Vol.2011-CVIM-176 No.7
2011/3/17

観戦者のセンサ情報を利用したスポーツ映像アノテーションと情報提示

大平茂輝[†1]　長尾確[†2]

スポーツ映像の検索や要約といった応用を行うための基礎技術として，様々な映像インデキシング手法やアノテーション手法が検討されている．本研究では，映像記録の事後処理としてではなく，映像中の事象が発生している瞬間に現場にいる人間の自然な行動から，映像アノテーションの一部となるメタデータを抽出することを目指す．具体的には，試合会場（スタジアム等）における観戦者の視線方向や身体動作に関する情報を，スポーツ観戦メタデータとして携帯情報端末を用いて抽出し，観戦中に撮影した画像と関連付けることで観戦コンテンツを作成する．本研究では，方位センサ情報から観戦者のフィールド上の視点を検出する手法について検討し，視線方向や身体動作を含む観戦コンテンツを活用したスポーツ中継映像の視聴システムを試作した．

Video Annotation and Information Presentation based on Sensor Data of Sport Spectator Activities

Shigeki Ohira[†1] and Katashi Nagao[†2]

Most studies that have provided fundamental techniques for retrieving and summarizing sport videos have focused on automatic indexing and semi-automatic annotations. In this paper, we propose a method of extracting metadata from spontaneous activities by spectators who remain on-site, when an event in a video is ongoing. Specifically, our system extracts the spectator's location and direction of his/her gaze in a stadium with various sensors connected to a personal digital assistant, and it creates the content for the spectator sport based on associating these metadata with still images and videos shot while the game is being watched. This method should contribute to advanced annotations of sports videos based on events of spectator interest.

1. はじめに

スポーツ映像の検索や要約といった応用を行うための基礎技術として，様々な映像インデキシング手法が検討されている．一般に，映像の内容（つまり映像内部の特定のシーン）を検索するためには，映像の内容を記述したアノテーションの付与が有効である[1)]．映像アノテーションは大きく 2 つの手法に分類できる．

一つは，映像に対して音声・画像・言語処理等の機械処理を行った上で，対象映像に対する専門家が修正を加えるオフラインアノテーション手法[2)3)4)] であり，高精度なアノテーションが可能なものの，人間に求められるコストの大きいことがデメリットとなっている．もう一方は，ネットワーク参加型のアノテーション，すなわち，機械処理と特定の専門家への負担を減らして，ネットワークに繋がっている多数のコンテンツ閲覧者が少ない労力でアノテーションを行うことを目的としたオンラインアノテーション手法[5)6)7)] である．シーン検索などで有用となる，利用者が注目する区間に対して集中的にメタデータを付与することが可能であるが，不特定多数の閲覧者からアノテーションされるために，データの信頼性や正確さが問題となっている．

上記 2 つの手法には，人間に求められるコストやデータの信頼性などの点において一長一短があるが，共通して言えることはすでに存在している映像データを対象にしているという点である．つまり，アノテーション時に対象とする映像は，放送局や個人による何らかの加工・編集が加えられたものであり，その意味でアノテーションは映像データ中で発生している事象とは非同期的に行われる事後処理に当たる．また，YouTube などの映像配信サイトで扱っているような数分程度の映像とは異なり，スポーツ映像のような大量の情報を含む映像を，記録時間全体にわたって事細かに解析し情報を付与することは，コスト面で現実的な解とは言い難い．

そこで本研究では，映像記録の事後処理としてではなく，「映像中の事象が発生しているまさにその瞬間にその現場にいる人間の自然な行動から，映像アノテーションの一部となるメタデータを抽出する」ことを目指す．具体的には，試合会場（スタジアム等）における観戦者の動作や視線の方向をスポーツ観戦メタデータとして抽出し，観戦中に撮影された画像

†1 名古屋大学 情報基盤センター
Information Technology Center, Nagoya University
†2 名古屋大学 大学院情報科学研究科
Graduate School of Information Science, Nagoya University

と関連付けることで観戦コンテンツを作成する．本研究では，方位センサ情報から観戦者のフィールド上の視点を検出する手法について検討し，取得された観戦メタデータを活用したスポーツ中継映像の視聴システムを試作した．

