

イベント会場を自由に視聴できるライブ映像配信システムの一検討

山口好江^{†1} 田中康暁^{†1} 深澤勝彦^{†1} 佐藤啓之^{†2} 土川仁^{†2}
木全英明^{†1} 下村道夫^{†2} 小島明^{†1}

あらまし イベント会場を自由に視聴できる「インタラクティブ・ライブ映像配信」のプロトタイプシステムを開発した。カメラ1台カメラワーク固定の映像は、人の視覚に比べて視聴範囲が狭いという問題がある。本システムは、ユーザが視聴位置を操作することで視聴範囲を仮想的に広くする機能を有する。この機能の有効性を確認するため、講演会のライブステージ全体の中から見たいところを選択、拡大する視聴について主観評価を行った。視聴対象物の移動範囲が少なくカメラ1台の画角に収まる講演会に対しては、本システムが十分に適用可能であることを確認した。また、本システムでは、イベント会場の複数個所にカメラを設置し、自由にカメラを選択した視聴が可能である。このカメラ選択方法について主観評価を行ったので、その結果についても報告する。

A Study on Live Video Streaming System for Free Viewing of Event Site

YOSHIE YAMAGUCHI^{†1} YASUAKI TANAKA^{†1}
KATSUHIKO FUKAZAWA^{†1} HIROYUKI SATOU^{†2}
MEGUMU TSUCHIKAWA^{†2} HIDEAKI KIMATA^{†1}
MICHIO SHIMOMURA^{†2} and AKIRA KOJIMA^{†1}

1. はじめに

モバイル環境における映像視聴のユーザが近年増加の傾向にある。情報通信白書によると動画共有サービスを約4割の人が家庭外、約3割の人が家庭内で利用していると回答している。また、スマートフォンの利用者が前年比に比べて約3倍と急速に普及が進んでいる。

私たちはライブ配信でいつでもどこでも視聴可能なモバイル端末に対応した映像視聴であって、かつイベント会場の臨場感の高い視聴を可能とする映像配信サービスの実現を目的に、イベント会場を自由に視聴できる「インタラクティブ・ライブ映像配信」のプロトタイプシステムを開発した。

この「インタラクティブ・ライブ映像配信」は、イベント会場を自由に視聴するために会場内に複数カメラを設置し、メイン会場には高精細なカメラを用い、会場の各所では小型カメラを用いることとした。このカメラ配置のもと開発した本システムでは、メイン会場を撮影するライブ映像全体の中から好きな部分を選択、拡大して視聴可能とする StageView モードと、複数台の小型カメラを放射状に設置して 360°方向からカメラ映像を選択して視聴可能とする WalkingView モードとを切り替えてタブレット端末での

視聴を可能としている。

タブレット端末による高解像度映像全体の中から見たいところの選択、拡大視聴は、既にビデオ・オン・デマンド配信で実現している。ビデオ・オン・デマンドでの高解像度映像は複数のハイビジョンカメラで撮影した各映像を1枚のパノラマ映像に合成したものであり、広範囲のパノラマ映像の中から自由に見たいところを選択、拡大視聴を可能としている。

一方近年、ハイビジョンの4倍の画素数を持つ4Kカメラの製品化が進み一般的に入手可能となってきた。そこで今回は、高精細な4Kカメラ映像1台での映像全体の中から見たいところの選択、拡大視聴を想定して「インタラクティブ・ライブ映像配信」のプロトタイプシステムを開発した。このシステムでは、パノラマ映像合成処理が不要となるため撮影から再生までにかかる時間の低減が期待できるが、4Kカメラ1台で撮影した映像では広範囲のパノラマ映像と違い画角が狭くなるという問題がある。そのため、画角が狭い映像の場合での見たいところの選択、拡大視聴に適するコンテンツを明らかにする課題がある。

本システムのコンテンツの適用範囲を確認するため、本稿では4Kカメラ撮影による講演会のライブ映像を対象として実際のイベント会場で評価実験を行った。さらにこれと同時に、本システムのユーザインタフェースの有効性を確認するための評価実験を行った。

