

VRMixer: 動画コンテンツと現実世界の融合と その適用可能性の検証

牧良樹^{†1†2} 中村聡史^{†1†2} 平井辰典^{†3} 湯村翼^{†4} 森島繁生^{†2†3}

本稿では、動画コンテンツと現実世界とを融合させ動画の世界の中に入り込むことを可能にするVRMixerをもとに改良を行い、その動画世界内で体の大小を変化させながら自在に楽しむことを可能とする仕組みを実現する。また、種々のコンテンツに対して本手法を適用し、ユーザ実験を実施することによって動画コンテンツに入り込む楽しさについて検証を行うとともに、その問題についても明らかにする。

VRMixer: Possible Applications of Mixing Video and Real World

YOSHIKI MAKI^{†1†2} SATOSHI NAKAMURA^{†1†2} TATSUNORI HIRAI^{†3}
TSUBASA YUMURA^{†4} SHIGEO MORISHIMA^{†2†3}

In this paper, we improve VRMixer which mixes the real world and a video clip in order to enable users to change the size of their body and position in a video clip by using simple gestures. Then, we clear possible applications and problems of our system by doing several experimental tests.

1. はじめに

YouTube やニコニコ動画に代表される動画共有サイトの登場によって、誰でも気軽に作成した動画を投稿し、共有し、視聴してもらうことができるようになった。こうした流れにより、これまでのテレビなど大手のマスコミから放送される動画だけではなく、一般のユーザの作成した動画を視聴できるようになり、楽しめる動画の種類が広がっている。しかし、映像視聴行為はそれほど変化しているわけではなく、視聴者は映像が呈示されるディスプレイと正対して視聴するのが一般的である。

森島らの Dive into the Movie[1]では視聴者の顔を映像の中の登場人物にはめこむことで、自身が登場人物の一人となって映像の中に入りこむことを可能としている。しかし、このシステムでは特別な装置が必要であり、使用する動画を作り込む必要があるため、動画に対して自由に入ることができなかった。

一方、我々はこうした映像の中に入りこむことを容易化し視聴を一つの体験として楽しむことを可能とするために、VRMixer[13]というシステムを提案および実装してきた。VRMixerでは既存の映像の中から人物を抽出し現実の映像から抽出された人物を重畳表示することにより、ユーザが

より深く動画の世界に入り込むことを可能としている。こうした仕組みにより、音楽動画内のダンサーと一緒に踊ることができるようになっていた。しかし、これまでのシステムでは、映像に入る際には自分の立ち位置の調整を行わなければならない、また映像の中に入った際にどの位置に自分を投影させるかといったことは変更できなかった。さらに、これまでの研究においてダンス練習において有効なのではというフィードバックをもらっていたが、十分にコンテンツ検証できておらずダンス練習において本当に有効であるかなど、その有効性や特徴などを十分には明らかにできていなかった。

そこで本研究では、まずVRMixerを発展させ、自身の縮尺や位置などの変更を可能にする手法を提案する。この手法により、ユーザは映像の中を自由に動くことができるようになるとともに自身の縮尺や位置を変更することによって映像の中身の大きさ感を変更し、映像の楽しみ方を拡張することを可能にする。また、様々なコンテンツに対する利用実験を行うことによりその有用性を明らかにする。最後に、VRMixerのダンス練習における有効性について検証し、本提案手法の適用可能性について議論する。

^{†1} 明治大学
Meiji University

^{†2} JST CREST

^{†3} 早稲田大学
Waseda University

^{†4} 北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

2. VRMixer

VRMixer では、動画像処理によって動画から人物を抽出し前後に配置することで現実の3次元空間と、動画の2次元空間の間にある2.5次元空間を構築する。こうした2.5次元空間にある動画と、現実世界の人物を組み合わせることにより現実世界の人物を動画の世界に入りこませている。また、同じ仕組みを逆に適用することで、動画内の人物を現実世界の映像に登場させることも可能となっている。VRMixer では、人物領域の抽出と2.5次元空間の構築の2つの工程でシステムを形成している。



図1 VRMixerのイメージ図

2.1 動画セグメンテーション

動画セグメンテーションは動画内にいる人物領域を抽出することにより、動画を人物領域と非人物領域に分けるものである。これを行うことにより、動画内の人物を現実世界に表示させることも可能となる。人物領域の検出は、これまでも多く提案をされている。VRMixer では導入コストを低くするため、セグメントを全自動化している。ここでは、人物領域を効率よく検出するため、顔検出手法を主な手法として用いている。顔検出は検出することのできるシーンや姿勢が幅広く、対象が未知の場合でも高い精度で人物領域の検出が精度を高くかつ自動で行うことができ、顔の特徴点を検出も行うことができる。この顔検出の結果を用いることで体の大きさなどを推定する。さらに体の範囲を求めるために顔から十分離れた範囲を背景領域と予測することで、人物抽出を行いやすくしている。さらに、体の領域を正しく検出するために人物検出も補助的に用いていく。顔検出では、横顔や後ろ向きの際には検出が難しいためHOG特徴量をSVMで学習させた人物検出[11]を用いることで横顔や後ろ向きの際にも検出可能性を増している。こうして検出した人物領域と非人物領域を基にして人物領域の精度向上を行う。

