

コート情報に基づくバレーボール映像の鑑賞支援と 戦術解析への応用の検討

板摺 貴大^{1,a)} 福里 司^{1,b)} 山口 周悟^{1,c)} 森島 繁生^{2,3,d)}

概要：我々はバレーボール映像を効率的に鑑賞するための一手法を提案する。効率的に鑑賞するために、我々は試合内容把握において重要なラリーシーンを自動抽出する方法と、ラリーシーンを解析・検索する機能を開発した。これらの機能を実現するため、我々はロバストなコート検出手法を紹介する。本手法は事前情報となるコートモデルの種類を変更することで、他スポーツ映像への適用も可能である。また鑑賞支援に留まらず、試合内容の解析や戦術解析技術としての拡張を検討する。

1. はじめに

スポーツ映像の鑑賞はテレビやウェブサイトを通じて行われる、世界中の人々に親しまれている娯楽の一つである。近年、放送システムの発展により、どこでも映像を視聴することが可能となっている。その一方で、バレーボール映像は試合時間が長いため、日常生活で映像鑑賞に充てることのできる時間で観たい試合全てを視聴することは困難である。この問題を解決するため、我々はバレーボール映像を効率的に視聴する手法を提案する。図1に提案手法の概要を示す。

鑑賞者が所望のシーンを探す際には、通常ビデオプレイヤーの SEEK BAR、早送りや巻き戻しのボタンを何度も操作するが、これらの機能は映像の内容を考慮していない。そこで、特定のルールに従って、映像を時間的に分割する手法が多く提案されている。時間的映像分割手法は、鑑賞者が望むシーンへの映像のスキップを実現する、効率的な鑑賞のための基礎的な手法である。この時間的分割のために、同一のカメラで連続的に撮影される「ショット」単位で分割する手法が提案されている [5], [8]。ショットによる分割は効率的な時間的分割ではあるが、より効率的な映像鑑賞のためには試合の流れを考慮する必要がある。特に、

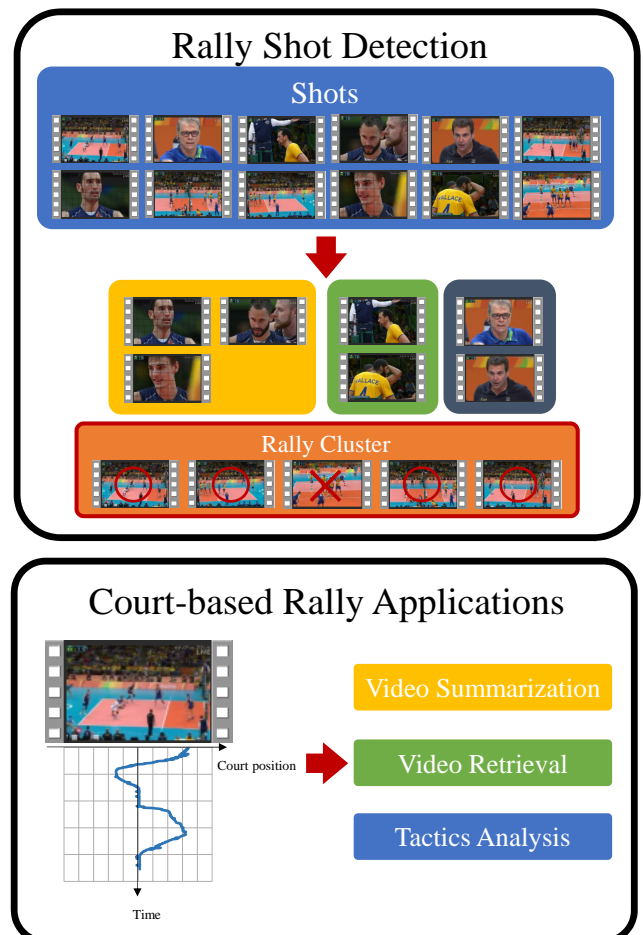


図1 提案手法の概要

バレーボールはラリーによる得点に従って試合が進行するため、「ラリー」単位で分割するのが適切である。近年、ラリーシーン検出手法がいくつか提案されている [3], [6]。Liu らはラケットスポーツ映像に対して、教師なしショッ

¹ 早稲田大学
Waseda University, 169-8555, Tokyo, Japan
² 早稲田大学理工学術院総合研究所
Waseda Research Institute for Science and Engineering
169-8555, Tokyo, Japan
³ JST ACCEL
^{a)} s132800732@fuji.waseda.jp
^{b)} tsukasa@moegi.waseda.jp
^{c)} wasedayshugo@suou.waseda.jp
^{d)} shigeo@waseda.jp

トクラスタリングとサポートベクタマシンによる音声情報の学習を用いた。この手法は高い検出精度を示すが、動画ごとに冒頭の三十分程度の音声情報に対してラベリングする必要があるため、実用性に欠けるといった問題があった。この問題を解決するため、Kawamuraらは白線検出を用いた学習なしかつ全自動でラリーシーン検出を実現した。これは、ラリーシーンが固定カメラで撮影されると仮定し、ショットの平均画像に対してコート上の白線が検出されたショットをラリーシーンとする手法である。しかしながら、バレーボール映像におけるラリーシーンには、パンによるカメラ移動が多く含まれる。したがって、Kawamuraらの手法をバレーボールに直接適用することは困難である。一般に、バレーボールはコート上で競技が行われるため、ラリーシーンにはコートが撮影される。そこで、我々はKawamuraらの仮定を用いず、ショット内のフレームに対してコートを検出することによってラリーシーン検出を行う。

