# 明　細　書

## 動画像予測方式

### 技術分野

この発明は、例えば、
テレビ電話・テレビ会議などの画像通信用途に用いられる携帯および据置型画像通信機器などで用いられる動画像符号化／復号装置、
デジタルVTR、ビデオサーバーなどの画像蓄積・記録装置などで用いられる動画像符号化／復号装置、
単独ソフトウエアもしくはDSP (Digital Signal Processor) のファームウエアの形で実装される動画像符号化／復号プログラム
などに用いられる動画像の予測に関するものである。

### 背景技術

従来の動画像の符号化／復号方式における予測符号化／復号の例として、ISO/IEC JTC1/SC29/WG11にて標準化作業がすすめられているMPEG-4 (Moving Picture Experts Group Phase-4)のビデオ符号化／復号参照方式(Verification Model、以下ＶＭ)があげられる。 VMはMPEG-4の標準化作業の進行に伴って方式の内容が変化しているが、ここではVM Version5.0を想定し、以下単にVMと表現する。

VMは動画像シーケンスを時間／空間的に任意の形状をとる画像オブジェクトの集合体としてとらえ、各画像オブジェクトを単位として符号化／復号を行う方式である。VMにおけるビデオデータ構造を図２９に示す。VMでは時間軸を含めた動画像オブジェクトをVideo Object(VO)と呼び、VOの各時刻の状態を表し符号化の単位となる画像データをVideo Obje

ct Plane(VOP)と呼ぶ。VOが時間的／空間的に階層性を持つ場合、特別にVOとVOPの間にVideo Object Layer(VOL)なる単位を設けてVO内の階層構造を表現するようになっている。各VOPは形状情報とテクスチャ情報とに分離される。ただし、動画像シーケンス中でVOが１つの場合、各VOPはフレームと同義となる。この場合は形状情報は存在せず、テクスチャ情報だけが符号化／復号される。

VOPは図３０に示すように、形状情報をあらわすアルファデータとテクスチャ情報をあらわすテクスチャデータからなる構造を持つ。各データは、それぞれ16x16サンプルからなるブロック（アルファブロック、マクロブロック）の集合体として定義される。アルファブロック内の各サンプルは8ビットで表現される。マクロブロックは、16x16サンプルの輝度信号に付随してそれに対応する色差信号を含む。動画像シーケンスからVOPデータを作り出す処理は本符号化装置外で行われるものとする。

図３１はVM符号化／方式によるVOP符号化装置の構成を示したものである。同図において、Ｐ１は入力となる原VOPデータ、Ｐ２はVOPの形状情報をあらわすアルファブロック、Ｐ３ａは入力された原VOPデータの形状情報の有無を伝えるためのスイッチ、Ｐ４はアルファブロックを圧縮符号化する形状符号化部、Ｐ５は圧縮アルファブロックデータ、Ｐ６は局所復号アルファブロック、Ｐ７はテクスチャデータ(マクロブロック)、Ｐ８は動き検出部、Ｐ９は動きパラメータ、Ｐ１０は動き補償部、Ｐ１１は予測画像候補、Ｐ１２は予測モード選択部、Ｐ１３は予測モード、Ｐ１４は予測画像、Ｐ１５は予測誤差信号、Ｐ１６はテクスチャ符号化部、Ｐ１７はテクスチャ符号化情報、Ｐ１８は局所復号予測誤差信号、Ｐ１９は局所復号マクロブロック、Ｐ２０はスプライトメモリ更新部、Ｐ２１はVOPメモリ、Ｐ２２はスプライトメモリ、Ｐ２３は可変

長符号化・多重化部、Ｐ２４はバッファ、Ｐ２５は符号化ビットストリームである。

また、図３２にこの符号化装置の動作を要約したフローチャートを示す。

図３１の符号化装置において、原VOPデータＰ１はまずアルファブロックＰ２、マクロブロックＰ７に分離され（ステップＰＳ２、ステップＰＳ３）、アルファブロックＰ２は形状符号化部Ｐ４へ、マクロブロックＰ７は動き検出部Ｐ８へ送られる。形状符号化部Ｐ４はアルファブロックＰ２のデータ圧縮を行う処理ブロック（ステップＰＳ４）であり、この発明は形状情報の圧縮方法に関わるものではないので、その処理の詳細はここでは省略する。

形状符号化部Ｐ４の出力は圧縮アルファデータＰ５と局所復号アルファブロックＰ６で、前者は可変長符号化・多重化部Ｐ２３に送られ，後者は動き検出部Ｐ８、動き補償部Ｐ１０、予測モード選択部Ｐ１２、テクスチャ符号化部Ｐ１６にそれぞれ送られる。

動き検出部Ｐ８（ステップＰＳ５）は、マクロブロックＰ７を受け取り、VOPメモリＰ２１に蓄積されている参照画像データと局所復号アルファブロックＰ６を用いて、マクロブロックごとにローカルな動きベクトルを検出する。ここで、動きベクトルは動きパラメータの一例である。VOPメモリＰ２１には、すでに符号化されたVOPの局所復号画像が蓄積される。VOPメモリＰ２１の内容は、マクロブロックの符号化が終わるごとに逐次その局所復号画像で更新される。さらに動き検出部Ｐ８は、原VOPのテクスチャデータ全体を受け取り、スプライトメモリＰ２２に蓄積されている参照画像データと局所復号アルファデータを用いてグローバルなワーピングパラメータを検出する機能も併せ持つ。スプライトメモリＰ２２については後で詳しく述べる。

動き補償部Ｐ１０（ステップＰＳ６）は、動き検出部Ｐ８で検出した動きパラメータＰ９と局所復号アルファブロックＰ６を用いて予測画像候補Ｐ１１を生成する。次いで、予測モード選択部Ｐ１２において、予測誤差信号電力と原信号電力とを用いて当該マクロブロックの最終的な予測モードＰ１３および予測画像Ｐ１４を決定する（ステップＰＳ７）。予測モード選択部Ｐ１２ではイントラフレーム符号化／インターフレーム符号化のいずれかの符号化かという判定も行われる。

テクスチャ符号化部Ｐ１６では、予測モードＰ１３に基づいて、予測誤差信号Ｐ１５または原マクロブロックそのものをDCT（Discrete Cosine Transform）、量子化し、得られた量子化DCT係数を予測後、または直接、可変長符号化・多重化部Ｐ２３へ送り符号化する（ステップＰＳ８，ステップＰＳ９）。可変長符号化部・多重化部Ｐ２３は、あらかじめ定められたシンタックスと可変長符号化コードとに従い、受け取ったデータをビットストリームに変換して多重化する（ステップＰＳ１０）。量子化DCT係数は、逆量子化、逆DCTを経て局所復号予測誤差信号Ｐ１８にされた後、予測画像Ｐ１４と加算され局所復号マクロブロックＰ１９を得る（ステップＰＳ１１）。局所復号マクロブロックＰ１９はVOPメモリＰ２１およびスプライトメモリＰ２２に書き込まれ、以降のVOPの予測に用いられる（ステップＰＳ１２）。

