

映像特徴に基づく自動映像分類システムの提案

服部 しのぶ[†] 高木 真一^{‡*} 小館 亮之[‡] 富永 英義^{‡‡}

[†] 早稲田大学 理工学部 電子・情報通信学科

〒 169-8555 東京都新宿区大久保 3-4-1

{sally,takagi,tominaga}@tom.comm.waseda.ac.jp

[‡] 早稲田大学 国際情報通信研究センター

* 早稲田大学 大学院 国際情報通信研究科

〒 169-0051 東京都新宿区西早稲田 1-3-10 早稲田大学 29-7 号館

akikod@giti.waseda.ac.jp

あらまし 近年、映像情報の大量化・一般化が進行するなか、ユーザの要望に応じたコンテンツの利用に関する検討の必要性が高まり、メタデータが注目されつつある。しかし現状のメタデータ付加作業は人手を介することが多いため、コスト軽減を実現するシステムの必要性が高まりつつある。そこで本稿では、ユーザの要望に応じたコンテンツ提供を行うサービスを低コストで実現する知的情報フィルタリングサービスについて検討する。その第一歩として、映像コンテンツへのメタデータ付加支援システムのうち、テレビ放送用スポーツ映像を対象としたジャンル分類を行うコンテンツベースの映像分類システムを提案・設計し、実装した結果について報告する。

キーワード 情報フィルタリング、映像分類、カメラパラメータ、キーフレーム、MPEG-7

Proposal of a Content-Based Video Classification System

Shinobu HATTORI[†] Shin'ichi TAKAGI^{‡*} Akihisa KODATE[‡] Hideyoshi TOMINAGA^{‡‡}

[†] Dept. of Elec. Info. and Comm. Engineering, WASEDA University
3-4-1 Ohkubo, Shinjuku-ku, Tokyo, 169-8555 JAPAN

{sally,takagi,tominaga}@tom.comm.waseda.ac.jp

[‡] GITI, WASEDA University

* GITS, WASEDA University

29-7 Building 1-3-10 Nishiwaseda, Shinjuku-ku, Tokyo, 169-0051 JAPAN

akikod@giti.waseda.ac.jp

Abstract With enlargement and generalization of video information of recent years, the necessity for investigating contents use, which meets users' demands, is increasing. Therefore, metadata is receiving attention. However, metadata is added by hand as for now that system, which reduces this cost, is in great need. For this, this papaer studys an Intelligent Information Filtering Service to meet users' demands. As for the first step, we propose a content-based genre classification system for TV sports programs. The system's design and some experimental results are also discussed in this paper.

Keywords Information Filtering, Content Classification, Camera Parameter, Key Frame, MPEG-7

1. はじめに

近年のマルチメディア技術の発展に伴い、映像情報の大量化・一般化が進行されるなか、国際標準化団体 MPEG に代表されるコンテンツ伝送技術に用いるコンテンツ表現方式や管理方式に対する議論が盛んに行われている。具体的には、現在 MPEG ではコンテンツの内容記述を規定する MPEG-7⁽¹⁾ やユーザ間のコンテンツ流通のための枠組みを規定する MPEG-21⁽²⁾ の標準化が活発に行われている。

コンテンツの表現・管理方式に関する検討が進むなか、ユーザの要望に応じたコンテンツの利用に関する検討の必要性が高まりつつある。例えば、MPEG-7 を用いて、視聴者の要望に応じて TV 番組に特化した各番組の特徴（放送時間、ジャンル etc）を記述したメタデータに基づき TV 放送を行うことを目的としている TV-Anytime Forum⁽³⁾ の活動が盛んである。しかし、これらのメタデータ付加作業は、現在人手を介しているため、それにかかるコストの問題が指摘され、コスト軽減を実現するシステムの必要性が高まりつつある。

これらの現状を踏まえ、将来ユーザの要望に応じて情報を提供するサービスが重要となる。本稿ではこれを“情報フィルタリングサービス”と定義する。このサービスは、ユーザの要望に応じたコンテンツの提供を目的とする。このサービスによりユーザの嗜好性等を踏まえたコンテンツの流通・管理が可能になると考えられる。

