

## 2つのDVマルチキャストストリームを用いた 多地点ビデオ会議システムの提案

武智 充司 都築 伸二 山田 芳郎

愛媛大学大学院理工学研究科 〒790-8577 松山市文京町3

E-mail: {taketi, tsuzuki}@sarah.ee.ehime-u.ac.jp

あらまし DV over IPによる多地点ビデオ会議用中継方式について述べている。従来の中継器では、各地点の様子を同時にモニタするために画面の分割合成を行うことが多いが、中継回線の容量が大きいことが前提となる。一方提案方式は発言している人の映像のみを中継するため所要回線容量は小さくてもよい(平均30Mbps, 最大60Mbps)という特長がある。しかし、発言者つまり自分の顔を見ながらしゃべる必要があった。その対策として本稿では、所要回線容量は小さいままで2つのDVストリームをマルチキャストする方式を提案している。

キーワード DV (デジタルビデオ), DV over IP, 動画像ストリーミング, ビデオ会議システム, マルチキャスト

### A Proposal of Multi-site Videoconferencing System Using Two DV Multicast Streams

Mitsushi TAKECHI Shinji TSUZUKI and Yosio YAMADA

Graduate School of Science and Engineering, Ehime University

3 Bunkyo, Matsuyama, 790-8577, Japan. E-mail: {taketi, tsuzuki}@sarah.ee.ehime-u.ac.jp

**Abstract** The relay systems for multi-site video conferencing by DV over IP are described. Although a monitor screen has been divided into multi-portions in order to watch the circumstances of remote places for the conventional systems, the required network bandwidth for the relay has been assumed to be large. On the other hand, the proposing system has the feature that the required bandwidth is small (its average is 30 Mbps, and the maximum is 60 Mbps) since only the video of those who have spoken is transmitted. However, they had to talk while looking at their own faces. To solve the problem, a system using two DV multicast streams is proposed in this paper, while the required bandwidth is still small.

**Keyword** DV (Digital Video), DV over IP, Streaming of moving picture, Video conferencing system, Multicast

#### 1. はじめに

近年、光ファイバーによる高速・大容量なIP網の整備が進んでおり、DV(Digital Video)のような高精細動画像でもIP網上で伝送(DV over IP)が可能となっている。DV動画像は高精細映像・低遅延といった特長を有していることから、例えばDV over IP用フリーソフトウェアであるDVTS[1]を用いて、様々な映像コミュニケーションイベントが行われるようになってきた。筆者らも、DV over IP用ソフトウェアの開発や、それを用いた多地点ビデオ会議システム用中継器の提案・開発を行ってきた。文献[2][3]で提案した方式は、複数地点(中継器用PCのCPU能力の都合で収容地点数は4地点)に設置したカメラからの動画像ストリームを一旦中継器に集約し、発言している人の動画像を選択中継する方式であり、中継器を設置する場所の所要回線容量は平均約30Mbps, 最大でも2地点分のDVデータに相当する60Mbpsでよいという特長を有して

いた。複数地点の様子を同時にモニタするために必要な画面合成機能は有していないものの、中継器設置場所に必要な回線容量は2地点分の60Mbpsに低減することができる特長を有している。

本論文では、[2][3]を実装した結果とその問題点を述べ、改善方法を提案する。

#### 2. 従来の多地点ビデオ会議システム

##### 2.1. アナログ式中継システム

図1に送信器・中継器・受信器からなる従来の多地点ビデオ会議システム例を示す。図はDVカメラが設置されている地点数が3地点の場合である。各地点から中継器に向かって、DVストリームがIPで送信される。中継器では、一画面を分割し受信した各地の動画像を挿入した合成画面を作る。コストパフォーマンスの面からDVデータを一旦アナログ信号にDA変換し

