

画像特徴とテキスト特徴を用いたWebスポーツニュース画像のイベント分類

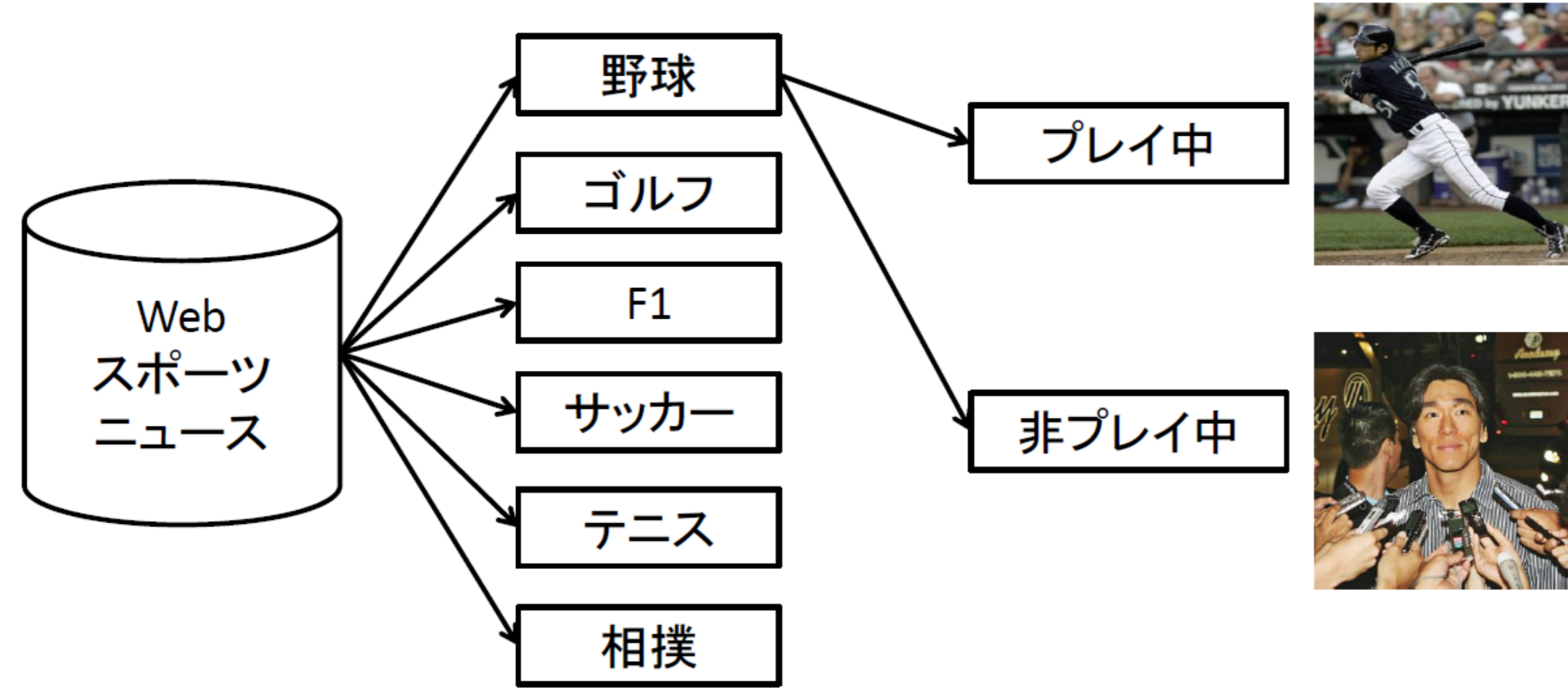
IS1-25 電気通信大学 情報工学専攻 北原 章雄 奥山幸也 柳井 啓司

背景

- Web上には大量のニュース画像が存在する
- それらを収集・蓄積することで有用なニュース画像データベースとなりうる
- しかし
- ジャンルや画像の内容ごとに分類されていない(国内、経済、スポーツなどの大雑把なレベル)
- 目的のニュース画像を探すのが困難(例)松坂が投げている画像ニュースが見たい

