

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年4月26日現在

機関番号：13101

研究種目：基盤研究（C）

研究期間：2009～2012

課題番号：21500161

研究課題名（和文） 可動多視点カメラ映像からの演技のパフォーマンスキャプチャ

研究課題名（英文） Performance Capture of Actor from Active Multiple Cameras

研究代表者

山本 正信（YAMAMOTO MASANOBU）

新潟大学・自然科学系・教授

研究者番号：00242397

研究成果の概要（和文）：多視点カメラによるモーションキャプチャの機能の拡張について報告する。新しい機能は次の2点である。（1）カメラの数や設置位置の制限により、どのカメラからでも身体が断片的にしか観測できないとき、断片的な観測情報から全身の動作を測定する手法を開発した。（2）背景のパノラマ画像とカメラをパンさせた画像とを照合させることによるカメラ較正法を提案した。この較正法により測定できる身体の可動範囲が拡大した。

研究成果の概要（英文）：This report describes two new functions of a motion capture system with multiple cameras. First, we develop a method to reconstruct a whole-body pose from a few body parts partially observed by multiple cameras. Second, we propose a camera calibration method which matches a panoramic image of the background with an image from panning camera. The new calibration method allows us to measure human action in a wider range of space.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,100,000	330,000	1,430,000
2010年度	600,000	180,000	780,000
2011年度	600,000	180,000	780,000
2012年度	700,000	210,000	910,000
総計	3,000,000	900,000	3,900,000

研究分野：情報学

科研費の分科・細目：人間情報学・知覚情報処理

キーワード：モーションキャプチャ，多視点カメラ，姿勢推定，機械学習，カメラ較正

### 1. 研究開始当初の背景

身体の動きを測定し、それをディスプレイ上でCG映像として再現するシステムはモーションキャプチャと呼ばれている。モーションキャプチャは物理的な動きの測定から始まったが、身体の動作を対象にしていることを強調する意味で、パフォーマンスキャプチャとも称されている。

市販のモーションキャプチャは、身体にマーカやセンサを取り付けることが多いため、演者に精神的、肉体的負担をかけること

になる。これに対しビデオ映像を使えば、演者に負担を掛けることなく動作の測定が可能である。特に、演者を多方向から撮影すれば、身体の3次元姿勢を正確に推定することができる。

このとき、多関節構造で表された身体モデルを多視点画像と照合させることにより身体の姿勢を測定することができる。多関節構造とは、身体を胴体、腕、脚、頭などの部位に分け、各部位を剛体とし、部位を関節で接続したリンク構造を指す。このモデルを身体

に一致させたとき、各関節の回転角で姿勢を表すことができる。動画像のコマごとに照合を行えば、動作の測定が得られる。

身体モデルの多視点画像への照合による動作測定は、1996年に Gavrila と Davis らが身体の超楕円体モデルを身体像に照合させた頃から盛んになった。多くの研究成果が発表される中で、2008年の SIGGRAPH では、Aguiar らは、身体の高精度な形状モデルを使ったパフォーマンスキャプチャを発表した。この研究では、レーザーレンジファインダのスキニングにより、身体の形状をあらかじめ多面体モデルとして求めておく。多面体モデルの表面は数多くのパッチにより覆われているが、パッチの頂点を移動させることにより、皮膚や衣服の微妙な動きも表現することが可能となった。

## 2. 研究の目的

多視点画像からのモーションキャプチャとして数多くのシステムが提案されている。しかし、いずれも限定されたスタジオでの測定に留まり、実際に、演劇やミュージカル等での演技や、リングや土俵などでの格闘技のキャプチャが行われた例は無い。

これまでのシステムの問題点は次の2点に集約される。

- (1) 測定できる動作や姿勢に制限がある。
- (2) カメラが固定されていることが多く測定範囲が狭い。

本研究では、これらの問題点を解決し、実際に舞台やリングで演技者の動作を測定することを目的とするものである。得られた測定結果から、演技のスキルの分析や、演技の再生による映画やアニメなどのデジタルコンテンツの制作が可能となる。

