

被写体追跡可能なスポーツグラフィックスシステムの試作 —複数の可動カメラを利用したボールの三次元リアルタイム追跡—

大久保 英彦[†] 高橋 正樹[†] 加納 正規[†] 池谷 健佑[†] 三科 智之[†]

[†] 日本放送協会 放送技術研究所 〒157-8510 東京都世田谷区砧 1-10-11

E-mail: [†] {ookubo.h-je, takahashi.m-ju, kanou.m-gc, ikeya.k-ec, mishina.t-iy} AT nhk.or.jp

あらまし 昨今のスポーツでは、競技解析に有用な情報として被写体、特に球技であればボールの正確な時系列位置情報取得に対する要求がある。実際に一部の中継放送で、ボールの移動データを解析し、スポーツグラフィックスという形で提示がなされている。本報告ではバレーボール中継を想定し、ボール追跡専用のカメラを必要とせずに、放送の本線映像を制作する複数の放送カメラの映像解析と、各カメラのパン・チルト・ズーム量のセンサー情報から、ボールの三次元位置をリアルタイムに算出し、実写映像に対して高精度にCGボールを重畳可能なグラフィックスシステムを構築した。

キーワード スポーツ, 被写体追跡, カメラキャリブレーション, リアルタイムグラフィックス

Prototyping of Live Sports Graphics System Equipped with Object Tracking Function —Real-time 3D Ball Tracking utilizing Multiple PTZ Cameras—

Hidehiko OKUBO[†] Masaki TAKAHASHI[†] Masanori KANO[†] Kensuke IKEYA[†] Tomoyuki MISHINA[†]

[†] Science & Technology Research Laboratories, Japan Broadcasting Corporation

1-10-11 Kinuta, Setagaya-ku, Tokyo, 157-8510 Japan

E-mail: [†] {ookubo.h-je, takahashi.m-ju, kanou.m-gc, ikeya.k-ec, mishina.t-iy} AT nhk.or.jp

Abstract In order to enrich the sports TV programs, there are various requests to obtain the useful information during the game for on-air graphics. Especially, for any ball game, time series 3D positional data of the ball is the one of the most important information. We developed a new on-air 3D graphics system for live sports broadcasting, it is equipped with object tracking function, to track a volleyball. The positional data of the ball is calculated in real-time, both from video analysis of PTZ cameras for broadcasting, and from PTZ sensor data. Our experiment shows that the system can produce accurate 3D ball compositing in real-time for on-air graphics.

Keyword Sports, Object Tracking, Camera Calibration, Real-time Graphics system

1. はじめに

本報告は、撮影映像中の特定の被写体を追跡する技術を利用した、スポーツ中継放送用の新しいリアルタイムオンエアグラフィックスシステムに関するものである。昨今のスポーツでは、映像制作技術の進化にともなって、生中継であっても高度な映像表現が求められるようになってきている。それに加え Hybridcast [1]など、放送映像とは別に、競技の関連情報をユーザーに提供するための粒度の細かい競技情報のニーズも高くなっている。競技中の情報として、特に球技では、ボールの正確な時系列位置情報は重要である。放送での利用を考えると、そこからボールの高さや速度が得られるだけでなく、軌跡としてグラフィックス表示されることにより、視覚効果的にわかりやすい表現をもたらすことが可能になる。本研究で我々は、特にバレーボ-

ール競技を対象に、そのような視覚効果を表現できる新たなオンエアグラフィックスシステムの試作を行った。そこで、システムに対する要求条件として以下の5つを設定した：1. 映像機器として番組に即応できる機器構成を備えること。特にスポーツ中継向けに、複数のカメラに対応するグラフィックスシステムとして稼働すること。2. 競技場を想定した複雑な背景下で運用し、その中を高速かつ複雑に移動するバレーボールを三次元で追跡できること。3. 追跡領域としては、ボールが存在する競技空間領域全体をカバーし、映像フレームレートで追跡できること。4. 追跡したボールを、撮影している全カメラの実写映像にCGボールとして合成可能とし、全カメラに対して違和感のない位置精度・時間精度での合成映像生成を行えること。5. 放送映像本線制作に必要とされる機器以外は可能