2. 体験記録としてのスポーツ観戦コンテンツ

コンピュータの小型化や磁気ディスクの大容量化，データの高圧縮技術を背景にして，ウェアラブル機器を利用した体験の常時記録に関する研究が行われている．スポーツ観戦を，個人の日常的な生活の一部と見なすと，ウェアラブルとユビキタスセンサを利用したライフログ[8] の取得を，スポーツ観戦というドメインに適用したものと捉えることができる．スマートフォンに代表される各種センサを搭載した携帯端末機器はますます身近になってきており，今後，ライフログの一部として，スタジアムで観戦した状況の記録は一般的に行われるようになると思われる．

また，スポーツ観戦メタデータの抽出を，スタジアムという場を共有する多くの人間によって協調的に行われる映像へのタギングと見なすと，本研究は，Web 上で多数のユーザが同時にアノテーションを付与する仕組みを，実環境下で行えるようにしたものと捉えることができる．

角ら[9] は，体験を「自らの身をもって何かを経験すること」とし，体験する人と対象物を観測対象とすることが必要である，と述べている．すなわち，人間と対象物の間でなされたインタラクションを観測し，データ化することを課題としている．さらに，体験は「行動一般に対する主観的解釈」と定義できる．つまり，目の前で起こっている事象に加えて，ユーザ自身の心情や感情を含むものである．従来の体験記録は，この行動一般を，体験にまつわる文脈情報として記録したものである．しかし，人間の内的状態を意味する主観的解釈を自動的に記録することは困難であることから，体験映像に対するコメントやブログ等の執筆によって，その代替手段としてきた．

スポーツ観戦という体験においても，ユーザの心情や感情を自動的に記録することは容易ではない．観戦コンテンツの一部として，観戦記録に対する個人的な意見や気持ちを言語的に記録し検索可能にすることは，体験の共有や追体験の観点からも意義のあることであると考える．しかし本論文では，勝敗の行方が心的状態を大きく左右するスポーツ観戦特有の性質上，体験を総括するような心情や感情は記録の直接的な対象とせず，試合観戦経過における観戦者の注目点と盛り上がりのみを対象とする．前者は試合中の撮影行為，後者は撮影以外の身体的な動作から抽出可能と考える．試合観戦中のカメラによる撮影は従来からも行われており，観戦状況の記録における最も基本的な手段である．本研究では，スポーツ観戦メタデータとして以下の情報を扱い，これに撮画像を組み合わせたものをスポーツ観戦コンテンツと定義する．

- 試合名
- 対戦チーム名
- 競技場名
- 天候
- 気温
- 湿度
- 座席位置（緯度・経度）
- 視線方向（方位角）
- 身体動作
- 写真・ビデオの撮影時刻

3. スポーツ観戦メタデータの抽出

筆者らは，スポーツ映像を対象として，オフラインアノテーション手法における機械処理の精度向上や専門家が修正する映像区間の絞り込み，また，オンラインアノテーション手法における多数の閲覧者の視点や注目度といった統計的情報の取得，これらを映像中の事象の発生と同期的に行うことによる新しい映像アノテーション手法の確立を目指している．

スポーツ観戦メタデータを獲得・共有することにより，現場 (スタジアム) で観戦する側の視点を取り込むことは，スポーツ中継映像の構造化を進める上で非常に重要であると同時に，さまざまな応用の可能性を含んでいると考える．

3.1 使用する機器

本研究で目指すスポーツ映像アノテーションは，放送局によるスポーツ映像の中継を見るのではなく，スタジアムに足を運び，試合を生で観戦・応援しながら持ち込んだデジタルカメラやデジタルビデオカメラで適宜撮影を行う，というごく自然な観戦スタイルを映像アノテーションに適用することによって行う．時間情報に加えて、観戦時の身体動作や視線の方向、観戦位置に関する情報を観戦メタデータとして獲得するために、電子コンパスを内蔵したデジタルカメラ (CASIO 社製 EX-H20G と SONY 社製 DSC-HX5V) と，各種センサを搭載した携帯情報端末（センサユニット）を導入した (**図 1**)．

具体的には，アクリル板上に各種センサとそれらセンサ情報を処理する携帯情報端末を

図 1 センサユニット
Fig. 1 Sensor unit.