本稿では、情報フィルタリングサービス実現の第一歩として、TV-Anytime Forum の実現に伴い、TV 番組に対するメタデータ付加支援システムの検討として、TV 番組のジャンル分類についての検討を行う。

本稿では、2.章で映像分類システムについて従来研究とともに述べる。3.章ではスポーツ映像分類システムについて述べるとともに本システムの要素技術について述べる。4.章では本システムの提案過程、5.章では提案分類アルゴリズムについて述べる。6.章で提案分類アルゴリズムの実装結果を示し、最後に7.章で結論と今後の課題について述べる。

2. 映像分類システム

2.1 映像の分類

情報の分類には様々な目的や要素が存在する。例えば、情報の共有を目的とするものや個人的なもの、情報の形状に基づいて映像や音声に分類するものなどがある。本稿において対象としている映像の分類も図1に示したように様々である。これらの映像の分類を一般的に映像のジャンルと定義できる。本稿の対象映像である TV 番組のジャンルは個人の意見に左右されることや、1つの番組が複数のジャンルに属する可能性があることなど、様々なパターンが存在し、1つの分類として表すことは困難であることが分かる。

また、人が考える番組のジャンルと画像の特徴に基づいて分けるジャンルは必ずしも一致しない。例えば、人間がある番組を見てそれをスポーツ系の番組だと判断することは容易である。しかし、画像的には“スポーツ”という大きな枠で区別することは難しい。よって、画像情報を用いた映像分類では、人間の考えるジャンルに沿った分類を行うことは困難である。そこで本稿では画像情報に基づいた分類を目指すものとする。

2.2 スポーツ映像

本稿では、様々な TV 番組の中でもスポーツ映像を対象コンテンツとする。その要因として、スポーツというジャンルは比較的視聴者にとって需要が高いジャンルで

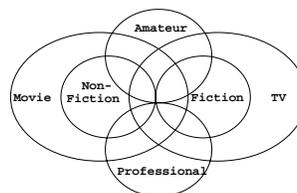


図 1: 映像の分類

あることや、スポーツは特定のルールに従って競技が行われるため、映像が構造化されている可能性が高く、比較的特徴の捉えやすい映像であることが挙げられる。

2.3 映像分類に関する従来研究

- 映像分類の従来研究には以下のものが挙げられる。
- 映像のデータベースを用いた分類に関する研究⁽⁴⁾
 - 画像・言語を用いたシーン分類に関する研究⁽⁶⁾
 - 複数の画像特徴量を用いた分類に関する研究⁽⁵⁾

しかしこれらの従来研究には、予め分類する映像データベース構築を要するため手間がかかることや、分類可能な映像に制限があること⁽⁴⁾⁽⁶⁾、また使用する特徴量の約6割が手入力であるため効率的でないこと⁽⁵⁾など、様々な問題点が存在する。

3. スポーツ映像分類システム

上記の問題点を踏まえ、本稿では自動抽出のみによる特徴量を用いた、予めデータベースの構築を必要としない映像分類システムを提案する。またメタデータの標準化に伴い、本システムの汎用性を踏まえ、用いる特徴量として MPEG-7 記述子である色、動き、エッジ等の特徴量を用いる。なお本システムは、以下に示す特徴を有するものとする。

1. 撮影者の意図を考慮した分類

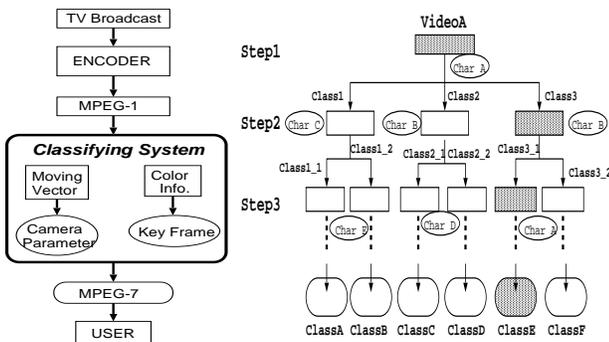
スポーツは特定の規則に沿って行われるため、その映像は構造化されていると考えられる。そこで、TV 放送用の映像は専門家によって撮影されているため映像は安定しており、この構造に基づいた意味を持った動きをしていると考えられる。そこで本稿ではこの“撮影者の意図”に関連する情報として、動き情報から得られるカメラパラメータを用いる。また、色情報から得られる映像のキーフレームとを組合わせて分類する手法を提案する。