また過去に、私たちが開発した別システム「マルチアングル映像ライブ配信システム」でのカメラワークがあるコ

^{†1} 日本電信電話株式会社 NTTメディアインテリジェンス研究所
NTT Media Intelligence Laboratories NTT Corporation

^{†2} 日本電信電話株式会社 NTTサービスエボリューション研究所
NTT Service Evolution Laboratories NTT Corporation

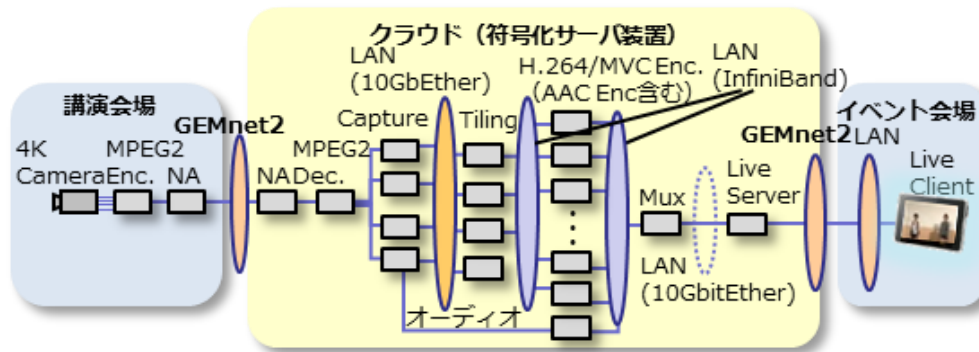


図 1 StageView モードシステム構成



図 2 講演会場の 4 Kカメラワーク

コンテンツのカメラ切り替えインタフェース評価実験を行ったことがある。その結果も参照しつつ報告する。

以下、第 2 章でインタラクティブ・ライブ映像配信の概要、第 3 章で実験結果について述べる。

2. インタラクティブ・ライブ映像配信

2.1 インタラクティブ・ライブ映像配信の概要

StageView モードは、高解像度映像の一部（見たい領域）を処理能力の低いタブレット端末で拡大視聴するために、複数解像度変換、タイル分割、符号化を高速処理する必要があり、各処理のクラスタ化を行い高速な分散処理を実現している。

また、WalkingView モードは、複数カメラの中から見たいカメラを自由に切り替える視聴を可能としている。

この 2 つのモードを 1 台のタブレット端末（Android4.0）上で切り替えて視聴できる。

以下に、StageView モード、WalkingView モードの各システム構成について詳しく述べる。

2.2 StageView モードのシステム構成

StageView モードのシステム全体は図 1 に示すように、講演会場（4 Kカメラ）、クラウド（符号化サーバ装置）、イベント会場（クライアント）で構成される。高速処理が必要な符号化サーバ装置はクラウド上に設置し、講演会場（4 Kカメラ）～クラウド（符号化サーバ装置）と、クラウド（符号化サーバ装置）～イベント会場（クライアント）を高速 NW（GEMnet2）で接続した。

GEMnet2 は、NTT の研究、開発用のネットワークであり、国内外の研究教育 NW と相互接続している。数 10Gbit/s の通信容量の基幹ネットワークである。

2.2.1 講演会場（4 Kカメラ）

講演者、説明画面がカメラ 1 台の画角に入るようにして講演会場のステージ全体を 4 Kカメラで撮影した（図 2）。また、カメラワークは固定とした。講演者の音声はマイクから直接と入力した。4 Kカメラからは 4 本に分割されたフル HD 映像が出力され、この 4 本のフル HD 映像を MPEG2（40Mbps×4 本）に一旦圧縮符号化し、GEMnet2 を用いてクラウド上の符号化サーバ装置に送信した。

2.2.2 クラウド（符号化サーバ装置）

図 1 に示すクラウド上では、

1. 4 本の映像をそれぞれに複数の解像度に変換し、さらに、複数解像度に変化した各映像をタイル状に分割する処理。

2. タイル状に分割した各映像を符号化する処理。また、タイル分割した映像の一部が 4 本映像のうちのどれかにまたがる場合は、1 つのタイルに結合して符号化する処理。

