マルチメディアと B-ISDN の将来ビジョンとその研究課題

富永　英義

早稲田大学 理工学部 電子通信学科

画像は高度に構造化されたものと捉え，計算機による画像の認識・理解という立場から構造を自動的に抽出するための研究が長年行なわれてきたが，符号化の分野においても画像の構造を利用する構造情報符号化の研究が盛んになってきた．これを動画像情報に拡張し，構造情報の時間軸方向への連続性を活かすためには，多フレームにわたって安定に画像の領域を追跡する必要がある．その基本的な考え方は動画像の構造抽出，すなわち，画像中に存在する物体(領域)の切り出しを行いその構造を解析して，構造記述データの符号化を行うことにある．しかし，自然動画像の構造抽出および解析は，いまだコンピュータビジョンの分野で発展途上の研究であり，動画像中の構造を安定に抽出することが，現在直面している研究課題として重要である．

動画像の符号化は必ずしも情報圧縮を主体とする符号化である必要はなくなりつつある．特に，これからの動画像符号化は，原信号を利用形態に最も適したフォーマットへ変換する技術に向けて発展していくと思われる．その利用形態として重要になるのは編集・加工であると考えられる．このような高度に柔軟な情報加工が可能となれば，マルチメディアを統一的に扱う環境が提供された場合など，メディア間で編集・加工処理が柔軟に行え，より豊かな情報伝達手段が与えられることになる．この様な観点から，動画像に対してより自由な情報の編集・加工を可能にするためのデータアーキテクチャおよび必要な技術の課題を整理している．

On the Studies of Multimedia Document Architecture for B-ISDN

Hideyoshi TOMINAGA

Dept. of Electronics & Telecommunication Eng., WASEDA University

3-4-1 Ohkubo, Shinjuku-ku, Tokyo, 169 Japan

Given a high performance computer and a high throughput digital network in the future, the main objective of the image coding must not be only to reduce the redundant of image. The next generation of video coding may be developed to transform the original video data into the data structure suitable for an effective use in reconstructing image. One of the most effective usages is thought to be the processability of the coded representation of video. When the flexible processing onto video image would be realized, one of more efficient media should be given to transmit any information. From such a point of view, the forth one of purposes is to propose the video data structuring model to realize such a flexibly processable video.

From the view point of both communication and information processing, I shall point out the problems in video coding scheme and representation method on considering the spatial and temporal structure of video data for efficient video manipulation.

1 はじめに

ISDN (Integrated Services Digital Network) の実用化に伴って，テレビ会議システムあるいはテレビ電話といった動画像を対象とした通信サービスのための符号化方式および蓄積媒体への蓄積を目的とした符号化方式の標準化作業は第一段階の勧告案 [1][2] が固まり実用化が始まっている．これらは，高能率な動画像符号化方式が開発されただけでなく，個体素子デバイス技術及び LSI 技術の飛躍的な発展によって高速演算処理プロセッサが開発されたことではじめて可能になってきた．つまり，複雑でかつ大量のデータをリアルタイムで処理することが可能となり，動画像の高能率符号化のような複雑な処理も実現できるようになったためである．

また，汎用 CPU の処理能力の向上とメモリ規模の拡大により，ワークステーションなど比較的小規模な計算機上で，画像・音声・データなどの複数の情報を容易に扱うことが可能になったという背景から，それらの情報を有機的に関連付けて扱うことを可能にするための「マルチメディア」システムに関する議論も活発化している．このマルチメディアシステムには，主としてワークステーション上で，アプリケーションの形でディジタル蓄積メディア (Digital Storage Media, DSM) に記録された動画・静止画・音声・データを扱うものが考えられている．中でも動画像は視覚に直接訴えるメディアとして非常に重要な位置を占める．人間は日常生活において映像に常に接しており，その中に含まれる情報をごく自然に吸収している．いわば最も馴染みの深いメディアである．