以下では、予測を行う部分、特に予測方式と動き補償、スプライトメモリＰ２２およびVOPメモリＰ２１の更新制御について詳しく説明する。

（１）VMにおける予測方式

VMでは通常、図３３に示すように４種類のVOPの符号化タイプがあり、それぞれのタイプごとに○で示される予測方式をマクロブロックごとに選択できる。I-VOPでは予測をまったく行わず、すべてイントラフレ

ーム符号化する。P-VOPは過去のVOPからの予測を行うことができる。B-VOPは、過去および未来のVOPを予測に使用できる。

以上の予測はすべて動きベクトルによる予測である。一方、Sprite-VOPはスプライトメモリを用いた予測が可能である。スプライトとは、下式

$$x' = (a\,x + b\,y + c)/(g\,x + h\,y + 1)$$

$$y' = (d\,x + e\,y + f)/(g\,x + h\,y + 1)$$

に示すワーピングパラメータセット

$$\vec{\alpha} = (a,\ b,\ c,\ d,\ e,\ f,\ g,\ h)$$

をVOP単位に検出し（→はベクトルを示す。以下、同じ）、これに基づいてVOPを順次混合していくことにより生成する画像空間であり、スプライトメモリＰ２２に蓄積される。

ここで、(x, y)は原VOPの２次元座標上での画素位置、(x', y')はワーピングパラメータによって(x, y)に対応付けられるスプライトメモリ中の画素位置である。Sprite-VOPの各マクロブロックでは、このワーピングパラメータセットを統一的に用いて、スプライトメモリ中の(x', y')を決定して予測画像を生成して予測を行うことができる。厳密には、スプライトには、予測に用いる「ダイナミックスプライト」と、予測に用いるとともに、復号側で近似的にVOPを合成する目的で用いる「スタティックスプライト」の区別があるが、以下に述べる図３４～図３７では、ダイナミックスプライトを「スプライト」の意味で用いる。

動き検出部Ｐ８では、以上の予測に用いる動きベクトルおよびワーピングパラメータを検出する。動きベクトルおよびワーピングパラメータを動きパラメータＰ９という言葉で総称する。

（２）.動き補償部

動き補償部Ｐ１０は例えば図３４に示すような内部構成をとる。同図

において、Ｐ２６はワーピングパラメータ、Ｐ２７は動きベクトル、Ｐ２８はグローバル動き補償部、Ｐ２９はローカル動き補償部、Ｐ３０はワーピングパラメータによる予測画像候補、Ｐ３１は動きベクトルによる予測画像候補である。予測画像候補Ｐ１１をワーピングパラメータによる予測画像候補Ｐ３０および動きベクトルによる予測画像候補Ｐ３１を総称する言葉とする。

動き補償部Ｐ１０の動作を要約するフローチャートを図３５のステップＰＳ１４からステップＰＳ２１に示す。

動き補償部Ｐ１０では、マクロブロックＰ７ごとに動き検出部Ｐ８で検出されたVOP全体のワーピングパラメータＰ２６またはマクロブロック単位の動きベクトルＰ２７を用いて予測画像候補Ｐ１１を生成する。グローバル動き補償部Ｐ２８においてワーピングパラメータＰ２６を用いた動き補償を行い、ローカル動き補償部Ｐ２９において動きベクトルＰ２７を用いた動き補償を行う。

I-VOPでは動き補償部Ｐ１０は動作しない（ステップＰＳ１４からステップＰＳ２１へ行く）。I-VOP以外のとき、ローカル動き補償部Ｐ２９が動作し、動きベクトルＰ２７を用いてVOPメモリＰ２１内の過去のVOP局所復号画像から予測画像候補(PR1)を取り出す（ステップＰＳ１５）。P-VOPではこの予測画像候補(PR1)のみを用いることになる。

ステップＰＳ１６でB-VOPであるとされた場合は、さらにローカル動き補償部Ｐ２９において、動きベクトルＰ２７を用いてVOPメモリＰ２１内の未来のVOP局所復号画像から予測画像候補(PR2)を取り出すとともに（ステップＰＳ１７）、過去および未来のVOP局所復号画像から得た予測画像候補を加算平均して予測画像候補(PR3)を得る（ステップＰＳ１８）。

また、ダイレクト予測(ITU-T勧告 H.263符号化方式におけるBフレー

ム相当の予測方式に基づく予測。組になるP-VOPのベクトルからBフレーム用のベクトルを作る。ここでは詳細の記述は省略する)についても同様に予測画像候補(PR4)を生成する（ステップＰＳ１９）。図３４において、動きベクトルによる予測画像候補Ｐ３１は上記予測画像候補PR1からPR4の一部またはすべてを総称するものとする。

I-VOPでもB-VOPでもない場合はSprite-VOPであり、VOPメモリから動きベクトルを用いて予測画像候補(PR1)を取り出すとともに、ステップＰＳ２０でグローバル動き補償部Ｐ２８においてワーピングパラメータＰ２６を用いてスプライトメモリＰ２２から予測画像候補Ｐ３０を取り出す。

グローバル動き補償部Ｐ２８はワーピングパラメータＰ２６からスプライトメモリＰ２２中の予測画像候補が存在するアドレスを計算し、これに基づいてスプライトメモリＰ２２から予測画像候補Ｐ３０を取り出して出力する。ローカル動き補償部Ｐ２９は動きベクトルＰ２７からVOPメモリＰ２１中の予測画像が存在するアドレスを計算し、これに基づいてVOPメモリＰ２１から予測画像候補Ｐ３１を取り出して出力する。

これらの予測画像候補Ｐ１１は、予測モード選択部Ｐ１２においてテクスチャデータＰ７のイントラフレーム符号化信号を含めて評価され、最も予測誤差信号電力の小さい予測画像候補と予測モードが選択される。

（３）メモリ更新

メモリ更新制御（ステップＰＳ１２）はメモリ更新部Ｐ２０で行われ、VOPメモリＰ２１およびスプライトメモリＰ２２の更新を行う。これらのメモリ内容の更新は、マクロブロック単位に選択された予測モードＰ１３に関係なく行われる。

メモリ更新部Ｐ２０の内部構成を図３６に、メモリ更新部Ｐ２０の動

作を示すフローチャートを図３７のステップＰＳ２２からステップＰＳ２８に示す。

図３６において、Ｐ３２は外部から与えられるVOP符号化タイプ、Ｐ３３はスプライトメモリを用いた予測をするかしないかを示す、外部から与えられるスプライト予測識別フラグ、Ｐ３４はスプライトメモリを用いた予測に用いられる、外部から与えられるブレンド係数、Ｐ３５はスイッチ、Ｐ３６はスイッチ、Ｐ３７はスプライト合成部、Ｐ３８はスプライト変形処理部、Ｐ３９はVOPメモリ更新信号、Ｐ４０はスプライト更新信号である。

まず、スプライト予測識別フラグＰ３３により、当該VOまたはVOLでスプライトを使用すると指定されているかどうかがチェックされ（ステップＰＳ２２）、スプライトを用いない場合、B-VOPかチェックされ（ステップＰＳ２７）、B-VOPならばVOPメモリＰ２１の更新は全く行わない。I-VOPまたはP-VOPでは、マクロブロックごとに局所復号マクロブロックＰ１９をVOPメモリＰ２１に上書きする（ステップＰＳ２８）。