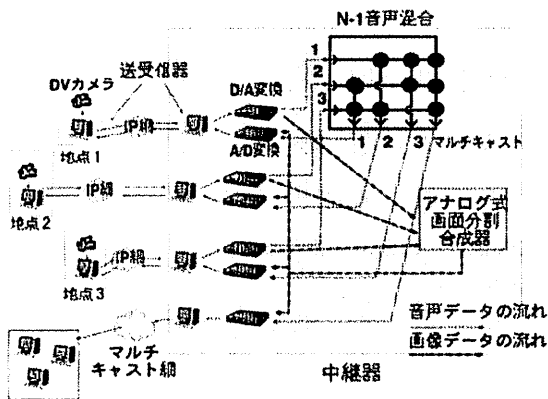


図1 アナログ式中継器

てこの分割合成を行った後再度 DV データに AD 変換することが多い。音声も同様に一旦、DA 変換を行い、アナログ信号で混合を行い、再度 DV に変換して中継する（本稿では、これらをアナログ式中継器と呼ぶ）。なお、図中の N-1 音声混合は、音声が悪くならないように送信地点の音声のみをミュートし、その他の地点の音声はすべて混合した音声を送り返すことを示している。ただし、N は送信地点数である。

N-1 音声混合を行うためには、高価な音声ミキサーが必要であり、配線が煩雑となる問題がある。また、DA・AD 変換により中継遅延時間が 0.1 秒程度増加する[2]。さらに中継地点には、30Mbps×N の DV ストリームが集中するため、十分な帯域を持った回線を用意する必要がある。また映像コミュニケーションイベント等では、画面分割、画面切替あるいは音声混合のため専門のオペレーターを中継地点に常駐させなければならないといった運用上の問題もある。

## 2.2. デジタル式中継器

DV ストリームを中継する際に前述のようなアナログ信号を用いず、デジタル信号のまま中継できる（デジタル式と呼ぶ）ビデオ会議システムとしては、ISID 社の QuallImage/Quatre がある[4]。Quatre は送受信クライアントとして、DVTS を用いており、収容可能な送信器数は 4 で、フルデジタルの画面分割合成・N-1 音声混合機能を備えている。しかし、中継器の設置場所には最大で 30Mbps×4 の回線容量が必要であり、使用する PC には高いスペックが要求されている (Xeon3.0GHz×2, Full duplex150Mbps 以上の NIC)。また、画面の中継地点の選択切替は手動であるので依然としてオペレータの介入が必要である。

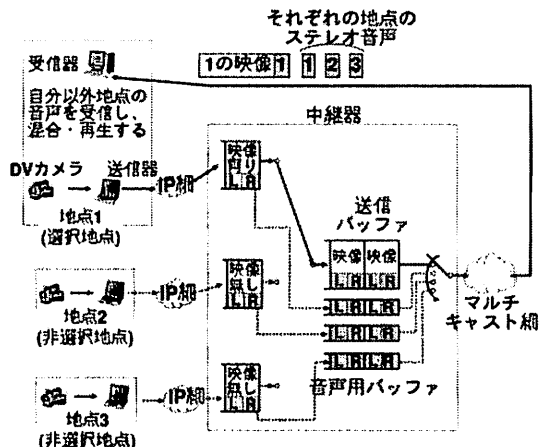


図2 多地点会議モード

## 3. 実装済中継システムの仕様

筆者らの実装済の中継システムでは、多地点間でのビデオ会議を想定した多地点会議モードと、複数のカメラ映像の中から 1 つを選択して中継する放送モードがある。以下に 2 つのモードを示す。

### 3.1. 多地点会議モード

多地点会議モードを図 2 に示す。多地点会議モードでは、有音である地点(発言している地点)を自動的に選択して送信する。各地点の送信器では 1 フレーム毎に音声エネルギーを算出し、そのエネルギーが閾値よりも高いか低いかで有音か無音かを判定している。無音か有音か判定する閾値は任意に設定することができる。無音であると判定された場合、各パケットの RTP ヘッドに無音であることを示すフラグが立つ。中継器では送信器から送られてくるパケットの無音フラグをモニタリングして、どの地点を中継するか選択する。3 秒中 2 秒以上有音であればその地点を選択・中継する。