近い将来，基盤技術として B-ISDN や大容量蓄積メディアが普及し計算機の計算速度が向上するのは確実であり，動画像の符号化も必ずしも情報圧縮だけを主目的とするものでない，多様な可能性が考えられる．これからの動画像符号化は，原信号を利用形態に最も適したフォーマットへ変換する技術に向けて発展していくと言われている．その利用形態として今後特に重要になるのは利用者における編集・加工の容易性であると考えられる．

特に，動画像をマルチメディア情報の一つとして捉えて，文書情報の一部として扱われることも考えられ，より高度な編集・加工が要求されるであろう．その要求の幅は広く，フレームを単位とした操作から，他メディアと協調させながらの操作，動画像中の被写体の構造を単位とする編集・加工等の操作までが考えられる．このような高度に柔軟な情報操作が可能となれば，マルチメディアを統一的に扱う環境が提供された場合に，メディア間で編集・加工処理がフレキシブルに行え，より豊かな情報伝達手段が実現される可能性がある．

こうした利用形態を想定した場合，従来の通信を主目的とした動画像情報の記述・符号化では不十分で，そのデータのモデリングと記述法ならびに符号化について見直しを行う必要がある．

通信と情報処理の両分野の視点から動画像情報の高度利用のために，その時間的・空間的性質，および構造を統一的に扱う，符号化手法と記述法についての課題を整理する．

2 汎用的な動画像符号化への要求条件

動画像信号をディジタルの形で伝送あるいは蓄積するには，もとの信号に含まれる冗長な情報を除き，できるだけ低いビットレートで符号化する必要がある．この符号化方式の確立を目指して，これまでに，次のような国際標準もしくはその最終案が存在する．

- CCITT 勧告 H.261
- MPEG1 標準
- CCIR 勧告 723

これらの発展形として，いずれもこれから必要な標準として CCIR 勧告 601 ディジタル TV 信号の高能率符号化をも指向している．また，上記の各既存標準の次世代の標準案として，より高ビットレートでしかも高品質映像のための符号化方式の確立に関する活動が以下の通り行われている．

- ISO/IEC JTC1/SC29/WG11(MPEG): DSM (digital Storage Media) のための汎用的な映像符号化，MPEG 1 の上位バージョンという意味で MPEG 2 と呼ばれている．
- CCITT SGXV ビデオ符号化専門家グループ: B-ISDN などの ATM 環境に適した映像符号化．H.261 の拡張という意味で H.26X と呼ばれる．
- TG CMTT/2 Special Rapporteur's Group for Secondary Ditribution of Digital TV

and HDTV: ディジタル TV ならびに HDTV を 2 次分配 (すなわち最終ユーザまで分配) するための映像符号化

- CCIR SG11: ディジタル TV 放送のための映像符号化

応用分野により，通信分野は CCITT, 放送分野は CCIR, 蓄積メディアの分野は ISO/IEC と異なる標準化機関が担当している．しかし，これら 3 分野の境界は連続的であり，また融合した使われ方が一般的になるであろうことから，類似の分野で異なる標準ができることがないよう，標準化機関の間の密接な連携が求められている．現在，共同作業や連絡文書の交換による努力が行われている．ISO/IEC と CCITT ではこれらの分野での応用に対する汎用的な符号化方式の標準方式を決定すべく MPEG2/H.26X と名付けて協調して作業が行われている．特に MPEG2/H.26X としているのは，ISO/IEC と CCITT が同一の技術仕様を共通のテキスト (Common Text) で，それぞれ MPEG2 標準，勧告 H.26X として標準化することを意味している．ここでの目標は以下の通り．

1. 現行の NTSC/PAL/SECAM 以上の品質をめざす: ディジタル画像への期待の一つは高画質であり，アナログ・コンポジット TV よりは高画質を実現する．具体的には，CCIR 勧告 601 で定義されている解像度 (720 画素 × 483/576 ライン ×30/25 フレーム / 秒) に近づけたい．