USB ハブを介して配置し，トレイルランニング用の薄いリュックに内蔵している．拡張バッテリを備えることで，5 時間程度の連続稼動が可能であり，試合観戦に十分耐えうる仕様となっている．体に密着するため，カメラ撮影や試合観戦への影響も非常に少ない．

また，デジタルカメラやデジタルビデオカメラといった撮影機器を，各種センサ情報を処理するセンサユニット本体から切り離すことにより，システムの低コスト・省電力化を図ると同時に，高画質な映像の記録を可能とし，個人が持ち込む撮影機器にも自由度を与えている．ただし，センサユニットで解析・記録される各種センサ情報と撮影データとの同期を行うため，センサユニット本体と撮影機器の内部時計を事前に合わせておく必要がある．センサユニットの構成および仕様を表 1 に示す．

本研究では，デジタルカメラで撮影された画像の Exif データから視線の方向を取得し，センサユニットの 3 軸加速度センサから身体動作に伴う Z 軸方向の加速度を検出する．なお，GPS はデジタルカメラにもセンサユニットにも搭載されているが，現状では 1m の精度が出ないこと，また屋根がある場合には位置情報の取得が困難であることから，現実的には RF-ID タグ等の非接触センサや電子チケット等による座席位置の自動設定が望ましいと考えられるため，本研究では観戦者の位置情報として利用せず，座席番号を事前情報として与えている．

表 1 センサユニットの構成と仕様
Table 1 Architecture and specification of sensor units.

携帯情報端末	品名	WILLCOM03	Advanced/W-ZERO3[es]
	メーカ	シャープ株式会社	
	OS	Windows Mobile 6.1 Classic	Windows Mobile 6 Classic
	CPU	Marvell PXA270 520MHz	
	メモリ	Flash 256MB/SDRAM 128MB	
	質量	約 135 g	約 156 g
3 軸加速度・3 軸地磁気・気圧センサ	型番	TDS01V	
	メーカ	株式会社バイテック	
	加速度センサ	加速度：±2 [g]，傾斜角：−90〜+90 [deg]，傾斜角分解能：1 [deg]，傾斜角精度：±3〜10 [deg]	
	地磁気センサ	磁束密度：±120 [μT]，方位分解能：1 [deg]，方位精度：±10 [deg]	
	気圧センサ	気圧：710〜1062 [hPa]，高度：−100〜2000 [m]，高度分解能：3 [m]，高度精度：±10 [m]	
GPS センサ	型番	BU-353	
	メーカ	GlobalSat 社	
	主な仕様	受信方式：並列 20 チャネル（SiRF StarIII/LP），測位更新間隔：毎秒，位置精度：10 [m]（単独測位, 2DRMS），速度精度：0.1 [m/s]，最高速度：515 [m/s]	
総重量 (電池含む)		約 870 g	約 860 g

3.2 観戦行為に基づくスポーツ中継映像の構造化

スポーツ観戦メタデータは，スタジアムにおける観戦者の多様な視点や注目の度合いといった中継映像には含まれていない情報を映像アノテーションとして付与することにより，従来型のアノテーションと共存する形でスポーツ映像の構造化に寄与することが可能である (**図 2**)．

特に，スタジアム観戦者の注目度や視点といったその場でしか得られない情報は，のちのオフラインアノテーションを効率良く行う手助けになることや，オンラインアノテーションが提供する統計的な注目度の初期値として利用し，視聴シーンを推薦することで，コメントやタグといったアノテーションをより多く収集するといった利点が期待される．

