

研究論文

映像の類似性を利用した ワイヤレスマルチビュービデオストリーミング

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

受付日 2016年1月15日, 採録日 2016年5月30日

概要: 複数の撮影機器の映像を同時に視聴者に提供するマルチビュービデオは、自由視点映像など将来の映像技術を支える技術として注目を浴びている。本稿では、スマートフォンなどの小型撮影機器を用いて、場所に制限されることなくマルチビュービデオを撮影できる仕組みとして、Domino Streaming++ を提案する。小型撮影機器を利用した自由なマルチビュービデオの撮影が可能になることで、マルチビュービデオの幅広い利用が可能となる。Domino Streaming++ は、各撮影機器が他の撮影機器の通信を傍受し、傍受した映像をエンコードに利用することで、映像品質を維持したまま、トラフィックの削減を実現する。さらに、映像の類似度を用いて映像の送信順を制御することで、より多くのトラフィックを削減する。MERL が提供している実ビデオシーケンスを用いた評価から、Domino Streaming++ が理想値に近いトラフィック量を達成していることを示す。

キーワード：マルチビュービデオ, カメラ間通信, 無線通信

Video Similarity Based Wireless Multi-view Video Streaming

SHIHO KODERA^{1,a)} TAKUYA FUJIHASHI^{2,b)} SHUNSUKE SARUWATARI^{2,c)} TAKASHI WATANABE^{2,d)}

Received: January 15, 2016, Accepted: May 30, 2016

Abstract: Multi-view video provides 3-D perception and free view navigation for viewers. In this paper, we propose Domino Streaming++ to capture multi-view video using mobile cameras. Domino Streaming++ captures a scene by multiple cameras without a limitation of place. Domino Streaming++ has two features. First, each mobile camera overhears other cameras' video packets, and encodes its own video frames using the overheard video packets. Second, the access point controls the transmission order of mobile cameras based on video similarity. Evaluations using multi-view video sequences show that the traffic volume of Domino Streaming++ approaches that of an ideal method with only a slight degradation in video quality.

Keywords: multi-view video, inter-camera communication, wireless communication

1. はじめに

4K, 8K に代表される超高精細映像の導入により、世界

中の人々が映像で同じ瞬間を共有することが可能となった。たとえば、8K で各オリンピック種目の配信が実現されることで、会場から遠く離れた場所でも会場の様子を細部まで視聴することができ、世界中でオリンピックの様子を共有できるようになる。しかしながら、実際に会場で観戦しているような臨場感あふれる映像を視聴者に提供するためには、超高精細映像だけではなく、実際のユーザの要求を反映した現実性の高い映像配信が必要である。これまでの映像配信は、実際のユーザの要求に関係なく、提供者側が選択したカメラ 1 台の映像を配信しているため、高い現実性を提供することはできない。

¹ 静岡大学大学院情報学研究科, 浜松市
Graduate School of Informatics, Shizuoka University, Hamamatsu, Shizuoka 432-8011, Japan

² 大阪大学大学院情報科学研究科, 吹田市
Graduate School of Information Science and Technology, Osaka University, Suita, Osaka 565-0871, Japan

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

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

c) saru@ist.osaka-u.ac.jp

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

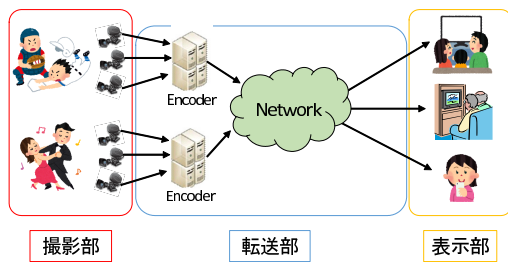


図 1 マルチビュービデオシステムの構成
Fig. 1 Overview of multi-view video system.

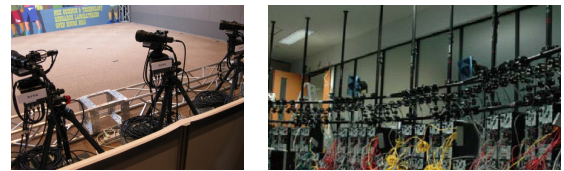
視聴者に現実性の高い臨場感あふれる映像を提供する技術として、立体映像、自由視点映像 [1], [2] がある。立体映像、自由視点映像を用いて、各視聴者の自由な視点選択を実現することで、より現実に近い、臨場感あふれる映像を提供することができる。マルチビュービデオは、映像に臨場感および現実性を与える技術の 1 つであり、超高精細映像に次ぐ、将来の映像技術である。

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、撮影機器が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流れを示す。マルチビュービデオシステムは、撮影部、転送部、表示部の 3 種類から構成される。まず撮影部では、被写体を複数の撮影機器を用いて同時に撮影する。撮影された映像はエンコーダへ伝送される。次に転送部では、撮影部から伝送された映像をエンコーダでエンコードする。エンコードした映像はネットワークを介して視聴者のデコーダまで伝送される。表示部では、転送部から伝送された映像をデコーダでデコードし、各視聴者の表示機器に表示する。

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) [3], [4], [5], YouTube など蓄積型配信に対応した Interactive Multi-view Video Streaming (IMVS) [6], [7], [8], [9], [10], ライブ配信に対応した User Dependent Multi-view Video Transmission (UDMVT) [11], [12], [13], 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) [14], [15], [16] があげられる。

表示部の研究には、転送された映像のデコード処理に対する研究と表示機器に対する研究が含まれる。デコード処理に対する代表的な研究として、映像の深さを含めたレンダリング処理 [17] や、3D warping [18] があげられる。表示機器に対する代表的な研究として、インテグラル立体テレビ [2] や立体テレビ [19] があげられる。

一方、現在のマルチビュービデオの撮影部では図 2 に示すマルチカメラアレイが用いられている。図 2(a) は直線上にカメラを配置したカメラアレイ [20] であり、図 2(b) は円形にカメラを配置したカメラアレイ [21] である。し



(a) ロボットカメラ (b) 自由視点映像の撮影システム [21]

図 2 既存マルチビュービデオ撮影部
Fig. 2 Conventional multi-view video acquisition.

かしながら、マルチカメラアレイは有線ネットワークでつながれた複数台の撮影機器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、マルチビュービデオを展開する際の障壁となっている。

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、撮影部を無線化するシステムを考える。撮影部を無線化することにより、場所を選ぶことなくマルチビュービデオが撮影できる。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、トラヒックの削減である。マルチビュービデオは複数台の撮影機器の映像を利用するためトラヒックは高いが、無線通信は有線通信と比較してデータレートが低い。トラヒックを削減せずに無線通信でマルチビュービデオを伝送すると、伝送遅延が発生する。2 つ目の要件は、映像品質の維持である。トラヒックの削減と映像品質の維持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、撮影部を無線化した際に、トラヒックの削減と映像品質維持の 2 つの要件を満たす方式として、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ヒックを削減する。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、アクセスポイントが撮影機器間の映像の類似度に従って送信順を制御する。JMVC エンコーダ上に実装した計算機シミュレーションと Mitsubishi Electronics Research Laboratory (MERL) が提供する実ビデオシーケンスを利用して Domino Streaming++ の性能を評価した。評価結果から、他の撮影機器の映像を利用しない場合と比較して、同じ映像品質下において、Domino Streaming++ が最大約 46.1% のトラヒック量を削減できることが分かった。同時に、Domino Streaming++ が無線マルチビュービデオ伝送におけるトラヒック量の理想値に漸近する性能を達成することが分かった。

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 エンコーダと実ビデオシーケンス

を用いて提案方式 Domino Streaming++ の性能を評価する。5章では、提案方式の応用例について議論する。6章では、本研究の関連研究について述べる。7章では、Domino Streaming++ の適用領域やトラフィック削減効果、遅延時間について議論する。最後に8章でまとめとする。

2. 要件

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に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。たとえば、屋外イベントの際にイベントの状況を各参加者が持っているスマートフォンなどの小型撮影機器で撮影し、その場にはいないユーザに対してマルチビュービデオとして提供することができる。

図3に、無線化した撮影部の通信モデルを示す。各撮影機器とアクセスポイントとの間は無線によって、アクセスポイントからエンコーダとの間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アクセスポイントまで伝送し、アクセスポイント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、トラフィックを少なくすることで、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撮影機器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多くなる。一方で、無線通信では、有限である電波を複数のノードによって共有しているため、有線通信と比べると伝送レートが低い。単純にマルチビュービデオを無線通信で伝送した場合、伝送レートの低さによって映像を撮影してから視聴者へ映像を届けるまでの遅延が大きくなり、結果として視聴者満足度が低下する。たとえば、視聴者が視点を切り替えたときに、所望の映像が届くことが遅延していることに起因して、映像が停止したりスキップしたりする。

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、実際と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。

1つ目の要件であるトラフィックの削減と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。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を大きくすることが考えられる。量子

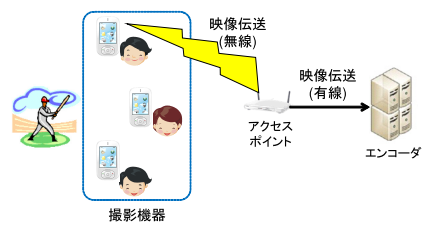


図3 撮影部の無線化モデル

Fig. 3 Model of wireless multi-view video acquisition.

表1 前提条件

Table 1 Assumption.