目的

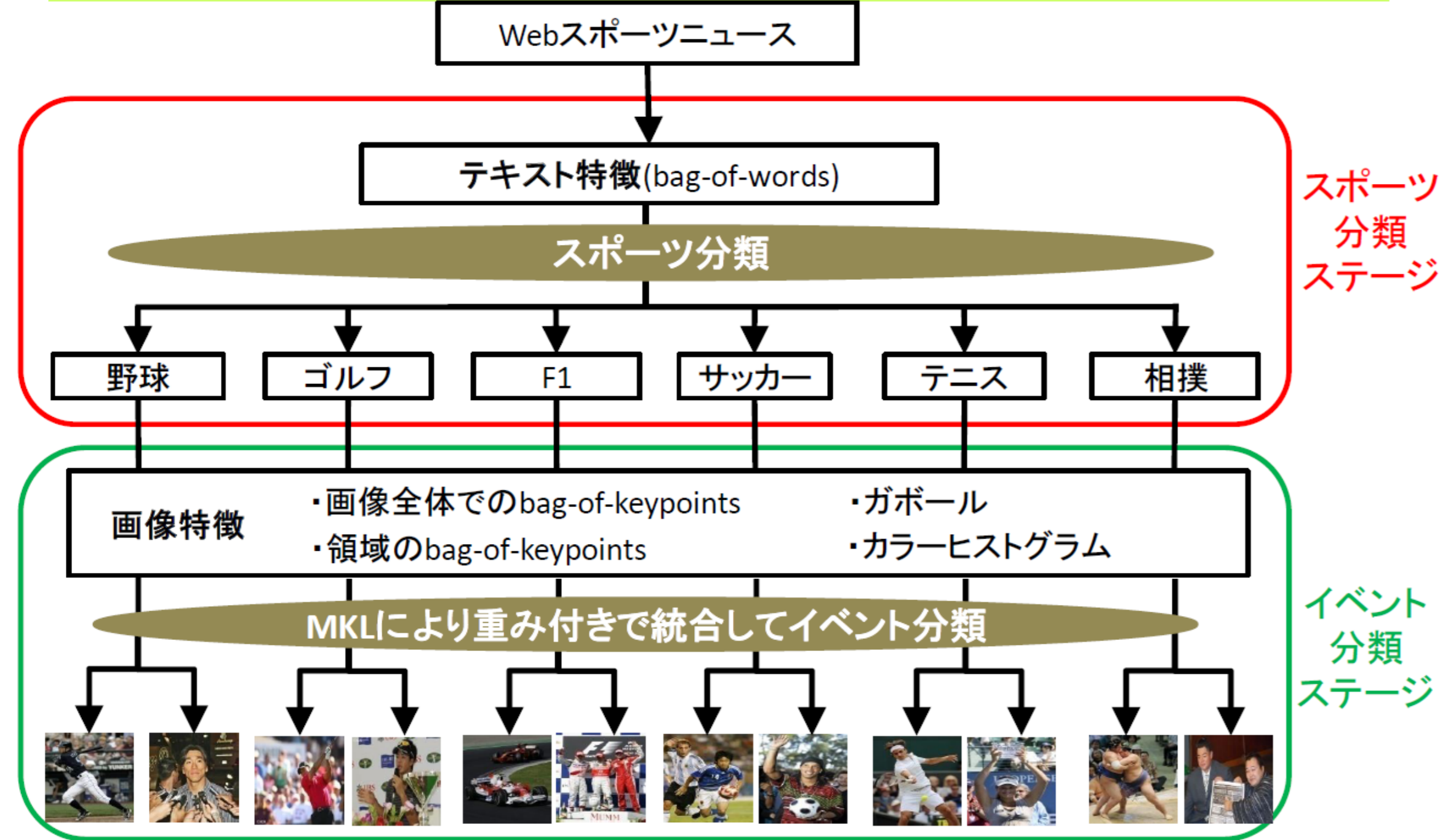
- スポーツニュース画像のイベント分類を行う
- イベント=プレイ中か非プレイ中か



