

## ■2群(画像・音・言語) - 11編(マルチメディア)

### 3章 マルチメディアシステム

(執筆者: 柴藤 稔) [2011年2月 受領]

#### ■概要

マルチメディアの本質である、音声、楽音、画像、映像、テキストを視聴者に伝達する系に関する技術を本章で解説する。

3-1節と3-2節は密接に関連している。マルチメディアデータがどのようにファイルや伝送パケットに格納されるかを解説したのち、どのように伝送されるかがプロトコルを中心に説明される。特に現在商用化されているファイル形式、伝送形式が説明され、用いられている技術が網羅されている。

3-3節はマルチメディアを記述するメタデータ標準 MPEG-7 の解説である。マルチメディアデータは非構造化データである。データがアノテーションにより厳密には定義されていない。MPEG-7はそのようなマルチメディアデータを記述するための構造を与える。XMLがどのように使われているかという点に注目してほしい。

3-4節のコンテンツ流通フレームワークでは、著作権管理を含めてどのように流通フレームワークが構成されているかという点を説明する。前節を含めて MPEG-7, MPEG-21 という二つの世界標準を例にとっている理由は、マルチメディアの構造記述と流通フレームワークを特定の商用システムに偏ることなく説明するためである。記述されている仕様ではなく機能に注目してほしい。議論されている機能を知れば、どのような技術が必要であるかが理解できる。また、この節では携帯電話と DVD における著作権管理を簡単に紹介しており、次節への導入になっている。

3-5節において、著作権管理技術についてその歴史から、暗号化・電子透かしの技術について説明する。

## ■2群 - 11編 - 3章

### 3-1 ファイルシステム/ファイル構造

(執筆者：遠間正真) [2011年2月 受領]

映像や音声などの符号化データは、フレームのサイズや復号・表示時刻などのヘッダ情報を付加し、システム多重化してから、伝送、あるいは蓄積されます。システム多重化の方式は、放送・通信・蓄積といったアプリケーションに応じて、MPEG (Moving Picture Experts Group) や IETF (Internet Engineering Task Force) などの国際標準、あるいは、サービスに固有の独自方式として規定されている。

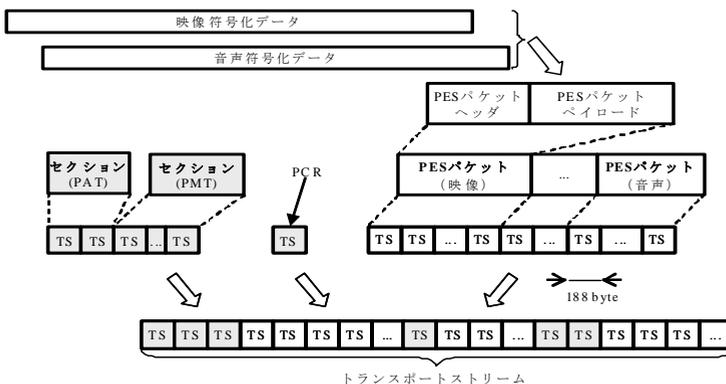
本節では、幅広いアプリケーションに適用されるシステム多重化方式の国際標準である、以下の3方式について説明する。

- (1) TS (Transport Stream)
- (2) RTP (Real-time Transport Protocol)
- (3) MP4 ファイルフォーマット

#### 3-1-1 TS (Transport Stream)

TS<sup>1)</sup>とは、主に放送向けの伝送フォーマットであり、地上デジタル放送、BS/CS デジタル放送、更には、BD (Blu-ray Disc) などで使用されている。

TSは、TS パケットと呼ばれる 188 バイトの固定長パケットのシーケンスであり、プログラム(番組)を構成する映像や音声を示すプログラム情報、映像や音声の符号化データ、及び、メディア間の同期を取るための基準クロック情報などを、それぞれ異なる ID 番号をもつ TS パケットに格納して伝送する。ここで、符号化データは、PES (Packetized Elementary Stream) パケットに、プログラム情報はセクション (Section) と呼ばれる格納単位にそれぞれ多重化した後に、TS パケットに格納する (図 3・1)。映像・音声の復号時刻や表示時刻などフレーム単位の情報は、PES パケットのヘッダにおいて示される。



PAT (Program Association Table) : プログラムの識別情報  
 PMT (Program Map Table) : プログラムを構成する映像・音声の識別情報  
 PCR (Program Clock Reference) : 基準クロック情報

図 3・1 トランスポートストリームの構造

### 3-1-2 RTP (Real-time Transport Protocol)

RTP<sup>2)</sup>とは、映像や音声などのリアルタイムデータをパケット化してIPネットワーク上で伝送するためのプロトコルであり、IETFにおいて規格化されている。RTPパケットはリアルタイム性を重視するため、下位層のプロトコルとしては、通常、フロー制御や再送制御を行わずにデータを転送するUDP (User Datagram Protocol) を用いている。RTPは、IPTV、第3世代携帯端末、PC (Personal Computer) 向けのストリーミングに広く採用されている。

RTPによるストリーミングでは、端末は、以下のような手順でサーバに蓄積されたコンテンツを受信する。

1. 映像や音声の符号化データを取得するためのサーバのアドレス情報を取得
2. サーバとのセッションを確立
3. 映像や音声の符号化データを格納した RTP パケットを受信

サーバ - 端末間の通信制御を行うためのプロトコル例としては、RTSP (Real Time Transport Protocol<sup>3)</sup>)があり、再生開始、終了、一時停止などの再生制御はRTSPのコマンドとして実行できる。RTSPデータの送受信にはRTPと別のチャンネルを使用し、下位層には信頼性の高いTCP (Transmission Control Protocol) が用いられている (図3・2)。

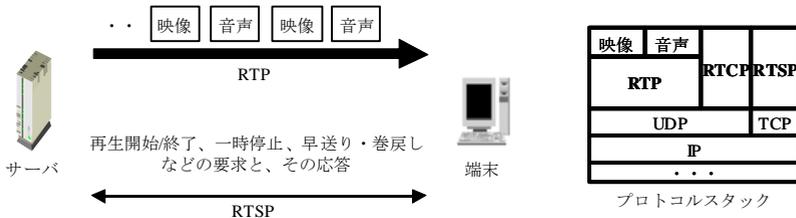


図3・2 RTPストリーミングの概要

図3・3にRTPパケットのデータ構造を示す。ヘッダ部は、ペイロードに格納される映像・音声符号化データの識別情報や再生時刻、RTPパケットのカウンタなどを示す。ペイロード部は、ペイロードヘッダとペイロードデータとから構成され、ペイロードヘッダには符号化方式に固有のヘッダ情報が、ペイロードデータには符号化データを格納する。ペイロード部におけるデータの格納方式はペイロードフォーマットと呼ばれ、H.264やAAC (Advanced Audio Coding) など符号化方式ごとのフォーマットが別途規定されている。

