

YouSubcon! (ユーさぶこん!): 配信型音楽ライブのCGM化に関する研究

増田慶士郎^{†1} 高島健太郎^{†1} 西本一志^{†1}

概要: 近年音楽コンサートや舞台において無観客でのオンライン配信, また有観客公演のオンライン同時配信が増加している. しかし配信される映像は主催者で用意した1種類であるため, 視聴者の多岐にわたるニーズに対応できていない. そこで映像編集ができるファン(ディレクター視聴者)に複数のカメラの収録映像を無編集で提供し, 編集してもらうことで配信映像のバリエーションを増やすことで, 多様なニーズにできる限り対応できるシステムを提案した. 従来の1種類の映像, 複数のカメラの収録映像をそのまま提供するマルチアングル映像と提案手法の3条件での比較実験を行ったところ, 満足度はマルチアングル映像に及ばなかったが, 複数の映像から選択可能であることは有益であること, 動画の選択方法に改善の余地があること示唆された.

1. はじめに

新型コロナウイルス感染症 (COVID-19) の流行により様々な文化活動に影響が出ている. 感染拡大初期では音楽コンサートや舞台の多くの公演が延期・中止された. その後, コロナ禍の中で興行を続けるため, 無観客でのオンライン配信という形態で開催する公演が増加した. また感染状況が落ち着くにつれて, 会場での有観客公演をオンラインで同時配信するという公演も増えつつある. オンライン配信であれば, 会場の収容人数による来場者数上限などの問題を緩和, もしくは解決することが可能である. ぴあ総研による調査では, 有料オンラインライブの国内市場規模が2020年4~6月には11億円であったのに対し, 同年10~12月には373億円となったと推計されている[1]. また年単位の調査でも, 2020年全体で448億円であった国内市場規模が2021年全体で512億円に増加していると推計されている[2].

現状のオンライン配信では, 会場の設置カメラが複数台あるにも関わらず配信映像はそれらのカメラの映像を主催者が編集した1種類のみである. しかし視聴者は個人ごとに視聴したい映像のニーズが異なると思われる. 例えば全景映像と呼ばれるステージ全体を定点で映した映像, 推しと呼ばれる好みの出演者を対象にその出演者をメインに映した映像, 各タイミングで歌唱中の出演者を映した映像などである. 予備的な調査として著者らが研究室内の学生2名に対し同じ曲をパフォーマンスしているアングルが違う3種類のアニメ映像を視聴してもらい好みの順番の回答を依頼した結果, 学生ごとに異なる回答となった. これらのことから, 現状の1種類の映像のみのオンライン配信では視聴者の多様なニーズに対応しきれていないという問題があるのではないかと考える.

複数種類の映像を作成し配信を行うことが望ましいが, これを行おうとすると主催者の編集の負担が増加してしまう. そこで本研究では, 主催者が複数台あるカメラの映像を編集せずに出演者のファンに渡し, 出演者のファンがリアルタイ

ムでこれを編集することで複数の配信映像の提供を実現することを考えた. 映像編集者にファンを選んだ理由としては, ファンパフォーマンスの見どころを理解している可能性が高い点, ファンごとに見てほしいポイントが異なることから多様な映像編集が期待される点が挙げられる.

本稿では, このようにライブのオンライン配信において, 多様な視聴者のニーズに応えるために, ファンが映像編集を行うことで複数種類の映像配信を行う手法を提案し, 効果検証を行った結果について報告する.

2. 先行事例

本研究に関連する先行事例として, 様々なアングルからの撮影を可能にする撮影技術, 視聴者が多様な映像を選択可能にする配信技術について述べる.

