

## 多視点映像視聴支援のためのマルチセンサを用いた

## 視点選択手法の検討

Study on selection view method with wearable multi-sensor for Multi-Video movie

角谷 昇一郎<sup>\*1</sup> 北出 卓也<sup>\*1</sup> 榎堀 優<sup>\*1</sup> 丸谷 宜史<sup>\*1</sup> 梶田 将司<sup>\*2</sup> 間瀬 健二<sup>\*1</sup>  
 Shoichiro Sumiya Takuya Kitade Yu Enokibori Takafumi Marutani Shoji Kajita Kenji Mase

<sup>\*1</sup> 名古屋大学大学院情報科学研究科<sup>\*2</sup> 名古屋大学情報連携統括本部

Graduate School of Information Science, Nagoya University Information and Communication Technology Services, Nagoya University

In recent year, movies taken from individual viewpoints are used for sports coaching, e.g., soccer and baseball because such multi-video content is necessary for such a wide area game like soccer to capture the whole wide field. However, it is troublesome to adequately choose a camera view from multiple sources. On the other hand, owing to the recent spread of smartphones with multi-sensor, it has become possible and easy to obtain position, acceleration, and direction of sports players by attaching smartphones them. In this study, we aim to support that players investigate their plays instantly with a multi-video soccer movie using the data acquired from sensors of smartphones attached to the players. In order to automate the selection of viewpoint, we examine wearable sensor data such as position or orientation data of the players.

## 1. はじめに

近年、ビデオカメラが安価に入手できるようになり、アマチュアのスポーツチームでも複数のカメラを使ってサッカーのような広い範囲を同期撮影して、自らプレーをチェックできるようになってきた。しかしながら、複数カメラの撮影された映像から自分が必要とするシーンを連続して視聴することは困難である。そのため、視点の選択の自動化は、多視点映像の視聴支援に繋がる。永井ら[永井 2010]は Quality of View (QoV) [Shen 2007]の考え方にに基づき、選手の位置情報や向きの情報を元に視点の評価値を計算し、適切な視点を自動選択する手法を提案している。また、丹羽ら[丹羽 2010]は、複数人の視聴ログより大衆の興味を抽出し、それにより視点選択をすることで映像ストリームを作成する手法を提案している。しかし、これらの研究では選手の位置や方位などのデータを手動で入力する必要があり、多大な労力が必要であった。一方で、現在ではスマートフォンなどに組み込まれるモーションセンサが充実してきたため、センサを着用してプレーし、そのデータを用いて多視点映像のアノテーションをする方法の可能性が広がってきた。センサを用いる利点として、選手の位置情報や向き情報の取得が容易であること、センサデータから選手行動を解析することで映像に自動でアノテーションを付加できることが挙げられる。そこで、本稿では、サッカー多視点映像を題材に、マルチセンサデータを用いて視聴対象選手が映っている視点の自動選択を行う手法を提案する。また、本手法の有効性を検証するための予備実験の結果を報告する。

## 2. 視点選択手法

本章では多視点映像において、視聴対象選手が映っている視点を、センサデータを利用して自動で選択する手法を提案する。視点選択の流れとして、まず各視点に対応するカメラ画像中に視聴対象が存在するかの判断を行う。3.1 節でその手法を説明する。次に対象が存在する各カメラ画像に対して、3.2 節で記す評価尺度をそれぞれ計算し、最終的に表示する視点を決定する。以下、カメラや選手の座標はフィールドの中心を原点と

