

局所的な選手人数分布に基づくサッカーイベント判別法

Soccer Event Classification Based on Analysis of Local Head Counts of Players

三須 俊枝[†] 苗村 昌秀[†] 藤井 真人[†] 八木 伸行[†]
 Toshie Misu Masahide Naemura Mahito Fujii Nobuyuki Yagi

1. まえがき

膨大な映像コンテンツから所望の番組やシーンを効率的に検索するためには、番組を意味的なまとまりに区分し、その内容を記述したメタデータの付与が必須となる。多様な状況やプレイ（以降、イベントと呼ぶ）が発生するスポーツ番組用メタデータ [1][2] は、映像検索のほか、サーバ型放送におけるダイジェスト視聴や、番組・オブジェクト連動データ放送のコンテンツ制作など幅広い応用も期待される。しかし、シナリオのないスポーツ映像に対しメタデータを全て手動付与するには莫大な時間、労力、コストを要することから、イベント種別の判別の自動化が望まれている。

そこで我々は、サッカー映像を対象とし、固定ハイビジョンカメラで撮影された映像からの選手の位置・速度・所属チームの推定、汎用的な低レベル特徴（局所人数と平均速度）の抽出、フィッシャー判別基準に基づく高レベル特徴への変換 [3]、ガウス混合モデルによる高レベル特徴分布表現を経てイベント種別を判別する手法を提案している [1]。

本論文では、競技における選手配置の偏在性に着目し、[1]の手法の低レベル特徴抽出において局所選手人数を計数するためのグリッドを不等間隔化することで、選手の集まる傾向のある領域を詳細に定量化することとした。以下、処理全体の概要を述べつつ、第3.1節および第4.1節で低レベル特徴抽出について詳述する。

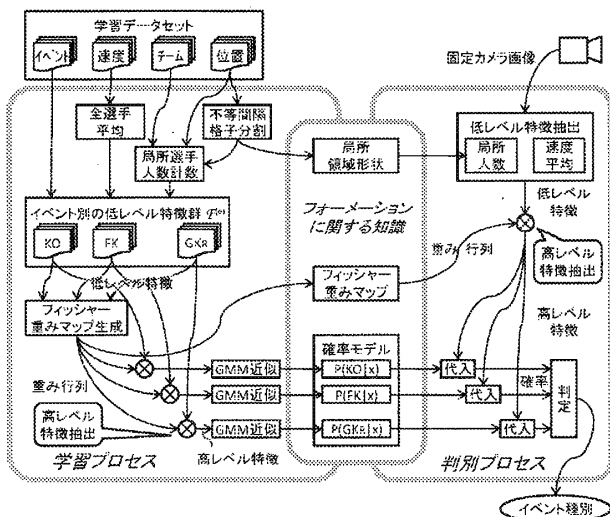


図 1: 学習および判別プロセスのブロック図。

2. 処理手順

学習および判別のプロセスを図1に示す。本手法ではサッカー競技における各種イベント（表1）のフォーメーションに関する知識を（1）サッカーコート領域の不等間隔分割、（2）フィッシャー判別基準に基づく重みづけ、（3）ガウス混合モデルによる確率密度関数のモデル化の3要素により表現する。

要素（1）では、選手の存在確率に基づきサッカーコートを格子状に不等間隔分割し、各格子に関して選手の存在人数を計数することで、選手配置を表現する。選手の存在確率が高い領域ほど細かく量子化されるため、選手配置の偏在性の強い競技においてはフォーメーションの差異を効率的に数値化することができる。これに、全選手の平均速度の情報を加えたものを、低レベル特徴量と定義する。

要素（2）は、低レベル特徴空間から、異イベント種別が分離し、同一イベント種別が集中するような高レベル特徴空間への変換のための基底群であり、フィッシャー判別基準に基づいて算出する。

