

Control Manually: 視聴者協力型ライブ演出システムによるコミュニケーションと演出効果の拡張

米澤 拓郎^{1,a)} 徳田 英幸^{1,b)}

概要: インターネット生放送の利点の一つに、配信者と視聴者間の密なコミュニケーションが可能である点が挙げられるが、現状ではコメントを通じた双方向性に限られている。生放送コンテンツとして楽器の演奏が挙げられるが、演奏中は集中しているため、コメントによる視聴者とのコミュニケーションは希薄となる。本研究では演奏配信に適した新たなコミュニケーション手段として、視聴者が演奏者の配信環境を制御し、演出を行うことでコミュニケーションを可能とする視聴者協力型ライブ演出システムを提案する。4週間に及ぶ配信実験の結果、同システムを用いることで演奏者と視聴者間、および視聴者間での非言語コミュニケーションが創発され、生放送参加者の一体感の拡大につながるとともに、ズームインやパン、カメラの切り替えなどのカメラワークによって演出効果が向上したことを示す。また実験結果に基づき、将来のインタラクティブ生放送支援システムのデザインの指針を示す。

1. はじめに

近年、インターネット及び情報技術の発展と普及に伴い、個人による情報発信が容易に可能となった。特に、これまで特定の放送事業者や高価な機器を有する専門家しか可能でなかった映像と音楽を含む生放送配信が、インターネットを利用することで個人の環境でも容易に可能となり、多くの個人が実際に生放送を行っている。生放送における重要なコンテンツとして、楽器の演奏配信が挙げられる。アマチュアミュージシャンによる演奏配信に関しては音楽著作権の問題が兼ねてから存在してきたが、Ustream [7]、YouTube [11] やニコニコ生放送 [21] では各著作権管理会社と包括契約を締結する [1] など、アマチュアが気軽に演奏配信が可能となるよう、その整備も進められてきている。

アマチュア演奏配信者は演奏技量こそプロに劣るものの、視聴者に演奏曲のリクエストを行ってもらうなどのコミュニケーションを通じ、演奏者と視聴者がともにリアルタイムに配信を作り上げていくことでエンターテインメント性の高い配信を行っている。またアマチュア演奏配信者の工夫の一つとして、インターネット上に投稿された音楽映像を再生しながら、それに自分の演奏を重ね合わせながら演奏配信するケース（以下、映像再生型セッション演奏配信）も見られるようになった（図1参照）。これは、例えばUstreamであればYouTubeの動画を再生しながら生放送の配信を許可・可能とさせたり、ニコニコ生放送であればニコニコ動画内に投稿された動画を再生しながらの配信を許可・可能とさせていることから、配信者が自分の演奏

を工夫を凝らして配信しようとする試みからなされてきている。近年、あるユーザが作成しインターネット上に投稿した映像や音楽に他ユーザが加工を加えて作品を発展させるマッシュアップ型コンテンツの作成が広がりを見せている。映像再生型セッション演奏配信は、このマッシュアップ型コンテンツのリアルタイム作成・配信であると言える。このようにアマチュア演奏者は、演奏者自身はもちろん、視聴者がより楽しめる配信となるよう様々な工夫を試みている。しかしその一方で、演奏者と視聴者のコミュニケーションはコメントを通じたものに限られており、演奏中には配信者がコメントを読んだりそれに応答することは難しいため、その双方向性は薄れてしまう。

本研究では、個人の生放送における多様な演出を、視聴者がインターネットを通じて協力・向上させるとともに、その演出を通じて配信者・視聴者間でコミュニケーションを行える、**視聴者協力型ライブ演出システム**を構築する。プロミュージシャンの演奏配信は、テレビ放送におけるライブ生中継などと同様、カメラマンや照明などの舞台演出家との協力のもと、臨場感のある映像を実現して行われる。例えば、1台のカメラによる映像でも、図2に示すように、ズームイン、パンなど多くのカメラワークによって映像の演出がなされている。一方、現在のアマチュアによる演奏配信は、主に楽器を演奏する様子を1台の固定Webカメラで撮影し、その映像と演奏を同時に配信する形式が主であり、プロミュージシャンの演奏配信と比較すると演出上の乏しさが目立つ。本研究では、視聴者が配信者の環境に存在するカメラや照明などを遠隔制御可能とさせることで、演出の向上を実現するだけでなく、演奏の際の演奏者と視聴者のコメントだけに頼らない新たなコミュニケーションの実現を目的とする。例えば、演奏者のリアルタイ

