

ダンスと場景を連動させた映像生成の提案と ダンス ONE プロジェクトへの実践

吉川 優依^{1,a)} 宍戸 英彦^{2,b)} 亀田 能成^{2,c)} 北原 格^{2,d)}

受付日 2022年2月18日, 再受付日 2022年4月25日/2022年7月5日,

採録日 2022年8月8日

概要: 近年, 新型コロナウイルス感染症 (COVID-19) の影響によりダンサーがパフォーマンスを披露する場はステージからメディアへと変化しており, 高校生のダンス部を応援するために, 2020 年よりテレビ局の企画として「ダンス ONE プロジェクト」が実施されている. ダンス ONE プロジェクトに参加する高校生は, 撮影および動画制作に不慣れであり, そのサポートや新たな教育実践法が不可欠である. 我々は, IT の活用体験を向上させることを目的として, コンピュータビジョンの最先端技術を活用した映像生成手法をサポートした. 研究協力校ダンス部の高校生とコンセプト作りや撮影作業などに取り組み, 音源のビート検出手法とオブジェクト検出アルゴリズムを活用したダンスと場景を連動させた「高揚感のある」映像生成法を提案した. 信号機の色を映像処理によって拡張し, 場景とダンスが一体化しているような効果を狙い, 違和感のない映像の生成方法を開発した. 動画編集の難しさを解決するためにカット割りとズーム設定の自動処理を提案した. また, ダンス ONE プロジェクトを通して, 高校探求学習科目「総合的な探究の時間」と同様の取り組みが, 部活動にも展開できることを示した. 「総合的な探究の時間」に向けて, 高校生に対しソフトウェアを用いた動画編集方法とコンピュータビジョン技術を活用した映像生成手法について講義した. IT の活用体験の向上に関する意向調査によって, ダンス ONE プロジェクトと本研究活動を通して参加者の IT の活用体験が向上したことを確認した.

キーワード: 教育, コンピュータビジョン, ダンスパフォーマンス, 動画編集

A Proposal of Dance and Scenery Linked Video Generation and Application to Dance One Project

YUI YOSHIKAWA^{1,a)} HIDEHIKO SHISHIDO^{2,b)} YOSHINARI KAMEDA^{2,c)} ITARU KITAHARA^{2,d)}

Received: February 18, 2022, Revised: April 25, 2022/July 5, 2022,

Accepted: August 8, 2022

Abstract: In recent years, due to the effects of COVID-19, the venues for dancers to perform have changed from the stage to the media. In order to support high school dancers, “Dance ONE Project” has been implemented as one of the TV stations’ projects since 2020. Our goal was to improve the IT utilization experience of the high school students participating in the Dance ONE Project by supporting the filming and video production. In this paper, we propose a video generation method that utilizes the latest technology in computer vision. We propose a video generation method that links dance and scenery by using a sound source beat detection method and an object detection algorithm. The color of the traffic lights is extended by video processing to create the effect that the scene and the dance are unified, and the video is generated without any sense of discomfort. Through the Dance ONE Project, we propose a new educational model by making the same approach as the inquiry study subject “Time for Comprehensive Inquiry” a part of club activities. In order to solve the difficulty of video editing, we propose an automatic process for detecting the cut and zoom settings. In the survey on the improvement of IT utilization experience, it was confirmed that the participants’ IT utilization experience was improved through the Dance ONE Project and this research activity.

Keywords: education, computer vision, dance performance, video editing

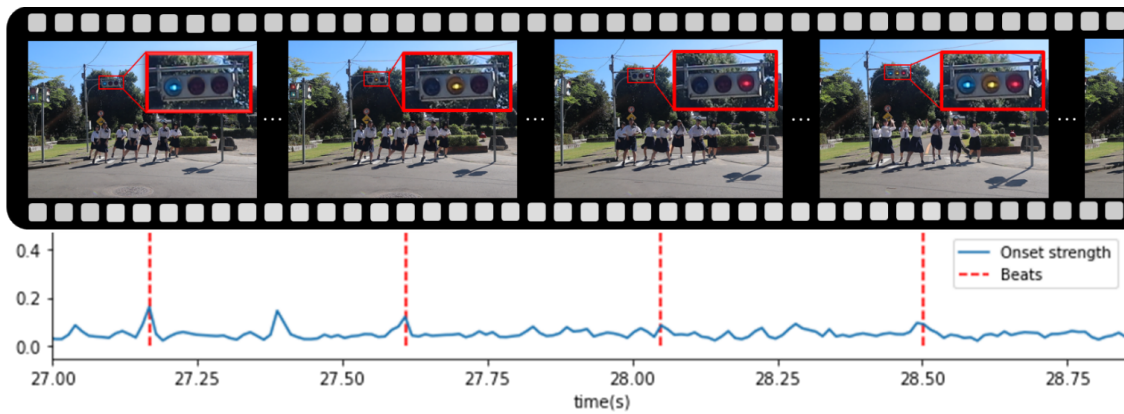


図1 上段：音源のビートに対応する信号機の変化 下段：取得されるオンセット強度とビート

Fig. 1 Upper row: Traffic light changes in response to the beat of the sound source.

Lower row: Onset strength and beat acquired from the sound source.