要素（3）は、イベント毎の高レベル特徴の複雑な分布形状をガウス混合モデルで近似したものである。

3. 判別プロセス

図1の右半分に示した判別プロセスに関し、各段階の処理について述べる。

3.1 低レベル特徴抽出

サッカーなどのチーム競技では、戦況やイベントに応じてコート上において典型的な選手配置をとる傾向がある。本手法では、サッカーコートを直交する格子状に分割し、各格子内に存在する選手数をチーム別に集計することにより、選手配置を表現することとした。

表 1: 判別対象の12イベント種別。

イベント大分類	動作主	場所	略号 e	
キックオフ	KO	—	中央	KO
フリーキック	FK	—	—	FK
コーナーキック	CK	左チーム	右奥	CK _{LF}
			右手前	CK _{LN}
		右チーム	左奥	CK _{RF}
		左手前	CK _{RN}	
スローイン	TI	左チーム	奥	TI _{LF}
			手前	TI _{LN}
		右チーム	奥	TI _{RF}
		手前	TI _{RN}	
ゴールキック	GK	左チーム	左	GK _L
		右チーム	右	GK _R

[†]日本放送協会, NHK

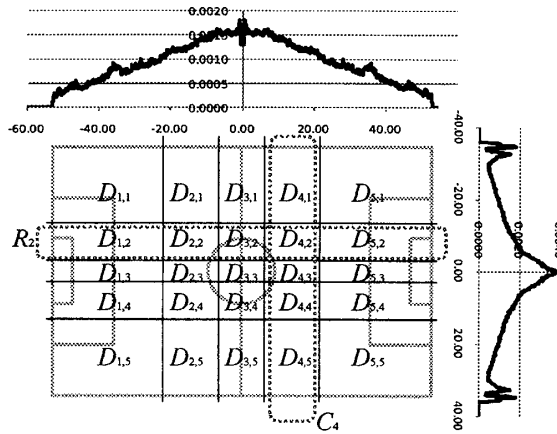


図 2: 選手の存在確率密度と格子分割の例.

ワールド座標系は、センターマークに原点をとり、ベンチ側からコートに向かって右方向にタッチラインと平行に p 軸を、ベンチ側からコートに向かって奥方向にハーフウェイライン上に q 軸をとる。コート p 軸方向および q 軸方向の分割数を L_p および L_q とし、コート左奥側から p 方向および q 方向に i 番目 ($i = 1, \dots, L_p$) および j 番目 ($j = 1, \dots, L_q$) の局所領域を $D_{i,j}$ とおく。以降、場内の局所領域 $D_{i,j}$ 内に存在する左側および右側チームの選手人数をそれぞれ $l_{i,j}$ および $r_{i,j}$ と書くこととする。

また、全体の動きの方向と速さも重要であることから、全選手に関する平均速度ベクトル $\boldsymbol{v} = [v_p, v_q]^T$ も採り入れ、低レベル特徴 \boldsymbol{f} を次式により定義した。

$$\boldsymbol{f} = [l_{1,1}, \dots, l_{L_p,1}; l_{1,2}, \dots; \dots, l_{L_p,L_q}; r_{1,1}, \dots, r_{L_p,1}; r_{1,2}, \dots; \dots, r_{L_p,L_q}; v_p, v_q]^T \quad (1)$$

3.2 高レベル特徴抽出

低レベル特徴の各成分には冗長性や相関性が含まれる。本ステップでは、低レベル特徴 \boldsymbol{f} に対して

$$\boldsymbol{x} = W_M \boldsymbol{f} \quad (2)$$

なる線形変換による次元圧縮を行い、より低次元の高レベル特徴 \boldsymbol{x} に変換する。

3.3 高レベル特徴分布の確率モデル

確率モデル $P(e | \boldsymbol{x})$ に、観測された高レベル特徴 \boldsymbol{x} を代入することにより、イベント種別 $e \in \mathcal{E}$ ($\mathcal{E} = \{K0, FK, CK_{LF}, \dots, GK_R\}$ は全イベント種別の集合) の生起する事後確率を算出する。本手法では、確率モデルを $p(\boldsymbol{x} | e)$ に関するガウス混合モデル (GMM) とイベント種別 e の事前確率 $P^{(e)}$ との積により表現する。

$$P(e | \boldsymbol{x}) \propto P^{(e)} \times \sum_{k=1}^{K^{(e)}} \alpha_k^{(e)} \frac{\exp \left\{ -\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu}_k^{(e)})^T [\boldsymbol{\Sigma}_k^{(e)}]^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_k^{(e)}) \right\}}{\sqrt{(2\pi)^{\dim(\boldsymbol{x})} |\boldsymbol{\Sigma}_k^{(e)}|}} \quad (3)$$

