

## マルチメディア遠隔提示システムの同期条件の検討

大野 隆一, 相田 仁, 齊藤 忠夫

東京大学工学部

マルチメディア情報を用いて多地点間での会議, 教育など様々な活動を支援するシステムに対する期待が高まっている。

本論文では, このようなマルチメディア遠隔提示システムを構築する場合に, 各ノードでの情報提示の際のメディア間の同期を維持するための条件について検討する。論文中では, まず, 蓄積型情報源から取り出されるメディア情報と参加者の声などの生のメディア情報との間の情報提示の際のタイミングに基づいた同期のセマンティクスを定める。次に, このセマンティクスを満たすために各ノードにおいて必要なバッファ量の算出法を示す。これを用いて, バッファサイズとコマンド応答時間の間の設計条件を明らかにすることができた。

キーワード マルチメディア, 同期条件, 遠隔提示システム, 会話型操作

## A Study of Conditions for Synchronization in Multimedia Remote Presentation Systems

Ryuichi Ohno, Hitoshi Aida, and Tadao Saito

Faculty of Engineering, the University of Tokyo

In this paper, we study conditions for synchronization in multimedia remote presentation systems, in which Retrieved Media Streams (RMS) such as media streams retrieved from video equipment and Live Media Streams (LMS) such as raw voices of some participants are presented at remotely located nodes. At first, semantics of synchronization between RMSs and LMSs is defined. Then, a method to calculate buffer sizes required for the semantics is presented. Compromise condition between buffer size and command response time is analyzed to give an appropriate design.

keywords multimedia, synchronization condition, remote presentation system, interactive control

## 1 まえがき

近年、ネットワーク技術及びコンピュータ技術の発展に伴い、ネットワークを介してマルチメディア情報を送ることで多地点間での会議、教育など様々な活動を支援するシステムの研究・開発が盛んに行なわれてきている。

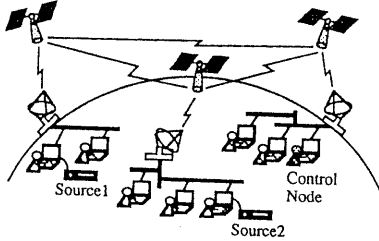


図 1: システム環境

このようなマルチメディア遠隔提示システムで用いられるネットワークとしてはB-ISDN、衛星回線など様々なものを考えることができるが、どのようなネットワークを用いるにしろネットワークの伝送遅延を避けることはできない。このようなネットワークの伝送遅延により以下のような問題点が生じる。

- ネットワーク中に分散して配置されているディスク、テープ、ビデオカメラなどの様々な情報源から送られてくる複数のメディア情報を各ノードで提示する際にメディア間のタイミングのずれが大きくなる。
- ディスク、テープなどの蓄積型情報源に対して一時停止などといった操作命令を出した際に、指示してから実際にその命令が情報提示に反映されるまでの遅延が大きくなる。

情報提示を行なうノードにおいてメディア情報をメモリ中にキャッシュすることは上記の2つの問題点のいずれにも有効な手法となる。

本論文では、複数のノードにおいてディスク、テープなどの蓄積型情報源から引き出されたメディア情報 (Retrieved Media Stream と名付ける。以下、RMS と略称する。) とその説明を行なう参加者の声などの生のメディア情報 (Live Media Stream と名付ける。以下、LMS と略称する。) が提示されるようなシステムにおける RMS と LMS との間のメディア間同期の問題を取り上げる。

論文では、まず2章で、想定するシステムの概略について述べる。次に3章で、RMS と LMS との間の情報提示の際のタイミングに基づいた同期 (以下、R&L 同期と略称する。) のセマンティクスを定め、4章でその

セマンティクスを満たすのに必要なバッファ量の算出法を示す。次に5章で幾つかの場合についての算出結果を示し、考察を加える。

## 2 システムの概要

### 2.1 システムの概念

各ノードにはディスプレイ装置などのマルチメディア表示用の装置、及び、メディア間のずれの吸収などに用いられるメモリが備わっている。また、システム中の幾つかのノードにはメディア情報の入ったディスク、テープなどの情報発生源となる装置が備わっている<sup>[3]</sup>。

