

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 16 日現在

機関番号：14501

研究種目：挑戦的萌芽研究

研究期間：2012～2014

課題番号：24650084

研究課題名(和文)映像の言語化と学習に関する研究

研究課題名(英文)Study on Language generation and learning from videos

研究代表者

有木 康雄(Ariki, Yasuo)

神戸大学・都市安全研究センター・教授

研究者番号：10135519

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：4台のHDカメラでサッカープレーを撮影し、選手とボールを自動抽出・追跡できるようになった。その結果得られた選手とボールの情報を基に、「ゴールキック」、「コーナーキック」、「フリーキック」、「スローイン」という4つの状況を、HMM(隠れマルコフモデル)で学習することにより、これらを高精度に認識できるようになった。この結果を用いて、デジタルカメラワークや解説生成の研究を行い、ルールを自動学習する展望を得た。

研究成果の概要(英文)：We developed a system which can extract and track a ball and players on videos of soccer sports games captured by four HD cameras. Based on their tracking results, four events such as "goal kick", "corner kick", "free kick" and "through-in" can be recognized with high accuracy by Hidden Markov Model (HMM). We carried out the researches of digital camerawork and commentary generation, and then got the prospects for automatic learning of game rules through this study.

研究分野：パターン認識

キーワード：状況理解 画像認識 音声認識 機械学習 HMM 抽出 追跡 メディア変換

1. 研究開始当初の背景

(1) サッカーなどのスポーツでは、人とボールの位置によって、状況は多様に変化する。我々はこれまで、スポーツ映像から状況を自動認識し、仮想的なカメラワークを使って、その状況を分かり易く映像表現する研究(図1の(A)、(D))を行ってきた。本研究では、理解した状況を言語表現し、さらに状況を解説する研究(B)、(C)に挑戦するものである。このような研究はまだ始まっていないが、これは、映像の言語化を含んでおり、極めて挑戦的な研究テーマである。

(2) 現在、スポーツ映像を対象とした研究では、映像解析が中心であり、その言語化や解説・映像の生成といった研究は行われていない。状況分析の方法を、実例から自動学習する研究(E)は、機械学習の分野で事例学習として研究されているが、多数の実況中継映像から分析方法を学習する研究は行われておらず、萌芽的な研究内容といえる。

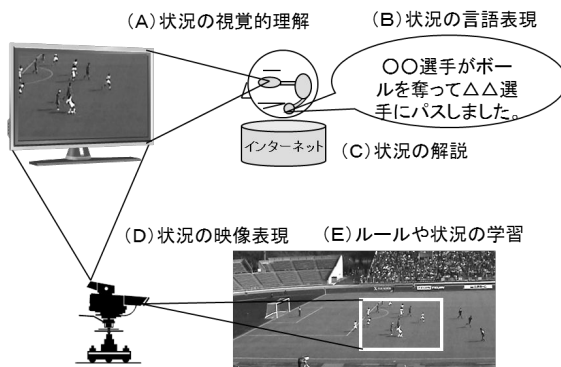


図1 映像の言語化と学習の研究内容

2. 研究の目的

(1) 現在見ている状況を、人が伝達するという行為には、その状況を理解し、言語化して発声するというプロセスが含まれている。このような情報伝達のメカニズムをコンピュータで行うには、

- (A) 現実に行っている状況を視覚的に理解する、
- (B) 理解した状況を言語化して音声で表現するといった処理が中心となる。さらに、

(C) 状況の原因や今後の予測を解説として言語化し、

(D) 理解した状況を図や映像として再表現する、といった機能が必要となる。

(2) また、コンピュータが状況を理解できるためには、

(E) 状況分析の方法を実例から自動学習する、といった機能も必要である。本研究では、この映像の言語化と学習のメカニズムを、スポーツ実況中継を対象として明らかにし、コンピュータ上でソフトウェアとして実装することを目的としている。

3. 研究の方法

(1) 情報伝達のメカニズムを解明するために、(A) 状況の視覚的な理解、(B) 状況の言語表現、(C) 状況の解説、(D) 状況の映像表現、(E) 状況分析の学習、といった5つの機能について研究する。