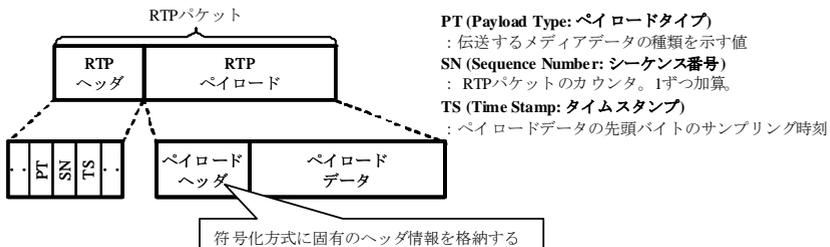


図3・3 RTPパケットのデータ構造

### 3-1-3 MP4 ファイルフォーマット

MP4 とは、映像や音声のデータを一つのファイルに多重化してダウンロード、あるいは、プログレッシブダウンロードするために MPEG で規定したファイルフォーマットの総称であり、符号化方式に依存しない基本構造を定める ISO Base Media File Format<sup>4)</sup> と、H.264 や AAC など符号化方式ごとの拡張を定める複数の規格から構成されている (図 3・4)。また、AMR (Adaptive Multi-Rate) など MPEG 規格外の符号化方式についても、3GPP (3rd Generation Partnership Project) などの応用規格において別途多重化の方法を規定することで対応している。

MP4 は、米国アップル社のマルチメディアコンテンツ配信サービスと共通性があり、第 3 世代携帯端末の動画ダウンロードサービスや、SD (Secure Digital) カードにおける蓄積フォーマットとして広く使用されている。

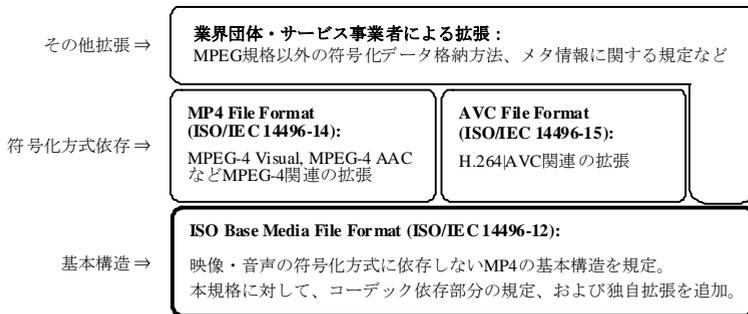


図 3・4 MP4 ファイルフォーマットの構成規格

MP4 ファイルは、ボックスと呼ばれる基本単位 (図 3・5) から構成され、各ボックスには、格納するデータの種類に応じて、アルファベット 4 文字からなる固有のタイプが割り当てられる。MP4 ファイルを構成する代表的な三つのボックスについて、図 3・6 を参照して説明する。

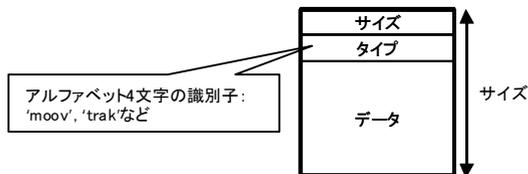


図 3・5 MP4 ファイルのボックス

まず、ムービーボックス (Movie Box : 'moov') は、符号化データのフレームサイズ、格納位置を示すアドレス情報、および、復号・表示時刻などのヘッダ情報を格納する。符号化データはメディアデータボックス (Media Data Box : 'mdat') と呼ばれるボックスに、フレーム単位で復号順に格納される。MP4 ファイルの再生時には、ムービーボックスを参照して所望の再生時刻に相当するフレームの格納アドレスを特定し、そのアドレスに基づいてメディ

アデータボックスからフレームデータを取得する。ファイルの先頭には、MP4 ファイルの再生互換性を示すブランド情報を示すファイルタイプボックス (File Type Box : 'ftyp') が配置され、端末は、ファイルタイプボックスを参照してファイルの再生可否を判定する。

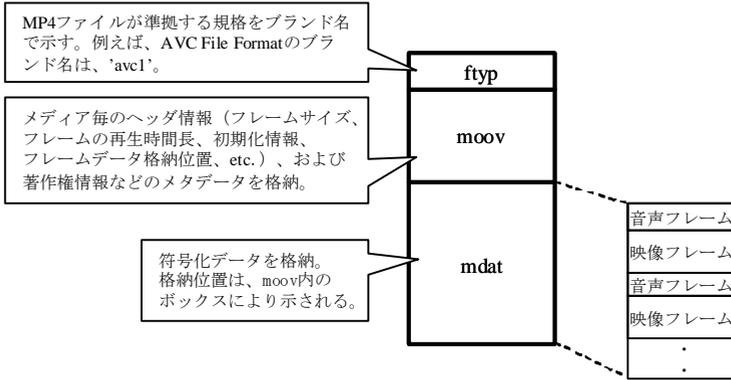


図 3・6 MP4 ファイルの構造

また、MP4 はファイルのランダムアクセスにも対応しており、ムービーボックス内のシンクサンプルボックス (Sync Sample Box : 'stss') によりランダムアクセス可能なフレームが示される。

現在普及している MP4 ファイルの構成としては図 3・6 の形式が一般的であるが、ヘッダ情報がすべてムービーボックスに格納されることから、コンテンツの再生時間長に比例してムービーボックスのサイズが増加する。その結果、プログレッシブダウンロードのように、ダウンロードしながら再生する場合でも、再生開始前にムービーボックスを取得する必要があり、再生時間までの待ち時間が長くなるなどの問題がある。このような問題を解決するために、コンテンツを任意の区間に分割し、分割した単位ごとにヘッダ情報を付加していく、ムービーフラグメントと呼ばれる方式も用意されている (図 3・7)。

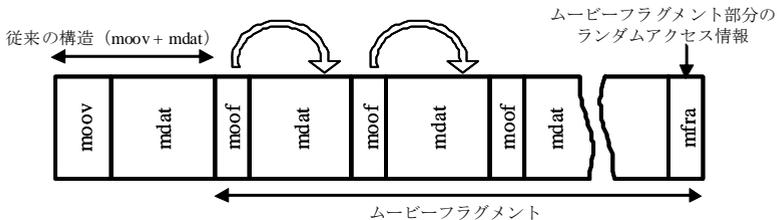


図 3・7 ムービーフラグメント

■参考文献

- 1) ISO/IEC 13818-1, “Generic coding of moving pictures and associated audio information: Systems”.
- 2) IETF RFC 3550, “RTP: A Transport Protocol for Real-Time Applications”.
- 3) IETF RFC 2326, “Real Time Streaming Protocol (RTSP)”.
- 4) ISO/IEC 14496-12, “ISO Base Media File Format”.

