

複数台の Azure Kinect を用いた モーションキャプチャデータ統合による 3D アバターの表示精度向上

I類(情報系) 成見研究室 学籍番号: 2010631 松浦怜央

1 はじめに

昨今、人や動物の体がどのように動いているかを追跡できるボディートラッキング技術、また、その追跡した動作をデータとして取り入れるモーションキャプチャ技術が普及している。これらの技術はスポーツにおける動作解析や映像コンテンツの作成には欠かせないものとなっている。

Microsoft 社が提供している Kinect は、トラッキング対象へのマーカーの付与および専用のスペースを必要としないモーションキャプチャデバイスである。従来のモーションキャプチャデバイスと比較して非常に安価(5万円未満)であることから、同 Microsoft 社が提供する専用の SDK や API を使用して今日までに様々なシステムが開発されている。

Magic Mirror は Kinect を用いて画角内の人たちのモーションキャプチャを行うことで、ディスプレイ上にキャプチャした人と同じ動作をするアバターを映し出すシステムである [1]。しかし Kinect の性質上、障害物に隠された部位があると精度が低下し、表示するアバターが不自然な姿勢を取ってしまう(図1)。

2 研究目的

本研究では、複数台の Azure Kinect を用いて死角となっている部分のモーションキャプチャデータを補完し合うことで、Magic Mirror を使用する際の課題である、アバターが明らかに不自然な体勢をしてしまう問題を改善することを目指す。また、障害物を想定した画角制限を行った状態でアバターに動作を反映させ、精度評価を行う。



図1: Magic Mirror 使用中の、モーションキャプチャ失敗例(左)と成功例(右)

3 Kinect

Azure Kinect[2] は、赤外線を利用することによってカメラの画角内に映る物体までの距離(深度)を計測し、複数人物の骨格情報をリアルタイムに推定することが可能である。Azure Kinect 以前に普及した Kinect v2 は、1台の PC で複数台を動かすことはできないため Kinect と同数の PC を準備する必要があった [3] が、後継機の Azure Kinect では、1台の PC で複数台動かすことができるようになった。

Kinect の種類にかかわらず、トラッキング対象の一部が家具などの物体で隠されている場合および自己遮蔽状態(トラッキングすべき部位の一部が自身の体によって隠されている状態)の場合は、モーションキャプチャ精度が低下する。

4 既存研究

4台の Kinect v2 を用いたモーションキャプチャ

森本らは4台の Kinect v2 を用いてモーションキャプチャを行うシステムを開発した [3]。

森本らが開発したシステムでは合計5台の PC を使用しているが、本研究では1台の PC のみを用いてモーションキャプチャを行うシステムを開発する。

4台の Azure Kinect を用いたモーションキャプチャ

Lee らは4台の Azure Kinect を用いてモーションキャプチャを行うシステムを開発した [4]。

Lee らは、6つの動作を行ってシステムの評価を行った。本研究では、障害物を想定した画角制限がある状況下で7種類のポーズをとってアバターを用いた評価実験を行う。

5 システム概要

開発したシステムの概略図を図2に示す。本システムは2台の Azure Kinect と1台の PC からなる。2台の Azure Kinect によるモーションキャプチャによって、リアルタイムにトラッキング対象の動作が PC 画面に映るアバターへ反映される。

本システムには「選択補完モード」と「自動補完モード」が存在する。選択補完モードでは、ユーザーが

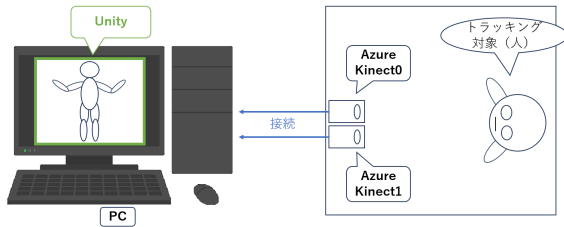


図 2: システム概略図

Azure Kinect1 のデータを使用する関節を選択することができる。自動補完モードでは、モーションキャプチャ中に随時、Azure Kinect によって算出される各関節の信頼度に応じてどちらの Azure Kinect のデータを使用するか、システムが自動選択する。

6 精度評価

精度評価方法

評価実験は図 3 の状況で行った。データ統合用の 2 台の Azure Kinect のうち 1 台は下半分を覆うことで障害物が下側にある状況を模擬し、もう 1 台は上半分を覆うことで障害物が上側にある状況を模擬している。

