

# レーザーキャナを用いたサッカー選手の半自動追跡

壁谷 勇磨<sup>1,a)</sup> 富安 史陽<sup>1,b)</sup> 間瀬 健二<sup>1,c)</sup>

**概要:** 近年、テレビ放送やインターネットの動画サイトなどの映像視聴の形式が多様化してきている。これを生かした視聴システムの研究が盛んに行われており、特にサッカーは人気が高く、研究対象としてよく挙げられている。これらの研究は、サッカー選手の移動情報を用いるものが多いが、画像特徴からこれを正確に推定するのは困難である。本研究では、レーザーキャナを用いたサッカー選手の追跡手法を提案する。レーザーキャナの測定情報を基にするため、正確な選手位置情報を得ることが出来る。手法は半自動的であり、多くの追跡処理を自動で行い、一部の修正を手動で行う。手動修正は、直感的なインタフェースを操作して行う。提案手法によってデータ処理実験を行った。その結果、正しい選手の移動軌跡を得ることが出来た。また、すべての処理のうち手動操作が必要な割合を 1.3% に抑えることが出来た。

## Semi-automatic Soccer Player Tracking by Laser Scanner

YUMA KABEYA<sup>1,a)</sup> FUMIHARU TOMIYASU<sup>1,b)</sup> KENJI MASE<sup>1,c)</sup>

**Abstract:** Formats of video viewing become diverse such as TV, video viewing website and so on. Many video viewing systems have been developed using these formats actively. In those studies, many of them focus on the popular sport, the soccer game. However, it is difficult to exactly estimate the tracking information of soccer players from image features which is important to these studies. Therefore, we propose a soccer player tracking method using laser scanner. The position information of players can be acquired since it is based on the scanned data. Our method is semi-automatic. Most of the tracking procedure runs automatically, only part of correction needs interactive operation. An intuitive interface is provided for the interactive procedure. We carried out an experiment processing the data by our method. As a result, we achieved exact player tracking and limit the interactive procedure at the rate of 1.3% within all procedure.

### 1. はじめに

地上デジタルテレビ放送(以下、地デジ)への移行や、高度な情報通信技術の普及に伴って、視聴者参加型番組などの新たな映像配信の形が実現している。例として、地デジにおけるデータ放送を挙げる。データ放送では、放送局側が番組と連動した情報を配信し、視聴者がそれを自由に参照することができる。また、近年のスマートフォンやSNSの普及によって、携帯電話やEメールの登場した時代に比べて容易かつ多様な遠隔コミュニケーションが可能になっている。これを生かしたテレビ番組も存在し、番組についてのソーシャルネットワークワーキングサービス(以下、SNS)上

の発言を、番組内でリアルタイムに取り上げている。さらにテレビ以外にも、インターネットでの動画配信も様々な形で普及している。配信形式が規定されておらず、視聴者の操作自由度が高いインターネットは、映像と視聴者が連動した映像配信に非常に適している。これまで挙げたように、視聴者が完全に受動的な過去のテレビ放送とは異なり、視聴者側に選択肢があるような映像コンテンツが広がってきている。

配信形式が進化する一方で、スポーツ観戦のような根強い人気を得ている映像ジャンルがある。2020年のオリンピック開催都市が東京に選ばれ、スポーツへの関心は今後も増え続けると考えられる。スポーツの中でも特に人気があるサッカーにおいては、視聴システムの提案や映像分析などの研究が盛んに行われている[1][2][3]。これらの研究では、重要な情報として選手の位置を用いているが、色領

<sup>1</sup> 名古屋大学 大学院情報科学研究科 社会システム情報学専攻

a) kabeya@cmc.ss.is.nagoya-u.ac.jp

b) tomiyasu@cmc.ss.is.nagoya-u.ac.jp

c) mase@nagoya-u.jp

域や形状などの画像特徴をもとに推定している。そのような方法では、オクルージョンや選手の姿勢変化等が大きく影響し、正確な位置が取得できない場合が多い。また、[4][5]のような映像視聴に関する研究では、選手一人ひとりのフィールド上の移動情報が必要である。村松ら [4] は、映像の視点切り替えを視聴者が自由に行えるシステムにおいて、視聴者個人の好みに基づいた視点推薦を提案した。この好みを決定する尺度の中には画面上の対象の位置や大きさが含まれており、サッカー映像の選手については、これらの尺度をフィールド上の移動軌跡を用いて計算している。[5]において我々は、見たい対象をつねに画面の中心に表示する視聴方式を提案した。この対象の画面上の位置は事前に必要な情報であり、サッカー映像における選手を見たい対象とするならば、選手の移動軌跡から画面上の位置を計算する。以上のように、映像視聴に関する研究においては、サッカー映像における選手の移動情報は重要である。しかし、選手の移動情報を得るには、選手の各フレームの位置情報に対して追跡処理などを行う必要がある。