そして、各ノードにおいて3章で述べる R&L 同期のタイミングを満たしながら同一のメディア情報が提示される。

### 2.2 想定する環境

ネットワーク ネットワークとしては予め資源を確保することで伝送遅延時間などがある範囲内に収まることを保証できる場合を考える。

端末 端末としてはこのシステムのための専用端末などのように本システムの正常な動作を保証するだけの資源があらかじめ確保できる場合を考える。

情報源 動画、音声などのような等間隔で情報が発生するメディアの入ったハードディスク、ビデオテープ装置などを考える。

### 2.3 同期に必要な付加情報

本システムでは、すべてのメディア情報源でのメディアフレーム発生間隔を33ミリ秒といった共通の値とする。各フレームにはフレーム番号が付けられ、同じフレーム番号を持つメディアフレームを同時に提示することでメディア間の同期を取る。

以下に同期情報を付加すべきメディア情報を示す。

情報源中に蓄積されている情報 等しい情報量のフレームが順番に並んでいるような場合には、1フレーム目の先頭の位置から他のフレームの位置もわかるので同期情報は必要ない。しかし、各フレームの情報量が異なる場合、フレーム番号と各フレームの先頭位置の対応関係を示す情報が必要となる。

RMS RMS 中の各フレームにはフレーム番号が付く。フレーム番号は1つのRMS中における各フレームと1対1に対応している。例えば、先頭のフレームから1,2,3,...といった順番で付けていく。

**LMS** LMS 中の各フレームを送信した際に提示されていた RMS 中のメディアフレームのフレーム番号をその LMS 中のフレームに付加する。各情報提示ノードでは LMS 中の各フレームもバッファリングし、RMS も含めて等しいフレーム番号のフレームを同時に提示することでメディア間同期を取る。ただし、3章の同期のセマンティクスにおいて sync\_with\_one, sync\_with\_multi でない場合には、厳密なメディア間同期を取ることはできないため、同期情報は必要ない。

**操作命令パケット** 操作命令を示す情報には停止位置、再生開始位置などを示すフレーム番号と命令が発行された時間を示す情報が付加される。

**バッファリングされている情報** バッファリングされている各フレームにもフレーム番号を付加する。

### 3 R&L 同期のセマンティクス

本章では、R&L 同期のセマンティクスを定める。

#### 3.1 システム条件

想定するシステムとしては図2のような、一つの操作者ノード、複数の情報発生源ノード及び情報提示ノードからなるものを考える。

ここで、操作者ノードを1、情報発生源ノードを  $i (2 \leq i < m)$ 、情報提示ノードを  $j (j \geq m)$  と表記し、1から2までのノード間遅延を  $T_{12}$  のように表記する。また、情報源を各命令に対してセットアップするのに必要な時間（ハードディスクのシークにかかる時間、ビデオテープ装置の機械的な動作にかかる時間など）を  $T_{set}$  とし、これら  $T_{12}$ 、 $T_{set}$  などの時間の上限及び下限はあらかじめわかっているものとする。

また、メディア情報の圧縮方式としてはフレーム内圧縮方式のみを考えて以下の議論を進める。フレーム間圧縮を用いる場合については6章で述べる。

メディア情報の圧縮伸長にかかる時間は LMS の場合  $T_c$  とする。RMS が圧縮された情報の場合、バッファ中には圧縮された情報をそのまま格納する。伸長は各ノードにおいて実際にメディアフレームが提示される際に行なわれる。この、伸長にかかる時間を  $T_{dcj} (2 \leq j < m)$  と表記する。また、情報源には圧縮された情報が格納されているものとする。（圧縮が必要な場合にも、圧縮にかかる時間は  $T_{set}$  に含めて考えることができる。）