2. 各ジャンルに特化した特徴量の定義

スポーツを競技場という観点で分類した場合、サッカーは室外の広い敷地で行うため、カメラは横方向の動きを多くすることが予想出来る。それに比べ、相撲等の室内で行う競技は比較的狭い場所で行われるため、カメラの横方向の動きは少なく安定した動きをすると予想される。また室外競技のようにフィールドで行われる競技は、フィールドが映ることが多いため色情報に特徴が表れると予想される。

このようにスポーツ映像は各競技に特化した特徴量が定義できると考えられる。そこで、本稿ではスポーツ映像の構造を様々な特徴量を用いて定義し、それを用いた分類手法を提案する。

本システムの概要図を図2に示す。図2(a)に示したシステムの流れは、まず TV 映像を encoder により MPEG-1 に符号化し、本システムに入力する。ここで映像からカメラパラメータやキーフレームを抽出し、これを用いて映像を分類する。分類結果は、本システムの汎用性を踏まえ MPEG-7 を用いて記述し、ユーザに提供する。



(a) システムの流れ (b) 分類の流れ

図 2: システムの概要図

また図 2(b) に示したように、本システムでは複数の段階により階層的に分類が行われていくものとする。階層的に分類を行うことにより一度で分類するより正確な分類が可能であると考えられる。例えばこの場合、VideoA を最初に CharA を用いて Class3 に分類し、Step2 で CharB を用いてさらに分類する。ここで必ずしも前段階と異なる特徴量を用いる必要はなく段階ごとに適した特徴量を用いる。これを繰り返して分類が行われる。

3.1 要素技術

3.1.1 カメラパラメータ抽出手法

本稿でのカメラパラメータ抽出手法には、映像から動きベクトルを抽出し、その方向と量をもとに 8 方向にクラスタリングし、そのヒストグラムの時間的変化によりカメラパラメータを判定する手法⁽⁷⁾を用いる。ここで抽出されるカメラパラメータは、Fix, Pan, Zoom, Shake の 4 種類である。なお Shake とは映像がぶれていると判定された場合のカメラパラメータである。

3.1.2 キーフレーム抽出手法

本稿におけるキーフレーム抽出手法には、映像のカット点により分割されたショット単位の映像をフレーム間の YCbCr 成分のヒストグラムを用いてさらに細かい区間に分割し、区間内の色空間での YCbCr 成分を用いたクラスタリングにより、区間毎のキーフレームを抽出する手法⁽⁸⁾を用いる。

4. 提案分類手法

本システムの分類手法の提案過程として 3.1 節に述べた要素技術を用いて Baseball(3 本), Basketball(6 本), Football(6 本), Soccer(11 本), Sumo(7 本), Volleyball(7 本) の、全 6 種類、30 本の映像から得られた結果の分析を行った結果を以下に示す。映像の符号化条件は表 1 に示した通りである。

4.1 カメラパラメータの抽出

4.1.1 カメラパラメータの抽出率

カメラパラメータの抽出を行った結果、以下に示す特徴が得られた。各スポーツ映像の平均カメラパラメータの抽出率を図 3 に示す。

1. 全種類の映像における Fix の高抽出率

その構造的な映像からスポーツ映像はカメラパラメータに特徴が表れることが予想されたが、図 3 より全ての映像に対して Fix が一番多く抽出されたことが分かる。よって、カメラパラメータの抽出率のみを用いた各スポーツの特徴の定義は困難であることが確認できた。