研究の方針

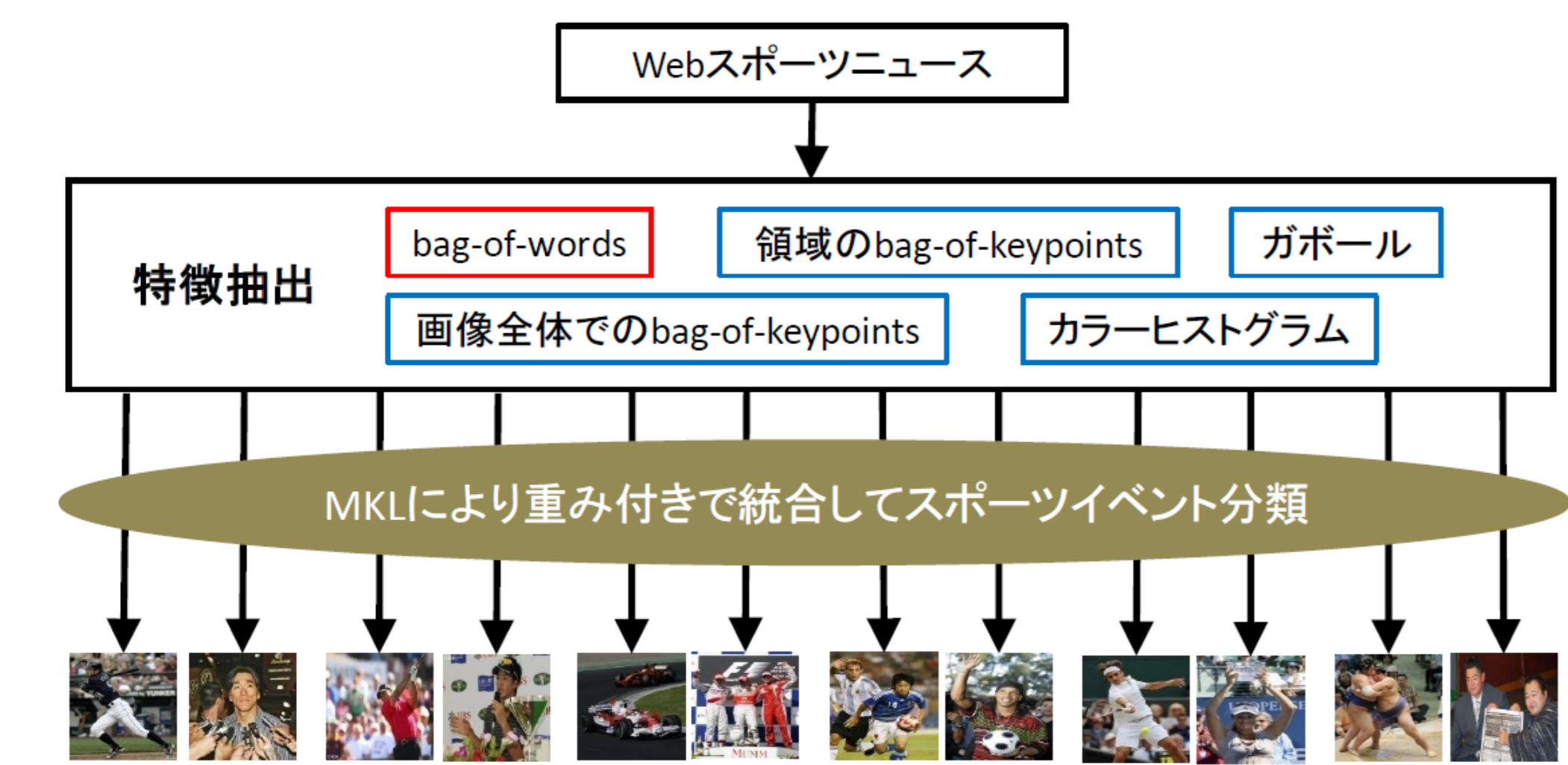
- 2つのスポーツイベント分類手法を提案
- 2段階手法
 - テキスト特徴を用いてスポーツ分類
 - 画像特徴を用いてイベント分類
- 1段階手法
 - テキスト特徴と画像特徴を同時に使ってスポーツイベント分類