また、再生中には future バッファを常に  $T_d$ （ネットワークの遅延変動を吸収するのに必要最小限の大きさ）以上に保つものとする。

再生点に関するずれは  $D_{12}$  などと表記する。また、各情報源でのメディア発生間隔を  $T_k$  と表記する。

まず、以下の議論を進めていく上での仮定を述べる。

仮定1:  $T_{1i} + T_{set} + T_{ij} \geq T_{1j} (2 \leq i < m, j \geq m)$  つまり、操作命令は必ずその命令により位置が変化したフレームよりも先に到着するものとする。

仮定2: 各ノードでは確実に (30frame/sec などの) 同じ速度で画面の走査が行われているものとする。

仮定3: 操作者ノードでの命令発行は、各走査が完了する時点 (33.3msec ごと等) においてしか発行されないとする。

仮定4: あるフレームはスキヤンされる直前の時点においてそのフレーム情報を完全にバッファ中に持っている必要があるものとする。

次に、各ノードでの状態と操作命令を図2に示す。

状態 命令	一時停止	再生	逆再生	早送り	巻き戻し
一時停止	×	◎	○	○	○
再生	◎	×	×	×	×
逆再生	○	×	×	×	×
早送り	○	○	○	×	○
巻き戻し	○	○	○	○	×
順コマ	◎	×	×	×	×
逆コマ	◎	×	×	×	×

図2: ノードの各状態で使用できる命令

この図において、◎または○で示してある部分はその状態に対して対応する命令が実行可能であることを示し、×に対応する命令は許されないとする。また、◎で示してある部分は後で述べる R&L 同期のセマンティクスを満たすものとする。

#### 3.2 R&L 同期のセマンティクス

LMS と RMS の情報提示時のタイミングに基づき R&L 同期のセマンティクスを以下のように定義する。ここで、RMS 間の同期はすべて厳密に取るものとする。

sync\_with\_one RMS と操作者からの LMS との同期。以下の式が満たされることが条件となる。

$$D_{i1} + T_{1i} + T_c \leq 0 \quad (i = 2, 3, \dots)$$

sync\_with\_all RMS と全ての参加者からの LMS との（厳密でない）同期。すべてのノードでの再生点が各時刻において等しいこととする。つまりすべての  $i, j (\geq 1)$  に対して  $D_{ij} = 0$ 。

**sync\_with\_multi** RMS と、操作者を含む複数の参加者（発信者ノード）からの LMS との同期。発信者ノードに対しては  $D_{ij} = 0$  ( $1 \leq i, j \leq N_s$ ) それ以外のノードに対しては  $D_{ji} + T_{ij} + T_c \leq 0$  ( $1 \leq i \leq N_s, j > N_s$ ) が満たされることが条件となる。（**sync\_with\_one**, **sync\_with\_all** は **sync\_with\_multi** の特別な場合と考えることもできる。）

**sync\_with\_none** RMS と LMS との間の同期は考えず、RMS 間の同期のみ保証する。

また、R&L 同期における各ユーザへの情報提示を以下のように規定する。

- 全てのノードにおいて少なくとも操作者ノードと同じだけの情報が提示される。（従って、各ノードで一時停止命令を受けとった際に、指定された位置までまだ情報提示が完了していない場合、その位置まで情報提示を行なってから停止する。）
- 再生命令に対しては、各ノードで R&L 同期のための再生点調整を行なった後、LMS との同期が取れた状態で情報提示が始められる。
- 一時停止、コマ送り命令に対して各ノードでは命令到着後、即座に反応する。コマ送り中には RMS と LMS との同期は保たれる必要はない。命令実行後、一定時間が経過した後、LMS との同期が回復する。