2. ビットレート: NTSC/PAL/SECAM レベルの品質を 3 ～ 5(Mbit/s) で，CCIR 勧告 601 レベルに近い品質を 8 ～ 10(Mbit/s) で実現する．動画像符号化では，ビットレートを変えると得られる画像品質が連続的に変わる．従って，一つの標準で，ある範囲のビットレート，ある範囲の品質を提供することが可能である．

3. 符号化対象 TV 方式: 当面現行 TV に集中するが，幅広画面や EDTV, HDTV への拡張を視野にいれる．また，インタレース信号のみでなく，フィルムソースなどノンインタレース信号 (Progressive format) も扱えるようにする．

4. アプリケーションによらない汎用的なアルゴリズムをめざす : MPEG はディジタル蓄積メディアを基盤にしているものの，蓄積メディアが単独で使われる応用のみでなく，例えば通信端末からネットワークを介して蓄積メディアの情報にアクセスするなど，多用な展開が想定される．動画像符号化の汎用化により，サービス性の向上を図るとともに，機器，デバイスの共通化による利点を期待している．

この汎用的な動画像符号化方式に特に求められるものは，アルゴリズムが汎用的であることである．しかし，汎用的であることと特定のアプリケーションに最適であることは，必ずしも両立しない．汎用性は，ビットストリームが異なるアプリケーションで互換性のあることが狙いであり，特定のアプリケーションに最適なシステムを構築する立場からは余分な要素を包含することになる．これを解決するために，共通で使用できる部分を最大限としそれ以外の多様なアプリケーションが必要とする特定要素をすべて盛り込んだ道具箱 (toolkit) のアプローチを目指して標準化が進んでいる．これまで，考えられるアプリケーションの中から主に以下の要求条件を規定している [3].

- CCIR Rec. 601 を中心とした飛び越し走査画像信号に対する処理
- ランダムアクセス，高速サーチ再生等の蓄積メディアの要求条件を満たす符号化 [4]
- MPEG1[2] および H.261[1] 等の既存の標準符号化方式との互換性
- 符号化データから一部分だけを取り出し意味のある画像を再生する機能 (スケーラビリティ : scalability)
- ATM 網でのセル廃棄に対する対策

3 画像符号化における色空間 (表色系) の扱い

近年のディジタルカラー複写機に代表されるカラー入出力技術の発達，パソコン・ワークステーションの普及 JPEG, JBIG 等の画像符号化技術の進展と標準化，ODA(Open Document Architecture) のカラー化等により，ファクシミリのカラー化を具体的に検討できる要素技術が整いつつある．一方，カラーの利用面から見ると，カラー印刷技術等のカラー技術の発達に

よりカタログ，雑誌等のカラー化，さらには新聞紙面へのカラーページの導入が急激に進んでいる．また，ビジネス分野でも，ディジタルカラー複写機やワークステーション等によるカラー画像処理が普及し，デザイン・広告・建築・設計の分野での利用が進みつつある．今後，専門的な利用を核に，広いビジネス領域で，カラーの利用が進展し，それにともない，カラー画像の高速高品質な伝送に関する要求が高まるものと予想される．これらの状況に対応して，カラーファクシミリを中心としたカラー画像処理のための標準化活動が活発化している．国際的には，1990年3月のCCITT SGVIIIブタペスト会合から検討が開始され，現在はSGVIII/WP1/Q4の中のG4ファクシミリのカラー化という形で検討が開始されている．国内では，標準化への国際的貢献と国内での先行的な検討を狙いとして，1989年頃から画像電子学会の中で検討体制について議論が始められ，1990年9月に画像電子学会テレマティクス研究専門委員会の下にカラーファクシミリ検討会が発足した．1991年からは，より具体的な検討を行うため，5つのワーキンググループを設けて検討を継続している．このワーキンググループのうちカラー画像の符号化・電送を見据えたカラー選択をするためのWGに色空間検討(WG2)がある．

