

コンサート自由視点映像視聴時の嗜好を考慮したカメラワーク生成手法の検討 A Camerawork Generation Method Considering Preferences for Viewing Free-Viewpoint Concert Video

瀬谷 諒太[†] 謝 淳[†] 北原 格[†]
Ryota Seya Xie Chun Itaru Kitahara

1. はじめに

通信技術の発展や動画配信プラットフォームの整備により、コンサートのオンライン配信の普及が進んでいる。配信映像制作において、カメラワークは作品の魅力を引き出す重要な要素の一つであり、カメラマンや演出家などの専門家によって決定されるのが一般的である。一方で、ステージに入り込んだような視点や、特定の演者に焦点を当てた視点など、オンライン配信ならではのカメラワークを求める意見が視聴者からあがっている [1]。視聴形態の多様化に伴い、専門家によって設定されたカメラワークが必ずしも全ての視聴者にとって最適とは限らず、視聴体験の好みの個人差を反映した映像メディアの実現への期待が高まりつつある。視聴者が好みに応じて視点を選択できる技術として、自由視点映像が注目を集めている。自由視点映像では、視聴者が任意の位置から被写体を観察することができ、没入感の高い映像視聴体験が可能となる。一方で、カメラワークの決定には、専門的な知識や経験が求められ、自由視点映像では配信映像のようなカメラワークによる演出効果が得られにくくなるという課題がある。

カメラワークの自動生成に関する研究例はいくつか存在するが、コンサートを対象としたものは多くない。一例としては、ダンスパフォーマンス撮影を対象に、音楽やダンスに基づいて単一のダンサーに焦点を当てたカメラワークを自動生成する手法が提案されている [2]。また、カメラワークと同様、視点の選択や切り替えを行う動画編集の手法として、複数台のカメラで撮影された映像に対し、ダンサーの動作情報や映像編集の原理に基づいて、自動的に映像を編集する手法が提案されている [3]。しかし、いずれも専門的な知識や経験を反映する手法であり、視聴者の嗜好を反映する手法の研究は例を見ない。我々は、コンサート映像視聴において、多様な嗜好を反映したカメラワークの自動生成が求められていると考えている。

コンサートのようにグループパフォーマンスが行われる場面では、視聴者の注目したいポイントは状況によって変化し、演者を個別に追いたい場合もあれば、ステージ全体を俯瞰して見たい場面もある。本研究では、図 1 に示すように、視聴者の嗜好に基づいたカメラワークの生成手法を提案する。コンサート自由視点映像の視聴中に視聴者によって操作されたカメラの位置、回転、視野角、ステージ上の基準点とカメラの距離の四つ要素を視聴者の嗜好として捉える。これらのカメラ操作情報、すなわち視聴者のカメラワークをコンサートの構成要素である音楽とダンスとともに深層学習に学習させることで、音楽とダンス情報の入

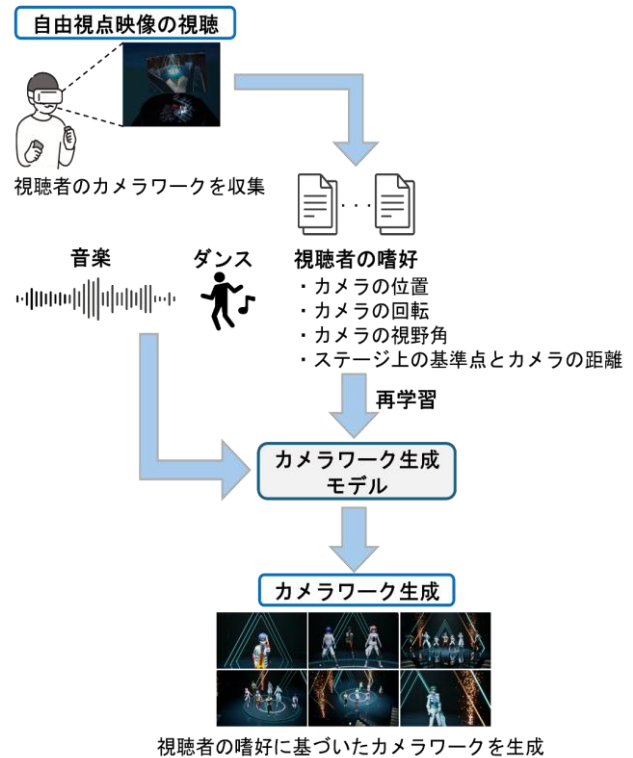


図 1 視聴者の嗜好に基づくカメラワーク生成手法

力から、個々の視聴者に応じたカメラワークを生成する。

2. コン서트自由視点映像

本研究で対象とするコンサート映像は、VR 環境 (Unreal Engine) と 3D アニメーションソフトウェア (MikuMiku-Dance (MMD) [4]) を用いて作成する。MMD から 7 体の 3D キャラクターモデルとダンスモーションデータを取得し、これらを Unreal Engine 上に導入することでステージでのパフォーマンスを再現した VR コン서트空間を構築する。ダンスモーションデータは毎秒 30 フレームで記述されており、フレーム毎に 3D キャラクターモデルのボーン的位置と回転が更新される。VR 空間内において、視聴者はカメラ位置や回転、視野角を自由に操作することができ、本研究ではこれをもって自由視点映像の視聴とする。

