明細書

画像符号化装置および画像符号化方法、画像復号装置および画像復号方法、並びに提供媒体

技術分野

本発明は、画像符号化装置および画像符号化方法、画像復号装置および画像復号方法、並びに提供媒体に関する。特に、例えば、動画像データを、光磁気ディスクや磁気テープなどの記録媒体に記録し、これを再生してディスプレイなどに表示したり、テレビ会議システム、テレビ電話システム、放送用機器、マルチメディアデータベース検索システムなどのように、動画像データを伝送路を介して送信側から受信側に伝送し、受信側において、受信された動画像データを表示する場合や、編集して記録する場合などに用いて好適な画像符号化装置および画像符号化方法、画像復号装置および画像復号方法、並びに提供媒体に関する。

背景技術

例えば、テレビ会議システム、テレビ電話システムなどのように、動画像データを遠隔地に伝送するシステムにおいては、伝送路を効率良く利用するため、画像データを、そのライン相関やフレーム間相関を利用して圧縮符号化するようになされている。

動画像の高能率符号化方式として代表的なものとしてMPEG（Moving Picture Experts Group）（蓄積用動画像符号化）方式がある。これはＩＳＯ－ＩＥＣ／ＪＴＣ１／ＳＣ２／ＷＧ１１において議論され、標準案として提案されたものであり、動き補償予測符号化とＤＣＴ（Discrete Cosine Transform）符号化を組み合わせたハイブリッド方式が採用されている。

ＭＰＥＧでは、様々なアプリケーションや機能に対応するために、いくつかのプロファイルおよびレベルが定義されている。最も基本となるのが、メインプロファ

イルメインレベル（ＭＰ＠ＭＬ（Main Profile at Main Level））である。

図１は、ＭＰＥＧ方式におけるＭＰ＠ＭＬのエンコーダの一例の構成を示している。

符号化すべき画像データは、フレームメモリ３１に入力され、一時記憶される。そして、動きベクトル検出器３２は、フレームメモリ３１に記憶された画像データを、例えば、１６画素×１６画素などで構成されるマクロブロック単位で読み出し、その動きベクトルを検出する。

ここで、動きベクトル検出器３２においては、各フレームの画像データを、Ｉピクチャ(フレーム内符号化）、Ｐピクチャ（前方予測符号化）、またはＢピクチャ（両方向予測符号化）のうちのいずれかとして処理する。なお、シーケンシャルに入力される各フレームの画像を、Ｉ，Ｐ，Ｂピクチャのいずれのピクチャとして処理するかは、例えば、予め定められている（例えば、Ｉ，Ｂ，Ｐ，Ｂ，Ｐ，・・・Ｂ，Ｐとして処理される）。

即ち、動きベクトル検出器３２は、フレームメモリ３１に記憶された画像データの中の、予め定められた所定の参照フレームを参照し、その参照フレームと、現在符号化の対象となっているフレームの１６画素×１６ラインの小ブロック（マクロブロック）とをパターンマッチング（ブロックマッチング）することにより、そのマクロブロックの動きベクトルを検出する。

ここで、ＭＰＥＧにおいては、画像の予測モードには、イントラ符号化（フレーム内符号化）、前方予測符号化、後方予測符号化、両方向予測符号化の４種類があり、Ｉピクチャはイントラ符号化され、Ｐピクチャはイントラ符号化または前方予測符号化のいずれかで符号化され、Ｂピクチャはイントラ符号化、前方予測符号化、後方予測符号化、または両方法予測符号化のいずれかで符号化される。

即ち、動きベクトル検出器３２は、Ｉピクチャについては、予測モードとしてイントラ符号化モードを設定する。この場合、動きベクトル検出器３２は、動きベクトルの検出は行わず、予測モード（イントラ予測モード）を、ＶＬＣ（可変長符号化）器３６および動き補償器４２に出力する。

また、動きベクトル検出器３２は、Ｐピクチャについては、前方予測を行い、その動きベクトルを検出する。さらに、動きベクトル検出器３２は、前方予測を行うことにより生じる予測誤差と、符号化対象のマクロブロック（Ｐピクチャのマクロブロック）の、例えば分散とを比較する。その比較の結果、マクロブロックの分散の方が予測誤差より小さい場合、動きベクトル検出器３２は、予測モードとしてイントラ符号化モードを設定し、ＶＬＣ器３６および動き補償器４２に出力する。また、動きベクトル検出器３２は、前方予測を行うことにより生じる予測誤差の方が小さければ、予測モードとして前方予測符号化モードを設定し、検出した動きベクトルとともに、ＶＬＣ器３６および動き補償器４２に出力する。

さらに、動きベクトル検出器３２は、Ｂピクチャについては、前方予測、後方予測、および両方向予測を行い、それぞれの動きベクトルを検出する。そして、動きベクトル検出器３２は、前方予測、後方予測、および両方向予測についての予測誤差の中の最小のもの（以下、適宜、最小予測誤差という）を検出し、その最小予測誤差と、符号化対象のマクロブロック（Ｂピクチャのマクロブロック）の、例えば分散とを比較する。その比較の結果、マクロブロックの分散の方が最小予測誤差より小さい場合、動きベクトル検出器３２は、予測モードとしてイントラ符号化モードを設定し、ＶＬＣ器３６および動き補償器４２に出力する。また、動きベクトル検出器３２は、最小予測誤差の方が小さければ、予測モードとして、その最小予測誤差が得られた予測モードを設定し、対応する動きベクトルとともに、ＶＬＣ器３６および動き補償器４２に出力する。

動き補償器４２は、動きベクトル検出器３２から予測モードと動きベクトルの両方を受信すると、その予測モードおよび動きベクトルにしたがって、フレームメモリ４１に記憶されている、符号化され、かつ既に局所復号された画像データを読み出し、この読み出された画像データを、予測画像データとして、演算器３３および４０に供給する。

演算器３３は、動きベクトル検出器３２がフレームメモリ３１から読み出した画像データと同一のマクロブロックをフレームメモリ３１から読み出し、そのマクロ

ブロックと、動き補償器４２からの予測画像との差分を演算する。この差分値は、ＤＣＴ器３４に供給される。

一方、動き補償器４２は、動きベクトル検出器３２から予測モードのみを受信した場合、即ち、予測モードがイントラ符号化モードである場合には、予測画像を出力しない。この場合、演算器３３（演算器４０も同様）は、特に処理を行わず、フレームメモリ３１から読み出したマクロブロックを、そのままＤＣＴ器３４に出力する。

ＤＣＴ器３４では、演算器３３の出力データに対して、ＤＣＴ処理が施され、その結果得られるＤＣＴ係数が、量子化器３５に供給される。量子化器３５では、バッファ３７のデータ蓄積量（バッファ３７に記憶されているデータの量）（バッファフィードバック）に対応して量子化ステップ（量子化スケール）が設定され、その量子化ステップで、ＤＣＴ器３４からのＤＣＴ係数が量子化される。この量子化されたＤＣＴ係数（以下、適宜、量子化係数という）は、設定された量子化ステップとともに、ＶＬＣ器３６に供給される。

ＶＬＣ器３６では、量子化器３５より供給される量子化係数が、例えばハフマン符号などの可変長符号に変換され、バッファ３７に出力される。さらに、ＶＬＣ器３６は、量子化器３５からの量子化ステップ、動きベクトル検出器３２からの予測モード（イントラ符号化（画像内予測符号化）、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれが設定されたかを示すモード）および動きベクトルも可変長符号化し、その結果得られる符号化データを、バッフ３７に出力する。

バッファ３７は、ＶＬＣ器３６からの符号化データを一時蓄積することにより、そのデータ量を平滑化し、符号化ビットストリームとして、例えば、伝送路に出力し、または記録媒体に記録する。

また、バッファ３７は、そのデータ蓄積量を量子化器３５に出力しており、量子化器３５は、このバッファ３７からのデータ蓄積量にしたがって量子化ステップを設定する。即ち、量子化器３５は、バッファ３７がオーバーフローしそうなとき、

量子化ステップを大きくし、これにより、量子化係数のデータ量を低下させる。また、量子化器３５は、バッファ３７がアンダーフローしそうなとき、量子化ステップを小さくし、これにより、量子化係数のデータ量を増大させる。このようにして、バッファ３７のオーバフローとアンダフローを防止するようになっている。

量子化器３５が出力する量子化係数と量子化ステップは、ＶＬＣ器３６だけでなく、逆量子化器３８にも供給されるようになされている。逆量子化器３８では、量子化器３５からの量子化係数が、同じく量子化器３５からの量子化ステップにしたがって逆量子化され、これによりＤＣＴ係数に変換される。このＤＣＴ係数は、ＩＤＣＴ器（逆ＤＣＴ器）３９に供給される。ＩＤＣＴ器３９では、ＤＣＴ係数が逆ＤＣＴ処理され、その処理の結果得られるデータが、演算器４０に供給される。

演算器４０には、ＩＤＣＴ器３９の出力データの他、上述したように、動き補償器４２から、演算器３３に供給されている予測画像と同一のデータが供給されている。演算器４０は、ＩＤＣＴ器３９の出力データ（予測残差（差分データ））と、動き補償器４２からの予測画像データとを加算することで、元の画像データを局所復号し、この局所復号された画像データ（局所復号画像データ）が出力される（但し、予測モードがイントラ符号化である場合には、ＩＤＣＴ器３９の出力データは、演算器４０をスルーして、そのまま、局所復号画像データとして、フレームメモリ４１に供給される）。なお、この復号画像データは、受信側において得られる復号画像データと同一のものである。

演算器４０において得られた復号画像データ（局所復号画像データ）は、フレームメモリ４１に供給されて記憶され、その後、インター符号化（前方予測符号化、後方予測符号化、量方向予測符号化）される画像に対する参照画像データ（参照フレーム）として用いられる。

次に、図２は、図１のエンコーダから出力される符号化データを復号する、ＭＰＥＧにおけるＭＰ@ＭＬのデコーダの一例の構成を示している。

伝送路を介して伝送されてきた符号化ビットストリーム（符号化データ）が図示せぬ受信装置で受信され、または記録媒体に記録された符号化ビットストリーム

（符号化データ）が図示せぬ再生装置で再生され、バッファ１０１に供給されて記憶される。

ＩＶＬＣ器（逆ＶＬＣ器（可変長復号器））１０２は、バッファ１０１に記憶された符号化データを読み出し、可変長復号することにより、その符号化データを、マクロブロック単位で、動きベクトル、予測モード、量子化ステップ、および量子化係数に分離する。これらのデータのうち、動きベクトルおよび予測モードは動き補償器１０７に供給され、量子化ステップおよびマクロブロックの量子化係数は逆量子化器１０３に供給される。

逆量子化器１０３は、ＩＶＬＣ器１０２より供給されたマクロブロックの量子化係数を、同じくＩＶＬＣ器１０２より供給された量子化ステップにしたがって逆量子化し、その結果得られるＤＣＴ係数を、ＩＤＣＴ器１０４に出力する。ＩＤＣＴ器１０４は、逆量子化器１０３からのマクロブロックのＤＣＴ係数を逆ＤＣＴし、演算器１０５に供給する。

演算器１０５には、ＩＤＣＴ器１０４の出力データの他、動き補償器１０７の出力データも供給されている。即ち、動き補償器１０７は、フレームメモリ１０６に記憶されている、既に復号された画像データを、図１の動き補償器４２における場合と同様に、ＩＶＬＣ器１０２からの動きベクトルおよび予測モードにしたがって読み出し、予測画像データとして、演算器１０５に供給する。演算器１０５は、ＩＤＣＴ器１０４の出力データ（予測残差（差分値））と、動き補償器１０７からの予測画像データとを加算することで、元の画像データを復号する。この復号画像データは、フレームメモリ１０６に供給されて記憶される。なお、ＩＤＣＴ器１０４の出力データが、イントラ符号化されたものである場合には、その出力データは、演算器１０５をスルーして、復号画像データとして、そのままフレームメモリ１０６に供給されて記憶される。

フレームメモリ１０６に記憶された復号画像データは、その後に復号される画像データの参照画像データとして用いられる。さらに、復号画像データは、出力再生画像として、例えば、図示せぬディスプレイなどに供給されて表示される。

なお、ＭＰＥＧ１および２では、Ｂピクチャは、参照画像データとして用いられないため、エンコーダまたはデコーダのそれぞれにおいて、フレームメモリ４１（図１）または１０６（図２）には記憶されない。

以上の図１、図２に示したエンコーダ、デコーダは、ＭＰＥＧ１／２の規格に準拠したものであるが、現在、画像を構成する物体などのオブジェクトのシーケンスであるＶＯ（Video Object）単位で符号化を行う方式につき、ＩＳＯ－ＩＥＣ／ＪＴＣ１／ＳＣ２９／ＷＧ１１において、ＭＰＥＧ（Moving Picture Experts Group）４として標準化作業が進められている。

ところで、ＭＰＥＧ４については、主として、通信の分野で利用されるものとして、標準化作業が進められていたため、ＭＰＥＧ１／２において規定されているＧＯＰ（Group Of Picture）は、ＭＰＥＧ４では規定されておらず、従って、ＭＰＥＧ４が蓄積メディアに利用された場合には、効率的なランダムアクセスが困難になることが予想される。

発明の開示

本発明は、このような状況に鑑みてなされたものであり、効率的なランダムアクセスをすることができるようにするものである。

請求の範囲第１項に記載の画像符号化装置は、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化する符号化手段を備えることを特徴とする。

請求の範囲第６項に記載の画像符号化方法は、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化する符号化ステップを備えることを特徴とする。

請求の範囲第１１項に記載の画像復号装置は、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを復号する復号手段を備えることを特徴とする。

請求の範囲第１６項に記載の画像復号方法は、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを復号する復号ステップを備えることを特徴とする。

請求の範囲第２１項に記載の提供媒体は、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られる符号化ビットストリームを提供することを特徴とする。

請求の範囲第２６項に記載の画像符号化装置は、秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成手段と、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成手段とを備えることを特徴とする。

請求の範囲第２９項に記載の画像符号化方法は、秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成ステップと、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成ステップとを備えることを特徴とする。

請求の範囲第３２項に記載の画像復号装置は、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、秒精度時刻情報および詳細時間情報に基づいて求める表示時刻算出手段を備えることを特徴とする。

請求の範囲第３５項に記載の画像復号方法は、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、秒精度時刻情報および詳細時間情報に基づいて求める表示時刻算出ステップを備えることを特徴とする。

請求の範囲第３８項に記載の提供媒体は、秒精度の時刻を表す秒精度時刻情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時

刻を表す情報として、秒精度時刻情報および詳細時間情報を、対応するＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加することにより得られる符号化ストリームを提供することを特徴とする。

請求の範囲第１項に記載の画像符号化装置においては、符号化手段が、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化するようになされている。

請求の範囲第６項に記載の画像符号化方法においては、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化するようになされている。

請求の範囲第１１項に記載の画像復号装置においては、復号手段が、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを復号するようになされている。

請求の範囲第１６項に記載の画像復号方法においては、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを復号するようになされている。

請求の範囲第２１項に記載の提供媒体においては、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られる符号化ビットストリームを提供するようになされている。

請求の範囲第２６項に記載の画像符号化装置においては、秒精度時刻情報生成手段は、秒精度の時刻を表す秒精度時刻情報を生成し、詳細時間情報生成手段は、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成するようになされている。

請求の範囲第２９項に記載の画像符号化方法においては、秒精度の時刻を表す秒精度時刻情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度よ

り細かい精度で表す詳細時間情報を生成するようになされている。

請求の範囲第３２項に記載の画像復号装置においては、表示時刻算出手段が、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、秒精度時刻情報および詳細時間情報に基づいて求めるようになされている。

請求の範囲第３５項に記載の画像復号方法においては、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、秒精度時刻情報および詳細時間情報に基づいて求めるようになされている。

請求の範囲第３８項に記載の提供媒体においては、秒精度の時刻を表す秒精度時刻情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、秒精度時刻情報および詳細時間情報を、対応するＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加することにより得られる符号化ストリームを提供するようになされている。

図面の簡単な説明

図１は、従来のエンコーダの一例の構成を示すブロック図である。

図２は、従来のデコーダの一例の構成を示すブロック図である。

図３は、本発明を適用したエンコーダの一実施の形態の構成例を示すブロック図である。

図４は、時刻によって、ＶＯの位置、大きさが変化することを説明するための図である。

図５は、図３のＶＯＰ符号化部3_1～3_Nの構成例を示すブロック図である。

図６は、空間スケーラビリティを説明するための図である。

図７は、空間スケーラビリティを説明するための図である。

図８は、空間スケーラビリティを説明するための図である。

図９は、空間スケーラビリティを説明するための図である。

図１０は、ＶＯＰのサイズデータおよびオフセットデータの決定方法を説明するための図である。

図１１は、図５の下位レイヤ符号化部２５の構成例を示すブロック図である。

図１２は、図５の上位レイヤ符号化部２３の構成例を示すブロック図である。

図１３は、空間スケーラビリティを説明するための図である。

図１４は、時間スケーラビリティを説明するための図である。

図１５は、本発明を適用したデコーダの一実施の形態の構成例を示すブロック図である。

図１６は、図１５のＶＯＰ復号部72_1～72_Nの他の構成例を示すブロック図である。

図１７は、図１６の下位レイヤ復号部９５の構成例を示すブロック図である。

図１８は、図１６の上位レイヤ復号部９３の構成例を示すブロック図である。

図１９は、スケーラブル符号化によって得られるビットストリームのシンタクスを示す図である。

図２０は、ＶＳのシンタクスを示す図である。

図２１は、ＶＯのシンタクスを示す図である。

図２２は、ＶＯＬのシンタクスを示す図である。

図２３は、ＶＯＰのシンタクスを示す図である。

図２４は、modulo_time_base と VOP_time_increment との関係を示す図である。

図２５は、本発明によるビットストリームのシンタクスを示す図である。

図２６は、ＧＯＶのシンタクスを示す図である。

図２７は、 time_code の構成を示す図表である。

図２８は、ＧＯＶ層の time_code、並びにＧＯＶの先頭のＩ－ＶＯＰの modulo_time_base と VOP_time_increment の符号化方法を示す図である。

図２９は、ＧＯＶ層の time_code、並びにＧＯＶの先頭のＩ－ＶＯＰよりも前に位置するＢ－ＶＯＰの modulo_time_base と VOP_time_increment の符号化方法を示す図である。

図３０は、modulo_time_base と VOP_time_increment の定義を変更しない場合のそれらの関係を示す図である。

図３１は、Ｂ－ＶＯＰの modulo_time_base と VOP_time_increment の第１の方法による符号化処理を示す図である。

図３２は、Ｉ／Ｐ－ＶＯＰの modulo_time_base と VOP_time_increment の第１および第２の方法による符号化処理を示すフローチャートである。

