

モバイルビデオのための多重化フォーマット設計に関する一検討

J-58 A study on configuration of multiplexing format for mobile video transmission

宮地 悟史
Satoshi Miyaji滝嶋 康弘
Yasuhiro Takishima和田 正裕
Masahiro Wada羽鳥 好律
Yoshinori Hatori

1. はじめに

第三代携帯電話が実用化され、モバイルによるマルチメディア通信が身近なものとなった。マルチメディアサービスの中でも、特に動画配信への期待は大きく、システムの設計においては、携帯電話上での実現性に加え、今後の発展性についても十分考慮することが重要である。

筆者らは、これまでに携帯電話向け動画配信フォーマットとして、ISO/IECのMP4をベースとしたコンテンツフォーマットを提案している[1]。ISOで規定されているMP4は、そのファイル内に含まれるメタデータ(コンテンツの情報を格納するための構造体)のシンタクス、階層関係および存在個数を規定しているのみであり、実際の動画ビットストリームの格納方法は実装依存となっており、従来定量的な検討が行われていない。

本稿では、MP4ファイルに関して、携帯電話での動画再生処理に最適なパラメータ設計について検討し、また、今後の展開として携帯電話動画ストリーミングも視野に入れた検討を行い、報告する。

2. モバイルビデオの要求条件

本検討では、モバイルビデオに要求される条件を、以下のよう

- ① 伝送路誤り(パケットロス等)に対して、耐性があること
- ② 伝送効率が十分高いこと(冗長度が低いこと)
- ③ 端末の実装インパクトが大きすぎないこと
- ④ 伝送路状態の変動によるスループット変動に対して、配信途中においても対応可能であること
- ⑤ 長時間コンテンツの配信が可能であること
- ⑥ ライブ配信が可能なこと
- ⑦ 蓄積コンテンツに対して、ランダムアクセス(ジャンプ、早送り、巻き戻し等)が可能なこと
- ⑧ 端末におけるコンテンツ要求からコンテンツ表示開始までの待ち時間が十分短いこと
- ⑨ ストリーミング拡張した際、ダウンロード専用端末でも一部受信可能なこと(後方互換性)
- ⑩ (広く普及させるため)国際標準に準拠し、オープンな仕様であること

3. メディア多重フォーマット

現在、マルチメディアの配信に用いられている、メディア多重フォーマットについて比較検討する。

3.1 MP4 フォーマット

MPEG-4 システム規格で規定されているメディア格納ファイルフォーマットである。従来、ファイルフォーマットは産業界で独自に制定されるのが一般的であったのに対し、MP4はISOで策定された国際標準フォーマットとなっている。

MPEG-1/2/4ビデオ・オーディオの圧縮形式に対応し、単一

ファイルのみならず複数分割ファイルやストリーミングにも対応する。メディアに関する情報を記述するヘッダと、実際のメディアデータとから構成され、各メディアはトラックとして管理される。また、情報ヘッダにはBoxと呼ばれる階層型の構造体が用いられており、トラックが多数ある場合でも、復号側において対象トラックの必要情報を容易に取得することができる。

3.2 RTP フォーマット

IETFで規定されているストリーミングのためのパケットフォーマットで、UDP/IPによる伝送を前提としている。

対象とする圧縮形式はMPEG-1/2/4ビデオ/オーディオ、H.261/263, G.711/723, JPEG, MPEG-2 TSパケット, IEEE 1394ビデオなど多岐にわたる。しかし、各圧縮形式に対して個別の格納フォーマットが規定されているため、RTPの中にも多様な形式が存在する。また、UDPでの伝送を前提としているため、RTPパケットのヘッダには、タイムスタンプや識別IDなどの情報が、ペイロード部分には圧縮方式依存のエラー回復用冗長情報が付加されている。

