

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 6 日現在

機関番号：13401

研究種目：基盤研究(C) (一般)

研究期間：2012～2015

課題番号：24500199

研究課題名(和文) 場を共有する手持ちビデオカメラ映像群からのカメラワーク合成

研究課題名(英文) Camerawork synthesis from multiple videos of handheld video cameras under space sharing situations

研究代表者

東海 彰吾 (TOKAI, SHOGO)

福井大学・工学(系)研究科(研究院)・准教授

研究者番号：50283627

交付決定額(研究期間全体)：(直接経費) 4,100,000円

研究成果の概要(和文)：本課題では、場を共有する複数の手持ちカメラで撮影されたシーン状況を分かりやすく映像化する手法としてカメラワーク要素の導入を考え、特に被写体の動きの可視化に注目した映像化手法について研究した。カメラワークの要素としては、被写体を捉える方向と画面内での位置・大きさの制御、さらにそれらの時間軸制御までを扱い、キーフレーム毎に設定した値からの補間によって生成する各フレームでのパラメータ値を獲得する手法を考案した。また、撮影映像のパノラマ化を経由した射影変換処理とトリミング処理によって、所望のフレーム群を生成する画像処理手法を開発・実装し、スポーツ状況でのカメラワーク映像生成により有効性を確認した。

研究成果の概要(英文)：In this theme, I considered employing the camerawork features to express a scene situation as attractive and intelligible video sequences combined from multiple handheld video cameras by spectators in the scene under space sharing situations. Here, I designed camerawork features as a setting of observing angles, position and size of a shooting object on the video, and their time series variations with considering to show its movement in the scene such as sports situation. I adopt key frames and interpolations of the camerawork parameters for final video sequence production. I developed a series of video processing method from original video sequences by handheld cameras to the final camerawork video sequence through the panoramic image synthesis by using configurations of camerawork features for each key frame. And, I also showed an effectiveness of the proposed method using in an actual sports scene.

研究分野：映像処理

キーワード：カメラワーク キーフレーム 画面構成制御 時間軸制御

## 1. 研究開始当初の背景

映画制作などでは、同時に複数カメラで撮影した映像群が利用される。シナリオや撮影計画に従って各カメラの役割やカメラ操作などのカメラワークが事前デザインされ、これに従って全カメラが共同で撮影し、映像群をつなぎ合わせることで、人の観賞に耐える映像コンテンツが生まれる。

一方、近年の高性能で安価なビデオカメラや携帯端末の普及により、高画質な映像を一般の人が自由に撮影できるようになり、個人での映像編集や、配信・公開が当たり前となってきた。さらに、スポーツ観戦やコンサートイベントなどで、同じ場所に集まった観客らが、一つの状況を同時に複数の視点から撮影した多視点映像が撮影・記録されている。

しかし、それらの映像の多くはそれぞれの撮影者の個人的意図や目的で撮影され、個別に利用されているに過ぎず、一般的なシーンでの多視点映像取得状況では、映画などと同じような事前デザインに基づく撮影・編集は容易ではない。

このような状況において、もしも非常に多くの視点からの多数の映像が撮影されていれば、これらから適切な視点や撮り方の映像を選択し、加工・編集することで、個人用途の単一視点映像素材からだけでは難しかった、それぞれのシーン状況に応じて、それらをわかりやすく見せるための多種多様なカメラワークを導入した映像コンテンツの制作が期待でき、カメラワークによる映像制作者の意図や、映像利用者の要求を反映したカメラワークの選択なども可能となる。

## 2. 研究の目的

前述の背景を踏まえて、本研究では、スポーツ観戦の観客状況のように、場を共有する複数の手持ちカメラによって、それぞれ個別に撮影された映像群にカメラワーク要素を付加しながら収集・加工・編集してそれらを統合することで、利用者にとって意味のある映像を制作するための計算機支援の方法を開発することを主な目的とする。

