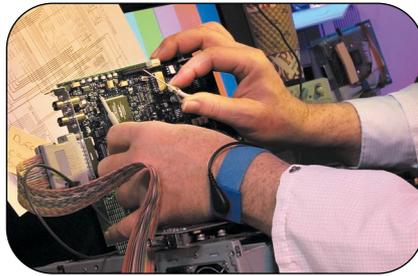


MPEGの基礎とプロトコル解析入門

～ DVB、ATSC、その他～



目次

第1章 MPEGの基本的な知識	1
1.1 技術の統合	1
1.2 圧縮が必要な理由	1
1.3 圧縮の原理	1
1.4 テレビのアプリケーションにおける圧縮	1
1.5 デジタル映像圧縮の初歩知識	2
1.6 音声圧縮の初歩知識	4
1.7 MPEGストリーム	4
1.8 監視と解析の必要性	5
1.9 圧縮に関する注意点	5
第2章 映像の圧縮	6
2.1 空間符号化と時間符号化	6
2.2 空間符合化	6
2.3 ウエイティング	7
2.4 走査	8
2.5 エントロピー符合化	8
2.6 空間コーダ	8
2.7 時間符合化	9
2.8 動き補正	9
2.9 双方向符合化	10
2.10 I/P/Bピクチャ	11
2.11 MPEG圧縮器	12
2.12 前処理	14
2.13 ウェーブレット	15
第3章 音声の圧縮	16
3.1 聴覚のメカニズム	16
3.2 サブバンド符合化	17
3.3 MPEGレイヤ1	17
3.4 MPEGレイヤ2	18
3.5 変換符号化	18
3.6 MPEGレイヤ3	18
3.7 MPEG-2オーディオ	18
3.8 MPEG-4オーディオ	19
3.9 AC-3	19
第4章 MPEGの規格	20
4.1 MPEGとは	20
4.2 MPEG-1	20
4.3 MPEG-2	21
4.3.1 MPEG-2のプロファイルとレベル	21
4.4 MPEG-4	22
4.4.1 MPEG-4の規格文書	23
4.4.2 オブジェクト符合化	23
4.4.3 映像および音声の符号化	23
4.4.4 スケーラビリティ	24
4.4.5 MPEG-4のその他の側面	25
4.4.6 MPEG-4の今後	26

4.5	MPEG-7	26
4.6	MPEG-21	27
第5章 エレメンタリ・ストリーム		28
5.1	ビデオ・エレメンタリ・ストリームの構文	28
5.2	オーディオ・エレメンタリ・ストリーム	29
第6章 パケット化エレメンタリ・ストリーム(PES)		30
6.1	ESパケット	30
6.2	タイムスタンプ	30
6.3	PTS/DTS	31
第7章 プログラム・ストリーム		32
7.1	録画と伝送	32
7.2	プログラム・ストリームの初歩知識	32
第8章 トランスポート・ストリーム		33
8.1	トランスポート・ストリームの役割	33
8.2	パケット	34
8.3	プログラム・クロック・リファレンス(PCR)	34
8.4	パケット識別子(PID)	35
8.5	番組特定情報(PSI)	35
第9章 デジタル変調		37
9.1	変調の原理	37
9.2	アナログ変調	37
9.3	直交変調	37
9.4	単純なデジタル変調方式	37
9.5	位相偏移変調	38
9.6	直交振幅変調(QAM)	38
9.7	残留側波帯変調(VSB)	39
9.8	符号化直交周波数分割多重(COFDM)	40
9.9	統合デジタル放送(ISDB)	40
9.9.1	ISDB-S衛星放送方式	40
9.9.2	ISDB-Cケーブル放送方式	40
9.9.3	ISDB-T地上波変調	41
9.9.4	ISDBのまとめ	41
第10章 DVBおよびATSCの初歩知識		42
10.1	概要	42
10.2	再多重化	42
10.3	番組配列情報(SI)	43
10.4	エラー訂正	43
10.5	チャンネル符合化	44
10.6	内部符合化	44
10.7	伝送術数	45

第11章	データ放送	46
11.1	アプリケーション	46
11.1.1	プログラム関連データ	46
11.1.2	オポチュニスティック・データ	46
11.1.3	ネットワーク・データ	46
11.1.4	エンハンスドTV	46
11.1.5	インタラクティブTV	46
11.2	コンテンツのカプセル化	47
11.2.1	MPEGのデータ・カプセル化	47
11.2.1.1	データ・パイピング	47
11.2.1.2	データ・ストリーミング	47
11.2.1.3	Digital Storage Medium Command and Control(DSMCC)	47
11.2.1.4	マルチプロトコルのカプセル化(MPE)	47
11.2.1.5	カルーセル	47
11.2.1.6	データ・カルーセル	48
11.2.1.7	オブジェクト・カルーセル	48
11.2.1.8	オブジェクト・カルーセルが放送される仕組み	48
11.2.1.9	MPEG-2データの同期化	49
11.2.2	DVBのデータ・カプセル化	50
11.2.3	ATSC A/90のデータ・カプセル化	50
11.2.4	ARIBのデータ・カプセル化	50
11.3	データ・コンテンツの伝送	50
11.3.1	DVBのアナウンス	50
11.3.2	ATSCのアナウンス	50
11.4	コンテンツの表示	50
11.4.1	セット・トップ・ボックスのミドルウェア	50
11.4.2	DVB Multimedia Home Platform (MHP)	51
11.4.3	ATVEF DASE	51
11.4.4	DASE	52
第12章	MPEGのテスト	53
12.1	テストの要件	53
12.2	トランスポート・ストリームの解析	53
12.3	階層表示	53
12.4	解析表示	55
12.5	構文およびCRC解析	55
12.6	フィルタリング	56
12.7	タイミング解析	56
12.8	エレメンタリ・ストリームのテスト	58
12.9	Sarnoff [®] 適合ビット・ストリーム	58
12.10	エレメンタリ・ストリームの解析	58
12.11	トランスポート・ストリームの生成	58
12.12	PCR誤差の発生	59
用語解説		60

第1章 MPEGの基本的な知識

MPEGは、最も普及している音声・映像圧縮技術の1つで、単一の規格ではありません。同様の原理に基づき、用途に合わせてさまざまな規格が策定されています。MPEGはMoving Picture Experts Groupの頭文字で、ISO(International Standards Organization: 国際標準化機構)とIEC(International Electrotechnical Commission: 国際電気標準会議)が設立した合同専門委員会JTC1に属しています。JTC1は情報技術を担当しており、JTC1内のサブグループSG29は、「音声、画像、マルチメディア、ハイパーメディア情報符号化」を担当しています。SG29内にはさらにいくつもの作業グループ(WG)があり、たとえばJPEG(Joint Photographic Experts Group)や、動画圧縮を担当するWG11などがあります。MPEGは、ISO/IEC JTC1/SG29/WG11です。

MPEGの説明には、さまざまな頭文字が使われます。たとえば、ETSIのように「CATとは、使用されるCAシステムに関連付けられたEMMを、IRDが見つけられるようにするためのポイントである」ということもできます。もちろん、これが理解できるのであれば本書は必要ありません。

1.1 技術の統合

デジタル技術が音声・映像の分野で急速な進歩を遂げたのには、いくつかの理由があります。デジタル情報は劣化に強く、符号化によるエラーがほとんどありません。そのため、録画・録音や伝送で生じる損失をなくすることができます。このことを、家電製品として、初めて実証したのがCDです。

CDによって、音質はレコード盤より向上しましたが、進歩したのはそれだけではありません。本当に重要なのは、デジタルの録画・録音技術や伝送技術によって、コンテンツの処理や編集が、アナログでは不可能だった段階にまで進んだことです。音声や映像をデジタル化すると、そのコンテンツはデータ形式になります。このようなデータは、他の種類のデータと同じように処理できます。そのため、デジタル・ビデオ(映像)、デジタル・オーディオ(音声)は、コンピュータ技術の一部になっています。

コンピューティングやPCM(Pulse Code Modulation: パルス符号変調)のような発明によって、コンピュータと音声・映像が統合されたのは必然の結果といえます。デジタル媒体はどのような種類の情報でも保存できるため、コンピュータの記憶装置をデジタル・ビデオに使用するのは簡単です。ノンリニア・ワークステーションは、アナログ世代にはなかった統合技術を応用した最初の例です。また、マルチメディアは、音声、映像、グラフィックス、文字、データなどを同じメディア上で組合せます。アナログには、マルチメディアに相当するものはありません。

1.2 圧縮が必要な理由

デジタル・ビデオが最初に普及したのは、重ね合せや効果を自在に使用できることで、高いコストに見合った利益が得られるポスト・プロダクション向けの用途です。しかし、プロダクションにおける標準デジタル・ビデオでは、200Mbps以上の速さでデータが生成されるため、このビット・レートでは、大容量の記憶装置と広い周波数帯域が必要になります。つまり、記憶容量と周波数帯域の条件を緩和できれば、さらに広い分野でデジタル・ビデオを使用できるわけです。これが圧縮の目的です。

圧縮とは、デジタル・オーディオやデジタル・ビデオのデータ量を少なくするための方法です。圧縮には次のようなメリットがあります。

- ある一定量の素材の保存に要する記憶容量を少なくできる。
- リアルタイム動作では、必要とする周波数帯域を狭くできる。また、テープとディスクのようなメディア間では、リアルタイムよりも速く転送できる。
- 圧縮記録フォーマットは低い記録密度を使用でき、レコーダが環境や保守の影響を受けにくくなる。

1.3 圧縮の原理

情報コンテンツを伝送するデータ量を小さくするのに、基本的に2種類の方法が使用されています。実際の圧縮方式では、通常、この2つはととも複雑に組み合わせられています。

1つは、符号化の効率を高めるものです。情報の符号化にはさまざまな方法があり、映像や音声のシンプルなデータ表現には、相当量の冗長性が含まれています。この冗長性とエントロピイの概念については、後述します。

冗長性の低減または除去には、さまざまな符合化技術を使用できます。たとえば、ラン・レングス符号化や、ハフマン符号のような可変長符号化方式です。これらは、正しく使用すれば、伸張(解凍)により圧縮前のデータを完全に復元できる可逆圧縮の技術です。この可逆圧縮は、ロスレス(無損失)とも呼ばれています。PKZipのようなアーカイブ・プログラムは、ロスレス圧縮を採用しています。

ロスレス圧縮が明らかに理想的ですが、一般に、ロスレス圧縮だけでは映像や音声のアプリケーションに必要なレベルのデータ圧縮率は得られません。しかし、損失がないため、圧縮方式のどの箇所にも使用でき、非可逆圧縮器のデータ出力にもよく使用されます。

冗長性を除いても、必要なレベルまでデータを削減できないと、情報の一部を捨てるを得なくなります。非可逆圧縮方式では、重要性のない、または重要性の低い情報を除いてデータを削減します。これは、どのようなデータ・ストリームにも応用できる一般的な技術でなく、各情報の重要性は、そのアプリケーションの範囲内で、データの内容や用途を検討して、はじめて判断できることです。テレビの場合では、人間の視覚と聴覚に対して画像や音声を示す上で、人的要因を十分に理解して、有効な圧縮システムを設計することが重要です。

ビデオ信号には、人間の視覚では認識できない情報も含まれており、この場合、そのような情報は、重要ではありません。このように、重要でない画像情報のみを捨てる方式は、視覚的なロスレス圧縮と呼ばれています。

1.4 テレビのアプリケーションにおける圧縮

アナログ、デジタルを問わず、テレビ信号は常に膨大な情報を表現するので、周波数帯域を削減する技術も、かなり早期から使用されてきました。そうした技術の中で最も古いものは、おそらくインタレースです。ある一定の走査線数と画像更新速度で、インタレースは必要な周波数帯域を2:1に削減します。これは非可逆処理です。インタレースは、垂直情報と時間情報の干渉によって、信号劣化を発生させ、その画像で使用できる垂直解像度を低下させます。それでも、この程度の犠牲はほとんど無視できるため、インタレースは当時としてはシンプルで有効な手段でした。

第1章 MPEGの基本的な知識

しかし、高度なデジタル圧縮方式になると、インタレースとその信号劣化による影響は深刻になります。MPEG-2で複雑になっている部分のほとんどは、インタレース信号に対応するためのもので、プログレッシブ信号と比べても、符号化の効率はかなり落ちます。

次の大きな転機は、カラー技術とともに訪れました。カラー・カメラはGBR信号を生成するため、理論的には情報量がモノクロ信号の3倍になりますが、その一方ではモノクロ信号と同じチャンネルで、カラー信号を送送する必要がありました。

これを実現するためにまず必要なのは、GBR信号をルミノランス信号(通常はYで表す)と2つの色差信号UとV、またはIとQに変換することでした。ルミノランス信号を発生できても、モノクロ受像機と互換性を保つには、まだ長い道のりがありました。しかし、色差信号の研究によって、周波数を最低限に抑えられる大きな進歩がありました。

人間の視覚には輝度を感じるセンサがあり、非常に高い解像度まで見分けられることがわかっています。色情報を感じるセンサもありますが、解像度はぐっと低くなります。その結果、ある一定の範囲内では、景色の輝度を表す鮮明なモノクロ画像に、ぼやけた(周波数帯域の低い)色情報を重ねても、鮮明なカラー画像に見えます。GBR信号では、各信号に輝度(ルミノランス)情報と色(クロミナンス)情報が両方とも含まれているため、この原理を活用できません。しかし、YUV領域であれば、ルミノランス情報のほとんどがY信号に変換され、色差信号にはほとんど含まれません。その結果、色差信号をフィルタリングして、伝送情報を大幅に削減できます。

これは重要でない情報(の大部分)を取り除く例です。想定された視聴条件下では、人間の視覚は色差信号の高周波情報にそれほど反応しないため、捨てることができます。NTSCテレビジョン伝送では、500kHz程度の色差信号しか伝送しませんが、多くのアプリケーションで十分に鮮明な画像が得られています。

NTSCやPALで周波数帯域を削減するための最終ステップは、モノクロ信号スペクトルの未使用部分に、色差信号を「見えないように挿入する」ことです。この処理は厳密にはロスレスではありませんが、信号の符号化効率を向上させる方法と考えることができます。

デジタル世界では、一部の技術は大きく異なるものの、同じ原理が応用できます。たとえばMPEGでは、信号を別の領域に変換することで、重要でない情報を分離しやすくします。色差信号領域への変換も使用されています。しかし、デジタル技術では、この色差信号をフィルタリングすることで垂直解像度を下げ、さらに削減を行います。

図1-1aのように、従来のテレビジョン方式では、制作・編集を行うためGBRカメラ信号をY、Pb、Prコンポーネント信号に変換し、

さらにアナログ・コンポジット信号にエンコードして伝送します。それが、最新の技術では図1-1bのようになります。Y、Pb、Pr信号はデジタル化され、SDI(Serial Digital Interface: シリアル・デジタル・インタフェース)形式のY、Cb、Cr信号で制作・編集された後、MPEGに符号化して伝送されます。MPEGは放送事業者にとって、コンポジット・ビデオよりも明らかに効率の高い手段と考えられます。さらにMPEGでは、アプリケーションに応じて必要なビット・レートを調整できるため、柔軟性も高くなります。また、ビット・レートや解像度を下げれば、MPEGをテレビ会議やテレビ電話に使用することもできます。

欧州デジタル・テレビ放送規格のDVB(Digital Video Broadcasting)と、米国デジタル・テレビ放送規格のATSC(Advanced Television Systems Committee)は、必要な周波数帯域が非常に高いため、圧縮技術なしでは運用できません。圧縮はDVD(Digital Video/Versatile Disk: デジタル・ビデオ/多用途ディスク)の再生時間を伸ばし、長編映画を1枚のディスクに収められるようにします。また、ENG(Electronic News Gathering: 電子ニュース取材)や、その他のテレビ制作に必要な作業コストを削減します。DVB、ATSCおよびDVDは、いずれもMPEG-2圧縮をベースにしています。

テープ録画の分野では、Digital Betacam^{*1}やDigital-S^{*2}では、圧縮率を低くすることによって、許容範囲を広げ、さらに信頼性を高めているのに対し、SX^{*1}、DVC、DVCPRO^{*3}、DVCAM^{*1}などは小型化を目指しています。ディスクベースのビデオサーバでは、圧縮でストレージ・コストを低減できます。圧縮すると周波数帯域も下るため、より多くのユーザがサーバにアクセスできるようになります。この性質は、VOD(Video On Demand: ビデオ・オン・デマンド)の用途でも重要です。

1.5 デジタル映像圧縮の初歩知識

実際の番組素材は、必ず2種類の成分で信号が構成されています。新しく予測できない成分と、予測できる成分です。前者はエントロピーと呼ばれ、これが真の信号情報です。後者は重要でないため、冗長性と呼ばれています。冗長性には、隣り合う画素がほぼ同じ値になる単色の広い部分のように空間的なものと、連続するピクチャ間に見られる時間的な相似性のものがあります。すべての圧縮方式は、エンコーダでエントロピーと冗長性を分離して圧縮します。記録・伝送されるのはエントロピーのみで、伝送された信号からデコーダが冗長性を算出します。図1-2aは、この概念を示しています。

理想エンコーダは、エントロピーをすべて抽出し、そのエントロピーのみをデコーダに伝送します。そして、理想デコーダは元の信号を復元します。ただし、これはあくまで理想であって、実

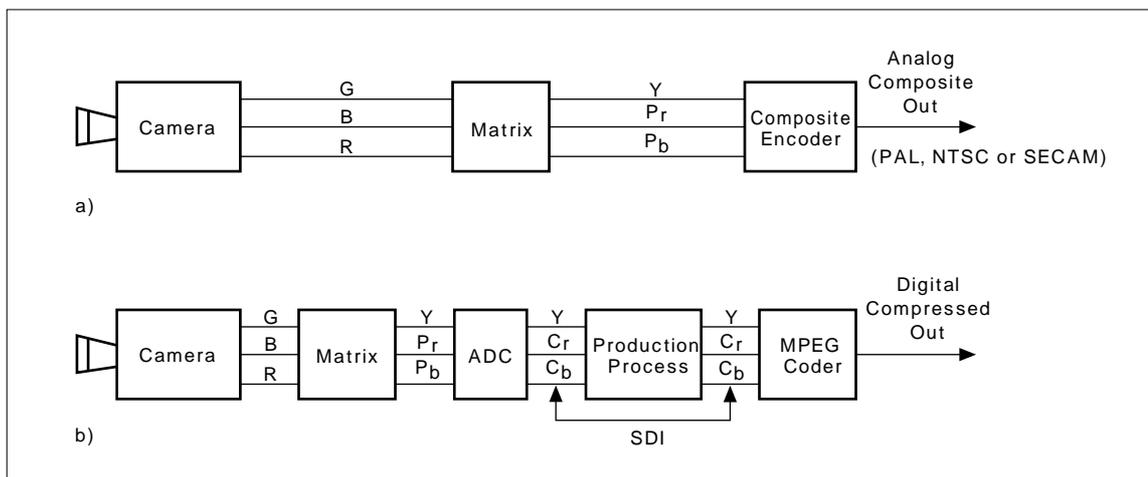


図1-1

際には不可能です。理想のコードは複雑で、時間的な冗長性を利用するために、非常に長い遅延を生じることがあります。録画・録音や放送のような用途では、多少の遅延は許されても、テレビ会議では許されません。また、複雑なコードは非常に高価になってしまいます。このように、単独で使用できる理想圧縮システムは存在しないといえます。

実際には、処理遅延や複雑さが異なるさまざまなコードが必要です。1つの圧縮フォーマットでなく、標準化されたさまざまな符号化ツールを柔軟に組合せて、多彩な用途に対応できるのがMPEGの強みです。どのように符合化されたかが圧縮データに含まれているため、コードがどのように判断して符合化したものでも、デコーダは自動的に復号できます。

MPEG-2とMPEG-4の符号化は、複雑さの異なる複数のプロファイルに分かれていて、各プロファイルは、入力画像の解像度に応じて別々のレベルで実行できます。プロファイルとレベルの詳細については、第4章で説明します。

デジタル・ビデオにはさまざまなフォーマットがあり、それぞれビット・レートが異なります。たとえば、HD方式のビット・レートは、SD方式の6倍になることもあります。したがって、コード出力のビット・レートだけを知っていても、あまり役に立ちません。重要なのは、入力ビット・レートと圧縮ビット・レートの比率を表す圧縮率で、たとえば2:1や5:1のように示されます。

ただし、数多くの変動要素が関わっているため、適切な圧縮率を決めるのは非常に困難です。図1-2aのような理想コードで、エントロピがすべて伝送されると、高い品質が保たれます。しかし、ビット・レートを下げするために圧縮率を上げると、エントロピーの一部が伝送されなくなるため、品質は低下します。圧縮率による品質低下は急勾配を描きます(図1-2b)。十分なビット・レートが得られない場合は、この品質低下を避けるため、入力画像のエントロピーを少なくした方が得策です。これには、フィルタリングを使用します。主観的には、フィルタリングによる解像度の低下のほうが、圧縮による信号劣化よりも許容できます。

エントロピーを完璧に識別できる理想的な圧縮器は、非常に複雑なものになってしまいます。実際の圧縮器は、経済的な理由か

らそれほど複雑にはせず、多めのデータを送ることで、すべてのエントロピーが伝送されるようにする必要があります。図1-2bは、コードの複雑さと性能の関係を示しています。圧縮率を高くしようとすると、それだけエンコーダを複雑にしなければなりません。

ビデオ信号のエントロピーは、様様ではありません。ニュースを伝えるアナウンサの録画は冗長性が多く、簡単に圧縮できます。これに対し、風に吹かれる木の葉や、絶えず動いているサッカーの観衆は冗長性が少ない情報、つまりエントロピーが多いため、圧縮は難しくなります。いずれの場合も、すべてのエントロピーが送信されないと、品質が低下します。このため、選択肢はビット・レートが一定で品質が変動するチャンネルと、品質が一定でビット・レートが変動するチャンネルのどちらかになります。放送事業者や通信事業者は、実用面を考えて、ビット・レートが一定の方を選ぶ傾向にあります。遅延が増大してもかまわなければ、バッファ・メモリを使用して、エントロピーの変動を平均化することもできます。録画の場合は可変ビット・レートの方が処理しやすいので、DVDでは可変ビット・レートを使用しています。バッファによる平均ビット・レートがディスク・システムの能力範囲内に収まるようにしています。

ピクチャ内符合化は、空間的な冗長性、つまり各ピクチャ内の冗長性を利用する技術で、ピクチャ間符合化は、時間的な冗長性を利用する技術です。ピクチャ内符合化は、JPEG静止画像規格の場合のように単独で使用したり、MPEGの場合のようにピクチャ間符合化と組合せて使用したりすることができます。

ピクチャ内符合化は、一般的な画像が持っている2つの性質を基本にしています。その1つはすべての空間周波数が同時に存在しないという性質で、もう1つは空間周波数が高くなるほど振幅は狭くなりやすい、という性質です。ピクチャ内符合化では、画像の空間周波数を解析する必要があります。これは、ウェーブレット変換やDCT(Discrete Cosine Transform: 離散コサイン変換)を用いた解析です。この変換によって、各空間周波数の大きさを表す係数が得られます。通常、多くの係数はゼロまたはほぼゼロになり、それらは除外できるため、ビット・レートが削減されます。

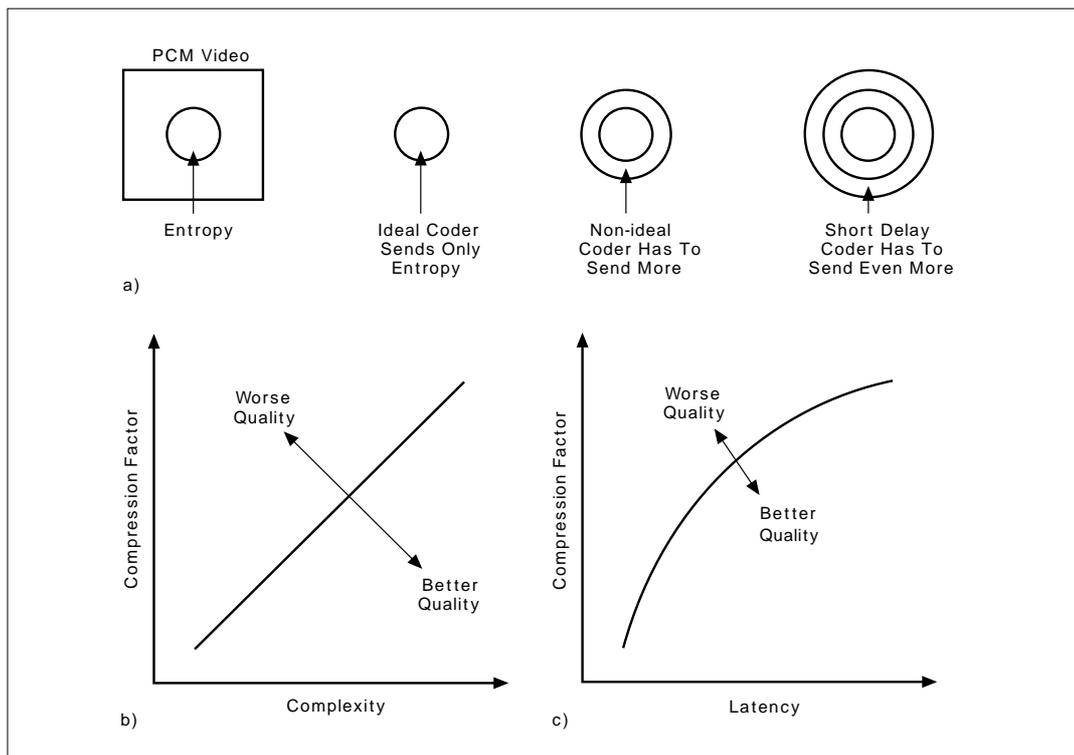


図1-2

第1章 MPEGの基本的な知識

ピクチャ間符合化は、連続するピクチャ間の類似性を見つけることを基本にしています。1枚のピクチャのデータがデコーダにある場合、その次のピクチャは差分を送るだけで作成できます。物体が動くとピクチャの差分は増大しますが、動いていても、全体的にはピクチャ間の外観がそれほど変わらないので、増大した分は動き補正によって相殺できます。動きが測定できれば、直前のピクチャの一部を新しい位置に移動することで、そのピクチャにより近いものを作成できます。この移動処理は、デコーダに送られる縦方向と横方向のずれ値(合せて「動きベクトル」と呼ぶ)によって調整されます。動きベクトルの伝送に必要なデータは、ピクチャ差分データよりも少なく済みます。

MPEGは、インタレース画像とプログレッシブ画像の両方に対応できます。フィールド、フレームに関係なく、時間軸上のある地点の画像を「ピクチャ」と呼びます。インタレースはそれ自体が圧縮技術なので、デジタル圧縮には向いていません。フィールド間で画素の位置が異なるため、時間符合化が複雑になります。

動き補正は、連続するピクチャ間の差を最小限に減らしますが、除去するわけではありません。ピクチャ差分自体は空間画像であり、前述の変換をベースにしたピクチャ内符合化によって圧縮できます。ピクチャ差分データの量を削減するのが動き補正です。

時間コーダは、動作可能な時間範囲の長さに伴って効率が高まります。図1-2cに示すように、高い圧縮率が必要な場合は、入力の時間範囲を長くする必要があるので、符号化の遅延も長くなります。時間符号化された信号は、少し前に伝送された画像データに基づいてピクチャが出力されることもあり、明らかに編集が難しくなります。編集システムでは、編集を可能にするのに、時間符号化の程度を制限する必要があります。この制限によって、使用可能な圧縮率も制限されます。

1.6 音声圧縮の初歩知識

PCMデジタル・オーディオ・チャンネルのビット・レートは1.5Mbps程度で、これは4:2:2デジタル・ビデオの0.5%ほどしかありません。Digital Betacam[®]のように圧縮率の低い映像圧縮方式では、音声圧縮は必要ありません。しかし、映像圧縮率が高くなると、音声圧縮も必要になります。

音声圧縮は、2つの事実を利用しています。1つ目は、通常のオーディオ信号では、すべての周波数が同時に存在することはないことで、2つ目は、マスキングによって人間の聴覚はオーディオ信号の細部をすべて聞き分けられないことです。音声圧縮では、音声スペクトルをフィルタリングや変換によって各帯域に分割し、レベルが低い帯域を表す場合に、含まれるデータ量を少なくします。マスキングによって、ある帯域の可聴性は低下するものの、送信するデータはさらに少なくできます。

聴覚の能力を考えると、音声圧縮は、映像圧縮ほど簡単ではありません。マスキングは、マスキングされた音声と空間的に一致した場合にのみ効果があります。モノラル録音では常に空間的に一致しますが、ステレオ録音ではその音階の別の部分でも低レベル信号が聞こえるため、一致しません。このため、ステレオ・システムやサラウンド・システムでは、ある一定の品質に対して許容される圧縮率が低くなります。また、性能の悪いスピーカでは、残響によって圧縮の信号劣化がマスキングされてしまうことも、音声圧縮を複雑にする要因になります。性能の悪いスピーカで圧縮器をテストすると不正確な結果しか得られず、一見満足できる信号のように思えても、性能の良いスピーカで聞くとがっかりしてしまうことがよくあります。

1.7 MPEGストリーム

単一のMPEG音声または映像コーダからの出力を、エレメンタリ・ストリームと呼びます。エレメンタリ・ストリームは、エンドレスな準リアルタイムの信号です。エレメンタリ・ストリームは、PES(Packetized Elementary Stream:パケット化エレメンタリ・ストリーム)という管理しやすいサイズのデータ・ブロックに分割されます。このデータ・ブロックには、パケット開始を表すヘッダ情報が必要です。また、パケット化によって時間軸が失われるので、タイムスタンプも必要です。

図1-3のように、すべてのコーダが共通のクロックにロックしていれば、1つのビデオPESと複数のオーディオPESを組合せて、プログラム・ストリームを形成できます。また、PESのタイムスタンプで映像と音声を同期できます。プログラム・ストリームは、ヘッダが付いた可変長パケットを持っています。プログラム・ストリームは、光ディスクやハードディスクのデータ転送のように、基本的にエラーがなく、ファイル・サイズが任意の分野で使用されます。DVDは、プログラム・ストリームを使用しています。

伝送やデジタル放送では、複数の番組とそれに関連するPESを1本のトランスポート・ストリームに多重化できます。トランスポート・ストリームがプログラム・ストリームと異なるのは、PESパケットが固定サイズの短いパケットに再分割される点や、異なるクロックで符号化した複数の番組を送ることができる点です。これは、トランスポート・ストリームにはPCR(Program Clock Reference:プログラム・クロック・リファレンス)というメカニズムがあり、複数のクロックを伝送して、デコーダでその中から1つを選び、再生できるためです。また、SPTS(Single Program Transport Stream:シングル・プログラム・トランスポート・ストリーム)も可能で、これはコーダとマルチプレクサの間で使用されます。トランスポート・ストリームは、デコーダ・クロックをエンコーダ・クロックにGENロックできるので、SPTSの方がプログラム・ストリームよりも普及しています。

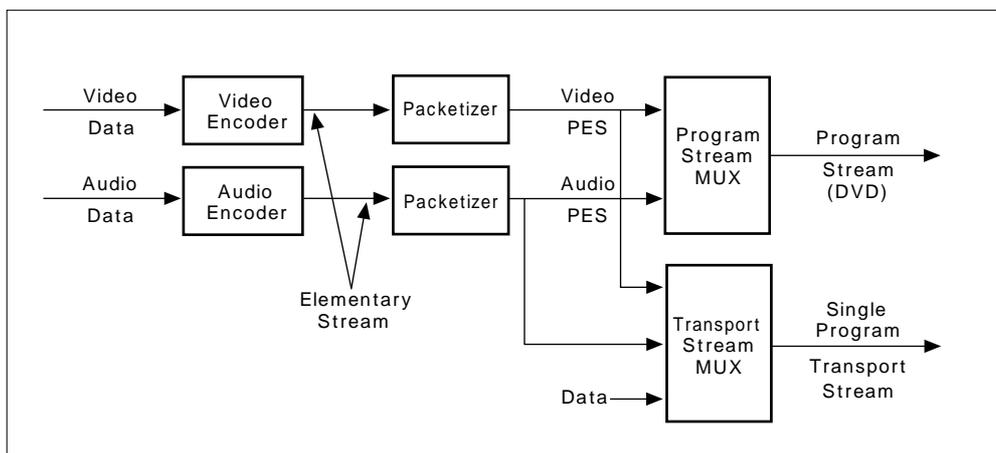


図1-3

トランスポート・ストリームは、単にオーディオPESとビデオPESを多重化したものではありません。圧縮されたオーディオ・データとビデオ・データのほかに、トランスポート・ストリームには、ビット・ストリームを表すメタデータが含まれています。このメタデータにあるPAT(Program Association Table: 番組関連テーブル)には、トランスポート・ストリームの番組がすべてリストされています。そしてPATの各項目は、それぞれの番組を構成するエレメンタリ・ストリームをリストしたPMT(Program Map Table: 番組マップ・テーブル)を指示しています。番組の中には自由に利用できるものもありますが、暗号化の対象になるものもあり、こうした情報もメタデータで伝えられます。

トランスポート・ストリームは、固定サイズのデータ・パケットで構成され、それぞれ188バイトになっています。各パケットには、PID(Program Identifier Code: 番組識別コード)があります。同じエレメンタリ・ストリームが搬送されるパケットのPIDはすべて同じなので、デコーダ(またはデマルチプレクサ)は、必要なストリームを選んで、それ以外を排除できます。パケットの連続性指標(Continuity Counts)により、ストリームの復号に必要なすべてのパケットを確実に受信できます。デコーダが各パケットの開始を正しく識別して、ビット・ストリームをワードにデシリアライズするには、適正な同期システムが必要です。

1.8 監視と解析の必要性

MPEGトランスポート・ストリームは非常に複雑な構造をしています。関連する各種テーブルとコード表示された各種識別子を用いて、複数の番組の中から特定のプログラムとエレメンタリ・ストリームを分離できます。各々のエレメンタリ・ストリームの内部は複雑な構造ですが、デコーダは各種ベクトルや係数、量子化テーブルなどを識別できます。

障害には、大きく分けて2つのカテゴリがあります。1つ目は、伝送システムは、ジッタの増加やビット・エラーを起こさずに、エンコーダ/マルチプレクサからデコーダへ正確に情報を伝送したにもかかわらず、エンコーダ/マルチプレクサ、あるいはデコーダに障害がある場合です。2つ目は、エンコーダ/マルチプレクサおよびデコーダに問題はありますが、双方間のデータ伝送に障害がある場合です。エンコーダ/マルチプレクサ、デコーダあるいはデータ伝送の中のどこに障害があるかを突き止めることが、迅速に解決策を見出す上で非常に重要です。

同期パターンの消失や破壊といった同期問題が生じると、トランスポート・ストリーム全体の受信が妨げられます。トランスポート・ストリーム・プロトコルに欠陥がある場合は、デコーダが番組データのすべてを見つけれないために、画像だけで音声は送られない、というようなことが起こります。また、データは正確に送られても、ジッタが極端に多いと、デコーダのタイミング問題が発生する可能性があります。

MPEGトランスポート・ストリームを使用しているシステムが障害を起こした場合、その障害はエンコーダ、マルチプレクサ、あるいはデコーダで発生していると考えられます。障害の切り分けは次のように行います。まず、トランスポート・ストリームがMPEG符号化規格に適合しているかどうか確認します。適合していない場合、デコーダが原因になっていることはまずありません。しかし、ストリームが適合している場合は、デコーダに注目する必要があります。

シグナル・ゼネレータ、波形モニタ、ベクトルスコープといった従来のビデオ計測器は、MPEGシステムを出入りするビデオ信号の品質が十分に確保されていることを確認する以外、MPEGシステムの解析には適しません。まず、受信機器やデコーダのテストには、有効なMPEGテスト信号を得られる信頼性の高いゼネレータが不可欠です。そして、適切なアナライザを使用することで、エンコーダ、伝送システム、マルチプレクサ、リマルチプレクサ

などの性能を高い信頼性で評価できます。当社は長年にわたり、高品質のビデオ計測器を開発・販売してきました。MPEGユーザが、複雑な圧縮システムが正常に機能していることを検証したり、問題がある場合には速やかに診断できるテスト/計測ソリューションを、今後も提供いたします。

1.9 圧縮に関する注意点

復号されたものがオリジナルと同一ではないという意味で、MPEG圧縮は非可逆です。エントロピは一樣ではなく、エントロピが高い素材を圧縮すると、復号したときに視覚的な信号劣化が残ってしまうことがあります。時間圧縮は、連続するピクチャ間に冗長性があることを前提にしています。そうでない場合は、圧縮がうまく行かないことがあります。たとえば、フラッシュが焚かれている記者会見の映像では、フラッシュが含まれている画像は隣同士でもまったく異なり、符号化による信号劣化がはつきり出てしまうこともあります。

不規則な動きや、別々の動きをする複数の物体がある場合は、大きなベクトル周波数帯域が必要になり、これを満たす唯一の方法は、画像データの周波数帯域を下げることです。この場合も、視覚的な信号劣化が出ることもあり、その程度は動きによって異なります。このような問題は、スポーツ中継の映像でよく発生します。

粗い量子化は、輝度のコンターリング(輪郭欠如)や色のボスタリゼーション(階調変更)の原因になります。これらは、単色の広い部分に斑点状の影やブロッキングとして表れることがあります。主観的には、アナログ・テレビの伝送システムで生じる比較的均一な品質低下よりも、圧縮による信号劣化の方が気になります。

圧縮率を下げるのが、これらの問題に対する唯一の解決策です。圧縮を利用する場合には、高い圧縮率による経済性と信号劣化とのバランスを考慮して、判断する必要があります。

時間符号化は、符号・復号の遅延が長くなるだけでなく、編集も困難にします。現に、MPEGビット・ストリームは任意の箇所での編集ができません。このような制約があるのは、時間符号化では、ピクチャの復号に前のピクチャのデータを必要としますが、編集によってそのデータがなくなってしまうためです。また、ピクチャが順序通りに送信されないことも、編集をより困難にしています。

適切な符号化を使用すれば、比較的ゆったりした間隔のスライス・ポイントに限り、編集を行うことができます。任意編集が必要な場合は、MPEGストリームに対して復号・修正・再符号化という処理を行う必要があるため、結果的にゼネレーション・ロスが生じます。

*1 ソニー株式会社の商標です。

*2 日本ビクター株式会社の商標です。

*3 松下電器産業株式会社の商標です。

第2章 映像の圧縮

本章では、人間の視覚を利用した映像圧縮について説明します。また、MPEGコーデカの仕組みを理解するため、使用されている変換や動き補正などの技術についても説明します。

2.1 空間符号化と時間符号化

第1章で説明したように、映像圧縮は時間的冗長性と空間的冗長性を利用しています。MPEGではまず、連続するピクチャ間の類似性を利用して、時間的な冗長性を減少させます。すでに伝送されたピクチャの情報から、できるだけ正確な現時点のピクチャを作成、つまり「予測」します。この方法では差分ピクチャを送るだけでよく、これによって実際のピクチャと予測との差をなくします。続いて、差分ピクチャは空間圧縮されます。ここでは便宜上、まず空間圧縮について説明してから、時間圧縮について説明します。

空間圧縮は、ピクチャ内の単色部分にある隣り合う画素間の類似性と、模様がある部分の主要な空間周波数に依存します。JPEG方式は静止画像を個別に伝送する技術なので、空間圧縮のみが使用されます。しかし、映像の連続するピクチャをJPEG方式で符号化することもできます。いわゆるMotion JPEGでは、時間符号化を使用した場合ほど圧縮率はよくありませんが、ビット・ストリームをピクチャ単位で自由に編集できます。

2.2 空間符号化

空間符号化ではまず、変換で空間周波数を解析します。変換とは、簡単にいえば波形を別の領域（この場合は周波数の領域）で表すことです。変換の結果、各周波数がどれだけ存在するかを表す一連の係数が得られます。また、逆変換で元の波形を復元できます。逆変換の際に、係数を十分な精度で処理すれば、元の波形と同じ波形が得られます。

最もよく知られている変換法がフーリエ変換です。フーリエ変換は、入力信号に含まれている周波数成分を求める方法で、入力波形に基底関数と呼ばれる各周波数のサンプルを掛け、その値を積分して、その周波数成分を求めます。図2-1のように、対象となる周波数が入力波形に含まれていない場合は積分した値がゼロになり、含まれている場合は積分した値がその周波数の振幅を表す係数になります。

周波数成分が基底関数と同位相の場合は、今述べた通りの結果になります。しかし、周波数成分が基底関数と直交している場合、積分した値はゼロになります。そのため、基底関数同士を直交させて、入力の全位相を検出できるようにして、各周波数を求める必要があります。

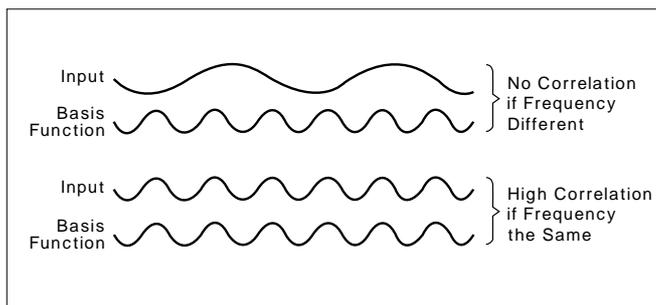


図2-1

フーリエ変換には、周波数ごとに正弦(サイン)成分と余弦(コサイン)成分の両方に、係数が必要になるという欠点があります。一方、余弦変換では、入力波形自体を時間的にミラーリングした後で基底関数を掛けます。このようにすると、図2-2のように、正弦成分はすべて打ち消され、余弦成分はすべて2倍になります。正弦成分用の正弦基底関数が不要になり、周波数ごとに必要な係数は1つだけになります。

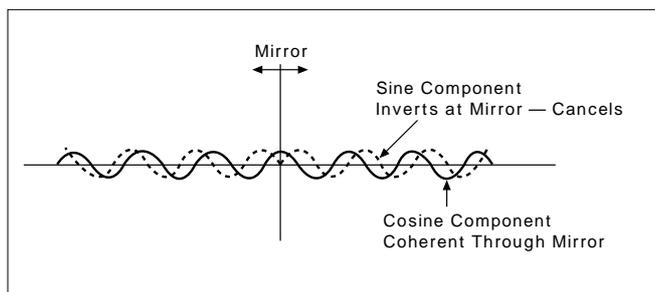


図2-2

DCT(離散コサイン変換)は代表的なコサイン変換方式で、MPEGの2次元形式で広く使用されています。8×8画素ブロックを、8×8係数ブロックに変換します。変換で分数を掛ける必要があるためワード長が伸び、係数値のワード長は画素値よりも長くなります。通常では、8ビットの画素ブロックが11ビットの係数ブロックになります。つまり、DCTは圧縮どころかデータを大きくしてしまうわけですが、元の画素が圧縮しやすい形式に変換されます。

DCT後の8×8ブロックの各係数を逆変換した結果を、図2-3に示します。ルミナンス信号の場合、左上の係数はブロック全体の平均輝度、つまりDC成分です。そこから右側へ行くほど水平空間周波数が高くなり、下側へ行くほど垂直空間周波数が高くなります。実際の画像では、垂直・水平のさまざまな空間周波数が同時に発生するため、ブロック内のある係数が表す水平・垂直周波数には、あらゆる組合せが考えられます。

また図2-3では、8つの係数が1次元の水平波形としても表されています。この波形にさまざまな振幅と片方の極性を組合せることで、8画素のあらゆる組合せを復元できます。2次元DCTの64個の係数を組合せると、元の8×8画素ブロックになります。もちろん、

