

■2群 (画像・音・言語) -5編 (画像符号化)

12章 システム標準

(執筆者：内藤 整) [2010年5月 受領]

■概要■

ビデオ、オーディオ、データにより構成される番組の伝送を行う場合、受信側においては共通のタイミング制御のもとで各種メディア間の同期が保証されなければならない。MPEGにおけるシステム標準は、主にMPEGの圧縮標準に準拠して符号化されたビデオ、オーディオの圧縮符号化ストリームを主な対象とし、データコンテンツの符号化データを含めた各種メディアの符号化ストリームの多重化ならびに受信側での同期制御を実現する。

各メディアデータの多重化方法は、ストリーム型とファイル型がある。ストリーム型はコンテンツの再生時刻順にデータが羅列される構成をとり、データを受信しながら、メディア再生を行うクライアント形態に適している。一方、ファイル型は再生を開始するタイミングにおいて、全データが揃っていることを前提としているため、ストリーム型のようなデータの配置への制約は少なく、むしろ任意のコンテンツ時刻まで頭出しするといったランダムアクセス性を重視した多重化機能が求められる。同期制御には、メディア同期とシステム同期の二つの意味がある。メディア同期は、話者のリップシンクのように、コンテンツを構成するメディア間のタイミング合わせを意味するのに対し、システム同期は送信側と受信側において、メディアのタイミング制御に矛盾が生じないように、基準時刻を共通化することを意味する。

【本章の構成】

本章では、まずシステム多重化(12-1節)の代表例としてMPEG-2システムのトランスポートストリームとプログラムストリームについて述べる。次にファイルフォーマット(12-2節)について、ISOメディアファイルフォーマットを中心に導入を行う。最後に、メディア同期及びシステム同期(12-3節)に関して、MPEG-2システムにおける同期方式を例にとり仕組みを解説する。

■2群 - 5編 - 12章

12-1 システム多重化

(執筆者：内藤 整) [2010年5月 受領]

システム多重化の代表例として MPEG-2 システム (MPEG-2 Systems) ¹⁾ で規定しているトランスポートストリーム (Transport Stream) 及びプログラムストリーム (Program Stream) について説明する. MPEG-2 システムにおいて, ビデオ, オーディオ, データといった各メディアを構成する符号化ストリームは, エレメンタリーストリーム (Elementary Stream) と呼ばれ, エレメンタリーストリームをアクセスユニット (Access Unit, 同一再生時刻を有するデータ単位) に分割し, 時刻情報を付与したものが PES (Packetized Elementary Stream) (詳細は 12-3 節) である.

トランスポートストリーム (以下, TS と呼ぶ) は, ビットエラーやパケットロスといった誤りの発生する可能性が比較的高い状況下での利用を想定しており, PES あるいは制御情報をパケット長が 188 バイト固定の TS パケットにより扱い多重化を行う. TS パケット (TS Packet) には, 16 バイト長のリードソロモン符号を付与して扱うモードもあり, この場合のパケット長は 204 バイトとなる. TS では複数の番組を同一ストリームで扱うことが可能で, 番組ごとに再生時刻の記述となるクロックを定義できる. TS に基づく MPEG-2 システムにおいて, 番組を構成するビデオ, オーディオの符号化ストリームと TS パケット PID (Packet ID) の対応表など, 各種番組制御情報は PSI (Program Specific Information) として伝送される.

TS パケットの符号規則 (以下, シンタックスと呼ぶ) は図 12・1 に従い, ヘッダ部分とペイロード部分から成る. ヘッダ部分は固定的に 32 ビットが割り当てられるフィールドと, 必要に応じて割り当てられるアダプテーションフィールド (Adaptation field) に分けられる. ヘッダ内のアダプテーションフィールドの有無, 及びパケット内のペイロード部分の有無は, ヘッダ内のアダプテーションフィールド制御 (図中 AF 制御, 2 ビット) により指示される. アダプテーションフィールドにより, クロック情報の伝送やスタッフィングが可能となる.

