

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月 8日現在

機関番号：82626

研究種目：若手研究（B）

研究期間：2009～2011

課題番号：21700214

研究課題名（和文） 3次元時空間データからの統計的特徴抽出に関する研究

研究課題名（英文） A study on statistical feature extraction from three-dimensional spatio-temporal data

研究代表者

小林 匠（KOBAYASHI TAKUMI）

独立行政法人産業技術総合研究所・情報技術研究部門・研究員

研究者番号：30443188

研究成果の概要（和文）：時空間データから認識に有効な特徴を抽出する手法を提案した。提案法では、時空間における局所近傍でのデータ記述子の共起性に注目し、様々な共起パターンの統計量を算出することで、対象データの特徴を抽出する。データ記述子としては、動画像では時空間勾配ベクトル、またより一般的な記述子として局所特徴量を採用することで効果的な動画像特徴や画像特徴が得られる。動作認識や画像認識の実験により、提案法の有効性を確認した。

研究成果の概要（英文）：In this study, we have proposed novel methods to extract features from spatio-temporal data for motion/image classifications. Those methods are based on co-occurrences of primitive descriptors in local neighborhoods in the spatio-temporal domain. By employing as the descriptors the space-time gradients and local feature vectors, we obtain the effective features to describe motions and images. In the experiments on motion/image classifications, the proposed methods exhibited favorable performances.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,200,000	360,000	1,560,000
2010年度	700,000	210,000	910,000
2011年度	600,000	180,000	780,000
年度			
年度			
総計	2,500,000	750,000	3,250,000

研究分野：パターン認識

科研費の分科・細目：情報学 知覚情報処理・知能ロボティクス（A）

キーワード：特徴抽出・時空間データ・動画像認識・画像認識

1. 研究開始当初の背景

(1) 記憶装置や情報基盤の発展に伴い大容量の情報が蓄積され、それら膨大な情報へのアクセスも容易となった。そのコンテンツも画像にとどまらず、動画像や3D-CAD、ポリ

ュームデータなどへと高次元（3次元）化しつつある。

(2) しかしながら、それらの3次元データから必要な情報を人手で抽出するには多大な

労力を必要とする。そのため、多量の3次元情報を効率的に処理できるよう、認識処理の自動化技術(自動認識)が望まれている。

(3) 自動認識を実現するためには、まず3次元データからその内容を示す特徴的な情報を抽出する、特徴抽出技術が必要となる。特徴抽出は自動認識において第一に行われる処理であるため、その性能が最終的な認識性能を大きく左右する重要な技術である。

2. 研究の目的

(1) 画像や動画などの時空間データから、認識に有効な特徴を抽出する手法を開発する。

(2) データを構成する記述子の局所近傍での共起性に注目し、確率・統計的な観点から特徴を抽出する。これにより、対象データをより詳細に特徴づけることが可能となる。

(3) 記述子として、動画像に対しては時空間勾配を用いる。動画像内での対象動作は、時空間での動作曲面として見なされ、時空間勾配によりその曲面の幾何学的特徴が効果的に抽出される。

(4) また、局所特徴量を記述子とするより一般的な共起特徴抽出の枠組みを提案する。

3. 研究の方法

(1) まず、提案法の一般的枠組みとして、データ記述子の局所近傍における共起特徴を次式により定義する。

$$f(\{a_i\}_{i=1}^N) = \int_{r \in D} w[m_r, \{m_{r+a_i}\}_{i=1}^N] \mathbf{h}_r \otimes_{i=1}^N \mathbf{h}_{r+a_i} dr \quad (式1)$$

ここで、 \mathbf{r} は時空間 D 内での位置ベクトル、 \mathbf{a}_i は i 番目の変位ベクトル、 \mathbf{h} は記述子のベクトル表現、 m は記述子の重み、 w は局所近傍記述子ペアの重みを決定する関数である。 \otimes は各ベクトル \mathbf{h} のテンソル積を計算する演算子である。局所近傍への変位ベクトル \mathbf{a}_i は隣り合うピクセル (ボクセル) 等として単純に決定される。一方、ここでは 1) 記述子 \mathbf{h} の表現、2) 重み関数 w の決定が重要となる。これらは対象とする時空間データの種類により様々なものが選択し得る。

(2) 動画像においては、記述子として時空間勾配に注目する。

① 時空間勾配ベクトルは動画像ピクセル値 I に対して、 xyt 軸に沿った微分値から成る 3次元ベクトル

$$\mathbf{g} = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t} \right]^T$$

として表現される。ここでは、勾配をベクトル \mathbf{g} そのものではなく、その (3次元) 方向 (θ, ϕ) と大きさ m に分解して表現する。

$$m = \sqrt{\frac{\partial I^2}{\partial x} + \frac{\partial I^2}{\partial y} + \frac{\partial I^2}{\partial t}}, \quad (式2)$$

$$\theta = \arctan\left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right), \quad \phi = \arcsin\left(\frac{\partial I}{\partial t} / m\right).$$