2段階手法概要



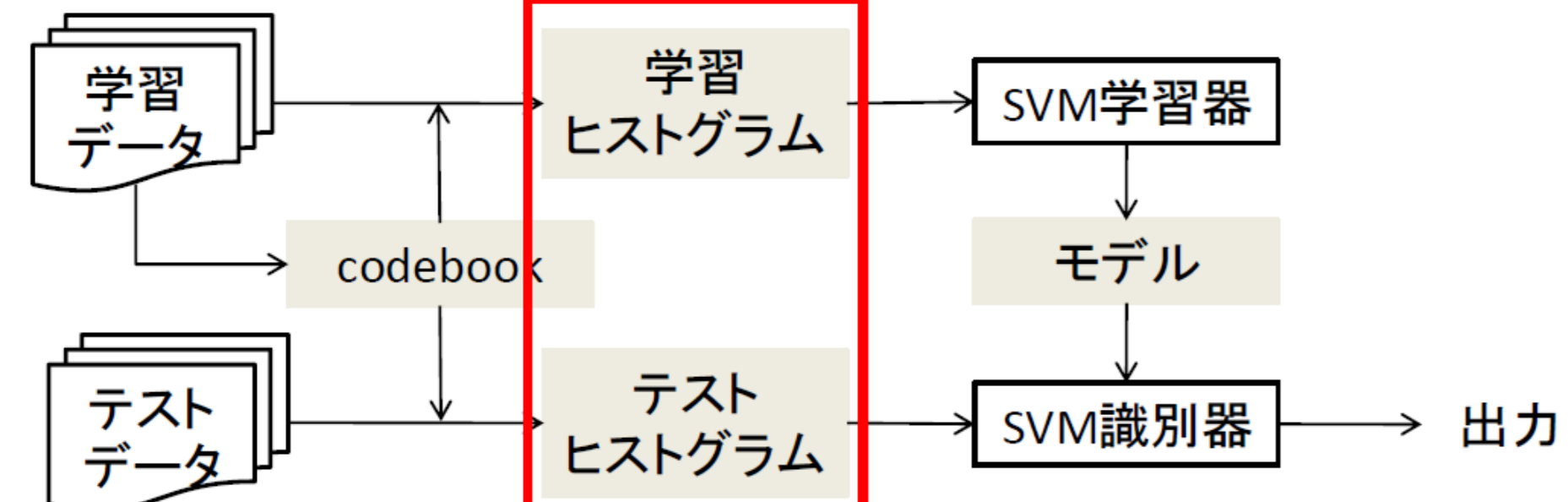
1段階手法概要

- テキスト特徴と画像特徴を同時に使ってスポーツイベントに分類



Bag-of-words

- テキストを品詞分解し、各テキストの単語頻度を得る
- 学習データの全テキスト中に出てくる単語からcodebookを作成
 - 全体の上位N個
 - 全体の上位と各スポーツの上位の組合せ
- codebookをもとに単語頻度ヒストグラムを作成
- SVMにより分類