一方、ステップＰＳ２２のチェックにおいてスプライトを用いるとされた場合は、まず上記と同様のVOPメモリＰ２１の更新が行われた後（ステップＰＳ２３，ステップＰＳ２４）、以下の手順に従ってスプライトメモリＰ２２の更新が行われる。

ａ）スプライトのワーピング（ステップＰＳ２５）

まず、スプライト変形処理部Ｐ３８において、スプライトメモリ上の領域（当該VOPの時刻をｔとしたときの、スプライトメモリ上の位置を原点とするVOPと同面積の領域）

$M(\vec{R}, t-1)$

を、ワーピングパラメータ

$\vec{\alpha}$ = (a, b, c, d, e, f, g, h)

によってワーピング（変形）させる。

b）スプライトのブレンド（ステップＰＳ２６）

上記a)の結果得られたワーピング画像を用いて、スプライト合成部Ｐ３７において、下式に従って新しいスプライトメモリ領域

$$M(\vec{R}, t) = (1-\alpha) \cdot W_b [M(\vec{R}, t-1), \vec{\alpha}] + \alpha \cdot VO(\vec{r}, t)$$

を求める。ここで、

$\alpha$：ブレンド係数Ｐ３４

$W_b [M, \vec{\alpha}]$：上記（a）の結果得られるワーピング画像

$VO(\vec{r}, t)$：局所復号VOPの位置$\vec{r}$及び時刻tにおける画素値

である。

ただし、局所復号マクロブロック中でVOPに属さない領域については、

$$VO(\vec{r}, t) = 0$$

とみなす。ブレンド係数$\alpha$はVOP単位で与えられるので、局所復号VOPは、VOP領域の内容の如何に関わらず、$\alpha$に基づく比重で、一括してスプライトメモリに合成される。

以上のような従来の符号化方式における予測方式では、動画像オブジェクトの予測に際して、動きベクトルの検出しか行わないメモリとワーピングパラメータの検出しか行わないメモリとを最大１画面ずつだけしか使用できない構成になっているため、予測に用いることのできる参照画像をごく限られた方法でしか使用しておらず、予測効率を十分に向上させることができない。

また、複数の動画像オブジェクトを同時に符号化するようなシステムにおいても、これらのメモリには、予測される動画像オブジェクト自身

の履歴を示す参照画像しか含まれないので、参照画像のバリエーションが限定されるとともに、動画像オブジェクト間の相関を利用して予測を行うことができない。

さらに、メモリの更新に際しても、動画像オブジェクトの内部構造や性質、履歴等に関係なくメモリが書き換えられるため、動画像オブジェクトを予測する上で重要な知識を十分にメモリに蓄積できておらず、予測効率の向上が図れないという問題があった。

この発明は前記問題点を解決するためになされたものであり、画像データを符号化／復号する場合、複数のメモリを設けて、動画像シーケンスの内部構造、性質などを考慮して、複数のメモリに効果的に動画像シーケンスの履歴を蓄積して、予測および符号化／復号の効率を高めるとともに、複数の動画像オブジェクト間で予測が行える構造を設けてより柔軟な予測方式を提供することを目的とする。

発明の開示

本発明は、予測に際して参照する画像データを蓄積するとともに、それぞれ個別の変形手法が対応づけられている複数のメモリと、被予測画像領域の動きを表現するパラメータを入力し、該被予測画像領域の予測に用いたメモリに蓄積された画像データを用いて、上記パラメータと、該メモリに対応づけられた変形手法に基づき予測画像を生成する予測画像生成部とを備え、少なくとも符号化装置および復号装置のいずれかに用いられることを特徴とする動画像を予測する動画像予測方式である。

上記符号化装置は、上記予測画像の生成に用いた該メモリを示す予測メモリ指示情報信号を生成し、復号側でも該メモリに蓄積された画像データを用いて、該メモリに対応づけられた変更手法に基づき上記予測画像が生成できるように、上記予測メモリ指示情報信号と上記パラメータ

を復号側へ送信することを特徴とする。

上記復号装置は、上記パラメータと上記予測画像の生成に用いたメモリを示す予測メモリ指示情報を符号化側から受け取り、上記予測画像生成部は、上記パラメータと該メモリに対応づけられた変形手法に基づき、該メモリに蓄積された画像データを用いて上記予測画像を生成することを特徴とする。

また、本発明は、予測に際して参照する画像データを蓄積するとともに、それぞれ個別のパラメータ有効値域が対応づけられている複数のメモリと、被予測画像領域の動きを表現するパラメータを入力し、該パラメータの値を含む上記パラメータ有効値域が対応づけられたメモリを選択し、選択されたメモリに蓄積された画像データを用いて予測画像を生成する予測画像生成部とを備え、少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する動画像予測方式である。

また、本発明は、予測に際して参照する画像データを蓄積する複数のメモリと、被予測画像領域の動きを表現するパラメータを入力し、該パラメータに基づいて上記複数のメモリに蓄積された画像データを用いて予測画像を生成する動き補償部と、上記複数のメモリの少なくとも１つのメモリに蓄積される画像データを任意のタイミングで更新するメモリ更新部とを有する予測画像生成部とを備え、少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する動画像予測方式である。

上記動画像予測方式は、第１と第２の動画像オブジェクトを有する動画像シーケンス上の動画像を予測し、上記複数のメモリは、上記第１と第２の動画像オブジェクトに対応する第１と第２の個別の複数のメモリを備え、上記予測画像生成部は、上記第１と第２の動画像オブジェクト

に対応する第１と第２の個別の生成部を備えており、上記第１の生成部は、上記第１のオブジェクトを予測する場合に、少なくとも上記第１と第２の複数のメモリのいずれかに蓄積されている画像データを用いて上記予測画像を生成するとともに、第２の複数のメモリを第１のオブジェクトの予測に使用しているか否かを示す情報を予測画像に付加するために生成することを特徴とする。

上記予測画像生成部は、動画像の各時刻における変化に応じて、上記複数のメモリの数とサイズのいずれかを増減することにより予測画像を生成することを特徴とする。

上記予測画像生成部は、動画像の各時刻における変化に応じて、予測に用いるメモリを限定することにより予測画像を生成することを特徴とする。

上記予測画像生成部は、上記複数のメモリに蓄積された上記画像データを用いてそれぞれに生成された複数の予測画像を演算することにより予測画像を生成することを特徴とする。

上記動画像予測方式は、被予測画像領域の重要度を示す特徴量パラメータを検出する重要度検出部を備え、上記予測画像生成部は、上記特徴量パラメータに基づいて、少なくとも複数の予測方式と、複数のメモリと、複数のメモリ更新方法のいずれかについて、複数の選択肢の中から少なくとも１つを選択することにより予測画像を生成することを特徴とする。

上記動画像予測方式は、少なくとも、被予測画像領域が利用可能な符号量と、各時刻における該画像領域の変化量と、該画像領域の重要度のいずれかを表すパラメータを検出する重要度検出部を備え、上記予測画像生成部は、上記パラメータに基づいて、少なくとも複数の予測方式と、複数のメモリと、複数のメモリ更新方法のいづれかについて、複数の