¹ 慶應義塾大学大学院政策・メディア研究科
Endou 5322, Fujisawa, Kanagawa 252-0882, Japan

^{a)} takuro@ht.sfc.keio.ac.jp

^{b)} hxt@ht.sfc.keio.ac.jp

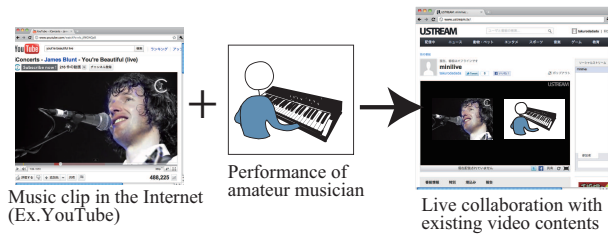


図1 映像再生型セッション演奏配信の概要

Fig. 1 Concept of Session Live with Video Playing

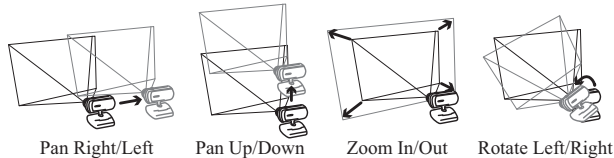


図2 カメラワークの種類

Fig. 2 Types of Camera Work

ムな演奏変化に視聴者が配信環境を追従させ演出効果を高めたり、逆に視聴者が主導となって配信環境を制御することで演奏者に対して演奏の変化を要求することが可能となると考えられる。なお既に製品としてネットワーク対応のカメラは多数販売されており、遠隔からズームイン、パンなどのカメラ制御は可能となっている。しかし、その使用用途は防犯用モニタリングであったり、交通状況のリアルタイム観測などが主であった。本研究では汎用的で安価なWebカメラを用い、遠隔カメラ制御をライブ生配信に適用させるとともに、その結果創出される演出効果、および配信者と視聴者のコミュニケーションの変化について焦点をあてる。本研究のもう一つの重要な動機は、手動機器制御に着目することである。これまでユビキタスコンピューティング技術ではコンテキストに応じた機器の自動制御手法が多く提案されてきた。また講義 [2], [17] やミーティング [8], [16] でのカメラの自動制御も多く研究報告されている。我々の目的は、自動制御ではなく、手動制御が持つ利点・新たなコミュニケーションの創発を顕在化させ、示すことである。本稿では、構築したシステムを利用した生放送配信実験を4週間に渡り行い、その効果・可用性を評価した。結果、視聴者協力型ライブ演出システムが演奏配信において演奏者と視聴者の新たなコミュニケーションを実現可能であるとともに、十分な視覚的演出効果を果たし、ライブ配信の臨場感を向上させることが分かった。

本稿の意義は、以下の3点である。

- 演奏者と視聴者が協力し、視覚的演出効果を高めることが可能なライブ演出システムを構築したこと
- 視聴者協力型ライブ演出システムが、演奏者と視聴者間、及び視聴者間における非言語コミュニケーションの創発に寄与することを示したこと
- 将来のインタラクティブ生放送システムのためのデザインの指針を示すこと

2. 視聴者協力型ライブ演出システム

2.1 概要

本研究が提案する視聴者協力型ライブ演出システムの概要を、図3に示す。配信者の環境では、配信者の演奏を撮

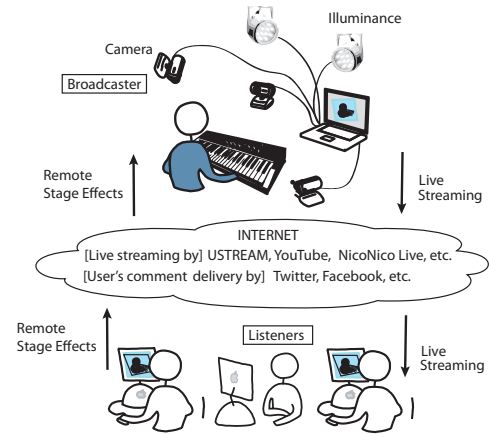


図3 視聴者協力型ライブ演出システムの概要

Fig. 3 Concept of Listener-Cooperative Live Production System

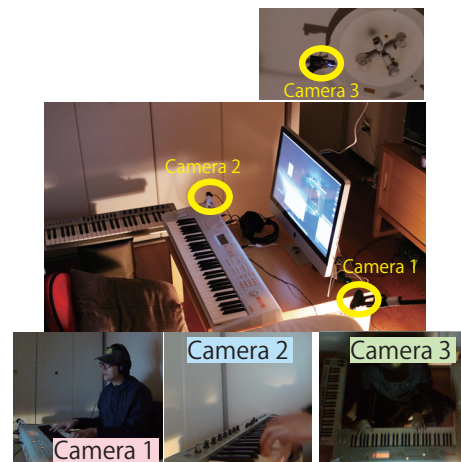


図4 実装システム：ハードウェア構成

影する複数のカメラ、雰囲気を変化させる照明などが設置されており、視聴者は配信者のリアルタイムな演奏を視聴し、演奏に応じて配信者側の環境を制御することができる。演奏者はUstreamやニコニコ生放送などの生放送配信Webサイトを通じて配信を行い、視聴者は同Webサイト上でのコメント投稿機能、及びそのWebサイトと連携したソーシャルメディア(TwitterやFacebook)などを通じてライブ環境制御命令を伝達する。

2.2 ハードウェア構成

図4に実装システムのハードウェア構成を示す。配信者の環境には3台のWebカメラ(Camera 1-3)、異なる音色を出す2台のキーボード(キーボード1,2)、及びそれらが接続されたコンピュータが存在する。Webカメラは、主にキーボード1と配信者を撮影できる位置(Camera 1)、キーボード2を弾く手を撮影できる位置(Camera 2)、キーボード1とキーボード2を含む上からの俯瞰映像が撮影できる位置(Camera 3)にそれぞれ設置した。

2.3 ソフトウェア構成

本研究では、配信者が配信に利用するコンピュータ上に視聴者のコメントを解析するソフトウェア、解析されたコメントに応じてカメラ制御を行うソフトウェアを実装し

表 1 カメラ制御コマンド一覧
Table 1 List of Camera Control Command

コマンド	機能
cam1	カメラ 1 へ切り替え
cam2	カメラ 2 へ切り替え
cam3	カメラ 3 へ切り替え
zoomin / zoomout	ズームイン/ズームアウト
zoomstop	ズーム処理の停止
panright / panleft	パンライト/パンレフト
panup / pandown	パンアップ/パンダウン
panstop	パン処理の停止
rotateright / rotateleft	右回転/左回転
rotatestop	回転処理の停止