また、検出した領域のピクセル値に対してVB-GMM[10]を適用し、色分布を学習させる。その後、色分布と検出し

た動画をシードとしてmin-cut/max-flowアルゴリズム[6]を用いた3次元グラフカットにより自然なセグメント境界を得る。さらに人物領域抽出をさらに正確に行うために先ほどのグラフカットを行う際に、Gunnar Farnebackアルゴリズム[8]を用いて、フレーム内の動きの大きさに重み付けを行うことでダンス動画のように人物の動きの大きい動画においてより正確な人物領域検出を行いやすくしている。このようにして人物領域と非人物領域に分けた動画を使って、動画と現実世界の融合を行う。

2.2 動画と現実世界の融合

2.1節で述べた動画セグメンテーションの結果を用いて人物領域を前に、非人物領域を後ろにくるように3次元空間内の異なる距離に配置する。そこにKinectでリアルタイムに抽出したユーザを混ぜ合わせ、Kinectで取得された距離を用い、ユーザを適切な位置に配置することで新たな2.5次元空間を構築する。

図2はVRMixerで実現されている2.5次元空間を、模式的に表したものである。図中で赤線が引かれている箇所が実際に合成されたオブジェクトである。現実世界から合成されるものはユーザの身体のみに限らず、背景領域に配置するものであれば合成結果に反映することができる。その結果、図3のように動画の中に実世界のユーザが入り込むことが可能となっている。

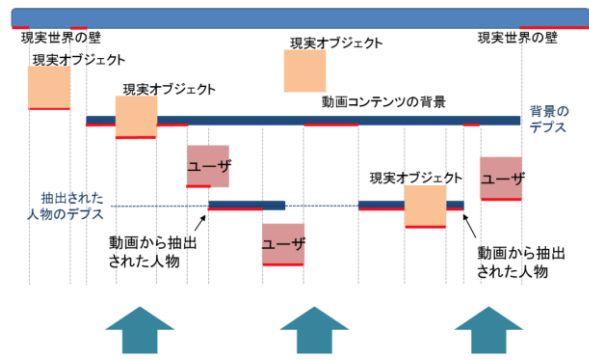


図2 動画と現実世界との融合の模式図

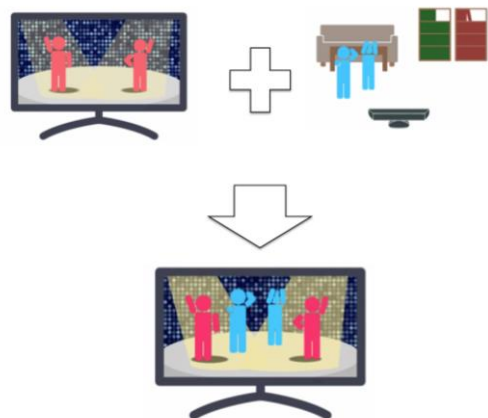


図3 VRMixerの概要

3. 改良手法とプロトタイプシステム

これまでの VRMixer では、動画内の人物と共演する際に、ユーザが自身を Kinect の前で上下左右前後に動くことで動画内の人物との位置を合わせる必要があった。そのため、ある一定以上の大きさや小ささになることが出来ず、体験可能な動画が限定的であった。また、ユーザは映像内での大きさが固定されており、新しい体験を創り出すという点では不十分であった。

ここで映像作品において、その中に現れるモノの大きさや小ささを表現するために人間などの比較対象を登場させることが多い。視聴者は、この映像内に現れる人間などから、その世界の大きさを把握することが可能である。つまり、映像内にユーザを大きく登場させたり、小さく登場させたりすることによって、新しい体験を創り出すことができるのではと考えられる。

そこで本研究では、VRMixer の改良として、ユーザ（視聴者）を任意の大きさに小さくしたり、大きくしたりして映像中に登場させる手法を実現する（図 4）。

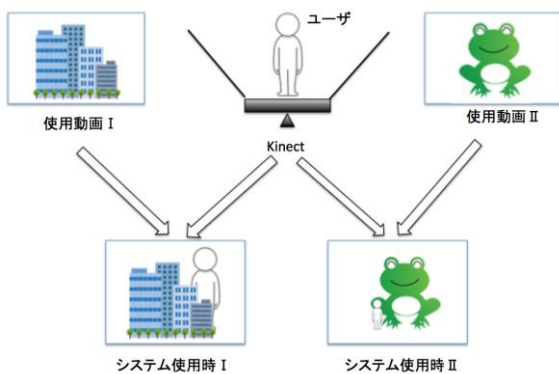


図 4 システム使用時イメージ

この大きさの変化により、ユーザは縮尺の違う様々な動画の中に入ることができるようになる。また、本来の人間サイズとは異なる大きさでの体験が可能になると期待される。例えば、怪獣映画内では自分が怪獣と一緒に建物を破壊することができる一方、逃げ回る人々となることで、本物の怪獣の大きさを体感し、恐怖を感じることができるようになると期待される。また、虫や鳥の世界に入り込んで、その大きさやインパクトも体験可能となると期待される。しかし、小さくなるだけでは、映像内のある部分でしか動きまわることが出来ず、十分な体験はできない。そこで、映像内を自由に動き回ってあらたなる体験を可能とするため、ジェスチャにより自身のサイズを大きくしたり小さくしたり、空間を移動できるようにするシステムを構築する。これにより、現実では得ることができない体験を、動画の中に入りながら体験ができるようになることが期待される。