選択肢の少なくとも１つを選択することにより予測画像を生成することを特徴とする。

上記動画像予測方式は、動画像を動画像オブジェクト単位で予測し、上記動画像予測方式は、少なくとも、被予測動画像オブジェクトが利用可能な符号量と、各時刻における該動画像オブジェクトの変化量と、該動画像オブジェクトの重要度のいずれかを表すパラメータを検出する重要度検出部を備え、上記予測画像生成部は、上記パラメータに基づいて、少なくとも複数の予測方式と、複数のメモリと、複数のメモリ更新方法のいづれかについて、複数の選択肢の少なくとも１つを選択することにより予測画像を生成することを特徴とする。

上記動画像予測方式は、動画像の符号化において、予測に係る情報を符号化する予測情報符号化部を備え、上記予測画像生成部は、予測に使用されるメモリの使用頻度数をカウントし、カウントした使用頻度数に基づいて上記複数のメモリのランクを決定し、上記予測情報符号化部は、予測に用いられるメモリのランクに応じて、予測に係わる情報に符号長を割り当てることにより予測に係わる情報を符号化することを特徴とする。

上記複数のメモリは、少なくとも、画像データをフレーム単位で蓄積するフレームメモリと、スプライト画像を蓄積するスプライトメモリとを有することを特徴とする。

上記スプライトメモリは、少なくとも、逐次更新を伴うダイナミックスプライトメモリと逐次更新を伴わないスタティックスプライトメモリのいずれかを有することを特徴とする。

上記複数のメモリは、少なくとも、平行移動、アフィン変換、遠近法変換の内のいずれかの変形手法が変更可能に対応づけられることを特徴とする。

また、本発明は、予測に際して参照する画像データを複数のメモリに蓄積する工程と、上記複数のメモリにそれぞれ個別の変形手法を対応づける工程と、被予測画像領域の動きを表現するパラメータを入力する工程と、上記被予測画像領域の予測に用いたメモリに蓄積された画像データを用いて、上記パラメータと該メモリに対応づけられた変形手法に基づき予測画像を生成する工程を備え、少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する方法である。

上記動画像を予測する方法は、さらに、上記予測画像生成に用いた該メモリを示す予測メモリ指示情報信号を生成する工程と、上記予測メモリ情報信号と上記パラメータを復号側へ送信する工程とを備えたことを特徴とする。

上記動画像を予測する方法は、さらに、予測画像の生成に用いたメモリを示す予測メモリ指示情報信号と被予測画像領域の動きを表現するパラメータを符号化側から受け取る工程を備え、復号に用いられることを特徴とする。

また、本発明は、予測に際して参照する画像データを複数のメモリに蓄積する工程と、上記複数のメモリにそれぞれ個別のパラメータ有効値域を対応づける工程と、被予測画像領域の動きを表現するパラメータを入力する工程と、上記パラメータの値を含む上記パラメータ有効値域が対応づけられたメモリを選択する工程と、選択されたメモリに格納された画像データを用いて予測画像を生成する工程を備え、少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する方法である。

また、本発明は、予測に際して参照する画像データを複数のメモリに蓄積する工程と、被予測画像領域の動きを表現するパラメータを入力する工程と、上記パラメータに基づいて、上記複数のメモリに蓄積された

画像データを用いて予測画像を生成する工程と、上記複数のメモリの少なくとも１つのメモリに格納される画像データを任意のタイミングで更新する工程を備え、少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する方法である。

図面の簡単な説明

図１は、この発明の実施の形態における動画像符号化装置を示す構成図である。

図２は、この発明の実施の形態における動画像符号化装置の動作を示すフローチャートである。

図３は、この発明の実施の形態における動画像符号化装置の動き補償部の構成を示す構成図である。

図４は、動き補償部の動作を示すフローチャートである。

図５は、この発明の実施の形態における動画像符号化装置のメモリ更新部の構成を示す構成図である。

図６は、メモリ更新部の動作を示すフローチャートである。

図７は、この発明の実施の形態における動画像符号化装置の動き補償部の構成を示す構成図である。

図８は、図７の動き補償部の動作を示すフローチャートである。

図９は、この発明の実施の形態における動画像符号化装置の動き補償部の構成を示す構成図である。

図１０は、図９の動き補償部の動作を示すフローチャートである。

図１１は、この発明の実施の形態における動画像符号化装置を示す構成図である。

図１２は、この発明の実施の形態における動画像符号化装置の動き補償部の構成を示す構成図である。

図１３は、図１２の動き補償部の動作を示すフローチャートである。

図１４は、この発明の実施の形態における動画像符号化装置のメモリ更新部の構成を示す構成図である。

図１５は、図１４のメモリ更新部の動作を示すフローチャートである。

図１６は、この発明の実施の形態における動画像符号化装置を示す構成図である。

図１７は、この発明の実施の形態における動画像符号化装置を示す構成図である。

図１８は、この発明の実施の形態における動画像符号化装置を示す構成図である。

図１９は、この発明の実施の形態における動画像符号化装置を示す構成図である。

図２０は、この発明の実施の形態１のビットストリーム２１を示す図である。

図２１は、この発明の実施の形態２のビットストリーム２１を示す図である。

図２２は、この発明の実施の形態３のビットストリーム２１を示す図である。

図２３は、この発明の実施の形態６のビットストリーム２１を示す図である。

図２４は、この発明の実施の形態における動画像復号装置を示す構成図である。

図２５は、この発明の実施の形態における動画像復号装置の動き補償部の構成を示す構成図である。

図２６は、動き補償部の動作を示すフローチャートである。

図２７は、内挿処理を説明する図である。

図２８は、この発明の実施の形態における動画像復号装置のメモリ更新部の動作を示すフローチャートである。

図２９は、VM符号化方式におけるビデオデータ構造を示す説明図である。

図３０は、VOPデータの構成を示す説明図である。

図３１は、VM符号化装置の構成を示す構成図である。

図３２は、図３１の符号化装置の動作を示すフローチャートである。

図３３は、VOP符号化タイプと対応する予測の種類を示す説明図である。

図３４は、図３１の符号化装置における動き補償部の構成を示す構成図である。

図３５は、図３４の動き補償部の動作を示すフローチャートである。

図３６は、図３１の符号化装置におけるメモリ更新部の構成を示す構成図である。

図３７は、図３６のメモリ更新部の動作を示すフローチャートである。

発明を実施するための最良の形態

実施の形態１.

