

カメラ間通信を用いた 無線マルチビューストリーミングの検討

小寺 志保^{1,a)} 藤橋 卓也^{3,b)} 猿渡 俊介^{2,c)} 渡辺 尚^{3,d)}

概要: 本稿では、マルチビュービデオを低トラフィック、高品質で無線伝送できる方式として Domino Streaming を提案する。Domino Streaming では、他の撮影機器がアクセスポイントに送信している映像を、近隣の撮影機器が傍受して自身の映像との差分のみをエンコードすることでトラフィック削減を達成している。また、一方向の撮影機器映像だけでなく、自身の前後または左右にある撮影機器の映像を利用できるように、各撮影機器の送信順を制御することで、より多くのトラフィック削減を実現する。MERL が提供しているビデオシーケンスを用いた評価から、Domino Streaming は各撮影機器が単純に映像を送信する場合と比較して、同じ PSNR で最大 52 % のトラフィック削減を達成することが分かった。

キーワード: マルチビュービデオ, カメラ間通信, 無線通信

A Study on Using Inter-Camera Communication on Wireless Multi-View Video Streaming

Abstract: Multi-view video consists of multiple video sequences captured simultaneously by multiple closely located cameras. Multi-view video enables users to freely change their viewpoints. Typical applications of multi-view video are video conference systems, free viewpoint TV (FTV), and 3D TV. This technical report discusses the multi-view video streaming over wireless networks.

Keywords: Multi-View Video, Inter-Camera Communication, Wireless Communication

1. はじめに

立体映像, 自由視点映像 [1, 2] に代表される超臨場感映像の発展は, 話し手と聞き手の間にある物理的距離を限りなくゼロに近づける。例えば, テレプレゼンス [3, 4] やオリンピック種目の自由視点映像配信などの超臨場感映像を用いたアプリケーションは, 人々に face-to-face による円滑なコミュニケーション, 会場さながらの熱狂を味わうことができるエンターテインメントを提供する。

マルチビュービデオは, 超臨場感映像の発展を担う一技術として注目を浴びている。マルチビュービデオは, ある被写体を

近隣に設置した複数台のカメラで同時に撮影し, 撮影した複数の映像を視聴者に提供する技術である。

図 1 にマルチビュービデオを撮影してから, 視聴者が映像を視聴するまでの構造を示す。マルチビュービデオシステムの構造は撮影部, 転送部, 表示部の 3 種類に分けることができる。撮影部では, 被写体を複数台のカメラを用いて同時に撮影し, エンコーダまで撮影した映像を送信する。転送部では, 撮影部から受信した映像をエンコードし, 視聴者のデコーダまでエンコードした映像を送信する。表示部では, デコーダに届いた映像をデコードし, 表示機器上に映像を表示する。

転送部の代表的な研究として, H.264/AVC によって標準化された Multi-view Video Coding (MVC) [5–7], YouTube など蓄積型配信に対応した Interactive Multi-view Video Streaming (IMVS) [8–12], ライブ配信に対応した User Dependent Multi-view Video Transmission (UDMVT) [13–15], 複数の視聴者への配信に対応した User dependent Multi-view video Streaming for Multi-user (UMSM) [16–18] が挙げられる。表示部の代表的

¹ 静岡大学情報学部

² 静岡大学大学院情報学研究所

³ 大阪大学情報学研究所

a) koder@aurum.cs.inf.shizuoka.ac.jp

b) fujihashi.takuya@ist.osaka-u.ac.jp

c) saru@inf.shizuoka.ac.jp

d) watanabe@ist.osaka-u.ac.jp

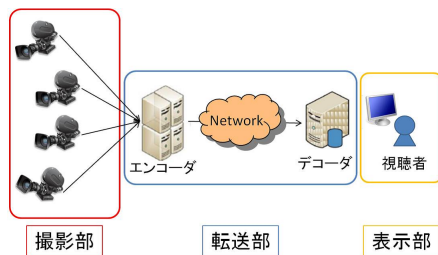


図 1 マルチビュービデオシステムの構成

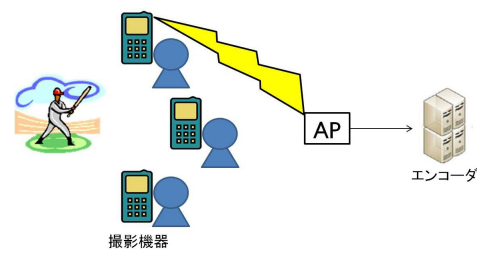


図 2 無線化した撮影部

な技術には、インテグラル立体テレビ [19], 自由視点テレビ [2] がある。

一方、マルチビュービデオの撮影部には従来、マルチカメラアレー [20] が用いられてきた。マルチカメラアレーの例としては、直線上にカメラを配置した直線カメラアレー、円形にカメラを配置したカメラアレー、直線カメラアレーを動かす移動カメラアレーなどが挙げられる。しかしながら、マルチカメラアレーは有線ネットワークで繋がれた複数台のカメラを利用することを想定しているため、カメラを設置できる環境が限定的になり、撮影可能な被写体が限られてしまう。

本研究では、デジタルカメラやスマートフォンといった小型撮影機器と無線通信を組み合わせることで、場所を選ぶことなくマルチビュービデオが撮影できるシステムを考える。撮影部に無線通信を用いることで、より多くの場面でマルチビュービデオが利用可能となる。撮影部を無線化するためには、2つの要件を満たす必要がある。1つ目の要件は、トラフィックの削減である。これはマルチビュービデオが複数台のカメラの映像から構成されることに加えて、無線通信が有線通信よりデータレートが低いためである。2つ目の要件は、映像品質の維持である。2つの要件を同時に満たさなければ、アプリケーションに対する視聴者満足度が低下する。