以上の処理をクラスタ化して高効率な分散処理を実現し、リアルタイム処理を可能としている。クラウド上の各装置の処理を以下に具体的に示す。

クラウド上の NA（受信装置）は、圧縮符号化された 4 Kカメラ映像である MPEG2（40Mbps×4 本）を GEMnet2 を介して受け取り、MPEG2 Dec（デコーダ装置）に送信する。

MPEG2 Dec（デコーダ装置）は、MPEG2（40Mbps×4 本）を復号し、復号した 4 本の映像を 4 台のキャプチャ PC に送信する。

Capture（キャプチャ PC）4 台は、受け取った 4 本の映像を Tiling（タイリングサーバ）に送信する。

Tiling（タイリングサーバ）4 台は、4 本の映像を複数解像度（本稿では 3 階層）に変換、さらに解像度毎にタイル状に分割（計約 200 個）処理をクラスタ化し高効率に分散処理する。

H.264/MVC Enc（圧縮符号化装置）22 台は、タイル分割（計約 200 個）した各映像の H.264/MVC 符号化処理

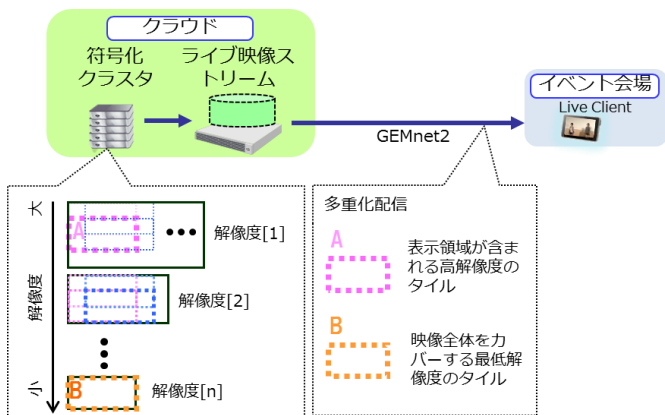


図 3 見たい領域映像と映像全体の多重化配信

をクラスタ化し高効率に分散処理する。

Mux (MUX サーバ) は、H.264/MVC 符号化された各タイル分割の映像を時系列にフレーム単位で多重化し、ライブ配信サーバに送信する。

LiveServer (ライブ配信サーバ) は、MUX サーバから多重化された全符号化データを受け取り、全符号化データの中からクライアントから要求された見たい領域が含まれる高解像度映像と映像全体をカバーする低解像度映像を取り出し、多重化 (2Mbps (高解像度)+2Mbps (低解像度)=4Mbps) して 1 本のストリームにしてクライアントに GEMnet2 を介して配信する。

見たい領域映像の配信を図 3 に示す。複数解像度の中から見たい領域が含まれる高解像度映像と全体をカバーする低解像度映像を多重化してストリーム配信する。

2.2.3 イベント会場 (クライアント)

ユーザが映像全体の中から選択した見たい領域の配信要求をライブ配信サーバに対して行い、LiveClient (クライアント) は見たい領域の配信要求をしている間は、全体をカバーする低解像度映像を用いて選択した領域を再生し、高解像度映像受信後に高解像度映像で見たい領域を拡大して視聴を可能としている。

2.3 WalkingView モードのシステム構成

WalkingView モードのシステムは図 4 に示すように、イベント会場の複数個所に WebHD カメラ、カメラサーバ、ライブ配信サーバ、クライアントで構成される。

2.3.1 Web HD Camera (WebHD カメラ)

会場の複数個所に WebHD カメラを設置し、撮影した WebHD カメラ映像を H.264 にリアルタイム符号化し CameraServer (カメラサーバ) へ送信する。