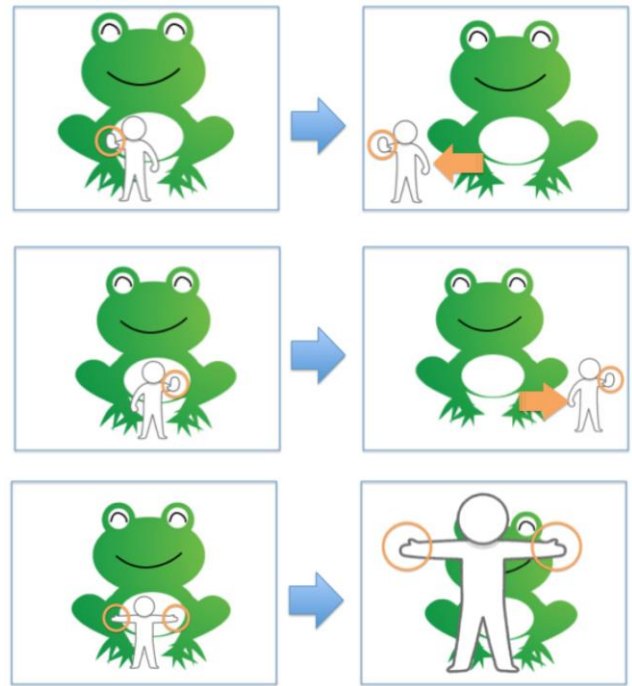


図 5 ジェスチャ操作のイメージ図

提案システムを、Processing と Kinect および、SimpleOpenNI などのライブラリを用いて実装した。実装したシステムの動作例は図 6 の通りである。



図 6 システムの動作例

4. コンテンツ体験実験

本システムを使うことでスピード感やスケール感などの感じ方を変更させることが可能なのか、またどのような動画において有効で、どのような動画において有効ではないのか、感じ方がどのように変わるのかなどを検証するため利用実験を行った。

4.1 実験内容

動画をただ試聴する場合と、提案手法によって動画内に入ることによって感じる事の変化を実験するため、動画コンテンツを用意した。ここで用意したコンテンツは、小さくなってカエルを見る、怪獣と一緒に暴れまわる、街並みに巨人化して入る、聴衆の前でプレゼンテーションする、

野球のバッティングする、バラエティ番組に入り込むという6つのパターンである。この6つのパターンについて、単純に動画を視聴した時と、VRMixerで入り込みながら視聴した時とでどのような違いがあるかを人によって手法を切り替えることで検証した。

実験では、大学3年生の男子学生11人を2つのグループに分け、グループによってコンテンツの視聴方法を変更した。また、それぞれの動画の体験後アンケートを行った。アンケートにおいてはカエルと怪獣、街並みについては自身のサイズ感は変わったのかを、バラエティ番組では面白さを、プレゼンテーションでは高揚感を、バッティングではピッチャーの投げる球のスピード感を評価してもらった。

なお、実験は明治大学中野キャンパスの体育館で行った。本システムによって重畳された動画は、プロジェクターによってスクリーンに映し出された。また、Kinectはユーザの正面またはユーザの背面に置き（コンテンツによって配置を変更した）、スクリーンに映し出されているユーザの身体は鏡のように左右反転させた。

4.2 実験結果

VRMixerを用いて体を小さくして自身よりも何倍も大きなカエルの前を通る動画体験は、単純に視聴した場合に比べ全体的に評価が高く、中でもある実験協力者は「カエルが動くのを見ると、自分が何かされるのではないかと思ひびっくりした」とコメントしていた。つまり、実際に映像の中に入る感覚を得ることができていると考えられる。このことから、様々な動画の中にドラえもんのスモールライトのように小さくなって入り込み世界を体験するようなことに本システムは有効であると考えられる。

怪獣の戦闘シーンに入る動画については、VRMixerを用いた際の実験協力者の評価は低かった。これは、この動画では比較対象が怪獣とその敵であり、自身とその怪獣の大きさが変わらず、あまり自分が大きくなった印象を得ることができなかったからだと考えられる。一方、街並みを撮影した動画の中に入ってもらったもの（図7）では、比較対象が道路を走っている車やビル、通行人となり、車や人と自身の大きさを比較することができるため、大きくなった感覚を得ることができるようになっていた。その結果、進撃の巨人に出てくる巨人のモノマネなど大きくなったとき特有の行為をする実験協力者もいた。つまり、スケール感の変更がある程度可能であったと考える。ただ、実際に映像の中で車を蹴飛ばしたりできるわけではないため、効果は限定的である可能性がある。