## ■2群 - 11編 - 3章

### 3-2 ストリーミング

(執筆者：宮川 和，谷崎文義) [2011年2月 受領]

ストリーミングとは、音声や映像などのマルチメディアデータをサーバから配信し、クライアントで再生する方式の一つである。通常、クライアントでデータを扱う場合は、データのすべてが受信されていることを前提とする。しかし、一般的に映像データはサイズが大きいため、受信完了を待って再生を開始した場合、非常に時間がかかってしまう。そこで、データの配信とほぼ同時にクライアントで受信したデータを再生することにより、再生待ち時間を大幅に短縮する。このような配信・再生方式をストリーミングと呼ぶ。

本節は、ストリーミングを利用した映像コンテンツ配信サービスとして IPTV サービスを概説する。また、ストリーミングに類似した配信方式であるプログレッシブダウンロード、および、分散型の配信方式である P2P について説明する。ストリーミングは高速・高品質な通信回線を前提とするが、プログレッシブダウンロードや P2P はインターネット網などの低速・低品質な通信回線で有効な方式である。

#### 3-2-1 IPTV サービス

IPTV (Internet Protocol Television) とは、IP ネットワークを用いて映像コンテンツを様々な形で配信するサービス、またはその技術の総称である。

IPTV サービスの試みは古くから行われてきたが、いわゆる TV 放送サービスに比肩するサービスとして認知され始めたのは 2002 年ごろからである。当初の IPTV サービスでは、各サービス提供会社が独自の方式を用いてシステム、端末を開発し提供してきた。しかし、新聞、ラジオ、TV に続く新たなメディアに発展するためには、端末の普及、コンテンツの流通を促進させる相互運用性の確保が不可欠であった。そこで 2005 年ごろから、IPTV サービスの仕様共通化に向けた取り組みが各国で開始された。

ITU-T (国際電気通信連合 電気通信標準化部門) は、2006 年 4 月に IPTV フォーカスグループ (FG-IPTV) を発足した。現在は、各研究委員会 (SG) と複数の研究委員会にまたがる合同会合 (IPTV-GSI : IP Television Global Standards Initiative) において IPTV の勧告化作業を進めている。日本国内においても、放送事業者や通信事業者、家電メーカーが参加する団体として 2008 年に「IPTV フォーラム<sup>1)</sup>」が設立され、技術仕様の第 1 版が 2008 年 8 月末に公開されている。

##### (1) IPTV における映像配信サービス

IPTV が提供する映像配信サービスとしては、(1)放送事業者による TV 放送サービスに相当する「IP 放送」、(2)見たいときに見たい番組が見られる「VOD」、(3)ファイル配信を用いた端末蓄積型の「映像ダウンロード」があげられる。本節では(3)については割愛する。

(1)は一对多通信である IP マルチキャストを、(2)は一对一通信である IP ユニキャストを用い、映像や音声を実時間再生するためのデータ伝送プロトコルである RTP (Real-time Transport Protocol) によって映像データを配信する (図 3・8 参照)。クライアント端末は、受信した映像データをバッファリングし、1~2 秒程度の遅延ののちに映像を再生する。スト

リーミングは、必要最小限な映像データのみをクライアント端末にバッファリングし、キャッシュデータなどを残さないため、その他の方式に比べて映像コンテンツを安全に扱える点で有利である。なお、VODではRTSP (Real Time Streaming Protocol) を用いてデータ伝送を制御し、早送りや巻き戻し、一時停止などの再生制御を実現する<sup>2)</sup>。

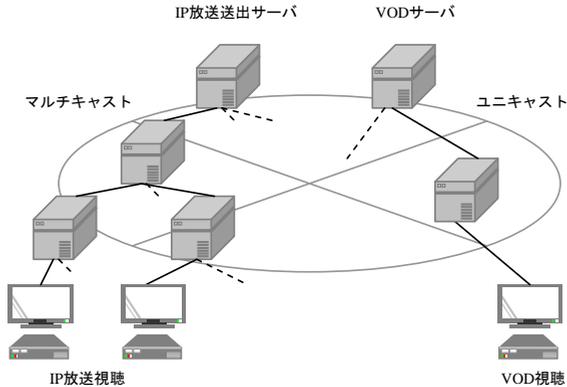


図 3・8 IPTV における映像配信サービスの概略図

リアルタイムでのデータ伝送では、データ欠損やネットワーク輻輳による遅延が想定される。FEC (Forward Error Correction : 前方誤り訂正) によるデータ誤り制御などが有効であるが、より高いサービス品質を維持するうえで、パケット優先制御やネットワーク帯域制御などの QoS (Quality of Service) 機能をもつ NGN (Next Generation Network : 次世代ネットワーク) と IPTV サービスとの親和性は高い。また、ハイビジョン画質、5.1 ch 音声、字幕などの魅力的なサービスを提供するうえでも、光ファイバを中心とした高速・大容量のブロードバンドネットワークは IPTV の普及に不可欠であると言える。

## (2) IPTV におけるストリーミングの今後の動向

「テレビ端末での TV 放送の視聴」が一般的であることから、IPTV は、「IP を用いた映像配信」と「従来の TV 放送サービスにおける QoE」を満たすべきサービスと言える。

QoE (Quality of Experience) とは、サービス利用者の経験値あるいは経験則のことである。例えば、テレビ端末の電源投入から映像が映るまでの時間や、ザッピング (チャンネルの切替え) 時間、数字ボタンによる直感的なチャンネル操作などがあげられる。地上デジタル放送が始まった当初は、アナログ放送に比べてザッピング時間が長いことが問題になった。また、地上デジタル放送の IP 再送信サービス向けのガイドライン<sup>3)</sup>では、QoE に関わる項目が多く規定された。これらは、映像配信サービスにとって QoE が重要であることの証左と言える。

IP 再送信サービスでは、ガイドラインを満たすため、QoE を含む様々な技術的課題を克服してきている<sup>4)</sup>。しかし、IPTV サービスは黎明期であり、今後も QoE の観点から様々な問題が生ずる可能性が高い。IPTV の標準化にあたっては利用者観点での QoE 要素とその指標が検討されており<sup>5)</sup>、今後は QoE の観点から、ストリーミングにおける更なる技術的要件の

精査やその実現が必要となっていくだろう。

### 3-2-2 プログレッシブダウンロード

プログレッシブダウンロードとは、ネットワークを使って動画や音声のファイル（以下、コンテンツファイル）を配信する方式の一つである。プログレッシブダウンロード以外の配信方式としては、ストリーミングやダウンロードがある。

プログレッシブダウンロードでは、配信サーバに WWW サーバを使用する。また、プロトコルに HTTP を使用する。使用する WWW サーバに特別な仕組みは必要ない。プログレッシブダウンロードの可否は、コンテンツファイルのデータ形式とその再生に利用するクライアント端末の再生用アプリケーションに依存する。

WWW サーバに置かれたコンテンツファイルを再生用アプリケーションから再生する場合、二つの方式がある。ひとつはダウンロード方式である。コンテンツファイルを WWW サーバから HTTP プロトコルによってクライアント側に伝送し、一旦、ハードディスクなどに保存して再生する。この方式では、コンテンツファイルをすべて伝送した後に動画を再生する。もうひとつの方式がプログレッシブダウンロードである。この方式は、ダウンロード方式と同様にコンテンツファイルを HTTP プロトコルによって伝送するが、再生を開始するために必要な最低限のデータを取得した時点で再生を開始する。また、追加のデータは再生と並行して伝送し、クライアントのキャッシュ領域などに保存して逐次再生する。

#### (1) プログレッシブダウンロードの利点

コンテンツファイルの伝送終了を待つことなくコンテンツが再生されるため、長尺なコンテンツや高品質なコンテンツでも再生までの待ち時間が少ない。また、コンテンツファイルのエンコードレートが大きくてもダウンロードできれば再生可能なため、クライアントの回線環境に依存せずにコンテンツファイルを再生できる。また、ストリーミングの場合はストリーミング専用のサーバアプリケーションが必要となるが、プログレッシブダウンロードの場合、既存の WWW サーバを利用して配信が可能である。