本稿では、トラフィックの削減と映像品質維持の2つの要件を満たして、撮影部の無線化を実現する方式として、Domino Streaming を提案する。Domino Streaming では、ある撮影機器は他の撮影機器が送信しているビデオを傍受し、傍受した映像と自身の映像との差分をエンコードすることでトラフィックを削減する。また撮影機器間にある冗長な情報をより多く削減するため、アクセスポイント (AP) によって送信順序を制御する。JMVC エンコーダと MERL が提供しているビデオシーケンスを利用した評価結果から、Domino Streaming は、他の撮影機器の映像を利用しない場合と比較して、同じ映像品質を維持したまま、最大 52 % のトラフィック削減を達成していることが分かった。

本稿の構成は以下のとおりである。2章では、撮影部の無線化を実現するための要件と、既存研究について述べる。3章では、提案方式である Domino Streaming について述べる。4章では、JMVC エンコーダにより、単純な伝送方式と提案方式である Domino Streaming を比較する。最後に5章でまとめとする。

2. 関連研究

本研究では、マルチビュービデオをより多くの場面で利用できるようにするため、撮影部を無線化したマルチビュービデオシステムの実現を目指す。撮影部を無線化することで、スタジオなどの屋内の限られた場所だけでなく、より多くの被写体をマルチビュービデオで提供することが可能となる。例えば、屋外において被写体を小型撮影機器で撮影し、マルチビュービデオとして提供できるようになる。

図2に無線化した撮影部のモデルを示す。AP間は無線によって接続され、APからエンコーダまでは有線によって接続されている。各撮影機器は自身が撮影した映像をAPまで伝送し、APは複数の撮影機器から受信した映像をエンコーダへ伝送する。

無線化したマルチビュービデオの撮影部を想定した場合、視聴者が、実際の被写体に限りなく近い映像を不自然に途切れることなく視聴できることが必要となる。そのためには、次の2つの要件を同時に満たす必要がある。

1つ目の要件は、トラフィックを低くすることで、伝送遅延による視聴者満足度の低下を防ぐことである。マルチビュービデオは複数のカメラからの映像を扱うため、シングルビデオよりトラフィックが高くなる。また無線通信は帯域幅が狭く、干渉などの影響もあるため、有線通信に比べてデータレートが低い。単純にマルチビュービデオを無線通信で伝送した場合、映像を撮影してから視聴者へ映像を伝送するまでの遅延が大きくなり、視聴者満足度の低下を招く。低下の要因として、視聴者が視点を切り替えたときに、映像の停止やスキップが発生し、視覚的に不自然な映像になることがある。

2つ目の要件は、映像品質を高く維持することである。映像品質とは、元の映像とデコードした映像との劣化の度合いを表す。品質が高い映像とは、劣化が小さく、再現率の高い映像を意味する。視聴者は、実際の被写体と同等の映像を求めため、映像品質の低下は視聴者満足度の低下を招く。

最も単純に撮影部の無線化を実現するための方法として、各撮影機器が無線通信で映像を個別に送信する方法がある。しかしながら、各撮影機器が送信する映像は高いトラフィックを持っているため、視聴者への伝送に時間がかかり、視聴者満足度の低下を招く。

トラフィックを削減する単純な方法として、マルチビュービデオの量子化パラメータ (QP) を上げる方法やフレームレートを落とす方法がある。しかしながら、映像品質が劣化し、視聴者満足度の低下を招く。

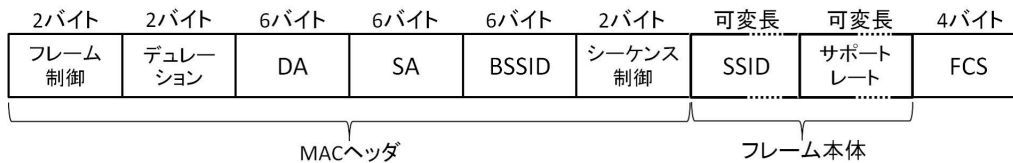


図 3 プローブ要求フレーム

表 1 前提条件

層	プロトコル
トランスポート	UDP
ネットワーク	IP
MAC	CSMA/CA
物理	802.11n

1つ目の要件であるトラヒックの削減と2つ目の要件である映像品質の維持はトレードオフの関係にある。例えば、QPを上げた場合、マルチビュービデオをエンコードするとき、量子化によって多くの映像データが捨てられるため、トラヒック量が減少する。一方で、マルチビュービデオをデコードするとき、元の映像データを復元しにくくなるため、映像品質が劣化する。

無線通信を介してトラヒックの削減を達成する方法として、DMVC [21-23] が挙げられる。DMVCでは各撮影機器が独立してSlepian-Wolf理論 [24] とWyner and Ziv理論 [25] を組み合わせた方法で自身の映像をエンコードする。その後、各撮影機器はエンコードした映像と共に、撮影機器間の情報について記載したside informationをAPへ送信する。side informationには撮影機器の設置位置や、撮影しているアングルなどが含まれる。APは各撮影機器から受信した各々の映像とside informationを有線通信を通してエンコーダに転送し、エンコーダでは各撮影機器から受信した映像とside informationを利用して、一旦映像のデコードを開始する。デコードを終えると、視聴者へ伝送するために再度、映像のエンコードを開始する。

DMVCは提案手法Domino Streamingとエンコード遅延とデコード遅延の観点から相補的な関係にある。Domino Streamingでは、各撮影機器は他の撮影機器の映像を元に、自身の映像との差分情報を算出してエンコードするため、DMVCと比較してエンコード遅延が大きくなる。一方、DMVCでは受信側が各撮影端末から受信した映像とside informationを利用した複雑なデコード処理をする必要があるため、Domino Streamingと比較してデコード遅延が大きくなる。