無音であると判定された非中継地点は、画像を送信し続ける必要がないので、音声のみを送信する。ここで、画像と音声の両方を送信するモードを AV 送信モードと呼び、音声のみ送信するモードを AO 送信モードと呼ぶ。この両モードは、送信器と中継器の間で確立してある UDP コネクションを用いて、中継器が送信器に対して切り替えを要求する[3]。非選択地点が選択されていないにも関わらず音声のみを中継し続ける理由は、割り込み発言など、複数地点で同時に喋る可能性があるからである。

中継地点切替り時のアルゴリズムを図 3 に示す。地

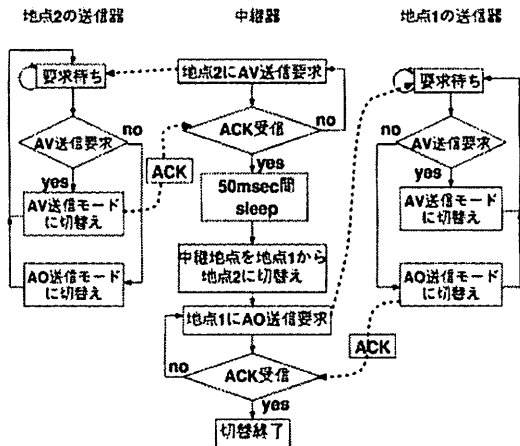


図 3 実装済中継システムの  
地点切替のアルゴリズム

表 1 中継システムの比較

|                   | アナログ中継システム                        | ISID社Quatre                       | 実装済中継システム             |
|-------------------|-----------------------------------|-----------------------------------|-----------------------|
| 遅延時間              | ×大(最大3秒程度)                        | ○小(約0.3秒)                         | △中(0.6秒程度, 中継器内で0.2秒) |
| 画質                | ○劣化(DV→NTSC変換による)                 | ○劣化無し                             | ○劣化無し                 |
| 多地点同時表示           | ○                                 | ○                                 | ○                     |
| 中継システム規模          | ×N個のPC(Pentium4×1程度)とアナログ機器が必要    | △PC(Xeon×2程度)1台                   | ○PC(Pentium4×1程度)1台   |
| 画面切替              | ×手動                               | ×手動                               | ○自動                   |
| 中継システム設置場所の所要回線容量 | ×(地点数N分のトラフィックを収容できる程度, 30Mbps×N) | ×(地点数N分のトラフィックを収容できる程度, 30Mbps×N) | ○(2地点分程度, 最大60Mbps)   |

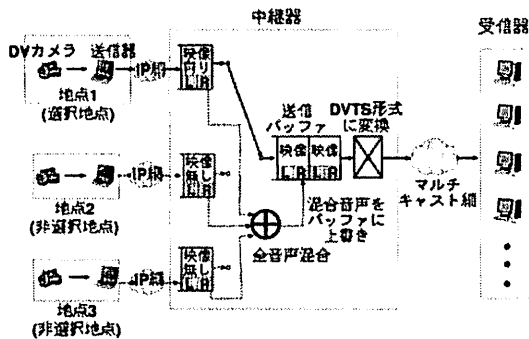


図 4 放送モードの構成

地点が1から2に替わるときを示している。まず、次の中継地点である地点2に30fpsで動画像を送信するように要求する。要求が届いたことを示すACKを受信したら、地点2の1フレームを受信するために50msの間sleepする。sleepした後に、地点を切替え、直前ま

で中継していた地点である地点1をAO送信モードに切り替える。

2地点以上が有音である場合は、有音の地点の映像をラウンドロビン式に選択中継する。画面切り替えの間隔を短くすると、画面が見苦しくなるため、切替え間隔を3秒としている。また、すべての地点が無音となった場合は最後に中継した地点の保持して中継する。

中継器は、受信した画像を1フレーム100packetに分割してマルチキャストで送信する。そして、選択地点の1フレームを送信し終えた直後に、全地点の音声データをマルチキャストで送信する。受信器では自分の地点以外の音声を受信し、混合・再生する。これにより、N-1音声混合と同等の再生がなされる。