層	プロトコル
トランスポート	UDP
ネットワーク	IP
MAC・物理	IEEE 802.11a

化パラメータとは、映像を量子化する際に元映像の情報をどれだけ残すかを表すパラメータである。たとえば、マルチビュービデオをエンコードするときに量子化パラメータを大きくした場合、量子化によって多くの映像データが捨てられるため、トラフィック量が減少する。しかしながら、映像をデコードするとき、捨てられた映像データの欠損によって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。

一方、高い映像品質を維持する単純な方法として、各撮影機器が撮影した映像を量子化せず伝送することが考えられる。量子化による映像データの欠損がないため、伝送された映像はすべて元の映像と同じものになる。しかしながら、多くの映像データを伝送することでトラフィックは高くなる。また、要件1で述べたとおり、複数の撮影機器が無線通信で映像を伝送する場合、1つの撮影機器あたりの伝送レートが低くなるため、視聴者への映像の伝送に遅延が発生する。

3. Domino Streaming++

無線化したマルチビュービデオ撮影部において低トラフィックと映像品質の維持を達成する Domino Streaming++ を設計した。Domino Streaming++ では、無線通信において近隣のノードが他のノードの通信を受信できるという特性を利用して、傍受した映像の差分をとることで高い圧縮率を実現する。

表1に Domino Streaming++ が各層で用いることを想定しているプロトコルを示す。物理層・MAC層は IEEE 802.11a [22]、ネットワーク層は IP、トランスポート層は UDP を用いることを想定している。

3.1 動作の全体像

Domino Streaming++ は初期化、特徴量算出、送信順決定、エンコード、映像伝送、デコードの流れで動作する。

- (1) Domino Streaming++に参加する撮影機器は、アクセスポイントの通信範囲内に入ると初期化を開始する。初期化の詳細については3.2節で述べる。
- (2) 各撮影機器は初期化を開始すると、自身の映像の特徴量を、ビデオフレームの画素値をもとに算出する。特徴量を算出すると、各撮影機器は算出した特徴量をアクセスポイントに対して送信する。特徴量算出の詳細については3.3節で述べる。
- (3) 各撮影機器の初期化が終了すると、アクセスポイントは初期化中に取得した各撮影機器の映像の特徴量から映像送信順を決定する。映像送信順を決定すると、アクセスポイントは各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3.4節で述べる。
- (4) 各撮影機器はアクセスポイント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3.5節で述べる。
- (5) 各撮影機器はアクセスポイント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アクセスポイントに対して伝送する。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。すべての撮影機器が映像を送信し終わると、アクセスポイントは再び各撮影機器の映像の特徴量に基づいて、映像送信順序を決定する。映像伝送の詳細については3.6節で述べる。
- (6) 各撮影機器とアクセスポイント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3.7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、アクセスポイントによって一意のIDを割り当てられる。アクセスポイントはアクセスポイント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。撮影機器はビーコンを受信すると、IDを取得するために、自身が撮影している映像の特徴量をアクセスポイントに対して送信する。映像の特徴量の算出方法については、3.3節で述べる。アクセスポイントは各撮影機器の映像の特徴量を受信すると、フレームを送信した撮影機器に対してID割当てフレームを送信する。ID割当てフレームにはアクセスポイントが各撮影機器に割り当てた一意のIDが格納される。

3.3 特徴量算出

各撮影機器はアクセスポイントの通信範囲内に入ると、自身の映像の特徴量を算出し、アクセスポイントに対して

送信する。映像の特徴量はInformation-Bound References (IBR) [23], [24]に基づいて算出する。映像の特徴量は、映像を構成する各ビデオフレームの周波数成分を利用して算出する。周波数成分は、各ビデオフレームに含まれる画素値をDCT変換することで求められ、大きく高周波成分と低周波成分の2種類に分けられる。高周波成分はビデオフレームの詳細な特徴を表している。高周波成分が変化しても、人間はほとんどビデオフレームに対する変化を視認できない。一方で、低周波成分はビデオフレームの大まかな特徴を表している。低周波成分が変化すると、ビデオフレームには人間が視認できる変化が生じる。

各撮影機器は、自身が撮影した映像の1 Group of Pictures (GOP) 分のビデオフレームのうち、先頭のビデオフレームから特徴量を算出する。GOPとは、複数のビデオフレームの集合体のことであり、通常は8枚のビデオフレームから構成される。まず、各撮影機器は先頭のビデオフレームの解像度を 128×128 画素に変換する。その後、撮影機器はビデオフレームの各画素値を4:2:2の輝度信号Y、色差信号Cb, Crに変換する。ビデオフレームの各画素値は合計24ビットのRGB信号から構成されているが、4:2:2の輝度信号および色差信号に変換することで、RGB信号と同じ色分解能を保ったまま、各画素値を8ビットの輝度信号Y、4ビットの色差信号Cr、4ビットの色差信号Cbの合計16ビットで表現することができる。撮影機器は輝度信号Yに対してDCT変換を実行し、輝度信号の低周波成分と高周波成分を取得する。

各撮影機器は輝度信号Yの低周波成分から、自身の映像の特徴量を求める。具体的には、 128×128 画素の輝度信号Yのうち、先頭の 8×8 画素の低周波成分を取り出す。撮影機器は取り出した合計64個の輝度信号Yの低周波成分から中央値を算出する。その後、取り出した各周波数成分が中央値を超えているかに応じて、1または0を割り当てる。最終的に、各撮影機器は特徴量として64個の0, 1列を取得する。

3.4 送信順決定

アクセスポイントは各撮影機器から受信した映像の特徴量をもとに、各撮影機器の映像送信順を決定する。具体的には、各撮影機器が自身の映像と最も類似した映像を利用して、映像をエンコード可能な映像送信順を決定する。アクセスポイントは撮影機器間の映像類似度を求めるために、2台の撮影機器の特徴量からハミング距離を算出する。ハミング距離が小さいほど、2台の撮影機器の映像が類似することを示している。

Algorithm 1に撮影機器がN台の場合の映像類似度に基づく送信順決定アルゴリズムを示す。送信順決定アルゴリズムは送信開始機器の決定と後続機器の送信順決定の2つから構成される。表2に、Algorithm 1で使用する変数と

表 2 Algorithm 1 で使用する変数, 関数

Table 2 List of variables and functions in Algorithm 1.

変数, 関数	説明
C	アクセスポイントの通信範囲内にある撮影機器の ID の集合
$order[i]$	i 番目に映像を送信する撮影機器の ID を格納する配列
$\arg \min_{a \in C} f(a)$	C に属する撮影機器のうち $f(a)$ が最小となる撮影機器 a の ID を取得する関数
I_b	撮影機器 b の映像の特徴量
$\text{calc_hamingdistance}(I_c, I_d)$	特徴量 c と d の間のハミング距離を算出する関数
$\text{delete}(C, e)$	撮影機器の集合 C から ID が e である撮影機器を削除する関数
$\text{size}(C)$	C に含まれる元の個数

Algorithm 1 Transmission order control

```

1:  $i \leftarrow 1$ 
2:  $x \leftarrow 1$ 
3:  $order[i] \leftarrow x$ 
4:  $i \leftarrow i + 1$ 
5:  $\text{delete}(C, x)$ 
6: while  $\text{size}(C) > 0$  do
7:    $y \leftarrow \arg \min_{y \in C} (\text{calc\_hamingdistance}(I_y, I_x))$ 
8:    $order[i] \leftarrow y$ 
9:    $i \leftarrow i + 1$ 
10:   $x \leftarrow y$ 
11:   $\text{delete}(C, y)$ 
12: end while
    
```

関数を示す。 C はアクセスポイントの通信範囲内にある撮影機器の ID の集合である。 $order[i]$ は i 番目に映像を送信する撮影機器の ID を格納する配列である。 $\arg \min_{a \in C} f(a)$ は C に属する撮影機器のうち $f(a)$ が最小となる撮影機器 a の ID を取得する関数である。 I_b は撮影機器 b が撮影している映像の特徴量を表す。 $\text{calc_hamingdistance}(I_c, I_d)$ は特徴量 c と特徴量 d の間のハミング距離を算出する関数である。 $\text{delete}(C, e)$ は撮影機器の集合 C から ID が e である撮影機器を削除する関数である。 $\text{size}(C)$ は C に含まれる元の個数であり、初期値は撮影機器数 N である。

1 つ目の処理である開始機器の決定では、 N 台の撮影機器の中で、最初に映像を送信する撮影機器を決定する。 Domino Streaming++ では、アクセスポイントに ID 1 を割り当てられた撮影機器を開始機器とする。具体的には、 Algorithm 1 の 1 行目から 5 行目で、アクセスポイントの通信範囲内にある撮影機器のうち、 ID 1 が割り当てられた撮影機器 x を開始機器とする。

2 つ目の処理である後続機器の送信順決定では、開始機器を除くすべての撮影機器の送信順を決定する。具体的に

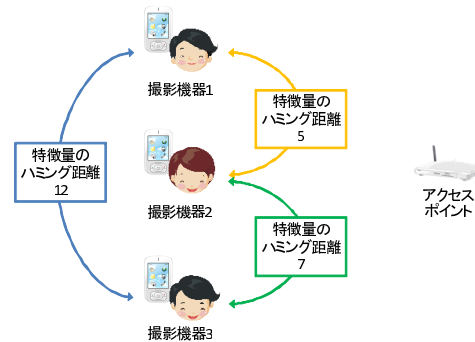


図 4 撮影機器が 3 台のときの配置例
Fig. 4 An example of topology of three nodes.

は、撮影機器間の特徴量のハミング距離が最小になる映像送信順を決定する。 Algorithm 1 の 7 から 11 行目で、まだ送信順を割り当てていない撮影機器の映像の特徴量の中で、直前に送信順を割り当てた撮影機器 x の映像の特徴量とのハミング距離が最小となる撮影機器 y を選択する。アクセスポイントは撮影機器 y に送信順を割り当て、撮影機器 y を撮影機器 x と見なし、 Algorithm 1 の 7 から 11 行目の処理を繰り返す。

例として 3 台の撮影機器がある場合の各撮影機器の配置例を図 4 に示す。各撮影機器は初期化後、撮影機器 1 に ID 1、撮影機器 2 に ID 2、撮影機器 3 に ID 3 が割り当てられている。図 4 のアクセスポイントは、各撮影機器から受信した映像特徴量から撮影機器間の特徴量のハミング距離を算出する。ここでは、ハミング距離が図 4 のようになったと仮定する。

まず、アクセスポイントは ID 1 が割り当てられた撮影機器 1 を開始機器とし、 $order[1]$ に撮影機器 1 の ID を格納する。次に、撮影機器 1 の映像の特徴量と撮影機器 2、3 の映像の特徴量とのハミング距離を算出する。撮影機器 1 と撮影機器 2 の間のハミング距離は 5、撮影機器 1 と撮影機器 3 の間のハミング距離は 12 であるため、アクセスポイントはハミング距離がより小さい撮影機器 2 の ID を $order[2]$ に格納する。その後、アクセスポイントは残った撮影機器 3 の ID を $order[3]$ に格納し、送信順決定アルゴリズムを終了する。最終的に、アクセスポイントが決定した送信順は撮影機器 1 → 撮影機器 2 → 撮影機器 3 となる。

3.5 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 ごとに自身の映像をエンコードする。 Domino Streaming++ では、撮影機器は自身の映像を低トラヒ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すべて傍受する。

図 5、図 6、図 7 に GOP のエンコードを示す。ここで撮影機器の台数は 3 台であり、映像送信順は撮影機器 1 →

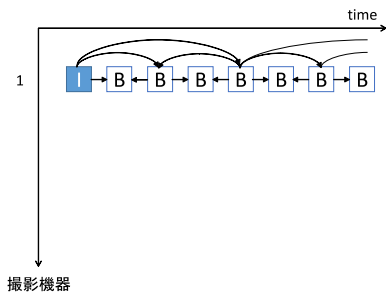


図 5 撮影機器 1 の映像のエンコード

Fig. 5 Prediction structure of camera 1.

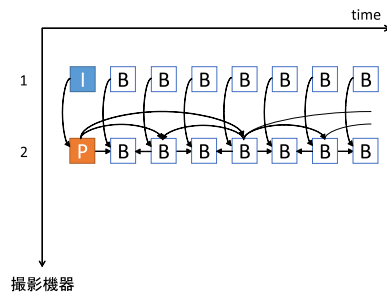


図 6 撮影機器 2 の映像のエンコード

Fig. 6 Prediction structure of camera 2.