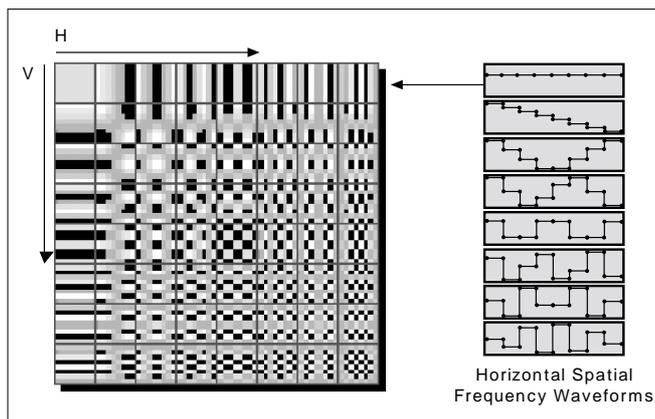


図2-3

カラー画像では、色差サンプルも処理する必要があります。Y、Cb、Crのデータを8×8の各配列に組入れて、それぞれ変換します。

実際の番組素材のほとんどでは、係数の多くがゼロや、ゼロに近い値ですので、伝送されません。これにより、実質的に損失のない圧縮が相当量行われます。高い圧縮率が必要な場合は、ゼロでない係数のワード長を短くする必要があります。これは係数の精度を下げて、損失を招くため、損失をできるだけ目立たないように注意します。

2.3 ウエイティング

図2-4のように、画像ノイズに対する人間の知覚は均一ではなく、空間周波数と関連して変化します。空間周波数が高くなるほど、多くのノイズが許容されます。また、詳細な画像部分では、画像ノイズは実質的にマスキングされますが、単色の部分では顕著になります。厳密な測定結果が主観的な結果により近くなるように、測定値をウエイティングするという方法は、従来からよく利用されています。

圧縮を行うと、係数の精度が下り、PCMでワード長の短いサンプルを使用したときと同じような影響があります。つまり、ノイズ・レベルが上がります。PCMでは、ワード長を短くすると、すべての周波数で等しくノイズ・レベルが上がります。DCTでは信号が複数の周波数に分割されるため、範囲ごとにノイズを管理することができます。実質的には、ウエイティングを行うことで、低い周波数域の係数のほうが高い周波数域よりも精度が高くなります。

図2-5のように、ウエイティング・プロセスでは、DCTブロックの係数を2次元周波数の関数である定数で割ります。低周波域の係数は小さい数字で割り、高周波域の係数は大きい数字で割ります。除算の結果は、端数を切捨てて最も近い整数にします。この切捨ては一種の再量子化です。ウエイティングされていないと、この再量子化の影響で、量子化ステップのサイズが均一に大きくなりますが、ウエイティングされていれば、除算に使用された数に

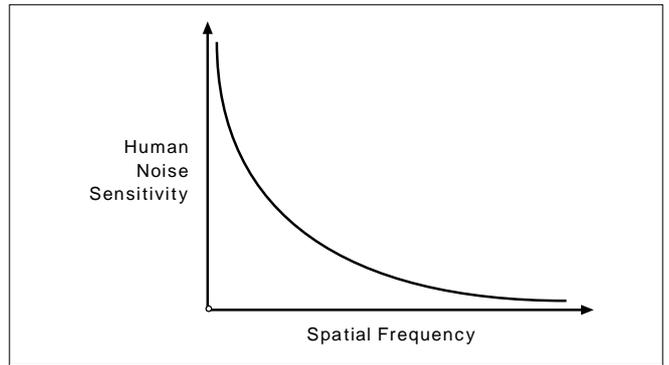


図2-4

じて大きくなります。

その結果、低い空間周波数の係数は比較的小さいステップで再量子化され、ノイズの増加が低く抑えられます。高い空間周波数の係数は大きいステップで再量子化され、ノイズが増加します。しかし、ステップ数が少ないということは、ステップの識別に要するビット数も少なくなるため、圧縮が得られます。

デコーダでは、低位のゼロを加えて、ウエイティング係数を正確な大きさに戻し、続いて逆ウエイティング係数を掛けます。高周波数域では、掛ける数が大きくなるため、再量子化ノイズも大きくなります。逆ウエイティングによって得られた係数は、元のDCTブロックの係数に再量子化エラーが付加されたものになりますが、このエラーは高い周波数で大きく、低い周波数では小さくなります。

切捨てを行う代わりに、ウエイティング係数を非直線的に再量子化することで、量子化ステップのサイズが係数の大きさに伴って大きくなるようにすることも可能です。この方法では、圧縮係数を高くできる反面、映像品質は悪化します。

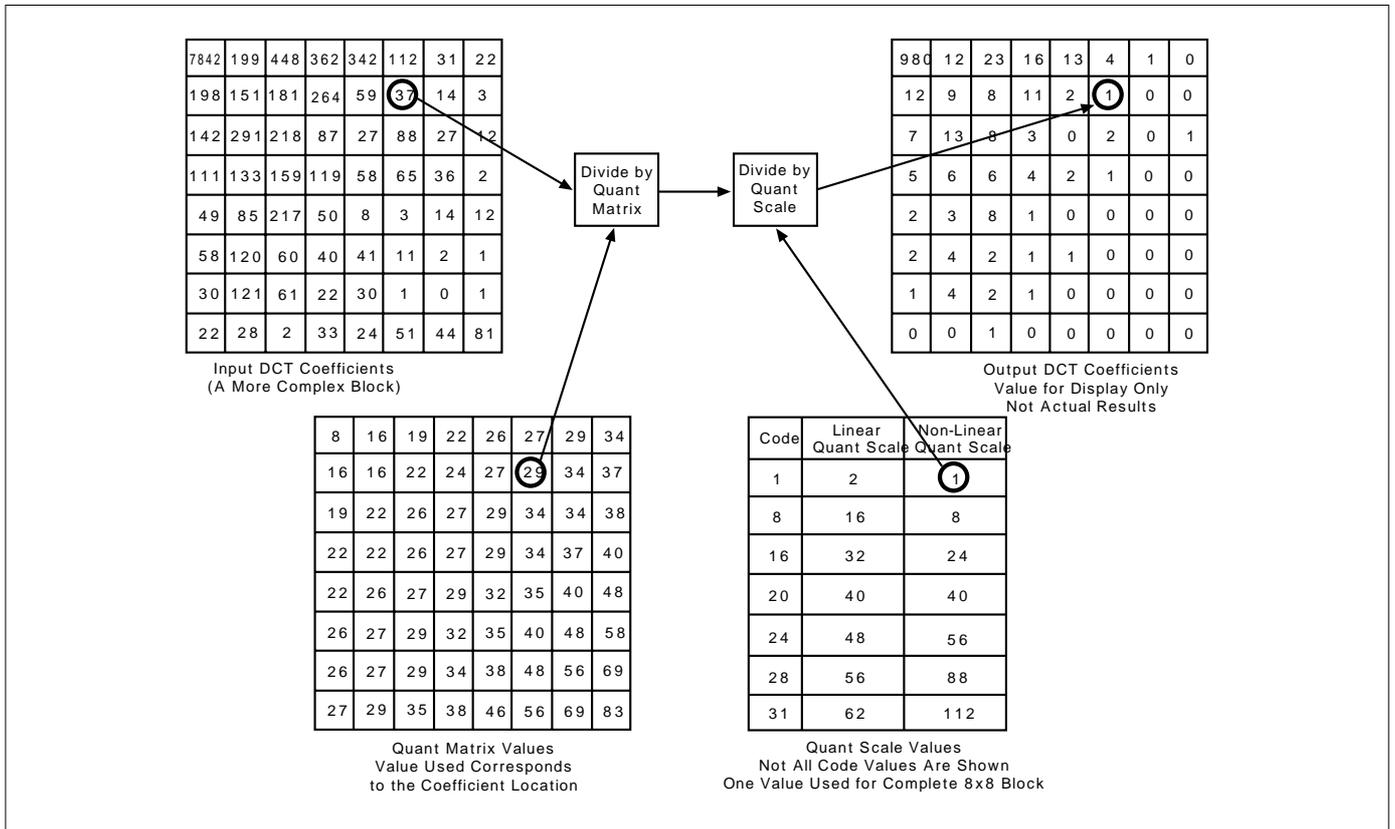


図2-5

得られる圧縮のレベルと、それによって得られるビット・レートが、再量子化の程度によって変わることは明らかです。そしてビット・レートが異なれば、必要なウエイティングテーブルも異なります。MPEGではさまざまなウエイティングテーブルを使用することができ、また、使用中のテーブルをデコーダに送ることで、正確な復号が行われるようにします。

2.4 走査

通常の番組素材では、有効なDCT係数の多くは、マトリクスの左上付近にあります。値の小さい係数はウエイティングで切捨てられてゼロになることがあります。ゼロでない係数をまず送り、続いてそれ以外はすべてゼロであることを示すコードを送れば、伝送の効率は高まります。走査は、こうした結果が得られる可能性を高めるために、有効な係数がある確率の高い部分から順に係数を送る技術です。図2-6aのように、プログレッシブ方式では、係数が大きい値を持っている確率が最も高いのは左上で、最も低いのは右下です。このような場合は、斜め45度のジグザグが最も効果的な走査順序です。

図2-6bは、インタレース方式で使用される交互走査パターンを示しています。インタレース画像の場合、各フィールドのDCT 8×8ブロックは、縦方向の長さが2倍あるため、一定の画像詳細度では、垂直周波数が水平周波数の2倍あるように見えます。したがってインタレース画像では、2倍の角度で走査するのが理想的です。図2-6bでは、垂直空間周波数を走査してから、その部分の水平空間周波数を走査しています。

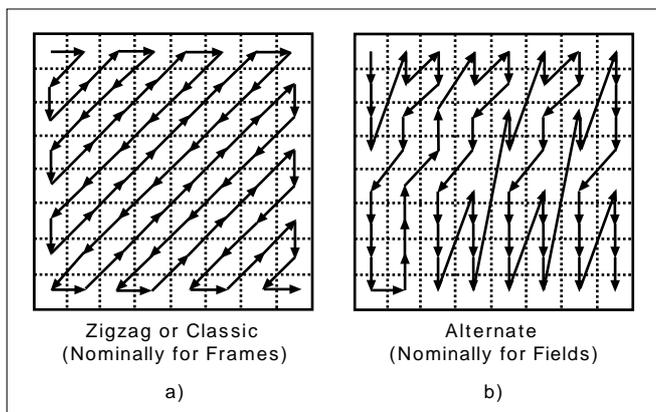


図2-6

2.5 エントロピー符合化

実際の映像では、すべての空間周波数が同時に存在することはないため、DCT係数のマトリクスにはゼロの部分があります。再量子化によって小さい値を切捨てると、ゼロの数は増加します。走査を行っても、有効な値の間にゼロの係数が現れます。RLC (Run Length Coding : ラン・レングス符号化) では、こうした係数をより効率的に処理します。ゼロの連続のように反復的な値が存在する場合、RLCでは、個々のビットではなく、そのゼロの数を送ります。

ある係数が実際の映像に現れる確率を調べることができます。実際に、非常に頻りに現れる値と、稀にしか現れない値があります。この統計情報に基づき、VLC (Variable Length Coding : 可変長符合化) を使用して、さらに圧縮を行うことができます。頻りに現れる値は短いワードに変換し、稀にしか現れない値は長いワードに変換します。復号しやすくするため、ワードを別のワードのプレフィックスにすることはできません。

2.6 空間コーダ

図2-7は、これまで述べてきた空間符合化の概念をすべて1つにまとめています。入力信号には、ワード長が8または10ビットの4:2:2 SDI (Serial Data Interface : シリアル・デジタル・インタフェース) を想定しています。MPEGで扱うのは8ビット分解能だけなので、SDI信号に10ビットのワードが含まれている場合は、丸め込み処理が必要になります。ほとんどのMPEGプロファイルでは4:2:0サンプリングが使用されているため、垂直ロー・パス・フィルタ/補間のプロセスが必要になります。丸め込みとカラー・サブサンプリングによる情報損失と、それに比例したビット・レートの減少は避けられません。ラスト走査入力フォーマットを保存して、8×8画素のブロックに変換できるようにしておく必要があります。

DCTの段階では、画像情報が周波数領域に変換されます。ただし、DCT自体では圧縮は一切行われません。DCTの後、係数をウエイティングして切捨てることで、最初の有効な圧縮が行われます。続いて、係数をジグザグ走査することにより、走査の早い段階で有効な係数が発生する確率を高めます。最後の非ゼロ係数の後は、EOB (End of Block : ブロック終了) コードが生成されます。

係数データは、RLC符号化とVLC符合化によってさらに圧縮されます。可変ビット・レート方式では、量子化が一定の場合もありますが、固定ビット・レート方式では、バッファ・メモリを使用して、符合化における変動を吸収します。非常に詳細な画像で

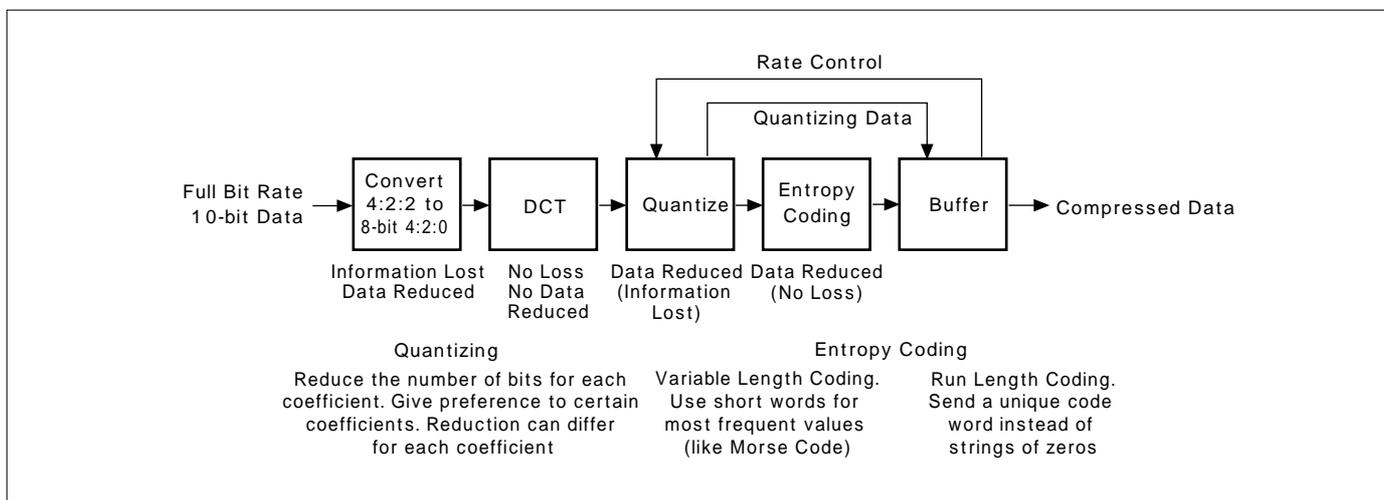


図2-7

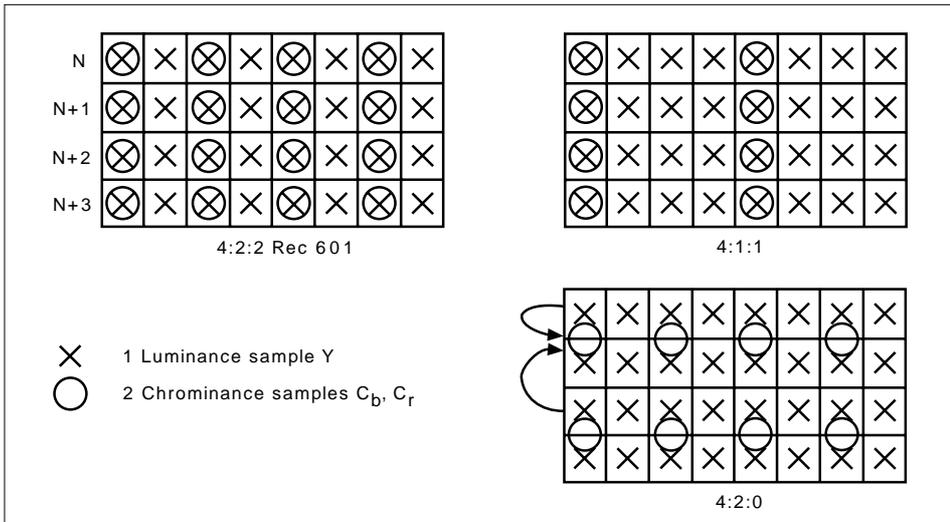


図2-8

はバッファが一杯になりやすく、単調な画像では空のままということもあります。バッファがオーバーフローする可能性がある場合は、再量子化のステップを大きくして、圧縮率を上げる必要があります。

デコーダでは、ビット・ストリームをデシリアライズし、エン트로ピィ符合化を逆に辿ってウエイティング係数を復元します。係数はジグザグ走査に従ってマトリクスに置かれ、逆ウエイティングによって、DCT係数のブロックが再現されます。そして逆変換によって、8×8画素ブロックが再現されます。ラスタ走査の結果を得るため、ブロックはRAMに保存され、1ラインずつ読み取られます。4:2:0データから4:2:2を得るには、図2-8のように垂直補間プロセスが必要になります。

4:2:0のクロミナンス・サンプルは、ルミナンス・サンプルの間に縦方向に置かれるため、インタレース方式を使用する場合は等間隔になります。

2.7 時間符合化

時間的な冗長性は、ピクチャ間の差分のみを符合化または伝送することによって利用できます。図2-9のように、1ピクチャ分の遅延と減算器を組合せると、ピクチャの差分を計算することができます。ピクチャの差分とは画像そのもので、前述のように空間コードでさらに圧縮できます。デコーダは空間符合化を逆にたどり、差分ピクチャを直前のピクチャに加えて次のピクチャを得ます。

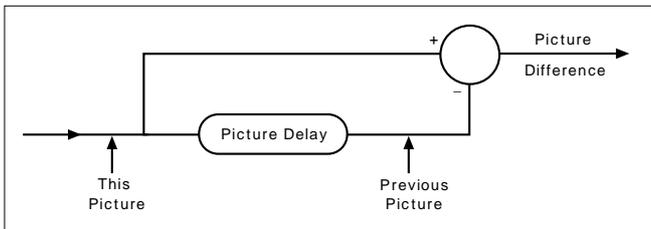


図2-9

この単純な方式にも、欠点はあります。第1に、送られるのが差分だけであるため、伝送を始めた後に復号をはじめることができません。こうした制約があるため、デコーダがビット・ストリームの切替(視聴者がチャンネルを変えたときに発生)に続いて、ピクチャを提供するのは困難です。第2に、差分データの一部でも不正確な場合、このエラーは制限なく伝搬します。

差分だけでない方式を使用することが、こうした問題に対する解決策です。図2-10では、周期的に完結したピクチャが送られています。これはピクチャ内符合化ピクチャ(1ピクチャ)と呼ばれ、空間圧縮のみで得られます。エラーやチャンネル変更があっても、次の1ピクチャで正確な復号を再開することができます。

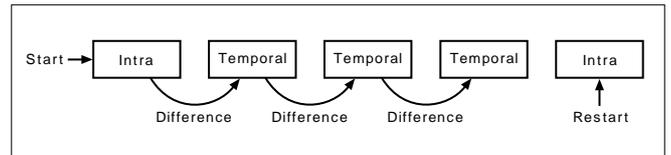


図2-10

2.8 動き補正

動きはピクチャ間の類似性を減少させ、差分ピクチャを作成するのに必要なデータ量を増大させます。この類似性を大きくするために、動き補正を行います。図2-11は、動き補正の原理を示しています。物体が画面上を横切ると、ピクチャごとに場所は異なりますが、外観はそれほど変わりません。そのため、エンコーダで動きを測定することで、ピクチャの差分を削減することができます。この差分は、ベクトルとしてデコーダへ送られます。デコーダでは、このベクトルを使って、直前のピクチャの一部を、新しいピクチャの適切な場所へ移動させます。

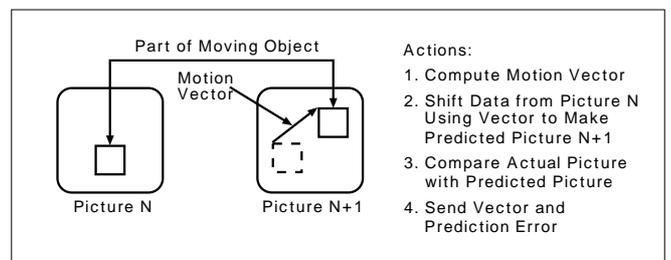


図2-11

1つのベクトルは、ピクチャを構成するマクロ・ブロックと呼ばれる単位の移動を調整します。マクロ・ブロックのサイズは、DCT符合化やクロミナンス・サブサンプリングの構造によって決まります。図2-12aのように、4:2:0方式では、クロミナンス・サンプルの縦・横の間隔は、ルミナンスの間隔のちょうど2倍になります。

第2章 映像の圧縮

8×8DCTクロミナンス・サンプル・ブロック1つは、8×8ルミナンス・ブロック4つと同じ広さです。つまり、これがベクトルによって移動できるピクチャの最低単位です。4:2:0のマクロ・ブロック1つの大きさは、ルミナンス・ブロックであれば4つ、CbとCrであれば1つです。

4:2:2プロファイルでは、クロミナンスは水平軸でのみサブサンプリングされます。図2-12bのように4:2:2では、8×8DCTクロミナンス・サンプル・ブロック1つは、ルミナンス・ブロック2つと同じ広さです。4:2:2のマクロ・ブロック1つの大きさは、ルミナンス・ブロックであれば4つ、CbとCrであれば2つです。

動き推定器は、連続する2つのピクチャからルミナンス・データを比較します。1枚目のピクチャにあるマクロ・ブロックが基準になります。範囲全体にわたって、あらゆるずれを半画素単位で測定することで、この基準と次のピクチャとの相関を測定します。そして、その中で最も大きい相関を、正確な動きとみなします。

動きベクトルには縦と横の成分があります。通常の番組素材では、動く物体が複数のマクロ・ブロックにわたることがあります。それらのベクトルの差分を送るようにすれば、より高い圧縮率が得られます。大きな物体が動くと、隣接するマクロ・ブロックのベクトルは同じになり、ベクトルの差分はゼロになります。

動きベクトルは、実際の物体ではなくマクロ・ブロックに関連付けられており、同一のマクロ・ブロック内で一部が動き、一部が動かないということが起こります。この場合、正しく補正することは不可能です。ベクトルを送って動体の動きを補正すれば、静止部分は不正確に移動し、差分データを修正する必要があります。ベクトルを送らなければ、静止部分は正確ですが、動く部分を修正するために差分データが必要になります。実際の圧縮器では、両方の手段を試して、必要なデータが少ない方を選択します。

2.9 双方向符合化

物体が動くと、前縁から背景が隠れ、後縁から背景が現れます。背景が現れた場合、その部分はそれまで隠されていて、以前のピクチャに情報がいないため、新たなデータを送る必要があります。

カメラをパンする場合も同様で、新たに視界に入ってきた部分に関する情報はありません。MPEGでは、この問題を最小限に抑えるため、双方向符合化によって、現時点のピクチャの前と後のピクチャから情報を得られるようにします。背景が現れつつある場合、その情報は後のピクチャに含まれますが、その情報を時間的に逆方向へ移動させることで、それよりも前のピクチャに加えることができます。

図2-13は、双方向符合化の概念を示しています。双方向符合化されたピクチャは、マクロ・ブロック単位で、動き補正されたデータを前と後のピクチャから得て、さらに、前と後のデータの平均を使用できます。双方向符合化では、予測レベルをできるだけ高めることで、必要な差分データの量を大幅に削減できます。MPEGでは、エンコードの構成については規定せず、適合するビット・ストリームのみを規定していますが、高性能な圧縮器であれば、これまでに述べた3つの符合化手段をすべて試行し、伝送するデータが最も少ないものを選ぶことができます。

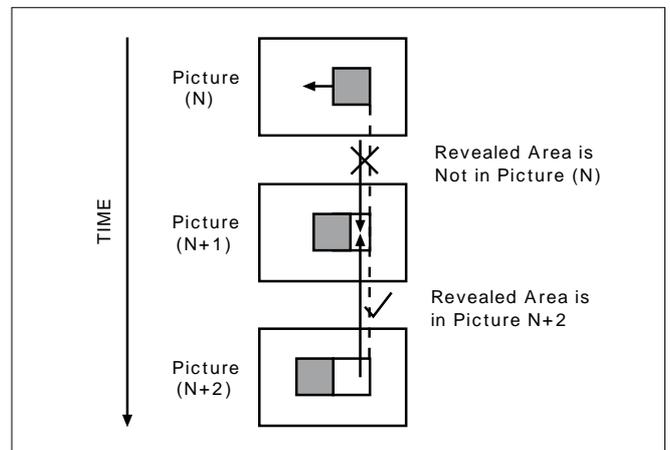


図2-13

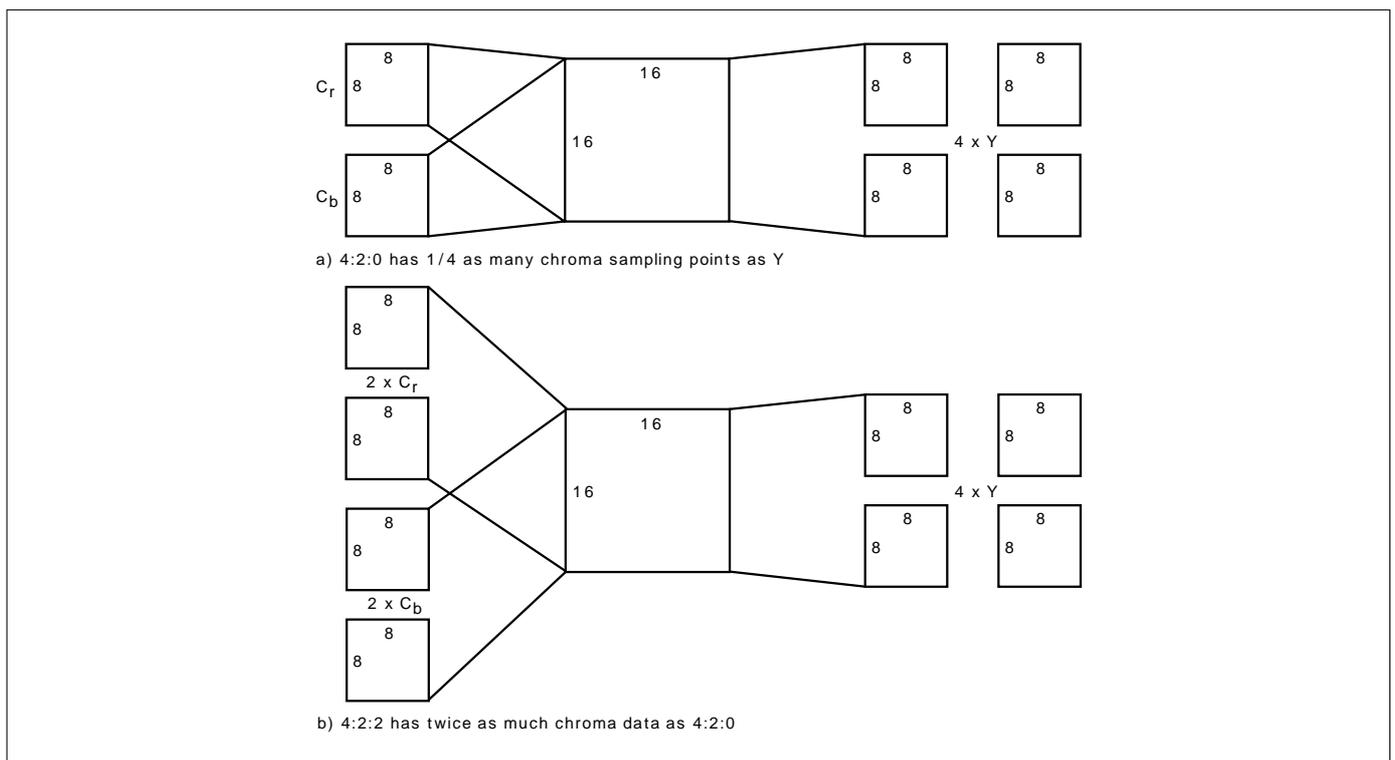


図2-12

2.10 I/P/Bピクチャ

MPEGでは、差分による符合化と双方向符合化に対応し、しかもエラー伝搬を最小限に抑えるには、I、P、Bという3種類のピクチャが必要です。

Iピクチャとは、ピクチャ内符合化ピクチャのことで、復号の際に、それ以上の情報を必要としないピクチャです。他の種類のピクチャと比べてデータ量が多いため、必要以上には送られません。主として変換係数で構成され、ベクトルは含まれていません。Iピクチャは、他のピクチャと無関係に復号されるため、視聴者がチャンネルを変えても、エラー伝搬を抑えられます。

Pピクチャとは、直前のI/Pピクチャから予測符合化を行う、順方向予測符合化ピクチャです。Pピクチャは、直前のピクチャのどこからマクロ・ブロックを取り出すかを表すベクトルと、そのマクロ・ブロックに付加する修正 / 差分データを表す変換係数で構成されています。動き補正で、マクロ・ブロックに対する一致が見つからない場合は、ピクチャ内データでそのマクロ・ブロックを符合化します。Pピクチャの必要データ量は、Iピクチャの約半分です。

Bピクチャとは、直前と直後のI/Pピクチャから予測符合化を行う、双方向予測符合化ピクチャです。Bピクチャは、直前または直後のピクチャのどこからデータを取り出すかを表すベクトルで構成されています。また、必要な修正を行うための、ピクチャ内符合化されたデータも含まれています。この場合も、動き補正において、マクロ・ブロックに対する一致が見つからない場合は、ピクチャ内データでそのマクロ・ブロックを符合化します。双方向予測は非常に有効なので、Bピクチャにあるマクロ・ブロックのほとんどは、動きベクトルによって符合化されます。また、Bピクチャは他のピクチャ符号化の基準として使用されることは決していないため、エラーが伝搬する可能性はありません。そのため、エンコーダで修正データを得る際に、再量子化の程度を上げることが可能です。通常、Bピクチャの必要データ量は、Iピクチャの約4分の1です。Bピクチャは、必ずしも両方向の予測を行う必要はなく、場合によっては、どちらか1方向だけを行うこともあります。クローズドGOP(Group of Picture)を構成する場合は、1方向だけになります。

図2-14は、GOPの概念を示しています。GOPはピクチャI、P、Bの連続構造で表されます。一般には、全体で同じ構造が繰返されますが、GOPの長さや構造はいつでも変更することができます。GOPの長さには正式な制限はありませんが、伝送目的であれば、通常の場合は12または15ピクチャです。

MPEGの時間圧縮の性質上、ピクチャの伝送順は表示順と同じではありません。Pピクチャは基準になったI/Pピクチャの次に来るため、特別な要件はありません。しかし、双方向符合化されるBピクチャは、両方の基準ピクチャを受け取って復号するまで復号できません。図2-14は、上が表示順に並んだGOPのピクチャで、下が伝送順です。伝送順では、Bピクチャは常に、予測の基準に使用するピクチャの次にきます。

GOPには、オープンとクローズドの2種類があります。

クローズドGOPは、それ自体の他に基準を必要としません。表示順はIピクチャで始まり、Pピクチャで終わります。伝送順では通常、最後のPピクチャの次にBピクチャがきますが、表示順では、Bピクチャは最後のPピクチャよりも前にきます。表示順をBピクチャで開始または終了することもできますが、この場合、開始と終了のBピクチャは、それぞれ1方向の予測のみで符合化します。クローズドGOPの開始では、逆方向予測しか使用できません。また、終了はPピクチャと同じで、順方向予測のみです。ただし、Bピクチャのルールは再量子化などにも利用できます。

オープンGOPには、予測ベクトルに関してこのような制約はありません。たとえばGOP最後のBピクチャは、最後のPピクチャからの順方向予測にも、次のGOPにある最初のIピクチャからの逆方向予測も使用できます。こちらの構造の方がやや効率は高いのですが、予測がピクチャの境界を越えてしまう可能性があります。ビット・ストリームのスプライスはクローズドよりもはるかに難しく、チャンネル変更のようなイベントによってエラー発生率も高くなります。

シーンが変わった場合、GOP構造はエンコーダで変更できます。シーン変更前と後のピクチャ間ではエントロピィが大きいので、変更にまたがって予測すると、通常は失敗します。エンコーダでシーン変更を検出する場合は、シーン変更までクローズドGOPを使用し、新しいシーンでは最初のピクチャをIピクチャにして新しいGOP(オープン/クローズド)を開始します。

ピクチャを順不同に送るには、エンコーダとデコーダに余分なメモリが必要で、遅延の原因にもなります。こうしたことが問題になる場合は、I/Pピクチャの間にあるBピクチャの数を制限することで、コストを削減し、遅延も抑えるようにします。

図2-15のように、圧縮率を上げようとする符合化遅延が大きくなるため、そのバランスを判断する必要があります。ある一定の品質において、Iピクチャのみを送るのに必要なビット・レートは、IBBPの順序で送る場合の2倍以上になります。

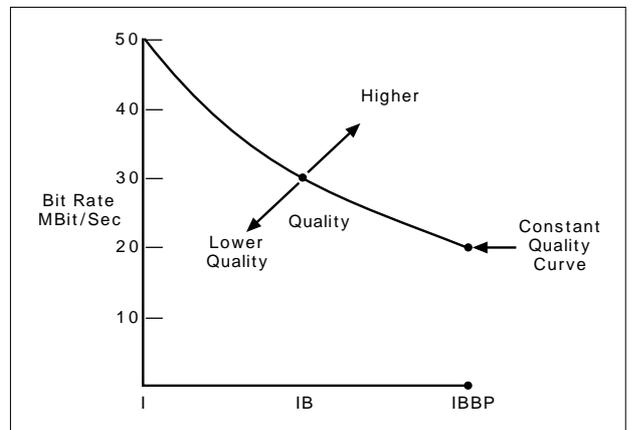


図2-15

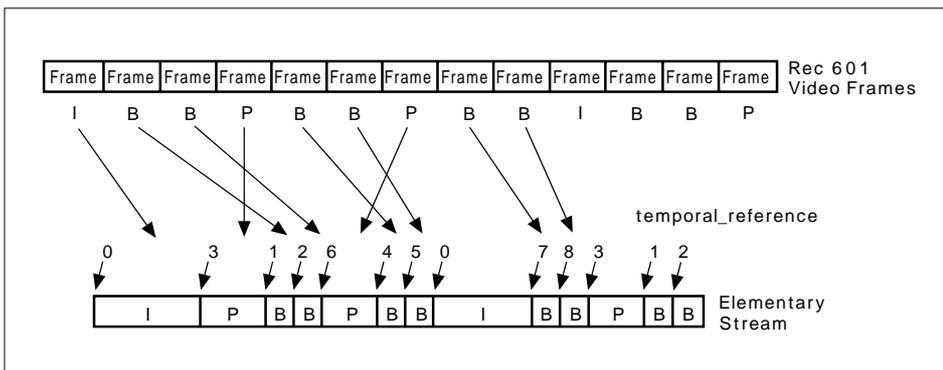


図2-14

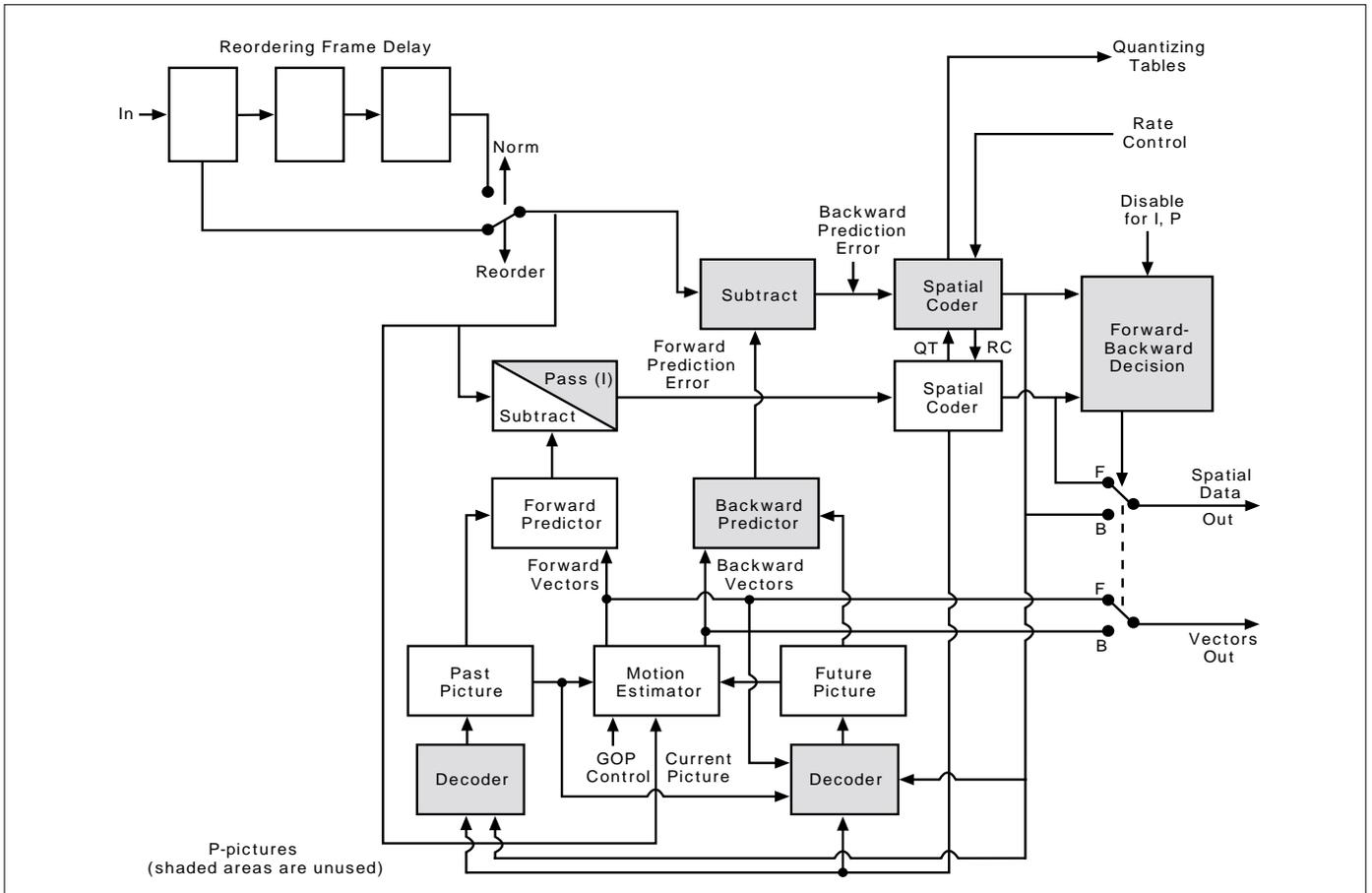


図2-16b

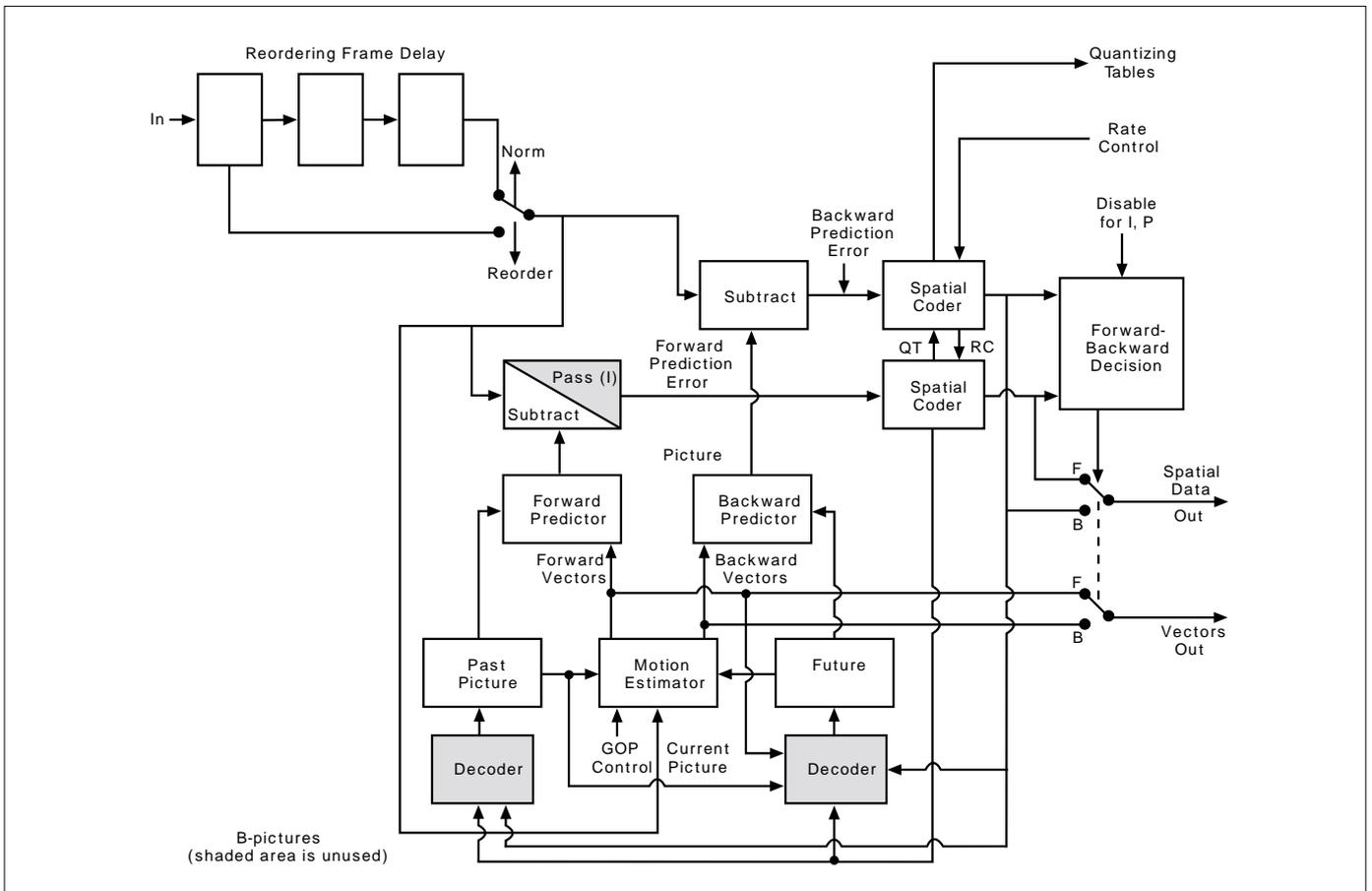


図2-16c

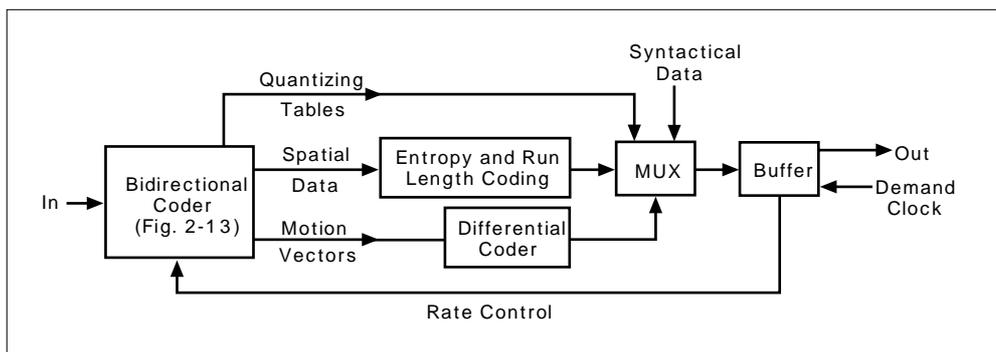


図2-17

図2-17は、MPEGコーダを示しています。動き補正器の出力は空間符合化され、マルチプレクサでベクトルが加えられます。ピクチャの種類(I、P、B)を識別する構文データも付加され、その他にも役立つ情報をデコーダに提供します(「第5章 エレメンタリ・ストリーム」を参照)。ビット・レートの時間的なばらつきに対処するため、出力データはバッファに入られます。平均ビット・レートが非常に高いと、バッファがオーバーフローしやすくなります。オーバーフローを防ぐには、量子化の程度を上げる必要があります。逆に、バッファがアンダフローの状態であれば、量子化を緩和して平均ビット・レートを維持します。