図３３は、Ｂ－ＶＯＰの modulo_time_base と VOP_time_increment の第１の方法による符号化処理を示すフローチャートである。

図３４は、第１および第２の方法により符号化したＩ／Ｐ－ＶＯＰの modulo_time_base と VOP_time_increment の復号処理を示すフローチャートである。

図３５は、第１の方法により符号化したＢ－ＶＯＰの modulo_time_base と VOP_time_increment の復号処理を示すフローチャートである。

図３６は、Ｂ－ＶＯＰの modulo_time_base と VOP_time_increment の第２の方法による符号化処理を示す図である。

図３７は、Ｂ－ＶＯＰの modulo_time_base と VOP_time_increment の第２の方法による符号化処理を示すフローチャートである。

図３８は、第２の方法により符号化したＢ－ＶＯＰの modulo_time_base と VOP_time_increment の復号処理を示すフローチャートである。

図３９は、modulo_time_base について説明するための図である。

図４０は、本発明を適用したエンコーダおよびデコーダの他の実施の形態の構成例を示すブロック図である。

発明を実施するための最良の形態

以下図面について本発明の実施例を詳述する。その前に、特許請求の範囲に記載の発明の各手段と以下の実施の形態との対応関係を明らかにするために、各手段の後の括弧内に、対応する実施の形態（但し、一例）を付加して、本発明の特徴を記述すると、次のようになる。

即ち、請求の範囲第１項に記載の画像符号化装置は、画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、画像を受信する受信手段（例えば、図１１や図１２に示すフレームメモリ３１など）と、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化する符号化手段（例えば、図１１や図１２に示すＶＬＣ器３６など）とを備えることを特徴とする。

請求の範囲第４項に記載の画像符号化装置は、イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、符号化開始秒精度絶対時刻を基準とする秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成手段（例えば、図３２に示すプログラムの処理ステップＳ３～Ｓ７、および図３７に示すプログラムの処理ステップＳ４３～Ｓ４７など）と、オブジェクトのグループに含まれるＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成手段（例えば、図３２に示すプログラムの処理ステップＳ８、および図３７に示すプログラムの処理ステップＳ４８など）と、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、秒精度時刻情報および詳細時間情報を、対応するＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する付加手段（例えば、図１１や図１２に示すＶＬＣ器３６など）とをさらに備えることを特徴とする。

請求の範囲第１１項に記載の画像復号装置は、画像を復号する画像復号装置であって、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを受信する受信手段（例えば、図１７や図１８に示すバッファ１０１など）と、符号

化ビットストリームを復号する復号手段（例えば、図１７や図１８に示すＩＶＬＣ器１０２など）とを備えることを特徴とする。

請求の範囲第１３項に記載の画像復号装置は、オブジェクトのグループには、その符号化を開始した、画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻が含められており、オブジェクトのグループに含まれるオブジェクトの表示時刻を、符号化開始秒精度絶対時刻に基づいて求める表示時刻算出手段（例えば、図３４に示すプログラムの処理ステップＳ２２～Ｓ２７および図３８に示すプログラムの処理ステップＳ５２～Ｓ５７など）をさらに備えることを特徴とする。

請求の範囲第２６項に記載の画像符号化装置は、画像を、その画像を構成するオブジェクトごとに符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成手段（例えば、図３２に示すプログラムの処理ステップＳ３～Ｓ７、および図３７に示すプログラムの処理ステップＳ４３～Ｓ４７など）と、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成手段（例えば、図３２に示すプログラムの処理ステップＳ８、および図３７に示すプログラムの処理ステップＳ４８など）と、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、秒精度時刻情報および詳細時間情報を、対応するＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する付加手段（例えば、図１１や図１２に示すＶＬＣ器３６など）とを備えることを特徴とする。

請求の範囲第３２項に記載の画像復号装置は、イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとともに、秒精度の時刻を表す秒精度時刻情報と、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報とが、表示時刻を表す情報として、対応するＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加されているとき、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、秒精度時刻情報および詳細時間情報に基づいて求める表示時刻算出手段（例えば、図３４に示すプログラムの処理ステップＳ２２～Ｓ２７、および図３８に示すプログラムの処理ステップＳ５２～Ｓ５７など）と、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰを、対応する表示時刻にしたがって復号する復号手段（例えば、図１７や図１８に示すＩＶＬＣ器１０２など）とを備えることを特徴とする。

なお、勿論この記載は、各手段を上記したものに限定することを意味するものではない。

図３は、本発明を適用したエンコーダの一実施の形態の構成例を示している。

符号化すべき画像（動画像）データは、ＶＯ（Video Object）構成部１に入力され、ＶＯ構成部１では、そこに入力される画像を構成するオブジェクトごとに、そのシーケンスであるＶＯが構成され、ＶＯＰ構成部2_1～2_Nに出力される。即ち、ＶＯ構成部１においてＮ個のＶＯ＃１～ＶＯ＃Ｎが構成された場合、そのＮ個のＶＯ＃１～ＶＯ＃Ｎは、ＶＯＰ構成部2_1～2_Nにそれぞれ出力される。

具体的には、例えば、符号化すべき画像データが、独立した背景Ｆ１のシーケンスと前景Ｆ２のシーケンスとから構成される場合、ＶＯ構成部１は、例えば、前景Ｆ２のシーケンスを、ＶＯ＃１として、ＶＯＰ構成部2_1に出力するとともに、背

景Ｆ１のシーケンスを、ＶＯ＃２として、ＶＯＰ構成部２$_2$に出力する。

なお、ＶＯ構成部１は、符号化すべき画像データが、例えば、背景Ｆ１と前景Ｆ２とを、既に合成したものである場合、所定のアルゴリズムにしたがって、画像を領域分割することにより、背景Ｆ１と前景Ｆ２とを取り出し、それぞれのシーケンスとしてのＶＯを、対応するＶＯＰ構成部２ｎ（但し、ｎ＝１，２，・・・，Ｎ）に出力する。

ＶＯＰ構成部２ｎは、ＶＯ構成部１の出力から、ＶＯＰ（VO Plane）を構成する。即ち、例えば、各フレームから物体を抽出し、その物体を囲む、例えば、最小の長方形（以下、適宜、最小長方形という）をＶＯＰとする。なお、このとき、ＶＯＰ構成部２ｎは、その横および縦の画素数が、例えば、１６の倍数となるようにＶＯＰを構成する。ＶＯ構成部２ｎは、ＶＯＰを構成すると、そのＶＯＰを、ＶＯＰ符号化部３ｎに出力する。

さらに、ＶＯＰ構成部２ｎは、ＶＯＰの大きさ（例えば、横および縦の長さ）を表すサイズデータ（VOP size）と、フレームにおける、そのＶＯＰの位置（例えば、フレームの最も左上を原点とするときの座標）を表すオフセットデータ（VOP offset）とを検出し、これらのデータも、ＶＯＰ符号化部３ｎに供給する。

ＶＯＰ符号化部３ｎは、ＶＯＰ構成部２ｎの出力を、例えば、ＭＰＥＧや、Ｈ.２６３などの規格に準拠した方式で符号化し、その結果得られるビットストリームを、多重化部４に出力する。多重化部４は、ＶＯＰ符号化部３$_1$～３$_N$からのビットストリームを多重化し、その結果得られる多重化データを、例えば、地上波や、衛星回線、ＣＡＴＶ網その他の伝送路５を介して伝送し、または、例えば、磁気ディスク、光磁気ディスク、光ディスク、磁気テープその他の記録媒体６に記録する。

ここで、ＶＯおよびＶＯＰについて説明する。

ＶＯは、ある合成画像のシーケンスが存在する場合の、その合成画像を構成する各オブジェクト（物体）のシーケンスであり、ＶＯＰは、ある時刻におけるＶＯを意味する。即ち、例えば、いま、画像Ｆ１およびＦ２を合成して構成される合成画像Ｆ３がある場合、画像Ｆ１またはＦ２が時系列に並んだものが、それぞれＶＯで

あり、ある時刻における画像Ｆ１またはＦ２が、それぞれＶＯＰである。従って、ＶＯは、異なる時刻の、同一オブジェクトのＶＯＰの集合ということができる。

なお、例えば、画像Ｆ１を背景とするとともに、画像Ｆ２を前景とすると、合成画像Ｆ３は、画像Ｆ２を抜くためのキー信号を用いて、画像Ｆ１およびＦ２を合成することによって得られるが、この場合における画像Ｆ２のＶＯＰには、その画像Ｆ２を構成する画像データ（輝度信号および色差信号）の他、適宜、そのキー信号も含まれるものとする。

画像フレーム（画枠）のシーケンスは、その大きさおよび位置のいずれも変化しないが、ＶＯは、大きさや位置が変化する場合がある。即ち、同一のＶＯを構成するＶＯＰであっても、時刻によって、その大きさや位置が異なる場合がある。

具体的には、図４は、背景である画像Ｆ１と、前景である画像Ｆ２とからなる合成画像を示している。

画像Ｆ１は、例えば、ある自然の風景を撮影したものであり、その画像全体のシーケンスが１つのＶＯ（ＶＯ＃０とする）とされている。また、画像Ｆ２は、例えば、人が歩いている様子を撮影したものであり、その人を囲む最小の長方形のシーケンスが１つのＶＯ（ＶＯ＃１とする）とされている。

この場合、ＶＯ＃０は風景の画像であるから、基本的に、通常の画像のフレームと同様に、その位置および大きさの両方とも変化しない。これに対して、ＶＯ＃１は人の画像であるから、人物が左右に移動したり、また、図面において手前側または奥側に移動することにより、その大きさや位置が変化する。従って、図４は、同一時刻におけるＶＯ＃０およびＶＯ＃１を表しているが、ＶＯの位置や大きさは、時間の経過にともなって変化することがある。

そこで、図３のＶＯＰ符号化部３ｎは、その出力するビットストリームに、ＶＯＰを符号化したデータの他、所定の絶対座標系におけるＶＯＰの位置（座標）および大きさに関する情報も含めるようになされている。なお、図４においては、ＶＯ＃０を構成する、ある時刻のＶＯＰ（画像Ｆ１）の位置を示すベクトルをＯＳＴ０と、その時刻と同一時刻における、ＶＯ＃１のＶＯＰ（画像Ｆ２）の位置を表すベ

クトルをＯＳＴ１と、それぞれ表してある。

次に、図５は、スケーラビリティを実現する、図３のＶＯＰ符号化部３ｎの構成例を示している。即ち、ＭＰＥＧでは、異なる画像サイズやフレームレートに対応するスケーラビリティを実現するスケーラブル符号化方式が導入されており、図５に示したＶＯＰ符号化部３ｎでは、そのようなスケーラビリティを実現することができるようになされている。

ＶＯＰ構成部２ｎからのＶＯＰ（画像データ）、並びにそのサイズデータ（VOP size）、およびオフセットデータ（VOP offset）は、いずれも画像階層化部２１に供給される。

画像階層化部２１は、ＶＯＰから、１以上の階層の画像データを生成する（ＶＯＰの１以上の階層化を行う）。即ち、例えば、空間スケーラビリティの符号化を行う場合においては、画像階層化部２１は、そこに入力される画像データを、そのまま上位レイヤ（上位階層）の画像データとして出力するとともに、それらの画像データを構成する画素数を間引くことなどにより縮小し（解像度を低下させ）、これを下位レイヤ（下位階層）の画像データとして出力する。

なお、入力されたＶＯＰを下位レイヤのデータとするとともに、そのＶＯＰの解像度を、何らかの手法で高くし（画素数を多くし）、これを、上位レイヤのデータとすることなども可能である。

また、階層数は、１とすることが可能であるが、この場合、スケーラビリティは実現されない。なお、この場合、ＶＯＰ符号化部３ｎは、例えば、下位レイヤ符号化部２５だけで構成されることになる。

さらに、階層数は、３以上とすることも可能であるが、ここでは、簡単のために、２階層の場合について説明を行う。

画像階層化部２１は、例えば、時間スケーラビリティ（テンポラルスケーラビリティ）の符号化を行う場合、時刻に応じて、画像データを、下位レイヤまたは上位レイヤのデータとして、例えば、交互に出力する。即ち、例えば、画像階層化部２１は、そこに、あるＶＯを構成するＶＯＰが、ＶＯＰ０，ＶＯＰ１，ＶＯＰ２，Ｖ

ＯＰ３，・・・の順で入力されたとした場合、ＶＯＰ０，ＶＯＰ２，ＶＯＰ４，ＶＯＰ６，・・・を、下位レイヤのデータとして、また、ＶＯＰ１，ＶＯＰ３，ＶＯＰ５，ＶＯＰ７，・・・を、上位レイヤデータとして出力する。なお、時間スケーラビリティの場合は、このようにＶＯＰが間引かれたものが、下位レイヤおよび上位レイヤのデータとされるだけで、画像データの拡大または縮小（解像度の変換）は行われない（但し、行うようにすることも可能である）。

また、画像階層化部２１は、例えば、ＳＮＲ（Signal to Noise Ratio）スケーラビリティの符号化を行う場合、入力された画像データを、そのまま上位レイヤまたは下位レイヤのデータそれぞれとして出力する。即ち、この場合、下位レイヤ並びに上位レイヤの画像データは、同一のデータとなる。

ここで、ＶＯＰごとに符号化を行う場合の空間スケーラビリティについては、例えば、次のような３種類が考えられる。

即ち、例えば、いま、ＶＯＰとして、図４に示したような画像Ｆ１およびＦ２でなる合成画像が入力されたとすると、第１の空間スケーラビリティは、図６に示すように、入力されたＶＯＰ全体（図６（Ａ））を上位レイヤ（Enhancement Layer）とするとともに、そのＶＯＰ全体を縮小したもの（図６（Ｂ））を下位レイヤ（Base Layer）とするものである。

また、第２の空間スケーラビリティは、図７に示すように、入力されたＶＯＰを構成する一部の物体（図７（Ａ）（ここでは、画像Ｆ２に相当する部分））を抜き出して（なお、このような抜き出しは、例えば、ＶＯＰ構成部２ｎにおける場合と同様にして行われ、従って、これにより抜き出された物体も、１つのＶＯＰと考えることができる）、上位レイヤとするとともに、そのＶＯＰ全体を縮小したもの（図７（Ｂ））を下位レイヤとするものである。

さらに、第３の空間スケーラビリティは、図８および図９に示すように、入力されたＶＯＰを構成する物体（ＶＯＰ）を抜き出して、その物体ごとに、上位レイヤおよび下位レイヤを生成するものである。なお、図８は、図４のＶＯＰを構成する背景（画像Ｆ１）から上位レイヤおよび下位レイヤを生成した場合を示しており、

また、図９は、図４のＶＯＰを構成する前景（画像Ｆ２）から上位レイヤおよび下位レイヤを生成した場合を示している。

以上のようなスケーラビリティのうちのいずれを用いるかは予め決められており、画像階層化部２１は、その予め決められたスケーラビリティによる符号化を行うことができるように、ＶＯＰの階層化を行う。

さらに、画像階層化部２１は、そこに入力されるＶＯＰのサイズデータおよびオフセットデータ（それぞれを、以下、適宜、初期サイズデータ、初期オフセットデータという）から、生成した下位レイヤおよび上位レイヤのＶＯＰの所定の絶対座標系における位置を表すオフセットデータと、その大きさを示すサイズデータとを計算（決定）する。

ここで、下位レイヤ並びに上位レイヤのＶＯＰのオフセットデータ（位置情報）およびサイズデータの決定方法について、例えば、上述の第２のスケーラビリティ（図７）を行う場合を例に説明する。

この場合、下位レイヤのオフセットデータＦＰＯＳ＿Ｂは、例えば、図１０（Ａ）に示すように、下位レイヤの画像データを、その解像度および上位レイヤの解像度の違いに基づいて拡大（アップサンプリング）したときに、即ち、下位レイヤの画像を、上位レイヤの画像の大きさと一致するような拡大率（上位レイヤの画像を縮小して下位レイヤの画像を生成したときの、その縮小率の逆数）（以下、適宜、倍率ＦＲという）で拡大したときに、その拡大画像の絶対座標系におけるオフセットデータが、初期オフセットデータと一致するように決定される。また、下位レイヤのサイズデータＦＳＺ＿Ｂも同様に、下位レイヤの画像を倍率ＦＲで拡大したときに得られる拡大画像のサイズデータが初期サイズデータと一致するように決定される。即ち、オフセットデータＦＰＯＳ＿ＢまたはサイズデータＦＳＺ＿Ｂは、それぞれのＦＲ倍か、初期オフセットデータまたは初期サイズデータと一致するように決定される。

一方、上位レイヤのオフセットデータＦＰＯＳ＿Ｅは、例えば、図１０（Ｂ）に示すように、入力されたＶＯＰから抜き出した物体を囲む最小長方形（ＶＯＰ）の、

例えば、左上の頂点の座標が、初期オフセットデータに基づいて求められ、この値に決定される。また、上位レイヤのサイズデータＦＰＯＳ＿Ｅは、入力されたＶＯＰから抜き出した物体を囲む最小長方形の、例えば横および縦の長さに決定される。

従って、この場合、下位レイヤのオフセットデータＦＰＯＳ＿Ｂおよびサイズデータ ＦＰＯＳ＿Ｂを、倍率ＦＲにしたがって変換し（変換後のオフセットデータＦＰＯＳ＿ＢまたはサイズデータＦＰＯＳ＿Ｂを、それぞれ、変換オフセットデータＦＰＯＳ＿Ｂまたは変換サイズデータＦＰＯＳ＿Ｂという）、絶対座標系において、変換オフセットデータＦＰＯＳ＿Ｂに対応する位置に、変換サイズデータＦＳＺ＿Ｂに対応する大きさの画枠を考え、そこに、下位レイヤの画像データをＦＲ倍だけした拡大画像を配置するとともに（図１０（Ａ））、その絶対座標系において、上位レイヤのオフセットデータＦＰＯＳ＿ＥおよびサイズデータＦＰＯＳ＿Ｅにしたがって、上位レイヤの画像を同様に配置すると（図１０（Ｂ））、拡大画像を構成する各画素と、上位レイヤの画像を構成する各画素とは、対応するものどうしが同一の位置に配置されることになる。即ち、この場合、例えば、図１０において、上位レイヤの画像である人の部分と、拡大画像の中の人の部分とは、同一の位置に配置されることになる。

第１および第３のスケーラビリティにおける場合も、同様にして、下位レイヤの拡大画像および上位レイヤの画像を構成する、対応する画素どうしが、絶対座標系において同一の位置に配置されるように、オフセットデータＦＰＯＳ＿ＢおよびＦＰＯＳ＿Ｅ、並びにサイズデータＦＳＺ＿ＢおよびＦＳＺ＿Ｅが決定される。