た、コメントは各 Web サイト及びソーシャルメディアから XMLSocket などを利用し、取得する。よって視聴者側にソフトウェアをインストールする必要はない。

視聴者は、表 1 に記すコマンドをライブ配信時に入力することで、各種カメラの制御が可能となる^{*1}。カメラの制御の種類としては、図 2 に示した代表的なカメラワークをデジタル画像処理で実現した。カメラ切り替えは入力ソースの切り替えで行い、ズームイン処理、パン処理、回転処理はいずれもキャプチャした映像を毎フレーム少しずつ変化させて行う。また視聴者は、各種カメラ制御コマンドを“:”でつなげることで複数のカメラの制御を同時に行うことが可能となっている。例えば“cam3:zoomin”と入力することで、カメラ 3 に切り替えると同時にズームイン処理が行える。図 5 に、カメラ制御コマンドとその結果処理された映像の対応シーケンスを示す。同一カメラを利用している場合、ズームイン、パン、回転は互いに独立して継続処理して行うようにしたため、コマンドを組み合わせることでズーム処理しながらパン処理、パン処理しながら回転処理、などのカメラワークを実現可能とした。これは、複数の視聴者の異なるカメラ制御の意図をできるだけ多く反映するためでもある。なお、カメラソースの切り替え処理が発生した際は、ズーム・パン・回転処理は初期化される仕様とした。コメント解析およびカメラ切り替え・デジタル処理部分は Java と OpenCV を利用して実装した。

操作要求の衝突解決は本研究の対象とはしないが、最後に送られたコマンドが常に優先され実行されるという単純なプライオリティポリシーを定めた。カメラ切り替えコマンドは排他的に実行されるが、ズーム処理、パン処理、回転処理はそれぞれ同時に処理可能とした。例えば、zoomin と rotateright というコマンドが同時に送られた際は、その効果がオーバーラップして実行される。これらの効果はカメラ切り替え時にキャンセルされる。

3. 評価

3.1 実験内容

視聴者協力型ライブ演出システムの効果、使いやすさを評価するため、ニコニコ生放送を利用した演奏配信実験を行った。配信者は鍵盤楽器演奏経験がある著者であり、実験は週 3 日程度、映像再生型セッション配信を中心

^{*1} ニコニコ生放送の同一コメント連続投規制を回避するためエイリアスコマンドも複数用意した (例: Cam1 と cam1)。

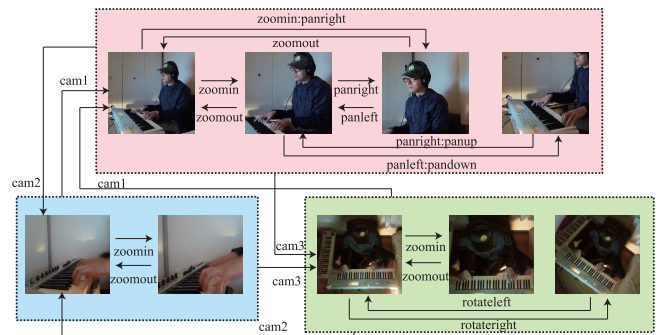


図 5 カメラ制御コマンドと実際撮影される映像の対応シーケンス
Fig. 5 Sequence of Captured Video According to Camera Control Command

に 4 週間に渡り行い (表 2 参照)、その放送内で見受けられたシステムの演出上の効果や視聴者の反応、演奏者と視聴者間のコミュニケーションの変化を調査した。また実験終了後、カメラ制御を行ったことのある視聴者にアンケートを行い、システムの使いやすさなどを調査した。以上の実験から、ユーザビリティの指標 [14] を表す学習しやすさ (Learnability)、効率性 (Efficiency)、記憶しやすさ (Memorability)、エラーがあるかまたは起こした際の回復 (Error Handling)、ユーザの主観的満足度 (User Satisfaction) の 5 つに関して評価する。

3.2 配信者と視聴者間のコミュニケーションの創発

本節では、配信実験の際に見受けられたカメラワークの具体例、及び配信者と視聴者間のコミュニケーションの変化について、実際の配信実験の様子を踏まえて検証する。実験例として、実験期間内に最も多くの視聴者数、コメント数を得られた 2011 年 4 月 22 日の配信を取り上げる。多数の視聴者数を得られた理由は、本配信がニコニコ生放送による公式の放送内で取り上げられたことが挙げられる。よって、半数以上の視聴者ははじめて本放送を視聴することとなり、多くの自然なフィードバックが得られた。放送内容は主にニコニコ動画内に投稿された動画と配信者の演奏を組み合わせた映像再生型セッション演奏配信であった。ライブ生放送は 90 分行われ、合計視聴者数は 658 人、総コメント数は 6323 であった。視聴者は配信者とは直接の面識のない一般視聴者である。総コメント数のうち、カメラ制御コマンドのコメント数は 560、うちカメラソース切り替えが 168 回、パン処理が 58 回、回転処理が 118 回、ズーム処理が 374 回行われた。カメラ処理の総回数 (718) がコメント総数 (560) を上回っている理由は、視聴者がコマンドを連結する “:” によって複数のカメラ処理を同一コメント内で表現する場合が多数あったためである。

実際の配信映像 (視聴者のコメントを含む) の 5 秒毎のフレームを図 6 に示す。図 6 には、異なる 2 つのシーン A、B それぞれのフレームにタイムラインにそって番号がふってあり、また視聴者のカメラ制御コマンドが入力されたタイミングに、吹き出し表示をしている。図 6 中からは視聴者が演奏者の演奏に追従してカメラを切り替えている様子がみとれる。例えばシーン A-7 で演奏者がキーボード 1 からキーボード 2 へと利用を変えようとするタイミング

表 2 配信実験の一覧

Table 2 List of experimental dates and the number of listeners, comments and control comments.