各ノードで再生命令を受け取った際の再生点調整を行なうのに必要な待ち時間を以下に示す。

$W_{rmaz} = \text{MAX}(0, D_{21} + T_{12}, D_{31} + T_{13}, \dots)$  とすると

$$W_{r1} = W_{rmaz} \quad (\text{ノード1})$$

$$W_{ri} = W_{rmaz} - (D_{i1} + T_{1i}) \quad (\text{ノード2} \sim N)$$

各ノードで一時停止、コマ送り命令実行後、LMS との同期回復に必要な待ち時間を以下に示す。

$$temp_i = T_{1i} \quad (D_{i1} + T_{1i} > 0)$$

$$temp_i = 0 \quad (D_{i1} + T_{1i} \leq 0)$$

$$W_{rmaz} = \text{MAX}(0, temp_2, temp_3, \dots) \text{ とすると}$$

$$W_{r1} = W_{rmaz} \quad (\text{ノード1})$$

$$W_{ri} = W_{rmaz} - T_{1i} \quad (D_{i1} + T_{1i} > 0)$$

$$W_{ri} = 0 \quad (D_{i1} + T_{1i} \leq 0) \quad (\text{ノード2} \sim N)$$

また、命令を受け取り、必要な時間待ってからすぐに同期回復とするのではなく、さらに各ノードで等しい時間待たせることで必要なバッファサイズを削減できる可能性がある。この際の待ち時間を  $W_s$  とする。

さらに、各ノードで情報提示する際に、伸長が必要なメディアの情報提示は伸長に必要な時間だけ遅れることになる。この遅れを吸収するために、伸長が必要でない

(または伸長にかかる時間が短い)メディアの情報提示を待たせる必要がある。この待ち時間  $W_{di}$  を以下に示す。

$$W_{dmaz} = \text{MAX}(T_{dc2}, T_{dc3}, \dots) \text{ とすると}$$

$$W_{dj} = W_{dmaz} - T_{dcj} \quad (2 \leq j < m)$$

## 4 R&L 同期を満たすために必要なバッファ量の算出式

本章では R&L 同期を満たすために必要なバッファ量を求めるための計算式を導き出す。

各ノードでは各情報源ごとにバッファを持っているので、図3のように、システムは情報発生源ノードごとに分割して考えることができる。

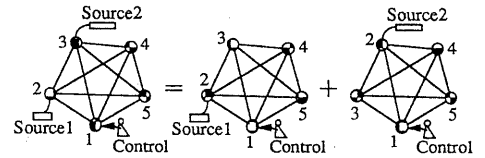


図3: システムの分割

そこで、以下、図3の分割後のような1つずつの操作者ノードと情報発生源ノード及び複数の情報提示ノードからなるシステムについて考えていく。

ここで、図3のように分割後のシステムで操作者ノードが1、情報発生源ノードが2、情報提示ノードが  $i$  ( $i \geq 3$ ) となるように番号を振り直す。また、 $T_{dcj}, W_{dj}$  は  $j$  を省き、 $T_{dc}, W_d$  と表記する。

以下に R&L 同期を満たすために必要なバッファ量を算出するための式を示す。

### 4.1 一時停止状態

一時停止状態はすべてのノードで再生点が同じ位置で止っている状態と規定する。

- 再生命令

それぞれのノードで正しい再生が開始できる条件は

$$T_{12} + T_{set} + T_{21} - W_{r1} - W_s - W_d + T_k < F_1 \quad (\text{ノード1})$$

$$T_{set} - W_{r2} - W_s - W_d + T_k < F_2 \quad (\text{ノード2})$$

$$T_{12} + T_{set} + T_{2i} - T_{1i} - W_{ri} - W_s - W_d + T_k < F_i \quad (\text{ノード3} \sim N)$$

さらに、情報源から現在の再生点よりも  $L$  先のデータから各ノードに情報が送られるとすると、

$$L \geq \max(T_{12} + T_{set} + T_{21} - W_{r1} - W_s - W_d + T_k, T_{set} - W_{r2} - W_s - W_d + T_k, \dots, T_{12} + T_{set} + T_{2i} - T_{1i} - W_{ri} - W_s - W_d + T_k, \dots)$$

が満たされることが必要となる。

この場合さらに、 $L \leq F_1, L \leq F_2, L \leq F_i$  ( $i \geq 3$ ) が満たされることが必要となる。

また、情報源からのデータが表示開始より先に到着する場合、先に到着する分のバッファも確保する必要がある。つまり、以下に示す式が正の値ならば、その大きさのバッファを  $L$  以外に確保する必要がある。

$$W_{r1} + W_s + W_d - T_{12} - T_{set} - T_{21} \quad (\text{ノード1})$$

$$W_{r2} + W_s + W_d - T_{set} \quad (\text{ノード2})$$

$$W_{ri} + W_s + W_d - T_{12} - T_{set} - T_{2i} + T_{1i} \quad (\text{ノード3} \sim N)$$

