

平成22年 5月31日現在

研究種目：若手研究(A)

研究期間：2007～2009

課題番号：19680010

研究課題名(和文) 高速ビジョンを用いた快適な入力インターフェースの研究

研究課題名(英文) Research on comfortable input interfaces using high speed vision

研究代表者

小室 孝 (KOMURO TAKASHI)

東京大学・大学院情報理工学系研究科・講師

研究者番号：10345118

研究成果の概要(和文)：本研究では、ビデオフレームレートを越える高フレームレートのカメラを用いて、各種情報端末における快適な入力インターフェースを実現することを目的とし、(1) 携帯端末向け空中マウス/キーボード(2) ズーミングタッチパネル(3) 読唇入力インターフェース(4) ハンディ三次元スキャナの各システムを構築した。

研究成果の概要(英文)：In this research, aiming to realize comfortable input interfaces for various information devices using high frame rate cameras which go over the video frame rate, and we built the following systems: 1) In-air mouse/keyboard for mobile devices, 2) Zooming touch panel, 3) Lip-reading interface, 4) Handy 3D scanner.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	5,500,000	1,650,000	7,150,000
2008年度	6,700,000	2,010,000	8,710,000
2009年度	6,800,000	2,040,000	8,840,000
年度			
年度			
総計	19,000,000	5,700,000	24,700,000

研究分野：総合領域

科研費の分科・細目：情報学—知覚情報処理・知能ロボティクス

キーワード：コンピュータビジョン

## 1. 研究開始当初の背景

従来、コンピュータの入力インターフェースには、マウスやキーボード、タッチパネルなどが使われてきた。ビジョンセンサを入力インターフェースに用いることで、非接触操作や広い操作空間などが実現できる。カメラを用いた入力インターフェースの研究例は多数あるが、快適性の向上を目的としたものは少ない。高フレームレートの視覚情報を用

いることで、従来のインターフェースでは得られなかった快適でストレスフリーな操作感を提供できると考えられる。

## 2. 研究の目的

本研究では、通常のビデオフレームレートより高い、毎秒100～1000フレームの動画をリアルタイムに取得することができるビジョンセンサを用いて、各種情報端末におい

て快適な入力インターフェースを実現することを目的とする。具体的には、以下の機能を実現する入力インターフェースシステムを作成する。(1) 携帯端末向け空中マウス／キーボード (2) ズーミングタッチパネル (3) 携帯端末向け読唇入力インターフェース (4) ハンディ三次元スキャナ

### 3. 研究の方法

#### (1) 携帯端末向け空中マウス／キーボード

- ・ ビジョンチップとポータブルPCを用いて、入力インターフェースシステムを構築する。
- ・ GUI環境で使用できるアプリケーションを構築する。
- ・ 手の姿勢・指の曲げを画像で認識するアルゴリズムをFPGAに実装する。
- ・ 上記FPGAが搭載されたボードをノートPCとPIOで接続し、通信が行えるようにする。
- ・ FPGAとノートPC、高フレームレートカメラにより、手の姿勢・指の曲げのリアルタイム認識実験を行う。
- ・ 画面上に表示したキーボードに対し、指の曲げによりタイピングを行うプロトタイプソフトウェアを作成する。
- ・ 作成したソフトウェアを、数人の被験者に使用してもらい、使いやすさの評価を行う。

#### (2) ズーミングタッチパネル

- ・ ディスプレイに取り付けた2台のカメラで指先の三次元位置を取得する実験を行う。
- ・ 指先の三次元位置情報を用いて画面を拡大縮小するプログラムを作成する。
- ・ カメラの最適配置や自動キャリブレーションについて検討する。
- ・ カメラを固定するフレームの作成を行う。
- ・ デモアプリケーションとして、地図や電車の路線図、美術品などを指先の三次元位置に基づいて拡大表示するプログラムを作成する。

#### (3) 読唇入力インターフェース

- ・ 唇の動きを高速カメラを用いて撮影し、唇の位置形状だけでなくその時間的変化の分析から、発話（特に子音）を識別する実験を行う。
- ・ PCと高フレームレートカメラを用いて読唇入力インターフェースの実験システムを構築する。

#### (4) ハンディ三次元スキャナ

- ・ カメラを手で動かし、表面形状を取得するアルゴリズムを考案し、オフラインで実験を行う。
- ・ 決定したアルゴリズムをPCに実装し、リアルタイムでの動作を実現する。
- ・ 復元した三次元形状をリアルタイムに画面に表示するプログラムを作成する

### 4. 研究成果

#### (1) 携帯端末向け空中マウス／キーボード

小型化が進む携帯機器において、機器表面上に広い操作領域を確保することは困難である。例えば、携帯電話に使われるキーボードやタッチパネルなどは操作領域が非常に狭く、操作性を損なっている。そこで本研究では、単眼カメラを用いて空中の指先の動きを検出し入力を行うインターフェースを提案する。