3. 自由視点映像視聴用インターフェース

本研究では、視聴者がコンサート映像視聴時に操作したカメラの位置、回転、視野角、およびステージ上の基準点との距離の変化を“視聴者の嗜好を示す情報源”とするため、カメラの操作方法が重要となる。一般に、Unreal Engine での操作は、キーボードおよびマウスによって行われるが、これらは 2 次元的な入力手段であり、3 次元空間

[†] 筑波大学 University of Tsukuba

内を動き回るカメラを操作するには直感性が十分とはいえない。そこで、図 2 に示すように VR を活用して自由視点映像視聴用のインターフェースを実装する。VR 空間内には、俯瞰的表現されたコンサート会場と会場内に設置されたカメラ映像を映し出すディスプレイを配置しており、視聴者は VR コントローラーを用いてカメラを操作する。視聴開始時の現実世界のコントローラーの位置を基準とし、基準の位置から上下左右、前後に動かした方向に仮想空間内のカメラを動かすことができる。カメラの回転操作については、プロのカメラワークでもロール方向の使用頻度が低いことを踏まえ、操作を簡略化する目的で、ヨーとピッチ方向に限定し、これらはコントローラーの回転と連動している。視野角やカメラの移動速度は、コントローラーの特定のボタンを押すことで変更する。

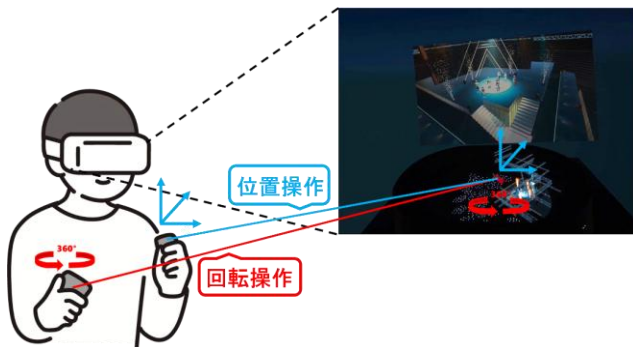


図 2 自由視点映像視聴用インターフェース

4. カメラワークの生成

カメラワークの学習および生成には、Transformer ベースの拡散モデルである DanceCamera3D を採用する[2]。DanceCamera3D では、音楽とソロダンスとカメラワークの関係を学習している。本研究でも同様に音楽とダンスを用いるが、カメラワークは視聴者の操作によるものを学習対象とする。しかし、一般の視聴者が常に適切なカメラワークを実現できるとは限らず、学習データとしての質に課題がある。そのため、視聴者によるカメラワークに加えて、

MMD 由来のカメラワークも学習に含めることで、専門的な知見を補完することを目指す。

自由視点映像視聴時におけるカメラの操作データを 15 件収集し、DanceCamera3D の再学習を行った。また、MMD 由来のカメラワーク 15 件を加えた合計 30 件のカメラワークで再学習を行い、カメラワークの生成を行った。学習設定は、バッチサイズ 15、エポック数 500 とした。図 3 に、視聴者によるカメラワークのみを用いて学習した場合と、MMD カメラワークを加えて学習した場合の生成結果、ならびに、視聴者によるカメラワークの一例を示す。生成したカメラワークは、元のカメラワークと必ずしも近いものになるわけではないが、学習用のカメラワークの各要素を取り入れたカメラワークが生成できることが確認できる。

5. おわりに

本研究では、コンサート自由視点映像撮影において、視聴者の嗜好に基づいたカメラワークの生成手法を提案した。視聴者が Unreal Engine 上に再現したコンサート自由視点映像を視聴中に操作したカメラワークをもとに、その要素を反映したカメラワークを生成できることを確認した。今後は、被験者を集め、視聴者によるカメラワークの収集および生成したカメラワークに対して視聴者の意見を調査する。

参考文献

- [1] 株式会社 SKIYAKI, “音楽ライブ配信についての意識調査レポート”, (2020).
- [2] Zixuan Wang, Jia Jia, Shikun Sun, Haozhe Wu, Rong Han, Zhenyu Li, Di Tang, Jiaqing Zhou, and Jiebo Luo “DanceCamera3D: 3D Camera Movement Synthesis with Music and Dance”, IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.7892-7901(2008).
- [3] Shuhei Tsuchida, Satoru Fukayama, and Masataka Goto “Automatic System for Editing Dance Videos Recorded Using Multiple Cameras”, ACE 2017, LNCS 10714, pp. 671-688,(2018).
- [4] “MikuMikuDance”, <https://sites.google.com/view/vpvp/>



図 3 カメラワークの生成結果

左：視聴者カメラワークのみ学習，中央：視聴者+MMD カメラワークで学習，右：視聴者によるカメラワークの一例