こうしたことから、どちらかと言うと映像内に対して影響をおよぼすことが出来ない小さいサイズになる方がコンテンツとしては面白く、大きいサイズになる方はコンテンツに工夫が必要であると考えられる。

審判視点で撮影されたピッチャーとバッターの対戦動画にバッターとして入る動画では、多くの人がVRMixerでのコンテンツ体験を楽しんでおり、動画をただ視聴したグ

ループに比べ、そのボールのスピードに驚いていた。これは、動画内に入ることでそのスピードをより強く体験出来ていることが理由だと考えられる。また、動画内のバッターが打ったときにタイミングよくスイングをすることでユーザが擬似的に打つことができるような体験を創り出せていた。これは、World Connector[2]で渡邊らが提案している映像の物体と現実の物体を一体化させ、映像の世界に入り込む感覚を得ることを可能とする手法と関連が有り、興味深いものである。

バラエティ番組の映像に入り込み、出演者と一緒になってネタをするというコンテンツにおいては、あまり評価が高くなかった。これは、ご当地番組特有のシュールな笑いを、番組内に入ることで感じることができなかった点が考えられる。また、今回の実験協力者にはこの番組を知っている人が少なかったことも原因と考えられる。

Apple社のスティーブ・ジョブズがiPhoneを発表した歴史的な場面に入り込み、一緒に説明を行う（図8）というものについては、実験協力者のコメントでは大観衆の前でアピールを行っているように感じたというコメントや、ジョブズという偉大な人物と共演をすることができるように感じて楽しかったというコメントなどがあった。このことから、有名なシーンに入り込むことはユーザの体験を高めることに繋がると考えられる。



図7 街の中への重畳



図8 Appleのカンファレンス

今回の実験により、本システムによる体の大きさを小さくしたときに動画の迫力を増加させることが可能であるこ

とがわかった。また、スピード自体も強く意識することができるようになることも分かった。

なお、今回の実験中に実験協力者が大きくジャンプをした際に、実験協力者が映像内でかなり上方に移動すると、実験協力者はまるで自身の能力が拡張され大きく跳ねることができるように感じていた。人間の能力が拡張されているかのように見せる仕組みとしては、Strato Jump[12]などがある。このシステムでは GoPro を取り付けたヘリウム気球を上空まで上下運動させて撮影した映像を、ユーザのジャンプ行動に合わせて Oculus Rift に写すことにより、ユーザがとても大きくジャンプをしたような感覚を得ることを可能にしているものである。我々の手法はこうした身体能力の拡張を手軽に実現することができる可能性がある。

5. ダンス練習における有効性の検証

VRMixer の初期のテストにおいて複数の体験者からダンスの練習において有効なのではとのコメントがあったが、その有効性を検証できていなかった。そこで、VRMixer がダンス練習において有効なのか、またどのようなユーザにおいて有効で、どのようなユーザにおいて有効ではないのかを検証するため、評価実験により明らかにする。

5.1 実験内容

ここでは、ダンス練習において鏡とディスプレイを用いる従来型のタイプと、VRMixer を用いるタイプとで比較し、そのダンスの覚えやすさなどから評価を行う。従来型のタイプでは、ディスプレイと鏡を図9のように配置し、ディスプレイ上にダンス映像を流すことで手本となるダンスを確認しながら鏡を使い自身のダンスを確認することを可能にしている。VRMixer を用いるタイプでは、図10のように一番奥にスクリーン、その手前にプロジェクターと Kinect、VRMixer 処理用の PC を配置している。Kinect をユーザの正面に配置することで、ユーザは自身の正面からの映像をみることができるようになっている。これにより、ユーザは動画の中の人物と同じ状況になって練習をすることができるようになる。

実験では、1分～1分半の2曲のダンス動画 ab を用意した。この2つの動画を選んだ理由は、VRMixer で使用しやすい正面からの固定カメラで撮られている「踊ってみた」動画であり、VRMixer の経験者でも30分の練習でギリギリ覚えることができるということである。



図9 鏡とディスプレイを使った実験の様子



図10 VRMixer を用いた実験の様子

実験はダンス経験者の女子大学生2人と、ダンス未経験者の男子大学生12人に依頼した。ダンスの練習では、この実験協力者を2つのグループに分け、グループによって従来型のタイプを使うダンス動画と、VRMixer タイプを用いるダンス動画を切り替えた。各ダンス練習においては、実験協力者は30分間の時間が与えられ、思い思いの方法でダンスの練習をしてもらう。また、30分のダンス練習終了後、1人ずつカメラの前で音楽だけを聞きながらダンスを思い出しつつ踊ってもらい、その様子を撮影した（撮影時は、プレッシャーがかからないように誰も見ないようにした）。さらに各実験終了後に実験協力者に対し、「リズム」「上半身のフリ」「下半身のフリ」「メリハリ」「滑らかさ」「左右のフリ分け」「楽しさ」の7項目について、「よくできた」「できた」「ややできた」「どちらともいえない」「ややできなかった」「できなかった」「まったくできなかった」の7段階で評価を行ってもらった。なお実験では、経験者は1人ずつダンス練習を、未経験者については2人ずつダンス練習を行ってもらった。