判別時には、 $\arg \max_{e \in \mathcal{E}} P(e | \boldsymbol{x})$ により最も確からしいイベント種別を求める。今回は、さらに判別対象フレームの前 120 フレームから後 30 フレームに至る 151 フレームにおいて多数決をとることとした。

4. 学習プロセス

図 1 の左半分の学習プロセスについて述べる。学習データセットは、全選手の位置、速度、およびチームをイベント種別 e ごとに整理したデータとする。イベント種別 e に関するデータ数を $n^{(e)}$ フレームとする。これに基づき、次節の手法により得た格子内に存在するチーム別人数を集計し、また平均速度を求めることで、低レベル特徴の学習データセット $\mathcal{F}^{(e)}$ を得る。

4.1 不等間隔格子分割

公式試合に用いられる典型的なサッカーコートは長さ 105m、奥行き 68m の大きさであるが、選手は必ずしも均等には存在せず、図 2 に示すように、とくに中央およびタッチライン付近に存在する確率が高い。偏在性の強い競技においては、選手の集中しがちな領域においてさまざまな種類のイベントが交錯するため、当該領域をより細かくサンプリングすることが望ましいと考えられる。そこで、学習データセットの選手分布を用い、以下の手順で縦横に領域分割することとした。

1. 各縦長領域 $C_p = \bigcup_{j=1}^{L_q} D_{p,j}$ 内の選手存在確率が、すべての p に関して均等となるよう p 方向に分割。
2. 同様に、各横長領域 $R_q = \bigcup_{i=1}^{L_p} D_{i,q}$ 内の選手存在確率を均等化するよう q 方向に分割。

図 2 左下のサッカーコート上に、実際の分割例を示す。中央付近が細かく分割されている。

4.2 フィッシャー重みマップ

式 (2) の重み行列 W_M による変換により、異イベント種別が互いに分離し、同イベント種別が密集するような高レベル特徴が得られることが望ましい。そこで、 W_M としてフィッシャー重みマップ [3] を採用することとした。学習データセットから得たイベント種別 e ごとの高レベル特徴群 $\mathcal{F}^{(e)}$ におけるクラス内分散およびクラス間分散をそれぞれ Σ_W および Σ_B とおく。

$$\Sigma_W = \left(\sum_{e \in \mathcal{E}} \sum_{\boldsymbol{f} \in \mathcal{F}^{(e)}} (\boldsymbol{f} - \bar{\boldsymbol{f}}^{(e)}) (\boldsymbol{f} - \bar{\boldsymbol{f}}^{(e)})^T \right) / \sum_{e \in \mathcal{E}} n^{(e)} \quad (4)$$

$$\Sigma_B = \left(\sum_{e \in \mathcal{E}} n^{(e)} (\bar{\boldsymbol{f}}^{(e)} - \bar{\boldsymbol{f}}) (\bar{\boldsymbol{f}}^{(e)} - \bar{\boldsymbol{f}})^T \right) / \sum_{e \in \mathcal{E}} n^{(e)} \quad (5)$$

なお、 $\bar{\boldsymbol{f}}^{(e)}$ はイベント種別 e に関する低レベル特徴の平均を、 $\bar{\boldsymbol{f}}$ は全イベント種別にわたる低レベル特徴の平均をそれぞれ表す。