2.12 前処理

圧縮器は、ピクチャ内とピクチャ間の冗長性を除去します。画像データ以外で、冗長性を減少させるものがあるのは好ましくありません。ノイズやフィルムの粒子状ノイズは、画像全体で発生することが多いので特に問題です。ノイズがあると、DCT処理によって、ゼロでない係数が増え、この情報と本当の画像データをコーダが区別できなくなります。すべての係数を符合化するには、量子化の程度を上げる必要があるため、画質は低下します。また、ノイズは連続するピクチャ間の類似性も減少させ、必要な差分データが増加します。

コンポジット・ビデオ信号を復号した映像に、残留サブキャリアがあると、コンポーネント信号の番組では通常レベルが低い空間周波数が高くなるため、影響は深刻になります。また、サブキャリアはピクチャと1つおきに同相になるため、差分データも増大します。コンポジットの復号によってMPEGコーダの入力に現れる信号劣化は、当然デコーダでも復元されやすくなります。

好ましくない動きを発生させるようなことは避けるべきです。カメラの取り付けが不安定だと、画像が揺れるだけでなく、ピクチャの差分や送らなければならないベクトルも増大します。テレシネ素材でも、送り穴の破損が原因でフィルムが縦や横にブレると、同じ問題が起こります。一般に、圧縮用の映像は可能な限り高い品質にする必要があります。高い品質が得られない場合は、ノイズ削減やその他の安定化技術を施します。

圧縮率が高くなると、信号劣化のレベルは上り、入力の品質が良くない場合はさらに深刻です。この場合はプレフィルタリングによって、コーダでのエントロピを減少させます。ビデオ信号には2次元のローパス・フィルタリングを行います。これは必要な係数の数を減らし、信号劣化のレベルを下げます。画像の鮮明度は落ちますが、信号劣化のレベルが高い場合は、鮮明度が低い方が都合良くなります。

MPEG-2アプリケーションの多くは4:2:0サンプリングを採用しており、ソースが4:2:2の場合はクロミナンスのダウンサンプリング処理が必要になります。MPEG-1では、さらにルミナンスとクロミナンスをダウンサンプリングして、横が352画素しかないCIF(Common Image Format)の入力画像を生成します。この技術では、さらに大きい倍率でエントロピが減少します。非常に圧縮率が高い場合は、横が176画素のQCIF(Quarter Common Image Format)の画像を使用します。ダウンサンプリングとは、空間ローパス・フィルタと補間器を組合せて行う処理です。インタレース信号をダウンサンプリングすると、垂直の細部情報が2フィールドに広がり、動きによって相関性がなくなることがあるため、好ましくありません。

ソース素材がテレシネの場合、ビデオ信号は通常と異なる特性を持ちます。50Hz方式のビデオでは、同じフィルム・フレームから1対の2フィールドを得られるので、その間に動きはありません。そのため、フィールド間の動きは、ゼロとフレーム間の動きを交互に繰返します。60Hz方式のビデオでは、24Hzのフィルムから60Hzを得るのに3:2プルダウンを使用します。1枚目のフレームは2フィールドに分割され、次のフレームは3フィールドに分割され、その後その繰返しになります。

そのため、5フィールドのうち1フィールドは完全に冗長になります。MPEGでは、3:2方式の3番目のフィールドを捨てることで、フィルム素材を最適に処理します。24Hzの符号が伝送され、デコーダに対し、フィールド・ストアの読み出しを繰返して、3:2のシーケンスを再現するように通知します。50Hz方式と60Hz方式のテレシネでは、各フィールド対をインタレース解除してフレームを作成し、フレーム間の動きを測定します。デコーダは、フレーム・ストアの走査線を1ラインおきに読み取ることで、インタレースを再現します。

カットがあると、多くの場合は予測がほぼ完全に失敗し、大量の修正データが必要になるため、圧縮器で処理するのは困難です。符合化の遅延が許される場合は、コーダで事前にカットを見つけ、1ピクチャがカットに重なって挿入されるようにGOPの構造をダイナミックに修正できます。この場合、非常にわずかな追加データでカットを処理できます。最後のBピクチャがこの1フレームの前にある場合は、ほぼ間違いなく順方向予測が必要になります。DVDマスタリングのようにリアルタイムでない一部のアプリケーションの場合、入力ビデオを処理するコーダは、困難な(エントロピの高い)部分を突き止めて符合化手段を策定する方法と、そのまま入力ビデオを圧縮するという方法の2つを選べます。

2.13 ウェーブレット

どのような変換方式も、不確定要素の影響を必ず受けます。なぜなら、周波数領域を正確に認識するほど時間領域について不正確になり、逆に、時間領域を正確に認識するほど周波数領域について不正確になるからです。DFT(Discrete Fourier Transform: 離散フーリエ変換)やDCT(Discrete Cosine Transform: 離散コサイン変換)など、多くの変換方式ではブロック長が固定されているため、時間分解能および周波数分解能も固定されます。周波数係数は、線形スケールの均等間隔の値を表しています。しかし、人間の知覚は対数的に変化するため、DFTやDCTの均等目盛では、一方の端では周波数分解能が不十分になり、他方の端では過剰になります。

ウェーブレット変換では、周波数分解能がオクターブの固定分解能で、対数的な特性を持っているため、この問題の影響を受けません。ブロック長を周波数の関数として変えることで、影響を回避します。周波数が下がるにつれて、ブロックは長くなります。つまり、基底関数の周期数がすべて同じで、その周期数を時間軸に沿って調整することによって、さまざまな周波数を探るのが、ウェーブレット変換の特徴です。図2-18は、固定ブロック長のDFT/DCTと、可変長のウェーブレットを対比しています。

ウェーブレット変換は、トランジェントの時間的な位置を正確に検出し、安定したトーンでピッチを正確に判断する、相反する要件に自動的に適応できるため、音声符合化には特に有効です。

映像符合化では、ウェーブレットには、分解能を調整できる信号を特別な手順なしで生成できるという利点があります。動画では、可変長のブロックに動きベクトルを割当てるのが困難なので、この利点は打ち消されてしまいますが、静止画つまりピクチャの符合化では、そのような問題はありせん。ウェーブレット変換は、とりわけビット・レートが非常に低いアプリケーションで強みを発揮します。ウェーブレット係数を過度に量子化することで生じる信号劣化は、通常「しみ」のように現れますが、これはDCT係数を過度に量子化することで生じる「濃淡のムラ」ほどは目立ちません。

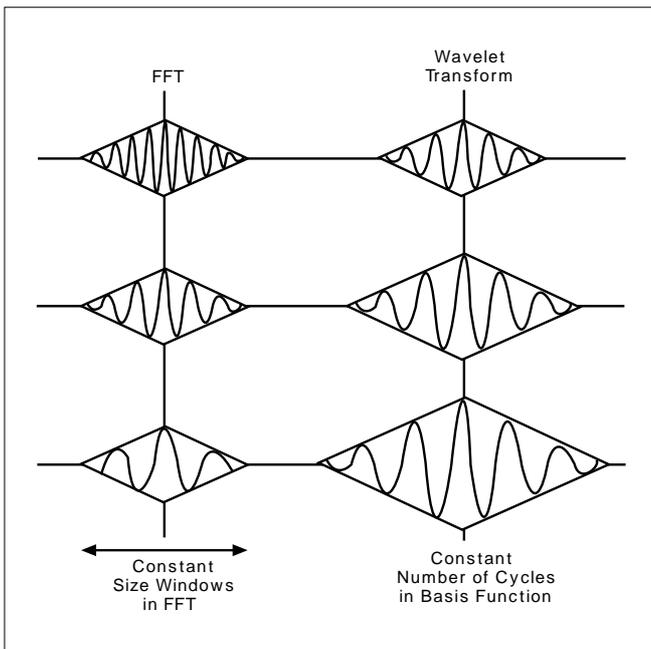


図2-18

第3章 音声の圧縮

非可逆の音声圧縮は聴覚の特性に基づいているので、圧縮について説明する前に、まず聴覚について考える必要があります。意外なようですが、人間の聴覚は視覚よりも識別力が高く、特にステレオ音質に対して優れています。そのため、音声圧縮は慎重に行う必要があります。また、映像圧縮の場合と同様、音声圧縮においても、必要な圧縮率に応じて複雑性のレベルが異なります。

3.1 聴覚のメカニズム

「聞く」とは、耳の物理的プロセスと神経/精神的プロセスが連動して、音声的な印象を受けることです。エントロピーの一部が失われるため、この印象は、外耳道における実際の音響波形と同じではありません。音声圧縮によって失われる部分が、聴覚メカニズムで失われるエントロピーと同じ部分になるようにすれば、よい結果が得られます。

聴覚の物理的メカニズムは外耳、中耳、内耳で構成されています。外耳には外耳道と鼓膜(ear drum)があります。鼓膜は、ちょうどマイクの振動板のように、入射音を振動に変換します。この振動は体液によって伝わり、内耳で感知されます。液体は気体よりもはるかにインピーダンスが高く、中耳は、電力伝達を向上させるインピーダンス整合トランスのような働きをします。

図3-1のように、振動は、前庭窓にはまり込んでいるアブミ骨(stirrup bone)によって内耳に伝わります。そして、蝸牛(cochlea)という頭蓋骨の螺旋状の空洞(図3-1では説明のため伸ばして表示)に達します。蝸牛は基底膜(basilar membrane)によって仕切られています。基底膜は先端に行くにつれて質量や硬さが変わります。前庭窓側の端では硬くて軽いため、共振周波数は高くなります。反対側の先端は重くて柔らかいため、低い周波数で共振します。この共振周波数の範囲によって人間の可聴周波数の範囲が決まり、通常は20Hz~15kHz前後です。

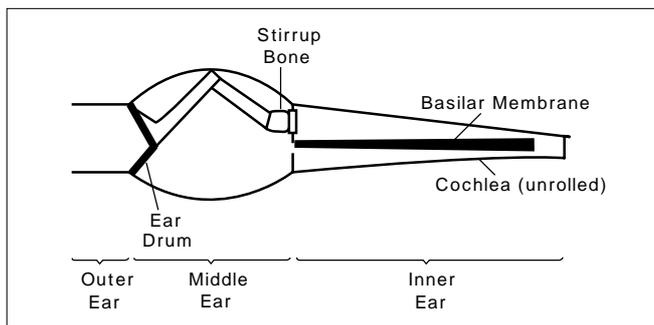


図3-1

入ってくる音の周波数によって、基底膜の振動する部分が異なります。そして、各部が別々の神経の末端につながっているため、音の高さが識別できます。また、神経の働きにより、基底膜の微細な筋肉が調和して正帰還システムのように機能することで、共振のQ値を高めます。

基底膜の共振動作は、変換アナライザの動作と似ています。変換の不確定性理論によると、信号の周波数領域の認識が正確になるほど、時間領域の認識は不正確になります。つまり、2つの周波数を識別する能力が高まるほど、2つのイベントの時間を識別する能力は低下します。人間の聴覚は、時間の識別と周波数の識別のバ

ランスを取るという一種の妥協をしながら発達してきました。バランスは取れていますが、どちらの能力も完璧ではありません。

このように聴覚は、周波数を完璧に識別できるわけではないので、間隔の狭い周波数を聞き分けられません。これは聴覚マスキングとって、「別の音が存在すると音に対する感度が低下すること」と定義されています。

図3-2aのように、聴力のスレッシュホールドは周波数によって変化します。最も感度が高いのは、やはり会話音域です。1種類の音が存在すると、スレッシュホールドは図3-2bのように変化します。音が存在する場合、スレッシュホールドは高い周波数で上り、低い周波数でもある程度まで上ります。音楽のように複雑な入力スペクトラムが存在する場合、スレッシュホールドは、ほぼすべての周波数域で上ります。アナログ・オーディオ・カセットで音楽を聴いているときに、静かな楽節になるとヒス・ノイズが聞こえるのも、こうした特性に起因しています。圧伸ではこの原理を利用して、低レベルのオーディオ信号を増幅してから録音/伝送し、後で正確なレベルに戻します。

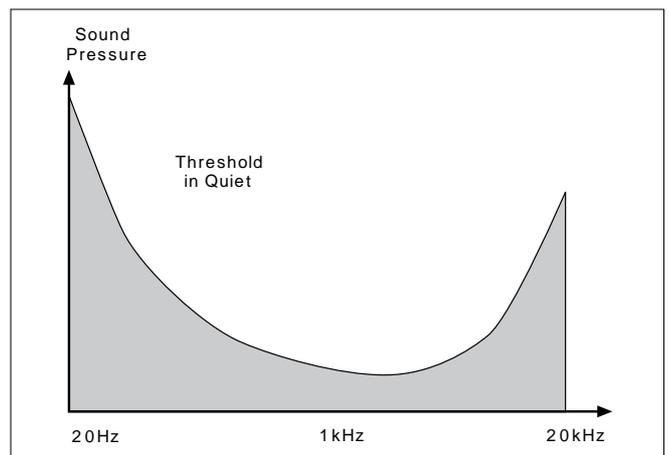


図3-2a

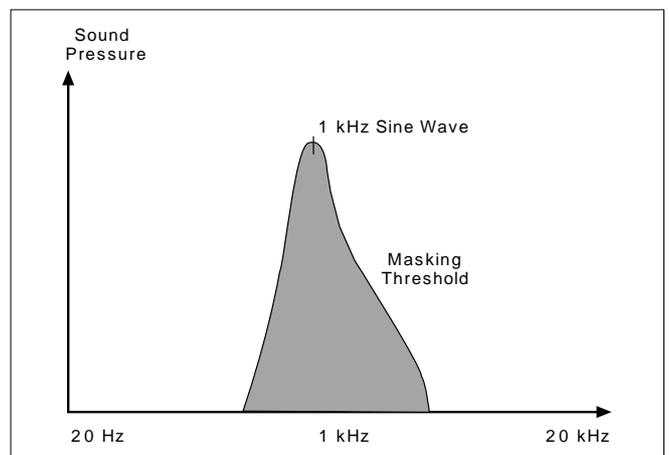


図3-2b

耳が時間を完璧に識別できないのは、共振応答が原因です。Q値によって、音は1ms以上存在しないと聞き取ることができません。このような応答の遅れがあるため、2つの信号が同時でもマスキングは生じます。実際のマスキング期間の前後に、より低いレベルで音のマスキングが続くと、その前後でもマスキングが生じます。図3-3は、この概念を示しています。

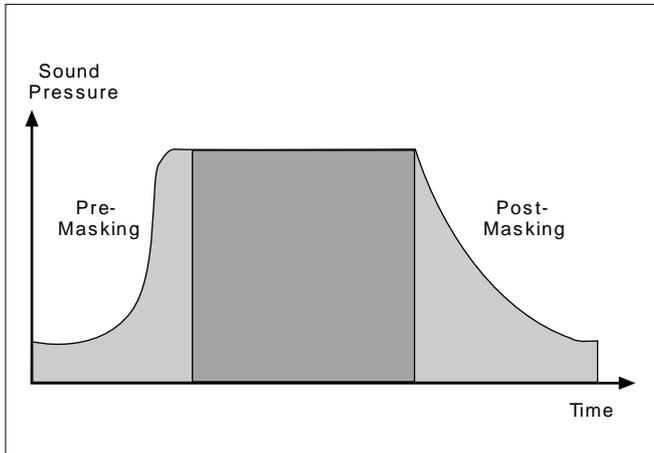


図3-3

マスキングは聴力のスレッシュホールドを上げ、圧縮器はこの効果を利用してノイズ・フロアを上げることで、オーディオ波形を表すのに必要なビット数を減らせます。ノイズ・フロアを上げられるのは、マスキング効果のある周波数のみです。マスキング効果を最大限に引き出すには、音声スペクトラムを複数の帯域(サブバンド)に分割し、帯域ごとに圧伸やノイズの量を変えられるようにする必要があります。

3.2 サブバンド符合化

図3-4は、帯域分割圧伸器を示しています。帯域分割フィルタは、同じ周波数帯域の中で重なり合う狭帯域・線形位相フィルタのセットで、各帯域からは波形を表すサンプルが出力されます。各帯域では、音声入力を最大レベルまで増幅してから伝送し、後で正確なレベルに戻します。帯域ごとに、伝送時に取り込まれたノイズが削減されます。ノイズ削減を聴力のスレッシュホールドと比べると、マスキング効果によって一部の帯域ではより大きなノイズが許容されることがわかります。そのため、各帯域で圧伸後にサンプルのワード長を削減できます。この方法により、分解能の低下で生じた量子化ノイズがマスキングされるので、圧縮が行えます。

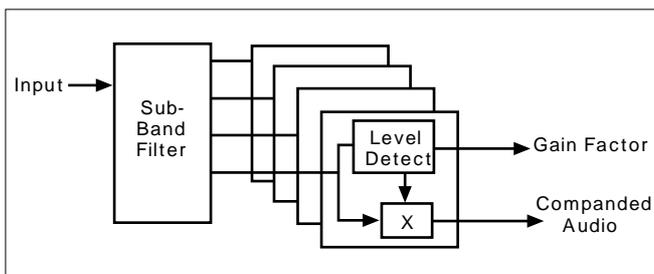


図3-4

図3-5は、MPEGレイヤ1で使用されるシンプルな帯域分割コーダです。デジタル・オーディオが入力されると、帯域分割フィルタはその信号のスペクトラムを複数の帯域に分割します。MPEGでは、この数は32です。時間軸は、長さの等しいブロックに分割されます。MPEGレイヤ1の入力サンプルは384なので、フィルタ出力では、32の各帯域のサンプル数は12になります。各帯域内では、レベルを最大限まで通倍することによって増幅します。必要なゲインは1ブロック期間中一定で、帯域ごとに各ブロックと一緒に送るスケール・ファクタを1種類にすることで、デコーダがプロセスを逆にたどれるようにします。

MPEGレイヤ1のフィルタ・バンク出力は、512ポイントのFFT (Fast Fourier Transform : 高速フーリエ変換)で解析され、入力信号のスペクトラムを判断します。この解析は、各帯域で予想されるマスキングの程度を判断するマスキング・モデルを駆動します。マスキングの程度が大きいくほど、サンプル精度の要件は低くなります。サンプル精度は再量子化でワード長を削減することで下ります。同じ帯域内では、すべてのワードで削減量が一定ですが、ワード長は帯域ごとに異なることがあります。各帯域のワード長をビット割当コードとして送ることで、デコーダがビット・ストリームを正しくデシリアライズできるようにする必要があります。

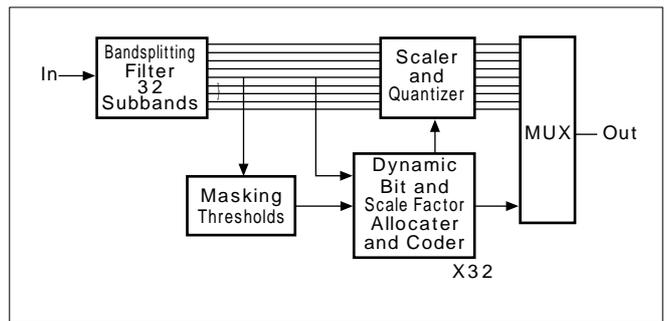


図3-5

3.3 MPEGレイヤ1

図3-6は、MPEGレイヤ1のオーディオ・ビット・ストリームを示しています。同期パターンとヘッダに続いて、4ビットずつの割当コードが32ビットあります。これらのコードには、各サブバンドのサンプルのワード長が記述されています。次に、各帯域の圧伸に使用する32個のスケール・ファクタがあります。このスケール・ファクタは、音声入力を正確なレベルに戻すためにデコーダに必要なゲインを決めます。そして、スケール・ファクタの次には、各帯域の音声データがあります。

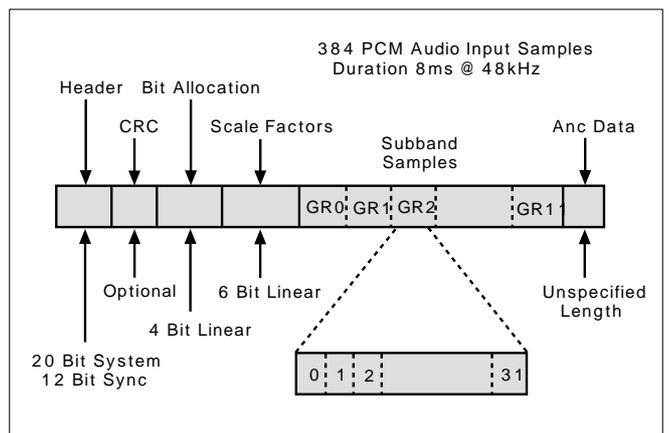


図3-6

図3-7は、レイヤ1のデコーダを示しています。タイミング・ゼネレータは同期パターンを検出し、ビット割当データとスケール・ファクタ・データをデシリアライズします。ビット割当データによって、可変長サンプルをデシリアライズすることができます。また、スケール・ファクタ・データによって、再量子化と圧縮を逆にたどり、各帯域を元の正確なレベルに戻します。そして、これら32の帯域を合成フィルタで1つにまとめて、音声として出力します。

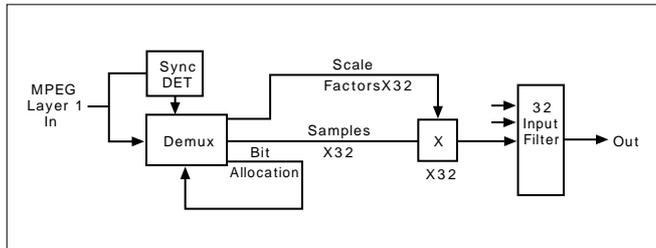


図3-7

3.4 MPEGレイヤ2

図3-8のように、帯域分割フィルタを使用してマスキング・モデルを駆動するとスペクトラム解析はあまり正確に行われませんが、これは、帯域が32しかなく、帯域内のどこにもエネルギーが存在する可能性があるためです。最悪のケースではマスキングできない場合もあるため、ノイズ・フロアはあまり上げられません。より正確にスペクトラム解析すれば、圧縮率を上げることができます。MPEGレイヤ2では、スペクトラム解析は個別のプロセスとして実施します。1024ポイントのFFTで入力から直接マスキング・モデルを駆動します。周波数をより正確に解析するには、変換の時間範囲を長くする必要があるので、ブロック・サイズを1152サンプルに増大しています。

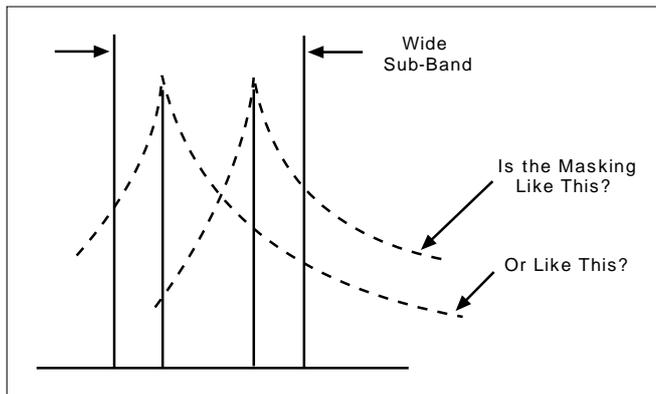


図3-8

ブロック圧伸方式はレイヤ1と同じですが、実際の番組素材では、ある程度の冗長性が含まれているため、すべてのスケール・ファクタが伝送されるわけではありません。同じ帯域内の連続するブロックでは、スケール・ファクタは2dB(時間の10%未満)ずつ異なり、連続するスケール・ファクタを3つずつまとめて解析することで、この特性を活かすことができます。サブバンドの内容がほぼ一定(静止画番組や準静止画番組)していれば、3つのスケール・ファクタのうち1つだけが送られます。過渡的なコンテンツが増加するにつれ、送られるスケール・ファクタも2つ、3つと増えます。また、スケール・ファクタ選択コードも送られるため、サブバンドごとに何が送られてきたかをデコーダが判断することができます。この方法により、スケール・ファクタのビット・レートは実質的に半減します。

3.5 変換符号化

レイヤ1と2は、信号を波形として扱う帯域分割フィルタをベースにしています。しかし、レイヤ3では、映像符号化で使用されるような変換符号化を採用しています。前述のように、人間の耳は入射音に対して周波数変換のような働きをし、また基底膜のQ値によって、応答速度は急激に上下することはありません。したがって、オーディオ波形を周波数領域に変換する場合、それほど頻りに係数を送る必要はありません。この原理が変換符号化の基礎になっています。高い圧縮率を得ようとして係数を再量子化すると、精度が落ちます。このプロセスによって発生するノイズは、マスキング効果が最も大きい周波数になります。変換符号化による副産物として、入力スペクトラムが正確にわかるため、厳密なマスキング・モデルを作成できます。

3.6 MPEGレイヤ3

このような複雑な符号化は、最も高い圧縮率が必要な場合にのみ使用します。レイヤ3は、レイヤ2とある程度共通性があり、ブロックごとに384の出力係数があるDCTが使用されます。この出力は入力サンプルを直接処理することで得られますが、マルチレベル・コーダでは、レイヤ1とレイヤ2における32帯域のフィルタリングをベースにした、ハイブリッド変換を使用することもできます。このようにした場合、QMF(Quadrature Mirror Filter: 直交ミラー・フィルタ)を通った32個のサブバンドは、さらに12帯域のMDCT(Modified Discreet Cosine Transform: 修正離散コサイン変換)で処理され、384の出力係数が得られます。

トランジェントにおけるプリエコー発生を防止するため、2種類のウィンドウ・サイズが使用されます。ウィンドウの切替は、心理聴覚モデルで行います。プリエコーは、平均値を上回る音声に含まれるエントロピイに関連していることがわかっています。最も高い圧縮率を得るため、係数の不均一量子化と、ハフマン符号化を組合せて使用します。この方法では、最も多く出現するコード値に、最も短いワード長を割当てます。

3.7 MPEG-2オーディオ

当初、MPEGオーディオのレベル1、2、3のように呼ばれていた方式は、現在ではMPEG-1レベル1などのように分類されています。MPEG-2は、MPEG-1オーディオの拡張機能とともに、さらに高度な符号化方式について定義しています。

MPEG-2では、MPEG-1よりも低いサンプリング・レートを使用できます。これは厳密には下位互換性はありませんが、MPEG-1デコーダでも、追加のテーブルだけで相互運用性が得られるようになっています。

MPEG-2 BC(Backward Compatible: 下位互換性)オーディオは、5.1チャンネル(5つの独立したチャンネルと、低域の低周波効果チャンネル1つ)について規定しています。MPEG-2 BCは、MPEG-1(2チャンネル)のビット・ストリームをコアにして、マルチチャンネル拡張機能を、MPEG-1デコーダが無視できるように追加しています。

MPEG-2 AAC(Advanced Audio Coding: 拡張オーディオ符号化)は、高解像度フィルタ・バンクと、追加の符号化ツールを備えた、さらに高度な方式です。符号化の効率は大幅に向上しますが、下位互換性はありません。

3.8 MPEG-4オーディオ

MPEG-4符合化は、オブジェクトに基づいています(4.4.2を参照)。MPEG-4オーディオ・オブジェクトは、自然音声から合成音声まで表すことができます。自然音声の符号化については、MPEG-4ツール・キットとして、MPEG-2 AACの他にもさまざまなツールが含まれています。その中には、非常に低いビット・レート用のパラメトリック符合化や、CELP(Code Excited Linear Predictive : 符号励振型線形予測)と呼ばれる、中間ビット・レートの音声符合化も含まれています。また、伝送システムに適用できるビット・ストリーム・スケラビリティなど、さまざまな形のスケラビリティに対応しています。

オブジェクト符合化を使用すれば、復号ポイントで選択肢が得られます。たとえば、協奏曲は、オーケストラとソロという2つのオブジェクトとして送ることができます。通常の復号であれば、ひとまとまりの作品として表現されますが、この場合、オーケストラのオブジェクトのみを復号し、ソロの部分を「生」で演奏することも可能です。また、同様のアプローチで番組を符合化することで、視聴者が「ミックス・マイナス」モードを選んで、スポーツ番組などの実況を消すこともできます。

今後、MPEG-4の合成音声の機能は確実に普及します。合成音声の主な機能には、Text-to-Speechや、SAOL(Structured Audio Orchestra Language : 構造化オーディオ・オーケストラ言語)を使ってダウンロードした楽器を音楽に合成できる「スコア・ドリブン」技術などがあります。

3.9 AC-3

Dolby Laboratories(ドルビー研究所)が開発したAC-3音声符合化技術は、MPEGの音声符号化方式の代わりにATSCシステムと組合せて使用されます。AC-3はATSC A/52として規格化され、その後も、オプションとしてDVBに採用されたり、Motorola Digicypher (モトローラ・デジサイファ)システムに採用されたりしています。AC-3は、周波数係数を再量子化することで符合化を得る、変換ベースのシステムです。

図3-9のように、AC-3コーダへのPCM入力は、ウィンドウが重なり合うブロックに分割されます。これらのブロックには512サンプルずつ含まれていますが、完全に重なり合っているため、冗長性は100%です。変換後も各ブロックに512の係数がありますが、冗長性があるため、TDAC(Time Domain Aliasing Cancellation : 時間領域エイリアス・キャンセル)という技術を用いて、256個まで大幅に削減することができます。

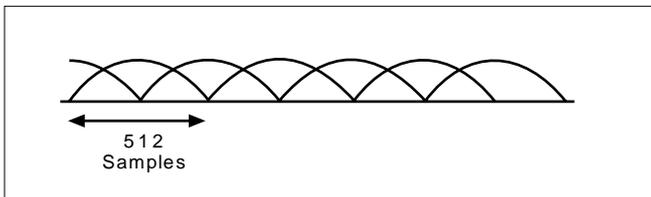


図3-9

入力波形を解析し、ブロックの後半部分に相当なトランジェントがある場合、波形を2つに分割してプリエコー発生を防止します。この場合、係数の数は変わりませんが、周波数分解能は半分になり、時間分解能は2倍になります。そして、この処理が行われたことを示すフラグがビット・ストリームに設定されます。

係数は、仮数と指数からなる浮動小数点で出力されます。これは科学的表記を2進数にしたものです。指数は実質的にスケール・ファクタになります。ブロックにある一連の指数によって、スペクトラム包絡線と呼ばれる対数目盛の精度まで、入力スペクトラム解析を行います。このスペクトラム解析はマスキング・モデルに入力され、各周波数においてどの程度までノイズを上げられるかを判断します。

マスキング・モデルによって再量子化プロセスが駆動され、仮数部が丸め込まれることで係数の精度は低下します。伝送されるデータのかなりの部分が仮数の値で構成されています。

また、指数も送られますが、利用できる冗長性がまだ含まれているため、直接送られるわけではありません。ブロック内で完全な形で送られるのは1番目(最も低い周波数)の指数のみです。それ以外の指数は差分が送られ、デコーダによって直前の値に付加されます。入力音声のスペクトラムがスムーズな場合、複数の周波数帯で指数が同じになっていると考えられます。指数は、2つ一組または4つ一組にまとめることができ、まとめたことを示すフラグが付けられます。

ブロックが6つ一組でAC-3同期フレームになります。フレームの最初のブロックは常に完全な指数データを持っていますが、定常信号の場合は、それ以降のブロックも同じ指数を持っている場合があります。

第4章 MPEGの規格

第2章および第3章では、映像・音声の圧縮技術や、MPEG規格で使用されるさまざまなツールについて紹介しました。本章では、MPEGの歴史や構成とともに、さまざまなMPEG規格の発展について説明します。

4.1 MPEGとは

MPEGとはMoving Pictures Experts Groupの頭文字で、国際標準化機構(ISO)と国際電気標準会議(IEC)が共同で組織した委員会です。IECは電気・電子技術の国際的な標準化を扱い、それ以外のほとんどすべてをISOが扱っています。情報技術(IT)時代の初期、ISOとIECはITの問題に対処するため、合同専門委員会(JTC1)を設置しました。JTC1にはいくつかの作業グループがあり、たとえば、JPEG(Joint Photographic Experts Group)やWG11などがあります。このWG11がMPEGです。

本委員会は、イタリアのLeonardo Chiariglione氏を委員長として1988年に設置されました。通常、年4回開催されるMPEG会議に出席する代表団の数も、1988年には15でしたが、2002年には300前後にまで増えました。これまで、MPEG-1、MPEG-2、MP3オーディオ圧縮規格(MPEG-1オーディオのレイヤ3)が広く採用されるなど、規格の策定に目覚ましい実績をあげています。この評判を落とすきっかけになったのがMPEG-4ですが、これは規格の欠陥ではなく、ライセンス条件の発表が大幅に遅れたことが原因で、2002年初頭によく発表された最初のライセンス条件に対しては強い反発もありました。

ただし、MPEGそのものはライセンス供与の問題には関係していません。ISO/IECの組織下にある本委員会は、規格に含まれる技術が「妥当かつ無差別的な条件」においてライセンス供与されるよう要求していますが、この「妥当」についての一般的な定義はありません。ライセンス供与は、関連特許の保有者の責任において行われ、これには、研究開発に貢献し、何らかの見返りを求めている世界中のさまざまな組織が該当します。

MPEG-2では、特許保有者が集まってMPEG-LA(MPEG Licensing Authority)を結成しました。重要な特許はすべてこのグループが認定し、規格の採用を希望する組織に対してブロック単位でライセンス供与します。この方法はMPEG-2ではうまく行っていました。MPEG-4では前述の問題に直面していました。

4.2 MPEG-1

MPEG-1(ISO/IEC 11172)は、動画圧縮用の初めての国際規格として1988年～1992年に開発されました。この方式では、JPEGと同様、DCT変換、係数量子化および可変長符号化が採用され、さらに、時間圧縮のための動き補正も取り入れています。

MPEG-1は次の3部で構成されています。

- ・システムISO/IEC 11172-1：多重化構成
- ・ISO/IEC 11172-2：映像符号化
- ・ISO/IEC 11172-3：音声符号化

MPEG-1は目覚ましい技術的成果を上げました。この規格は、SIFサイズの画像ストリーム(PALでは352×288(25Hz)、NTSCでは352×240(30Hz))および対応する音声を圧縮して、全体で1.5Mbps程度のデータ・レートで送ることを意図しています。この速度はT1データ回線での転送や、CD-ROMの再生に適しており、一般向けビデオ・レコーダの解像度にもほぼ対応しています。オーディオCDに関する数値を比較することで、どの程度の技術的成果が得られるのかを推定できます。通常の2チャンネル・オーディオCDは、解像度16ビット、サンプリング・レート44.1kHzで、データ転送速度は1.5Mbps程度です。MPEG-1では、映像と音声を圧縮し、両方を合せて1.5Mbps以内で送ることが可能になりました。

CIFフォーマットは、欧州と米国のSIF(Source Input Format)フォーマットを折衷させたもので、空間解像度は625SIF(352×288)、時間解像度は525SIF(29.97Hz)です。CIFはテレビ会議システムのベースになっています。

MPEG-1はCIF画像に合せて設計されており、インタレース画像の処理にツールを必要としないため、テレビ放送の分野で使用しても、目立った影響はほとんどありません。

最後に、このMPEG-1規格の内容と、得られる相互運用性について説明します。この規格は、ツール・セット、ビット・ストリームの構文、およびデコーダの操作について定義しています。エンコーダの操作については定義されておらず、適合したデコーダで復号できる構文的に有効なビット・ストリームを生成するデバイスは、すべて有効なMPEGエンコーダということになります。また、画像や符号化の品質についても定義していません。そのため、規格に変更を加えなくても符号化技術を発展させることができ、使用しているデコーダを新しくする必要もありません。この戦略はすべてのMPEG規格で採用されています。現在のDVB伝送方式において、映像にはMPEG-2が使用されていても、MPEG-1レイヤ2オーディオが引き続き主要な音声圧縮方式として使用されていることを見ても、この戦略が成功であることは明らかです。

4.3 MPEG-2

MPEG-1は1991年に凍結されました(以降は編集上の変更のみ可)。その同じ年にMPEG-2のプロセスが開始し、1994年に規格化されました。当初の目標は、放送品質の映像の幅に対応するという簡単なもので、そのためには「フル・サイズ」の標準解像度画質(704×480(29.97Hz)、704×576(25Hz))の符合化と、インタレース映像の効率的な符合化を実現する必要がありました。

MPEG-2は、MPEGがさまざまな意味で「一人前」の規格になったことを示しています。柔軟性を高めたMPEG-2と、LSIの普及によって、MPEG-2を数多くのアプリケーションに使用できるようになりました。MPEG-2の成功を最もよく表しているのは、高解像度テレビ(HDTV)向けに開発されていたMPEG-3が廃止されたことです。MPEG-2がHDTVにも容易に対応できることが明らかになって間もなく、MPEG-3の開発は打ち切られました。MPEG-2は、ATSCとDVB放送規格のベースになっているだけでなく、DVDの圧縮方式にも使用されています。

MPEG-2では、対象を移行することも可能になりました。プロファイルとレベル(後述)を使用することにより、あるアプリケーションに対して規格を完成させ、その後、より要求の高いアプリケーションへ発展的に移行させることができます。MPEG-2を拡張する作業は2002年まで続きます。

MPEG-2はISO/IEC 13818として規格化され、現時点で10のパートがあります。重要なのは以下のパートです。

- ISO/IEC 13818-1: システム(トランスポート/プログラム・ストリーム)、PES、T-STDバッファ・モデル、基本的なPSIテーブル-CAT、PAT、PMT、NIT
- ISO/IEC 13818-2: 映像符合化
- ISO/IEC 13818-3: 音声符合化
- ISO/IEC 13818-4: テストおよび適合性
- ISO/IEC 13818-6: データ放送およびDSMCC

MPEG-2の主な功績の1つ、13818-1のトランスポート・ストリームについては、その柔軟かつ堅牢な設計により、MPEG-4およびMPEG-7のデータ転送を含め、さまざまなアプリケーションでの使用を可能にしました。

注: DVBおよびATSCのトランスポート・ストリームは、ビデオ/オーディオPESを「プログラム(番組)」として運びますが、これはDVDやCDに使用される「プログラム・ストリーム」とはまったく異なります。

通常、MPEGトランスポート・ストリームのビット・レートは一定ですが、プログラム・ストリームは可変ビット・レートです。

4.3.1 MPEG-2のプロファイルとレベル

例外はありますが、MPEG-1は固定サイズの画像と対応する音声を、従来使用されてきた1.5Mbpsのビット・レートになるように符合化するという1つの課題に合せて設計されています。MPEG-1のツールや構文はそれ以外の目的にも使用可能であり、実際それ以外の目的に使用されてきたのも確かですが、そうした使用は規格外であり、それぞれ独自のエンコーダやデコーダが必要になります。ただし、MPEG-1規格に適合したデコーダは1種類しかありません。

MPEG-2にも当初は同様の目標があり、公称525/60および625/50のインタレース・テレビ方式における放送画像と音声の符号化を意図していました。しかし、設計作業が進むにつれ、開発中のツールでさまざまな画像サイズや広範囲のビット・レートに対処できることが明らかになりました。また、スケーラブルな符合化方式のために、より複雑なツールも開発されました。こうし

て、MPEG-2デコーダを1種類に限定することが現実的にも難しくなりました。さまざまなツールを使って符合化された高速ビット・ストリームを、1種類の適合デコーダで処理できるようにしなければならぬとすると、中心的なアプリケーションにとって経済的なデコーダとはいえなくなります。たとえば、20Mbpsの高解像度信号を復号できるデコーダは、5Mbps前後の標準加須ウド信号の復号に限定したデコーダよりもはるかに高価になります。単純なアプリケーションに高価な装置の使用が要求されるようでは、優れた規格とはいえません。

MPEGは、ビット・ストリームやデコーダを分類するために、プロファイルとレベルという2次元的な体系を考案しました。プロファイルは、使用可能なツールについて定義します。たとえば、双方向符合化(Bフレーム)はMainプロファイルでは使用できませんが、Simpleプロファイルでは使用できません。レベルは、規模について定義します。Highレベルのデコーダは高速ビット・ストリームを受信するため、デコーダ・バッファやフレーム・ストアをMainレベルのデコーダよりも大きくしなければなりません。しかし、MP@HL(MainプロファイルのHighレベル)と、MP@ML(MainプロファイルのMainレベル)では、使用する符号/復号ツールおよび構文エレメントは、まったく同じになります。

図4-1は、MPEG-2で定義されているプロファイルとレベルの組合せを示しています(横軸がプロファイル、縦軸がレベル)。このように、すべての組合せが有効というわけではなく、規格で定義されているのは完全な組合せのみです。これは、あるプロファイル/レベルのデコーダは、それよりも低いプロファイルやレベルを復号できなければならないという適合の要件によるものです。たとえば、MP@MLのデコーダは、MP@LL(MainプロファイルのLowレベル)やSP@ML(SimpleプロファイルのMainレベル)のビット・ストリームを復号できるものでなければなりません。

Simpleプロファイルは双方向符合化に対応しておらず、出力されるのはピクチャとPピクチャのみです。これにより符号化・復号化の遅延が低減されるため、ハードウェアを簡素化することができます。Simpleプロファイルは、Mainレベルとの組合せのみが定義されています。

Mainプロファイルは、アプリケーションの大部分で使用できるように設計されています。Lowレベルでは水平方向が352画素程度の低解像度の入力を使用されます。放送アプリケーションの大半では、SDTV(標準解像度テレビ)に対応したMP@MLが必要です。

High-1440レベルは、Mainレベルと比べて解像度が2倍の高解像度方式です。また、Highレベルでは解像度が2倍になるだけでなく、その解像度を16:9フォーマットで維持するため、1440の水平サンプルを1920に増やしています。

空間変換と再量子化を使用している圧縮方式では、スケーラブルな信号を生成することができます。スケーラブルなプロセスとは、入力によってメイン信号とヘルパ("helper")信号が得られるプロセスのことです。メイン信号は単独で復号しても、一定品質のピクチャが得られますが、ヘルパ信号の情報を加えると、品質を部分的に向上させることができます。

たとえば、従来のMPEGコードでは、係数に対する程度の高い再量子化によって、ピクチャが中程度の信号対ノイズ比で符合化されます。しかし、そのピクチャは元のピクチャからピクセル単位で局所的に復号・減算されているため、量子化ノイズのピクチャが生じます。このようなピクチャは、ヘルパ信号として圧縮・伝送することができます。単純なデコーダでは、ノイズの多いメイン信号のビット・ストリームしか復号できませんが、複雑なデコーダであれば、両方のビット・ストリームを復号して組合せ、低ノイズのピクチャを生成することができます。これは信号対ノイズ比(SNR)スケーラビリティの原理を利用しています。

HIGH HIGH – 1440 MAIN LOW		4:2:0 1920x1152 80 Mbps I, P, B	4:2:2 1920x1080 300 Mbps I, P, B			4:2:0, 4:2:2 1920x1152 100 Mbps I, P, B
		4:2:0 1440x1152 60 Mbps I, P, B			4:2:0 1440x1152 60 Mbps I, P, B	4:2:0, 4:2:2 1440x1152 80 Mbps I, P, B
	4:2:0 720x576 15 Mbps I, P	4:2:0 720x576 15 Mbps I, P, B	4:2:2 720x608 50 Mbps I, P, B	4:2:0 720x576 15 Mbps I, P, B		4:2:0, 4:2:2 720x576 20 Mbps I, P, B
		4:2:0 352x288 4 Mbps I, P, B		4:2:0 352x288 4 Mbps I, P, B		
LEVEL / PROFILE	SIMPLE	MAIN	4:2:2 PROFILE	SNR	SPATIAL	HIGH

図4-1

別の方法として、HDTVピクチャの低いほうの空間周波数のみを符合化して、SDTV受像機で復号できるようなメイン・ビット・ストリームを生成することもできます。この低解像度ピクチャを元のピクチャから局所的に復号・減算すると、解像度を高めるピクチャが得られます。このピクチャは符号化してヘルパ信号にすることができます。適切なデコーダを使用すれば、メイン信号とヘルパ信号を組合せて、HDTVピクチャを再生成することができます。これは空間スケーラビリティの原理を利用しています。