3.3 PC向けフォーマット

QuickTimeは、Apple社が提唱しているファイルフォーマットで、前述のMP4策定のベースになったものである。技術仕様のには、MP4の基本仕様(バージョン1)とほぼ同様であり、バージョン2で拡張された分割格納などは含まれていない。

AVIおよびASFは、Microsoft Windowsで利用されているファイルフォーマットで、ファイル保存、ストリーミングの両方に対応している。さまざまな圧縮形式データが格納できるが、仕様は公開されていない。

4. モバイルビデオの実現

4.1 伝送プロトコルと多重フォーマット

ビデオストリーミングを携帯電話に対して行う場合、安定して受信できるという点が最も重要となる。有線ネットワークと比較して、外乱の影響を格段に受けやすいモバイルネットワークにUDP/IPを適用した場合、

- エラー耐性を有するためのデータ冗長度(MPEG-4+Error Resilienceの場合15%程度[2])や、RTPフォーマットによる冗長ビットが必要となり、伝送効率(要求条件②)が低下する
- 上記エラー対策でも復元できないようなデータロス発生がある[2]

これらいずれも受信品質(要求条件①)に大きく影響を与える可能性があること、さらにはテレビ電話のような低遅延性は要求されないことから、本検討では伝送プロトコルにはTCP/IPを用いることとする。

TCP/IPの場合、HTTPのようなファイル転送プロトコルを用いることで複雑なサーバクライアント間制御が不要となり、ダウンロード動作とストリーミング動作との間の後方互換性(要求条件⑨)が実現でき、またRTPのような冗長情報も不要となり、伝送

効率および実装インパクトへの条件(要求条件②および③)を満足することができる。さらに国際標準性(要求条件⑩)を考慮すると、モバイルビデオにおいては、MP4 ファイルフォーマットを用いるのが最適であるといえる。

4.2 MP4 ファイルの最適構成

MP4 ファイルフォーマットでは、全ての情報は「Box」と呼ばれる構造体に格納されており、多重化された音声/動画ビットストリーム(メディアデータ)ならびに、これらメディアデータに対する管理情報(メタデータ)から構成されている。

コンテンツメタデータは「moov」Box に格納されており、Box 内の各構造体は階層関係を有している。図 1 に一般的な MP4 ファイルフォーマットの構成を示す。

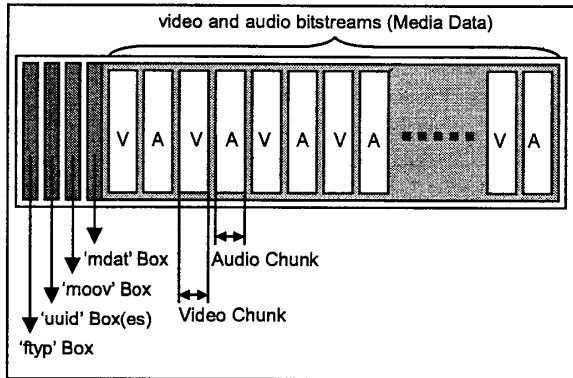


図 1: MP4 ファイルフォーマットの構成

MPEG-4 システム規格によれば、MP4 ファイルの構成法については、

- ファイル識別 Box ('ftyp') は先頭になければならない
- 上記以外の Box 配置順序は自由
- メディアデータの多重形態については自由
- ファイル分割(フラグメント)の構成については自由

など、実際のファイル構成に際しての自由度が大きく残されている。

ビデオストリーミングを考えた場合、まず、再生までの待ち時間を短くする必要性(要求条件⑧)から、コンテンツメタデータ('moov' Box)はファイル先頭で、かつ大きすぎないことが必要となる。従って、長時間コンテンツ(要求条件⑤)の場合には、ファイルを分割(フラグメント)しヘッダを分散配置する必要がある。さらに、MP4 では、分割した場合の最初のフラグメントは、分割しない場合の単一ファイルと同形式となるため、後方互換性(要求条件⑨)も同時に満足できる。