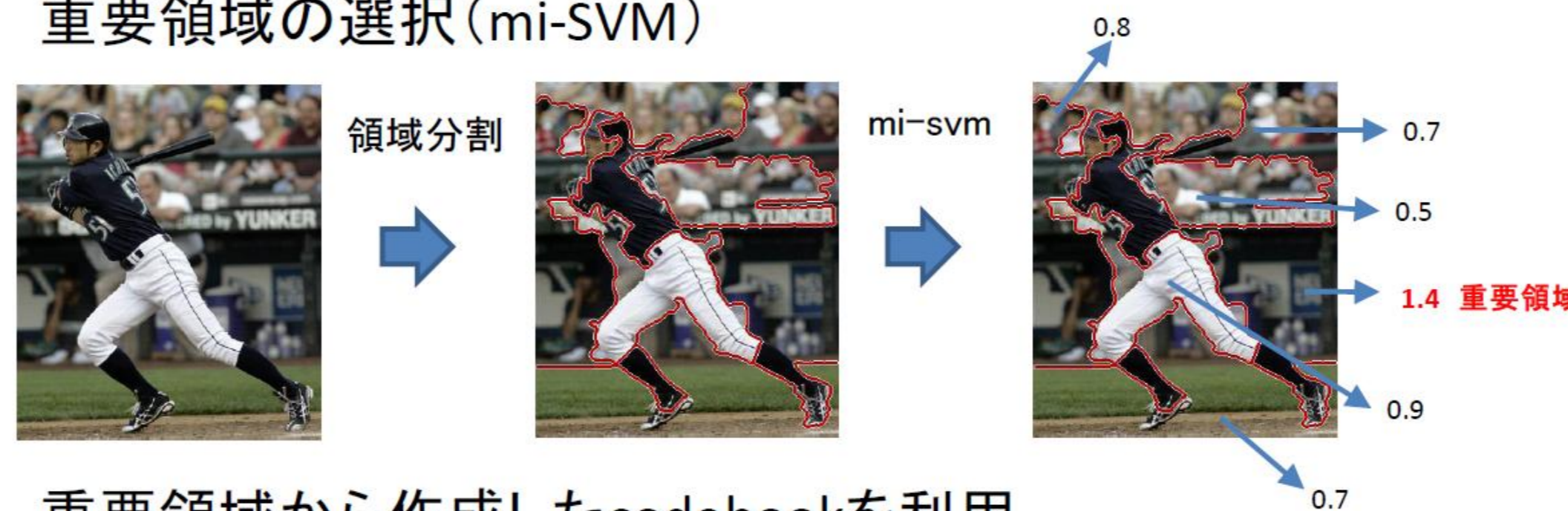
Bag-of-keypoints

- 画像を局所特徴をヒストグラム表現する
- 局所特徴: SIFT特徴
 - スケール変化, 回転・視点変化にロバスト

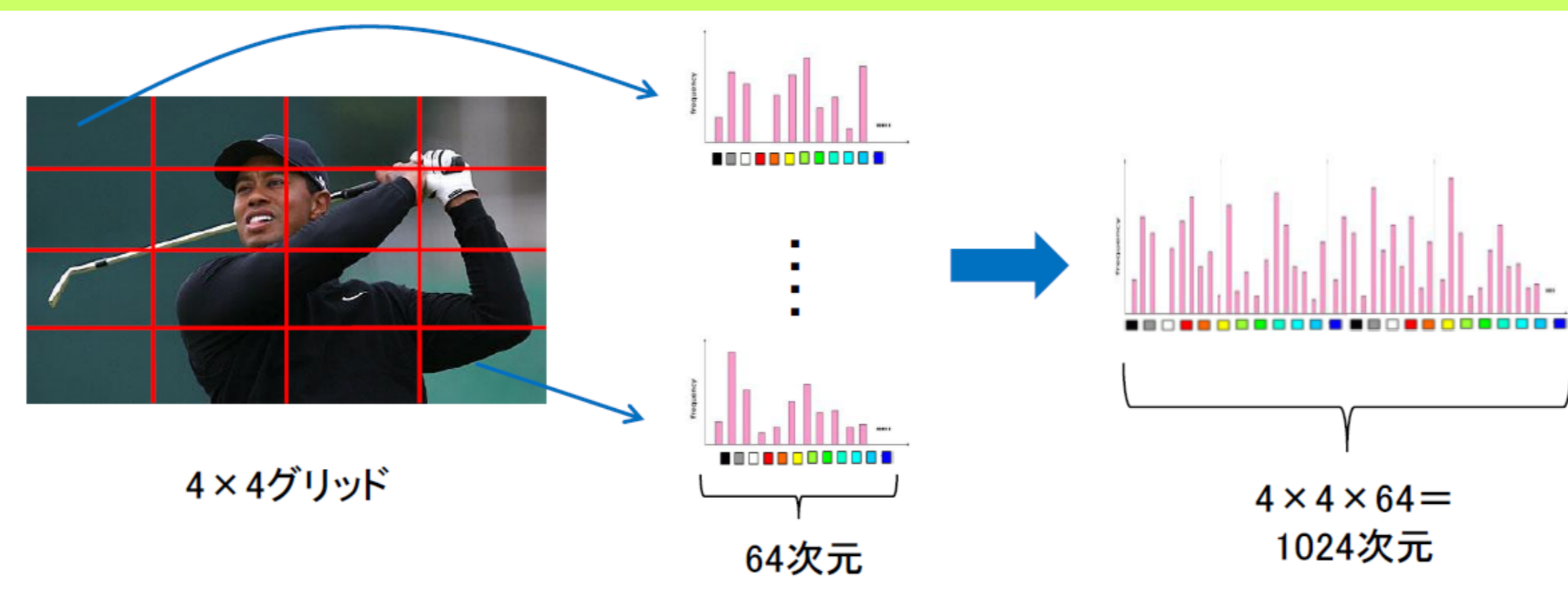


領域でのbag-of-keypoints

- 画像を領域分割(JSEG)
- 重要領域の選択(mi-SVM)
- 重要領域から作成したcodebookを利用

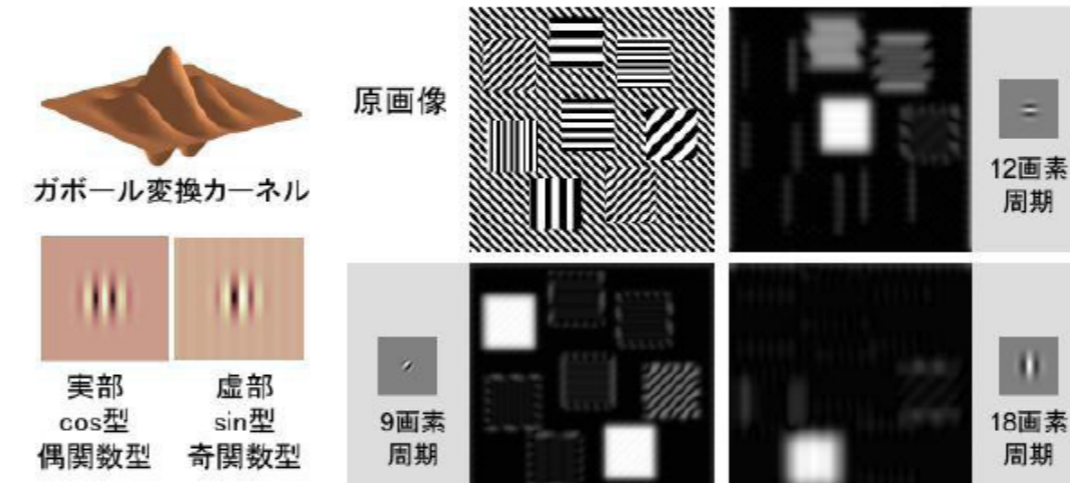


カラーヒストグラム



ガボール

- ガボールフィルタカーネルを用い、画像の局所的な濃淡情報を表現
- 画像を4x4グリッドに分割
 - 6x4x4x4=384次元



特徴統合(MKL-SVM)

- Multiple Kernel Learning(MKL)
 - 複数のカーネルを重み付きで統合、新たなカーネルを作成
 - 最適な重みパラメータ β_j を求める

$$K_{combined}(x, x') = \sum_{j=1}^K \beta_j k_j(x, x'), \quad \sum_{j=1}^K \beta_j = 1, \quad \beta_j \geq 0$$

- MKL-SVM: サブカーネルにSVMのカーネルを使う
 - サブカーネルは x^2 カーネル

$$k(x, x') = \exp(-\gamma \sum_i \frac{\|x - x'\|_i^2}{x + x'})$$

実験

- 野球, ゴルフ, F1, サッカー, テニス, 相撲の6種類を分類
- データセット: Yahoo!JAPAN写真ニュース
 - 学習用にプレイ中, 非プレイ中100枚ずつ
 - テスト用にも同様に100枚ずつ

