

手軽に利用できるハンドジェスチャーユーザインタフェースの 考察と提案

Considerations and Proposals of easily usable Hand Gesture User Interface

坪田 尚洋[†] 中島 克人[†]
Naohiro Tsubota Katsuto Nakajima

1. はじめに

快適な操作環境を提供するための非接触型ユーザインタフェースの開発が続けられている。中でも手振りを利用した入力操作を可能にするハンドジェスチャーユーザインタフェース(HGUI)は、人間が日常から手振りを用いて意思疎通を行っているという背景から人に馴染み深いと言われている。90年代からユーザへの負担や設備費などを考慮して、コンピュータビジョンによるHGUIの開発が試みられてきた。我々もその方針に賛同している。

コンピュータビジョンによるHGUIの実現結果は既に幾つか報告されている[1,2,3]。しかし、それらは特定の装置やプロセッサに依存していたり[1,3]、多数のカメラを用いたりする[2]という点で使用環境が限定されてしまっている。そこで我々はそれらを克服した手軽に利用できるHGUIを開発したいと考えている。

本稿は2章で手軽に利用できるHGUIについて考察した後、3章で我々が開発を試みているHGUIの仕様を述べ、その実現への対応を示す。4章は開発に先立って行った検証実験を示し、5章で結びとする。

2. 手軽に利用できるHGUIの考察

この章では手軽に利用できるHGUIとは何かを考える。その結果、我々が開発するHGUIに必要な仕様を示す。

2.1 使用環境

我々は従来のHGUIの最大の弱点は使用環境にあると考えている。自己オクルージョンへの対応として複数のカメラを用いる手法[2]は認識精度の向上が可能だが、例えば一般住居内にカメラを都合のよい位置に多数配置するのは現実的ではない。また、複雑な処理を実時間で実行するために特別なハードウェアを導入する[3]ことは設備費の増加やプラットフォームの限定を招く。つまり、我々が目指す手軽に利用できるHGUIは1台のカメラと一般的なハードウェアのみを用いた実装を目指したい。

2.2 使用感

使用環境について好ましいHGUIが完成しても実際に有用でなければならない。機能性や利便性はアプリケーションやGUI設計によるが、安定性と応答性、すなわち認識精度と処理速度は技術的な課題である。HGUI技術の構成要素である手領域抽出と手形状認識、そして、ハンドジェスチャー認識の手法は多く提案されている。その中から高速かつ高精度の手法を厳選する必要がある。ただし、ここで重要視すべきは認識精度よりも高速性である。なぜならば、HGUIはあくまでOSやアプリケーションへのUIであるので、共存する当該プロセスの処理時間も必要だからである。

2.3 柔軟性

既存のHGUIの弱点の1つとして認識するジェスチャーの数が少ないという点が挙げられる。これは識別能力の問題でもあるが、それ以前にHGUIの仕様の問題でもある。HGUIに予め厳選されたジェスチャーが設定されていても、ユーザがそれで満足するとは限らないし、慣れるのが困難で使い切れないかも知れない。この観点からジェスチャーの登録とジェスチャーと命令(コマンド)の対応付けはユーザが任意に行える事が望ましい。なおかつ、登録できるジェスチャーの数に十分な自由度を与えるために、手形状認識とハンドジェスチャー認識は注意深く設計されるべきである。

3. 提案するHGUIの構築

この章では、我々のHGUIの仕様とそれを実現するための技術選択や工夫を述べる。

3.1 仕様

ここで我々が開発するHGUIの仕様を明確にする。ひとまず、我々のHGUIは多くの家庭で使用されている標準的な性能のPC上での利用を想定する。プロセス性能が劣るためにより難易度が高くなる家電製品への組み込みも将来は目指したいが、複雑背景への対応やストレスを感じさせない実行速度、カメラの設置自由度などの面からは通常のPCであってもかなり厳しい想定であると言えよう。このUIが実現できれば多くのPCユーザが手軽にHGUIを利用出来るようになる。

提案するHGUIのシステムフローを図1に示す。このシステムにおける応答はウィンドウの最小化等のOSに働きかけるものを想定しているが他のアプリケーションに組み込んでも良い。学習モードでユーザは任意のハンドジェスチャーをシステムに学習させ、それに応答を対応させる。ここで応答が対応付けられたジェスチャーをコマンドと呼ぶ。認識モードで固定カメラに向かってコマンドを見せると対応する応答が実行される。カメラは近年普及してきているUSBカメラを想定している。

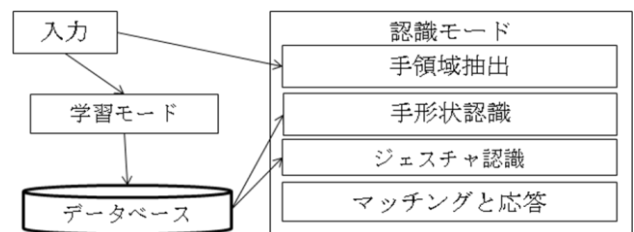


図1 システムフロー

[†] 東京電機大学大学院未来科学研究科
Graduate School of Science and Technology for Future Life,
Tokyo Denki University

3.2 実現技術

3.2.1 手領域抽出

手領域抽出はシステムの仕様上、複雑背景から行わなければならない。前景領域を背景から分離する技術は古くから研究されているが、精度、実行速度、そして、対環境変化の全てにおいて完璧な手法は確立されていない。しかし、我々の場合は対象が手に限定され、分離精度も後段の処理に耐えうる程度で良いため、既存技術の組合せで実現可能と考えている。