### 3.2. 放送モード

放送モードを図4にその構成を示す。有音である地点が自動的に選択される。手動で中継地点を選択することも可能である。中継器ではフレーム単位で全地点の音声データが混合され、送信バッファの音声データ部分にその混合音声が入り書きされる。全地点の音声データが混合された動画像データはDVTSフォーマットに変換され、マルチキャストで送信される。受信はDVTSを用いて受信・再生する。

表1にそれぞれの中継システムの比較を示す。提案の中継システムでは、中継システム規模が小さく、Pentium4 一個程度のPCで動作することと、所要回線容量が少ないのが特長である。

### 3.3. 実装済中継システムの入力ストリーム量

Quatreのようなフルデジタルの多地点会議システムではすべての地点が30fps動画像を送信しているため、3地点を収容すれば中継器に入力されるストリーム量は90Mbpsを超える。一方、実装済中継システムでは、発言していない地点は音声のみを送信しているため、中継器に入力されるストリーム量を抑えることができる。ここで、中継器の入力ストリーム量を明らかにするために、DVのフレームレートと伝送レートとの関係を示す。フレームレート(fps; frame per second)変換は画像1フレームを間引き、画像データの代わりに音声データを送信することによって行う。つまり、mフレーム中1フレームだけは間引かずフレーム内のデータを送信し、それ以外のフレームは、音声データのみを送信する。このmを間引き係数とよぶ。

次に伝送レートの計算式を示す。29.97[frame/s]はフルレート(m=1)時のDV画像の1秒間に送信されるフレーム数である。120,000[Byte]は1フレームのデータ

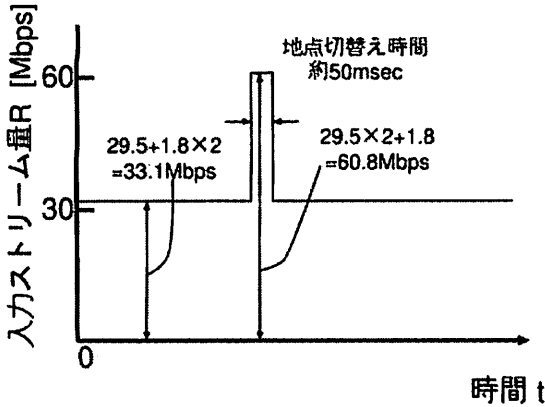


図 5 実装済中継器の入カストリーム量 (N=3)

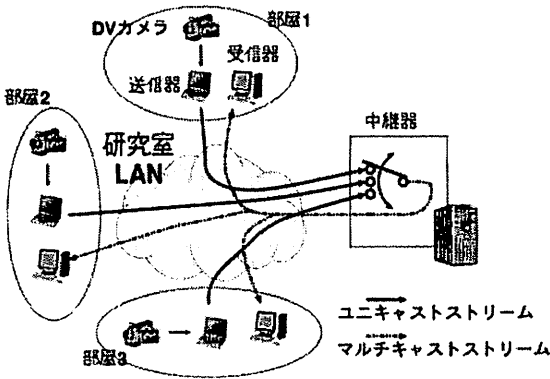


図 6 多地点会議モードでの使用例

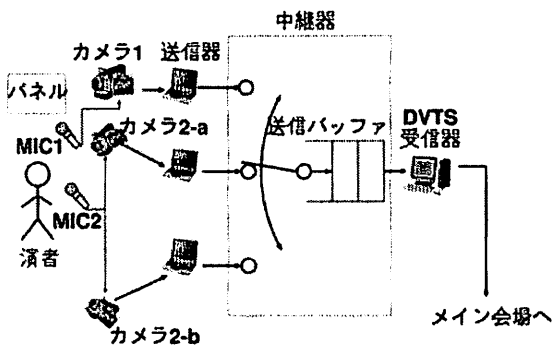


図 7 放送モードでの使用例

サイズ, 32[Byte]はヘッダサイズ, 100 は画像 1 フレームの分割パケット数, 7,200[Byte]は 1 フレーム内の音声データサイズ, 10 は音声データ 1 フレーム分の分割パケット数である。