鑑賞支援として、映像要約や映像検索に関する研究が多く行われている。映像要約は、映像の内容を可能な限り失うことなく、概要を生成することを目的としている。Hanjalicは、動き量、カット密度、音量を用いて映像の盛り上がり量を定義し、盛り上がり量が大きくなる映像部分を切り取り要約映像を生成した [2]。この手法は汎用性が高い反面、試合内容を考慮した要約が困難である。そこで、試合内容を取得するために、各スポーツに特化したイベント検出手法が提案されている [4], [10]。一般的にイベント検出手法は選手やボールを追跡した結果を用いている。しかしながら、これらの情報はカメラの移動速度や映像の解像度、フレームレートに依存することに加えて、特にバレーボール映像では選手同士や選手とボールが激しく交錯するために、選手やボールを追跡することは困難である。そこで我々はカメラの動きに着目する。カメラは注目領域を映すように操作されるため、バレーボールのような球技においては、ボールが映像中に映るように撮影される。したがって、我々はボールの大域的な動きをカメラの動きとみなすことができ、カメラの動きの情報によって試合内容を取得可能であると考えられる。提案手法ではカメラの動きとして、画像中のコート中心の推移を用いる。これを実現するために、我々はFarinらのコート検出手法 [1] を改良し、ノイズに頑健なアルゴリズムを提案する。このように提案手法はコート情報のみを使用するため、コートモデルを変更することで他のスポーツへの適用が可能となっている。また映像検索において、Xuらは映像のアノテーションデータを用いて、テキストを入力した映像検索手法を提案している [9]。しかしながら、スポーツの専門的知識がない場合や、アノテーションデータが不十分な場合では、十分な映像検索は困難である。我々の手法は、先程のコートの推移情報の類似度に基づいて、専門的な知識を必要と

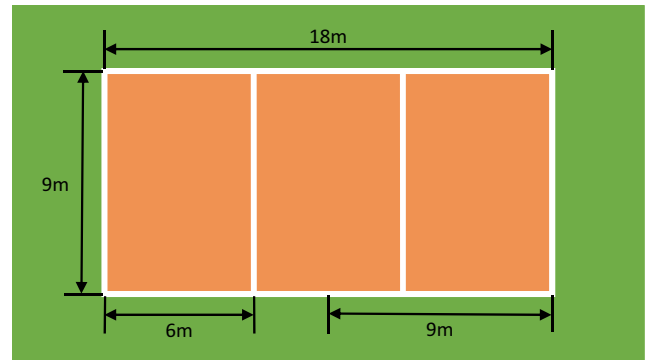


図2 コートモデル

せず、映像自体をクエリとした検索が可能である。

また近年では、鑑賞支援に加えて、戦術解析が注目されている。戦術解析によって、自分のチームや対戦相手の分析が盛んに行われ、そこで得られた情報がテレビで多く放送されている。しかしながら、戦術解析は選手やボールの追跡結果を用いることが多く、上述した理由から、映像に対して手付けで情報を付加しているのが現状である。そこで、我々は先述のコート検出手法を応用し、戦術解析や情報付加を行う手法を検討する。

2. 提案手法

2.1 コート検出と追跡

バレーボールのコートは平面であるため、世界座標系から画像座標系への変換はホモグラフィ変換によって可能である。ここでは、実際のコートと同じ縮尺のコートモデル (図2) から画像座標系の変換を考える。ホモグラフィ変換行列は 3×3 の行列で記述され、コートモデル上の点 \vec{p} から、画像座標系の点 \vec{p} に射影される。同次座標系で記述すると、変換式 $\vec{p}' = H\vec{p}$ は次のようになる。

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} \quad (1)$$

同次座標系はスケールに対して不変であるため、ホモグラフィ変換行列の自由度は8になる。8つのパラメータを求めるには、コートモデルと画像の4点の対応が必要である。この対応点として特徴的なコートのコーナーを使用することが挙げられるが、選手によるオクルージョンに脆弱である。そこで、提案手法ではコートの線を検出し、それらの交点を用いることで、選手のオクルージョンに頑健である。さらに、画像中にある交点を利用できるため、より頑健にコートを検出できる。この対応付けでは、組み合わせ最適化によって、コート線候補をコートモデルの線に割り当てる。それぞれの対応付けによって、コートモデル座標系から画像座標系へのホモグラフィ変換行列が求まる。そして、最適な対応付けから得られるホモグラフィ変換によって、画像座標系に投影されるコートモデルをコートと



図3 入力画像

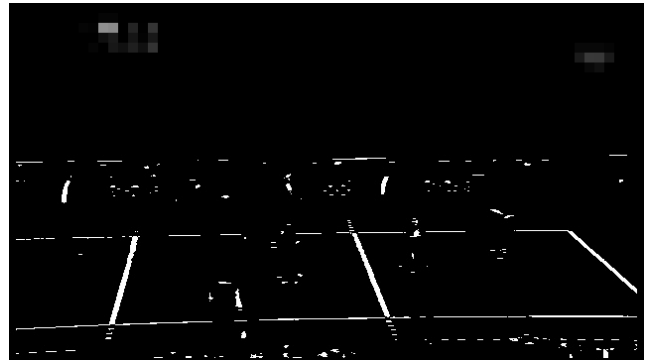


図4 二値画像

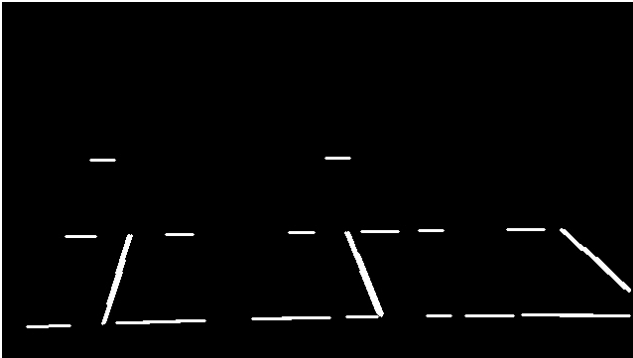


図5 確率的ハフ変換 (1回目)