#### (2) プログレッシブダウンロードの欠点

クライアントのキャッシュ領域などにコンテンツファイルが残るため、コンテンツファイルが不正に二次利用されたり、改竄される恐れがある。このような不正利用を避けるためには、デジタル著作権管理 (Digital Rights Management : DRM) を併用する。また、動画や音声をリアルタイムに伝送するライブ配信には利用できない。

#### (3) 外部動向

Microsoft 社の WindowsMediaPlayer, Apple 社の QuickTimePlayer, Real 社の RealPlayer, Adobe 社の Adobe Flash Player など、多くのユーザに利用されているであろう再生用アプリケーションは、いずれもプログレッシブダウンロードに対応している。また、プログレッシブダウンロードを利用した代表的なサービスとしては、YouTube などの動画共有サービスがあげられる。多くの動画共有サイトでは一般的な WWW サーバアプリケーションと Adobe 社 Adobe Flash Player が利用されている。

### 3-2-3 P2P

P2P (Peer to Peer) とは、クライアントからの要求に対してサーバが処理を行い応答するクライアントサーバモデルと異なり、ネットワークに参加したコンピュータ (以下ノード) 間で自律的に直接通信を行い、相互に情報を交換しあうコンピュータネットワークモデルのことである。また、最近では、P2P モデルを利用しデータ伝送を行うアプリケーション群のことを P2P と呼ぶ場合がある。

#### (1) P2P モデルの分類

クライアントサーバモデルでは、利用者の欲しい情報はサーバに配置されるが、P2P では各ノードに分散して配置されている。そのため、P2P を、情報を検索するための方式に応じていくつかのモデルに分類している。情報の所在場所を蓄積して、ノードからの検索要求に応答するサーバ (インデックスサーバ) をもつ方式をハイブリッド P2P モデルと呼ぶ。インデックスサーバをもたず、ノード間の通信のみで検索を行う方式をピュア P2P 方式と呼ぶ。また、インデックスサーバをもたないかわりに、あるノード (スーパーノード) がインデックスサーバのように振る舞い検索要求に応答する方式をスーパーノード P2P モデルと呼ぶ。

#### (2) P2P の利点と欠点

現在、P2P モデルは、巨大なファイルを多くの端末に伝送するような場面で多く利用されている。例えば、オープンソースの OS のインストールイメージファイルや、コンピュータゲームのパッチファイルなどは、ファイルが巨大になることが多い。このようなファイル伝送をクライアントサーバモデルを利用して行おうとすると、多数のサーバと十分な容量の回線が必要となり、結果的に多くの費用が必要となる。しかし、P2P モデルを利用した場合は、各ノード間でデータが伝送されるため多数のサーバは必要なく、これらに関する様々なコストを削減できる。また、基本的にサーバを必要としないモデルであるため、アクセス集中によるサーバダウンなど、サーバに起因する故障が発生しない。一方で、サーバをもたないということは、ネットワーク参加のための認証やノードの振る舞いを記録する管理機能を組み込みにくいということになる。この欠点を補うために、それらの部分にのみクライアントサーバモデルを用いる場合が多い。

#### (3) P2P を取り巻く動向

P2P を利用した代表的なアプリケーションやサービスを紹介する。まず、世界的にみて最も多く利用されているアプリケーションは Bittorrent プロトコルを利用したファイル伝送アプリケーションである。Bittorrent プロトコルを実装したアプリケーションは米国の BitTorrent 社で開発されているが、Bittorrent プロトコルは公開されており、これを実装したクライアントアプリケーションが多数存在する。これ以外に、Skype Technologies 社が開発したインターネット電話アプリケーションの Skype や、インターネットテレビの Joost などがある。

日本国内では、P2P に関する社会的問題の影響で「P2P=違法」という見方がなされているが、「ネットワークの中立性に関する懇談会 WG2 P2P ネットワークの在り方に関する作業部会」の最終報告書 (2007 年 6 月 29 日)<sup>9)</sup> の結果を受け、P2P ネットワーク実験協議会を中心に、P2P に関する社会的理解の促進を目的とした「P2P を利用したサービス/ソフトウェアに

関するガイドライン<sup>7)</sup>」が策定され、また、同協議会主催による P2P を利用した実証実験が行われている。

#### ■参考文献

- 1) <http://www.iptvforum.jp/>
- 2) 亀山 渉, “デジタル・コンテンツ流通教科書,” インプレス, pp.57-87, 2006.
- 3) <http://www.nab.or.jp/shinsakai/>
- 4) 日経ニューメディア, “通信・放送融合時代の新ビジネス大全,” 日経 BP 社, pp.90-100, 2007.
- 5) 川添雄彦, 岸上順一, “今が楽しい IPTV サービス,” 信学誌, vol.91, no.6, pp.484-488, 2006.
- 6) [http://www.soumu.go.jp/s-news/2007/070629\\_11.html](http://www.soumu.go.jp/s-news/2007/070629_11.html)
- 7) [http://www.fmmc.or.jp/p2p\\_web/guideline.html](http://www.fmmc.or.jp/p2p_web/guideline.html)

## ■2群 - 11編 - 3章

### 3-3 マルチメディアメタデータ (MPEG-7)

(執筆者：國枝孝之) [2011年2月 受領]

MPEG-7 (エムペグセブン) は、ISO/IEC JTC1 SC29/WG 11 Moving Picture Experts Group (MPEG) によって 2003 年に制定されたマルチメディア情報へのメタデータ表記方法の国際標準 (正式名称: 「マルチメディアコンテンツ記述インタフェース Multimedia Content Description Interface」) である。MPEG-1, MPEG-2, MPEG-4 がマルチメディア情報 (映像・音声) の圧縮に関する標準であるのに対して, MPEG-7 は, XML をベースとしたメタデータ記述をマルチメディア情報に付加する方法・体系を定めた規格となる。これにより映像などのマルチメディア情報もテキスト情報と同様の検索が行えるようになる<sup>1)</sup>。

#### 3-3-1 メタデータとマルチメディア情報

映像・音声や画像といったマルチメディア情報は, 実際には数値の羅列でありそれ自体は意味情報をもたない。そのためテキスト情報に対してキーワード検索ができるようにマルチメディア情報を効率良く検索する手段として, 対象とするマルチメディア情報から検索対象となる特徴量をあらかじめ抽出しておき, これら特徴量を直接の検索対象とする。これら特徴量は元のマルチメディア情報 (データ) に対するデータであることから, メタデータ (情報への情報) と呼ばれる<sup>2)</sup>。

#### 3-3-2 MPEG-7 のアプローチ

MPEG-7 は, マルチメディアコンテンツを検索するために特徴を記述する規格であるが, 記述体系として二つのアプローチを規定している。

- (1) **低レベル記述**: コンピュータで処理可能な信号レベルの特徴量をメタデータ記述しておき, マルチメディア情報が与えられたときに抽出した特徴量と照合するための記述方法

図 3・9 に示した国旗の例では, (1)色ヒストグラム (色の構成が共通のもの), (2)輪郭線分布 (形の傾向が共通のもの), (3)色分布 (色の配置が共通のもの), (4)オブジェクト形状 (特徴的な形が共通のもの), といった様々な視点の特徴量を抽出して記述するための記述子 (Descriptor) を規定している。これらを組み合わせることで必要な情報を絞り込んで探し出すことができるようになる<sup>3)</sup>。