3.2.2 手形状認識

HGUI において、どれだけ多くの形状のクラスを識別可能であるかがコマンド設定の自由度に影響してくる。例えば、1つのコマンドは2つの手形状間の遷移であるとし、ユーザは20個程度のコマンド登録を希望すると仮定すると、コマンド設定の自由度を考慮し、希望個数の10倍の200コマンドは実現可能とすべきであろう。これは14種類(200の平方根)程度の手形状識別の必要性を意味する。

手形状認識は既に多くの手法が提案されている。それらは利用する特徴量の種類で大きく分けて、画像空間、スケルトン、輪郭の3つに分類できる。画像空間から直接形状特徴を得る方法は、情報落ちがないだけに精密な特徴抽出が可能ではある。しかし、この方法は対象が手に限定されている利点を活かすにも、処理が重くなる傾向がある。スケルトンはいくらかの情報落ちを伴っているものの、それ自身が手の骨格に相当するため、精密に手形状を表現できる。しかし、指先のスケルトンが性格に取れていないと誤認識の可能性が高くなる。また、スケルトンは指関節の回転角などのパラメータが詳細に分かる利点があるが、その数値的な細かさはハンドジェスチャ認識に必ずしも必要ではない。なぜならば、細かい形状変化は意図せずに発生する可能性が高く、そういったコマンドを設定すること自体が不自然だからである。輪郭は前者と比べて多くの情報を落としているが、それでも手形状の識別は可能である。輪郭ベースの認識の長所は、計算量が小さい所にある。計算量の小ささはHGUIに大変重要な応答速度に貢献する。これらの理由から、我々は古くから輪郭ベースの形状認識に利用され、ノイズに耐性があり、拡大や回転に対する頑健性を持ち、特徴をコンパクトに記述できるフーリエ記述子[4]を採用する事にした。

3.2.3 手の裏表と穴の有無の識別

フーリエ記述子を用いた輪郭ベースの手形状認識は純粋に輪郭の形状のみを記述できる。しかし、フーリエ記述子の利用だけでは、例えば手の裏表や穴の有無が識別できない。我々はそれらを輪郭内の別の特徴を見る事で解決しようと考えている。現状では、計算コストを考慮し、輪郭内の輝度値の統計データを追加利用して識別を行うつもりである。3.2.2と合わせて言うと、我々はフーリエ記述子をベースに他の情報を補助的に活用する手形状認識手法を採用する。

3.2.4 ジェスチャ認識

ジェスチャ認識はフレーム毎にその時点での手形状認識を行うと同時に、前フレームでの結果との比較により、

手の移動量や回転量も計算することで行う。既存研究では隠れマルコフモデルをベースにして実現しているものがあるので我々もそれに倣う。なお、移動量や回転量の計算の基となる手の位置や回転角は、それぞれ手領域抽出と回転角の正規化の時に得ることが出来る。

3.2.5 マッチング

認識されたジェスチャを登録済みのコマンドに対応するジェスチャとマッチングを取るには識別器が必要であるが、今回は議論しない。ここでは、より高速な動作のための工夫の一案として、手の運動量に応じた手形状認識の簡素化を提案する。即ち、移動量や回転量が大きく、手が激しく動いている間はそのも鮮明に輪郭が得られないという事を前提として、その様なシーンでは手形状認識を簡略化し、むしろ、ジェスチャの終了状態を見落とさないために、フレームレートを重視するという試みである。幸いにも、フーリエ記述子を用いたマッチングに適用するならば運動量の数値に応じてマッチングに用いるフーリエ係数を減らせば良いだけである。

4. 輪郭内の輝度値情報の比較実験

本格的な開発の着手に先立って3.2.3で述べた手法の可能性を調べるために実験を行った。図2に裏表の例として「グー」、穴の有無の例として「横を向いたグー」の画像とその輝度値のX軸上およびY軸上への線積分画像を示す。どちらも表向きと裏向き、穴が有るのと無いのではヒストグラムの形が異なり、識別可能だと予想される。

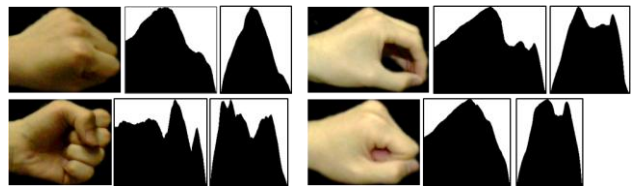


図2 線積分の適用結果。左)裏表の例、右)穴の有無の例、

5. まとめ

我々は広く普及が期待できるHGUIの開発を計画している。本稿ではそのために手軽に利用できるHGUIについて考察し、そのために必要な制約や課題を示した。その上で我々が開発するHGUIの仕様を明らかにし、その実現のための技術検討を行った。今後は、手領域抽出や識別器の構成などの現段階で方式を決定していない部分の検討を進め、それに基づいてHGUIの開発を進める。

参考文献

- [1]石淵 耕一, 岩崎 圭介, 竹村 治雄, 岸野 文郎 “画像処理を用いた実時間手振り推定とヒューマンインタフェースへの応用”, 信学技報, Vol.J79-D-II, No.7, (1996).
- [2]内海 章, 大谷 淳, 中津 良平, “多数カメラを用いた手形状認識とその仮想空間インタフェースへの応用”, 情処論, Vol.40, No.20 (1999).
- [3]坂本 圭, 大竹 敏史, 池 司, 藤田 将洋, “ハンドジェスチャインタフェース技術”, 東芝レビュー. 63, 11, pp.58-62, (2008).
- [4]Charles T. Zahn, Ralph Z. Roskies, "Fourier Descriptors for Plane Closed Curves", IEEE Transactions on Computers, Vol.C-21, No.3, (1972).