し、ゴールとゴールを結ぶ直線を  $x$  軸とする世界座標系で表す。

## 2.1 カメラ内の対象存在判断

本節では視聴対象選手がある時刻にあるカメラに映っているかを判断するための手法について述べる。

まず、多視点映像におけるそれぞれのカメラ  $Cam_i$  について、カメラがカバーするフィールドの領域(以下カメラカバー領域)を4点  $A_i, B_i, C_i, D_i$  とする(図 2)。これは、カメラに写っているフィールドの画像点4点を透視投影変換しフィールド上の世界座標系の点に変換することで得られる。時刻  $t$  におけるプレーヤー  $j$  の位置  $P_j(t)$  とし、ベクトル  $\vec{b}, \vec{c}, \vec{p}$  を以下のように定義する。

$$\vec{b} = A_i B_i, \quad \vec{c} = A_i C_i, \quad \vec{p} = A_i P_j[t]$$

ここで、

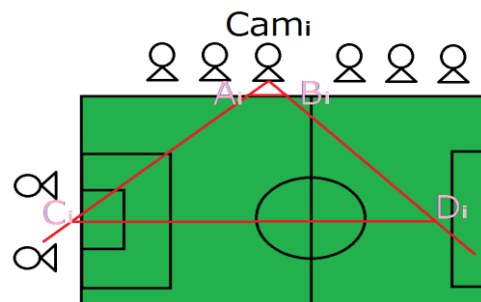
$$\vec{p} = s\vec{b} + u\vec{c}$$

を満たす実数  $s, u$  を連立 1 次方程式を解くことで求める。ここで求めた  $s, u$  が以下の条件を満たすとき、 $P_j(t)$  は三角形  $A_i B_i C_i$  の内部に存在する。

$$s \geq 0, u \geq 0, s + u \leq 1$$

同様に  $P_j(t)$  が 3 点に囲まれた三角形  $B_i C_i D_i$  の内部に存在するかの判定を行い、どちらかの三角形の内部に存在した場合  $P_j(t)$  は台形  $A_i B_i C_i D_i$  の内部に存在する。つまりカメラカバー領域  $A_i B_i C_i D_i$  に存在するかどうかを判定できる。ここで、 $f_{ij}(t)$  を以下のように定義する。

$$f_{ij}(t) = \begin{cases} 1: P_j(t) \text{ が台形 } A_i B_i C_i D_i \text{ の内部に存在} \\ 0: \text{ otherwise} \end{cases}$$



連絡先: 角谷昇一郎, 名古屋大学大学院情報科学研究科,  
sumiya@arch.nagoya-u.ac.jp

図 1 カメラ Cam<sub>i</sub> のカバー領域

$T$  : 設定上限フレーム数

## 2.2 選択基準尺度

一般的に視聴対象選手が映っている視点は複数存在する。その中でどの視点が視聴者にとって最適な視点なのかを判断するための選択基準尺度を考える。Shen ら[Shen 2010]は、複数カメラからの視点選択の際に用いる評価尺度として QoV を定義した。これは被写体の距離と角度により計算される。さらに永井[永井 2011]は QoV に視聴者の興味をモデル化して一般化した QoV を提案している。本研究ではこの考え方を援用して、距離尺度、正面尺度を定義した上で、新たに映像としての評価尺度として連続尺度を追加した。いずれも値域は  $[0,1]$  である。以下ではそれぞれの尺度について説明する。

### (1) 距離尺度

この尺度は選手の画像中の見かけの面積を考慮した尺度である。一般にカメラに映る選手の縦横比は、選手とカメラ間の距離に反比例する。そこでカメラ  $i$  の時刻  $t$  における距離尺度値  $Dist_{ij}(t)$  を以下のように定義する。この値が大きいほど選手が大写しになっている。

$$Dist_{ij}[t] = \frac{\min_k dis(p_j[t], c_k)}{dis(p_j[t], c_i)}$$

$p_j(t) = (p_{jx}(t), p_{jy}(t))$  : 時刻  $t$  における選手  $p$  の座標

$c_i (i_x, i_y)$  : カメラ  $i$  の座標

$dis(a, b)$  : 2 点  $a, b$  のユークリッド距離

### (2) 正面尺度

この尺度は、あるカメラに対して選手の体がどれだけ正面を向いているかを、腰部のセンサにより得られた方位データを用いて数値化した尺度である。時刻  $t$  における選手の向きを  $\vec{d}$  として、カメラ  $i$  の時刻  $t$  における正面尺度値を以下のように定義する。

$$Front_{ij}[t] = \frac{|\theta_{ij}(t)|}{\pi}$$

$$\theta_{ij}(t): \vec{b} = \overrightarrow{p_j(t)c_i} \text{ と } \vec{d} \text{ のなす角 [rad]}$$