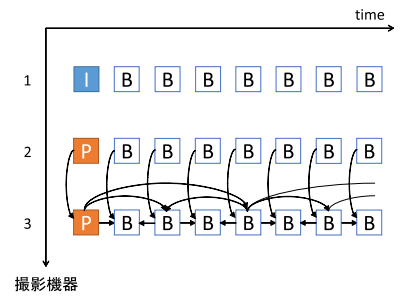


図 7 撮影機器 3 の映像のエンコード

Fig. 7 Prediction structure of camera 3.

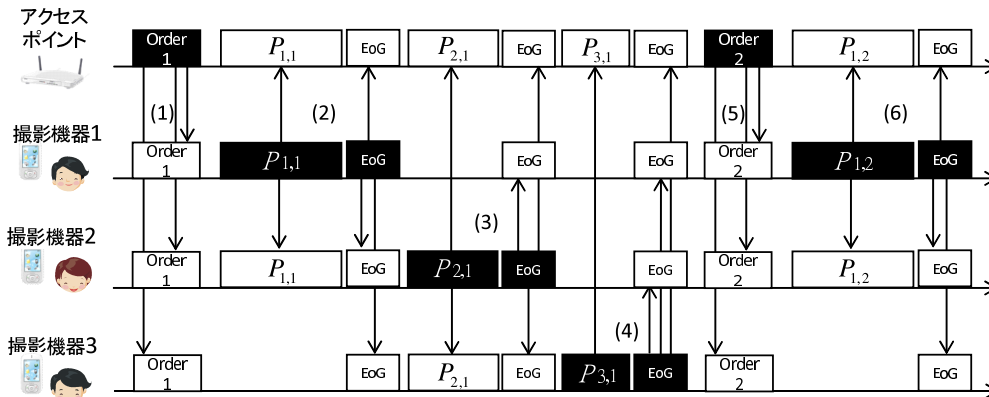


図 8 映像伝送タイムシーケンスチャート

Fig. 8 Timing diagram of Domino Streaming++.

撮影機器 2→撮影機器 3 の順であると想定する。図 5 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 6 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 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 2 の先頭フレームを P フレームとしてエンコードする。P フレームは他の撮影機器の I フレームをもと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 7 に撮影機器 3 の予測構造を示す。撮影機器 3 は、撮影機器 2 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 3 の先頭フレームを P フレームとしてエンコードする。

3.6 映像伝送

各撮影機器はアクセスポイントが割り当てた送信順に従って、3.5 節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に対して送信する。図 8 に映像伝送のタイムシーケンスチャートを示す。図 8 では、アクセスポイントが撮影機器 1→撮影機器 2→撮影機器 3 の順に送信順を決定した

表 3 ビデオパケット $P_{i,j}$

Table 3 Video packet $P_{i,j}$.

フィールド	バイト
映像類似度	8
映像	可変長

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 3 に $P_{i,j}$ のフォーマットを示す。映像類似度フィールドには、3.4 節で述べた方法に従って算出した各撮影機器の映像の特徴量を格納する。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) アクセスポイントは 3.4 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すべての撮影機器にブロードキャストする。

(2) アクセスポイント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 はエンコードした GOP1 の映像を映像フィールド、GOP2 の先頭のビデオフレームから算出した特徴量を映像類似度フィールドに格納し、 $P_{1,1}$ をアクセスポイントに対して送信する。撮影機器 2 は $P_{1,1}$ を傍受して撮影機器 1 の映像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 [22]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3は、次はまだ自身の送信順ではないと判断し、撮影機器2の通信が終了するまで待機する。

- (3) 撮影機器2は、エンコードしたGOP1の映像を映像フィールド、GOP2の先頭のビデオフレームの特徴量を映像類似度フィールドに格納し、アクセスポイントに対して $P_{2,1}$ を送信する。撮影機器3は $P_{2,1}$ を傍受することで撮影機器2の映像を取得する。その後、撮影機器3は受信した映像のデコードを開始する。撮影機器2は $P_{2,1}$ の送信が終了すると、EoGパケットをアクセスポイントへ送信する。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2の映像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3は、エンコードしたGOP1の映像を映像フィールド、GOP2の先頭のビデオフレームの特徴量を映像類似度フィールドに格納し、アクセスポイントに対して $P_{3,1}$ を伝送する。撮影機器3は $P_{3,1}$ の送信が終了すると、EoGパケットをアクセスポイントへ送信する。
- (5) アクセスポイントは受信した $P_{1,1}$ 、 $P_{2,1}$ 、 $P_{3,1}$ 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特徴量から、3.4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、決定した送信順をすべての撮影機器にブロードキャストする。
- (6) GOP2に対する映像送信順を受信した各撮影機器は、映像送信順に従ってGOP2に対する映像伝送を開始する。ここでは、GOP2に対する映像送信順が撮影機器1→撮影機器2→撮影機器3の順であると仮定する。GOP1のときと同様に、撮影機器1はエンコードしたGOP2の映像を映像フィールド、GOP3の先頭のビデオフレームの特徴量を映像類似度フィールドに格納し、 $P_{1,2}$ をアクセスポイントに対して送信する。

Domino Streaming++ではすべてのGOPの送信が終了するまで、(2)から(6)の動作を繰り返す。

3.7 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。撮影機器とアクセスポイント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もとに作成された差分映像である。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。エンコーダではすべ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すべ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に対して伝送することで、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 [25]上に実装した計算機シミュレーションとMERLが提供しているテストビデオシーケンス [26]によって、各撮影機器の映像に対する特徴量と映像品質に対する映像ビットレートを測定した。Domino Streaming++の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アクセスポイントに対して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) Ideal

Idealは各撮影機器が理想的な組合せでエンコードした場合のトラフィック量である。Idealは無線マルチビュービデオにおけるトラフィック量の理想値である。Idealでは、各撮影機器が撮影している映像は全知であると仮定する。アクセスポイントは各撮影機器の映像に応じて、GOPごとに理想的な映像送信順を決定する。Idealは、Domino Streaming++が無線マルチビュービデオ伝送における性能限界にどの程度近づいたかを示す尺度となる。

(3) Domino Streaming++

Domino Streaming++は、3章で述べた提案方式である。Domino Streaming++では、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、アクセスポイント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

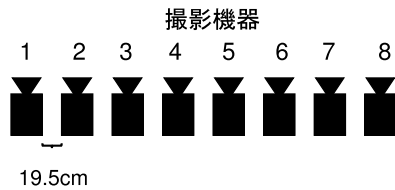


図 9 撮影機器の配置

Fig. 9 Arrangement of cameras.

表 4 評価パラメータ

Table 4 Simulation parameters.

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

表 5 PSNR に対する MOS 評価結果

Table 5 Correlation between MOS and PSNR.

PSNR [dB]	MOS
≥ 45	5
≥ 33 & < 45	4
≥ 27.4 & < 33	3
≥ 18.7 & < 27.4	2
< 18.7	1

ケンスには, Ballroom, Exit, Vassar を利用した. 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. 図 9 に, Ballroom, Exit, Vassar における撮影機器の配置を示す. 各ビデオシーケンスは, 1 列に並んだ 8 台の撮影機器によって撮影されている [26]. 撮影機器間の距離は 19.5 cm である. 8 台の撮影機器の位置および角度は固定されており, 同じ視点から映像を撮影し続ける. 本評価では, マルチビュービデオの無線化に向けた提案手法の基礎性能を明らかにしている. 具体的には, 少数数のユーザが同じ対象物を並んで撮影している環境での提案手法の性能を評価している.

表 4 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) 時のトラヒック量を取得した. 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. また, PSNR と視聴者満足度との関係は, 主観品質評価法の 1 つである Mean Opinion Score (MOS) によって表される. MOS とは, 動画像に対する主観的評価方法である. 表 5 に, 文献 [27], [28] から得られた PSNR と MOS の関係を示す. 文献 [27], [28] では, PSNR が異なる 175 種類の映像を 25 名の被験者が実際に視聴し, 動画像の品質を 5 段階で評価している. 5 段階の評価は, 非常に良い (5), 良い (4), 普通 (3), 悪い (2), 非常に悪い (1) となっている.

4.2 映像の特徴量算出

マルチビュービデオにおける映像類似度の特性を評価するために, 3.3 節に示した方法で撮影機器間の映像類似度を評価した.

図 10 に, 利用したビデオシーケンスが Ballroom, 各撮影機器の映像の中で GOP1 の先頭フレームを特徴量算出に使用した場合の 2 撮影機器間の映像類似度を示す. 図 10 (a) から (h) はそれぞれ撮影機器 1 から 8 の映像に対する他の撮影機器の映像の類似度を示す. 各図に対して, 横軸は類似度算出に用いたもう一方の撮影機器番号, 縦軸は 2 撮影機器の特徴量のハミング距離である. 本評価では, 撮影機器どうしの番号が近いほど, 撮影機器間の距離が近くなることを意味している.

図 10 より, 類似度算出に用いた撮影機器どうしの距離が近いほど, 映像間の類似度を表すハミング距離が小さくなっていることが分かる. 利用したビデオシーケンスが Exit, Vassar の場合も, 類似度算出に用いた撮影機器どうしの距離が近いほど, 映像間の類似度を表すハミング距離が小さくなった. マルチビュービデオは複数台の撮影機器が同じ対象物を撮影するため, 撮影機器間の距離が近づくほど, 映像間の相関が高くなるという特性を持つ. 各撮影機器の特徴量に基づく類似度算出を用いることで, アクセスポイントによるマルチビュービデオの特性をふまえた, より正確な送信制御が可能になるといえる.

4.3 理想的なトラヒック量との比較

Domino Streaming++ の基本性能を評価するために, 映像品質を変化させた場合のトラヒック量を理想的なトラヒック量と比較した.