また、すべてのダンス収録が終わった後、実験協力者が踊った動画と元々のオリジナルのダンスを横に並べた動画を作成し、ダンス経験者の女子大学生1人、未経験者の女子大学生1人、男子大学生2人に見てもらい、12人の2回分のダンスに対する評価を行ってもらった。ここでの評価も、先述のアンケートの内容と同じにした。なお、本評

a【みりん】ハッピーシンセサイザを踊ってみた
<http://www.nicovideo.jp/watch/sm15302883>

b【恋愛サーキュレーション Full ver.】を踊ってみた *馬琴
<http://www.nicovideo.jp/watch/sm10767335>

備では、どのダンスにおいて VRMixer を使用したのかわからないよう、提示順番を入れ替えるなどした。

5.2 実験結果（ユーザ評価）

実験結果をまとめたものが表 1~5 である。アンケートは「よくできた」を+3, 「まったくできなかった」を-3 としてシステムの評価を行っている。なお, VRMixer を使用したものを「VRMixer」と、鏡とディスプレイを使用したものを「既存手法」と表記する。

表 1 は, VRMixer を使用して練習行った場合と、ディスプレイを使用して練習行った場合のそれぞれ項目に対する平均と分散を示したものである。この結果から、上半身の振り付けと楽しさについて、VRMixer 使用時の方が既存手法に比べスコアの平均が高く、分散も低いことがわかる。つまり、上半身の振りと楽しさについては VRMixer がやや高く評価されていると言える。しかし、楽しさについては、既存手法は分散が大きく、これは動画コンテンツによって評価が割れた可能性がある。また、他の項目についても、分散がかなり大きい個人差が大きいと考えられる。

表 1 VRMixer と既存手法の平均と分散

	VRMixer使用時		既存手法	
	平均	分散	平均	分散
リズム	1.071	3.302	1.214	3.258
フリ(上半身)	0.786	0.951	0.143	3.516
フリ(下半身)	-0.429	2.110	-0.643	2.863
アクセント	0.071	2.995	-0.714	2.220
滑らかさ	-0.071	2.379	-0.643	2.401
左右	0.857	4.132	-0.571	2.725
楽しさ	1.929	2.841	1.214	4.951

そこで、それぞれの手法を用いたときの個々の評価を整理したものが表 2 および表 3 である。この結果より、経験者と未経験者の差、さらに未経験者内でもかなり差があることが分かる。一方、リズム、アクセントについて経験者の VRMixer の評価が低くなっている。これはアンケートの回答などから、VRMixer 上に提示される自身の映像に遅延があったことが原因であると考えられる。実際に VRMixer を用いた場合踊りとお手本とを照らし合わせると、タイミングがずれている箇所があった。つまり、この遅延をなくすことが VRMixer をダンス練習に応用するための最優先事項であると考えられる。

表 2 および表 3 によると、上半身の振りについては、経験者は VRMixer を低く、未経験者は VRMixer を高く評価していることがわかる。これは練習中の経験者の発話から、経験者が左右反転されているダンス動画に慣れており、VRMixer でわざわざ左右反転させているものを、気持ち悪く感じたことが理由として考えられる。一方、鏡を使った練習をしたことがない未経験者には、左右反転して動画の中に入るのが有効であると言える。

下半身の振りについて、未経験者はいずれも低い評価となっているが、これは 30 分という限られた実験では十分にこなすことが出来ず、下半身に追いつけなかったことが原因であると考えられる。今後は、この時間を長くすることによりさらなる検証を行う予定である。

コピーの楽しさについて、VRMixer の方が、スコアが高くなっている。これは、VRMixer により動画の中の人と一緒に踊ることができたことが影響していると考えられる。しかし、楽しさはコピーした曲に依存する可能性もある。そこで、曲別に分けたアンケート結果を表 4 に表す。

表 2 VRMixer

VRMixer	経験者			未経験者													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N			
リズム	1	1	2	3	2	2	-2	-1	3	-3	2	1	2	2			
フリ(上半身)	1	0	2	0	1	1	1	1	1	-2	1	1	2	1			
フリ(下半身)	2	-2	0	0	-1	1	-2	1	1	-2	-1	-2	-2	1			
アクセント	1	-3	2	0	-1	2	-2	0	1	-2	3	0	1	-1			
滑らかさ	1	-2	1	0	-2	2	1	-1	-1	-2	1	2	-2	1			
左右	3	-2	1	-1	2	3	1	3	-2	-3	1	2	2	2			
楽しさ	3	-1	3	3	2	3	-2	3	3	0	3	2	3	2			
合計	12	-9	11	5	3	14	-5	6	6	-14	10	6	6	8			