図6 確率的ハフ変換 (2回目)

して検出する。

2.1.1 コート線候補抽出

まず画像中のコート線候補ピクセルを抽出する。白色領域を抽出したあと、線の幅や長さ、連続性に対して制約を設けることで、コート線候補ピクセルを抽出する。提案手法は、確率的ハフ変換 [7] を反復的に適用することによって、長さや連続性に対する制約を与えることができ、ノイズに頑健にコート線候補ピクセルを抽出することができる。そして、得られたコート線候補ピクセルに対して、ハフ変換を行うことでコート線候補を抽出する。

具体的な手順としては、まず映像の各フレーム (640×360 [pixels]) に対して、輝度値が θ_d 以上となるピクセルを抽出する (本稿では $\theta_d = 180$)。ここでは、公式のバレーボールのコート線が白色であることを利用している。しかし、抽出したピクセルには、会場や観客、選手のユニフォームといったノイズが含まれている。そこで、線の幅や長さ、連続性に対して制約を設ける。線の幅制約として、コート線の幅を細いことに着目し、各ピクセル (x, y) に対して上下のピクセル $(x, y \pm \tau)$ または左右のピクセル $(x \pm \tau, y)$ より輝度値が θ_l 大きいピクセルを残す (本稿では、 $\tau = 4, \theta_l = 20$)。上記の制約を通過したピクセルによる二値画像を生成する (図 4)。次に線分検出手法である確率的ハフ変換を反復的に適用することで、線の長さや連続性の制約を設ける。確率的ハフ変換はパラメータとして、投票数の閾値 T 、長さの閾値 L 、最大間隔の閾値 G を持つ。本稿では、確率的ハフ変換を 2 回適用する。1 回目では、

最大間隔の閾値を小さく設定することによって、連続性の制約をかける (本稿では、 $T = 15, L = 15, G = 3$)。この操作によって非常に小さく、不連続な候補ピクセルの除去が可能である (図 5)。2 回目では、長さの閾値を大きく設定する (本稿では、 $T = 70, L = 100, G = 60$)。この操作によって、同一直線上に存在する線同士を接続して長い線を取得し、一方で、独立する短い線は除去される (図 6)。

最後に、ハフ変換によって直線を検出する。しかしながら、ハフ変換は、入力画像上で太い線に対して、複数本の近接した直線が検出される。この問題を解決するために、解像度の高い投票空間でハフ変換を行った後に、投票数の極大値を取る線を抽出する。極大値を選択する手順としては、(1) ランダムに線の一つを選択し、(2) 投票空間上で近傍を探索する (本稿では、近傍範囲を $\Delta d \leq 70 \wedge \Delta \theta \leq 40$ [degree] と設定した。ただし、 Δd と $\Delta \theta$ はそれぞれ原点からの距離の差分と角度の差分である。)、(3) 極大値を取る線以外を除外し、(4) 極大値を取る線のみが残るまで (1) から (3) を繰り返す。極大値を取る線を抽出した後、それぞれの線を RANSAC によって精錬する。

2.1.2 コートモデルフィッティング

コートフィッティングでは、得られた直線の交点を利用し、コートモデル座標系と画像座標系のホモグラフィ変換行列を算出する。コート線は垂直に交わる線しか存在しないため、まず得られた直線を縦線と横線に分類する。直線の法線を (n_x, n_y) とし、 $|\tan^{-1}(n_y/n_x)| < 35$ [degree] となる直線を横線とし、そうでないものを縦線とする。次に、検