表 1: 実験映像の MPEG 符号化における条件

MPEG1 Sequence Factors	Parameters
Frame Rate	30[frame/s]
I/P Frame Distance	3[frames]
Frame Size	352[pixel]*240[line]
Bit Rate	1.5[Mbps]
Telerecording Device	Sony Giga Pocket
Encoder	Canopus MPEG Station
Total Sequence Time	15[min]
The Number of Frames	26967[frames]

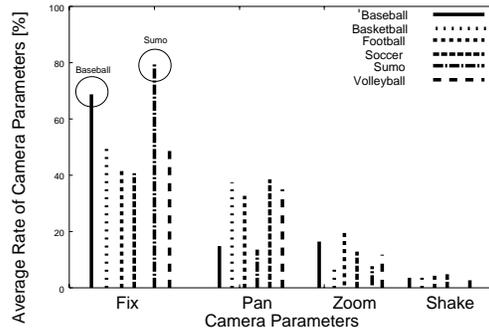


図 3: カメラパラメータの抽出率

2. Sumo と Baseball における Fix の抽出率の特徴

Sumo と Baseball の結果に着目すると、2 つの Fix の抽出率が他のジャンルより大幅に高いことが分かる。これは 2 つとも比較的カメラが固定されており、映像が大きく変化しないことが原因だと考えられる。そこで、Sumo と Baseball の特徴として Fix の高抽出率を定義できることが確認できた。

以上のことより、カメラパラメータの抽出率を用いて各スポーツの特徴を定義することは困難であるが、その抽出率の違いに着目し特徴の 1 つとして Fix の高抽出率を定義付けられることが確認できた。

4.1.2 カメラパラメータ区間の定義

上に述べたように、カメラパラメータの抽出率のみを用いてスポーツの特徴を定義するのは困難である。そこで、カメラパラメータの時間的変化に着目した結果、以下に示す特徴が得られた。

1. カメラパラメータの連続性

図 4 と図 5 の結果を比較すると、Basketball の結果では Fix で連続している区間が Volleyball より長いことが分かる。これは Fix の線の疎らさに着目すると明らかである。線が疎らであるとは、それだけ連続して同じカメラパラメータが抽出されたことを意味する。そのためグラフに変化が少なくなる。これらのスポーツは図 3 の結果に示したカメラパラメータの抽出率では顕著な特徴は得られず同様な値が得られているが、このように時間的変化に着目すると、そのカメラパラメータの連続性に特徴が表れることが確認できた。

2. カメラパラメータ区間の有用性

本稿では上記のことを踏まえ、カメラパラメータの時間的変化を用いて、特定のカメラパラメータが連続して抽出された場合、それを 1 つの区間として“カメラパラメータ区間”と定義する。

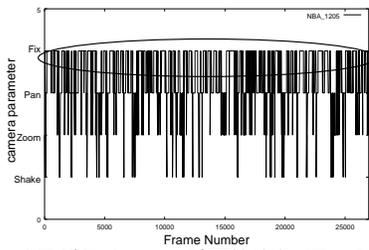


図 4: カメラパラメータの時間的変化 (Basketball)

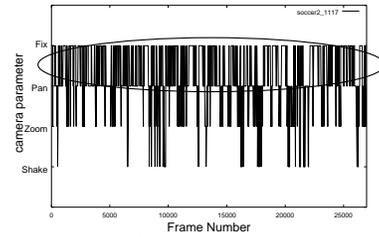


図 7: カメラパラメータの変化 (Soccer)

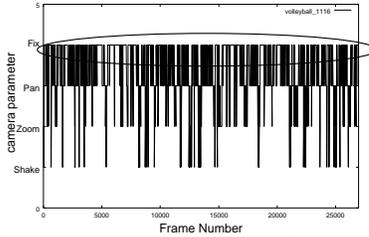


図 5: カメラパラメータの時間的変化 (Volleyball)

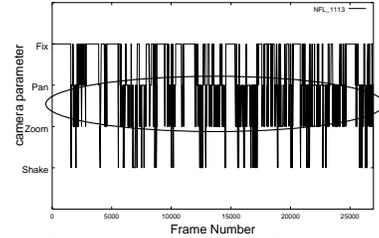


図 8: カメラパラメータの変化 (Football)

