

デジタル画像圧縮の種類としくみ

YITコンサルティング 石川 安則

従来放送局など制作現場においては、デジタルVTRの導入により映像・音声信号のデジタル化がすでに行われていた。近年、家庭用ビデオへのDV方式導入に始まり、パソコンの急速な普及に伴ったDVD、デジタルオーディオの浸透、デジタル放送の本格的開始と、民生用途においてもデジタル化が急激に進行している。

膨大な情報量を有するデジタル映像・音声データを効率的に扱うためには、デジタル圧縮技術の利用が必須である。たとえば2時間の映画は、映像データだけでも200Gバイト以上の情報量をもち、記録容量4.7GバイトのDVDに納めるために、MPEG 2を利用して1/50以上に圧縮している。また、音声データではMP3などの圧縮技術を用いて1/10以下のデータ量に圧縮し、携帯デジタルオーディオ機器でもCDと同等な録音時間を確保している。

この項目では、こうした状況を鑑み、さまざまな映像・音声圧縮方式を概観したあと、圧縮技術の前提となる映像・音声信号のデジタル化の手順やメリットを簡単に説明する。そして、動画像圧縮方式の主流であるフレーム間圧縮方式の基本的考え方と、これを応用した標準動画像圧縮方式であるMPEGについて詳述する。最後にDVやJPEGなどフレーム内圧縮方式の動画像への応用について解説する。

デジタル圧縮の方式と用途

ビデオ制作に用いられる映像・音声信号の1次入力信号は、アナログ信号である。映像信号のほとんどは、CCDなどの固体撮像素子からアナログ信号として出力され、通常はカメラ内で直ちにAD変換器によってデジタルデータに変換され、SDI出力などに非圧縮データとして出力されるか、または圧縮動

画像データとしてディスクまたはテープに記録される。音声信号は比較的狭帯域なので、マイクロホンから出力されたアナログ信号は、ケーブル経由でレコーダーなどの装置に入力された後AD変換されるケースや、カメラ一体型のレコーダー内でAD変換されて非圧縮デジタルデータとして記録されるケースなどがある。

ビデオ制作においては、データ量が大きく高い圧縮率が必要なときは、MPEG 2が利用されることが多いが、画質優先の場合はDV方式やMPEG 2のフレーム内圧縮モードが用いられ、音声データは、非圧縮のまま伝送・記録される場合が多い。

また、コンテンツ配信においても、放送やパケットメディアではMPEG 2が利用されているが、ネットワーク配信などには、より高い圧縮率が得られるMPEG 4が主流となっている。この場合、音声データも、MP3やAACなどの高能率な圧縮方式が利用される。

表1 代表的な圧縮方式(動画像)

方式名	技術・特徴	用途	その他
H.261	フレーム間圧縮 DCT、ハフマン符号化	テレビ電話/会議	最初の動画像圧縮方式
H.263	フレーム間圧縮 H.261の効率向上	テレビ電話/会議	MPEG 4の基礎となる
MPEG 1	フレーム間圧縮 双方向予測	Video-CD インターネット	蓄積メディア用に開発
MPEG 2	フレーム間圧縮 標準TV方式に対応	DVD デジタル放送	放送/通信/蓄積メディア用
MPEG 4	フレーム間圧縮 オブジェクト符号化	インターネット 携帯電話	合成画像(CGなど)にも対応
DV	フレーム内圧縮 DCT	デジタルビデオコーダー	テープメディア向け
M-JPEG ※	フレーム内圧縮 DCT、ハフマン符号化	インターネット デジタルカメラ(動画)	AVIファイルで利用
M-JPEG2000 ※	フレーム内圧縮 DWT、算術符号化	インターネット	
Nancy	—	携帯電話	利用技術非公開