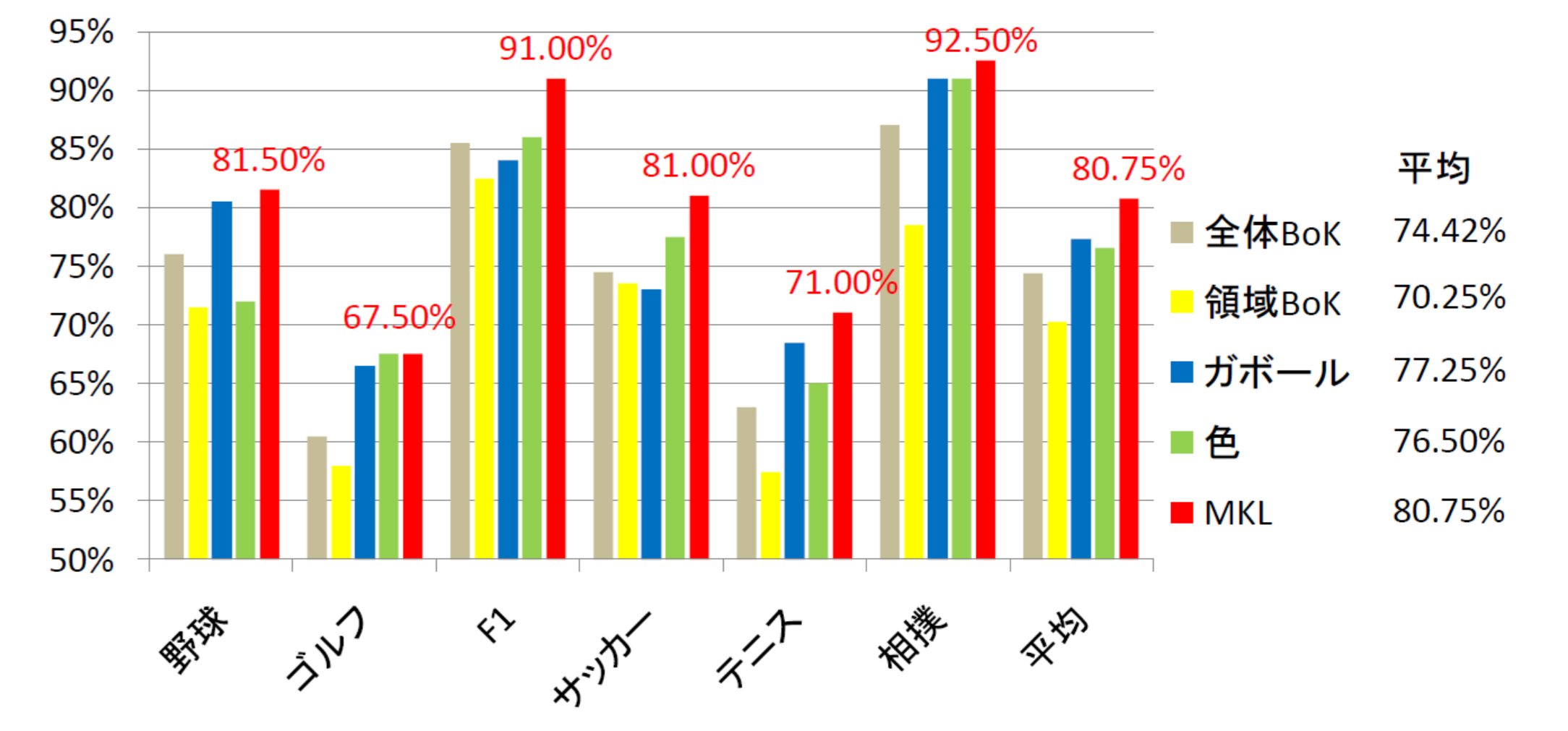
- 評価指標: 分類率 = $\frac{\text{正しく分類されたニュースの数}}{\text{分類するすべてのニュースの数}}$

2段階手法スポーツ分類結果

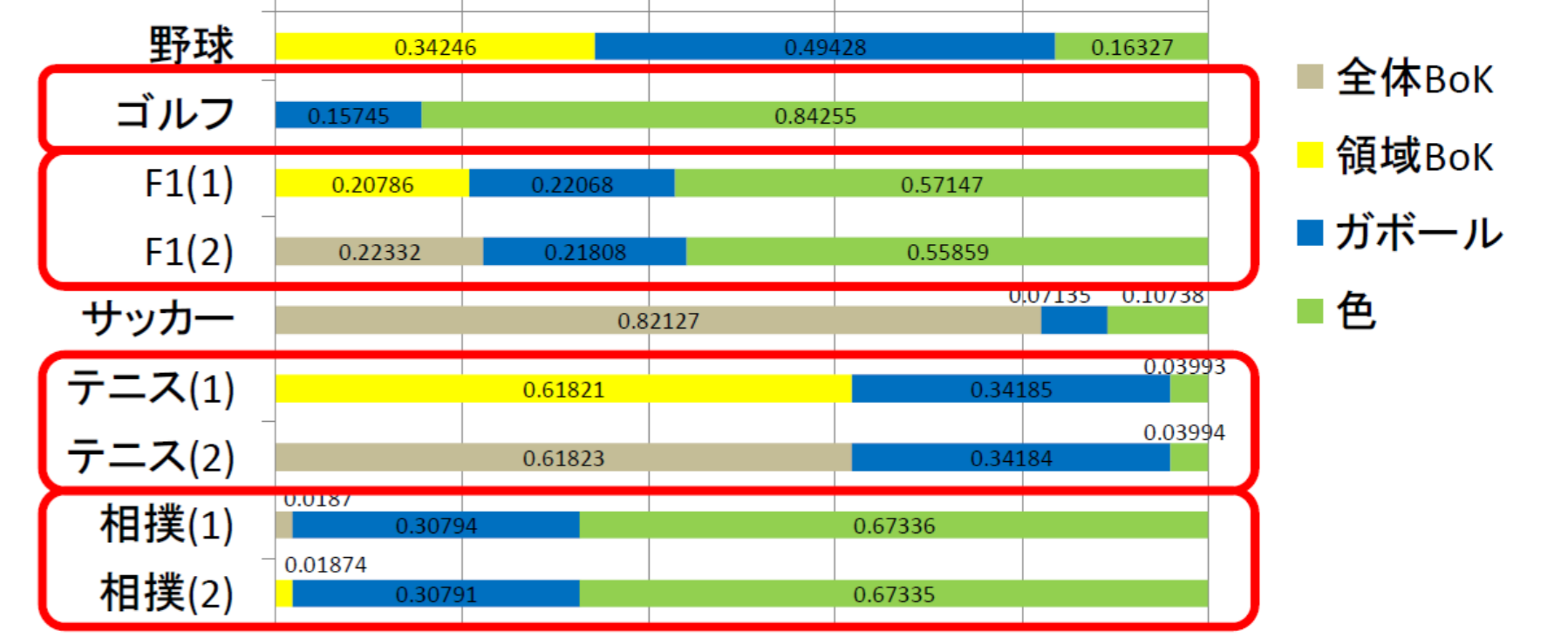
codebookサイズ	分類率
1000: 全学習データの上位1000単語からcodebookを作成	99.00%
2000	99.17%
1000-100: 全学習データの上位1000単語と各スポーツ学習データの上位100単語からcodebookを作成	99.25%
1000-300	99.33%

