

2021年度 プレ卒論

任意方向のカメラ移動を伴うモーションキャプチャ
及び端末の自己位置推定の手法

2022年2月10日

システム工学科
(学生番号: 60246336)

高野 源太

和歌山大学システム工学部

概要

スマートフォンなどの端末でモーションキャプチャーを行う場合、カメラが動くことを考えるとモーションキャプチャーと同時に端末自身の自己位置を推定し、位置関係を適応する必要がある。本稿では、卒業研究で開発するシステムについて関連研究である「単眼カメラを用いたモーションキャプチャの手法」と「障害物を考慮した自己位置推定 (v-SLAM) の手法」をもとにシステムの流れ図を作成し、研究手法の提案並びに実装の計画を行う。

目次

第1章	はじめに	1
第2章	関連研究	2
2.1	単眼カメラを用いたモーションキャプチャの手法	2
2.2	障害物を考慮した自己位置推定 (v-SLAM) の手法	2
第3章	使用機材・実装環境・研究計画	3
第4章	提案手法	4

第1章 はじめに

現在のスマホなどの単眼カメラが搭載された端末で行う光学式モーションキャプチャはカメラ位置を固定することが前提となっているものがほとんどであり，端末を動かしながらモーションキャプチャを行うといった利用方法は想定されていない．そのため，屋内での端末の位置を推定し，端末の移動・カメラの向き等を考慮したモーションキャプチャの手法を確立することで，カメラ画角等にとらわれないモーションキャプチャを行うことができ，また，より簡易的にモーションキャプチャを利用することができるのではないかと考えた．

この研究では，部屋等の閉空間において一人の人間がカメラを持ちながらモーションキャプチャを行うといったような，単眼カメラの移動を考慮したモーションキャプチャの手法について検討し，スマートフォン（と PC の接続）での実用を目指す．通常単眼カメラによるモーションキャプチャにおいては，背景が固定となり，人間の動作から関節の動きを，もしくは人間の目の位置を判定し表情を画像認識等で抽出するという手法などがとられている．このため，カメラを動かすことを考えると人間と背景の境界の判定が精度面において必須になると考えられる．また，本手法では境界の判定を行ったあと人間の部分を除いた背景の情報及びスマホの回転角からスマホの自己位置推定を行う．

第2章 関連研究

本章では従来から研究されている単眼カメラを用いたモーションキャプチャの手法並びに障害物を考慮した自己位置推定 (v-SLAM) の手法について述べる.

2.1 単眼カメラを用いたモーションキャプチャの手法

本研究では, 人の動きをデータとして取得するためにモーションキャプチャの手法について検討する.

単眼カメラを用いたリアルタイムでのモーションキャプチャの手法の一つとして, 人体のシルエット画像並びに人体の肌色部分を抽出し, 3次元モデルデータの骨格とマッチングさせるという手法がある [1]. この手法は, 主に「顔位置の検出」, 「肩位置の推定」, 「人物領域・肌色領域の抽出」, 「3次元位置の推定」の4ステップの工程によって成り立っており, フレーム毎に処理を行うと同時に, キャプチャが終了するまで4工程を繰り返し続けるものになっている. また, 文献 [1] では腕などの関節が多数あり, 動きが多い部分の推定には, 対象の動きを予測したサンプルを用いて対象の動きを逐次的に推定するパーティクルフィルタを適応するという手法がとられている.

さらに, 動画像の場合 (リアルタイムでは), 処理を行った直前の画像から整合性がとれるように骨格位置を修正する手法や近年では AI に人の骨格及びその動きの情報を与え学習させ, 様々な姿勢において検出ができるようにするといった手法 [2] も取られており, 後者においては SDK やツールとして販売もされている.

2.2 障害物を考慮した自己位置推定 (v-SLAM) の手法

Visual SLAM(v-SLAM) は, 画像からカメラの位置姿勢と環境を同時に推定する手法である. このうち特徴点に基づく方法は, 画像から検出した特徴点を追跡することによってカメラ位置の推定を行う.

このとき, 移動する障害物 (歩行者等) によって一部の特徴点が見えなくなる場合があると, 照合ミスを起こし自己位置推定の破綻を起こす. このため, 未知の障害物がある場合の位置推定を行う手法として 2.1 節と同様にパーティクルフィルタを用いた手法が提案されている [3]. このパーティクルフィルタを用いた手法では, 複数の特徴点を用いて推定される自己位置の候補 (パーティクル) を複数仮定し, カメラの動きの情報などからそれぞれの候補の妥当性を調べて候補を絞るということを行う. また, 自己位置推定を行う事前に, 推定を行うエリアについて環境地図を作成し, 壁などの既知の障害物の情報を取得するという行っており, 実際の測定でマッチングを行う際, 取得した情報と環境地図の位置が合うように同次変換を行った後, ピクセルの差異を求めパーティクルの確からしさを判定する.