※Motion-JPEG、Motion-JPEG2000

方式名	技術・特徴	用途	その他
MPEG 1/2 Audio Lavel I	聴覚心理モデル 適応ビット割り当て	Video-CD CSデジタル放送	MPEG 2 Audioは、 MPEG 1 Audioの拡張。 5.1chなどに対応
MPEG 1/2 Audio Lavel II	聴覚心理モデル 適応ビット割り当て	Video-CD CSデジタル放送	
MPEG 1/2 Audio Lavel III (MP3)	MDCT、聴覚心理モデル ハフマン符号化	インターネット 携帯プレーヤー	
MPEG 2 Audio AAC	MDCT、聴覚心理モデル 予測モデル、ハフマン符号化	デジタル放送(日本) 携帯電話	MPEG 4にも適応
MPEG 4 Audio	AAC、TwinVQ CELP、TTS	インターネット 携帯電話	種々の技術を含む
Dolby Digital (AC-3)	MDCT、聴覚心理モデル ハフマン符号化	DVD デジタル放送(米国)	Dolby独自方式
ATRAC 3	MDCT スペクトル分離符号化	MD (Mini Disk)	MD専用方式

表2 代表的な圧縮方式(音声)

表1、2に代表的な動画および音声データ圧縮方式の一覧をまとめる。

映像・音声信号のデジタル化と圧縮

■ デジタル化とその利点

デジタル化は、

- ①標本化(sampling)：一定間隔で信号値を取り出す
 - ②量子化(quantizing)：ある範囲内の整数値に丸める
 - ③符号化(coding)：2進数で表現する
- という手順により行われる。

音声信号は、時間軸に沿って一定の時間間隔で標本化を行い、得られたアナログ標本値を、必要なビット数のデジタル値に等間隔に分割する線形量子化が行われ、デジタル値を2進数に割り当てることで符号化が

行われる(図1)。一方、映像信号は、従来のアナログビデオ信号でも時間軸方向に標本化を行っている。たとえばNTSC方式では、1/29.97秒(正確には1001/30000秒)の間隔で標本化されたアナログ静止画(これをフレームと呼ぶ)を順に表示する。デジタル化は、各フレームの2次元平面上的アナログ値を、

- ①縦・横に等間隔に分割して最小単位の「画素」を得る(空間領域における標本化)
 - ②画素値(アナログ値)の線形量子化を行ってデジタル値を得る(カラー画像はRGBの3原色ごとに行う)
 - ③デジタル値の2進数を符号とする
- という手順で行われる(図2)。

デジタル化によるメリットとしては、一般に以下のような点が挙げられる。

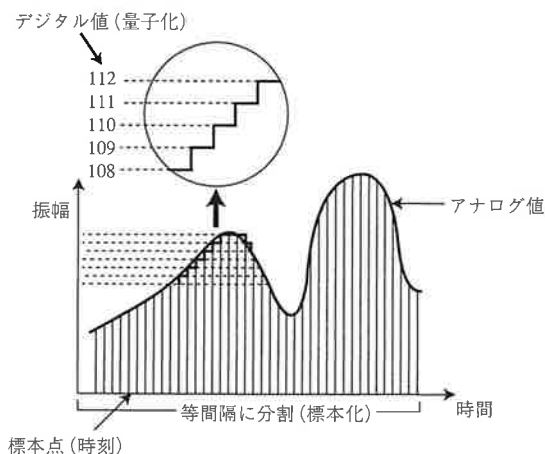


図1 音声信号のデジタル化

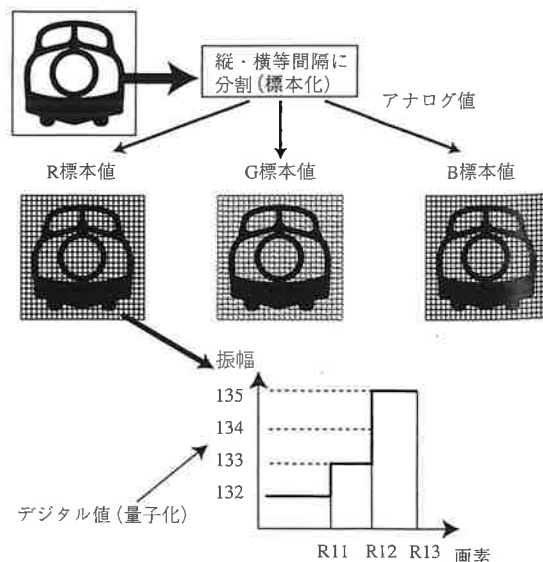
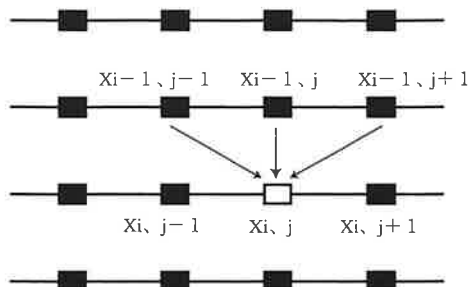