1. はじめに

新型コロナウイルス感染症（COVID-19）パンデミックが宣言された2020年3月以降，人前でステージパフォーマンスを披露する機会が失われている．なかでも，ダンスパフォーマーは生のパフォーマンスと観客の歓声が一体となる演出を重要視しているが，近年の状況では実現が困難である．特に現役高校生においては，パフォーマンスの重要な機会である文化祭が中止され，ダンスに対するモチベーションが低下している．

そこで，高校生のダンス部を応援するために，2020年よりテレビ局の企画の1つとして「ダンス ONE プロジェクト」が実施された [1]．参加高校が課題曲に対してダンス動画を制作することで，テレビ放送やYouTube配信によりダンスを多くの人に披露する機会を提供するプロジェクトである．参加する高校生にとっては，作り上げた作品を発表する数少ない機会である一方で，文化祭のステージパフォーマンスよりも広範囲な視聴層をターゲットとできる．実際に，2020年に実施されたダンス ONE プロジェクトでは，全国で153の高校が参加し，多くのダンス部の高校生および学校関係者から注目を集めた．参加者の高校生はダンス動画の制作に関して未経験であるが，視聴者の目を引くような衣装，特色ある地元のロケーション，ダンスの躍動感を際立たせるカメラワークなど工夫して制作に取り組んでいる．

このように，2020年に実施されたダンス ONE プロジェクトは高校生が積極的に参加し，YouTube配信では学校関係者以外の方々からも称賛や励ましのコメントが溢れ

た．しかし，高校生としては高揚感のある映像の制作方法が不明であることや，特色のある映像の制作に難しさを感じている．本稿における高揚感とは，気持ちが高ぶり興奮した感覚と表現できるものと捉えている．このことは高校生らとの意見交換で確かめている．

本研究では，高揚感のある映像制作および映像編集に焦点を当て，図1に示すようにダンスと場景が連動する映像生成手法を提案する．ダンス映像の主体はダンサーであり，ダンスミュージックである．ダンサーより目立った違和感のある変化はこの意識を妨げてしまう．したがって，高校生の要求を満たすような意識を妨げずかつ高揚感のある映像生成を狙う．

我々は，ダンス ONE プロジェクトに参加する高校生の撮影および動画制作をサポートする形で協力し，高校生のIT活用体験を向上させることを狙った教育活動に取り組んでいる．具体的には，日常的に利用しているスマートフォンによる動画撮影ではなく一眼レフカメラとカメラスタビライザーを使用した撮影体験，動画編集ソフトウェアを活用したスマートフォンよりも高度な動画編集の体験を提供する．さらに最先端のコンピュータビジョン技術やその活用による映像生成手法に関する講義およびディスカッションを実施し，最先端の技術に触れる機会を提供する．

高等学校では，探求学習科目として「総合的な探究の時間」があり，生徒が主体的に課題を設定し，情報の収集や整理・分析をしてまとめるといった能力の育成を目的としている．このような教育活動の中で，研究協力校の茨城県立竹園高等学校では探究を「探Q」（Qはqualia（「感覚質」）の意味）と呼び，論理的思考力，知的好奇心，表現力，情報発信力を育成するための学術的な活動として，最先端の研究内容に触れながら，週1回，探究型学習を行っ

¹ 筑波大学大学院 理工情報生命学術院 システム情報工学研究群 知能機能システム学位プログラム

University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

² 筑波大学 計算科学研究センター

Center for Computational Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan

a) yoshikawa.yui@image.iit.tsukuba.ac.jp

b) shishido@ccs.tsukuba.ac.jp

c) kameda@ccs.tsukuba.ac.jp

d) kitahara@ccs.tsukuba.ac.jp

ている。ダンス ONE プロジェクトの活動においてダンスパフォーマンス動画の制作過程は、ダンスパフォーマンスの振り付け、撮影方法、動画編集などの情報を調査し実践するものであり、探求学習科目「総合的な探究の時間」と同様の経験が部活動の一環として可能である。我々は、スポーツ・文化を通じた教育活動において、最先端のコンピュータビジョン技術を活用した映像生成手法を提案し、高校生がダンス ONE プロジェクトへ取り組む活動をサポートすることで、高度な教育活動を行い、新たな教育モデルを実践する。

2. ダンスパフォーマンス映像表現の高度化

ダンサーのパフォーマンスをより魅力的に視聴者に伝えられるようダンスパフォーマンス映像表現の高度化を狙い、日常的变化に見慣れている信号機を変化させることでダンスと場景が連動した映像を生成する。高校生への意向調査により、高揚感があり、かつ違和感のない映像生成の実現を確認した。

2.1 ダンス ONE プロジェクト

ダンス ONE プロジェクト [1] とは、全国の高校生を対象に、参加する高校生が課題曲に応じたダンス動画を制作することで、テレビ放送や YouTube 配信を通して多様な人へ披露できる機会を提供するプロジェクトである。部分的な見本となる振り付けダンス（決まりパート）が提供されており、決まりパートの振り付けは全身を使った高校生の元気を表現できるものとなっている。決まりパート以外は各校が自由に振り付け可能である。全体的に動画撮影・編集に制限はなく、自由に考えられるダンスパフォーマンス動画のコンセプトが各校の個性が際立つポイントとなる。課題曲は YOASOBI の「群青」に設定され、昨年と同様にアップテンポなリズムの曲調である。