図５に戻り、画像階層化部２１において生成された上位レイヤの画像データ、オフセットデータＦＰＯＳ＿Ｅ、およびサイズデータＦＳＺ＿Ｅは、遅延回路２２で、後述する下位レイヤ符号化部２５における処理時間だけ遅延され、上位レイヤ符号化部２３に供給される。また、下位レイヤの画像データ、オフセットデータＦＰＯＳ＿Ｂ、およびサイズデータＦＳＺ＿Ｂは、下位レイヤ符号化部２５に供給される。また、倍率ＦＲは、遅延回路２２を介して、上位レイヤ符号化部２３および解像度変換部２４に供給される。

下位レイヤ符号化部２５では、下位レイヤの画像データが符号化され、その結果得られる符号化データ（ビットストリーム）に、オフセットデータＦＰＯＳ＿Ｂおよびサイズデータ ＦＳＺ＿Ｂが含められ、多重化部２６に供給される。

また、下位レイヤ符号化部２５は、符号化データを局所復号し、その結果局所復号結果である下位レイヤの画像データを、解像度変換部２４に出力する。解像度変換部２４は、下位レイヤ符号化部２５からの下位レイヤの画像データを、倍率ＦＲにしたがって拡大（または縮小）することにより、元の大きさに戻し、これにより得られる拡大画像を、上位レイヤ符号化部２３に出力する。

一方、上位レイヤ符号化部２３では、上位レイヤの画像データが符号化され、その結果得られる符号化データ（ビットストリーム）に、オフセットデータＦＰＯＳ＿ＥおよびサイズデータＦＳＺ＿Ｅが含められ、多重化部２６に供給される。なお、上位レイヤ符号化部２３においては、上位レイヤ画像データの符号化は、解像度変換部２４から供給される拡大画像をも参照画像として用いて行われる。

多重化部２６では、上位レイヤ符号化部２３および下位レイヤ符号化部２５の出力が多重化されて出力される。

なお、下位レイヤ符号化部２５から上位レイヤ符号化部２３に対しては、下位レイヤのサイズデータＦＳＺ＿Ｂ、オフセットデータＦＰＯＳ＿Ｂ、動きベクトルＭＶ、フラグＣＯＤなどが供給されており、上位レイヤ符号化部２３では、これらのデータを必要に応じて参照しながら、処理を行うようになされているが、この詳細については、後述する。

次に、図１１は、図５の下位レイヤ符号化部２５の詳細構成例を示している。なお、図中、図１における場合と対応する部分については、同一の符号を付してある。即ち、下位レイヤ符号化部２５は、基本的には、図１のエンコーダと同様に構成されている。

画像階層化部２１（図５）からの画像データ、即ち、下位レイヤのＶＯＰは、図１における場合と同様に、フレームメモリ３１に供給されて記憶され、動きベクトル検出器３２において、マクロブロック単位で動きベクトルの検出が行われる。

下位レイヤ符号化部２５では、下位レイヤの画像データが符号化され、その結果得られる符号化データ（ビットストリーム）に、オフセットデータＦＰＯＳ＿Ｂおよびサイズデータ ＦＳＺ＿Ｂが含められ、多重化部２６に供給される。

また、下位レイヤ符号化部２５は、符号化データを局所復号し、その結果局所復号結果である下位レイヤの画像データを、解像度変換部２４に出力する。解像度変換部２４は、下位レイヤ符号化部２５からの下位レイヤの画像データを、倍率ＦＲにしたがって拡大（または縮小）することにより、元の大きさに戻し、これにより得られる拡大画像を、上位レイヤ符号化部２３に出力する。

一方、上位レイヤ符号化部２３では、上位レイヤの画像データが符号化され、その結果得られる符号化データ（ビットストリーム）に、オフセットデータＦＰＯＳ＿ＥおよびサイズデータＦＳＺ＿Ｅが含められ、多重化部２６に供給される。なお、上位レイヤ符号化部２３においては、上位レイヤ画像データの符号化は、解像度変換部２４から供給される拡大画像をも参照画像として用いて行われる。

多重化部２６では、上位レイヤ符号化部２３および下位レイヤ符号化部２５の出力が多重化されて出力される。

なお、下位レイヤ符号化部２５から上位レイヤ符号化部２３に対しては、下位レイヤのサイズデータＦＳＺ＿Ｂ、オフセットデータＦＰＯＳ＿Ｂ、動きベクトルＭＶ、フラグＣＯＤなどが供給されており、上位レイヤ符号化部２３では、これらのデータを必要に応じて参照しながら、処理を行うようになされているが、この詳細については、後述する。

次に、図１１は、図５の下位レイヤ符号化部２５の詳細構成例を示している。なお、図中、図１における場合と対応する部分については、同一の符号を付してある。即ち、下位レイヤ符号化部２５は、基本的には、図１のエンコーダと同様に構成されている。

画像階層化部２１（図５）からの画像データ、即ち、下位レイヤのＶＯＰは、図１における場合と同様に、フレームメモリ３１に供給されて記憶され、動きベクトル検出器３２において、マクロブロック単位で動きベクトルの検出が行われる。

但し、下位レイヤ符号化部２５の動きベクトル検出器３２には、下位レイヤのＶＯＰのサイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂが供給されるようになされており、そこでは、このサイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂに基づいて、マクロブロックの動きベクトルが検出される。

即ち、上述したように、ＶＯＰは、時刻（フレーム）によって、大きさや位置が変化するため、その動きベクトルの検出にあたっては、その検出のための基準となる座標系を設定し、その座標系における動きを検出する必要がある。そこで、ここでは、動きベクトル検出器３２は、上述の絶対座標系を基準となる座標系とし、その絶対座標系に、サイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂにしたがって、符号化対象のＶＯＰおよび参照画像とするＶＯＰを配置して、動きベクトルを検出するようになされている。

なお、検出された動きベクトル（ＭＶ）は、予測モードとともに、ＶＬＣ器３６および動き補償器４２に供給される他、上位レイヤ符号化部２３（図５）にも供給される。

また、動き補償を行う場合においても、やはり、上述したように、基準となる座標系における動きを検出する必要があるため、動き補償器４２には、サイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂが供給されるようになされている。

動きベクトルの検出されたＶＯＰは、図１における場合と同様に量子化係数とされてＶＬＣ器３６に供給される。ＶＬＣ器３６には、やはり図１における場合と同様に、量子化係数、量子化ステップ、動きベクトル、および予測モードが供給される他、画像階層化部２１からのサイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂも供給されており、そこでは、これらのデータすべてが可変長符号化される。

動きベクトルの検出されたＶＯＰは、上述したように符号化される他、やはり図１における場合と同様に局所復号され、フレームメモリ４１に記憶される。この復号画像は、前述したように参照画像として用いられる他、解像度変換部２４（図５）に出力される。

なお、ＭＰＥＧ４においては、ＭＰＥＧ１および２と異なり、Ｂピクチャ（Ｂ－ＶＯＰ）も参照画像として用いられるため、Ｂピクチャも、局所復号され、フレームメモリ４１に記憶されるようになされている（但し、現時点においては、Ｂピクチャが参照画像として用いられるのは上位レイヤについてだけである）。

一方、ＶＬＣ器３６は、図１で説明したように、Ｉ，Ｐ，Ｂピクチャ（Ｉ－ＶＯＰ，Ｐ－ＶＯＰ，Ｂ－ＶＯＰ）のマクロブロックについて、スキップマクロブロックとするかどうかを決定し、その決定結果を示すフラグＣＯＤ，ＭＯＤＢを設定する。このフラグＣＯＤ，ＭＯＤＢは、やはり可変長符号化されて伝送される。さらに、フラグＣＯＤは、上位レイヤ符号化部２３にも供給される。

次に、図１２は、図５の上位レイヤ符号化部２３の構成例を示している。なお、図中、図１１または図１における場合と対応する部分については、同一の符号を付してある。即ち、上位レイヤ符号化部２３は、フレームメモリ５２が新たに設けられていることを除けば、基本的には、図１１の下位レイヤ符号化部２５または図１のエンコーダと同様に構成されている。

画像階層化部２１（図５）からの画像データ、即ち、上位レイヤのＶＯＰは、図１における場合と同様に、フレームメモリ３１に供給されて記憶され、動きベクトル検出器３２において、マクロブロック単位で動きベクトルの検出が行われる。なお、この場合も、動きベクトル検出器３２には、図１１における場合と同様に、上位レイヤのＶＯＰの他、そのサイズデータＦＳＺ＿Ｅおよびオフセットデータ ＦＰＯＳ＿Ｅが供給されるようになされており、動きベクトル検出器３２では、上述の場合と同様に、このサイズデータＦＳＺ＿ＥおよびオフセットデータＦＰＯＳ＿Ｅに基づいて、絶対座標系における上位レイヤのＶＯＰの配置位置が認識され、マクロブロックの動きベクトルが検出される。

ここで、上位レイヤ符号化部２３および下位レイヤ符号化部２５における動きベクトル検出器３２では、図１で説明したように、予め設定されている所定のシーケンスにしたがって、ＶＯＰが処理されていくが、そのシーケンスは、ここでは、例えば、次のように設定されている。

即ち、空間スケーラビリティの場合においては、図１３（Ａ）または図１３（Ｂ）に示すように、上位レイヤまたは下位レイヤのＶＯＰは、例えば、Ｐ，Ｂ，Ｂ，Ｂ，・・・またはＩ，Ｐ，Ｐ，Ｐ，・・・の順でそれぞれ処理されていく。

そして、この場合、上位レイヤの最初のＶＯＰであるＰピクチャ（Ｐ－ＶＯＰ）は、例えば、同時刻における下位レイヤのＶＯＰ（ここでは、Ｉピクチャ（Ｉ－ＶＯＰ））を参照画像として用いて符号化される。また、上位レイヤの２番目以降のＶＯＰであるＢピクチャ（Ｂ－ＶＯＰ）は、例えば、その直前の上位レイヤのＶＯＰおよびそれと同時刻の下位レイヤのＶＯＰを参照画像として用いて符号化される。即ち、ここでは、上位レイヤのＢピクチャは、下位レイヤのＰピクチャと同様に他のＶＯＰを符号化する場合の参照画像として用いられる。

なお、下位レイヤについては、例えば、ＭＰＥＧ１や２、あるいはＨ．２６３における場合と同様に符号化が行われていく。

ＳＮＲスケーラビリティは、空間スケーラビリティにおける倍率ＦＲが１のときと考えられるから、上述の空間スケーラビリティの場合と同様に処理される。

テンポラルスケーラビリティの場合、即ち、例えば、上述したように、ＶＯが、ＶＯＰ０，ＶＯＰ１，ＶＯＰ２，ＶＯＰ３，・・・で構成され、ＶＯＰ１，ＶＯＰ３，ＶＯＰ５，ＶＯＰ７，・・・が上位レイヤとされ（図１４（Ａ））、ＶＯＰ０，ＶＯＰ２，ＶＯＰ４，ＶＯＰ６，・・・が下位レイヤとされた場合においては（図１４（Ｂ））、図１４に示すように、上位レイヤまたは下位レイヤのＶＯＰは、例えば、Ｂ，Ｂ，Ｂ，・・・またはＩ，Ｐ，Ｐ，Ｐ，・・・の順でそれぞれ処理されていく。

そして、この場合、上位レイヤの最初のＶＯＰ１（Ｂピクチャ）は、例えば、下位レイヤのＶＯＰ０（Ｉピクチャ）およびＶＯＰ２（Ｐピクチャ）を参照画像として用いて符号化される。また、上位レイヤの２番目のＶＯＰ３（Ｂピクチャ）は、例えば、その直前にＢピクチャとして符号化された上位レイヤのＶＯＰ１、およびＶＯＰ３の次の時刻（フレーム）における画像である下位レイヤのＶＯＰ４（Ｐピクチャ）を参照画像として用いて符号化される。上位レイヤの３番目のＶＯＰ５

（Ｂピクチャ）も、ＶＯＰ３と同様に、例えば、その直前にＢピクチャとして符号化された上位レイヤのＶＯＰ３、およびＶＯＰ５の次の時刻（フレーム）における画像である下位レイヤのＶＯＰ６（Ｐピクチャ）を参照画像として用いて符号化される。

以上のように、あるレイヤのＶＯＰ（ここでは、上位レイヤ）については、ＰおよびＢピクチャを符号化するための参照画像として、他のレイヤ（スケーラブルレイヤ）（ここでは、下位レイヤ）のＶＯＰを用いることができる。このように、あるレイヤのＶＯＰを符号化するのに、他のレイヤのＶＯＰを参照画像として用いる場合、即ち、ここでは、上位レイヤのＶＯＰを予測符号化するのに、下位レイヤのＶＯＰを参照画像として用いる場合、上位レイヤ符号化部２３（図１２）の動きベクトル検出器３２は、その旨を示すフラグｒｅｆ＿ｌａｙｅｒ＿ｉｄ（階層数が３以上存在する場合、フラグｒｅｆ＿ｌａｙｅｒ＿ｉｄは、参照画像として用いるＶＯＰが属するレイヤを表す）を設定して出力するようになされている。

さらに、上位レイヤ符号化部２３の動きベクトル検出器３２は、ＶＯＰについてのフラグｒｅｆ＿ｌａｙｅｒ＿ｉｄにしたがい、前方予測符号化または後方予測符号化を、それぞれ、どのレイヤのＶＯＰを参照画像として行うかを示すフラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅ（参照画像情報）を設定して出力するようにもなされている。

即ち、例えば、上位レイヤ（Enhancement Layer）のＰピクチャが、その直前に復号（局所復号）される、それと同一のレイヤに属するＶＯＰを参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「００」とされる。また、Ｐピクチャが、その直前に表示される、それと異なるレイヤ（ここでは、下位レイヤ）（Reference Layer）に属するＶＯＰを参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「０１」とされる。さらに、Ｐピクチャが、その直後に表示される、それと異なるレイヤに属するＶＯＰを参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「１０」とされる。また、Ｐピクチャが、それと同時刻における、異なるレイヤの

ＶＯＰを参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「１１」とされる。

一方、例えば、上位レイヤのＢピクチャが、それと同時刻における、異なるレイヤのＶＯＰを前方予測のための参照画像として用い、かつ、その直前に復号される、それと同一のレイヤに属するＶＯＰを後方予測のための参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「００」とされる。また、上位レイヤのＢピクチャが、それと同一のレイヤに属するＶＯＰを前方予測のための参照画像として用い、かつ、その直前に表示される、それと異なるレイヤに属するＶＯＰを後方予測のための参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「０１」とされる。さらに、上位レイヤのＢピクチャが、その直前に復号される、それと同一のレイヤに属するＶＯＰを前方予測のための参照画像として用い、かつその直後に表示される、それと異なるレイヤに属するＶＯＰを後方予測のための参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「１０」とされる。また、上位レイヤのＢピクチャが、その直前に表示される、それと異なるレイヤに属するＶＯＰを前方予測のための参照画像として用い、かつその直後に表示される、それと異なるレイヤに属するＶＯＰを後方予測のための参照画像として用いて符号化される場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは「１１」とされる。

ここで、図１３および図１４で説明した予測符号化の方法は、１つの例であり、前方予測符号化、後方予測符号化、または両方向予測符号化における参照画像として、どのレイヤの、どのＶＯＰを用いるかは、例えば、上述した範囲で、自由に設定することが可能である。

なお、上述の場合においては、便宜的に、「空間スケーラビリティ」、「時間スケーラビリティ」、「ＳＮＲスケーラビリティ」という語を用いたが、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅによって、予測符号化に用いる参照画像を設定する場合、空間スケーラビリティや、テンポラルスケーラビリティ、ＳＮＲスケーラビリティを明確に区別することは困難となる。即ち、逆にいえば、フラグｒｅｆ＿ｓｅ

ｌｅｃｔ＿ｃｏｄｅを用いることによって、上述のようなスケーラビリティの区別をせずに済むようになる。

ここで、上述のスケーラビリティとフラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅとを対応付けるとすれば、例えば、次のようになる。即ち、Ｐピクチャについては、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅが「１１」の場合が、フラグｒｅｆ＿ｌａｙｅｒ＿ｉｄが示すレイヤの同時刻におけるＶＯＰを参照画像（前方予測のための参照画像）として用いる場合であるから、これは、空間スケーラビリティまたはＳＮＲスケーラビリティに対応する。そして、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅが「１１」の場合以外は、テンポラルスケーラビリティに対応する。

また、Ｂピクチャについては、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅが「００」の場合が、やはり、フラグｒｅｆ＿ｌａｙｅｒ＿ｉｄが示すレイヤの同時刻におけるＶＯＰを前方予測のための参照画像として用いる場合であるから、これが、空間スケーラビリティまたはＳＮＲスケーラビリティに対応する。そして、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅが「００」の場合以外は、テンポラルスケーラビリティに対応する。

なお、上位レイヤのＶＯＰの予測符号化のために、それと異なるレイヤ（ここでは、下位レイヤ）の、同時刻におけるＶＯＰを参照画像として用いる場合、両者の間に動きはないので、動きベクトルは、常に０（（０，０））とされる。

図１２に戻り、上位レイヤ符号化部２３の動きベクトル検出器３２では、以上のようなフラグｒｅｆ＿ｌａｙｅｒ＿ｉｄおよびｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅが設定され、動き補償器４２およびＶＬＣ器３６に供給される。

また、動きベクトル検出器３２では、フラグｒｅｆ＿ｌａｙｅｒ＿ｉｄおよびｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅにしたがって、フレームメモリ３１を参照するだけでなく、必要に応じて、フレームメモリ５２をも参照して、動きベクトルが検出される。

ここで、フレームメモリ５２には、解像度変換部２４（図５）から、局所復号された下位レイヤの拡大画像が供給されるようになされている。即ち、解像度変換部

２４では、局所復号された下位レイヤのＶＯＰが、例えば、いわゆる補間フィルタなどによって拡大され、これにより、そのＶＯＰを、ＦＲ倍だけした拡大画像、つまり、その下位レイヤのＶＯＰに対応する上位レイヤのＶＯＰと同一の大きさとした拡大画像が生成され、上位レイヤ符号化部２３に供給される。フレームメモリ５２では、このようにして解像度変換部２４から供給される拡大画像が記憶される。

従って、倍率ＦＲが１の場合は、解像度変換部２４は、下位レイヤ符号化部２５からの局所復号されたＶＯＰに対して、特に処理を施すことなく、そのまま、上位レイヤ符号化部２３に供給する。

動きベクトル検出器３２には、下位レイヤ符号化部２５からサイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂが供給されるとともに、遅延回路２２（図５）からの倍率ＦＲが供給されるようになされており、動きベクトル検出器３２は、フレームメモリ５２に記憶された拡大画像を参照画像として用いる場合、即ち、上位レイヤのＶＯＰの予測符号化に、そのＶＯＰと同時刻における下位レイヤのＶＯＰを参照画像として用いる場合（この場合、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは、Ｐピクチャについては「１１」に、Ｂピクチャについては「００」にされる）、その拡大画像に対応するサイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂに、倍率ＦＲを乗算する。そして、その乗算結果に基づいて、絶対座標系における拡大画像の位置を認識し、動きベクトルの検出を行う。