図2 映像信号のデジタル化

図 3 予測モデル

(a) 隣接画素との差分を求める



(b) 差分値の分布 (模式図)

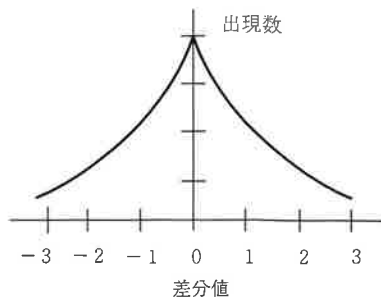
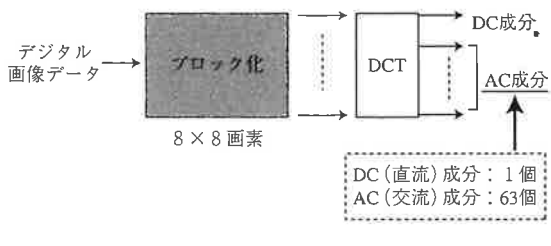
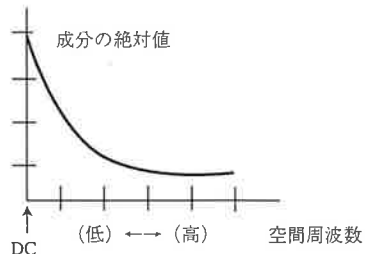


図 4 変換モデル

(a) 周波数成分への変換



(b) 周波数成分の分布 (模式図)



- ・伝送、蓄積特性が優れる (劣化がない、多様なメディアに対応)
- ・信号のリアルタイム性からの解放 (伝送方法の自由度、編集性)
- ・コンピュータ処理が容易 (データ加工・多重化が容易)

■ デジタル圧縮の基本的なしくみ

デジタル情報圧縮は、原理的に「モデリング」と「符号割当」に分けられる。

情報系列から有意な情報のみを抽出する操作がモデリングであり、デジタル画像のある画素と、その左隣の画素の値は、通常非常に近い値をとり、画像全体で差分値の分布をとるとゼロ付近に値が集中する (図 3)。これを「予測モデル」といい、時間軸方向に利用すれば、隣り合ったフレーム間で同じ位置の画素値の差分をとるフレーム間予測モデルが得られる。

また、DCT (離散コサイン変換) や DWT (離散ウェーブレット変換) といった信号変換処理により得られる周波数分布は、画像データでは周波数 (空間周波数) が “0” (直流成分) 付近に集中する (図 4)。これを「変換モデル」といい、音声データの場合は周波数 (時間周波数) がピッチ周波数 (基本周波数) 付近に集中する。

さらに、モデリングのステップで、人の知覚特性を利用した情報量の削減ができる。たとえば、人の聴覚特性には絶対可聴域が存在し、音声周波数により知覚できる信号レベルの最低値

が決まっている。また、強い信号 (大きな音) に近い周波数の信号は知覚されないという特性のマスキング効果を利用し、不要な情報を削減していくことが可能である (符号化が不要)。映像信号においても、視覚特性上、種々の情報削減ができる。たとえば、画像データにおける高周波数成分も視覚的感度が低いため削減が可能となる。この結果、低周波数成分への集中がより強まる。情報の集中とは、いい換えると事象の出現確率が事象ごとに偏った分布をすることである。

「符号割当」においては、この性質を利用し、出現確率の高い事象に短い符号を、低い事象に長い符号を割り当てるようにして平均符号量を少なくする。簡単

表 3 符号割当の簡単な例

数	2進数符号	出現確率	符号	符号ビット数
0	000	1/2	1	1
1	001	1/4	01	2
2	010	1/24	001	3
3	011	1/24	0001	4
4	100	1/24	00001	5
5	101	1/24	000001	6
6	110	1/24	0000001	7
7	111	1/24	00000001	8