な限り利用しないこと。

これら5つの要求を満たすシステムを実現するため、ボールを対象とした被写体追跡、パン・チルトおよびズームレンズデータ取得可能な複数の雲台マウントカメラ用のカメラキャリブレーション、ボールの三次元位置算出とCG描画を中心的な要素技術に、すべてリアルタイム処理可能とすべく研究開発を進め、それらを統合したシステムを構築してきた。本報告では、オープンなデモ環境（9m×6mの模擬バレーボールコート）に4台のカメラを設置した実験で、ズームを含めた現実的な演出に即したカメラワークと競技中のさまざまなプレイに対して、提案システムがボールを頑健に追跡し、かつ実写のボールに正確にCGをリアルタイム合成可能であることを確認した。以下、その概要を報告する。

2. オンエアグラフィックスの背景技術と課題

放送番組で利用されるオンエアグラフィックスの中で、実写映像に対して三次元幾何的に整合したCGの合成を行うバーチャルセットと呼ばれる映像表現は、本物の被写体と、自由に制御できるCGオブジェクトを融合することで、高い演出効果を生み出す。提案するグラフィックスシステムでは、さらにその“本物の被写体”を追跡することで、被写体の動きに関連付けたCGオブジェクトの演出を可能とする。これらを実現するために必要な技術は、まず1つ目が被写体追跡技術である。被写体追跡は追跡対象や用途によってさまざまな手法が提案されている。その中でも特にバレーボールを複雑な背景下で追跡するものとしては、[2]や[3]などが提案されているが、処理速度の問題や追跡用カメラの配置条件などから、中継放送での利用にはなじまない。次に必要な技術はカメラキャリブレーションである。これは実際の撮影カメラが、実世界において、どこからどういった姿勢で、またどのような画角で撮影をしているのか、という情報（カメラパラメータ）を取得する技術で、これはそのままCG描画の際の仮想カメラに適用させるパラメータとして利用される。特に放送では、安定したカメラワークとダイナミックな演出を両立させるために、雲台にマウントされたズームレンズ搭載カメラを利用するのが一般的である。さらに、雲台やレンズにロータリーエンコーダを搭載することで、パン・チルト・ズームの情報を取得可能な仕組みも広く利用されている。その場合、ロータリーエンコーダを利用することで、リアルタイムに間接的なデータ取得は可能になるが、そのデータと実際のカメラパラメータを正確に関連付ける必要がある。それを目的としたカメラキャリブレーション手法については、[4]や[5]などがあるが、現場での実地運用が可能なPTZカメラのキャリブレーション手法として

は要件として不足する部分がある。3つ目は、三次元位置算出である。複数の撮影カメラのカメラパラメータが算出され、それらのカメラの映像ごとに画像座標としてボールの位置が得られれば、いわゆる三角測量の要領により、2つのカメラからボールの三次元位置算出が可能となる。この条件下であれば、OpenCVをはじめとするツールキットを利用することで、オンライン処理での三次元位置算出を行うことができる。しかし、バレーボール競技においては、ボールがネットを挟んで行き来し、選手によるオクルージョンも頻繁に生じる上、プレイによって求められるカメラワークが変化していくため、三次元位置算出のためにあらかじめ2つのカメラを指定しておく運用は現実的ではない。また精度面を考慮すると、3つ以上のカメラによるボール位置算出を行うことが求められてくる。すなわち、競技状況で時々刻々ボールをとらえるカメラが変わっていくことへの対応が必要になる。このように、変則的なデータ対応を行いながら、高精度な三次元位置算出を実現することが課題となってくる。