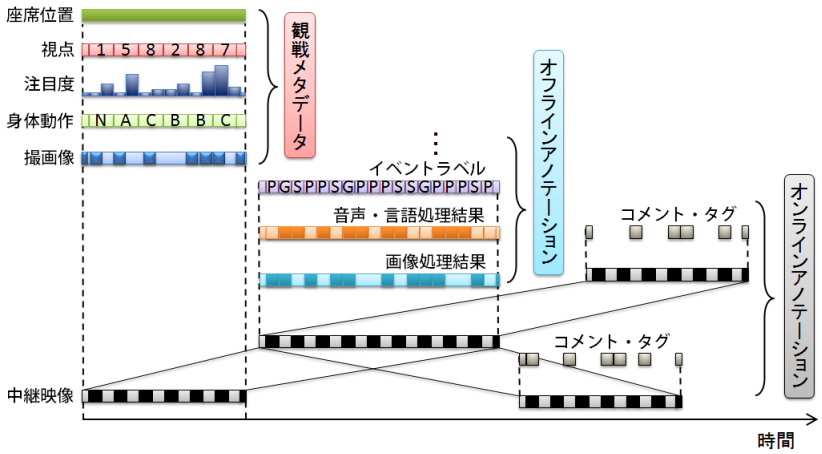

図 2 観戦メタデータに基づくスポーツ中継映像の構造化
Fig. 2 Structualization of live sport game TV with metadata of spectator activities

4. スポーツ観戦メタデータの分析と利用

4.1 観戦者の撮影頻度と使用機器

本研究では，観戦者の一部が試合中に撮影を行うことを前提にしている．また，個人の視点によって撮影されたスポーツ映像・写真を独立したデータとして扱うのではなく，スタジアム全体を一つの巨大なスポーツ観戦コンテンツとみなすことで，映像アノテーションに限らない幅広い応用が可能になると考えている．

そこで，観戦中に撮影を行う人数と撮影枚数について調査を行った．フクダ電子アリーナ (千葉市) 内のスタジアム通路で隔てられた 1 ブロック (16 × 12 列) 約 190 名を対象に，試合中の撮影の有無，撮影機器の種類を目視により確認したところ，11 名が携帯カメラ，16 名がデジタルカメラによる撮影を行った．さらに，デジタルカメラの利用者 16 名に対して，1 試合の撮影枚数についてアンケート調査を行ったところ，平均 101 枚 (最大 300 枚，最小 2 枚，無回答 5) という回答が得られた．撮影枚数にばらつきはあるものの，1 割程度の観戦者が試合中に撮影をしていることから，様々な視点からの撮画像の共有と観戦メタデータの統計的な処理に基づく映像アノテーションは現実的な手法と言える．

4.2 視線方向検出

4.2.1 使用する方位情報付き撮画像データ

上述のセンサユニットおよびデジタルカメラを用いて，港サッカー場 (名古屋市) において方位情報付きの撮画像データの収集を行った．2 種類のデジタルカメラを各 3 台ずつ用意し，データ収集に協力した大学院生 6 名を 3 名ずつの 2 グループ (A，B) に分け，グループごとに同一の機種を使用するように配布した．座席位置は，メインスタンドに 2 名，バックスタンドに 2 名，サイドスタンドに 2 名ずつを**図 5** のように配置し，被撮影者が移動するフィールド上の 54ヶ所にマーカーを設置した．

データ収集は，以下の 4 つの手順で行った．E1 と E2 は全 54 マーカーに対して，E3 は重複なしのランダムな 30 マーカーに対して行った．手旗を下ろすまでの間は何度撮り直してもよく，ズームとピント合わせは各自の判断に任せた．

E0) キャリブレーション用に，座席位置から最も近いタッチラインもしくはゴールラインの左右 2 本のコーナーフラッグを撮影

E1) 撮影者はカメラを構えずに待ち，被撮影者がマーカー位置へ移動したのち，手旗の合図 (10 秒) 中にカメラを構えて撮影

E2) 撮影者はカメラを目の前に構えていつでも撮影できる体勢を整え，被撮影者がマーカー位置へ移動し手旗の合図 (3 秒) 中に撮影

E3) 撮影者はカメラを目の前に構えていつでも撮影できる体勢を整え，被撮影者がランダムに選択したマーカー位置へ移動し手旗の合図 (3 秒) 中に撮影

4.2.2 撮影方位の補正

一般的に，電子コンパスは 10 度程度の誤差を含んでおり，周辺の磁界の影響も受けやすい．そこで，以下の手順により撮画像が保持する撮影方位情報の補正を行った．なお，撮影者の座席位置とマーカー位置の緯度・経度は，Google Maps API を用いて取得しており，その値を国土交通省国土地理院が定める平面直角座標系（愛知県は VII 系）の X, Y 座標として求めてある．座席位置からマーカーへの方位角は，Bowring[10] の式を用いて計算した．