(2) (A)では、状況を視覚的に理解するための映像解析手法について研究する。特に、選手の姿勢推定・動作認識、選手とボールの関係、複数選手の関係を抽出するための映像処理について研究する。(E)では、状況分析の方法を映像から学習するメカニズムについて研究する。特に、用語に関連する状況とカメラワークを多数の映像から集め、クラスタリングすることにより、用語と状況・カメラワークの対応表を作成する。この対応表を基に映像を解析して、(B)状況を言語表現し、(D)仮想的なカメラワークを付与して映像を再生成する研究を行なう。更に、(C)状況を引き起こした原因や今後の予測について、解説を可能とする研究を実施する。

4. 研究成果

(1) (A)状況の視覚的な理解について研究を行った。試合の状況は、選手とボールの関係、選手と選手の関係で表現されるが、複数選手

がオクルージョンを起こしている場合には、選手の追跡に失敗し易い。そこで、オクルージョンに口バスタな選手追跡を行うために、時間状況グラフにガイドされたパーティクルフィルタによる新しい選手追跡手法を提案した。従来のパーティクルフィルタによる選手追跡では、映像のフレーム間で複数選手の位置情報を用いていないため、一度対象を見失うと再度発見するのが困難であるという欠点があった。そこで、複数選手の位置情報を時間状況グラフとして表現しておき、これにガイドされる形でパーティクルフィルタを実行すれば、オクルージョンが起こっても選手の誤追跡を大幅に減らすことができる。

評価実験では、実際の視点固定単眼サッカー映像に対して追跡実験を行い、時間状況グラフを用いないパーティクルフィルタによる選手追跡と、提案手法の時間状況グラフを用いたパーティクルフィルタによる選手追跡を比較した。その結果、従来手法に比べて提案手法の追跡精度は7.35 ポイント向上した。図2は、入力映像から、選手とボールを抽出処理する流れである。この処理の後、上に述べた時間状況グラフとパーティクルフィルタにより、選手とボールを追跡している。

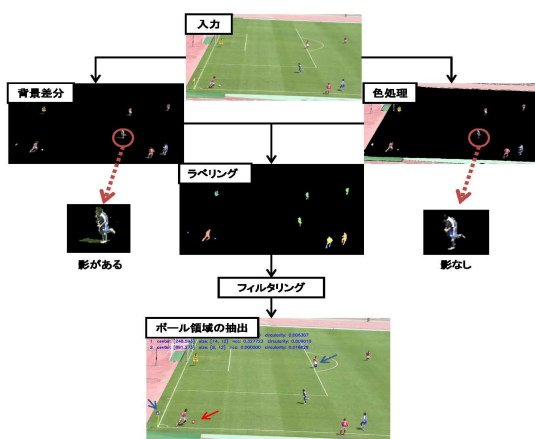


図2 選手とボールを抽出する処理の流れ

(2)(E)状況分析の学習について研究を行った。「ゴール」や「コーナーキック」といった試合のイベントに関する用語をリストアップし、この用語が用いられている映像を多

数収集した。これを基に、イベントが発生した時点、ホイッスルの存在とボールがラインを割ったかどうかの判定により検出する方法を提案し、適合率94.3%、再現率 96.1%を得た。

(3)(B)状況の言語表現について研究を行った。「ゴールキック」、「コーナーキック」、「フリーキック」、「スローイン」といったイベントをHMMで学習して認識を行い、言語出力する方法を提案した。この結果、「ゴールキック」、「コーナーキック」、「フリーキック」、「スローイン」それぞれを、81.6%、72.0%、80.9%、82.5%の精度で認識することができた。

(4)(D)の状況の映像表現について研究を行った。昨年度までの研究により、「ゴールキック」、「コーナーキック」、「フリーキック」、「スローイン」に関して状況認識ができていたので、その状況を分かり易く再現するために、仮想的なカメラワークを付与する研究を行った。

(5)(C)状況の解説について研究を行った。状況を形成している個々の選手とボールの関係や、どの選手からどの選手にパスが渡されたか、などの情報が重要である。この情報は、「ゴールキック」などの状況認識で用いているボールの位置と速度、選手の存在位置の情報だけでは、取り出すことができない。そこで、選手とボール、選手と選手の関係、ボールの移動について、グラフ表現する研究を行った。

(6)この挑戦的萌芽研究を通して、4台のHDカメラでサッカープレーを撮影し、選手とボールを自動抽出・追跡できるようになった。また、その結果得られた選手とボールの情報を基に、「ゴールキック」、「コーナーキッ

