

# 身体動作の重畳表示による動画上での一体感共有

吉田 有花<sup>†</sup> 宮下 芳明<sup>†,††</sup>

アーティストやDJなどのライブイベントは、パフォーマンスと観客が一体となって作り上げるものである。現在、動画共有サイトでは、様々なライブイベントを非同期に鑑賞することができる。しかし、ライブイベントにおける醍醐味である「観客同士が一緒に体を動かして盛り上がる」ことは、非同期メディアゆえに実現していない。そこで本稿では、動画共有サイトにおける動画コンテンツ上に、抽象化・可視化された身体動作情報を重畳することで、非同期メディアであっても、一体感を共有できるエンタテインメントシステムを提案する。また、2つの学会の懇親会で展示・データ共有を行い、一体感を得られるかどうかの調査と考察を行った。

## Video Sharing System that Overlays Body Movements for the Sense of Unity

ARIKA YOSHIDA<sup>†</sup> and HOMEI MIYASHITA<sup>†,††</sup>

In the live show by artists and DJs, the audience and artists come together to dance. Using video-sharing website, we can enjoy those live shows asynchronously. However it is not possible to share a feeling of unity because of that asynchronous characteristic. Therefore in this paper we propose a system that overlays abstracted physical movement of the user on the video at video-sharing site. We also used the proposed system at convivial parties of two academic conferences as an experiment, and discussed the possibility of proposed method.

### 1. はじめに

近年、YouTube<sup>1)</sup> やニコニコ動画<sup>2)</sup> などの動画共有サービスが人気を博し、様々な動画がアップロードされている。ニコニコ動画では、動画上に表示できるコメントをユーザが自由に投稿でき、テキストコミュニケーションが活発に行われている。そのさまは、非同期メディアであるにも関わらず、同期メディア（チャット）のような一体感を形成している。

動画共有サイトでは、様々な場所で開催されるライブイベントを、好きな場所・好きな時間に鑑賞できる。実際のアーティストやDJなどのライブイベントでは、パフォーマンスと観客の両者が同一空間を共有し、観客はパフォーマンスを応援するため、あるいは自ら楽しむために体を動かしている。また、テレビなどでスポーツ観戦をする際には、選手には直接伝わらないにも関わらず、選手を応援するために体を動かしたり声を発したりしている。ニコニコ動画を見ながらコメントを書きこむのと同様に、映像を鑑賞しながらの身体動作は同

期コミュニケーションのように作用するのではないかと我々は考えた。

アーティストやDJなどのライブイベントにおいて、観客が体を動かしたり声かけをすることを本稿では“ノリ”と呼ぶことにする。これは、一体感を高めるとともに参加の実感を与える醍醐味ともいえるものである。それにも関わらず、これらの情報を非同期メディアで共有するような提案はこれまでなされていない。

本研究では、動画上で現在の自分のノリと、異なる時間、異なる空間における他の観客のノリを重畳表示し、観客同士の一体感を実現することを目的とした。今回は Kinect および Web カメラを用いて、観客のノリ情報の取得・記録・再生といった機能を実装することで、目的とするシステム“ノリ乗り♪”の開発を行った。

また、異なる2つの学会の懇親会で展示・データ共有を行い、一体感を得られるかどうかの調査と考察を行った。

### 2. 背景

#### 2.1 既存サービスとしてのニコニコ動画

ニコニコ動画は、ユーザが動画再生中に入力したコメントを動画上に重畳する非同期コミュニケーション

<sup>†</sup> 明治大学理工学部情報科学科

Department of Computer Science, Meiji University

<sup>††</sup> 独立行政法人科学技術振興機構, CREST

JST, CREST

が可能な動画共有サービスである。コメントの投稿に時間差があっても、動画内の時間軸においては常に書き込まれた時と同じタイミングで表示される。これによってユーザは、時間・空間を超越した擬似的な時間共有を体感することができる。ニコニコ動画が2006年に公開されてからというもの、その発展は目覚ましい。現在では、一般会員登録者数が約2250万人、有料会員は約130万人、モバイル会員は623万人(2011年7月12日現在)<sup>3)</sup>にも上っている。また、2011年4月に英語版の「Niconico β」が公開されている。ニコニコ動画が公開される以前から、ユーザが自由に動画をアップロード・共有できるシステムは存在していた。しかし、動画上にコメントを同期させて重畳表示するだけで、新たなメディアコミュニケーションが生まれた。不特定多数が動画に関わることにより、大量のコメントを同じタイミングで書きこんで表示する「弹幕」などの新たな文化も生まれた。「踊ってみた」「歌ってみた」といったUGC (user-generated content) も生まれ、多数の動画がアップロードされている。