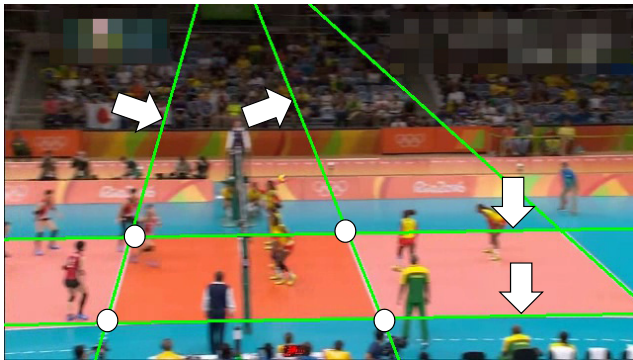


図7 直線の選択と交点の算出

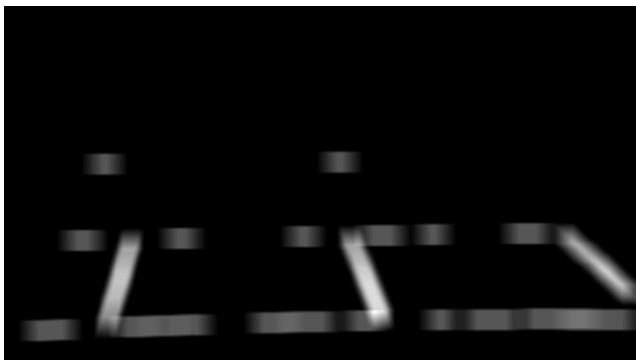


図8 スコアマップ

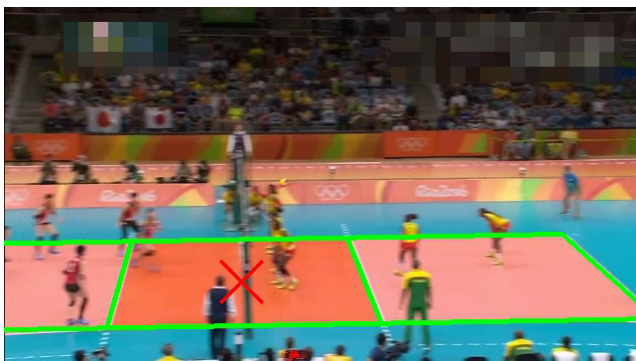


図9 コートモデルフィッティング

出した直線の縦線と横線からそれぞれ2本ずつ線を選択し、4交点 \mathbf{p} を算出する。同様にコートモデル座標系の4交点 \mathbf{p}' を取得し、画像座標系とコートモデル座標系の対応点を4つ得る。その4つの対応点から次式によって、ホモグラフィ変換行列を算出する。

$$\begin{pmatrix} 0 & \mathbf{p}^T & -\mathbf{p}'^T \mathbf{p}^T \\ \mathbf{p}^T & 0 & -\mathbf{p}'^T \mathbf{p}^T \end{pmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \quad (2)$$

ただし、 \mathbf{p} と \mathbf{p}' は同次座標系で記述し、

$$\mathbf{h}^i = \begin{pmatrix} h_{i0} \\ h_{i1} \\ h_{i2} \end{pmatrix} \quad (3)$$

である。得られたホモグラフィ行列によってコートモデル

を画像上に投影し、スコア（重なり度）を算出する。スコアの算出のために、確率的ハフ変換を1回適用した画像に対してガウシアンフィルタをかけた画像を生成し、これをスコアマップとする（図8）。スコアは次式で算出する。

$$\text{Score} = \sum_i SM(\mathbf{p}_i'') \quad (4)$$

SM はスコアマップ、 \mathbf{p}_i'' は投影されたコートモデルが通るピクセルである。スコアが最大かつ閾値以上となるホモグラフィ変換行列によって投影されたコートモデルを、コートとして検出する（図9）。

2.2 ラリーショット検出

本節では、学習を用いない全自動ラリーショット検出を行う。ラリーシーンにコートが撮影されることに着目し、コートが検出されるショットをラリーショットとして検出する。

2.2.1 ショットクラスタリング

入力動画の全フレームに対する HSV ヒストグラム ($H \times S \times V = 16 \times 4 \times 4$) に基づいて、ショット分割を行う。次に、Zhang らの教師なしショットクラスタリング手法 [11] を適用する。この手法は、クラスタ数を自動で決定することができる。概要としては、(1)HSV ヒストグラムに基づくショット（またはクラスタ）間の距離を計算し、(2)最近傍の2つのショット（またはクラスタ）を統合する。(3) 後述する基準に到達するまで、(1)と(2)を繰り返す。

具体的な手順を以下に示す。 i 番目のショットに対して等間隔にキーフレーム $k_l^i (l = \{1, L\})$ を取得する（本稿では、 $L = 13$ ）。HSV ヒストグラムに基づくショット（またはクラスタ）間の距離 SD を次式で定義する。

$$SD(s_i, s_j) = \sum_{n=1}^N \text{sort}_n \min \left(d(k_{l_1}^i, k_{l_2}^j) \right) \quad (5)$$

$$d(k_{l_1}^i, k_{l_2}^j) = \frac{1}{256} \sum_{b=1}^{256} |H_{l_1}^i(b) - H_{l_2}^j(b)| \quad (6)$$

s_i と s_j はそれぞれ i 番目と j 番目のショット（またはクラスタ）、 $\text{sort}_n \min(\cdot)$ は n 番目に小さい値を示す。本稿では、 $d(k_{l_1}^i, k_{l_2}^j)$ は、キーフレーム $k_{l_1}^i$ と $k_{l_2}^j$ の HSV ヒストグラムの距離、 H_l^i はキーフレーム k_l^i の HSV ヒストグラム、 b はヒストグラムのビンである。上述の定義におけるショット（またはクラスタ）間距離で最近傍となる2つのショット（またはクラスタ）を統合し、新しいクラスタ n とする。新しいクラスタ n におけるキーフレームのヒストグラムを次のように更新する。

$$H_l^n = \frac{N_i \times H_l^i + N_j \times H_l^j}{N_i + N_j} \quad (7)$$

N_i と N_j はそれぞれ i 番目と j 番目のクラスタのショットの数である。以上の操作を反復することでクラスタリング

を行う。反復回数は次式における E_n が最小となるように決定する。

$$E_n = J_n + k_n \quad (8)$$

$$J_n = \frac{\sum_{c=0}^{K_n} J_w^c}{J_{inter}} \quad (9)$$

$$= \frac{\sum_{c=0}^{K_n} \sum_{i=0}^{N_c} |H_{L/2}^i - H_{mean}^c|}{\sum_{i=0}^N |H_{L/2}^{i_{min}} - H_{mean}|}$$

