

自然な手振りによる ハンドジェスチャ ユーザーインターフェース

Hand Gesture User Interface Using Natural Hand Motions

池 司 中洲 俊信 岡田 隆三

■IKE Tsukasa ■NAKASU Toshiaki ■OKADA Ryuzo

近年、パソコン(PC)などの電子機器をユーザーの手の動き(以下、ハンドジェスチャと記す)で離れたところから操作する、ハンドジェスチャ ユーザーインターフェース(HGUI: Hand Gesture User Interface)があいついで実用化されている。

東芝は、ハンドジェスチャで操作可能なPCを2008年にいち早く商品化したが、これを発展させ、人の自然なコミュニケーション手段である音声とジェスチャで操作するユーザーインターフェース“てぶらナビ”を開発し、2011年9月からPCに搭載している。てぶらナビのHGUIは、上下左右に手を払う自然な動作で操作できることが特長である。今回当社は、このようなジェスチャを高精度に認識するため、手のすばやい動き、姿勢変化に対応した手追跡技術、及び手の払い動作の認識技術を開発した。

Hand gesture recognition technology, which makes it possible for users to remotely control digital devices such as PCs by hand motions, has entered practical application in a succession of commercial products.

Toshiba has been commercializing gesture-controlled PCs since 2008. We have also developed "Tebura-navi," a novel hand gesture interface that introduces voice and gesture control with natural sweeping hand motions for up/down/left/right commands as an effective means of natural communication, emulating that used between humans. Tebura-navi was realized through the following technological advancements: (1) a hand-tracking algorithm for swift hand motions, and (2) a gesture recognition algorithm for sweeping hand motions. We released PCs equipped with Tebura-navi in September 2011.

1 まえがき

近年、リモコンなどのデバイスを用いず手の動きで機器を操作できる、ハンドジェスチャ ユーザーインターフェース(HGUI: Hand Gesture User Interface)技術を搭載した機器があいついで実用化されている。HGUI技術により、手元にリモコンがないときや、手が汚れていたり濡れていたりして機器やリモコンに触れられないときにも機器を操作できるようになる。

東芝は、2008年にPC向けHGUI“ハンドジェスチャリモコン”をいち早く商品化したが、PCを操作するために、長時間握った姿勢を保ちながら手を動かさなければならず、疲労感が大きかった。そこで、音声とハンドジェスチャによる操作を実現した“てぶらナビ”を開発し、2011年に商品化したREGZA PC D731を皮切りに、順次PCへの搭載を進めている(図1)。てぶらナビでは、手を自然に払ったり振ったりする動作で、表示されている画像の切替えや、上下左右へ移動させる操作などをできるようにし、操作性を向上させた。

手を払ったり振ったりする動作は、手の姿勢や光の当たり方が大きく変化するうえ、動きも速い。また、振りかぶりなどユーザーの無意識な手の動きにより、意図しない動作を引き起こすこともある。そこで、これらの動作を高精度に認識できるようにするため、すばやい手の動きや姿勢変化に対応した手追跡技術を開発し、自然な手の動きを追跡できるようにした。



図1. てぶらナビ—自然な手の動きでPCを操作できる。
Tebura-navi voice and gesture integrated user interface

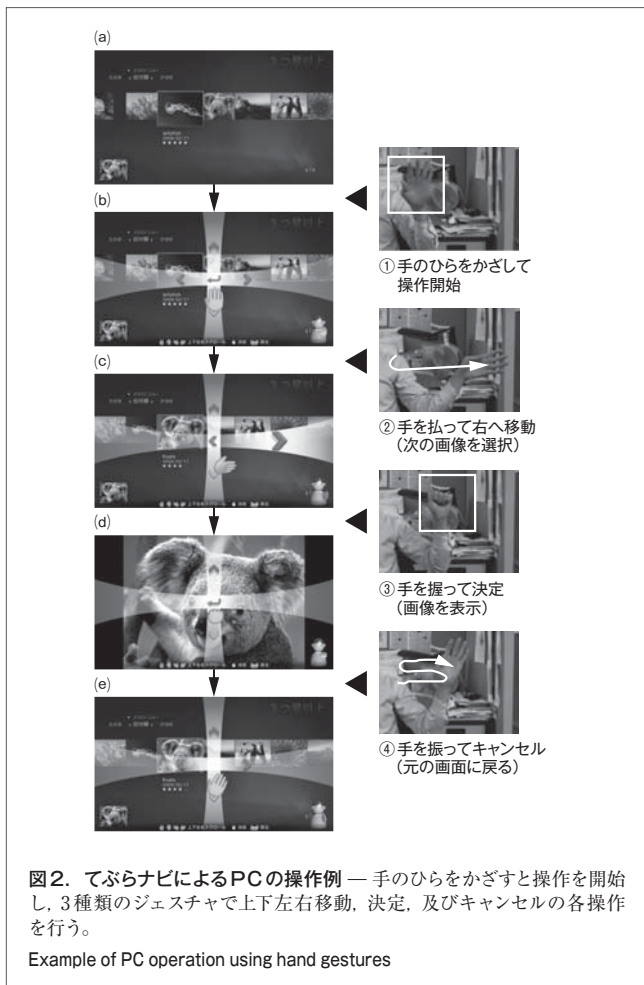
また、手の動きの大きさや速さの遷移パターンから、ユーザーの意図する動きだけを認識するジェスチャ認識技術を開発し、ジェスチャの誤認識を低減した。