#### ● 順コマ送り命令

順コマ送りの場合、一コマあたり  $T_k$  時間進むとすると、一コマだけなら

$$T_k < F_1$$

が満たされればよい。

次に、連続  $M$  回、 $T_b$  ( $T_b > T_k$ ) 時間ごとにコマ送りボタンを操作者が押す場合について考える。この場合も再生の場合と同様、 $F_1, F_2, F_3, \dots$  に以下に示す式を満たす  $L$  以上の情報が蓄えられていることが必要である。

$$L \geq \max(((T_{12} + T_{set} + T_{21})/T_b + 1)T_k, (T_{set}/T_b + 1)T_k, \dots, ((T_{12} + T_{set} + T_{2i} - T_{1i})/T_b + 1)T_k, \dots)$$

#### ● 逆コマ送り命令

逆コマ送り命令も順コマ送り命令の場合と同様に、必要なバッファ量を求めることができる。

### 4.2 再生状態

再生状態はそれぞれのノードで適当な再生点を取りながら再生が行なわれている状態と規定する。

#### ● 一時停止命令

それぞれのノードで一時停止が行なえる条件は

$$\text{無条件} \quad (\text{ノード1})$$

$$D_{21} + T_{12} + T_{dc} + T_k < P_2 \quad (W_{r2} + W_s + W_d < T_{set})$$

$$\text{無条件} \quad (W_{r2} + W_s + W_d \geq T_{set}) \quad (\text{ノード2})$$

$$D_{1i} + T_{1i} + T_{dc} + T_k < P_i$$

$$(W_{ri} + W_s + W_d < T_{12} + T_{set} + T_{2i} - T_{1i})$$

無条件

$$(W_{ri} + W_s + W_d \geq T_{12} + T_{set} + T_{2i} - T_{1i})$$

(ノード3 ~ N)

また、再生状態では各ノードへの情報源からのメディア情報の到着のずれを吸収するためのバッファが必要である。そのためには、以下の2つの条件が満たされる必要がある。

$$X - D_{i2} - T_{2i} < F_i \quad (i = 1, 2, 3, \dots)$$

$$T_d < F_i \quad (i = 1, 2, 3, \dots)$$

ここで、 $X$  は全てのノードで共通の値で、2つ目の条件が満たされるように設定する。

## 5 システム構成に関する考察

本章では R&L 同期を満たすために必要なバッファ量を求めるための計算式を使って幾つかの具体的な数値例を示し、算出結果について考察を加える。

### 5.1 算出条件

3章で述べた R&L 同期のセマンティクスを満たすために最低限必要なバッファ量を図3の構成のネットワークについて求める。

ここで、情報源としては動画情報の入ったハードディスク ( $T_{set} = [0ms, 30ms]$ ) を想定する。

また、ネットワークとしては衛星回線を想定し、その基本的な遅延は、図3(の分割前)を正五角形として短い方の辺の遅延 ( $T_{12}, T_{23}$  など) を  $[270ms, 300ms]$ 、大きい方の辺の遅延 ( $T_{13}, T_{14}$  など) を  $[432ms, 480ms]$  とした。

伸長・圧縮に伴う遅延は  $T_c = 0ms$  とし、 $T_{dc}$  に関しては伸長が必要な場合  $33ms$ 、必要でない場合  $0ms$  とした。

また、 $T_k = 33ms, T_i = 100ms, T_d = 33ms$  とした。

$sync\_with\_one$  の場合、再生命令実行時に各ノードで再生点調整に必要な待ち時間、及び、一時停止、コマ送り命令実行後 LMS との同期回復に必要な待ち時間は共に  $[0ms, 15ms, 24ms, 24ms, 15ms]$  となる。 $sync\_with\_all$  の場合は  $[480ms, 195ms, 24ms, 24ms, 195ms]$  となる。