なお、動きベクトル検出器３２には、下位レイヤの動きベクトルと予測モードが供給されるようになされており、これは、次のような場合に使用される。即ち、動きベクトル検出部３２は、例えば、上位レイヤのＢピクチャについてのフラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅが「００」である場合において、倍率ＦＲが１であるとき、即ち、ＳＮＲスケーラビリティのとき（但し、この場合、上位レイヤの予測符号化に、上位レイヤのＶＯＰが用いられるので、この点で、ここでいうＳＮＲスケーラビリティは、ＭＰＥＧ２に規定されているものと異なる）、上位レイヤと下位レイヤは同一の画像であるから、上位レイヤのＢピクチャの予測符号化には、下位レイヤの同時刻における画像の動きベクトルと予測モードをそのまま用いること

ができる。そこで、この場合、動きベクトル検出部３２は、上位レイヤのＢピクチャについては、特に処理を行わず、下位レイヤの動きベクトルと予測モードをそのまま採用する。

なお、この場合、上位レイヤ符号化部２３では、動きベクトル検出器３２からＶＬＣ器３６には、動きベクトルおよび予測モードは出力されない（従って、伝送されない）。これは、受信側において、上位レイヤの動きベクトルおよび予測モードを、下位レイヤの復号結果から認識することができるからである。

以上のように、動きベクトル検出器３２は、上位レイヤのＶＯＰの他、拡大画像をも参照画像として用いて、動きベクトルを検出し、さらに、図１で説明したように、予測誤差（あるいは分散）を最小にする予測モードを設定する。また、動きベクトル検出器３２は、例えば、フラグｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅやｒｅｆ＿ｌａｙｅｒ＿ｉｄその他の必要な情報を設定して出力する。

なお、図１２では、下位レイヤ符号化部２５から、下位レイヤにおけるＩまたはＰピクチャを構成するマクロブロックがスキップマクロブロックであるかどうかを示すフラグＣＯＤが、動きベクトル検出器３２、ＶＬＣ器３６、および動き補償器４２に供給されるようになされている。

動きベクトルの検出されたマクロブロックは、上述した場合と同様に符号化され、これにより、ＶＬＣ器３６からは、その符号化結果としての可変長符号が出力される。

なお、上位レイヤ符号化部２３のＶＬＣ器３６は、下位レイヤ符号化部２５における場合と同様に、フラグＣＯＤ，ＭＯＤＢを設定して出力するようになされている。ここで、フラグＣＯＤは、上述したように、ＩまたはＰピクチャのマクロブロックがスキップマクロブロックであるかどうかを示すものであるが、フラグＭＯＤＢは、Ｂピクチャのマクロブロックがスキップマクロブロックであるかどうかを示すものである。

また、ＶＬＣ器３６には、量子化係数、量子化ステップ、動きベクトル、および予測モードの他、倍率ＦＲ、フラグｒｅｆ＿ｓｅｒｅｃｔ＿ｃｏｄｅ，ｒｅｆ＿ｌ

ａｙｅｒ＿ｉｄ、サイズデータＦＳＺ＿Ｅ、オフセットデータＦＰＯＳ＿Ｅ、も供給されるようになされており、ＶＬＣ器３６では、これらのデータがすべて可変長符号化されて出力される。

一方、動きベクトルの検出されたマクロブロックは符号化された後、やはり上述したように局所復号され、フレームメモリ４１に記憶される。そして、動き補償器４２において、動きベクトル検出器３２における場合と同様にして、フレームメモリ４１に記憶された、局所復号された上位レイヤのＶＯＰだけでなく、フレームメモリ５２に記憶された、局所復号されて拡大された下位レイヤのＶＯＰをも参照画像として用いて動き補償が行われ、予測画像が生成される。

即ち、動き補償器４２には、動きベクトルおよび予測モードの他、フラグｒｅｆ＿ｓｅｒｅｃｔ＿ｃｏｄｅ，ｒｅｆ＿ｌａｙｅｒ＿ｉｄ、倍率ＦＲ、サイズデータＦＳＺ＿Ｂ，ＦＳＺ＿Ｅ、オフセットデータＦＰＯＳ＿Ｂ，ＦＰＯＳ＿Ｅが供給されるようになされており、動き補償器４２は、フラグｒｅｆ＿ｓｅｒｅｃｔ＿ｃｏｄｅ，ｒｅｆ＿ｌａｙｅｒ＿ｉｄに基づいて、動き補償すべき参照画像を認識し、さらに、参照画像として、局所復号された上位レイヤのＶＯＰ、または拡大画像を用いる場合には、その絶対座標系における位置と大きさを、サイズデータＦＳＺ＿ＥおよびオフセットデータＦＰＯＳ＿Ｅ、またはサイズデータＦＳＺ＿ＢおよびオフセットデータＦＰＯＳ＿Ｂに基づいて認識し、必要に応じて、倍率ＦＲを用いて予測画像を生成する。

次に、図１５は、図３のエンコーダから出力されるビットストリームを復号するデコーダの一実施の形態の構成例を示している。

このデコーダには、図３のエンコーダから伝送路５または記録媒体６を介して提供されるビットストリームが供給される。即ち、図３のエンコーダから出力され、伝送路５を介して伝送されてくるビットストリームは、図示せぬ受信装置で受信され、あるいは、記録媒体６に記録されたビットストリームは、図示せぬ再生装置で再生され、逆多重化部７１に供給される。

逆多重化部７１では、そこに入力されたビットストリーム（後述するＶＳ

(Video Stream)) が受信される。さらに、逆多重化部７１では、入力されたビットストリームが、ＶＯごとのビットストリームＶＯ＃１，ＶＯ＃２，・・・に分離され、それぞれ、対応するＶＯＰ復号部７２ｎに供給される。ＶＯＰ復号部７２ｎでは、逆多重化部７１からのビットストリームから、ＶＯを構成するＶＯＰ（画像データ）、サイズデータ（VOP size）、およびオフセットデータ（VOP offset）が復号され、画像再構成部７３に供給される。

画像再構成部７３では、ＶＯＰ復号部72_1～72_Nそれぞれからの出力に基づいて、元の画像が再構成される。この再構成された画像は、例えば、モニタ７４に供給されて表示される。

次に、図１６は、スケーラビリティを実現する、図１５のＶＯＰ復号部72_Nの構成例を示している。

逆多重化部７１（図１５）から供給されるビットストリームは、逆多重化部９１に入力され、そこで、上位レイヤのＶＯＰのビットストリームと、下位レイヤのＶＯＰのビットストリームとに分離される。上位レイヤのＶＯＰのビットストリームは、遅延回路９２において、下位レイヤ復号部９５における処理の時間だけ遅延された後、上位レイヤ復号部９３に供給され、また、下位レイヤのＶＯＰのビットストリームは、下位レイヤ復号部９５に供給される。

下位レイヤ復号部９５では、下位レイヤのビットストリームが復号され、その結果得られる下位レイヤの復号画像が解像度変換部９４に供給される。また、下位レイヤ復号部９５は、下位レイヤのビットストリームを復号することにより得られるサイズデータＦＳＺ＿Ｂ、オフセットデータＦＰＯＳ＿Ｂ、動きベクトル（ＭＶ）、予測モード、フラグＣＯＤなどの、上位レイヤのＶＯＰを復号するのに必要な情報を、上位レイヤ復号部９３に供給する。

上位レイヤ復号部９３では、遅延回路９２を介して供給される上位レイヤのビットストリームが、下位レイヤ復号部９５および解像度変換部９４の出力を必要に応じて参照することにより復号され、その結果得られる上位レイヤの復号画像、サイズデータＦＳＺ＿Ｅ、およびオフセットデータＦＰＯＳ＿Ｅが出力される。さらに、

上位レイヤ復号部９３は、上位レイヤのビットストリームを復号することにより得られる倍率ＦＲを、解像度変換部９４に出力する。解像度変換部９４では、上位レイヤ復号部９３からの倍率ＦＲを用いて、図５における解像度変換部２４における場合と同様にして、下位レイヤの復号画像が変換される。この変換により得られる拡大画像は、上位レイヤ復号部９３に供給され、上述したように、上位レイヤのビットストリームの復号に用いられる。

次に、図１７は、図１６の下位レイヤ復号部９５の構成例を示している。なお、図中、図２のデコーダにおける場合と対応する部分については、同一の符号を付してある。即ち、下位レイヤ復号部９５は、基本的に、図２のデコーダと同様に構成されている。

逆多重化部９１からの下位レイヤのビットストリームは、バッファ１０１に供給され、そこで受信されて一時記憶される。ＩＶＬＣ器１０２は、その後段のブロックの処理状態に対応して、バッファ１０１からビットストリームを適宜読み出し、そのビットストリームを可変長復号することで、量子化係数、動きベクトル、予測モード、量子化ステップ、サイズデータＦＳＺ__Ｂ、オフセットデータＦＰＯＳ__Ｂ、およびフラグＣＯＤなどを分離する。量子化係数および量子化ステップは、逆量子化器１０３に供給され、動きベクトルおよび予測モードは、動き補償器１０７と上位レイヤ復号部９３（図１６）に供給される。また、サイズデータＦＳＺ__Ｂおよびオフセットデータ ＦＰＯＳ__Ｂは、動き補償器１０７、画像再構成部７３（図１５）、および上位レイヤ復号部９３に供給され、フラグＣＯＤは、上位レイヤ復号部９３に供給される。

逆量子化器１０３、ＩＤＣＴ器１０４、演算器１０５、フレームメモリ１０６、または動き補償器１０７では、図１１の下位レイヤ符号化部２５の逆量子化器３８、ＩＤＣＴ器３９、演算器４０、フレームメモリ４１、または動き補償器４２における場合とそれぞれ同様の処理が行われることで、下位レイヤのＶＯＰが復号され、画像再構成部７３、上位レイヤ復号部９３、および解像度変換部９４（図１６）に供給される。

次に、図１８は、図１６の上位レイヤ復号部９３の構成例を示している。なお、図中、図２における場合と対応する部分については、同一の符号を付してある。即ち、上位レイヤ復号部９３は、フレームメモリ１１２が新たに設けられていることを除けば、基本的に、図２のエンコーダと同様に構成されている。

逆多重化部９１からの上位レイヤのビットストリームは、バッファ１０１を介してＩＶＬＣ器１０２に供給される。ＩＶＬＣ器１０２は、上位レイヤのビットストリームを可変長復号することで、量子化係数、動きベクトル、予測モード、量子化ステップ、サイズデータＦＳＺ＿Ｅ、オフセットデータＦＰＯＳ＿Ｅ、倍率ＦＲ、フラグｒｅｆ＿ｌａｙｅｒ＿ｉｄ，ｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅ，ＣＯＤ，ＭＯＤＢなどを分離する。量子化係数および量子化ステップは、図１７における場合と同様に、逆量子化器１０３に供給され、動きベクトルおよび予測モードは、動き補償器１０７に供給される。また、サイズデータＦＳＺ＿Ｅおよびオフセットデータ ＦＰＯＳ＿Ｅは、動き補償器１０７および画像再構成部７３（図１５）に供給され、フラグＣＯＤ，ＭＯＤＢ，ｒｅｆ＿ｌａｙｅｒ＿ｉｄ、およびｒｅｆ＿ｓｅｌｅｃｔ＿ｃｏｄｅは、動き補償器１０７に供給される。さらに、倍率ＦＲは、動き補償器１０７および解像度変換部９４（図１６）に供給される。

なお、動き補償器１０７には、上述したデータの他、下位レイヤ復号部９５（図１６）から、下位レイヤの動きベクトル、フラグＣＯＤ、サイズデータＦＳＺ＿Ｂ、およびオフセットデータＦＰＯＳ＿Ｂが供給されるようになされている。また、フレームメモリ１１２には、解像度変換部９４から拡大画像が供給される。

逆量子化器１０３、ＩＤＣＴ器１０４、演算器１０５、フレームメモリ１０６、動き補償器１０７、またはフレームメモリ１１２では、図１２の上位レイヤ符号化部２３の逆量子化器３８、ＩＤＣＴ器３９、演算器４０、フレームメモリ４１、動き補償器４２、またはフレームメモリ５２における場合とそれぞれ同様の処理が行われることで、上位レイヤのＶＯＰが復号され、画像再構成部７３に供給される。

ここで、以上のように構成される上位レイヤ復号部９３および下位レイヤ復号部９５を有するＶＯＰ復号部７２ｎにおいては、上位レイヤについての復号画像、サ

イズデータＦＳＺ＿Ｅ、およびオフセットデータＦＰＯＳ＿Ｅ（以下、適宜、これらをすべて含めて、上位レイヤデータという）と、下位レイヤについての上位レイヤについての復号画像、サイズデータＦＳＺ＿Ｂ、およびオフセットデータＦＰＯＳ＿Ｂ（以下、適宜、これらをすべて含めて、下位レイヤデータという）が得られるが、画像再構成部７３では、この上位レイヤデータまたは下位レイヤデータから、例えば、次のようにして画像が再構成されるようになされている。

即ち、例えば、第１の空間スケーラビリティ（図６）が行われた場合（入力されたＶＯＰ全体が上位レイヤとされるとともに、そのＶＯＰ全体を縮小したものが下位レイヤされた場合）において、下位レイヤデータおよび上位レイヤデータの両方のデータが復号されたときには、画像再構成部７３は、上位レイヤデータのみに基づき、サイズデータＦＳＺ＿Ｅに対応する大きさの上位レイヤの復号画像（ＶＯＰ）を、オフセットデータＦＰＯＳ＿Ｅによって示される位置に配置する。また、例えば、上位レイヤのビットストリームにエラーが生じたり、また、モニタ７４が、低解像度の画像にしか対応していないため、下位レイヤデータのみの復号が行われたときには、画像再構成部７３は、その下位レイヤデータのみに基づき、サイズデータＦＳＺ＿Ｂに対応する大きさの上位レイヤの復号画像（ＶＯＰ）を、オフセットデータＦＰＯＳ＿Ｂによって示される位置に配置する。

また、例えば、第２の空間スケーラビリティ（図７）が行われた場合（入力されたＶＯＰの一部が上位レイヤとされるとともに、そのＶＯＰ全体を縮小したものが下位レイヤとされた場合）において、下位レイヤデータおよび上位レイヤデータの両方のデータが復号されたときには、画像再構成部７３は、サイズデータＦＳＺ＿Ｂに対応する大きさの下位レイヤの復号画像を、倍率ＦＲにしたがって拡大し、その拡大画像を生成する。さらに、画像再構成部７３は、オフセットデータＦＰＯＳ＿ＢをＦＲ倍し、その結果得られる値に対応する位置に、拡大画像を配置する。そして、画像再構成部７３は、サイズデータＦＳＺ＿Ｅに対応する大きさの上位レイヤの復号画像を、オフセットデータＦＰＯＳ＿Ｅによって示される位置に配置する。

この場合、上位レイヤの復号画像の部分が、それ以外の部分に比較して高い解像

度で表示されることになる。

なお、上位レイヤの復号画像を配置する場合においては、その復号画像と、拡大画像とは合成される。

また、図１６（図１５）には図示しなかったが、上位レイヤ復号部９３（ＶＯＰ復号部７２ｎ）から画像再構成部７３に対しては、上述したデータの他、倍率ＦＲも供給されるようになされており、画像再構成部７３は、これを用いて、拡大画像を生成するようになされている。

一方、第２の空間スケーラビリティが行われた場合において、下位レイヤデータのみが復号されたときには、上述の第１の空間スケーラビリティが行われた場合と同様にして、画像が再構成される。

さらに、第３の空間スケーラビリティ（図８、図９）が行われた場合（入力されたＶＯＰを構成する物体ごとに、その物体（オブジェクト）全体を上位レイヤとするとともに、その物体全体を間引いたものを下位レイヤとした場合）においては、上述の第２の空間スケーラビリティが行われた場合と同様にして、画像が再構成される。

上述したように、オフセットデータＦＰＯＳ＿ＢおよびＦＰＯＳ＿Ｅは、下位レイヤの拡大画像および上位レイヤの画像を構成する、対応する画素どうしが、絶対座標系において同一の位置に配置されるようになっているため、以上のように画像を再構成することで、正確な（位置ずれのない）画像を得ることができる。

次に、図３のエンコーダが出力する符号化ビットストリームのシンタクスについて、例えば、MPEG4 規格の Video Verification Model(Version6.0)（以下、適宜、VM6.0 と記述する）を例に説明する。

図１９は、VM6.0 における符号化ビットストリームの構成を示している。

符号化ビットストリームは、ＶＳ（Video Session Class）を単位として構成され、各ＶＳは、１以上のＶＯ（Video Object Class）から構成される。そして、ＶＯは、１以上のＶＯＬ（Video Object Layer Class）から構成され（画像を階層化しないときは１のＶＯＬで構成され、画像を階層化する場合には、その階層数だけ

のＶＯＬで構成される）、ＶＯＬは、ＶＯＰ（Video Object Plane Class）から構成される。

なお、ＶＳは、画像シーケンスであり、例えば、一本の番組や映画などに相当する。

図２０または図２１は、ＶＳまたはＶＯのシンタクスをそれぞれ示している。ＶＯは、画像全体または画像の一部（物体）のシーケンスに対応するビットストリームであり、従って、ＶＳは、そのようなシーケンスの集合で構成される（よって、ＶＳは、例えば、一本の番組などに相当する）。

図２２は、ＶＯＬのシンタクスを示している。

ＶＯＬは、上述したようなスケーラビリティのためのクラスであり、video_object_layer_id で示される番号によって識別される。即ち、例えば、下位レイヤのＶＯＬについての video_object_layer_id は０とされ、また、例えば、上位レイヤのＶＯＬについての video_object_layer_id は１とされる。なお、上述したように、スケーラブルのレイヤの数は２に限られることなく、１や３以上を含む任意の数とすることができる。

また、各ＶＯＬについて、それが画像全体であるのか、画像の一部であるのかは、video_object_layer_shape で識別される。この video_object_layer_shape は、ＶＯＬの形状を示すフラグで、例えば、以下のように設定される。

即ち、ＶＯＬの形状が長方形状であるとき、video_object_layer_shape は、例えば「００」とされる。また、ＶＯＬが、ハードキー（０または１のうちのいずれか一方の値をとる２値（Binary）の信号）によって抜き出される領域の形状をしているとき、video_object_layer_shape は、例えば「０１」とされる。さらに、ＶＯＬが、ソフトキー（０～１の範囲の連続した値（Gray-Scale）をとることが可能な信号）によって抜き出される領域の形状をしているとき（ソフトキーを用いて合成されるものであるとき）、video_object_layer_shape は、例えば「１０」とされる。

ここで、video_object_layer_shape が「００」とされるのは、ＶＯＬの形状が長方形状であり、かつ、そのＶＯＬの絶対座標形における位置および大きさが、時間

