情報通信機器やロボットが操作可能な ICT 技術の研究開発 (102303006)

Information communication technology for gesture-guided control of ICT equipments and robots without sensors and controllers

研究代表者

星野 聖 筑波大学

Kiyoshi Hoshino University of Tsukuba

研究分担者

向井利光 株式会社クレセント

Toshimitsu Mukai Crescent,inc.

研究期間 平成 22 年度～平成 23 年度

概要

本課題は、入力装置として、ヒト手指の非接触モーションキャプチャシステムの開発と、出力装置として、臨場感や没入感を与える広視野角ヘッドマウントディスプレイの開発の、2つから構成される。前者の3次元手指形状推定については、画像特徴量の低次元化により、指の太さや長さ、動かし方の癖などの個人差に対応した膨大な照合用手指データベースを作成し、第一段階での粗い絞り込みと第二段階の精緻な類似度照合により、関節角度の推定誤差2～3度程度で100fps以上の高速の手指形状推定を実現した。とくに拇指の形状推定を正確に行えるようにデータベースを工夫し、物体操作動作における拇指対向動作を正確に推定できるようにした。後者の立体視可能な広視野角ヘッドマウントディスプレイについては、映像デバイスの改良と機構の抜本的改良により、高解像度、広視野角、1.7kgまでの軽量化を実現した。

Abstract

For the development of motion-capture system for human hands and fingers with one or two cameras, we construct the enormous hand and finger database for users' individual differences such as thickness and length, habit of the movement way of the finger, by decreasing of the dimension of the image characteristics quantity, and we adopted rough searching as the first step and detailed fitting as the second step. Our hand pose estimation system works with the processing speed over 100fps, with two or three degrees of estimated error in joint angle. Especially, the joint angle of the thumb can be accurately estimated in our system. For the development of wide range head-mounted display for output, we designed wide range, high resolution, 1.7kg of weight HMD, by improvement of the image device and drastic improvement of the mechanism.

1. まえがき

本研究開発課題は、センサ類の装着やコントローラの使用なしに、日常動作と同じように振る舞うことで、情報通信機器、家電製品、ロボットなど身の回りのものを操作できる技術の創出を目指した。本提案は2つの技術要素から構成される。第一は、2眼カメラにより、ユーザの手指動作を高速かつ高精度に推定・認識するマシンビジョンの開発。第二は、立体視可能な広視野角ヘッドマウントディスプレイ(HMD)の開発である。これら2つの技術を組み合わせ、第一に、手指動作により3次元自由造形情報が入力可能なバーチャル粘土細工システム。第二に、観視映像に対して動作で働き掛けることで正確な物体操作が実現されるジェスチャー駆動ロボットの開発を目指した。

2. 研究内容及び成果

第一の小課題「不特定ユーザ対応の3次元手指形状推定システム」実現を目的として、第一に、手指画像中の尾根線を利用する方法を開発した。すなわち、手の形(骨の長さ、太さ、掌と指の比率)や、手の動かし方(関節可動域、その中での姿勢の取り方)などには大きな個人差が存在する。年齢や性別による個人差もあるし、人種間の差も存在する。そういう問題に対して、たとえば、手指3次元モデルをあらかじめコンピュータ内に構築しておき、入力された手指画像と照合することにより手指形状推定を行うシステムでは、事前のキャリブレーションにより個人差対応

も可能かもしれない(現実には、対応できているシステムはまだ存在しない)。一方、本課題のように、事前に照合用の手指データベースを構築しておき、入力画像との照合を行うシステムでは、個人差を持った膨大な規模のデータベースを事前に構築しておくことが不可欠であり、かつ、その中から、入力画像と最も類似する手指画像を高速に見つけなければならない。しかし、さまざまな個人差を含む膨大なデータベースの中から、操作対象のロボットやコンピュータが実時間で操作できるほどの高速処理を実現することは簡単ではない。手指に存在するこのような個人差を解決しようとするのが本項目であり、具体的には、従来技術での個人差が生じる原因であったと思われる手の輪郭線情報による推定でなく、前景画像である手画像の各ピクセルそれぞれの背景画像からの距離を高さとし手画像を山として見ることによってそこに引くことのできる尾根線情報を推定に用いた。これにより、手の形や動かし方による影響を大きく減らした情報で形状推定を行うことを可能になる。また、尾根線情報を特徴量化する際にベクトル化によって特徴量化することで照合用データベースの各データセットが持つ特徴量次元を大幅に低減することにより、逆に次元圧縮した分のデータセットを増やしても、従来と同じ高速性を維持できる。さらに、尾根線情報は個人差に影響されにくいので、新たに個人差対応した照合用の大規模データベースを加える必要も低減できる。