表 3 既存手法

既存手法	経験者			未経験者													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N			
リズム	2	1	1	2	3	0	3	2	3	-3	-2	1	2	2			
フリ(上半身)	2	1	-2	-1	2	-2	2	1	1	-3	-2	-1	2	2			
フリ(下半身)	2	1	-2	-1	-1	-3	0	-1	1	-3	-2	-2	2	0			
アクセント	2	1	-2	-2	0	-3	-1	0	-1	-3	-1	-1	0	1			
滑らかさ	2	-2	0	-2	0	-3	1	1	0	-2	-2	-2	1	-1			
左右	3	1	0	-2	-1	0	-2	-1	1	-3	-2	-2	1	-1			
楽しさ	3	2	-2	2	-1	1	3	3	3	-2	-3	2	3	3			
合計	16	5	-7	-4	2	-10	6	5	8	-19	-14	-5	11	6			

表 4 各曲の平均と分散

	アニメソング		ボーカロイド楽曲	
	平均	分散	平均	分散
リズム	1.500	2.269	0.786	4.027
フリ(上半身)	0.714	2.066	0.214	2.489
フリ(下半身)	-0.571	2.264	-0.500	2.731
アクセント	-0.214	3.566	-0.429	1.956
滑らかさ	-0.286	2.989	-0.429	1.956
左右	-0.214	3.566	0.500	4.115
楽しさ	2.143	1.670	1.000	5.692

表 4 より、各曲の平均と分散の差が少ないことから、曲による踊りやすさの差はあまりなかったと考えられる。一方、楽しさでは大きな差がでており、分散も高い。これは、ボーカロイドの楽曲を知っている人、知らない人で楽しみ方が大きく変わったことが考えられる。また、実験協力者の「アニメソングは聞き取りやすく乗りやすかったが、ボーカロイド楽曲は歌詞が聞き取りにくく乗ることができなかった」というコメントから、聞き取りやすさが影響した可能性もある。今後は楽曲のタイプを揃えるなど工夫する予定である。

5.3 他者評価実験の結果

他者による評価実験において経験者と、未経験者でありながらダンスがうまい人、ダンスがあまりうまくない人と

の間にかなり差があったため、経験者と他者評価による項目の合計値で上位4人(高得点者)と中位4人(中得点者)、下位4人(低得点者)の3つのグループに分けた。この各グループに対する各項目の評価平均について、VRMixerに対する評価から既存手法に対する評価を引いた結果を図11に表す(棒グラフが上に伸びているものはVRMixerが良く、下に伸びているものは既存手法が良い)。また、評価者による評価値平均から各グループのダンス練習の協力者の自身の評価の平均を引いたグラフを図12に表す(棒グラフが上に伸びているものは評価者のスコアがよく、下に伸びているものはダンス練習者のスコアがよい)。

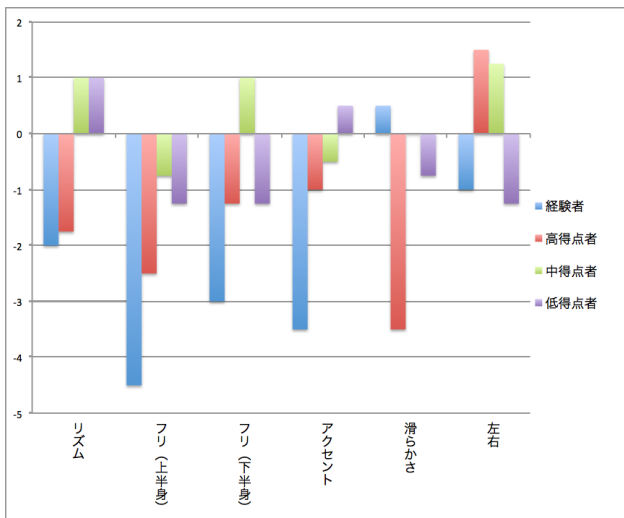


図 11 評価実験アンケート結果

図 11 から、経験者、高得点者、低得点者においては全体的に既存手法の方が点数は高くなっていることがわかる。しかし、中得点者においては、VRMixer の点数がとて高くなっている。このことより、中程度の未経験者においては VRMixer が有効であると考えられる。一方、図 12 を見ると、中得点者は、評価者のスコアより、自身の評価の方が高いことがわかる。これは、VRMixer を使って覚えた自身が覚えたフリ付けをうまく踊れなかったのではと考える