このようなインターフェースを実現するためには、高速に動く指の位置姿勢の安定した認識やキーストローク動作の正確な検出といった課題がある。これらの課題に対し、本研究では二値化した指先画像をテンプレートとして用い、並進、回転、スケールに関してトラッキングすることにより、指先の三次元位置を安定して取得している。携帯機器では、カメラと指先の距離が近い場合、画像上で指先は速く移動する。そこで、高フレームレートカメラを用いることで、速い指先の動きにも追従することができるようにした。また、指先画像のスケールの変化に対して周波数フィルタを適用することで、単眼カメラのみで指先の微小な奥行き位置の変化を検出している。

提案インターフェースの有効性を検証するために、プロトタイプシステムを構築した。本システムは小型カメラに広角レンズを取り付け、PCで処理を行い、小型ディスプレイに表示した。入力画像の取得レートは154fps(frames per second)、画像サイズは400×160pixelsとした。図1にシステムの外観を示す。

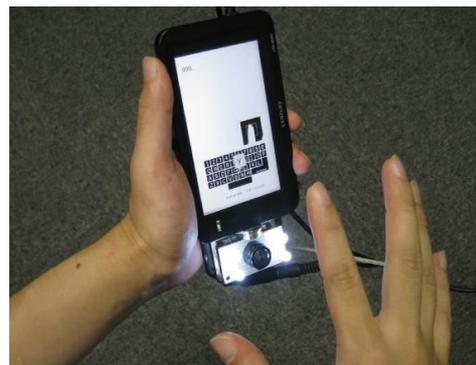


図1 携帯入力インターフェース

本システムを用いていくつかのアプリケーションを実現した。一つ目は、指先の位置に合わせて動くカーソルを画面に表示し、画面上のソフトウェアキーボードを空中でタイプして文字入力を行うというものである。空中で指を軽く前後に動かすと、キーストロークアクションとして検出される。二つ目は、

指の三次元位置が取得できることを利用し、写真をズームしたりスクロールしたりできる写真ビューアである。キーストローク動作が検出されると、写真を動かすことができるようになる。再度、キーストローク動作が検出されると、写真が画面に固定される。三つ目は三次元空間に直接線を描画できる三次元ペインタである。描画した絵は、指の動きにより、視点を変えて表示させることができる。

このように、単眼で三次元情報が取得できることにより、携帯機器の操作領域の問題を解決するだけでなく、これまでとは異なった新たな携帯アプリケーションを実現することができるようになった。

このほか、複雑な背景や変化する光源環境において、安定した手指姿勢を認識するアルゴリズムの開発を行った。リアルタイム処理のために、変形テンプレートマッチング処理を高速に演算するハードウェアを導入し、空中でのタイピング動作に対する実時間認識を 138fps で行うシステムを試作した。図 2 にリアルタイム認識実験の様子を示す。

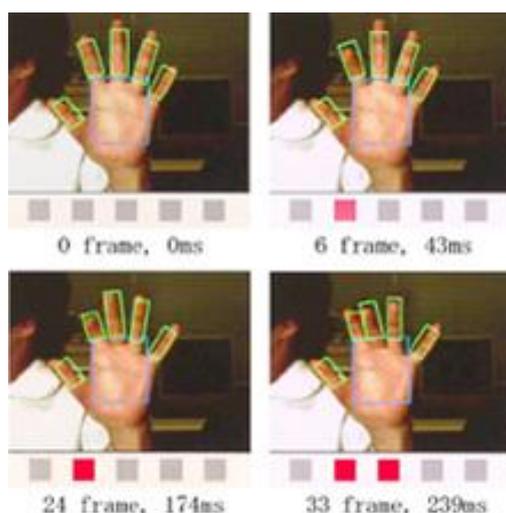


図 2 手指姿勢のリアルタイム認識

## (2) ズーミングタッチパネル

銀行の ATM や駅の券売機、飲食店のオーダー端末など、タッチパネルを用いた端末は身近に存在するが、タッチパネルには細かい操作が難しいことや、衛生面での問題などが存在する。そこで、小型カメラを 2 つ用いて、パネルに近づけた指の 3 次元位置を計測し、パネル上にカーソルを表示したり、指の奥行き位置により画面をズームさせたりするなどの新しいインターフェースを備えたタッチパネルを製作した。図 3 にシステムの外観を示す。駅の券売機への応用を想定し、地図上で直接目的地を選択して切符を買うことのできるアプリケーションを実装した。この

ように、本システムは現行のタッチパネルに比べ細かい入力を可能にし、広い表示範囲と精細な表示を両立させることができる。



図 3 ズーミングタッチパネル

## (3) 読唇入力インターフェース

視覚情報を用いた音声認識の研究は数多く行われているが、日本語の子音発話においては口唇の形状にほとんど差がなく、識別が難しかった。そこで、口唇の形状が時間変化する際のタイミング構造が、識別のための重要な情報を含んでいると予想し、独自の特徴量を提案した。口唇の外側輪郭の上下の画像上の幅を高さ  $h$  として取得し、その時間変化を多項式フィッティングによりパラメータ化した (図 4)。