3. Domino Streaming

2章での議論を基に、無線化したマルチビュービデオ撮影部において低トラヒックと映像品質の維持を達成するDomino Streamingを設計した。表1にDomino Streamingが各層で用いるプロトコルについて示す。物理層は802.11nを想定し、MAC層はCSMA/CAを用いる。ネットワーク層はIP、トランスポート層はUDPを用いる。

3.1 全体像

Domino Streamingは以下の流れで動作する。

- (1) Domino Streamingに参加する撮影機器は、APの通信範囲内に入ると初期化を開始し、APによって送信順が割り当てられる。初期化の詳細については3.2節で述べる。
- (2) 各撮影機器の初期化が終了すると、APは初期化中に取得した各撮影機器の位置情報から映像送信順を決定し、撮影機器に決定した送信順をブロードキャストする。送信順決定の詳細については3.3節で述べる。
- (3) 各撮影機器はAPが決定した送信順に従って、自身の映像をすでに傍受した他の撮影機器の映像を用いて一定量エンコードする。エンコードの詳細については3.4節で述べる。
- (4) 各撮影機器はAPが決定した送信順に従ってエンコードした自身の映像をAPへ伝送する。また他の撮影機器は通信を傍受することで、送信中の撮影機器の映像を受信する。全ての撮影機器が映像を送信し終わると、APは再び各撮影機器の位置情報に基づいて、映像送信順序を決定する。映像伝送の詳細については3.5節で述べる。
- (5) 各撮影機器とAPから映像を受信しているエンコーダは、新たに撮影機器の映像を受信するとデコードを開始する。デコードの詳細については3.6節で述べる。

3.2 初期化

Domino Streamingでは、各撮影機器は映像を伝送する前に、APによって一意のIDを割り当てられる。APは各撮影機器にIDを割り当てるため、APの通信範囲に入ったことを撮影機器に通知するビーコンを周期的に送信する。APの通信範囲に入った撮影機器はビーコンを受信すると、IDを取得するために、サポートレートフィールドに位置情報を含んだプローブ要求フレームをAPへ送信する。図3にプローブ要求フレームのフォーマット [26] を示す。サポートレートフィールドは通常、撮影機器がサポートしているデータレートを通知するために利用する。Domino Streamingでは、撮影機器がサポートする全てのデータレートを記述し、続けて自身の位置情報を8バイト分付加することで、位置情報をAPへ伝達する。また位置情報には、撮影機器に搭載されたGPS情報を利用する。

APは撮影機器からプローブ要求フレームを受信すると、フレームを送信した撮影機器へID割り当てフレームを送信する。ID割り当てフレームのフォーマットは2バイトのIDフィールドで構成される。IDフィールドには、APが各撮影機器に割り当てた一意のIDが格納される。

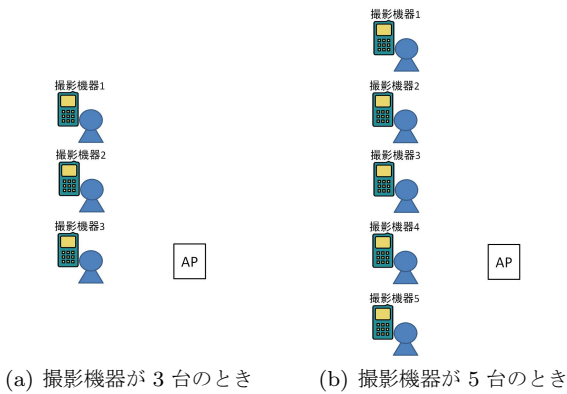


図 4 撮影機器の配置例

3.3 送信順決定

AP は撮影機器間にある冗長な情報を効率的に削除するため、各撮影機器の位置情報を元に、各撮影機器の映像送信順を決定し、全ての撮影機器に決定した送信順をブロードキャストする。送信順序は H.264/AVC における双方向予測を利用できるように設定される。双方向予測とは、ある撮影機器が自身の映像をエンコードするとき、自身の前後または左右の撮影機器が撮影した映像を利用することで、より大幅なトラフィック削減を達成する技術である [5-7]。

以下に、AP の通信範囲内に N 台の撮影機器が存在する場合の送信順序決定処理を示す。また各撮影機器は 1 列に並んでおり、3.2 節の方法に従って初期化されているとする。

- (1) AP は自身の通信範囲内で最も自身からの直線距離が長い撮影機器 A の送信順を 1 番目に決定する。また AP は双方向予測を考慮した送信順を決定するため、前方向の撮影機器として撮影機器 A を設定する。
- (2) AP は自身の通信範囲内でまだ送信順を決定していない撮影機器が存在するか確認する。
 - (a) 送信順を決定していない撮影機器が 2 台以上存在する場合は、AP は前方向の撮影機器から最も直線距離が短い撮影機器 B と、前方向の撮影機器を除いて、撮影機器 B から最も直線距離が短い撮影機器 C を検出する。撮影機器 C は撮影機器 B の映像を双方向予測する場合、後方向の撮影機器となる。AP はすでに決定している送信順の末尾に、撮影機器 C、撮影機器 B の順に送信順を追加することで、双方向予測を考慮した送信順を決定する。その後、AP は新たな前方向の撮影機器として撮影機器 C を設定し、再び (2) を開始する。
 - (b) 送信順を決定していない撮影機器が 1 台存在する場合は、すでに決定している送信順の末尾に、まだ送信順を決定していない撮影機器の ID を追加する。その後、送信順決定処理を終了する。
 - (c) 送信順を決定していない撮影機器が存在しない場合は、全ての撮影機器に送信順を決定したとして、送信順決定処理を終了する。