本研究では、レーザスキャナをフィールド付近に設置し、選手のフィールド上の位置を取得する。これによって、フィールド上での選手位置が正確に測定できる。しかし、この方法で測定したデータを選手軌跡として扱うためには、複数点の統合や追跡などの事後処理が必要である。レーザスキャナで人物を追跡した研究として [6][7] が挙げられる。これらの研究における人物追跡手法では、手の位置などの姿勢情報を用いているが、本研究の対象であるサッカー選手に対してはこの情報はうまく測定できないため、この方法を適用するのは困難である。本研究では、サッカー選手の測定データに対する処理アルゴリズムを提案する。処理アルゴリズムは、完全自動でレンジデータに適用すると正確な選手軌跡が得られないので、一部を手動修正する必要がある。本研究では、この手動修正の負担を抑えるための手動インタフェースも併せて提案する。

2 節では、提案する処理アルゴリズムについて説明する。3 節では、提案ツールの手動修正インタフェースについて説明する。4 節では、本ツールを使ってサッカーの試合を測定したデータに対する処理について記す。最後に 5 節では本研究をまとめる。

## 2. レンジデータ処理

初めに、本研究におけるレーザスキャナおよびレンジデータについて定義する。レーザスキャナは、周辺の平面上の有効範囲に存在する点を測定する。レーザスキャナは、測定範囲のある角度にレーザを照射し、物体から反射してくる光の到達時間によってその角度の対象との距離を測定する。図 1 に、レーザスキャナによる測定イメージを示す。レンジデータは、レーザスキャナによって測定した対象の平面上の位置を表す時系列データである。以降、レ

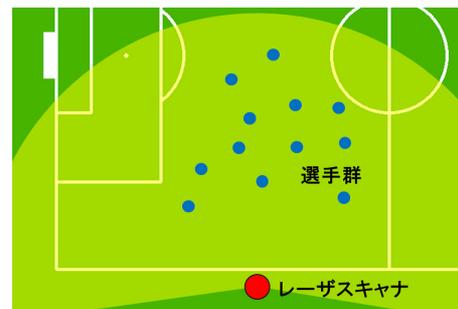


図 1 レーザスキャナによる測定イメージ

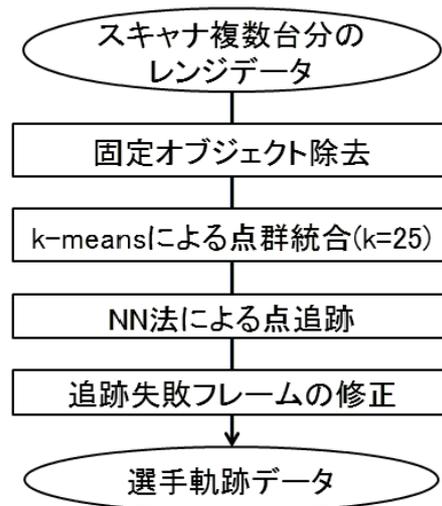


図 2 レンジデータ処理の流れ

ンジデータ中にある一つの位置データを測定点と呼ぶ。各測定点は、直交座標系  $(x,y)$  で表される。

処理アルゴリズムでは、複数台のレーザスキャナで同じサッカーフィールドを測定した複数のデータを扱う。これは、ある角度からではオクルージョンによって測定されなかった選手を、別角度から測定することでデータ欠損を補うためである。

以降、レンジデータ処理の詳細を述べる。データ処理の流れを図 2 に示す。

### 2.1 固定オブジェクト除去

レンジデータには、ゴールポストやコーナーポストなどの固定オブジェクトの測定点も含まれる。そこでまず、固定オブジェクトのレンジデータを除去する処理を行う。固定オブジェクトは静止しており、決まった位置で測定される。この性質を利用して、長期間、同じ位置で測定されている点を削除する。これによって固定オブジェクトの測定点を除去できる。

