

# 4次元データに対する高次局所自己相関特徴を用いた 3次元動画モーション認識

広瀬 大† 森 裕紀† 浅田 稔†

† 大阪大学大学院 知能・機能創成工学専攻 〒565-0871 大阪府吹田市山田丘 2-1

E-mail: †{dai.hirose,hiroki,asada}@ams.eng.osaka-u.ac.jp

あらまし 高次局所自己相関特徴 (Higher-order Local Auto-Correlation) をボクセル時系列データ (4次元配列) 形式の3次元動画に適用することによって, 3次元動画を用いた認識対象を限定しないモーション認識を実現する.  
キーワード モーション認識, 高次局所自己相関特徴

## 1. はじめに

モーション認識技術は, ジェスチャ認識や不審者の検出などのあらゆるアプリケーションが考えられるためその社会的貢献は大きい. 特にステレオビジョンやレーザーレンジファインダなどによる3次元空間の情報を用いたモーション認識には, 立体情報を用いない通常のカメラによるモーション認識に比べより現実に忠実な情報が得られるため, 高い認識能力が期待される. 本研究ではHLAC[1]に注目し, ボクセル時系列データに対する特徴抽出手法を提案する.

HLAC(Higher-order Local Auto-Correlation:高次局所自己相関)[1]は画像などの配列状データに適用可能な, 位置不変性・加法性等の優れた性質を持つ特徴量である. 画像データのみならず, 過去にボクセルデータ  $(x, y, z)$  や動画像  $(x, y, t)$  などの3次元配列データの特徴量を計算するため Cubic HLAC(CHLAC)が提案されており [2], 人の歩容認識で優れた性質をもつことが実証されている.

そこで本研究では, 対象のモデルが必要なく対象を限定しない3次元情報を用いたモーション認識を行うため, ボクセル時系列データ (4次元配列データ) の特徴量抽出の手法として4次元データのためのHLAC(以下, 4D HLAC)を提案し, その有用性を実験によって示す.

## 2. 4D HLAC

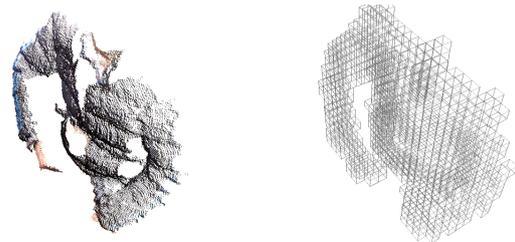
関数  $I(\mathbf{r})$  の  $N$  次の HLAC 特徴量  $R_N$  の一つは次式で表される.

$$x_N(\mathbf{a}_1, \dots, \mathbf{a}_N) = \int f(\mathbf{r})f(\mathbf{r} + \mathbf{a}_1) \dots f(\mathbf{r} + \mathbf{a}_N) d\mathbf{r} \quad (1)$$

ここで,  $\mathbf{r}$  は位置ベクトル,  $\mathbf{a}_n$  ( $n = 1, \dots, N$ ) は変位ベクトルである.  $R_N$  は  $\mathbf{a}_1, \dots, \mathbf{a}_N$  の組み合わせをかえることにより異なる複数の値が得られるので, それらを並べることで特徴ベクトルを構成する. 一般的に  $N = 0, 1, 2$ ,  $\mathbf{a}_i = -1, 0, 1 (i = x, y)$  とすることが多く, その条件下でのマスク数を表1に示す. HLACは積分特徴であるため, 位置不変性や加法性などの性質がある. 特に位置不変性は3次元情報を用いる場合, 奥行き方向にも成り立つ. 本研究では3次元動画をボクセル形式に変換 (図1, 図2) することで, HLACを4次元配列状の関数  $I(x, y, z, t)$  として3次元動画に適用し, 3次元動画

表1 マスク数

Type	Gray scale data	Binary data
HLAC	35	25
CHLAC	279	251
4D HLAC	2563	2481



(a) 点座標データ

(b) ボクセルデータ

図1 3次元動画の形式

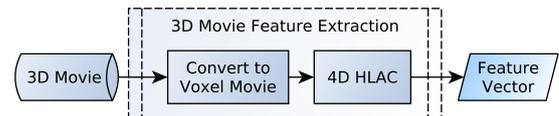


図2 3次元動画の特徴抽出プロセス

の特徴量抽出およびパターン認識を行う.

## 3. 評価実験

腕を「前方向に回す (Forward)」、「後ろ方向に回す (Backward)」、「上下させる (Up Down)」という3つの動作 (図3) を正面から採取したデータを用いた性能評価実験を行った. この3つの動作は通常の色動画では判別が困難であるため, 3次元情報を用いた認識の確認に有効である. 実験用データとして, Microsoft社のKinectを用い10人の被験者から合計41分40秒 (75000フレーム) の動画を採取した.