ク」、「フリーキック」、「スローイン」という4つの状況を、HMM(隠れマルコフモデル)で学習することにより、これらを高精度に認識できるようになった。図3は、これを実現するためのシステムの構成図である。この結果を用いて、デジタルカメラワークや解説生成の研究を行い、ルールを自動学習する展望を得ることができた。

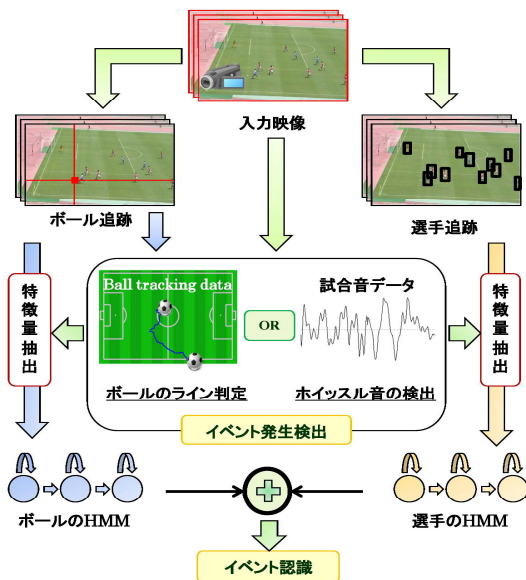


図3 イベント検出・認識システムの構成

(7)本研究の内容を体系化することにより、アナウンサー・エージェントとカメラマン・エージェントをソフトウェアとして構成できる。これにより、今後、人の社会生活を映像として記録しておくだけで、そこでの状況を自動認識して言語化できるだけでなく、プレーなどの分析やアドバイスもできる知能カメラが構築できる。このため、映像のアノテーションや検索が可能になるという効果が期待できる。

<引用文献>

Y.Ariki, T.Takiguchi, K.Yano, Digital Camera Work for Soccer Video Production with Event Recognition and Accurate Ball Tracking by Switching Search Method, ICME 2008, WE-PM2-L4.1, 2008, 889-892

P.J.Figueroa, N.J.Leite, R.M.L.Barros, Tracking soccer players aiming their

kinematical motion analysis, Computer Vision and Image Understanding, Volume 101 Issue 2, 2006, 122-135

5. 主な発表論文等

[雑誌論文](計2件)

有木康雄、ボールと選手のHMMを統合したサッカー映像のイベント認識、画像ラボ、査読無、9月号、2015

伊藤拓紀、滝口哲也、有木康雄、単眼サッカー映像における時間状況グラフを用いた選手追跡、電子情報通信学会論文誌D、査読有、Vol.J96-D、2013、1854-1864

[学会発表](計5件)

王鶴錦、伊藤拓紀、滝口哲也、有木康雄、ボールと選手のHMMを統合したサッカー映像のイベント認識、画像の認識・理解シンポジウムMIRU2014、2014.07.31、岡山コンベンションセンター(岡山県)

伊藤拓紀、滝口哲也、有木康雄、サッカー映像におけるホイッスル音声情報を利用したイベント検出、画像の認識・理解シンポジウムMIRU2013、2013.08.1、国立情報学研究所(東京)

Hiroki ITOH, Tetsuya TAKIGUCHI, Yasuo ARIKI, Event Detection and Recognition Using HMM with Whistle Sounds, The 9th International Conference on Signal Image Technology, 2013.12.02, Kyoto Terrsa Conference Center(Kyoto)

伊藤拓紀、滝口哲也、有木康雄、単眼サッカー映像における時間状況グラフを用いた選手追跡、画像の認識・理解シンポジウムMIRU2012、2012.08.08、福岡国際会議場(福岡)

Hiroki ITOH, Tetsuya TAKIGUCHI, Yasuo ARIKI, 3D Tracking of Soccer Players Using Time-Situation Graph in Monocular Image Sequence, International Conference on Pattern Recognition 2012, 2012.11.14, Tsukuba International Congress Center (Ibaraki)

6. 研究組織

(1)研究代表者

有木 康雄 (ARIKI, Yasuo)
神戸大学・都市安全研究センター・教授
研究者番号：10135519