

## スポーツ記事の自動判別と類似シーンの検索

6 G-7

渋谷 亜紀 杉山 善明 有木 康雄

龍谷大学 理工学部 電子情報学科

## 1 はじめに

ニュース映像をデータベース化する場合、膨大な記事の量を扱うことから、計算機によって記事の内容を自動分類できることが望まれる。本研究では、画像認識の観点からスポーツ記事を対象として、記事内容を自動分類した結果を示し、さらに類似シーンを検索する方法と結果について示す。

## 2 部分空間法を用いたシーンの判別

## 2.1 部分空間法

部分空間法は、各カテゴリ毎にそのカテゴリの特徴をよく表現する部分空間を、学習データよりあらかじめ設計しておく。入力ベクトルが与えられると、各カテゴリの部分空間への射影量を求め、この射影量が最大となるカテゴリに入力ベクトルを判定する方法である<sup>(1)</sup>。本研究では、この部分空間法 (CLAFFIC 法) を用いているが、判別精度を上げるため、スポーツ名 (カテゴリ) を判定する際に、スポーツ映像の細分化 (クラスタ) を行ない、1つのカテゴリに複数の部分空間を設定する多重部分空間法を用いている。

## 2.2 多重部分空間法

従来、部分空間法では各カテゴリに1つの部分空間を設定していた。本研究では、1カテゴリあたりの部分空間を複数個に設定することにより、判別能力を上げる。1つのカテゴリに複数個の部分空間を設定するという意味で、この手法を多重部分空間法と呼んでいる<sup>(2)</sup>。射影行列の計算方法などは部分空間法と同様である。部分空間法との違いは、もとの1つの部分空間を複数の小さな部分空間に分割する点にある。このため1カテゴリの学習データを、複数のクラスタに分割しておく必要がある。この方法として、学習データにベクトル量子化を施してクラ

スタに分割し、各クラスタ毎に部分空間を作り多重部分空間を構成する。分割されたクラスタ毎に部分空間法と同様、射影行列を求めておく。判別の際には、入力ベクトルをすべての部分空間に射影し、射影量の最も大きい部分空間を含むカテゴリに入力ベクトルを分類する。

## 3 判別手法

## 3.1 前処理

本研究では、ニュース映像から自動的にシーンの変わり目 (カット点) を検出し、ニュース記事の論理構造に基づいて記事の切り出しを行う手法を用いている<sup>(3)</sup>。この手法によって取り出されたカット点のフレーム画像を、スポーツシーンの学習および認識に用いた。

## 3.2 特徴量の抽出

特徴量は各フレーム画像の DCT 成分を用いる。フレーム画像を  $4 \times 4$  のブロックに分割し、各ブロックにおいて DCT を行ない、直流、第 1(2 個)・第 2(3 個) 交流成分を計算する。これとは別に画像全体を 1 ブロックとして DCT を行ない、直流、第 1(2 個)・第 2(3 個) 交流成分を計算する。合計 102 ( $17 \times 6$ ) 次元の入力パターンが得られる。

## 3.3 判別と評価方法

入力パターンを多重部分空間法で判別する。判別の結果は認識率と適合率で評価する。認識率は判別誤りの少なさを、適合率は判別の正確さを表し、次式で定義される。

$$\text{認識率} = \frac{\text{正しく認識されたデータ数}}{\text{実際に存在するデータ数}} \quad (1)$$

$$\text{適合率} = \frac{\text{正しく認識されたデータ数}}{\text{認識されたデータ総数}} \quad (2)$$

## 4 実験結果

## 4.1 実験データ

スポーツ・ニュース映像は、計 8 スポーツ、98 記事、1816 カット点フレームを使用した。この内約半

Automatic Discrimination of Sports News and Retrieval of Similar Scenes

Aki Shibutani and Yoshiaki Sugiyama and Yasuo Arika  
Ryukoku University

1-5, Yokotani, Oe-cho, Sete, Otsu-shi, 520-21 Japan

分を学習に、残り半分を判別に用いた。学習・判別において放映日、チャンネルとも違う番組から収集した。

## 4.2 フレーム判別

フレーム毎にシーン判別を行って、認識率と適合率を求めた結果を表1に示す。各スポーツカテゴリの部分空間の数は8つとしている。全体の平均認識率は67.2%であった。

## 4.3 記事判別

記事毎の判別では、フレーム毎に求まっている射影量をもとに、各カテゴリ毎の記事内平均射影量を求め、この値が最大となるカテゴリに対象記事を判別する。記事判別の認識率を表1に示す。唯一、Volleyballの1記事をAmerican Footballに誤判別した。全体では、98.6%の認識率を得た。