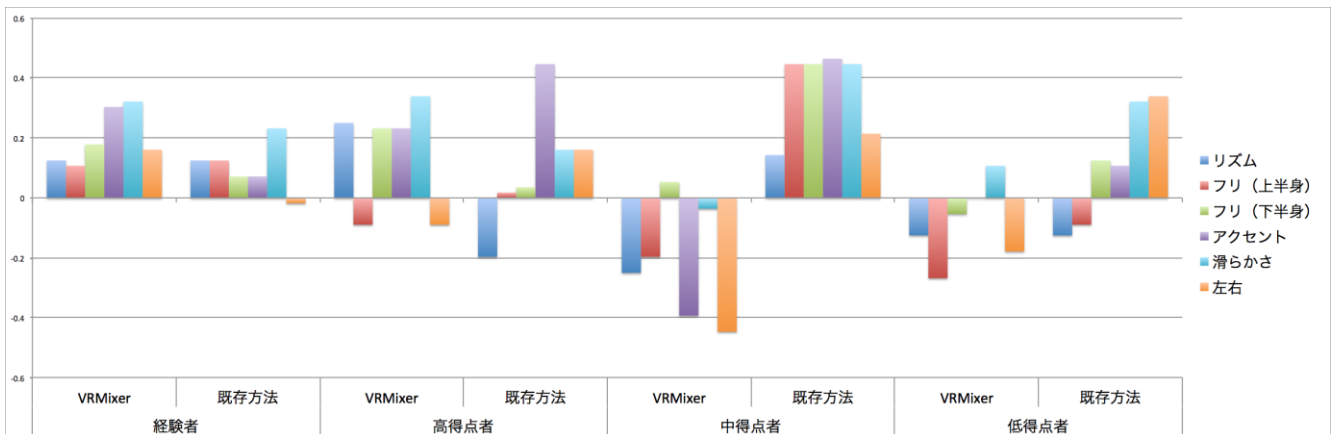


図 12 他者評価と自己評価の比較

ことができる。この傾向は、低得点者にも見ることができると、中得点者や低得点者が VRMixer を使いさらに練習

を積むことで、ダンス練習者が自身の覚えた振り付けをうまく踊れるようになり、得点が伸びると期待される。

ダンスの自己評価と、他者評価に特に差があったのが、上半身のフリである。これは、手本が常に横で踊っていることで、ダンスを覚えることよりも振り付けを真似することに注力してしまい、ダンスを忘れてしまいうまくダンスを踊ることができなかったことが考えられる。

左右の振り付けにおいて、高得点者で差が大きくでいた。これは、高得点者は、既存手法ではダンスのフリをしっかりと覚えることができるが、左右の振り分けは経験者のようには覚えることができていない。そのことで、左右が反転されている VRMixer では左右をしっかりと分けて覚えることが可能であったのでは考えられる。高得点者の自身のスコアと評価者の比較では、どちらも評価者のスコアの方が高くなっており、高得点者は自身が思っていたよりも、うまく踊ることができていたことがわかる。

6. 考察

ユーザがジャンプすると、ユーザ自身の映像が画面上を大きく上下に移動することができるようになると映像の体験はさらに拡張することができるようになることが実験により確認することができた。さらに Strato Jump[12]のようにユーザを上空に飛ばすことができるようにすることでこれまで見たことのない世界と自身との比較を動画内で行うことができるようになる。また、ユーザがジェスチャを行うことで、自身の拡大、縮小や左右を移動することを可能にした。これをさらに応用し、ジェスチャによって、ユーザの画面上の位置を変更することをできるようにすることで、動画内を自由自在に飛び回ることができるようになり、空を飛んでいるような感覚や魔法を使っているような感覚を得ることも可能である。このことで、ユーザは自身の体を自由に動かすことが可能となり、アクション映画をユー

ザが追体験することができるようになることが考えられる。

VRMixerでは、顔検出による人物の抽出を行っているが、カエルや怪獣などのように、人間ではないものを対象として処理することが出来ない。また、生物のみならず木やビルなどの後ろに回り込む事ができるようになると、より映像の中に入っている感じを演出することができると考えられる。今後はこうした任意の対象に対する前後関係の抽出を行い、その世界の中に入りこむ技術を実現する事が考えられる。

今回のダンス練習実験では、各曲に対して30分という短い期間での実験であった。そのため、実験協力者がVRMixerでの練習方法にまだ慣れていなかった可能性もある。長期的な実験によって結果が変わる可能性があると考えられるため、VRMixerを長期的に使用して練習したユーザと既存手法で練習したユーザがどのような変化があるのかを明らかにする予定である。また、今回の実験では単純にダンスを真似するというものであったが、本システムの特徴である深度を使ったフォーメーションダンスへの応用が可能であるかなどの可能性も考えていく予定である。

ダンス練習において何度も同じシーンを確認するときや動画を止めるときに、動画のシークバーをわざわざ操作する必要があった。こうした操作はダンス練習において手間であるため、今後はジェスチャによって動画のシークバーを操作可能とする手法を実現する。

本システムでは個人でのみ楽しむシステムになっている。しかし、WaaZam![3]のように遠隔地にいる人物と同じ動画の中に入ることが可能になれば、ダンス練習の際にユーザ同士がお互いのフリを、手本を見ながら確認することが可能になる。また、フォーメーションダンスを練習する際に人数が足りないときや、手本を見ながら自身の立ち位置などを練習したい際などの練習方法としてVRMixerを使用することで練習はさらに高度化に行うことができるようになる。こうした方法により、さらに新しい体験を得ることができるようになると期待される。

これまで、テレビやYouTubeなど様々なメディアで動画は消費されているが、見るという消費方法は変わっていない。我々は、VRMixerによって、動画の消費方法を変えていくことができるのではと期待している。

7. 関連研究

ノリ乗り♪システム[4]では、ユーザの骨格を抽出したものをノリとして表現し映像に重畳している。ノリを共有することにより視聴者はライブ会場で実際に他の視聴者と一緒に観客として一体感を楽しんでいる感覚を得ることを可能にしている。しかしこのシステムでは、観客側の一体感を得ることのみを目的としておりユーザ自身が出演者となって楽しむことはできない。