2.2 ダンスパフォーマーとの連携によるコンセプト作り

研究協力校の茨城県立竹園高等学校では第 1 回のダンス ONE プロジェクトへ参加した実績を持つ。このときは初めての参加ということもあり、撮影場所は高校の敷地内のみ、撮影方法・動画編集方法はすべてスマートフォンで実施し、即興で作られたものであった。このような経験を踏まえ、第 2 回の参加では、探求学習科目「総合的な探究の時間」を意識した活動に方針を定め、撮影方法・動画編集方法について調査した。ダンスパフォーマーである高校生は、考案した振り付けとダンスパフォーマンス動画の完成イメージを我々と共有し、ダンスパフォーマンス動画の基礎や可能な撮影方法を学習し、ディスカッションを行うことで完成イメージをより具体的なものとした。前年のダンス ONE プロジェクトの撮影より高度な方法として、同一シーンにおける様々な角度からの撮影や高校の敷地内のみ

ならず近隣の施設での撮影を決定し、撮影方法を模索しながら完成イメージを修正した。

2.3 映像表現高度化のための提案

我々は、茨城県立竹園高等学校ダンス部と協力し、ダンス ONE プロジェクト '21 に参加する。そこで、ダンス映像の主体への意識を妨げず、かつ高揚感のある映像の 1 つとして、ダンスと場景が連動する映像生成手法を提案する。

ダンス ONE プロジェクトでは、地域の特色が出るような映像が求められ、画面上でキャラクターや企業の商標・ロゴが見えないように撮影しなければならない制約がある。また県内に参加する高校は多数あり、近隣の観光施設は他校にとっても地元の特徴である。そこで、高校生と我々はつくば市の中でも竹園高校の近隣に位置し、オランダのアムステルダム駅舎をアレンジしたレンガ造りの建物や、自転車の練習場所となる交差点のミニチュア版を保持しているさくら交通公園を貸し切り、図 2 のように 1 台の移動カメラを用いてダンサーと信号機が映るよう撮影を実施した。交差点のミニチュア版は、横断歩道や信号機（歩行者用含む）のサイズが通常より一回り小さく作られており、ダンスパフォーマンスを撮影した場合、図 1 上段に示すように信号機が通常よりも大きく映るメリットがある。このシーンの中でダンス映像の主体への意識を妨げない場景として、日常生活において色の変化に見慣れている信号機とする。

2.4 ダンスと場景を連動させた映像生成手法

ダンスと連動した映像を生成するために、ダンスモーションと連動している音源のビートを自動で検出する。近年、音源のビート検出手法が提案されており、対数パワースペクトラムの時間方向の差分より計算されたオンセット（音の始まり）強度の相関関係よりテンポを推定し、推定されたテンポと一致するオンセット強度のピークを選択することでビート検出を実現した [2]-[4]。ダンス ONE プロ



図 2 さくら交通公園での撮影の様子

Fig. 2 Shooting at Sakura Transportation Park.

プロジェクト'21の課題曲（ファイル形式：wav）に対して Librosa [2] を用いてビート検出を適用し、手動でアノテーションを付与した値と比較した結果、許容誤差範囲（2フレーム）以下で一致することを確認した。したがって、本研究においても従来手法を採用し、映像と同期した音声クリップを入力し、図3に示すように楽曲信号のオンセット強度の時系列データよりビート情報を取得する。Librosaのonset検出にはlibrosa.onset.onset_strength関数を使用し、ビート検出にはlibrosa.beat.beat_track関数を使用する。次に図4に示すようにダンス映像（ファイル形式：mp4）から静止画像（ファイル形式：jpeg）を切り出し、フレームごとに信号機を検出する。セマンティックセグメンテーション[5]を用いた場合、信号機は検出されるが、セグメンテーション領域が不完全であった。そこでより検出精度の高いオブジェクト検出アルゴリズムであるYOLO（You Only Learn One Representation）を使用する[6]-[8]。検出された信号機領域から点灯する円形領域を抽出する[9]-[11]。点灯領域内においても画素値が異なるため誤検出が発生するが、パラメーターの調節により許容範囲内の検出精度を確認した。cv2.HoughCircles関数において各パラメーターは、円同士が最低限離れる距離：10 pixel、円の中心を検出する際の閾値：30、検出する円

の半径の範囲：5-50 pixelと設定した。意図した色の位置に点灯もしくは消灯のテンプレート画像を貼り付け、ビートに合わせて様々なパターンに変化させることで、ダンスと連動した映像を生成し、音声付き動画クリップを出力する。提示パターンは図5のような青、黄、赤それぞれが単独で点灯する一般的な提示に加え、図6のように複数点灯や全色消灯といった拡張したパターンを含める。

2.5 信号機の色を拡張したパターン提示

信号機の色を拡張したパターン提示について、図8に示すように、場景がダンスと一体化しているような効果、違和感のない提示を検討した。提案手法を適用した交通公園のシーンは1ビートごとにリズムをとるような曲調であり、高校生が考案したダンスの振り付けも1ビートごとに変化する。また、緩やかな曲調になるほど振り付けも緩やかとなる。そこで、序盤では単独で点灯する一般的な提示を1ビートごとに変化させ、その後拡張したパターンを1ビートごとに変化させる。次に、緩やかな曲調になるほど拡張したパターンを含めて2ビートごとに変化させる。このような信号機のパターン提示によって、場景がダンスと一体化しているような効果を狙った。また、全体で2分15秒の作品のうち、提案手法を適用した交通公園のシー

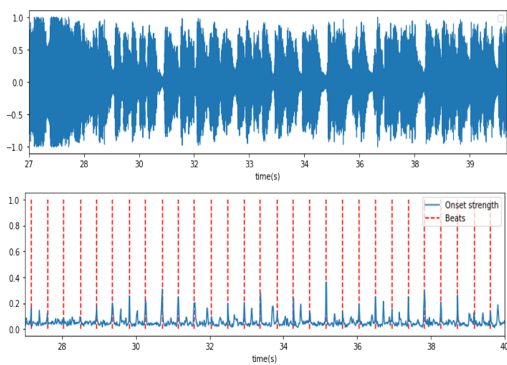


図3 上段：音源の音声波長
下段：取得されるオンセット強度とビート
Fig. 3 Upper row: Audio wavelength of the sound source.
Lower row: Onset strength and beat.

