

機関番号：13901

研究種目：若手研究(B)

研究期間：2009～2010

課題番号：21700104

研究課題名(和文) 携帯情報端末によるスポーツ観戦メタデータの抽出・共有と映像アノテーションへの利用

研究課題名(英文) Extraction and Sharing of Metadata for Spectator Sports with Personal Digital Assistant and its Application to Video Annotation

研究代表者：大平 茂輝 (OHIRA SHIGEKI)

名古屋大学・情報基盤センター・助教

研究者番号：60339695

研究成果の概要(和文)：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している。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等を用いて抽出し、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した。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。

研究成果の概要(英文)：We propose a method for extracting metadata from the spontaneous eye and body movements of spectators at sporting events and combining the metadata with video shot while the event is being watched. In particular, our system extracts the spectator's body movement and direction of their gaze in a stadium using multi-sensor equipment, which is made up of a personal digital assistant and a digital camera, and this is used to create the spectator sport content based on associating these metadata with the still images and videos shot while the game is being watched. We developed a method for detecting the spectator's view area on the field by using a direction sensor and created a prototype of our sport video viewer using the spectator's content, which includes the gaze direction and body movement.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	2,000,000	600,000	2,600,000
2010年度	1,200,000	360,000	1,560,000
総計	3,200,000	960,000	4,160,000

研究分野：メディア科学

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：モバイルシステム、コンテンツ、映像アノテーション

1. 研究開始当初の背景

(1) 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、映像の内容を検索するためには、映像の内容(つまり映像内部の特定のシーン)を記述したアノテーションの付与が有効である。研究代表者のグループは、映像コンテンツの

検索や要約に必要なメタデータを半自動的に付与する映像アノテーション技術の研究・開発を行ってきた。

(2) 映像アノテーションは大きく2つの手法に分けられる。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。

2. 研究の目的

(1) オフラインアノテーションとオンラインアノテーションという2つの映像アノテーション手法にはそれぞれ一長一短があるが、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。

(2) 本研究の着想は、「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できないか」という点に基づいている。特に、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。

(3) スポーツ中継を行う放送局は、スタジアム内の多数のカメラを切り替えることで配信用の映像を作成しているが、すべての重要なシーンを漏れなく最良の角度で記録することは、熟練のカメラマンやスイッチャを擁しても難しい。本研究は、スタジアム内の観客数だけカメラが存在するという仮定の下でのスポーツ映像の新しい構成方法に着目している。多視点テレビ放送に見られるような視聴者の多様性に応えるためには、より多くの視聴ポイントを実際の視聴者の視点で押さえる必要がある。携帯情報端末を利用してスポーツ観戦状況をリアルタイムに記録することは、スポーツ観戦という本来の目的を阻害しない形で自然に情報を記録し、且つ、その情報を映像アノテーションと

いうコストの高い情報処理に適用する最初の意味付けという意義がある。

3. 研究の方法

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ハンディカムで適宜撮影を行う、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組み込む手法を提案する。しかし、デジタルカメラやハンディカムで撮影した静止画や動画が保持するメタデータは、一般的に時間情報のみである。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状況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)を導入する(図1)。また、取得したメタデータをスタジアム内の観戦者同士が共有することにより、観戦位置などの条件によって欠落した試合内容の補完や、理解促進の手助けになる情報提示手法について検討する。



図1：センサユニット

4. 研究成果

(1) 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。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、様々な視点から撮影された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考えられる。

(2) 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。視線の方向(撮影方位)は、デジタルカメラで撮影された画像のExifデータから取得した。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A・Bの2グループに分け、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2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。撮影方位角と真の方位角との差を調査した結果、単独のデジタルカメラでは、比較的精度の良い機種でも5度程度の方位誤差を含んでおり、撮影方位角の補正処理を行わないと実用的とは言い難いということが分かった。

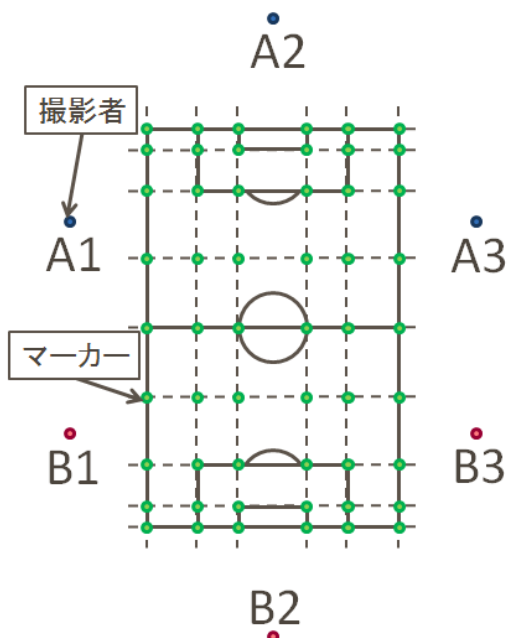


図2：座席位置とマーカーの配置

(3) 単独のデジタルカメラを用いる場合、方位誤差の影響を無視することはできず、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ことか