1) E0 によって撮影されたフラッグの方位角の値が小さい方を θ_{min}，大きい方を θ_{max} とする．両者の差分 $\theta_{diff} = \theta_{max} - \theta_{min}$ が，
・180 度より小さいならば，$\theta_L = \theta_{min}, \theta_R = \theta_{max}, \theta_{range} = \theta_{diff}$
・180 度より大きいならば，$\theta_L = \theta_{max}, \theta_R = \theta_{min}, \theta_{range} = 360 - \theta_{diff}$

2) 各座席位置より，左右のフラッグの真の方位角 θ_{L0} と θ_{R0} が得られる．
$\theta_{L0} < \theta_{R0}$ の場合は，$\theta_{range0} = \theta_{R0} - \theta_{L0}$

$\theta_{L0} > \theta_{R0}$ の場合は，$\theta_{range0} = 360 - (\theta_{L0} - \theta_{R0})$

3) ある時点に撮影された画像の方位角が θ のとき，補正後の方位角 θ' を次式で表す．

$$\theta' = \theta_{L0} + (\theta - \theta_L) \cdot \frac{\theta_{range0}}{\theta_{range}}$$

撮影方位角と真の方位角との差の平均値および標準偏差を，撮影者 (A1〜A6)，撮影手順 (E1〜E3)，撮影方位の補正前後で比較したものを**図 3** に示す．

A1〜A3 および B1 の 4 名については，補正による効果が顕著に表れている，もしくは補正前後で差がないと判断できるが，B2 については補正の結果，真の方位角とのずれが大きくなっている．これは，フラッグ撮影時の方位角が真の方位角と 3 度程度しか違わないことに起因すると思われる．そのため，上記 3) による方位角の補正は，$|\theta_L - \theta_{L0}| < 5°$ かつ $|\theta_R - \theta_{R0}| < 5°$ の場合にのみ行うという条件を加えることが考えられる．また，B3 の E3 からは非常に標準偏差が大きいという結果が見られるが，機器の不具合によるものなのか，データのほぼ全体にわたって異常値が含まれているため，本評価からは外す．

撮影手順による違いについては，常にカメラを構えている方がセンサの値が安定すると予想したが，E1 と (E2,E3) の間に有意な差は見られなかった．また，マーカー間の移動先がある程度予想できる (E1,E2) と異なり，E3 は予測が難しくかつ撮影時間も短いことから最も精度が悪くなると考えたが，こちらについても両者に有意な差は見られなかった。

以上より，A グループでは約 5 度，B グループでは約 10 度のずれが見られることから，撮影方位の検出を単独で行う場合，デジタルカメラの性能に左右されると言える．

また，**図 4** では，E1 による撮影の 10 秒間に複数枚撮影した場合の，1 枚目と 2 枚目の方位角の違いを示している．こちらも撮影機器による違いのみが表れており，2 枚目の撮影の方が方位角の精度が良いとは必ずしも言えない．

4.2.3 複数人による撮影に基づくフィールド上の視点検出

単独のデジタルカメラでは，比較的良い精度を出す機種でも 5 度程度の方位誤差を含んでいる．また，仮に正確な方位が分かったとしても，フィールド上の位置を特定することは難しい．そこで，複数人による撮影によって得られる複数の撮影方位から，観戦者の視点の検出を行う (図 5)．具体的な手順は以下の通りである．

