

動画像動作認識の為の移動領域図形抽出(第2報)

—時間的重なりを考慮した移動領域図形抽出法—

北海道大学工学部 ○国米祐司 洪川勝久 岸浪建史

要旨

動画像から動作認識を行う際、各画像フレーム中からの運動物体(移動領域図形)抽出は重要である。これまでは連続差分画像を用いる事により移動領域図形を求める事が可能であったが、その際、時間的に連続したフレーム間で運動物体が重なっていると精度の良い移動領域図形を抽出できないという問題があった。そこで本報告では前処理として大体の図形重心を求め、それを利用して精度よく移動領域図形を抽出する方法を提案する。

1. はじめに

様々な状況下における機械(ロボット)との協調作業において、非接触型ユーザインタフェースの実現は極めて重要であり、もしコンピュータによる動きを識別できる視覚が実現されたならば、それに必要な情報を得ることができる。特に、手振り動作は、直感的かつ幅広い表現力を持つため、非接触型ユーザインタフェースにとって有効であると考えられる。

本研究では手振り動作が腕全体の形状(移動領域図形)の変化と、その位置変化により表されると考え、その時間的な変化情報を用いて動作をモデル化し、動画像から動作認識を行うことを目的とする。その際、より多種類の動作を認識・識別するには動画像から動作部分の詳細な図形抽出を行うことが重要である。今までは詳細形状を得るために、フレーム間時間平均を用いたノイズ削減、連続差分画像を用いることにより、背景・衣服・ノイズ(手振り時に起こる体の微妙な動き等)の影響を受けない輪郭図形を得ることができた([1])。しかし、動作部分が時間的に重なると精度が落ちる問題があった。そこで本報告では、その問題解決の為にまず大体の図形重心を求め、次にそれを利用した連続差分画像を行う方法を提案し、その有効性について検討する。

2. 詳細な移動領域(動作部分)図形抽出法

本手法では体の微妙な動きなどを含まない大まかな移動領域図形と詳細な移動領域(不完全な移動領域図形)を別々に求めた後、両者の論理積を取り、詳細な移動領域図形を求める。

従来手法では大まかな移動領域図形は連続差分画像を利用して求めるが、動作部分が時間的に重なると移動領域図形が欠けてしまい、精度良く抽出できなかった。そこで本手法(図1)では、まず大体の図形重心を求め、続いてそれを用いて時間的重なりのない連続差分画像を行い精度のよい大まかな移動領域図形を求め、詳細な移動領域との論理積をとる。

3. 大体の移動領域図形重心抽出

動作部分の時間的重なりを避ける為に各フレーム中の動作部分の位置を求める。その為に以下の処理により大体ではあるが移動領域図形の重心を求める。

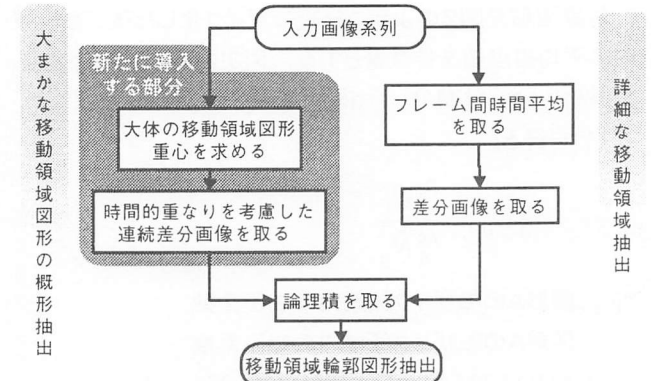


図1 詳細な移動領域図形抽出の流れ

3.1. 入力画像系列

本研究では市販のビデオカメラで撮影した動画像を、1画素256階調のRGB(赤緑青)表色系(320×240ピクセル)のカラービットマップをとって保存し、入力画像系列とする。

3.2. モザイク化

動きを大まかにとらえ、画面中の微妙なちらつき等の影響を減らす為にRGBそれぞれについて4×4ピクセルのモザイク化を行う。

3.3. 差分画像

画像フレーム間の差分画像を求める。2枚のフレーム間の差から、画像を定常領域と変化領域に分割する。RGBそれぞれについて差分画像を2値化(定常領域は0、変化領域は1に)し、最後に論理和をとる。差分画像は計算が容易であるという利点がある。

3.4. 連続差分画像([2])

