

J-016

# 画像特徴に基づくトリミング用メタデータの簡易抽出手法

## Simplified Extraction Method of the Metadata for Trimming Based on Image Features

沼田 誠† 妹尾 宏† 鹿喰 善明† 金次 保明†  
Makoto Numata Hiroshi Senoo Yoshiaki Shishikui Yasuaki Kanatsugu

### 1. まえがき

放送受信機能が付いた PDA や携帯電話などのモバイル機器の普及に伴い、低解像度ディスプレイで放送番組を視聴する機会が増えてきている。しかし、放送番組はハイビジョンテレビなど高解像度テレビでの視聴を前提として制作されているため、低解像度ディスプレイでは画像内容の視認性に課題がある。

我々は視聴環境適応型放送サービス(AdapTV)のアプリケーションの一つとして、受信機のディスプレイ解像度に応じたトリミングを行うシステム[1]の研究を進めている。このシステムで映像の一部をトリミング表示させることにより、低解像度ディスプレイでの視聴における視認性の改善や視聴者の興味に応じた部分の表示を行うことができる。この手法では、メタデータとして登場人物などのオブジェクトの名称、位置・サイズを利用するが、これらは放送局側で作成され、伝送されてくることを前提としていた。

本研究では、受信機側でトリミングシステムに必要なメタデータを自動生成するための手法を提案する。これまで、オブジェクトを検出する手法として、サッカー映像上のボール検出[2][3]などが提案されている。これらの手法では、高精度な検出が可能であるが、計算量が大きく、民生用放送受信機器などの能力に限られたハードウェアに実装するのは困難である。また、数フレームでも変換を失敗すると非常に不自然な映像になってしまうので高い信頼性が必要である。本稿では、信頼性の高いトリミング用メタデータを放送映像の画像特徴のみから抽出する簡易手法を開発し、実験を行ったので報告する。

### 2. サッカー映像のトリミングとメタデータ

放送番組は多種多様なジャンルがあり、各々で構図なども異なる。本稿では、サッカー番組に適応したトリミング用メタデータ作成ルールを提案する。

低解像度ディスプレイでの視聴でトリミングによる視認性の改善効果が大きいのは、フィールド部分を広く映したショット「ロングショット」と数人の選手の様子が分かるようにフィールドの一部分を映した「ミディアムショット」の場合である。そこで、これらのショットについて、トリミングで選手部分を拡大表示させることができるメタデータを自動抽出することを目的とする。

まず画像特徴から「ロングショット」「ミディアムショット」を判別し、トリミングの対象となる選手部分をメタデータとして抽出する。これ以外のショット(以下、「その他のショット」)は、選手のアップや観客席等が含まれるが、そのままの映像でも視認性は良いので、拡大表示するようなトリミングは行わない。画像特徴だけで

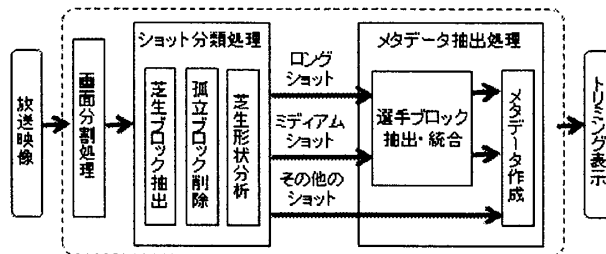


図1. 簡易抽出手法の概要

は判断がつかないショットについても、トリミングによって不自然な映像にならないようにする。

### 3. メタデータの簡易抽出手法

提案するメタデータ簡易抽出手法の概要を図1に示す。本手法は、画面分割処理、ショット分類処理とメタデータ抽出処理から構成され、映像フレーム毎に処理を行う。

#### 3.1 画面分割処理

放送ハイビジョン映像(1920x1080)を 240x135 解像度に縮小し、さらに 16x16 画素の 120 ブロックに分割する。

#### 3.2 ショット分類処理

各ブロックの画像特徴を用いて、サッカー映像を「ロングショット」、「ミディアムショット」、「その他のショット」の3種類に分類する。

##### [1] 芝生ブロック抽出

ブロック毎に RGB 平均値を算出し、色情報からフィールド(芝生)の緑色部分(以下、芝生ブロック)を抽出し、それ以外のブロックを非芝生ブロックとする。

##### [2] 孤立ブロック削除

同種類ブロックと上下左右斜め方向とも隣接せずに孤立している芝生または非芝生ブロックは、ノイズと見なし、もう一方のブロック種別に変更する。

##### [3] 芝生形状分析

得られたブロック情報を基に、芝生ブロック集合の形状の複雑度  $C$ 、画面下半分の全 60 ブロックに対して芝生ブロックが占める割合  $R_d$  から、映像のショット分類を行う。芝生ブロック集合の周囲長を  $L$ 、面積を  $S$ 、画面下半分の芝生ブロックの数を  $N_d$  とし、以下の式で計算する。

$$C = \frac{L^2}{S} \dots(1) \quad R_d = \frac{N_d}{60} \dots(2)$$

「ロングショット」は、芝生ブロック集合が矩形形状に近くなる(複雑度  $C$ : 小)のに対し、「ミディアムショット」では選手が比較的大きく映っているため、芝生ブロック集合が複雑な形状(複雑度  $C$ : 大)をとる傾向がある。芝生ブロック自体が少ない場合(芝生割合  $R_d$ : 小)は「その他のショット」とする。