表1: 判別結果 (%)

| スポーツ名             | フレーム判別 |      | 記事判別  |
|-------------------|--------|------|-------|
|                   | 認識率    | 適合率  | 認識率   |
| American Football | 66.0   | 63.6 | 100.0 |
| Baseball          | 75.9   | 75.9 | 100.0 |
| Golf              | 31.3   | 70.3 | 100.0 |
| Ice Hockey        | 96.6   | 62.8 | 100.0 |
| Soccer            | 82.0   | 83.3 | 100.0 |
| Sumou             | 63.3   | 74.0 | 100.0 |
| Tennis            | 70.8   | 53.1 | 100.0 |
| Volleyball        | 51.2   | 66.2 | 88.9  |
| Average           | 67.2   | 68.7 | 98.6  |

## 5 検索

### 5.1 検索方法

8スポーツのカット点フレームから特にイベントと思われるシーンを複数人に選んでもらい、そのイベントについて検索を行なった。イベントとは、たとえば野球ではピッチャーのアップのシーンや、ホームランシーンを指す。スポーツによっては、何をイベントとするか難しいが、検索したいと思うカット点フレーム55枚を選んだ。

検索対象である55カット点フレームの入力ベクトルを多重部分空間に各々射影し、距離が最も近くなる部分空間に含まれているカット点フレーム(学習データ)を検索した。検索されたフレーム複数枚に対してさらに入力ベクトルとの類似度を求めた。入力ベクトルを $x$ 、部分空間に含まれている学習データのカット点フレームの特徴ベクトルをそれぞれ $y$ とす

る。次式により $x$ と $y$ の類似度を求めた。

$$\cos \theta = \frac{x \cdot y}{\|x\| \|y\|} \quad (3)$$

検索されたカット点フレームの特徴ベクトル $y$ のうち類似度が最も高い第1候補に対して類似シーンであるかどうかを目視により判定した。更に類似度の高い第5候補までの検索結果に対しても類似シーンであるかどうかを判定した。判定方法はその枚数までの検索結果に対して1枚でも類似シーンが含まれているかどうかでその割合を求めた。

### 5.2 検索結果

第1候補のみの検索結果では20.0%の精度しか得られなかった。原因として、フレームで検索を行なったためフレーム判別率67.2%以上にはならないことと、最大射影量をもつクラスに判定しても、それが誤っていた場合、候補を増やしても結果が良くならないためである。今後は間違ったクラスに判定した場合でも検索を改善できるように、射影量が最大ではないクラスについても検索を行なう予定である。

表2: 検索結果 (%)

| 候補数    | 検索率  |
|--------|------|
| 第1候補のみ | 20.0 |
| 第3候補まで | 32.7 |
| 第5候補まで | 61.8 |

## 6 おわりに

テレビで放映されているスポーツ・ニュース映像を対象として、検出されたカット点でのフレームを代表フレームとみなし、自動的にシーン判別する手法を提案した。各スポーツのカテゴリをさらに細分化して複数のクラスで構成し、多重部分空間法を用いて記事判別を行なった。また、フレーム毎に判別するより、記事単位で判別する方が有効であることを実験的に示した。また、類似シーンの検索方法と結果を示した。

### 参考文献

- [1] エルッキ・オヤ 著, 小川 英光, 佐藤 誠 訳: “パターン認識と部分空間法”, 産業図書 (1986).
- [2] 茂木 祐治, 有木 康雄: “手書き漢字認識のための多重部分空間法について”, 電子情報通信学会秋期大会, D-304, pp.312 (1994).
- [3] 齊藤 陽子, 有木 康雄: “ニュース映像のデータベース化に向けて -ニューススタジオの映像検出と記事切り出し-”, 第149回画像電学研究会, 95-04-04, pp.13-16 (1995-04).