様々なアングルからの撮影を可能にする撮影技術としてはボリュメトリックビデオ技術[3], EyeVision[4]などがある. ボリュメトリックビデオ技術は撮影スタジオの全体にカメラを設置し, 同時に撮影することで空間を3Dデータ化し, 即座に3D映像を生成できる技術である[3]. しかし3D映像を生成することで好きなアングルを捉えられるものの撮影した映像を編集するという3D映像の生成処理が必要なうえ, 大掛かりな設備が必要なため, ライブでの普及は発展途上である. EyeVisionは2001年にアメリカのスーパーボウル中継で導入された3Dリプレイシステムである[4]. 会場に等間隔で設置されたカメラ同士で死角を補い合うことで会場のどこでスーパープレイが起こっても映像として捉えることができる. しかしEyeVisionは死角を補い映像補完することを主に想定したシステムであり, 本研究の主旨である各出演者や会場の様子を捉えるという点には合致しない. また両方に言えることとして, 映像の編集には専用の設備及び技術が必要であり, 場合によっては専用の場所などが必要であるため容易に使用できるものではない.

一方, 視聴者に多様な映像を選択可能にする配信技術としてはマルチアングル映像[5]などがある. マルチアングル映像とは複数のカメラでアングルを変えて同時に撮影した各映像

^{†1} 北陸先端科学技術大学院大学 先端科学技術研究科
Graduate School of Advanced Science and Technology, Japan Advanced
Institute of Science and Technology

を1つのタイトルの中に収録し、各映像を視聴者の好みに応じて切り替えて視聴できるようにしたものである。視聴者は常にカメラを切り替えられるため、見たいアングルの映像を視聴することができる。しかし視聴者自身が映像の切り替えを行う必要があり、「映像を選択する」「映像を視聴する」という2つのタスクが生まれるため、視聴に集中することができない。そのため視聴者は映像を選択するための小画面ばかりを注視することになり、映像を大画面で楽しめない恐れがある。

3. 提案手法

本手法に係る関係者として、アーティストおよび主催者、映像編集ができる出演者のファン（以下「ディレクター視聴者」）、映像編集を行わない視聴者（以下「一般視聴者」）がある。提案手法の概要を図1に示す。

アーティストおよび主催者はステージでのパフォーマンスを行い、その様子を複数のカメラで撮影する。そして撮影したそれらの映像をディレクター視聴者に対し無編集の状態で提供する。

ディレクター視聴者は提供された無編集の映像を受け取り、リアルタイムで編集する。編集を行った映像は一般視聴者にそのまま配信される。その際、編集した映像に対し視聴者が選ぶ上で参考になるキャッチフレーズをつけ公開する。キャッチフレーズはディレクター視聴者がどのような目的の映像を編集したかを示すものであり、「出演者のうちAさんを中心に映している映像」「舞台セットを含めた全体を映している映像」といったものを想定している。

そして一般視聴者はディレクター視聴者が作成したキャッチフレーズをもとに映像を選択、視聴する。

4. 予備実験

4.1 概要

本手法を実現するにあたり、ディレクター視聴者は映像編集を行うことができるか、ディレクター視聴者ごとに編集された映像が変化するかを確認するために予備実験を行った。

予備実験では作成した実験システムを用いて提供映像から配信映像を作成してもらう作業を行った。編集はカメラのスイッチングのみを行ってらっている。

図2は本研究で使用する実験システムの動作画面である。

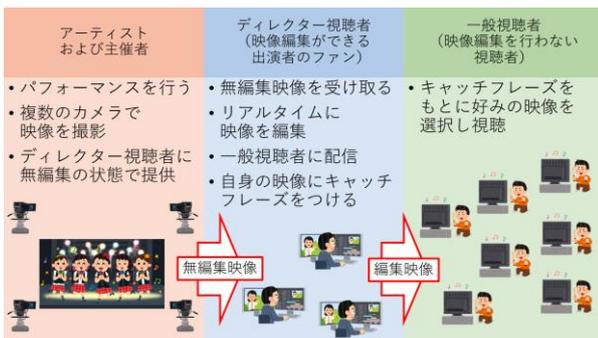


図1 提案手法の概要



図2 予備実験システム 動作画面

②は主催者から提供されるカメラ映像を想定したものであり、複数のカメラアングルの映像がそれぞれ同タイミングで再生されている。その映像のうち視聴したいカメラアングルのアイコンをクリックすることによりその映像が①の選択映像ウィンドウに大きく表示される。①は視聴者に配信される映像を想定している。リアルタイム編集を模するために各映像は巻き戻せないようになっている。またライブの音は操作に関わらず映像に合わせて流れ続ける。①に投影する映像を変更すると、変更した映像の番号および再生開始からの経過時間がシステムにログとして記録される。システムはHTML及びJavaScriptを用いて作成している。