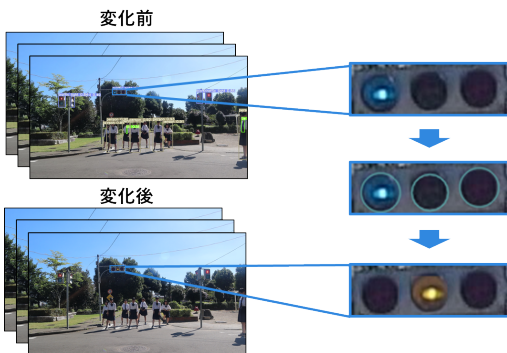


図4 信号機点灯領域の色変化の流れ
Fig. 4 Flow of color change in the traffic light lighting.



点灯：青 点灯：黄 点灯：赤
図5 車両用信号機の一般的な色パターン
Fig. 5 General color patterns for vehicle traffic light.



点灯：青, 黄, 赤 点灯：青, 赤 点灯：なし
図6 車両用信号機拡張した色パターン
Fig. 6 Expanded color patterns for vehicle traffic light.

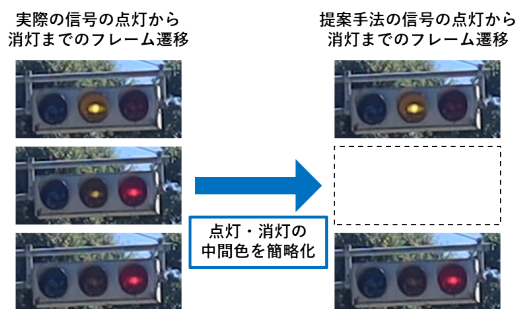


図7 実際の信号と提案手法の信号の点灯から消灯までのフレーム遷移の比較
Fig. 7 Actual frame and edited frame from light on to light off.



図8 ビートに合わせて変化した信号機のパターン提示

Fig. 8 Presented traffic light pattern that changes with the beat.

ンは、14秒（1/9程度）とすることで、違和感の低減を狙った。

次に、信号機の点灯・消灯の提示方法について検討した。図7に示すように実際の信号機は、30fpsの動画では3フレームにかけて点灯から消灯へ変化する。1フレーム0.03秒であることから点灯・消灯の中間色は簡略化が可能であり、点灯および消灯の変化を容易に生成できることが事前調査より判明した。したがって、色パターンの変化は、点灯・消灯する領域の2種類のみで提案手法は適用可能である。

2.6 映像生成手法の効果に関する意向調査

ダンスと場景を連動させた映像生成手法を適用したシーンに関して、本研究の協力者であるダンス部員24名に対して映像生成手法の効果に関する意向調査を行った。提案手法の効果において「高揚感」「違和感」といった感性に基づく評価が必要であることから、意向調査を選択した。図9のアンケート結果より、生徒全員が高揚感に対して

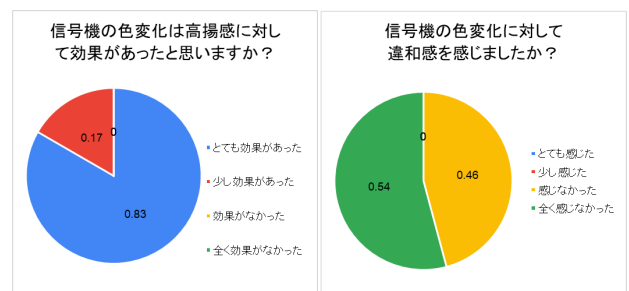


図9 映像生成手法の効果に関する意向調査

Fig. 9 Results of an intention survey on the effectiveness of image generation methods.

効果があり、違和感がないと感じていることを確認した。「信号機の色が変わった前と後の動画を並べて見たときに、変わった後のほうが明るく楽しげなようにみえた」というコメントや「実は信号機の色が変わっていたことは知らなかった」など、高校生の要求を満たすような高揚感があり、かつ違和感のない映像生成が実現されていることを確認した。ダンサーの動きに合わせて周囲の風景（本提案対

象の信号機)が変化することで、映像中のダンサーだけでなく、映像全体が踊っているようなより迫力のある映像の生成を実現した。

今回の意向調査では対象者を参加者の高校生に限定していることから、一般的な視聴者に対する意向結果と異なる可能性がある。したがって、次のプラクティスにおいてはダンス経験のない一般的な視聴者も含めた意向調査を実施し、異なる調査結果やコメントの獲得を期待する。

3. ダンス ONE プロジェクトにおける映像制作支援と得られた知見

高校生の未経験デバイスの利用抵抗感を払拭するため、撮影や編集作業を高校生と共同で行い、ダンスパフォーマンス動画制作の支援を行う。高校生が普段使用することのない一眼レフカメラとカメラスタビライザーの利用体験やPCを用いた動画編集の機会を提供した。その中で、動画編集作業において労苦がみられたことから、動画編集の難しさを解決するための自動処理の提案を行う。

