

J-001

リアルタイム IP 通信に適した擬似同期型マルチメディア再生方式 Pseudo Synchronus Media Presentation for Real-time IP Communication

馬場 昌之[†] 黒川 弘幸[†]
Masayuki Baba Hiroyuki Kurokawa

横山 幸雄[†] 小川 文伸[†]
Yukio Yokoyama Fuminobu Ogawa

1. はじめに

近年、インターネットの普及とブロードバンド IP 網の広域配備等により、テレビ電話やストリーミングなどのリアルタイムマルチメディア IP 通信が盛んになってきた。リアルタイム通信では送受間でのメディアのクロック同期が必要となるが、従来の方式では、揺らぎの大きい IP 網や、再生用クロックの周波数調整機能を持たないパソコンや PDA などの情報機器への適応が困難である。

本稿では厳密なクロック同期を必要としない擬似同期型マルチメディア再生方式を提案する。

2. 従来のメディア再生方式

2.1 メディアクロック同期の必要性

リアルタイム通信においては、図 1 のように送信側のメディアサンプリングクロックと受信側の再生用クロックが同期する必要がある。もし送受間でこのクロックが同期していないと、受信側で再生時にサンプルの過不足が発生し、音声の場合には音飛びや遅延が起きる。

2.2 従来方法

回線交換系の通信装置はネットワークから伝送用クロックを供給されているため、そのクロックを分周しメディア用クロックとすることで、送受装置間での同期を取っていた。一方、パケット交換系の通信装置では、ネットワークに頼らず通信装置間で同期を取るメカニズムを具備する。例えば、MPEG-2 システム規格 [1] では、送信側がクロック情報をメディアデータと一緒に送信し、受信側でそのクロック情報からメディア用クロックを生成して送受間でのクロック同期を達成している。

2.3 課題

揺らぎの大きい IP 網では、従来の厳密なクロック同期方式は適応できない。また、IP 通信機能を備えるパ

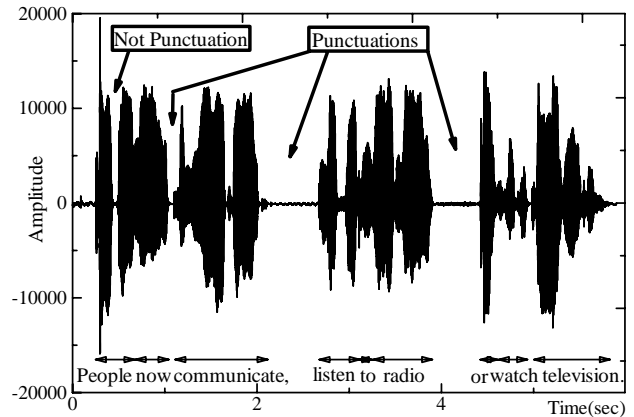


図 2: 音声波形例

ソコンや PDA などの情報機器は、一般にメディア再生用クロックの周波数を調整する機能を持たない。従って、送受間での厳密なメディア再生用クロックの同期・調整を行なうことなく、メディア再生を行なうための手段が必要となる。

3. 擬似同期型メディア再生方式の提案

本章では、厳密なクロック同期を取らない擬似同期型マルチメディア再生方式を提案する。

3.1 メディアストリームの特性

ビデオや音声などのメディアストリームは、時間的に連続するものであるが、その内容的には必ずしも連続とは限らない。例えばビデオのシーンチェンジ、音声の句読点が無音部分となる。図 2 は “People now communicate, listen to radio or watch television.” と言っている音声波形例である。文中に存在する無音部分のいくつかは句読点にあたり不連続点とみなせる。しかし単語中に存在する無音部分は、不連続点とみなすべきではない。

不連続点で囲まれたストリーム部分をストリームブロック (Stream Block: SB) と称し、この SB をメディアの内容的な連続性を保証すべき最小の単位とする。再生時に連続する 2 つの SB 間の時間的間隔が本来の間隔から多少変わっても、内容的には完全に変わっているため、メディアの再生品質はほとんど影響を受けない。

3.2 ストリームブロック単位のメディア再生

クロック同期を行わない通信環境では、メディアの再生は自走クロックに基づいて行われる。提案方式では、メディアストリームの特徴を考慮して各 SB を保護するように、受信側では各 SB 単位で再生を行う。送受間のクロック差を吸収するために、SB 間の再生間隔を各不連続点において調整する。こうすることでメディアの再生品を保ちながら擬似的に同期が可能となる。

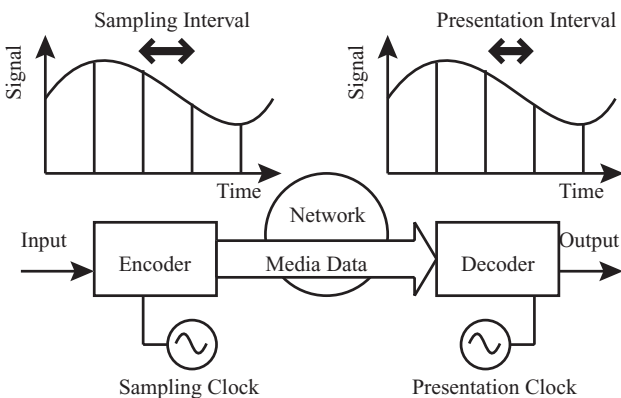


図 1: メディアクロックの同期

[†]三菱電機 (株), Mitsubishi Electric Corporation

4. システム動作

4.1 ストリームブロックの生成

SBの生成には、音声認識技術やビデオのシーンチェンジ検出技術を利用することにより、句読点やシーンチェンジを含む分割ポイントの検出や認識が可能となる。メディアストリームは、Access Unit(AU) と呼ばれる単位で符号化・伝送されるため、SBは図3に示すように1つ以上のAUから構成される。