### (3) 連続尺度

この尺度は、あるカメラの映像中に選手がどれだけ連続して存在しているかを数値化した尺度である。プレーの最中に頻繁にカメラが切り替わるとプレーを集中して見ることの妨げになってしまう。そこで、連続尺度を導入することにより、各カメラで連続して対象が映っている度合いを評価する。時刻  $t$  におけるカメラ  $i$  の連続尺度値を以下のように定義する。なお、本研究では  $T=50$  とし 5 秒 (=50 フレーム) の間で対象がフレームアウトするまでのフレーム数の割合を評価している。映っているかの判定は 3.1 節の方法を用いる。

$$Cont_{ij}[t] = \begin{cases} \frac{g_{ij}(t)}{T} & (g_{ij}(t) \leq T) \\ 1 & (g_{ij}(t) > T) \end{cases}$$

$$g_{ij}(t) = t_{ij} - t - 1$$

$t_{ij}$  :  $t < t', f_{ij}(t') = 0$  を満たす最少の  $t'$

以上で定義した 3 つの評価尺度値の線形和として、選手  $j$  を映すカメラ  $i$  の、時刻  $t$  における総合評価値  $V_{ij}(t)$  を決定する。

$$V_{ij}(t) = w_1 Dist_{ij}[t] + w_2 Front_{ij}[t] + w_3 Cont_{ij}[t]$$

この値が最大となる視点を選択することで、視点の自動選択を実現する。

## 3. 予備実験

実際に収集したデータからインタフェースを作成し、3.2 節で定義した評価尺度から注目選手を追跡する視点の自動選択を行った。

### 3.1 データ収集

モーションセンサを内蔵しているスマートフォンを用いて、時刻、GPS、加速度、方位を常時記録し続けるアプリケーションを実装した。センサのサンプリングレートは 10[Hz] で、値は CSV 形式でファイルに保存される。このスマートフォンを複数の選手の腰後部にバンドでつけてもらいデータを収集した。一方、多視点映像はサッカー場のスタンドにほぼ等間隔に 16 台のデジタルカメラ(カシオ EX-F1)を設置し、無線 LAN による遠隔制御によって同期ビデオ撮影を行った。各カメラは標準でズームレンズを用いて最も広角に設定して撮影した。1 台のカメラはサッカー場の約 1/4~1/2 のみをカバーしており、1 台のカメラ映像から全プレーを確認することは不可能である。

### 3.2 多視点視聴インタフェース

多視点映像において、一つ一つのカメラ映像を見ていくだけではフィールド全体の状況を即座に判断することは難しい。そこで、センサデータを用いてフィールド上の選手の情報を可視化する事でフィールド全体の状況を把握可能にする。本研究では、フィールド上の選手の状況を容易に把握できるインタフェースを作成した(図 1)。図左上の枠では、センサより得られた選手の位置データと方位データを OpenGL により選手を簡単な 3 次元モデルとして表示している。マウス操作で自由な角度・距離からフィールドを俯瞰したり、選手の軌跡を調べたりすることができる。中央の枠内のバーは時間軸の操作とカメラの選択を行うことができ、選択された視点からの実映像は左下の枠に表示される。右の枠内のバーは視聴対象選手の指定を行う。