2.3.2 CameraServer (カメラサーバ)

カメラサーバは、複数の WebHD カメラから符号化したカメラ映像を受け取り、ライブ配信サーバへは要求された符号化カメラ映像を送信する。

2.3.3 LiveServer (ライブ配信サーバ)

ライブ配信サーバは、要求した符号化カメラ映像をカメ

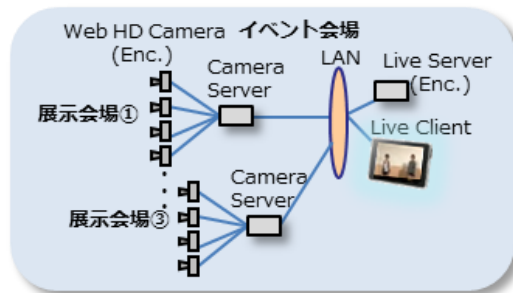


図 4 WalkingView モードシステム構成



図 5 WalkingView モードのカメラ設置

ラサーバから受け取る。さらに、WebHD カメラで符号化した映像は複数カメラ切り替え視聴に適していないため再符号化 (H.264) し、クライアントから要求された再符号化したカメラ映像をストリーム配信する。カメラ切り替えのためイベント会場で符号化、配信処理を行った。

2.3.4 LiveClient (クライアント)

ユーザがカメラ設置位置を示す MAP 上から直接カメラ映像の選択、または順送りにカメラを切り替えて選択したカメラ映像の配信要求をライブ配信サーバに行い、カメラ切り替え視聴を可能としている。

2.4 カメラ設置

2.4.1 StageView モードのカメラ設置

講演会場と 4 K カメラ設置の様子を図 2 に示すように、ステージ全体 (講演者、説明図画面) がカメラ 1 台の画角に入るようにステージ中央後方から撮影した。カメラワークは固定とした。

2.4.2 WalkingView モードのカメラ設置

WebHD カメラ設置の様子を図 5 に示す。多角度から視聴できるように 1 箇所カメラ 4 台を設置し、また展示会場全体の様子が分かるように俯瞰撮影 (WebHD カメラを上部に設置) した。

2.4.3 カメラ MAP

図 6 に示すように StageView モードは講演会場に 1 箇所、WalkingView モードはイベント会場 4 箇所の計 10 箇所にカメラ (カメラ 4 台 × 10 箇所 = 計 40) を設置した。

2.5 イベント会場での視聴の様子

イベント会場での視聴の様子を図 7 に示す。図 6 のカメラ MAP に示すパティオ (BF1) の各テーブルにタブレット

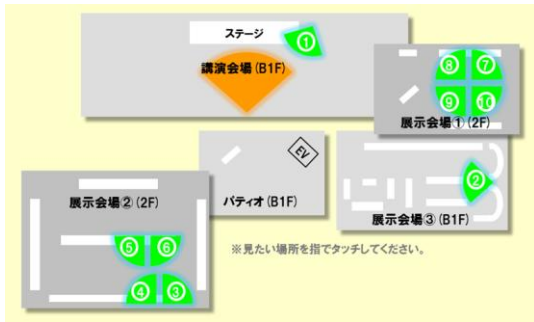


図 6 カメラ MAP



図 8 StageView モードの再生画面



図 7 イベント会場での視聴の様子

端末を設置し、ライブ講演会やイベント会場を自由に視聴できるようにした。

3. イベント会場における評価実験

本システムのコンテンツの適用範囲とユーザインタフェースの有効性を確認するために、実際のイベント会場において主観評価を行った。被験者は、14人である。

3.1 主観評価方法

デモの説明をしながら質問に回答して貰った。回答方法は選択形式とし、選択肢は各自の主観に左右されないように曖昧性のない表現にし4択とした。展示説明をしながらの質問のため、質問にかかる時間を考慮し質問項目は4つとした。

3.2 StageView モードの主観評価

カメラ1台、カメラワーク固定による撮影で講演会が拡大視聴のコンテンツとして適しているか明確にするために、イベント会場でシステムの説明をしながら2つ質問した。図8のStageViewモードの再生画面に示すように、画面全体の中に講演者、説明画面が入り、見たい領域を自由に移動して選択、拡大の視聴を体験してもらった。