Dates	Time length (min)	# of L	# of AC	# of CC
Mar. 25	60	222	596	156
Mar. 26	60	141	336	79
Mar. 30	90	266	651	177
Apr. 1	90	210	403	138
Apr. 3	30	95	110	49
Apr. 4	90	499	1210	152
Apr. 5	60	256	674	263
Apr. 6	90	343	763	216
Apr. 9	90	328	836	181
Apr. 13	90	327	1226	194
Apr. 14	60	196	672	119
Apr. 16	60	192	346	152
Apr. 18	90	193	494	148
Apr. 20	60	181	234	89
Apr. 21	90	250	1148	212
Apr. 22	90	658	6323	560
Apr. 24	90	352	1436	311

of L = number of listeners, # of AC = number of all comments,

of CC = number of comments for camera control

で、キーボード 2 を主に撮影するカメラ 2 へと切り替えるため”cam2” コマンドを視聴者が入力している。また B-24 で演奏が終了したタイミングで、演奏後の配信者の顔を捉えるカメラ 1 へと切り替えを行っている。このように、視聴者は演奏者の意図 (キーボード 1 を弾く、キーボード 2 を弾く、キーボード 1/2 を同時に弾く) をくみ取り、それに適したカメラ制御を行う様子が多く確認された。また、よく見られた視聴者主導の演出方法として、変化のない映像が続いた場合、カメラを切り替えたり、ズームイン/ズームアウト制御を行うことが挙げられる。これらの演出は、映像再生型セッションでなくとも同様に見受けられた。

視聴者が配信者の演奏に合わせてカメラワークを切り替えるだけでなく、逆に視聴者が主導でカメラを切り替え、演奏者がそれに対応して演奏する場合もあった。図 7 のシーン C-1 では演奏者はキーボード 1 を弾いているが、シーン C-2 ではキーボード 2 を撮影するよう視聴者がカメラを切り替えている。その後、シーン C-3 では、演奏者はキーボード 2 の映像が撮影されていることからキーボード 2 を弾くように演奏を変化させている様子がわかる。一方、同様に視聴者がカメラを切り替えても演奏者が追従して演奏を変化させない場合は、再度視聴者が演奏者に合わせたカメラに切り替え直すなどの場合も見受けられた。このように、カメラ切り替え・演出機能を通じて演奏者と視聴者は互いの意図を伝達し、これまでにないコミュニケーションの創発が確認された。更に、ある視聴者が切り替えたカメラ映像に別の視聴者がズームインやズームアウト処理を加えるなど、視聴者間でよりよいカメラワークを行おうとするケースも非常に多くみられた。このように、放送に参加する多くのユーザ間で非言語コミュニケーションが発生したことが認められた。映像再生型セッション演奏配信において、プロミュージシャンによる演奏映像のカメラワークに合わせて視聴者がカメラワークを変化させる例も多く

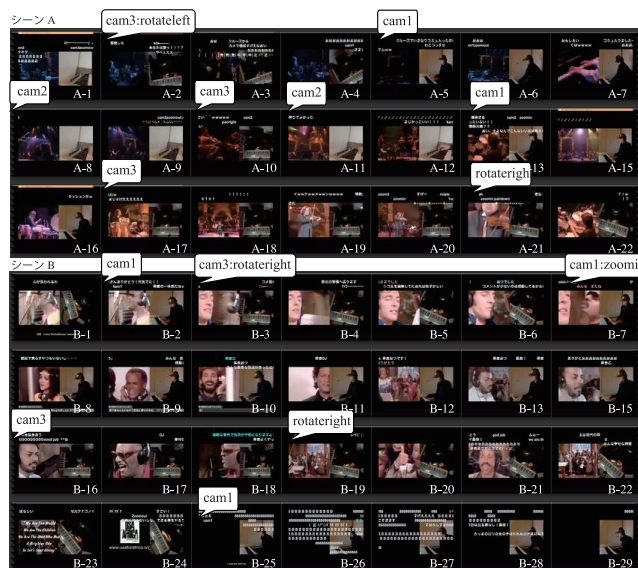


図 6 実験内容 (2011 年 4 月 22 日放送分)

Fig. 6 Live Performance Streaming on April 22, 2011



図 7 視聴者によるカメラワークに合わせて演奏を変化させる例

Fig. 7 Timeline of the live broadcast showing non-verbal communication between the broadcaster and listeners.

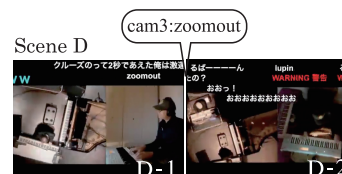


図 8 映像に合わせたカメラワークの例

Fig. 8 Timeline of the live broadcasting showing camera control according to a professional music clip.