まず、撮影の状況として、必ずしも意図が事前に調整されていない複数の撮影者によって撮影された映像群を収集・加工・編集することを考え、「利用の際に意味のある」、「シーン状況を正しく伝えられる」、「分かりやすく見飽きない」、「制作者の意図が伝わる」映像コンテンツの制作を考える。このとき、特に、複数視点の位置関係と、それらの映像で捉えられた同一被写体を抽出し関係付けることで、視点切替えを含む多様なカメラワー

クを付加し、シーン状況理解が容易になるような提示映像を獲得するための一貫した処理の流れを設計し、映像編集に対する計算機支援の方法を開発する。

## 3. 研究の方法

本課題では特に「対象となる被写体がそれぞれの撮影映像のどの位置に、どのように捉えられているか」に着目することとし、シーン内での同一の被写体を、画面内での位置や大きさ、さらに、カメラの方位情報から算出して対応づけるアプローチをとることとした。これは、各映像に対する処理を軽減し、より簡易な画像処理・認識技術で実現できること、また、カメラワークを付加する際に、被写体やその動きに対する画面構成やカメラ操作を適切に設定する際に直接的に関わると考えられたためである。

そこで、(1) それぞれのカメラの動的な位置や向き、ならびに、撮影時刻の情報を獲得すること、(2) 映像に写った被写体を抽出し、他の映像の被写体と関係づけること、(3) 同じ被写体が撮影されている映像群を利用して、適切なカメラワークを付加することの3点について、実際のカメラ群での撮影から最終的なカメラワーク付加映像の制作までの一貫した機能の整備を考えることとした。ここでは、バスケットボールなどのスポーツ状況を具体例として、選手やボールを被写体とするような状況設定の中で研究を進めた。

具体的には、まず項目(1)として、手持ち撮影された多視点映像に対して、パノラマ映像合成を利用した仮想広角画像の生成と撮影フレームの統合によるカメラ姿勢の抽出を行った。続く項目(2)については、パノラマ合成像との背景差分による被写体領域抽出と、多視点間の情報統合による被写体位置推定、ならびに、逆投影に基づく撮影映像上の被写体領域の確定、さらに項目(3)として、キーフレーム毎に設定したカメラワーク要素に基づいて、多視点映像からのカメラの切り替え選択、および、指定の画面構成で被写体をとらえるような射影変換とトリミングを組み合わせた画像の加工を行い最終的なカメラワーク込の提示映像を得る方法を開発するものである。

## 4. 研究成果

本研究課題の成果として、撮影後のカメラ校正処理、被写体抽出処理、カメラワーク要素の設計とキーフレームによる設定方法、ならびに、それらの手法を用いた撮影実験とそ

の結果についてそれぞれ説明する。

### ①カメラ校正と被写体との関係付け

まず、手持ち撮影された多視点映像について、それぞれのカメラの位置、姿勢の校正を考える。ここでは、スポーツ状況の観客視点を想定し、カメラの位置はほぼ動かないとみなし、視線方向のみが変化するものとする。視線方向のみが変化する一連の首振り撮影映像から、映像の各フレームの特徴点对应付けに基づいた射影変換によってパノラマ画像を生成する。被写体が明らかに映っていない部分画像を組み合わせることで、被写体が含まれない背景パノラマ画像を得る。さらに、シーン内の3次元座標が既知の点の投影位置に着目してカメラ位置を推定する。また、被写体を含む首振り撮影映像の各フレームについて、画像上の特徴点の対応付けに基づいた背景パノラマ画像との重畳処理を行い、パノラマ画像に当該フレームを埋め込むパノラマ映像化処理を行う。被写体を含むパノラマ映像のフレームと背景パノラマ画像との差分を利用して被写体候補領域を抽出し、それらの逆投影によって、各フレーム時刻における被写体の存在位置を推定する。パノラマ画像合成の処理例を図1に示す。



(a) 撮影フレーム



(b) 背景パノラマ



(c) パノラマ統合結果

図1 カメラ校正とフレーム抽出のためのパノラマ合成処理の例

なお、撮影映像の全フレームをパノラマ化しておくことで、手持ちにより首振り撮影さ

れた映像を固定カメラとして扱えるようになり、首振り撮影による視線方向の変化が無視できる。さらに、後述する画面構成制御を含む画像の加工の際に、撮影映像の周囲を背景パノラマで補間することができるため、提示映像の加工の際の違和感の低減にも有効に働く。