このWG2では，カラーファクシミリの標準としてどのカラースペースを採用すべきかの評価を行っている．色彩工学で行われている色空間に対する検討の成果と実際の画像工学で得られた画像符号化の成果を工学的な見地から融合するという意味では全く新しい非常に画期的な研究機関であり，今後の他の分野例えば，画像符号化，CG，ODA，マルチメディアなど各分野に与える寄与は非常に大きなものがある．

この活動の各色空間に対する評価の目的は，複数の色空間を標準として備えることを許しているISO/IEC JTC1/SC18のODA Color Addendumでの目的とは異なり，複数の色空間を認めずできれば一つに絞りたいという観点から行っている．このため，ODAでの検討とは大きく異なり，色空間の評価をできるだけ定量的に行おうとしており，非常に詳細で客観的な検討を行ってきた．

CCITTでは，当初，ヨーロッパのビデオテックスでのYCbCrの採用を背景として，YCbCrを必須，CIE_L*a*b*をオプションとする提案がなされた．しかし，これに対し，

- デバイスインディペンデント性を重視すべきで，RGBとγが1つに決まっていないデバイス依存性の高いYCbCrを必須とするのは問題がある．
- カラースペースの選択は客観的評価により行うべきである．

とする，日本からの提案，及び米国からも，カラースペース表現可能範囲が狭いなどの理由から，YCbCrを必須とすることへの反対意見が出された．日本からは，各色空間で圧縮・復元を行ったサンプルを含む先述のWG2での検討結果を報告し，さらに，米国からはCIE_L*a*b*を支持する奇書が提出された．これらに基づいた議論により，カラー画像表現にはCIE_L*a*b*がもっとも有望であるとのコンセンサスが得られている[5][6]．

しかし，CIE_L*a*b*においても画像符号化を念頭においた場合，大まかには以下のような問題点が存在する．

- 各色軸間に相関が存在し，軸毎に独立に符号化する一般的な画像符号化方式では，冗長度が残されている．
- gamutの形状から不必要な点を表現できてしまい，冗長である．
- 実際のカラー画像上では，画素値の分布が空間中の一部分に片寄っており，これを考慮した最適な部分色空間を構成する可能性が残されている．

4 動画像の形状に着目した符号化

画像は高度に構造化されたものであり，人間の視覚もこの構造情報に基づいて画像を物体単位で認識を行っていると考えられる．計算機による画像の認識・理解という立場から構造を自動的に抽出するための研究が長年行われてきたが，符号化の分野においても画像の構造を利用する構造情報符号化の研究が盛んに行われるようになってきた．

構造抽出符号化では，画像の生成源である物体の構造に基づく適切な構造記述モデルを作ることが必要となる．このモデルとしては，3次元構造を持つ3次元構造記述モデルや，物体を2次元平面パッチとしてそれらの重なりで画像を表す2次元構造記述モデルなどが考えられる．画像の認識・理解を行うためにはなるべく正確な3次元構造記述モデルを構築することが要求

されるが，画像の符号化・処理・検索などの目的では2次元構造記述モデルで画像を記述するだけでも効率性・柔軟性をかなり向上できるとされている．これらの観点から，比較的符号化への応用が容易であると考えられる2次元構造記述モデルを想定し，動画像から2次元層構造を自動的に抽出する試みもなされている．

画像を構造的に記述すれば，同一物体上で輝度・色度の変動は微小であるという性質を利用してフレーム内の画素値情報の符号量の低減を図ることができる．構造抽出符号化のさらに大きな利点は，動画像においてフレーム間の動き補償を実空間の物体の動きに対応した形で行うことができる点である．

構造モデルの時間軸方向への連続性を活かすためには，多フレームにわたって安定に画像の領域を追跡する必要がある．その基本的な考え方は，動画像の構造抽出，すなわち，画像中に存在する物体 (領域) の切り出しを行い，その構造を解析して，構造記述データの符号化を行うことにある．しかし，自然動画像の構造抽出および解析は，いまだコンピュータビジョンの分野で発展途上の研究であり，動画像中の構造を安定に抽出することは非常に難題な課題として今もなお存在する．