i) 各座席位置からセンサで取得した方位へ線分を描画する

ii) A グループの線分の交点が形成する三角形を T_A，B グループの線分の交点が形成する三角形を T_B とする

iii) T_A と T_B の位置関係とグループ優先領域 (P_A, P_B) により，以下の場合分けを行う

(a) T_A が P_A の内部にのみ存在する場合，iv) の処理を T_A に対して行う

(b) T_B が P_B の内部にのみ存在する場合，iv) の処理を T_B に対して行う

(c) T_A と T_B が P_A と P_B にまたがるように混在する場合，T_A と T_B の面積を比較して小さい方に対して iv) の処理を行う

iv) 得られた三角形の重心 P_c を求め，これを視点の検出結果とする．

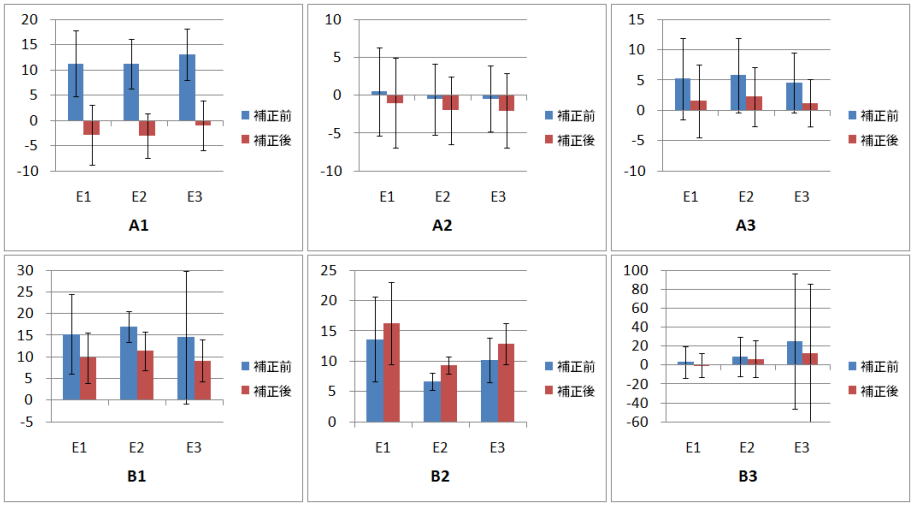

図 3 撮影方位角のずれの比較 (撮影者別，撮影手順，補正前/後)(±SD)
Fig. 3 Difference between true azimuth and gaze orientation with digital compass.

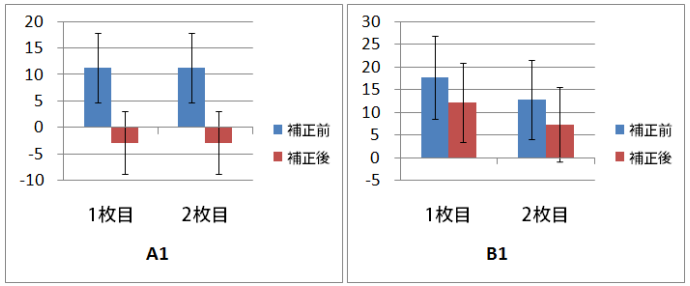

図 4 連続的な 2 枚の撮影による撮影方位角の違い (±SD)
Fig. 4 Difference between two pictures by continuous shooting.

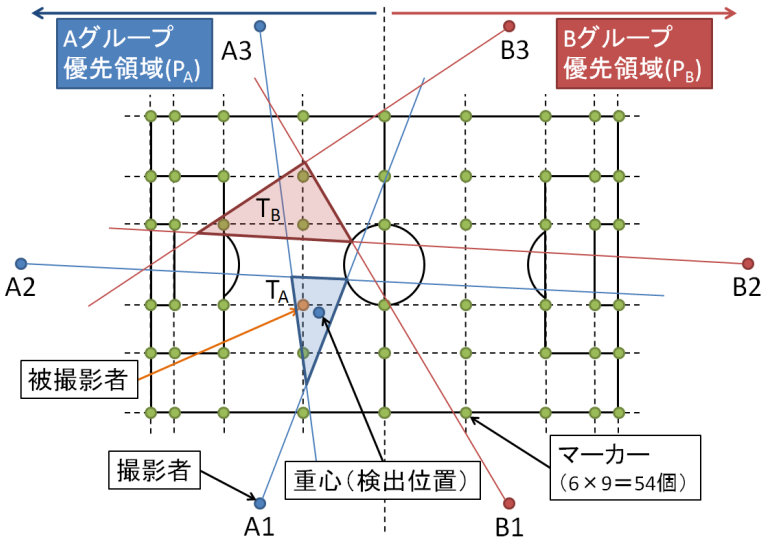

図 5 複数の撮影方位角に基づく視点検出
Fig. 5 Detection of gaze based on multiple azimuth angles.