図 11 に, 撮影機器数は 3 台, 量子化パラメータを 24 から 40 に変化させた場合の映像ビットレート, 図 12 に, 撮影機器数は 4 台, 量子化パラメータを 24 から 40 に変化させた場合の映像ビットレート, 図 13 に, 撮影機器数は 6 台, 量

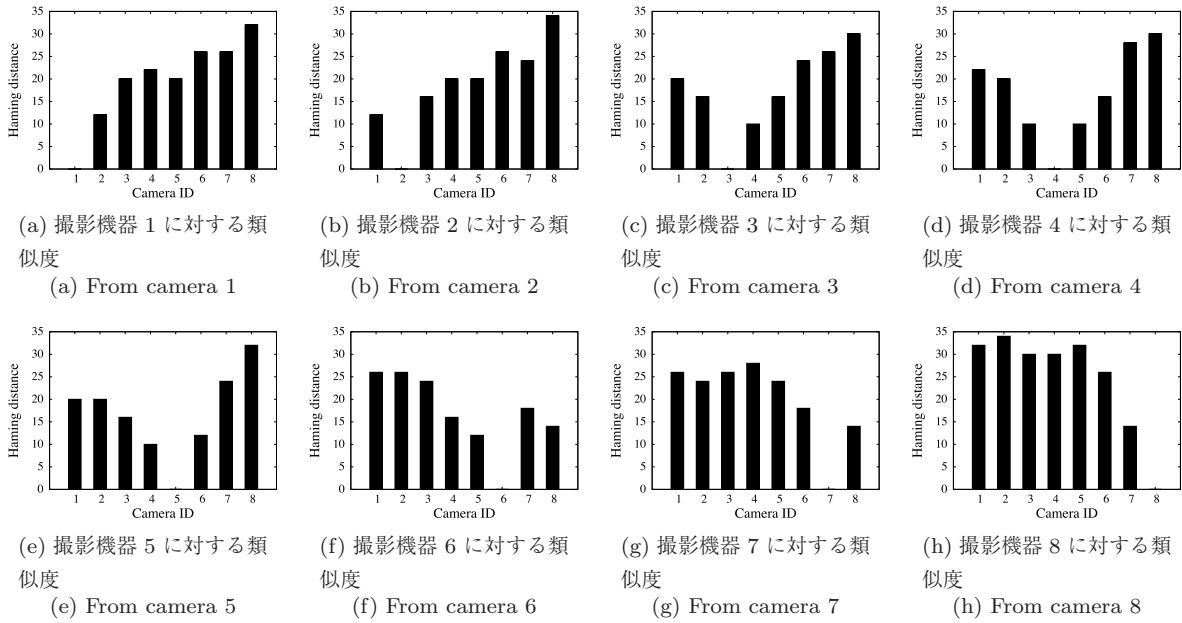


図 10 特徴量のハミング距離

Fig. 10 Hamming distance between cameras.

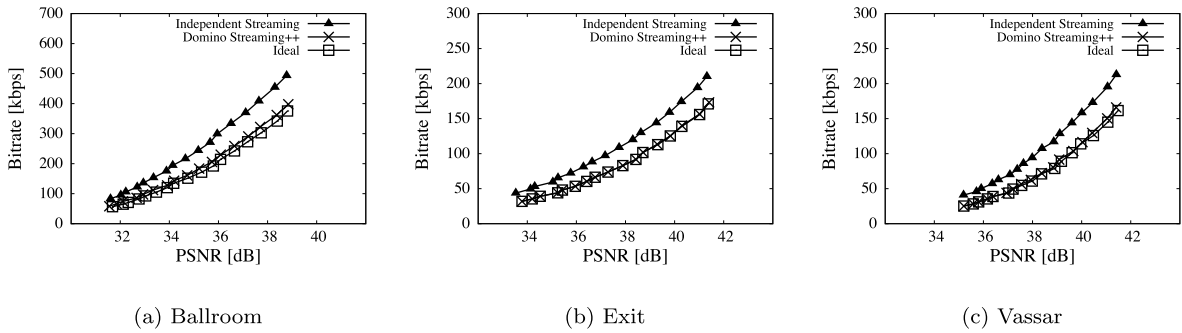


図 11 理想値との比較 (撮影機器が 3 台の場合)

Fig. 11 Comparison with ideal at three cameras.

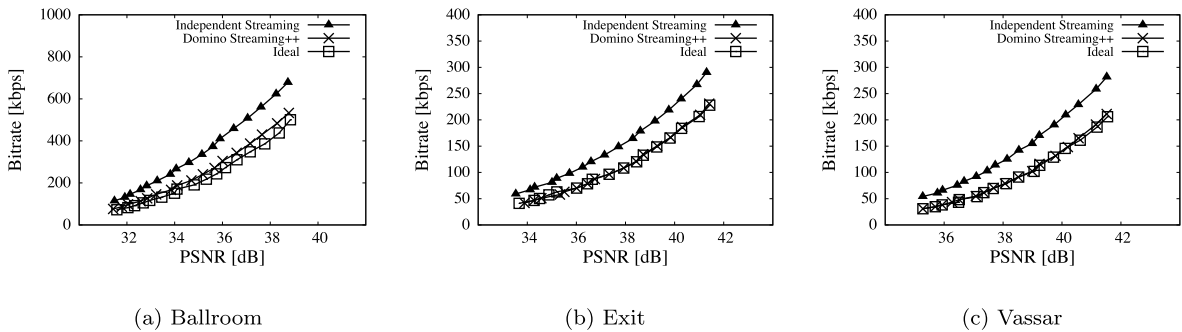


図 12 理想値との比較 (撮影機器が 4 台の場合)

Fig. 12 Comparison with ideal at four cameras.

子化パラメータを 24 から 40 に変化させた場合の映像ビットレートを示す。それぞれ、図 11 (a), 図 12 (a), 図 13 (a) はビデオシーケンス Ballroom を用いた場合、図 11 (b), 図 12 (b), 図 13 (b) はビデオシーケンス Exit を用いた場合、図 11 (c), 図 12 (c), 図 13 (c) はビデオシーケンス Vassar を用いた場合の映像ビットレートである。各ビデオシーケンスは、撮影機器数が 3 台の場合は図 9 の撮影機器 1 から

3 の映像、撮影機器数が 4 台の場合は図 9 の撮影機器 1 から 4 の映像、撮影機器数が 6 台の場合は図 9 の撮影機器 1 から 6 の映像を用いた。すべて、横軸は PSNR [dB], 縦軸は映像ビットレート [kbps] である。図 11, 図 12, 図 13 より、次の 3 つのことが分かる。

1 つ目は、撮影機器の台数やビデオシーケンスの種類にかかわらず、Domino Streaming++ は、Independent Stream-

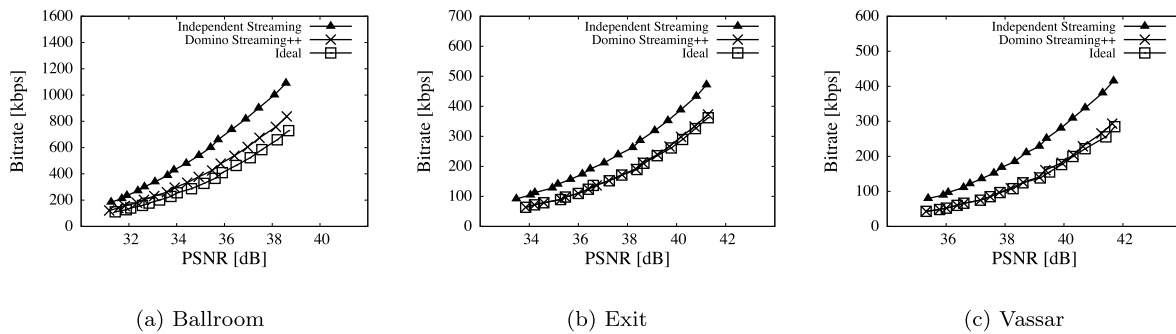


図 13 理想値との比較 (撮影機器が 6 台の場合)

Fig. 13 Comparison with ideal at six cameras.

ing 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。たとえば、図 12(a)において撮影機器が 4 台、ビデオシーケンスが Ballroom, PSNR が 36 [dB] のとき, Domino Streaming++ は Independent Streaming と比較して, 約 250 [kbps] のトラフィック量を削減している。また, 図 13(c)において撮影機器が 6 台, ビデオシーケンスが Vassar, PSNR が 36 [dB] のとき, Independent Streaming のビットレートは約 96.9 [kbps], Domino Streaming++ のビットレートは約 52.1 [kbps] であり, 両者の差分からトラフィック削減率を算出すると約 46.1 [%] となる。図 10 から図 12 のすべての結果に対して同様に削減率を算出した後, 約 46.1 [%] が最大の削減率であることを確認した。Domino Streaming++ では, 他の撮影機器の映像を通信傍受によって取得し, 自身の映像のエンコードに利用することで, 撮影機器間に存在する冗長な情報を削減している。また, 図 11(a), 図 12(a), 図 13(a) から, Domino Streaming++ はより低い伝送レート下において, 高い視聴者満足度を達成できることが分かる。たとえば, 図 13(a) および 4.1 節の表 5 から, Independent Streaming が MOS 4 を達成するためには, 約 339 [kbps] の伝送レートが必要となる。一方で, Domino Streaming++ における必要な伝送レートは約 231 [kbps] となる。

2 つ目は, 撮影機器の台数やビデオシーケンスの種類にかかわらず, Domino Streaming++ のトラフィック量が, トラフィック量の理想値である Ideal に漸近していることである。たとえば, 図 12(a)において撮影機器が 4 台, ビデオシーケンスが Ballroom, PSNR が 36 [dB] のとき, Domino Streaming++ は Ideal と比較して, 約 30 [kbps] だけ離れている。また, 図 12(c)において撮影機器が 4 台, ビデオシーケンスが Vassar, PSNR が 36 [dB] のとき, Domino Streaming++ のトラフィック量と, Ideal のトラフィック量は同じ値になる。Domino Streaming++ では, 映像類似度に基づく送信順制御を利用することで, 撮影機器間に存在する冗長な情報をより多く削減することができ, 理想値に近いトラフィック量を達成している。

3 つ目は, ビデオシーケンスが Ballroom の場合, PSNR

が高くなるにつれて, Domino Streaming++ と Ideal との差が広がっている点である。たとえば, 図 12(a)においてビデオシーケンスが Ballroom, PSNR が 32 [dB] のとき, Domino Streaming++ は Ideal と比較して, 約 6 [kbps] しか離れていないが, PSNR が 36 [dB] のときは約 30 [kbps] 離れている。PSNR が高くなると, ビデオフレームの低周波成分だけでなく, 特徴量には現れない高周波成分にも変化が生じる。高周波成分の変化によって, アクセスポイントが理想的な映像送信順とは異なる映像送信順を決定するため, Domino Streaming++ と Ideal とのトラフィック量の差が大きくなっていると考えられる。一方, Exit や Vassar を利用した場合, PSNR が高くなっても, Domino Streaming++ と Ideal との差が小さいことが分かる。Exit や Vassar は Ballroom と比較して, 動きの少ない映像であるため, 映像の周波数は低周波成分に集中する。よって, PSNR が高くなるにつれて生じる高周波成分の変化は, 映像送信順に大きな影響を与えず, Domino Streaming++ と Ideal とのトラフィック量の差は広がらないと考えられる。

5. 応用

表 6 に, 各分野において, 複数の携帯型カメラ, ウェアラブルデバイスカメラ, ビークル搭載カメラなどの小型撮影機器を利用した応用例をまとめた。

教育の分野では, 携帯型カメラの映像から作成した立体映像を教材に利用することができる。たとえば, 体育の実技指導に立体映像を教材として提供することで, 教師の実技技術にかかわらず授業を進めることが可能となる。ウェアラブルデバイスカメラで撮影した映像は, 人間の視点から見た車両整備, 理容などの技術習得の支援に利用することができる。ビーグル搭載カメラで撮影した映像は, 車両, 航空機などの免許取得シミュレータに利用することで, 屋内にいながら屋外で実際に車両を運転しているように錯覚する没入型の訓練が可能となる。

エンタテインメントの分野では, 有名人の撮影やライブイベント, バイクツーリングなど, 多様な場面での利用が考えられる。特にスポーツ観戦では, 携帯型カメラやウェア

表 6 無線化したマルチビュービデオの応用例
Table 6 Applications of wireless multi-view video systems.