5 動画像の編集・加工を目的とするデータアーキテクチャ

動画像の符号化は必ずしも情報圧縮を主目的とするものである必要はなくなりつつある．特に，これからの動画像符号化は，原信号を利用形態に最も適したフォーマットへ変換する技術に向けて発展していくと思われる [7]．その利用形態として今後重要になるのは，編集・加工であると考えられる．

文書情報については DTP(Desk Top Publishing) の普及など，すでにその必要性が広く認められており，さらに編集・加工の自由度を高めるため，文書情報を構造的に捉えて記述するためのデータ構造に関する国際標準化作業は CCITT / ISO で進められ，ODA (Open / Office Document Architecture[8]) として勧告案がまとまっている．

マルチメディアの中で動画像を考える場合，文書情報の一部として扱われることも考えられ，動画像に高度な編集・加工処理が要求されると思われる．その要求の幅は広く，MPEG で可能となるフレームを単位とした加工から，他メディア（文書・静止画など時間的変化を伴わない静的メディア，あるいは音声など動画像と同じように時間的変化を伴う動的メディア）とリンクさせながらの処理，またその枠を超えて，動画像中の被写体の構造を単位とする編集・加工等の処理までが考えられる．このような高度に柔軟な情報処理が可能となれば，マルチメディアを統一的に扱う環境が提供された場合など，メディア間で編集・加工処理が柔軟に行え，より豊かな情報伝達手段が与えられることになる．例えば，一般ユーザが既存の動画像を簡単に再利用 (編集・加工) できれば，動画像を用いてドキュメントを作成するようなことも可能になる．

このためには，従来の通信を主目的としたフレームの時間順に従った予測を行う動画像情報の記述・符号化では不十分で，そのデータ構造と記述法，ならびに符号化について根本的な見直しを行う必要がある．

一方，最近は，動画像中の被写体の構造を抽出して符号化する研究が盛んに行われている．例えば，顔画像の特徴点の動きだけを抽出・符号化する分析合成符号化 [9] や，画像内のまとまった領域ごとに符号化を行う領域分割符号化 [10] などがある．また，静止画像に対しても領域単位で閉じた符号化 [11] や，自然界の複雑さを表すといわれるフラクタルを用いた符号化 [12] など，構造に着目した符号化が数多く研究されており，これらが動画像に適用されることも考えられる．

6 スケーラビリティの課題

スケーラビリティは，ビットストリームから符号化データを部分的に取り出して再生できる機能のことを示している．つまり，符号化ビットストリームから復号化器側において必要な部分のみを処理し要求する画像フォーマットで再生画像を得るというものである．スケーラビリティ機能には以下に挙げる特徴がある．

まず第一に再生時において解像度選択が行える．例えば，最高解像度層を CCIR Rec.601 フォーマットシーケンス，2層目を SIF フォーマットシーケンス，3層目を QCIF フォーマットシーケンスとして，ユーザが再生時に解像度選択を行える構造になっている．また，各フォーマットで符号化レートを柔軟に変化させることが可能であり，様々な伝送路に適用可能である．

他の特徴としては，MPEG1 同様イントラ画像をエントリポイントとして順方向・逆方向に

ランダムアクセスが可能である．また高速サーチ再生においては原則としてGOP毎にイントラ画像を時間順または逆順にランダムアクセスすることで実現できる．データスキップ時のメディアへのアクセス速度及び復号化時間がデータ転送速度に比べて無視できる程度であると仮定すれば，通常多くのビット量を割り当てるイントラ画像の符号化ビット量が再生速度のボトルネックになる．例えば，階層構造をとらない場合，エントリポイントであるイントラ画像が15フレーム毎に用意されたとしてもその符号量が全体の1/3を占めるのであれば，15倍速度の再生は実現できず，高々3倍の高速再生が実現できるだけである．しかし，低解像度層の画像を再生すれば，イントラ画像の低解像度層の実質符号化ビット量のGOP全体のビット量に占める割合の逆数倍までのスピードで高速再生を実現できる[13].