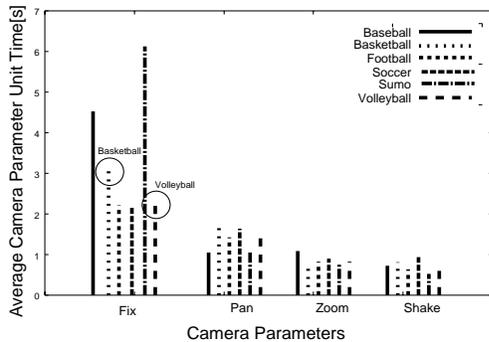


図 6: カメラパラメータ平均区間時間

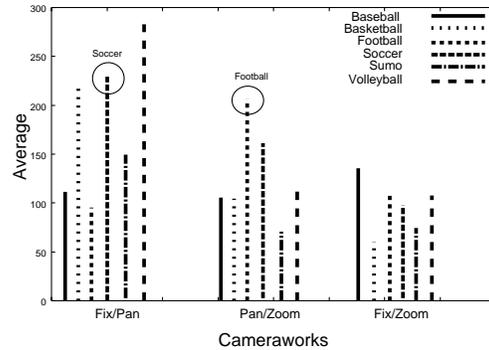


図 9: Camerawork Transition の抽出結果

カメラパラメータ区間を抽出した結果，各スポーツのカメラパラメータ区間とその区間時間に特徴が得られた．各スポーツのカメラパラメータ区間の平均区間時間を図 6 に示す．ここで，図 4 と図 5 に示した Basketball と Volleyball に着目すると，値に大きな差が得られていることが分かる．よって，これを用いてこの 2 つの分類が可能であることから，カメラパラメータ区間の有用性が明らかになった．

4.1.3 Camerawork Transition(CT) の定義

カメラパラメータ区間に加え，カメラパラメータの時間的変化には以下に示す特徴が得られた．

1. カメラパラメータの変化量

図 4，図 5 と同様に Soccer と Football のカメラパラメータの時間的変化を図 7，図 8 に示す．この結果を比較すると Soccer では Fix と Pan 間に線が密集しているのに対して，Football では Pan と Zoom 間に密集した線になっている．

これは，Soccer では Fix と Pan 間，Football では Pan と Zoom 間のカメラパラメータの変化量が多いことを示している．これらのスポーツは図 3 と図 6 に示した結果では顕著な特徴は得られず同様な値が得られているが，このように時間的変化に着目した場合，そのカメラパラメータの変化量に特徴が表れることが確認できた．

2. Camerawork Transition の有用性

本稿では上記のことを踏まえ，カメラパラメータの時間的変化上のカメラパラメータの変化を“Camerawork Transition: 遷移カメラワーク”(以下 CT) と定義する．各スポーツにおける CT の抽出結果を図 9 に示す．なお CT としては，Fix/Pan, Pan/Zoom, Fix/Zoom の 3 種類を用いる．ここで，図 7 と図 8 に示した Soccer と Football の結果に着目すると，一番多く抽出されている CT が異なっていることが分かる．よって，この 2 つの分類が可能であることから CT の有用性が明らかになった．

4.2 キーフレームの抽出

4.2.1 映像の典型的なシーンの定義

映像のキーフレームを抽出することは，その映像の代表的なシーンのフレームを抽出することである．そこで本稿では，映像の典型的なシーンのキーフレームのみを用いることにする．これにより，各映像により特化した情報を用いた特徴定義が可能になると考えられる．

映像の分類と同様に，人間が考える映像の典型的なシーンは画像情報から得られるものと必ずしも一致しない．そこで各スポーツの典型的なシーンを定義した上でキーフレームを抽出した結果，両者に関連性が表れた．表 2 に Baseball と Soccer の一映像から得られたキーフレームを主観評価により 4 種類に分類したものとその平均区間時間を示す．

表 2: キーフレーム抽出結果

Genre	Total Key Frame	Class	Class Key Frame	Ave Unit Time
BASEBALL	323	box	23[frames]	5.89[s]
		player	60	4.10
		field	13	3.55
		other	70	2.78
SOCCER	304	field	69	5.83
		non-filed	31	2.62
		player	70	2.75
		other	69	2.04