### 2.2 k-means による点群統合

レンジデータは、測定対象である選手に体表の複数ヶ所の測定点を含んでいる。さらに本手法では、複数のレーザ

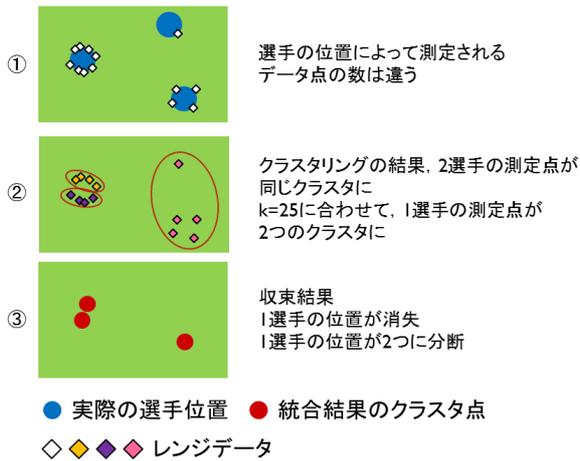


図3 クラスタリングにおける消失および分断の発生過程

スキャナのレンジデータを合和せて扱う。したがって、一人の選手に対して複数の測定点がある。そこでこれを統合し、フレーム毎に選手数分の位置データを得る必要がある。

統合アルゴリズムには  $k$ -means を利用する。これは、測定対象である選手の人数が一定であり、各選手の測定点は密集している、というデータの性質に基づく方法である。 $k$ -means のクラスタ数は両チームの選手数 (22 人) と審判 (3 人) の数を合わせた  $k = 25$  とする。初めに、あるフレームの全ての測定点に  $k$ -means を用い、25 個のクラスタ中心点を得る。次に、実際の測定結果を参照するため、各クラスタ中心の最近傍点を元の測定点群から選び、これらをそのフレームの選手位置とする。以上の処理を全てのフレームに対して行う。

### 2.2.1 分断補正処理

$k$ -means 統合処理によって不正確なデータが得られる現象がある。本項ではこれを修正するための後処理について述べる。現象についての説明と後処理について、図3に示す。

$k$ -means によって処理すると、選手一人に対して一つの位置が得られない場合がある。これは以下の状況で起こる。

選手同士が重なることでスキャナで測定できなくなるオクルージョンにより、ある選手の測定点の数が少なくなる。このような点は、 $k$ -means のアルゴリズム中で他の選手の測定点からなるクラスタに属してしまう。その結果、総クラスタ数を  $k=25$  とするために、別の選手の測定点群を2つのクラスタとしてしまう。このようにして、ある選手位置が消失し、同時に別の選手位置が2つに分断する。

分断した選手位置を1つにするために、後処理を行う。これを分断補正処理と呼ぶ。この処理では、2点の距離が閾値以下となる点の組を分断とみなし、2点の平均を正しい選手位置とする。これによってあるフレームに存在する選手位置の数は25個以下となるが、消失した選手位置は後述の補間処理によって補う。

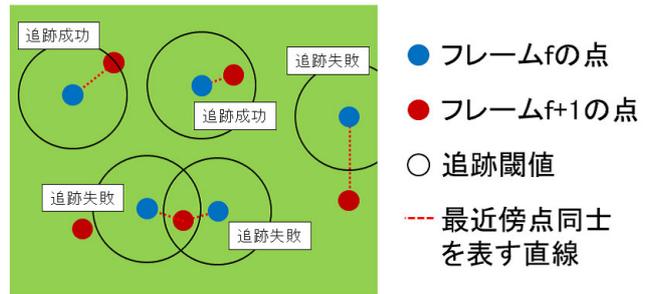


図4 フレーム f での各選手位置の追跡処理

### 2.3 最近傍法 (NN 法) による各選手点追跡

$k$ -means により統合したデータはフレーム間で選手の対応付けがされていないので、目的とする選手軌跡を得るために選手位置に対して追跡処理を行う。あるフレーム  $f$  での各選手位置の追跡処理について、図4に示す。追跡処理は以下のような手順で行う。

- (1) 初期フレームでは25個の選手位置が得られているものとし、各選手位置に任意の選手IDを付ける。
- (2) ある選手位置 (以降、追跡点と呼ぶ) について、次のフレームで最も距離が近い選手位置 (以降、最近傍点と呼ぶ) を求める。
- (3) 追跡点と求めた最近傍点の距離が追跡閾値内であれば、2点を接続する。このとき、閾値外の場合と、求めた最近傍点が他の追跡点の最近傍点でもある場合を追跡失敗と呼ぶ。ただし、後者の場合は接続を行うとする。
- (4) (2),(3) の追跡処理を、各選手、各フレームについて行う。