$$k_n = \frac{K_n}{N} \quad (10)$$

J_{inter} はクラスタリング前のクラス間分散, J_w^c はクラスタ c のクラス内分散, N と N_c はそれぞれクラスタリング前とクラスタ c のショットの数である。 $\|\cdot\|$ はユークリッド距離, $H_{L/2}^{i_{min}}$ と H_{mean} はそれぞれクラスタリング前の i 番目のショットの中心キーフレームのヒストグラムと平均ヒストグラム, $H_{L/2}^i$ と H_{mean}^c はそれぞれ c 番目のクラスタの中心キーフレームのヒストグラムと平均ヒストグラムである。 k_n はクラスタリング前のショットの総数とクラスタの数 K_n の割合を示す。 図 10 はクラスタ数に対する評価関数 E_n , J_n , k_n の推移を示す。 クラスタ数が減少するに従って, J_n は増加し, k_n は減少する。 E_n が最小となるとき, クラスタの統合を終了し, クラスタ数を決定する。 クラスタリングによって HSV ヒストグラムが類似したクラスタが得られ, 図 1 のようにラリーやリプレイ, 選手のズームといったクラスタが生成される。

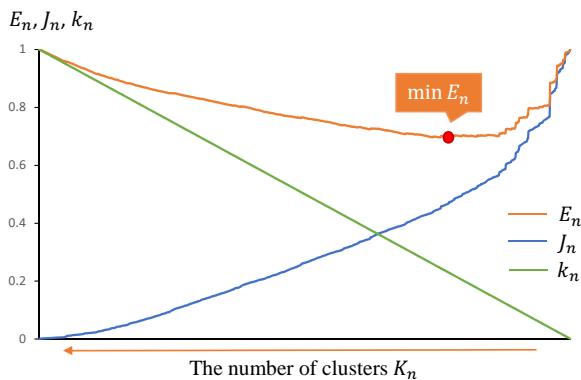


図 10 クラスタ数に対する評価関数 (E_n , J_n , k_n) の推移

2.2.2 ラリークラスタの決定

ショットクラスタリングで得られたクラスタ群の中から, ラリーショットが含まれるクラスタを決定する。 ラリーシーンにはコートが撮影されることに着目し, コート検出を行う。 またバレーボール映像では, センターラインの延長線上から撮影されることを利用し, 横向きのコートのみを検出する (図 11)。 ここでは, それぞれのクラスタの全ショットの中心フレームに対して, コート検出を行う。 コート検出率が閾値 (本稿では 50%) 以上となるクラスタをラリークラスタと決定する。

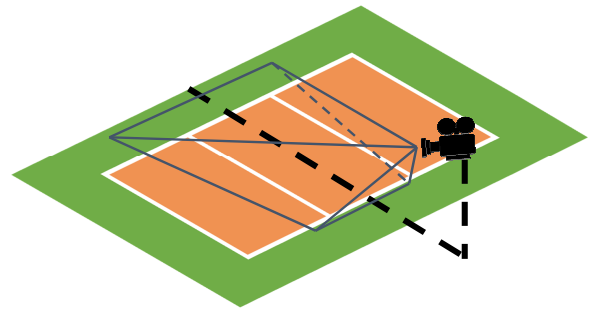


図 11 カメラ位置の仮定

2.2.3 ラリーショットの決定

Kawamura らは, ラリークラスタに含まれる全てのショットをラリーショットとして検出した [3]。 しかしながら, クラスタは HSV ヒストグラムの類似度によって生成されているため, ラリークラスタには HSV ヒストグラムが類似したラリーシーンでないショットが含まれる。 そこで, 我々はラリークラスタを決定した後, ラリークラスタ内の全ショットの複数フレームに対して, 再度コート検出を行う。 そして, コート検出率が閾値 (本稿では 50%) 以上となるショットをラリーショットと決定する。

2.3 鑑賞支援

我々は, 映像要約や映像検索に必要な試合内容を取得するために, ラリーシーンにおけるカメラの動きに着目する。 カメラは注目領域を映すように操作されるため, バレーボールのような球技においては, ボールが映像中に映るように撮影される。 したがって, 我々はボールの大域的な動きをカメラの動きとみなすことができ, カメラの動きの情報によって試合内容を取得可能であると考え。 バレーボール映像におけるラリーシーンはカメラパンによって撮影されているため, カメラの動きとして画像中のコートの中心位置 (x, y) の x 座標の推移を用いる。 コートの推移のグラフの例を図 12 に示す。 縦軸はコートの中心位置の x 座標であり, 画像の幅で正規化している (画像の左端が -1 , 右端が 1)。 横軸は時間である。

2.3.1 映像要約

各ラリーショットに対して, コートの推移の情報に基づいて, ラリーの重要度を定義する。 r 番目のラリーに対して, 重要度 I_r を次式で定義する。

$$I_r = \alpha m_r + \beta v_r + \gamma l_r \quad (11)$$

ただし, m と v はそれぞれコート中心の総移動量と平均移動速度, l はラリーショットの時間長さである。 また α と β と γ は重み係数である (本稿では, $\alpha = \beta = \gamma = 1.0$)。 ラリーの重要度に対して, 閾値を設定し, 閾値以上の重要度のラリーシーンのみを含む映像を要約映像として生成する。 各ラリーにおける重要度の例を図 13 に示す。