送信順決定処理が終了すると、AP は自身が決定した送信順を

送信順通知フレームを用いて全ての撮影機器にブロードキャストする。送信順通知フレームのフォーマットは 32 バイトの送信順フィールドで構成される。送信順フィールドには AP が決定した送信順が格納される。

図 4 を用いて、AP の通信範囲内に 3 台、5 台の撮影機器が存在する場合の送信順決定処理の具体例を示す。例えば図 4(a) のように、3 台の撮影機器が配置されているとする。まず、AP は自身の通信範囲内で、最も自身からの直線距離が長い撮影機器 1 の送信順を 1 番目とし、前方向の撮影機器として撮影機器 1 を設定する。次に、AP の通信範囲でまだ送信順が決定していない撮影機器数が 2 台以上であるため、AP は前方向の撮影機器である撮影機器 1 から最も直線距離が短い撮影機器 2 と、撮影機器 2 から最も直線距離が短い撮影機器 3 を検出する。その後、AP はすでに決定している送信順の末尾に撮影機器 3、撮影機器 2 の順に送信順を追加する。AP は新たな前方向の撮影機器として撮影機器 3 を設定するが、すでに全ての撮影機器に対して送信順を決定しているため、送信順決定処理を終了する。最終的に、AP が決定した送信順は撮影機器 1, 3, 2 となる。

図 4(b) のように、5 台の撮影機器が配置されているとする。撮影機器 1 から 3 までの送信順は、撮影機器が 3 台のときと同じ方法で決定する。撮影機器 1, 2, 3 の送信順が決定したときには、新たな前方向の撮影機器として撮影機器 3 が設定されている。まだ送信順が決定していない撮影機器数が 2 台以上であるため、AP は前方向の撮影機器である撮影機器 3 から最も直線距離が短い撮影機器 4 と、撮影機器 4 から最も直線距離が短い撮影機器 5 を検出する。その後、AP はすでに決定している送信順の末尾に撮影機器 5、撮影機器 4 の順に送信順を追加する。AP は新たな前方向の撮影機器として撮影機器 5 を設定するが、すでに全ての撮影機器に対して送信順を決定しているため、送信順決定処理を終了する。最終的に、AP が決定した送信順は撮影機器 1, 3, 2, 5, 4 となる。

3.4 エンコード

送信順が決定すると、各撮影機器は送信順に従って自身の映像をエンコードする。各撮影機器は H.264/AVC に基づき、GOP (Group Of Picture) ごとに自身の映像をエンコードする。GOP とは、複数のフレームの集合体のことであり、通常は 8 枚のフレームから構成される。Domino Streaming では、撮影機器は自身の映像を低トラフィックで送信するために、自身に割り当てられた送信順より前の通信を全て傍受する。

図 5 に GOP のエンコードを示す。ここで撮影機器の台数、位置関係、送信順は図 4(a) と同じであることを想定する。図 5(a) に、撮影機器 1 の予測構造を示す。撮影機器 1 は単独でエンコードするため、先頭のフレームが I フレームとなる。I フレームは、JPEG によって符号化された静止画情報であり、データ量が大きい。

図 5(b) に、撮影機器 3 の予測構造を示す。撮影機器 3 は、撮影機器 1 の映像を傍受しているため、撮影機器 1 と自身の映像との差分情報をエンコードする。具体的には、撮影機器 3 の先頭フレームを P フレームとしてエンコードする。P フレームは

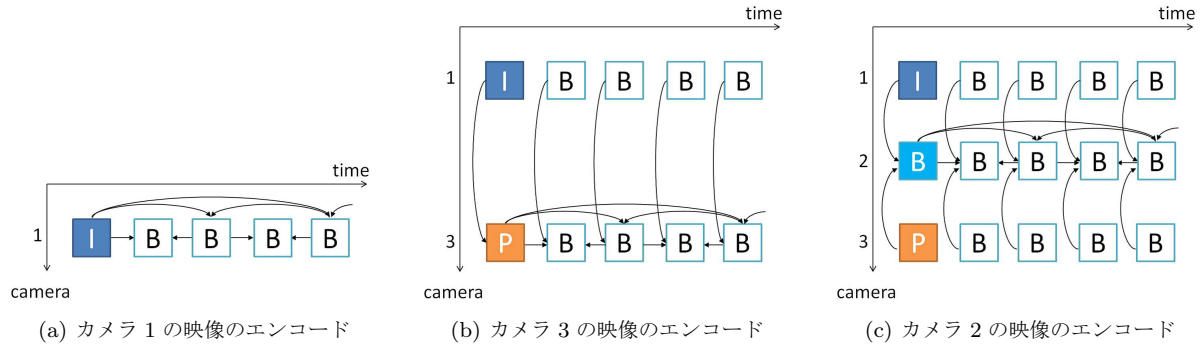


図 5 エンコード方法

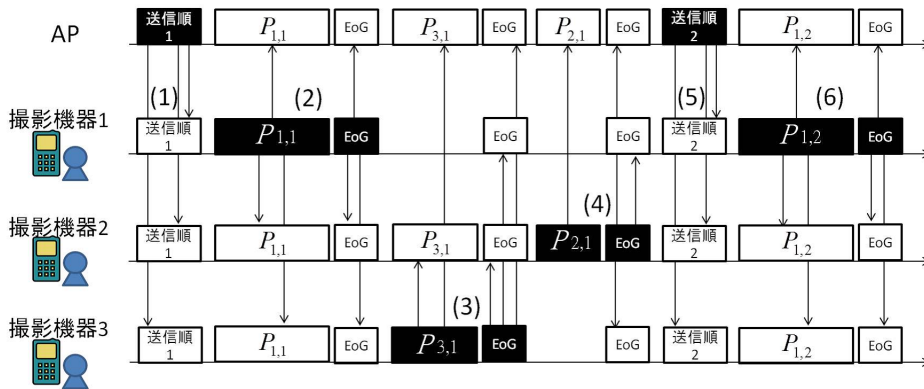


図 6 映像伝送タイムシーケンスチャート

表 2 ビデオパケット $P_{i,j}$

フィールド	バイト
カメラ位置	8
映像	可変長

他の撮影機器の I フレームを基に作成した自身の映像との差分情報であるため、I フレームよりデータ量が少なくなる。

図 5(c) に撮影機器 2 の予測構造を示す。撮影機器 2 は、撮影機器 1 と 3 の映像を傍受しているため、撮影機器 1, 3 の映像と自身の映像との差分情報をエンコードする。具体的には、撮影機器 2 の先頭フレームを B フレームとしてエンコードする。B フレームは複数の撮影機器の映像を基に作成した自身の映像との差分情報であるため、最もデータ量が少なくなる。