以上の処理により、フレーム間で同一選手の対応付けをすることができる。(3) で接続を行わなかった点は、次の2.4節の修正処理で接続を行う。

### 2.4 追跡失敗フレームの修正

2.3節までの処理が完了したデータで不完全な部分を手動で修正する。修正対象は以下の通りである。

**修正対象 a** 2.3節で追跡失敗した選手位置

- 追跡点と最近傍点との距離が閾値外で、接続がなされていない場合
- 追跡点の最近傍点が他の追跡点の最近傍点でもある場合

**修正対象 b** 2.2節で消失した選手位置

修正対象 a に該当する点に対して、手動操作によって接続の修正をする。実際には、点の座標やフィールドの実映像などを参照しながら、追跡失敗した点と正しい追跡先の点を選択し、対応付ける。選択した2点間のフレーム差が2フレーム以上である場合は、点データがないフレームを2点間の線形補間によって補う。この線形補間処理が、修正対象 b の修正に相当する。

### 3. レンジデータ処理ツール

2章で説明したレンジデータ処理を行うツールを提案する。

#### 3.1 ツールの処理

このツールが行う処理の流れを説明する。

(1) ユーザから以下の設定の入力を受け付け、処理を開始する。

- 各レンジデータのファイルパス
- 各レーザスキャナの位置と角度
- フィールドの縦横の長さ
- 初期フレームにおける各選手 ID
- スキャナ間の時間同期情報

(2) 2.1 節から 2.3 節までの処理を実行する。

(3) 2.4 節における修正対象を、初期フレームから順に探索する。修正対象を見つけると手動修正インタフェースを展開する。ユーザはこれを用いて正しい接続先の点を手動操作によって選択する。修正が完了すると、次の修正対象を探索する。

(4) (3) における修正対象の探索が最終フレームまで完了したら、追跡した選手軌跡データを出力する。

以上より、本ツールにおけるユーザの操作は、流れ (3) の手動修正だけとなる。さらに、修正が必要な点を自動的に探索するので、全フレームにわたりユーザがチェックする負担を軽減することができる。

#### 3.2 手動修正インタフェース

手動修正インタフェースは、以下の 3 つのウィンドウで構成される。

- 追跡グラフウィンドウ
- 俯瞰フィールドウィンドウ
- 実映像ウィンドウ

以下に各ウィンドウの説明を述べる。

##### 3.2.1 追跡グラフウィンドウ

選手位置の追跡状況を表す追跡グラフを表示するウィンドウ。ユーザはこのウィンドウに対してマウスとキーボードによるアクションをし、手動修正を行う。追跡グラフウィンドウを図 5 に示す。追跡グラフは、選手の点を表すノードと点同士の接続を表すリンクで構成される。同一フレームの点が横に並んでおり、フレーム数は上から下にかけて進む。フレーム数は右端に表記している。赤い帯で表すが、後述の残り 2 つのウィンドウに対応するフレーム (以降、参照フレームと呼ぶ) である。ユーザはキーボード操作によって参照フレームを上下選択することができる。

ユーザは追跡グラフに対してクリック操作を行い、2.3 項で述べた追跡失敗の修正を行う。残り 2 つのウィンドウはその操作に連動して表示内容が変わる。以下、ユーザが

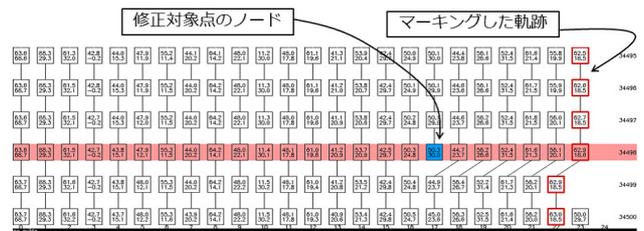


図 5 操作ウィンドウ



図 6 俯瞰フィールドウィンドウ

行う 2 種類のクリック操作を説明する。

##### ● 接続点の選択

修正対象の次に接続すべき選手位置を、次のフレーム以降から選択する操作。修正対象を表すノードは青く塗りつぶされており、このノードの接続先をウィンドウ中のノードをクリックして選択する。選択が完了すると次の修正対象が検索され、ノードが青く塗りつぶされる。