予備実験ではこのログに加え、作成された映像の保存および実験後のインタビューを行った。実験は著者らの研究室の博士前期課程学生2名である。映像に関しては本研究室の博

表1 予備実験 実験結果

被験者A		被験者B	
経過 (秒)	動画番号	動画番号	経過 (秒)
0.000	未選択	未選択	0.000
		全身	0.276
		表情	2.153
4.353	表情	鍵盤	5.995
13.733	鍵盤	表情	16.643
		鍵盤	28.053
51.384	全身	表情	51.481
		ペダル	59.397
67.631	表情	鍵盤	71.547
		表情	79.280
90.188	鍵盤	鍵盤	91.133
		全身	103.020
		ペダル	113.409
114.784	全身	表情	122.557
		鍵盤	130.370
		表情	147.233
150.500			150.500

士前期課程学生1名がピアノを演奏している様子を「全体の様子」「演奏者の表情」「鍵盤」「足元のペダル」の4ヶ所同時に撮影した映像を無編集の状態で使用している。

4.2 結果と考察

実験結果を表1に示す。行は実験開始からの経過秒数であり、左列が被験者A、右列が被験者Bの編集操作である。

被験者Aと被験者Bの大きな違いとして、被験者Bはすべてのカメラアングルを使用した一方で、被験者Aは足元のペダルを映した映像は使用しなかった。また、被験者Aはペダル以外のカメラを同じ回数用いたが、被験者Bは演奏者の表情を映したカメラを他のカメラよりも頻繁に用いた。

またインタビューで「作成映像のコンセプト」「カメラアングルを切り替えたきっかけ」「カメラアングルの切り替え先を決めたきっかけ」について聞いたところ、作成映像のコンセプトでは「飽きず、視聴者の興味対象を意識した映像（被験者A）」「音楽番組でのピアノ演奏の映像をイメージした映像（被験者B）」という回答がそれぞれ得られた。カメラアングルを切り替えたきっかけに関しては「演奏者が視聴者の興味対象だと考えた（被験者A）」「演奏者が座りなおすなど、演奏者の動きは映すようにした（被験者B）」という回答が得られた。またカメラアングルの切り替え先を決めたきっかけに関しては「全体の様子と演奏者の表情を多めにするようにした（被験者A）」「同じ映像ばかりで飽きが来ないようにした（被験者A）」「後半以降は使用していない映像を優先的に使用するようにした（被験者B）」という回答が得られた。

これらのことから本手法で提案するリアルタイムでの映像編集及びキャッチフレーズの作成は可能であることが分かった。またディレクター視聴者が想定する視聴者のニーズは同一でないことが改めて確認できた。そのため、今後ディレクター視聴者を増加させた場合も、すべてが同一の傾向の映像になるとは考えにくく、多様な視聴者ニーズに対応できるのではないかと考えられる。

5. 本実験

5.1 概要

予備実験を踏まえ、実際に本手法が視聴者のニーズに対応



図3 ディレクター視聴者側実験システム・一般視聴者側実験システム（マルチアングル映像用）動作画面

することで満足感の向上につながるかを検証するために実験を行った。実験はディレクター視聴者側と一般視聴者側それぞれで実施している。実験に用いたシステムと実験内容について下記に述べる。

5.2 ディレクター視聴者側

ディレクター視聴者には実験システムを用いカメラ映像を編集してもらった。実験手順としては初めに操作方法、カメラアングルの説明を行い、被験者が操作を理解した段階でチュートリアル用の映像で練習を行い、その後、3曲分の映像を編集してもらった。各映像の編集前に「ステージへの出演人数」「編集する映像に使用する曲の概要」のみを被験者に伝えた。各映像編集後、被験者には作成した映像にキャッチフレーズをつけてもらった。

図3は使用したシステムの画面である。4.1節で述べた予備実験で使用したシステムと基本的には同じだが、予備実験では実験プログラムを開くと同時に再生していた映像を、実験上の利便性から③の再生ボタンを押すことによる再生へと変更した。