	携帯型カメラ	ウェアラブルデバイスカメラ	ビーグル搭載カメラ
教育	3D 学校教材	スキルトレーニング (車両整備, 理容, 陶芸など)	運転免許シミュレータ教材
エンターテインメント	有名人の撮影	ライブイベント グループスカイダイビング グループスキューバダイビング	気球大会, バイクツーリング ドローン編隊
スポーツ観戦	野球, サッカーなど		F1・ワールドフランス
医療	遠隔からの介護支援	遠隔からの医療支援 過去の映像を利用した診断や治療	内視鏡検査システム
防犯	360 度監視システム	複数の警備員による警備の強化	車上荒らしの防止
交通安全	3D 交通安全マップ	ヒューマンビューモニター	自動・手動運転支援 飛び出し事故の防止
資源探査	定点カメラを用いた自然資源の調査	宇宙飛行士による惑星探査	AUV やヘリによる資源探査 無人探査機による調査
防災・復旧支援	詳細な避難経路マップ		災害の予兆探知や調査 原子炉建屋内の調査・被害状況記録
生態系調査	定点カメラを用いた生態系の調査	ダイバーによる生態系の調査	人が立ち入れない場所の生態系調査
デジタルアーカイブ	定点カメラを用いた建造物構造の保存	個人の活動記録 (3D アルバム)	事故時の裁判資料

ウェアラブルデバイスカメラを用いた観客視点, または, 競技者の車両に設置されたカメラを用いた競技者視点のマルチビュービデオを提供することができる。

医療の分野では, 携帯型カメラを用いることで, 遠隔からの介護支援が可能になると考えられる。医師や看護師が装着したウェアラブルデバイスカメラの映像を組み合わせることで, 別の医者による遠隔からの医療支援や過去の映像からの変化を用いた病気の早期発見・治療が可能となる。小型のビーグルカメラを用いて, 自由に観察したい場所を変更できる内視鏡カプセルを実現することができると考えられる。

防犯の分野では, 複数の携帯型カメラやウェアラブルデバイスカメラを組み合わせることで監視カメラとして用いることで, 死角のない監視システムを提供することができる。ほかに, 複数の駐車車両の車載カメラを組み合わせることで, 車上荒らしの早期発見などに利用することができる。

交通安全の分野では, 携帯カメラからの映像を用いることで, 人間の視線を利用した 3D 交通安全マップが作成できると考えられる。ウェアラブルデバイスカメラの映像を用いることで, 助手席や後部座席からの映像を利用したヒューマンビューモニタの提供が可能となる。また, 各車両のカメラ映像を利用することで, 手動運転の支援だけでなく, より高精度な自動運転が可能になると考えられる。

資源探査の分野では, 複数の小型定点カメラからの映像を用いて, 有線通信を用意することができない場所での資源調査ができる。また, 宇宙飛行士が装着したウェアラブルデバイスカメラの映像を用いることで, より円滑な惑星探査が可能になると考えられる。ほかに, Autonomous Underwater Vehicle (AUV) や惑星探査機に搭載したカメ

ラを用いることで, 人間の立ち入りが困難である水中や惑星での探査を支援することができる。

防災の分野では, 携帯カメラやウェアラブルデバイスカメラを用いて撮影した映像を避難経路マップに利用することで, 災害時の円滑な避難を支援することができる。また, ビーグル搭載カメラで撮影した地上, 上空, 海底からの映像を災害の予兆探知や以後の災害への対策に利用できる。

生態系調査の分野では, 複数の小型定点カメラからの映像を用いて, 有線通信を用意することができない場所での生態系の調査が可能になると考えられる。飼育環境に複数の定点カメラを設置することで, 飼育している動物の健康状態をチェックすることも可能となる。また, 複数のダイバが装着したウェアラブルデバイスカメラからの映像を用いることで, 広域に及ぶ海中での生態系調査が可能となる。一方で, ビーグル搭載カメラを利用することで, 人が立ち入ることができない深海などでの生態系調査も可能となる。

デジタルアーカイブの分野では, 複数の小型定点カメラを利用することで, 有線通信を用意することができない場所での建造物の構造や自然の景観保存が可能となる。たとえば, 日本唯一の崩壊中の無人都市である長崎県の軍艦島では, 崩壊中の建造物を立体映像化することで, 建造物の構造をより忠実に保存することができる [29]。ほかに, ウェアラブルデバイスカメラを利用することで, 子供の成長などを立体映像として記録したり, ビーグル搭載カメラを利用することで, 事故発生時の周辺車両の映像を収集し, 事故の原因究明に利用したりすることが考えられる。

6. 関連研究

本研究は映像間の類似度算出技術、無線通信を用いた複数映像の伝送技術、分散マルチビュービデオ技術に関する。

6.1 映像の類似度算出技術

映像の類似度算出技術には、ビデオフレームのカラーヒストグラムをもとに算出する方式、ビデオフレーム内の特徴点をもとに算出する方式、ビデオフレームの画素値をもとに算出する方式がある。

ビデオフレームのカラーヒストグラムをもとに算出する方式として、文献 [30] があげられる。本方式では、各ビデオフレームを構成する色の分布から類似度を算出する。まず、ビデオフレーム内のピクセルを色の種類、明暗に応じて分類する。次に、縦軸を色の種類、横軸をピクセル数とする赤成分、緑成分、青成分のカラーヒストグラムをそれぞれ出力する。最後に、2枚のビデオフレームから算出したカラーヒストグラムの形状を比較することで、映像間の類似度を算出できる。

ビデオフレーム内の特徴点を用いる方式として Scale-Invariant Feature Transform (SIFT) [31], GIST [32] があげられる。SIFTはあるビデオフレームに対して、回転、スケールの変化、明度の変化などの操作を加えても変化が現れない特徴点を算出する。まず、スケールの異なるビデオフレームを平滑化させることで、スケールの変化に影響されない特徴点を算出する。次に、算出した特徴点の中から、明度の変化に影響しない特徴点を算出するため、コントラストの変化が大きい点を削除する。さらに算出した特徴点の向きを正規化する。GISTはあるビデオフレームに対して、被写体の構造的な特徴点を算出する。具体的には、ビデオフレームの周波数スペクトルを用いて、入力されたビデオフレームの特徴点を算出する。SIFT, GISTは、2枚のビデオフレームと各ビデオフレーム内の特徴点を用いることで、映像間の類似度を算出できる。

ビデオフレームの画素値を用いる方式として Information-Bound References (IBR) [23], [24] があげられる。IBRは、あるビデオフレームを構成する画素値から算出される周波数成分を利用する。IBRを用いることで、92 [byte] で表される映像の特徴値を算出できる。その後、2枚のビデオフレームにおける特徴値を比較することで、2映像の類似度を算出できる。

Domino Streaming++ では、各撮影機器がアクセスポイントに対して、IBRを用いた映像の特徴値を送信することで、撮影機器間の映像類似度に基づいたトラフィック削減を達成する。各ビデオフレームのカラーヒストグラムを用いる場合、アクセスポイントはビデオフレームを構成する色情報のみで映像の類似度を判断する。このとき、映像の被

写体は考慮されないため、類似度が正確に判断できない場合が生じる。たとえば、上半分が白色、下半分が黒色で構成されたビデオフレームと下半分が白色、上半分が黒色で構成されたビデオフレームは視覚的に類似していないが、カラーヒストグラムを類似度算出に用いた場合、白色と黒色を構成するピクセル数は同じであるため、類似していると判断される。一方で、IBRを用いた場合、ビデオフレームの特徴を表す周波数値からビデオフレームの特徴値を算出するため、被写体の構造も考慮した映像の類似度算出が可能となる。

各ビデオフレーム内の特徴点を用いた場合、アクセスポイントにビデオフレームと特徴点を送信する必要があり、伝送遅延が大きくなる。一方で、IBRを用いた場合、アクセスポイントに92 [byte] で表される特徴値のみを送信すればよいため、ビデオフレームと特徴点を送信する場合と比較して伝送遅延は小さくなる。

6.2 無線通信を用いた複数映像の伝送技術

無線通信を用いた複数映像の伝送技術として、撮影機器間で同期をとり、映像を伝送する方式 SimCast [33], 映像品質に応じたスケジューリングを用いる方式 [34] があげられる。

SimCastはMulti-User Multiple Input Multiple Output (MU-MIMO) を用いて複数の撮影機器が同時に映像を伝送することで、効率的な映像アップロード・ダウンロードを実現する。具体的には、ネットワーク状況に応じた冗長な映像伝送の抑制、撮影機器間の類似性を用いた映像データの圧縮を実現する。

SimCastは2台の撮影機器による映像伝送に焦点を当てている。一方で、Domino Streaming++ では、より多くの撮影機器による映像伝送に焦点を当てている。

文献 [34] では、複数の撮影機器から映像を伝送する場合、伝送後の映像品質の劣化量が最小となる撮影機器間のスケジューリング方式が提案されている。

本スケジューリング方式では、各撮影機器は集約機器に対して、他の撮影機器の映像を参照せずに自身の映像を送信している。一方で、提案手法 Domino Streaming++ は、各撮影機器の映像間には相関が存在すること、撮影機器は他の撮影機器の通信を傍受できることを利用し、自身の冗長な映像データを削減することで、効率的な映像伝送を達成している。

6.3 分散マルチビュービデオ技術

マルチビュービデオにおいて、映像品質を維持したまま無線通信時のトラフィックを削減する方法として、Distributed Multiview Video Coding (DMVC) [35], [36], [37] があげられる。DMVCは、マルチビュービデオの無線伝送において、映像間の相関を利用して符号化レベルでアプローチし

ている研究である。DMVCでは、分散情報源符号化を適用するとともに、撮影機器の設置位置や撮影しているアングルを映像データとともにside informationとして送信する。分散情報源符号化とは、相関を有する2つのデータに独立に符号化しても参照したときと同程度の圧縮率を達成できる符号化手法であり、Slepian-Wolf理論[38]やWyner-Ziv理論[39]などが存在する。Domino Streaming++は映像の前処理部分に着目したアプローチであるため、符号化レベルでのアプローチであるDMVCと組み合わせて使用することもできる。