### 5.2 算出結果

算出結果を以下に示す。ただし、ここで示すグラフは全て横軸が  $W_s$ 、縦軸が各ノードで必要なバッファの大きさとなっている。また、バッファの大きさは時間 (何ミリ秒分相当か) で表す。

● 例 1

図 4 は情報源がハードディスクで source1,  $T_{dc} = 33ms$  の場合に sync\_with\_one のセマンティクスを満たすために各ノードで最低限必要なバッファの大きさを示している。

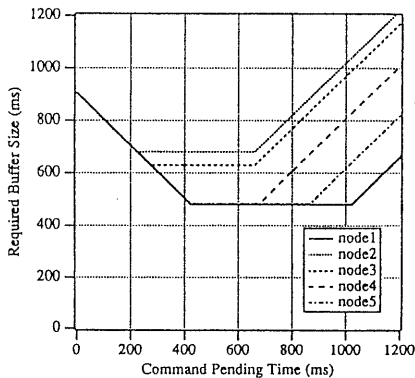


図 4: 必要なバッファ量 (例 1)

● 例 2

図 5 は例 1 と同じ条件で sync\_with\_all のセマンティクスを満たすために必要なバッファの大きさを示している。

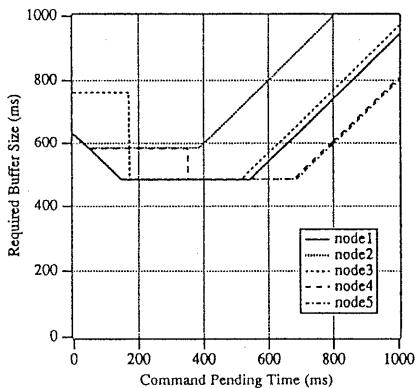


図 5: 必要なバッファ量 (例 2)

5.3 考察

前節で示したグラフから以下のことがわかる。

- $W_s$  をある程度増やすことにより必要なバッファの大きさを削減できる。しかし、 $W_s$  をある程度以上増やすと、情報提示より先に到着する分のバッファによりかえって必要なバッファ量が増加する。
- $W_r$  も含めて考えると、操作者ノードに近いノードほど必要なバッファ量が小さく、情報源ノードに近いノードほど必要なバッファ量は大きい。
- sync\_with\_all の場合、ノード 3,4 などでは再生状態からの一時停止に必要なバッファのために  $W_s$  が小さい領域である程度の大きさのバッファを必要とする。
- バッファをある程度大きく取ることによって、応答性を良くすることができる。

例えば例 1 の場合、sync\_with\_one のセマンティクスを満たすために最低限必要なバッファを用意した場合よりも、さらに大きなバッファを用意することで  $W_s$  を 500ms 程度削減することができる。つまり、大きめのバッファを用意することで 15 フレーム分程度、応答性を良くすることができる。

6 むすび

本稿では、マルチメディア遠隔提示システムを構築する場合に、各ノードでの情報提示の際のメディア間の同期を維持するための条件について検討した。論文中では、まず、蓄積型情報源から取り出されるメディア情報と参加者の声などの生のメディア情報との間の情報提示の際のタイミングに基づいた同期のセマンティクスを定めた。次に、このセマンティクスを満たすために各ノードにおいて必要なバッファ量の算出法を示した。さらに、本手法を用いてバッファ量の算出を行なった結果、一時停止などの命令に対する応答時間の設定値をある程度まで大きくすることで必要なバッファ量を削減できることを確認した。

今後の課題としては、本稿で述べたようなマルチメディア提示システムを様々な環境で構築することが挙げられる。

参考文献

- [1] Little T.D.C.: "Multimedia Synchronization Protocols for Broadband Integrated Services", IEEE J. Sel. Areas Commun., 9, 9, pp.1368-1382(1991).
- [2] Steinmetz R.: "Synchronization Properties in Multimedia Systems", IEEE J. Sel. Areas Commun., 8, 3, pp.401-412(1990).
- [3] 大野, 相田, 齊藤: "会話型操作を伴う分散マルチメディアシステムにおける同期機構", 情報処理学会マルチメディアと分散処理研究会, 66-21, 1994.