

オプティカルフローを用いた身振り認識

4K-6

—人-カメラ間距離に基づくフロー検出パラメータの適応的選択法—

○西川 敦 西村 正典 平野 敦士 小荒 健吾 宮崎 文夫

大阪大学大学院 基礎工学研究科 システム人間系専攻

1. はじめに

近年、動画像に基づく人間の身振り動作の認識手法の研究開発が盛んに行われている。我々は、既報告において、連続動画像の局所相関演算に基づくオプティカルフローを用いた身振り認識法を提案し、予め定められた位置で行った身振り動作に対しては、被験者によらず高い認識率を達成した¹⁾。本報告では、オプティカルフローを検出する際のいくつかのパラメータを“身振りを行う手とカメラ間の距離”に基づき適応的に調整する手法を新たに提案する。本手法を適用すれば、人-カメラ間距離が大きく変化(1m~8m)しても、身振りの認識率を常に高いレベル(不特定の動作者に対して90%程度)に保持できることを実験により示す。

2. 身振り認識手法の概要

我々は、1台のCCDカメラにより撮影される連続動画像のオプティカルフローから算出される手の移動方向の変化率に基づいて、身振り動作を予め用意されたいくつかの「基本動作」の連続として分割・記号化することにより、不特定の動作者に対しても高い認識率を達成できる身振り認識手法を提案した。この詳細については文献1)に譲るが、その概略は次のとおりである：(i)画像中に設定された動作検出用ウィンドウ内に格子点を等間隔に設置し、(ii)各格子点を基準とした連続するフレーム間での局所相関演算に基づきオプティカルフローを求め(ただし相関演算の際必要に応じて画素を間引くものとする)、(iii)その大きさがある閾値 V_{large} 以上のフローベクトル群の総和をとることにより手の移動方向を時々刻々推定し、(iv)求めた手の移動方向の“変化率(曲率)”の時系列をベースに一連の身振り動作を手を曲げる、折り返す、回転させる、弧を描くといった「基本動作」の系列として記号化し、(v)得られる記号列が認識対象モデルの記号列と完全に一致した場合のみ正しく認識できたとみなす。この手続きにより高い認識率を得るためには、“手の動作に起因する”フローベクトルが安定して多く求まることが不可欠である。具体的には、①身振りを行う手の見掛け上の投影面積が十分に大きく、手を動かす範囲が動作検出ウィンドウ内に収まっていること、②単位サンプリングあたりの手の見掛け上の移動量がフローの探索範囲内に収まっていること、が必要不可欠である。既報告²⁾では、ある一定の“身振り動作者-カメラ間距離”を想定し、その距離関係においてうまくフローベクトル(身振り動作)が検出されるように各

種パラメータ(格子点数やその間隔、画素の間引き率など)を予め調整した上で、これを基準に距離に応じて単純な比例関係に基づき上記パラメータを再設定することにより①,②の問題に対処したが、その調整法は経験的であり、また、身振り動作者とカメラ間の距離が初期パラメータ設定時に比べて大きく変動すると(具体的には、同距離が4mを越えると)認識率が大きく低下することも明らかになった³⁾。ここでは、これらの問題に対処するため、いくつかのフロー検出パラメータを手の動作範囲や手の画像上での大きさおよびその平均的な移動速度に関する拘束条件に基づいて“手-カメラ間距離”の関数として定める新たな手法を導入する。

3. フロー検出パラメータの適応的選択法

以下で、本稿で新たに提案する(1)動作検出ウィンドウ内に設置される格子点の数とその間隔の設定方法、および、(2)局所相関演算の際の画素の間引き率と画像を取り込む時間周期の設定方法について順番に説明する。

3.1 格子点の数と格子点間隔

格子点の数(水平方向 I 、垂直方向 J)と格子点間隔(水平方向 Δ_x 、垂直方向 Δ_y)が満たすべき条件は動作検出領域の大きさおよび手の画像上での大きさに関する拘束条件から導出される。

3.1.1 動作検出領域の大きさから導かれる条件

いま身振りを行う手とカメラ間の距離を D (mm)とおく。身振り動作者は、距離 D の位置にある画像平面と平行な横 X (mm)、縦 Y (mm)の矩形領域の中で身振りを行うと仮定する。ここでは、この矩形領域の投影が動作検出領域にほぼ一致するように格子点の数と格子点間隔を調整することを考える。動作検出領域は、その定義より、横 $(I \cdot \Delta_x)$ (pixel)、縦 $(J \cdot \Delta_y)$ (pixel)の大きさの矩形となるので、ただちに、 (I, J) と (Δ_x, Δ_y) が満たすべき次の関係を得る：

$$I \cdot \Delta_x \approx F_x \cdot X/D, \quad J \cdot \Delta_y \approx F_y \cdot Y/D \quad (1)$$