これらは図1に示すように、勾配を構成する幾何学量となっている。

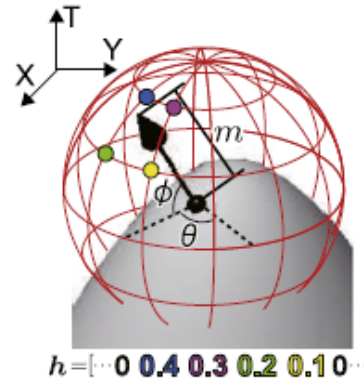


図1：時空間勾配

② 次に、勾配方向 (θ, ϕ) を量子化する。角度 (θ, ϕ) は球面上での方向を表す経度・緯度に相当しており、球面を図1のように分割 (赤線の交点が量子化点) することで、勾配方向を量子化する。ここでは、最近傍の量子方向のみではなく、4近傍の量子方向で勾配を線形表現することに注意されたい。これにより量子化誤差を軽減することができる。図1に示すように、各量子方向の重みを要素とする勾配方向の記述子ベクトル \mathbf{h} が得られる。

③ 以上①②により得られた、記述子 \mathbf{h} 及び重み m (式2) を用いて (式1) により時空間勾配共起特徴=動作特徴を算出する。なお重み関数はノイズの影響を極力軽減できるように $w(\cdot) \equiv \min(\cdot)$ と設定する。また、近傍数は $N=1$ として、変位ベクトル \mathbf{a}_i は図2に示すような時空間での13近傍を採用する。近傍としては、これらの対称パターンも存在しているが、それらは (式1) の積分により重複するため、除外している。

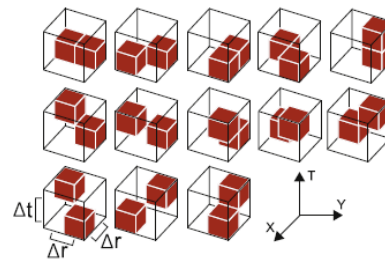


図2：変位ベクトル

④ 上述の時空間勾配共起特徴は、(式1)の $\mathbf{h} \otimes \mathbf{h}$ により時空間勾配方向の近傍共起性という統計量を抽出しているだけでなく、下記のような幾何学的特性も反映している。

動画はフレームが時系列に並んだデータとして得られるが、一方で時空間 xyt での3次元ボリュームデータとも見なせ、そこでは動作は曲面を成す(図3)。

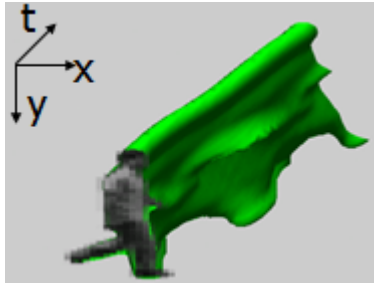


図3：時空間での動作曲面

勾配共起特徴は、この動作曲面の局所曲率(や勾配)をパターン化して特徴づけていることに相当する。曲率や勾配は曲面を構成する重要な幾何学量であるため、曲面として表現される動作の特徴も効果的に抽出される。定性的には、曲率は加速度、勾配は速度に対応している。

(3) 動画では基本的な幾何学量として時空間勾配に注目したが、より一般的な記述子として局所特徴量(ベクトル)を採用することも可能である。

① 近年認識手法として注目されている Bag-of-Features 法では、画像や動画から局所特徴量(SIFTやHOG)を多数抽出し、その集合として対象データを表現している。ここでは、そのような局所特徴量から記述子 \mathbf{h} を構築し、(式1)によりその共起特徴を抽出する。

② 局所特徴量は一般にベクトルとして抽出されるため、(2)②に記載した時空間勾配ベクトルの場合と同様に、局所特徴ベクトルを量子化することで記述子 \mathbf{h} を構成する。局所特徴ベクトルを k-means 法などによりクラスタリングすることで、各ベクトルは近傍クラスターを用いて(線形)量子化表現され、各量子クラスターの重みを要素とする記述子 \mathbf{h} が得られる。ここでも、最近傍の量子クラスターのみだけでなく、複数(例えば3つ)の近傍量子クラスター群により線形に量子化表現されることに注意されたい。また、このとき各局所特徴量が均等に扱われるように記述子の重みは $w=1$ と設定する。

③ 上記記述子と重みを用いて(式1)により一般的な局所特徴ベクトルに対する共起特徴が得られる。例えば、2次元画像上で、格子状に局所特徴ベクトルが抽出された場合には、図4に示すように共起特徴が算出される。