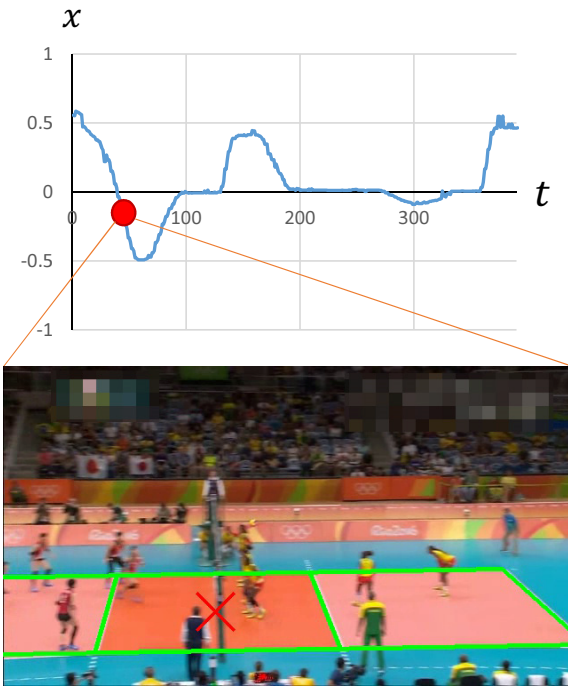


図 12 コートの中心位置の x 座標の推移.

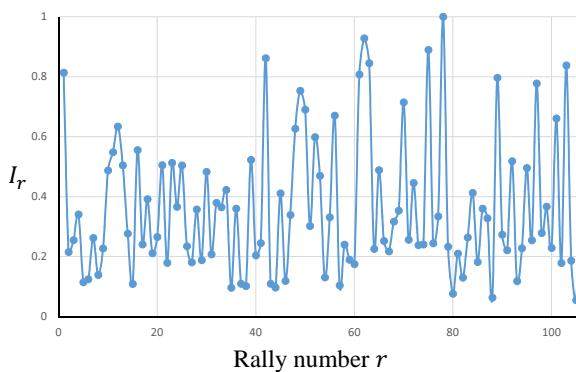


図 13 r 番目のラリーの重要度.

2.3.2 映像検索

ラリーショットを入力として、コートの推移が類似したラリーショットを検索する。ラリーショット間のコスト関数 E を次式で定義する。

$$E = \sum_t |x_i(t) - x_j(t)| + w\delta t \quad (12)$$

ただし、 x_i と x_j は i 番目と j 番目のラリーショットのコートの推移、 t は時間、 δt は2つのラリーショットの時間長さの差分、 w は重み係数である。このコスト関数 E の小さい方から順番に、類似するラリーショットとして出力する。

2.4 戦術解析への応用の検討

戦術解析に必要とされる情報は、手作業で付加されているのが現状である。そこで、我々はコートの推移の情報を用いて、戦術解析のための情報付加の支援及び戦術解析の自動化を検討する。図 14 と図 15 はコートの推移とある瞬間における映像である。図 14 のようにコートの移動方向

が変わる瞬間において、返球時の映像が現れることがわかる。特に最初の返球はサービスで、最後の返球が得点シーンであることがわかる。バレーボールにおいては返球時にサービス、スパイク、ブロックといったイベントが発生するため、いイベント情報を付加する瞬間に映像を素早く合わせるができ、情報付加の支援に繋がると考えられる。また図 15(a) のように、コート位置の絶対値が大きな値を取る時は、ボールが速くにはじき出されているシーンが含まれ、図 15(b) のようにコート位置が 0 付近で留まっているときは、ネット付近でプレイが行われているシーンが含まれる。これらは通常のラリーとは異なった情報を含んでおり、ボールを粘り強く繋いでいるかや攻撃に繋がっているかといったことがわかる。またコートの推移の瞬間的な速度から、サービスやスパイクのおおよその速度が推測される。以上のように戦術の解析に利用可能な情報がコートの推移の情報から多く抽出できることがわかり、戦術解析の自動化に繋がると考えられる。