↑
符号長 = 3

↑
平均符号長 = 2.38

平均符号長 = $\sum \{(\text{出現確率}) \times (\text{符号ビット数})\}$

$$\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{24} \times 3 + \frac{1}{24} \times 4 + \frac{1}{24} \times 5 + \frac{1}{24} \times 6 + \frac{1}{24} \times 7 + \frac{1}{24} \times 8 = 2.38$$

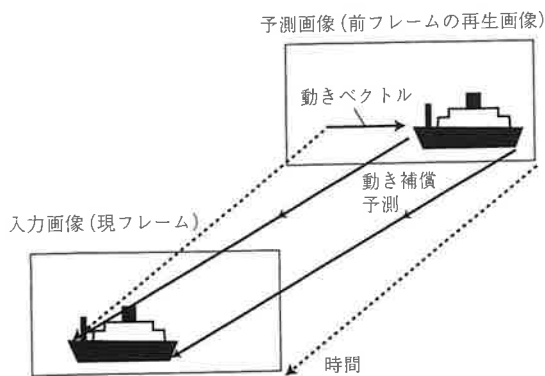


図4 動き補償予測

な例を表3に示す。

フレーム間圧縮を利用した動画圧縮方式

■フレーム間圧縮の基本的考え方

前述したように、動画は従来のアナログ的表現（テレビジョンや映画）においても、時間軸上で離散的な間隔で標本化されており、標本化された画像は1枚の静止画とみなすことができる。一般にそれぞれの静止画をフレームと呼び、また1秒あたりに表示されるフレームの数をフレームレートという。なお、動画がなめらかに動いて見えるフレームレートの下限は、20フレーム/秒（fps）程度といわれている。フレームレートは、NTSC方式では29.97fps、ヨーロッパにおいて主流のPAL/SECAM方式では25fps、また映画は、周知のように24fpsである。

ある程度以上のフレームレートを有する連続した静止画の間には、フレーム間予測モデルが成り立つことはすでに述べた。極端な場合として、同一の静止画を連続して表示すれば、最初のフレームだけが個々の画素値として意味をもち、その後のフレームはすべて同じという情報を送ればよい。つまり一般には、動きのある画像では、動いた部分だけ画素値の情報を送り静止部分は前フレームと同じという情報を送ればよいことになる。

MPEGなど動画圧縮方式においては、上記したフレーム間予測に加え、さらに動き補償予測を行う。図4に示すように、動物体（画像の一部）が剛体であって、フレームの間で平行移動したと仮定すると、その物体は1つのベクトル情報を送るだけで、つぎのフレームを再現できる。実際には、あるフレームのブロックと呼ばれる矩形領域の画素値と、隣接するフレームのブロックの画素値を比較して、最も誤差の少ないブロックに移動したとして、ベクトル情報を得る。

■国際標準動画圧縮方式～MPEG

MPEG 2の圧縮手順を以下に詳細に説明する（図5）。

①動き補償フレーム間予測とGOP

MPEGでは連続するフレームをGOP（Group of Picture）と呼ぶ1つの連続フレーム単位に分割する。各フレーム（ピクチャーという）は、フレーム間予測を行わないIピクチャー、前のフレームからの予測を行うPピクチャー、および前後のフレームからの予測を行うBピクチャーで構成されている（図6）。GOPの先頭フレームは、かならずIピクチャーであり、途中で生じた誤差のフレーム間予測による伝搬を止めている。

GOPのピクチャー構成法は、動画の圧縮率（すなわち画像品質）に多大な影響を与え、対象となる動画データごとの最適なピクチャー構成法設計がエンコーダー性能の指標となる。

②DCT（離散コサイン変換）（図7）

動き補償フレーム間予測で得られた画像データは、つぎに8×8画素ブロックに細分化され、DCTが施される。一般の画像データでは、非常に大きい直流成分（DC成分）と空間周波数が高くなるに従って急激に振幅が減衰する交流成分（AC成分）に変換される。

③量子化

DCTは、元の画素値が離散値（一般に8ビット）であっても実数計算が行われるため、計算結果の周波数成分（DC成分、AC成分）は、アナログ値となる。これを再び、量子化し、離散的なデジタル値を得る。この際に視覚特性を考慮して、低周波数成分は細かく、高周