3.1 ダンスパフォーマンス映像クリップの編集

ダンス ONE プロジェクトへ投稿する動画のオープニングでは、高校生のメッセージ動画クリップから開始される。一般的な曲の構成は、イントロ、A メロ、B メロ、サビ、間奏と続く。課題曲の「群青」では、イントロはなく、A メロが穏やかな曲調から始まる。B メロは曲調が大きく変わりリズムに手拍子が入り、ビートを強調する曲調である。わずか13秒のB メロからサビへと移る。我々は、このリズムを強調する曲調が、視聴者に躍動感を感じ

させる最適な箇所と判断し、図8に示すように、ダンスと場を連動させた映像を適用した。

撮影は、竹園高校の校内、つくば市さくら交通公園、つくばセンター広場の中の合計8箇所で行われ、各シーンにはダンサー全体が映る映像と中心のダンサーが大きく映る映像の引きと寄り2種類を各々2,3クリップ撮影した。編集は、ダンサーが良く踊っているクリップと良く撮影できているクリップを厳選し、各シーンにおいて引きと寄りの映像を1つずつ用意した。次に、引きの映像と寄りの映像をどのように切り替えるかについて、高校生に映像を提示しながら協議した。この際、複数のパターンの中から複数の切り替えタイミングを決定する作業には多くの時間を要した。動画編集には、膨大なクリップの管理、それぞれのクリップと音響の同期、クリップ切り替えなどが、大画面で容易に可能であるPCのソフトウェアのAdobe Premiere Proを使用した。2.4節の手法にて取得されたビート情報を利用して、高校生との協議を踏まえた動画編集を実施した。撮影した映像クリップと音響データをAdobe Premiere Proにアップロードし、図10に示す編集画面において、複数のクリップを管理し(左下)、クリップ映像(右上)を確認しながら音響の同期とクリップの切り替え(右下)を実施した。撮影箇所や撮影方法の検討に3日、撮影の実施に3日、クリップの厳選に2時間、切り替えタイミングの協議に2時間、動画編集作業に4時間を要し、本プロジェクトへの投稿までに2ヵ月を要した。このように、ダンスパフォーマンス映像の制作期間は多くの時間を必要とする。ダンスパフォーマンス動画の制作の目的はダンス披露であり、客前でのステージパフォーマンスの目的

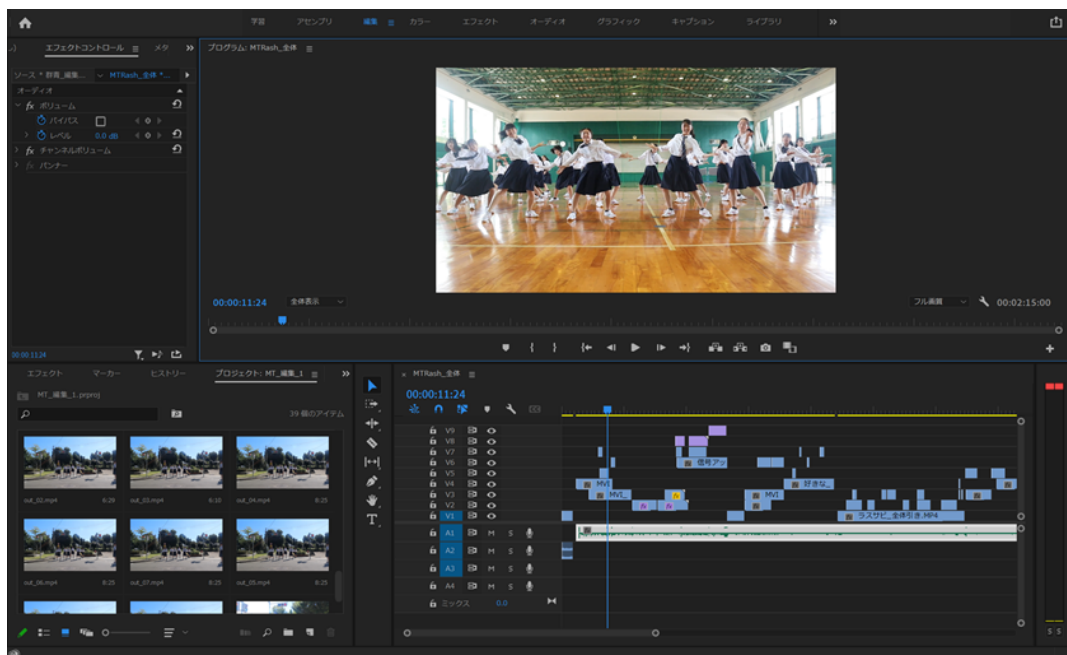


図10 Adobe Premiere Proの編集画面

Fig. 10 Editing screen of Adobe Premiere Pro.