## 2.2 非言語行動

人間は無意識のうちに非言語行動を行なっていることがある。非言語行動とは、言語を用いない行動のことで、外見的特徴・ジェスチャと動作・表情と視線行動・音声行動・空間・接触到に分類される<sup>4)</sup>。

人間は昔から、言語以外の行動でもコミュニケーションを行なっている。Birdwhistellは人の間で交わされるメッセージの65%は、非言語が担っているとしている<sup>5)</sup>。動物のコミュニケーションは、養育行動・威嚇・性的ディスプレイ・闘争など、ほとんどが非言語のコミュニケーションで行われている<sup>6)</sup>。

しかし、今日におけるインターネット上でのコミュニケーションは、ほとんどが言語を用いて行われている。身体行動などの非言語を用いてコミュニケーションが行われていることは少ない。

本研究では、非言語行動のうち、ジェスチャと動作に注目し、動画上で身体動作を用いて表現・コミュニケーションを行う。これにより、今までできなかった非言語行動の共有が可能になる。

## 2.3 本研究での“ノリ”の定義

本研究における“ノリ”とは、動画の前に立つ観客による身体動作全てを包括することとする。例えば、「上下運動・手を振る・歓声・踊る・拍手」などのような動作である。これらの動きを、“ノリ”と定義する。これらのノリを可視化した情報を同じ動画の時間軸上に乗せ、非同期に共有する。

## 2.4 本研究での“一体感”の定義

システムを利用しているユーザと、コンテンツを共有している他の複数ユーザのノリを重畳表示し、時間・空間的に離れた複数のユーザとあたかもその場に一緒にいるかのように感じることができるインタラクションが存在するシステム中において、臨場感やコミュニケーションから受けることができる感情を“一体感”と定義する。

## 3. システム

提案システムでは、観客のノリ情報の取得・処理にKinectおよびWebカメラを利用している。以下に概要を述べる。

### 3.1 Kinectを利用した観客のノリの取得

動画の観客はKinectの前に立ち、人体の骨格情報を図1のように取得する。Kinectで取得できる情報は、頭・手・体・足・関節の位置情報であり、各点を線でつないで動画上に描画する。この観客のノリ情報は、観客が見ている動画の時間軸と同じ時間軸で、観客のノリ情報として取得する。

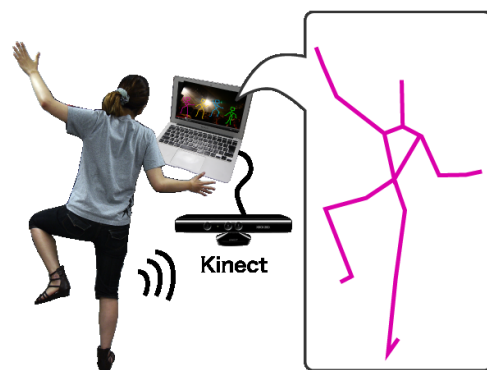


図1 Kinectによる人体の骨格情報の取得

### 3.2 Webカメラを利用した観客のノリの取得

Webカメラの前にいる観客の顔や手を認識し、それらが動いた情報を図2のように取得する。それらの位置情報を動画上に描画する。この観客のノリ情報は、観客が見ている動画の時間軸と同じ時間軸で、観客のノリ情報として取得する。

### 3.3 観客のノリ情報の記録

上記で取得した観客のノリ情報を、観客が視聴している動画時間軸で記録する。まず、起動時にすでに記録されているデータをすべて読み込み、初期化を行う。動画の前に観客がいる場合に、ノリ情報を各フレーム毎に記録していく(図3)。



図 2 Web カメラによる身体動作の取得

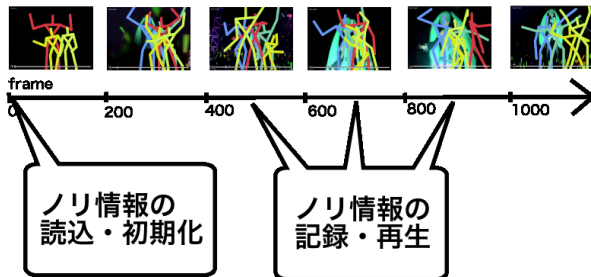


図 3 ノリ情報の初期化・記録・再生の流れ

### 3.3.1 記録データ

記録データについては、動画を1秒10フレームに振り分け、1フレーム毎に、ユーザID・フレーム番号・骨格情報の各点をサーバにアップロードし、共有する。特定フレーム、特定ユーザ毎にデータを保存する。保存したデータは、図4のように表示できる。横軸がxであり、縦軸がyである。左上を始点とし、今回は横640px・縦480pxとした。