ライブ映像には3DCG作成ソフトウェアであるMikuMikuDance[6]を使用し作成したキャラクターのダンス映像を使用した。楽曲はチュートリアルが「ロウワー」、実験使用曲が「トラフィック・ジャム」「グッバイ宣言」「ヒビカセ」である。「ロウワー」及び「トラフィック・ジャム」は2人でのステージ、「グッバイ宣言」は1人でのステージ、「ヒビカセ」は4人でのステージとした。キャラクターモーションおよびステージのモデルデータに関してはインターネット上で無料公開されているデータを使用した。また、キャラクターのモデルデータに関してはMikuMikuDanceに収録されている「初音ミク」「鏡音リン」「鏡音レン」「巡音ルカ」「MEIKO」「カイト」のデータを使用した。映像のカメラアングルは「正面（通常距離）」「斜め22.5度（左側）」「斜め45度



図4 正面（通常距離）



図5 斜め22.5度（左側）



図6 斜め45度（左側）



図7 斜め22.5度（右側）



図8 斜め45度（右側）

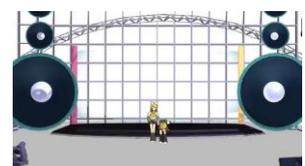


図9 正面（遠距離）

(左側)「斜め 22.5 度 (右側)」「斜め 45 度 (右側)」「正面 (遠距離)」の 6 種類である。実際のカメラアングルを示した図を図 4~9 に示す。

被験者は著者らが所属する大学の大学院生 6 名とした。評価方法としては、各映像の編集終了毎に編集した映像のキャッチフレーズを聞くほか、実験終了後にアンケートとインタビューを実施し、編集環境についての質問を行った。編集された映像は保存し、キャッチフレーズとともに後述する一般視聴者側実験にて使用する。

5.3 一般視聴者側

一般視聴者側の実験では、従来通りの「1 種類の映像 (以下「指定映像」)」、複数のカメラ映像をそのまま見せる「マルチアングル映像」、ディレクター視聴者が編集した映像を選択して視聴する「本手法の映像(以下「本手法」)」の 3 種類を視聴してもらい、満足度をアンケートおよびインタビューで調



図 10 一般視聴者側実験システム (指定映像・本手法用) 動作画面

査した。

指定映像は各カメラをまんべんなく用いて第一著者が編集した映像を使用した。実験システムの外観を図 10 に示す。マルチアングル映像は先述のディレクター視聴者側実験で提示した映像と同じカメラ映像・システムを使用した。実験システムの外観は図 3 と同じである。被験者は映像を視聴途中で変更することができる。本手法の映像では事前にディレクター視聴者側実験で作成したキャッチフレーズを表示したリストを提示し、その中から見たい映像を選択してもらった。マルチアングル映像と違い選択した映像は途中で変更することはできない。システムの外観は図 10 と同じである。

被験者は著者らが所属する大学の大学院生 17 名を 3 つのグループ(A,B,C)に分け、順序効果を排除するため、グループごとに同内容の映像に対する提示方法を変えたうえ、実験を行ってもらった (図 11)。実験終了後にアンケートとインタ

	Aグループ	Bグループ	Cグループ
1回目 1種目の映像	マルチアングル映像	本手法	指定映像
2回目 2種目の映像	指定映像	マルチアングル映像	本手法
3回目 3種目の映像	本手法	指定映像	マルチアングル映像

図 11 グループ毎の実験順番

表 2 ディレクター視聴者が作成したキャッチフレーズの一覧

曲名	被験者番号	キャッチフレーズ
トラフィック・ジャム	1	ダンスがよく見えるようにした映像
	2	全体中心に映した映像
	3	MEIKO (女子) を中心に映した映像①
	4	カメラから出演者が近いようにしたうえ斜めからのアングルを重視した映像
	5	MEIKO (女子) を中心に映した映像②
	6	ダンスしている人の顔がよく見えるようにした映像
グッバイ宣言	1	曲の小節を意識して切り替えた映像
	2	全体中心に映した映像①
	3	ミュージックビデオの雰囲気にした映像
	4	中央を中心に映した映像①
	5	全体中心に映した映像②
	6	中央を中心に映した映像②
ヒビカセ	1	個人個人のダンスを被らないようにした映像
	2	全体を中心に映した映像①
	3	全体を中心に映した映像②
	4	4人全員を映るようなアングルにした映像
	5	左右カメラの使用割合を均等にした映像
	6	4人全員が均等に映るようにした映像