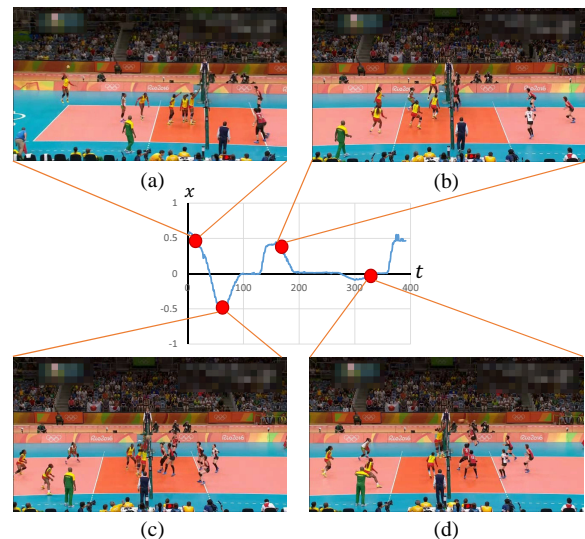


図 14 コートの推移とイベントシーン. (a) サービスシーン, (b)(c) 返球シーン, (d) 得点シーン

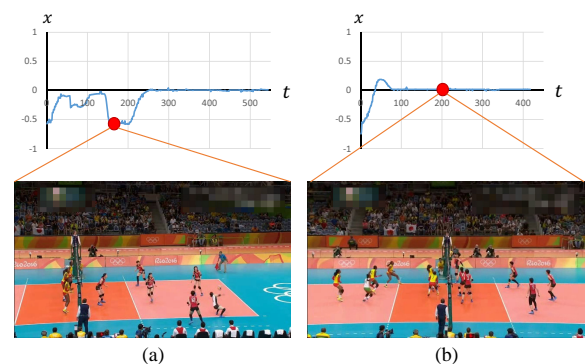


図 15 コートの推移と特殊なシーン (a) ボールが大きくはじき出されたシーン (b) ネット付近でのプレイシーン

3. 結果

3.1 ラリーショット検出

ラリーショット検出の精度評価のために、次式で定義される「Precision」、「Recall」、「F-measure」を用いた。

$$\text{Precision} = \frac{C}{D} \quad (13)$$

$$\text{Recall} = \frac{C}{T} \quad (14)$$

$$\text{F-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (15)$$

ただし、 C は検出した正解ラリーショット数、 D は検出したラリーショット数、 T は正解ラリーショット数である。表 1 は提案手法によるラリーショット検出精度を示す。提案手法は、バレーボール映像に加えて、ラケットスポーツ映像においても、Precision, Recall, F-measure は高い値を示している。バレーボール映像において、カメラ位置の仮定を参考に横向きのコートのみを検出した。そのため、縦向きに撮影されたラリーシーンを検出することができなく、全てのバレーボール映像において、Recall より Precision が高い値を示した。ラケットスポーツ映像においては、縦向きのコートのみを検出したが、ラケットスポーツのラリーシーンは全て縦向きに撮影されていたため、Precision と Recall の値に差はなかった。表 2 は、提案手法と既存手法 [3] における検出精度の比較を示す。提案手法では、ラリークラスタから色が類似したラリーでないショットを除去したため、精度が向上した。

表 1 バレーボール映像とラケットスポーツ映像におけるラリーショット検出精度 (P : Precision, R : Recall, F : F-measure)

入力映像	P	R	F
バレーボール 1	1.000	0.924	0.960
バレーボール 2	1.000	0.903	0.949
バレーボール 3	0.993	0.834	0.907
テニス	0.971	1.000	0.985
バドミントン	1.000	0.991	0.995
卓球	1.000	1.000	1.000

表 2 既存手法 [3] とのラリーショット検出精度の比較 (P : Precision, R : Recall, F : F-measure)

手法	入力映像	P	R	F
提案手法	バレーボール	0.998	0.887	0.939
	ラケットスポーツ	0.998	0.955	0.976
既存手法 [3]	バレーボール	0.417	0.887	0.569
	ラケットスポーツ	0.864	0.992	0.924

3.2 鑑賞支援

3.2.1 映像要約

重要度の高いラリーには、長く続くラリーや激しくボールが行き来するラリーが含まれていた。一方で、重要度の低いラリーには、短いラリーが含まれていた。しかしながら、サービスミスとサービスエースが同様の重要度で評価されていたが、サービスエースは試合の流れが大きく変わるプレイであるため、高く評価されるべきである。バレーボールは得点したチームが次のサービスを行うため、次のサービスチームをコート位置から判定することで、サービスエースとサービスミスを区別することができる。表 3 は、入力映像に対する、ラリーシーンのみを含む映像の時間短縮率を示す。入力映像の 30%程度で全てのラリーシーンを視聴可能となった。また重要度に対して閾値を設けることによって映像を要約することが可能となっているため、更なる視聴時間の短縮が可能である。また視聴したい時間に合わせて、閾値を決定することが可能である。

表 3 元映像に対する、ラリーシーンのみを含む映像の時間短縮率

	入力映像 [s]	ラリーのみ [s]	短縮率 [%]
バレーボール 1	6275	785	13.3
バレーボール 2	4823	1095	22.6
バレーボール 3	6743	1514	22.4
テニス	11530	2401	21.3
バドミントン	5522	1344	24.3
卓球	3490	1151	33.0

3.2.2 映像検索

映像検索結果の例を図 16, 図 17, 図 18 に示す。これらの例 (通常のラリー, サービスエース, ネット付近でのプレイ) は特に定義したコスト関数が小さくなったラリーであり、バレーボールにおいて頻繁に出現するパターンである。一方コスト関数が大きくなったラリーは、非常に長いラリーであり、出現頻度の低い特殊なラリーである。

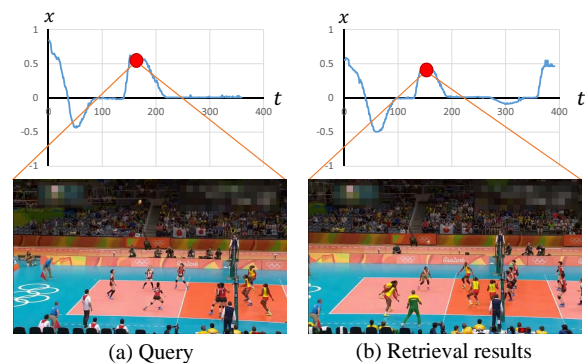


図 16 映像検索結果. 通常のラリー.