図１は実施の形態１及び後述する実施の形態における符号化装置の構成を示すブロック図である。同図において、１は入力動画像信号、２はテクスチャデータ、３は動き検出部、４は動きパラメータ、５は動き補償部、６は予測画像候補、７は予測モード選択部、８は予測モード、９は予測画像、１０は予測誤差画像、１１はテクスチャ符号化部、１２は量子化DCT係数、１３は局所復号された予測誤差画像、１４は局所復号

画像、１５はメモリ更新部、１６はメモリａ、１７はメモリｂ、１８はメモリｃ、１９は可変長符号化・多重化部、２０は送信バッファ、２１はビットストリーム、８０はシーンチェンジ検出部、８１はタイマーである。このうち動き補償部５およびメモリ更新部１５が予測方式を実現する予測画像生成部１００である。また、メモリａ，ｂ，ｃがメモリ領域２００である。図中この実施の形態で取り上げない部分については後の実施の形態で説明する。図２は、この符号化装置の動作の流れを示すフローチャートである。

この実施の形態では複数、例えば３つのメモリ、を入力動画像の動き量・色度等の画像の特徴に基づく重要度に応じて使い分け、かつ任意の時間間隔で任意のメモリ（領域）、例えばメモリａ、の内容を更新する仕組みを設ける。また、動画像シーケンスはフレーム単位に入力されるものとする。

（１）入力信号

上述のごとく、上記符号化装置は動画像シーケンスの各時刻の画像を表すフレームを単位に入力し、フレームは被予測画像領域の一例である符号化単位となるマクロブロックに分割される（ステップＳ１）。

（２）メモリの使い分け方

メモリには、過去に復号済みの画像や、またはあらかじめ固定的に与えられた画像が蓄積されるが、この実施の形態では、フレーム中の部分領域の重要度に応じて、3つのメモリを以下のように使い分ける。

メモリａ：重要度小の画像領域（＝動きが静止または均一で、かつテクスチャが均一な背景的画像領域）を蓄積する。

メモリｂ：重要度中の画像領域（＝被写体の動きで、比較的動きの小さい画像領域）を蓄積する。

メモリｃ：重要度大の画像領域（＝被写体の動きで、動きの大きな画像

領域）を蓄積する。

メモリａに蓄積される重要度小の画像領域は、テレビ会議のシーンなどに出てくる背景画像領域と考えることができる。また、ある微少な動きを伴う被写体を含み、カメラ操作によって画面全体の均一な動きが含まれるようなシーンにおける背景領域にも相当する。これらの動きはマクロブロックという小さな単位よりもあらかじめフレーム全体の動きを求めて、それをマクロブロックの動きとして代用すると効率的である。すなわち、従来例に述べたスプライトにおけるワーピングパラメータに相当する変形パラメータを求め、フレーム全体の変形パラメータを各マクロブロックの動きパラメータとみなすことにする。動きパラメータは、単なる平行移動パラメータ（＝動きベクトルと同義）でもよいし、変形を含んだアフィン、遠近法動きパラメータなどであってもよい。ここで、動きベクトルは動きパラメータの一例である。

メモリｂに蓄積される重要度中の画像領域は例えばテレビ会議のシーンにおいて発言をせずに体だけが動いているような人物の画像領域で、画面中の注目度が低い被写体領域と考えることができる。メモリｃに蓄積される重要度大の画像領域は例えばテレビ会議のシーン中、発言者などもっとも注目度が高い被写体領域と考えることができる。

メモリｂ、ｃに蓄積される領域は被写体固有の動きを持つため、マクロブロックごとに異なる動きパラメータを持つと考えるのが自然である。この際の動きパラメータは、単なる平行移動パラメータ（＝動きベクトルと同義）でもよいし、変形を含んだアフィン、遠近法動きパラメータなどであってもよい。

（３）動き検出（ステップＳ２）

この実施の形態における動き検出部３は、従来例のような動きベクトルとワーピングパラメータという区別を取り去り、３つのメモリすべて

について任意の変形パラメータをマクロブロック単位に検出できるようにする。また、動き検出部３はメモリａを用いてフレーム全体の変形パラメータを検出するグローバル動きパラメータ検出と、メモリａからｃを用いてマクロブロックごとの変形パラメータを検出するローカル動きパラメータ検出の機能を備える。

（４）動き補償（ステップＳ３）

この実施の形態における動き補償部５の内部構成を図３に示す。同図において、２２は予測画像メモリアドレス算出部、２３は予測画像メモリアドレス、２４はメモリ読み出し部、２５は外部から与えられる参照メモリ指示信号である。ここでは、参照メモリ指示信号２５は、メモリａ，ｂ，ｃを使用することを指示しているものとする。この動き補償部５の動作を記したフローチャートを図４のステップＳ１１からステップＳ１６に示す。

まずI(Intra)-フレームならば、動き補償を行わない（ステップＳ１１）。I-フレームでなければ、動き検出部３で検出されたグローバル動きパラメータならびに各メモリに基づくローカル動きパラメータをもとに予測画像候補を生成する（ステップＳ１２からステップＳ１５）。具体的には、予測画像メモリアドレス算出部２２において、参照メモリ指示信号２５で指示されるメモリ中の予測画像候補が存在する予測画像メモリアドレス２３を動きパラメータ４に基づいて計算し、予測画像メモリアドレス２３によってメモリ読み出し部２４が対応するメモリから予測画像候補６を取り出して出力する。

この実施の形態においてはグローバル動きパラメータもローカル動きパラメータも同一の変形方式に基づくので、いずれの予測画像生成も図３の動き補償部５を共有して用いることができる。なお、グローバル動きパラメータにより予測画像候補６を生成する場合（ステップＳ１５）

は、常にメモリａが参照メモリとして用いられる。

（５）予測モードの選択（ステップＳ４）

この実施の形態における予測モードは、

（ａ）メモリａを使用するモード、

（ｂ）メモリｂを使用するモード、

（ｃ）メモリｃを使用するモード、

（ｄ）イントラフレーム符号化信号を使用するモード

があるものとする。予測モード選択部７は、例えば従来例に示したように、動き補償部５によって生成されるすべての予測画像候補６とイントラフレーム符号化信号を含めて、最も予測誤差信号の電力（振幅）の小さい予測画像候補６を選択して予測画像９として出力する。また、選択した予測画像９に対応する予測モード８を出力する。この予測モード８の情報の中には、選択した予測画像９を予測したメモリを示すメモリ選択情報も含まれている。予測モード８は可変長符号化・多重化部１９へ送られ、割り当てられた符号長で予測メモリ指示情報８００としてビットストリーム２１の中に符号化される。

（６）メモリの更新

メモリの更新はメモリ更新部１５で制御される。この実施の形態におけるメモリ更新部１５の内部構成を図５に示す。同図において、２６はメモリａの更新に用いるアクティビティ（詳細は後述する）、２７はメモリａ更新判定部、２８は参照メモリ選択部、２９、３０はスイッチ、３１はメモリａを更新する画像データ、３２はメモリｂを更新する画像データ、３３はメモリｃを更新する画像データ、３４はメモリａを更新するグローバル予測画像データである。また、メモリ更新の動作フローを図６に示す。

この実施の形態におけるメモリ更新は以下の手順を踏む。メモリ更新

部１５は局所復号画像１４を入力すると、メモリａ更新判定部２７において該局所復号画像１４が属するフレームに対するメモリａ更新の必要性を判定し（ステップＳ１７）、参照メモリ選択部２８において予測モード８に基づき予測に使用したメモリを選択し（ステップＳ１８、Ｓ１９）、選択されたメモリ内の参照画像を該局所復号画像１４のメモリａを更新する画像データ３１、メモリｂを更新する画像データ３２、メモリｃを更新する画像データ３３、メモリａを更新するグローバル予測画像データ３４のいずれかを用いて下記のルールで更新する。なお、該メモリ更新は予測単位（マクロブロック単位）でフレーム毎に行われる。