3.2.1 StageView モードの主観評価結果と考察

No. 1の質問の「映像の中を自由に移動して見たいところを選択、拡大視聴は良いと思いますか?」の質問に対して、図9に示すように、86%が「十分良いと思う」と回答した。「ほぼ良い」が14%となり、合わせて100%が見たいところの選択、拡大視聴を認める結果となった。

これは、ライブ映像でステージ全体を見渡せ、講演者の

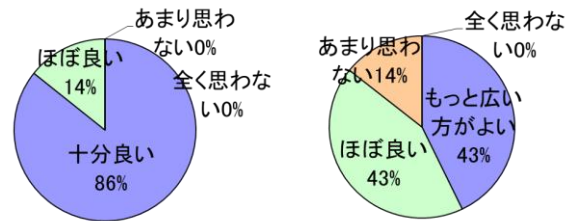


図 9 質問 1 の回答

図 10 質問 2 の回答

声を聴きながら、自分のタイミングで見たいところ(説明図画面や講演者)を選択、拡大視聴できるのが受け入れられたためと考えられる。

また、無駄がない(図面を拡大、声が聞こえる)、会場の雰囲気が伝わってくる、プレゼンやセミナーに良い、求める場所が見えるので良い、スポーツにも良い等の意見を得ることができた。

No. 2の質問の「移動できる横の範囲はもっと広い方が良いと思いますか?」の質問に対しては、図10に示すように、もっと広い方が良いと「全く思わない」、「あまり思わない」、「ほぼこれで良い」を合わせて57%となり、約6割の人が移動範囲はこの範囲で良いと回答した。

以上のことから、出演者が少なく、移動範囲も少ない講演会等において本システムが十分適用可能と考えられる。

しかし、「もっと広い方が良いと思う」が43%になった。理由として、講演会はこの程度で良いが人数が多い場合や演劇の舞台、スポーツ等をもっと広い方が良い等の意見が得られた。これは、スポーツ等の人数や移動量が多い場合では、カメラ1台の画角から外れてしまうためと考えられる。人数、移動量、構図によっては4Kカメラ1台では不十分と考えられる。

3.3 WalkingView モードの主観評価

カメラ切り替えにカメラ設置位置を示すMAP上からユーザが直接カメラ選択か、または順送りにカメラを切り替えて視聴する方法の有効性について、イベント会場で視聴の説明をしながら2つ質問した。タブレット端末上に表示される図6と同じカメラMAP上から視聴したいカメラを選択(画面を押す)してもらった。図11に、WalkingViewモードでのイベント会場の再生画面を示す。



図 11 WalkingView モードの再生画面

順送りのカメラ切り替えは、図 1 1 の再生画面の左右淵の<>箇所を押すことで、順送り（先に進む）、逆送り（もどる）再生が可能となっている。実際に<>を押して体験してもらった。

3.3.1 WalkingView モードの主観評価結果と考察

№. 3 の質問の「カメラ MAP から直接カメラ切り替え視聴は良いと思いますか？」の質問に対して、

図 1 2 に示すように、“十分良い”、“ほぼ良いと思う”合わせて 43%，“あまり思わない”、“全く思わない”合わせて 57% となり、約 6 割が否定的な回答になった。これは、サムネイルのような映像一覧から直接選びたい、カメラ MAP から選択してもどこを見ているのかわからない等の意見が得られ、カメラ MAP でカメラ選択は分かり易いが選択後に映像が分かるもどかしさ、全体の中のどこを再生しているのか知りたい欲求のためと考えられる。

№. 4 の質問の「指定された順送りでのカメラ切り替え視聴は良いと思いますか？」の質問に対して、

図 1 3 に示すように、“十分良いと思う”、“ほぼ良いと思う”合わせて 28% となった。一定時間で順番に切り替わると良いとの意見が得られた。

しかし、“あまり思わない”、“全く思わない”合わせて 72% となった。これは、順送り再生は今どこを再生しているのか分からない、視聴カメラ映像とカメラ MAP とが同時に表示されると良い、カメラ MAP 上に映像が表示されていて選択できると良い等の意見が得られ、質問 3 と同様、全体の中で今どこを再生しているのか知りたい欲求があるためと考えられる。