さらに各フレームを階層構造にすることで解像度毎に優先順位をつけ，ATM 伝送でのセル廃棄への耐性を確保できる．例えば，2段階の優先制御機能を備えたATM網では，高解像度層の画像を再生するためのデータを非優先情報(lower priority)とし，低い解像度の画像を優先情報(high priority)として伝送することで対応する．

以上，スケーラビリティの利点をまとめると，

- 解像度選択を自由に行える
- ビット誤り，ATM網におけるセル廃棄に耐性を有する
- 縮小・拡大，順方向・逆方向再生など柔軟な再生を効率良く実現できる

等が主な特徴として挙げられる．

スケーラビリティの手法は，空間スケーラビリティ，SNRスケーラビリティ，周波数スケーラビリティの大きく3つに分けられる．

6.1 空間スケーラビリティ

空間スケーラビリティは空間解像度方向の階層符号化である．

この方式は解像度が下がる毎に画像の縦横を半分に縮小している．よって，低解像度層においては符号化の前処理としてまず原画像からその階層のフォーマット(2層の場合，SIFフォーマット)の原画像を作成する．そして，その画像を利用して符号化を行う．高解像度層は原画像に対して符号化を行う．低解像度層画像作成のためのダウンサンプル，また低解像度層から高解像度層画像への動き予測のためのアップサンプルの縮小・拡大の操作には，フィルタリングを用いる．その処理ではできるだけ画像の周波数遮断特性が優れていることが望ましい．

各階層の符号化プロセスでは，階層間での予測に用いるためのアップサンプルされた一階層下の局所復号画像(低解像度局所画像)，動き補償予測に用いる同階層の局所復号画像，動きベクトルを入力としている．

各画像はマクロブロックを単位として許される予測方式を適用し，その中から最も誤差信号電力の小さい予測誤差信号を選択してそのデータを符号化する．そのため，マクロブロック毎に選択された予測方向を復号化器へ送るための付加情報が必要である．このマクロブロック毎の予測モードをマクロブロックタイプと呼ぶ．それぞれの方式でのマクロブロックタイプはあらかじめ用意しておくものである．

この方式では，各階層毎，フレーム単位で符号化を行っているので，各々の階層においてマクロブロックタイプを決定し，予測モード，動きベクトルの付加情報を各々の階層で符号化する．よって，復号化器においても階層毎にマクロブロックタイプに従って復号化を行う．

符号化データは各レイヤ構造にしたがってビットストリームに書き込まれる．復号化器ではそのレイヤ構造のシンタックスに従ってビットストリームを解釈する．

この空間的階層符号化方式では低解像階層は空間解像度が劣化しているため，被写体の詳細な動きはとらえないが，大まかな動きを捉えて符号化できる．一方，高解像度階層では被写体の動きを局所的にしか追ず，検出動き量の空間解像度を高くして詳細な領域での符号化を行うことができる．

6.2 SNR スケーラビリティ

SNRスケーラビリティは量子化誤差信号を階層毎に量子化の精度を向上させるためのビットを付加して符号化する方式である．つまり，符号化，復号化の量子化器部分における階層化であり，符号化器，復号化器の両方において階層構造をとる．

この方式は階層が下がっても扱う画像サイズは変わらないという大きな特徴がある．よって，符号化の前処理としての画像フォーマット変換

は必要ない．また，復号化の際にも予測画像を得るための拡大処理が不要となる．

SNRスケーラビリティの動き予測方式は，時間方向における予測モードだけを持ちそれらを切り替えて行う．よって，予測アルゴリズムとしては非階層の場合と同一である．

動きベクトルは時間方向において検出し，すべて高解像度画像において求める．よって，マクロブロックタイプの決定もすべて高解像度層において行う．高解像度層，低解像度層ともに共通の動きベクトルを利用し，マクロブロックタイプも同一のものを利用する．ただし，このマクロブロックタイプやベクトル情報は低解像度画像でも復号化時に必要となるので，付加情報はすべて低解像度層において符号化する．