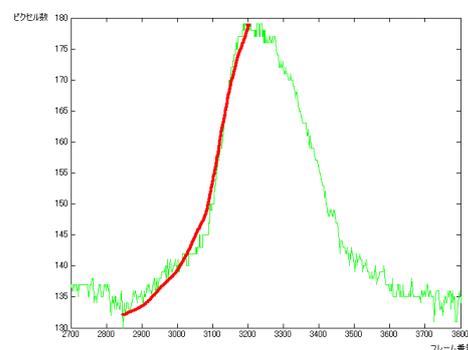


図 4 口唇時間変化の多項式フィッティング

被験者に 10 子音(子音なし, k, s, t, m, h, m, y, r, w) を 5 回ずつ発話をしてもらい、上記のパラメータを Fuzzy c-means によりクラスタリングしたところ、クラスタリングの純度は 0.54 となった。この結果から、口唇の時間変化が子音の識別に有効であることが示唆された。

## (4) ハンディ三次元スキャナ

PC に接続したカメラを手で動かして、手軽に三次元形状を取得できる装置の実現を目指し、実験システムを試作した (図 5)。最新フレームを基準とし、過去  $N$  フレームとのステレオマッチングにより、三次元位置を計算

する。カメラの位置姿勢はマーカを用いて取得し、マーカ座標系を基準とした復元を行う。処理の高速化のため、基準フレームにおける画像中心付近のみを復元するようにした。



図5 試作したシステム

実験では、直近の 100 フレームを用いて、画像中心 1 ピクセルの三次元位置を復元した。結果をリアルタイムに画面に描画することで、ペイントアプリのような自分の手で描く感覚が得られるとともに、ユーザーがエラー箇所を撮り直しによって更新できることが確認された。本研究により、インタラクティブシステムとしての有効性は確認できたが、高フレームレートによる実験までは残念ながら至らなかった。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 0 件)

[学会発表] (計 11 件)

Takashi Komuro, Bjorn Werkmann, Takashi Komai, Masatoshi Ishikawa, Shingo Kagami: A High-Speed and Compact Vision System Suitable for Wearable Man-machine Interfaces, IAPR 10th Conference on Machine Vision Applications, 2007.5.17

寺嶋一浩, 岩下貴司, 小室孝, 石川正俊: 高速ビジョンを用いた空中タイピング動作の認識, 第 13 回画像センシングシンポジウム, 2007.6.7

寺嶋一浩, 小室孝, 石川正俊: 高フレームレートカメラと FPGA による空中タイピングシステムの構築, 動的画像処理実利用化ワークショップ 2008, 2008.3.7

山本啓太郎, 山口光太, 渡辺義浩, 小室孝, 石川正俊: 高速ビジョンを用いた特徴点追跡による 3 次元形状復元, 電子情報通信学会

2008 年総合大会, 2008.3.21

山口光太, 渡辺義浩, 小室孝, 石川正俊: 遠隔カメラ映像のための覗き込みインタフェース, 第 13 回映像メディア処理シンポジウム (IMPS2008), 2008.10.31

山口光太, 小室孝, 石川正俊: 顔追跡による PTZ 操作と魚眼パノラマへの応用, 情報処理学会インタラクション 2009, 2009.3.5

Kazuhiro Terajima, Takashi Komuro, Masatoshi Ishikawa: Fast Finger Tracking System for In-air Typing Interface, The 27th Annual CHI Conference on Human Factors in Computer Systems (CHI2009), 2009.4.7

Kota Yamaguchi, Takashi Komuro, Masatoshi Ishikawa: PTZ Control with Head Tracking for Video Chat, The 27th Annual CHI Conference on Human Factors in Computer Systems (CHI2009), 2009.4.7

廣部祐樹, 船橋一訓, 新倉雄大, 渡辺義浩, 小室孝, 石川正俊: 単眼カメラを用いた携帯機器向け空中タイピングインターフェース, 映像情報メディア学会 情報センシング研究会・メディア工学研究会, 2009.6.16

山本啓太郎, 小室孝, 石川正俊: 高速ビジョンを用いたビデオモザイクング, 映像情報メディア学会 情報センシング研究会・メディア工学研究会, 2009.6.16

Yuki Hirobe, Takehiro Niikura, Yoshihiro Watanabe, Takashi Komuro, Masatoshi Ishikawa: Vision-based Input Interface for Mobile Devices with High-speed Fingertip Tracking, 22nd ACM Symposium on User Interface Software and Technology (UIST2009), 2009.10.5

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 1 件)

名称: 画像表示装置  
発明者: 小室孝、福岡功慶、石川正俊、廣川新、高橋真  
権利者: 国立大学法人東京大学、株式会社 PFU  
種類: 特願  
番号: 2007-49062  
出願年月日: 平成 19.2.28  
国内外の別: 国内

○取得状況 (計 0 件)

[その他]

ホームページ等

<http://www.k2.t.u-tokyo.ac.jp/vision/index-j.html>

## 6. 研究組織

### (1) 研究代表者

小室 孝 (KOMURO TAKASHI)

東京大学・大学院情報理工学系研究科・講師

研究者番号：10345118