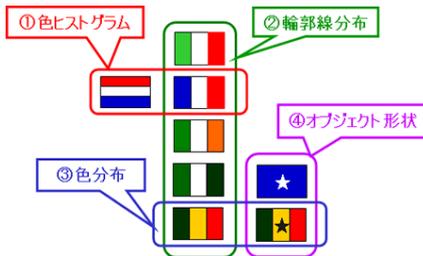


図 3・9 信号処理で抽出される特徴量

音声データの場合は、音声特徴量を抽出しておくことで、ハミングで歌って入力すれば、その旋律をもった音声データが探し出せるようになる。

- (2) **高レベル記述**：人間が判断した特徴量（意味情報）をメタデータ記述としてマルチメディア情報に関連付けておく方法

例えば、野球の映像に対して、それぞれのシーンごとに試合日、イニング、球場名、選手名、球種、バッティング結果などを映像の構造情報と合わせてメタデータとして映像に関連付けておき、キーワードを指定することで映像の中から見たい選手のホームランシーンだけを選んで視聴することが可能である。また、番組として構成された映像ではなく、情報の受け手である視聴者側が自分の嗜好に応じて検索し視聴できるようになる。

このように MPEG-7 は自動的に抽出できる特徴量を記述するための低レベル記述子と、人間が判断した意味情報をメタデータとしてマルチメディア情報に関連付けておくための高レベル記述子の二つの記述子を規定し、更に映像・音声の論理構造や視聴者の嗜好情報などを統合的な記述スキーム（Description Scheme）として規格を定めているのが大きな特長である。更に、これらメタデータ記述を映像や音声と多重化して送出するために XML をバイナリに自動変換する方式 BiM<sup>4)</sup>（Binary Format for MPEG-7）も規定している。

### 3-3-3 MPEG-7 の現状と今後の展開

MPEG-7 の応用は多岐にわたるが、実在する MPEG-7 のメタデータ記述を用いたシステム例として電子図書館システム<sup>5)</sup>を紹介する。電子図書館システムでは、蔵書、論文、ジャーナルなどの情報を提供するだけでなく、講義も一つのコンテンツとしてビデオ収録しインターネット上の公開講座や e ラーニング教材として提供している。これらの講義映像に対し MPEG-7 準拠のメタデータを付加することで、蔵書から講義映像までを横断的に検索できる仕組みを利用者に提供できるようになった。更に講座名や講師名での検索だけではなく、講義内容に関しての検索や目的のシーンの頭出しといったサービスを提供している。

MPEG-7 で定められている記述の項目（タグの種類）は、420 種類以上にも及び、メタデータ記述のための作業には膨大なコストがかかる。今後、低レベル記述の自動抽出・照合技術のさらなる進化が望まれると同時に、人手による意味情報付加が必要な高レベル記述に対するソーシャルタギングのようなアプローチも期待されている。

#### ■参考文献

- 1) MPEG-7 Overview, <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
- 2) MPEG-7 Japan Domestic Home Page, <http://www.itscj.ipsj.or.jp/mpeg7/index.html>
- 3) 國枝孝之, 脇田由喜, 高橋 望, “MPEG-7 と映像検索,” CQ 出版, 2004.
- 4) オーディオビジュアル複合情報処理シンポジウム 2001 論文集, 情報処理学会.
- 5) NAIST Digital Library, <http://library.naist.jp/mylibrary/portal>

## ■2群 - 11編 - 3章

### 3-4 コンテンツ流通フレームワーク (MPEG-21)

(執筆著：妹尾孝憲) [2011年2月 受領]

安全で便利なコンテンツ流通のためには、単に映像や音声データを送るのみでは不十分で、欲しいコンテンツが簡単に検索できる機能や、不正使用やコピーを防ぐ著作権管理機能、暗号化コンテンツとその解読鍵を別々に流通させることにより、プロバイダの権利を守りながら、ユーザが自由にコンテンツを視聴できる超流通の機能などが不可欠である。

本節では、これらの機能を備えたコンテンツ流通のためのフレームワークとして、マルチメディア符号化標準 (MPEG-21)<sup>1)</sup> を概説する。MPEG-21 は網羅的な記述であり、その一部の機能を実現している例として携帯電話の著作権保護方式 (OMA DRM)<sup>2),3)</sup> と光ディスクの著作権管理システム (DVD DRM)<sup>2)</sup> を紹介する。

#### 3-4-1 MPEG-21

国際標準化機構 (ISO) と国際電気標準会議 (IEC) の共同技術委員会 (JTC1) の下部組織である動画専門家グループ (MPEG: 正式名称は ISO/IEC JTC1/SC29/WG11) では、コンテンツ流通のために各種の標準を作成している。MPEG-1, MPEG-2, MPEG-4 は、映像・音声圧縮符号化標準であるが、MPEG-7 は、コンテンツの検索や選択に有用な映像や音声の特徴を付随情報 (メタデータ) として記述するための標準である。MPEG-21 は、マルチメディアフレームワークと名付けられ、コンテンツ流通に必要な識別や著作権保護のためのメタデータ標準である。更にこれらの標準を、映像配信や音楽配信などの具体的なアプリケーションごとに、最適に組み合わせた標準セットとして、MPEG-A が標準化されている (図 3・10)。MPEG-B、C、D は、その後新しく標準化された AV 多重化のためのシステム標準、映像符号化標準、音声符号化標準であり、MPEG-E は、これらをミドルウェアとして組み込むための標準である。最近では更に、仮想世界間や現実世界とのインタフェース記述と、風・熱・振動・湿度・光などの多感覚情報の記述を標準化する MPEG-V も進んでいる。



図 3・10 MPEG 標準一覧

これらの標準の中で、MPEG-21 は、MPEG-7 のコンテンツ記述を拡張したもので、マルチメディアコンテンツの構造や、識別子、著作権保護情報、コンテンツの再生に関する記述などを標準化している (図 3・11)。以下に主なものを概説する。

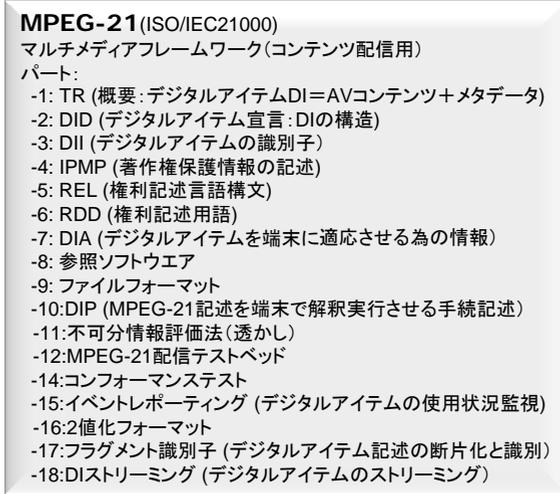


図 3・11 MPEG-21 標準体系

### (1) DID (デジタルアイテム宣言)

MPEG-21 では、コンテンツをデジタルアイテム (DI) と呼ぶが、DID は、コンテンツの先頭に置かれて、その構造を示す。アイテムは複数あっても良く、コンテナと呼ばれるフォルダに入れる。アイテムは、その中に AV コンテンツをコンポーネントとして含む。これらはすべて、コンテンツの属性を XML で記述したものであり、実際の映像や音声データとは、リソースと呼ばれるポインタでリンクされている。超流通では、先ず軽い DID だけをダウンロードし、内容をチェックして気に入れば、購入などの手続き後、アイテムの再生を始めると、AV データがダウンロードまたはストリーミングされる。