フィールド内に三角形を形成できない，あるいは，優先領域内に別グループの三角形のみが形成される場合は，各グループから 2 点ずつ計 4 点を選出し，四角形の重心を求めて検出結果とする．

図 6 に示す結果から，基本的な三角形の重心を求める方法により，約 9m(サッカー場のセンターサークル程度) の範囲内で視点の検出が可能であり，マーカーに隣接するブロック内に検出位置が入る割合は約 78%であることが分かった．また，1 点を追加して四角形の重心を求めた場合には，約 6m の範囲内で検出可能となり，マーカーの隣接ブロックに入る割合はほぼ 100%となった．単独のデジタルカメラでは，撮影方位角の補正処理を行わないと実用的とは言いがたいが，多地点からの複数撮影に基づく視点検出では，補正処理を行わなくても重心の計算に利用する点を増やすことで精度を上げることが可能であることを示唆している．

4.2.4 複数の観戦者による視点の共有

前述の視点検出手法は，複数地点の観戦者が同時刻に撮影を行うことが前提となっているため，その実現可能性について調査するために，スポーツ観戦コンテンツ 1 試合分を対象に

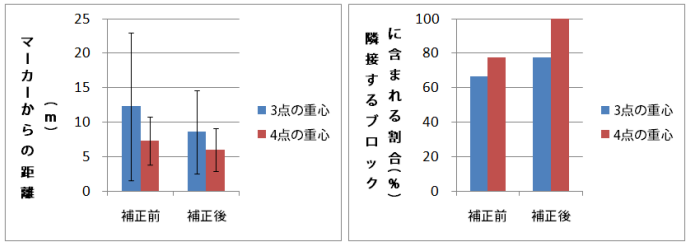

図 6 複数撮影情報を利用したフィールド上の視点検出結果
左：マーカーからのずれ，右：隣接ブロックとの一致度
Fig. 6 A result of gaze detection in a stadium field based on multiple azimuth angles.

検証を行った。本データでは，被験者 3 名が同じ試合を観戦し，各自のタイミングで撮影を行っており，撮影回数はそれぞれ 167 回，105 回，183 回である．その結果，1 試合で 2 名以上が同時に撮影した回数は 16 回，うち同じシーンを撮影していた割合は 94%であり，2 名以上が 1 秒違いで撮影した回数は 26 回，うち同じ (連続する) シーンを撮影した割合は 92%であった．これより，撮影する観戦者数が増えれば，視点の共有は可能と予想される．

4.3 身体動作による盛り上がりの検出

スポーツ観戦時の身体動作としては，カメラ撮影行為以外にも様々なものがある．ここでは，ゴールシーンなどの試合の盛り上がりに応じた典型的な身体動作として，上下運動を検出する．センサユニットで取得した Z 軸 (上下) 方向の加速度について，設定した振幅閾値 0.2 を超える箇所を盛り上がりとして検出した．同一試合の中継映像をもとに，該当時刻の前後 5 秒間とゴールシーン 3 箇所およびシュートシーン 16 箇所とを比較したところ，ゴールシーンの検出率は 100%であり，シュートシーンについては再現率 62.5%，適合率 83.3%という結果が得られた．

シュートシーンの誤検出の理由としては，シュートとは関係ないシーンで興奮して立ち上がったり，試合終了直後に選手をねぎらうために立ち上がって拍手で迎えた，といった現象が起こっていたためであり，盛り上がりという意味においては正しく検出されたと言える．また，シュートシーンの再現率が低くなった理由としては，座席から遠くてよく見えない位置でのシュートに対して，強い興奮が得られなかったため，と推測される．

4.4 観戦メタデータを活用したスポーツ中継映像視聴システム

獲得した観戦メタデータは，試合終了後にスポーツ中継映像とともに自らの体験を振り返る際の手掛かりにすることができる．また，試合中にメタデータを共有することができれ

図 7 スポーツ中継映像視聴システム
Fig. 7 Sport video browser with metadata of spectator activities.