3.5 映像伝送

各撮影機器は AP が割り当てた送信順に従って、3.4 章に示した方法でエンコードした映像を送信する。各撮影機器は 1 GOP ずつ自身の映像をアクセスポイントへ送信する。図 6 に映像伝送のタイムシーケンスチャートを示す。図 6 では AP が撮影機器 1, 3, 2 の順に送信順を決定したと仮定する。また $P_{i,j}$ は、撮影機器 i の GOP j のビデオフレームから構成されるパケットを表している。

表 2 に $P_{i,j}$ のフォーマットを示す。カメラ位置フィールドに

は、GPS によって取得した撮影機器の位置情報を格納する。映像フィールドには、撮影機器 i がエンコードした GOP j の映像を格納する。

- (1) AP は 3.3 節の方法に従って決定した GOP1 に対する各撮影機器の送信順を送信順通知フレームに格納し、全ての撮影機器にブロードキャストする。
- (2) AP から受信した送信順に基づいて、撮影機器 1 から映像の伝送を開始する。撮影機器 1 は自身の位置情報をカメラ位置フィールド、エンコードした自身の映像を映像フィールドに格納し、 $P_{1,1}$ を AP へ送信する。撮影機器 2 と 3 は $P_{1,1}$ を傍受して撮影機器 1 のデコードを開始する。撮影機器 1 は $P_{1,1}$ を送信し終わると、続けて EoG (End of GOP) パケットを送信する。EoG パケットは、1 GOP 分の映像の送信が終了したことを他の撮影機器に通知するために利用され、IEEE 802.11 における ACK フレーム [26] と同じフォーマットを利用する。EoG パケットを傍受した撮影機器 3 は、次が自身の送信順だと判断し、デコードが終了した撮影機器 1 の映像を利用して自身の映像のエンコードを開始する。一方、撮影機器 2 は、次はまだ自身の送信順ではないと判断し、デコードが終了した撮影機器 1 の映像を保持し、送信機器 3 の通信が終了するまで待機する。
- (3) 撮影機器 3 は、自身の位置情報をカメラ位置フィールド、エンコードした自身の映像を映像フィールドに格納し、AP へ $P_{3,1}$ を送信する。撮影機器 2 は $P_{3,1}$ を傍受することで

撮影機器 3 の映像と位置情報を取得する。その後、撮影機器 2 は受信した映像のデコードを開始する。撮影機器 3 は $P_{3,1}$ の送信が終了すると、EoG パケットを AP へ送信する。EoG パケットを傍受した撮影機器 2 は、次が自身の送信順だと判断し、デコードが終了した撮影機器 3 の映像と保存していた撮影機器 1 の映像を利用して自身の映像をエンコードする。

- (4) 撮影機器 2 は、自身の位置情報をカメラ位置フィールド、エンコードした自身の映像を映像フィールドに格納し、AP へ $P_{2,1}$ を伝送する。撮影機器 2 は $P_{2,1}$ の送信が終了すると、EoG パケットを AP へ送信する。
- (5) AP は受信した $P_{1,1}$, $P_{2,1}$, $P_{3,1}$ を有線通信によってそのままエンコーダへ転送するのと同時に、 $P_{1,1}$, $P_{2,1}$, $P_{3,1}$ に含まれている各撮影機器の位置情報から、3.3 節の方法に従って、GOP2 に対する映像伝送順を決定する。その後、決定した送信順を送信順通知フレームに格納し、全ての撮影機器にブロードキャストする。
- (6) AP から受信した送信順に基づいて、撮影機器 1 は自身の位置情報をカメラ位置フィールド、エンコードした自身の映像を映像フィールドに格納し、AP へ $P_{1,2}$ を送信する。その後、撮影機器 2, 3 は受信した映像のデコードを開始する。

Domino Streaming ではすべての GOP の送信が終了するまで、以上の動作を繰り返す。

3.6 デコード

Domino streaming のデコードには特別な処理を必要としない。撮影機器と AP から各撮影機器の映像を受信しているエンコーダは標準的な H.264/AVC デコーダを利用して、それぞれ受信した映像をデコードする。各撮影機器とエンコーダが最初に受信する映像は、先頭フレームが静止画である I フレームによってエンコードされているため、撮影機器とエンコーダは映像を受信すると同時にデコードを開始する。1 番目に送信された撮影機器の映像以降、撮影機器とエンコーダが受信する映像は、すでに受信した映像を元に作成された差分映像であるため、撮影機器とエンコーダはすでに受信した映像のデコードが終了すると同時に、新たに受信した映像のデコードを開始する。エンコーダでは全ての撮影機器の映像をデコードすると、転送部で利用する方式に従って全ての撮影機器の映像を再びエンコードする。最後に、エンコードした映像を視聴者へ伝送することで、視聴者は複数台の撮影機器が撮影したマルチビュービデオを視聴することができる。

