

1-24 ヘッドマウントカメラを用いた日常生活動作の抽出法 The Method to Detect One's Activities of Daily Living Using Head Mounted Camera

飯島 俊匡[†]
Toshimasa Iijima

川嶋 稔夫[‡]
Toshio Kawashima

青木 由直[†]
Yoshinao Aoki

1. はじめに

我々は人物の日常生活の情報を蓄積・再利用することで記憶や思考の補助とする研究 [1]-[2] を行っている。以前よりコンピュータビジョンの研究においては、人物の移動や手足の動きを認識する研究は多くなされており、監視カメラの映像を解析することにより人間の行動を認識する研究などが盛んに行われている。しかしこれらのアプローチは人物の行動範囲が制限された状況を想定して実現されているため、日常生活のすべての場面にこれらの手法を適応するのは困難である。そこで我々はシステムを環境に用意するのではなく、ユーザ側に用意する。つまりウェアラブルコンピュータを用いることでこれらの問題の解決を図る。この場合人物の行動範囲に制限は無くなり、使用者ごとにシステムが独立するためプライバシーの問題も生じない。

また、これまでの実験により人間の行動は常に視覚による認識を経て動作へと移ることがわかっているため、ユーザへの負担を最小限にする上で入力をヘッドマウントカメラによる画像のみとし、得られた視点映像列を解析することでユーザの動作を推定することが可能である。本稿では、人物の顔や手などの対象領域を抽出し、頭部の動揺を補正しつつその動きを追跡することで動作の判別を行う手法について述べる。

2. 動作の検出

2.1 自分の動きの推定

視点映像は人物の頭部の動揺により、不安定な状態であるのでそれを補正する技術が必要となる。ここではオプティカルフローを用いて頭部の動揺の影響を軽減させる。オプティカルフローはノイズに極めて敏感であるため、ここでは前処理としてガウシアンフィルタによるノイズの低減処理を行い、より安定したフローベクトルを得るために多重解像度戦略をとる。

ここで、得られた速度分布を縦3、横3の計9個所の区画に分け、それぞれの区画ごとにフローベクトルの平均を計算し、各区画のフローベクトルの向きと大きさから自分の動きを推定する。ここでは左右の縦3ブロックの平均ベクトルが吸出し状になっていた場合、人物は前進移動していると判断する。また、それ以外の場合は6ブロック以上の平均ベクトルが同じ方向である時、その平均ベクトルを首振りによる視線の移動ベクトル v_e とする。また、頭部の動揺がおきても体の向きは変わらないという特徴を利用すると、 $|\sum_i v_e| < \text{threshold}$ という条件式により、頭部は動揺しており人物は移動していないと判断できる。

2.2 肌色領域の抽出

次に入力画像列から人間の手や顔領域を抽出するために、色情報による特徴量を用いて肌色領域を抽出する。先にも述べたとおり、本稿では日常生活における映像から抽出を行うことを目的としているため、照明条件の変化が激しい。そこで特に暗くなることによるコントラストの低下、白熱灯などの光源の色によるカラーバランスの崩れをイコライズ処理により補正する。ここでは、コントラストの補正は線形補完を、イコライズはヒストグラム中心法を用いる。

人間の肌の色は人種、民族の多様性にもかかわらず色相がほぼ一定で、一般的な背景に対して比較的彩度が高いという特徴が知られている。そこで補正された入力画像をHSV色空間に変換し、色相と彩度を元にした肌色類似度マップを利用して人物の顔や手領域を取得する。

また、人間の手は凹凸のある複雑な立体構造であるため、そこには若干ながら陰影が生じる。そこで、肌色領域でかつ微量の輝度変化が存在する領域を手領域として取得する。この時、得られた領域が一定の大きさ以下である場合は、ノイズとして除去する。図1に元画像、HSV色空間による肌色領域画像、輝度変化画像、それらを組み合わせて得られた手領域の抽出結果の例を示す。

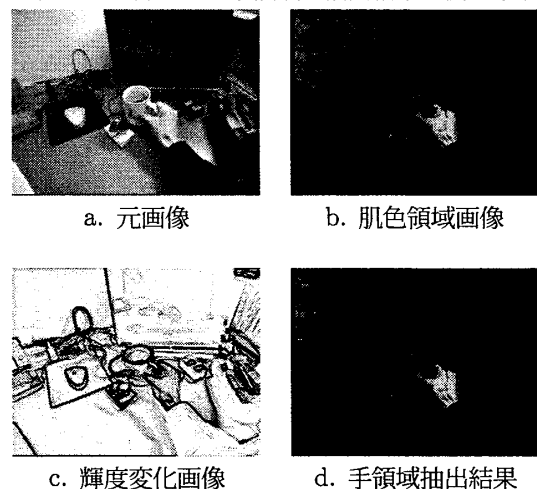


図1: 手領域抽出

2.3 手の動きのトラッキング

手のトラッキングは何らかの動作が行われてから、つまり視界に手が入ってきた時点で開始となる。また、手を使った動作はそのほとんどが視界の下部から手が進入してくるので、視点映像列の下部の外縁領域を常時監視して手の検出を行う。その際視点と両手の位置関係より画面右側より侵入してきた領域を右手、左側を左手としてトラッキングを行う。トラッキングは前フレームで得られた手領域を囲むウィンドウと現フレームとのブロッ