とともに変化しない、即ち、一定の場合である。なお、この場合、その大きさ（横の長さと縦の長さ）は、video_object_layer_width と video_object_layer_height によって示される。video_object_layer_width および video_object_layer_height は、いずれも１０ビットの固定長のフラグで、video_object_layer_shape が「００」の場合には、最初に、一度だけ伝送される（これは、video_object_layer_shape が「００」の場合、上述したように、ＶＯＬの絶対座標系における大きさが一定であるからである）。

また、ＶＯＬが、下位レイヤまたは上位レイヤのうちのいずれであるかは、１ビットのフラグである scalability によって示される。ＶＯＬが下位レイヤの場合、scalability は、例えば１とされ、それ以外の場合、scalability は、例えば０とされる。

さらに、ＶＯＬが、自身以外のＶＯＬにおける画像を参照画像として用いる場合、その参照画像が属するＶＯＬは、上述したように、ref_layer_id で表される。なお、ref_layer_id は、上位レイヤについてのみ伝送される。

また、図２２において、hor_sampling_factor_n と hor_sampling_factor_m は、下位レイヤのＶＯＰの水平方向の長さに対応する値と、上位レイヤのＶＯＰの水平方向の長さに対応する値をそれぞれ示す。従って、下位レイヤに対する上位レイヤの水平方向の長さ（水平方向の解像度の倍率）は、式 hor_sampling_factor_n/hor_sampling_factor_m で与えられる。

さらに、図２２において、ver_sampling_factor_n と ver_sampling_factor_m は、下位レイヤのＶＯＰの垂直方向の長さに対応する値と、上位レイヤのＶＯＰの垂直方向の長さに対応する値をそれぞれ示す。従って、下位レイヤに対する上位レイヤの垂直方向の長さ（垂直方向の解像度の倍率）は、式 ver_sampling_factor_n/ver_sampling_factor_m で与えられる。

次に、図２３は、ＶＯＰ（Video Object Plane Class）のシンタクスを示している。

ＶＯＰの大きさ（横と縦の長さ）は、例えば、１０ビット固定長の VOP_width と

VOP_height で表される。また、ＶＯＰの絶対座標系における位置は、例えば、１０ビット固定長の VOP_horizontal_spatial_mc_ref と VOP_vertical_mc_ref で表される。なお、VOP_width または VOP_height は、ＶＯＰの水平方向または垂直方向の長さをそれぞれ表し、これらは、上述のサイズデータＦＳＺ＿ＢやＦＳＺ＿Ｅに相当する。また、VOP_horizontal_spatial_mc_ref または VOP_vertical_mc_ref は、ＶＯＰの水平方向または垂直方向の座標（ｘまたはｙ座標）をそれぞれ表し、これらは、上述のオフセットデータＦＰＯＳ＿ＢやＦＰＯＳ＿Ｅに相当する。

VOP_width, VOP_height, VOP_horizontal_spatial_mc_ref、および VOP_vertical_mc_ref は、video_object_layer_shape が「００」以外の場合にのみ伝送される。即ち、video_object_layer_shape が「００」の場合、上述したように、ＶＯＰの大きさおよび位置はいずれも一定であるから、VOP_width, VOP_height, VOP_horizontal_spatial_mc_ref、および VOP_vertical_mc_ref は伝送する必要がない。この場合、受信側では、ＶＯＰは、その左上の頂点が、例えば、絶対座標系の原点に一致するように配置され、また、その大きさは、図２２で説明した video_object_layer_width および video_object_layer_height から認識される。

図２３において、ref_select_code は、図１９で説明したように、参照画像として用いる画像を表すもので、ＶＯＰのシンタクスにおいて規定されている。

ところで、VM6.0 では、各 VOP(Video Object Plane:従来の Frame に相当する)の表示時刻は、modulo_time_base と、VOP_time_increment（図２３）によって、次のように定められる。

即ち、modulo_time_base は、エンコーダのローカルな時間軸上における時刻を、１秒（1000ms（ミリ秒））の精度で表す。modulo_time_base は、VOP ヘッダの中で伝送されるマーカ（marker）で表現され、必要な数の「1」と、１の「0」とで構成される。modulo_time_base を構成する「1」の数が、最後に（現在から遡って、最も最近に）（直前に）符号化／復号された modulo_time_base によって示された同期点（１秒精度の時刻）からの累積時間を表す。即ち、modulo_time_base が、例えば、「０」の場合は、直前に符号化／復号された modulo_time_base によって示さ

れた同期点からの累積時間が０秒であることを表す。また、modulo_time_baseが、例えば、「１０」の場合は、直前に符号化／復号されたmodulo_time_baseによって示された同期点からの累積時間が１秒であることを表す。さらに、modulo_time_baseが、例えば、「１１０」の場合は、直前に符号化／復号されたmodulo_time_baseによって示された同期点からの累積時間が２秒であることを表す。以上のように、modulo_time_baseの「１」の数が、直前に符号化／復号されたmodulo_time_baseによって示された同期点からの秒数になっている。

なお、VM6.0では、modulo_time_baseについて、「This value represents the local time base at the one second resolution unit (1000 milliseconds). It is represented as a marker transmitted in the VOP header. The number of consecutive "1" followed by a "0" indicates the number of seconds has elapsed since the synchronization point marked by the last encoded/decoded modulo_time_base.」と記載されている。

VOP_time_incrementは、エンコーダのローカルな時間軸上における時刻を、1msの精度で表す。VM6.0では、VOP_time_incrementは、I-VOPおよびP-VOPについては、直前に符号化／復号されたmodulo_time_baseによって示された同期点からの時間を表し、B-VOPについては、直前に符号化／復号されたI-VOPまたはP-VOPからの相対時間を表す。

なお、VM6.0では、VOP_time_incrementについて、「This value represents the local time base in the units of milliseconds. For I and P-VOP's this value is the absolute VOP_time_increment from the synchronization point marked by the last modulo_time_base. For the B-VOP's this value is the relative VOP_time_increment from the last encoded/decoded I- or P-VOP.」と記載されている。

そして、VM6.0では、「At the encoder, the following formula are used to determine the absolute and relative VOP_time_increments for I/P-VOP's and B-VOP's, respectively.」と記載されている。

即ち、エンコーダにおいて、以下の式を使って、I-VOPおよびP-VOPと、B-VOPとについて、それぞれの表示時刻を符号化する旨が規定されている。

$$tGTB(n) = n \times 1000ms + tEST$$

$$tAVTI = tETB(I/P) - tGTB(n)$$

$$tRVTI = tETB(B) - tETB(I/P)$$

・・・（1）

但し、式（1）において、tGTB（n）は、n番目に符号化されたmodulo_time_baseによって示された同期点の時刻（上述したように、秒精度）を表し、tESTは、エンコーダにおけるVOの符号化開始時刻（VOの符号化が開始された絶対時刻）を表す。また、tAVTIは、I-VOPまたはP-VOPについてのVOP_time_incrementを表し、tETB（I/P）は、エンコーダにおけるI-VOPまたはP-VOPの符号化開始時刻（VOPの符号化が開始された絶対時刻）を表す。さらに、tRVTIは、B-VOPについてのVOP_time_incrementを表し、tETB（B）は、エンコーダにおけるB-VOPの符号化開始時刻を表す。

なお、VM6.0では、式（1）におけるtGTB（n），tEST，tAVTI，tETB（I/P），tRVTI，tETB（B）について、「tGTB（n） is the encoder time base marked by the nth encoded modulo_time_base, tEST is the encoder time base start time, tAVTI is the absolute VOP_time_increment for the I or P-VOP, tETB（I/P） is the encoder time base at the start of the encoding of the I or P-VOP, tRVTI is the relative VOP_time_increment for the B-VOP, and tETB（B） is the encoder time base at the start of the encoding of the B-VOP.」と記載されている。

また、VM6.0では、「At the decoder, the following formula are used to determine the recovered time base of the I/P-VOP's and B-VOP's, respectively:」と記載されている。

即ち、デコーダ側では、以下の式を使って、I-VOPおよびP-VOPと、B-VOPについ

て、それぞれの表示時刻を復号する旨が規定されている。

$$tGTB(n) = n \times 1000ms + tDST$$
$$tDTB(I/P) = tAVTI + tGTB(n)$$
$$tDTB(B) = tRVTI + tDTB(I/P)$$

・・・（2）

但し、式（2）において、tGTB（n）は、n番目に復号されたmodulo_time_baseによって示された同期点の時刻を表し、tDSTは、デコーダにおけるVOの復号開始時刻（VOの復号が開始された絶対時刻）を表す。また、tDTB（I／P）は、デコーダにおけるI-VOPまたはP-VOPの復号開始時刻を表し、tAVTIは、I-VOPまたはP-VOPについてのVOP_time_incrementを表す。さらに、tDTB（B）は、デコーダにおけるB-VOPの復号開始時刻（VOPの復号が開始された絶対時刻）を表し、tRVTIは、B-VOPについてのVOP_time_incrementを表す。

なお、VM6.0では、式（2）におけるtGTB（n），tDST，tDTB（I／P），tAVTI，tDTB（B），tRVTIについて、「tGTB（n） is the encoding time base marked by the nth decoded modulo_time_base, tDST is the decoding time base start time, tDTB（I／P） is the decoding time base at the start of the decoding of the I or P-VOP, tAVTI is the decoding absolute VOP_time_increment for the I or P-VOP, tDTB（B） is the decoding time base at the start of the decoding of the B-VOP, and tRVTI is the decoded relative VOP_time_increment for the B-VOP.」と記載されている。

図２４は、以上の定義に基づいて、modulo_time_baseとVOP_time_incrementとの関係を示した図である。

図２４において、VOは、Ｉ１（Ｉ－ＶＯＰ），Ｂ２（Ｂ－ＶＯＰ），Ｂ３，Ｐ４（Ｐ－ＶＯＰ），Ｂ５，Ｐ６，・・・というＶＯＰのシーケンスで構成されている。いま、VOの符号化／復号開始時刻（絶対時刻）をｔ０とすると、

modulo_time_baseは、時刻ｔ０からの経過時間を、１秒精度で表すから、ｔ０＋１秒、ｔ０＋２秒，・・・という時刻（同期点）を表す。なお、図２４において、表示順は、Ｉ１，Ｂ２，Ｂ３，Ｐ４，Ｂ５，Ｐ６，・・・であるが、符号化／復号順は、Ｉ１，Ｐ４，Ｂ２，Ｂ３，Ｐ６，・・・である。

図２４では（後述する図２８～図３１、および図３６においても同様）、各ＶＯＰについてのVOP_time_incrementを、四角形で囲んだ数字（単位はms）で示してあり、modulo_time_baseによって示される同期点の切り替わりを、▼印で示してある。従って、図２４では、Ｉ１，Ｂ２，Ｂ３，Ｐ４，Ｂ５，Ｐ６についてのVOP_time_incrementが、３５０ｍｓ，４００ｍｓ，８００ｍｓ，５５０ｍｓ，４００ｍｓ，３５０ｍｓとそれぞれされており、Ｐ４およびＰ６において、同期点が切り替わっている。

いま、図２４において、Ｉ１のVOP_time_incrementは、３５０ msであるから、Ｉ１の符号化／復号時刻は、直前に符号化／復号されたmodulo_time_baseによって示された同期点から３５０ｍｓ後の時刻となる。なお、符号化／復号の開始直後は、その開始時刻（符号化／復号開始時刻）ｔ０が同期点となるので、Ｉ１の符号化／復号時刻は、符号化／復号開始時刻ｔ０から３５０ｍｓ後の時刻ｔ０＋３５０ｍｓということになる。

そして、ＢＢ２またはＢ３の符号化／復号時刻は、直前に符号化／復号されたI-VOPまたはP-VOPから、VOP_time_incrementだけ経過した時刻であるから、いまの場合、最後の符号化／復号されたＩ１の符号化／復号時刻ｔ０＋３５０ｍｓから、４００ｍｓまたは８００ｍｓ後の時刻ｔ０＋７５０ｍｓまたはｔ０＋１２００ｍｓということに、それぞれなる。

次に、Ｐ４についてであるが、Ｐ４では、modulo_time_baseによって示される同期点が切り替わっており、従って、同期点は時刻ｔ０＋１秒となる。その結果、Ｐ４の符号化／復号時刻は、時刻ｔ０＋１秒から５５０ｍｓ後の時刻（ｔ０＋１）秒＋５５０ｍｓということになる。

Ｂ５の符号化／復号時刻は、直前に符号化／復号されたI-VOPまたはP-VOPから、

VOP_time_increment だけ経過した時刻であるから、いまの場合、最後の符号化／復号されたＰ４の符号化／復号時刻（ｔ０＋１）秒＋５５０ｍｓから、４００ｍｓ後の時刻（ｔ０＋１）秒＋９５０ｍｓということになる。

次に、Ｐ６についてであるが、Ｐ６では、modulo_time_base によって示される同期点が切り替わっており、従って、同期点は時刻ｔ０＋２秒となる。その結果、Ｐ６の符号化／復号時刻は、時刻ｔ０＋２秒から３５０ｍｓ後の時刻（ｔ０＋２）秒＋３５０ｍｓということになる。

なお、VM6.0では、modulo_time_base によって示される同期点の切り替わりは、Ｉ－ＶＯＰとＰ－ＶＯＰとに対してだけ許されており、Ｂ－ＶＯＰに対しては許されていない。

また、VM6.0において、VOP_time_increment が、I－VOP と P－VOP については、直前に符号化／復号された modulo_time_base によって示された同期点からの時間を表すのに対し、B-VOP についてだけは、直前に符号化／復号された I-VOP または P-VOP からの相対時間を表すこととされているのは、主として、次のような理由による。即ち、B-VOP は、表示順で、その B-VOP を挟む I－VOP または P－VOP を参照画像として予測符号化されるので、その予測符号化時に参照画像として用いる I－VOP または P－VOP に対する重みを、B-VOP から、それを挟む I－VOP または P－VOP までの時間的距離に基づいて決めるために、その時間的距離を、B-VOP についての VOP_time_increment としたことによる。

ところで、上述した VM6.0 の VOP_time_increment の定義では、不都合が生じる。即ち、図２４では、B-VOP についての VOP_time_increment が、その B-VOP の直前に符号化／復号される I-VOP または P-VOP からの相対時間ではなく、直前に表示される I-VOP または P-VOP からの相対時間を表すものとしてある。これは、次のような理由による。即ち、例えば、Ｂ２やＢ３に注目した場合、その直前に符号化／復号される I-VOP または P-VOP は、上述した符号化／復号順からいって、Ｐ４である。従って、B-VOP について、VOP_time_increment が、その B-VOP の直前に符号化／復号された I-VOP または P-VOP からの相対時間を表すとした場合、Ｂ２やＢ３につい

てのVOP_time_incrementは、Ｐ４の符号化／復号時刻からの相対時間を表すこととなり、負の値になる。

一方、MPEG4規格では、VOP_time_incrementは、１０ビットとされており、０以上の値のみをとるものとすれば、０～１０２３の範囲の値を表現することができるから、隣接する同期点の間の位置を、時間的に前（図２４において左方向）に位置する同期点を基準として、1ms単位で表すことができる。

しかしながら、VOP_time_incrementが、０以上の値だけでなく、負の値もとることを許すと、例えば、隣接する同期点の間の位置が、時間的に前に位置する同期点を基準として表されたり、また、時間的に後に位置する同期点を基準として表されたりすることになるため、ＶＯＰの符号化時刻や復号時刻を求める処理が煩雑になる。

従って、VM6.0では、上述したように、VOP_time_incrementが、「This value represents the local time base in the units of milliseconds. For I and P-VOP's this value is the absolute VOP_time_increment from the synchronization point marked by the last modulo_time_base. For the B-VOP's this value is the relative VOP_time_increment from the last encoded/decoded I- or P-VOP.」と定義されているが、最後の文の“For the B-VOP's this value is the relative VOP_time_increment from the last encoded/decoded I- or P-VOP”は、“For the B-VOP's this value is the relative VOP_time_increment from the last displayed I- or P-VOP”と変更するべきであり、これにより、VOP_time_incrementが、直前に符号化／復号されたI-VOPまたはP-VOPからの相対時間ではなく、直前に表示されるI-VOPまたはP-VOPからの相対時間を表すものと定義すべきである。

VOP_time_incrementを、このような定義にすることにより、B-VOPについての符号化／復号時刻の計算の基準が、B-VOPよりも過去の表示時刻を持つI/P-VOP（I-VOPまたはP-VOP）の表示時刻になるので、B-VOPについてのVOP_time_incrementは、それが参照するI-VOPが、そのB-VOPよりも先に表示されない限り、常に、正

の値をとることになり、従って、I/P-VOPのVOP_time_incrementも、常に正の値をとることになる。

また、図２４では、さらにVM6.0の定義を変更して、modulo_time_baseおよびVOP_time_incrementによって表される時刻が、符号化／復号時刻ではなく、VOPの表示時刻であるとしてある。即ち、図２４では、VOPのシーケンス上の絶対時刻を考えた場合に、式（１）におけるｔＥＳＴ（Ｉ／Ｐ）および式（２）におけるｔＤＴＢ（Ｉ／Ｐ）は、ＩまたはP-VOPが位置するシーケンス上の絶対時刻を、式（１）におけるｔＥＳＴ（Ｂ）および式（２）におけるｔＤＴＢ（Ｂ）は、B-VOPが位置するシーケンス上の絶対時刻を、それぞれ表すものとしてある。

次に、VM6.0では、式（１）における符号化開始時刻ｔＥＳＴ（the encoder time base start time)は符号化されず、その符号化開始時刻ｔＥＳＴと、各VOPの表示時刻（VOPのシーケンス上の各VOPの位置を表す絶対時刻）との差分情報としてのmodulo_time_baseおよびVOP_time_incrementが符号化される。このため、デコーダ側では、modulo_time_baseおよびVOP_time_incrementを用いて、各VOPの間の相対的な時間関係は定めることができるが、各VOPの絶対的な表示時刻、即ち、各VOPが、VOPのシーケンスの中のどの位置にあるものなのかを定めることはできない。従って、modulo_time_baseおよびVOP_time_incrementだけでは、ビットストリームの途中にアクセスすること、つまり、ランダムアクセスを行うことはできない。

一方、単に符号化開始時刻ｔＥＳＴを符号化すると、デコーダでは、それを用いて、各VOPの絶対時刻を復号することはできるが、常に、符号化ビットストリームの先頭から、符号化開始時刻ｔＥＳＴと、各VOPの相対的な時間情報であるmodulo_time_baseおよびVOP_time_incrementを復号しながら、それを累積して、絶対時刻を管理する必要があり、これは面倒であり、効率的なランダムアクセスができない。

そこで、本実施の形態では、容易に、効率的なランダムアクセスを行うことができるように、VM6.0の符号化ビットストリームの構成（階層）の中に、VOPのシー

ケンス上の絶対時刻を符号化する階層（この階層は、スケーラビリティを実現する階層（上述の下位レイヤや上位レイヤ）ではなく、符号化ビットストリームの階層である）を導入する。この階層は、符号化ビットストリームの先頭だけでなく、適当な位置に挿入できるような符号化ビットストリームの階層とする。