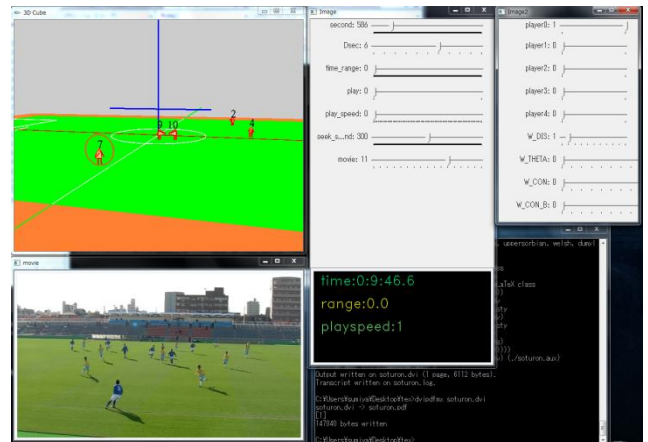


図 2 多視点視聴インタフェース

### 3.3 実験方法

3.2 節で定義した 3 つの評価尺度に対する重みの組を

$$W = (w_1, w_2, w_3)^T$$

とすし、図 3 のように、この値を視聴しながら自由に変えられるバーを作成し、25 秒間の映像を、 $W$  の値を以下のように変えながら視聴を行った。また、その際尺度を評価するための要素の一つとしてカメラの切り替え回数も同時に記録した。

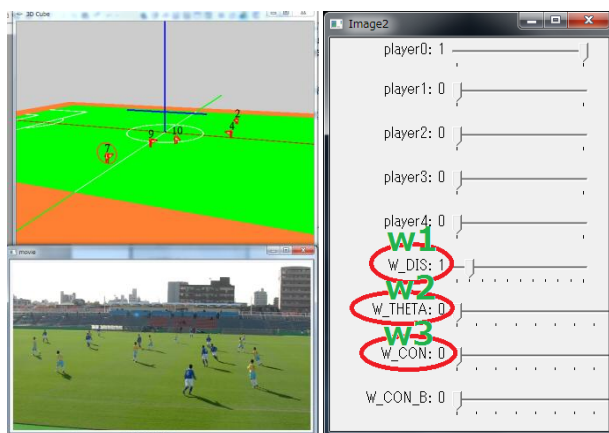


図 3 対象追跡・重み決定バー

### 3.4 結果

まず  $w_1$  の重みを大きくして ( $W=(1.0, 0.2, 0.2)^T$ ) 視聴を行った。選手の動きに合わせて細かくカメラが切り替わったり、時折カメラが瞬時に連続して切り替わる部分はあったが、対象が大きく映る見やすい映像が選ばれた。次に  $w_2$  の重みを大きくした ( $W=(0.2, 1.0, 0.2)^T$ ) と、選手が正面を向いているカメラが選択されたが、頻繁にカメラが切り替わり、対象を追うのが難しく、対象選手のプレーを見るのに支障があった。最後に  $w_3$  の重みを大きくして ( $W=(0.2, 0.2, 1.0)^T$ ) 視聴を行ったところ、カメラの切り替えは比較的少なく、プレーを安定して見ることはできたが、対象選手が遠くに映るカメラを優先して選択することもあった。

表 1 に、各重みのつけ方に対して生成映像 25 秒間の対象追跡時のカメラ切り替え回数を示す。

表 1 カメラの自動切替え回数の変化

表各評価尺度の重み	カメラ切り替え回数
距離尺度重視 ( $W=(1.0, 0.2, 0.2)^T$ )	17
中心尺度重視 ( $W=(0.2, 1.0, 0.2)^T$ )	208
連続尺度重視 ( $W=(0.2, 0.2, 1.0)^T$ )	14