以上のことから、視聴カメラ映像とどのカメラを視聴しているのか分かるサムネイルやカメラ MAP が同時に表示される必要があることを確認できた。

また過去に、私たちが開発した別システム「マルチアングル映像ライブ配信システム」でのカメラワークがあるコンテンツのカメラ切り替えインターフェース評価実験を行ったことがある。このコンテンツのカメラワークは、カメラ 3 台で出演者 3 人をそれぞれに追いかける映像である。

このカメラワークがあるコンテンツでの評価結果は、“順送り操作したいと思う”が 45% となった。

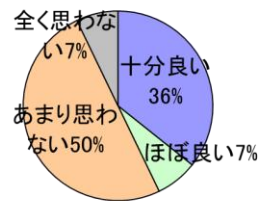


図 12 質問 3 の回答

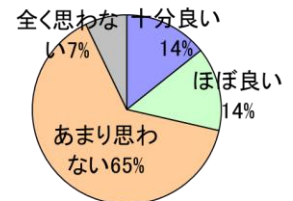


図 13 質問 4 の回答

理由として、カメラ台数は 3 台と少ないため、各カメラ映像を覚えてしまい何が再生されるか分かるのでカメラ台数が少ない場合には良い、

また、このシステムの他の機能としてボタン操作（リモコン操作）でカメラ番号直接選択があり、それと比較して画面を見たまま操作できるのが良い等が得られている。

以上のことから、「マルチアングル映像ライブ配信システム」評価では、カメラ台数の少ない場合やボタン操作と比較した場合に、約半数近くが“順送り操作したいと思う”を確認できた。

しかし、「マルチアングル映像ライブ配信システム」評価でも、“順送り操作したいと思わない”25%，“どちらとも言えない”30%を合わせると半数以上となった。

思わない理由として、どのカメラに何が映っているのか不安になる、分かりづらい、面倒くさい、直接選べた方が良い等の意見があり、

どちらともいえない理由は、場合によるカメラ台数が多いと不便、全部で何チャンネルあるかによる、3 視点くらいだったら見たいと思う等の意見が得られた。

以上の結果からみても、順送りカメラ選択が効果的でない結果が得られており、カメラワークの有無にかかわらず、カメラ切り替え視聴に順送りインターフェースは適切ではないと考えられる。

4. あとがき

本稿では、コンテンツの適用範囲、ユーザインターフェースの有効性を確認するために実際のイベント会場において主観評価を行い、StageView モードでは講演者の声を聴きながら説明図画面や講演者を選択、拡大視聴できるインタラクティブ・ライブ映像配信が受け入れられる評価が得られ、カメラ 1 台、カメラワーク固定での好きな部分を選択、拡大視聴のコンテンツ適用範囲を確認できた。

WalkingView モードではイベント会場の各地移動のためには、サムネイルとカメラ MAP を表示してユーザの選択を補助するインターフェースの重要性を確認できた。

コンテンツの適用範囲、ユーザインターフェースについての評価を行ったが、今後は、講演会の様な人が少なく、移動範囲も少ない場合に、どこの領域を視聴しているかの評価を行う予定である。また、見たい領域（高解像度映像）の配信要求から再生までの時間短縮、画質の向上の検討も

行ってきたい。

参考文献

- 1) 技術ジャーナル 下村道夫, 他 “研究開発用テストベッドネットワーク「GEMnet2」” NTT 技術ジャーナル 2012.8
<http://www.ntt.co.jp/journal/1208/files/jn201208042.pdf>
- 2) Hideaki Kimata, Daisuke Ochi, Akio Kameda, Hajime Noto, Katsuhiko Fukazawa, and Akira Kojima,
"Mobile and Multi-device Interactive Panorama Video Distribution System," IEEE GCCE 2012, Oct., 2012.
- 3) 山口好江, 深澤勝彦, 木全英明, 小島明 “インタラクティブ 3Dマルチアングル映像ライブ配信の一検討” 電子情報通信学会総合大会講演論文集 2012年_基礎・境界, pp.250-250, Mar.(2012)