ここでは、この階層として、例えば、MPEG1/2で用いられているGOP(Group of Picture)層と同様に規定されるものを導入する。これにより、MPEG4に独自な符号化ストリームの階層を用いる場合に比べて、MPEG4と、MPEG1/2とのコンパチビリティ（Compatibility）を高めることができる。この新規に導入する階層を、ここでは、ＧＯＶ（またはＧＶＯＰ）（Group Of Video Object Plane）と呼ぶ。

図２５は、VOPのシーケンス上の絶対時刻を符号化するGOV層を導入した符号化ビットストリームの構成例を示している。

GOV層は、ビットストリームの先頭だけでなく、符号化ビットストリームの任意の位置に挿入することができるように、VOL層とVOP層との間に規定されている。

これにより、あるVOL#0が、VOP#0, VOP#1, ・・・, VOP#n, VOP#(n+1), ・・・, VOP#mといったVOPのシーケンスで構成される場合において、GOV層は、その先頭のVOP#0の直前だけでなく、VOP#(n+1)の直前にも挿入することができる。従って、エンコーダにおいて、GOV層は、例えば、符号化ストリームの中の、ランダムアクセスさせたい位置に挿入することができ、従って、GOV層を挿入することで、あるVOLを構成するVOPの一連のシーケンスは、GOV層によって、複数のグループ（以下、適宜、GOVという）に分けられて符号化されることになる。

GOV層のシンタクス（Syntax）は、例えば、図２６に示すように定義される。

同図に示すように、GOV層は、グループスタートコード（group_start_code）、タイムコード（time_code）、クローズドGOP（closed_gop）、ブロークンリンク（broken_link）、ネクストスタートコード（next_start_code()）が順次配置されて構成される。

次に、GOV層のセマンティクス（Semantics）について説明する。なお、GOV層のセマンティクスは、基本的には、MPEG2のGOP層と同様であり、従って、特に記述

しない部分については、MPEG2Video 規格(ISO/IEC13818-2)を参照されたい。

group_start_code は、000001B8 （１６進数）で、GOV の開始位置を示す。

time_code は、図２７に示すような、１ビットの drop_frame_flag、５ビットの time_code_hours、６ビットの time_code_minutes、１ビットの marker_bit、６ビットの time_code_seconds、および６ビットの time_code_pictures の合計２５ビットで構成される。

time_code は、IEC standard publication 461 で規定されている「time and control codes for video tape recorders」に相当する。ここで、MPEG4 では、ビデオのフレームレート（Video の Frame Rate）の概念がないので（従って、VOP は、任意の時刻に表示することができる）、ここでは、time_code がドロップフレームモード（drop_frame_mode）で記述されているか否かを示す drop_frame_flag を利用せず、その値は、例えば、0 に固定する。同様の理由で、time_code_pictures も利用せず、その値は、例えば、0 に固定する。従って、ここでは、time_code は、時刻の時間の単位を表す time_code_hours、時刻の分の単位を表す time_code_minutes、および時刻の秒の単位を表す time_code_seconds によって、GOV の先頭の時刻を表す。その結果、GOV 層の time_code（符号化開始秒精度絶対時刻）は、秒精度で、その先頭の時刻、即ち、その GOV 層の符号化が開始された、VOP のシーケンス上の絶対時刻を表現することとなる。このため、本実施の形態では、秒より細かい精度の時刻（時間）（ここでは、ミリ秒）は、VOP 毎に設定する。

なお、time_code の marker_bit は、符号化ビットストリームにおいて、０が２３個以上連続しないように１とされる。

closed_gop は、MPEG2Video 規格(ISO/IEC 13818-2)における close_gop の定義の記載の中のＩ，Ｐ、またはＢピクチャを、I-VOP, P-VOP、または B-VOP にそれぞれ置き換えたものを意味し、従って、ある GOV の中の B-VOP が、その GOV を構成する VOP だけでなく、他の GOV を構成する VOP を参照画像として符号化されているかどうかを表す。ここで、以下に、MPEG2Video 規格(ISO/IEC 13818-2)における

close_gop の定義について、上述のような置き換えを行った文を示す。

This is a one-bit flag which indicates the nature of the predictions used in the first consecutive B-VOPs (if any) immediately following the first coded I-VOP following the group of plane header. The closed_gop is set to 1 to indicate that these B-VOPs have been encoded using only backward prediction or intra coding. This bit is provided for use during any editing which occurs after encoding. If the previous pictures have been removed by editing, broken_link may be set to 1 so that a decoder may avoid displaying these B-VOPs following the first I-VOP following the group of plane header. However if the closed_gop bit is set to 1, then the editor may choose not to set the broken_link bit as these B-VOPs can be correctly decoded.

broken_link も、MPEG2Video 規格(ISO/IEC 13818-2)における broken_link の記載について、closed_gop における場合と同様の置き換を行ったものを意味し、従って、GOV の先頭の B-VOP が正確に再生することができるかどうかを表す。ここで、以下に、MPEG2Video 規格(ISO/IEC 13818-2)における broken_link の定義について、上述のような置き換えを行った文を示す。

This is a one-bit flag which shall be set to 0 during encoding. It is set to 1 to indicate that the first consecutive B-VOPs (if any) immediately following the first coded I-VOP following the group of plane header may not be correctly decoded because the reference frame which is used for prediction is not available (because of the action of editing). A decoder may use this flag to avoid displaying frames that cannot be correctly decoded.

next_start_code()は、次の GOV の先頭の位置を与える。

以上のような GOV 層を導入し、GOV の符号化を開始する、GOV のシーケンス上の絶対時刻（以下、適宜、符号化開始絶対時刻という）を、GOV のタイムコード

time_code に設定する。さらに、上述のように、GOV 層の time_code は秒精度なので、ここでは、各 VOP の、VOP のシーケンス上の絶対時刻の、さらに細かい精度の部分を、VOP 毎に設定する。

即ち、図２８は、図２６の GOV 層を導入した場合の time_code と、modulo_time_base および VOP_time_increment との関係を示している。

図２８において、GOV は、その先頭から、表示順で、Ｉ１，Ｂ２，Ｂ３，Ｐ４，Ｂ５，Ｐ６が配置されて構成されている。

いま、例えば、GOV の符号化開始絶対時刻を、0h:12m:35sec:350msec（０時１２分３５秒３５０ミリ秒）とすると、GOV の time_code は、上述したように、秒精度（秒単位）なので、0h:12m:35sec とされる（time_code を構成する time_code_hours, time_code_minutes、または time_code_seconds が、それぞれ０，１２、または３５とされる）。一方、Ｉ１の、VOP のシーケンス上の絶対時刻（図２８の GOV を含む VS の符号化前（または復号後）の VOP のシーケンスの絶対時刻）（これは、VOP のシーケンスが表示されるときの、Ｉ１が表示される時刻に相当するので、以下、適宜、表示時刻という）が、例えば、0h:12m:35sec:350msec である場合には、その表示時刻の、秒精度より細かい精度である 350ms は、Ｉ１についての I-VOP の VOP_time_increment に設定されて符号化されるように（Ｉ１についての VOP_time_increment＝350 とされて符号化されるように）、VOP_time_increment のセマンティクスを変更する。

即ち、図２８において、GOV の、表示順で先頭の I-VOP（Ｉ１）の VOP_time_increment は、GOV の time_code と、I-VOP の表示時刻の差分値とする。従って、秒精度による time_code で表された時刻が、GOV の最初の同期点（ここでは、秒精度の時刻を表す点）となる。

なお、図２８において、GOV の２番目以降に配置された VOP であるＢ２，Ｂ３，Ｐ４，Ｂ５，Ｐ６についての VOP_time_increment のセマンティクスは、図２４で説明したように、VM6.0 の定義を変更したものと同様である。

従って、図２８において、Ｂ２またはＢ３の表示時刻は、直前に表示される I-

VOP またはP-VOP の表示時刻から、VOP_time_increment だけ経過した時刻であるから、いまの場合、直前に表示されるＩ１の表示時刻 0h:12m:35s+350ms から、400ms または 800ms 後の時刻 0h:12m:35s:750ms または 0h:12m:36s:200ms ということに、それぞれなる。

次に、Ｐ４についてであるが、Ｐ４では、modulo_time_base によって示される同期点が切り替わっており、従って、同期点は時刻 0h:12m:35s から１秒経過した 0h:12m:36s となる。その結果、Ｐ４の表示時刻は、時刻 0h:12m:36s から 550ms 後の時刻 0h:12m:36:550ms ということになる。

Ｂ５の表示時刻は、直前に表示される I-VOP またはP-VOP から、VOP_time_increment だけ経過した時刻であるから、いまの場合、直前に表示されるＰ４の表示時刻 0h:12m:36:550ms から、400ms 後の時刻 0h:12m:36s:950ms ということになる。

そして、Ｐ６についてであるが、Ｐ６では、modulo_time_base によって示される同期点が切り替わっており、従って、同期点は時刻 0h:12m:35s+2 秒、即ち、0h:12m:37s となる。その結果、Ｐ６の表示時刻は、時刻 0h:12m:37s から 350ms 後の時刻 0h:12m:37s:350ms ということになる。

次に、図２９は、表示順で、先頭の VOP が B-VOP になっている場合の、GOV についての time_code と、modulo_time_base および VOP_time_increment との関係を示している。

図２９において、GOV は、その先頭から、表示順で、Ｂ０，Ｉ１，Ｂ２，Ｂ３，Ｐ４，Ｂ５，Ｐ６が配置されて構成されている。即ち、図２９では、図２８において、Ｉ１の前にＢ０が追加されて、GOV が構成されている。

この場合、GOV の先頭のＢ０についての VOP_time_increment を、その GOV を構成する I/P-VOP の表示時刻を基準として定めることとすると、即ち、例えば、Ｉ１の表示時刻を基準として定めることとすると、その値は負になり、上述したように、都合が悪い。

そこで、GOV の中の、I-VOP よりも先に表示される B-VOP（GOV の中で、最初に表

示されるI-VOPよりも先行して表示されるB-VOP）のVOP_time_incrementについては、そのセマンティクスを、以下のように変更する。

すなわち、そのようなB-VOPのVOP_time_incrementは、GOVのtime_codeの時刻と、B-VOPの表示時刻との差分値とする。この場合、図２９に示すように、Ｂ０の表示時刻が、例えば、0h:12m:35s:200msであり、GOVのtime_codeが、例えば、0h:12m:35sであるときには、Ｂ０のVOP_time_incrementは、350ms（＝0h:12m:35s:200ms－0h:12m:35s）になる。このようにすることで、VOP_time_incrementは、常に正の値になる。

以上のような、VOP_time_incrementについてのセマンティクスの２つの変更により、GOVのtime_codeと、VOPのmodulo_time_baseおよびVOP_time_incrementとを関係付けることができ、さらに、これにより、各VOPが表示される絶対時刻（表示時刻）を特定することができる。

次に、図３０は、I-VOPの表示時刻と、それから予測されるB-VOPの表示時刻との間隔が１秒（正確には、１．０２３秒）より大きい場合の、GOVについてのtime_codeと、modulo_time_baseおよびVOP_time_incrementとの関係を示している。

図３０において、GOVは、表示順で、Ｉ１，Ｂ２，Ｂ３，Ｂ４，Ｐ６が順次配置されて構成されており、Ｂ４が、直前に表示されるI-VOPであるＩ１の表示時刻よりも、１秒より後の時刻において表示されるようになされている。

この場合、上述のようにセマンティクスを変更したVOP_time_incrementによって、Ｂ４の表示時刻を符号化しようとしても、VOP_time_incrementは、上述のように１０ビットであるため、1023までしか表現できず、1.023秒より長い時間を表現することはできない。そこで、VOP_time_incrementのセマンティクスをさらに変更するとともに、modulo_time_baseのセマンティクスをも変更し、このような場合であっても対応できるようにする。

ここでは、例えば、次のような第１または第２の方法のいずれかによって対応する。

即ち、第１の方法では、I/P-VOPの表示時刻と、それから予測されるB-VOPの表

示時刻との間の時間を、ミリ秒精度で求め、その時間を、秒の単位までは、modulo_time_base で表現し、残りのミリ秒の精度を、VOP_time_increment で表現する。

図３０に示した場合において、第１の方法にしたがって、modulo_time_base および VOP_time_increment を符号化した場合の、GOV についての time_code と、modulo_time_base および VOP_time_increment との関係を、図３１に示す。

即ち、第１の方法では、modulo_time_base の付加を、I-VOP および P-VOP だけでなく、B-VOP に対しても許可する。そして、B-VOP に付加されている modulo_time_base は、同期点の切り替わりではなく、直前に表示される I/P-VOP の表示時刻からの秒単位の繰り上がりを表すものとする。

さらに、第１の方法では、B-VOP に付加される modulo_time_base によって示される、直前に表示される I/P-VOP の表示時刻からの秒単位の繰り上がり後の時刻を、その B-VOP の表示時刻から減算した値を、その VOP_time_increment として設定する。

従って、第１の方法によれば、図３０において、例えば、Ｉ１の表示時刻を、0h:12m:35s:350ms とするとともに、Ｂ４の表示時刻を、0h:12m:36s:550ms とすると、Ｉ１とＢ４との表示時刻の差は、１秒以上の 1200msec であるから、Ｂ４には、図３１に示すように、直前に表示されるＩ１の表示時刻からの秒単位の繰り上がりを示す modulo_time_base（図３１において、▼印で示す）が付加される。具体的には、Ｂ４に付加される modulo_time_base は、1200ms の１秒の位の値である１秒の繰り上がりを表す「１０」とされる。そして、Ｂ４の VOP_time_increment は、図３１に示すように、Ｉ１とＢ４との表示時刻の差の、１秒未満の値（Ｂ４の表示時刻から、その modulo_time_base によって示される、直前に表示される I/P-VOP であるＩ１の表示時刻からの秒単位の繰り上がり後の時刻を減算した値）であるである２００とされる。

以上のような、第１の方法による modulo_time_base と VOP_time_increment についての処理は、エンコーダ側では、例えば、図１１および図１２に示したＶＬＣ器３６において、デコーダ側では、例えば、図１７および図１８に示したＩＶＬＣ器

１０２において、それぞれ行われる。

そこで、まず、図３２のフローチャートを参照して、ＶＬＣ器３６が行うI/P-VOPのmodulo_time_baseおよびVOP_time_incrementに関する処理について説明する。

ＶＬＣ器３６は、VOPのシーケンスを、GOVごとに分けて処理を行うようになされている。なお、GOVは、少なくとも１のイントラ符号化されるVOPを含むように構成される。

ＶＬＣ器３６は、GOVを受信すると、例えば、その受信時刻を、そのGOVの符号化開始絶対時刻とし、その符号化開始絶対時刻の秒精度まで（秒の桁までの符号化開始絶対時刻）を、time_codeとして符号化して、符号化ビットストリームの中に含める。その後、ＶＬＣ器３６は、GOVを構成するI/P-VOPを受信するごとに、そのI/P-VOPを注目I/P-VOPとして、図３２のフローチャートにしたがい、注目I/P-VOPのmodulo_time_baseおよびVOP_time_incrementを求めて、符号化する。

即ち、ＶＬＣ器３６では、まず最初に、ステップＳ１において、modulo_time_baseに０Ｂ（Ｂは２進数を表す）がセットされるとともに、VOP_time_incrementに０がセットされることにより、modulo_time_baseおよびVOP_time_incrementがリセットされる。

そして、ステップＳ２に進み、注目I/P-VOPが、処理対象としているGOV（処理対象GOV）の中で、最初に表示されるI-VOP（First I-VOP）であるかどうかが判定される。ステップＳ２において、注目I/P-VOPが、処理対象GOVの中で、最初に表示されるI-VOPであると判定された場合、ステップＳ４に進み、処理対象GOVのtime_codeと、注目I/P-VOP（ここでは、処理対象GOVの中で、最初に表示されるI-VOP）の表示時刻の秒精度との差分、即ち、time_codeと、注目I/P-VOPの表示時刻の秒の桁までとの差分が求められ、変数Ｄにセットされて、ステップＳ５に進む。

また、ステップＳ２において、注目I/P-VOPが、処理対象GOVの中で、最初に表示されるI-VOPでないと判定された場合、ステップＳ３に進み、注目I/P-VOPの表示時刻の秒の桁までと、その直前に表示されるI/P-VOP（処理対象GOVを構成する

VOPのうちの、注目I/P-VOPの直前に表示されるI/P-VOP）（Last display I/P-VOP）の表示時刻の秒の桁までとの差分値が求められ、その差分値が、変数Ｄにセットされて、ステップＳ５に進む。

ステップＳ５では、変数Ｄが０に等しいかどうか、即ち、time_codeと、注目I/P-VOPの表示時刻の秒の桁までとの差分、または注目I/P-VOPの表示時刻の秒の桁までと、その直前に表示されるI/P-VOPの表示時刻の秒の桁までとの差分値が０秒であるかどうかが判定される。ステップＳ５において、変数Ｄが０に等しくないと判定された場合、即ち、変数Ｄが１以上である場合、ステップＳ６に進み、modulo_time_baseのＭＳＢ（Most Significant Bit）として、１が付加される。即ち、この場合、modulo_time_baseが、例えば、リセット直後の０Ｂであるときには、１０Ｂとされ、また、modulo_time_baseが、例えば、１０Ｂであるときには、１１０Ｂとされる。

そして、ステップＳ７に進み、変数Ｄが１だけデクリメントされ、ステップＳ５に戻る。その後、ステップＳ５において、変数Ｄが０に等しいと判定されるまで、ステップＳ５～Ｓ７の処理が繰り返される。即ち、これにより、modulo_time_baseは、time_codeと、注目I/P-VOPの表示時刻の秒の桁までとの差分、または注目I/P-VOPの表示時刻の秒の桁までと、その直前に表示されるI/P-VOPの表示時刻の秒の桁までとの差分値に対応する秒数と同一の数だけ「１」が連続し、その最後に０が付加された値とされる。

そして、ステップＳ５において、変数Ｄが０に等しいと判定された場合、ステップＳ８に進み、VOP_time_incrementに、注目I/P-VOPの表示時刻の秒精度より細かい精度の時刻、即ち、ミリ秒単位の時刻がセットされ、処理を終了する。

以上のようにして求められた注目I/P-VOPのmodulo_time_baseおよびVOP_time_incrementは、ＶＬＣ回路３６において、注目I/P-VOPに付加され、これにより、符号化ビットストリームの中に含められる。

なお、modulo_time_baseおよびVOP_time_increment、並びにtime_codeは、ＶＬＣ回路３６において可変長符号化される。

次に、ＶＬＣ器３６は、処理対象GOVを構成するB-VOPを受信するごとに、そのB-VOPを注目B-VOPとして、図３３のフローチャートにしたがい、注目B-VOPのmodulo_time_baseおよびVOP_time_incrementを求めて、符号化する。