ば，他者の視点を知ることで情報の欠落を補完したり，より深く試合内容を理解することにも結び付く可能性がある．

本研究では，スポーツ中継映像を視聴する際に観戦メタデータを活用するシステムを試作した．本システムは，**図 7** に示すような Web ブラウザ上で動くインタフェースを有しており，中継映像とともに観戦者の視点情報や撮影した画像，身体動作の様子や撮影枚数の時系列情報を閲覧することができる．共有可能な観戦メタデータが存在する場合には，他の観戦者の視点で撮影された画像も一緒に閲覧することができる．

5. まとめと今後の課題

観戦者の視線方向や身体動作を観戦メタデータとして抽出することにより，スポーツ映像アノテーションの一部として利用可能となることを示した．電子コンパスを利用した視線方向の検出精度は十分に高いとは言えないが，スタジアムにいる複数観戦者の視線情報を統合することにより，フィールド上の視線領域の絞り込みを行うことができることがわかった．また，観戦メタデータを活用した中継映像の視聴システムを試作したが，視点検出精度のさらなる向上，観戦時の身体動作の分類・抽出，システムの評価は今後の課題である．観戦コンテンツを記録・共有することによって，スポーツ映像の視聴方法に変化が起こるのか，観戦者の視点が映像アノテーションに加わることにより，従来のアノテーション手法やスポーツ映像要約などの処理結果にどのような影響を与えるのか，といった点についても分析を進めていきたいと考えている．

謝辞 本研究は科研費（21700104）の助成を受けたものである．

参 考 文 献

1) Nagao, K., Shirai, Y. and Squire, K.: Semantic Annotation and Transcoding: Making Web Content More Accessible, *IEEE MultiMedia*, Vol.8, No.2, pp.69–81 (2001).
2) Davis, M.: An Iconic Visual Language for Video Annotation., *Proceedings of the IEEE Symposium on Visual Language*, pp.196–202 (1993).
3) Nagao, K., Ohira, S. and Yoneoka, M.: Annotation-Based Multimedia Summarization and Translation, *Proceedings of the Nineteenth International Conference on Computational Linguistics (COLING-02)*, pp.702–708 (2002).
4) Smith, J.R. and Lugeon, B.: A Visual Annotation Tool for Multimedia Content Description, *Proceedings of the SPIE Photonics East, Internet Multimedia Management Systems*, pp.49–59 (2000).
5) 山本大介，増田智樹，大平茂輝，長尾 確：映像を話題としたコミュニティ活動支援に基づくアノテーションシステム，情報処理学会論文誌，Vol.48, No.12, pp.3624–3636 (2007).
6) Yamamoto, D., Masuda, T., Ohira, S. and Nagao, K.: Video Scene Annotation based on Web Social Activities, *IEEE MultiMedia*, Vol.15, No.3, pp.22–32 (2008).
7) Masuda, T., Yamamoto, D., Ohira, S. and Nagao, K.: Video Scene Retrieval Using Online Video Annotation, *Lecture Notes on Artificial Intelligence (LNAI 4914: JSAI 2007 (K.Satoh, et al. Ed.))*, Springer-Verlag, pp.54–62 (2008).
8) Tancharoen, D., Puangpakisiri, W., Yamasaki, T. Aizawa, K.: Life Log Platform for Continuous and Discrete Recording and Retrieval of Personal Media, *Trans. ECTI-EEC*, Vol.5, No.2, pp.165–173 (2007).
9) 角康之, 保呂毅, 三木可奈子, 西田豊明：体験共有コミュニケーションを促すガイドシステム, 第 19 回人工知能学会全国大会 (2005).
10) Bowring, B. R.: TOTAL INVERSE SOLUTIONS FOR THE GEODESIC AND GREAT ELLIPTIC, *Survey Review*, Vol.33, No.261, pp.461–476 (1996).