

テンソル分解に基づく多視点映像を用いた人物の動作認識の一検討

A Study on Human Motion Recognition Based on Tensor Decomposition using Multiple Video Cameras

堀 隆之[†] 大谷 淳[†] 糊沢 順[‡]
Takayuki HORI[†] Jun OHYA[†] Jun KURUMISAWA[‡]

1. はじめに

画像処理技術において動画像から対象の動作を認識することは重要な研究テーマである。店舗や人の往来の激しい場所には防犯などの目的で監視カメラが設置され、画像処理を用いた人物の行動認識技術の発展は監視システムの自動化やシーンの定常性評価などへの応用が期待されている。人物の行動などの動画像には非常に有益な情報が含まれているが、画像処理技術を用いてそれらを自動的に処理するためには、人物の動作に伴う姿勢変化や身体方位の問題、オクルージョンの問題など、基本的で重要な課題が残されている。また画像処理に基づく人物の動作認識において、人物の身体方位が変わると人物の動作の見え方が変わることから、動作を安定的に認識するためには被写体の身体方位に依存しない動作認識手法が求められる。

Vasilescu ら[1]によると、人物の動作には特有の個性が存在し、それを利用することで個人や動作を特定することが可能である。Vasilescu らの取り組みはベクトル空間で扱われてきた人物の動作認識を多次元テンソルに拡張し、動作の個性を求めることで動作を認識する取り組みである。彼らは HOSVD により多次元テンソルにおける人物の動作の個性を求めている。また彼らはその手法を応用し顔認識における顔の方位の問題と光源の変化に起因する照明変化の問題を解決した[2,3]。そこではテンソル空間に人物、方位、照明の各次元を配置することで、顔の方位と照明変化に頑健な顔認識手法を提案している。

我々はこれまで、Vasilescu らのテンソル分解に基づく手法を用いた動作認識に取り組んできた。これまでの取り組みではテンソル空間に人物、動作、時系列の各次元を設定した動作認識である。ただしそこで扱われた内容は被写体の身体方位が一定であり、身体方位について制約があった。ところで Vasilescu らは顔認識において、顔の方位や光源の変化に関する問題に対し、テンソル空間に方位、照明環境の次元を与えることで顔の向きや光源に関わる顔認識の課題を解決している。そこではテンソル空間に顔認識に関わる変数について、各次元を与えている。

Vasilescu らが顔認識において取り組まれたように、動作認識において、テンソル空間に身体方位の次元を設定することで、人物の動作認識における身体方位の変化に対応可能な動作認識手法が考えられる。本研究では Vasilescu らのテンソル分解に基づく動作認識、ならびに顔認識の手法を参考に、人物の身体方位の変化に対応可能な動作認識手

法について検討する。本手法ではテンソル空間に身体方位の次元を用意し、人物の身体方位の自由度を獲得することで、被写体とカメラとの位置関係に依存しない動作認識手法について検討する。そして身体方位のバリエーションを有する人物動作の動画像について動作の認識実験を行い、その有効性を確認する。

2. テンソル分解に基づく動作認識

2.1 手法の概要

Vasilescu の取り組みはベクトル空間で扱われてきた人物の動作情報を多次元テンソル空間へ拡張したものであり、多次元テンソル空間におけるテンソル分解[4]によって求められる。Vasilescu の手法は N-mode SVD によって、人の動作からその個性を算出した。本研究は Vasilescu の手法を用いて画像処理に基づく動画像から動作の認識を行う。我々がこれまで取り組んできた手法では、テンソル空間は人物・動作・時間についての次元を有している。本研究では人物の動作について、身体方位が自由度を有する場合について取り組むため、テンソル空間に方位の次元を追加することで、身体方位の自由度に対応するテンソル空間を設計する。すなわちテンソル空間には人物・動作・時間・方位の次元を有している。ここで $P \times A \times T \times V$ (P は人物, A は動作の種類, T は時系列, V は方位) のスケールを持つテンソル \mathcal{D} を定義すると、コアテンソル \mathcal{Z} は式(1)によって与えられる

$$\mathcal{D} = \mathcal{Z} \times_1 \mathbf{P}_1 \times_2 \mathbf{A}_2 \times_3 \mathbf{T}_3 \times_4 \mathbf{V}_4 \quad (1)$$

2.2 動作認識アルゴリズム

認識過程において、人物の動作について定義したテンソル \mathcal{D} がデータベース内の既存のデータと最も類似しているものを動作の認識結果とする。本研究では動作を識別するための評価方法として、コアテンソル \mathcal{Z} を用いる。

本手法において、入力されるデータのスケールに対してデータベースのスケールが異なるため、人物の動作認識において入力データとデータベースを単純に比較することはできない。そこでデータベース内の一人と入力データを入れ替えることで、データベースとサイズの等しいテンソル \mathcal{D} を算出する。この作業は人物、動作、方位の各次元について行い、それらの組み合わせの数だけ繰り返される。

ここで元のデータと入力データが似ている場合、互いの値を入れ替えてもコアテンソルの値は類似する。つまり元のデータベース内の各動作に関するコアテンソルの値と入力データによって新しく算出されたコアテンソルの差が最も小さくなる場合を探し出し、その結果を人物の動作の判定に利用する。

[†] 早稲田大学大学院国際情報通信研究科
Graduate School of Global Information and
Telecommunication Studies, Waseda University
[‡] 千葉商科大学政策情報学部
Dept. of Policy Informatics, Chiba University of Commerce