表 3: Typical Key Frame の抽出結果

	Total TKF	True TKF Rate	Rate
BASKETBALL	24.2[frames]	84.7[%]	0.09[%]
FOOTBALL	22.9	60.1	0.08
SOCCER	21.0	76.7	0.08
VOLLEYBALL	13.4	61.5	0.04

表 2 より, その区間時間が長いものほど人間が考える典型的なシーンのキーフレームとして抽出されることが分かる. 例えば Baseball の人間が考える典型的なシーンを選手がバッターボックスに立っているシーン, Soccer をフィールド全体が映っているシーンとすると, それぞれのキーフレームの平均区間時間は他より長いことが分かる. よって, 一回に映っている映像の区間が長いものほど人に強い印象を与えるということから, その映像の典型的なシーンと定義付けられることが分かる.

4.2.2 Typical Key Frame(TKF) の定義

上記のことを踏まえ, この典型的なシーンのキーフレームを “Typical Key Frame: 典型的キーフレーム” (以下 TKF) と定義する. TKF を用いることで, より厳選されたキーフレームの情報を用いるため, 映像により特化した分類が可能となり人間が考えるジャンルの分類に近い分類が可能になると考えられる.

各スポーツ映像のキーフレームを抽出した結果, 区間時間と TKF に関連性が表れた. 抽出したキーフレームと典型的なシーンの適合率を分析した結果, 区間が 10 秒以上のキーフレームが一番適合率の良い結果が得られた. よって, 本稿では TKF を抽出されたキーフレームの区間時間が 10 秒以上のもので定義する.

そこで, TKF の抽出結果を表 3 に示す. なおここで, True TKF Rate とは正しく抽出された TKF の割合を示す適合率, Rate は合計フレーム数に対する TKF の割合のことを示す. 表 3 に示した結果より, 全体に対して TKF が占める割合が小さいことや, TKF の適合率より TKF の有用性が明らかになった. しかし, 正しく抽出されなかった TKF の原因として, スポーツでは選手が怪我をした場合など, 例外的に長く映されるシーンがあることが挙げられる. そのため, TKF の抽出に影響が出たと考えられる.

4.2.3 dominant color の定義

TKF の有用性を踏まえ, これを用いた分類手法として各フレーム内の色情報を用いた手法を提案する.

例えばフィールド競技とそれ以外の室内競技の間には色情報に異なる特徴が得られると予想できる. そこで, YCbCr を用いて TKF のフレーム平均値を抽出した.

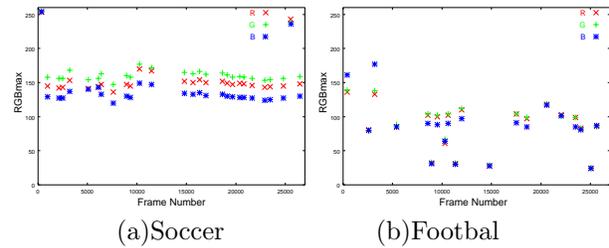


図 10: TKF の dominant color(フィールド競技)

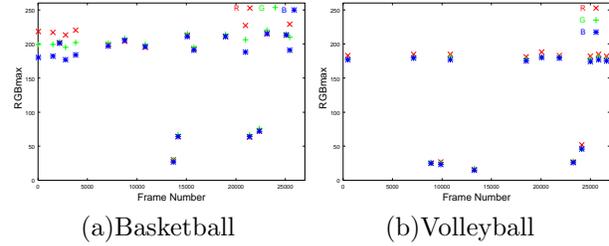


図 11: TKF の dominant color(室内競技)

その結果, 予想に反して, 全ての YCbCr 値において同様な値が得られ, あまり顕著な特徴が得られなかった. この手法では各フレーム内の色情報を平均しているため, 例えばフィールドと観客席が同じフレーム内に映っているものが抽出された場合, 2 つを合わせてしまっていることが原因の 1 つだと考えられる.