7. 議論

7.1 適用領域に関する議論

Domino Streaming++を利用するための条件として、以下の4つがある。

- (1) アクセスポイントが設置可能な環境であること
- (2) アクセスポイントの通信範囲内に撮影機器が存在すること
- (3) アクセスポイントと各撮影機器が直接通信できること
- (4) 各撮影機器の通信範囲内に他の撮影機器が存在して通信を傍受できること

4つの条件を今すぐすべて揃えるのは困難であると考えている。たとえば、現在のスマートフォンで通信を傍受するのは不可能ではないものの困難である。そこで著者らはDomino Streaming++を段階的に展開していくことを想定している。

一番最初に想定しているのは現在のマルチビュービデオの無線化である。現在のマルチビュービデオの撮影部では、有線を用いているが、現在のマルチビュービデオ撮影システムが有線で接続されていることに起因する問題を緩和するため、配線の手間やカメラの動きの制約から無線化したいという要望がある。移動の制約やケーブルの引き回しなどが解消されることで、たとえば、携帯撮影機器から集めた映像とクラウドサービスを用いた新たな立体・自由視点映像配信サービス[40]が可能になると考えている。

現在のマルチビュービデオ撮影部の有線を無線に置き換えることだけ考えた場合、無線通信を独自のシステムとして構築可能であると考えている。専用機器を用いる場合では、各撮影機器が無線通信を傍受できるようにハードウェアなどを設定することは比較的容易だと考えられる。たとえば、IEEE 802.11gの無線LANチップのアドホックモードを利用して独自システムを構築することを想定する。

筆者らが最終形態として考えているユーザのモバイル端末を利用したマルチビュービデオを実現するためには、他の端末の通信の傍受方法が課題となる。傍受方法の課題を解決するためには、現在さかんに研究されている、他の端末の通信を傍受することを前提とした無線通信方式と合わせ

て取り組んで行く必要があると考えている。CISCO Visual Networking Index 2013-2018[41]によると、全世界の無線トラフィックは2018年までに現在の11倍になるといわれている。近年の無線研究では、このような需要に応えるため、周辺の無線機器どうしが互いの通信を傍受し、機器間で協調して通信する手法が考えられている[42],[43],[44]。たとえば、協調通信は、各無線機器が独立して通信するよりも効率的であることが示されている。このように将来的には、通信効率の向上を目指して、各端末が互いに通信を傍受することを前提とした無線通信方式が前提となるのではないかと予想している。

また、Domino Streaming++は、モバイル端末が爆発的に増加した将来の環境を想定した方式である。文献[45]で述べられているような近年の爆発的なモバイル端末数の増加傾向を考慮すると、複数のアクセスポイントの分散配置あるいはチャンネルの分散利用をした場合であっても、1つのチャンネルあたりのモバイル端末数は多くなる。このような環境下においてマルチビュービデオ伝送を利用した場合、複数のモバイル端末が同一チャンネルを用いて映像を無線送信する場面が増加すると予想している。たとえば、スタジアムでの応用を考えた場合、スタジアム内において多数のモバイル端末が同一チャンネル内でマルチビュービデオ伝送することが想定される。

7.2 評価パラメータが変化した場合の考察

4.1節の表4で示した、解像度、フレームレート、GOPサイズ、カメラ数、量子化パラメータの各評価パラメータと、4.2節で評価した映像の類似度について、それぞれのパラメータの変化が結果に与える影響を考察した。

(1) 1ビデオフレームの解像度

解像度は1つのビデオフレームを構成する縦横の画素の量である。解像度が向上した場合、図9から図11に示した結果と同様に、撮影機器間で冗長な情報が増加するため、Domino Streaming++のトラフィック削減量が増加する。一方、解像度が低下した場合、撮影機器間で冗長な情報が減少するため、Domino Streaming++のトラフィック削減量が減少する。

(2) フレームレート

フレームレートは1秒あたりに再生されるビデオフレーム数である。フレームレートが増加した場合、1秒間に表示するビデオフレーム数が増加することで撮影機器間の冗長な情報が増加するため、Domino Streaming++のトラフィック削減量が増加する。一方、フレームレートが減少した場合、1秒間に表示するビデオフレーム数が少なくなるため、各ビデオフレームの冗長な情報を削減することで得られる利得が小さくなり、Domino Streaming++のトラフィック削減量も減少する。

(3) GOP サイズ

GOP サイズは1つの GOP を構成するビデオフレーム数である。GOP サイズが増加した場合、各撮影機器は全体のビデオフレームのうち、より多くのビデオフレームを、他の撮影機器の映像を用いてサイズの小さい B フレームとしてエンコードするため、Domino Streaming++ のトラフィック削減量が増加する。一方、GOP サイズが減少した場合、各撮影機器は全体のビデオフレームのうち、より多くのビデオフレームを、他の撮影機器の映像を用いることなくサイズの大きい I フレームとしてエンコードするため、Domino Streaming++ のトラフィック削減量が減少する。

(4) カメラ数

カメラ数は被写体を撮影している撮影機器の台数である。カメラ数が増加した場合も減少した場合も、新たに加わった撮影機器が撮影している映像と他の撮影機器の映像との間の相関に応じて、Domino Streaming++ のトラフィック削減量が増減する。映像間の相関が強い場合、Domino Streaming++ のトラフィック削減量が増加し、相関が弱い場合、Domino Streaming++ のトラフィック削減量が減少する。

(5) 量子化パラメータ

量子化パラメータは映像をエンコードする場合の量子化の度合いを表す数値である。量子化パラメータが増加すると、量子化によって、撮影機器間で冗長な情報を含む多くの元映像情報が削除されるため、Domino Streaming++ のトラフィック削減量が減少する。量子化パラメータが減少すると、量子化後にも、撮影機器間で冗長な情報を含む多くの元映像情報が残っているため、Domino Streaming++ のトラフィック削減量が増加する。

(6) 映像の類似度

量子化パラメータは映像をエンコードする場合の量子化の度合いを表す数値である。映像の類似度が高いほど、すなわち撮影機器間の特徴量のハミング距離が小さいほど、Domino Streaming++ のトラフィック削減効果は増加する。一方で、類似度が低くなるほど、Domino Streaming++ のトラフィック削減効果は小さくなる。

表 7 から、異なるビデオシーケンス間での平均ハミング距離は、いずれも 30 を超えていることが分かる。図 14 に、各ビデオシーケンスにおける 2 台目のカメラ映像を、各ビデオシーケンスの 1 台目のカメラ映像を用いてエンコードしたときの Independent Streaming に対するトラフィック削減率を示す。図 14 から、いずれのビデオシーケンスにおいても、異なるビデオシーケンスを用いた場合のトラフィック削減率は最大で約 1% 程度であることが分かる。すなわち、映像間のハミング距離が 30 を超えるとき、Domino

表 7 同じビデオシーケンスあるいは異なるビデオシーケンス間のハミング距離

Table 7 Hamming distance between same/different video sequences.

ビデオシーケンス	ハミング距離
Ballroom	23.9
Exit	25.8
Vassar	15.2
Ballroom と Exit	32.4
Ballroom と Vassar	31.3
Exit と Vassar	32.5

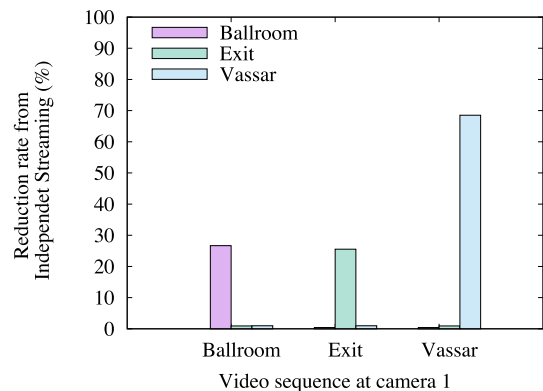


図 14 撮影機器間で異なるビデオシーケンスを参照した場合のトラフィック削減効果

Fig. 14 Traffic reduction rate with different reference video sequences.

Streaming++ によるトラフィック削減はほとんど期待できないと考えられる。

7.3 無線伝送レートと提案手法の関係に関する議論

Domino Streaming++ によるトラフィック削減は、ある任意の解像度・無線伝送レート下において、より視聴者満足度の高い無線撮影部の実現につながっている。たとえば、MERL が提供しているデータセットを用いた図 10(a) の結果から、解像度が QCIF (176×144)、無線伝送レートが約 800 [kbps] のとき、Domino Streaming++ は既存手法と比較して、品質を約 2 [dB] 高く維持したまま、映像を収集できることが分かる。高品質映像を用いたシステムは高い視聴者満足度を達成する [27], [28].

このような Domino Streaming++ によるトラフィック削減は、2K (1920 × 1080), 4K (3840 × 2160) などの高解像度映像、IEEE 802.11n や IEEE 802.11ac などの高伝送レートにも適用できる。本稿では評価の再現性の観点から MERL が研究用に提供しているデータセットのみを用いて評価している。しかしながら、Domino Streaming++ 自体は解像度や伝送速度に依存しない仕組みで実現している。たとえば、Domino Streaming++ で前提としている映像間の相関情報を用いた映像符号化は 2K, 4K などの高解像度

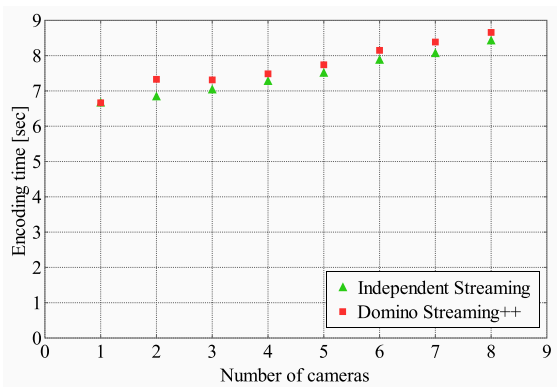


図 15 各撮影機器のエンコード時間
Fig. 15 Encoding delay in each camera.

映像でも用いられている [46].