##### ● 軌跡のマーキング

操作ウィンドウのノードをクリックすることで、軌跡をマーキングする。クリックしたノードから、現時点で接続している全てのノードの枠線に色がつけられる。これによって、選択する点と、これまでの軌跡と接続できているか視覚的に確認できる。この色は複数選択でき、後述の俯瞰フィールドウィンドウと対応している。

##### 3.2.2 俯瞰フィールドウィンドウ

俯瞰フィールドウィンドウを図 6 に示す。俯瞰フィールドウィンドウは、参照フレームにある全選手位置をサッカーフィールドに投影した画像を表示する。参照フレームを上下させながらこのウィンドウを参照し、修正対象の次に接続すべき正しい選手位置を目視で確認することができる。修正対象を表す青く塗りつぶされたノードとマーキングの色は、俯瞰フィールドウィンドウ上の点に対応しており、操作ウィンドウとの対応を取りやすくしている。

##### 3.2.3 実映像ウィンドウ

俯瞰フィールドウィンドウを参照するだけでは、正しい接続先が決定できない時は、実映像ウィンドウを参照する。

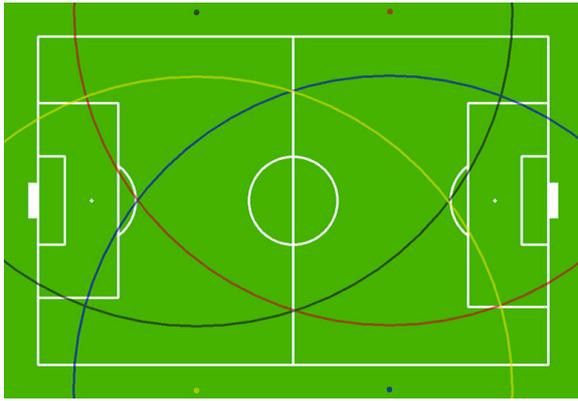


図 7 レーザスキャナの設置

このウィンドウには、参照フレームに対応するサッカーの実画像を表示する。本ウィンドウを使用するためには、レンジデータ取得時に実映像の撮影が必要となる。参照フレームを上下させ、実映像中の選手の動きを参照することができる。

#### 4. データ処理実験

本研究で提案した処理ツールを使って、実際にサッカーの試合を測定したレンジデータを処理し、選手軌跡データを生成した。本実験では、愛知県の豊田スタジアムで行われた大学生のサッカーの試合を測定したレンジデータを処理した。レンジデータは4台のレーザスキャナで測定した。選手の胴体あたりを測定するために三脚を用いて設置した。スキャナはそれぞれ、フィールドに対して図7のように設置した。4色の点は各スキャナの位置を表し、対応する色の円は、各スキャナの測定範囲を表す。

処理したデータは、試合中の1分間であり、処理した選手位置数は37500点(=1500フレーム×25人)である。手動修正にかかるおおよその時間を評価した。ただし、インタフェースに慣れるために対象区間の300フレーム前から操作を開始した。この300フレームの処理にかかった時間は含めない。手動修正のミスがないか確認するため、開始フレームと終了フレームの手動修正結果と実映像の対応をとった。

##### 4.1 レーザスキャナ

用いたレーザスキャナは、SICK社製LMS511である。LMS511は、範囲角度および有効距離に従う扇形の範囲に存在する点を測定する。出力されるレンジデータは、フレーム、時間ごとに、スキャナから測定対象への距離で表される。この距離値に対して、極座標系からxy座標系への変換を行い、処理を適用する。座標変換は事前に行っておくものとする。LMS511の設定は、表1のとおりである。

表 1 LMS511 の設定

範囲角度	190°
角度分解能	0.167°
照射頻度	25Hz
測定範囲	65m

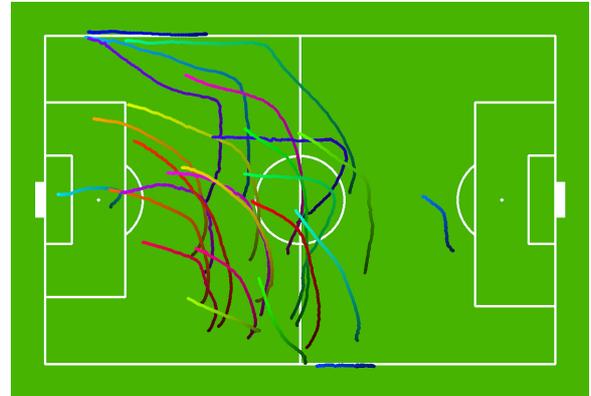


図 8 選手軌跡データ

#### 4.2 実験結果

処理の結果、1500フレーム間の手動修正にかかった時間はおおよそ8時間であった。手動修正ミスは1箇所だけあったが、これは1500フレームの区間よりも前のフレームでのミスであった。また、1500フレーム内のすべての選手点の中で手動修正を加えた点の割合は1.3%であった。