4.2 ストリームブロックの伝送

提案方式におけるメディアストリーム伝送においては、受信側でのSB識別のための情報を通信プロトコルに埋込む必要がある。具体的な方法としては、1SBを1パケットとして伝送する方法や、SBを複数のパケットに分割して転送する際にブロックの先頭を含むパケットのヘッダに不連続を示す情報を埋め込む方法、等がある。実装にはRFC3550[2]に規定されるRTP(Real-time Transport Protocol)を利用し、不連続情報としてはMビットを用いることができる。

4.3 ストリームブロックの擬似同期再生

本節では、送受信装置が自走クロックで動作する場合の、提案方式による擬似同期再生動作を示す。

図4は、受信側のメディアクロック周波数が送信側よりも高い場合のシステム動作である。受信側では再生すべきデータが不足するため、一時的にメディア再生を停止する必要が生じる。提案方式では、メディア再生停止をSBの境界で発生させるようにする。そのために、受信データのある程度バッファに蓄積しながら再生を行い、バッファ残量が少なくなった時点で一時的に再生を停止する。これにより、再生品質の劣化を最小限に抑制することができる。

図5は受信側のメディアクロック周波数が送信側より低い場合のシステム動作である。この場合、受信側での再生タイミングが徐々に遅れていく。提案方式では、再生タイミング遅延を取戻すために、SBの最後のAUを廃棄する。そのため送信側では最後のAUが廃棄される可能性があることを考慮してSBを生成する。

5. 本提案の有効性の検証

本提案の有効性を検証するため音声再生に関するシミュレーションを行った。20ms単位の音声データを1AUとして、データ過不足による音飛びの影響を主観評価した。

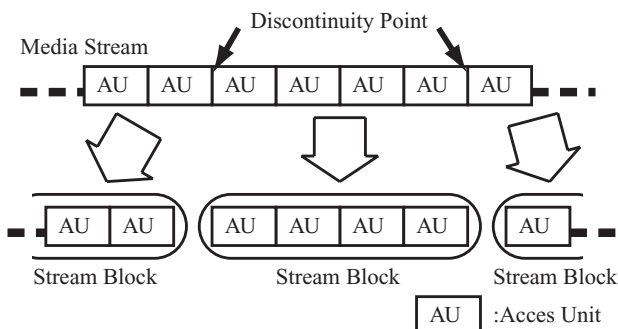


図3: ストリームブロック

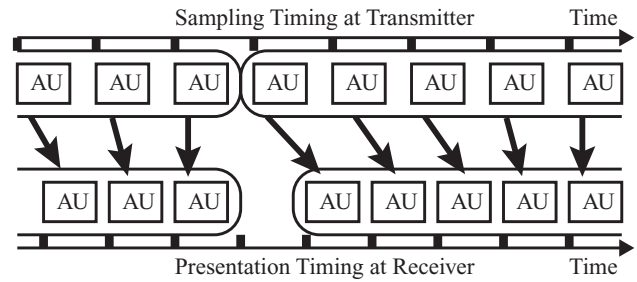


図4: 再生タイミング(1)

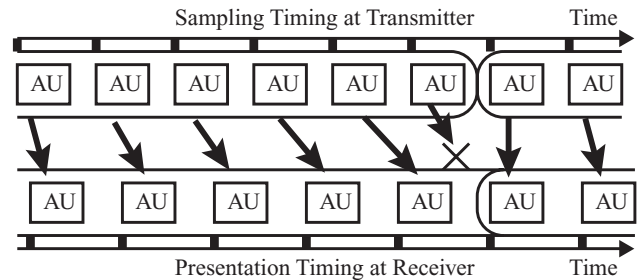


図5: 再生タイミング(2)

評価は以下の2点について行い、さらに音声に雑音を加えた場合の評価も行った。

(A) 不連続点でのAUの廃棄および無音挿入

(B) 連続点でのAUの廃棄および無音挿入

その結果、(A)での音声の劣化は認識できなかったが、(B)では音飛びを認識できる程度に音声の劣化を感じた。また、雑音を加えた音声に関しても同様の結果が得られた。これにより、本提案による擬似同期型メディア再生方式の有効性が確認できた。

6. おわりに

メディアストリームの内容的な特性を利用することにより、送受信間で厳密なクロック同期を取ることなくメディア再生を行うことが可能な擬似同期型マルチメディア再生方式を提案した。本方式は、揺らぎにより厳密なクロック同期が困難なIP網や、再生用クロックの周波数調整機能を持たないパソコンやPDAなどの情報機器への適用が可能である他、無線LANやBluetooth等の伝送方式への対応も可能である。

参考文献

- [1] ISO/IEC 13818-1, "Generic coding of moving pictures and associated audio: systems," ISO/IEC, 2000
- [2] IETF RFC3550, "RTP: A Transport Protocol for Real-Time Applications," IETF, July 2003
- [3] A. K. Anandakumar, A. McCree, and E. Paksoy, "An Adaptive Voice Playout Method for VOP Applications," GLOBECOM '01, November 2000