一般化固有値問題 $\Sigma_B \boldsymbol{w} = \lambda \Sigma_W \boldsymbol{w}$ を解き、固有値 λ の大きい方から順に M 個の固有ベクトル $\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_M$ を求める。フィッシャー重みマップ W_M は、これら M 個の固有ベクトルを並べた行列として定義される。

$$W_M = [\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_M]^T \quad (6)$$

約 160 分の学習データセットにより求めたフィッシャー重みマップ W_{11} (12 クラスの判別に有効な M の

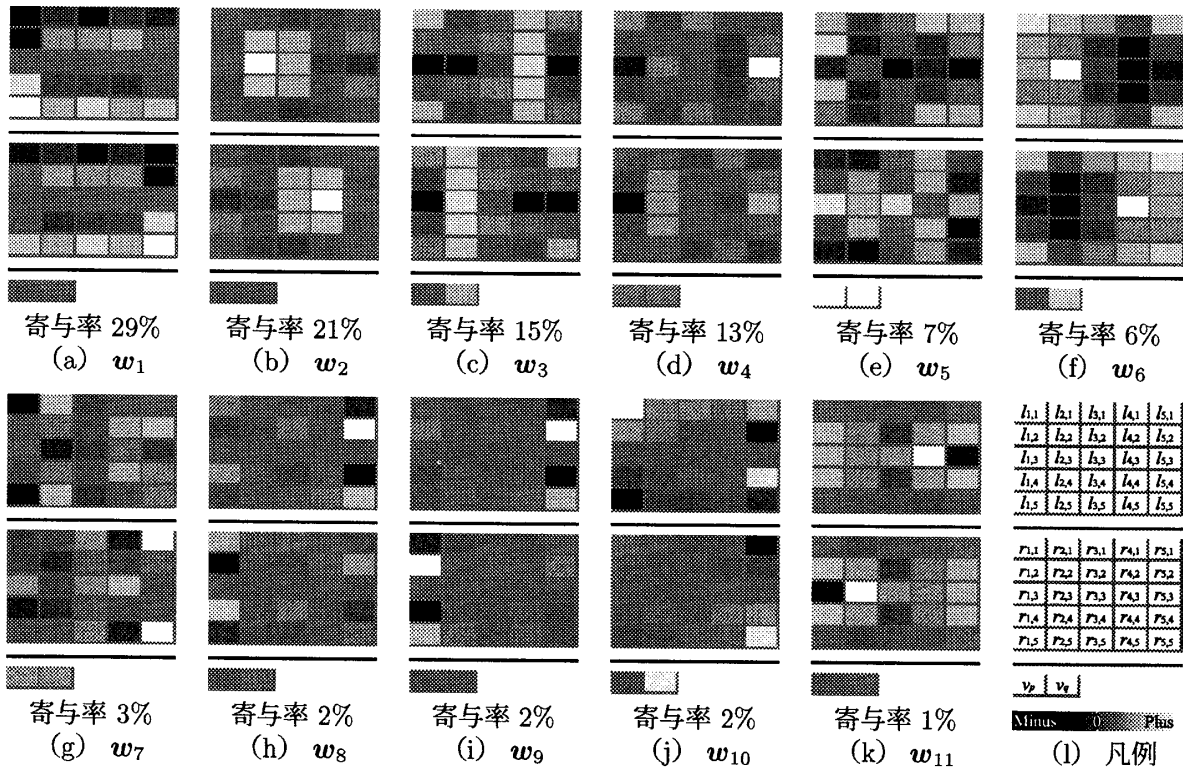


図3: フィッシャー重みマップ W_{11} の例 ($\Sigma_W^{-1} \Sigma_B$ の各固有ベクトルの各成分を輝度値で可視化).