（１）メモリｂ、ｃの毎フレーム更新（ステップＳ２０、Ｓ２１）

当該画像の予測に用いたメモリｂまたはメモリｃに、局所復号画像１４を書き込む。

（２）メモリａの任意フレーム更新（ステップＳ２２、Ｓ２３）

当該画像の予測に用いたメモリａにステップＳ１７のメモリａ更新判定結果１０００に基づき、任意のフレームについてのみ、または任意の時間間隔で局所復号画像１４を書き込む。

メモリａの内容はいわば背景画像などの時不変画像領域の履歴であり、シーンチェンジや画面全体の大きな動きなどによって領域の内容がおおきく変化することでもない限り、メモリ内容の更新の必要がない。

上述のように被写体などの動領域についてはフレーム単位で逐次更新し、メモリａの内容はより長い時間間隔で更新する仕組みを備えておけば、被写体の動きによって見え隠れする背景画像などの予測に効果的である。

この実施の形態では、以上の観点からメモリａの更新を任意の時間間隔で行う仕組みを設ける。具体的には例えば、

ａ．グローバル動きパラメータの大きさに基づき、動きが大きい場合は

画面全体の内容を一斉に更新し、動きが小さい場合は内容を更新しない方法、

ｂ．フレーム間の時間間隔に限定されず、ある一定時間ごとに画面全体の内容を一斉に更新する方法、

ｃ．シーンチェンジを検出した場合、シーンチェンジ直後のフレームで画面全体の内容を一斉に更新する方法

などの任意更新判定基準が考えられる。

この実施の形態では、以上のような任意更新判定基準になるデータをメモリａの更新に用いるアクティビティ２６という言葉で総称する。まず、メモリａ更新判定部２７はアクティビティ２６を用いてメモリａの内容を更新するかどうかを判定する（ステップＳ１７）。上記任意更新判定基準ａの場合は動き検出部３で検出されたグローバル動きパラメータの値がアクティビティとなり、任意更新判定基準ｂの場合はタイマー８１からの当該フレームのタイムスタンプが、任意更新判定基準ｃの場合はシーンチェンジ検出部８０から出力されるシーンチェンジ検出を知らせるフラグがアクティビティ２６に相当する。

メモリａの内容を更新すると判断された場合は、局所復号画像１４の内容がグローバル予測画像データ３４として出力されてメモリａの内容を書きかえる（ステップＳ２３）。メモリａの内容を更新すると判断されなかった場合は、メモリａの更新を行わない。

なお、当該フレームにおけるメモリａ更新判定結果１０００は復号側で同様の任意更新を行うことができるよう、ビットストリーム２１に多重化されて復号側に伝送される。

図２０は、この実施の形態のビットストリーム２１を示す図である。

図２０においては、フレームデータが順に符号化されて転送される様子を概念的に示している。各フレームデータの先頭には、フレーム単位

の付加情報としてヘッダ情報が付け加えられている。また、ヘッダ情報には、メモリａ更新判定結果１０００が多重化されて復号側に伝送される。ヘッダ情報の後からは、そのフレームを構成するマクロブロックデータが続いている。マクロブロックデータの内部には、そのマクロブロックデータを予測するために用いられたメモリを示す予測メモリ指示情報８００が含まれている。復号装置においては、各マクロブロックデータの予測メモリ指示情報８００に基づいて、予測画像を予測するメモリを特定して予測画像を生成する。

また、図示していないが、メモリａ更新判定結果１０００とともに、或いは、その代わりに、メモリｂ更新情報、及び／又はメモリｃ更新情報を復号側に転送するようにしても構わない。

以上述べた符号化装置により、動画像シーケンスの内容に応じて複数のメモリを効率的に使い分ける仕組みが供給され、予測効率を向上させることができる。すなわち、動画像シーケンスの予測に際して、動画像シーケンスの内容や性質によって複数のメモリを使い分けて任意の変形パラメータに基づいた予測を行うので、複雑な動きまでもカバーして局所的な画像の性質に追随した効率的な動画像予測が行え、予測効率が向上し、符号化画像の品質を保ちつつ符号化データの発生量を低減させる符号化装置を構成することができる。また、この発明の予測方式に基づいて符号化されたビットストリームを復号する復号装置においても同様の予測方式により構成することができる。

この実施の形態ではフレームごとに符号化する装置について述べたが、任意の形状を持つ動画像オブジェクト（VOP）を符号化する装置についても同様の効果が期待できる。

また、この実施の形態では被予測画像領域の一例としてマクロブロックを単位とした符号化装置について述べたが、任意の形状を持つ部分画

像などの画像単位または複数個の固定サイズブロックの組み合わせによって構成される可変形状のブロックなどの画像単位で画像を符号化する装置についても、同様の効果が期待できる。

また、この実施の形態では、メモリａを用いたグローバル動きパラメータ検出を用いたが、これを用いずにローカル動き検出だけを用いる構成でも適用可能であることは言うまでもない。グローバル動き検出を行わない場合は、予測モードとしてグローバル／ローカル予測の判別情報は伝送する必要がない。

また、この実施の形態において、あらかじめ動画像シーケンスの内容に基づいて生成した参照画像データを蓄積して符号化中に更新を行わないメモリを設け、それを予測に用いる構成もとることができる。

また、この実施の形態においては、メモリａ，ｂ，ｃに対してそれぞれ部分領域を蓄積するようにしてメモリ更新部１５によるメモリ更新は、メモリａ，ｂ，ｃのいずれか１つのメモリに対して行う場合を示したが、メモリａ，ｂ，ｃが画像の一部分、或いは、全部を共有して画像を蓄積するような場合には、メモリ更新部１５は、メモリａ，ｂ，ｃのうち２つのメモリ、或いは、全てのメモリを更新する。例えば、メモリａが１フレームの参照画像データを蓄積するフレームメモリであり、メモリｂが逐次更新を伴わないスタティックスプライトメモリであり、メモリｃが逐次更新を伴うダイナミックスプライトメモリである場合には、スタティックスプライトメモリは、予め固定的なデータを参照画像データとして蓄積しているため、メモリｂはメモリ更新部１５によっては更新されないが、メモリａとメモリｃが同一領域の参照画像データを蓄積している場合には、メモリ更新部１５は、メモリａとメモリｃを同時に更新することになる。このように、メモリａ，ｂ，ｃが参照画像データを重複して蓄積している場合には、重複した領域をそれぞれメモリ更新

部１５が更新する。

以上のことは、以下に述べる実施の形態においても同様である。

また、この実施の形態では、３つのメモリａ，ｂ，ｃを用いる場合を示したが、いずれか２つのメモリを用いる場合であっても構わない。

また、この実施の形態で述べた動き補償部５、メモリ更新部１５とまったく同一の部材を有する予測画像生成部１００を用いた復号装置を構成することもできる。復号装置に用いられる場合は、動き補償部は３つの全ての予測画像候補を生成する必要はなく復号された動きパラメータに関わる予測画像だけを生成すればよい。

実施の形態２.