ここでは、HGUIの概要、手の検出と追跡、及びジェスチャの認識技術について述べる。

2 HGUIの概要

2.1 てぶらナビ

てぶらナビは、手を払う、握る、振るという3種類のハンドジェスチャを認識する。ユーザーは、これらのハンドジェス



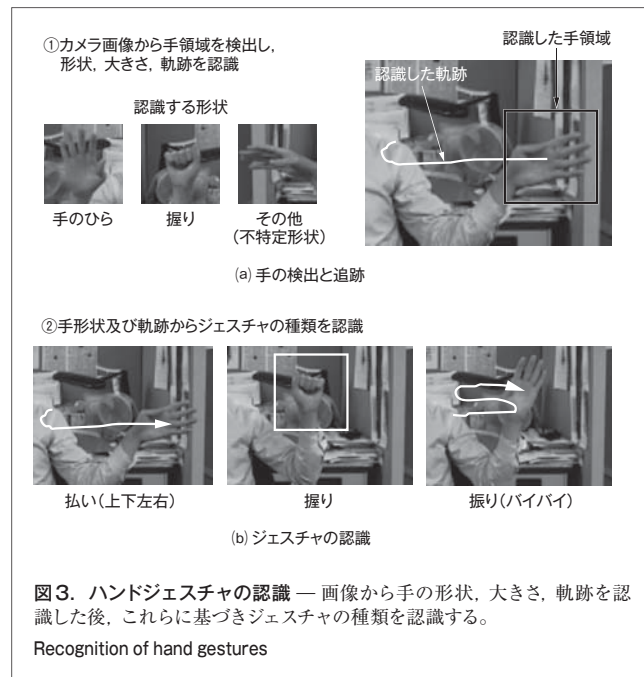
チャを用いて、方向移動、決定、及びキャンセルの操作を行う。図2を用いて、画像ビューアの操作例について述べる。

まず、①手のひらをかざすと画面に操作ガイダンスが表示(図2(b))され、ジェスチャによる操作が可能になる。次に、②手を右に払うと画面のカーソルが右に移動し、右の画像が選択(図2(c))され、③手を握ると選択された画像が全画面表示(図2(d))される。最後に、④手を振ることで元の画面(図2(e))に戻る。

2.2 ハンドジェスチャの認識

ハンドジェスチャの認識は、PCに搭載されたWebカメラで撮影したユーザーの映像を用いて行う。図3を用いて、カメラ映像からハンドジェスチャを認識する処理について述べる。

まず、①得られた画像から手の検出・追跡処理を行い、手領域及びその軌跡を認識する(図3(a))。その際、検出した手領域内の画像について、手のひら形状及び握り形状であるかどうかを識別することで、手形状も同時に認識する。次に、②得られた軌跡と手形状からジェスチャ認識処理を行い、ジェスチャの種類を認識する(図3(b))。手の形状、移動量、速さ、及び移動方向の変化を検出するとともに、ユーザーが意図した手の軌跡だけを抽出する。これらに基づき、払い、握り、振