7.4 エンコード処理時間に関する議論

撮影機器数の増加に対する各撮影機器のエンコード処理時間の変化を評価するため、JMVC エンコーダを用いて各撮影機器における 1 GOP 分のエンコードに要する処理時間を計測した。評価のために、ビデオシーケンスは Ballroom, GOP サイズを 8 フレームとした。計測機器として Panasonic CF-B11QWHBR (Intel Core i7-3635QM) を使用する。また、撮影機器の配置や撮影機器の間隔は、4.1 節の図 9 と同様の環境を想定している。撮影機器が 1 台のときは撮影機器 1 の位置に撮影機器が設置されており、撮影機器の台数が増えるに従って、撮影機器の番号順に設置されていく。

図 15 に、撮影機器数を増やした場合の各撮影機器におけるエンコード時間を示す。横軸が撮影機器数、縦軸が新たに追加された機器がエンコードに要する処理時間 [sec] を表す。ここで、Independent Streaming とは、他の撮影機器の映像を利用せず、自身の映像をエンコードする場合の処理時間を表す。図 15 から、撮影機器数が増加したとき、他の撮影機器の映像を利用しない場合に対して、映像間の相関を利用することによるエンコード処理時間の増加量はほぼ一定であることが分かる。Domino Streaming++ では、各撮影機器は直前にオーバーヒアした撮影機器 1 台の映像のみを参照し、それ以前に他の撮影機器の映像は参照しない。すなわち、機器数がいくら増えてもエンコード処理時間の増加は直前にオーバーヒアした映像との差分をとる処理時間のみとなる。

具体例として、図 16 (a) に撮影機器数 2 の場合の Domino Streaming++ のエンコード予測構造、図 16 (b) に撮影機器数 3 の場合の Domino Streaming++ のエンコード予測構造を示す。図 16 (a) における撮影機器 2 は、撮影機器 1 の映像を参照してエンコードする。図 16 (b) における撮影機器 3 は撮影機器 1、撮影機器 2 に続いて映像を送信するが、参照するのは撮影機器 2 の映像のみを利用する。

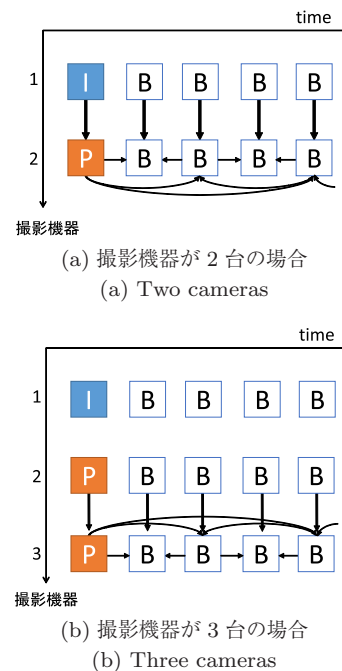


図 16 Domino Streaming++ エンコード予測構造
Fig. 16 Prediction structure of our proposed Domino Streaming++.

また、図 15 から、撮影機器数が増加したとき、Domino Streaming++ では、新たに加わった撮影機器のエンコード処理時間が増加しているように見える。しかしながら、Independent Streaming も同様の傾向を示しているため、エンコード処理時間の増加は提案手法によるものではないと考える。両手法で処理時間が増加した要因としては、評価に用いた Ballroom では、各撮影機器のデータサイズが撮影機器 1 から順に大きくなっており、エンコード後の出力処理などに時間を要したためだと考えられる。

7.5 伝送遅延に関する議論

Independent Streaming と Domino Streaming++ における遅延量は、GOP あたりのデータ量、伝送速度、エンコード処理時間に依存する。式 (1), (2) に、撮影機器数 N における Independent Streaming ならびに Domino Streaming++ の伝送開始時から終了時までの遅延量 $T_{\text{independent}}$ [sec], T_{domino} [sec] を示す。

$$T_{\text{independent}} = N \frac{D}{R} \quad (1)$$

$$T_{\text{domino}} = \{1 + (N - 1)\alpha\} \frac{D}{R} + (N - 1)\beta T_{\text{encode}} \quad (2)$$

D [KB] は 1 GOP あたりのデータ量、 R [KBps] は伝送速度、 T_{encode} [sec] は各撮影機器が独自にエンコードした場合にかかる処理時間を表す。 α は撮影機器数増加時のデータ量の削減率、 β は撮影機器数増加時のエンコード処理時間の増加率である。 α は、4.3 節中で示した撮影機器数 3 台、4 台、6 台時のトラヒック量に加えて、撮影機器数 8 台のときまでの Independent Streaming および Domino Streaming++

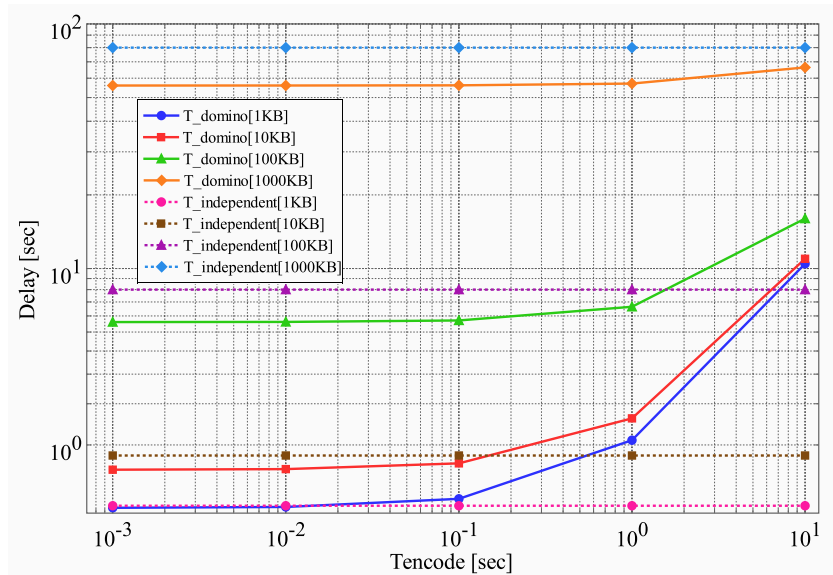


図 17 エンコード時間に対する処理遅延

Fig. 17 Analysis of total delay with different encoding delay.

のトラフィック量を評価した後、これらのトラフィック量をもとに算出する。具体的には、撮影機器数が i 台のときの Independent Streaming および Domino Streaming++ のトラフィック量をそれぞれ $B_i^{(independent)}$, $B_i^{(domino)}$ としたとき、以下の式で α を算出する。

$$\alpha = 1 - \left\{ \frac{1}{N_\alpha - 1} \sum_{i=2}^{N_\alpha} \frac{B_i^{(independent)} - B_i^{(domino)}}{B_i^{(independent)}} \right\}$$

ここで N_α は α 算出時に用いた最大の撮影機器数であり、本評価では 8 としている。

β は、7.3 節中の図 13 で示した撮影機器数に対する Independent Streaming および Domino Streaming++ のエンコード処理時間から算出している。具体的には、撮影機器数が i 台のときの Independent Streaming および Domino Streaming++ のエンコード処理時間をそれぞれ $E_i^{(independent)}$, $E_i^{(domino)}$ としたとき、以下の式で β を算出している。

$$\beta = 1 - \left\{ \frac{1}{N_\beta - 1} \sum_{i=2}^{N_\beta} \frac{E_i^{(independent)} - E_i^{(domino)}}{E_i^{(independent)}} \right\}$$

ここで N_β は β 算出時に用いた最大の撮影機器数であり、本評価では 8 としている。ここでは、各撮影機器のデータサイズならびにエンコード処理時間は同じであると仮定する。

Independent Streaming では、伝送前に全撮影機器が並列に自身の映像をエンコードする。エンコード後は、各撮影機器が自身の映像を逐次伝送する。Domino Streaming++ では、伝送前に撮影機器 1 台が自身の映像をエンコードする。後続の撮影機器はエンコード時に他の撮影機器 1 台の映像を利用する。映像間の相関を利用することで、2 台目以

降のデータ量は小さくなる。エンコード後は、Independent Streaming と同様に、各撮影機器が映像を逐次伝送する。

Domino Streaming++ が有効に動作するのは、式 (3) に示すように、 T_{domino} が $T_{independent}$ を下回る場合といえる。

$$N \frac{D}{R} > \{1 + (N - 1)\alpha\} \frac{D}{R} + (N - 1)\beta T_{encode} \quad (3)$$

式 (3) から、 $N = 1$ のとき、Independent Streaming と Domino Streaming++ の遅延量は等しくなる。 $N > 1$ のとき、Domino Streaming++ が Independent Streaming より有効である場面は、以下のように求めることができる。

$$\frac{D}{R} > \frac{\beta}{(1 - \alpha)} T_{encode} \quad (4)$$

式 (4) より、撮影機器 1 台分の伝送遅延が撮影機器 1 台分のエンコード処理時間の $\frac{\beta}{(1 - \alpha)}$ 倍よりも大きいとき、Domino Streaming++ が有効であるといえる。

上述の式の意味を議論するため、異なる伝送遅延、エンコード処理時間下での Independent Streaming および Domino Streaming++ の遅延量を評価した。図 17 は、各撮影機器における GOP あたりのデータ量 D とエンコード処理時間 T_{encode} を変化させたときの各方式の遅延量 $T_{independent}$ ならびに T_{domino} を示す。事前評価より、 α は 0.7、 β は 1.04 としている。横軸は T_{encode} [sec]、縦軸は伝送開始時から終了時までの遅延量 [sec] である。このとき、伝送速度は 100 [KBps]、 D は 1 [KB]、10 [KB]、100 [KB]、1,000 [KB] の 4 種類を利用する。図 17 から以下の 2 つのことが分かる。

1 つ目は、伝送速度に対する GOP あたりのデータ量が大きくなるにつれて、すなわち、伝送遅延が大きくなるにつれて、Domino Streaming++ が有効である場面が広がる点である。GOP あたりのデータ量は送信映像の品質と関

係する。したがって、各撮影機器が高品質な映像を送りたいとき、Domino Streaming++は効果的であるといえる。

2つ目は、伝送速度に対してGOPあたりのデータサイズが小さくなるに従って、Domino Streaming++が有効である場面が狭まる点である。データサイズが小さくなるにつれて、他の撮影機器を利用して削減できるデータサイズと伝送遅延が小さくなる。結果として、データサイズとエンコード処理時間がともに小さい場合にはDomino Streaming++とIndependent Streamingの差はほとんどなくなる。さらに、データサイズが小さく、エンコード処理時間が大きい場合にはIndependent Streamingが効果的になる。

8. 終わりに

本稿では、マルチビュービデオの撮影部を無線化する方式として、Domino Streaming++を提案した。Domino Streaming++では、傍受し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することで、トラフィック削減を達成する。また、アクセスポイントが撮影機器間の映像類似度に基づいて、各撮影機器の映像送信順を制御することで、撮影機器間の冗長な情報を多く削減でき、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++が無線マルチビュービデオにおけるトラフィック量の理想値に近い性能を達成していることが分かった。今後は、実環境に近い場面での提案手法の効果を明らかにするため、手ブレや移動による振動を含む映像、撮影機器ごとに解像度が異なる映像、撮影機器ごとに注目対象が変化する映像のようなテストビデオシーケンスを用意して評価することを考えている。