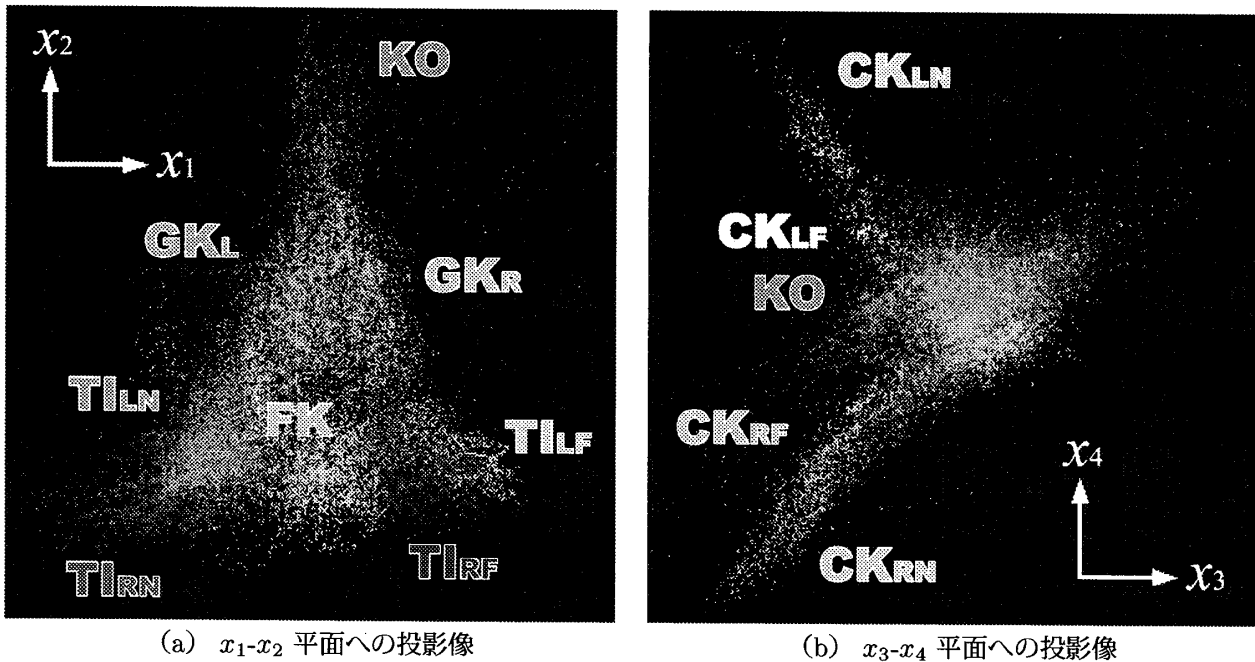


図4: イベント種別 e ごとの高レベル特徴分布 $\mathcal{X}^{(e)} = \{W_{11}f \mid f \in \mathcal{F}^{(e)}\}$.

最大値 $M = 11$ を採用) の例を図3に示す. (a) はコートタッチライン寄りに正負の重みが表れており, $TI_{LN} \leftrightarrow TI_{LF}$ といったスローインの手前/奥の区別に有効であることが示唆される. (b) はチーム別に正の重みの領域がそれぞれ分離しており, 各チーム選手が自陣内に入るキックオフの判別に有効と考えられる.

イベント種別 e ごとの高レベル特徴 f を図3のフィッシャー重みマップ W_{11} により射影した高レベル特徴 x の散布図 (2次元に投影) を図4に示す. 前述のとおり, (a) の x_1 軸方向にスローインの奥/手前が分離し, x_2 軸上方にキックオフが分離している. また, (b) の x_4 軸はコーナーキックの左右判別に効いている.

表 2: イベント判別結果のコンフュージョンマトリクスと精度.

(a) 均等間隔分割

		判別結果					再現率		F 値 (調和平均)
		KO	FK	CK	TI	GK			
真値	KO	7	3	0	0	1	64%	(7/11)	0.67
	FK	1	26	3	8	8	57%	(26/46)	0.58
	CK	0	1	10	5*1	1	59%	(10/17)	0.51
	TI	0	7	4	66	32*1	57%	(66/115)	0.58
	GK	2	6	0	7	47	2*2	73%	(47/64)
適合率		70%	60%	45%	58%	73%	正答率		
		(7/10)	(26/43)	(10/22)	(66/114)	(47/64)	62 %		(156/253)

(b) 不等間隔分割 (本手法)