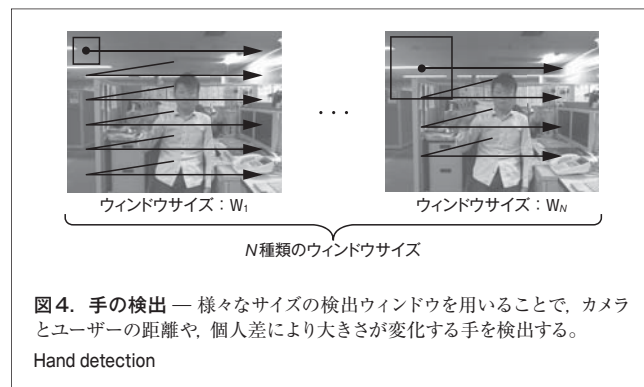
り(バイバイ)の各ジェスチャを認識する。

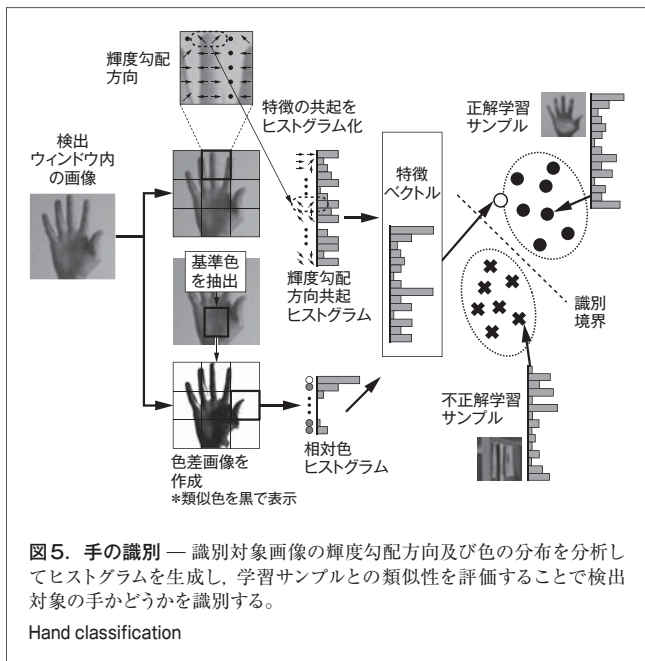
3 手の検出と追跡

3.1 手の検出

かざした手のひらを検出するため、図4に示すように、様々な大きさのウィンドウで画面を走査し、検出ウィンドウ内の画像に対して手のひらの識別処理を行う。これにより、手の大きさの個人差や、カメラからの距離に関わらず手のひらを検出できる。図5を用いて、各ウィンドウの識別処理について述べる。

まず、検出ウィンドウ内の画像を部分領域に分割し、各部分領域内で輝度勾配方向(各画素に対して、より高い輝度の画素が存在する方向)、及び相対色(基準色に対する各画素の色差)を算出する。次に、各部分領域での輝度勾配方向の共起、及び相対色の各出現頻度からヒストグラムを生成し、特徴ベクトルとする。多数の正解及び不正解サンプルの特徴ベクトルから、あらかじめ学習により求めておいた識別境界に基づ





き、どちらのサンプルに近いかを判定することで、手のひらかどうかを識別する。

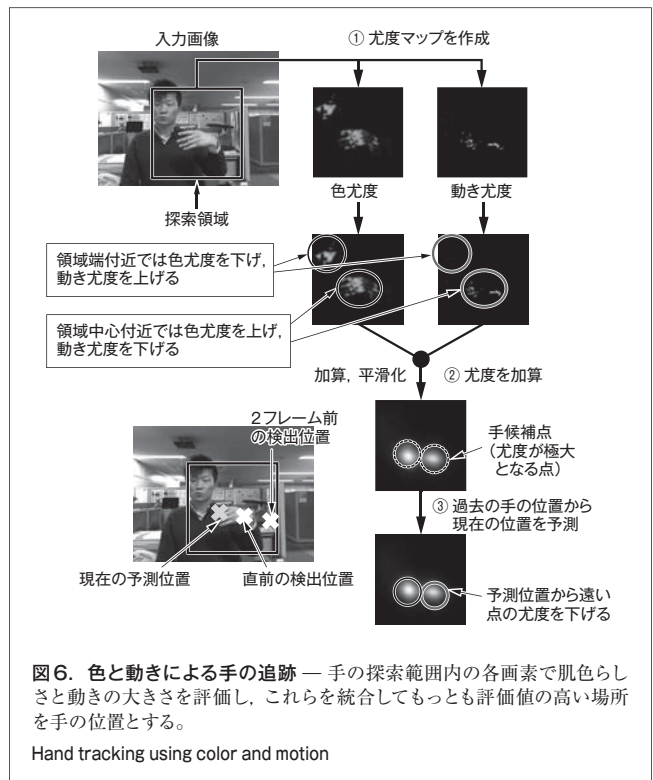
輝度勾配方向を使うことで、照明環境による輝度差変化にロバスト（頑健）な処理となり、手の輪郭などの形状情報を抽出できる。また、共起を使うことで、例えば指の先端の曲線などで観測される、隣接画素間での輝度勾配方向の変化も特徴に使えるため、識別性能が向上する⁽¹⁾。更に、相対色を使うことで、照明環境による色成分への影響を抑え、安定して類似色パターンを抽出できる。これらに加え、ヒストグラム化により手の形状や姿勢の個人差を吸収して安定に識別できる。