と同様である。したがって、高校生は制作過程においてもダンスパフォーマンスを向上させるための重要な時間と認識している。しかし、動画編集作業においてはITを活用した経験の少なさから、PCの扱いや編集作業の基本操作に対し労苦がみられた。本研究の提案手法では、ダンスと場景の連動により高揚感のある映像生成を実現したが、動画編集作業の難しさに関しては解決できていない。そこで、動画編集作業を少しでも簡単にするための解決法を提案し、ITを活用した経験を高校生へ幅広く持たせる取り組みを実施した。

3.2 動画編集の難しさを解決するための自動処理の提案

ダンスパフォーマンス動画制作は、複数のロケーションにて複数のパターンを撮影する。したがって、動画編集において、膨大なクリップを管理する必要があり、クリップを切り替えるタイミングと使用するクリップを決定しなければならない。この作業は動画編集者の経験スキルに依存するため初心者にとっては難しく、我々が初めて動画を編集した際も3分のダンスパフォーマンス動画の制作に2週間を要した。そこで、見本となるダンスパフォーマンス動画のカット割りとズーム設定の自動検出を提案し、動画編集の難しさの大幅な改善を狙う。多数存在する編集作業の部分的な自動化によって編集者の負担を軽減し、高校生は好きなパフォーマンス動画と類似した動画を高品質かつ容易に制作することが可能となる。

図11に示すようにカットの検出は、隣接するフレーム間差分画像を用いる。差分画像から取得される平均絶対誤差に閾値処理を施すことによってカット情報を検出する。

ズームの検出は、「引き」で撮影された映像か「寄り」で撮影された映像かを判定するに留める。ダンスパフォーマンス映像ではドリー・イン/ドリー・アウトと呼ばれる撮影方法（メインの被写体に、カメラが近寄りたり、遠ざかったりする撮影方法）が主流であることによる。図12に示すようにズームの検出は、従来手法である骨格推定ライブラリ（OpenPose）[12]を使用し、ダンサーの骨格位置を取得する。画像上の首と肩の距離に閾値処理を施すことによって「引き」、「寄り」のクリップを検出する。

以上の提案により、任意のダンスパフォーマンス映像からカット情報とズーム情報を検出することが可能となり、制作時間の大幅な短縮が期待できる。

4. ダンス ONE プロジェクトを通じた新たな教育モデル

ダンスパーフォーマーの高校生に対しIT活用体験を提供する。ITへの興味を向上させるために、研究協力校を訪問し動画編集方法やコンピュータビジョン技術を活用した映像生成手法についての講義を実施する。ITの活用体験の向上に関する意向調査において、高校生へコンピュータ

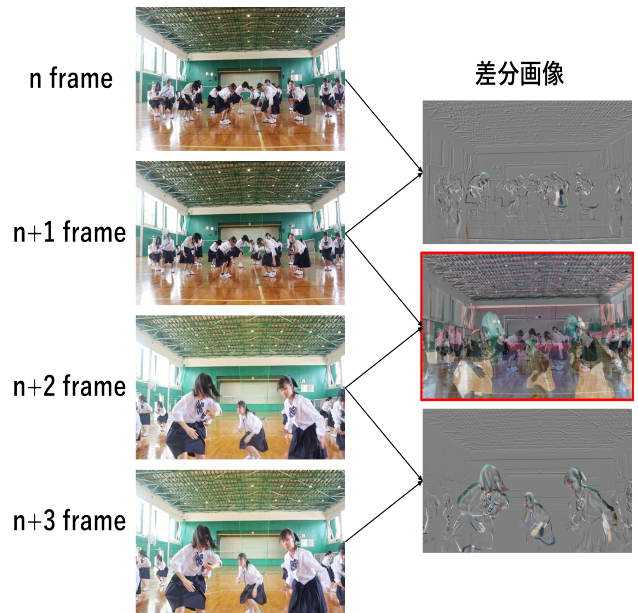


図11 カット割り検出のための差分画像出力
Fig. 11 Difference image output for cut-split detection.



図12 ズーム設定検出のための骨格位置情報検出例
上段：寄りの場合 下段：引きの場合

Fig. 12 Example of skeletal position information detection for zoom setting detection.

Upper image: zoom in, Lower image: zoom out.

ビジョン技術に触れる、理解する機会を提供することにより、高校生は実体験として新しいIT技術を学習する。

4.1 高校探究学習科目「総合的な探究の時間」に向けて新たな教育モデルを従来は別々であった探求学習科目



図 13 講義の様子

Fig. 13 Scene of lecture.

「総合的な探究の時間」と「スポーツ・文化活動」を一体化させることと定義し、ダンス ONE プロジェクトを通して実践する。

我々は図 13 に示すように、研究協力校を訪問し動画編集方法やコンピュータビジョン技術を活用した映像生成手法についての講義を実施した。コンピュータビジョン技術は、一般的な高等学校では触れる機会が少ない学問領域であり、コンピュータビジョンを知らない生徒が大半を占める。そこで、内容の理解を深めるためにできるだけ簡単な言葉に置き換えて説明した。映像の生成方法や目的の説明、生成映像の提示などにより生徒の興味向上を心掛けた。講義後、高校生から「自分自身の映像が変化してからこそその驚きや、コンピュータビジョンが視聴体験に与える影響を実感した」という意見があった。このように高校生は、大学院で学ぶコンピュータビジョン技術を実体験として学習できる。総合的な探究の時間の目的である“どのように社会と関わり、よりよい人生を送るか”、“何を理解しているか・何ができるか”、“理解していること・できることをどう使うか”について、効率的に学習できる。したがって、学びに向かう力、思考力、判断力、表現力は、自然な形で身に付けることが可能であり、新しい IT 技術を学習できる高度な教育活動と位置付けられる。

4.2 IT の活用体験の向上に関する意向調査

動画編集方法やコンピュータビジョン技術を活用した映像生成手法についての講義終了後、ダンス部員 24 名に対して IT の活用体験の向上に関する意向調査を図 14 のようにフォームを用いて実施した。講義の目的は生徒の IT への興味向上であることから、体験した生徒の意向調査により IT の活用体験の向上に関する評価が可能である。

参加者は、以下の 2 項目の設問に回答した。

- ダンス ONE プロジェクトや動画編集方法について、講義を通じた動画作成への興味の上
- コンピュータビジョン技術を活用した映像生成手法に関する講義について、コンピュータビジョンの理解や興味の上

図 14 Google Form を用いた意向調査画面

Fig. 14 Screen of the intention survey.