見受けられた。図 8 のシーン D-1 では、プロミュージシャンの映像 (左) が頭上からの俯瞰ズームアウト映像に切り替わると、同様の演出を行うよう視聴者がカメラ 3 に切り替え、ズームアウト処理を行っている。

3.3 演出効果に関する評価

次に、実際の放送において視聴者協力型ライブ演出システムが実現した演出に対してユーザが受けた印象について述べる。表 3 に、放送時にユーザが入力したコメントのうち、提案システムに関する印象と思われるものを原文のまま抜粋した。カメラワークの演出は、多くの視聴者にとって好印象に受け入れられたことが分かる。視聴者自身が演出に協力する形式に関しても、放送を共に作り上げていることによる一体感が創出される、という肯定的な印象を持って受け入れられた。

また、カメラ制御機能を構築する前 (1 台の固定カメラ

表3 配信時に得た視聴者協力型ライブ演出システムに関するコメントの例
Table 3 Comments about Listener-Cooperative Live Production System While Live Streaming

内容の種類	コメント
カメラワークに関して	プロが撮影してるの？/映像がすげー迫力/カメラ機能すげー/pan 観がすごい
	神カメラワークww/カメラ職人すげえええ/カメラワークいいなあ 十分楽しめる
	カメラ操作うま過ぎワロタww/リスナーにプロのカメラマンいるなこりゃ
視聴者協力型に関して	音楽の新しい楽しみ方/なんだこれ楽しすぎんぞw/驚異の一体感だねw
	放送してる人も見てる人も幸せになれる放送/すげー視聴者参加型w
	だっておかしいだろ。コメントも参加できて主も演奏してるんだぜ？なんだよこの放送

による演奏配信)と後での演出の向上に関し、カメラ制御機能を追加する前から放送を視聴していたユーザを対象として、WEBを通じたアンケートを行った。質問内容とその結果を表3.3に示す。アンケートでは表3.3のスケールでの評価に加え、なぜそう思ったか、その理由を自由記述で回答してもらった。回答者は計44名で、内訳は10代もしくはそれ以下が4名、20代が20名、30代が12名、40代が6名、50代が2名であった。理由として挙げられたものとして、「色々な角度から演奏者を見れるから」「曲が転調する際にカメラが切り替わったりした時が、演出効果として良かった」「複数のカメラマンがいるみたいで、その人の存在を感じさせることが、手作りされたライブという感じで良くなった」「本当のライブのように感じたから」など、好意的な理由が多かった。一方で、「多くのカメラ制御が一度にきた場合は多くのカメラ切り替えが一瞬のうちに発生する際は、演出的にはおかしいのかなと思いました(もちろん、それはそれでライブ感があって面白いですが!)」と、演出として失敗している場合のコメントも見受けられた。全体的には、本提案システムにより演出が強く向上したと感じたユーザが多くいることがわかった。また、「導入前より視聴者と配信者の垣根がなくなった」「見る、聞くだけでなく、放送に参加できるようになってそれぞれの楽しみが増えた」と、視聴者と演奏者の距離に関する記述も見受けられ、一体感の向上につながっていることがわかった。

3.4 ユーザビリティ評価

コメントによるカメラ制御手法のユーザビリティについて、実際にカメラ制御を行ったことのある視聴者を対象に、自由記述を含むアンケート調査を行った。アンケートはWebを通じて行ってもらい、37名の回答を得た。回答者の内訳は年齢が10代もしくはそれ以下が8名、20代が15名、30代が7名、40代が5名、50代が2名であった。アンケート項目とその結果を表5に示す。

- 学習しやすさ:「使い方を習得するのは簡単だった」という質問項目に対する平均得点は3.78であり、ある程度学びやすいと評価された。自由記述では、「慣れればある程度使用可」「Pan操作を理解するのに時間がかかった」「コマンドがわかりやすく直感的に操作出来た」「コマンドライン育ちなため苦勞せず操作を覚えれた」「2回目の操作からは普通にできた」などの意見が得られた。複数のユーザにとっては操作に慣れが必要であったことも分かる。

- 効率性:「カメラ制御システムは思うように動作した」という質問項目に対する平均得点は3.65であり、カメラ制御コマンドがある程度効率よく操作可能であったことが見受けられる。自由記述では、「:」によるマルチコマンドにより使いやすさが向上している」というコマンドに対する評価もあったが、一方で「TVのスイッチャーマンのように映っている絵から選ぶのは簡単だが、演奏者の動きを予想して完成形を頭の中に描き操作することは頭を使った」など、コマンドのみでのカメラ制御の難しさも指摘された。
- 記憶しやすさ:「使い方は記憶しやすい」という質問項目に対する平均得点は3.46であった。自由記述には、「英語のコマンドではなく、”上”, ”回転”などの日本語にも対応してもらえると良い」という意見も見受けられた。一方、「一年経っても使い方は忘れないと思う」というユーザも存在し、ユーザによって記憶しやすさは少々のばらつきが見受けられた。
- エラーがあるかまたは起こした際の回復:自由記述において指摘された意図しない制御を行った例として、パンライトとレフト、右回転と左回転の方向を逆にしてしまったという意見があった。カメラの制御方向を逆にしてしまった例は、実際の放送時にも多く見受けられたエラーであった。一方、「自分のコマンドが間違っていたと思ったら、直後に他ユーザが別の制御を行っていただけだった」という感想もあり、複数ユーザの制御により意図がうまく反映できていないケースも見受けられた。また、多かった意見として、制御コマンドを入力してから実際にカメラが切り替わるまでの遅延の存在が指摘された。配信者によるライブ演奏映像は、ニコニコ生放送の映像配信サーバの状況によって、3秒から最大で10秒近い遅延が発生して届けられていた。よって、演奏者の動きを事前に予測してカメラ制御コマンドを行う必要があり、その点が難しかったという意見が見受けられた。一方、タイミングよくカメラ制御を行うためには予測が行う必要がある、という点を逆にゲーム感覚で捉え、その点の面白さを感じる視聴者も存在した。
- ユーザの主観的満足度「カメラ制御システムを利用するのは楽しい」という質問項目に対する平均得点は4.27であり、全体的に高い満足度を得ていることがわかる。自由記述にも、「面白い機能」「楽しい」「盛り上がった」「うまく切り替えられた時の感動とリスナーの反応に嬉しくなる」「演奏者に反応してカメラを切り替