ビューを実施し、一番満足した動画、満足していない動画、それらの理由について質問した。なお、本稿執筆時点でまだ被験者を募集している状況であり、以降これまでに得られたデータについて述べる。

6. 結果と考察

6.1 ディレクター視聴者側

ディレクター視聴者が作成したキャッチフレーズの一覧を表2に示す。

アンケートでは主に「映像編集に対しキャッチフレーズ以外で意識した点」「キャッチフレーズを考慮することが負担となるか」「システムへの追加機能の要望」について聞いた。

映像編集に対しキャッチフレーズ以外で意識した点では、「曲のリズムに合わせるようにした」「サビなどの盛り上がるシーンでは正面からのカメラアングルとした」「出演者が複数人いる場合は人同士が重ならないようなカメラアングルを選択した」などの回答が得られた。

キャッチフレーズを考慮することが負担となるかに関しては、なんと答えた人とならないと答えた人に分かれたほか、条件によると答えた人もいた。負担となるという回答に関しては「テーマを持つことはできてそれをキャッチフレーズのように言語化できない」という意見があった。一方ならないという回答に関しては「そもそもライブ自体にテーマがあるためそれをもとに考えられる」という意見があった。条件によるという回答に関しては「人数や内容による」「容易に考えられるがライブ中にそれを貫き続けるのは難しい」という意見が得られた。

システムへの追加機能の要望に関しては、「出演者を追従するようなカメラアングルの映像」「斜め下や斜め上、真上、背面など今回の実験では用意していないカメラアングルの映像」「映像に対するズームインやズームアウトなど映像の距離感を変える機能」「映像の上からエフェクトをかける機能」などが挙げられた。

6.2 一般視聴者側

これまでに得られたアンケートの結果を表3に示す。一番満足した動画として回答者が多かったのはマルチアングル映像であり17人のうち10人が選択した。一方本手法は3人と全手法の中で一番少なかった。一番満足していない動画とし

て回答者が多かったのは指定映像であり17人のうち11人が選択した。本手法は2人とこちらは全手法の中で1番少なかった。

自由記述およびインタビューにて聞いた動画順位の選択理由としては、マルチアングル映像は「好きなアングルで視聴することができる」という好印象の意見が多かった一方で「自身でアングルを調整するのが少し面倒だと感じた」「カメラが多くても使わなかった」などという意見が得られた。一番満足した動画として挙げられた一方、4名が満足していない動画として答えているのは、この手間によるものが大きいと考えられる。

従来手法である指定映像が一番満足していない動画として最も回答数が多かった。「引きで見るべきでないタイミングで引きのアングルになったりしたため見づらかった」など単調なカメラ切り替えの動画を視聴せねばならず、自分の嗜好に合った動画の選択ができないことが満足度を下げる大きな要因だと考えられる。

本手法が一番満足した動画にも一番満足していない動画にも回答が少ない結果となった。動画を選択可能にしたことは、指定映像より良い評価に結び付いたと推測される。一方、一番満足した動画では指定映像よりも劣ってしまった理由として、インタビューより、選んだキャッチフレーズと実際の動画でイメージの齟齬があったことが考えられる。アンケートでも「ダンスの見やすさで選んだはずなので切り替わるとやや見づらい印象を受けた」という意見があった。また、本手法では映像を切り替えることはできないが、他の映像が存在するという事は伝わるので、不本意な動画であった場合に別の映像はどうなっていたのかが気になり、満足度を下げた可能性もある。