HighプロファイルではSNRおよび空間のスケーラビリティに対応しているだけでなく、4:2:2サンプリングのオプションも使用できます。

4:2:2プロファイルは、デジタル制作機器との適合性を高めるために開発されました。このプロファイルでは、高度なHighプロファイルを使用しなくても4:2:2によるオペレーションが可能です。たとえば、HP@MLデコーダはSNRスケーラビリティをサポートする必要がありますが、これは制作上の要件ではありません。4:2:2プロファイルは、GOP構成に対して他のプロファイルと同じような自由度がありますが、実際には、編集が簡単な短いGOPと一緒に使用されるのが一般的です。4:2:2によるオペレーションには、4:2:0よりも高いビット・レートが必要で、短いGOPを使用すると、同じ品質を得るために必要なビット・レートはさらに高くなります。

プロファイルとレベルの概念もMPEG-2の成果であり、堅牢かつ拡張可能であることが証明されています。MPEG-4では、プロファイルとレベルの配列がはるかに複雑で、これについては後述します。

4.4 MPEG-4

国際的な標準化には時間がかかるため、その間の技術的進歩が開発中の規格に反映できることもあります。多くの場合これは好ましいことですが、規格の改良を続けていけば、いつまでたっても完成しない、使用できないという事態も起こり得ます。規格を完成させるため、標準化プロセスでは、ある時点以降は実質的な変更を禁止する厳しい規則があります。そのため、規格が一般に採用されるときには、未反映の改善点が蓄積していることもよくあります。MPEG-2においても、こうしたことがありました。前述のように、MPEG-3は標準化に取りかかった後で廃止されたため、次のプロジェクトはMPEG-4になりました。MPEG-4ではすでに2つのバージョンが完成しており、さらに拡張を行うための作業が続いています。

MPEG-4は当初、非常に低いビット・レートでの映像・音声の符号化を主眼に置いていました。実際、MPEG-4は以下の3つのビット・レート範囲に合せて、明示的に最適化されています。

- 64kbps未満
- 64～384kbps
- 384kbps～4Mbps

低ビット・レートでの性能は依然として主要な目標の1つですが、この目標には非常に独創的なアイデアが貢献しています。また、エラー回復力にも大きな注目が集まったことで、MPEG-4は、携帯端末への送信のように、エラーの起きやすい環境での使用に

も非常に適したのになりました。しかし、他のプロファイルやレベルでは、最大38.4Mbpsのビット・レートを使用するものもあり、最大1.2Gbpsのデータ・レートを使用するスタジオ品質のプロファイルやレベルに関する開発作業も進められています。

そして、それ以上に重要なのは、MPEG-4が単なる圧縮の方式ではなく、インタラクティブ性や多彩なアプリケーションに対応するパワフルなツールを使用するマルチメディア符合化という、まったく新しい概念へと発展したことです。このMPEG-4については、公式な「概要」でさえ67ページにも及ぶため、本書ではこの規格のあらましを手短かに紹介するにとどめます。

4.4.1 MPEG-4の規格文書

MPEG-4の主要なパートは以下の通りです。

- ISO/IEC 14496-1 : システム
- ISO/IEC 14496-2 : ビジュアル
- ISO/IEC 14496-3 : オーディオ
- ISO/IEC 14496-4 : 適合性テスト
- ISO/IEC 14496-6 : DMIF(Delivery Multimedia Integration Framework)
- ISO/IEC 14496-10 : AVC(Advanced Video Coding)

4.4.2 オブジェクト符合化

従来の伝送方式と比べた場合の最も重要な進展は、オブジェクトの概念を取り入れたことです。最終的なシーンの各部分を個別に符合化して、映像オブジェクトや音声オブジェクトとして伝送し、それらをデコードによって1つにまとめる、つまり合成することができます。さまざまな種類のオブジェクトを、その役割に最適のツールを使って符号化します。オブジェクトを個別に作成したり、シーンを解析して、たとえば前景と背景のオブジェクトに分けることもできます。この技術でサッカー試合の中継ビデオを処理して、ボールのみをシーンと別にするとした面白いデモも紹介されていました。この背景(ボール抜きのシーン)を流して、

ペイ・パー・ビューの視聴者を勧誘します。見えるのは選手と試合場だけで、料金を支払えばボールも見えるようになるという仕組みです。

このオブジェクト指向アプローチにより、MPEG-4ストリームの3つの主要な特徴が生まれました。

- 複数のオブジェクトをさまざまな技術で符合化し、デコーダで合成できる。
- オブジェクトは、カメラで撮った自然の素材でも、テキストのような合成の素材でもよい。
- ビット・ストリームで運ばれる命令やユーザの選択によって、1つのビット・ストリームで何通りもの表示ができる。

MPEG-4のオブジェクト符合化のために一般化されたシステムを図4-2に示します。この図では、MPEG-4システム内でのユーザ対話の可能性も強調されています。これは、特にテレビ・ゲーム設計者にとって役に立つ機能です。

ただし、これらの機能は使用しなくても構いません。MPEG-4では従来の映像・音声の符号化も提供されており、効率の向上やエラー回復力などによって、MPEG-2よりもさらに高い機能を提供しています。ただし、MPEG-4の真価は今述べたアーキテクチャにあります。オブジェクトを個別に符合化することには、数々の利点があります。オブジェクトごとに最も効率的な方法で符合化でき、さまざまな空間/時間スケーリング(4.4.3を参照)を必要に応じて使用することができます。

4.4.3 映像および音声の符号化

MPEG-4の映像符合化ツールの多くは、MPEG-2と似ていますが、予測符合化やより効率の高いエントロピー符合化を効果的に取り入れることで、強化が図られています。また、ツールの適用が先行規格と大きく異なります。

MPEG-4では映像オブジェクトを符合化します。簡素化したモデルで表すと、符合化はMPEG-2と非常によく似ていますが、MPEG-4では映像が長方形の映像オブジェクトとして表されます。このような画像の表現は、テクスチャ符合化と呼ばれています。

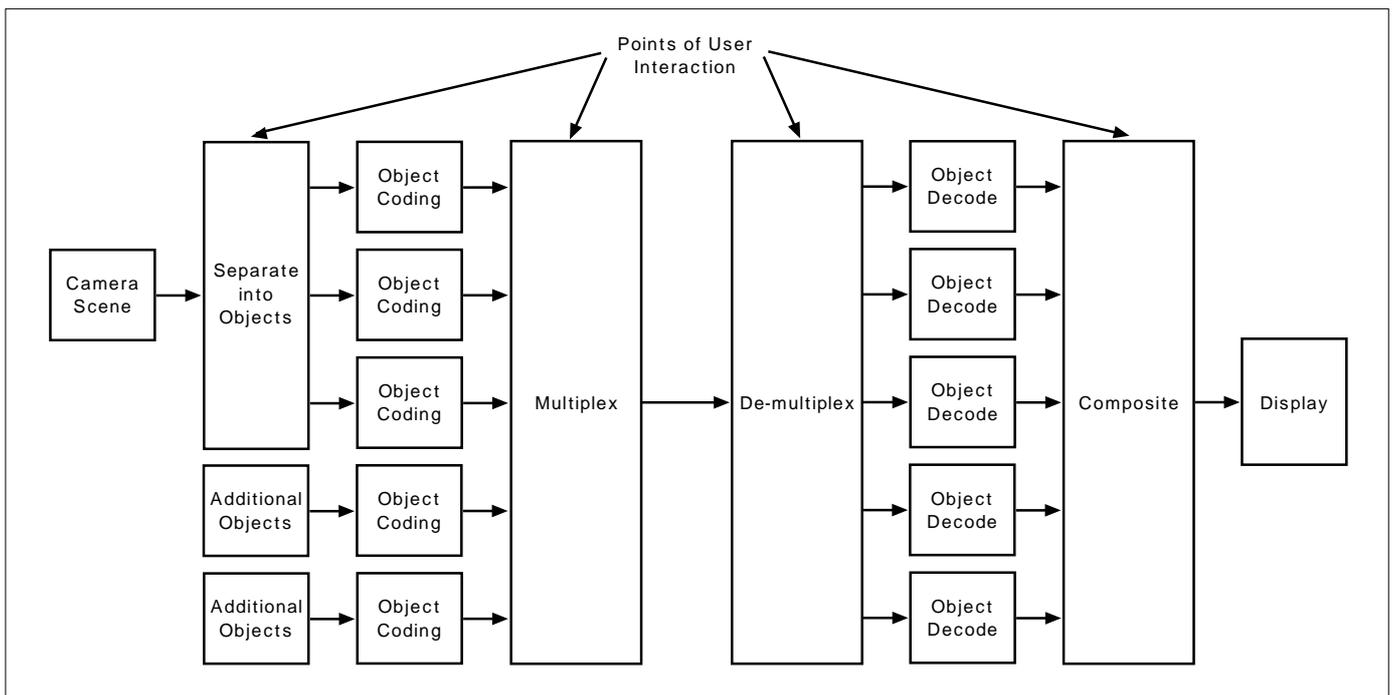


図4-2

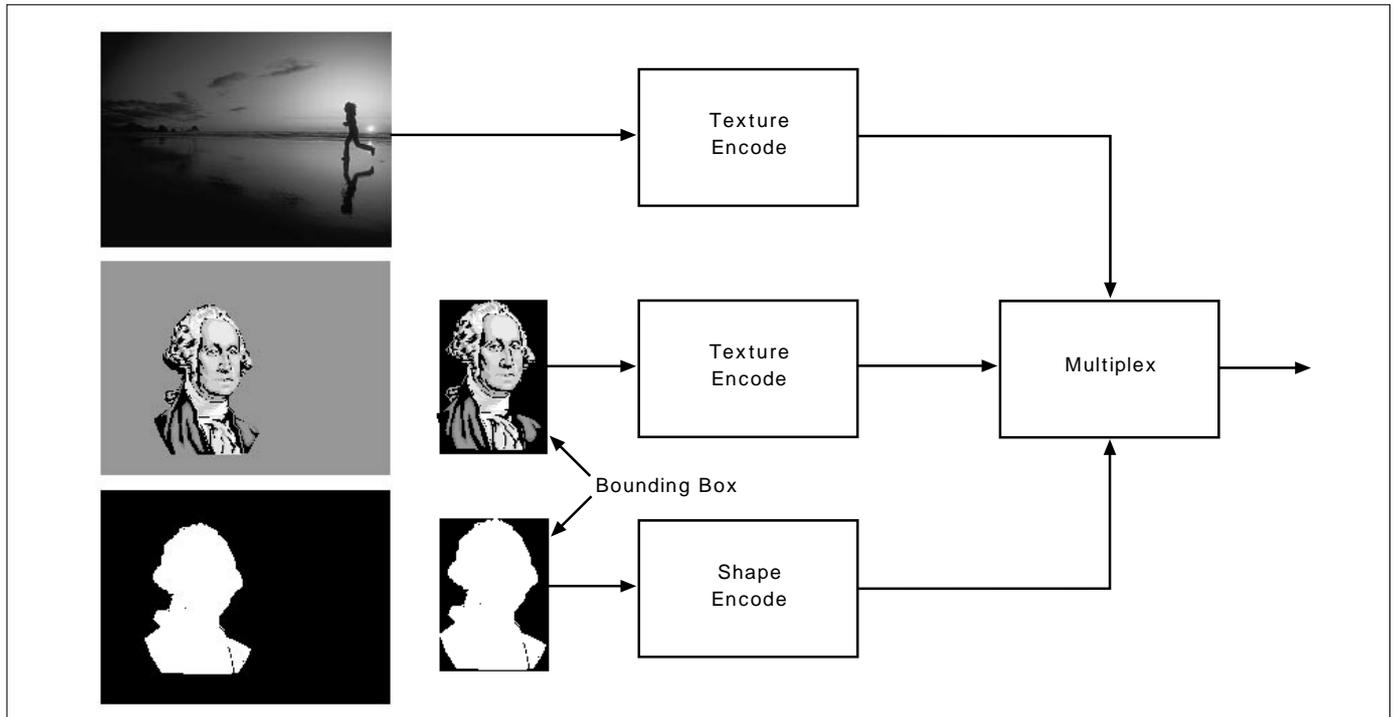


図4-3

映像オブジェクトが複数ある場合、中には形状が不規則なものもあり、一般にはどれもフル画面の背景オブジェクトよりも小さくなります。つまり、符号化が必要なのはオブジェクトの有効な部分だけですが、形状や位置も表す必要があります。この規格には、長方形や不規則なオブジェクトを2進表現やグレー・スケール表現(アルファ・チャンネルに類似)で形状符合化するためのツールも含まれています。図4-3にこの概念を示します。

また、ライブ音声の符号化にも、MPEG-1やMPEG-2と同様のツールを使用しており、AACによってさらに高い効率を得られます。複数の音声「オブジェクト」を個別に符号化し、デコーダで合成できます。映像の場合と同じく、自然の音声から合成の音声までオブジェクトで表すことができます。

4.4.4 スケーラビリティ

メディア圧縮の分野におけるスケーラビリティとは、1つのビット・ストリームの中に、複数の品質レベルのコンテンツを分布させる機能を指します。図4-4のように、MPEG-2とMPEG-4では、エンコーダで1つのベース・レイヤと1つ以上のエンハンスメント・レイヤを生成するという従来のモデルによって、スケーラブルなプロファイルを提供します。伝送や復号においてリソースが不十分な場合は、エンハンスメント・レイヤを捨てることができます。この方法は有効ですが、品質のレベルに関する判断はすべて符合化の時点で行う必要があります。実際には、エンハンスメント・レイヤの数は厳しく(通常は1つに)制限されます。

MPEG-4の新しいバージョンには、ファイン・グレイイン・スケーラビリティ(FGS)プロファイルが含まれています。この技術では、品質レベルが最高のビット・ストリームが生成されますが、品質の低いバージョンも抽出されるようになっています。FGSでは、図4-5に概念的に示したようなビット・プレーン符号化を使用しま

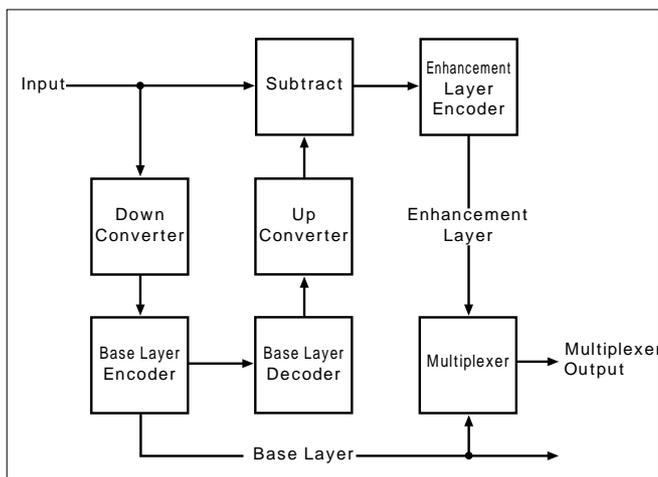


図4-4

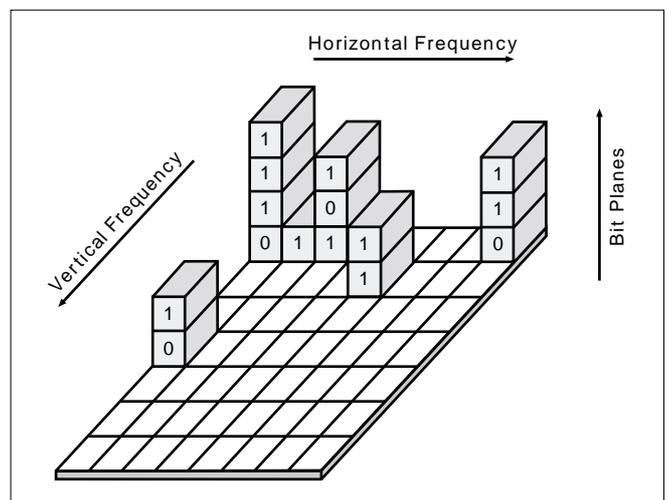


図4-5

す。量子化された係数は、最上位ビットから1ビットずつ「スライス」されます。これにより、まず最大(かつ最上位)係数のおおよその値が表現されます。そしてさらに、最上位係数がより正確に表現され、その次に大きい係数のおおよその値が表現される...というようにスライスされていきます。

リソースが十分でない場合は、FGSなどの空間スケーリングを、低いフレーム・レートでの伝送や復号が可能な時間スケーリングと組合せることもできます。前述のように、オブジェクトは個別にスケーリングできるため、前景の重要なオブジェクトは時間解像度を十分に維持し、背景は低いレートに更新することも可能です。

4.4.5 MPEG-4のその他の側面

MPEG-4規格の内容は膨大で、これまでの説明でも、さまざまな側面のごく一部に触れたにすぎません。高品質符号化用のStudioプロファイルでは、オブジェクト符号化と組合せることで、映像として合成するさまざまな個別の要素を構造化ストリージで管理することもできます。今後さらに拡張されることで、デジタル・シネマに適した品質レベルが得られる可能性もあります。図4-6は、現在のMPEG-4プロファイルの定義を示しています。(この図に示したのはプロファイルのみであり、通常はプロファイルごとに複数のレベルも定義されます。)

MPEG-4で定義されているオブジェクトの中には興味深いものがあります。その一例がスプライト(sprite)です。スプライトとは

一般に、ディスプレイ装置の表示域に収まりきらない静的な背景オブジェクトのことをいいます。たとえば、テレビ・ゲームの動きは背景シーンの手前にあります。スプライトを使用すると、大きな静的背景を一度送れば、ゲームの進行中、表示域の動きに応じて背景の適切な部分を表示させることができます。

MPEG-4では、顔および胴体のアニメーション・プロファイルが定義されています。いずれの場合も、デフォルトの顔や胴体があり、コマンドを送ってこのオブジェクトに動きを与えます。あるいは、デフォルトのオブジェクトにビット・ストリームで変更を加えることもできます。たとえば、具体的な表情の顔を送って動きを与えます。言語に関連した高度なアニメーション・コマンドでは、保存されている顔に、さまざまな言語で文字を「読ませる」こともできます。

MPEG-4はテレビ・ゲームの規格と形容されることもあり、確かに、この業界にとって最適ともいえる要素を数多く備えています。しかし、ざっと見ただけでもこれほど豊富な機能があり、しかもそれぞれに十分な奥行きがあることを考えると、さまざまな分野での利用の可能性が無限にあることがわかります。

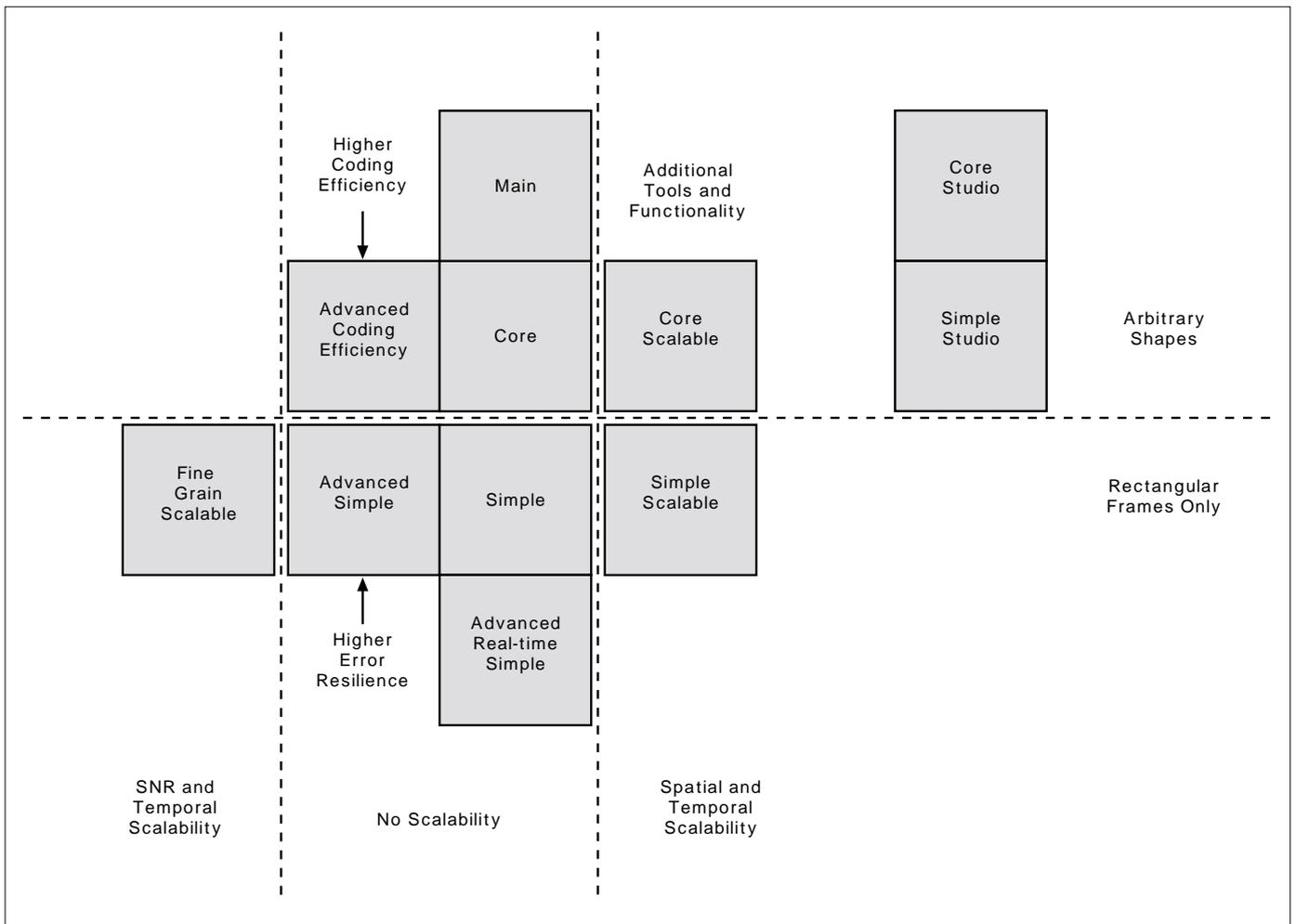


図4-6

第4章 MPEGの規格

4.4.6 MPEG-4の今後

以上のように、MPEG-4は、さまざまなアプリケーションに豊富な機能を提供するための多彩な規格で構成されています。ただし、これはあくまで理論上の話です。かつては、MPEG-4が短期間のうちにインターネットで伝送されるAV素材の主要な符号メカニズムになり、現在使用されている各社開発のコーデックに取って代わるという見方が大勢を占めていました。このような現状には、主に2つの原因があります。

原因の1つは、技術とそれによって得られる性能です。MPEG-4は、ITUが90年代初頭に開発したH.26xに基づく映像圧縮技術を使用しています。インターネット上での音声・映像配信は競争の激しいビジネスであり、Apple、Microsoft、RealNetworksなどの大手は、MPEG-4コーデックよりも性能面で優れた独自の符号化方式を開発しています。

もう1つの原因は、特許許諾の状況です。2002年初頭までの間、MPEG-4を採用したい企業は、特許保有者に対してどれだけの使用料を払えばよいのかわかりませんでした。現在、MPEG-4の基本レベルに対するライセンス供与体系の案が発表され、産業界からの強い反発にありました。

しかし、将来的な希望はあります。ITUとMPEGの共同組織である合同ビデオ・チーム(JVT)が、H.26Lと呼ばれるコーデックの開発に取り組んでいます。ITUによれば、「H.26Lは、ブロック・ベースの動き補正ハイブリッド変換コーデックで、先行規格と比べると、本質的には同じですが、細部に多くの違いがあります。H.26Lでは使用可能なブロック・サイズや、動きの推定に使用できる基準ピクチャ数を大幅に向上させています」。この新たなコーデックは、動きの推定の精度を大きく向上させており(分野によっては1/8画素単位)、主要なブロック・サイズも、ほとんどのMPEG方式で使用されている8×8ではなく4×4です。

H.26Lは符号化の効率を著しく向上させると期待されており、参加組織は、インターネット・ストリーミングに適した基本レベルについては無償でライセンスすることを目標としています。JVTの第1段階は2002年に完了し、MPEG-4 Part 10として発表されました。

4.5 MPEG-7

MPEG-3の規格化が中止されたため、実際の規格はMPEG-1、MPEG-2、MPEG-4の順になっています。委員会の参加組織の中には、次の規格をMPEG-5にしたいという意見や、1、2、4と倍数で来たので、次はMPEG-8がいいという意見もありました。しかし、最終的に、ただ順番通りではMPEG-1～MPEG-4と根本的に違うことを示せないという結論に至り、MPEG-7になりました。

MPEG-7は圧縮ではなく、「ビットに関するビット」とも呼ばれるメタデータの規格です。メタデータとは、他のデジタル・データの内容について記述したデジタル情報です。最近では、番組の素材やコンテンツ、実際の画像、映像、音声、データなど、情報を伝えるオブジェクトをデータ・エッセンス(本質)と呼んでいます。メタデータを見れば、データの本質が何かを知るために必要な情報がすべてあります。

ビデオ・テープ、書籍、音楽など、情報を保存する作業に携わった経験があれば、正確な目録作成や索引付けの重要性や難しさが理解できるはずです。保存された情報は、そこに保存されていることがわかり、必要なときにすぐに取り出すことができ初めて有益なものになります。

こうした問題は常にあり、アナログの世界では、ラベル、カタログ、索引カードなどを使用して対処してきました。近年、コンピュータ技術がもたらした、経済的で効率の高いリレーショナル・データベースでは、強力なサーチ・エンジンによって、保存情報に画期的な方法でアクセスできるようになりました。ただしこの場合、情報はサーチ・エンジンが対応できる形式で存在していなければなりません。

ここに重要な問題があります。世界中で新たなメディア・コンテンツがめまぐるしいペースで生み出され、そのペースは一段と加速しています。そして、デジタル記憶媒体の大容量化と低価格化が進み、こうしたコンテンツを一層多く保存できるようになっています。LANやWANのネットワークによって、そうしたコンテンツに対するアクセスや配信が可能になります。しかし、サーチ・エンジンで必要な情報を見つけたり、素材自体にデータベースをリンクさせたりするためには、サーチ・エンジンに適した形式で、データベースにすべての索引情報を登録する必要があります。

これまでの規格から推測すると、MPEG委員会は、データ生成のメカニズムについては必要以上に関与しないと思われます。標準的な体系を作り、そこに市場のニーズがあれば、技術的なギャップは自ずと埋められる、という立場をとっています。これまでのMPEG規格では、構文とデコーダについて指定していました。ところが、図4-7に示すように、MPEG-7で規格化されているのは構文のみです。メタデータの作成や、それを使用するアプリケーションについても指定されていません。指定されているのはメタデータの表現の仕方です。つまり、データベースを作成するフィールドについて指定しているため、サーチ・エンジンを設計する際に、存在する記述エレメントの種類や符号化の仕方がわかります。

MPEG-7では、記述子および記述方式を定義することで、ほぼどのような特徴でも表現できるようにしています。少なくとも、色相ヒストグラムや形状などの基本エレメントを組合せて、顔のような複雑な要素を表すことは理論的に可能です。また、素材に自動的に索引を付けることで、たとえば、クリントン大統領と連邦準備制度理事会のグリーンズパン議長が一緒に映っているシーンをデータベースから検索するということもできます。機能は画像だけに限定されるわけではありません。音声サンプルを使って、パラボロティの録音データや画像を検索したり、キーボードを弾いて、それと一致または類似するメロディを検索することも可能になるでしょう。

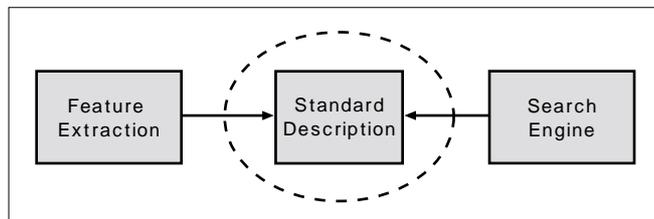


図4-7

記憶装置やネットワーク・システムの急速な進歩により、膨大な量のデジタル・コンテンツにアクセスできるようになります。技術が進歩してMPEG-7の要求を満たせるようになれば、数年前には想像もできなかったような方法によるデータの索引付けや取り出しが可能になります。そこで次に必要になるのは、こうしたコンテンツに関わるアクセス、プライバシー、商取引などを管理するシステムです。これはMPEG-21が対象にしている領域です。

4.6 MPEG-21

MPEG-21も、これまでのMPEG規格と本質的に異なります。基本的な概念はきわめて簡潔ですが、カバーしようとしている範囲は広大です。MPEG-21は、商取引や権利の管理に対するインフラ面でのさまざまな対応をはじめとした、デジタル資産を管理・使用するための包括的な枠組みの構築を目指しています。MPEG-21の構想は「マルチメディア資源を、さまざまなネットワークやデバイスにわたって、透過的かつ一層有効に利用できるようにすること」です。

MPEG-21標準化作業の範囲は、テクニカル・レポートの原案で定義された7つの構成要素によって表されます。

1. デジタル・アイテム宣言では、「デジタル・アイテムを定義するための、均一かつ柔軟な抽象概念と相互運用可能な体系を確立」するものと予想されます。この体系は、あらゆる種類のメディア資源や記述方式に合わせて拡張でき、また、検索やナビゲーションが簡単な階層構造にも対応できなければなりません。
2. デジタル・アイテムの表現では、コンテンツを符合化したり、コンテンツのさまざまなエレメントを同期化するのに必要なメカニズムを提供する技術について規定します。このレイヤは、少なくともMPEG-4を参照して策定されるものと予想されます。
3. デジタル・アイテムの識別および記述では、デジタル・アイテムを識別・記述する(コンテンツのエレメントをリンクさせる)ための枠組みについて規定します。ここにはMPEG-7の記述方式が加えられるものと考えられますが、「効果的かつ正確な自動化されたイベントの管理および報告(ライセンスに関する処理、使用規則、監視や追跡など)をサポートする、新世代の識別システム」も含まれている必要があります。あらゆるクラスのMPEG-21ユーザのニーズを満たせるものでなければなりません。
4. コンテンツの管理および取扱では、MPEG-21のデジタル・アイテムや記述を保存・管理するためのインタフェースやプロトコルについて定義します。コンテンツの保存や登録に対応する際には、ユーザの権利も維持されるようにし、アイテムや記述への変更を追跡する機能にも対応できるようにする必要があります。また、消費者が個人情報と引き換えにコンテンツへのアクセス権を得る一種の「取引」や、「パーソナル・チャンネル」あるいは同等のメカニズムの確立にも対応することになると考えられます。
5. 知的財産の管理と保護は、不可欠な要素です。MP3オーディオ・ファイルをめぐる現在の論争からも明らかのように、デジタル世界を認識した新たな著作権の仕組みが必要です。コンテンツは、保護されなければ価値がないともいえます。MPEG-21は、規格化作業が進んでいるMPEG-4やMPEG-7がベースになりますが、新しい種類のデジタル・アイテムや送信メカニズムに対応するための拡張も必要です。
6. MPEG-21の端末とネットワークでは、さまざまなネットワークでのアイテムの送信と、さまざまな端末へコンテンツを提供する機能に対処します。概念的には、映画館への動画送信では、十分なデジタル・シネマ品質が維持されなければならない、通信速度の遅いネットワークで一般ユーザの端末へ送る場合は、それよりも品質が下る(価格も別)ということになります。いずれの場合も、種類や利用回数には何らかの制約が設けられることになります。送信や提供に関連する問題や複雑性をユーザが意識せずにする必要があります。

7. 最後に、イベント報告において「報告が必要な各種イベントを実行するためのメトリクスやインタフェースを標準化」する必要があります。たとえば、システムが保護アイテムへのアクセスをユーザに許可する場合は、同時に適切な支払いが行われるようにする必要があります。

第5章 エレメンタリ・ストリーム

エレメンタリ・ストリームとは、エンコーダからの未処理出力であり、基本的にはデコーダが元の画像や音声を復号するのに必要なデータ以外は含まれていません。MPEGでは、圧縮信号の構文が厳密に定義されているため、それに基づいたデコーダの動作が保証されます。エンコーダについて定義されているのは、方法を問わず正しい構文を生成するものでなければならないということのみです。

このアプローチは、大抵はエンコーダの数よりもデコーダの数のほうが多くなる実際の状況に適しています。デコーダを規格化すれば、コストを低減することができます。それに対し、エンコーダは複雑で高価なものにしても、大きなコストの負担はなく、複雑なものになるほど画質が向上する可能性があります。エンコーダとデコーダの複雑さが異なる場合、その符合化システムは非対称になっているといえます。

MPEGのアプローチでは、符合化アルゴリズムの高度化によって品質を高めつつ、先行モデルのデコーダにも理解可能なビット・ストリームを生成できます。また、このアプローチでは、公開する必要のない独自開発の符合化アルゴリズムを使用することもできます。

5.1 ビデオ・エレメンタリ・ストリームの構文

図5-1は、ビデオ・エレメンタリ・ストリームの構造を示しています。画像情報の基本単位は、画素が 8×8 に配列されたY、Cb、CrなどのDCT(離散コサイン変換)ブロックです。まずDC係数が送られ、他の係数よりも正確に表現されます。そして、それ以外の係数が送られた後に、ブロック終了(EOB)コードが送られます。

ブロックは組合されてマクロ・ブロックになります。これは動き補正できるピクチャの基本単位です。各マクロ・ブロックのヘッダには2次元の動きベクトルがあります。Bピクチャでは、順方向だけでなく逆方向のベクトルもあります。動き補正はフィールドまたはフレーム単位で行われ、どちらかが示されます。また、係数の再量子化に使用するスケールも示されます。デコーダはベクトルを使用して、直前と直後のピクチャから情報を得て、予測ピクチャを作成します。そして、ブロックを逆変換して補正ピクチャを生成し、これを予測ピクチャに加えることで復号を行います。

す。4:2:0の符合化では、マクロ・ブロックごとにYブロックが4つと色差ブロックが2つになります。どのブロックがどのコンポーネントを記述しているかを識別できるようにするため、ブロックは指定された順序で送られます。

マクロ・ブロックは組合されてスライスになります。スライスは常に、ピクチャの左から右への水平の帯になります。MPEGでは、スライスはどこから開始してもよく、サイズも任意ですが、ATSCでは常にピクチャの左端から開始しなければなりません。画面の横方向に複数のスライスが存在することができます。スライスは可変長符合化や差分符合化における同期化の基本単位です。スライスの最初のベクトルは必ず送られますが、それ以外は差分によって送られます。Iピクチャでは、スライスの最初のDC係数は必ず送られますが、それ以外のDC係数は差分によって送られます。差分ピクチャでは、これらのDC係数による相関性が見込まれていないため、この技術は適していません。

エレメンタリ・ストリームにビット・エラーがあると、可変長記号の非直列化が失敗するか、後に続く差分符合化した係数やベクトルが不正確になります。このようなスライス構造では、ビット・ストリームに再同期化ポイントを設けることで復旧できます。

複数のスライスを組合せて、フィールドあるいはフレームの有効部分のピクチャを作成します。ピクチャ・ヘッダは、そのピクチャがI、P、Bのいずれによって符合化されたかを定義するとともに、ピクチャが正確な時間に表示されるようにするための時間基準が含まれています。パンやおりの場合は、すべてのマクロ・ブロックのベクトルが同じになります。ピクチャ全体のグローバル・ベクトルが送られ、個々のベクトルは、このグローバル値との差分になります。

ピクチャを組合せて、Iピクチャ(伝送順)で始まるGOPを生成することができます。GOPは時間符合化の基本単位です。MPEG規格ではGOPの使用は任意ですが、実際には必要です。第2章で説明したように、Iピクチャの間に可変数のPピクチャやBピクチャを配置することができます。GOPはオープン、クローズドのどちらも可能で、クローズドGOPでは、最後のBピクチャは、復号のために次のGOPのIピクチャを必要とせず、そのGOPの終わりでビット・ストリームを切ることも可能です。

GOPを使用する場合、複数のGOPを組合せて、ビデオ・シーケ

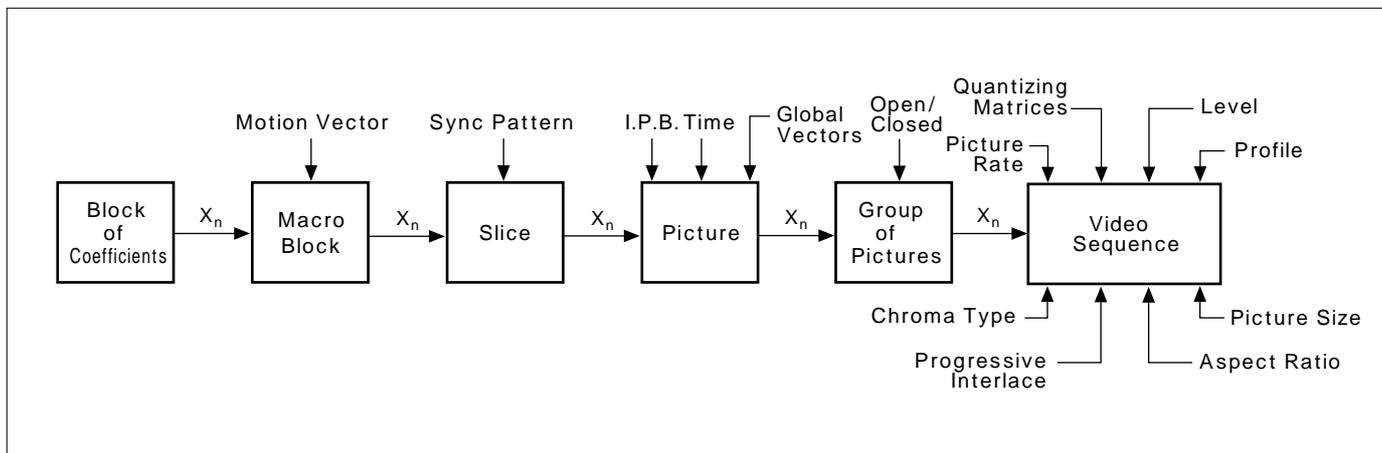


図5-1

ンスを作成することができます。シーケンスは、シーケンス開始コードで始まり、次にシーケンス・ヘッダが来て、最後にシーケンス終了コードがきます。また、追加のシーケンス・ヘッダを、シーケンス全体に配置することができます。このアプローチにより、DVDやカセット・テープを再生する時のように、シーケンスの途中の部分から復号を始めることができます。シーケンス・ヘッダは、ピクチャの縦と横のサイズ、アスペクト比、色度サブサンプリング・フォーマット、ピクチャ・レート、プログレッシブ/インタレース走査の使用、プロファイル、レベル、ビット・レート、ピクチャ内/ピクチャ間符合化ピクチャで使用する量子化マトリックスなどについて指定します。

シーケンス・ヘッダのデータがないと、デコーダはビット・ストリームを理解できないため、シーケンス・ヘッダは、デコーダが正確な動作を開始するためのエントリ・ポイントになります。エントリ・ポイントの間隔は、視聴者がテレビのチャンネルを変えた際に起こる復号の遅延に影響します。

5.2 オーディオ・エレメンタリ・ストリーム

MPEG-2多重化では、MPEGレイヤ1、2、3やAC-3で符合化された音声を含め、さまざまな種類の音声を埋め込むことができます。使用する符合化の種類を記述子に含めることで、デコーダがその記述子を読み取り、適切な種類の復号が実行されるようにする必要があります。

音声圧縮のプロセスは映像圧縮と大きく異なります。音声フレームにはI、P、Bのような種類がなく、含まれている音声データの量は常に同じです。また、双方向符合化に相当するものもなく、伝送時にフレームの順序が変わることもありません。

MPEG-2オーディオでは、音声の圧縮に使用したレイヤと、使用した圧縮の種類(ジョイント・ステレオなど)、および元のサンプリング・レートが、シーケンス・ヘッダの記述子に含まれます。オーディオ・シーケンスは、符合化されて音声フレームになるアクセス・ユニット(AU)で構成されます。

ATSCのようにAC-3符合化を使用すると、そのことがシーケンス・ヘッダに反映されます。音声のアクセス・ユニット(AU)は、セクション3.7で説明したようなAC-3同期フレームです。AC-3同期フレームは、音声サンプル1536個分の時間範囲を表し、48kHzのサンプリングでは32ms、32kHzでは48msになります。

第6章 パケット化エレメンタリ・ストリーム(PES)

実用上、圧縮器から出力される音声や映像を運ぶエレメンタリ・ストリームは、パケットに分割する必要があります。分割されたパケットは、同期化のためのタイムスタンプを含むヘッダによって識別されます。このPESパケットを使用して、プログラム・ストリームやトランスポート・ストリームを生成することができます。

MPEGでは、エンコーダではなく、トランスポート・ストリームについてのみ定義しているため、設計者はエレメンタリ・ストリームをワンステップでトランスポート・ストリームに変換できるマルチプレクサを構築することもできます。この場合、PESパケットは識別可能な形式では存在せず、トランスポート・ストリームのペイロードに論理的に存在することになります。

6.1 ESパケット

PESでは、エンドレスなエレメンタリ・ストリームが、アプリケーションに合せて適当なサイズのパケットに分割されます。サイズはアプリケーションによって異なりますが、数百キロバイト程度です。

各パケットの先頭にはPESパケット・ヘッダがきます。図6-1は、ヘッダの内容を示しています。まず、24ビットの開始コード・プレフィックスと、パケットの中身が映像か音声かを識別して音声符合化の種類についても指定するストリームIDがあります。これら2つのパラメータ(開始コード・プレフィックスとストリームID)が、パケットの始まりを識別するパケット開始コードを構成します。PESのパケットと、トランスポート・ストリームで使用される(はるかに小さい)パケットは名前が同じであるため、混同しないようにしてください。

6.2 タイムスタンプ

圧縮されたピクチャは、双方向符合化の関係で順序を変えて送られます。必要なデータの量は一定ではなく、多重化や伝送によって、遅延も変動します。音声と映像を一緒に固定しておくため、ピクチャごとに定期的にタイムスタンプを加えます。

タイムスタンプは33ビットの数字で、90kHzのクロックによって駆動されるカウンタのサンプルです。このクロックは27MHzのプログラム・クロックを300で割ることによって得られます。表示時間は均等なので、すべての表示ユニットにタイムスタンプを加える必要はありません。タイムスタンプはデコーダによって補間することができます。ただし、プログラム・ストリームやトランスポート・ストリームにおいて、間隔を700ms以上空けてはいけません。

タイムスタンプは、アクセス・ユニットが属している時間的な位置を示します。ビデオおよびオーディオPESパケットのヘッダにタイムスタンプを組み込むことで、リップ・シンクを行うことができます。デコーダは、選択されたPESパケットを受信すると、アクセス・ユニットごとに復号してRAMに入れます。タイムラインのカウンタ数がタイムスタンプの値に到達すると、RAMが読み出されます。このオペレーションによって、望ましい結果が2つ得られます。1つ目は、各エレメンタリ・ストリームで実質的にタイム・ベース補正が行われることであり、2つ目は、ビデオおよびオーディオのエレメンタリ・ストリームを同期させて番組を作成できることです。