### 3.1 識別実験

3動作を識別する実験を行った. 判別は, フィッシャーの線形判別による次元削減の後に, 各クラス中心からのユークリッド距離を比較することで行った. 評価は採取した  $P$  人分のデータにおいて  $(P-1)$  人分を学習データ

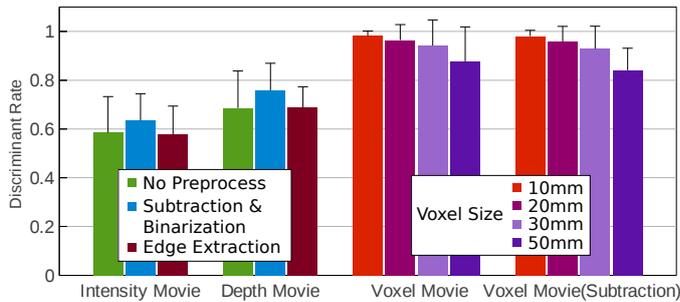


図 5 判別実験結果

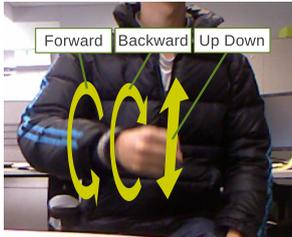


図 3 認識対象の 3 つの動作



図 6 動作数推定実験環境

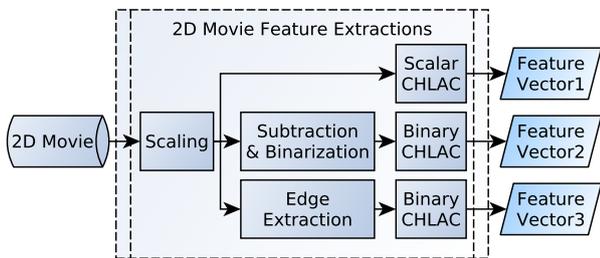


図 4 比較対象に用いた特徴抽出プロセス

として用い残りの 1 人分を検証データとして使用した場合の正答率の平均を識別率として用いた。

比較対象として輝度動画および深度動画を用いた識別を行う。これらは、両者とも 3 次元の配列データとして取り扱うことが可能であるため、従来手法である CHLAC による特徴抽出が可能である。特徴抽出プロセスには図 4 を用い、他のプロセス・条件は等しくする。正面からの輝度動画では判別し難い 3 つの動作を取り上げたため、輝度動画による識別の識別率は低いことが予想されるが、深度動画は 3 次元情報を含むため、深度動画による識別の識別率は高くなると考えられる。

比較結果を図 5 示す。本手法による識別率は 98.2% となった。一方、深度動画による識別率は 75.8% と提案手法に比べ低くなった。その理由の一つとして、深度画像よりボクセル形式の方がデータ構造として現実に忠実であり、3 次元形状の特徴を抽出する形式としてふさわしいことが考えられる。

### 3.2 動作数推定実験

4D HLAC は HLAC と同様に、位置不変性や加法性の性質を有している。そのため、 $N$  種類の複数の対象が含まれた動画全体の特徴量  $x$  は、個々の種類の対象動作の一つあたりの特徴量  $m_i$  とそれぞれの動作対象の数  $a_i$  により

$$x \approx a_1 m_1 + a_2 m_2 + \dots + a_N m_N = \sum_{i=1}^N (a_i m_i) \quad (2)$$

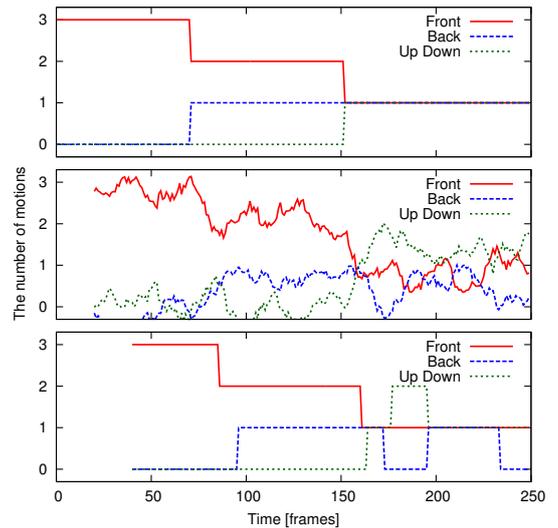


図 7 動作数推定結果 (上: 正答値, 中: 推定結果, 下: 推定結果の単純移動平均および整数化)

と表される。 $m_i$  は学習データから既知とすると、対象領域を区切ることなく動画全体の特徴量  $x$  から  $a_i$  を推定することが可能であると考えられる。ここでは、図 6 のように、腕の 3 つの動作 ( $N = 3$ ) を 3 人が同時に行う状況で各動作を行う人数  $a_i$  を推定する実験を行った。

動作数  $a_i$  を求めるために、表 1 に示されるマスク数と等しい高次元の特徴ベクトル  $m_i, x$  を、フィッシャーの線形判別による  $N$  次元まで削減する写像により  $m'_i, x'$  に変換した後、

$$(a_1 \ a_2 \ \dots \ a_N)^T = (m'_1 \ m'_2 \ \dots \ m'_N)^{-1} x' \quad (3)$$

で各動作数  $a_i$  を求めた。なお、フィッシャーの線形判別は  $(N-1)$  次元に削減する手法であるため、上式で解くために零ベクトルを新たなクラスとして加えることで、次元を一つ増加させ、 $N$  と等しくしている。

図 7 に、正答値、推定値  $a_i$  および推定値の移動平均の後離散化した結果を示す。この結果より、推定結果には多くのノイズ成分が含まれているが、およその数は推定できることが示された。

## 4. おわりに

本研究では、3 次元時系列データによるパターン認識を行うために高次局所自己相関特徴を 4 次元データに適用し、それを用いたパターン認識システムを構築し、実験によってその有用性を確認した。

## 文 献

- [1] N.Otsu and T.Kurita, "A new scheme for practical flexible and intelligent vision systems," Proc. IAPR Workshop on Computer Vision, pp.431 - 435, Tokyo, Japan, Oct.1988.
- [2] T.Kobayashi and N.Otsu, "Action and simultaneous multiple-person identification using cubic higher-order local auto-correlation," Proc. 17th ICPR, pp.741-744, Cambridge, England, UK, Aug.2004.