より満足度を上げるための改善点としては、本実験で6種類であったバリエーションを増やすことが有効であると考えられる。しかし、ただ増やすだけでは映像の選択が大変になるため選択を行いやすくするための工夫も同時に必要である。また動画を切り替えたいというアンケートでの要望も多く、マルチアングル同様に途中での切り替え機能を足すことも考えられる。しかし切り替え機能を足すことによりマルチアングル映像同様に映像を選択するというタスクが増えてしまうため十分な検討が必要である。

表3 一般視聴者側実験結果

	一番満足した動画			一番満足していない動画		
	指定映像	マルチアングル	本手法	指定映像	マルチアングル	本手法
Aグループ	0	4	2	5	1	0
Bグループ	3	4	0	3	2	2
Cグループ	1	2	1	3	1	0
合計	4	10	3	11	4	2

7. 実現時の課題

本研究をビジネスとして実現する場合に生ずる主要な課題として「権利的問題」「金銭的問題」について述べる。

権利的問題は、「ライブという肖像権などが関わる映像の提供をどのようにして行うか」「編集後の映像の、著作権などの権利はどのように扱われるか」などがある。ディレクター視聴者の画面では画面キャプチャを制限する必要があるほか、事前の確認や規約の記載が重要であると思われる。

金銭的問題としては、ライブ開催・配信にかかる費用、そして収益をどのように分配するかがあげられる。通常のライブ配信では興行・配信に係る費用は主催者が負担し、視聴者からのチケット収益を得てそれを賄う。しかしディレクター視聴者による配信は想定されていないため仕組みが存在しない。著者としては、ディレクター視聴者は一般視聴者より少し高額なチケット料金を支払い、その後編集した映像の視聴数に応じてグッズ、キャッシュバックなど何かしらのインセンティブを受け取れることにすることで、充実した映像を作成するようになるのではないかと考えている。

8. おわりに

本研究では近年増加する音楽コンサートや舞台などのオンライン配信について、幅広い視聴者ニーズに対応するため、映像編集ができるファン(ディレクター視聴者)にカメラの収録映像を無編集で提供し編集してもらうことで映像のバリエーションを増やして配信するシステムを提案した。

従来の1種類の映像と、複数のカメラ映像を無編集で提供するマルチアングル映像との3条件での比較実験を行ったと

ころ、本稿執筆時点では一番満足した動画にはマルチアングル映像が選ばれ、提案手法はこれに大きく及ばない結果となった。一方で、一番満足していない動画としての回答数は少なく、マルチアングル映像よりも好ましい結果となった。これは、映像をある程度選択できる余地を与えつつ、操作をしながら視聴しなければならないという問題点を解決しているからではないかと考えらえる。

今後は実験を継続するとともに、判明した改善点をもとに、より満足感を向上させることができるシステムの構築に取り組みたいと考えている。

参考文献

- [1] “2020年の有料型オンラインライブ市場は448億円に急成長。～ポスト・コロナ時代は、ライブ・エンタテインメントへの参加スタイルも多様化へ / ぴあ総研が調査結果を公表 | ぴあ株式会社” . https://corporate.pia.jp/news/detail_live_enta_20210212.html, (参照 2022-11-30).
- [2] “2021年のオンラインライブ市場は前年比14.4%増の512億円に成長 / ぴあ総研が調査結果を公表 | ぴあ株式会社” . https://corporate.pia.jp/news/detail_live_enta20220615_3.html, (参照 2022-11-30).
- [3] “ポリュメトリックビデオ技術：空間を3Dデータ化し即時に3D映像を生成できる技術 | 知財図鑑” . <https://chizaizukan.com/property/382/>, (参照 2022-11-30).
- [4] “金出武雄 - 武田計測先端知財団調査報告書” . <http://takeda-foundation.jp/reports/pdf/ant0210.pdf>, (参照 2022-12-20).
- [5] “マルチアングルとは何？わかりやすく解説 Weblio 辞書” . <https://www.weblio.jp/content/%E3%83%9E%E3%83%AB%E3%83%81%E3%82%A2%E3%83%B3%E3%82%B0%E3%83%AB>, (参照 2022-11-30).
- [6] “Vocaloid Promotion Video Project” . <https://sites.google.com/view/vpvp/>, (参照 2022-12-20).