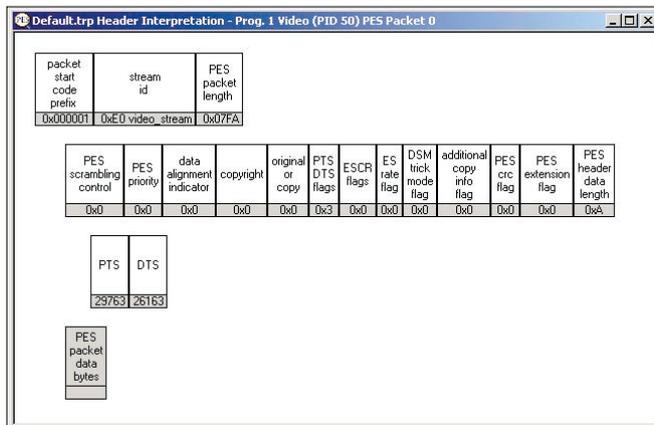


図6-1

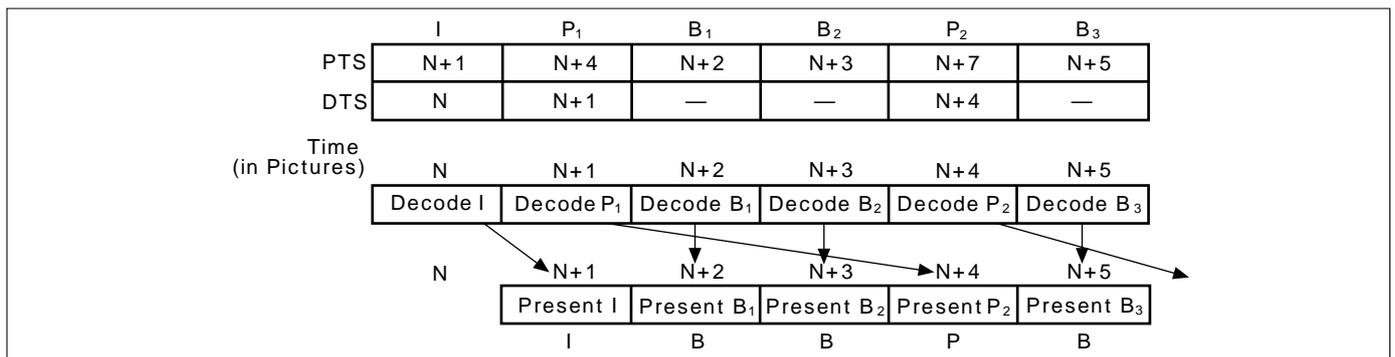


図6-2

6.3 PTS/DTS

双方向符号化を使用する場合、ピクチャを表示される少し前に復号して、Bピクチャのデータ源にしなければならない場合があります。たとえば、表示される時はIBBPの順でも、伝送の時はIPBBの順になることがあります。そのため、2種類のタイムスタンプが存在します。復号タイムスタンプ(DTS)は、ピクチャを復号する時間を示し、表示タイムスタンプ(PTS)は、デコーダから出力して表示する時間を示します。

Bピクチャは復号と同時に表示されるため、含まれているのはPTSのみです。IPBBシーケンスを受け取った場合は、1つ目のBピクチャの前に、IピクチャとPピクチャを復号する必要があります。デコーダは一度に1枚ずつしか復号できないため、まずIピクチャを復号して保存します。そしてPピクチャを復号している間に、復号済みのIピクチャを出力することで、続いてBピクチャを復号できるようにします。

図6-2のように、Iピクチャが含まれるアクセス・ユニットを受け取った場合、ヘッダにはDTSとPTSの両方があり、これらのタイムスタンプは1ピクチャ期間ごとに分けられます。双方向符号化を使用している場合は、次にPピクチャが来る必要があり、このピクチャにもDTSとPTSの両方がありますが、この2つのタイムスタンプは、Bピクチャを挿入できるように、3つのピクチャ期間で分けられます。そのため、IPBBを受け取った場合、Iは1ピクチャ期間遅れ、Pは3ピクチャ期間遅れ、2つのBにはまったく遅れがなく、表示順はIBBPとなります。IとPの間にBが多くなるようにGOP構造を変更すると、PピクチャにおけるDTSとPTSの差は当然大きくなります。

パケット・ヘッダのPTS/DTSフラグは、PTSタイムスタンプのみ、あるいはPTSとDTSの両方のタイムスタンプが存在していることを示すように設定されます。音声パケットに複数のアクセス・ユニットが含まれている場合、パケット・ヘッダにはPTSが含まれます。ただし、決して転送順序が変わることはないため、音声パケットにDTSはありません。

第7章 プログラム・ストリーム

プログラム・ストリームは、複数のPESパケット・ストリームをまとめる方法の1つであり、DVDのような録画のアプリケーションに強みを発揮します。

7.1 録画と伝送

画質を一定とした場合、圧縮した映像のデータ・レートはピクチャの中身によって変動します。そのため、最もよい結果を得られるのは可変ビット・レートのチャンネルです。しかし、実際のチャンネルはほとんどビット・レートが固定されており、全体のビット・レートはスタッフィング(意味のないデータ)を使って一定に保たれています。

DVDでは、スタッフィングを使うと記憶容量が無駄になります。しかし、記憶媒体では、物理的に、あるいは(ディスク・ドライブのように)データ転送レートを変えることによって、加速/減速できます。このアプローチによって、容量を犠牲にすることなく、可変レートのチャンネルを得ることができます。媒体の再生時、ダイナミックに変動する実際のビット・レートに関係なく、スピードを調整してデータ・バッファの使用率を半分ほどに保つことができます。デコーダがバッファから読み出すレートが上げれば、バッファは空になりやすく、ドライブ・システムは、アクセス速度を高めてバランスを回復します。ただし、この方法が有効なのは、音声と映像が同じクロックで符合化されている場合のみです。そうでない場合は、音声と映像が録画データ全体にわたってずれてしまいます。

録画と伝送の相反する要件に対応するため、プログラム・ストリームとトランスポート・ストリームが互いに代替手段として使用されてきました。プログラム・ストリームは、録画環境における可変ビット・レートの単一の番組に適しているのに対し、トランスポート・ストリームは伝送環境における固定ビット・レートの複数の番組に適しています。

DVDプレーヤーでは、ソースへのGENロックの問題は起きません。DVDプレーヤーは、ローカルの同期化パルス・ゼネレータ(内部または外部)によって、映像のタイム・ベースを判断し、そのタイム・ベースでピクチャを提供するためにディスクからデータを得ます。伝送では、デコーダはエンコーダにおけるタイム・ベースを再現する必要があり、再現されないと、オーバーフローやアンダーフローが起こります。そのため、トランスポート・ストリームはプログラム・クロック・リファレンス(PCR)を使用しますが、プログラム・ストリームでは、プログラム・クロックは必要ありません。

7.2 プログラム・ストリームの初歩知識

プログラム・ストリームとは、PESパケットを多重化したもので、同じマスタ・クロックまたはシステム・タイム・クロック(STC)を使って復号した複数のエレメンタリ・ストリームが運ばれます。プログラム・ストリームは、ビデオ・ストリームとそれに対応するオーディオ・ストリームで構成される場合、あるいはオーディオのみのマルチチャンネル番組の場合もあります。ビデオのエレメンタリ・ストリームはアクセス・ユニット(AU)に分割され、各ユニットに1つのピクチャを記述した圧縮データが含まれています。これらのピクチャはI、P、Bのいずれかに識別され、それぞれ正しい表示順を示すAU番号が含まれています。1つのビデオAUが、1つのプログラム・ストリーム・パケットになります。映像では、このパケットのサイズは一定ではありません。たとえば、1ピクチャ・パケットはBピクチャ・パケットよりもはるかに大きくなります。デジタル・オーディオのアクセス・ユニットは一般に同じサイズで、複数が組合さって1つのプログラム・ストリーム・パケットになります。このプログラム・ストリーム・パケットと、トランスポート・ストリーム・パケット(はるかに小さく、固定サイズ)を混同しないようにしてください。ビデオAUとオーディオAUの境界が時間軸上で重なることはほとんどありませんが、これは各境界が独自のタイムスタンプ構造を持っているためであり、重ならないこと自体は問題ではありません。

第8章 トランスポート・ストリーム

トランスポート・ストリームは、単に多数のPESパケットが多重化されたものではありません。プログラム・ストリームでは、音声と映像が共通クロックに固定されるため、タイムスタンプがあれば時間軸は十分に再現できます。しかし、データ・ネットワークで離れた地点へ伝送するとなると、デコーダで各番組のクロックを再現するという要件が追加されます。そのため、PCR信号を提供する構文のレイヤを追加することが必要になります。

8.1 トランスポート・ストリームの役割

トランスポート・ストリームではさまざまな番組が運ばれますが、それぞれに圧縮率が異なったり、全体としては一定のビット・レートもダイナミックに変化している場合があります。こうした状態は統計的な多重化と呼ばれ、素材の処理が困難な番組が、素材の処理が簡単な番組から周波数帯域を借りることができます。ビデオPESごとに、関連付けられたオーディオ/データPESの数は異なります。このような柔軟性があるにもかかわらず、デコーダは番組ごとに、音声およびデータの適切なチャンネルを選ぶことができなければなりません。一部の番組を、加入者や料金を払った視聴者のみが見られるよう保護することができます。この保護を管理するため、トランスポート・ストリームにはCA情報が含まれています。そして、このタスクを処理するためのPSIが含まれています。

トランスポート・レイヤは、PESデータを(必要であれば、スタッフィング・ビットを加えて)自己完結的な一定サイズの小包に分割します。これらの小包がデコーダに到着した際、タイミング的にジッタを起すことがあります。時分割多重を行っても遅延が生じますが、遅延度が一定でないのは、番組ごとに割当てるビット・ストリームの割合が一定でないためです。タイムスタンプは解決策の1つですが、安定したクロックが得られな

ければ効果がありません。トランスポート・ストリームには、安定したクロックを再現できるようにするための追加データが必要です。

デジタル・ビデオ制作機器の動作は、安定したシステム・クロックの配信によって同期が得られるかどうかによって大きく左右されます。ビデオ制作ではGENロックを使用しますが、距離が長くなると、個別のクロックを配信するのは実用的ではありません。トランスポート・ストリームでは、出所の異なるさまざまな番組があり、それらが同期化されているとは限りません。そのため、トランスポート・ストリームは、番組ごとに同期の手段を個別に提供する必要があります。

この追加的な同期手段はPCRと呼ばれ、再現される安定した基準クロックはデコーダで分割されてタイム・ラインを生成するため、各番組のエレメンタリ・ストリーム用のタイムスタンプも有効になります。そのため、1つの番組のエレメンタリ・ストリームは、時間基準がすべて同じになります。

シングル・プログラム・トランスポート・ストリーム(SPTS)では、1つのPCRチャンネルが、音声と映像の両方に対して1つのプログラム・クロックを再現します。SPTSは通常、オーディオ/ビデオ・コードとマルチプレクサの間の通信に使用されます。

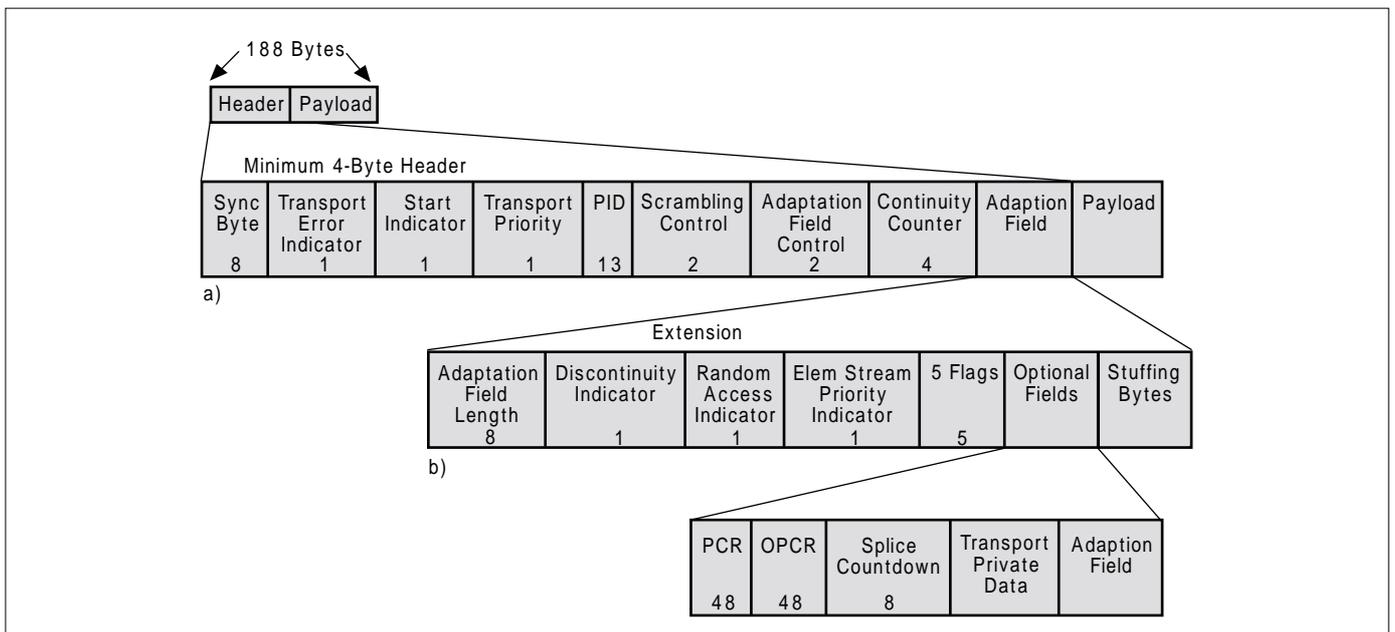


図8-1

8.2 パケット

図8-1は、トランスポート・ストリーム・パケットの構造を示しています。サイズは常に188バイトで、ヘッダ部とペイロード部に分けられます。図8-1aでは、4バイトの最小ヘッダになっています。最小ヘッダにおいて最も重要な情報は次の通りです。

同期バイト。デコーダによって認識され、ヘッダとペイロードが非直列化されるようにします。

トランスポート・エラー・インジケータ。トランスポート・レイヤの上位にあるエラー訂正レイヤが、訂正不能なほど高いビット・エラー率(BER)を経験した場合に設定され、パケットにエラーが含まれている可能性があることを示します。エラー訂正レイヤの詳細については、「第10章 DVBおよびATSCの初歩知識」を参照してください。

パケット識別子(PID)。13ビットのコードで、パケットの種類を識別するのに使用されます。詳細については後述します。

連続性カウンタ。4ビットの値で、同じPIDのパケットが新たに送られるたびに、マルチプレクサによって値が増加されます。パケットの消失や繰返し、順序の乱れなどがないかどうかの判断に使用されます。

これよりも多くのヘッダ情報が必要な場合もあり、その場合には、アダプテーション・フィールドの制御ビットを設定して、ヘッダが通常よりも大きいことを示します。図8-1bのように、追加されたヘッダ長は、アダプテーション・フィールド長コードで表します。ヘッダが拡張された場合、一定のパケット長を維持するため、ペイロードは小さくなります。

8.3 プログラム・クロック・リファレンス(PCR)

番組に使用されるエンコーダには、プログラム・クロックが27MHzのものがあります。SDI(シリアル・デジタル・インタフェース)入力の場合は、ビット・クロックを10で割って、このエンコーダ・プログラム・クロックを発生させます。同じ制作施設で作られた番組が複数ある場合、それらのクロックがすべて同じということも考えられます。アナログ・ビデオ入力の場合は、H-SYNC(水平同期)期間に、位相ロック・ループの定数を掛けて、27MHzのクロックを発生させる必要があります。

パケット・ヘッダのアダプテーション・フィールドは、定期的使用されてPCRコードを加え、デコーダで固定クロックが生成されるようにします。エンコーダまたはマルチプレクサがソースを切替える際、PCRが不連続になることがあります。また、連続性カウンタが妨げられることもあります。このイベントは非連続インジケータによって処理され、障害が予期されることをデコーダに知らせます。それ以外の場合、不連続はエラー条件になります。

図8-2は、デコーダがPCRを使用して、番組ごとに27MHzクロックをリモートで再現する仕組みを示しています。エンコーダ・クロックは、常時実行される2進カウンタを駆動し、カウンタの値を定期的にサンプリングして、ヘッダのアダプテーション・フィールドにPCRとして置きます。PCRは42ビットの数で、33ビットのPCR基底部と、解像度を高めるための9ビットPCR拡張部で表されます(PCR基底部はPTSと同じく、33ビットの数字で、90kHzのクロックで駆動されるカウンタのサンプルです)。各エンコーダで生成されたパケットには、それぞれ別のPIDが付けられます。デコーダは選択した番組ごとに正しいPIDのパケットを認識し、それ以外を無視します。デコーダでは、電圧制御発振器(VCO)が公称27MHzのクロックを発生させ、これによってローカルのPCRカウンタを駆動します。ローカルのPCRは、パケット・ヘッダのPCRと比較され、その差分がPCR位相エラーになります。このエラーをフィルタリングしてVCOを調整することで、ローカルPCRのカウントが、ヘッダのPCRと一致するようにします。このVCOフィルタリングのレベルを高くすると、PCR伝送のジッタによってクロックが変調されなくなります。非連続インジ

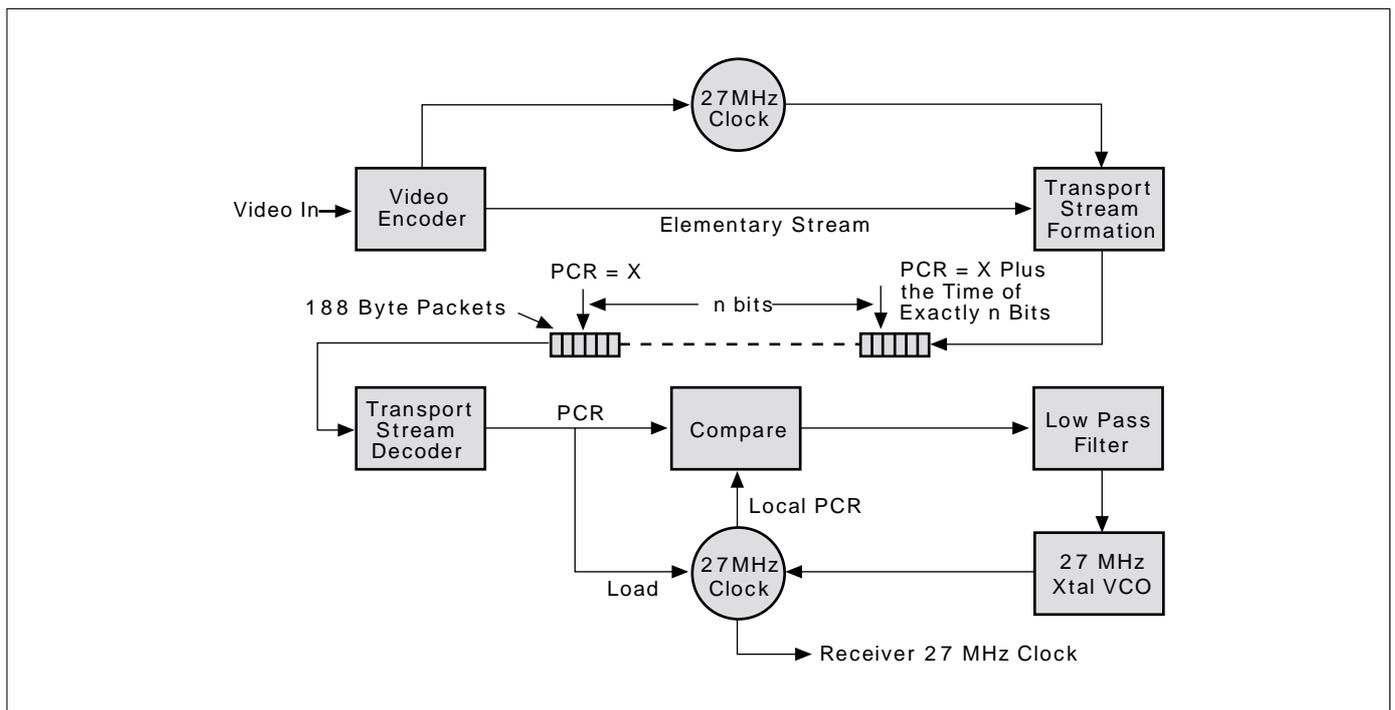


図8-2

データはローカルのPCRカウントをリセットしますが、これを使用してフィルタリングのレベルを下げ、システムが新しいタイミングに速やかにロックされるようにすることもできます。

MPEGでは、PCRの送信レートを最低10個/秒としていますが、DVBでは最低25個/秒と定めています。

8.4 パケット識別子(PID)

トランスポート・パケットのヘッダにある13ビットのフィールドには、パケット識別子(PID)コードが含まれています。PIDは、デマルチプレクサが、さまざまな種類の情報を含むパケットを識別するために使用します。トランスポート・ストリームのビット・レートは一定にする必要がありますが、ストリームのビット・レートの合計は一定にならない場合があります。このような場合は、ヌル・パケットを使用します。ペイロードの割合が低下した場合に、その分だけヌル・パケットが挿入されるようになります。ヌル・パケットのPIDは8191(2進表記で1が13個)で常に同じです。

トランスポート・ストリームでは、同じエレメンタリ・ストリームに属しているパケットのPIDはすべて同じになります。そのため、デマルチプレクサは、PIDが正しいパケットを受け入れるだけで、そのエレメンタリ・ストリームのすべてのデータを簡単に選び出すことができます。ビデオ/オーディオ/データ(字幕や文字放送など)のストリームのPIDを使用すれば、番組全体のデータを取り出せます。ただし、デマルチプレクサがパケットを正確に選び出すことができるのは、パケットを、属しているエレメンタリ・ストリームと正しく関連付けられる場合のみです。そして、正しく関連付けるには、デマルチプレクサが正しいPIDを認識していなければなりません。これはPSIの機能です。

8.5 番組特定情報(PSI)

PSIは固有のPIDを持つパケットで運ばれますが、これらのパケットには規格化されているもの他に、番組関連テーブル(PAT)、限定受信テーブル(CAT)、トランスポート・ストリーム記述テーブル(TSDT)などで指定されているものがあります。これらのパケットは、すべてのトランスポート・ストリームに定期的に含まれている必要があります。PATのPIDは常に0、CATのPIDは常に1、TSDTのPIDは常に2です。これらの値と、ヌル・パケットPIDの8191のみが、MPEG規格で定められているPIDです。それ以外のPIDについては、デマルチプレクサが該当するテーブルにアクセスして判断する必要があります。ただし、ATSCおよびDVBでは、PIDの使用について多少の制約があります。こうした点を含めて、MPEGとDVB/ATSCの間には、互換性が完全でない部分があります。DVBおよびATSCのトランスポート・ストリームはすべて、MPEG-2(ISO/IEC 13818-1)への適合が要求されていますが、MPEG-2のトランスポート・ストリームは、ATSC(A/65A)やDVB(EN 300 468)の規格にすべて適合しているわけではありません。

トランスポート・ストリームに存在する番組は、番組関連テーブル(PAT)パケット(PID=0)にリストされており、このパケットによって、各PMTパケットのPIDが運ばれます。PATの最初の項目である番組0はネットワーク・データ用に確保されており、ネットワーク情報テーブル(NIT)パケットのPIDが含まれています。NITの使用はMPEG-2では任意ですが、DVBでは必須です。

限定受信共通情報(ECM)および限定受信個別情報(EMM)のPIDは、限定受信テーブル(CAT)パケット(PID=1)にリストされています。

図8-3のように、同じ番組に属しているビデオ/オーディオ/データのエレメンタリ・ストリームのPIDは、番組マップ・テーブル(PMT)パケットにリストされます。通常、PMTごとに独自のPIDがありますが、これはMPEG-2では義務付けられていません。PMTに含まれている番組No.が、各PMTを一意に定義します。

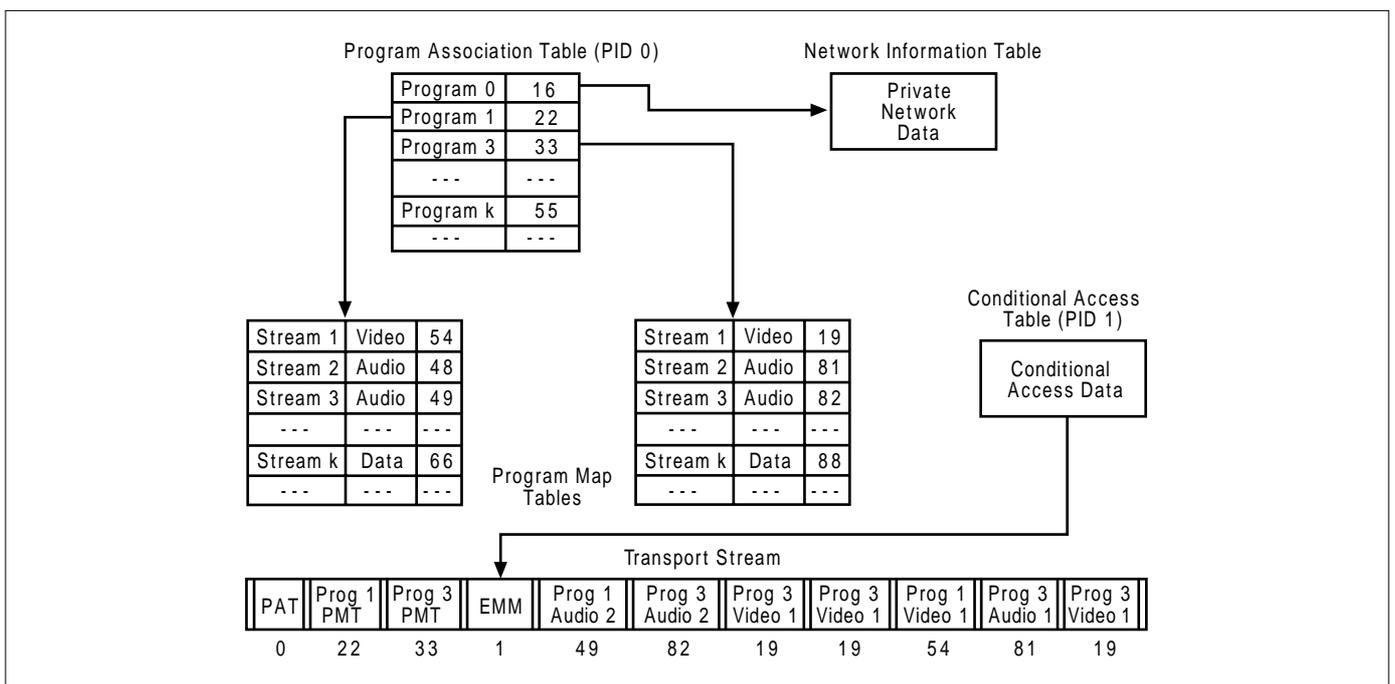


図8-3

第8章 トランスポート・ストリーム

ネットワーク情報テーブル(NIT)には、それを運ぶトランスポート・ストリームについての情報以外にも詳細な情報が含まれています。また、たとえば同一のデコーダにおいて、別のRFチャンネルに変えたり、パラボラ・アンテナを別の人工衛星の方に向けたりすることで得られる、他のトランスポート・ストリームの情報も含まれています。NITには他のトランスポート・ストリームが複数リストされていることがあり、それぞれに無線周波数、軌道位置などを指定する記述子が必要です。DVBでは、DVB-SIと呼ばれる追加のメタデータが含まれており、NITはDVB-SIの一部と見なされます。このオペレーションについては「第10章 DVBおよびATSCの初歩知識」で説明します。一般にこのテーマについて話す場合は、PSI/SIという用語が使われます。

トランスポート・ストリームを受け取り次第、デマルチプレクサは、パケット・ヘッダでPIDが0か1かを確認する必要があります。PIDがすべて0のパケットにはPATが含まれています。PIDがすべて1のパケットにはCATデータが含まれています。

デマルチプレクサはPATを読み取ることで、NITと番組マップ・テーブル(PMT)のPIDを見つけることができます。そしてPMTを見つけることで、各エレメンタリ・ストリームのPIDを見つけることができます。

そのため、ある番組を復号する場合、PATを参照し、続いてPMTを参照するだけで、その番組のエレメンタリ・ストリームにあるPIDをすべてを見つけることができます。番組が暗号化されている場合は、CATへのアクセスも必要になります。PATがないと逆多重化できないため、ロックアップ速度は、PATパケットが送られる頻度に関連して変化します。MPEGでは、PATパケットおよびそのPATパケットで参照されるPMTパケットの頻度について、間隔を最大0.5秒と定めています。DVBおよびATSCでは、NITは特定のPIDを持つパケットの中にあります。

第9章 デジタル変調

MPEGシステムは、映像、音声、その他のデータを符号化して圧縮します。得られたデータ・ストリームは、ハードディスク・ドライブやDVDなどに保存することができます。ある場所から別の場所へ、あるいは多くの場所へMPEGデータを送る場合、通常はトランスポート・ストリーム形式で送ります。テレビはケーブル、衛星、および地上波をさまざまな方法で使用して放送や配信を行います。いずれの伝送メカニズムでも、データを変調して搬送波(キャリア)に乗せる必要があります。本章では、こうした目的で使用するデジタル変調方式について紹介します。

本章では、変調技術のみについて説明します。データ・ストリームを伝送に適した状態にするのに必要な処理は他にもありますが、選択した変調方式やチャンネルの特性によって異なります。そうした技術については第10章で説明します。

9.1 変調の原理

搬送波は絶えず流れていますが、伝送する情報によって何らかの変化を加えない限り、情報は運ばれません。搬送波の変化は、振幅、周波数、位相のいずれかに対して行います。周波数と位相はいうまでもなく密接に関連しています。変調の種類として、両者は一般に別のものとして扱われますが、区別がはっきりしない場合もあり、「周波数変調」方式の中には「位相変調器」を使って実施されるものもあります。

9.2 アナログ変調

アナログでは、搬送波の振幅や周波数の変化(変調)は、オーディオ/ビデオ信号の振幅に基づき、通常は線形法則に従って行われます。また、位相変調もアナログ方式で使用されていますが、代表的な例としては、NTSC/PALテレビジョン方式での、色差信号から色サブ搬送波への変調があります。

この変調は非常にわかりやすい例です。色情報というのは2通りに考えることができます。まず色情報は、位相角(色相に關係)と振幅(飽和に關係)で表すことができます。また、2つの色差信号B-YとR-Yで表すこともできます。色差成分をx、yとしてグラフ上に表すと、この2つは、図9-1のように密接に関連しています。このグラフは、ベクトルスコープの表示にも似ています。

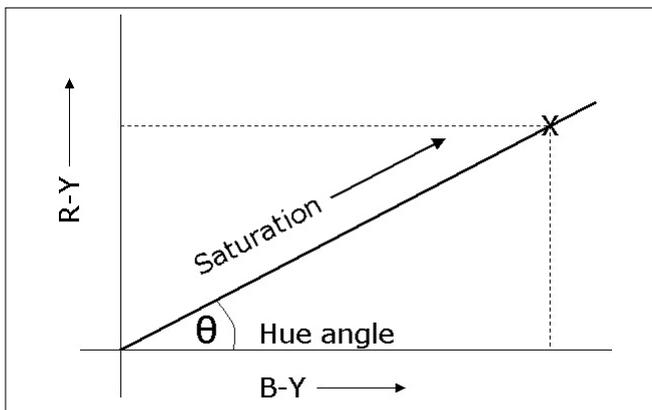


図9-1

9.3 直交変調

図9-2は、「直交変調」と呼ばれる技術を使い、2種類の信号によって搬送波を変調する方法を示しています。まず、1本の搬送波が2本に分割され、うち1本はその搬送波の周期の4分の1に相当する時間だけ遅延されます。これにより、周波数は元の搬送波と同じで、位相が90度偏移した搬送波が発生します。これら2本の搬送波はそれぞれ適切な信号によって振幅変調された後、合成されます。こうして、2つの変調信号の振幅で決められた振幅と位相を持つ信号が発生します。

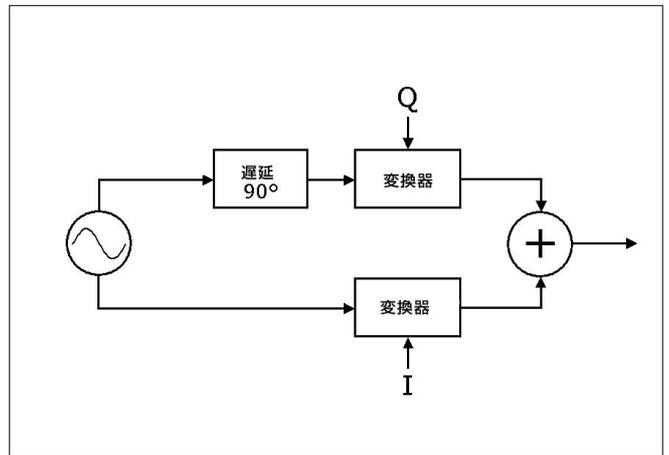


図9-2

復調もほぼ同じプロセスで行われます。まず、受信した信号を2つの復調器へ送ります。そしてそれぞれの復調器内で、その信号に同期化させた発振器信号を掛けます。この2つの発振器信号は90度離れており、各復調器は、直交変調の1本の軸を回復させます。

9.4 単純なデジタル変調方式

多くのデジタル変調方式では直交変調が使用されており、通常、2つの軸はそれぞれI軸、Q軸と呼ばれます。変調軸が1本だけの場合もあります。

デジタル変調方式では、周波数帯域の効率、堅牢性、および複雑性の間でバランスを取る必要があります。伝送信号の周波数帯域を決める際の最大要因はシンボル・レートです。シンボル・レートとは、変調を切替える速度のことで、変調信号の周波数帯域と同じになります。

単純なデジタル変調方式の中には、シンボルごとに運ばれる情報が1ビットしかないものもあります。つまり、シンボルごとに、2進0と2進1という2通りの状態が想定され、そのどちらかになります。この方式の場合、ビット・レートはシンボル・レートと同じになります。しかし、他の方式では、シンボルごとに想定される状態が多くあり、シンボルごとの情報も1ビットより大きくなります。通常、状態の数は2の累乗になるため、各方式のビット・レートは、シンボル・レートの整数倍になります。

デジタル変調方式の種類には、シンボルごとの状態の数を前に付けることがよくあります。たとえば、4QAMはシンボルごとに4通りの状態が想定される直交振幅変調を表します。4つの各状態は2ビットの情報(00、01、10、11)を伝えることができるため、4QAM方式のビット・レートはシンボル・レートの2倍になります。

最も単純な変調方式では、シンボルごとに1ビットの情報が伝えられます。シンボルごとに2進0と2進1という2通りの状態が想定されます。これらの状態は振幅、周波数、あるいは位相の変調によって得られますが、周波数と位相の変換が最も一般的です。

2進周波数偏移変調(BFSKまたは2FSK)では、ある搬送波の周波数で2進0を表し、別の周波数で2進1を表します。周波数の差が非常に小さい場合もあり、その場合は位相変調器を使用します。

2進位相偏移変調(BPSKまたは2PSK)では、搬送波(振幅は一定)の一方の位相を使って2進0を表し、その逆(180度の位相偏移)を使って2進1を表します。シンボルごとに想定されるさまざまな状態は通常、コンスタレーション・ダイアグラムで表され、変調信号IとQの許容値から得られるさまざまな組合せを示します。図9-3のように、BPSKのコンスタレーション・ダイアグラムは非常にシンプルあり、使われている軸は1本のみで、許容値も2つしかないため、この方式は非常に堅牢です。受信機は信号(または信号対ノイズ比)だけで、シンボルごとに、2つの状態のうちどちらが伝送されたかを判断できます。ただし、スペクトルの使用は効率的ではなく、周波数帯域がビット・レートと同じになります。この方式は深宇宙の遠隔測定のように、非常に困難な伝送パスで使用されます。

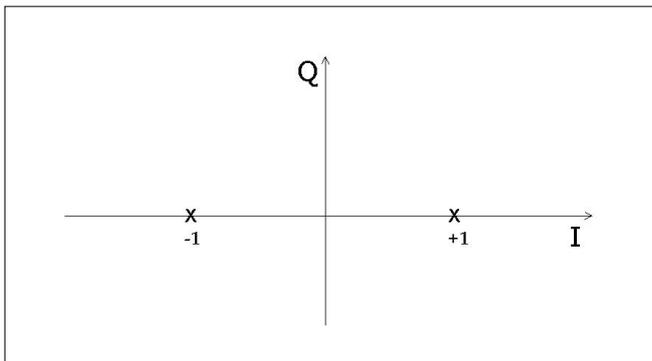


図9-3

9.5 位相偏移変調

前セクションでは、BPSK(2PSK)について説明しました。I軸とQ軸を使用するPSK変調は他にもあります。最も一般的なのは4位相偏移変調(QPSK、直交位相偏移変調とも呼ばれる)で、各軸で2つの値を使用します。図9-4にコンスタレーション・ダイアグラムを示します。QPSKはシンボルごとに4通りの状態が想定されるため、シンボルごとに2ビットの情報が送信されます。この図では、想定される状態を2進値にマッピングしています。QPSKは衛星通信で広く使用されています。

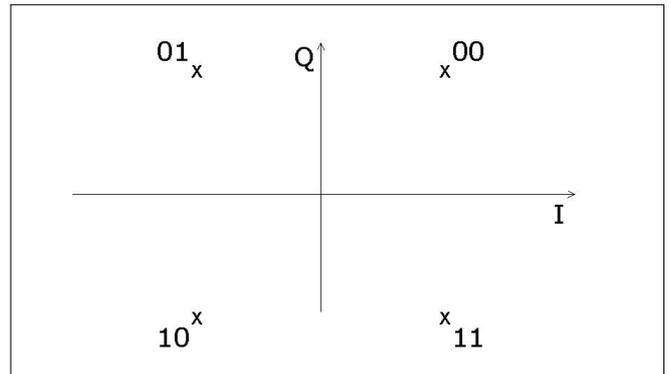


図9-4

8PSKはQPSKほど普及していませんが、特に日本の衛星通信などで使用されています。図9-5にコンスタレーション・ダイアグラムを示します。8PSKではシンボルごとに3ビットの情報が送信されるため、ビット・レートはシンボル・レートの3倍になります。

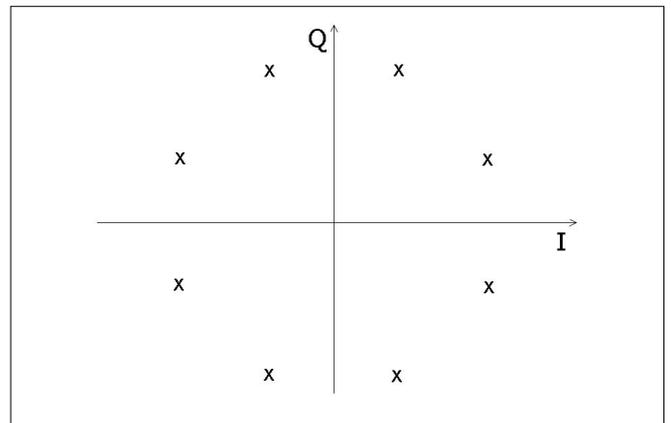


図9-5

9.6 直交振幅変調(QAM)

直交振幅変調(QAM)は、多くの伝送方式の基礎になっています。変調にはI軸とQ軸の両方が使用され、QAMの種類によっては、軸あたり2つ以上の振幅レベルが許容されます。

最も単純な4QAMでは、使用される値は軸あたり2つのみで、シンボルごとに4通りの状態が想定されます。図9-6にコンスタレーション・ダイアグラムを示していますが、4QAMはQPSKと似ており、シンボルごとに2ビットが送信されます。

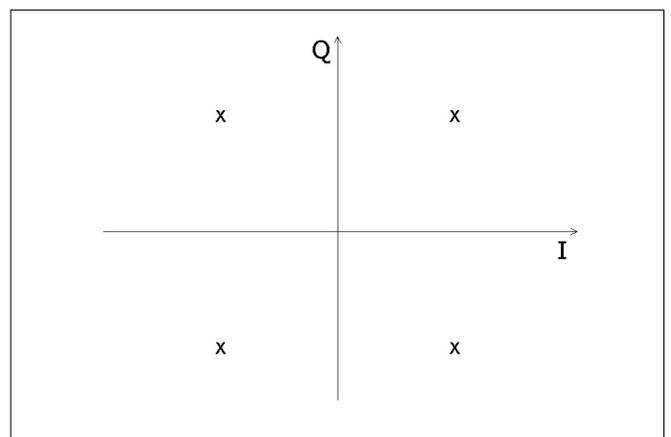


図9-6

16QAMでは軸あたり4つの値を使用し、16通りの状態が想定されます。また、シンボルごとに4ビットが送信されます。変調軸ごとに6つの値が許容される場合は、全体で36通りの状態が想定されます。また、5ビットを送信する場合は32通りの状態しか使用できず、使用されない状態が4つあります。図9-7に16QAM、図9-8に32QAMのコンスタレーション・ダイアグラムを示します。32QAMでは、4つの「角」にあたる状態は使用されませんが、この4つでは振幅が最も大きくなり、したがって送信機の電力も最大になります。

図9-7と図9-8は、ビット・レートと堅牢性の兼ね合いについても、わかりやすく示しています。ノイズやジッタがある場合、32QAMの状態の間隔を(同じ伝送電力のまま)狭めると、復号エラーが起こりやすくなります。また、シンボルごとに想定される状態が多くなると、エラー率を一定にするために必要な信号対ノイズ比が高くなるという見方もできます。

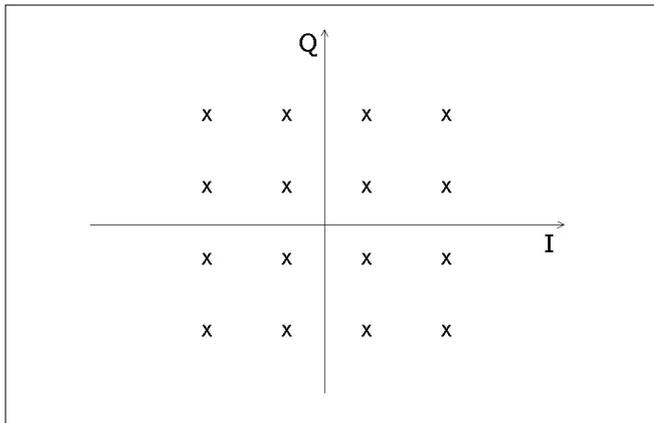


図9-7

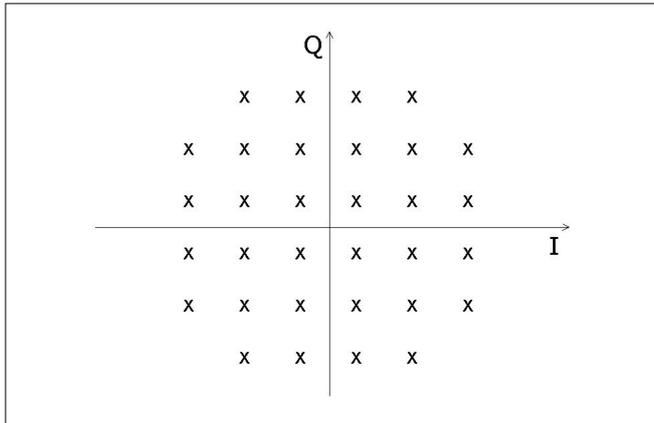


図9-8

高い信号対ノイズ比が確保できる場合は、さらに高い密度のコンスタレーションを使用することができます。64QAMでは、軸あたり8つの値を使用し、シンボルごとに6ビットを送信します。64QAMは、堅牢性をそれほど落とすことなく、従来のケーブル・インフラとの互換性も得られるため、ケーブルの変調方式として世界で最も広く使用されています。最新のケーブル・テレビ方式の一部で使用されている256 QAMは、変調軸あたり許容値が16個あり、シンボルごとに8ビットを送信します。

9.7 残留側波帯変調 (VSB)

搬送波を変調すると、その周波数の上と下に側帯波が発生します。たとえば、シンボル・レートが3メガシンボル/秒のQAM方式では、上と下の側帯波の幅はいずれも3MHzになり、6MHzのチャンネル周波数帯域が必要ということになります。

振幅と位相情報の両方(直交変調の両軸)を回復するには、受信機で両方の側帯波を回収する必要があります。

残留側波帯方式では、一方の側帯波は伝送前にほとんど除去されてしまうため、回収できる変調軸は1つだけです(一方の変調軸は、好ましくない側帯波を抑えるために使用されたという解釈もできます)。2VSBのコンスタレーションはBPSKと同じになります。図9-9に4VSB、図9-10に8VSBのコンスタレーション・ダイアグラムを示しており、それぞれシンボルあたりのビット数は2ビットと3ビットです。