3. 多視点映像を用いた動作認識実験

3.1 実験概要

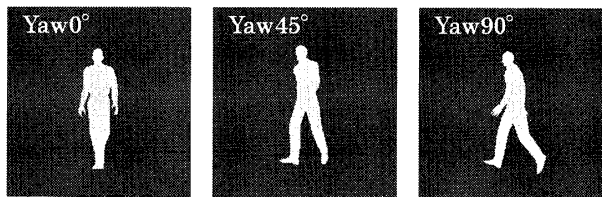


図 1(a). 実験で用いた動画像の方位の種類

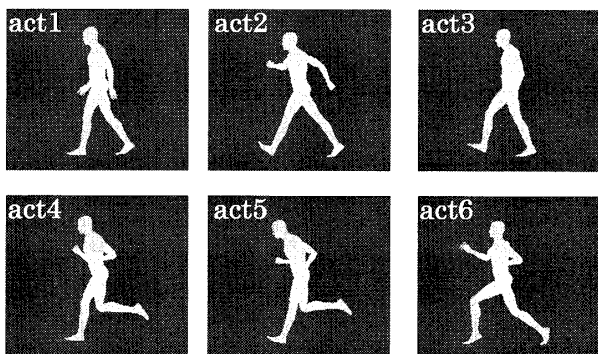


図 1(b). 実験で用いた動画像動作の種類

動作認識実験では3つの方位から撮影された6種類の動作クラスを識別する。認識実験では各動作について各方位について動作の認識率を確認する。動作認識実験ではCGソフト POSER で作成した人物のCGモデルを使用した。人物の動作について3方向から撮影する(図1(a))。ここで扱う動作は歩く(act1), 強い調子で歩く(act 2), 腕を振らず歩く(act 3), 走る(act 4), 早く走る(act 5), 不審な歩き(act 6)の計6種類の動作クラスである。また動作周期は3周期分の動作を30フレームの動画像として扱う(図1(b))。

本研究では動画像から人物の動作を扱うことから、ここではシルエット画像から画像特徴をベクトル化しテンソル空間で扱う。画像特徴の記述方法として、ここでは Mesh 特徴量を用いた。Mesh 特徴量とは人物のシルエット画像を囲む領域に対して、水平・垂直方向に小領域を作成し、その領域内に占めるシルエット画像の占有率を利用した特徴量である。ここでは人物のシルエットを囲むバウンディングボックスを関心領域とし、水平方向に5分割、垂直方向に6分割した計30の小領域について、シルエット画像の占有率を求めている(図2)。

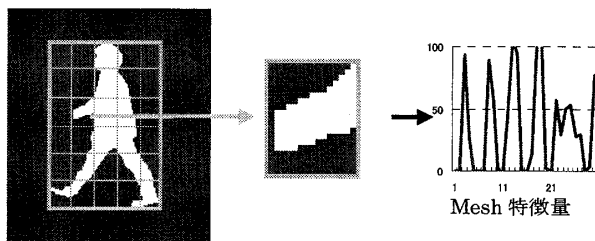


図 2. シルエット画像と Mesh 特徴量

3.2 実験結果と考察

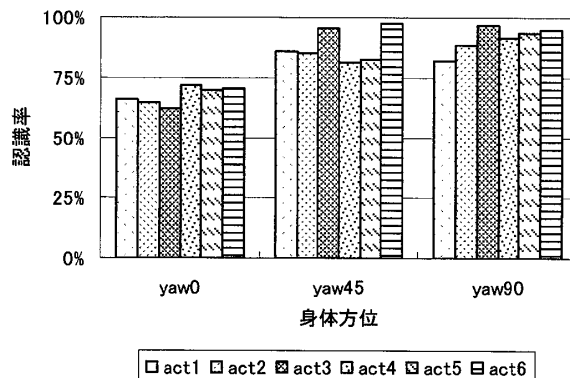


図 3. 各認識手法による認識率の比較

6種類の動作について、3つの方位ごとにみた各動作の認識結果を図3に示す。提案手法を用いた動作認識において、身体方位ごとにみた各動作の認識率は70%を上回っており、いずれの場合でも安定的に動作を識別可能であることが確認できる。また身体方位が0°, 45°, 90°の場合を比べると、yaw0°の場合が最も認識率が悪く、yaw45°, yaw90°と視点の角度が広がるにつれて認識率が向上していることが確認できる。これは本研究で扱った歩行系動作について、人物を撮影する視点が人物の身体の側面であるほど高い認識率で動作を識別可能であることが確認できる。

4. まとめ

本研究ではテンソル空間に人物の身体方位の次元を用意し、人物の身体方位の自由度を獲得することで、被写体とカメラとの位置関係に依存しない動作認識手法について検討した。テンソル空間に人物の動作に関わる次元として、人物、動作、時系列、身体方位の各次元を設定することで、人物の身体がいずれの方向に向いている場合でも動作を安定的に認識することが可能であることを示した。ここで扱った身体方位は3方向である。実際に映像に写る人物の動作を観察すると、歩きながら向きを変えるシーンが確認できる。つまり曲がり角などで歩きながら向きを変えるような、動作に伴い身体方位が変化する場合は考えられる。今後は本手法を発展させ、動作に伴う身体方位の変化について取り組む予定である。また身体方位の自由度について、今回扱ったものは水平方向(Yaw方向)に限られる。今後は垂直方向(Pitch方向)の自由度について扱い、上方から見下ろす視点からの動作認識について取り組む予定である。

参考文献

- [1] M.A.O. Vasilescu, "Human Motion Signatures: Analysis, Synthesis, Recognition", International Conference on Pattern Recognition (ICPR'02).
- [2] M.A.O. Vasilescu, "Multilinear image analysis for facial recognition", International Conference on Pattern Recognition (ICPR'02).
- [3] M.A.O. Vasilescu, D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces", In Proc. European Conf. on Computer Vision (ECCV 2002).
- [4] T. G. Kolda, "Orthogonal tensor decompositions", SIAM J. on Matrix Analysis and Applications, 23(1):243-255, 2001.