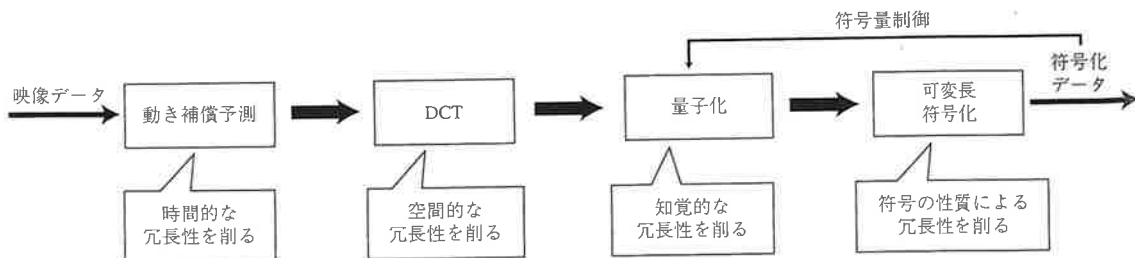


図5 MPEG方式のデータ圧縮の流れ

波数成分は粗く量子化する非線形量子化を行う。量子化の結果、圧縮データから元の画像を復元しても完全には元に戻らない非可逆圧縮が行われる。

実際の画像データについて非線形量子化した結果は、低周波数側の数個の成分を除くと、ほとんどゼロとなり、大幅な情報量削減が行われることがわかる。ただし、細かい絵柄の画像など、高い周波数成分を多く含む画像データでは、削減しすぎてはずみの目立つ復元画像になってしまう。

④ハフマン符号化(符号割当)

DC成分については 8×8 ブロックごとに左隣のブロックからの差分値に対して定められた符号を割り当て、AC成分については、ブロック内で低周波数成分から高周波数成分の順に列に並べ替え、現れる成分の値に従って符号を割り当てる(図8)。

より詳細には、成分ブロックに数種類ある符号化モードの1つが決められ、符号化モードによって成分の符号割り当て方法が少しずつ異なる。また、動きベクトル値や符号化モードについても符号が割り当てられており、これらの符号と成分の符号を合わせて圧縮データとなる。

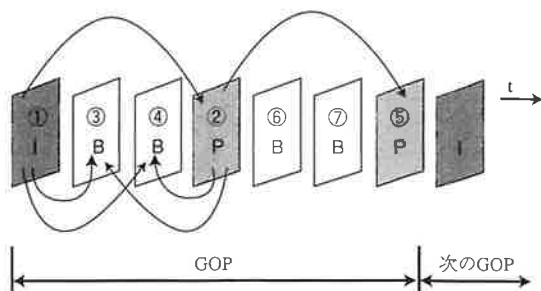


図6 GOPの構成例

MPEG 1やMPEG 4についても、基本技術はほぼ同じである。MPEG 4では、動き補償範囲の拡大、フレーム内予測モードの拡張、ハフマン符号化の拡張、オブジェクト符号化(1フレーム内を分割して異なる符号化モードで圧縮する)などの拡張が行われ、圧縮効率や機能性が向上している。