次に、図１に示した構成の符号化装置において、動き補償部５の構成だけ別の構成とした符号化装置の実施形態を示し、実施の形態２の動き補償部５ａの構成と動作について述べる。

この実施の形態における動き補償部５ａの内部構成を図７に示す。同図において、３５は参照メモリ決定部である。またこの動き補償部５ａの詳細な動作を記したフローチャートを図８に示す。

まず、I-フレームならば動き補償を行わない（ステップＳ２４）。I-フレームでなければ、参照メモリ決定部３５は動きパラメータ４の値に基づいて参照メモリを決定する（ステップＳ２５）。参照メモリ決定部３５は各メモリａ，ｂ，ｃに割り当てられている有効動きパラメータ値域（詳細は後述する）を保持しており、動きパラメータ４の値と比較することによって、当該動きパラメータ４がどのメモリを指しているかを判断し、メモリａ，ｂ，ｃを識別する参照メモリ指示信号２５ａを出力する。

有効動きパラメータ値域とは、例えば動きベクトルを検出する場合、

その探索範囲が±１５画素だとすると、±０から３画素の範囲ではメモリａを、±４から８画素の範囲ではメモリｂを、±９から１５画素の範囲ではメモリｃを予測に用いるようにするというような、各メモリごとの有効探索範囲を意味する。ただし、グローバル動きパラメータで予測画像を生成する場合は参照メモリはメモリａと決まっているので、参照メモリ決定部３５を起動するのはローカル動きパラメータを用いるときのみとする。このように、動きベクトルの値で予測に用いるメモリを特定するのは、背景画像は動きが少なく、注目度が高い画像ほど動きが大きくなるはずであるという前提に基づくものである。このように、動きベクトルの値で予測に用いるメモリを特定する場合は、予測モードを符号化して伝送する必要がない。

次いで、選択された参照メモリ指示信号２５ａにしたがって予測画像候補６が生成される（ステップＳ２６からステップＳ３０）。具体的には、予測画像メモリアドレス算出部２２において、参照メモリ指示信号２５ａで指示されるメモリ中の予測画像候補６が存在する予測画像メモリアドレス２３を動きパラメータ４に基づいて計算し、予測画像メモリアドレス２３によってメモリ読み出し部２４が対応するメモリから予測画像候補６を取り出して出力する。

この実施の形態においてはグローバル動きパラメータもローカル動きパラメータも同一の変形方式に基づくので、いずれの予測画像生成も図７の動き補償部５ａを共有して用いることができる。なお、グローバル動きパラメータにより予測画像候補６を取り出す場合（ステップＳ３１）は、常にメモリａが参照メモリとして用いられる。

有効動きパラメータ値域は、動画像シーケンス毎に一定値域でもよいが、たとえば、各フレーム毎に変更してもよい。各フレーム毎に変更する場合は、当該フレームにおける各メモリの有効動きパラメータ値域は

、復号側で同様のメモリ選択を行うことができるよう、ビットストリームに多重化されて復号側に伝送される。

図２１は、この実施の形態のビットストリーム２１を示す図である。

このビットストリームの先頭には、動画像シーケンス単位に付加されたヘッダ情報がある。このヘッダ情報は、各メモリの有効動きパラメータ値域指示情報を有している。このように、動画像シーケンスの先頭に有効動きパラメータ値域指示情報を指定することにより、その動画像シーケンスに対して一定値域の有効動きパラメータ値域を用いた予測方式を復号装置で行うことができる。

フレーム毎に有効動きパラメータ値域を変更する場合は、フレーム単位に付加されたヘッダ情報の中に有効動きパラメータ指示情報を含ませればよい。

以上述べた動き補償部５ａの構成を持つ符号化装置により、フレームの局所的な動きの程度に応じて複数のメモリを効率的に使い分ける仕組みが供給され、予測効率を向上させることができる。

この実施の形態ではフレームごとに符号化する装置について述べたが、任意の形状を持つ動画像オブジェクト（VOP）を符号化する装置についても同様の効果が期待できる。

また、この実施の形態ではマクロブロックを単位とした符号化装置について述べたが、任意の形状を持つ部分画像などの画像単位または複数個の固定サイズブロックの組み合わせによって構成される可変形状のブロックなどの画像単位で画像を符号化する装置についても、同様の効果が期待できる。

また、この実施の形態では、メモリａを用いたグローバル動きパラメータ検出を用いたが、これを用いずにローカル動き検出だけを用いる構成でも適用可能であることは言うまでもない。グローバル動き検出を行

わない場合は、予測モードとしてグローバル／ローカル予測の判別情報は伝送する必要がない。

また、この実施の形態で述べた動き補償部５とまったく同一の部材有する予測画像生成部１００を用いた復号装置を構成することもできる。復号装置に用いられる場合は、動き補償部は復号された動きパラメータに関わる予測画像だけを生成すればよい。

実施の形態３．

次に、図１に示した構成の符号化装置において、動き補償部５の構成だけをさらに別の構成とした符号化装置の実施形態を示し、動き補償部５ｂの構成と動作について述べる。この実施の形態における動き検出部３ａは、動きパラメータ４ａとして平行移動量、アフィンパラメータ、遠近法パラメータをそれぞれ出力するものとする。

また、この実施の形態におけるメモリａは参照画像を１フレーム分記憶するフレームメモリであり、メモリｂはスタティックスプライトメモリであり、メモリｃはダイナミックスプライトメモリであるものとする。

この実施の形態における動き補償部５ｂの内部構成を図９に示す。同図において、３６は平行移動量（＝動きベクトル）、３７はアフィンパラメータ、３８は遠近法パラメータ、３９は平行移動量に基づく予測画像メモリアドレス算出部、４０はアフィンパラメータに基づく予測画像メモリアドレス算出部、４１は遠近法パラメータに基づく予測画像メモリアドレス算出部である。またこの動き補償部５ｂの詳細な動作を記したフローチャートを図１０に示す。

まず、I-フレームならば動き補償を行わない（ステップＳ３３）。I-フレームでない場合、予測画像メモリアドレス算出部３９から４１はそ

れぞれ動きパラメータ４ａの値に基づいて予測画像メモリアドレス２３を計算する（ステップＳ３４）。

各予測画像メモリアドレス算出部は、対応する各メモリに割り当てられている画像変形方式に基づいてアドレス計算を行う。この実施の形態においては、メモリａに平行移動、メモリｂにある程度回転や拡大・縮小などの単純な変形を伴うアフィンパラメータ、メモリｃに3次元的で複雑な動きを伴う遠近法パラメータを用いる。これらの変形方式は以下のような変換式で表せる。

〔平行移動〕

平行移動量(a, b)：

$x' = x + a$

$y' = y + b$

〔アフィン変換〕

アフィンパラメータ(a, b, c, θ)：

$x' = a(\cos\theta)x + a(\sin\theta)y + b$

$y' = a(-\sin\theta)x + a(\cos\theta)y + c$

〔遠近法変換〕

遠近法パラメータ(a, b, c, d, e, f)：

$x' = (a\ x + b\ y + c)/(g\ x + h\ y + 1)$

$y' = (d\ x + e\ y + f)/(g\ x + h\ y + 1)$

ここで、２次元座標上の(x, y)は原マクロブロックの画素位置、(x', y')は各パラメータによって(x, y)に対応付けられるメモリ中の画素位置である。すなわち、これらのパラメータをもとに（x', y'）のメモリ上の位置を計算する。この仕組みをもつことにより、各マクロブロックごとに動きの性質がもっとも適合するメモリから予測を行うことができるようになる。各動きパラメータ３６、３７、３８から計算した予