### (2) DII (デジタルアイテム識別子)

DII は、コンテンツの識別コードであり、我が国のコンテンツ ID フォーラムが策定したフォーマット (cIDf) が基本となっているが、ISRC (International Standard Recording Code) や、ISBN (International Standard Book Number) なども取り入れ、どのコードが使われているかを識別コードで区別する構造になっている。

### (3) IPMP (著作権保護情報)

IPMP (Intellectual Property Management & Protection) は、著作権保護の互換性を確保するために、権利記述の他にどのような暗号化ツールをダウンロードすれば良いかを記述したものである。その情報は、IPMP コンテナの中の IPMP 情報を記述する IPMP 情報記述子の中のツールエレメントに記述されている。この中には、コンテンツ保護に使われた暗号などの復号ツールの仕様記述やツールの参照先が記述されている。更に、入手したツールの初期設定条件や、入手したコンテンツの視聴権利範囲の記述子があり、これらの内容証明のためのシグ

ネチャが最後に添付される。

#### (4) その他のパート

REL (権利記述言語構文) は、著作権情報記述言語の文法と意味を規定したものであり、そこで使われる用語は、RDD (権利記述用語) で規定されている。DIA (デジタルアイテム適合) は、サーバとクライアント間で、コンテンツ形式の適合化を行うための記述であり、端末の機能・能力、ユーザ特徴、ネットワーク機能・能力などの記述ツールと、ビットストリーム構造やチャンネル QoS などのリソースを適合させるための記述ツールと、セッションモビリティやデジタルアイテム構成などを適合させるための記述ツールから成る。

### 3-4-2 OMA DRM

OMA DRM は、通信事業者や端末機器メーカーで構成された国際標準化団体 OMA (Open Mobile Alliance) が策定した、携帯電話用の著作権保護仕様である。OMA DRM 1.0 は、比較的簡易な著作権保護管理方式 WAP (Wireless Application Protocol) である。この中で、Forward Lock モードでは、コンテンツは暗号化されず権利管理メッセージも送られないが、携帯電話自身がコンテンツの他への出力を禁止する。Combined Delivery モードでは、非暗号化コンテンツとコンテンツ再生回数などの権利管理メッセージとがセットで端末に送られる。Separate Delivery モードでは、暗号化されたコンテンツと、その鍵と権利管理メッセージとが別々に配信されるので、鍵の入ったメッセージを購入するたびに、コンテンツの視聴が何回でもでき、超流通が実現される。OMA DRM 2.0 では、携帯電話以外の著作権保護機能をもつ機器でもコンテンツの視聴が可能のように、コンテンツ視聴が許可された機器をドメインで管理する機能をもつ。コンテンツの暗号鍵と権利情報は、ROAP (Rights Object Acquisition Protocol) と呼ばれるプロトコルで送られる。

### 3-4-3 DVD DRM

初期の DVD DRM は CPTWG (Copy Protection Technical Working Group) が作成した、コピー禁止に主眼を置いた著作権保護方式で CSS (Contents Scramble System) と呼ばれ、ディスク内のコンテンツを暗号化し、メディアであるディスクと、デバイスであるプレーヤがそれぞれもつ秘密鍵で相互認証を行った後に、コンテンツの再生を許可する方式であったが、著作権保護を不正に解除する方法が広まったため、その後、プロテクション機能を強化した CPPM (Contents Protection for Pre-recorded Media) と、記録メディアに使われる CPRM (Contents Protection for Recordable Media) が作成された。これらでは、機器の秘密鍵が解読された場合、その機器以外の機器でのみコンテンツ再生を許可する MKB (Media Key Block) を、メディアにもたせるようにして、不正視聴を防いでいる。Blu-Ray ディスクでは、業界 8 社が作成した更に高度な AACS (Advanced Access Content System) を用いており、権利保護範囲を、コンピュータネットワークやデジタル放送にまで拡張できるようになっている。

#### ■参考文献

- 1) 妹尾孝憲, “MPEG-21, MPEG-A の概要とその目的,” 情報処理, vol.48, no.10, pp.1118-1122, 2007.
- 2) 安田 浩, 小暮拓世, “DRM の技術動向,” 信学誌, vol.91, no.3, pp.225-236, 2008.

- 3) 上野秀俊, 住田正臣, 石川憲洋, “OMA における DRM の標準化動向,” NTTDoCoMo テクニカルジャーナル, vol.12, no.4, pp.52-59, 2005.

## ■2群 - 11編 - 3章

### 3-5 著作権管理技術

(執筆著者：小暮拓世) [2011年2月 受領]

#### 3-5-1 著作権管理技術概要

出版、映画や音楽に代表される創造的な著作物については、それを利用してなにがしかの対価や利益を得る権利があると法律（著作権法）で保護されている。創造的な制作物の定義はあいまいさがあるものの、一般常識として単なる数値データの並び替えや、ニュース報道には著作権が生じないとされる。

故に著作権の管理とは創造的、個性的な制作物に対する権利行使に関して許諾や提供に係る管理手法に係る。デジタル化された著作物の流通に関しては、特別な管理技術が必要となる。デジタル著作物は、基本的にコピー（複製）、加工の精度が向上し大量複製が可能である。対価や許諾を伴うデジタル著作物の流通過程に必要な管理技術は暗号化技術や複製保護システムであり、それらのもつ秘匿性や独占性から、技術内容の詳細はあまり公表されない。

デジタル著作物の著作権管理技術を DRM (Digital Rights Management) と総称する場合がある。DRM 技術は様々なコンテンツ配信のシステムや要素技術に使われているが、その実態は、単独の暗号ツールの適用だけの簡単なものから、電子すかし、複雑な認証システムまで含むシステムまでを含み、それぞれ使われている要素技術も様々である。

#### 3-5-2 DRM 技術発展の経緯

DRM 技術で扱うデジタルコンテンツは、デジタルアイテム (Digital Item) と総称される部類の一部である。デジタルコンテンツで扱う DRM は、デジタルコンテンツの複製回数や複製そのものを管理者側で制限する意図で開発され発展してきた技術であり、配信者の視点で、映画や音楽など有料著作物の無秩序な消費を防ぐために開発された技術である。したがって、基本的には、オリジナルコンテンツを非公開の秘匿方式によって暗号化し、管理情報をメタデータ化して記録しコンテンツ制御する目的に利用される。配信された暗号化デジタルコンテンツの再生には、特定のソフトウェアあるいはハードウェアのみで再生する。DRM 技術の導入で、一般のコンテンツ消費者が複製や再利用を難しくする事は可能になる。これは一種のコピーガード技術であり、このコピーガード技術を導入したコンテンツ配信のシステムを DRM システムともいう場合もあり、あるいは、コピーそのものの制限を目的に設計されたシステム技術とも言える。