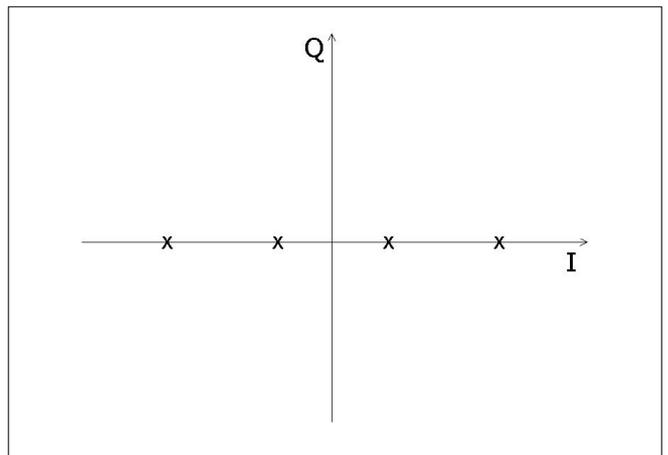


図9-9

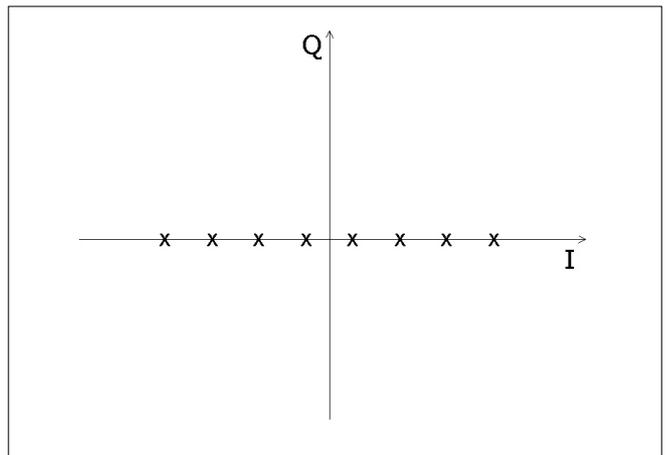


図9-10

8VSB変調は、米国のATSCデジタル・テレビ規格で採用されています。当初はシンボルあたり2ビットの4VSBが提案されましたが、8VSBは、トレリス符合化(セクション10.6を参照)および畳み込み内部エラー訂正と組合せることで、同じデータ・レートのまま、信号対ノイズ比が向上することが明らかになりました。

9.8 符号化直交周波数分割多重(COOFDM)

前述の方式では、搬送波が1本(シングル・キャリア)の変調器にベースバンド信号が供給されることで側帯波が発生しています。広帯域方式に代わるものとして、間隔が綿密に調整された多数の狭帯域搬送波を発生させる方式があります。図9-11aでは、デジタル変調した搬送波の両側がスペクトル的にヌルになっています。また図9-11bのように互いに直交させれば、同じ搬送波を加えても干渉が起きません。これはOFDM(直交周波数分割多重)の原理を利用しています。実際、OFDMの一種であるCOOFDM(符号化直交周波数分割多重)では、次章で述べるように、ビタビ量み込み符合化を使用することで、理想的でないチャンネル状態でも、パフォーマンスが大幅に向上します。COOFDMは、DVBの地上波方式(DVB-T)で採用されています。

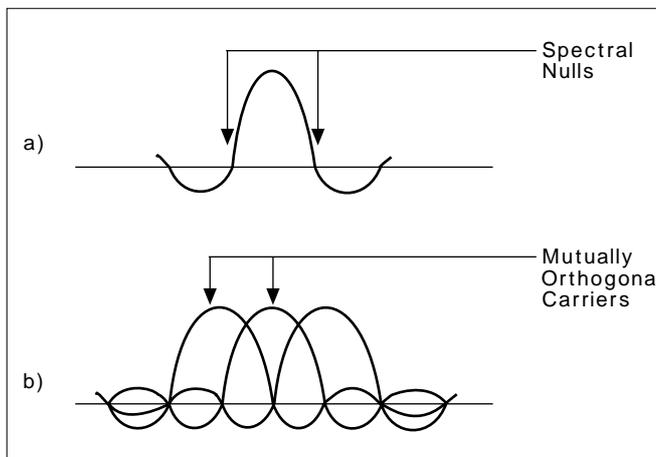


図9-11

OFDM方式の搬送波は、本章で述べているどの方法でも変調できますが、実際にはQAMが使用され、16QAMと64QAMが一般的です。搬送波をすべて使用する必要はありません。たとえば、チャンネルの一部が高レベルの干渉を受けることがわかれば、影響を受ける搬送波を省いても構いません。

OFDMでは搬送波の数が非常に多くなります。そのためDVB-Tでは、搬送波1705本と6817本に代わるオプション(それぞれ2k、8k)があります。各搬送波に割当てられる周波数帯域は小さいため、それに呼応してシンボル・レートも低くなり、シンボルあたりの伝送時間が長くなります。これが、マルチパス干渉に対するOFDMの耐性のカギになっています。8VSBのようなシングル・キャリア方式では、データ・レートが高いと、シンボル時間は非常に短くなります。ATSCデジタル・テレビ方式の場合、毎秒およそ1100万個のシンボルが伝送されるため、シンボル1つの所要時間は100ns未満になります。そのため、マルチパスにおいて、あるシンボルの信号が遅れて後のシンボルの受信時に到着すれば、非常に短い遅延であってもシンボル間の干渉を起こします。

それに対し、搬送波が数千本のOFDM方式では、シンボル時間は数百マイクロ秒になります(データ・レート、搬送波の本数、使用する変調方式などによって異なります)。各シンボルに「ガード・バンド」を加えて、意図的にシンボルを必要以上に長くすることで、シンボル間の干渉はほとんど解消されます。この処理によって、シンボル・レートはごくわずかですが低下します。たとえば、シンボルの所要時間が200 μ sで、ガード・バンドが50 μ sだとすると、シンボル・レートの低下は20%にすぎず、シンボル間の干渉を解消してより上位のコンスタレーションを使用できることを考えると、損失よりも利益のほうが大きいといえます。ただし、この方法は広帯域のシングル・キャリア方式では実用的ではありません。たとえば、ガード・バンドが同じ50 μ sでも、シンボル時間

が100nsでは、データ・レートはほとんど低下しません。

COFDM方式にはこうしたマルチパス干渉に対する耐性があるため、同期化された2つ以上の送信機が同じ信号を出す、単一周波数のネットワークにも適しています。受信機は、場所やアンテナ方式に応じて、複数の送信機からの信号を異なる時間に受信できます。パス長が極端に違う場合、おそらくメイン信号は2次信号よりもはるかに強くなり、干渉は最小限になります。パス長と信号強度が同じである場合は、ガード・バンドによってシンボル間干渉を防ぎます。

COFDM方式は非常に柔軟性が高く、多様な伝送の要件に合わせて「チューニング」することもできますが、常にいえるのは、堅牢性を高めるとデータ・レートが犠牲になるということです。シングル・キャリア方式とマルチ・キャリア方式の相対的なメリットについてはさまざまな議論がありますが、一般的には、単純なチャンネル条件でサービス・エリアとデータ・レートが同じであれば、COFDMはVSBよりも多くの電力が必要になると考えられています。また、COFDMは送信側でのピーク対平均電力比も高く、他のサービスへの干渉も多くなります。

しかし、大都市中心部の「ビル谷間」のように複雑なマルチパスの状況では、COFDMのほうがしっかりした受信状態が得られると考えられています。

9.9 統合デジタル放送(ISDB)

統合デジタル放送(ISDB)は、日本のデジタル・テレビ・サービス用に開発された規格で、いくつもの変調方式が使用されます。ISDBは、多くのレベルからなる階層システムに対応できるようになっています。たとえば、移動体通信の特に困難な受信条件には低いデータ・レート、受信不良地域には中程度のデータ・レート(SDTV)、良好な受信条件には高いデータ・レート(HDTV)を、同時に提供することができます。

日本では現在、3種類のISDB変調方式が使用されています。

9.9.1 ISDB-S衛星放送方式

2000年12月から運用が始まったISDB-Sにより、2つの放送事業者で衛星トランスポンダを共用することが可能になりました。ISDB-Sは、放送衛星ではBSデジタル、通信衛星ではCSデジタルとも呼ばれています。

トランスポンダを共用する事業者間で合意した方法で、最高8本のトランスポート・ストリームを一緒に使用することができます。総ビット・レートは、トランスポンダの周波数帯域と使用する変調モードによって異なります。たとえば、34.5MHzのトランスポンダの場合、最大レートは、前方エラー訂正を含めて56.610Mbpsになります。

階層変調を使用すると、48パケットのフレーム内において、パケット単位で変調モードを変えることができます。各パケットには変調スロットが割当てられます。BSPK(1/2)、QPSK(7/8)、TC8PSKの変調モードに対応しています。スロットの数は、使用するモードによって異なります。

9.9.2 ISDB-Cケーブル放送方式

この方式の最も大きな特徴は、複数のトランスポート・ストリームを1本の64QAM搬送波で伝送する点です。この方式は、ISDB-S信号で運ばれてきた情報を効率よく伝送するために開発されました。通常、BSデジタル搬送波の情報伝送量は最大で52.17Mbpsです。64QAM/6MHz信号の情報レートは29.162Mbpsです。したがって、BS搬送波の情報を転送するには、ケーブル・テレビ・チャンネルが少なくとも2つ必要になります。BSデジタル・サービス全体は4つの放送事業者で構成され、ガード・バンドを含めて約174MHzを使用しています。これだけの情報を運ぶのに、ISDB-Cであればケーブル・チャンネルが8つで済むのに対

し、1搬送波あたり1トランスポート・ストリームという従来のケーブル伝送であれば、29チャンネルも必要になります。

TSMF(トランスポート・ストリーム多重化フレーム)には、52の変調スロットと、同期化ヘッダ用の1スロットがあります。

9.9.3 ISDB-T地上波変調

ISDB-Tチャンネルは13のセグメント(通常、400~500kHz)に分割され、セグメントごとに個別のCOFDM伝送を行います。堅牢性に影響を及ぼすパラメータ(搬送波の本数、ガード・バンド長、変調の種類、畳み込み符合化)はすべて、階層のレイヤごとに選択できます。たとえば、最も堅牢なセグメントでは、長いガード・バンド、QPSK変調、1/2畳み込み符合化を使用することもあります。最高レベルでは、短いガード・バンド、64QAM、7/8畳み込み符合化を使用し、堅牢なセグメントの何倍ものデータ・レートを実現することができます。

中間セグメントは、狭帯域の受信機でこのセグメントのみを受信するように設計された、部分受信に使用します。

通常のOFDMでは、独立したチャンネル周波数帯域が、1つのレイヤを表します。使用される搬送波は、ある周波数の一定の倍数の間隔で、その帯域全体に配置されます。ISDB-Tでは、5.6MHzのチャンネル周波数帯域をそれぞれ429kHzの13のセグメントに分割します。ISDB-Tでは、伝送パラメータが異なるOFDMセグメント・グループを送ることで、階層伝送を行います。こうしたレイヤのグループがレイヤを構成しています。非階層変調では、13セグメントすべてに対して、同じ変調方式を使用します。

9.9.4 ISDBのまとめ

ISDB-Sは、複数のトランスポート・ストリームに対してさまざまな変調モードを適用し、1本の搬送波に乗せて34.5MHzで伝送する手段です。

ISDB-Cは、複数のトランスポート・ストリームを1本の搬送波に乗せて6MHzの単一チャンネルで伝送する手段で、変調モードは共通です。

ISDB-Tは、6MHzの単一チャンネルの同じトランスポート・ストリームを、最大3つのレベルで伝送します。

第10章 DVBおよびATSCの初歩知識

MPEG圧縮はすでに放送で使用されており、今後も重要性は高まります。本章では、2つの主要なデジタル・テレビ放送規格で実施されている追加的な要件について説明します。

10.1 概要

ATSC(Advanced Television Systems Committee)は、地上波デジタル放送の規格を定めた米国の機関です。DVBは、Digital Video Broadcastingプロジェクト、およびこのDVBプロジェクトによって定められた規格や基準を指します。本来は欧州の団体ですが、作成された規格やガイドは世界の広い地域で採用されています。これらの規格やガイドは、衛星、ケーブル、および地上波の放送を含むあらゆる通信メディアをカバーしています。

図10-1のように、デジタル放送では、配信や伝送の要件はさまざまです。放送事業者は、複数のテレビ番組を含むトランスポート・ストリームを作成します。トランスポート・ストリームではエラーに対する保護がないため、圧縮データでは、エラーは深刻な影響を及ぼします。トランスポート・ストリームは、送信機、衛星アップリンク、およびケーブル・ヘッド・エンドヘエラーのない状態で届ける必要があります。この場合、エラーがないとは、ビット・エラー率(BER)が 10^{-11} 以下の状態をいいます。通常、この役割は通信ネットワーク事業者に委ねられ、追加のエラー訂正レイヤが必要に応じて使用されます(エラー訂正の手法は伝送チャンネルに応じて選択)。このレイヤは送信先に対して透過的であるようにします。

送信機やケーブルの事業者ごとに、トランスポート・ストリームで送られる番組がすべて必要というわけではありません。リマルチプレクサを使用すれば、複数のトランスポート・ストリームを受け取ってチャンネルを選択し、1本のトランスポート・ストリームに再多重化して出力することができます。構成をダイナミックに変えることもできます。

デジタル領域の放送では、トランスポート・ストリーム全体が視聴者に送られます。チャンネルがケーブル、衛星、地上波のいずれにおいても問題はほぼ同じです。そのため、伝送内容について記述したメタデータを標準化された方法で符合化して、トラン

スポート・ストリームで送る必要があります。DVBではこのメタデータを番組配列情報(DVB-SI)と呼び、そのストリーム自体や他の多重化ストリームで送られている番組の詳細や、文字放送などのサービスが含まれます。

放送時には信号品質を十分に管理できないため、ノイズや干渉が起きる可能性があります。そのため、エラー訂正(FEC)レイヤが必要になります。通信ネットワーク事業者が使用している独自開発の(あるいは、ヨーロッパ電気通信規格協会(ETSI)がSDH/PDHネットワークにおけるDVB伝送について規格化した)FECとは異なり、放送で使用するFECは、受信機で処理できるように標準化されていなければなりません。

送信機やケーブルに関する限り、エラー訂正を加えれば、ビット・レートは間違いなく増加します。データを直列化するだけでは、経済的で信頼性のある無線/有線でのデータ伝送はできません。実際のシステムではチャンネル符合化が必要になります。

10.2 再多重化

再多重化は、他のビット・ストリームを組合せて適合ビット・ストリームを出力するという複雑なプロセスです。必要な番組のPIDが示される番組関連テーブル(PAT)や番組マップ・テーブル(PMT)を参照することで、入力トランスポート・ストリームから必要なデータを選び出します。2つの入力トランスポート・ストリームで同じPIDが使用されている可能性もあるため、いくつかのエレメンタリ・ストリームのPIDを変更しなければならない場合もあります。また、終端のデコーダが27MHzクロックを再現するため、パケット・ヘッダでプログラム・クロック・リファレンス(PCR)を送る必要があります。新たに多重化したストリームではPCRを含むパケットの位置が変わることがあるため、リマルチプレクサがPCR値を編集して、新しい位置を時間軸上に反映させる必要があります。

限定受信テーブル(CAT)と同じように、PMTおよびPATを編集して、新しいトランスポート・ストリームの構造を反映させる必要があります。

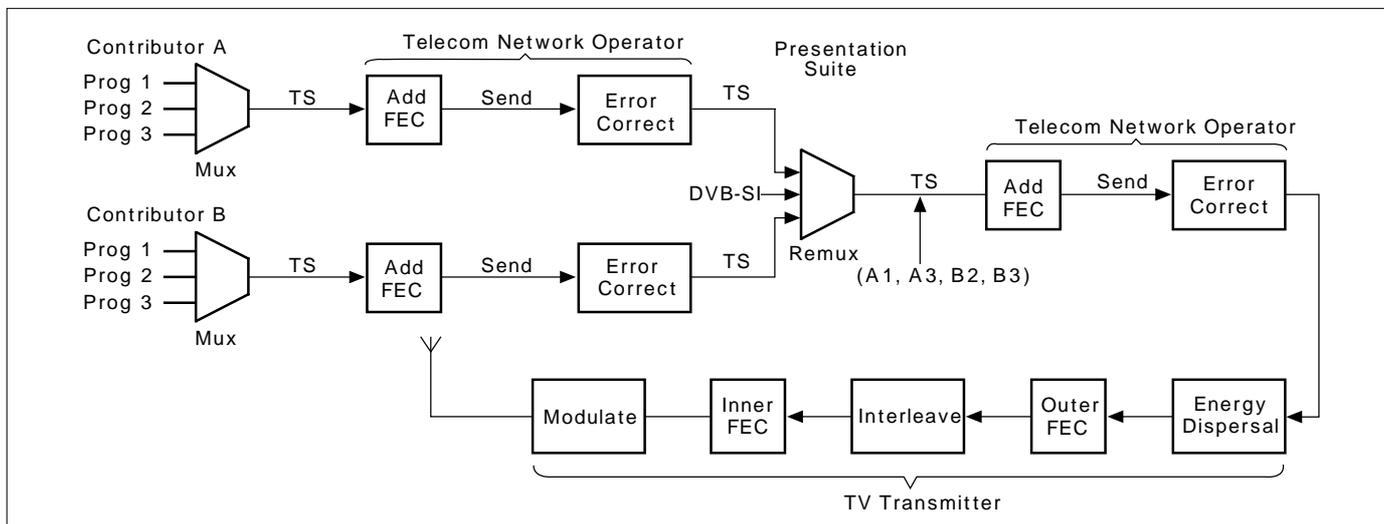


図10-1

選択したプログラム・ストリームのビット・レートの合計が、出力ビット・レートよりも小さい場合、リマルチプレクサは、適切なPIDのスタッフィング・パケットを生成します。しかし、統計マルチプレクサからのトランスポート・ストリームの場合、新しいトランスポート・ストリームの瞬間ビット・レートがチャンネル容量を超えてしまう可能性があります。異なるトランスポート・ストリームから選択された複数の番組のエントロピーがどれも高いと、こうした状況が発生します。この場合は、再圧縮してビット・ストリームに短い係数を作成することでビット・レートを削減するのが唯一の解決策です。

10.3 番組配列情報(SI)

今後、デジタル送信によって、多くの番組や文字放送などのサービスが視聴者に提供されるようになりますが、それらは複数のトランスポート・ストリームに分散している場合があります。そのため、視聴者側のデジタル放送受信機(IRD)で、サービス内容の表示や、選んだサービスの出力を簡単にできるようにする必要があります。こうした機能はMPEG-PSI(番組特定情報)の範囲ではなく、DVB-SI(番組配列情報)と呼ばれるメタデータが使用されます。DVBでは、このSIにNITも含まれると見なされますが、MPEGトランスポート・ストリームでは、NITの使用は任意です。

DVB-SIは、固有のPIDを持つ追加のトランスポート・パケットとしてトランスポート・ストリームに埋め込まれ、IRDに必要な技術情報を運びます。また、番組の性質や時間、その番組があるチャンネル、利用できる国といった電子番組ガイド(EPG)情報も含まれています。また、番組をレーティング(ランク付け)して、保護者が判断できるようにすることも可能です。

MPEG-PSIの範囲を超える機能について、DVB-SIでは次のようなテーブルが必要です。

ネットワーク情報テーブル(NIT)、1つのトランスポート・ストリームに含まれ、多数のトランスポート・ストリームについて記述した情報。ネットワークで送られる多重化トランスポート・ストリームの物理的な構成や、そのネットワーク自体の特性に関連する情報を伝えます。トランスポート・ストリームは、元のネットワークIDとNITのトランスポート・ストリームIDの組合せで識別されます。

サービス記述テーブル(SDT)、DVBトランスポート・ストリームのサービスにはそれぞれサービス記述子があり、これらの記述子をまとめたのがサービス記述テーブルです。サービスとは、テレビ、ラジオ、文字放送などを指します。サービス記述子には、サービス・プロバイダの名前が含まれています。

イベント情報テーブル(EIT)、DVB用のテーブルで、番組名、開始時間、所要時間などが含まれています。

日時テーブル(TDT)、UTC(協定世界時)の日時スタンプをトランスポート・ストリームに埋め込むためのテーブルです。

DVB-SIでは、その他にも任意のテーブルとして、ブーケ関連テーブル(BAT)、進行状態テーブル(RST)、時間オフセット・テーブル(TOT)、スタッフィング・テーブル(ST)などを定義しています。

ATSCもDVBと同様に、MPEG-2専用セクション・テーブルを使って、新たなテーブルを定義しています。ATSCがA/65Aで定義したこれらの新しい必須テーブルは、Program and System Information Protocol(PSIP)の一部になります。MPEG-PSIの範囲を超える機能について、ATSC PSIPでは次のようなテーブルが必要です。

地上仮想チャンネル・テーブル(TVCT)、少なくとも、このテーブルのあるトランスポート・ストリームに埋め込まれるMPEG-2番組を定義します。

マスタ・ガイド・テーブル(MGT)、トランスポート・ストリームにある、システム時間テーブル(STT)を除くすべてのPSIPテーブルについて、種類、パケット識別子、バージョンを定義します。

レーティング地域テーブル(RRT)、トランスポート・ストリームで送られるコンテンツ通知記述子で参照する、テレビ保護者ガイドライン・システムを定義します。

システム時間テーブル(STT)、現在の日付と時間を定義します。イベント情報テーブル(EIT-n)、12時間のイベント(テレビ番組)について記述した4つのイベント情報テーブル(EIT-0、EIT-1、EIT-2、EIT-3)を順に定義します。4つのテーブルは、それぞれ3時間ずつカバーし、TVCTにリストされた仮想チャンネルはすべて含まれます。

10.4 エラー訂正

伝送パスが長いと状態を管理できないため、エラー訂正が必要になります。エラー検出だけで、再伝送を要求できる方式もあります。しかし、この方法はテレビのようなリアルタイムの信号では有効ではありません。FECでは、冗長性と呼ばれる十分な量の追加ビットをデータに加えることで、デコーダがリアルタイムで訂正できるようにします。

最新の方式で使用されているFECは、通常、リード・ソロモン(R-S)コードをベースにしています。このコードについての詳しい説明は本書の範囲外ですが、簡潔に言えば、データへの冗長性によってコード・ワードに変更を加え、各シンボルを少なくとも2つの連立方程式の項として使用した場合に、エラーがなければ、SUM(シンドローム)が常にゼロになるようにすることです。このゼロ状態は、データに関係なく得られるため、チェックが簡単になります。トランスポート・ストリームでは、エラー訂正データを加える前のパケット長は常に188バイトです。16バイトのR-S冗長性を加えると、パケット長は204バイトになります(実際にはFECが行われない場合でも、204バイトのパケットがトランスポート・ストリームで送られます。この16スタッフィング・バイトによって、FECの追加や削除時におけるストリームの再ロックが防止されます)。

シンドロームがゼロ以外の場合、連立方程式を解くと、エラー訂正に必要な、エラーの場所とエラーの性質という2つの値が得られます。しかし、追加した冗長性の半分を超えるサイズのエラーは訂正できません。一般的な伝送チャンネルでは、信号品質は統計的に表されます。つまり、1つ1つのビットがノイズによってエラーになる場合もあれば、バーストのように多数のビットが一度に破壊されることもあります。なお、こうした破壊は、落雷や他の電気機器による干渉などが原因です。

バーストはそれほど頻繁に発生するものではないため、すべてのコード・ワードを保護するのは非経済的です。解決策として、

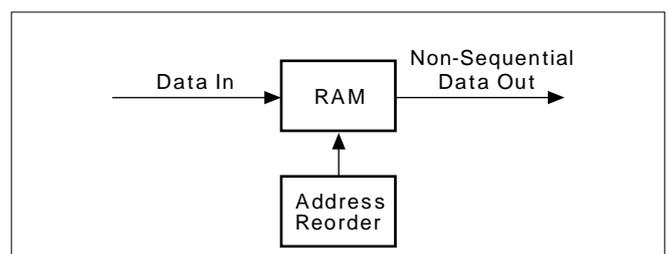


図10-2

インターリーブと呼ばれる技術を使用します。図10-2のように、インターリーブを使用すると、ソース・データは送信前にFEC符号化され、RAMバッファに入れられます。図10-3では、データを行でRAMに入れ、列で読み出すという方法を示しています。並べ替えられたデータは伝送され、受信時に2つ目のRAMによって元の順序に戻される、つまりデインターリーブされます。デインターリーブ後は、バーストが多数のシングル・ビット・エラーになるため、簡単に訂正できます。

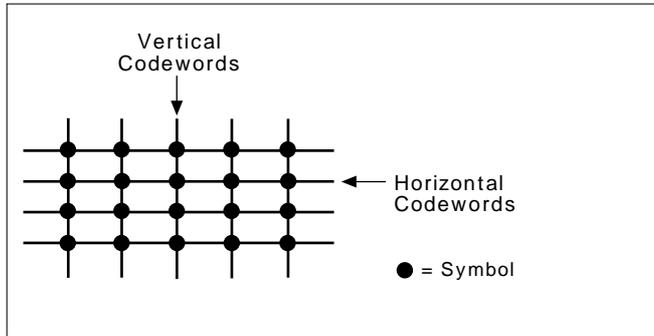


図10-3

バースト・エラーのサイズが訂正可能な限度に近づくと、コード・ワードを訂正不能にするランダム・ビット・エラーが起きやすくなります。インターリーブ後に適用され、デインターリーブ前に訂正された内部コードを使用することで、ランダム・エラーがデインターリーブ・メモリに入るのを防ぐことができます。

図10-3のように、ブロック・インターリーブ構造でこの方法を使用すると、プロダクト・コードが得られます。また、図10-4のように、各行に異なる量の遅延を加えてデータ配列を傾斜させる、畳み込みインターリーブを行うこともできます。この畳み込み、つまりクロス・インターリーブには、インターリーブおよびデインターリーブに必要なメモリを少なくできるという利点があります。

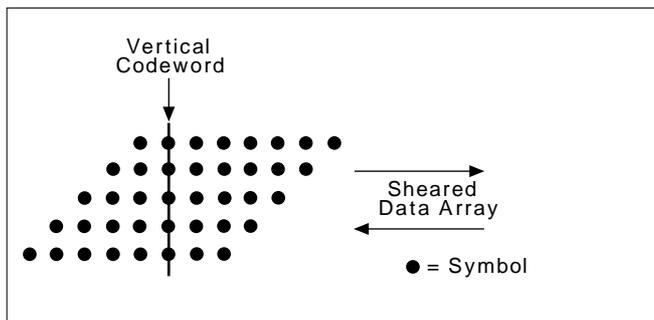


図10-4

10.5 チャンネル符号化

未処理のシリアル2進データは伝送に適しておらず、これにはいくつかの理由があります。同じビットを何度も実行するとDCオフセットを起し、ビット・クロックが不足します。また、スペクトルは管理されておらず、必要な周波数帯域が非常に大きくなります。実際の無線/有線のシステムでは、チャンネル・コードと呼ばれる変調方式が必要になります。デジタル変調方式については、第9章で説明しています。図10-5では、この原理を64QAM変調器に応用しています。

上記の方式では、信号の伝送スペクトルは信号ごとに異なります。スペクトルには、高エネルギーが含まれていて他のサービスへの干渉の原因になる部分と、エネルギーがわずかであるために

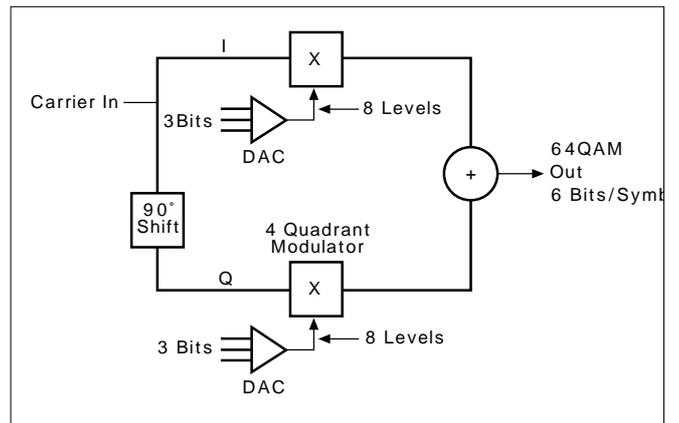


図10-5

他からの干渉を受けやすい部分とがあります。実際には、ランダム化によって、伝送スペクトルとデータ・コンテンツとの相関をなくす必要があります。図10-6のように、ランダム化によるエネルギー分散を行うと、疑似ランダム・シーケンスがシリアル・データに加えられてから、変調器に入力されます。その結果、統計的にかなり定常な、ノイズのようなスペクトルとして送信されます。そして同一の同期シーケンスは、図に示したように受信側で取り除く必要があります。ランダム化は同期パターンには適用できず、適用しても検出することができません。

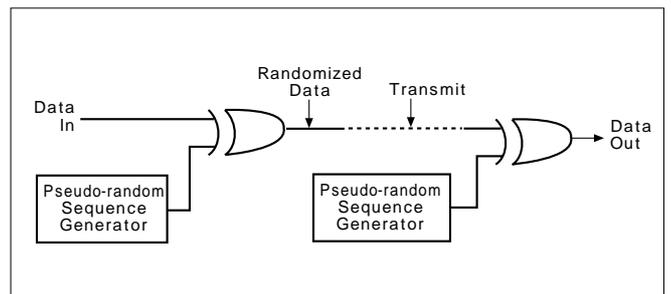


図10-6

10.6 内部符号化

FEC符号の内部符号化は、ランダム・エラーによってインターリーブ方式の効果が低下するのを防ぎます。適切な内部コードを使用すれば、伝送のSNRが確実に高まり、ランダム・エラーが防止されます。マルチレベルの信号方式とともに使用できるトレリス符号化では、複数のマルチレベル・シンボルを関連付けてグループ化します。こうしたシンボルのグループから得られる波形をトレリスと呼びます。各シンボルに8つのレベルがある場合、3つのシンボルからなるグループでは512通りのトレリスが可能です。

トレリス符号化では、特定のトレリス波形のみが有効なデータを表すようにデータを符号化します。64種類のトレリスのみがエラーのないデータを表すとすると、シンボルごとに送ることのできるデータは、3ビットではなく2ビットになります。64のトレリス以外はエラーが原因しているため、残りのビットは一種の冗長性になります。シンボルのレベルの1つがノイズによって曖昧なトレリスを受け取っても、レベルが正確でなければ有効なトレリスは得られないはずなので、この曖昧さは解決します。この手法を最尤復号と呼びます。

低い信号対ノイズ比でシステムの動作を続けるには、64種類の有効トレリスをできる限り異なるものにします。トレリス・コードがエラーを起こした場合は、外部のコードによって訂正されます。

DVBでは、Viterbi畳み込み符号化を使用することができます。図10-7では、インターリーブされたデータはシフト・レジスタへ送られます。そして、シフト・レジスタは入力データに対するパリティ・チェックの値を2つ出力することで、ビット・エラーを訂正できるようにします。1ビット入力されるごとに、2ビットが出力されるため、図のコードはレートが1/2のコードとして表されています。1/1から1/2のレートであれば、元のデータを伝送することができますが、冗長性の量は変動します。1/2出力の全体を伝送しないことを「間引き」といい、これにより、ビット・レートと訂正機能との間で必要なバランスを得ることができます。

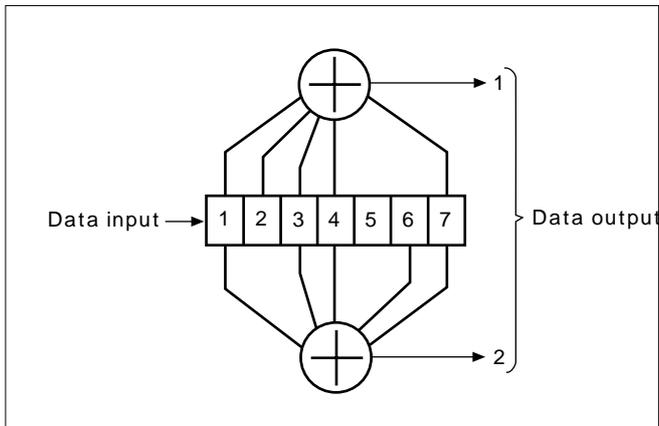


図10-7

10.7 伝送桁数

図10-8は、ATSCデジタル送信機の構成要素を示しています。トランスポート・ストリームには、伝送内容について記述した番組配列情報が加えられます。このストリームはランダム化された後に、外部のR-Sエラー訂正コードに送られ、そこでデータに冗長性が追加されます。続いて、畳み込みインターリーブ処理でデータを並べ替えることで、トランスポート・ストリームで隣り合うデータが伝送で隣り合わないようにします。そして、内部のトレリス・コードを使用して、マルチレベル信号を発生させ、残留側波帯(VSB)変調器に送ります。

図10-9は、DVB-T送信機を示しています。前図と同じように番組配列情報が加えられ、続いて、ランダム化によってエネルギーを分散させます。続いて、外部のR-Sチェック・シンボルが加えられ、インターリーブされます。インターリーブの後、内部の符号化プロセスを実行し、符号化したデータをCOFDM変調器に送ります。そして、変調器からの出力をアップコンバートしてRFを出力します。

受信機では、ビット・クロックが抽出され、システム全体のタイミングを制御するために使用されます。チャンネル符号化を逆にたどって、生データと伝送エラーを得ます。内部コードは、ランダム・エラーを訂正し、大きいエラーを識別することで、デインターリーブ後の外部コードのプロセスを助けます。ランダム化が解除され、元のトランスポート・ストリームになります。受信機は、PAT、番組配列情報(SI)およびPATが指示しているPMTを識別できなければならず、それによりトランスポート・ストリームの内容が視聴者にわかり、選んだ番組を多重化ストリームの中で見つけることができます。

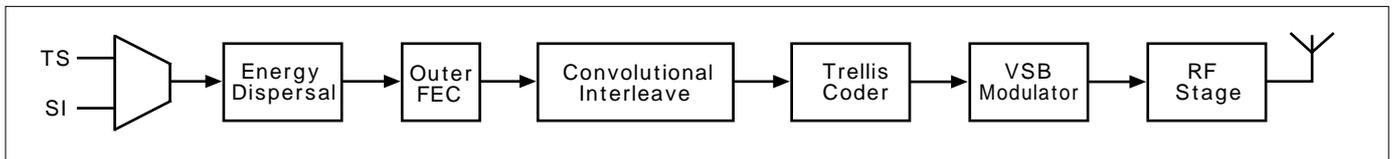


図10-8

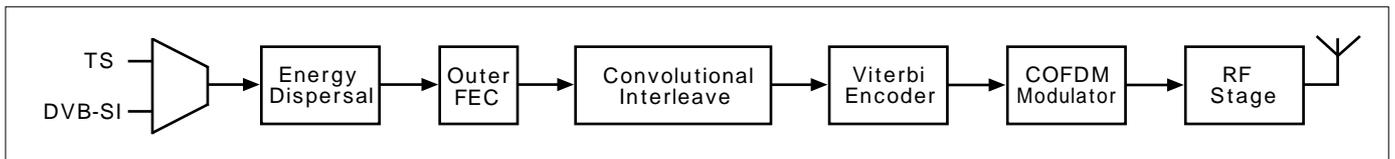


図10-9

第11章 データ放送

これまでの章では、MPEG-2トランスポート・ストリームの基礎と、圧縮した映像や音声のストリームを従来のアナログ放送のように伝送する主要なアプリケーションについて見てきました。しかし、MPEG-2トランスポート・ストリームの重要な強みの1つは、映像や音声だけでなくデータも伝送できることです。アナログ・テレビ方式でもデータ伝送は使用されていますが、トランスポート・ストリームと比べると、利用可能なデータ周波数帯域が著しく制限されます。

本章では、トランスポート・ストリームでのデータ・カプセル化のために、MPEG-2が提供するさまざまな方法と、地域的な標準化(DVB、ATSC、ARIB(電波産業会))についての概略を述べます。そして、次のセクションでは、データがセット・トップ・ボックスによって視聴者に表示される仕組みについて説明します。

11.1 アプリケーション

トランスポート・ストリームでのデータ放送には、さまざまな種類のアプリケーションがありますが、それぞれに必要なデータの種類の異なり、タイミングの要件も異なります。

たとえば、インターネット送信に必要なデータの種類の、セット・トップ・ボックスのファームウェア更新に必要なデータとは大きく異なります。また、価格情報の非リアルタイム更新におけるタイミングの要件は、問題と回答を映像/音声と緊密に同期化させて送らなければならないクイズのアプリケーションとは大きく異なります。

MPEG-2は、非常に多彩なデータ送信方法を提供します。どの方法にするかについては、アプリケーションのタイミング要件を満たしつつ、周波数帯域の最適化(と、それによって決まるサービス提供のコスト)のバランスを考えて決める必要があります。

以降のセクションで述べますが、アプリケーションの種類は、リアルタイムの要件とインタラクティブ性のレベルによって、大まかにグループ分けできます。

11.1.1 プログラム関連データ

MPEG-2の基本規格では、テレビ・チャンネルや、トランスポート・ストリームで送られる個々の番組の情報を提供するEPGについて一切規定していません。地域的な標準化団体であるATSC、DVB、ARIBでは、EITのようなMPEG-2専用テーブル構造を使用して、番組に関する追加のデータを提供してきました。こうしたテーブルは、一定の間隔で送信し、開始時間、番組の概要などのサービスや番組に関する情報を提供する必要があります。

しかし、こうした情報だけでは十分とはいえないため、独自開発のEPGによって利用可能な情報の量を増やしたり、上級検索、お気に入りチャンネルといった機能や、その他の拡張サービスを提供しています。このような拡張EPGは、通常はセット・トップ・ボックスに常駐して、標準的なテーブル情報と独自のデータ・フォーマットを組合せて使用します。この種の番組情報は即時性が要求されるものではないため、安価な低ビット・レート技術が使用されています。多くの場合、データはセット・トップ・ボックスにキャッシュされるようになっているため、実際の総伝送時間はかなり長くても、ユーザには即座に表示されるように見えます。

11.1.2 オポチュニスティック・データ

多重化やテーブル挿入などの複雑な制約に対処する必要性を考えると、ビデオ/オーディオ・ストリームで、トランスポート・ストリームの周波数帯域をすべて使用できることは稀です。オポチュニスティック・データ方式では、ヌル・パケットを有益なデータと置き換えることで、この空き容量を利用します。しかし、データの周波数帯域は保証できず、非常に小さくなってしまいう場合もあります。したがって、使用はリアルタイム制約のないアプリケーションに限られます。

このようなデータを使用するアプリケーションとしては、価格リストなどのファイル転送や、人工衛星による全拠点への在庫レベル・データの配信などが考えられます。こうした業務には、夜間に転送を行わなければならないという制約がありますが、そのため、低周波数で安価なオポチュニスティック・データ・サービスを使用できます。

11.1.3 ネットワーク・データ

放送用トランスポート・ストリームは、ネットワーク・データの伝送にもごく普通に使用されています。たとえば、地理的に離れた2つの拠点を、衛星を介して非常に高い周波数帯域で相互接続することができます。このようなネットワーク・リンクは、ほぼあらゆる種類のネットワーク・プロトコルに使用することができます。

しかし、現在最も多い用途は、IPプロトコルによる高速のインターネット・ダウンロードです。この場合、コマンドやページ要求を、低速インターネット接続を使用してインターネット・サーバへ送ります。しかし、インターネット・サーバは、ISDNや平均的なブロードバンド接続よりもはるかに高速に、要求されたデータを衛星経由で転送します。

利用可能なバック・チャンネルが低速のみの場合、これと異なるアプリケーションが使用されます。たとえば、放映中のテレビ番組に関連するデータを、伝送信号によってオン・デマンドで送ることができ、インターネットのサイトにアクセスしていることを視聴者に意識させません。

11.1.4 エンハンスドTV

エンハンスドTVのアプリケーションにはバック・チャンネルがないため、必要なデータはすべてトランスポート・ストリームで送る必要があります。そのため、設計やアプリケーションの範囲に明白な制約があるものの、非常に多彩なゲーム、クイズ、インフォーマショナルなどを制作でき、すべてテレビのみとのインタラクティブであることを視聴者にまったく意識させません。特に事前収録素材では、クイズの問題と回答のようなデータ・ポップアップを、音声/映像と同期化させるのは比較的簡単です。

11.1.5 インタラクティブTV

エンハンスドTVとインタラクティブTVの最大の違いは、インタラクティブTVにはバック・チャンネルがあり、個人性の高いデータが送受信されるということです。そのため、個人性の高いフィードバックはもとより、多人数参加ゲームや投票など、他の視聴者とのリアルタイム・インタラクションを含め、対応できるアプリケーションの範囲が大きく広がります。

アプリケーションの範囲は、高速インターネット接続とバック・チャンネルを組合せることでさらに広がります。個人性の高いデータとテレビ放送がシームレスに統合され、真のVODによって番組やその他の素材を提供することも可能になります。

11.2 コンテンツのカプセル化

データ放送ではまず、トランスポート・ストリームによる伝送に適した形式にデータをカプセル化します。非常に多くの種類のデータがあるため、MPEG-2規格ではさまざまなカプセル化方法が提供されます。DVBやATSCなどの規格では、地域的な要件を満たすために、MPEG-2の基本オプションを拡張していますが、これらはすべてコアのMPEG-2規格をベースにしています。

11.2.1 MPEGのデータ・カプセル化

11.2.1.1 データ・パイピング

データ・パイピングは、ネットワーク上にある受信側セット・トップ・ボックスへの、単純な非同期送信に使用します。主として非公開のシステムでの専有データの送信に使用され、受信側のセット・トップ・ボックスは、特定のPIDを持ったデータを探すようにあらかじめ設定されています。タイミング情報はなく、データはMPEG-2 TSパケットのペイロードとして直接送られます。地域内にある全販売拠点の製品在庫レベルを夜間に更新する業務などが、代表的なアプリケーションです。

11.2.1.2 データ・ストリーミング

データ・ストリーミングは、同期または非同期での終端間のデータ送信に使用します。データはPESパケットとして、ビデオ/オーディオ・サービスと同様の方法で送られます。実際に、ビデオやオーディオは、同期データ・ストリーミング・サービスの代表例です。PESパケットは、番組配列情報(SI)テーブルで通知され、タイミング情報を運ぶことができるため、データ・ストリーミングはデータ・パイピングよりも柔軟ですが、実際には専有の非公開システムで同じように使用されています。

11.2.1.3 Digital Storage Medium Command and Control(DSMCC)

MPEG-2 DSM-CC規格(ISO/IEC 13818-6)では、標準的なMPEG-2専用テーブルのセクションでデータを放送する追加的な方法について規定しています。当初は、トランスポート・ストリ

ームを使ったネットワークでの番組素材のVOD送信に対応する方法として考案されましたが、オン・デマンド送信(MPEのパラダイム)だけでなく、複数のネットワーク・プロバイダ間でのデータの定期的送信(カプセルのパラダイム)の両方に対応できるよう、プロトコルが拡張されています。

11.2.1.4 マルチプロトコルのカプセル化(MPE)

MPE(マルチプロトコルのカプセル化)を使用すれば、任意の通信プロトコルのデータグラムを、トランスポート・ストリームを介して、DSM-CCテーブルのセクションで伝送することができます。データグラムとは、サイズ、内容、送信先、送信方法など、データに関するあらゆる定義情報を含む論理構造のことです。

最も一般的なアプリケーションは、TCP/IPデータグラムで送信元と送信先(受信側)の論理(IP)アドレスや、送信先の媒体アクセス制御(MAC)アドレスについての情報を運ぶインターネット送信です。しかし、MPEはほぼすべてのネットワーク・プロトコルに対応しており、TCP/IPデータのみ限定されるわけではありません。

11.2.1.5 カプセル

カプセルは、トランスポート・ストリームでの定期的な情報の送信に使用します。カプセルのコンテンツは、受信側ユーザの要求に応じて変更できますが、受信側が待機しているか、そのデータがそのときに必要かといったことに関係なく、定期的に送信するというのが一般的です。データが必要な受信側は、送信されるまで待つことになります。