第二に、手指画像中の爪の個数とそれぞれの重心位置と

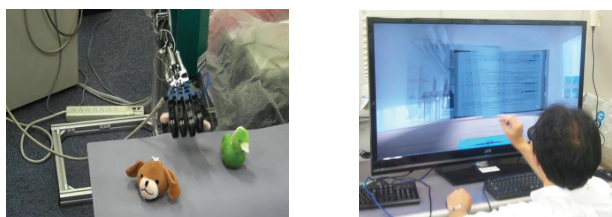
を高速かつ正確に算出するシステムを開発した。すなわち、爪の情報を活用できると、画像中に写る手の左右側や、顔や腕といった手以外の肌色領域と手指領域とを、簡単に区別できるようになる。しかし爪は、掌側の皮膚のうちの拇指丘や指腹と非常に似た色相を持っているため、手の甲側の肌と爪との分離は比較的簡単であるが、掌側の肌と爪との分離は極めて難しかった。とくに手指形状推定では、3次元空間におけるユーザの自由な動きを許容しなければならないため、似た色相を持つ拇指丘や指腹との分離が鍵であった。本課題では、爪と、爪と類似の色相を持つ爪以外の領域とでは、爪と類似する色相を持つ画素の分布の仕方が異なることに着目し、注目する領域の広さを変え、局所的な画素の分布状態を見ることにより、簡単なキャリブレーションで、皮膚の色に個人差があるユーザに対しても、写っている爪の個数とそれぞれの重心位置とを正確に推定する手法を開発した。

第二の小課題「立体視可能な広視野角 HMD」のために、802.11a 準拠の伝送チップ利用のメカ設計に着手し、そのプロト基板製造を行った。ここでのチャレンジは、フルハイビジョン（解像度 1920x1080）映像を左右眼 2 系統のデータを遅延なく伝送を行うこと、かつ、伝送に伴う圧縮による画像劣化を極力抑えること、そして、基板自体をできる限り小型化を行うこと、の 3 点にあった。次の優位性、すなわち伝送データを非圧縮で行うことで、高画質を担保できる。圧縮プロセスがないゆえ、通常発生する処理遅延が 0 である。低消費電力チップにより、ヒートシンク等の部品を必要とせず軽量化に貢献できる。将来の 5 GHz 帯域にも対応しており、より広帯域で安定した伝送を担保できる等を勘案し、Amimon 社チップセットをベースにした。

また、色収差補正、歪み補正をより高速に高精度に処理を行うための 1920×1080 画素用高解像度ソフトの開発を行った。同開発は、この HMD が世界で一番広い視野角、水平 140 度、垂直 90 度を有しているがために必要な処理工程である。とくに、なるべく光学的に色収差、歪みの無いものを作りためには、通常収差の少ない硝材を使用して凹凸レンズを製造、実装する。ところが、このような硝材は、重量が非常にあり、HMD に搭載した場合は、小型、軽量化の大きな妨げとなってしまう。したがって、一部のレンズをプラスチックで製造し、そのために発生してしまう色収差および歪みを、高解像度かつ遅延なしにデータ処理できるアルゴリズムを考案した。