† NHK 放送技術研究所

### 3.3 メタデータ抽出処理

分類されたショットに応じて、トリミング用メタデータを抽出するための処理を行う。「ロングショット」と「ミディアムショット」に分類された場合は、トリミング表示の対象となる選手部分(以下、選手ブロック)を各ブロック内の輝度分散値を用いて抽出する。

#### [1] 選手ブロックの抽出・統合

芝生みのブロックに比べて選手を含むブロックは輝度分散値が大きくなる。したがって、「ロングショット」の場合は、観客席等を省くために芝生ブロックでかつ輝度分散値が大きいブロック、「ミディアムショット」の場合は輝度分散値が大きいブロックを選手ブロックとして抽出する。選手ブロック間の距離(上下左右で隣接するブロック間の距離を1と定義)に近い場合は、選手同士が接近しているものとして、該当するブロックを一つの集合として統合する。

#### [2] トリミング用メタデータ作成

「ロングショット」「ミディアムショット」の場合、3.3 [1] で求めた選手ブロック集合をトリミングの対象となるようなメタデータを生成する。これまでの研究[1]で提案したトリミングシステムは、トリミング対象の重心座標とサイズ(縦幅)を指定し、自動的にディスプレイ解像度に応じたトリミング領域を決定することが可能なので、選手ブロック集合の重心座標、縦幅をメタデータとして記述する。「その他のショット」では、映像をそのまま表示させたいので、放送ハイビジョン映像の重心座標(960,540)と縦幅(1080)をメタデータとして記述する。

## 4. 実験

サッカーの放送映像(30分間)を使用しメタデータ抽出実験を行った。図2は抽出したメタデータに応じて行ったトリミング表示例である。ショットに応じて選手部分の拡大表示(図2(a),(b))、画面中央をそのままトリミング(図2(c))することが可能であった。なお抽出手法の各種パラメータは、予備実験として別のサッカー映像(135分間)を検証した結果、表1に示す値に設定した。同様に、芝生ブロックを抽出するための判別式をブロックRGB平均値を用いて  $G > R + 5$ ,  $G > B + 25$ 、選手ブロックを抽出するためのブロック輝度分散値を100以上、選手ブロックの統合距離を2以下とした。

### 4.1 ショット分類結果

映像から任意に収集した静止画像181枚を正解例として人手でショット分類を行い、その後、提案手法を用いてショット分類を行った。結果を表2に示す。「ロングショット」「その他のショット」の正解率はそれぞれ90%、96.7%であった。「ミディアムショット」は背景の一部に観客席が見えている場合などがあり、正解率が50%と低かったが、誤った場合はすべて「その他のショット」と判定されていた。「その他のショット」では画面中央をトリミングするので、不自然な映像にはならない。

### 4.2 選手ブロック抽出結果

4.1で「ロングショット」と判定された画像において抽出された選手ブロック1300個のうち、実際に選手が映っているブロックは1084個、正解率は約83.4%であった(表3)。フィールド上の白線がはっきり映っている場合などで誤検出が起きていた。

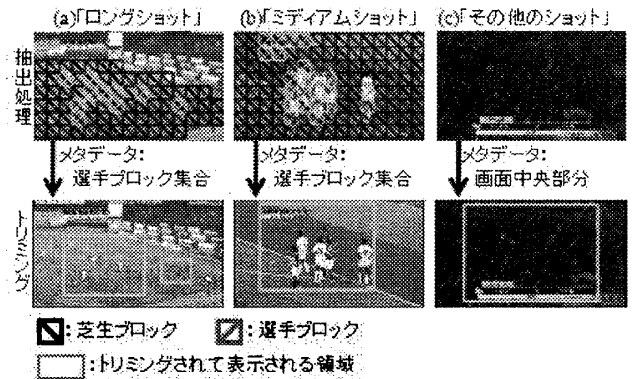


図2. メタデータ抽出とトリミング表示例

表1. ショット分類閾値

	ロング	ミディアム	その他
複雑度 $C$	35 以下	35 超	—
芝生割合(下半分) $R_d$	0.8 以上		0.8 未満

表2. ショット分類結果

	ロング	ミディアム	その他
静止画像総数	110	10	61
判定: ロング	99	0	2
判定: ミディアム	5	5	0
判定: その他	6	5	59
正解率	90.0%	50.0%	96.7%

表3. 選手ブロック抽出結果

抽出された選手ブロック数	実際に選手が映っていた選手ブロック数	正解率
1300	1084	83.4%

## 5. まとめ

本稿では、放送サッカー映像の画像特徴からトリミング用メタデータの簡易抽出する手法について述べた。実験により、簡易な方法で信頼度の高いトリミング用メタデータを抽出できることが確認できた。

今後は、各種パラメータ・判定条件の精査によるメタデータの信頼性向上、別ジャンルの番組への適用、家庭用機器への実装を想定した最適化を行う予定である。

## 参考文献

- [1] 沼田, 澤井, 松村, 上野, 金次, 八木: “オブジェクトの意味と視聴者環境に適応した放送番組の画面再構成手法の検討” 第5回情報科学技術フォーラム FIT2006, K-073(2006).
- [2] 山田, 姫路, 山内, 関: “サッカー試合ビデオにおけるキーフレーム検出のためのサッカーボール認識” 第4回情報科学技術フォーラム FIT2005, I-031(2005).
- [3] 宮本, 上垣, 泉, 福永: “カメラの動きを用いたサッカーの中継映像におけるボールの追跡” 2005年電子情報通信学会総合大会, D-12-13(2005).