3. 提案システムの構成

提案システムは、前章で述べた各課題に対処した要素技術を統合したものになっている。それに加え、後述するデモンストレーション環境や放送利用に適合するための要素技術も盛り込んだ構成をとっている。以下に各要素の概要を述べる。

3.1. カメラシステム

ロータリーエンコーダ搭載の遠隔操作可能なズームレンズ搭載雲台マウントカメラ（PTZカメラ）として、多視点映像制御用のロボットカメラ[6]を利用した。本来このカメラシステムは、取り囲むように設置した20台程度のPTZカメラに対して、一人のカメラマンの操作により、指定した被写体をフレーミングすることを集中制御で行えるように開発されたものである。今回は放送本線用演出カメラ（兼ボール追跡用カメラ）として4台を使用するコンパクトな構成をとった。

3.2. 被写体（バレーボール）追跡

追跡処理では、パン・チルト・ズーム操作を自由に行う4台のカメラそれぞれにおいてボールの二次元画像座標値を算出する。後述する三次元位置算出のプロセスにおいてカメラパラメータが常に利用できるため、動くボールを動くカメラで捉えながら、その画像座標値を取得することになる。ボールは図1に示す青と黄色で構成されるバレーボールのFIVB公式試合球（Mikasa MVA200）を対象とする。そのボール追跡には、複雑な背景下における頑健な追跡を可能にするため、機械学習を利用する。まず学習フェーズにおいては、事前にボール・非ボールのRGB画像を収集し、RGB, HSV色空間ヒストグラム、およびLBP（ローカ

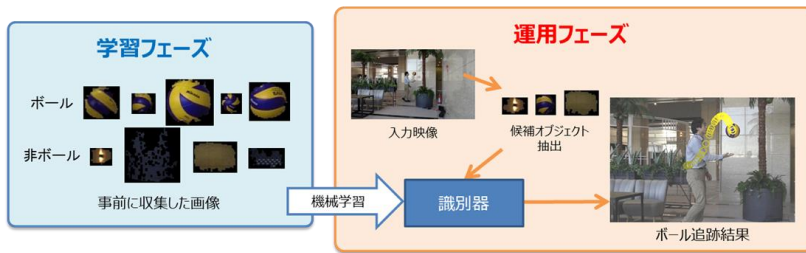


図 1. バレーボール追跡処理の概要

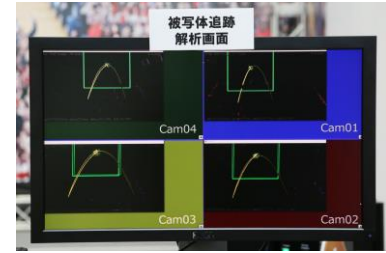


図 2. ボール追跡処理モニタ画面（4台分）

ルバイナリーパターン）からなる特徴量を構成して SVM（サポートベクターマシン）による学習を行う。特に特徴量は、バレーボールの持つ2色の情報を有効に利用するよう構成する。実際に追跡を行うフェーズでは、入力される各映像フレームから色情報をベースにボール候補となるオブジェクト領域を抽出し、それを識別器にかけることでボールを決定、得られた領域から尤もボールの中心と思われる位置を算出して画像座標とする。その上でボールの探索ウィンドウを導入、前時刻フレームでのボール位置をベースに、画面上に占めるボールの大きさや時系列位置変化に合わせてウィンドウ位置とサイズを自動調整する機構を加えることで、処理の高速化や外乱抑止効果を図った。（図 2）

3.3. カメラキャリブレーション

本システムで利用しているキャリブレーション手法は、PTZ カメラを数値モデル化し、市松模様などのキャリブレーションパターンを撮影した画像と、パン・チルト・ズームの各エンコーダデータを総合して、非線形最適化処理によりモデル化したパラメータを算出する手法[7]となっている。まず事前にズームによる焦点距離変化を含めたカメラモデルの各パラメータのデータベース作成を行う。撮影現場にカメラシステムを持ち込んだ際には、それぞれの位置決めを行った後に、キャリブレーションパターンを利用した“設置校正”処理[5]をすべてのカメラに対して同時に行うことで、現場設置したカメラの位置・姿勢に関するパラメータのみをアップデートする。オンエア時は、カメラに映っている映像内容に依存することなく、パン・チルト・ズームのエンコーダデータから、常に高精度なカメラパラメータをリアルタイムに算出可能となる。