## 3. 研究の方法

多視点画像から胴体の姿勢を精密に測定するためには、胴体の輪郭が各カメラから観測されていることが必要である。そのためには腕部の取るべき姿勢にかなり厳しい条件が課せられてきた。例えば、Aguiar らの研究では、演技者のスカートの襷の揺れまで測定できているが、胴体のシルエットがカメラから明瞭に観測できるように、腕が胴体に接触するような動作は禁じられている。また、Sundaresan と Chellappa らの研究は、多視点画像から視交差により得られた身体のボクセルデータに対し、ラプラシアン固有空間法を適用し胴体の体軸（背骨）を求めているが、腕は胴体に原則として接触しないことが前提である。仮に接触したとしても手先を腰に当てることは許されるが、肘は胴体と離れループを構成することが条件とされている。これらが、測定できる動作や姿勢に対する制限となっている。

身体の胴体部は腕により部分的に隠されることが多い。加えて多くの衣服は胴体を覆い隠し、外見から生身の胴体形状を知ることは難しい。一方、腕や脚は比較的その形状や体軸を見て取ることが可能である。

このように、身体は部位によっては、その形状を捉えることに難度が異なっている。これは、多視点カメラの配置や動作の種類にも依存する。本研究では、部分的に観測できる身体部位の断片情報から、身体の姿勢や動作を推定する手法を提案する。

二つ目の目標として、身体動作の測定範囲を拡大する方法について述べよう。測定範囲を拡大するためには、

- (1) カメラを動かす
- (2) 広角カメラを利用する
- (3) カメラの台数を増やす

の三通りが考えられる。

(1)の方法はカメラの位置・姿勢の制御が必要であり、そのためにはカメラを可動させるシステムに費用が嵩むという欠点がある。

(2)の方法は、広角レンズあるいは全方位カメラを利用することにより、カメラを可動させずに広範囲の撮像空間を確保するものである。短所として身体像の解像度を十分に確保できないことがあげられる。(3)の方法は、1990年代にカーネギーメロン大学で多視点カメラの研究が始まった頃から試みられたが、これまた費用が嵩むこと、カメラ校正の保守に負担がかかること等が短所である。

本研究では、(1)の方法を採用するが、低コストでカメラ校正を行う方法を提案する。提案手法では、カメラの動きをパンのみとする。演技者のいない撮影空間に対し、カメラをパンさせた動画像からパノラマ画像を作成しておく。パノラマ画像の画素点にカメラ光軸が向かったとき、画素点位置からパン角度すなわちカメラ姿勢を知ることができる。演者を撮影しているとき、カメラは演者が常に画像の中心付近に写るように動かされている。したがって、身体の写っていない画像の周辺部をパノラマ画像に照合すれば、照合した位置からカメラの姿勢を知ることができる。

## 4. 研究成果

### (1) 部分的情報からの全身の姿勢推定

断片的にしか知ることのできない身体部位情報から、全身の姿勢を推定する手法を開発した。

身体の部位のうち、腕や脚は他の部位によって隠されることが少なく、衣服は体に密着していることが多いので、その体軸を知ることが容易である。これに対し、胴体部はルーズな衣服に覆われることも多く、腕などに隠されることもあり、その生身の形状を知ることは難しい。

本研究では胴体を胸部と腰部の二つの部に分け、その連結点を四肢の体軸から推定することにした。

対象とする身体の部位サイズ（四肢は長さ、胸部腰部は高さ）を予め測定しておく。測定が難しい場合には、公開されている人体寸法データを利用した。まず、多視点カメラから観測された上腕と下腕の体軸から、その交点として肘の位置が得られる。肘から上腕の長さ分遡った位置が肩の位置となる。同様に股関節位置も求まる。肩や股関節の位置は着衣や筋肉に覆われ、正確に知ることは難しい。これに対し、上腕や下腕などの四肢の部位はその見かけが矩形であり、その主軸として体軸が得られる。体軸から肩関節や股関節の位置を決定する方法は、それらの位置を画像上から見つけるよりは正確であると考えられる。なお、腕や脚が真直ぐに伸びている場合は、肘や膝の位置を決定することが難しいので、手首や足首の位置から身体を遡って肩や股関節位置を求める。