さらに、色収差補正、歪み補正を行いながらも、画像劣化を最低限に抑え、画質向上、コントラスト向上を図るための光学ハードウェア的視点からの設計も同時に行った。HMD の映像表示原理は、プリズムで結像させた映像をミラーで屈折させながら眼前に広がるスクリーンに投射し、それをスクリーンの裏面から接眼レンズで広視野角に広げて見る。そのため、どうしても画像が暗くなり、コントラストが落ち、クッキリ感が出づらくなる。これらを光学的見地から再検証を行い、最適解を導いた。

第三に、「バーチャル粘土細工システム」の実現と、情報端末操作への拡張として「バーチャル積み木遊び」の試



ジェスチャーによる遠隔ロボット（左）と PC 操作（右）

作を行った。そして第四に、遠隔ロボットが把持物体を落下させず、ユーザが意図した位置に物体を移動できる性能を持った「ジェスチャー駆動ロボット」を実現した。

3. むすび

現在の情報通信端末のように、表示が 2 次元なら、タッチパネル操作を基本とする情報入力には便利である。しかし、表示が 3 次元になった時、入力はタッチパネルやマウスではなく「ジェスチャー」である。2010 年は立体元年と呼ばれ、3D コンテンツが家庭に入ってきた。また立体表示対応のノートパソコンなども販売されるようになってきた。近い将来、各家庭には大型の立体テレビが普及するようになることが予想される。ただ単に 3D 映画を観るだけでなく、情報機器操作もジェスチャーで、とくに手指の複雑で細かい動作で操作できるようになれば素晴らしい。

通常の手指のジェスチャー認識による情報通信機器操作は動きベクトル解析技術をもとにしているため、いわば高機能のポインティングデバイスしか実現できない。ポインティングデバイスでは、ユーザは事前に、たとえばファンクションキー 1 番の機能はジャンケンのチョキ動作、ファンクションキー 2 番の機能はジャンケンのパー動作といった動作と機能との対応づけに習熟しなければならない。しかも 2 次元映像の動きベクトル解析が土台なので、たとえば手首運動のような z 軸まわりの動きを検出できないのが現状である。それに対して、本提案はカメラ画像から「手指の全関節角度」を推定する技術を基礎にするので、真の意味での「身ぶりで操作できるパソコン」「身ぶりによって形成される 3 次元造形物のデジタル入力」「身ぶりと同じように動く遠隔ロボット」などを実現することが可能になる。情報端末を操作するために、ユーザは特別な操作方法を習得する必要がなく、日常動作と同じ動作をすれば良いため、高齢者や障害者にとってのデジタルディバイド（情報格差）低減にも繋がる。

【誌上発表リスト】

- [1] K.Hoshino, T.Kasahara, M.Tomida, and T.Tanimoto: "Gesture-World Environment Technology for Mobile Manipulation - Remote Control System of a Robot with Hand Pose Estimation -," Journal of Robotics and Mechatronics, 24, 1, pp.180-190, 2012
- [2] K.Hoshino and W.D.Gayan.K: "Control of air cylinder actuator with common bias pressure," Journal of Robotics and Mechatronics, 23, 6, pp.991-998, 2011
- [3] K.Hoshino, T.Kasahara, N.Igo, M.Tomida, T.Tanimoto, T.Mukai, G.Brossard, H.Kotani: "Mobile robot control using 3D hand pose estimation," 10th Intl. Conf. on Quality Control by Artificial Vision (QCAV), Vol.8000, No.28, pp.1-10, 2011

【申請特許リスト】

- [1] 星野 聖, 以後直樹: アクチュエータ、アクチュエータの制御方法及びプログラム, 特願 2011-273230, 出願日 2011 年 12 月 14 日
- [2] 星野 聖, 藤嶋教彰: 爪領域抽出装置、爪領域抽出方法、及び爪領域抽出プログラム, 特願 2011-223493, 出願日 2011 年 12 月 14 日
- [3] 星野 聖, 豊原遥平: 手指形状推定装置、手指形状推定方法、及び手指形状推定プログラム, 特願 2011-223492, 出願日 2011 年 10 月 7 日

【本研究開発課題を掲載したホームページ】

<http://www.youtube.com/watch?v=UjbZYN1Db14>