4. 性能評価

4.1 評価環境

Domino Streaming の有効性を確認するために、JMVC エンコーダ [27] と MERL が提供しているテストビデオシーケンス [28] によって映像品質に対するトラフィック量、異なるビデオシーケンスを利用した場合のトラフィック量、撮影機器の位置を変えた場合のトラフィック量を測定した。Domino Streaming の

表 3 評価パラメータ

解像度	176 × 144
フレームレート	15 fps
フレーム数	250
GOP サイズ	8 フレーム
カメラ数	8 台
量子化パラメータ (QP)	24~40

性能を相対的に評価するために、次の 3 つのアプローチを比較した。

(1) Independent Streaming

Independent Streaming は、各撮影機器が無線通信で自身の映像を AP へ個別に送信する方式である。Independent Streaming は、他の撮影機器の通信を傍受しないため、Domino Streaming において他の撮影機器の映像を利用することによって得られた性能を示す尺度となる。

(2) Domino Streaming w/o order control

Domino Streaming w/o order control は、各撮影機器が AP によって割り当てられた ID 順に自身の映像を送信する方式である。Domino Streaming w/o order control は撮影機器間で双方向予測によるエンコードをしないため、Domino Streaming において AP による送信順制御によって得られた性能を示す尺度となる。

(3) Domino Streaming

Domino Streaming は、3 章で述べた提案方式である。Domino Streaming は、近隣の撮影機器の映像を利用して自身の映像をエンコードすることで、トラフィックを削減する。また AP による送信順制御によって、より大幅なトラフィック削減を達成する。

共通の評価パラメータを以下に示す。テストビデオシーケンスには、Ballroom, Exit, Vassar を利用する。Ballroom は舞踏会の映像で、動きが多い。Exit は出入り口の映像で、Ballroom ほどではないが動きがある。Vassar は風景の中を車が走る映像で、動きは少ない。撮影機器と AP 間の無線通信路はロスレスを想定している。また各撮影機器間の距離は 19.5 cm とする。

表 3 に JMVC エンコーダのパラメータを示す。フレームレートは 15 fps、各テストビデオシーケンスのフレーム数は 250 フレームである。1 GOP あたりのフレーム数は 8 フレームとし、撮影機器数は 8 台とした。各評価にあたって、量子化パラメータを 24 から 40 まで変化させ、異なる Peak Signal-to-Noise Ratio (PSNR) 時のトラフィック量を取得した。PSNR とは映像品質を表すために一般的に用いられている指標であり次式で表される。

$$PSNR = 20 \log_{10} \left(\frac{MAX}{\sqrt{MSE}} \right)$$

MAX は元画像がとりうる最大画素値、MSE は元映像とデコード後の映像との平均二乗誤差 (Mean Square Error) である。

4.2 映像品質に対するトラフィック量の評価

Domino Streaming の基本性能を評価するために、映像品質

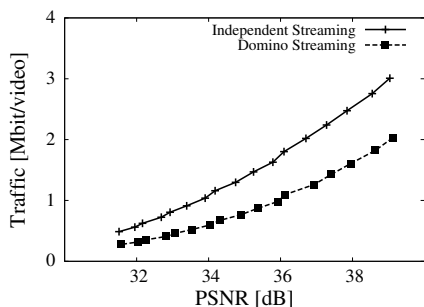


図 7 映像品質に対するトラフィック量

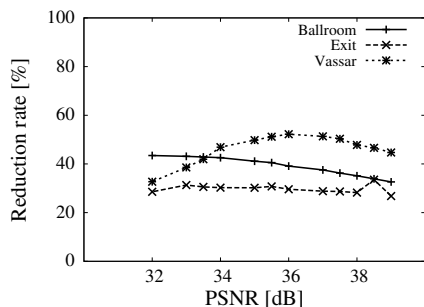


図 8 ビデオシーケンスごとのトラフィック削減率

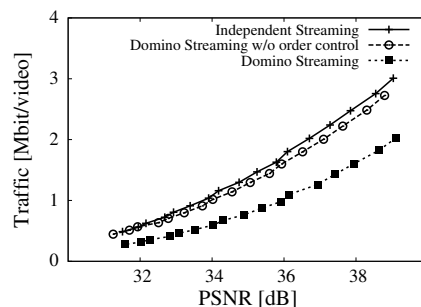


図 9 撮影機器の位置を変えた場合のトラフィック量

を変化させた場合のトラフィック量を評価した。

図 7 に、利用したビデオシーケンスが Ballroom、各撮影機器の位置は固定、量子化パラメータを 24~40 に変えた場合のトラフィック量を示す。横軸は PSNR [dB]、縦軸は 1 映像あたりのトラフィック [Mbits/video] である。

図 7 より、2 つのことが分かる。1 つ目は、Domino Streaming は、Independent Streaming より映像品質を維持したまま、トラフィック削減を達成していることである。例えば、PSNR が 36 [dB] のとき、Domino Streaming は Independent Streaming と比較して、約 700 [Kbits/video] のトラフィックを削減している。Domino Streaming がトラフィックを削減しているのは、他の撮影機器の通信を傍受して得られた映像を利用して自身の映像をエンコードすることで、撮影機器間に存在していた冗長な情報を削減しているためだと考えられる。2 つ目は、PSNR が高くなるにつれて Independent Streaming と Domino Streaming のトラフィック量の差が大きくなっていることである。例えば、PSNR が 32 [dB] のとき Domino Streaming は Independent Streaming と比較して約 240 [Kbits/video] のトラフィックを削減しており、PSNR が 39 [dB] のとき約 980 [Kbits/video] のトラフィックを削減している。PSNR が高くなると、各撮影機器が送信する映像は元映像に近づくため、高いトラフィックが発生する。その結果、Independent Streaming では、PSNR の上昇とともにトラフィック量が急激に増加している。一方、Domino Streaming では、各撮影機器が自身の映像をエンコードするときに、利用する他の撮影機器の映像が元映像に近づくほど、自身の映像との間の冗長な情報が増加する。その結果、2 映像から得られる差分情報が小さくなり、より大幅なトラフィック削減を達成していると考えられる。