そこで本稿では, MPEG-7 記述子である “dominant color” に着目する. “dominant color” とは, フレーム内の主要な色を記述する記述子である. これを用いて TKF の dominant color を抽出した場合, フィールド競技はフレーム内で芝生が占めている部分が多いため, フィールドの特徴を dominant color として抽出でき, 正確にフィールド競技の分類が可能になる.

そこで, YCbCr を用いて TKF の dominant color を抽出した結果, Cb と Cr の値はフィールド競技, 室内競技に関わらずほぼ同じ値が得られ, Y の値のみに特徴が得られた. よって, 本手法では特徴が表れやすい Y の値のみを用いることになる. しかし, これでは Cb と Cr の値を全く無視してしまうことになるため, 良好な分類結果が得られない可能性がある. そこで, 色情報には YCbCr, RGB, HSV など様々なフォーマットが存在することから, RGB を用いて dominant color の抽出を行った. 図 10 にフィールド競技, 図 11 に室内競技における抽出結果を示す.

図 10, 図 11 で RGB の値がともに共通する領域内に収まっていることから, フィールド競技と室内競技の間に共通した特徴を定義することが可能だと考えられる. つまり, YCbCr を用いる手法より良好な結果が得られると予想出来る. 以上のことを踏まえ, 本稿では TKF を用いた分類手法として TKF における RGB フォーマットの色情報を用いて抽出した dominant color を用いてフィールド競技を分類する手法を提案する.

ここで重要となるのがフィールドの特徴定義である. 信憑性の高い特徴定義をするには, 画像から得られる全ての情報を用いることである. そこで, RGB 変換時に YCbCr 全ての値を用いる G の値を利用することにする. 図 10 と図 11 より, フィールド競技の G の値が 150 以下のものが多く, 室内競技は 200 付近のものが多く 150 以下はほとんど存在しないことが分かる.

表 4: スポーツ映像分類結果

Genre	Total Cnt.	Success Cnt.	Class	Rate
BASEBALL	3	2	class1_3	66.7[%]
BASKETBALL	13	10	class4_1	76.9
FOOTBALL	17	16	class3_2	94.1
SOCCER	14	12	class3_1	85.7
SUMO	11	7	class1_1	63.6
VOLLEYBALL	10	8	class4_1	80.0

ここで、Rate とは正しく分類されたコンテンツの割合を表す。また、表の Class とは各特徴量の分析結果より各スポーツの最終的に分類されると予想される Class のことである。

表 4 より、比較的良好な結果が得られていることが分かる。よって、本提案分類アルゴリズムの有用性が示された。しかし、最終的に同じ分野に分類されているスポーツや、コンテンツの数が圧倒的に少ないものや、他のスポーツに比べて成功率が低いスポーツが存在することなど残された課題が存在することも明らかになった。

7. まとめ

本稿では、情報フィルタリングサービス実現の第一歩として、スポーツ映像を対象としたジャンルの分類手法を提案した。その際用いる特徴量として、カメラパラメータから得られるカメラパラメータ区間、Camera-work Transition を定義した。またキーフレームからは Typical Key Frame と dominant color を定義した。そして、提案分類アルゴリズムの確認実験を行い、その有用性を検証した。

今後の課題としては、コンテンツの増量、Baseball の分類手法の改善、Basketball と Volleyball の分類手法の検討、TKF の定義の改善、dominant color を用いたフィールド・室内競技分類手法の改善、映像の内容構造に対応した分類手法の検討、そしてスポーツ映像以外の映像に対応した分類手法の検討などが挙げられる。