ここに、 F_x, F_y はそれぞれ水平方向、垂直方向の等価焦点距離(単位: pixel)である。我々の用いたシステムでは $F_x = 912, F_y = 864$ である。

3.1.2 手の画像上での大きさから導かれる条件

次に手の画像上での大きさについて考察するために、その基準の姿勢として、“親指を上にして手の全指先をカメラの正面に向ける”場合を考える。このときの手の幅および高さをそれぞれ H_w (mm)、 H_h (mm)とおく。いま、この手を取り囲む矩形領域 $(H_w \times H_h)$ (mm)の画像上への投影がある程度数 $(K \times L)$ 個以上の格子点を占めるように格子点間隔を制御することを考える。手とカメラ間の距離を D (mm)とすると、ただちに、 (Δ_x, Δ_y) が満たすべき次の関係を得る。

$$K \cdot \Delta_x \leq F_x \cdot H_w/D, \quad L \cdot \Delta_y \leq F_y \cdot H_h/D \quad (2)$$

Gesture Recognition Using Optical Flow -Adaptive selection of flow detection parameters based on the distance between the gesturer's hand and camera-

Atsushi Nishikawa, Masanori Nishimura, Atsushi Hirano, Kengo Koara, and Fumio Miyazaki
Department of Systems and Human Science, Graduate School of Engineering Science, Osaka University

3.1.3 格子点数と格子点間隔の決定

ここで、式(1)と式(2)を同時に満たす格子点数(I, J)と格子点間隔(Δ_x, Δ_y)の組を選択することを考える。数学的には、この組合せは無数に存在するので、本稿では次の方策をとる：まず、式(1)を満たす任意の(Δ_x, Δ_y)について式(2)が常に満たされることを保証する(I, J)の条件を求め、その中で I, J とも最小となるものを選択する。こうして定まった格子点数(I, J)と式(1)に基づいて、格子点間隔(Δ_x, Δ_y)を手とカメラ間の距離 D の関数として定める。最小の I, J を選ぶ理由は、オプティカルフローの計算コストをできる限り小さくするためである。また(Δ_x, Δ_y)のみを距離 D の関数とする理由は、距離 D によらずフローの計算量を一定にするためである。この方針に従えば、まず、式(1),(2)より、ただちに、格子点数(I, J)が満たすべき次の条件を得る：

$$I > C \cdot K \cdot X / H_w \quad \text{かつ} \quad J > C \cdot L \cdot Y / H_h \quad (3)$$

ただし、 C は1より大きな正数であり、ここでは $C = 1.1$ とした。いま、実測に基づき、身振りを行う矩形領域ならびに手の幅および高さをそれぞれ $X = 350(\text{mm})$ 、 $Y = 300(\text{mm})$ 、 $H_w = 50(\text{mm})$ 、 $H_h = 100(\text{mm})$ と設定する。定数 K, L については、手の幅と高さの比率も考慮し、ここでは、 $K = 2.5$ 、 $L = 5$ とする。このとき、式(3)より、 $I > 19.3$ 、 $J > 16.5$ となるので、 $I = 20$ 、 $J = 17$ と定める。次に、式(1)に実際のパラメータおよび上記(I, J)を代入して(Δ_x, Δ_y)について整理すると、 $\Delta_x \approx 16000/D$ および $\Delta_y \approx 15200/D$ となるので、ここでは、簡単のため、両者の平均をとって、 $\Delta_x = \Delta_y = [15600/D + 0.5]$ と定める。ただし、 $[\]$ はガウスの記号である。

3.2 画素の間引き率と画像サンプリング周期

次に、単位サンプリング当たりの身振り動作者の手の見掛け上の移動量がフロー探索範囲内に常に収まるように画素の間引き率と画像サンプリング周期を制御することを考える。いま、 T フレーム周期(ただし1フレーム $= 1/30$ 秒)で身振り動画像を取り込むものとし、それぞれの画像において各格子点を基準に間引き率 Q で(つまり Q 画素おきに抜き出した画素のみを用いて)局所相関演算を行うものとする。ここで、 $Q = 0$ かつ $T = 1$ のときに手の見掛け上の移動量の平均値が V_{std} 画素となる距離を基準距離 D_{std} と定め、次の問題を解くことを考える。

問題 1 身振りを行う手とカメラ間の距離 D が与えられたとき、 $\{T/(1+Q)\} \cdot (D_{std}/D) \approx 1$ が成立するように非負の整数 Q および正の整数 T を制御する。