### ②キーフレームを用いたカメラワーク要素の設定

一般の映像作品で用いられるカメラワークには様々なバリエーションがある。ここでは、カメラワーク要素として、被写体を捉える画面構成と被写体を捉える方向の要素を考えることにする。具体的には以下の通りである(図2)。

- アングル ( $\theta$  [deg]): 被写体をとらえる方向を設定する。被写体正面からのなす角  $\theta$  により設定し、 $\theta=0$  なら正面、 $\theta=90$  なら被写体左側から捉える設定となる。また、多視点撮影環境下では、カメラの切り替え選択を伴う。
- ポジション ( $x$ [%],  $y$ [%]): 提示映像画面の縦横サイズを基準として、被写体の画面上の位置を設定する。(50%, 50%) なら画面中央に、(80%, 20%) なら画面右上に被写体を配置する画面構成となる。
- フレームサイズ ( $s$ [%]): 被写体が画面に占める大きさを、提示映像画面の縦サイズを基準とした割合で設定する。200% なら半身を拡大したショットに、50% なら全身を捉えたロングショットの設定となる。

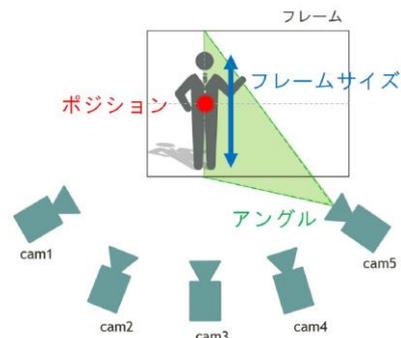


図2 カメラワーク要素

さらに、これらの時間変化を制御することで多様なカメラワーク設定を可能とするため、複数のキーフレーム毎に上記のカメラワークパラメータを設定し、その補間によって最終的な提示映像の各フレームのカメラワークパラメータを算出し、後述するカメラワーク生成処理を行う。なお、キーフレームでは、最終映像上のフレーム数での設定(シー

ン時間)と被写体シーンの実際の経過時間(実時間)の両方を設定することとし、早回しやスローモーションなどの時間軸制御を伴うカメラワークにも対応することとした。時間軸設定の例を図3に示す。図3の①の部分では、実時間とシーン時間が同じで通常の撮影映像提示と同様であるが、②の部分ではシーン時間に対して実時間が短いスロー再生部分、③の部分では逆に実時間に対してシーン時間が短い高速再生部分、④の部分ではシーン時間の進行に対して実時間の進行を止め、静止したシーン状況に対して各種のカメラワーク変化のみを施すような映像が得られることになる。

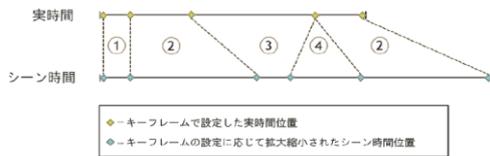


図3 時間設定と再生速度制御の関係

このように設定したキーフレームと次のキーフレームの間の時区間が提示映像を構成するサブシーケンスとなる。それぞれのサブシーケンスにおいて、どのパラメータをどのように変化させるかをキーフレームで設定することで、連続性を持つカメラワークが実現できる。

### ③カメラワーク映像生成処理

各フレームに対して算出されたカメラワークパラメータに基づいて、提示用フレームの画像を処理し、最終的な映像を構築する。まず、今処理したいフレームに対応する映像上のフレームを抽出する。実時間とシーン時間の設定を反映するため、 $i$ 番目のキーフレームの実時間設定を  $t_i$ 、シーン時間設定を  $t'_i$  とすると、当該フレームのシーン時刻  $t'$  に対する撮影映像のタイムスタンプは、

$$\frac{t_{i+1} - t_i}{t'_{i+1} - t'_i} (t' - t'_i) + t_i$$

で与えられる。

次に、当該フレームにおける被写体の3次元位置情報と各カメラの設置位置情報、および、アングル設定に基づいて、カメラの選択を行う。ここでは、被写体の動きを可視化するカメラワークを考えることとし、被写体の移動経路  $\mathbf{P}(j)$  から算出される当該時刻の被写体の移動方向  $\mathbf{d}(j) = \mathbf{P}(j) - \mathbf{P}(j-1)$  ベクトルと、当該フレームのために補間されたアング