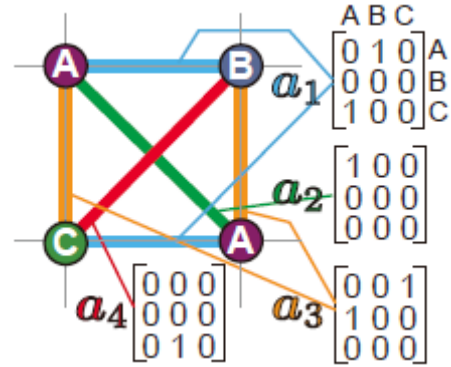


図4：局所特徴の共起

ここでは、局所特徴ベクトルが3つのクラスター(A, B, C)で量子化されている。

④ 従来の Bag-of-Features 法の枠組みでは、量子化された局所特徴量の生起ヒストグラムを対象データの特徴としていたが、ここではより高次の統計的特徴として共起特徴が得られる。そのため、局所特徴量が対象物を構成する各パーツを表現する場合には、それらパーツの関係性も共起性により特徴づけることができる。これにより、判別力の高い特徴が得られる。

⑤ 局所特徴量を用いた共起特徴を算出する一般的な枠組みを示した。ここでは、任意の局所特徴量を採用することができる。そのため、拡張手法として、ここで算出した「局所特徴量の共起特徴」をさらに局所特徴量と考えることで、階層的に共起特徴を抽出することもできる。

4. 研究成果

動画認識や画像認識の実験により、提案法の有効性を確認した。

(1) 動画特徴として時空間勾配共起特徴[3(2)]を用いて、動作認識を行った。

①まず、時空間勾配共起特徴の特性について解析する。時空間勾配は、“静止”物体に対しては時間方向の立ち上がりがなく($\phi=0$)、“動き”の大きい物体に対してはその立ち上がりが大きくなる($|\phi| \gg 0$)。このように、“静止”から“動き”情報までを一貫して扱っているが、動作認識においては、“静止”情報よりも“動き”情報に注目することが重

要であると考えられる。そこで、時空間勾配の量子方向を、“静止”方向と“動き”方向に分け、それぞれで構成した共起特徴の動作認識性能を比較した(表1)。

表1：共起特徴の解析 (KTH dataset)

“動き”x“動き”	○	○	○
“動き”x“静止”	×	○	○
“静止”x“静止”	×	×	○
認識率 (%)	92.8	95.6	94.6

共起特徴では“動き”と“静止”に対して3パターンの組み合わせがある。表1から、「動きx動き」と「動きx静止」パターンからなる共起特徴が最も性能が高いことが分かる。これは、“静止”のみからなる特徴は動作認識には寄与しないことを示している。つまり、動作認識の場合には“動き”に関連した共起特徴だけで十分である。

② 次に、従来法との性能比較を表2に示す。

表2：動作認識結果 (KTH dataset)

手法	認識率 (%)
Laptev et al. (2008)	91.8
Wang et al. (2009)	92.1
Ballan et al. (2009)	92.1
時空間勾配共起特徴	95.6

従来法に比べ、高い認識性能が得られていることが分かる。これにより、提案法が動作認識に非常に有効な特徴を抽出していることが示された。

(2) 画像特徴として局所特徴量の共起特徴[3(3)]を用いて画像認識を行った。

① まず、局所特徴量を用いた従来法としてBag-of-Features法との比較を行う(表3)。

表3：Bag-of-Features法との比較 (Caltech101 dataset)

手法	認識率 (%)
Bag-of-features	41.8
共起特徴	56.6
階層的共起特徴	59.8

Bag-of-Features法では量子化した局所特徴量の生起ヒストグラムを画像特徴としている。提案法では共起性に注目することで、その性能が大きく向上していることが分かる。さらに、「局所特徴量の共起特徴」を局所特徴量として階層的に共起特徴に抽出する[3(3)⑤]ことで、性能はさらに向上する。

② 他の従来法との比較を表4に示す。

表4：画像認識結果 (Caltech101 dataset)

手法	認識率 (%)
Zhang et al. (2009)	52.0
Lazebnik et al. (2009)	56.4
Zhang et al (2006)	59.1
階層的共起特徴	59.8

従来法と比べても良好な認識性能を示している。これにより、画像認識においても、共起特徴が有効であることが示された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

- ① 小林匠、大津展之、Motion recognition using local auto-correlation of space-time gradients、Pattern Recognition Letters、査読有、Vol. 33、No. 9、2012、pp. 1188-1195、DOI: 10.1016/j.patrec.2012.01.007

[学会発表] (計2件)

- ① 小林匠、大津展之、Bag of Hierarchical Co-occurrence Features for Image Classification、International Conference on Pattern Recognition (ICPR)、2010年8月26日、Istanbul, Turkey
- ② 小林匠、動画像による人物動作認識、第10回計測自動制御学会システムインテグレーション部門講演会、2009年12月25日、芝浦工業大学(東京都)

6. 研究組織

(1) 研究代表者

小林 匠 (KOBAYAHSI TAKUMI)

独立行政法人産業技術総合研究所・情報技術研究部門・研究員

研究者番号：30443188