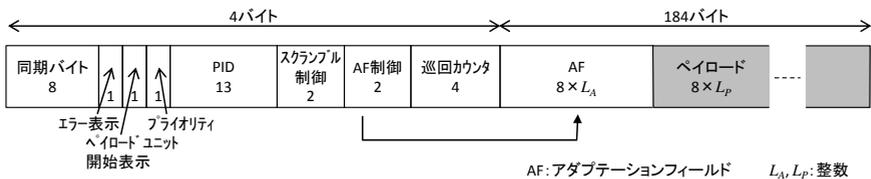


図 12・1 TS パケット

PID はパケットの識別を行うための番号であり, 13 ビットで表される. PES パケット (PES packet) をトランスポートストリームで伝送する場合, 一つの PES パケットは, 同一の PID をもつ複数の TS パケットのペイロードに分けて格納される. ここで PES パケットの先頭は, TS パケットのペイロードの先頭から開始するように格納され, TS パケットが PES パケットの先頭を含むか否かをペイロードユニット開始表示フラグにより示す. また, 同一 PID をもつ TS パケットに対しては, パケットの連続性を通知する目的で巡回カウンタ (4 ビット) に先頭のパケットから順に 1 ずつインクリメントした値を設定することで, 受信側でのパケッ

トロスの検知を可能としている。更に、TS パケットのペイロード部分に対してはスクランブルリングの適用が可能であり、スクランブル制御（2 ビット）により適用の有無が示される。

プログラムストリーム（以下、単に PS と呼ぶ）は、パッケージメディアのように誤りの発生しない状況下での伝送を想定しており、複数の PES パケットをまとめて形成されるパックという単位で多重化を行う。PS で扱える番組は一つに限定され、当然基準となるクロックも一つしか定義されていない。PS パック（PS pack）のシンタックスは図 12・2 に従い、パックヘッダ、システムヘッダ、及び複数の PES パケットが格納されるペイロード部分から成る。パックヘッダはパックの先頭に常に割り当てられ、SCR（System Clock Reference）というクロック情報やビットレートなどが記述される。システムヘッダの付与は先頭のパックにのみ義務付けられ、ビットレート、オーディオチャネル数、ビデオチャネル数、フラグ類、全個別ストリームに関する必要バッファメモリ量などが記述される。



図 12・2 PS パック

■2群 - 5編 - 12章

12-2 ファイルフォーマット

(執筆者：内藤 整) [2010年5月 受領]

ダウンロード型でのマルチメディアサービスにおいて、サーバ側ではファイル形式でコンテンツを扱うこととなる。このファイル形式はISO/IECにて、ISOメディアファイルフォーマット (ISO Media File Format)²⁾ (以下、ISOファイルフォーマット) として標準化されている。ISOファイルフォーマットは、ボックス形式のデータを木構造に配置した構成をとり、代表的なボックスにftyp, mdat, moovがある。ftypはファイルの先頭のみ存在し、ファイル種別を示す。mdatには、例えばビデオであれば、MPEG-4によるビデオエレメンタリーストリームなど、メディアデータが格納される。moovはファイルのメタデータを格納しており、複数のボックスにより構成される。メタデータとしては、mdatに含まれるメディアデータ(トラックと呼ばれる)ごとに、再生に必要なタイミング情報、ランダムアクセスのためのファイル内位置情報、コンテンツ属性などを記述する。ISO/IECでは、ISOファイルフォーマットに対し、エレメンタリーストリームとしてMPEG-4による動画・音声を扱う際の拡張定義を含む形で、MP4ファイルフォーマット (MP4 File Format)³⁾ が標準化されている。MP4ファイルフォーマットでサポートするビデオデータはMPEG-4ビデオ符号化形式に限定されるわけではなく、MPEG-1, MPEG-2といった他の標準符号化方式にも対応している。参考までにISOファイルフォーマットに対し、ビデオエレメンタリーストリームとしてMPEG-4 AVCを扱うための拡張が、AVCファイルフォーマット (AVC File Format)⁴⁾ として標準化されている。MP4ファイルフォーマットは携帯型のゲーム機やメディアプレーヤで幅広く採用されている。更に携帯電話サービスとして、主に動画コンテンツを扱うためのファイルフォーマットとして、3GPPファイルフォーマット (3GPP File Format) 及び3GPP2ファイルフォーマット (3GPP2 File Format) があり、いずれもMP4ファイルフォーマットをベースに個別の拡張定義を含んだものである。

■2群 - 5編 - 12章

12-3 メディア同期, システム同期

(執筆著: 内藤 整) [2010年5月 受領]