肩関節位置が決まれば、胸部は両肩を通る直線を回転軸とした1自由度に束縛される。腰部も両股関節を通る直線を回転軸とした1自由度に束縛される。したがって、胸部と腰部の連結位置は一意に決定される。しかしながら、この連結位置は正しくないかもしれない。実際、得られた連結位置から胸部と腰部の高さを求めると、予め与えられた高さとも一致しないことが多い。これは、肩関節や股関節の位置が正確ではないことが原因である。そこで、肩関節や股関節の位置を微修正することにより正しい連結位置を求めた。この修正法は局所最適化なので、局所解に陥る恐れがある。そこで、手首と肩関節及び股関節位置の関係から、身体が前屈みか反り返りかを予め判定し初期連結位置を定めておき、局所解に収束しないようにした。

提案手法を使ってレスラーの姿勢推定を行った。リングサイドに6台のカメラを設置しレスラーを撮影した。図1は首投げを掛けているレスラーの体軸を示す。2名のレスラーは互いに絡み合っており、腕や脚がどちらのレスラーのものか判定が難しい。

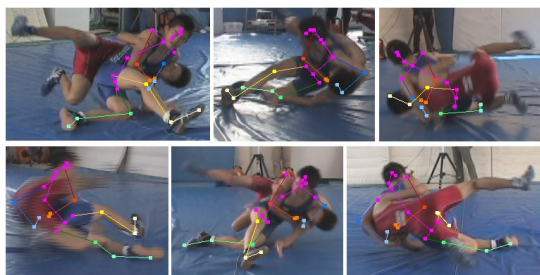


図1 レスラーの体軸

図2にはレスラーを3次元CGとして再構成した結果を示す。



図2 レスラーの3次元CG

動画のフレーム毎に姿勢推定を行えば、動作を測定することができる。しかし、フレーム毎の姿勢推定はコストがかかる。そこで、対象の追跡手法を利用する。初期フレームで身体姿勢が推定されたら、連続するフレーム間の姿勢の変化分は、画像間差分と時空間勾配法により容易に計算することができる。したがって、この姿勢の変化分を初期姿勢に累積して行けば、以降のフレームの姿勢推定が可能である。しかし、誤差も同時に累積されるためドリフトが生じる。そこで、最終フレームでも提案手法で姿勢推定を行い、累積姿勢が最終姿勢と一致するように姿勢の変化分を修正する。実際は、最終フレームだけでは間に合わず、途中の幾つかのフレームで姿勢の推定を行う必要がある。これらのフレームをキーフレームと呼ぶ。

## (2) 身体部位の識別

図1に示される体軸のうち、四肢については手動で与えている。キーフレームは数が少ないとはいえ、手動による体軸指定は負担がかり自動化が望まれる。

身体の写っているサンプル画像に対し、その身体の姿勢が与えられているとき、クエリ画像がサンプル画像の身体像と一致すれば、サンプル画像の身体の姿勢をクエリ画像の身体の姿勢とする。

このような姿勢付のサンプル画像のデータベースを用意しておき、クエリ画像をこの画像データベースと照合し身体部の3次元姿勢を得ることとする。このときデータベース中のサンプル画像とクエリ画像との逐次照合には膨大な計算量を必要とする。そこで、サンプル画像を幾つかのクラスに分類しておき、まずクラスと照合し、ついでクラス内のサンプル画像と照合すれば計算量を節約することができる。

本研究では、サンプル画像を身体姿勢に基づき姿勢クラスに分類する。クエリ画像に対し、姿勢クラス識別器により特定の姿勢クラスが選ばれる。この姿勢クラスの中から最もクエリ画像に近いサンプル画像を選ぶ。サンプル画像数には限りがあるため、選ばれた画像に対応する姿勢がクエリ画像中の身体姿勢からずれている可能性があり、修正が必要である。選ばれた姿勢は実姿勢に近いのでこ

の修正は微調整で済む。

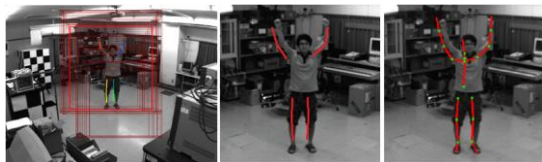


図3 腕や脚の体軸推定

図3左は画像から身体領域を抽出した結果である。身体像を正例、背景を負例としてAdaBoostにより学習した識別器を使用した。図3中では左の身体像の姿勢クラスを判定し、クラスの中から最も近い姿勢の体軸を身体像に重ねている。図3右は中の体軸をチャプターマッチングにより修正し、全身の姿勢を推定した結果である。