差分画像1枚では2画面間の変化部分のみが求められ、動作部分の概形(移動領域図形)は求められない。その為、式(1)のように時間間隔を変えた複数の差分画像の論理積をとる(連続差分画像)。そうすることにより時刻 t の動作部分の概形が求められる。

$$f_{SD}(x,y,t) = f_D(x,y,t,t+\Delta t) \cap \dots \cap f_D(x,y,t,t+n \times \Delta t) \quad (1)$$

$$0 \leq x \leq 320, 0 \leq y \leq 240, n=1,2,3,\dots,N$$

但し、 $f_D(x,y,t,t')$ は時刻 t と t' との差分画像とする。

3.5. 図形融合

連続差分画像を用いて求められた移動領域図形には、まだノイズ、体の微妙な動きも多く含まれる。その除去のために収縮・拡散の図形融合を複数回行う。そして最後に大体の図形重心を求める。

4. 時間的重なりを考慮した連続差分画像

求められた大体の図形重心から、時刻 t の時のフレームと動作部分が時間的に重なっていない2つのフレーム(t_1, t_2)を見つけ連続差分画像をとる。フレーム間距離を式(2)のように評価する。

$$d(t, t') = (x_G(t) - x_G(t'))^2 + (y_G(t) - y_G(t'))^2 \quad (2)$$

但し、 $(x_G(t), y_G(t))$: 時刻 t におけるフレームの図形重心 と
 $(x_G(t'), y_G(t'))$: 時刻 t' におけるフレームの図形重心とする。

$d(t, t')$ が最大値を取るときの t' を t_1 とし、2番目に大きいときの t' を t_2 とする。但し、 t_1 と t_2 が時間的に重なりと精度よく抽出できないので、もし、式(3)が成り立つならば、式(4)が最小となる時の t' を t_2 とする。

$$0 < \frac{tt_1 \cdot tt_2}{|tt_1| |tt_2|} < 1 \quad (3)$$

但し、 $tt_1 = (x_G(t_1) - x_G(t), y_G(t_1) - y_G(t))$ とする。
 $tt_2 = (x_G(t_2) - x_G(t), y_G(t_2) - y_G(t))$

$$d'(t') = (x'_G - x'_G(t'))^2 + (y'_G - y'_G(t'))^2 \quad (4)$$

但し、 $x'_G = \frac{x'_G(t) + x'_G(t_1)}{2}$, $y'_G = \frac{y'_G(t) + y'_G(t_1)}{2}$ とする。

次に、こうして求めた時刻 t, t_1, t_2 の画像をモザイク化し、 t と t_1, t_2 の連続差分画像を取る(式(5))。

$$f_{SD}(x, y, t) = f_D(x, y, t, t_1) \cap f_D(x, y, t, t_2) \quad (5)$$

$$0 \leq x \leq 320, 0 \leq y \leq 240$$

最後に体の微妙な動き等を消す為に図形融合を行う。

5. 詳細な移動領域抽出([1])

詳細な移動領域を求めるにはモザイク化や図形融合によりノイズの消去を行うことはできない。そこでフレーム間時間平均画像(式(6))を用いて差分画像を求める(式(7))。体のおれなどの不必要な領域は消去できないがランダムノイズは十分消去可能である。

$$f_A(x, y, t) = \frac{f(x, y, t) + f(x, y, t + \Delta t) + \dots + f(x, y, t + n \times \Delta t)}{n+1} \quad (6)$$

$$f_{AD}(x, y, t) = f_A(x, y, t) - f_A(x, y, t + \Delta t) \quad (7)$$

$$0 \leq x \leq 320, 0 \leq y \leq 240, n = 1, 2, 3, \dots, N$$

6. 詳細な移動領域図形抽出

4により詳細ではないが大まかに抽出された移動領域図形の概形と5により詳細であるが完全に移動領域図形を抽出していない移動領域の論理積を取ることにより詳細な移動領域図形抽出を行う。

7. 移動領域図形抽出実験

(1) 入力画像

デジタルビデオカメラにより0.2秒毎に1フレームの割合で撮影した画像を入力画像系列として用いた。
 (2) 実験結果

実験結果を図2に示す。結果を見ると従来手法では精度よく抽出できなかったフレーム(主に動作部分の動きが遅く、時間的に動作部分が重なってしまうフレーム)において、時間的重なりを考慮する事により精度よく抽出できた。

8. 終わりに

本研究ではデータグローブ、カラー手袋等の器具を用いずに動画像から動作部分(移動領域図形)を抽出する手法を提案した。その際に、時間的重なりを考慮した連続差分画像を用いて詳細な移動領域図形を抽出する手法により、従来では抽出できなかったフレームに対しても精度よく抽出できる事を示した。

今後はエッジ情報等を利用して動作部分の内部形状を抽出する手法について検討して行く予定である。



図2 実験結果

【参考文献】

- [1] 国米祐司: "動画像動作認識の為の移動領域図形抽出-フレーム間時間平均によるノイズ削減-", 1997年度精密工学会春季大会学術講演会論文集
- [2] 谷内田正彦: "コンピュータビジョン", 丸善株式会社, pp.167-190 (1990)