参加したダンス部員は 24 名中、1 年生が 15 人、2 年生が 9 人、文系が 16 人、理系が 8 人、ダンス歴 3 年未満が 20 人、3 年以上が 4 人である。図 15 に意向調査の結果を示す。

ダンス動画の制作について未経験者が 8 割の中で、ダンス ONE プロジェクトに参加する前は動画制作に興味があり制作してみたいと思っていた生徒が 2 割、興味はあるが制作してみようとは思わなかった生徒が 6 割であった。ダンス ONE プロジェクトに参加後、制作してみたいと思った生徒が 2 割から 8 割にまで増え、ダンス ONE プロジェクトを通して動画制作への興味が向上したことが分かった。動画編集方法についての講義を実施した後、スマートフォンを用いてダンスパフォーマンス動画を制作してみたいと思った生徒が講義前よりも 1 割増加した一方で、PC を用いて高度な動画を制作してみたいと思った生徒が講義前よりも 1 割減少した。動画編集方法の実演によって普段から使用していない PC での動画制作の難しさを感じ、意欲が減少したと考察する。また、講義ではスマートフォンでの編集の限界について多く言及しなかったことから、スマートフォンにはない PC の高度な編集要素の説明が不十分であったと考察する。図 15 下に示すように、生徒属性の比較では、1 年生よりも 2 年生、ダンス歴が短い生徒よりも長い生徒のほうがダンス動画制作への興味が向上している割合が大きく、ダンス経験と動画制作への興味との相関がみられた。

コンピュータビジョンに関する意向調査では、コンピュータビジョンについて全く知らなかった生徒が 8 割、単語は聞いたことがあるが詳しく知らない生徒が 2 割と、24 名全員がコンピュータビジョンについて知らなかった。理解が深まった回答や興味が沸いた回答が多く得られたことは、コンピュータビジョン技術について触れ、理解する機会を提供できたことを示し、参加者は実体験として新し



図 15 IT の活用体験の向上に関する意向調査

Fig. 15 Results of an intention survey on improvement of IT utilization experience.

い IT 技術を学習したと考察した。

このように、スポーツ・文化を通じた教育活動において、最先端のコンピュータビジョン技術を活用した映像生成手法を提案し、高校生がダンス ONE プロジェクトへ取り組む活動をサポートすることで、探求学習科目「総合的な探究の時間」と同様の経験が可能になる。このように、探求学習科目「総合的な探究の時間」と同様の取り組みを部活動の一環とすることで、新たな教育モデルを実践した。

5. 本プラクティスの評価と次のダンス ONE プロジェクトへの展望

本研究では3つのプラクティスすなわち、ダンスパフォーマンス映像表現の高度化、映像制作における編集作業の支援、高校生が探究的に学習し IT 活用体験を向上させることのできる教育を実践した。

ダンスパフォーマンス映像表現の高度化では、ダンスと場を連動させた映像生成手法を提案した。提案手法の音響処理はオンセットのずれを想定するテンポで補正される

ため、テンポ変更やリズム変更が曲中でされないことを前提とする。したがって、対象音以外の音が含まれる混合音ではオンセットが不明瞭になる。テンポ変更やリズム変更が曲中でされる場合には正常に機能しない可能性が高い。音楽情報処理や人と共演する音楽ロボットの研究 [13], [14] では性能のよいビートトラッキングが提案されているため、提案手法の拡張によってより広範囲のユーザの獲得が期待できる。

映像制作における編集作業の支援では、動画編集の難しさを解決するための自動処理を提案した。出力された情報を Adobe Premiere Pro へ組み込むプラグインを実装することで、クリップの切り替わりを体感的に理解させることが可能となり、動画編集への理解をさらに深められる。

高校生が探究的に学習し IT 活用体験を向上させることのできる教育では、動画編集方法やコンピュータビジョン技術を活用した映像生成手法についての講義を実施し、高校生へコンピュータビジョン技術に触れる、理解する機会を提供した。高校生には各々の時代においてトレンドがある。たとえば、新しい日本語や制服のアレンジ方法、特定

の SNS アプリケーションによる発信などが挙げられ、一つの時代も高校生は流行を発信してきた。そこで、現在の高校生の流行あるいは流行の源となる題材とコンピュータビジョンをコラボレーションさせる。したがって、これまでに実施したコンピュータビジョン技術に触れる、理解する機会は、高校生のオリジナリティを強調する機会へ広げた教育へ発展する。

6. おわりに

本稿では、ダンス ONE プロジェクトに参加する高校生の撮影および動画制作をサポートする形で協力し、IT の活用体験を向上させることを目的とした、コンピュータビジョンの最先端技術を活用した映像生成手法を提案した。ダンス ONE プロジェクトを通して、探求学習科目「総合的な探究の時間」と同様の取り組みを部活動の一環とし、新たな教育モデルを実践した。