3.4. ボールの三次元位置算出

前記 3.2, 3.3 の処理から、4台すべてのカメラに関する高精度なカメラパラメータと、被写体であるバレーボール位置の画像座標が、全カメラ分、かつ毎フレーム取得できることになる。ここでの処理としては、まずキャリブレーションによって得たレンズ歪パラメータから、歪のない場合のボール位置座標算出を行う。そして、ボールを映像フレーム内にとらえている2～4台（可変）のカメラパラメータとカメラに対応するボールの無歪二次元画像座標値から、多焦点テンソル

のソルバを利用してボールの三次元位置を算出する。本処理ではソルバとして、“最適補正による測量”[8]に基づいたボールの三次元位置算出処理をオンライン処理向けに実装した。これにより、ボール捕捉カメラの変化に対応した高精度な三次元位置のリアルタイム算出を可能とした。なお、ボールは三次元位置で算出されるので、ボールがオクルージョンなどでとらえられていないカメラにも CG 合成することが可能になる。

3.5. システムの時間的同期処理

正確な三次元位置算出と CG 合成のためには、処理する映像の時刻に対して、適用するデータが同時刻であることが保証されなければならない。マルチカメラを利用する中継放送の場合、必ずすべてのカメラは元となる映像同期信号でロックされるため、パン・チルト・ズームの値取得も同じ映像同期をトリガに取得するように設計を行った。また、映像・データに SMPTE タイムコードを付与させることによりコンピュータ上の処理では、これをガイドにバッファリングをコントロールし、絶対的な時間の整合をとるようにしている。

3.6. CG描画と実写映像への合成

実写映像への CG 合成は、HD-SDI 映像信号の入出力機能を備えたグラフィックスカードを搭載した PC 上で開発した OpenGL アプリケーションで行った。まず実写映像をライブテキストチャとして背景描画し、前景として被写体となる CG ボールを算出した三次元位置に配置、実写映像を撮影しているカメラのカメラパラメータを利用して描画を行うことで実写合成を完了させた。なお前景 CG シーン描画の際には、キャリブレーションで得られたレンズ歪パラメータに基づいた光学歪をオフスクリーンレンダリング経由で加えることで、全画面領域における実写との幾何的整合をとった。合成映像は HD-SDI で出力可能であり、そのままオンエアグラフィックス用映像システムとして利用できる。

4. 実験と考察

提案システムを NHK 技研公開 2017 にて実機のデモ展示を行った。バレーボールの実演エリアとして、9m × 6m のエリアに小さなバレーコートを展開し、実際に選手たちが試技を行う中、4台の PTZ カメラを操作して撮影（図 3）、ボールの追跡から映像合成までをリアルタイムで行った。



図 3. デモ環境およびシステム外観



図 4. 映像合成結果出力例．実ボール位置には NHK ロゴ CG を重畳，軌跡は 3 フレーム間隔でワイヤーフレーム描画した．右下に計測スピードを表示した．

なお，実験で使用した主な機材スペックや実験パラメータを表 1 に示す．

映像	1920x1080 59.94i (シャッター 1/250)
ズームレンズ	4.7mm - 94mm (20 倍)
エンコーダ パルス	パン・チルト: 800,000 パルス / 360° ズーム: 1,025 パルス (0-1024)
ボール 学習データ	正例: 400 枚, 負例: 400 枚