参考文献

- (1) MPEG7: “MPEG-7 Japan,” <http://www.itscj.ipsj.or.jp/mpeg7/>
- (2) ISO/IEC JTC1/SC29 WG11: “MPEG-21,” <http://www.darmstadt.gmt.de/mobile/hm/projects/MPEG7/Mpeg21.html>
- (3) TV-Anytime Forum: “TV-Anytime Forum,” <http://www.tv-anytime.org/>
- (4) 孟 洋 他: “事例を用いた映像シーン分類手法とその評価,” 情処全大, 6P-1, March 1998
- (5) 佐藤 隆 他: “映像コーパスの構築と分析,” 信学論, D-II, Vol.J82-D-II, No.10, pp1552-1560, Oct 1999
- (6) 新田 直子 他: “放送型スポーツ映像の構造を考慮した常用シーンへの自動アノテーション付け,” 信学論, D-II, Vol.J84-D-II, No.8, pp1838-1847, Aug 2001
- (7) 土橋 健太郎 他: “手ぶれを考慮した MPEG2 からのカメラワーク検出に関する検討,” 信学総大, D-12-56, p223, 2001
- (8) 大串 亮平 他: “動画像からのキーフレーム抽出に関する検討,” 信学総大, D-12-53, p220, 2001

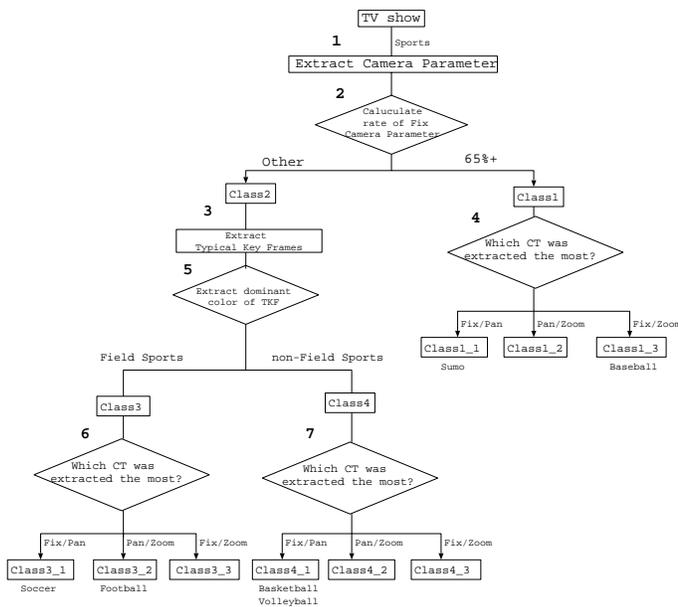


図 12: 提案分類アルゴリズム

そこで、閾値を用いて各映像から得られた TKF の内、条件を満たすものが半分以上存在した場合をフィールド競技と分類し、それ以外を室内競技とする手法を用いて評価した結果、閾値を 170 で設定する方法が一番最適であることが分かった。よって、フレームの RGB フォーマットを用いて抽出した dominant color の G の値が 170 以下のものをフィールドの特徴を有するフレームと定義する。

5. 提案分類アルゴリズム

上記のことを踏まえ、図 12 に示す映像の分類アルゴリズムを提案する。また、以下にその流れを示す。

1. 映像からカメラパラメータを抽出する。
2. Fix の抽出率を求め、65% 以上のものを class1, それ以外を class2 に分類する。
3. class2 に分類されたものは、映像から区間 10 秒以上の Typical Key Frame を抽出する。
4. class1 に分類されたものは、抽出されたカメラパラメータから CT を抽出し、一番多く抽出されたものをもとに class1_1(Fix/Pan),class1_2(Pan/Zoom), class1_3(Fix/Zoom) に分類する。
5. class2 において抽出された TKF を用いて dominant color を抽出し、それをもとにフィールド競技とそれ以外の class3 と class4 に分類する。
6. ステップ 4 と同様に、抽出されたカメラパラメータより CT を抽出し、一番多く抽出されたものをもとに class3_1(Fix/Pan),class3_2(Pan/Zoom), class3_3(Fix/Zoom) に分類する。
7. ステップ 4,6 と同様に、抽出されたカメラパラメータより CT を抽出し、一番多く抽出されたものをもとに class4_1(Fix/Pan),class4_2(Pan/Zoom), class4_3(Fix/Zoom) に分類する。

6. 実験・評価

図 12 に示した提案分類アルゴリズムの実装を表 1 の条件で符号化した全 6 種類、68 本の映像に行った結果を表 4 に示す。