4.3 異なるビデオシーケンスを利用した場合のトラフィック量の評価

4.2 節でビデオシーケンスが Ballroom の場合を評価した。しかしながら、マルチビュービデオではアプリケーションによって撮影対象が変化すると考えられる。このような観点から、ビデオシーケンスを変えた場合のトラフィック削減率を評価した。

図 8 に各撮影機器の位置は固定、ビデオシーケンスには Ballroom, Exit, Vassar を利用し、それぞれのビデオシーケンスに対して量子化パラメータを 24~40 に変えた場合のトラフィック削

減率を示す。各ビデオシーケンスの特徴は、4.1 節で述べたとおりである。横軸は PSNR [dB]、縦軸は Independent Streaming に対する Domino Streaming のトラフィック削減率 [%] である。

図 8 から 2 つのことが分かる。1 つ目は、Domino Streaming は映像品質が変化しても、被写体に関係なく Independent Streaming よりトラフィック削減を達成していることである。Vassar におけるトラフィック削減率は、PSNR が 36 [dB] のとき 52 [%] であり、最もトラフィック削減率が大きい。また、各ビデオシーケンスの削減率には、ビデオの特徴に応じて差が生まれているものの、全てのビデオシーケンスにおいて、少なくとも約 20 [%] 以上の削減率を達成している。この結果から、マルチビュービデオの撮影対象が変化しても、Domino Streaming は他の撮影機器の映像を利用して撮影機器間の冗長な情報を削除することで、トラフィック量を削減することができると言える。2 つ目は、PSNR が低い時、Vassar の削減率が急激に低下していることである。例えば、PSNR が 32 [dB] のとき、Independent Streaming のトラフィック量に対する Domino Streaming のトラフィック削減率は 25 [%] であるが、PSNR が 36 [dB] のとき、Independent Streaming のトラフィック量に対する Domino Streaming のトラフィック削減率は 58 [%] である。Vassar では PSNR が低いとき、各撮影機器はエンコード時に映像情報のほとんどを量子化によって捨ててしまう。具体的な値としては、PSNR が 32 [dB] のとき、Independent Streaming における各ビデオシーケンスのトラフィック量は、56 [Kbits/video] (Ballroom), 14 [Kbits/video] (Exit), 7 [Kbits/video] (Vassar) となっており、Vassar のトラフィック量を Ballroom と Exit のトラフィック量と比較すると少ないことが分かる。Domino Streaming では、各撮影機器はほとんどの情報が捨てられた他の撮影機器の映像を利用して、自身の映像をエンコードしようとするが、撮影機器間で冗長だった情報も事前に捨てられてしまっているため、トラフィック削減率が低下したと考えられる。

4.4 撮影機器の位置を変えた場合のトラフィックと映像品質の評価

4.2, 4.3 節で各撮影機器の位置が固定の場合を評価した。しかしながら、無線化した撮影部を利用する場合、マルチビュービデオを撮影中に、各撮影者の移動によって撮影者同士の位置が変わることが考えられる。例えば、ゴルフの試合で、ある選

手やグループを無線化した撮影部を用いて撮影する場合、コースが終了するたびに撮影者は対象を撮影しながら位置を移動する必要があり、移動中に撮影者同士の位置が変わることが考えられる。このような観点から、各撮影機器の位置を変えた場合のトラフィック量を評価した。

図9に、利用したビデオシーケンスが Ballroom、各撮影機器の位置は1ビデオ終了ごとにランダムに変化させ、量子化パラメータを24~40に変えた場合のトラフィック量を示す。本評価では、各撮影機器の位置を変化させて100回結果を取得し、その平均値をトラフィック量として取得した。横軸はPSNR [dB]、縦軸は1映像あたりのトラフィック [Mbits/video] である。

図9から、各撮影機器の位置が変化しても、Domino Streamingが最も低いトラフィック量を達成していることが分かる。具体的な値として、Domino StreamingはDomino Streaming w/o order controlと比較して、最大700 [Kbits/video]のトラフィック削減を達成している。Domino Streaming w/o order controlでは、撮影機器間の位置が変化すると、ある撮影機器が自身の映像をエンコードするときに、自身とは遠く離れた場所にある撮影機器の映像を利用する場合がある。しかしながら、自身の映像と遠く離れた場所にある撮影機器の映像の間にはほとんど冗長な情報が含まれていないため、Domino Streaming w/o order controlはそれほどトラフィックを削減することができなかったと考えられる。一方で、Domino Streamingでは、もし撮影機器の位置が変化しても、撮影機器間の冗長な情報を効率的に削減できるように、APが次のGOP送信の前には各撮影機器の位置情報を元に、新たな送信順を決定する。したがって、各撮影機器の位置が変化する環境下であっても、Domino Streamingに参加する各撮影機器は、自身の映像をエンコードする際、最も多くの冗長な情報を持つ他の撮影機器の映像を利用できるため、Domino Streamingが最も低いトラフィック量を達成できたと考えられる。

5. おわりに

本稿では、マルチビュービデオ撮影部の無線化を実現するための伝送方式として、Domino Streamingを提案した。Domino Streamingでは、各撮影機器は近隣の撮影機器の映像を用いて自身の映像をエンコードすることで、トラフィック削減を達成する。またAPが各撮影機器の送信順を制御することで、より大幅なトラフィック削減を達成する。性能評価から、提案方式Domino Streamingが既存方式と比較して、映像品質の劣化を防ぎながら、トラフィック削減を達成していることが分かった。

謝辞

本研究は電気通信普及財団研究調査助成を受けて行なった。

参考文献

[1] Wojciech, M. and Hanspeter, P.: 3D TV: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes, *ACM Transactions on Graphics*, Vol. 23, No. 3, pp. 814–824 (2004).
[2] Masayuki, T.: Overview of free viewpoint television, *Signal Processing: Image Communication*, Vol. 21, No. 6, pp.