フレーム内圧縮の動画像への応用

■フレーム内圧縮とフレーム間圧縮

DVDやデジタル放送におけるMPEG 2、インターネットや携帯電話のムービー配信に使われるMPEG 4は、

図7 DCTの基底画像

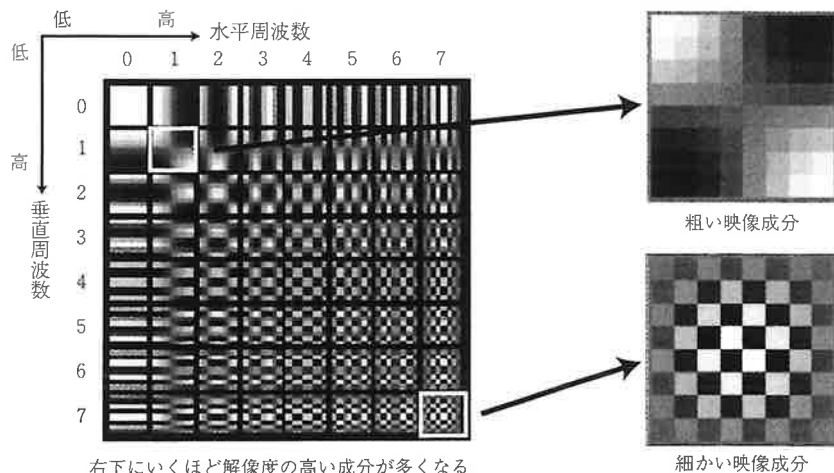


図8 ハフマン符号化

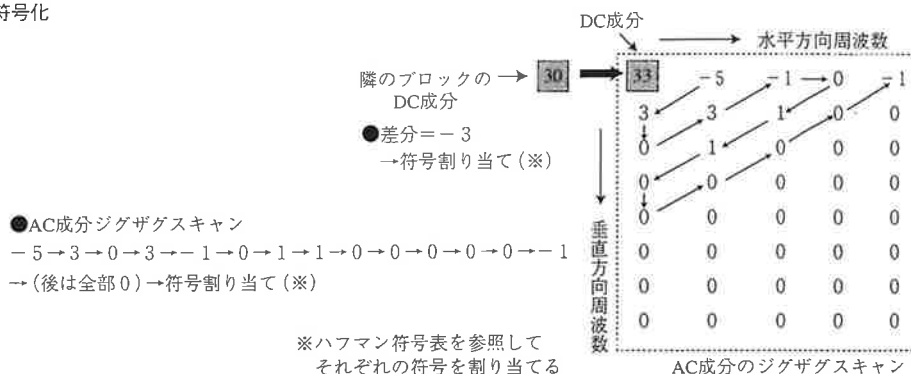


表4 フレーム内圧縮とフレーム間圧縮の利点比較

フレーム内圧縮	フレーム間圧縮
・圧縮・伸長が低コスト ・編集作業が容易・低コスト (独立フレーム単位で処理) ・遅延が少ない (リアルタイム性) ・証拠性の問題がない (前後のフレームからの予測を使わないため)	・圧縮率が高い (同じビットレート・ファイル量で高画質)

フレーム間圧縮方式が基本である。これに対し、民生用として発展してきたDV方式や、デジタルカメラの動画記録に用いられるMotion-JPEG (M-JPEG) 方式は、フレーム間予測を用いずフレーム内圧縮画像を静止画として連続的に表示することで動画画像圧縮を達成している。

一般に、動画画像圧縮方式としては、フレーム間圧縮を利用するほうが圧縮効率が高いとされ、同じビットレートでの配信やファイル容量ではフレーム間圧縮方式のほうが画質が良くなる傾向にある。しかしながらフレーム間圧縮方式は、差分予測のための大量のフレームバッファや高速処理リソースを必要とし、高コストとなってしまうがらである。さらに、任意の時間位置フレームを編集しようとした場合、一端完全に復元画像を再生する必要がある、編集用の記憶領域確保などのコストアップ要因が生じる。

また、特定の用途においては、短いフレーム単位で前後の情報をいわずに画像情報を保存できることが、情報の証拠性の確保に必要とされ、このような場合、フレーム間圧縮では信頼性が乏しいとされている。さらに再生の遅延の問題も指摘されている。このため、MPEG 2のフレーム内圧縮モード(MPEG IMX方式)や、携帯電話における動画配信のようなMPEG 4のフレーム内圧縮モードなど、圧縮率を犠牲にしたMPEG方式の利用形態もある。

一方、DV方式やMotion JPEGなど、フレーム内圧縮方式は、低コストなうえ、編集の容易性などの利点がある。表4に、フレーム間圧縮とフレーム内圧縮方式の利点比較を示す。

■フレーム内圧縮の手法～Motion JPEG2000

従来のJPEG方式の問題点を改善し、新たな要求を満たすべく定められた静止画圧縮国際標準方式JPEG2000が注目されている。以下に、その特徴を示す。

①高圧縮率時の画像品質

JPEGは、高圧縮時においてブロック歪みやモスキ-

ト歪みと呼ばれる人工的歪みが目立つ。JPEG2000で採用されたDWTは、本質的にブロック歪みが生じない。また、符号化方式にはEBCOT (Embedded Block Coding with Optimized Truncation) と呼ばれる処理と算術符号というそれぞれ複雑な処理が必要だが、圧縮効率の高い方式が採用され、高圧縮時の画質改善に貢献している。