表 4 演出効果に関するアンケート結果
Table 4 Result of Questionnaire about Usability

質問	強くそう 思わない (1)	そう 思わない (2)	普通 (3)	そう 思う (4)	強くそう 思う (5)	平均
本システムにより演出効果が向上した	0	0	4	15	25	4.47

表 5 ユーザビリティに関するアンケート結果
Table 5 Result of Questionnaire about Usability

質問	強くそう 思わない (1)	そう 思わない (2)	普通 (3)	そう 思う (4)	強くそう 思う (5)	平均
使い方を習得するのは簡単だった	0	2	11	17	7	3.78
利用するのは楽しい	0	1	3	10	20	4.27
思うように動作した	0	3	9	19	3	3.65
使い方は記憶しやすい	0	4	16	13	4	3.46

えたり、逆に演奏者が反応するようカメラをあえて切り替えて要求を出すなどのコミュニケーションがとれて楽しい」など、肯定的な意見が目立った。また、「生放送をみんなで作っていく感覚が楽しい」という意見や、「一人がcam3zoomin, もう一人がrotateleftと操作したりして、その気の合い方を楽しむといったことが、生放送ならではの楽しさだった」という意見もあり、複数人で同時にカメラ制御を行うことで発生する楽しさや喜びも多く指摘された。このように、カメラ制御によって生まれる副次的なコミュニケーションが、視聴者の満足度に大きく貢献していることが分かった。

また、自由記述欄に見受けられたユーザのコメントとして、「人数が多くなった際にカメラの奪いあいになって変な争いが起こったら心配」「人数が増えた時に荒らしに対応できるか不安」との懸念も示された。今回の実験ではいたずら目的の操作(例えばカメラの切り替えを激しく行うなど)はなく、カメラ切り替え機能、またカメラ切り替えを行おうとしコメント入力に失敗するユーザに対しても否定的なコメントは見受けられなかった。むしろ、カメラ操作を行い放送をより良くしようとする姿勢を持つユーザに対し、好意的なコメントが目立った。現在の実装システムではニコニコ生放送のコメント連投規制/NG ユーザ登録機能などが存在するため、それがいたずらに対する抑止力として働いた可能性も考えられる。今後より悪質ないたずらを想定し、それに対応可能な機能も必要であると考えられる。

配信者は著者の一人であったため公平な観点で配信者側のコメントを述べるのは難しいが、実験の理解を深めるためその印象を報告する。まず、本システムが演出とコミュニケーションの向上に十分寄与できたと実感した。特に、視聴者が効果的に演奏を演出できた際には、演奏者として大変楽しい感覚を得た。また、本システムによってこれまでにない演奏の表現方法を発想することができた。例えば、視聴者が3つのカメラを一定間隔で交互に切り替える際に、両方のキーボードを同時に演奏するというパフォーマンスを行うようになった。このような演奏方法は本システムを利用する以前にはなかったものである。視聴者主導でパフォーマンスを切り替える点についても、好意的な感覚を得た。配信全ての時間を通じ、演奏者が効果的なパフォーマンスを考えることは労力を要する。視聴者の要求

に応えるようパフォーマンスをすることは、放送を組み立てる上で助けられた点が多い。一方、演奏者によっては演奏者の意図と視聴者の意図が大きく異なった際に、視聴者主導のパフォーマンス決定に不満を感じることもあるかもしれない。この点に関しては、2つの解決方法が考えられる。第一に実験で得られた経験として、演奏者が自分の意図をパフォーマンスとして視聴者に伝え続けることで、視聴者が演奏者の意図にあわせるよう演出を変更してくれるということである。第二の方法として、本システムでは視聴者による機器制御を停止するインタフェースを用意している。実験ではこの機能を利用したいと思う状況には遭遇しなかったが、前述した荒らし行為などに対しても本機能は有効であると考えられる。

4. 議論およびデザインに対する示唆

4.1 より効果的な演出に向けて

本稿ではカメラの制御に焦点をおいたが、照明や特殊効果の制御も期待できる。評価アンケートでも、「ミラーボールみたいな機器も制御できたら楽しい」との意見が見受けられた。制御機器の種類を増やすことで、よりプロフェッショナルな演出に近づくことが可能となると考えられる。カメラ自体に関しても、簡易クレーンカメラやクアッドヘリコプターなどを導入することでよりダイナミックな映像が撮影可能になると考えられる。

機器制御をより直感的に行うためのインタフェースの開発も期待される。例えば、マルチタッチジェスチャを用いたカメラ制御を可能とすることで、直感性が増すと考えられる。動画配信の遅延を考慮した機器制御手法も有効である。これにより、視聴者の期待したタイミングで機器制御を行うことが可能となるであろう。一方で、機器を制御しようとする視聴者の存在を隠さないことは重要である。本実験結果から、演出に参加する視聴者に対して、他の視聴者からの賛辞のコメントが多数見受けられた。このポジティブなコミュニケーションを保つことは、優先すべき重要な点であると考えられる。

4.2 他分野への応用

本稿では音楽演奏に着目したが、提案システムはアート、

ダンス、手品、ボードゲーム、スポーツなど様々なライブパフォーマンスに適用可能である。ユビキタスコンピューティングの研究分野では、生活の支援を目的とし、多数の機器がネットワークに接続されたスマート空間が多く実装されてきた [13]。これらの空間の機器を対象とし本システムを導入することで、生活の効率化だけでなく、他者とのコミュニケーションが可能となると考えられる。またコンテキスト認識技術を本システムに適用することで、配信者の意図を多様な形式で視聴者に伝えることが可能となり、機器制御の指針とすることが可能であろう。本システムはホーム環境だけでなく、屋外やライブハウスなどでの適用も可能である。特にライブハウスに適応することで、新たな観客の獲得につながり、ミュージシャンと観客の関係性をより強いものにする可能性もある。