最近の DRM システム開発の動きでは、互換性や相互運用性を意識したシステム提案やその標準化の試みもある。

デジタルアイテム（著作物）の流通環境では、コンテンツ自体の流通という価値の連鎖 (Value Chain) と連携した、対価（お金）の連鎖 (Money Chain) を伴う統合システムがソリューションの一つで、その事業化も始まっている。

DRM システムの標準化事例としては、任意の国際標準化団体 DMP (Digital Media Project) での活動がある。DMP では、統合 DRM 技術をベースにしたコンテンツ流通モデルの概念を詳述し、それに基づく抽象化流通モデルを具象化している。その DMP モデルを用いたディ

デジタルコンテンツ流通の相互運用性の定義と流通参照モデルを定義した。現在、互換性ターミナル第3フェーズの定義まで進み Money Chain の枠組みを入れた相互運用性 DRM モデルも提唱されている。

### 3-5-3 初期の DRM 技術例

初期の DRM の例は DVD に見る事ができる。DVD が市場に登場したのは 1996 年頃で、扱い易さと記録容量の妥当性から成長が期待されたが同時にコピー制御の方式と動作に消費者の関心が集まった。

## DVDビデオのDRM

### 目的 記録済みDVDのコンテンツを保護する

- ①コンテンツ事業者からデジタルコピー防止対策を求められる
  - 端末機器事業者が中心となりコンテンツ暗号方式CSSを開発 DVDに導入
  - 第3者機関による「鍵」管理機構を立ち上げる
- ③アナログ映像出力の録画禁止も考慮
  - マクロビジョン社開発のマクロビジョン信号(録画妨害信号)をコンテンツに挿入



図 3・12 DVD に適応する DRM の例

図 3・12 は、初期の DVD の DRM の仕組みをアナログ方式と比較して示した図である。実際のコピー制御は CSS (Content Scrambling System) というデジタル暗号化方式を採用したコンテンツ暗号化によるコピー防止システムである。すなわち、記録するコンテンツデータを CSS で暗号化し CSS 解読装置を内蔵したプレーヤーでしか再生できない仕組みである。本方式では、マスター、ディスク、タイトルという 3 種類の鍵データがある。タイトルの鍵は DVD に記録されるタイトルごとに設定される鍵であり、コンテンツの暗号化時に使用される。ディスク鍵は DVD ディスクごとに設定される鍵であり、当該ディスクで使用されるタイトル鍵の暗号化時に用いる。マスター鍵は DVD 機器の製造者に個別に割り当てられる鍵であり、これを使ってディスク鍵を暗号化する。この鍵は DVD ディスクの導入部に記録する。

暗号解読は DVD 機器や再生ソフトがもつマスター鍵で DVD ディスク鍵を解読し、ディスク鍵でタイトル鍵を解読してコンテンツの暗号を解読し、オリジナル コンテンツを復元する。

DVD 市販後、CSS 暗号解読ソフトが出現し、インターネットなどを通じて知れ渡った。このため現在では CSS は使われていないが、本格的な DRM 技術の量産メディアへの採用は、CSS が初めてであり、その後の DRM 技術の基礎を築いた例と言えよう。当時は未だアナロ

グVTR との共存の過度期であったため、蓄積メディア環境では、アナログ信号経由のコピーにも制御がかかる方式が採用され、図 3・12 に示した DVD コンテンツのアナログコピーを禁止する信号の導入も存在する。PC (Personal Computer) ネットワーク環境においても PC バス (Bus) を用いたバス認証方式で DVD メディアの DRM に適用している。

### 3-5-4 実際の DRM 技術

前述の CSS が解読されてしまった以降、DVD 業界では、DRM の強化策として導入したのが CPRM (Content Protection for Recordable Media) 及び CPPM (Content Protection for Pre-recorded Media) である。CPRM は記録メディア、CPPM は、再生専用メディアに適用する。図 3・13 は CPRM 方式のコピー制御方式であり、コンテンツ配信事業者、メディア事業者、端末機器事業者の間で合意された暗号化方式と暗号解読鍵の管理、機器間の相互認証がシステムの中心である。

暗号化方式などは、CSS を踏襲しているが、強化策として、メディア固有の識別 ID を DVD の BCA (Burst Cutting Area, 最内周部分) に書き込み、コンテンツデータは暗号化して記録する方式である。再生側は、メディア識別 ID を識別しなければ復号できない。すなわち、再生機が CPRM に対応 (デバイスキーを再生機がもつ) する必要がある。再生時は、このメディア ID と別の MKB (Media Key Block) によって作られる暗号鍵と機器のもつデバイスキーで復号する。仮に暗号鍵が破られてもメディア側の MKB データを更新すれば解読できなくなる。

記録されたコンテンツは、すべて暗号化されており、再生時は、機器側がもっている機器鍵 (デバイスキー) とメディア側に記録されている MKB を使用して暗号を解くために必要となる「メディア鍵」を生成しコンテンツの暗号化を解き、再生を行う。

## メディア (DVD) の著作権制御保護技術

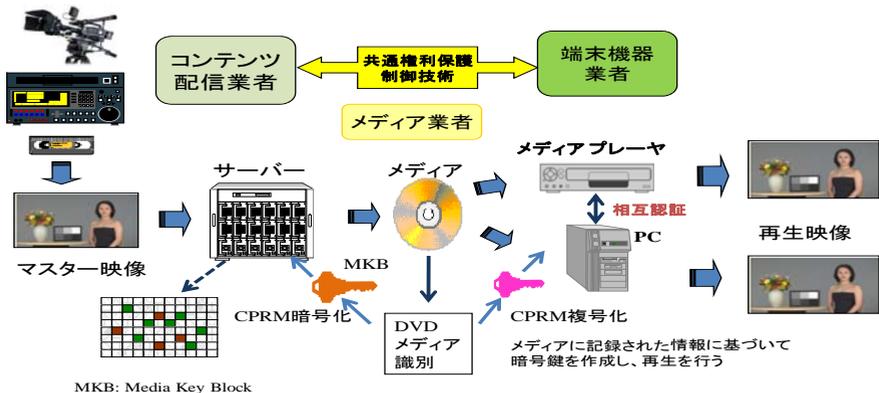


図 3・13 DVD メディアにおける著作権保護の概要

DRM は暗号化データの生成とそのデータを再生する再生機器が所定の暗号化方式に対応

して初めて成り立つシステムであることから、特定のシステムに依存したシステムになりやすい傾向がある。すなわち、公開型の標準には馴染まない方式と言える。

### 3-5-5 電子透かし

DRM 技術は大別すると暗号化・認証技術と電子透かし技術とがある。電子透かし技術は、コンテンツに一義的に特化して摘要し、コンテンツの流通経路のトレースに必要な情報をコンテンツに直接埋め込む技術である。システムの秘匿性から、電子透かし情報の埋め込み場所や、埋め込み情報、その技術内容の詳細は公開されない。電子透かし信号情報は、コンテンツに直接埋め込むので、オリジナルコンテンツ情報にとってはノイズになる。このノイズを検知されないレベルでの埋め込み情報量を確保することが、埋め込み技術の要点であり、DRM システム技術としてではなく、単独の DRM 要素技術として扱われることが多い。