第3章 使用機材・実装環境・研究計画

今回，研究をするにあたって利用を想定している環境及び機材を表 3.1 に示す．

表 3.1: 使用する機材・実装環境

言語	Java (Kotlin も検討)
開発 OS	Windows
統合開発環境 (IDE)	AndroidStudio
端末	Pixel 10(研究室にあるものの予定)

今回は，Android 端末の利用を想定しているため言語は Java としているが，Android アプリ開発の記事などを見ると，Kotlin を利用して開発を行っている企業等が多かったため，春休みの期間を使い Kotlin の利用も検討する．（そのため，来年度初頭に実装言語・環境を確定する．）

また，簡単な研究計画を表 3.2 に示す．

表 3.2: 研究計画

始業前	Kotlin に触れる，方針確定
4 月，5 月	境界判定の実装
5 月～8 月	モーションキャプチャ部実装
夏季休業	室内マッピング，特徴点検知部実装
夏季休業～10 月	自己位置推定部実装，統合
11 月～	卒論執筆開始

第4章 提案手法

図 4.1 に現在考案しているモーションキャプチャ手法の流れ図を示す。

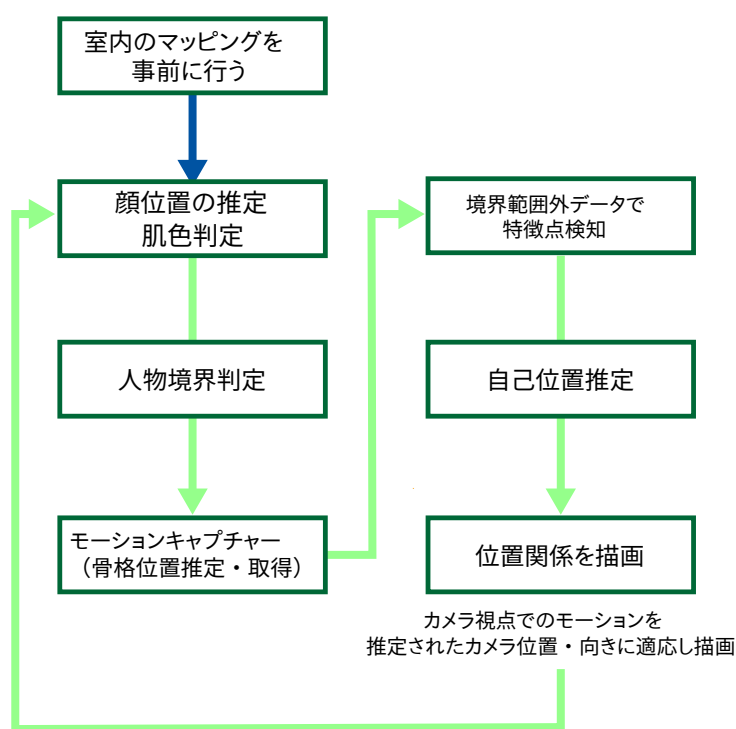


図 4.1: 提案手法の流れ（仮）

本研究では、カメラ移動を伴うモーションキャプチャ及びキャプチャ端末の自己位置推定を行う上で、事前にモーションキャプチャを行う環境（室内）について、事前に v-SLAM を用いてマッピングを行う。これは、骨格位置推定後のスマホの自己位置推定時に、人物の境界外のデータから特徴点を見つけ出し、事前に作成した環境地図と比較することによって推定を行うためである。また、自己位置推定について、文献 [3] では未知の障害による特徴点の消失をパーティクルフィルタを用いた手法によって解決していたが、本研究ではこの未知の障害をモーションキャプチャを行う人物と置き換えることで人物の境界外のデータのみでもスマホの自己位置推定ができるのではないかと考えている。

カメラの移動を含むモーションキャプチャーの精度について、人物がカメラ内に移っている場合は人物の境界判定の精度に依存すると考えられ、最終的に骨格位置推定とスマホの自己位置推定の結果を合成して描画を行うため、現状では原則として人物がカメラに写っていることを前提とする。（解決法を現在検討中）

なお、今回、顔位置の推定や肌色の判定、人物の境界判定を行う際に OpenCV の利用を想定し

ている。(Kotlin の場合は検討)

全体的な課題として、v-SLAM、モーションキャプチャー部分の機能をどのように実装するかということについて詰められていないため、使用言語の決定と共にこちらも来年度初頭までに目途をつける。

参考文献

- [1] 北 昭雄, 萩原 良信, 今村 弘樹, and 崔 龍雲.
"人体構造を考慮したパーティクルフィルタを用いた単眼カメラによる 3 次元リアルタイム
モーションキャプチャ."
映像情報メディア学会技術報告 35.9 巻 (2011), 社団法人映像情報メディア学会, 2011
- [2] 株式会社ネクストシステム, 姿勢推定 AI エンジン「VisionPose (ビジョンポーズ)」,
<https://www.next-system.com/visionpose>.
- [3] 富沢 哲雄, 村松 聡, 平井 雅尊, 佐藤 晶則, 工藤 俊亮, and 末廣 尚士.
"グリッドマップのマッチングに基づく未知障害物にロバストな自己位置推定."
日本ロボット学会誌 30 巻 (2012) 3 号, 一般社団法人 日本ロボット学会, 2014