符号化の原理は低階層で画像データをDCT係数領域で量子化・逆量子化した後の周波数成分を高階層へ渡し，原周波数成分との誤差成分データを符号化する．つまり，解像度が上がる度に予測誤差の周波数成分を符号化して量子化誤差を低減していく．言い替えると，周波数成分での量子化誤差を量子化を細かくして符号化し，量子化誤差を階層毎で修正し量子化歪を低減している．よって，高解像度層になるほど量子化の精度が高まる．符号化時の量子化 VLCテーブルは各階層とも同一のものを用いる．

この方式は符号化プロセスをすべて原画像サイズで行い，符号化器側では，各階層で独立の動き補償を必要としDCT計算量も2階層の場合は非階層時の 2倍となるので，計算量の増加と必要なフレームメモリ量の増加が問題となる．さらに，復号化器側ではどの階層においても原画像サイズで再生されるので，小さいサイズの再生画を得るためには何らかの画像フォーマット変換を行う必要がある．符号化器では階層数分の回数の動き補償が必要なのに対し，復号化器では1回の動き補償で復号化できる (single loop decoder)．また，階層化をしたときのそれぞれの階層に割り当てるビット量は，アルゴリズムは変更せずパラメータの変更のみで行い，各階層のレートを自由に分配することが可能である．

この方式の最大の特長は量子化の部分において階層化を行っている点である．つまり，符号化予測モードは非階層時と同じで符号化予測モデルとは独立に扱うことができる．よって，DCTだけではなく，DPCM, Wavelet, VQ など様々な情報源符号化へも対応する可能性があり，またどのような予測構造にも適用可能な柔軟な符号化方式であるといえる．

6.3 周波数スケーラビリティ

一般的に画像符号化を行う場合直交変換を用いることが多いが，例えばDCTを使った場合，その変換によって空間領域から周波数領域への変換が行われる．その周波数領域において符号化するときに扱う領域を分割し，階層化して符号化する方法が周波数スケーラビリティである．

周波数スケーラビリティには，サブバンド方式とピラミッド方式の 2つの方法がある．サブバンド方式は，のようにDCT係数領域においてその周波数成分毎で分離して符号化し，復号化器でそれらを統合する方式である．また，ピラミッド方式は，周波数領域で分けられた領域のうちの低周波数の領域において低解像度層との差分を取り符号化する方式である．この方式は，前に述べた空間領域におけるスケーラビリティを周波数空間で行う方式である．

8×8 DCTの周波数領域全体を扱うとき高解像度画像が得られ，またそのうちの 4×4 の部分周波数領域で扱えば低解像度画像が得られる．この方式では周波数領域において扱う範囲を変化させ階層化することから，周波数領域から空間領域へ戻したとき，つまり復号化器から出てきたとき低解像度層からは縮小画像が得られる．このとき低解像度階層で得られる画像は周波数領域において扱う範囲が低周波数よりに偏っているため，ローパスフィルタのかかった画像が得られる．

ちなみに3階層時では，低解像度階層では2×2 IDCTを，中間解像度階層では4×4 IDCTを，高解像度階層では 8×8 IDCTを行う．

この方法の利点は，情報源符号化部分の変換・逆変換の計算が符号器・復号器においてそれぞれ1度の計算で行えることである．階層符号化で問題となる計算量を低減できる．再生側では一種類の逆変換復号器を持てばよく，復号器側が容易であるという利点もある．また，IDCTを 2×2, 4×4, 8×8 と変えれば，画像フォーマットも原画像に対して1/16, 1/4, 1倍と変換でき柔軟な再生が可能である利点を持っている．

一方，各階層により復号化器において扱う画像サイズが異なるので，各階層用の動き補償を考える必要がある．例えば，画像サイズが縦横で半分とすれば，1/2画素精度の動き補償を 1/4画素精度の動き補償に変更する必要がある．