図8に、修正した区間中のある300フレームの選手軌跡を投影した画像を示す。軌跡の色相が各選手を表し、明度のグラデーションが時間を表す。時間が進むほど明度が上がる。これは、生成した選手軌跡に、単純移動平均による平滑化処理を行ったものである。画像と比較し、正しい軌跡が得られていることを確認できた。

#### 4.3 考察

自動処理に対する手動処理の割合は十分に抑えられたが、手動修正にかかる時間の短縮は不十分である。手動コストをさらに抑えるには、統合処理(2.2節)の改善が有効である。2.2.1項で述べた消失および分断の発生が、2.3節の追跡失敗を招き、手動修正のコストを高くしているためである。統合処理の改善のために、以下の二つの方法を検討する。まず、k-meansの初期ラベルを、前のフレームのクラスタ中心をもとに決定する。これは、フレーム間のレンジデータが非常に近い距離で測定されるという性質を利用するものである。もう一つは、k-meansクラスタリングを、複数フレームの点群に対して行う。これは、オクルージョン等によって欠損する位置データを、隣接フレームのもので補うものである。

## 5. まとめ

本研究では、サッカー映像を用いた研究において重要とされるフィールド上の選手軌跡情報を得るために、レーザスキャナを用いた選手の追跡を提案した。レーザスキャナにより得られたレンジデータを、選手軌跡情報に変換するアルゴリズムを提案した。レンジデータの処理における修正処理を手動で行うためのツールも提案した。

提案したツールによって、実際に取得したレンジデータに対して処理を行った結果、正確な選手軌跡情報を得ることが出来た。また、アルゴリズムによって手動修正の負担を抑えることが出来た。具体的には、37500点 (=1500フレーム × 25人) の選手位置に対して、手動修正を行った点の割合は1.3%であった。

今後の課題は、手動修正にかかる時間を短縮することが挙げられる。このために、4.3項で述べた2つの方法を検討、実践する。また、スキャナの位置や角度の設定のずれが、統合処理の正確性を妨げている可能性もあるため、正確な位置合わせの方法も検討する。

**謝辞** 本研究の一部は、独立行政法人情報通信機構(NICT)の委託研究「革新的な三次元映像技術による超臨場感コミュニケーション技術の研究開発」による。

## 参考文献

- [1] 稲本奈穂, 斎藤英雄: 視点位置の内挿に基づく3次元サッカー映像の自由視点観賞システム (小特集: 3次元映像情報メディア技術), 映像情報メディア学会誌: 映像情報メディア, Vol. 58, No. 4, pp. 529-539 (2004).
- [2] 丸尾二郎, 岩井儀雄, 谷内田正彦, 越後富夫, 飯作俊一: サッカー映像からの特定映像イベントの抽出, 電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎, Vol. 99, No. 183, pp. 31-38 (1999).
- [3] Hayashi, K. and Saito, H.: Synthesizing Free-Viewpoint Images from Multiple View Videos in Soccer StadiumADIUM, *Computer Graphics, Imaging and Visualisation, 2006 International Conference on*, pp. 220-225 (2006).
- [4] Muramatsu, Y., Hirayama, T. and Mase, K.: Video generation method based on user's tendency of viewpoint selection for multi-view video contents, *Proceedings of the 5th Augmented Human International Conference*, ACM, p. 1 (2014).
- [5] Mase, K., Niwa, K. and Marutani, T.: Socially assisted multi-view video viewer, *Proceedings of the 13th international conference on multimodal interfaces*, ACM, pp. 319-322 (2011).
- [6] 神田崇行, Glas, D. F., 塩見昌裕, 萩田紀博: 移動する人にサービス提供するロボットのための環境情報構造化, 日本ロボット学会誌, Vol. 27, No. 4, pp. 449-459 (2009).
- [7] Glas, D., Miyashita, T., Ishiguro, H. and Hagita, N.: Laser tracking of human body motion using adaptive shape modeling, *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 602-608 (2007).