伝送レート

$$= \left\{ \frac{29.97[\text{frame/s}]}{m} \times (120,000[\text{Byte}] + 32[\text{byte}] \times 100) + (29.97[\text{frame/s}] - \frac{29.97[\text{frame/s}]}{m}) \right.$$

$$\left. \times (7,200[\text{Byte}] + 32[\text{Byte}] \times 10) \right\} \times \frac{8[\text{bit}]}{10^6} \text{ Mbps}$$

すなわち、フルレートの時の伝送レートは、29.5Mbpsである。

また、音声のみ送信しているときの伝送レートは  $29.97[\text{frame/s}] \times (7,200[\text{Byte}] + 32[\text{Byte}] \times 10)$

$$\times \frac{8[\text{bit}]}{10^6} = 1.8\text{Mbps}$$

となる。

以上から、中継器に集中する DV データトラフィックは平均的に、 $29.5 + 1.8 \times N[\text{Mbps}]$  となり、画面切替時の 50m 秒の sleep 中は、 $29.5 \times 2 + 1.8 \times (N-1)[\text{Mbps}]$  となる。図 5 に地点数 N が 3 の時の中継器の入カストリーム量を示す。図が示すように、中継器の入カストリーム量は地点切替の 50m 秒の間は  $29.5 \times 2 + 1.8 = 60.8\text{Mbps}$  となる。

## 4. 実装済中継システムの使用結果

### 4.1. 多地点会議モード

多地点会議モードで使用し、それによって明らかとなった問題点を述べる。使用例を図 6 示す。研究室の 3 部屋それぞれに DV カメラ、送信器、中継器を設置した。それぞれの部屋の送信器はユニキャストで DV パケットを中継器に送信し、中継器は選択地点の動画をマルチキャストで送信した。各受信器はマルチキャストで中継動画・音声を受信する。

3 部屋間でそれぞれ発言しあい、動作の確認を行ったところ、発言者が自動的に選択中継され、音声は自分の声以外が混合されており、ビデオ会議システムとして機能することを確認した。

しかし、本来であれば、発言者は聴衆の顔を見ながら発言するのであるが、中継器では発言者の動画像を選択中継するため、自分の顔を見ながら発言することになり、カメラ目線で発言しにくい。また、地点数が 2 地点しかないときは、自分が発言すれば自分の顔が映し出されてしまい、テレビ電話のような使い方ができない。