ル設定角  $\theta$  から、設定アングルで捉えるための理想的なカメラ観察方位直線を導き、その直線に最も近いカメラを選択する方法をとっている(図4)。

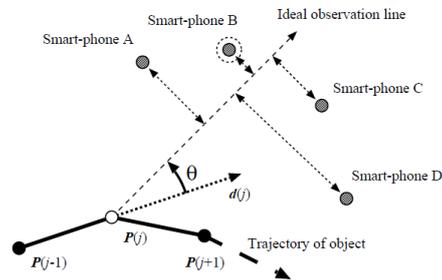


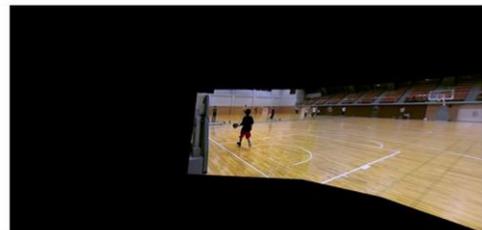
図4 被写体の移動軌跡とアングル設定を考慮したカメラ選択

次に、選択されたカメラに対して、当該時刻のフレームを選択する。各フレームはあらかじめ前述の方法でパノラマ統合処理されているものとし、以下の2段階の手順でカメラワーク要素を持つフレームを生成する。

(1)被写体中心のパノラマ画像変換：被写体の3次元位置と当該カメラの位置、およびパノラマ合成画像として姿勢情報から、パノラマ画像上の被写体像の位置を求める。この位置を画面中央に捉えなおすように、射影変換を施す。射影変換では、投影中心位置が不変なパン・チルト回転のみを想定する(図5)。



(a) 変換前



(b) 変換後

図5 被写体を中央に捉えるパノラマ画像の射影変換処理例

(2) 画面構成設定に基づくトリミング処理：被写体中心のパノラマフレームに対して、トリミング処理によって最終的なフレームを生成する。ここでは、ポジションとサイズの設定に基づいてトリミングのパラメータを決定する。サイズについては、当該被写体の3次元空間中での大きさを既知と考え、カメラと被写体との距離、および、投影変換の際に利用される焦点距離の値に基づいて、拡大縮小のパラメータを決定することで行う。

これらの処理によって、カメラワーク要素設定を反映した各フレームが生成され、それらを連続的につなぐことで最終的なカメラワーク込みの提示映像を獲得する。

#### ④実験とカメラワーク生成結果

これらの処理を、実際のスポーツ状況に適用し、手法の有効性を確認した。具体的には、バスケットボールコート周囲に20人の撮影者を配置し、それぞれにAndroid端末（SAMSUNG 製 Galaxy Camera(EK-GC100)とSHARP 製 AQUOS PHONE (SH-12C)を半数ずつ使用）を操作してもらいながら、コート全体を8の字ドリブルする選手を多視点撮影した。映像は1280x720画素の30FPSの3GP映像として獲得された。実験状況を図6に示す。



図6 バスケットボール状況の撮影実験の様子

次に、それらの映像を元に、各カメラ視点毎に背景となるパノラマ画像を合成し各カメラの校正を行った。さらに、撮影された被写体映像のフレームを統合することで、撮影映像のパノラマ化を行い、背景との差分を利用して被写体の位置とその動きを抽出した。図7にカメラ配置の校正結果と被写体の移動軌跡を示す。ほぼ正確なカメラ位置推定が行われており、また、実際の被写体選手の移動軌跡が十分な精度で獲得できていることが分かる。この結果に基づいて、各カメラにおける被写体の画面内の位置や大きさを算出して、カメラワーク付加処理に利用する。