定量的評価のため、データ統合用 Azure Kinect、正解データ用 Azure Kinect のそれぞれと対応付けた 1 体のアバターを Unity を通して同一座標をルートとして 1 つの画面に表示させる。トラッキング対象は Azure Kinect から 210cm 離れた地点で事前に決められた図 4 の合計 7 種類のポーズを 300 フレームの間行う。Unity 側では 1 フレームごとに、表示された 2 体のアバターの 13 個の関節¹ごとの距離を計算する。

データ統合用 Azure Kinect の画角制限をなくした状態でも同様の実験を行い、合計 28 個²の関節ごとの距離データのセットを取得する。取得したデータの、関節ごとの平均値およびデータの遷移状況により本システムの総合的な評価を行う。

結果

表 1 は正解のアバターとの誤差の平均をまとめたものである。表 1 から、自動補完モードがより誤差が小さいことが分かった。ただし、画角制限があるときの誤差は、画角制限がないときの誤差より少し大きい。

7 まとめと今後の課題

まとめ

2 台の Azure Kinect を用いることで、障害物があってもアバターが不自然にならないようにモーションキャプチャできるシステムを開発した。精度評価によって、本

¹頭、首、左肩、右肩、左ひじ、右ひじ、左手、右手、腰、左ひざ、右ひざ、左足、右足

²画角制限ありかつ選択補完モードで 7 ポーズ、画角制限ありかつ自動補完モードで 7 ポーズ、画角制限なしかつ選択補完モードで 7 ポーズ、画角制限なしかつ自動補完モードで 7 ポーズ

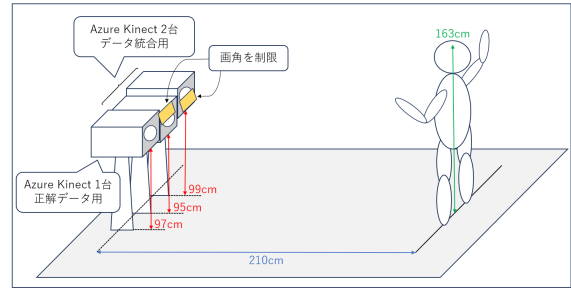


図 3: 精度評価実験の際の状況

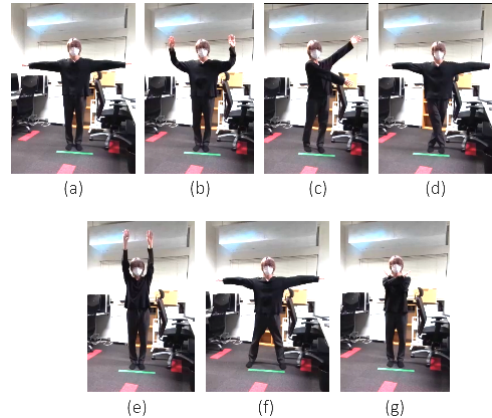


図 4: 精度評価実験の際の 7 種類のポーズ

表 1: 正解のアバターとの誤差 (単位: cm)

	画角制限あり	画角制限なし
選択補完モード	11	7.6
自動補完モード	10	6.3
両モード総合	10	7.0

システムによって表示されるアバターは、1 台の Azure Kinect によって表示するアバターと比べて 10cm 程度の誤差であることが分かった。また、画角制限の有無にかかわらず、自動補完モードの方が高精度であることが分かった。

今後の課題

モーションキャプチャ可能人数を増加させる。2 台の Azure Kinect の位置を変更可能にする。また、障害物によってトラッキング対象が隠れている場合に生じる、アバターの乱れの抑制を行うなどが挙げられる。

参考文献

- [1] 明石禎紀, 富井陸矢, 真鍋光希, 高見太基, 帆山遼, 羽賀聡希, 寺崎葉月, 武井友里恵, 鈴ヶ嶺聡哲, 山中佑紀, 峯水延浩, 成見哲, 等身大の仮想、手の中のリアル。エンターテインメントコンピューティングシンポジウム, (2019).
- [2] Azure Kinect DK について. <https://learn.microsoft.com/ja-jp/azure/kinect-dk/> (閲覧日 2024-01-28).
- [3] 森本隆志, 満上育久, 複数 Kinect の姿勢データの時空間統合によるモーションキャプチャシステム。情報処理学会研究報告, Vol.2019-CVIM-217 No.25, pp1-8, (2019).
- [4] Lee, S. H., Lee, D. W., Jun, K., Lee, W., Kim, M. S. "Markerless 3d skeleton tracking algorithm by merging multiple inaccurate skeleton data from multiple rgb-d sensors". Sensors, 22(9), 3155, (2022).