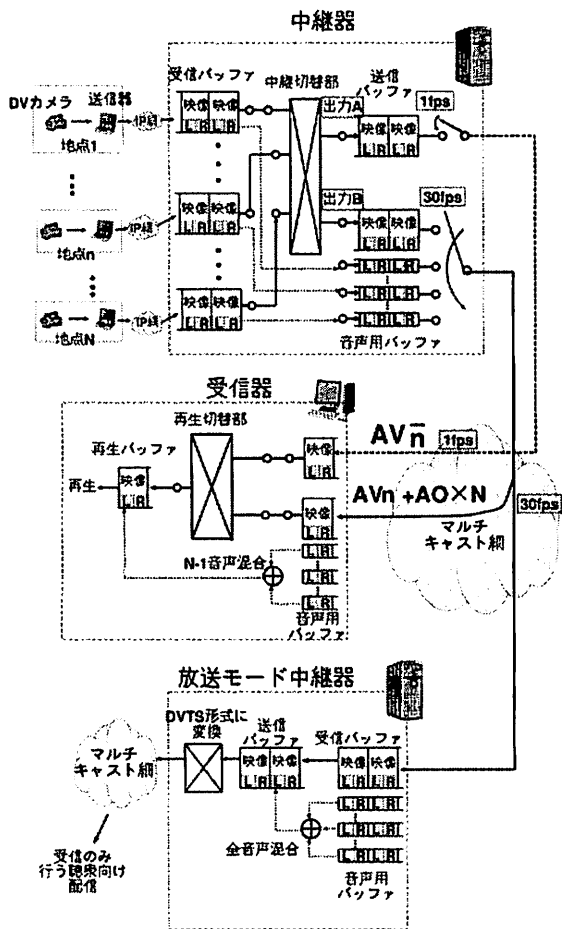


図 8 提案する 2 つのマルチキャストストリームを用いた中継システム

## 4.2. 放送モード

実装済中継システムを放送モードで、実用した結果を示す。構成を図 7 示す。

2006 年 11 月 22 日に行われた四国 JGN2 セミナーにおいて、愛媛大学からの映像をメイン会場に伝送するとき用いた。愛媛大学には 3 台のカメラを設置する。1 台はパネルのみを、残り 2 台は 2 方向から人物を撮影する。パネル用カメラには 1 本のマイク (MIC1) を接続し、人物用カメラには共通のマイク (MIC2) をそれぞれのマイク端子に接続する。送信器は 3 台とも無音検出を有効としている。したがって、パネルを説明したい場合は、MIC2 は OFF にし、MIC1 を使って発言しつづければパネルの映像が自動的に選択中継される。また、人物の映像を送信したい場合は、MIC2 を用いて発言する。カメラ 2-a とカメラ 2-b には MIC2 から同じ

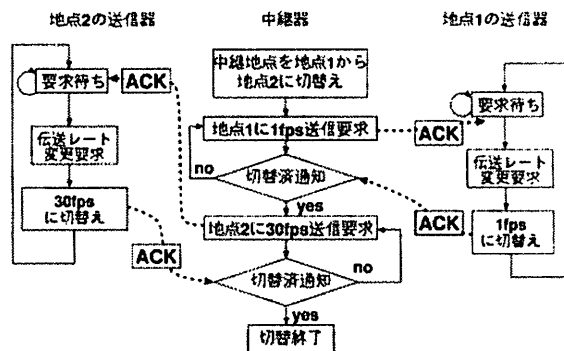


図 9 提案手法における地点切替のアルゴリズム

音声が入力されるので、両方とも有音となる。したがって、MIC2 で発言をしている間は、カメラ 2-a とカメラ 2-b の映像が 3 秒間隔で交互に切り替わって中継される。選択映像は会場に送信する前に一旦、DVTS 受信器で受信した。

図 7 の構成で実際に運用してみたところ、3 カメラによるマルチアングルの中継伝送が実現できた。イベントにおける映像伝送は 1 地点に対してカメラ 1 台が一般的であり、カメラアングルは固定になってしまう。1 地点にカメラを数台設置し、マルチアングルで映像を送送しようとする、アナログの映像スイッチャー、音声ミキサーやオペレーターが必要となる。それに対し、図 7 では、セットアップさえ行ってしまうと、発言するのにマイクを代えるだけでカメラアングルが変わるので簡単であった。

## 5. 2 つのマルチキャストストリームを用いたビデオ会議システムの提案

### 5.1. 概要

多地点会議モードを実使用した際に判明した問題に対する対策として以下の手法を新たに提案する。その構成を図 8 に示す。

発言している地点は実装済中継システムと同様に 30fps で動画を送信する。また、発言していない地点もフレームレートを落として動画を送信する。ここでは、例として 1fps とする。中継器には 1 出力 A (1fps ストリーム) と出力 B (30fps ストリーム) の 2 本を設ける。出力 A には発言していない地点の動画を出力する。出力 B には発言している地点の動画を出力する。出力 A には、発言者が複数の地点の様子を見ながら発言できるようにするために、発言していない地点の動画をラウンドロビン式に出力する。出力 A, 出力 B