即ち、ＶＬＣ器３６では、まず最初に、ステップＳ１１において、図３２のステップＳ１における場合と同様に、modulo_time_baseおよびVOP_time_incrementがリセットされる。

そして、ステップＳ１２に進み、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOP（First I-VOP）よりも先に表示されるものであるかどうかが判定される。ステップＳ１２において、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOPよりも先に表示されるものであると判定された場合、ステップＳ１４に進み、処理対象GOVのtime_codeと、注目B-VOP（ここでは、処理対象GOVの中で、最初に表示されるI-VOPよりも先行して表示されるB-VOP）の表示時刻との差分が求められ、変数Ｄにセットされて、ステップＳ１５に進む。従って、ここでは、変数Ｄには、ミリ秒精度の時間（ミリ秒の桁までの時間）がセットされる（これに対して、図３２における変数Ｄには、上述したように、秒精度の時間がセットされる）。

また、ステップＳ１２において、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOPよりも後に表示されるものであると判定された場合、ステップＳ１４に進み、注目B-VOPの表示時刻と、その直前に表示されるI/P-VOP（処理対象GOVを構成するVOPのうちの、注目B-VOPの直前に表示されるI/P-VOP）（Last display I/P-VOP）の表示時刻との差分値が求められ、その差分値が、変数Ｄにセットされて、ステップＳ１５に進む。

ステップＳ１５では、変数Ｄが１より大であるかどうか、即ち、time_codeと、注目B-VOPの表示時刻との差分値、または注目B-VOPの表示時刻と、その直前に表示されるI/P-VOPの表示時刻との差分値が、１秒より大であるかどうかが判定される。ステップＳ５において、変数Ｄが１より大であると判定された場合、即ち、modulo_time_baseのＭＳＢとして、１が付加され、ステップＳ１７に進む。ステッ

プＳ１７では、変数Ｄが１だけデクリメントされ、ステップＳ１５に戻る。そして、ステップＳ１５において、変数Ｄが１より大でないと判定されるまで、ステップＳ１５～Ｓ１７の処理が繰り返される。即ち、これにより、modulo_time_base は、time_code と、注目 B-VOP の表示時刻との差分値、または注目 B-VOP の表示時刻と、その直前に表示される I/P-VOP の表示時刻との差分値に対応する時間の秒数と同一の数だけ「１」が連続し、その最後に０が付加された値とされる。

そして、ステップＳ１５において、変数Ｄが１より大でないと判定された場合、ステップＳ８に進み、そのときの変数Ｄの値、即ち、time_code と、注目 B-VOP の表示時刻との差分値、または注目 B-VOP の表示時刻と、その直前に表示される I/P-VOP の表示時刻との差分値の秒の位より下の位（ミリ秒単位の時間）が、VOP_time_increment にセットされ、処理を終了する。

以上のようにして求められた注目 B-VOP の modulo_time_base および VOP_time_increment は、ＶＬＣ回路３６において、注目 B-VOP に付加され、これにより、符号化ビットストリームの中に含められる。

次に、ＩＶＬＣ器１０２は、上述のようにして、ＶＬＣ器３６が、VOP のシーケンスを、GOV ごとに分けて処理を行うことにより出力する符号化ストリームの中の VOP の表示時刻を、各 VOP についての符号化データを受信するごとに、その VOP を注目 VOP として処理を行うことにより認識し、その表示時刻に VOP が表示されるように、可変長復号を行うようになされている。即ち、ＩＶＬＣ器１０２は、GOV を受信すると、その GOV の time_code を認識し、その GOV を構成する I/P-VOP を受信するごとに、その I/P-VOP を注目 I/P-VOP として、図３４のフローチャートにしたがい、注目 I/P-VOP の modulo_time_base および VOP_time_increment に基づき、その表示時刻を求める。

即ち、ＩＶＬＣ器１０２では、まず最初に、ステップＳ２１において、注目 I/P-VOP が、処理対象 GOV の中で、最初に表示される I-VOP (First I-VOP) であるかどうかが判定される。ステップＳ２１において、注目 I/P-VOP が、処理対象 GOV の中で、最初に表示される I-VOP であると判定された場合、ステップＳ２３に進み、変

数Tに、処理対象GOVのtime_codeがセットされ、ステップＳ２４に進む。

また、ステップＳ２１において、注目I/P-VOPが、処理対象GOVの中で、最初に表示されるI-VOPでないと判定された場合、ステップＳ２２に進み、注目I/P-VOPの直前に表示されるI/P-VOP（処理対象GOVを構成するVOPのうちの、注目I/P-VOPの直前に表示されるI/P-VOP）（Last display I/P-VOP）の表示時刻の秒の桁までが、変数Tにセットされて、ステップＳ２４に進む。

ステップＳ２４では、注目I/P-VOPに付加されているmodulo_time_baseが０Ｂに等しいかどうかが判定される。ステップＳ２４において、注目I/P-VOPに付加されているmodulo_time_baseが０Ｂに等しくないと判定された場合、即ち、注目I/P-VOPに付加されているmodulo_time_baseに１が含まれる場合、ステップＳ２５に進み、そのmodulo_time_baseのＭＳＢの１が削除され、ステップＳ２６に進む。ステップＳ２６では、変数Tが１秒だけインクリメントされ、ステップＳ２４に戻り、以下、ステップＳ２４において、注目I/P-VOPに付加されているmodulo_time_baseが０Ｂに等しいと判定されるまで、ステップＳ２４～Ｓ２６の処理を繰り返す。これにより、変数Tは、注目I/P-VOPに、最初に付加されていたmodulo_time_baseの１の数に対応する秒数だけインクリメントされる。

そして、ステップＳ２４において、注目I/P-VOPに付加されているmodulo_time_baseが０Ｂに等しいと判定された場合、ステップＳ２７に進み、変数Tに、VOP_time_incrementが表すミリ秒精度の時刻が加算され、その加算値が、注目I/P-VOPの表示時刻として認識されて、処理を終了する。

次に、ＩＶＬＣ器１０２では、GOVを構成するB-VOPを受信した場合には、図３５のフローチャートにしたがい、注目B-VOPのmodulo_time_baseおよびVOP_time_incrementに基づき、その表示時刻が求められる。

即ち、ＩＶＬＣ器１０２では、まず最初に、ステップＳ３１において、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOP（First I-VOP）よりも先に表示されるものであるかどうかが判定される。ステップＳ３１において、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOPよりも先に表示されるもの

であると判定された場合、ステップＳ３３に進み、以下、ステップＳ３３～Ｓ３７において、図３４のステップＳ２３～Ｓ２７における場合とそれぞれ同様の処理が行われることにより、注目 B-VOP の表示時刻が求められる。

一方、ステップＳ３１において、注目 B-VOP が、処理対象 GOV の中で、最初に表示される I-VOP よりも後に表示されるものであると判定された場合、ステップＳ３２に進み、以下、ステップＳ３２，Ｓ３４～Ｓ３７において、図３４のステップＳ２２，Ｓ２４～Ｓ２７における場合とそれぞれ同様の処理が行われることにより、注目 B-VOP の表示時刻が求められる。

次に、第２の方法では、I-VOP の表示時刻と、それから予測される B-VOP の表示時刻との間の時間を、秒の桁まで求め、その値を、modulo_time_base で表現し、B-VOP の表示時刻のミリ秒精度を、VOP_time_increment で表現する。即ち、VM6.0 では、上述したように、B-VOP の予測符号化時に参照画像として用いる I－VOP または P－VOP に対する重みを、B-VOP から、それを挟む I－VOP または P－VOP までの時間的距離に基づいて決めるのに、その時間的距離を、B-VOP についての VOP_time_increment としており、このため、I－VOP および P－VOP についての VOP_time_increment が、直前に符号化／復号された modulo_time_base によって示された同期点からの時間を表すのと異なっているが、B-VOP の表示時刻と、それを挟む I－VOP または P－VOP の表示時刻が分かれば、それらの間の時間的距離は、差分をとるだけで求めることができ、従って、B-VOP についての VOP_time_increment だけを、I－VOP および P－VOP についての VOP_time_increment と異なる取り扱いとする必要性は小さい。むしろ、処理効率の観点からは、Ｉ，Ｂ，Ｐのすべての VOP の VOP_time_incrment（詳細時間情報）、さらには、modulo_time_base（秒精度時刻情報）は、同一の取り扱いとするのが望ましい。

そこで、第２の方法では、B-VOP についての modulo_time_base および VOP_time_increment を、I/P－VOP についてのものと同一の取り扱いとしている。

例えば、図３０に示した場合において、第２の方法にしたがって、modulo_time_base および VOP_time_increment を符号化した場合の、GOV についての

time_code と、modulo_time_base および VOP_time_increment との関係を、図３６に示す。

即ち、第２の方法でも、modulo_time_base の付加を、I-VOP および P-VOP だけでなく、B-VOP に対しても許可する。そして、B-VOP に付加される modulo_time_base も、I/P-VOP に付加される modulo_time_base と同様に、同期点の切り替わりを表すものとする。

さらに、第２の方法では、B-VOP に付加されている modulo_time_base によって示される同期点の時刻を、その B-VOP の表示時刻から減算した値が、その VOP_time_increment として設定される。

従って、第２の方法によれば、図３０において、GOV の time_code が表す時刻である、GOV の最初の同期点から、時刻 time_code＋１秒で示される同期点までの間に表示されるＩ１またはＢ２の modulo_time_base は、いずれも０Ｂとされるとともに、それぞれの VOP_time_increment は、Ｉ１またはＢ２の表示時刻の秒の位より下の位のミリ秒単位の値がセットされる。また、時刻 time_code＋１秒で示される同期点から、時刻 time_code＋２秒で示される同期点までの間に表示されるＢ３またはＢ４の modulo_time_base は、いずれも１０Ｂとされるとともに、それぞれの VOP_time_increment は、Ｂ３またはＢ４の表示時刻の秒の位より下の位のミリ秒単位の値がセットされる。さらに、時刻 time_code＋２秒で示される同期点から、時刻 time_code＋３秒で示される同期点までの間に表示されるＰ５の modulo_time_base は１１０Ｂとされるとともに、その VOP_time_increment は、Ｐ５の表示時刻の秒の位より下の位のミリ秒単位の値がセットされる。

図３０において、例えば、上述したように、Ｉ１の表示時刻を、0h:12m:35s:350ms とするとともに、Ｂ４の表示時刻を、0h:12m:36s:550ms とすると、Ｉ１またはＢ４の modulo_time_base は、上述したように、それぞれ０Ｂまたは１０Ｂとされる。また、Ｉ１またはＢ４の VOP_time_increment は、それぞれ、表示時刻のミリ秒単位である 350ms または 550ms とされる。

以上のような、第２の方法による modulo_time_base と VOP_time_increment につ

いての処理も、例えば、第１の方法による場合と同様に、図１１および図１２に示したＶＬＣ器３６、並びに図１７および図１８に示したＩＶＬＣ器１０２において行われる。

即ち、ＶＬＣ器３６では、I/P-VOPについては、図３２における場合と同様にして、modulo_time_baseおよびVOP_time_incrementが求められる。

また、B-VOPについては、ＶＬＣ器３６において、GOVを構成するB-VOPを受信するごとに、そのB-VOPを注目B-VOPとして、図３７のフローチャートにしたがい、注目B-VOPのmodulo_time_baseおよびVOP_time_incrementが求められる。

即ち、ＶＬＣ器３６では、まず最初に、ステップＳ４１において、図３２のステップＳ１における場合と同様にして、modulo_time_baseおよびVOP_time_incrementがリセットされる。

そして、ステップＳ４２に進み、注目B-VOPが、処理対象としているGOV（処理対象GOV）の中で、最初に表示されるI-VOP（First I-VOP）よりも先行して表示されるものであるかどうかが判定される。ステップＳ１２において、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOPよりも先行して表示されるものであると判定された場合、ステップＳ４４に進み、処理対象GOVのtime_codeと、注目B-VOPの表示時刻の秒精度との差分、即ち、time_codeと、注目B-VOPの表示時刻の秒の桁までとの差分が求められ、変数Ｄにセットされて、ステップＳ４５に進む。

また、ステップＳ４２において、注目B-VOPが、処理対象GOVの中で、最初に表示されるI-VOPよりも後に表示されるものであると判定された場合、ステップＳ４３に進み、注目B-VOPの表示時刻の秒の桁までと、その直前に表示されるI/P-VOP（処理対象GOVを構成するVOPのうちの、注目B-VOPの直前に表示されるI/P-VOP）（Last display I/P-VOP）の表示時刻の秒の桁までとの差分値が求められ、その差分値が、変数Ｄにセットされて、ステップＳ４５に進む。

ステップＳ４５では、変数Ｄが０に等しいかどうか、即ち、time_codeと、注目B-VOPの表示時刻の秒の桁までとの差分、または注目B-VOPの表示時刻の秒の桁までと、その直前に表示されるI/P-VOPの表示時刻の秒の桁までとの差分値が０秒で

あるかどうかが判定される。ステップＳ４５において、変数Ｄが０に等しくないと判定された場合、即ち、変数Ｄが１以上である場合、ステップＳ４６に進み、modulo_time_base のＭＳＢとして、１が付加される。

そして、ステップＳ４７に進み、変数Ｄが１だけデクリメントされ、ステップＳ４５に戻る。その後、ステップＳ４５において、変数Ｄが０に等しいと判定されるまで、ステップＳ４５～Ｓ４７の処理が繰り返される。即ち、これにより、modulo_time_base は、time_code と、注目 B-VOP の表示時刻の秒の桁までとの差分、または注目 B-VOP の表示時刻の秒の桁までと、その直前に表示される I/P-VOP の表示時刻の秒の桁までとの差分値に対応する秒数と同一の数だけ「１」が連続し、その最後に０が付加された値とされる。

そして、ステップＳ４５において、変数Ｄが０に等しいと判定された場合、ステップＳ４８に進み、VOP_time_increment に、注目 B-VOP の表示時刻の秒精度より細かい精度の時刻、即ち、ミリ秒単位の時刻がセットされ、処理を終了する。

一方、ＩＶＬＣ器１０２では、I/P-VOP については、上述の図３４における場合と同様にして、modulo_time_base および VOP_time_increment に基づき、その表示時刻が求められる。

また、Ｂ－VOP については、ＩＶＬＣ器１０２において、GOV を構成する B-VOP を受信するごとに、その B-VOP を注目 B-VOP として、図３８のフローチャートにしたがい、注目 B-VOP の modulo_time_base および VOP_time_increment に基づき、その表示時刻が求められる。

即ち、ＩＶＬＣ器１０２では、まず最初に、ステップＳ５１において、注目 B-VOP が、処理対象 GOV の中で、最初に表示される I-VOP (First I-VOP) よりも先行して表示されるものであるかどうかが判定される。ステップＳ５１において、注目 B-VOP が、処理対象 GOV の中で、最初に表示される I-VOP よりも先行して表示されるものであると判定された場合、ステップＳ５２に進み、変数Ｔに、処理対象 GOV の time_code がセットされ、ステップＳ５４に進む。

また、ステップＳ５１において、注目 B-VOP が、処理対象 GOV の中で、最初に表

示されるI-VOPよりも後に表示されるものであると判定された場合、ステップＳ５３に進み、注目B-VOPの直前に表示されるI/P-VOP（処理対象GOVを構成するVOPのうちの、注目B-VOPの直前に表示されるI/P-VOP）（Last display I/P-VOP）の表示時刻の秒の桁までが、変数Ｔにセットされて、ステップＳ５４に進む。

ステップＳ５４では、注目B-VOPに付加されているmodulo_time_baseがＯＢに等しいかどうかが判定される。ステップＳ５４において、注目B-VOPに付加されているmodulo_time_baseがＯＢに等しくないと判定された場合、即ち、注目B-VOPに付加されているmodulo_time_baseに１が含まれる場合、ステップＳ５５に進み、そのmodulo_time_baseのＭＳＢの１が削除され、ステップＳ５６に進む。ステップＳ５６では、変数Ｔが１秒だけインクリメントされ、ステップＳ５４に戻り、以下、ステップＳ５４において、注目B-VOPに付加されているmodulo_time_baseがＯＢに等しいと判定されるまで、ステップＳ５４～Ｓ５６の処理を繰り返す。これにより、変数Ｔは、注目B-VOPに、最初に付加されていたmodulo_time_baseの１の数に対応する秒数だけインクリメントされる。

そして、ステップＳ５４において、注目B-VOPに付加されているmodulo_time_baseがＯＢに等しいと判定された場合、ステップＳ５７に進み、変数Ｔに、VOP_time_incrementが表すミリ秒精度の時刻が加算され、その加算値が、注目B-VOPの表示時刻として認識されて、処理を終了する。

以上のように、符号化ビットストリームの構成（階層）の中に、符号化開始絶対時刻を符号化するGOV層を導入し、このGOV層を、ビットストリームの先頭だけでなく、適当な位置に挿入できるようなものとするとともに、VM6.0で規定されていたmodulo_time_baseとVOP_time_incrementの定義を、上述のように変更したので、VOPのピクチャタイプの並びや、隣接するVOPの時間間隔などによらず、すべての場合において、各VOPの表示時刻（絶対時刻）を求めることが可能となる。

従って、エンコーダにおいて、符号化開始絶対時刻を、GOV単位で符号化するとともに、各VOPのmodulo_time_baseとVOP_time_incrementを符号化し、符号化ビットストリームに含めることで、デコーダでは、符号化開始絶対時刻を、GOV単位

で復号するとともに、VOPのmodulo_time_baseおよびVOP_time_incrementを復号し、それらから、各VOPの表示時刻を復号することができるので、ランダムアクセスを、GOV単位で、効率的に行うことが可能となる。

なお、modulo_time_baseに付加する１の数を、同期点の切り替わりにしたがって、単純に増加していくと、time_codeが示す時刻から、例えば、１時間（３６００秒）経過後は（但し、GOVが、それだけの時間に相当するVOPで構成されるとする）、modulo_time_baseは、３６００ビットの１と、１ビットの０とで構成されるから、３６０１ビットという莫大なビット数になることになる。

そこで、MPEG4では、modulo_time_baseは、同期点の切り替わり後に最初に現れるI/P-VOPにおいてリセットされるように規定されている。

従って、例えば、図３９に示すように、GOVが、そのtime_codeが表す時刻である、GOVの最初の同期点から、時刻time_code＋１秒で示される同期点までの間に表示されるＩ１およびＢ２、時刻time_code＋１秒で示される同期点から、時刻time_code＋２秒で示される同期点までの間に表示されるＢ３およびＢ４、時刻time_code＋２秒で示される同期点から、時刻time_code＋３秒で示される同期点までの間に表示されるＰ５およびＢ６、時刻time_code＋３秒で示される同期点から、時刻time_code＋４秒で示される同期点までの間に表示されるＢ７、並びに時刻time_code＋４秒で示される同期点から、時刻time_code＋５秒で示される同期点までの間に表示されるＢ８で構成される場合には、GOVの最初の同期点から、時刻time_code＋１秒で示される同期点までの間に表示されるＩ１およびＢ２のmodulo_time_baseは、０Ｂとされる。