②豊富な機能性

DWT、EBCOTの採用により、以下のようなJPEGにない機能性が得られる。

・解像度スケーラビリティ

DWTの特性により、1/2、1/4、1/8といった解像度の復元画像が瞬時に再生できる。また、プログレッシブ再生も容易。

・ロスレス圧縮(可逆圧縮)とロッキー圧縮(非可逆圧縮)がシームレス

高い圧縮率から低い圧縮率の非可逆圧縮へ、さらに可逆圧縮まで連続的に制御できる。

・レート制御

要求される圧縮率を得るために、JPEGでは2パス以上の処理が必要だが、JPEG2000は1パスで可能。

・ランダムアクセス

圧縮データのまま、画面上の任意位置にアクセス可能である。JPEGでは復元画像を再生しないと任意位置にはアクセスできない。

③動画画像圧縮への対応

JPEGは国際標準規格ではあるが、動画画像圧縮の規定は存在せず、現状ではM-JPEGと呼ばれる圧縮方式は、各ベンダーによってファイルやストリームに互換性がない場合が多い。これに対して、M-JPEG2000は、JPEG2000規格のPart. 3として標準化されており、規格上の互換性が完全に保証されている。

M-JPEG2000は、基本的な圧縮システムはすべて静止画の規格(JPEG2000規格Part. 1およびPart. 2)に依存しており、これらの特徴や機能性がすべて継承されている。フレーム内圧縮方式の優位性に加え、こうした特徴や機能性が利用できれば、動画画像圧縮方式として非常に利用価値の高い方式となる可能性がある。

※参考文献

情報圧縮技術(IDGジャパン)

音声符号化(電子情報通信学会)

国際標準画像符号化の基礎技術(コロナ社)

日経エレクトロニクスNo.783「きれいな画像にJPEG-2000」(日経BP)

Inside Macintosh "QuickTime File Format" (Publication Draft/Apple Computer Inc.)

Contents

The Basics of Video Technique

Introduction／

- 5 テレビ番組制作の現状とプロセス

須川 貴夫

Chapter 1／テレビ放送のしくみとビデオ信号

- 12 テレビの基礎～アナログ放送からデジタル放送へ
16 NTSC信号の基礎
19 ビデオ信号の基礎
26 SDTVとHDTV

藤田 淳
藤田 淳
藤田 淳
石原 渉

Chapter 2／ビデオカメラのしくみと使い方

- 32 ビデオカメラのしくみ
38 ズームレンズのしくみと特性
43 ロケーションにおけるハンディカメラの調整方法と使い方
50 スタジオカメラの調整方法と使い方

本郷 俊行
岡村 征夫
和田 正夫
川崎 淳

Chapter 3／ビデオレコーダーのしくみと記録フォーマット

- 56 VTRのしくみ
64 ハードディスク&光ディスクレコーダーのしくみ
69 デジタル画像圧縮の種類としくみ
75 VTRの変遷と種類

龍口 健二
石川 安則
石川 安則
服部 多栄子／小浦 壽三

Chapter 4／モニターのしくみと映像管理

- 86 カラーモニターのしくみ
91 波形モニター&ベクトルスコープの種類と特徴
96 波形モニター&ベクトルスコープの使い方

森 千重
久保木 啓介
大谷 敬亮

Chapter 5／撮影システムを理解する

- 104 ENGの撮影システムを理解する
109 スタジオ収録システムを理解する
115 ドラマ撮影のスタイルを理解する
121 撮影システムに使用するケーブル&コネクター
126 システム運用における同期信号の基礎
132 HD/SD一体化制作の実際

安部 裕
小山 雅典
市川 隆男
木川 豊
徳永 朱美
小笠原 洋一

ビデオ技術マニュアル

発行人●岩瀬 雄一

編集人●竹村 元子

印刷●図書印刷 株式会社

発行所●株式会社写真工業出版社

〒101-0051東京都千代田区神田神保町 2-10-31

電話03 (3556) 2321(代) FAX03 (3238) 2123

振替00120-5-14046

発行●2004年 7 月31日