カプセルには、オブジェクト・カプセルとデータ・カプセルの2種類があり、主な違いは以下の通りです。

データ・カプセルに含まれているのは、コンテンツとして未指定のデータのモジュールのみで、受け取ったデータの処理は受信側が決めます。

オブジェクト・カプセルには画像、テキスト・ファイル、実行可能プログラム・ファイルなどの識別可能なデータ・オブジェクトと、それらをすべてリスト化したディレクトリが含まれています。

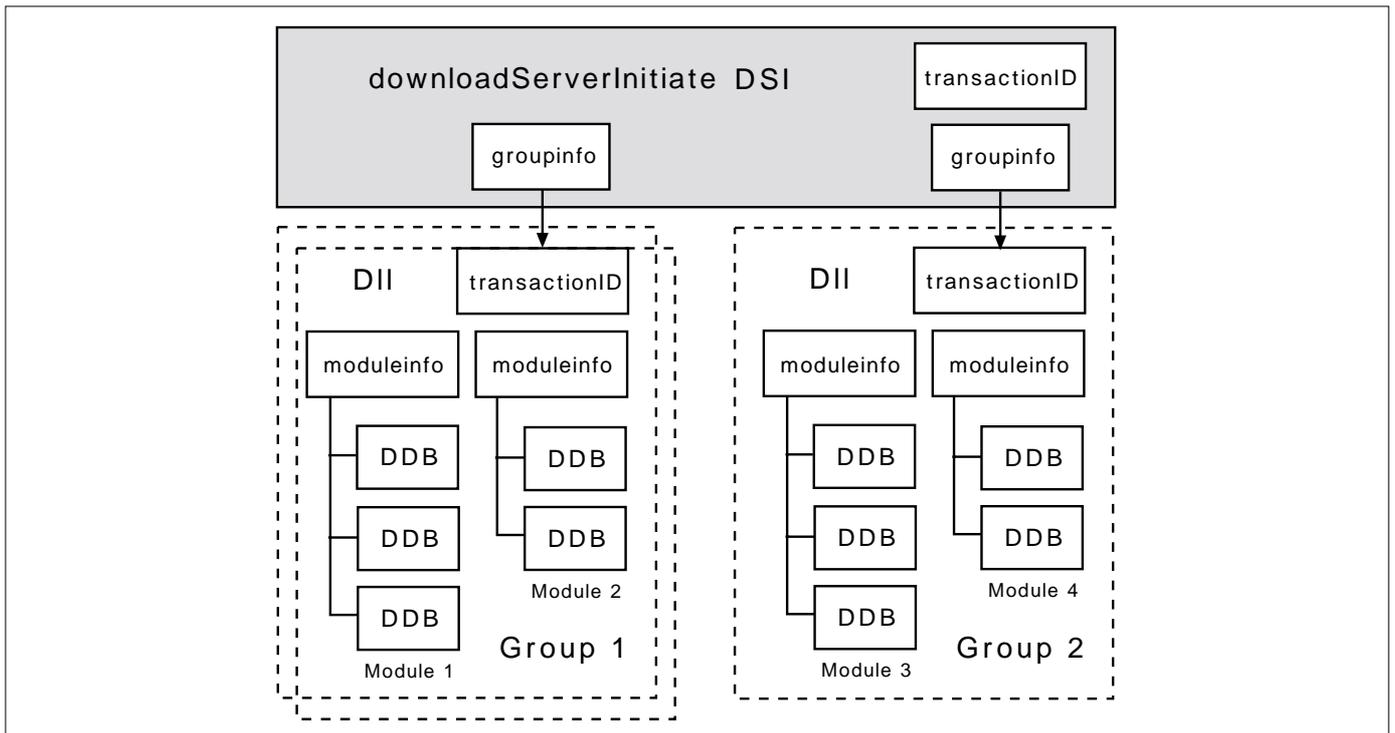


図11-1

データ・カールセルは、セット・トップ・ボックスへの新しいシステム・ソフトウェアのダウンロードに使用されることが多いのに対し、オブジェクト・カールセルは、ショッピング・サービスやEPG、さらにはアプリケーションやゲームの送信などに使用されます。

どちらのカールセルも、定期的に繰り返し送信が行われますが、オブジェクト・カールセルでは、オブジェクトごとに反復率が簡単に変えられるようになっています。たとえば、数時間先までのEPGは、翌月のEPGよりもはるかに高い頻度で送信することができます。オブジェクトの反復率は、サービス・プロバイダが周波数帯域を最大限に有効利用する目的で、商業ベースで決めます。

オブジェクト・カールセル、データ・カールセルはともに、MPEG-2の拡張規格であるDSM-CC(ISO/IEC 13818-6)と、DVB、ARIB、ATSC方式それぞれの拡張規格に基づいています。

11.2.1.6 データ・カールセル

データ・カールセルには、ひとまとまりのデータが含まれているだけで、データ項目やディレクトリ構造などはありません。データの内容や用途についての判断は受信側のユーザに委ねられています。

構造を図11-1に示します。ひとまとまりのデータを「モジュール」と呼びます。伝送モジュールは1つ以上のブロックに分かれます。各ブロックは、MPEGで定義された専用テーブル構文に従ったDownloadDataBlock(DDB)メッセージのペイロードのセクションとして送られます。DDBメッセージは、どのような順序や周期でも送ることができるため、各DDBがどのモジュールに属しているかを識別するメカニズムが必要です。

DownloadInfoIndication(DII)メッセージは、DDB同士をモジュールとしてリンクさせるのに使用します。1つのDIIメッセージには複数のモジュールの情報を入れることができ、これがグループを形成します。通常、グループには論理的に関係のあるデータ・モジュールが含まれています。

関連するデータ・モジュールが多く、1つのDIIメッセージでグループ化できない場合は、複数のDIIメッセージによってスーパーグループを形成できます。これらのグループはDownload ServerInitiate(DSI)メッセージによって参照されます。

1レイヤのデータ・カールセルには、1つのDIIで参照される程度の少数のモジュールが含まれています。

2レイヤのデータ・カールセルには、1つ以上のDIIメッセージを参照するDSIメッセージが含まれています。単一のカールセルよりも小さい場合と大きい場合があります。

2レイヤ・カールセルは多言語サポートに使用されるのが一般的です。たとえば、1番目のグループでは、1つ目のモジュールで実行可能プログラムを運び、2つ目のモジュールに英語のテキストを付けます。2番目のグループでは、モジュールをフランス語テキスト用の1つのみにする事で、プログラムのモジュールを繰り返すことによるオーバーヘッドを削減します。

11.2.1.7 オブジェクト・カールセル

オブジェクト・カールセルは、識別可能な個々のデータ項目を、サーバから受信機へ放送するのに使用します。この項目はオブジェクトと呼ばれ、画像、テキスト・ファイル、プログラム、映像PIDへのポインタ、カールセルにある項目のディレクトリ・リスト、つまりサービス・ゲートウェイなどです。関連するオブジェクトはグループ化され、サービス・ドメインという単一のカールセルとして送信されます。オブジェクトは必要に応じた頻度で送ることができ、オブジェクトの種類によって反復率を大きく変えることもできます。

オブジェクト・カールセルの重要な特徴の1つは、すべてのオブジェクトがBIOP(Broadcast Inter-ORB Protocol)パラダイムによって送られることです。従来のソフトウェア開発者は、長年にわたりORB(Object Request Brokerage)を使用してきました。

BIOPはこの基本的な方法を拡張して、さまざまなサービス・プロバイダのネットワークにまたがる放送環境でのオブジェクトの識別や使用をサポートします。BIOPとは、カールセルで放送されているオブジェクトに関する情報を交換するための方法です。BIOPには、オブジェクトが含まれている場合と、オブジェクトへのポインタのみの場合があります。また、オブジェクトを使用するのに必要なアプリケーション・ソフトウェアをダウンロードするリンクなど、オブジェクトの使用法について示している場合もあります。

オブジェクトのグループを組合せてモジュールを形成する点では、オブジェクト・カールセルはデータ・カールセルと似ています。そして、データ・カールセルの基本的な方法では、ブロック、モジュール、およびDIIを使用してそのデータを送信します。ここで最も重要な違いは、DSIを使用して、サービス・ゲートウェイのディレクトリ・オブジェクトを直接指示し、そこから、カールセルにある他のすべてのオブジェクトを見つけるようになっている点です。この仕組みを図11-2に示します。

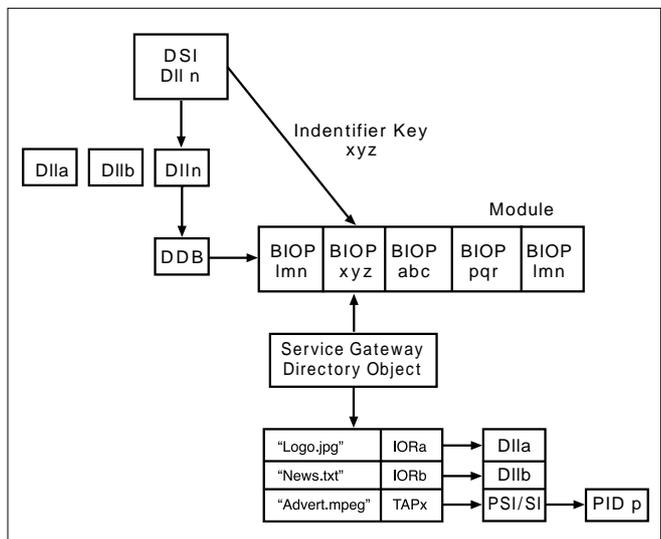


図11-2

11.2.1.8 オブジェクト・カールセルが放送される仕組み

詳しい説明は本書の範囲外であり、以下にごく簡潔に概要を説明します(図11-3も参照)。ディレクトリ、ファイル、およびストリームのオブジェクトは、データ・カールセルと同じ方法で送られます。つまり、モジュールをブロックに分け、Download Data Block(DDB)のペイロードのセクションとして送られます。

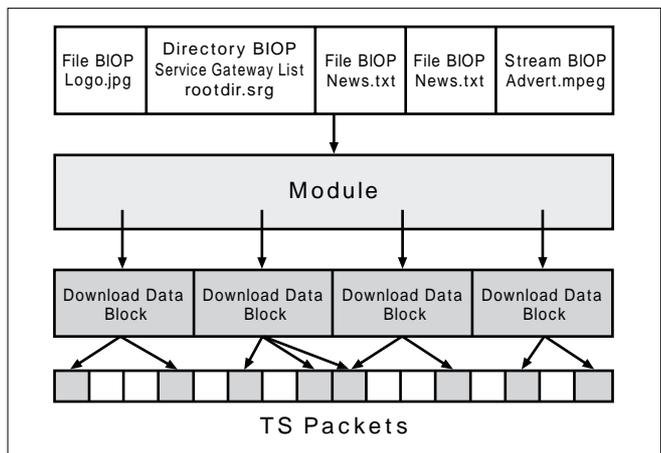


図11-3

DownloadServerInitiate(DSI)メッセージには、サービス・ゲートウェイと呼ばれる特別なディレクトリ・オブジェクトの場所があります。DSIメッセージはSIで参照されるため、特定のオブジェクト・カルーセルにどのような項目が含まれているかを理解する起点になります。DSIはDownloadInfoIndication(DII)を参照し、DIIが参照するDDBには、サービス・ゲートウェイ・オブジェクトが送られるモジュールが含まれています。

オブジェクトは、ディレクトリ・オブジェクトにおいてIOR (Interoperable Object Reference)によって参照されます。ここには、同じサービス・ドメイン内あるいは別のオブジェクト・カルーセル(別のトランスポート・ストリームで放送されるカルーセルを含む)のオブジェクトへアクセスするのに必要な情報がすべて含まれています。

IORに含まれていてオブジェクトの位置を表す構造は、プロファイル本体と呼ばれ、次の2種類があります。

- BIOPプロファイル本体 - 同じサービス・ドメイン内のオブジェクトのみに使用
- ライト・オプション・プロファイル本体 - 他のサーバやトランスポート・ストリームのオブジェクトに使用

あるオブジェクトが複数のカルーセルに含まれている場合、IORに複数のプロファイル本体があれば、セット・トップ・ボックスは最も簡単 / 迅速にアクセスできるオブジェクトを選択することができます。

DII、DDB、およびビデオ / オーディオ・ストリームの放送に実際に使用されるPIDが、伝送直前までわからないという問題を解決するために、タップが使用されます。そのため、カルーセル内では、PIDの参照はすべてタップによって行われます。タップとPIDの関連付けはSIで行います。これにより、複数のネットワークにわたるストリームの再多重化が大幅に簡素化されます。

11.2.1.9 MPEG-2データの同期化

データ放送は、放送されている番組と何らかの方法で同期化させる必要があります。スポーツ番組での点数の更新のように、数秒以上のエラーが実用面で問題にならないリアルタイム・アプリケーションを除き、リアルタイムのデータ送信を同期化の方法として使用するの、あまり現実的ではありません。しかし、たとえばクイズでは、1、2秒の違いでも非常に大きな影響があり、回答を出すのが早過ぎれば、深刻な結果を招くことになります。

MPEG-2では、データ・カプセル化の種類に応じて、さまざまなタイミング・メカニズムを提供します。データ・パイピングおよびMPEは、近リアルタイムのデータ送信以外のタイミング・メカニズムには一切対応していません。

データ・ストリーミングPESパケットには、ストリームのPCR値に由来するシステム・クロックと同期させるために、表示タイムスタンプ(PTS)と、場合によっては復号タイムスタンプ(DTS)を加えることができます。このメカニズムは、映像や音声の同期化や復号とまったく同じです。

MPEG-2のデータ・カルーセルにはタイミング情報はありません。しかし、オブジェクト・カルーセルには「ストリーム・イベント」と呼ばれる特殊なオブジェクトがあり、各テレビ番組の通常再生時間(NPT)に関するタイミング情報が含まれています。NPTクロックは、コマーシャルのときなどに一時停止することができ、PCRと同じではありません。つまり、番組のNPTは、番組をいつ伝送しても、どのようにコマーシャルを入れて分割しても、番組タイムラインとの完全な同期が維持されます。

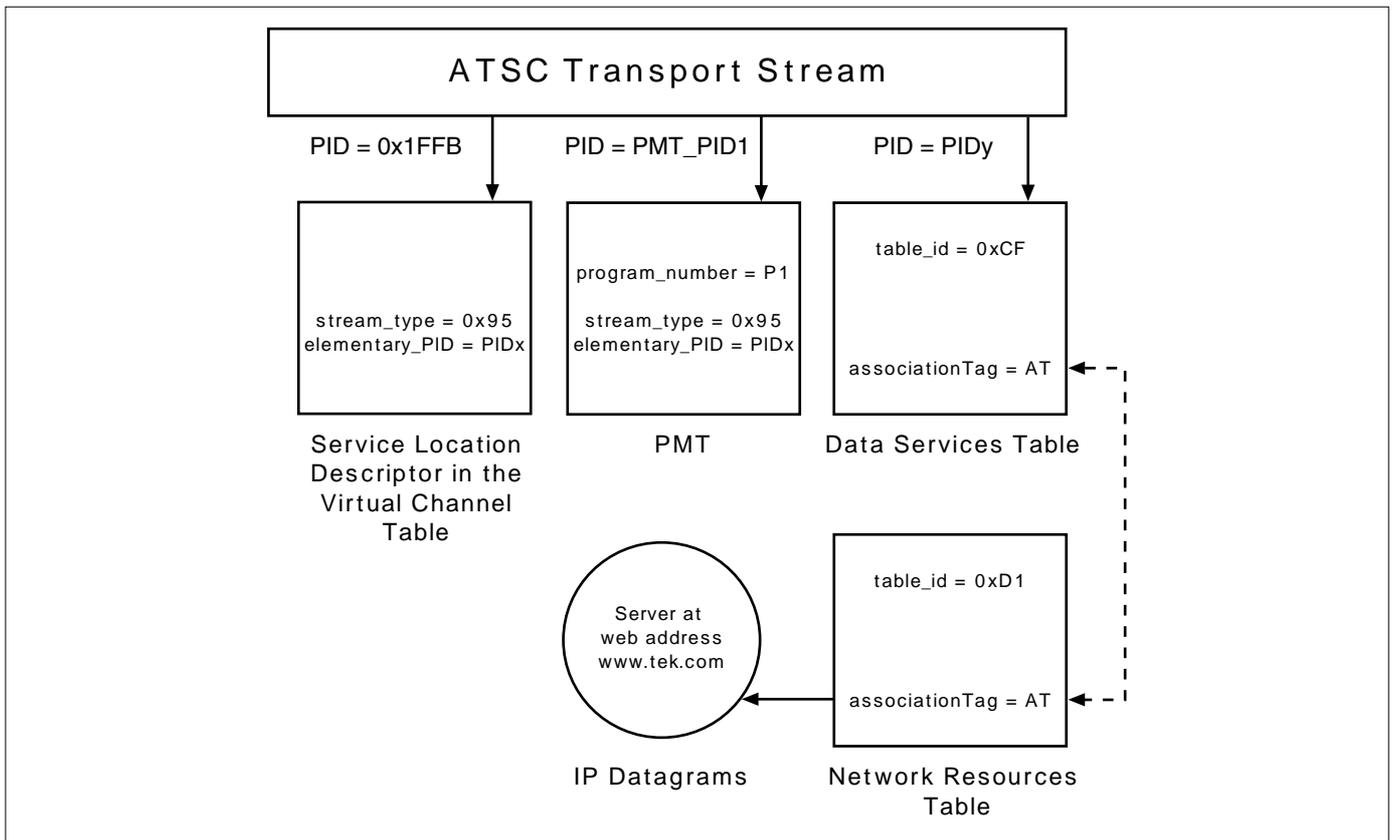


図11-4

11.2.2 DVBのデータ・カプセル化

DVBは、主として元の規格(EN 301 192)による曖昧さをなくす目的で、MPEG-2のカプセル化の方法を、ごくわずかな変更を加えて採用しています。変更には、わずかに異なるMPEフォーマットの指定や、データPESパケットの最大サイズに関する制限などがあります。また、DVBは、固有のネットワークIDとサーバIDを割当てることで、IORを使用するデータ・サービスに曖昧さがないようにしています。

DVBでは、DVB字幕とDVB文字放送のための2種類の専用データ・ストリーミングPESフォーマットについて定義し、セット・トップ・ボックス(STB)ファームウェア更新を送信するためのデータ・カプセルのフォーマットも規格化しています。また、データ記述子や追加のテーブルについても定義していますが、これについては、信号方式および発表に関するセクションで後述します。こうした追加の中で最も有益なのは、記述子を使用して、MPEデータグラムのMACアドレスを簡単にマスクできる機能です。これにより、ナローキャスト、マルチキャスト、ブロードキャストのMPEサービスに簡単に対応できます。

11.2.3 ATSC A/90のデータ・カプセル化

ATSC A/90規格はDVBの数年後に策定され、DVBやMPEG-2と比べてかなりの相違点があります。第一に、A/90では、オブジェクト・カプセルやCORBA/IOR方式を一切使用しない代わりに、ネットワーク・リソース・テーブル(NRT)という単独のテーブルを使用して、被参照データと実際の場所をリンクさせる情報を提供します。このようにしてシステム操作を簡単にしたことに加え、DVBやMPEGでは対応していないインターネットURLにも直接対応できるようになっています。このメカニズムを図11-4に示します。

第二に、A/90では、オブジェクトが1つしか含まれていないデータ・カプセルのみを使用します。データ・カプセルのフォーマットは、PTSによる同期データ送信も可能なDSM-CCアダプテーション・ヘッダを任意で加えられるように拡張されています。ただし、MPEGにおける番組のNPTの概念には対応していません。

A/90では、独自のDSM-CC MPEのバージョンや、PESフォーマットを使ったIPデータのストリーミングなど、追加的なデータの種類も定義しています。また、さまざまな種類のデータに合せて、DVBよりも豊富なタイミング・モデルに対応しています。

11.2.4 ARIBのデータ・カプセル化

日本のARIB標準規格は、A/90規格よりも後に策定されており、これまでの方式の中で最もシンプルなものといえます。データ・バイピングやオブジェクト・カプセルには対応していませんが、データ・カプセルのフォーマットを使用して1つ以上のエンティティを送るために、データ・カプセルのコンテンツに特定のディレクトリ構造を課しています。ARIBは、PMTで直接参照されないエンティティへの参照を認めていないため、CORBA(Common Object Request Broker Architecture)やNRTのようなタイプのリンクは必要ありません。

11.3 データ・コンテンツの伝送

データ・サービスはまず該当する国内/国際規格に従ってカプセル化された後、トランスポート・ストリームによってSTBへ伝送されます。

しかし、データ・サービスを利用するには、まずSTBがデータ・サービスの存在と、利用可能になる時期(アナウンス)を認識していなければなりません。さらに、実際に放送されているときに、データ・サービスのさまざまな成分(合図)を発見・識別できなければなりません。MPEG-2で定義されているのはPATとPMTのみであるため、エレメンタリ・ストリームの種類とごく基

本的な記述子について調べるだけで、すべてのアナウンスと合図を認識することができます。

PATとPMTしか使用しないというのは便利ではありますが、データ・サービスやイベントのスケジュールを作成できないことを考えても、数多くの制約があります。そこでDVBとATSCは、データ・サービスについての情報提供量を大幅に増加させる追加のテーブルを定義しました。ただし、いずれの場合も、少なくとも1つのPMTにある項目によって、データ伝送PIDが存在することを通知する必要があります。

11.3.1 DVBのアナウンス

DVBでは、データ・サービスは個々のイベント(クイズ番組など)と関連付けられているか、スポーツ・チャンネルのようなサービスの一部と考えられています。実質的には、データ・サービスが従来のオーディオ/ビデオのストリームと論理的に区別されていないMPEG-2のパラダイムに適合しています。

そのためDVBは、データ・サービスやイベントが利用可能になることをアナウンスするEITやSDTテーブルに加える追加の記述子を定義しています。そして、データ・サービスやコンポーネントの実際の場所は、前述の方法によって通知されます。

唯一新たに定義されたテーブルは、MHPサービスのみで使用するアプリケーション情報テーブルです。このテーブルには、カプセル内の全アプリケーションのリストやブート・クラスへのポインタが含まれており、チャンネル変更時にアプリケーションを自動で起動/停止させることもできます。

11.3.2 ATSCのアナウンス

DVBとは対照的に、A/90規格では、データのイベントやサービスを映像や音声のサービスとまったく別に扱っており、このパラダイムに対応するため、データ・サービス・テーブル(DST)やデータ・イベント・テーブル(DET)などのテーブルが追加されています。DSTには送信中のデータ・サービスがリストされています。DETは従来の番組予定情報におけるEITとまったく同じで、使用する構文も同じです。3つ目は長期間サービス・テーブルで、これはDETと似ていますが、DETが提供する16日間よりも先の長期的な予定情報を提供します。

11.4 コンテンツの表示

データ放送では最終的に、STBでアプリケーションを実行して情報を視聴者に表示します。コンテンツが映像、音声、データの正確なプロトコルでカプセル化されている限り、どのようなコンテンツを運んでいようと、トランスポート・ストリームには一切関係がありません。つまり、トランスポート・ストリームは、STBで使用するためのアプリケーションやデータを運んでいるだけなので、その次のセクションは、MPEG-2、DVB、ATSCなどのデータ放送規格とはほぼ完全に切り離されます。

11.4.1 セット・トップ・ボックスのミドルウェア

インターネット・データのように、使用する環境がそれぞれに定義されたり、それほど重要でなかったりする場合もありますが、通常、ゲームやアプリケーションなどでは、受信側のSTBで正しく実行されるようにデータ・アプリケーションを構築することが重要です。STBには多くの場合、アプリケーション・プログラム・インタフェース(API)があり、アプリケーションからの高レベルの機能呼び出しによってソフトウェア・ドライバに接続し、STBのハードウェアの制御や、トランスポート・ストリームで送られてきたデータへのリンクを実行します。このアプリケーションとドライバの間に存在するソフトウェア層を「ミドルウェア」と呼びます。

しかし、各国で使用されているSTBは種類が非常に豊富です。性能や特徴についても、SDTVの画質を復号して表示するだけのきわめて基本的なものから、パワフルな処理機能を備えた最新の装置までさまざまです。最新の装置の中には、高性能グラフィックス用のチップ・セットを備え、LinuxやWindowsのような本格的なOSを使用できるものもあります。また、番組録画用のハードディスクや、DVD再生用のDVD ROMなども搭載されており、MicrosoftやStar Officeパッケージのような従来のプログラムも実行できます。

最新の高性能セット・トップ・ボックスで表示するようにデザインされたグラフィックスは、単純なものでも、それ以前の装置では表示さえできないことがあります。アプリケーション開発業者は、STB、ミドルウェア、OSなど、種類ごとに異なるプログラミング言語を使ってまったく別々にプログラムを開発しなければならなくなり、明らかに非効率で、開発コストも高くなります。

そのため、STBや最低限の機能だけでも標準化しようとする試みが何度もなされてきましたが、STBの機能は各放送事業者のごく限られた範囲内で明確に定義され管理されています。

また、ミドルウェアと、適合プログラムを開発するためのアプリケーション開発ツールを組合せたソリューションを提供するというアプローチも各社によって実施されています。こうした方式はこれまで以上にオープンになり、基本的な機能要件さえ満たせば、性能や特徴の異なるさまざまなセット・トップ・ボックスで、ライセンス契約したミドルウェアを使用することができます。

これまで、各社独自のデータ伝送方式と「壁で仕切られた」インタラクティブ・サービスが広く利用され、バック・チャンネルによる真のインタラクティブ性も提供されてきました。その他にも、欧州全域で普及している各社独自のサービスがあり、標準的なオブジェクト・カルーセルでアプリケーションを送信することもできます。各社が開発した方式は他にも数多くあります。

よりオープンな規格としては、Multimedia Hypermedia Experts Group(MHEG)が定義したアプリケーション環境があります。その中でもMHEG-5は、性能が低くメモリの小さいアプリケーションやセット・トップ・ボックスに合せて最適化されています。MHEG-5は、英国で地上波放送に採用されるなど、すでにある程度の成功を収めています。

したがって、現在の傾向としては、APIのプログラミングを明確に定義したオープン規格のミドルウェアに基づいた規格化が進んでいます。これは、ユーザにとってはセット・トップ・ボックスの選択肢が広がり、アプリケーション開発者にとってはコードを再使用できるという利点があります。最近では、DVBのMultimedia Home Platform(MHP)のような、グローバル・スタンダードの最有力候補が出てきています。米国では、Advanced Television Enhancement Forum(ATVEF)およびDigital TV Application Software Environment(DASE)が主要なオープン規格です。

11.4.2 DVB Multimedia Home Platform(MHP)

Multimedia Home Platform(MHP)が定義するAPIは、性能やコストの異なるSTBやサービスに対応するさまざまなプロファイルやレベルを備えています。エンハンスドTVやインタラクティブTV(バック・チャンネル付き)に対応しており、他のDVB規格では直接対応していないインターネット接続にも対応しています。

MHPのアプリケーションは、標準的なDVBオブジェクト・カルーセルを使って送信され、前述のAITテーブルを使って発表されます。そのため、アプリケーションの起動、休止、停止が自動にでき、アプリケーションのブート時間を速めるためにリソースを明示的にキャッシュすることもできます。MHP規格はJava技術をベースにしており、JavaTV、HAVI(Home Audio/Video interoperability)、DAVIC(Digital Audio Visual Council)といった既存の規格に由来する一連のAPIをカバーしています。

アプリケーションはJavaベース、HTMLベースのどちらでもよく、MHP Software Development Kit(SDK)が複数のサプライヤから提供されています。システムの相互運用性やセキュリティを確保するため、アプリケーションの適合性テストと電子署名が必要です。

MHPロゴも保護されており、STBにロゴを付けて販売するには、適合性プロセスを実施する必要があります。大手メーカから出ている各種のMHPセット・トップ・ボックス上で、すべてのMHPアプリケーションが実行できることを確認するため、2年前から相互運用性テストが定期的に行われています。

MHPは欧州におけるデファクト・スタンダードです。フィンランドでは2001年8月に正式にサービスを開始しており、現在では欧州全域で数多くの試験放送が行われています。ドイツは2002年半ばにMHPを採用する予定で、他の国々でも今後数年の間に採用される見通しです。

しかし、MHPは欧州だけの規格ではなく、世界各国で採用されています。たとえば、韓国では衛星データ放送、オーストラリアでは地上波放送で、MHPの使用が規定されています。また、2002年2月には、米国のOpen Cable Application Platform(OCAP)が、米国のデジタル・ケーブル放送規格としてMHPを採用しました。その他にも、今後数年の間に多くの国や機関がMHPを採用すると予想されます。

MHPが世界的に及ぼすインパクトの大きさは相当なものになります。現在のインタラクティブTV市場の主要企業は、自社開発のミドルウェア上で使用できるMHPプラグインの提供を2002年中に開始すると表明しており、事実上、各社独自のセット・トップ・ボックスがグローバル市場に公開されることとなります。すでに欧州各国がMHPの採用に取り組んでいることを考えると、それだけで10億人近い潜在市場があることになり、このオープン規格の将来的な成功へのよい前兆といえます。

11.4.3 ATVEF DASE

ATVEF(Advanced Television Enhancement Forum)とDASE(Digital TV Application Software Environment)の違いは、前者がHTMLを推奨してJavaの使用を認めているのに対し、後者はJavaを推奨してHTMLの使用を認めている点だといわれていますが、それ以外にも基本的な違いがあります。

ATVEFは異業種間の企業グループで、インタラクティブなコンテンツをセット・トップ・ボックスやPCベースの受像機、インタラクティブTVなどで実行するのに必要な標準化機能を規格として定義します。このATVEFの取り組みにより、Enhanced Content Specification(ECS)規格が策定されています。

ATVEF ECSでは、HTMLをインタラクティブTVコンテンツを作成するための基本と定義していますが、JavaScriptの使用も可能です。その他にも、Portable Network Graphics(PNG)フォーマットを標準的な画像交換フォーマットとして使用するなどの基本的な機能が要求されています。したがって、ATVEFはインターネットのウェブ・ページのデザインとかなりの共通性があり、アプリケーション開発者にとっては明らかなメリットがあります。

ATVEFの強みの1つは、伝送のメカニズムが規格によって厳密に定義、想定されていないことです。伝送に依存しないコンテンツ・フォーマットを定義し、参照バインディングにIPを使用することで、映像コンテンツのあるなしに関係なく、アナログとデジタル間の伝送を申し分なく実行できます。そのため、IPを使用するあらゆる伝送方式を使用できます。したがって、ATVEFはMPEGトランスポート・ストリームだけでなくNTSC方式でも使用することができ、シンプルなりターン・パスのメカニズムも提供されます。ATVEFは、米国のアナログおよびデジタルの両テレビ方式で採用されています。

第11章 データ放送

11.4.4 DASE

これに対しDASEは、米国のATSCが策定した、はるかに複雑なミドルウェアの規格です。ATSCはA/90データ放送規格も策定していますが、DASE規格はA/90伝送方式と関連付けられており、DASEアプリケーションを実行するAPIについても定義しています。2002年初めには、候補規格としてDASE-1がリリースされています。

DASEには、信号方式および発表のスキームとともに、実際のアプリケーションのデータ・コンテンツ自体も盛り込まれており、データ・エッセンスと呼ばれています。2種類のデータ・エッセンスが認められており、使用可能な環境や機能についても併せて規定されています。

XDMLおよびXMLのマークアップ・テキストやスクリプトをベースにした宣言型データ・エッセンス
JavaTvのXletをベースにしたプロシージャ型データ・エッセンス

DASEの受信機は、これらのデータ・エッセンスとともに、2つの方式で共用されるJPEGやPNGコーデックのような一般的なコア機能にも対処できます。また、伝送される音声/映像のコンテンツに直接リンクを提供し、HDTV解像度など、インタラクティブ・コンテンツを表示するグラフィカル画像フォーマットについても定義しています。

アプリケーションは、MHPアプリケーションのライフサイクルのパラダイムと同じように、初期化、起動、休止、初期化解除することができます。必要に応じてリソースをキャッシュすることもできます。

DASEはATSCの影響力に支えられていますが、比較的后発ということもあり、現時点では普及には至っていません。

第12章 MPEGのテスト

テストでは、使用しているトランスポート・ストリームの適合性を解析する機能はもちろんですが、既知の適合したトランスポート・ストリームを発生させる機能も必要になります。

12.1 テストの要件

MPEGの技術はこれまでの技術と大きく異なりますが、テストの要件は基本的に同じです。オペレーション・ベースでは、全体が問題なく動作していることを定期的に確認できる簡単なチェック方法が求められています。障害の発生時は、発生場所を速やかに特定する必要があります。また、装置の設計においては、問題の性質をある程度詳しく調査する必要があります。方法としては、さまざまな信号テストと同様、既知の有効な信号を発生させてシステムに送り、各地点で測定を行います。

MPEGの特長の中で従来の放送映像機器と最も異なることの1つは、下位のレイヤに対して透過的な複数の情報レイヤを備えていることです。障害のあるレイヤを特定できるということは、効率的な障害検出のためにも重要です。

たとえば、MPEGデコードのモニタに視認できる欠陥が表示されている場合は、いくつかの原因が考えられます。まず、エンコーダに障害があり、その誤った情報をトランスポート・ストリームが忠実に伝えている場合があります。その逆に、エンコーダに問題がなくても、トランスポート・レイヤでデータの破壊が発生することもあります。DVBには、エネルギー分散、エラー訂正、インターリーブなど、さらに多くのレイヤがあります。このように複雑な場合には、適切なツールを使った体系的な障害検出方法が必要になります。方法を考える上で、圧縮データのプロトコル解析に関する説明が参考になります。また、テレビの圧縮方式における画質の評価について知っておくことも重要です。

12.2 トランスポート・ストリームの解析

MPEGトランスポート・ストリームの構造はきわめて複雑ですが、AD953A型のようなアナライザを使うと、構造が論理的に分解され、ユーザが必要な部分を詳しく観察できるようになります。また、ライブのトランスポート・ストリームにおいて、各種の一般的な解析をリアルタイムで実施することができます。たとえば、トランスポート・ストリームにおける番組の階層や、各ストリームに割当てられたビット・レートの割合などを表示できます。

トランスポート・ストリームの一部を記録して、後で個別に取り出すようにできれば、さらに詳しい解析も可能です。こうした手法はデファードタイム・テストと呼ばれ、タイムスタンプの内容を調べる場合などに使用できます。

MPEGトランスポート・ストリーム・アナライザをデファードタイム・テストに使用すると、MPEG専用のデータ解析ツールとしてロジック・アナライザのように機能します。他のロジック・アナライザと同様、データ・ストレージを行う時間や条件を決めるには、リアルタイム・トリガ・メカニズムが必要です。図12-1のように、アナライザにはリアルタイム解析、データ・ストレージ、デファードタイム解析の各セクションがあります。リアルタイム解析ではリアルタイム・セクションのみが動作し、また、信号源を接続する必要があります。データ・ストレージにおいては、リアルタイム・セクションを使用して、いつトリガするかを決めます。アナライザはフィルタと呼ばれるツールを備えており、デ

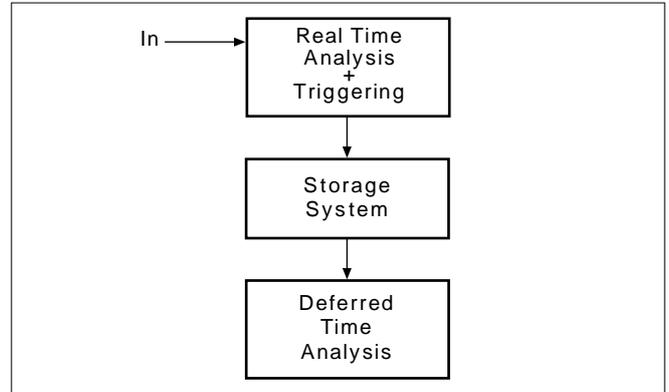


図12-1

ータ・ストレージの前または後に選択的に解析を行えます。

データ・ストレージが完了したら、デファードタイム・セクションがストレージ・データを処理できるようになるため、入力信号は不要になります。リアルタイム入力を直接表示したり、保存して後で調べる機能は、ストレージ・オシロスコープとよく似ています。

12.3 階層表示

未知のトランスポート・ストリームを解析する場合は、まず階層表示によって、ストリームの全成分をグラフィック表示しておくことをお勧めします。図12-2は、当社のMTX100型で得られる階層表示の例です。まず、トランスポート・ストリーム全体が左上で分割され、ストリーム成分がすべてアイコンで表されています。

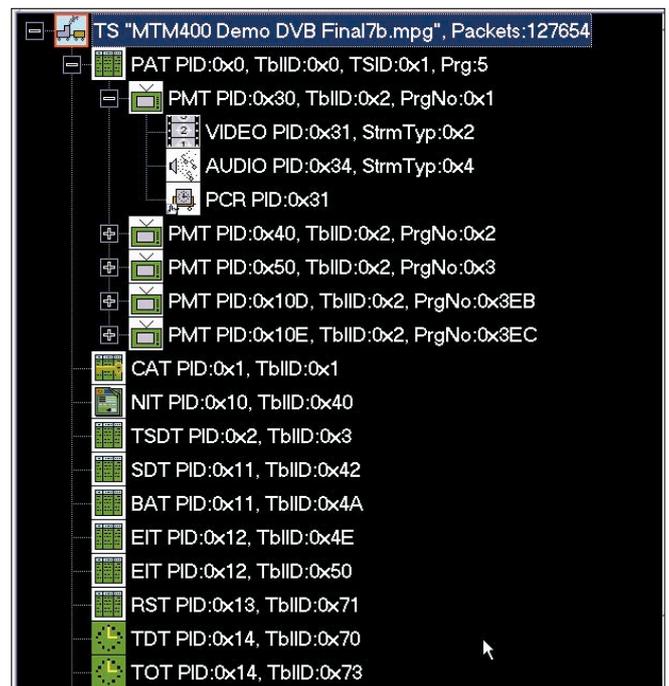


図12-2

第12章 MPEGのテスト

表12-1は、階層表示で使用される各種のアイコンとその意味を示しています。プログラム・ストリームがいくつ存在し、各コンテンツが映像か音声かについても簡単にわかるようになっています。各アイコンは最上位のレイヤを表し、その下位に解析や情報などの複数のレイヤがあります。

アナライザは、トランスポート・ストリームのPSIデータにあるPATとPMTを使用して、階層表示を作成します。これらのテーブルのPIDは、各アイコンの下に表示されます。PATとPMTのデータは、デマルチプレクサやデコーダが動作するための基礎になります。アナライザが階層表示できなかつたり、明らかに間違っただけの表示をする場合は、テスト対象のトランスポート・ストリームにPAT/PMTエラーがあります。さらに上位機種のアナライザであっても、このストリームを解析するのは不可能です。

デマルチプレクサやデコーダがトランスポート・ストリームにロックできるかどうかは、PSIデータを送る周波数によって決まります。図12-3に示したPSI/SIレートのオプションでは、システム情報を挿入する周波数が示されています。PSI/SI情報には、ビット・ストリームの実際の内容との整合性が必要です。たとえば、あるPIDをPMTで参照した場合、その値のPIDがビット・ストリームでも見つかるはずですが、この比較は、整合性チェック機能が行います。図12-4は、2つの未参照パケットを含むストリームによる整合性エラーを示しています。

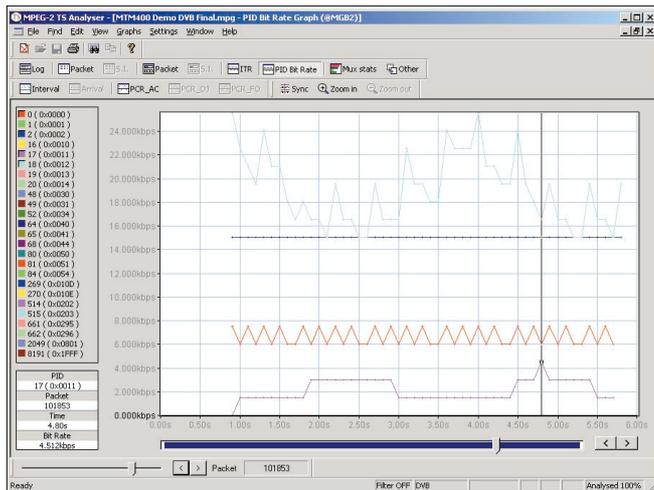


図 12-3

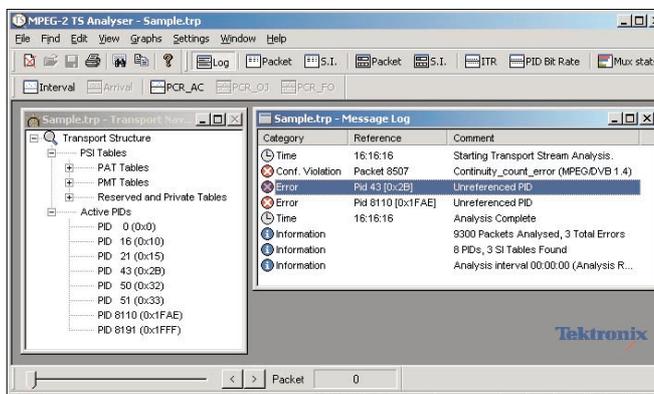


図 12-4

表 12-1. 階層表示アイコン

アイコン	エレメントの種類
	多重トランスポート・パケット。ストリームを構成するすべて（188バイトと204バイトのトランスポート・パケット）を表します。トランスポート・ストリームを列車にたとえると、このアイコンは、構造（長物車、有蓋貨車、ホッパー車など）や貨物の種類に関係なく、列車の各車両を表します。
	特定のPID（番組ID）を持つトランスポート・パケット。それ以外のエレメント（テーブル、クロック、PESパケット）は、トランスポート・パケットに含まれている「ペイロード」、つまり同じPIDを持つ複数のトランスポート・パケットのペイロードで構成されています。PID番号はアイコンの下に表示されます。階層表示では、このアイコンの右にあるアイコンが、同じPIDを持つパケットのペイロードです。
	個別のPCRクロックを含むトランスポート・パケット。PIDはアイコンの下に表示されます。
	PAT（番組関連テーブル）セクション。常にPIDが0のトランスポート・パケットに含まれています。
	PMT（番組マップ・テーブル）セクション
	NIT（ネットワーク情報テーブル）。SelectionメニューのPSI/SIコマンドによって、SIテーブルへアクセスします。Privateセクションにも使用されます。DVBオプション（Optionsメニュー下）を選択した場合は、このアイコンがSDT、BAT、EIT、TDTのセクションも表します。
	PES（パケット化エレメンタリ・ストリーム）。あるエレメンタリ・ストリームを構成するすべてのパケットを表します。個々のPESパケットは、複数のトランスポート・パケットのペイロードから構成されます。
	ビデオ・エレメンタリ・ストリーム
	オーディオ・エレメンタリ・ストリーム
	データ・エレメンタリ・ストリーム
	ECM（限定受信共通情報）セクション
	EMM（限定受信個別情報）セクション

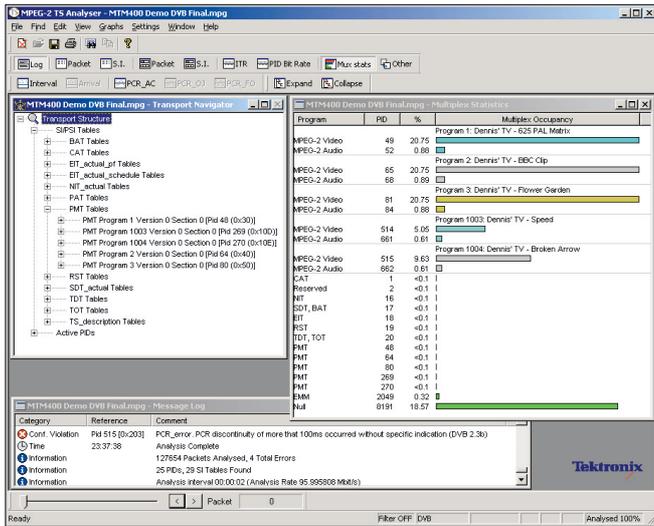


図 12-5

MUX割当てチャートは、各PIDつまり番組に割当てられたトランスポート・ストリームの割合をグラフ表示します。図12-5はMUX割当てチャートの表示例です。階層表示とMUX割当てチャートによって、トランスポート・ストリームのエレメント数と割当てられた周波数帯域の割合を示しています。