### 3.5 考察

距離尺度は選手とカメラの距離が近い視点を優先して選択するよう定義したものである。距離尺度の重み  $w_1$  を大きくした際、選手が大きく映っている視点が表示され、またカメラ切り替え回数が少なく安定して見ることで生成された映像が生成されたことより、評価尺度として有効に機能していると考えられる。次に、正面尺度は選手が正面を向いて映っているカメラを優先して選択するよう定義したものであるが、正面尺度の重み  $w_2$  を大きくした時、選手が正面を向いたカメラは選ばれたものの頻繁にカメラ切り替えが起こってしまった。サッカープレー中では選手が頻繁に方向転換を行うため、評価値が大きく揺れることが原因と考えられる。これでは視聴に困難が生じてしまい、現在のままでは有効性があるとは言えない。最後に連続尺度は、カメラ切り替えが少なく安定して選手のプレーを見られるカメラを優先して選択するよう定義したものであり、表 1 を見ると、連続尺度の重み  $w_3$  を大きくした時にカメラ切り替え回数が一番少なくなっていることが分かる。このことから、連続尺度の導入により煩雑なカメラ切り替

えを防ぐことができることが分かり、連続尺度の定義は評価尺度として有効に機能していると言える。ただし、カメラの切り替え回数こそ低いものの選手が遠くに映っている画面が優先されて表示されたことから、この尺度の重要度が最も高いとは言えない。遠くに映る選手が優先的に表示されたのは、カメラから遠くにいるということにより、等速で動いていても単位時間における映像中の移動距離が小さくなることで画面からフレームアウトしにくくなることが原因と考えられる。

## 4. おわりに

本研究は、多視点サッカー映像を題材に、マルチセンサを利用することで多視点映像の視聴支援を行うことを目的とし、視聴対象選手が映っている視点の自動選択手法を提案した。そのためにまずカメラ選択の評価尺度として距離尺度、正面尺度、連続尺度の 3 つの尺度を定義した。次に、マルチセンサから得たデータを 3 次元空間上に可視化し映像と同期することで、多視点サッカー映像視聴インタフェースを作成し、それぞれの尺度値を実際に収集したセンサデータから計算してカメラ自動切替えを行った。この際、それぞれの尺度に対する重みを変えながら、対象追跡結果にどのような影響を及ぼすかを検証し尺度定義の有効性を考察したところ、正面尺度の有効性は確認できなかったが広域尺度と連続尺度の有効性は確認できた。正面尺度が有効でなかった理由として、選手が頻繁に方向転換を行うことが考えられ、今後の課題としては、選手が頻繁に方向転換することを考慮した正面尺度に代替する尺度の考案や、より適切な視点選択のための評価尺度の重み決定が挙げられる。

## 謝辞

本研究の一部は独立行政法人情報通信研究機構(NICT)の委託研究「三次元映像通信・放送のための中核的要素技術」により行われた。

多視点映像は東工大の藤井研究室の協力を得て収集したので感謝する。

## 参考文献

- [丹羽 2011] 丹羽 宏介, 東海 彰吾, 川本 哲也, 藤井 俊彰, 丸谷 宜史, 梶田 将司, 間瀬 健二, "多視点映像視聴支援の為のソーシャルな視聴履歴の利用", 情報処理学会技術研究報告. EC, エンターテインメントコンピューティング, 2011.
- [永井 2011] 永井 有希, 丸谷 宜史, 梶田 将司, 間瀬 健二, "視聴者の関心を考慮した Quality-of-View に基づく多視点画像の評価", 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, 2011.
- [永井 2011] 永井 有希, 丸谷 宜史, 梶田 将司, 間瀬 健二, "プレーに着目したスポーツ多視点映像の評価尺度", 情報処理学会技術研究報告. EC, エンターテインメントコンピューティング, 2011.
- [間瀬 2011] 間瀬 健二, 永井 有希, 丹羽 宏介, 藤井 俊彰, 丸谷 宜史, 川本 哲也, "視聴者の興味モデリングを用いた多視点映像視聴インタフェース", 情報処理学会 HCI 研究会研究報告, 2011.5.
- [Shen 2007] Shen, C., Zhang, C. and Fels., S.S.: A Multi-Camera Surveillance System that Estimates Quality-of-View easurement, The IEEE International Conference on Image Processing (ICIP), pp. 193-196, (2007).