4.3 手動操作と自動操作のバランス

本実験を通じ、我々は手動操作と自動操作のバランスの重要性について気づきを得た。もし完全な自動化制御のアプローチをとれば、本実験のようなコミュニケーションの創発にはつながらなかったであろう。一方で、我々はカメラ制御に関していくつかの制限を設けた。例えば、ズームインエフェクトは、150%までの拡大に制限している。もし視聴者が zoomstop コマンドを送らなかったとしても、映像の乱れを防ぐためズームエフェクトは自動で停止するよう設計されている。

Parasuraman らは、10 段階からなるオートメーションの形式を提案している [15] (表 6 参照)。演出効果のための機器の自動化はどのレベルでも行えると考えられるが、効率さとコミュニケーションの可能性のトレードオフを考慮することは重要である。もう一つの重要な要素として、制御方法のシンプルさが挙げられる。提案システムは非常にシンプルな機器制御の仕組みを提供したが、これは一方で視聴者自身が様々な工夫を行える余地がある。1 ヶ月を通じた実験の終わりには、ある視聴者は複数のコマンドを効果的に組み合わせ、巧みなカメラ制御を行うようになった。他の視聴者はこういった技術を身につけた視聴者を、賛辞を込めて「カメラ職人」と呼ぶようになった。通常のライブ配信では、配信者のみにスポットライトが当たっていた。一方、本実験を通じ、配信者だけでなく視聴者の存在が注目されるようになったのは興味深い事実である。我々はシンプルさが視聴者の存在、参加、そして人間の持つ能力の強調に重要な役割を果たすと考えている。

5. 関連研究

インタラクティブなインターネット配信を目的としたいくつかの研究がなされている。Interactive television [9] は、経験の共有を目的とし、視聴者と配信者間でインタラクティブな機能の提供を目的としている。しかし、これらの研究例の多く [3], [6], [20] はテキストやボイスチャットなど、言語コミュニケーションを主に用いており、本研究とは異なる。本研究では機器操作を媒体とした非言語コミュニケーションに着目しており、演奏パフォーマンスに直結

表 6 決定と実行の自動化のレベル

Table 6 Levels of automation of decision and action selection [15].

(High)
10. The computer decides everything, acts autonomously, ignoring the human.
9. informs the human only if it, the computer, decides to
8. informs the human only if asked, or
7. executes automatically, then necessarily informs the human, and
6. allows the human a restricted time to veto before automatic execution, or
5. executes that suggestion if the human approves, or
4. suggests one alternative
3. narrows the selection down to a few, or
2. The computer offers a complete set of decision/action alternatives, or
1. The computer offers no assistance: human must take all decisions and actions.
(Low)

したインタラクティブリティを提供する。結果として、配信者・視聴者間の一体感が創出された。斉藤ら [18] は、視聴者が配信側のカメラの制御要求を容易に伝達することができるインタフェースを提案しているが、最終的にカメラを制御するのは配信者側となっている。我々の研究は、実際のインターネットライブ配信サービス上において、多くの視聴者によって同時に演出が試みられた最初のケースとして報告される。

HCI 分野においては、カメラ制御の半自動および全自動制御の研究が広くなされてきた。半自動制御の手法として、TVML (TV program Making Language) [12] はテレビ番組を記述できるテキストベースの言語であり、TVML で番組台本を書くことで、TVML プレイヤーでテレビ番組として再生することができる。道家らは番組に必要な情報を入力することで自動に必要な TVML スクリプトを構築するシステムを開発し、カメラワークの自動切り替えを可能としている [22]。FlySPEC [10] はパノラマカメラと PTZ カメラを組み合わせ、講義中の興味がある箇所を複数のユーザに選択してもらい、その要求に応じてカメラを制御する手法を提案している。また全自動制御の手法として、講義 [2], [17] やミーティング [8], [16] の様子を記録する研究が挙げられる。Ranjan ら [16] は、モーショントラッキングシステムとマイクロフォンを用い、テレビ制作の原理を適用して自動的にミーティングの様子を捉えるカメラを制御する手法を実現している。これらの研究は効率的な映像の制作・配信に寄与するが、我々はカメラの制御を用いたコミュニケーションの拡張を試みている点で異なる。我々は実験を通じ、単純なマニュアル制御が経験の共有、新たなコミュニケーションの創発、一体感の創出に寄与することを示した。

本研究は CSCW 分野とも関連している。近年複数ユーザでのビデオ制作システムの提案、およびそれを用いた際の人間の行動様式の実験調査、実験に基づいた協調システムのためのデザインの示唆が報告されている [4], [5], [19]。Engström らは [4] ダンスクラブの VJ パフォーマンスを分析し、観客が撮影した映像を VJ に応用可能な SwarmCam を提案している。また、Engström [5] らは Instant Broadcasting

System というシステムの利用を通じ、複数のアマチュアユーザ間での協調ビデオ編集作業がどう行われるか、実験を行っている。これらの研究は協調作業支援システムにおいて複数のユーザらがどのようにインタラクションを行うか興味深い考察を行っており、本研究と関連する。本研究とこれら関連研究との違いは、協調作業システムを実際のライブストリーミングサービスの演出に適用するだけでなく、コミュニケーションの創発に着目し、実験を通じてその効果を検証した点である。