メディア同期はアクセスユニットごとのタイムスタンプ制御により実現される。MPEG-2 システムの PES を例にとり、タイムスタンプ制御を説明する。PES パケットは図 12・3 のシンタックスに従う可変長のパケットである。パケット長では PES パケットの総バイト数を示し、このうちペイロード部分を除いたヘッダ部分の総バイト数をヘッダ長で示す。16 ビットのフラグ部分には PES パケットの優先度を示す PES プライオリティ、再生・表示時刻の指定があるか否かを示す PTS・DTS フラグ、CRC による誤り検出符号を使用するか否かを示す CRC フラグなどが割り当てられている。ここで、PTS (Presentation Time Stamp) と DTS (Decoding Time Stamp) はそれぞれ、受信側での再生時刻、復号時刻を意味し、受信側では PTS と一致する時刻に復号されたユニットの再生を、DTS に一致する時刻に符号化ストリームの復号を実行する。拡張データには、上述のフラグ部分の指定に応じて必要となるデータ

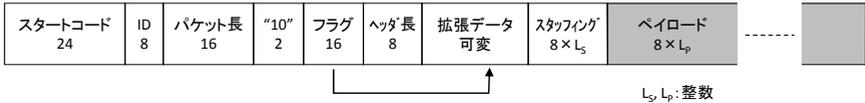


図 12・3 PES パケット

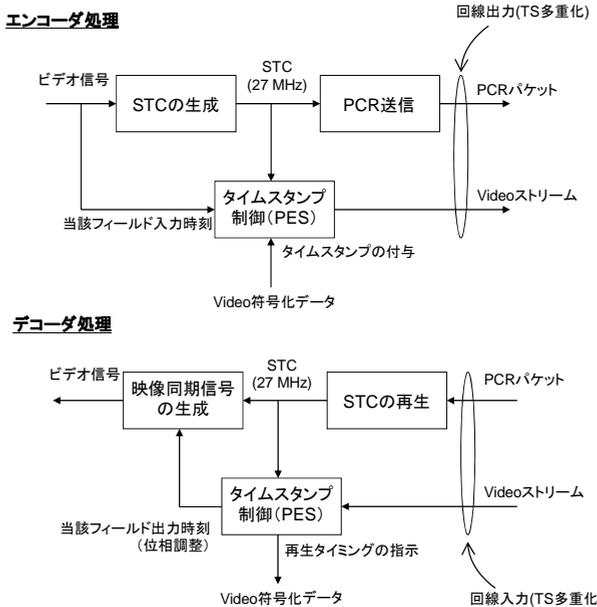


図 12・4 システム同期メカニズム

が格納される。例えば、PTS・DTS フラグが“11”に設定された場合、PTS と DTS をともに記述することを意味し、それぞれ 33 bit 表現されたタイムスタンプが PTS、DTS の順番で格納される。フレームの再生時刻と復号時刻が一致している場合には、DTS の記述は省略される。

次にシステム同期の仕組みを、MPEG-2 システムの TS を例に取り説明する。PTS 及び DTS は、送信側と受信側で共通のクロックを基準として管理される必要がある。TS ではこのクロックを 27 MHz の STC (System Time Clock) によって管理する。ビデオ伝送の場合を例にとり、TS のシステム同期メカニズムを図 12・4 に示す。送信側では、十分に高い頻度で STC を参照し、読み取ったカウンタを PCR (Program Clock Reference) として、当該時刻に対応する TS パケットのアダプテーションフィールドに重畳する。受信側では、TS パケットの受信時刻と PCR のカウンタ値を対応付けることによって STC を再生する。これにより、送信側と受信側でシステムが基準とするクロックを共通化でき、システム同期が確立される。

■参考文献

- 1) ISO/IEC 13818-1:2007, “Information technology - Generic coding of moving pictures and associated audio information: Systems,” 2007.
- 2) ISO/IEC 14496-12:2008, “Information technology - Coding of audio-visual objects - Part 12: ISO base media file format,” 2008.
- 3) ISO/IEC 14496-14:2003, “Information technology - Coding of audio-visual objects - Part 14: MP4 file format,” 2003.
- 4) ISO/IEC 14496-15:2004, “Information technology - Coding of audio-visual objects - Part 15: Advanced Video Coding (AVC) file format,” 2004.