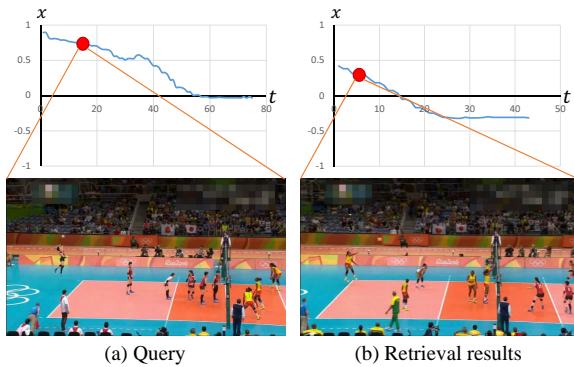


図 17 映像検索結果. サービスエース.

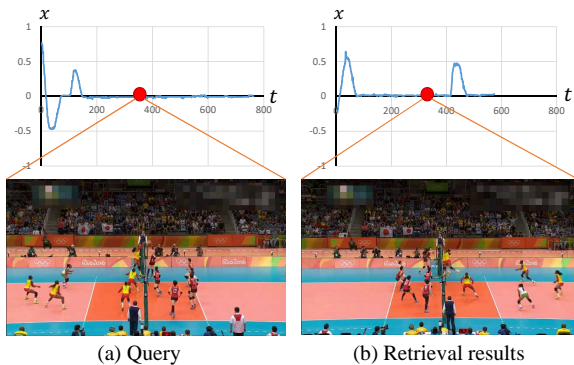


図 18 映像検索結果. ネット付近でのプレイ.

4. まとめと今後の課題

我々は, Farin らの手法 [1] を改良したコート検出を使用して, カメラの動きに頑健なラリーショット検出手法を提案した. その結果, ラケットスポーツ映像に加えて, バレーボール映像における高い精度でラリーショット検出を実現した. また鑑賞支援手法として, 映像要約と映像検索の機能を提案した. 映像要約では, 選手やボールの追跡に基づくイベント検出を行わず, コートの推移の情報を用いることで, 試合内容を考慮した要約映像の生成が可能となった. 映像検索では, テキストではなく映像を入力とし, スポーツに関する専門的な知識を必要としない検索が可能となった. また頻出パターンや出現頻度の低いラリーの知見を得た. 戦術解析への応用の検討では, コートの移動方向が変化する瞬間で返球が行われており, 情報を付加する際の支援が考えられる. またコートの推移から得られる情報を用いることで, 自動戦術解析に繋がると考えられる.

一方, 今後の課題としては, コートの情報のみだけではなく, 得点情報や背景差分に基づく選手領域の概形から得られる情報を用いて, より高度な鑑賞支援や戦術解析を行いたい. また映像要約における, ラリー重要度の重み係数を重回帰分析によって, 重要度への寄与が大きい要素を明らかにすると共に, より視聴者の関心に沿った要約映像の生成を目指したい.

謝辞 本研究の一部は, JST ACCEL 及び JST CREST の

支援を受けた.

参考文献

- [1] Farin, D., Krabbe, S., Effelsberg, W. and de With, P. H.: Robust Camera Calibration for Sport Videos using Court Models, *Proceedings of SPIE Storage and Retrieval Methods and Applications for Multimedia*, Vol. 5307, pp. 80–91 (2004).
- [2] Hanjalic, A.: Generic approach to highlights extraction from a sport video, *Proceedings of International Conference on Image Processing (ICIP2003)*, Vol. 1, pp. 1–4 (2003).
- [3] Kawamura, S., Fukusato, T., Hirai, T. and Morishima, S.: RSViewer: An Efficient Video Viewer for Racquet Sports Focusing on Rally Scenes, *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP2016)*, Vol. 2, pp. 249–256 (2016).
- [4] Li, B. and Sezan, I.: Event detection and summarization in American football broadcast video, *Proceedings of SPIE Storage and Retrieval for Media Databases*, Vol. 4676, pp. 202–213 (2002).
- [5] Lian, S., Dong, Y. and Wang, H.: Efficient temporal segmentation for sports programs with special cases, *Proceedings of the 11th Pacific Rim Conference on Advances in Multimedia Information Processing - PCM 2010 PART1*, Vol. 6297, pp. 381–391 (2010).
- [6] Liu, C., Huang, Q., Jiang, S., Xing, L., Ye, Q. and Gao, W.: A Framework for Flexible Summarization for Racquet Sports Video Using Multiple Modalities, *Computer Vision and Image Understanding*, Vol. 113, No. 3, pp. 415–424 (2009).
- [7] Matas, J., Galambos, C. and Kittler, J.: Robust detection of lines using the progressive probabilistic hough transform, *Computer Vision and Image Understanding*, Vol. 78, No. 1, pp. 119–137 (2000).
- [8] Singh, R. D. and Aggarwal, N.: Novel research in the field of shot boundary detection - A survey, *Advances in Intelligent Informatics*, pp. 457–469 (2015).
- [9] Xu, C., Wang, J., Lu, H. and Zhang, Y.: A novel framework for semantic annotation and personalized retrieval of sports video, *IEEE Transactions on Multimedia*, Vol. 10, No. 3, pp. 421–436 (2008).
- [10] Zawbaa, H. M., El-Bendary, N., Hassanien, A. E. and hoon Kim, T.: Event detection based approach for soccer video summarization using machine learning, *International Journal of Multimedia and Ubiquitous Engineering (IJMUE)*, Vol. 7, No. 2, pp. 1–18 (2012).
- [11] Zhang, W., Ye, Q., Xing, L. and Gao, W.: Unsupervised sports video scene clustering and its applications to story units detection, *Proceedings of SPIE Visual Communications and Image Processing*, Vol. 5960, pp. 446–455 (2005).