また、ビットレートの途中切り替え、ライブ配信における途中からの受信参加、またはランダムアクセス(要求条件④⑥⑦)を柔軟に行うためには、

- ランダムアクセスやレート切り替えの単位を細かくすることができる
- ランダムアクセスやライブ配信受信開始時、開始するフラグメント位置と、指定位置との誤差を小さくできる

といった点からフラグメント単位は出来るだけ小さいほうが望ましいが、一方でこれはヘッダによるオーバーヘッドとなり要求条件②と相反する。図 2 にフラグメント間隔とオーバーヘッドとの関係を示す。

4 種類のプロットは、それぞれビデオビットレート、オーディオ符号化方式の違いを示す。本図から、フラグメント間隔が 10 秒以上でヘッダの冗長度が飽和するのがわかる。

さらに、ランダムアクセス等の柔軟性(要求条件④⑥⑦)に関してはビデオフレームの Intra 符号化フレーム(I-VOP)挿入間隔も考慮しなければならない。最低限フラグメント先頭に 1 つ必要となる。I-VOP 間隔が小さくなるとフラグメント内のランダムアクセス可能ポイントが増えるというメリットがある反面、特にモバイルのような低ビットレートでは、符号化効率の低下を招く。従ってビットレートとフラグメント間隔とから最適値を決定する必要がある。

次に、メディアデータの多重間隔について検討する。多重間隔は、ストリーミング時の再生待ち時間(要求条件⑧)、および端末の実装インパクト(要求条件③)を決定付ける。多重間隔を小さくすると、

- 短い時間で音声・動画像の両方が揃って受信できるため、ストリーミング再生の待ち時間を小さくできる
- 同時に読み込む音声・動画像データ量が小さくなるため、端末のメモリを圧迫しない
- MP4 では、多重の単位ごとにデータ位置を示すポインタを持っており、多重間隔が小さいとデータアクセスが頻繁になり、端末の処理負荷が増大する

といった一長一短がある。このため、端末の処理を圧迫しない範囲で、なるべく小さな多重間隔が最適となる。

以上の考察の結果から、

- 'moov' Box はファイルの先頭に配置する
- フラグメント間隔は 10 秒程度
- Intra 符号化フレーム間隔は 5 秒程度(1 フラグメントに 2 つ)
- メディアデータの多重間隔は 1 秒程度

が最適な配信フォーマットの形態となる。

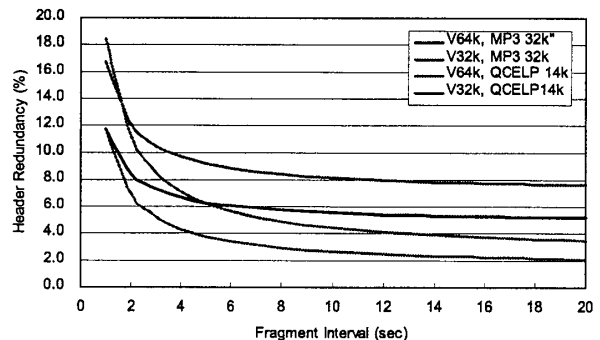


図 2: フラグメント間隔とヘッダ冗長度との関係

5. まとめ

モバイルビデオ配信のための、最適ファイルフォーマットについて検討を行った。その結果、要求条件を満足するため、MP4 を使用し、またモバイルビデオに最適ないくつかのパラメータを示した。

なお、本研究は、通信・放送機構より受託した「IP ネットワーク上の音声・動画像配信(Webcasting)における QoS 制御に関する研究開発」に基づき行われた。

参考文献

- [1] 宮地 他: “携帯電話ビデオ配信用コンテンツ制作システムの開発—コンテンツフォーマット—”, ITE 冬季大会, 1-5, 2001
- [2] 宮地 他: “ワイヤレス IP 環境下におけるビデオストリーミングに関する一検討”, ITE 年次大会, 6-2, pp.78-79, 2001