codebookサイズ 1000-300 の混合行列	野球	ゴルフ	F1	サッカー	テニス	相撲
野球	195	0	4	0	0	1
ゴルフ	0	200	0	0	0	0
F1	0	0	200	0	0	0
サッカー	0	0	0	199	0	1
テニス	1	1	0	0	198	0
相撲	0	0	0	0	0	200

2段階手法イベント分類結果



MKLによって得た各特徴の重み



- F1, テニス, 相撲は最もよい重みが2種類存在
- 全体BoKと領域BoKは特徴として類似

プレイ中と認識された画像 (上段: 正解, 下段: 不正解)



1段階手法結果

分類率と各特徴の重み

MKL_C	分類率	テキスト	全体BoK	領域BoK	ガボール	色
0	68.00%	0.61687	0	0.16136	0.08261	0.13916
10	71.08%	0.56720	0.06240	0.11919	0.11919	0.13202
20	73.17%	0.49127	0.12718	0.12718	0.12718	0.12718
30	75.92%	0.41920	0.14520	0.14520	0.14520	0.14520
40	77.00%	0.35616	0.16096	0.16096	0.16096	0.16096
50	76.75%	0.30392	0.17402	0.17402	0.17402	0.17402
60	77.00%	0.26121	0.18470	0.18470	0.18470	0.18470
70	76.92%	0.22611	0.19347	0.19347	0.19347	0.19347
80	77.08%	0.2	0.2	0.2	0.2	0.2
90	77.08%	0.2	0.2	0.2	0.2	0.2
100	77.08%	0.2	0.2	0.2	0.2	0.2

1段階手法の混合行列

分類後のクラス												
	b_p	b_n	g_p	g_n	f_p	f_n	so_p	so_n	t_p	t_n	su_p	su_n
b_p	68	32	0	0	0	0	0	0	0	0	0	0
b_n	7	93	0	0	0	0	0	0	0	0	0	0
g_p	0	0	56	44	0	0	0	0	0	0	0	0
g_n	0	0	23	77	0	0	0	0	0	0	0	0
f_p	0	0	0	0	81	19	0	0	0	0	0	0
f_n	0	0	0	0	7	93	0	0	0	0	0	0
so_p	0	0	0	0	0	0	85	15	0	0	0	0
so_n	0	0	0	0	0	0	22	78	0	0	0	0
t_p	0	0	0	0	0	0	0	0	17	83	0	0
t_n	0	0	0	0	0	0	0	0	7	93	0	0
su_p	0	0	0	0	0	0	0	0	0	0	89	11
su_n	0	0	0	0	0	0	0	0	0	0	9	91

スポーツ: b 野球, g ゴルフ, f F1, so サッカー, t テニス, su 相撲
イベント: p プレイ中, n 非プレイ中

まとめ

- Web画像スポーツニュースをプレイ中/非プレイ中に分類した
- 2段階手法
 - スポーツ分類: 99.33%
 - イベント分類: 平均80.75%
- 1段階手法
 - 全体分類率77.08%(スポーツ分類率だけみると100%)

今後の課題

- スポーツの種類を追加
- 特徴量の追加
- 選手のポーズ認識
- イベントの追加
- 各スポーツに特化した手法の検討