		判別結果					再現率		F 値 (調和平均)
		KO	FK	CK	TI	GK			
真値	KO	11	0	0	0	0	100%	(11/11)	0.96
	FK	1	12	8	13	12	26%	(12/46)	0.33
	CK	0	2	12	3*1	0	71%	(12/17)	0.60
	TI	0	7	0	77	28*1	67%	(77/115)	0.65
	GK	0	6	0	3	53	2*2	83%	(53/64)
適合率		92%	44%	52%	64%	76%	正答率		
		(11/12)	(12/27)	(12/23)	(77/121)	(53/70)	65 %		(165/253)

*1: イベント大分類は正しいが, チームまたは場所を誤ったケース

*2: イベント大分類は正しいが, チームを誤ったケース

4.3 高レベル特徴分布の確率モデル

式 (3) において, 学習すべきパラメータは, 事前確率 $P^{(e)}$, GMM 混合数 $K^{(e)}$, GMM の各ガウシアン の重み $\alpha_k^{(e)}$, 平均 $\mu_k^{(e)}$ および共分散 $\Sigma_k^{(e)}$ である. まず, $P^{(e)}$ は, 学習データセット $\mathcal{F}^{(e)}$ に含まれるサンプル数 $n^{(e)}$ に比例して $P^{(e)} = n^{(e)} / \sum_{e \in \mathcal{E}} n^{(e)}$ のように設定した.

一方, $\alpha_k^{(e)}$, $\mu_k^{(e)}$ および $\Sigma_k^{(e)}$ は, GMM がイベント種別 e ごとの高レベル特徴分布 $\mathcal{X}^{(e)} = \{W_M \mathbf{f} \mid \mathbf{f} \in \mathcal{F}^{(e)}\}$ を近似するよう, EM アルゴリズムにより最適化を行った. なお, GMM 混合数はすべてのイベント種別 e に対して $K^{(e)} = 5$ とし, 共分散 $\Sigma_k^{(e)}$ はすべて対角行列に限定した.

5. 実験

サッカー映像 3 試合 5 ハーフを学習用 4 ハーフと評価用 1 ハーフとに分け, 判別精度の評価を行った. なお, 各ハーフには約 50 イベントが含まれているため, 学習用データセットには約 200 イベント, 評価用データセットには約 50 イベントが含まれている.

学習用ハーフと評価用ハーフの組み合わせを変えることで交差検定を行った結果を表 2 のコンフュージョンマトリクスで示す. 均等間隔 (5 × 5) の場合を (a), 不等間隔 (5 × 5) の場合を (b) に示す. 不等間隔の分割により, FK の場合を除くイベント種別では, 再現率で 9~36%, 適合率で 2~22% 向上している.

本手法では, 低レベル特徴の次元を抑えるため, イ

ベントの有無や種別によらず全映像区間において選手の分布の統計をとった. したがって, 全般的な偏在性は考慮できるため概ね精度が向上するものの, 個々のイベントに注目すると傾向が一致しないものが生じる可能性がある. 本例の FK は偏在性が薄く, むしろ均等間隔分割が有利であったといえる.

6. まとめ

フィッシャー判別基準とガウス混合モデルに基づくサッカーイベントの判別法 [1] において, 低レベル特徴の抽出段階において選手配置を不等間隔格子で計数することにより, 偏在性のあるイベントの判別精度向上が可能であることが示された.

今後, イベント毎にそれぞれ分割を変えた格子で計数を行い, その全パターンを低レベル特徴として用いるなど, イベントによる偏在性の違いを考慮して, FK を含む全イベントへの低レベル特徴の最適化を図りたい.

参考文献

- [1] 三須, 苗村, 藤井, 八木: 選手フォーメーション解析によるサッカーイベント判別法, 動的画像処理実用化ワークショップ DIA2007, O1-1, pp.1-6, 2007.
- [2] 瀧, 長谷川: チームスポーツにおける集団行動解析のための特徴量とその応用, 信学論 D-II, vol. J81-D-II, no. 8, pp. 1802-1811, 1998.
- [3] 篠原, 大津: フィッシャー重みマップを用いた顔画像からの表情認識, 信学技報, PRMU2003-269, vol. 103, no. 737, pp. 79-84, 2004.