ら、複数人の撮影によって得られる複数の撮影方位をもとに、観戦者の視点の検出を行った(図3)。フィールドを2つの優先領域に分割し、誤差を含んだ撮影方位によって形成される三角形の重心を求める手法により、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となることがわかった。また、1名分の撮影方位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った(図4)。

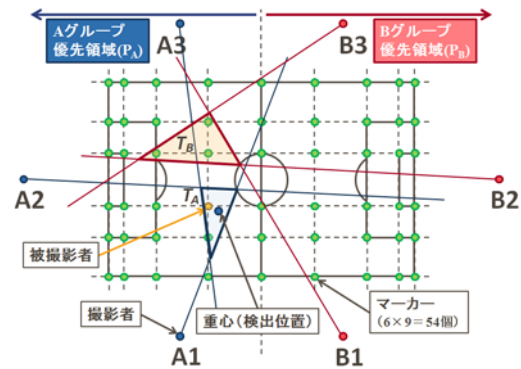


図3：複数の撮影方位角に基づく視点検出

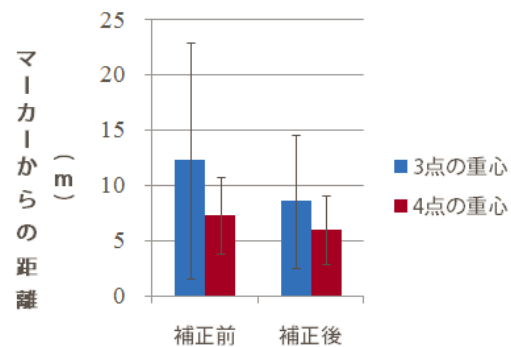


図4：複数撮影情報を利用したフィールド上の視点検出結果

(4) 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、上下運動を検出した。センサユニットの3軸加速度センサから取得したZ軸(上下)方向の加速度について、設定した振幅閾値を超える箇所を盛り上がりとして検出した(図5)。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には、シュート以外の盛り上がりの現象が発生したことが関係しており、再現率が低い理由としては、座席から遠くて見づらい位置

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。

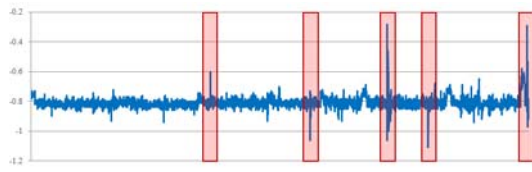


図5：身体動作による盛り上がりの検出

(5) 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中継映像の構造化に寄与することが可能となった(図6)。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、試合後の従来アノテーションを効率良く行う手助けになることや、統計的な注目度の初期値として利用することにより視聴シーンの推薦を行うことで、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。

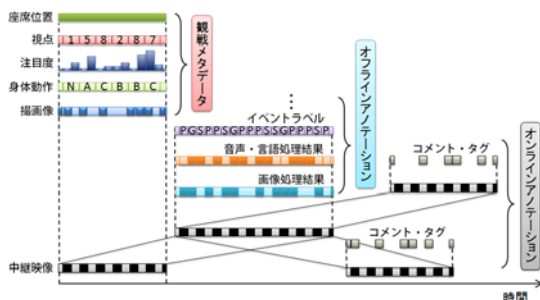


図6：スポーツ中継映像の構造化

(6) 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、他者の視点を知ることによって情報の欠落を補完すること、より深く試合内容を理解することにも結び付く可能性を示した。本システムは、中継映像とともに、観戦という体験を振り返る手掛かりとしても利用可能である。



図7：スポーツ中継映像視聴システム

(7) 本研究では、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し、スポーツ映像アノテーションの一部として利用可能であることを示した。今後の展望としては、まず、視点検出精度のさらなる向上や、観戦時の身体動作の分類・抽出が挙げられる。また、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有意な影響を与えるのかといった点について、収集データを増やしてさらに分析を進めていく必要がある。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計2件)

- ① 大平茂輝、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、情報処理学会コンピュータビジョンとイメージメディア研究会、2011年3月17日、東京工業大学
- ② 大平茂輝、スポーツ観戦中の情報抽出と映像アノテーションへの利用、情報処理学会第73回全国大会、2011年3月3日、東京工業大学

6. 研究組織

(1) 研究代表者

大平 茂輝 (OHIRA SHIGEKI)
名古屋大学・情報基盤センター・助教
研究者番号：60339695

(2) 研究分担者 なし

(3) 連携研究者 なし