また、時刻time_code＋１秒で示される同期点から、時刻time_code＋２秒で示される同期点までの間に表示されるＢ３およびＢ４のmodulo_time_baseは、１０Ｂとされる。さらに、時刻time_code＋２秒で示される同期点から、時刻time_code＋３秒で示される同期点までの間に表示されるＰ５のmodulo_time_baseは、１１０Ｂとされる。

そして、Ｐ５は、GOVの最初の同期点から、時刻time_code＋１秒で示される同

期点に切り替わった後に、最初に表示されるP-VOPであるから、modulo_time_baseは、０Ｂにリセットされ、その後に表示されるＢ６のmodulo_time_baseは、Ｐ５の表示時刻を求めるときに参照される同期点、即ち、いまの場合、時刻time_code＋２秒で示される同期点を、GOVの最初の同期点とみなして設定される。従って、Ｂ６のmodulo_time_baseは、０Ｂとされる。

その後、時刻time_code＋３秒で示される同期点から、時刻time_code＋４秒で示される同期点までの間に表示されるＢ７のmodulo_time_baseは、１０Ｂとされ、時刻time_code＋４秒で示される同期点から、時刻time_code＋５秒で示される同期点までの間に表示されるＢ８のmodulo_time_baseは、１１０Ｂとされる。

図３２、図３３、および図３７で説明したエンコーダ側（ＶＬＣ器３６）の処理は、上述のようにして、modulo_time_baseを設定するようになっている。

また、この場合、デコーダ側（ＩＶＬＣ器１０２）では、同期点の切り替わり後に最初に表示されるI/P-VOPを検出した場合は、それに付加されているmodulo_time_baseによって示される秒数を、time_codeに累積加算して、表示時刻を求める必要があるが、即ち、例えば、図３９に示した場合においては、Ｉ１～Ｐ５の表示時刻は、time_codeに、各VOPに付加されているmodulo_time_baseに対応する秒数と、VOP_time_incrementとを加算して求めれば良いが、同期点の切り替わり後に最初に表示されるＰ５の後に表示されるＢ６～Ｂ８の表示時刻は、time_codeに、各VOPに付加されているmodulo_time_baseに対応する秒数と、VOP_time_incrementとを加算する他に、さらに、Ｐ５のmodulo_time_baseに対応する秒数である２秒を加算して求める必要があるが、図３４、図３５、および図３８で説明した処理は、そのようにして、表示時刻を求めるようになされている。

次に、以上説明したエンコーダおよびデコーダは、それ専用のハードウェアによって実現することもできるし、コンピュータに、上述したような処理を行わせるためのプログラムを実行させることによっても実現することができる。

図４０は、図３のエンコーダまたは図１５のデコーダとして機能するコンピュータの一実施の形態の構成例を示している。

ＲＯＭ（Read Only Memory）２０１は、例えば、ブートプログラムなどを記憶している。ＣＰＵ（Central Processing Unit）２０２は、例えば、ＨＤ（Hard Disk）２０６に記憶されたプログラムを、ＲＡＭ（Read Only Memory）２０３上に展開して実行することで、各種の処理を行うようになされている。ＲＡＭ２０３は、ＣＰＵ２０２が実行するプログラムや、ＣＰＵ２０２の処理上必要なデータを一時記憶するようになされている。入力部２０４は、例えば、キーボードやマウスなどでなり、必要なコマンドやデータを入力するときなどに操作される。出力部２０５は、例えば、ディスプレイなどでなり、ＣＰＵ２０２の制御にしたがったデータを表示する。ＨＤ２０６は、ＣＰＵ２０２が実行すべきプログラム、さらには、エンコード対象の画像データや、エンコード後のデータ（符号化ビットストリーム）、デコード後の画像データなどを記憶するようになされている。通信Ｉ／Ｆ（Interface）２０７は、外部との通信を制御することにより、例えば、エンコード対象の画像データを、外部から受信したり、また、エンコード後の符号化ビットストリームを外部に送信したりするようになされている。また、通信Ｉ／Ｆ２０７は、外部でエンコードされた符号化ビットストリームを受信したり、また、デコード後の画像データを、外部に送信するようにもなされている。

以上のように構成されるコンピュータのＣＰＵ２０２に、上述したような処理を行うためのプログラムを実行させることにより、このコンピュータは、図３に示したエンコーダや、図１５に示したデコーダとして機能することになる。

なお、本実施の形態では、VOP_time_increment は、VOP の表示時刻を、1ms 単位で表すものとしたが、VOP_time_increment は、その他、例えば、次のようなものとすることも可能である。即ち、１の同期点から、その次の同期点までの間を、Ｎ個に分割し、VOP の表示時刻に対応する分割点が、１の同期点から何番目の分割点かを表す値を、VOP_time_increment とすることが可能である。このように VOP_time_increment を定義した場合、Ｎ＝１０００とすると、VOP_time_increment は、VOP の表示時刻を、1ms 単位で表すものとなる。なお、この場合、デコーダでは、１の同期点から、その次の同期点までの間を幾つに分割したかという情報が必

要となるが、同期点の間の分割数は、あらかじめ定めておいても良いし、あるいは、GOV層よりも上位の階層に含めて、デコーダに提供するようにしても良い。

請求の範囲第１項に記載の画像符号化装置および請求の範囲第６項に記載の画像符号化方法によれば、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものが、複数のグループに分けて符号化される。従って、その符号化結果に対して、グループ単位でのランダムアクセスを行うことが可能となる。

請求の範囲第１１項に記載の画像復号装置および請求の範囲第１６項に記載の画像復号方法によれば、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームが復号される。従って、符号化ビットストリームに対して、グループ単位でのランダムアクセスを行い、復号することが可能となる。

請求の範囲第２１項に記載の提供媒体によれば、画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られる符号化ビットストリームが提供される。従って、符号化ビットストリームに対して、グループ単位でのランダムアクセスを行うことが可能となる。

請求の範囲第２６項に記載の画像符号化装置および請求の範囲第２９項に記載の画像符号化方法によれば、秒精度の時刻を表す秒精度時刻情報が生成され、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報が生成される。従って、秒精度時刻情報および詳細時間情報に基づいて、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を認識し、その認識結果に基づいて、ランダムアクセスを行うことが可能となる。

請求の範囲第３２項に記載の画像復号装置および請求の範囲第３５項に記載の画像復号方法によれば、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻が、秒精度時刻情報および詳細時間情報に基づいて求められる。従って、その表示時刻に基づいて、ランダムアクセスを行うことが可能となる。

請求の範囲第３８項に記載の提供媒体によれば、秒精度の時刻を表す秒精度時刻

情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成し、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、秒精度時刻情報および詳細時間情報を、対応するＩ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加することにより得られる符号化ストリームが提供される。従って、秒精度時刻情報および詳細時間情報に基づいて、Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を認識し、その認識結果に基づいて、ランダムアクセスを行うことが可能となる。

産業上の利用可能性

本発明は、動画像データを、光磁気デイスクや磁気テープなどの記録媒体に記録し、これを再生してデイスプレイなどに表示する画像情報記録再生装置や、動画像データを伝送路を介して送信側から受信側に伝送し、受信側において、受信された動画像データを表示したり、編集して記録したりするテレビ会議システム、テレビ電話システム、放送用機器、マルチメデイアデータベース検索システムなどに利用できる。

請求の範囲

1．　画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、

上記画像を受信する受信手段と、

上記画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化する符号化手段と

を備えることを特徴とする画像符号化装置。

2．　上記オブジェクトがイントラ符号化またはインター符号化される場合において、

上記符号化手段は、上記オブジェクトのシーケンスを、少なくとも１のイントラ符号化されるオブジェクトを含むグループに分ける

ことを特徴とする請求の範囲第１項に記載の画像符号化装置。

3．　上記符号化手段は、上記オブジェクトのグループに、その符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を含める

ことを特徴とする請求の範囲第１項に記載の画像符号化装置。

4．　イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

上記符号化開始秒精度絶対時刻を基準とする秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成手段と、

上記オブジェクトのグループに含まれる上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時

刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成手段と、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、上記秒精度時刻情報および詳細時間情報を、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する付加手段と

をさらに備えることを特徴とする請求の範囲第３項に記載の画像符号化装置。

５．　上記秒精度時刻情報生成手段は、所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものを生成する

ことを特徴とする請求の範囲第４項に記載の画像符号化装置。

６．　画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化方法であって、

上記画像を受信する受信ステップと、

上記画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化する符号化ステップと

を備えることを特徴とする画像符号化方法。

７．　上記オブジェクトがイントラ符号化またはインター符号化される場合において、

上記符号化ステップにおいて、上記オブジェクトのシーケンスを、少なくとも１のイントラ符号化されるオブジェクトを含むグループに分ける

ことを特徴とする請求の範囲第６項に記載の画像符号化方法。

８．　上記符号化ステップにおいて、上記オブジェクトのグループに、その符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を含める

ことを特徴とする請求の範囲第６項に記載の画像符号化方法。

９．　イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

上記符号化開始秒精度絶対時刻を基準とする秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成ステップと、

上記オブジェクトのグループに含まれる上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成ステップと、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、上記秒精度時刻情報および詳細時間情報を、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する付加ステップと

をさらに備えることを特徴とする請求の範囲第８項に記載の画像符号化方法。

１０．　上記秒精度時刻情報生成ステップにおいて、所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものを生成する

ことを特徴とする請求の範囲第９項に記載の画像符号化方法。

１１．　画像を復号する画像復号装置であって、

上記画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを受信する受信手段と、

上記符号化ビットストリームを復号する復号手段と

* C L M S *

を備えることを特徴とする画像復号装置。

１２．　上記オブジェクトがイントラ符号化またはインター符号化されている場合において、

上記オブジェクトのシーケンスが、少なくとも１のイントラ符号化されるオブジェクトを含むグループに分けられている

ことを特徴とする請求の範囲第１１項に記載の画像復号装置。

１３．　上記オブジェクトのグループには、その符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻が含められており、

上記オブジェクトのグループに含まれる上記オブジェクトの表示時刻を、上記符号化開始秒精度絶対時刻に基づいて求める表示時刻算出手段をさらに備える

ことを特徴とする請求の範囲第１１項に記載の画像復号装置。

１４．　イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

上記符号化開始秒精度絶対時刻を基準とする秒精度の時刻を表す秒精度時刻情報と、上記オブジェクトのグループに含まれる上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報とが、上記表示時刻を表す情報として、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加されており、

上記表示時刻算出手段は、上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、上記符号化開始秒精度絶対時刻、並びに上記秒精度時刻情報および詳細時間情報に基づいて求める

* C L M S *

ことを特徴とする請求の範囲第１３項に記載の画像復号装置。

１５．　所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものが用いられている

ことを特徴とする請求の範囲第１４項に記載の画像復号装置。

１６．　画像を復号する画像復号方法であって、

上記画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化することにより得られた符号化ビットストリームを受信する受信ステップと、

上記符号化ビットストリームを復号する復号ステップと

を備えることを特徴とする画像復号方法。

１７．　上記オブジェクトがイントラ符号化またはインター符号化されている場合において、

上記オブジェクトのシーケンスが、少なくとも１のイントラ符号化されるオブジェクトを含むグループに分けられている

ことを特徴とする請求の範囲第１６項に記載の画像復号方法。

１８．　上記オブジェクトのグループには、その符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻が含められており、

上記オブジェクトのグループに含まれる上記オブジェクトの表示時刻を、上記符号化開始秒精度絶対時刻に基づいて求める表示時刻算出ステップをさらに備える

ことを特徴とする請求の範囲第１６項に記載の画像復号方法。

１９．　イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符

* C L M S *

号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

上記符号化開始秒精度絶対時刻を基準とする秒精度の時刻を表す秒精度時刻情報と、上記オブジェクトのグループに含まれる上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報とが、上記表示時刻を表す情報として、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加されており、

上記表示時刻算出ステップにおいて、上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、上記符号化開始秒精度絶対時刻、並びに上記秒精度時刻情報および詳細時間情報に基づいて求める

ことを特徴とする請求の範囲第１８項に記載の画像復号方法。

２０．　所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものが用いられている

ことを特徴とする請求の範囲第１９項に記載の画像復号方法。

２１．　画像を符号化して得られる符号化ビットストリームを提供する提供媒体であって、

上記画像を受信し、

上記画像を構成するオブジェクトのシーケンスを１以上の階層に分けたものを、複数のグループに分けて符号化する

ことにより得られる上記符号化ビットストリームを提供する

ことを特徴とする提供媒体。

２２．　上記オブジェクトがイントラ符号化またはインター符号化される場合にお

* C L M S *

いて、

上記オブジェクトのシーケンスが、少なくとも１のイントラ符号化されるオブジェクトを含むグループに分けられている

ことを特徴とする請求の範囲第２１項に記載の提供媒体。

２３．　上記オブジェクトのグループには、その符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻が含められている

ことを特徴とする請求の範囲第２１項に記載の提供媒体。

２４．　イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

上記符号化開始秒精度絶対時刻を基準とする秒精度の時刻を表す秒精度時刻情報と、上記オブジェクトのグループに含まれる上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報とが、上記表示時刻を表す情報として、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加されている

ことを特徴とする請求の範囲第２３項に記載の提供媒体。

２５．　所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものが用いられている

ことを特徴とする請求の範囲第２４項に記載の提供媒体。

* C L M S *

２６．　画像を、その画像を構成するオブジェクトごとに符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、

イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成手段と、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成手段と、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、上記秒精度時刻情報および詳細時間情報を、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する付加手段と

を備えることを特徴とする画像符号化装置。

２７．　上記秒精度時刻情報生成手段は、上記オブジェクトのシーケンスを、１以上のグループに分けて符号化する場合の、そのグループのオブジェクトの符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を基準とした秒精度の時刻を、上記秒精度時刻情報として生成する

ことを特徴とする請求の範囲第２６項に記載の画像符号化装置。

２８．　上記秒精度時刻情報生成手段は、所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものを生成する

ことを特徴とする請求の範囲第２７項に記載の画像符号化装置。

２９．　画像を、その画像を構成するオブジェクトごとに符号化し、その結果得られる符号化ビットストリームを出力する画像符号化方法であって、

イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

秒精度の時刻を表す秒精度時刻情報を生成する秒精度時刻情報生成ステップと、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成する詳細時間情報生成ステップと、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、上記秒精度時刻情報および詳細時間情報を、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する付加ステップと

を備えることを特徴とする画像符号化方法。

３０．　上記秒精度時刻情報生成ステップにおいて、上記オブジェクトのシーケンスを、１以上のグループに分けて符号化する場合の、そのグループのオブジェクトの符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を基準とした秒精度の時刻を、上記秒精度時刻情報として生成する

ことを特徴とする請求の範囲第２９項に記載の画像符号化方法。

３１．　上記秒精度時刻情報生成ステップにおいて、所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものを生成する

ことを特徴とする請求の範囲第３０項に記載の画像符号化方法。

* C L M S *

３２．　画像を、その画像を構成するオブジェクトごとに符号化して得られる符号化ビットストリームを復号する画像復号装置であって、

イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとともに、

秒精度の時刻を表す秒精度時刻情報と、上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報とが、上記表示時刻を表す情報として、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加されているとき、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、上記秒精度時刻情報および詳細時間情報に基づいて求める表示時刻算出手段と、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰを、対応する表示時刻にしたがって復号する復号手段と

を備えることを特徴とする画像復号装置。

３３．　上記秒精度時刻情報は、上記オブジェクトのシーケンスが、１以上のグループに分けて符号化されている場合の、そのグループのオブジェクトの符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を基準とした秒精度の時刻を表す

ことを特徴とする請求の範囲第３２項に記載の画像復号装置。

３４．　所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒

* C L M S *

精度で表したものが用いられている

ことを特徴とする請求の範囲第３３項に記載の画像復号装置。

３５．　画像を、その画像を構成するオブジェクトごとに符号化して得られる符号化ビットストリームを復号する画像復号方法であって、

イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとともに、

秒精度の時刻を表す秒精度時刻情報と、上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報とが、上記表示時刻を表す情報として、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加されているとき、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻を、上記秒精度時刻情報および詳細時間情報に基づいて求める表示時刻算出ステップと、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰを、対応する表示時刻にしたがって復号する復号ステップと

を備えることを特徴とする画像復号方法。

３６．　上記秒精度時刻情報は、上記オブジェクトのシーケンスが、１以上のグループに分けて符号化されている場合の、そのグループのオブジェクトの符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を基準とした秒精度の時刻を表す

ことを特徴とする請求の範囲第３５項に記載の画像復号方法。

３７．　所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表

したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものが用いられている

ことを特徴とする請求の範囲第３６項に記載の画像復号方法。

３８．　画像を、その画像を構成するオブジェクトごとに符号化して得られる符号化ビットストリームを提供する提供媒体であって、

イントラ符号化されるオブジェクトをＩ－ＶＯＰ（Intra-Video Object Plane）と、イントラ符号化または前方予測符号化のうちのいずれかで符号化されるオブジェクトをＰ－ＶＯＰ（Predictive-VOP）と、イントラ符号化、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれかで符号化されるオブジェクトをＢ－ＶＯＰ（Biderectionally Predictive-VOP）と、それぞれするとき、

秒精度の時刻を表す秒精度時刻情報を生成し、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰそれぞれの表示時刻の直前の上記秒精度時刻情報から、それぞれの表示時刻までの時間を、秒精度より細かい精度で表す詳細時間情報を生成し、

上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰの表示時刻を表す情報として、上記秒精度時刻情報および詳細時間情報を、対応する上記Ｉ－ＶＯＰ，Ｐ－ＶＯＰ、またはＢ－ＶＯＰにそれぞれ付加する

ことにより得られる上記符号化ストリームを提供する

ことを特徴とする提供媒体。

３９．　上記オブジェクトのシーケンスを、１以上のグループに分けて符号化する場合の、そのグループのオブジェクトの符号化を開始した、上記画像のシーケンス上の絶対時刻であって、秒精度の時刻である符号化開始秒精度絶対時刻を基準とした秒精度の時刻を、上記秒精度時刻情報として生成する

ことを特徴とする請求の範囲第３８項に記載の提供媒体。

４０．　所定のオブジェクトについての秒精度時刻情報として、上記符号化開始秒精度絶対時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表

* C L M S *

したもの、または上記所定のオブジェクトの直前に表示されるＩ－ＶＯＰもしくはＰ－ＶＯＰの表示時刻から、上記所定のオブジェクトの表示時刻までの時間を、秒精度で表したものを生成する

ことを特徴とする請求の範囲第３９項に記載の提供媒体。