12.4 解析表示

未指定の場所に特定のデータがないかチェックする代わりに、特定の場所でトランスポート・ストリーム・パケット、テーブル、PESパケットなどの未指定のデータを解析することもできます。この解析は、アナライザがデータを自動で構文解析して復号し、その意味を表示するため、解析表示と呼ばれています。図12-6は、MPEGトランスポート・パケットの16進表示および解析表示の例です。選択した項目が変わると、ストリームの開始に関連したパケット番号が表示されます。

図12-7は、解析表示におけるPATの例を示しています。

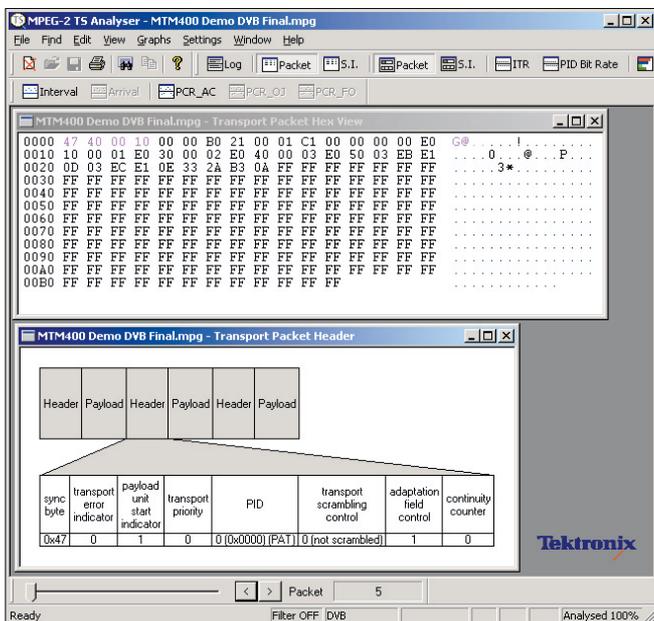


図 12-6

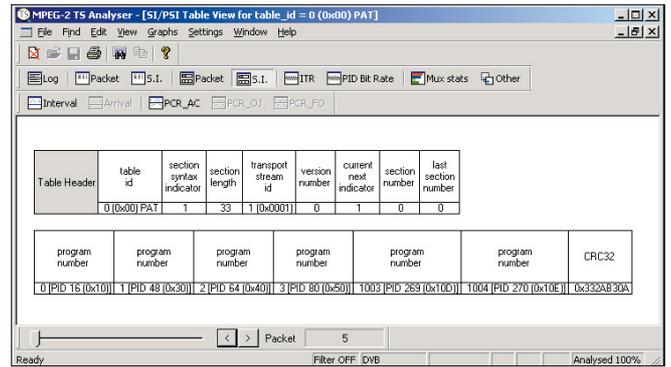


図 12-7

12.5 構文およびCRC解析

トランスポート・ストリームによる番組素材の伝送は、エンコーダが構文を正確に使用するかどうかですべて決まります。固定フラグ・ビット、同期パターン、パケット開始コード、パケット・カウントなどが正しく設定されていないと、デコーダがビット・ストリームを誤って解析するおそれがあります。構文検査機能は、番組素材以外のビットをすべて調べ、矛盾があれば表示します。スプリアスな矛盾は伝送エラーが原因と考えられ、一貫した矛盾はエンコーダまたはマルチプレクサの障害を示しています。図12-8は、構文エラーと巡回冗長検査 (CRC) エラーを示しています。

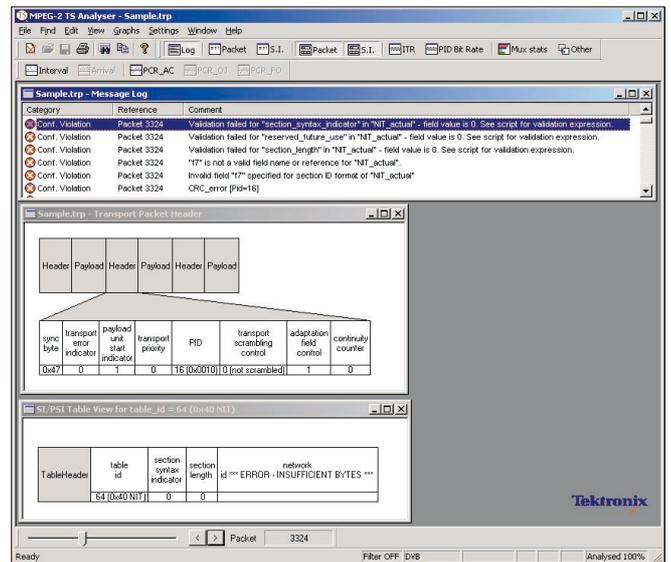


図 12-8

多くのMPEGテーブルには、エラー検出のためにチェックサム、つまりCRC文字が付加されています。アナライザはチェックサムを再計算し、実際のチェックサムと比較します。この場合も、CRCのスプリアスな不一致はストリーム・ビットのエラーに原因があり、一貫したエラーはハードウェア障害を示しています。

12.6 フィルタリング

トランスポート・ストリームには大量のデータが含まれており、実際の障害状態でも、深刻な問題が存在しない限りデータの大部分は有効で、影響を受けるのはごく一部のエレメンタリ・ストリームや番組であると考えられます。こうした場合は、フィルタリングの機能を使って、選択的にテストを実施したほうが効果的です。

基本的にはアナライザのフィルタリング機能によって、トランスポート・ストリームを選択的に検査することが可能になります。すべてのビットを受け入れるのではなく、ある一定の条件を満たすデータのみを解析できます。

ある条件でパケット・ヘッダをフィルタリングすれば、特定のPIDのパケットのみが解析されます。このアプローチでは、PID 0を選択するだけで簡単にPATを検査することができます。このようにして、それ以外のPIDもすべて読み出すことができます。階層表示などから疑わしいストリームのPIDがわかれば、そのPIDのみを選び出して解析することも簡単です。

12.7 タイミング解析

前述のテストは、トランスポート・ストリームに正しいエレメントや構文が存在するかどうかを検査するものです。しかし、音声と映像をリアルタイムで正確に提供するには、デコーダへ正確なタイミングが送信されていなければなりません。これは、PCRとタイムスタンプ・データの解析によって確認できます。

プログラム・クロック・データは復号プロセス全体のタイミングを制御するものであるため、このデータを正確に転送することはきわめて重要です。PCR解析によって、PCRデータが、規格に適合する十分なレートや精度で送られているかどうかを、番組ごとに示すことができます。

マルチプレクサからのPCRデータは正確でも、再多重化の際に、ある番組のパケットが時間軸上の別の場所に置かれ、リマルチプレクサによる編集が必要になる場合があります。そのため、データの再多重化後は、PCRに誤差がないかテストすることが重要です。

図12-9のPCR表示は、平均クロックを基準にしてPCRが受信された位置を示しています。図12-10のように、次の表示レベルでは、PCRを開いてPCRデータを表示できます。アナライザは誤差を測

定するため、直前のPCRとビット・レートを使って補間PCRを作成してPCR値を予測します。この予測PCRから実際のPCRを引くことで、誤差を推定します。

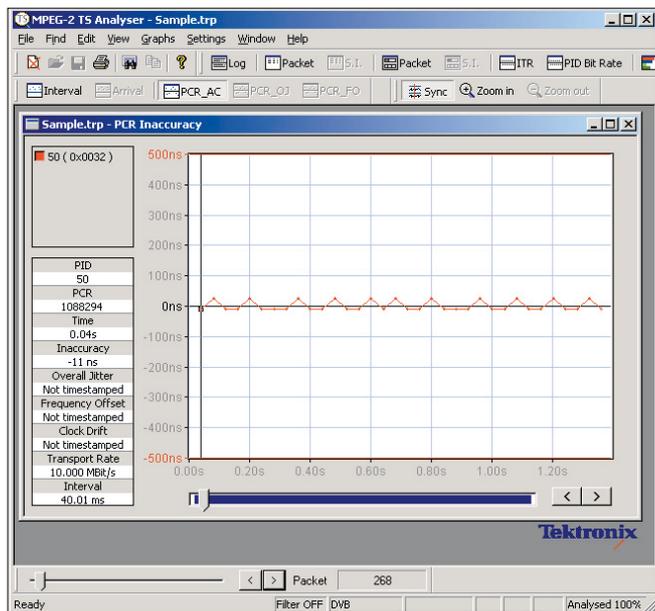


図12-10

図12-11に示した別のアプローチでは、PCRインターバル、誤差、ジッタ、周波数のオフセットやドリフトがグラフ表示され、リアルタイムで更新されています。

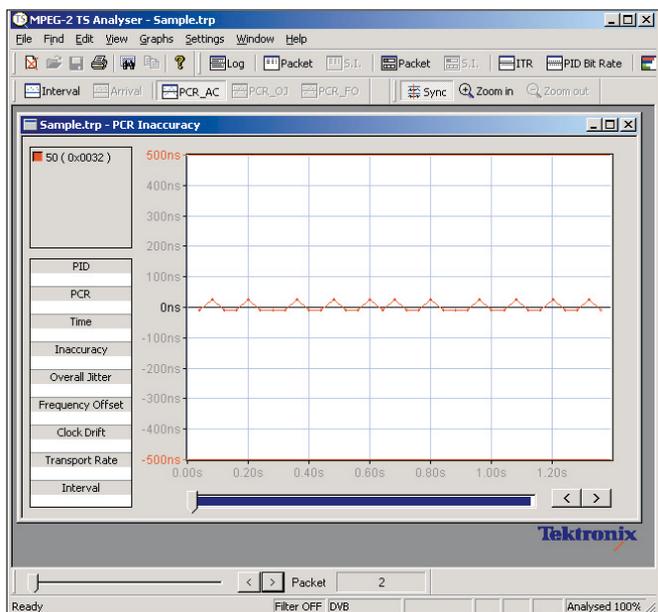


図12-9

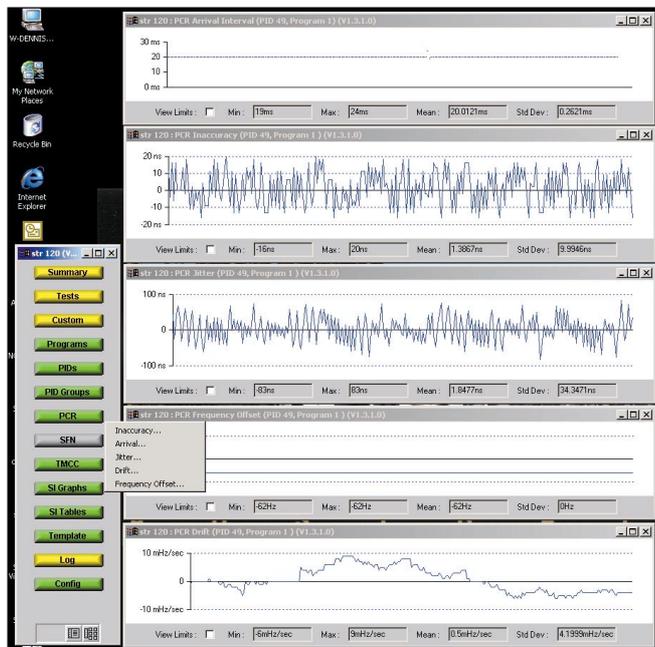


図12-11

図12-12は、選択したエレメンタリ・ストリームのタイムスタンプ表示です。アクセス・ユニット、表示時間、および該当する場合は番号時間が表示されます。

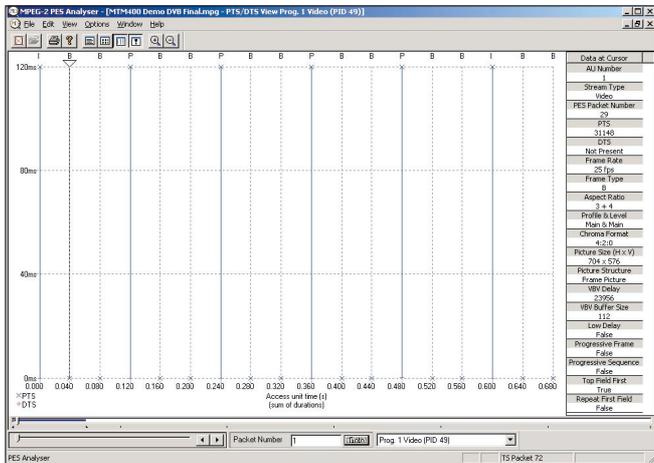


図 12-12

MPEGでは、種類の異なるピクチャを並べ替えて使用すると遅延の原因になり、エンコーダとデコーダの両方でバッファが必要になります。エレメンタリ・ストリームは、デコーダでバッファを利用できる範囲で符合化する必要があります。MPEGではT-STD (Transport Stream System Target Decoder) と呼ばれるモデル・デコーダが定義されており、エンコーダやマルチプレクサは、T-STDのバッファ能力を超えるような歪みをデータ・フローに与えてはいけません。トランスポート・ストリームには、エレメンタリ・ストリームに必要なバッファの量を指定した、VBV (ビデオ・バッファ・ベリファイ) というパラメータが含まれています。

T-STD解析により、バッファの使用率がグラフ表示されるため、オーバーフローやアンダーフローが起きていても簡単にわかります。図12-13はバッファの表示です。

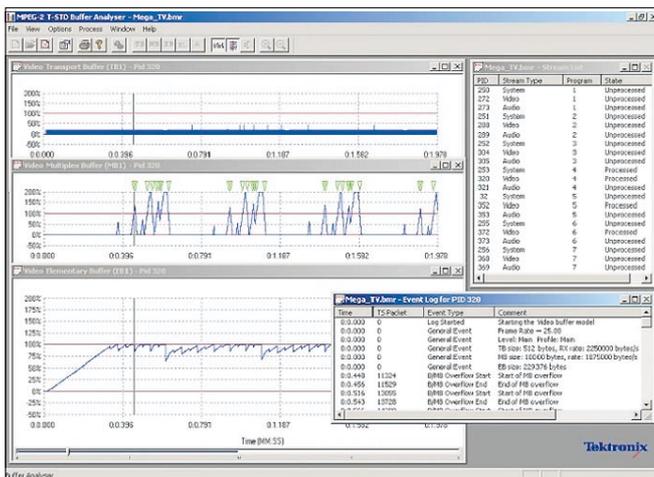


図 12-13

通常の圧縮器やマルチプレクサの出力は決定的なものではないので、限定的にしか使用されません。デコーダの欠陥が観察されたとしても、同じビデオ信号によって、同じトランスポート・ストリームが得られるわけではないので、テストを繰り返して同じ欠陥が見られるという保証はありません。こうした場合、調査や修正のために何度でも発生させられる、完全に再現可能なトランス

ポート・ストリームが不可欠です。

トランスポート・ストリーム・ジッタは一定の範囲内に抑えられるべきですが、優れた設計のデコーダでは、信頼性の高いオペレーションを保証するため、この範囲を超えて番組を回復させることができるようになってきました。存在するトランスポート・ストリームが適合していれば、デコーダのテストは行われなため、そのトランスポート・ストリームを使ってこの機能があるかどうかを調べる方法はありません。障害があっても復元できなければ、それがジッタによるものなのか、それ以外の不適合によるものを明らかにできないことがあります。この場合の解決策として、あらゆる面で適合したトランスポート・ストリームを発生させ、限定的な量の誤差を加えることで、誤差が不適合の唯一の原因であることがわかるようにします。AD953型のエディタ機能は、この信号を発生するように設計されています。

12.8 エレメンタリ・ストリームのテスト

MPEGビット・ストリームは柔軟な性質を持っているため、含まれている可能性や組合せは数え切れないほど多くあります。エンコーダが定義されてないため、エンコーダのメーカはあらゆる可能性を想定するよう要求されてはいませんが、コストの問題を考えると、実際にそのようにしていることはまずありません。そのため、あるエンコーダとデコーダの組合せが正しく動作したからといって、適合性の証明にはならず、テストは非常に困難になります。エンコーダが、デコーダに障害を起こすようなモードを使用していなかっただけという可能性もあります。

また、エンコーダは決定的なものではなく、映像や音声の入力を繰返しても同じビット・ストリームは発生しないため、この問題はさらに複雑になります。I、P、Bのピクチャと映像フレームの間に同じ配列が存在する可能性はわずかしかなりません。1回目のテストでデコーダに障害があっても、次のテストで障害があるとは限らないため、障害検出は困難です。エンコーダに障害があっても、その障害の原因がエンコーダにあるのか、デコーダにあるのかは判断できません。符合化の困難さは番組素材の性質に大きく左右され、番組素材の符合化範囲全体で、すべてのパラメータが実行されているとは限りません。

テストによって有効な結果を得るには、次の2つのツールが必要になります。

符合化範囲全体を意図的に調べるための、既知のテスト用適合信号源。この信号は、デコーダ障害によって現象を反復的に発生させる、決定的な信号であることが必要です。Sarnoff適合ビット・ストリームは、このタスクを実行するように設計されています。

エンコーダからの構文全体の適合性を検査できる、エレメンタリ・ストリーム・アナライザ

12.9 Sarnoff[®] 適合ビット・ストリーム

Sarnoff社は、デコーダの適合性テスト専用ビット・ストリームを開発しました。そのビット・ストリームは、多重化してトランスポート・ストリームとしてデコーダに送ることができます。

デコーダの内部機構にアクセスする必要はありません。デコーダ出力の解析に時間をかけなくてもすむように、解析が完了した時点で単色の画像が作成されるようになっているため、画像モニタをデコーダの出力端子に接続するだけで、結果が表示されます。

単色の画像は何枚か出力されます。図12-14はグレーの検証画面です。この検証画面を調べ、グレーの背景と対照的な矛盾が表示されていないか確認します。検証画像には、グレー以外の色もありません。



図12-14

テストを実施して障害があった場合、画像がまったく出力されません。また、テストが完了した際には、画面に「VERIFY」という文字が表示されます。

次のテストでは、画像上の可動エレメントのスムーズな動きを検査します。タイミングや順序に問題があると、視覚的なジッターの原因になります。

Sarnoffのテスト一式を使用すれば、MPEG構文のすべてのエレメントを順に検査していくことができます。あるテストでは、まずIピクチャのみのビット・ストリームから始め、次いでPピクチャ、Bピクチャを加えることで、全種類のMPEGピクチャに対応でき、正確に並べ替えられるかをテストします。MPEG-1との下位互換性も証明できます。さまざまなGOP構造の使用をテストするためのビット・ストリームもあります。また、値の全範囲で動きベクトルを検査するテスト、スライスのサイズやスタッフィングの量を変えるテストもあります。

デコーダのテストの他、Sarnoffストリームには、正常なデコーダで標準的なビデオ・テスト信号を発生させるシーケンスが含まれており、DAQ(デジタル・アナログ・コンバータ)、信号レベル、コンポジットやY/Cのエンコーダなどを検査することができます。デコーダはこのシーケンスによって、ゾーン・プレート、ランプ、カラー・バーなどの従来のビデオ信号を発生させることができる、ビデオ・テスト・パターン・ゼネレータに変わります。

12.10 エレメンタリ・ストリームの解析

エレメンタリ・ストリームは、トランスポート・ストリームが透過的に送信しなければならないペイロードです。トランスポート・ストリームにとって、エレメンタリ・ストリームが適合しているかどうかは問題ではありません。つまり、トランスポート・ストリームの適合性テストでは、エレメンタリ・ストリームが変化せずに運ばれているかどうかを確認します。エレメンタリ・ストリームが最初から正しく組合されているかどうかは関係ありません。

エレメンタリ・ストリームの構造、つまり構文は圧縮器の影響を受けます。したがって、エレメンタリ・ストリームのテストは、基本的には圧縮器のテストです。圧縮器が適合する構文を生成できても、音声や映像の品質が十分でないこともあります。しかし、構文が不正確であれば、デコーダでエレメンタリ・ストリームを解釈できなくなります。圧縮器は決定的ではなく、アルゴリズムに基づいているため、一般的でない動作モードが正確に実施されないと、エレメンタリ・ストリームが断続的に不適合になることもあります。

トランスポート・ストリームに含まれている各番組は、コーダが同じでないことも多いため、エレメンタリ・ストリームの問題

が1つの番組に限定されやすいのに対し、トランスポート・ストリームの問題は、すべての番組に影響を及ぼしやすくなります。ある特定のデコーダからの出力で問題に気付いた場合は、そのデコーダに対してSarnoff適合性テストを実施します。そして、テストに問題がなければ、障害は入力信号にあると考えられます。トランスポート・ストリーム構文をすでにテストしており、他の番組にも障害がない場合、エレメンタリ・ストリームの解析が必要になります。

エレメンタリ・ストリームの解析は、構文の最上位から下位へ向かって実施します。シーケンス・ヘッダは、圧縮で使用されるすべての関連モードやパラメータをデコーダに知らせるため、非常に重要です。セクション5.1と5.2のエレメンタリ・ストリームの構文を参考として使用してください。図12-15は、AD953型に表示されるシーケンス・ヘッダと拡張部を示しています。図12-16のように、低いレベルのテストでは、復号されたBフレームが動きベクトルと一緒に、画像に重ねて表示されます。

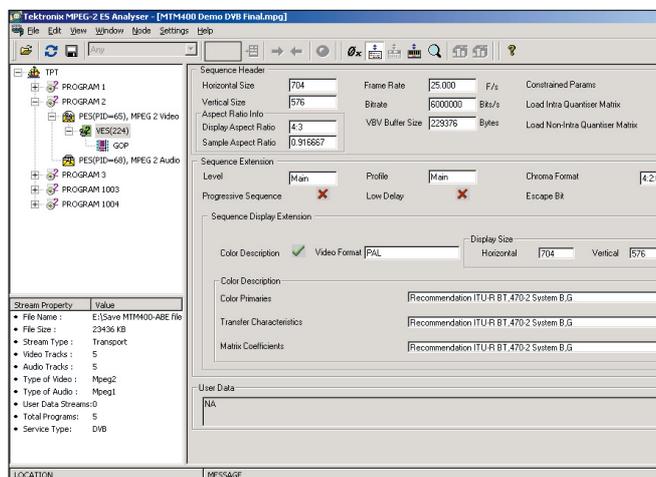


図12-15

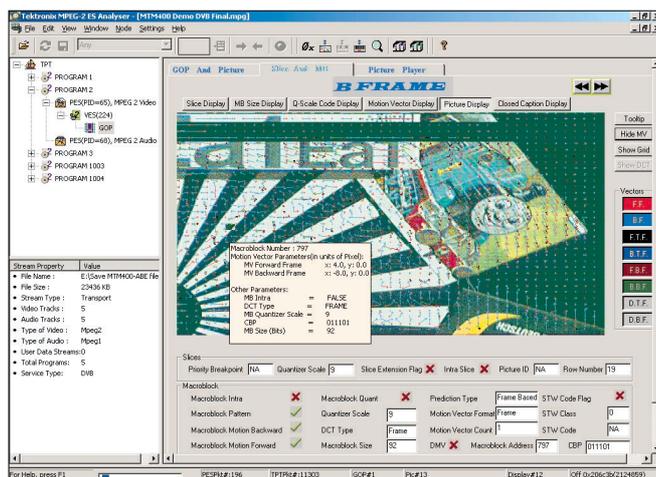


図12-16

12.11 トランスポート・ストリームの生成

デコーダが疑わしい場合、既知の品質のテスト信号を発生させることができれば便利です。図12-17のように、MPEGトランスポート・ストリームには、1つ以上のプログラム・ストリームを記述したPAT、PMT、NITなどの番組特定情報(PSI)が含まれていなければなりません。各プログラム・ストリームには、独自のPCRと、周期的にタイムスタンプがあるエレメンタリ・ストリームが含まれています。

DVBのトランスポート・ストリームには、さらにBAT、SDT、EITテーブルなどの番組配列情報が含まれています。PSI/SIエディタを使用すれば、適合するPSI/SIの組合せを選んでカスタムのテスト・ストリームに挿入することができます。

トランスポート・ストリームのレートは、各項目に分配する必要があります。マルチプレクサはレート・ゲージによって、使用されているビット・レート合計を表示します。残ったビット・レートには、すべて1のPIDを持つスタッフィング・パケットを挿入します。このパケットはデコーダで拒否されます。

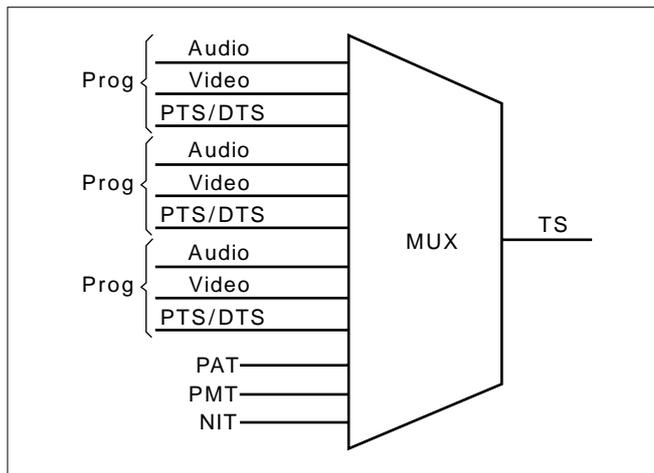


図12-17

12.12 PCR誤差の発生

MPEGデコーダは、PCRデータのクロック・サンプルによって連続クロックを再現し、位相ロック・ループを駆動する必要があります。ループは、フィルタリングとダンピングを行い、PCRデータ到着時のジッタによって、クロックが不安定にならないようにする必要があります。

位相ロック・ループの動作をテストするには、既知の誤差を持つ信号が必要で、それがないとテストは無効になります。AD953A型では、この目的で擬似的な誤差を発生させることができます。AD953A型はリファレンス・ゼネレータであるため、クロック回路はきわめて安定度が高く、実際の実出力ジッタもごくわずかです。ジッタの効果を発生させるために、PCRデータのタイミングは一切変えません。その代わりに、PCRの値に変更を加えることで、含まれているPCRカウントが理想の値から少し外れるようにします。値を変えると、デコーダで位相エラーが発生しますが、これは本物のジッタと見分けが付きません。

この方法の利点は、PCRデータに変更を加えるだけで、それ以外のデータには一切触れなくても、必要な規模のジッタを簡単にプログラム・ストリームに付加できることです。トランスポート・ストリームにある他のプログラム・ストリームにはジッタを加える必要はありません。実際、安定したプログラム・ストリームがあれば、基準として使用するのに最適です。

テストの目的に合わせて、ループ位相エラーの範囲を決める複数の方法でタイム・ベースを変調して、ループ・フィルタリングをテストします。

方形波ジッタは、増減が均一なジッタを発生させます。

正弦波ジッタの値は、サンプリングされた正弦波の位相エラーを発生させます。

ランダム・ジッタは、ノイズに似た位相エラーを発生させます。

用語解説

< A >

- AAC Advanced Audio Coding
 AAU オーディオ・アクセス・ユニット。アクセス・ユニットを参照。
 AC-3 Dolby LaboratoriesがATSCデジタル・テレビ規格用に開発した音声圧縮方式。一般の電化製品ではドルビー・デジタルと呼ばれている。
 A/D アナログ - デジタル・コンバータ
 AES Audio Engineering Society
 ANSI アメリカ規格協会
 API アプリケーション・プログラム・インタフェース
 ARIB 社団法人電波産業会
 ATM(非同期転送モード) ブロードバンド・デジタル・ネットワークにおいて、定速およびバースト性の情報を効率よく転送するためのデジタル信号プロトコル。ATMデジタル・ストリームは、「セル」と呼ばれる53バイト(1バイト=8ビット)の固定長パケット(5バイトのヘッダと、48バイトの情報ペイロード)で構成される。
 ATSC Advanced Television Systems Committee
 ATVEF Advanced Television Enhancement Forum
 AU アクセス・ユニット

< B >

- BAT ブーケ関連テーブル
 BER ビット・エラー率
 BFSK 2進周波数偏移変調
 BIOP Broadcast Inter-ORB Protocol
 BPSK 2進位相偏移変調

< C >

- CA 限定受信情報。番組にスクランブルがかけられているかどうかを示す情報。
 CAT 限定受信テーブル。PIDコード(「第8章 トランスポート・ストリーム」を参照)が1で、スクランブル方式についての情報が含まれているパケット。ECMおよびEMMを参照。
 CD コンパクト・ディスク
 CELP 符号励振型線形予測
 CIF Common Interchange Format。テレビ会議用の每秒30フレーム、352×240画素のフォーマット
 CORBA Common Object Request Broker Architecture
 COFDM 符号化直交周波数分割多重。OFDMの改良型。きわめてデータ・レートが低いキャリア波を非常に多く使用するデジタル変調方式。DVB-Tで使用される。
 CRC 巡回冗長検査

< D >

- DAC デジタル - アナログ・コンバータ
 DASE Digital TV Application Software Environment
 DAVIC Digital Audio Visual Council
 DCT 離散コサイン変換
 DDB ダウンロード・データ・ブロック
 DET データ・イベント・テーブル
 DFT 離散フーリエ変換
 DII Download Info Indication
 DSI Download Server Initiate
 DSMCC Digital Storage Media Command and Control

- DST データ・サービス・テーブル
 DTS 復号タイムスタンプ。アクセス・ユニットをいつ復号するかを示すPESヘッダの一部。
 DVB Digital Video Broadcasting。放送事業者、製造業者、規制機関などで構成され、デジタル・テレビやデータ・サービスの送信に関する標準化を進める欧州の団体。デジタル放送の規格には、DVB-C(ケーブル)、DVB-S(衛星)、DVB-T(地上波)の3種類がある。
 DVB-SI DVB番組配列情報。DVB多重化ストリームで運ばれ、他のストリームのコンテンツについて記述した情報。NIT、SDT、EIT、TDT、BAT、RST、STなどがある(「第10章 DVBおよびATSCの初歩知識」を参照)。
 DVC デジタル・ビデオ・カセット
 DVD デジタル・ビデオ・ディスク、またはデジタル多用途ディスク

< E >

- ECM 限定受信共通情報。制御ワードやその他ストリーム特有のスクランブル・パラメータを指定する限定受信情報。
 ECS Enhanced Content Specification
 EIT イベント情報テーブル。DVB-SIを参照。
 EMM 限定受信個別情報。特定のデコーダの認証レベルやサービスを指定する限定受信情報。個別のデコーダおよびデコーダ・グループの両方に対応できる。
 ENG 電子ニュース取材。ニュース報道において、フィルムに代わりにビデオ録画を使用することを表す用語。
 EOB ブロック終了
 EPG 電子番組ガイド。紙への印刷ではなく、データ転送によって提供される番組表。
 ETSI ヨーロッパ電気通信規格協会

< F >

- FEC 前方エラー訂正。メッセージに冗長性を付加することで、受信側でエラーをダイナミックに訂正できるようにする方式。
 FGS ファイン・グレイン・スケラビリティ

< G >

- GOP Group of Picture。伝送順では、GOPは1ピクチャで始まり、最後のピクチャの次にまた1ピクチャが来る。

< H >

- HAVI Home Audio/Video interoperability

< I >

- IEC 国際電気標準会議
 IOR Interoperable Object Reference
 IP Internet Protocol
 Iピクチャ ピクチャ内符合化ピクチャ
 IRD Integrated Receiver Decoder(デジタル放送受信機)。テレビ受信機をデジタル放送に対応させるために使用する、RF受信機とMPEGデコーダを組合せたもの。
 ISDB 統合デジタル放送。日本で開発されたデジタル放送方式。
 ISO 国際標準化機構
 ITU 国際電気通信連合

< J >

JPEG Joint Photographic Experts Group
 JTC1 IECの合同専門委員会
 JVT 合同ビデオ・チーム

< M >

MAC 媒体アクセス制御
 MDCT 修正離散コサイン変換
 MGT マスタ・ガイド・テーブル
 MHP Multimedia Home Platform
 MP@HL MainプロファイルのHighレベル
 MP@LL MainプロファイルのLowレベル
 MP@ML MainプロファイルのMainレベル
 MPE マルチプロトコル・カプセル化
 MPEG ISO/IEC JTC1/SC29/WG11のMoving Picture Experts Group、またはこのGroupが策定した各種の規格
 MPEG-LA MPEG License Agreement

< N >

NIT ネットワーク情報テーブル。1本のトランスポート・ストリームにあり、多くのトランスポート・ストリームについて記述した情報。
 NPT 通常再生時間
 NRT ネットワーク・リソース・テーブル

< O >

OCAP Open Cable Application Platform
 OFDM 直交周波数分割多重
 ORB Object Request Brokerage

< P >

PAL Phase Alternate Line
 PAT 番組関連テーブル。PID(「第8章 トランスポート・ストリーム」を参照)コードが0の packets にあり、MPEGデコーダがトランスポート・ストリームに存在する番組を判断するのに使用するデータ。PATはPMTを指示し、PMTが各番組の映像/音声/データのコンテンツを指示している。
 PCM パルス・コード変調。周期的な数値サンプルとして表される音声/映像信号などのアナログ・ソース波形に対して用いられる用語。PCMによって得られるデータは非圧縮デジタル信号。
 PCR プログラム・クロック・リファレンス。デコーダ・クロックと同期をとるためにプログラム・ヘッダで送られるエンコーダ・クロック・カウンタのサンプル。
 PES パケット化エレメンタリ・ストリーム
 PID 番組識別子。トランスポート・パケットのヘッダにある13ビットのコード。PID 0は、PATのPIDが含まれるパケットであることを表す(「第8章 トランスポート・ストリーム」を参照)。PID 1は、CATが含まれるパケットであることを表す。PID 8191(すべて1)は、ヌル(スタッフィング)パケットであることを表す。同一のエレメンタリ・ストリームにあるパケットのPIDはすべて同じになる。
 PMT 番組マップ・テーブル。トランスポート・ストリームの映像/音声/データのコンテンツを指示しているPATのテーブル。
 PS プログラム・ストリーム
 PSI 番組特定情報。MPEGトランスポート・ストリームのさまざまな番組や、各番組のエレメンタリ・ストリームを管理するための情報。PAT、PMT、NIT、CAT、ECM、EMMなどのPSIがある。
 PSI/SI MPEG PSIとDVB-SIを組合せた場合の、一般的な呼称
 PSIP Program and System Information Protocol

PTS 表示タイムスタンプ。表示ユニットが視聴者に表示される時間。
 PU 表示ユニット。1つ1つの圧縮画像や音声ブロック。

< Q >

QAM 直交振幅変調。デジタル変調方式の一種。
 QCIF Quarter Common Interchange Format。Common Interchange Format(CIF)の4分の1の解像度。CIFを参照。
 QMF 直交ミラー・フィルタ
 QPSK 4位相偏移変調(直交位相偏移変調とも呼ばれる)。特に衛星通信に適したデジタル変調方式。
 QSIF Quarter Source Input Format。Source Input Format(SIF)の4分の1の解像度。SIFを参照。

< R >

RLC ラン・レングス符合化。同様のビットを個別に送るのではなく、同様のビットがいくつあるかを送る符合化方式。
 RRT レーティング地域テーブル
 RST 進行状態テーブル
 R-S リード・ソロモン。トランスポート・パケットごとに最高8バイトを保護するため、DVBで採用されている多項式関数。

< S >

SAOL Structured Audio Orchestra Language
 SDI シリアル・デジタル・インタフェース。制作用のデジタル・ビデオ信号向けに開発された、シリアル同軸ケーブル・インタフェース規格。
 SDK Software Development Kit
 SDT サービス記述テーブル。トランスポート・ストリームに含まれるサービスのプロバイダをリストしたテーブル。
 SDTV 標準精細度(標準画質)テレビ
 SI DVB-SIを参照
 SIF Source Input Format。MPEG-1で使用される、解像度が半分の入力信号。
 SMPTE 映画テレビ技術者協会
 SNR 信号対ノイズ比
 SP@ML SimpleプロファイルのMainレベル
 SPTS シングル・プログラム・トランスポート・ストリーム
 ST スタッフィング・テーブル
 STB セット・トップ・ボックス
 STC システム・タイム・クロック。同じ番組の映像と音声の符合化に使用する共通のクロック。
 STT システム時間テーブル

< T >

TCP/IP Transmission Control Protocol/Internet Protocol
 TDAC 時間領域エイリアス・キャンセル。AC-3音声圧縮で使用される符合化技術。
 TDT 日時テーブル。DVB-SIで使用される。
 TOT 時間オフセット・テーブル
 T-STD Transport Stream System Target Decoder。エンコーダを想定した一定量のバッファ・メモリを持っているデコーダ。
 TVCT 地上仮想チャンネル・テーブル

< V >

VAU ビデオ・アクセス・ユニット。プログラム・ストリームで送られる1つ1つの圧縮ビクチャ。
 VBV ビデオ・バッファ・ペリファイ
 VCO 電圧制御発振器
 VLC 可変長符合化。生起率が高い値に短い符号を割当て、生起率が低い値に長い符号を割当てる圧縮技術。
 VOD ビデオ・オン・デマンド。テレビ番組や映画を、各ユーザが見たい時に配信するシステム。

用語解説

VSF 残留側波帯変調。ATSCで採用されているデジタル変調方式。

<Y>

Y/C 輝度と色度

<ア行>

圧縮 データを表現するためのビット数を減らすこと

アクセス・ユニット ピクチャや音声ブロックとそれに続くスタッフティング(ヌル値)の符合化データ

アンカ・フレーム 予測に使用するビデオ・フレーム。アンカ・フレームには、一般にIフレームとPフレームが使用され、Bフレームは決して使用されない。

インターリーブ パースト・エラーを多数のエラーに細分するエラー訂正で使用する技術

ウェーブレット 固定長でなく、周波数の低下につれて長くなる基底関数を用いた変換

動きベクトル 予測の基準ピクチャの縦方向と横方向のずれを表す一組の数字

エレメンタリ・ストリーム 1種類のビデオ/オーディオ信号を運ぶ、圧縮器からの未処理出力

エントロピー符合化 信号のデジタル表現を可変長ロスレス符号化して、冗長性を減らすこと

<カ行>

加重 前もって値を掛けておくことで、切捨てによるノイズの分布を変える方法

画素 画像素子、ピクセル。画像の最小単位で、単体あるいはGBRやYCrCbのように組合せて表される。

切捨て サンプルや係数の低位ビットを取り去って、ワード長を短くすること

クローズドGOP 最後のピクチャが、双方向符合化のために次のGOPのデータを必要としないGOP。クローズドGOPを使用して、ビット・ストリームにスプライス・ポイントを作成する。

係数 周波数の振幅を指定する数字、または変換における基底関数

<サ行>

シンδροーム エラー検査計算の最初の答え。一般に、シンδροームがゼロの場合はエラーがないと見なされる。

スケーラビリティ 何層もの映像データによって複数の品質レベルを提供するMPEG-2の特性。データが多層化されているため、複雑なデコーダは多くの層のデータを使用して品質の高い画像を生成でき、単純なデコーダは第1層のデータだけでも画像を生成できる。

スタッフティング 一定のビット・レートを維持するために加えられる意味のないデータ

スライス 連続するマクロ・ブロックのシーケンス

<タ行>

チャンネル符合化 生データを、有線/無線で録画または伝送できる信号に変換すること

トランスポート・ストリーム(TS) パケットで運ばれる複数のプログラム・ストリームを多重化したもの。逆多重化は別のパケットID(PID)で行う。PSI、PAT、PMT、PCRを参照。

ドルビー・デジタル AC-3を参照

<ハ行>

ヌル・パケット データは含まれていないが、ペイロードが変動するビット・レートを一定に保つために必要な「スタッフティング」パケット。ヌル・パケットのPIDは常に8191(すべて1)になる(「第8章 トランスポート・ストリーム」を参照)。

パケット 2通りの意味で用いられ、プログラム・ストリームでは、1つ以上の表示ユニットが含まれている単位、トランスポート・ストリームでは、一定のサイズに細分されたデータ量を指す。

ハフマン符合化 長さの異なるコードを使用して発生確率が不均等なシンボルを表す、ソース符合化の一種

ピクチャ間符合化 連続するピクチャ間の冗長性を利用する圧縮。時間符合化とも呼ばれる。

ピクチャ内符合化 完全に1つのピクチャ内で行う圧縮。空間符合化とも呼ばれる。

ビット・レート 圧縮ビット・ストリームが、チャンネルからデコーダ入力端子へ送られる速度

ブーケ 運ばれる番組がネットワークIDとPIDの組合せて識別される、トランスポート・ストリームのグループ(DVB-SIの一部)

プログラム・ストリーム 圧縮された映像/音声/タイミング情報が含まれているビット・ストリーム

ブロック 画素値またはDCT係数を、通常は縦横8つずつ(8×8)配列し、輝度や色度の情報を表したもの

プロファイル 使用する符合化構文を指定する。

<マ行>

前処理 MPEG符合化の前に行われるビデオ信号処理。たとえば、ノイズ削減、ダウン・サンプリング、カット・エディット、3:2ブルダウンなどが前処理として行われる。

マクロ・ブロック 複数の輝度/色差DCTで表され、1つの動きベクトルによって動き補正される画面の領域

マスキング ある一定の音が別の音の存在によって聞こえなくなる心理聴覚現象

<ラ行>

量子化 所定の値の1つを割当てて、信号値の近似値を求める処理

レベル プロファイルと組合せて使用する入力ピクチャのサイズ(「第2章 映像の圧縮」を参照)

Tektronix お問い合わせ先:

アメリカ 1 (800) 426-2200
アメリカ (輸出販売) 1 (503) 627-1916
イタリア +39 (02) 25086 1
インド (91) 80-2275577
英国およびアイルランド +44 (0) 1344 392400
オーストリア +43 2236 8092 262
中央ヨーロッパおよびギリシャ +43 2236 8092 301
オランダ +31 (0) 23 569 5555
カナダ 1 (800) 661 5625
スウェーデン +46 8 477 6503/4
スペイン +34 91 372 6055
大韓民国 82 (2) 528-5299
台湾 886 (2) 2722-9622
中華人民共和国 86 (10) 6235-1230
デンマーク +45 44 850 700
ドイツ +49 (221) 94 77 400
東南アジア諸国/オーストラリア/パキスタン (65) 6356-3900
日本 81 (3) 6714-3010
ノルウェー +47 22 07 07 00
フィンランド +358 (9) 4783 400
ブラジルおよび南米 55 (11) 3741 8360
フランスおよび北アフリカ +33 (0) 1 69 86 80 34
ベルギー +32 (2) 715 89 70
ポーランド +48 (0) 22 521 53 40
香港 (852) 2585-6688
南アフリカ +27 11 254 8360
メキシコ、中米およびカリブ海諸国 52 (55) 56666-333
ロシア、その他の旧ソ連共和国およびバルト海諸国 +358 (9) 4783 400
その他の地域からのお問合せ: Tektronix, Inc., USA 1 (503) 627-7111

詳細について

当社ホームページ(www.tektronix.com または www.tektronix.co.jp)をご参照ください。



Copyright © 2003, Tektronix, Inc. All rights reserved.
Tektronix製品は、米国およびその他の国の取得済みおよび出願中の特許により保護されています。本書は過去に公開されたすべての文書に優先します。製品の仕様と価格は予告なく変更する場合があります。TEKTRONIXおよびTEKは、Tektronix, Inc.の登録商標です。その他本書に記載されている商品名は、各社のサービス・マーク、商標または登録商標です。

12/03 HB/PT 25W-11418-4

Tektronix

Enabling Innovation

日本テクトロニクス株式会社

東京都港区港南2-15-2 品川インターシティ B棟6階 〒108-6106
製品についてのご質問・ご相談は、お客様コールセンターまでお問い合わせください。

TEL 03-6714-3010 FAX 0120-046-011

電話受付時間 / 9:00 - 12:00・13:00 - 18:00 月曜 - 金曜(祝日は除く)

当社ホームページをご覧ください。 www.tektronix.co.jp
お客様コールセンター ccc.jp@tektronix.com