454–461 (2006).
[3] Wei-Chao, W., Herman, T., Lars, N., Greg, W. and Henry, F.: Toward a Compelling Sensation of Telepresence: Demonstrating a portal to a distant (static) office, *IEEE Visualization*, pp. 327–333 (2000).
[4] Ramesh, R., Greg, W., Matt, C., Adam, L., Lev, S. and Henry, F.: The office of the future: A unified approach to image-based modeling and spatially immersive displays, *ACM GRAPHITE*, pp. 179–188 (1998).
[5] Vetro, A., Pandit, P., Kimata, H., Smolic, A. and Wang, Y.-K.: *Joint Draft 8.0 on Multi-view Video Coding* (2008).
[6] K. Muller, P. Merkle, H. Schwarz, T. Hinz, A. Smolic and T. Wiegand: Multi-view video coding based on H. 264/AVC using hierarchical B-frames, *IEEE PCS* (2006).
[7] Text Of ISO/IEC 14496-10:2008/FDAM 1 ISO/IEC JTC1/SC29/WG11: Multiview Video Coding (2008).
[8] Xiaoyu, X., Gene, C. and Jie, L.: Frame structure optimization for interactive multiview video streaming with bounded network delay, *IEEE ICIP*, pp. 593–596 (2011).
[9] Ngai-Man, C., Antonio, O. and Gene, C.: Distributed source coding techniques for interactive multiview video streaming, *IEEE PCS*, pp. 1–4 (2009).
[10] Gene, C., Antonio, O. and Ngai-Man, C.: Interactive streaming of stored multiview video using redundant frame structures, *IEEE Transactions on Image Processing*, Vol. 20, No. 3, pp. 744–761 (2011).
[11] Zhi, L., Gene, C. and Yusheng, J.: Unified distributed source coding frames for interactive multiview video streaming, *IEEE ICC*, pp. 2048–2053 (2012).
[12] Huan, H., Bo, Z., S-HG, C., Gene, C. and Pascal, F.: Coding and replication co-design for interactive multiview video streaming, *IEEE INFOCOM*, pp. 2791–2795 (2012).
[13] Pan, Z., Ikuta, Y., Bandai, M. and Watanabe, T.: User Dependent Scheme for Multi-view Video Transmission, *IEEE ICC* (2011).
[14] Pan, Z., Ikuta, Y., Bandai, M. and Watanabe, T.: A user dependent system for multi-view video transmission, *IEEE AINA*, pp. 732–739 (2011).
[15] Pan, Z., Bandai, M. and Watanabe, T.: A USER DEPENDENT SCHEME FOR MULTI-VIEW VIDEO LIVE STREAMING, *International Journal of Computational Information Systems*, Vol. 9, No. 4, pp. 1439–1448 (2013).
[16] Fujihashi, T., Pan, Z. and Watanabe, T.: Traffic Reduction for Multiple Users in Multi-view Video Streaming, *IEEE ICME* (2012).
[17] Fujihashi, T., Pan, Z. and Watanabe, T.: UMSM: A Traffic Reduction Method on Multi-View Video Streaming for Multiple Users, *IEEE Transactions on Multimedia*, Vol. 16, No. 2, pp. 1–14 (2014).
[18] Fujihashi, T., Pan, Z. and Watanabe, T.: Traffic Reduction on Multi-View Video Live Streaming for Multiple Users, *IEICE Transactions on Communications*, Vol. 96, No. 7, pp. 2034–2045 (2013).
[19] Takayuki, I.: 技研における立体テレビの研究結果 (2010).
[20] 横井孝紀, 福嶋慶繁, 圓道知博, パナヒブル テヘラニメヒルダド, 藤井俊彰, 谷本正幸: 移動カメラアレーを用いた時空間の自由視点画像生成 (研究速報, 映像メディア処理, 特集画像符号化・映像メディア処理レター), 電子情報通信学会論文誌 D 情報・システム, Vol. 94, No. 12, 一般社団法人電子情報通信学会, pp. 1996–1999 (2011).
[21] Guo, X., Lu, Y., Wu, F., Gao, W. and Li, S.: Distributed Multi-view Video Coding, *VCIP*, Vol. 38, No. 11, pp. 1917–1921 (2006).
[22] Xavi, A., Egon, A. and Luis, T.: Side information generation for multiview distributed video coding using a fusion approach, *IEEE NORISIG*, pp. 250–253 (2006).
[23] Frederic, D., Mourad, O. and Touradj, E.: Recent advances in multiview distributed video coding, *DSS*, pp. 1–11 (2007).
[24] Slepian, D. and Wolf, J. K.: Noiseless coding of correlated information sources, *IEEE Transactions on Information Theory*, Vol. 19, pp. 471–480 (1973).
[25] D. Wyner, A. and Ziv, J.: The rate-distortion function for source coding with side information at the decoder, *IEEE Transaction on Information Theory*, Vol. 3, No. 4, pp. 45–49 (1976).
[26] Gast, M. S.: 802.11 無線ネットワーク管理, オライリー・ジャパン (2003).
[27] Joint Video Team Of ITU-T VCEG And ISO/IEC MPEG: *JMVC (Joint Multiview Video Coding) Software* (2008).
[28] ISO/IEC JTC1/SC29/WG11: *Multiview Video Test Sequences from MERL* (2005).