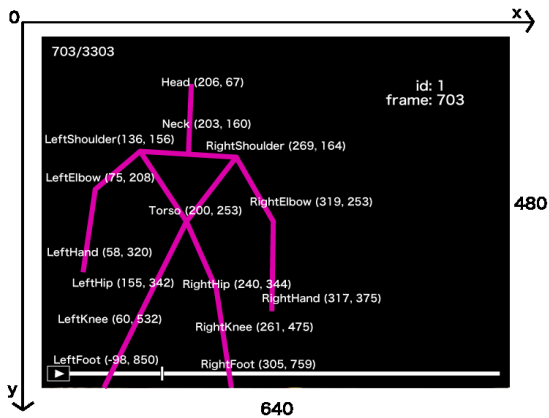


図 4 記録データの一例

### 3.4 観客のノリ情報の再生

再生は観客が動画を見ている時に、現在の観客の動き情報と、異なる時刻で記録された別の観客のノリ情報を利用するデバイスによって、図5・図6のように再生する。観客のノリ情報はいくつも重なりあうことができる。

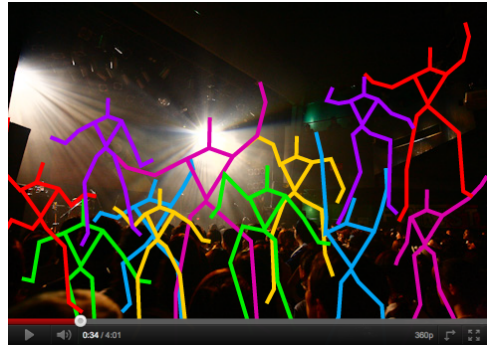


図 5 Kinectによる観客のノリ情報の共有イメージ

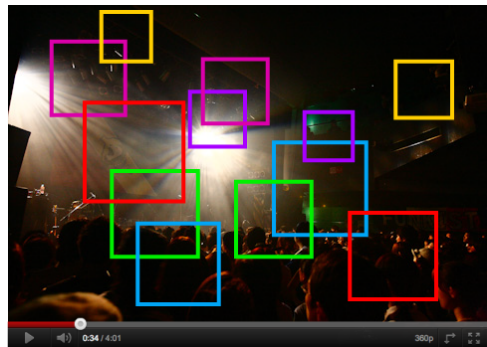


図 6 Web カメラによる観客のノリ情報の共有イメージ

## 4. EC2011・WISS2011における展示(実験)と今後の課題

2011年10月7日～9日に行われたエンタテインメントコンピューティング2011(以下EC2011)<sup>8)</sup>と、約2ヶ月後の2011年12月1日～3日に行われた第19回インタラクティブシステムとソフトウェアに関するワークショップ(以下WISS2011)<sup>9)</sup>の両懇親会で本研究のKinectを用いたシステムを展示し、ノリを取得した。両学会の懇親会の”ノリ”を動画上に重畳することにより、異なる時間・空間でインタラクションができるかどうかの実験を行った。

図7・図8のように、テーブルの上にモニタを設置し、そこに特定の動画をループ再生するブースを設置した。今回利用した動画は、クラブイベントにいるパフォーマと観客が盛り上がっている約5分の動画であ



る。設置したブースの動画の前にいる人達の情報を重ねつけた。

まず、EC2011の懇親会で、ノリ情報を取得した(図9)。WISS2011では、着色によってEC2011のノリ情報との区別を行って重畳表示した(図10)。両学会合計で短時間のものも含めたの延べID数486のノリを取得し、ニコニコ動画における弾幕の様な表示をすることができた。



図7 EC2011での展示の様子



図8 WISS2011での展示の様子

#### 4.1 アンケート

WISS2011において、アンケートを行った。「EC2011の懇親会にいた人たちとつながったかのような“一体感”を感じることができましたか」というアンケートを行ったところ、16人中12人が、一体感を感じることはできませんでした、と回答した。また、一体感を感じるには何が足りないと感じたのか、ということも聞いてみたところ、「共同作業がしたい」「リアスティックな形が欲しい」「ECの人たちとの相互交流ができていない」「人間が多すぎて個性が感じられない」「骨格情報とムービーとの違和感がある」「ヘッドホン・真っ暗な部屋が欲しい」「自分が動きたくなくなるようなフィードバックが欲しい」などの意見・感想を得た。