3.2 手の追跡

手のひらを検出すると、次のフレームからは検出した手領域が画像上のどこへ移動したかを追跡することで、手の移動の軌跡を得る。追跡処理の流れを図6に示す。

まず、直前に手が検出された位置を基準に探索領域（手を探索する領域）を設定し、①探索領域内におけるフレーム間の輝度値変化量の分布を示す動き尤度（ゆうど）マップ、及び画素値の肌色らしさの分布を示す色尤度マップを生成する。次に、②これら二つの尤度マップを加算と平滑化した後、極大点を検出して手候補点とする。最後に、③過去の手の位置から、等速直線運動モデルに基づき現在の位置を予測し、予測位置からの距離に応じて各手候補点の尤度を下げた後、もっとも尤度の高い点を手の位置とする。これにより、手候補点が複数検出された場合でも、動きの連続性に基づき手を安定に追跡できる。

色尤度は、手の検出時に手のひら領域から取得した肌の色を基準に、各画素の色の類似性から算出する。このとき、様々な環境で撮影した手画像の肌色から、事前に手の表裏の



色の違いや、光の当たり具合による色の変化を学習し、これらの変化を考慮して色尤度を算出することで、払いや振りによって手の姿勢が変化しても安定に手を追跡できる。

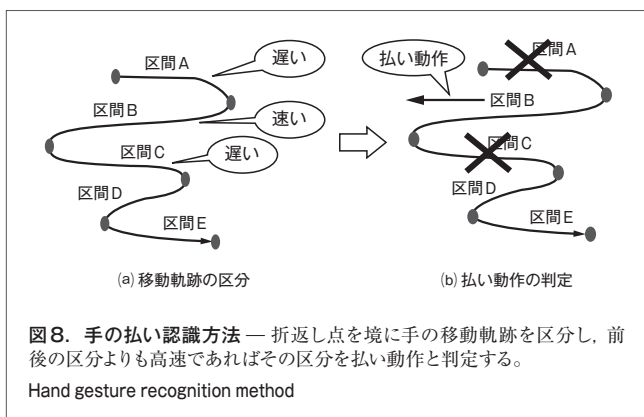
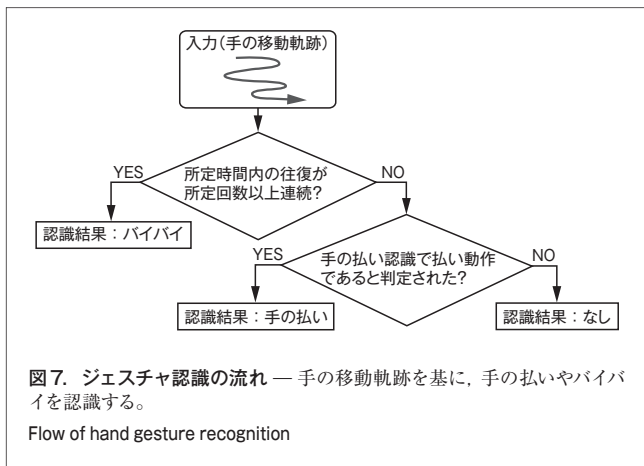
また、手順②で尤度を加算する際、探索領域中心からの距離に応じて色尤度の重みを下げ、動き尤度の重みを上げる。これにより、手の動きが速いときに画像がぶれて、手領域に背景の色が混じり肌色が観測されにくくなっても、手の移動量が大きいときは動きを重視して手を安定に追跡できる。

4 ジェスチャの認識

3.2節で追跡された手の移動軌跡を基に、それが上下左右方向の手の払いであるかバイバイであるかを認識する。ジェスチャ認識の流れを図7に示す。

まず、所定時間内の往復が所定回数以上連続しているかどうかを判定し、一定回数以上連続していれば“バイバイ”として認識する。それ以外の場合は、手の移動軌跡が手の払いであるか否かを判定するとともに、払い方向についても認識する。