#### (1) 電子透かしの要求仕様

画像データや音声データなどのデジタルコンテンツにおいて、それらのデータの冗長性などを利用して、他のデータ（例：著作権者名や管理番号など）を埋め込み（データの一部を置き換える）「見えない」あるいは「聞こえない」状態で秘匿させて、必要に応じて開示し、コンテンツ著作権などの管理に利用する著作権管理技術を、電子透かし技術「Digital Watermark」という。

実際に電子透かし技術を用いて何らかの情報をコンテンツに埋め込む場合、埋め込まれたコンテンツは何らかの影響を受ける。したがって、それらを最小限にし、かつできるだけ多くの情報量を埋め込み、コンテンツの編集によって、埋め込まれた情報が改変あるいは欠落しないことが要求される。

電子透かし技術に要求される仕様は、

- (1) コンテンツの中に秘匿すべき情報を電子透かしで埋め込んだ状態は、コンテンツ利用者が埋め込まれた情報をそのままでは認識できないこと、すなわちコンテンツの利用時に埋め込んだことが見えない、あるいは、聞こえないこと
- (2) 埋め込まれた電子透かし情報は、必要な時に抽出することができること、かつ、コンテンツに埋め込まれた情報は確実に認識することができること
- (3) コンテンツを多少編集加工しても電子透かしで埋め込まれた情報は確実に残りそれを抽出することができること。
- (4) コンテンツの品質を保ったままで、透かし情報を除去することは著しく困難であること。

ここで、電子透かしを埋め込む対象物（画像や音声など）をキャリア（Career）といい、あるいはコンテナ（Container）ともいうことがある。電子透かしによって埋め込まれた情報は、通常の人間では知覚されないことが原則である。すなわち、キャリアが画像であれば眼で見ても外見的には何も変化がないこと、キャリアが音の場合は、音を聞いても判別できないことが要求される。

更に、電子透かし技術によって透かし情報を埋め込んだ場合でも、ファイルの大きさが変わらないことも大切である。

## (2) 電子透かしに利用できるキャリアデータの特徴

実際の電子透かしは、複数のキャリアの特質を利用した組み合わせ技術により実現する。利用可能なコンテンツ視聴の性質は

### (1) キャリアデータの一部をあるルールに従い変更する

キャリアデータの一部を別のデータに変更すると、当然、元のデータとは異なるのでコンテンツの視聴に影響がある。その影響が人間の知覚限界内であれば、コンテンツ品質に問題ないとする。

### (2) デジタルコンテンツの冗長的性質を利用する

画像や音声のデータは冗長性が高いので画像や音声のデータを一部置き換えても、全体的なコンテンツの品質に影響しないと考えられる。この性質を利用して、目立たない部分に透かしデータを埋め込む。

冗長性とは、コンテンツ(対象画像)のコア部ではなく例えば、背景に相当するような補助データともいえる領域である。

### (3) 人間の知覚的な特徴を利用する

人間の視覚には共通する生理的な特徴があり、細かい図柄の変化は検知されにくい、色の変化は輝度の変化に比べて検知されにくいなど、電子透かしによる情報埋め込みに都合の良い特徴的な性質がある。これらの性質を利用することで、電子透かしの埋め込みを行う。

## (3) 実際の電子透かし埋め込み技術

### (a) 輝度情報を利用した電子透かし

画像情報のうち、輝度情報の冗長性を利用した電子透かしがある。この方式はビット置換型、あるいは、画素置換型ともいわれている。実際の電子透かしによるデータの埋め込みには、画素データの規則的な変換が必要である。ここでいう規則性とは画素値のデータ列で、テキストデータ、あるいは2値画像などのデータもある。テキストデータであれば、バイナリーデータに変換の後にビット列に変換し、埋め込みデータ(文字コード変換のビット列)とする。具体的な変換例では、

\*透かし情報が ビット0のときは画素データを偶数に変換する

\*透かし情報が ビット1のときは画素データを奇数に変換する

変換は、近い値で増加規則として最大値のみ減少規則を適用する

これらの簡単な変換を複雑化し、埋め込み箇所を拡散し、暗号化するなど、秘匿性を高めることは可能である。

### (b) 画像の周波数特性を利用した電子透かし

画像を周波数領域に変換するには、フーリエ変換を行い、 $\sin$   $\cos$  成分に分けてその成分構成比で表現できる。この三角関数の成分に分解すると、入力信号がどの成分で構成されているかを分析し特徴抽出に利用できる。これを利用した電子透かしの埋め込み方法は、いくつか考えられる。その中で、ウェーブレット変換で用いられる電子透かしでは、スペクトル拡散方式がある。

電子透かしで埋め込む情報は、文字情報などのコード表現ができる例が多い。その場合、

周波数スペクトルは特定領域に集中する傾向がある。他方、一般の画像情報は周波数成分が比較的拡散する傾向がある。したがって、透かし情報を拡散させると画像周波数成分の中に透かし信号成分が拡散し、視覚的に目立たなくなる。

#### (c) 画像圧縮を利用した電子透かし

一般に、画像情報は大容量で情報に冗長度が多いとされ、この冗長度を利用した静止画像圧縮として JPEG 画像圧縮がある。JPEG の基本原理は原画像を周波数成分に分解することにある。MPEG は、離散的コサイン変換 (DCT) する。DCT 利用の簡単な電子透かし埋め込みの実際の処理は以下の方法がある

- \* 電子透かしデータのビットが 1 であれば DCT 係数の至近の奇数に変換
- \* 電子透かしデータのビットが 0 であれば DCT 係数を至近の偶数に変換
- \* 最初から上記条件を満たす場合はそのままにする
- \* これらの埋め込み変換操作で係数が 0 になる場合は、埋め込み操作を止める。

#### (d) 色情報を利用する電子透かし

カラー画像の色情報を操作して情報を埋め込む電子透かし方法がある。RGB などのカラー信号から輝度信号を作り (YUV 変換) 輝度情報として電子透かしを埋め込む方法は輝度情報利用の場合と同様である。

カラー情報を直接利用する電子透かし技術としては

- \* 全画素の色情報をわずかに変化させる方法
- \* 上記の全画素の代わりにブロック分割する方法
- \* 透かし部分のパラメータ変換を行い透かし情報を得る方法

などがある。

### 3-5-6 電子透かし埋め込み技術の実際

#### (1) 幾何学的改変に対する耐性をもつ電子透かし

電子透かしの耐性検証ツールとして、StirMark モデルがある。StirMark による攻撃は、幾何学的な改変 (拡大, 縮小) を含む。そこで、このような、最新の攻撃に対する耐性をもつ電子透かしとしては、近接画素の連鎖を抽出し、その連鎖から透かし鍵を求める方法がある。

#### (2) 任意形状領域の切り貼りに対する耐性をもつ電子透かし

任意形状領域の切り貼りに対する耐性をもつ電子透かしとは、画像の一部を切り抜き、別の画像に貼り付ける画像処理方式がある。すなわち、2 値のテキスト画像を用いて透かし情報とする。対象とする画像の最下位ビットプレーンを当該画像で置き換えることで透かし画像を得る。このテキスト画像を解析すれば、埋め込み画像が得られる。基本はテキスト画像の特徴から埋め込み領域を特定する方法と言える。