また提案システムに対して、「楽しい」「酒が入ると最高に楽しい」「映像エフェクトと音さえあれば十分ノリノリ」などの肯定的な感想も得られた。



図9 EC2011で取得したノリ

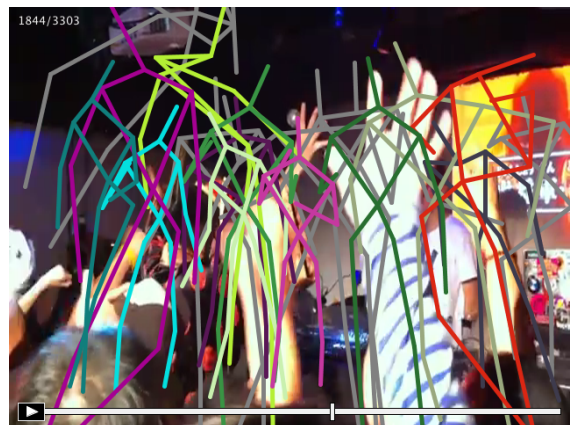


図10 WISS2011で取得したノリを重ねた様子

#### 4.2 新たなエンタテインメント体験

今回の実験で、興味深い知見も得られた。我々が最初に想定していた動作とは異なる動作を、ユーザが始めたことを観察できた。しばらく動画の前にいたユーザ2名が、Exile - Choo Choo Trainのダンスのような動きを始めた。通常は5~6名で縦に並び、順番に回る動作をするものであるが、提案システムであれば、少ない人数でこの動きを表現することが可能である。その他にも、重畳された群衆の中でも目立とうと前に出る人がいたり、激しい動きをする人がいた。また、「組体操のような表現もできそうだ」という意見も得られた。これらは今回の実験で得られた、メディアの新たな楽しみ方の可能性を予見しうる事例であると我々は考える。

### 4.3 明らかになった課題

今回、一体感を感じられなかった理由は、画面内の自身の骨格が認識困難であったこと、過去や現在のノリとのインタラクションができていなかったこと、映像エフェクトが実装されていなかったこと、などである。これらの問題を解消するようにシステムを改良していく必要がある。

## 5. 関連システム

動画上にコンテンツを重畳する研究として、Harrison らの CollaboraTV<sup>10)</sup> は、動画に対するユーザのコメントをその動画上に用意したアバタのリアクションを表現することで、鑑賞者のコミュニケーションの活性化を試みている。また、棟方らの ExiTube<sup>11)</sup> は、他者の存在を感じながら動画を鑑賞することができる動画鑑賞システムである。他者を模したアバタを動画とともに表示することで、アバタを通して他者を感じることができる。ユーザの興奮は、手掌の精神性発汗 (skin conductance) を検知する。

映像メディアに対する時間的および空間的インタラクション手法の提案<sup>12)</sup>として、高嶋らは (TbVP: Time-based Visual Presentation) を利用して、映像の時間的・空間的な同期を行っている。また、時間軸という概念を飛び越えて、同一空間上に現在の自分が過去にその場にいたユーザとインタラクションを行うことができるシステムとして、竹内らは PRIMA<sup>13)</sup>を開発している。これは、4台の Kinect を利用して体の3次元ユーザデータを取得し、過去のユーザとのインタラクションを行うことにより、Parallel Reality (並行現実感) を実装している。他にも、ご近所知るえっと<sup>14)</sup>が同種の感覚を生成するメディアであると考えられる。同じ建物に住んでいる人たちに向けて、エレベーターを待つ人の影を取得し、これをスクリーンに投影し、影を自動でアニメーションさせるシステムである。

twitter アイコンやコメントやバーチャルグッズなどを動画の下に表示し、配信者に投げ銭ができるシステムとして、HikeTheater<sup>15)</sup>がある。他にも、自分が今聴いている曲を共有し、聴いている人たちのアバタを表示するシステムとして、turntable.fm<sup>16)</sup>がある。さらに、Web上のあらゆるページの上に twitter アイコンとチャットを表示し、評価・共有するシステムとして、CHeeRZ<sup>17)</sup>がある。

## 6. おわりに

本研究では、ニコニコ動画のような、コメントを重

畳するシステムではなく、観客のノリである、身体動作を抽象化・可視化し、動画に重畳・共有することにより、時間・空間を超えて一体感を得られるシステムを開発した。しかし、異なる学会での実験を行ったところ、現行システムでは、一体感を感じることはできないという回答を得られた。今後システムの改良が必要であることが明らかになった。

しかし、提案システムである“ノリ乗り♪”は、コメントだけではできなかった新たなコミュニケーションを生み出す可能性があり、世界中の人々が国境や言語の壁を超えて一緒にノリを楽しむことができる可能性のあるシステムであると我々は考えている。