[†]北海道大学大学院 工学研究科

[‡]公立はこだて未来大学 情報アーキテクチャ学科

クマッチングにより行い、対応の取れた各領域の重心を手の中心とし、手の移動ベクトル v_p とする。しかし、ここで得られた手の移動ベクトル v_p は頭部の動揺の影響を受けた見かけの手の移動ベクトルである。頭部の移動ベクトルは前節で得た v_e であるので、真の手の移動ベクトル v_h は $v_h = v_p + v_e$ で得られる。これらの移動ベクトルは動作が完了するまで、つまり手が視点映像から出て行くまで、ベクトル列 $v_h(t), v_e(t)$ として保存しておく。また、動作対象の抽出とその動作の判定を行うために、その間の視点映像列 $I_{tmp}(t)$ 、および手が最後に静止した視点映像上の座標 (x_{stat}, y_{stat}) 、右手か左手かの情報を一時的に保管しておく。

3. 動作の分類と判別

これまで得られた人物の移動情報、およびその手の運動から動作の分類を行う。それらの組み合わせとして以下のようなものが考えられる。

1. 移動 ⇒ 短時間の動作 ⇒ 移動：物の操作
2. 移動 ⇒ 長時間の動作 ⇒ 移動：会話
3. 静止 ⇒ 短時間の動作 ⇒ 静止：物の取捨
4. 静止 ⇒ 長時間の動作 ⇒ 静止：デスクワーク

日常生活ではこれらの動作が繰り返して現れ、1. や 3. のような状況は一瞬しか現れないため記憶に残りづらい。しかしこれらの動作はそのときのユーザの思考や意識の流れを示す重要なものであると考えられる。そこで、ユーザが使った物の履歴を取るため物を置く、拾うといった動作の検出を考える。物を置いたり拾ったりする動作は、

1. 手が置く(拾う)場所まで移動し
2. 対象物を置き(拾い)
3. 手が視線映像から出る

という一連の動きによって行われる。そこで前節で得た手の移動ベクトル列から手の動きを取得し、置く拾う動作に適合するかどうか判定する。

まず、手が視点映像に入ってきた時刻を t_{in} 、出て行った時刻を t_{out} 、最後に静止した時刻を t_{st} とする。時刻 t_{st} が存在しない場合は手が横切っただけとし、検索対象からははずす。また、 $t_{out} - t_{in}$ が閾値より大きい場合は、そのときの手の動作は停滞であると考えられるため、これも置く拾う動作の検索対象からははずす。

次に時刻 t_{st} の手の視点映像上の座標 (x_{st}, y_{st}) から、右手の場合は左に、左手の場合は右にオフセットさせた動作の対象領域 I_{obj} の中心座標 (x_{obj}, y_{obj}) を得る。また、両手が同時に現れている場合は両手の座標の中点を対象領域の中心座標とする。

$$(x_{obj}, y_{obj}) = (x_{st}, y_{st}) \pm offset$$

すると時刻 t_{in} 、 t_{out} におけるその動作の対象領域 I_{in}, I_{out} の中心座標 $(x_{in}, y_{in}), (x_{out}, y_{out})$ は次のようにして得られる。

$$(x_{in}, y_{in}) = (x_{obj}, y_{obj}) + \sum_{t=t_{in}}^{t_{st}} v_e(t)$$

$$(x_{out}, y_{out}) = (x_{obj}, y_{obj}) + \sum_{t=t_{st}}^{t_{out}} v_e(t)$$

この3つの領域 $(I_{in}, I_{st}, I_{out})$ のどの組に同じ対象物が存在するか調べることで、そのときの動作の種類を得ることが出来る。

- $I_{st} = I_{in}$ かつ $I_{st} \neq I_{out} \Leftrightarrow$ 置く動作
- $I_{st} \neq I_{in}$ かつ $I_{st} = I_{out} \Leftrightarrow$ 拾う動作
- $I_{st} = I_{in}$ かつ $I_{st} = I_{out} \Leftrightarrow$ 操作 / 何もしない

これをイメージであらわすと図2のようになる。

ある時刻の動作対象領域の組が同一対象物を含む場合、RGB 各要素に対するヒストグラムの分布はほぼ同じように現れると考えられる。そこで、各対象領域それぞれの RGB 各要素に対する多次元ヒストグラムを取り、その分布の距離が閾値以下であれば同一対象物を含むとして動作の判定を行う。

4. 実験

実際に本手法を利用して、ユーザの3日間の行動履歴を取得する実験を行った。実験の結果については発表の際に紹介する。

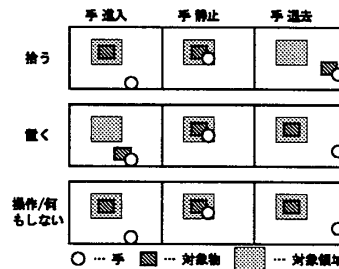


図2: 動作の判定

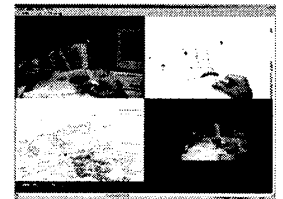


図3: 解析の様子

5. おわりに

本稿では環境による制限を受けずに人物の動作の検出を行う手法を提案した。その際の技術的な課題である視点映像列からの手の動きの追跡と動作の認識についてその手法を述べ、実験を行った。

現段階における問題点としては、対象物が小さい場合の検出が困難である、頭部の動きと目の動きの差により手の動きを見失う場合がある、背景人物などの存在により手を誤検出する、といったことがあげられる。

今後の課題として、前述の問題点の解決、手による動作の判別手法として、手の形状を識別することでより詳細な動作の分類を行うことなどがあげられる。

参考文献

- [1] 飯島俊匡, 川嶋稔夫, 青木由直, “日常視点映像列の解析による動作抽出手法”, 信学技報, PRMU2001-50, Jul. 2001.
- [2] T. Iijima, T. Kawashima, Y. Aoki, “Personal history management by monitoring action in everyday life”, Proc. of IWSCI'99, pp.383-388, June 1999.