### (3) 可動カメラの較正と姿勢推定

カメラを雲台に載せパンさせることにより、広範囲に動き回る演技者の動作を測定することができる。撮影前にカメラをパンさせ背景の動画像を撮影する。この動画像から背景のパノラマ画像を作成する。このとき、カメラの回転中心と視点が一意している必要がある。しかし、カメラモデルの視点はカメラ内部にありその位置は明確ではない。そこで、カメラを雲台上で前後にスライドさせることが可能なプレートを挟み、カメラをパンさせても前方に置いた物体が背景上を移動しないような位置にカメラをスライドさせた。この位置が回転中心と視点が一致する位置である。さらに、透視変換による視野周辺部の歪を無くすために、画像を円筒画像に変換し、パノラマ画像を作成した。パノラマ画像の横軸上の画素位置から、カメラの向きを知ることができる。図4に、4台のカメラからのパノラマ画像をそれぞれ示す。



図4 パンしたカメラのパノラマ画像

図5は歩行中の演技者の一コマである。この画像の周辺画像をパノラマ画像と照合させカメラ較正を行った。なお、画像間照合にはSURFを使用した。

歩行者像から四肢の体軸を取り出し、3次元姿勢を推定することができる。約30mの直線歩行を測定することができた。歩行を捉えるために各カメラは約45度回転している。



図5 歩行動作中の多視点画像

## 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)

[雑誌論文] (計1件)

- ① 渡辺直人、糸雅亮祐、佐藤真悟、三浦裕樹、山本正信、フルハイビジョン映像からの音楽演奏時の身体・手・顔のモーションキャプチャ、映像情報メディア学会誌、査読有、65巻、(2011)、1647-1651  
<http://dx.doi.org/10.3169/itej.65.1647>

[学会発表] (計12件)

- ① M.Yamamoto, Y.Wada, S.Suzuki, A 3D motion capture for outdoor pedestrian, Symposium on Progress in Information & Communication Technology, 査読有, 2010.12.13, Kuala Lumpur
- ② 山本正信、フェロー記念講演：動画像処理30年、電子情報通信学会パターン認識とメディア理解研究会、2010.5.14、中部大学
- ③ 五十嵐崇、鈴木慎也、山本正信、パノラマ画像によるパン可動カメラの較正、電子情報通信学会総合大会、2011.3.17、東京都市大学
- ④ 上田達也、堀澤健太、山本正信、磯野誠也、Asymmetric AdaBoostを用いた多視点カメラからの身体姿勢推定、電子情報通信学会総合大会、学生ポスターセッション、2012.3.21、岡山大学
- ⑤ Y.Oka, T.Kida, M.Yamamoto, A cartoon-character costume with facial expression, The 4<sup>th</sup> ACM SIGGRAPH in Asia, 査読有, 2011.12.14, Hong Kong
- ⑥ 佐藤真悟、山本正信、三浦裕樹、磯野誠也、モーションキャプチャによる熟練技能の教材アニメーションの作成、電子情報通信学会信越支部大会、2012.10.13、新潟大学
- ⑦ M.Yamamoto, S.Isono, Y.Wada, Motion capture in outdoors, The 2<sup>nd</sup> International Workshop on Networking, computing,

- Systems, and Software, 2012.12.5, Okinawa
- ⑧ 三浦裕樹、大沢哲史、山本正信、熟練手作業のモーションキャプチャ、電子情報通信学会ヒューマン情報処理研究会、2013. 3. 14、沖縄
  - ⑨ 佐藤真悟、山本正信、航空機の翼の研磨作業における熟練度の評価、電子情報通信学会総合大会、2013. 2. 21、岐阜大学

[その他]

ホームページ等

新潟大学山本研究室HP

<http://www.vision.ie.niigata-u.ac.jp/>

## 6. 研究組織

### (1) 研究代表者

山本 正信 (YAMAMOTO MASANOBU)

新潟大学・自然科学系・教授

研究者番号：00242397