周波数スケーラビリティの動き補償予測方式

は，時間方向における予測モードだけをとり，非階層の場合と同一である．動きベクトルは時間方向において検出し，すべて高解像度画像において求める．よって，マクロブロックタイプの決定も高解像度画像を用いて行う．高解像度層，低解像度層ともに共通の動きベクトルを利用し，マクロブロックタイプも同一のものを利用する．ただし， SNR スケーラビリティと同様に付加情報はすべて低解像度層において符号化する．動き補償予測は高解像度層で予測ループを1回としている．その他の予測方式として動き補償予測を複数にする方法 (multiple encoder loop) もある．しかし，計算量の面から考えると，1回の予測ループで実現する方が処理量が少なく効率が良い．また，符号化器側では階層符号化を行わず符号化ビットストリームから復号化器側だけでの処理で周波数スケーラビリティを行う方法もある．

参考文献

[1] CCITT STUDY GROUP XV, Recommendation H.261: "Video Codec for Audiovisual services at p x 64 kbit/s".

[2] ISO-IEC JTC1, CD 11172: "Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s – Part2:Coding of moving pictures information" (1991).

[3] S. Okubo: "Requirements for high quality video coding standards", Image Communication, **4**, pp. 141–151 (1992).

[4] 米満 潤: "MPEG 標準案ビデオパート (ISO 11172 VIDEO)", 画電学誌, **20**, 4, pp. 306–316 (1991).

[5] 池上 博章: "カラーファクシミリの標準化動向 — 色空間の選択 –", 色彩工学コンファレンス 講演予稿集 (1992).

[6] 酒井 善則, 岸本 重治, 河原 厚, 池上 博章, 小野 文孝, 大町 隆夫, 松木 真: "カラーファクシミリの動向", 画像電子学会テレマティクスシンポジウム講演予稿集 (1992).

[7] 原島 博: "画像情報圧縮のこれから", TV 学会誌, **Vol.44**, No.6, pp. 693–700 (1990).

[8] CCITT, Recommendation T.400 series: "Terminal Equipment and Telematic Services".

[9] 原島 博, 森川博之, 青木幸代: "3次元構造モデルを用いる画像の符号化・処理・表示", 信学春季全大, SD–3–12 (1989).

[10] M.Hotter and J.Ostermann: "Analysis–Synthesis Coding Based on Planar Rigid Moving Objects", International Workshop on 64Kbps Coding of Moving Video, 8–1 (1988).

[11] U.Franke and R.Mester: "Region Based Image Representation with Variable Reconstruction Quality", SPIE Visual Communication and Image Processing, pp. 178–186 (1988).

[12] M.F.Branslay and A.D.Sloan: "A Better Way to Compress Images", BYTE, pp. 215–223 (1988).

[13] 花村 剛, 関口 俊一, 亀山 渉, 富永 英義: "MPEGII に対する階層的動画像符号化方式の提案", 信学技報, IE91–113 (1992 年).

[14] 花村 剛: "動画像の高機能処理のためのアーキテクチャ構築と符号化方式の研究", PhD thesis, 早稲田大学大学院 理工学研究科 (1993 年 3 月 受理).

[15] T. HANAMURA, W. KAMEYAMA and H. TOMINAGA: "Hierarchical Coding Scheme of Video Signal with Scalability and Compatibility", EURASIP: Signal Processing, IMAGE Communication, **Vol.5**, nos.1–2, pp. 159–184 (1993).

[16] 花村 剛, 関口 俊一, 亀山 渉, 富永 英義: "解像度選択および互換性を有する階層的動画像符号化", 信学論 B–I, **Vol.J71–B–I**, 3, pp. 299–310 (1993 年).

[17] 関口 俊一, 花村 剛, 杉原 明, 富永 英義: "ベクトル量子化を用いた動画像の階層符号化方式", 信学技報, IE91–123 / PRU91–143 (1992 年).

[18] 花村 剛, 原 真由美, 富永 英義: "カラー画像の符号化・特徴解析を目的とした色空間変換", TV 学技報, IPU93–1 (1993 年).