数学的には同一の解を与える(Q, T)の組合せが無数に存在するので、ここでは次のような方策をとる。

1. $D \leq D_{std}$ ならば、サンプリング周期 $T = 1$ と固定し、間引き率 $Q = [D_{std}/D - 0.5]$ と定める。
2. $D > D_{std}$ ならば、間引き率 $Q = 0$ と固定し、サンプリング周期 $T = [D/D_{std} + 0.5]$ と定める。

ただし $[\]$ はガウスの記号である。基準距離は、 $D_{std} = F' \cdot H_v / V_{std}$ で近似的に与える。ここに、 $F' = (F_x + F_y)/2$ であり、 H_v は複数の被験者に実際にいくつかの身振りを行って得られた結果得られる身振りの“平均的な”速度である。実測により、 $H_v = 15(\text{mm}/\text{frame})$ とした。また V_{std} は、ここでは、フロー探索範囲の半径(これは使用する専用ハードウェアの制約から7 pixelである)の1/2に設定するものとする；すなわち、 $V_{std} = 3.5(\text{pixel})$ 。これらより、基準距離 $D_{std} = 3800(\text{mm})$ となる。

Table 1 実験に用いたフロー検出パラメータ

フロー検出用パラメータ	D			
	1000	2000	4000	8000
格子点数(水平) I	20			
格子点数(垂直) J	17			
格子点間隔(水平) Δ_x	16	8	4	2
格子点間隔(垂直) Δ_y				
フロー探索範囲半径	7			
間引き率 Q	3	1	0	
サンプリング周期 T	1		1	2
動作検出用閾値 V_{large}	2			
相関演算ブロックサイズ	16 × 16			

Table 2 手-カメラ間距離 $D = 1000, 2000, 4000, 8000(\text{mm})$ における身振り認識率(単位: [%])

$D(\text{mm})$	被験者						合計
	A	B	C	D	E	F	
1000	93.3	98.3	89.2	90.0	98.3	100	94.9
2000	90.0	94.2	90.8	90.8	97.5	93.3	92.8
4000	84.2	90.0	88.3	89.2	97.5	90.8	90.0
8000	87.5	92.5	87.5	85.0	91.7	96.7	90.1

なお、相関演算ブロックのサイズは、ここでは、簡単のため、 D によらず一定とし、基準距離における基準姿勢(3.1.2節参照)での手の投影幅と使用ハードウェアの制約(8の倍数でなければならない)をもとに、 16×16 (pixel)と設定する。また、動作者の無意識的な小さな動きに起因するフローを除去するための閾値 V_{large} (2.節(iii)参照)は、手の動きによるフローを V_{std} 近傍に押えていることも考慮し、 $V_{large} = 2$ (pixel)と設定した。

4. 実験結果

手-カメラ間距離 $D = 1000, 2000, 4000, 8000(\text{mm})$ とし、各距離において6人の被験者に次の4種類の身振り(①手を左右に2回振る、②ギリシャ文字 Ω を描く、③ワイパーのように手を振る、④四角形を描く、の4つ；文献1)参照)を各30回ずつ行ってもらい、身振り認識率を測定した。オプティカルフローの計算には富士通社製のカラートラッキングビジョン⁴⁾を利用(Linuxにより制御⁵⁾)した。提案手法により定めたフロー検出パラメータをTable 1に、認識実験の結果(各距離における被験者毎の身振り認識率)をTable 2に示す。本稿で示した手法の導入により、手-カメラ間距離1m~8mの広い範囲にわたり、身振り認識率の高いレベルの保持(被験者平均で90%以上)に成功していることがわかる。

参考文献

- 1) 西川ら：“連続動画像からのオプティカルフローを用いた身振り動作の曲率に基づく分割と認識,” 画像の認識・理解シンポジウム(MIRU'98)論文集, vol.II, 35-40 (1998).
- 2) 西川ら：“オプティカルフローを用いた身振りインタフェース-人とカメラの位置関係が大きく変動する場合への対応-,” 平成10年電気関係学会関西支部連合大会論文集, G358 (1998).
- 3) 西川ら：“局所相関演算に基づくオプティカルフローを用いた身振り認識-画素の間引き率と画像サンプリング周期の制御による身振り認識可能な距離範囲の拡大-,” 1999年電子情報通信学会総合大会論文集 (1999).
- 4) 森田ら：“カラートラッキングビジョン,” 第14回日本ロボット学会学術講演会予稿集, 279-280 (1996).
- 5) 松本ら：“PCベースのハイパーマシンの活用カーネル,” 第15回日本ロボット学会学術講演会予稿集, 979-980 (1997).