WaaZam![3]では、遠隔地にいるユーザ同士のコミュニケ

ーションを支援するためKinectで相互のユーザを切り出し、同一空間にマッピングする手法を提案している。この手法により遠隔地にいる人物同士がひとつの場所に集まりながら遊んだりプレゼンテーションしたりすることを可能にしている。しかしこのシステムは、コミュニケーションを支援するシステムであり、既存の動画に入りこんで楽しむことは提案されていない。

Dive into the Movie[1]は、ユーザが映画の登場人物の一人になることにより、その映画への没入感も増すことを可能にしつつ、楽しみを増幅させるものであり愛地球博で展示、披露されたものである。またこのプロジェクトの一貫で、ユーザを擬似的に映画の中に登場させるFuture Cast System[7]も提案している。このシステムでは、顔のみを抽出し決まったストーリーの中で登場することができる。ユーザは決められたストーリーの中でしか映像に入ることができず、ユーザは自由に動き回ることができない。

DDMixer2.5D[9]は、あらかじめ深度情報をもっている2.5次元の動画のオブジェクトとKinectから取得した現実世界の映像を融合させることで、動画と人物の融合を可能としている。しかし、DDMixer2.5Dは深度情報をもった動画にしか使用することができず、好きな動画の中に自由に入ることができない。我々の手法は、任意の動画に対して入り込むことを可能にするものである。

山内ら[5]は、Kinectとワイヤレスマウスを併用したダンスを併用したダンス学習支援システムを提案している。Kinectの骨格抽出によって判定したダンスの振り付けをうまくできているかの判定と、マウスクリックによるリズムをうまくとれているかの判定を行っている。しかし、手本はあらかじめ骨格抽出をされた動画を使用しなければならず、自身の好きな動画を使って練習を行うことができない。我々の手法は、ユーザの好きな動画の中に入ることが可能にするため様々な動画を使うことができるよう、必要な情報は最小限に抑えている。

8. おわりに

本研究では、VRMixerを改良し、ユーザの画面上でのジェスチャによるサイズ変更や位置変更を可能にすることでユーザのスケール感を変更することを可能にし、どのような動画において本システムが有効であるかを実験とアンケートによって確認した。また、ダンス経験者と未経験者にVRMixerを用いたダンス練習実験を行ってもらうことでダンス練習への応用の有用性と問題点を確認した。

VRMixerには、まだまだ発展の余地が残っており、本システムをさらに応用させていくことで既存の動画の楽しみ方を拡張できると考えられる。また、様々なコンテンツに対しても適応可能とし、その有効性について確認していく。

参考文献

- 1) Morishima, S., Yagi, Y., and Nakamura, S.: Instant Movie Casting with Personality Dive into the Movie System, *Journal on Virtual and Mixed Reality*, pp.187-196 (2011)
- 2) 渡邊恵太, 中村聡史.: WorldConnector: カメラの身体性付与による映像世界へ入り込むインターフェース, 第 19 回日本バーチャルリアリティ学会大会論文集(2014)
- 3) Hunter, S., Maes, P., Tang, A., Inkpen, K., and Hessey, S.: WaaZam!: supporting creative play at a distance in customized video environments, *Proc. of CHI2014*, pp.1197-1206 (2014)
- 4) 吉田有花, 宮下芳明: 身体動作の重畳表示による動画上での一体感共有, *インタラクション 2012 論文集*, pp.527-532(2012)
- 5) 山内雅史, 篠本亮, 西脇絵里子, 小野澤理沙, 北原鉄朗: Kinectとワイヤレスマウスを併用したダンス練習支援システムの試用, *エンターテインメントコンピューティング 2013*.
- 6) Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on PAMI*, Vol. 26, No. 9, pp.1124-1137(2004)
- 7) Maejima, A., Wemler, S., Machida, T., Takebayashi, M., Morishima, S.: *The IEICE Transaction on Information and System*, Vol.E91-D (4), pp.1135-1148 (2008)
- 8) Farneback, G.: Two-Frame Motion Estimation Based on Polynomial Expansion, *Image Analysis*, pp.363-370 (2003)
- 9) Kurihara, T., Okabe, M., Onai, R.: DDMixer2.5D:drag and drop to mix 2.5D video objects, *UIST2013 Adjunct*, pp.69-70 (2013)
- 10) Corduneanu, A., Bishop, C.: Variational Bayesian Model Selection for Mixture Distributions, *AI Statics2001*, pp.1-8 (2001)
- 11) Dalal, N., Triggs, B.: Histogram of oriented gradients for human detection. *CVPR2005*, pp.8886-8893 (2005)
- 12) 築瀬洋平, 笠原俊一, 新田慧, 伊藤周, 樋口啓太, 暦本純一: StratoJump 成層圏気球飛行全天球映像を用いた超跳躍体験, *WISS 2014* (2014)
- 13) Hirai, T., Nakamura, S., Yumura, T., Morishima, S.: VRMixer: Mixing Video and Real World with Video Segmentation, *Processing of the 11th Advance in Computer Entertainment Technology Conference* (2014)