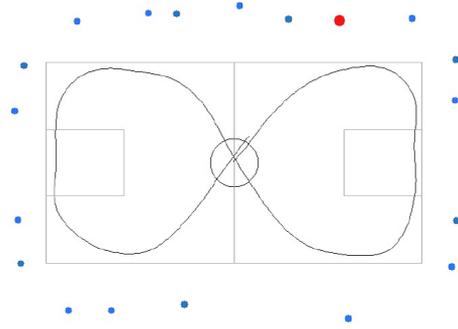


図7 カメラ配置の校正結果と被写体軌跡の抽出結果の例

このシーンに対して、いくつかのキーフレームに対して、前述のカメラワーク要素を設定した。この設定に基づいたキーフレーム間の補間によって、最終的な映像の各フレームについて、カメラの選択と、画面構成加工のパラメータを確定し、射影変換を行うことで、最終的な提示映像のフレームが生成される。具体的な例を図8に示す。

図8(a)は、アングル $\theta=180$ に設定して作成した映像である。被写体となるドリブルする選手の後方から追跡するようなカメラワークが実現されていることが分かる。選手の移動方向の変化に応じて、より適切な位置のカメラへと切り替えながら映像が構成されている。図8(b)は、ポジションのパラメータを(90%, 50%)から、(10%, 50%)に変化させることで、画面内での被写体の位置を右から左に移動させるようなカメラワークとなっている。図8(c)では、サイズのパラメータを80%から20%に変更することで、ズームアウトしていきようなカメラワークとなっている。



(a) 被写体の後方からのカメラワーク例 (アングル設定によるカメラ切り替え)



(b) 画面内位置変化のカメラワーク例 (ポジション設定による仮想的なパン)



(c) ズームアウトを伴うカメラワーク例 (フレームサイズ設定での画面構成変化)

図8 生成されるカメラワーク映像の例

## ⑤まとめ

これらの成果により、場を共有する複数の撮影者（観客）の操作によって手持ち撮影された映像群から、利用者がキーフレームで指定したカメラワーク設定に基づいて、最終的なカメラワーク映像の生成に至る一連の処理が具体的に構築され、スポーツ状況を対象とした撮影映像からの実カメラワーク映像生成を通じて、提案手法が有効に機能することを確認した。

今後の課題として、より複雑なシーン状況に対して、効率的に機能するアルゴリズムの開発や、カメラワークデザインを行う段階での適切な作業支援の方法が必要と考えられる。

## 5. 主な発表論文等

〔学会発表〕（計5件）

① 山本詩織, 東海彰吾, キーフレーム設定に基づくシーン状況に応じた多視点映像からのカメラワーク映像生成の一手法, 電子情報通信学会スマートインフォメディアシステム3月研究会, 信学技報, Vol. 115, No. 505, SIS2015-57, pp. 55-58, 2016年3月10日, 東京都市大（東京都世田谷区）

② 山本詩織, 東海彰吾, カメラワーク映像生成におけるパラメータ設定法の改良, 平成27年度電気関係学会北陸支部連合大会, 発表番号 F2-30, 2015年9月13日, 金沢工業大学（石川県野々市市）

③ 山本詩織, 東海彰吾, キーフレーム設定に基づく多視点映像群からの仮想カメラワーク映像合成, 映像情報メディア学会メディア工学研究会 10月研究会, ME2014-9, 映像学技報, Vol. 38, No. 39, pp. 9-12, 2014年10月2日, 下呂温泉小川屋（岐阜県下呂市）

④ S. Yamamoto, S. Tokai, H. Hase, Camera-worked Video Generation from Multiple View Videos with Handheld Smart-phones, Nicograph International 2014, paper\_3\_5.pdf, 2014年5月30日, Uppsala Univ. Campus Gotland (Sweden)

⑤ R. Takeba, T. Shibata, S. Tokai, H. Hase, A Method of Camera Position Alignment on a Gimbal Head for Fixed Viewpoint Swiveling, IWAIT2013, pp. 72-77, 2013年1月8日, 名古屋大学（愛知県名古屋市）

## 6. 研究組織

### (1) 研究代表者

東海彰吾 (TOKAI, Shogo)  
福井大学大学院・工学研究科・准教授  
研究者番号：50283627