ニコニコ動画は言語を利用している動画共有サービスであるので、動画に重畳されている言語の意味が分からない場合、意思疎通ができず、あまり楽しむことができないであろう。言語によるコミュニケーションは、機械翻訳を利用すればある程度は可能になるが、うまく翻訳できないことが多い。しかし、提案システムは身体動作の新しい抽象化・可視化により、文化によらないゆるやかなコミュニケーションを生み出す可能性があると我々は考えている。また、非言語情報は、文化によって表現の方法が異なる<sup>18)</sup>が、言語情報ほど理解できないものではないと思われる。むしろ異文化間の差異がわかることによって、異文化交流が進んでいくのではないかと考えられる。

今回の実験で観察できたことは、当初我々が予想していなかった動作をユーザが始めたことである。そこから予想されることは、提案システムを利用するユーザによって新たなメディアの創造が生まれ続けるのではないかと、ということである。ニコニコ動画は、ただ動画にコメントを重畳し、右から左に流すだけのシステムであるにも関わらず、様々な楽しみ方が生まれ、日本中に大きな影響を与えている。それと同じく、提案システムも、ユーザ同士によって様々な楽しみ方が生まれるのではないかと期待している。例えば、ラジオ体操、パラパラ、エアロビクスなどを複数の人々が非同期に踊るといったことが考えられる。表示の部分では、ニコニコ動画における弹幕のような身体動作を大量に重畳表示することが考えられる。Kinect や Web カメラを通常とは違う方向に設置して横立ちや逆立ちすることも考えられる。インタラクティブな例としては、ユーザ同士が抱擁をする、じゃんけんをする、ゲームのように怪獣と戦う、なども想定できる。

将来的には動画だけではなく、Web ページや twitter や電子書籍など、様々なコンテンツにも応用できるのではないかと我々は考えている。これにより、Web 上

の様々な場所で、身体動作を重畳でき、異文化コミュニケーション・新たなメディアの誕生が起こるのではないかと期待している。

## 参 考 文 献

- 1) YouTube, <http://www.youtube.jp/>.
- 2) ニコニコ動画, <http://www.nicovideo.jp/>.
- 3) ニコ動のライブハウス「ニコファーレ」完成, <http://www.yomiuri.co.jp/net/news/20110712-OYT8T00813.htm>
- 4) Virginia P. Richmond, James C. McCroskey, 山下 耕二. 非言語行動の心理学—対人関係とコミュニケーション理解のために, p.12-13 (2006).
- 5) Birdwhistell, R.L : Systeme de ka Mode,Seuil (1967).
- 6) 黒川 隆夫. ノンバーバルインタフェース, オーム社, pp.8-11 (1994).
- 7) 石井 裕. CSCW とグループウェア, オーム社, pp.105-112 (1994).
- 8) EC2011, <http://ec2011.entcomp.org/>.
- 9) WISS2011, <http://www.wiss.org/WISS2011/>.
- 10) Harrison, C. and Amento, B. CollaboraTV: Using asynchronous communication to make TV social again, Adjunct Proceedings of EuroITV2007,pp.218-222 (2007).
- 11) 棟方 渚, 代藏 巧, 小野 哲雄, 松原 仁. ExiTube : 他者の存在を感じられる動画観賞システム, CEDEC2011 (2011).  
<http://cedil.cesa.or.jp/session/detail/693>.
- 12) 高嶋 章雄, 山本 恭裕, 中小路 久美代. 映像メディアに対する時間的および空間的インタラクション手法の提案, 情報処理学会研究報告. HI, ヒューマンインタフェース研究会報告, vol.2003,No.9,pp.63-70 (2003).
- 13) 竹内 俊貴, 中島 統太郎, 西村 邦裕, 谷川 智洋, 廣瀬 通孝. PRIMA—異なる時間軸上のユーザとのインタラクションを実現するシステム—, 第20回エンタテインメントコンピューティング研究発表会, pp.1-6 (2011).
- 14) 中森 玲奈, 青木 貴司, 椎尾 一郎. ご近所知るえっと—身近な他人との緩やかなコミュニケーション支援, WISS2010 論文集, pp.35-40 (2010).
- 15) HikeTheater, <http://theater.hiking.to/>.
- 16) turntable.fm, <http://turntable.fm/>.
- 17) CHeeRZ, <http://www.cheerz.st/>.
- 18) 石井 敏, 岡部 郎一, 久米 昭元. 異文化コミュニケーション [改訂版], 有斐閣 (1994).