一般に、自然な手の払い動作には、振りの大きさや速さに個人差が生じる。また、ユーザーの意図した方向への“払い動作”のほかに、意図しない方向への振りかぶりや振り戻しが生じる。このように個人差やノイズを含んだ手の移動軌跡から払い動作だけを認識することが課題となる。複数の被験者を対象とした予備実験により、払い動作は振りかぶりや振り戻しよりも高速であるという性質が共通して見られた。この手法で

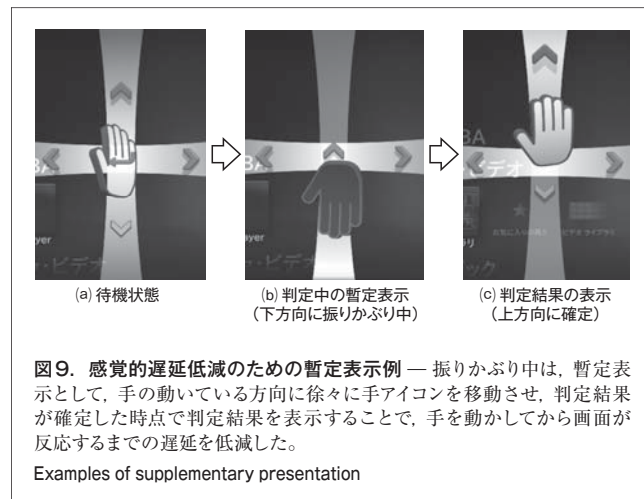


は、この知見を利用する。

具体的には、まず、手の移動方向が変化する点(折返し点)を検出し、これを境に手の軌跡を区分する(図8の区間A~E)。次に、各区間の手の速さをその前後の区間と比較し、いずれの区間よりも速い場合に、その区間を払い動作と判定する(図8(b))。これにより、知見に基づき払い動作(区間B)を検出するとともに、直後の振り戻し(区間C)、及び直前の振りかぶり(区間A)を棄却する。またこの手法では、一人のユーザーの手振りにおける相対的速度で比較するため、振り位置や、振りの大きさや速さといった個人差も吸収できる。

なお、この手法では、手の払いが上方向であるか下方向であるかを認識する処理と、手の払いが右方向であるか左方向であるかを認識する処理が並列に動いている。手の動きの上下方向成分と左右方向成分のうち、折返し点からの移動距離や速さの大きいほうを優位であると認識し、最終的に上下左右のどちらか一方向に決定する。

また、直後の状況も考慮しているため、原理的に手の払い方向が確定するまでに遅延が生じる。そこで図9に示すように、手の払い方向が確定するまでの間は、手の動いている方向に手アイコンを徐々に移動させることでユーザーの手の振りに反応していることをフィードバックし(図9(b))、払い方向が確定したら当該方向への手アイコン移動(図9(c))と確定音提



示を行うような工夫により、ユーザーの感覚的な遅延を低減し、操作性を損なわないようにした。

5 あとがき

手の姿勢や光の当たり方の変化に対応することができ、速い動きも追跡できる手追跡技術、及びユーザーの意図した動きだけを認識するジェスチャ認識技術を開発した。これらの技術により、手の払いなどの自然な動きでPCを操作できるようにし、操作性を向上させた。

今後は、より操作の手間を減らすような操作方法を検討するとともに、操作中の状態をユーザーにより適切にフィードバックする方法を検討し、いっそうの操作性向上に取り組んでいく。

文献

- (1) Watanabe, T. et al. "Co-occurrence Histograms of Oriented Gradients for Pedestrian Detection". Proc. 3rd Pacific Rim Symposium on Image and Video Technology. Tokyo, 2009-01, National Institute of Informatics, Germany, Springer, 2009, p.37 - 47.



池 司 IKE Tsukasa, Ph.D.

研究開発センター インタラクティブメディアラボラトリー研究主務、博士(情報科学)。画像認識技術及びユーザーインタフェースの研究・開発に従事。IEEE、電子情報通信学会会員。Interactive Media Lab.



中洲 俊信 NAKASU Toshiaki, Ph.D.

研究開発センター インタラクティブメディアラボラトリー、博士(学際情報学)。ユーザーインタフェースの研究・開発に従事。映像情報メディア学会、日本顔学会会員。Interactive Media Lab.



岡田 隆三 OKADA Ryuzo, Ph.D.

研究開発センター インタラクティブメディアラボラトリー主任研究員、博士(工学)。画像認識技術の研究・開発に従事。電子情報通信学会会員。Interactive Media Lab.