6. まとめ

インターネット上のコンテンツにおける個人の生放送配信は、今後ますます拡大していくと考えられる。本研究では、個人の演奏家によるライブ生配信において、視聴者による遠隔演出により配信者と視聴者のコミュニケーションを可能とするシステム、視聴者協力型ライブ演出システムを提案した。本稿では特に複数のカメラワークを視聴者がコメント投稿により実現可能とするシステムを構築し、ニコニコ生放送にて4週間にわたった配信実験を行った。実験の結果、視聴者協力型ライブ演出システムは、(1) 視聴者と配信者間、及び視聴者間でこれまでにない非言語コミュニケーションを創発するとともに、(2) ライブ演奏における演出の向上に大きく寄与することが確認された。視聴者は演出に参加することで、演奏生放送の主体として参加することができ、共に放送を作り上げていくという一体感の向上が実現された。これらのことから、視聴者協力型ライブ演出システムは、演奏者と視聴者両者にとって、大きな利点を与えられることがわかった。更に本稿では実験結果から、将来のインタラクティブ生放送システムのためのデザインの指針を示した。

謝辞 インターネットライブ放送に参加して頂いた視聴者の皆様に感謝する。本研究の一部は、独立行政法人情報通信研究機構に支援頂いた。

参考文献

- [1] Ustream asia 日本の著作権管理事業者と楽曲の利用許諾に関する包括契約を締結。 http://www.softbank.co.jp/ja/stc/group_sup/20100706.01.pdf.
- [2] M. Bianchi. Automatic video production of lectures using an intelligent and aware environment. In *Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia*, MUM '04, pp. 117–123, New York, NY, USA, 2004. ACM.
- [3] T. Coppens, L. Trappeniers, M. Godon, and F. Wellesplein. Amigotv: towards a social tv experience. *Proceedings from the Second European Conference on Interactive Television Enhancing the experience University of Brighton EuroITV*, 36(11):1–4, 2004.
- [4] A. Engström, M. Esbjörnsson, and O. Juhlin. Mobile collaborative live video mixing. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, MobileHCI '08, pp. 157–166, New York, NY, USA, 2008. ACM.
- [5] A. Engström, M. Perry, and O. Juhlin. Amateur vision and recreational orientation:: creating live video together. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, pp. 651–660, New York, NY, USA, 2012. ACM.
- [6] D. Geerts. Comparing voice chat and text chat in a communication tool for interactive television. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, NordiCHI '06, pp. 461–464, New York, NY, USA, 2006. ACM.
- [7] U. Inc. Ustream.tv: You're on. <http://www.ustream.tv/>.
- [8] T. Inoue, K.-i. Okada, and Y. Matsushita. Learning from tv programs: application of tv presentation to a videoconferencing system. In *Proceedings of the 8th annual ACM symposium on User interface and software technology*, UIST '95, pp. 147–154, New York, NY, USA, 1995. ACM.
- [9] J. F. Jensen. Interactive television - a brief media history. In M. Tscheligi, M. Obrist, and A. Lugmayr eds., *EuroITV*, Vol. 5066 of *Lecture Notes in Computer Science*, pp. 1–10. Springer, 2008.
- [10] Q. Liu, D. Kimber, J. Foote, L. Wilcox, and J. Boreczky. Flyspec: A multi-user video camera system with hybrid human and automatic control. In *IN ACM MULTIMEDIA 2002*, pp. 484–492. ACM Press, 2002.
- [11] Y. LLC. Youtube - live. <http://www.youtube.com/live>.
- [12] T. K. M Ueda Hayashi, H. Tvml (tv program making language) automatic tv program generation from text-base script.
- [13] S. Meyer and A. Rakotonirainy. A survey of research on context-aware homes. In *Proceedings of the Australasian information security workshop conference on ACSW frontiers 2003 - Volume 21*, ACSW Frontiers '03, pp. 159–168, Darlinghurst, Australia, Australia, 2003. Australian Computer Society, Inc.
- [14] J. Nielsen. *Usability Engineering*. Morgan Kaufmann Publishers Inc., 1995.
- [15] R. Parasuraman, T. B. Sheridan, and C. D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems Man and Cybernetics Part A Systems and Humans*, 30(3):286–297, 2000.
- [16] A. Ranjan, J. Birnholtz, and R. Balakrishnan. Improving meeting capture by applying television production principles with audio and motion detection. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pp. 227–236, New York, NY, USA, 2008. ACM.
- [17] Y. Rui, A. Gupta, J. Grudin, and L. He. Automating lecture capture and broadcast: technology and videography. *ACM Multimedia Systems Journal*, 10:3–15, 2004.
- [18] Y. Saito and Y. Murayama. An empirical study of audience-driven interactive live television on the internet. *EuroITV'09*, 2009.
- [19] S. Vihavainen, S. Mate, L. Seppälä, F. Cricri, and I. D. Curcio. We want more: human-computer collaboration in mobile social video remixing of music concerts. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pp. 287–296, New York, NY, USA, 2011. ACM.
- [20] J. D. Weisz, S. Kiesler, H. Zhang, Y. Ren, R. E. Kraut, and J. A. Konstan. Watching together: integrating text chat with video. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pp. 877–886, New York, NY, USA, 2007. ACM.
- [21] 株式会社ニワンゴ. ニコニコ生放送. <http://live.nicovideo.jp/>.
- [22] 道家, 林, 牧野. Tvml を用いた番組情報からのニュース番組自動生成. 映像情報メディア学会誌: 映像情報メディア, 54(7):1097–1103, 2000-07-20.