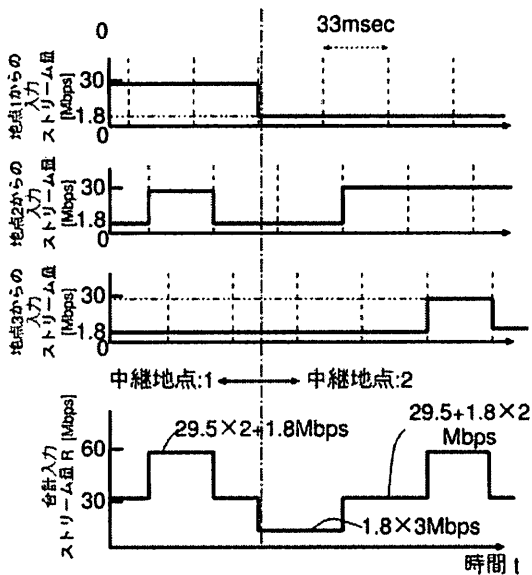


図 10 提案手法における中継器の入力ストリーム量(N=3)

それぞれの出力切替は中継切替部で行う。

各パケットの RTP ヘッドには、どの地点の動画像か区別できるようにするため、地点番号を挿入する。受信器は、出力 B の動画像が自分の地点から送出したものであれば出力 A を再生し、そうでなければ、出力 B を再生する。受信器の再生切替は再生切替部で行う。これによって、発言している地点では、自分の顔を見ながら発言しなくなるので、カメラ目線でも発言しやすくなる。

また、受信器の代わりに放送モードの中継器を設けることにより、受信のみ行う聴衆向けにストリームを配信することも可能となる。

## 5.2. 提案手法の地点切替のアルゴリズムと中継器の入力ストリーム量

地点切替時のアルゴリズムを図 9 に示す。実装済中継システムは、次の中継地点の 1 フレームを受信するために 50msec の sleep があり、地点切替の反応を鈍くさせる要因となっていた。しかし、提案手法では発言していない地点も 1fps ではあるが定期的に動画像を送送しているので、即座に地点を切替えることができる。

提案手法における中継器の入力ストリーム量を図 10 に示す。収容地点数は 3 地点である。図 9 に示すアルゴリズムと同様に、地点が 1 から 2 に切替るときで示している。また、タイムスロットを 1 フレームの送

信時間である 33m 秒で区切り、33m 秒間平均での入力ストリームとしている。発言していない地点は 1fps であるが、30 フレーム中 1 回は動画像データを伝送するので、33m 秒は瞬間的に 29.5Mbps(30fps)となる。また、発言している地点は常に 29.5Mbps で伝送しているので瞬間的に約 60Mbps の入力ストリーム量となる。また、切替り時にはすべての地点が 1fps となるので瞬間的に約 5Mbps となる。このことから、発言していない地点からも 1fps で動画像を送送しても、実装済中継システムと同様の入力ストリーム量で済むことわかる。

## 6. おわりに

本稿では、文献[2][3]で提案した 2 つの中継モードの仕様と使用結果を述べた。放送モードでは良好な結果を得たものの、多地点会議モードではカメラ目線で発言しにくいという問題があった。その対策として、実装済中継システムと同様に帯域を節約しつつ、中継する動画像を 2 ストリームとする方式を提案した。

今後の課題は提案手法の実装と検証である。

また、現在の実装では有音検出のためのサンプリング時間を 3 秒としており、発言し始めてから地点が切替るまでの時間が長いと感じられるため、改善する必要がある。

## 文 献

- [1] "DVTS", <http://www/sfc/wide.ad.jp/DVTS/index-j.html>
- [2] 武智, 都築, 山田:"デジタルビデオ多地点会議の地点自動選択方式およびエコー防止方式", マルチメディア, 分散, 協調とモバイル(DICOMO2006)シンポジウム, pp.597-600, July 2006.
- [3] 武智, 都築, 山田:"多地点 DV カメラ映像の選択中継方式に関する基礎検討", 平成 18 年度電気関係学会四国支部連合大会, 12-26, Sep. 2006.
- [4] 株式会社電通国際情報サービス, "QualImage/Quatre", <http://www.isid.co.jp/QualImage/quatre.html>