表 1. 実験時の主な機材スペックやパラメータ

デモにおける合成結果例を図 4 に示す．本システムは，複雑な背景の下，映像本線制作用の PTZ カメラのみでボールを追跡し，合成までをおこなうことができた．すなわち，要求条件 2 と 5 は満たすことができた．システムとしても映像システムとして動作し，要求条件 1 にも適合している．精度面では，事前のカメラキャリブレーションにおいて，全 4 カメラでサンプルしたキャリブレーションパターンの再投影誤差は，平均で 1.2 ピクセル以下という高い精度結果であった．一方，ボールの合成に関しては，カメラの操作によって，2 台以上のカメラでバレーボールをとらえられていれば，被写体のボールと CG が，画面位置上そして時間方向にもずれることなく視覚的にはぴったりと合った合成映像として出力できた．このことから，要求条件として挙げた違和感のない合成は達成できていると考える．数値的には，収録した映像から，合成対象カメラやズームの画角，ボール位置がさまざまになるように映像フレームを適当にサンプルし，実写ボール位置と CG ボール位置をピクセル座標で比較すると，30 サンプル平均で約 4.3 ピクセルの誤差であった．これは，ボールの一部が映っていないフレームがあったり，またズームレンジがテレ（望遠）側でキャリブレーション精度が落ちる傾向があること，さらにボールをとらえるカメラの数が 4 台の時に比べ，2 台の場合には特にベースラインに直行する奥行方向の精度低下が生じるため，ピックアップする画像サンプルによって数値上では比較的大きな誤差として観測されたためと考えられる．映像システムとしては，画像座標での評価が重要になるが，用途によっては厳密に要求条件 3 を評価するために，三次元的な位置での誤差評価が今後必

要になると考えられる．なお，CG 合成出力は，入力映像に対して 10 フレーム（約 0.3 秒）程度のディレイで出力され，約 5 日間のデモを通して，描画についてフルフレーム（59.94fps）でコマ落ちも観測されず，安定に映像出力ができた．

5. まとめ

実験結果から，バレーボールを対象とした三次元追跡・合成システムとして見た場合，精度やシステムの信頼性を含め，スポーツのオンエアグラフィックスとしての利用においては，十分適用可能なシステムが構築できたと思われる．今後は実際の試合会場などでの適用実験を通して，さらなる被写体追跡の頑健化と精度向上などを進め，高い実用性をもったシステムとして発展させていきたいと考えている．

文 献

- [1] Hybridcast : <http://www.iptvforum.jp/hybridcast/>
- [2] M. Takahashi, K. Ikeya, M. Kano, H. Okubo, and T. Mishina, "A Robust Volleyball Tracking System using Multi-View Cameras," In Proceedings of International Conference on Pattern Recognition (ICPR), pp.2741-2746 (2016).
- [3] X. Cheng, N. Ikoma, M. Honda, T. Ikenaga, "Multi-View 3D Ball Tracking with Abrupt Motion Adaptive System Model, Anti-Occlusion Observation and Spatial Density Based Recovery in Sports Analysis", IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences Vol.E100-A No.5 pp.1215-1225 (2017).
- [4] Willson R. G., "Modeling and Calibration of Automated Zoom Lenses", PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, (1994).
- [5] 加納正規 他, 「現場設置を考慮した可動式多視点カメラの校正手法」, 映情学年次, Vol.24A, No.2, (2015).
- [6] K. Ikeya, Y. Iwadate, "Multi-Viewpoint Robotic Cameras and their Applications", ITE Transactions on Media Technology and Applications Vol. 4, p. 349-362,(2016).
- [7] 加納正規 他, 「センサを用いた複数のパン・チルト・ズームカメラの校正手法」, 第 20 回画像の認識・理解シンポジウム(MIRU2017), PS3-03, (2017) (to be published).
- [8] 矢野直樹 他, 「直交射影による複数画像からの最適な三角測量」, 情報処理学会研究報告, 2008-CVIM-165-6, 2008.11.27-28, 35-42 (2008).