研究協力校ダンス部の高校生は我々の支援のもとコンセプト作りや撮影作業などに取り組んだ。その中で我々は音源のビート検出手法とオブジェクト検出アルゴリズムを活用したダンスと場を連動させた映像生成手法を提案した。信号機の色を映像処理によって拡張し場とダンスが一体化しているような効果を狙い、違和感のない映像を生成した。高校生は撮影・生成された映像を使用してダンスパフォーマンス映像クリップの編集を我々と行い、ダンス ONE プロジェクトへ動画を投稿した。

動画編集の難しさを解決するためにカット割りの検出とズーム設定の検出の自動処理を提案した。高校探究学習科目「総合的な探究の時間」に向けて、高校生に対してソフトウェアを用いた動画編集方法とコンピュータビジョン技術を活用した映像生成手法について講義した。IT の活用体験の向上に関する意向調査では、ダンス ONE プロジェクトと本研究活動を通して参加者の IT の活用体験が向上したことを確認した。

本研究の実施にあたり、筑波大学倫理委員会の承認を得た。

参考文献

- [1] ひとつになろう!ダンス ONE プロジェクト'21: <https://www.ntv.co.jp/sukkiri/dance/index.html> (参照 2021-07-13).
- [2] McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E. and Nieto, O.: librosa: Audio and music signal analysis in Python, *Proc. of the 14th python in science conference*, pp.18-24(2015).
- [3] Duxbury, C., Bello, J. P., Davies, M. and Sandler, M.: Complex domain onset detection formusical signals, *Proc. Digital Audio Effects Workshop (DAFx)*, London: Queen Mary University, p.6-9(2003).
- [4] Davies, M. E. P., Plumbley, M. D. and Eck, D.: Towards a musical beat emphasis function, *IEEE Workshop on Applications of Signal Processing to Audio and Acous-*

tics, pp.61-64(2009).

- [5] Yuan, Y., Chen, X. and Wang, J.: Object-Contextual Representations for Semantic Segmentation. European conference on computer vision, *Springer, Cham*. pp.173-190(2020).
- [6] Wang, C., Yeh, I. and Liao, H.: You Only Learn One Representation: Unified Network for Multiple Tasks, arXiv preprint, aeXiv: 2105.04206(2021).
- [7] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A.: You only look once: Unified, real-time object detection, *In Proc. of the IEEE conference on computer vision and pattern recognition*, pp.779-788(2016).
- [8] Abraham, A., Purwanto, D. and Kusuma, H.: Traffic Lights and Traffic Signs Detection System Using Modified You Only Look Once, *In 2021 International Seminar on Intelligent Technology and Its Applications (ISITIA) IEEE*. pp.141-146(2021).
- [9] Li, Q. and Wu, M.: An Improved Hough Transform for Circle Detection using Circular Inscribed Direct Triangle, 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp.203-207(2020).
- [10] Cuevas, E., Zaldivar, D., Pérez-Cisneros, M. and Ramírez-Ortegón, M.: Circle detection using discrete differential evolution optimization. *Pattern Analysis and Applications*. pp.93-107(2011).
- [11] Cuneyt, A. and Cihan, T.: EDCircles: A real-time circle detector with a false detection control. *Pattern Recognition*, Vol.46, Issue 3, pp.725-740(2013).
- [12] Cao, Z., Hidalgo, G., Simon, T., Wei, S. -E. and Sheikh, Y.: OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, *IEEE Trans. on Pattern Analysis and Machine Intelligence*. Vol.43, No.1, pp.172-186(2021).
- [13] Yoshii, K. et al.: A Biped Robot that Keeps Steps in Time with Musical Beats while Listening to Music with Its Own Ears, *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp.1743-1750(2007).
- [14] Otsuka, T. et al. Real-Time Audio-to-Score Alignment using Particle Filter for Co-player Music Robots, *EURASIP Journal of Advances in Signal Processing*(2011).



吉川 優依 (学生会員)

2021年筑波大学理工学群工学システム学類卒業。同大学大学院理工情報生命学術院システム情報工学研究群知能機能システム学位プログラム博士前期課程在学中。オーディオビジュアル複合情報処理の研究に従事。



宍戸 英彦 (正会員)

2016年筑波大学大学院システム情報工学研究科修了。同年、(独)国立スポーツ科学センター研究員を経て、2017年より筑波大学計算科学研究センター助教、2018年、英国サリー大学客員研究員兼務。コンピュータビジョン、多視点映像解析の研究に従事。博士(工学)。



亀田 能成 (正会員)

1991年京都大学工学部情報工学科卒業、1993年同大学大学院工学研究科卒業、1999年同大学博士(工学)受領。1996年より京都大学助手。2001年マサチューセッツ工科大学人工知能研究所客員研究員(兼任)。2002年筑波大学講師着任。現在、筑波大学計算科学研究センター教授。研究テーマは、コンピュータビジョン、3次元データ処理、VR、AR、複合現実感、3次元ユーザインタフェースなど。スポーツ科学や福祉技術への応用も研究テーマとしている。



北原 格 (正会員)

1996年筑波大学大学院理工学研究科修了。同年シャープ(株)入社。2000年筑波大学先端学際領域研究センター助手、2003年(株)国際電気通信基礎技術研究所(ATR)研究員、2005年筑波大学大学院システム情報工学研究科講師、2008年同准教授を経て、2019年より同大学計算科学研究センター教授。自由視点映像、複合現実感の研究に従事。博士(工学)。