謝辞 本研究はJSPS 科研費 15K12018 の助成を受けている。

参考文献

- [1] Müller, K., Schwarz, H., Marpe, D., Bartnik, C., Bosse, S., Brust, H., Hinz, T., Lakshman, H., Merkle, P., Rhee, H. et al.: 3D high efficiency video coding for multi-view video and depth data, *IEEE Trans. Image Processing*, Vol.22, No.9, pp.3366–3378 (2013).
- [2] Tanimoto, M. and Kazuyoshi, S.: Global view and depth (GVD) format for FTV/3DTV, *Three-Dimensional Imaging Visualization And Display*, pp.1–10 (2013).
- [3] Vetro, A., Pandit, P., Kimata, H., Smolic, A. and Wang, Y.-K.: *Joint Draft 8.0 on Multi-view Video Coding* (2008).
- [4] Muller, K., Merkle, P., Schwarz, H., Hinz, T., Smolic, A. and Wiegand, T.: Multi-view video coding based on H. 264/AVC using hierarchical B-frames, *IEEE Picture Coding Symposium (PCS)*, Vol.2006 (2006).
- [5] Text Of ISO/IEC 14496-10:2008/FDAM 1 ISO/IEC JTC1/SC29/WG11: Multiview Video Coding (2008).
- [6] Xiaoyu, X., Gene, C. and Jie, L.: Frame structure optimization for interactive multiview video streaming with bounded network delay, *IEEE International Conference on Image Processing (ICIP)*, pp.593–596 (2011).
- [7] Ngai-Man, C., Antonio, O. and Gene, C.: Distributed source coding techniques for interactive multiview video streaming, *IEEE Picture Coding Symposium (PCS)*, pp.1–4 (2009).
- [8] Gene, C., Antonio, O. and Ngai-Man, C.: Interactive streaming of stored multiview video using redundant frame structures, *IEEE Trans. Image Processing*, Vol.20, No.3, pp.744–761 (2011).
- [9] Liu, Z., Cheung, G. and Ji, Y.: Unified distributed source coding frames for interactive multiview video streaming, *IEEE International Conference on Communications (ICC)*, pp.2048–2053 (2012).
- [10] Huang, H., Zhang, B., Chan, S.-H., Cheung, G. and Frossard, P.: Coding and replication co-design for interactive multiview video streaming, *IEEE International Conference on Computer Communications (INFOCOM)*, pp.2791–2795 (2012).
- [11] Pan, Z., Ikuta, Y., Bandai, M. and Watanabe, T.: User dependent scheme for multi-view video transmission, *IEEE International Conference on Communications (ICC)*, pp.1–5 (2011).
- [12] Pan, Z., Ikuta, Y., Bandai, M. and Watanabe, T.: A user dependent system for multi-view video transmission, *IEEE International Conference on Advanced Information Networking and Applications (AINA)*, pp.732–739 (2011).
- [13] 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).
- [14] Fujihashi, T., Pan, Z. and Watanabe, T.: Traffic reduction for multiple users in multi-view video streaming, *IEEE International Conference on Multimedia and Expo (ICME)*, pp.182–187 (2012).
- [15] Fujihashi, T., Pan, Z. and Watanabe, T.: UMSM: A traffic reduction method on multi-view video streaming for multiple users, *IEEE Trans. Multimedia*, Vol.16, No.2, pp.1–14 (2014).
- [16] Fujihashi, T., Pan, Z. and Watanabe, T.: Traffic reduction on multi-view video live streaming for multiple users, *IEICE Trans. Commun.*, Vol.96, No.7, pp.2034–2045 (2013).
- [17] Fehn, C.: *A 3D-TV approach using depth-image-based rendering (DIBR)*, Vol.3, No.3 (2003).
- [18] Mark, W.R., McMillan, L. and Bishop, G.: *Post-rendering 3D warping*, pp.7–16 (1997).
- [19] Hisatomi, K., Ikeya, K., Katayama, M., Iwadate, Y. and Aizawa, K.: Depth estimation based on stereo camera pairs of color and infrared using cross-based local midpoint filter, *Three Dimensional Systems and Applications (3DSA)*, Vol.1, p.3 (2013).
- [20] Ikeya, K., Hisatomi, K., Katayama, M., Mishina, T. and Iwadate, Y.: Bullet time using multi-viewpoint robotic camera system, *ACM Conference on Visual Media Production (CVMP)*, p.1 (2014).
- [21] Tanimoto, M.: Overview of free viewpoint television, *Signal Processing: Image Communication*, Vol.21, No.6, pp.454–461 (2006).
- [22] IEEE Computer Society: IEEE Standard for Information technology-Telecommunications and information exchange between systems Local and metropolitan area networks-Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications (2012).

- [23] Anand, A., Akella, A., Sekar, V. and Seshan, S.: A case for information-bound referencing, *ACM Special Interest Group on Data Communication (SIGCOMM) Hot Topics in Networks workshops (HotNets)*, p.4 (2010).
- [24] Anand, A., Balachandran, A., Akella, A., Sekar, V. and Seshan, S.: Enhancing video accessibility and availability using information-bound references, *ACM Special Interest Group on Data Communication (SIGCOMM)*, pp.345–356 (2013).
- [25] Joint Video Team Of ITU-T VCEG And ISO/IEC MPEG: JMVC (Joint Multiview Video Coding) Software (2008).
- [26] ISO/IEC JTC1/SC29/WG11: Multiview Video Test Sequences from MERL (2005).
- [27] Zinner, T., Abboud, O., Hohlfeld, O., Hossfeld, T. and Tran-Gia, P.: Towards QoE management for scalable video streaming, *International Teletraffic Congress Specialist Seminar on Multimedia Application, Performance and QoE*, pp.64–69 (2009).
- [28] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Processing*, Vol.13, No.4, pp.600–612 (2004).
- [29] 軍艦島モニタリングプロジェクト, 入手先 (<http://sarulab.inf.shizuoka.ac.jp/battleship/>) (2014).
- [30] Swain, M.J. and Ballard, D.H.: Color indexing, *International journal of computer vision*, Vol.7, No.1, pp.11–32 (1991).
- [31] Lowe, D.G.: Object recognition from local scale-invariant features, *IEEE International Conference on Computer Vision (ICCV)*, Vol.2, pp.1150–1157 (1999).
- [32] Oliva, A. and Torralba, A.: Building the gist of a scene : The role of global image features in recognition, *Progress in brain research*, Vol.155, pp.23–36 (2006).
- [33] Wang, G., Wu, K., Zhang, Q. and Ni, L.M.: SimCast: Efficient video delivery in MU-MIMO WLANs, *IEEE International Conference on Computer Communications (INFOCOM)*, pp.2454–2462 (2014).
- [34] Chakareski, J.: Viewpoint-popularity-driven uplink scheduling of multi-camera sensor arrays, *IEEE GLOBECOM*, pp.1638–1643 (2013).
- [35] Xun, G., Yan, L., Feng, W., Wen, G. and Shipeng, L.: Distributed multi-view video coding, Vol.6077, pp.1–8 (2006).
- [36] Artigas, X., Angeli, E. and Torres, L.: Side information generation for multiview distributed video coding using a fusion approach, *IEEE Nordic Signal Processing Symposium (NORSIG)*, pp.250–253 (2006).
- [37] Frederic, D., Mourad, O. and Touradj, E.: Recent advances in multiview distributed video coding, *Defense and Security Symposium (DSS)*, pp.1–11 (2007).
- [38] Slepian, D. and Wolf, J.K.: Noiseless coding of correlated information sources, *IEEE Trans. Information Theory*, Vol.19, pp.471–480 (1973).
- [39] Wyner, A. and Ziv, J.: The rate-distortion function for source coding with side information at the decoder, *IEEE Trans. Information Theory*, Vol.3, No.4, pp.45–49 (1976).
- [40] Chakareski, J.: Adaptive multiview video streaming: challenges and opportunities, *IEEE Communications Magazine*, Vol.51, No.5, pp.94–100 (2013).
- [41] CISCO Visual Networking Index 2013-2018: *Cisco visual networking index: white paper* (2013).
- [42] Stefanov, A. and Erkip, E.: Cooperative coding for wireless networks, *IEEE Trans. Commun.*, Vol.52, No.9, pp.1470–1476 (2004).
- [43] Sheng, Z., Leung, K.K. and Ding, Z.: Cooperative wireless networks: from radio to network protocol designs, *IEEE Communications Magazine*, Vol.49, No.5, pp.64–69 (2011).
- [44] An, D., Woo, H., Yoon, H. and Yeom, I.: Enhanced cooperative communication MAC for mobile wireless networks, *Computer Networks*, Vol.57, No.1, pp.99–116 (2013).
- [45] The Growth Of The Internet Of Things – platform, available from (<https://www.ncta.com/platform/industry-news/infographic-the-growth-of-the-internet-of-things/>)
- [46] Hetul, S.: *Low power architecture for motion compensation in a 4K Ultra-HD AVC and HEVC video codec system*, pp.400–404 (2013).



小寺 志保 (正会員)

2014年静岡大学情報学部情報科学科卒業。2016年同大学大学院情報学研究科情報学専攻修了。映像伝送、計算機ネットワークに関する研究に従事。2013年情報処理学会 MBL 研究会優秀論文賞。2015年情報処理学会山下

記念研究賞。



藤橋 卓也 (正会員)

2016年大阪大学大学院情報科学研究科情報ネットワーク学専攻博士後期課程修了。博士(情報科学)。2014～2015年 Mitsubishi Electric Research Laboratories 滞在。2014年より日本学術振興会特別研究員。映像伝送、計算機ネットワークに関する研究に従事。2012年 IEEE ICME (International Conference on Multimedia & Expo) ベストペーパー最終候補。2014年電気通信普及財団テレコムシステム技術賞。電子情報通信学会, IEEE 各会員。



猿渡 俊介 (正会員)

2007年東京大学大学院博士課程修了。科学博士。2003～2004年IPA未踏ソフトウェア創造事業, 2006～2008年日本学術振興会学振特別研究員, 2007～2008年イリノイ大学客員研究員, 2008年東京大学先端科学技術研究センター

助教, 2012年静岡大学大学院情報学研究科助教。2015年同大学講師。2013年より理化学研究所と産業技術総合研究所にて客員研究員。2016年より大阪大学大学院情報科学研究科准教授。専門はワイヤレスネットワーク, センサネットワーク, システムソフトウェア等。2009年電子情報通信学会論文賞。2010年情報処理学会山下記念研究賞。電子情報通信学会, IEEE, ACM各会員。



渡辺 尚 (正会員)

1982年大阪大学工学部通信工学科卒業。1987年同大学大学院博士後期課程修了。工学博士。同年徳島大学工学部情報工学科助手。1990年静岡大学工学部情報知識工学科助教授。2001年同大学大学院情報学部情報科学科教授。

1995年文部省在外研究員(カリフォルニア大学アーバイン校)。計算機ネットワーク, 分散システムに関する研究に従事。2005年情報処理学会モバイルコンピューティングとユビキタス通信研究会主査。2011年情報処理学会理事。2013年電子情報通信学会知的環境とセンサネットワーク研究会副委員長, 訳書「計算機設計技法」, 「802.11無線ネットワーク管理」等。電子情報通信学会, IEEE各会員。本会フェロー。