測画像メモリアドレス２３によって、メモリ読み出し部２４が対応するメモリから予測画像候補６を取り出して出力する（ステップＳ３５からステップＳ３９）。

なお、当該フレームにおける各メモリの変形方式のタイプは、復号側で同様の動き補償を行うことができるよう、動き検出手法指示情報としてビットストリーム２１に多重化されて復号側に伝送される。

図２２は、この実施の形態のビットストリーム２１を示す図である。

動画像シーケンスの先頭に付加されるヘッダ情報の中には、動き検出手法指示情報が含まれている。符号化装置では、各メモリで使用する変形方式のタイプが変更可能であり、この対応関係を示す動き検出手法指示情報を動画像シーケンスのヘッダ情報として復号装置に送る。こうして復号装置においては、各メモリで使用する変形方式のタイプを識別することができる。

復号装置においては、この識別された変形方式のタイプが動的に各メモリに対応づけられるようになっている。

以上述べた動き補償部５ｂの構成を持つ符号化装置により、フレームの局所的な動きの性質に応じて複数のメモリを効率的に使い分ける仕組みが供給され、予測効率を向上させることができる。

この実施の形態ではフレームごとに符号化する装置について述べたが、任意の形状を持つ動画像オブジェクト（VOP）を符号化する装置についても同様の効果が期待できる。

また、この実施の形態ではマクロブロックを単位とした符号化装置について述べたが、任意の形状を持つ部分画像などの画像単位または複数個の固定サイズブロックの組み合わせによって構成される可変形状のブロックなどの画像単位で画像を符号化する装置についても、同様の効果が期待できる。

また、この実施の形態では、メモリａを用いたグローバル動きパラメータ検出を用いたが、これを用いずにローカル動き検出だけを用いる構成でも適用可能であることは言うまでもない。グローバル動き検出を行わない場合は、予測モードとしてグローバル／ローカル予測の判別情報は伝送する必要がない。

また、この実施の形態では、メモリａ，ｂ，ｃを用いる場合を示したが、メモリａとｂだけを用いる場合、或いは、メモリａとｃを用いる場合、或いは、メモリｂとｃだけを用いる場合であっても構わない。

また、この実施の形態で述べた動き補償部５ｂとまったく同一の部材を有する予測画像生成部１００を用いた復号装置を構成することもできる。復号装置に用いられる場合は、動き補償部は復号された動きパラメータに関わる予測画像だけを生成すればよい。

実施の形態４．

次に、形状情報を持った複数、例えば２つ、の異なる動画像オブジェクトが混在する動画像シーケンスを対象とし、これらの動画像オブジェクトを一括して符号化する装置の実施形態を説明する。図１１にこの実施の形態における符号化装置の構成を示す。

同図において、４２は入力画像フレーム、４３はオブジェクト分離部、４４ａ、４４ｂはオブジェクトデータ、４５ａ、４５ｂは形状ブロック、４６ａ、４６ｂはスイッチ、４７ａ、４７ｂは形状符号化部、４８ａ、４８ｂは圧縮形状ブロックデータ、４９ａ、４９ｂは局所復号形状ブロック、５０ａ、５０ｂはテクスチャデータ(マクロブロック)、５１ａ、５１ｂは動き検出部、５２ａ、５２ｂは動きパラメータ、５３ａ、５３ｂは動き補償部、５４ａ、５４ｂは予測画像候補、５５ａ、５５ｂは予測モード選択部、５６ａ、５６ｂは予測モード情報、５７ａ、５７

bは予測画像、５８ａ、５８ｂは予測誤差信号、５９ａ、５９ｂはテクスチャ符号化部、６０ａ、６０ｂは圧縮テクスチャデータ、６１ａ、６１ｂは局所復号予測誤差信号、６２ａ、６２ｂは局所復号マクロブロック、６３はメモリ更新部、６４はメモリａ、６５はメモリｂ、６６はメモリｃ、６７はメモリｄ、６８はメモリｅ、６９はメモリｆ、７０ａ、７０ｂは可変長符号化部、７１は多重化部、７２はバッファ、７３はビットストリーム、９４はメモリ部、８８ａはＡオブジェクトを符号化するＡオブジェクト符号化部、８８ｂはＢオブジェクトを符号化するＢオブジェクト符号化部である。オブジェクト符号化部８８ａ、８８ｂは同一の部材による同一の内部構成を有する。

この符号化装置は画像フレーム４２を入力し、これがオブジェクト分離部４３において符号化単位のオブジェクトに分けられる。オブジェクト分離部４３の処理は任意の方法をとることができるものとする。

オブジェクトの形状情報は形状ブロック４５ａ、４５ｂの形式で形状符号化部４７ａ、４７ｂに送られて符号化され、圧縮形状ブロックデータ４８ａ、４８ｂとして可変長符号化部７０ａ、７０ｂに送られる。

動き検出部５１ａ、５１ｂはVM符号化方式と同様、局所復号形状ブロック４９ａ、４９ｂを考慮して動きパラメータの検出を行う。動きパラメータ検出はマクロブロック単位にメモリａからｆのすべてを用いて行うことができる。

ただし、原則としてＡオブジェクト符号化部８８ａにおいて符号化されるＡオブジェクトについてはメモリａからｃを、Ｂオブジェクト符号化部８８ｂにおいて符号化されるＢオブジェクトについてはメモリｄからｆをそれぞれ用いるものとする。

また、動きの種類としては、動きベクトルとワーピングパラメータという区別を取り去り、メモリ部９４のすべてのメモリについて任意の変

形パラメータをマクロブロック単位に検出できるものとする。

動き補償部５３ａ、５３ｂで各動きパラメータ５２ａ、５２ｂに基づいてすべての予測画像候補５４ａ、５４ｂを生成した後、予測モード選択部５５ａ、５５ｂで予測画像５７ａ、５７ｂおよび予測モード情報５６ａ、５６ｂを得る。原信号またはテクスチャデータ５０ａ、５０ｂと予測画像５７ａ、５７ｂとの差分をとって予測誤差信号５８ａ、５８ｂを得、これがテクスチャ符号化部５９ａ、５９ｂで符号化されて可変長符号化部７０ａ、７０ｂに送られる。また、局所復号された予測誤差信号６１ａ、６１ｂが予測画像５７ａ、５７ｂと加算されて局所復号マクロブロック６２ａ、６２ｂとなり、メモリ更新部の指示にしたがってメモリａからｆに書き込まれる。

上記Ａ／Ｂオブジェクト符号化部８８ａ、８８ｂにおいて符号化されたＡ／Ｂオブジェクトデータは多重化部７１で１つのビットストリーム７３に多重化され、バッファ７２を介して伝送される。

以下、この実施の形態における予測について主要な動作を行う動き補償部５３ａ、５３ｂを中心に説明する。

この実施の形態における動き補償部５３ａ、５３ｂは、動き検出部５１ａ、５１ｂで検出した動きパラメータ５２ａ、５２ｂに基づいて予測画像候補を生成する。動き補償部５３ａの内部構成を図１２に示す。また、Ａオブジェクト符号化部８８ａにおける動き補償部5.3ａの動作のフローチャートを図１３に示す。なお、動き補償部５３ｂでも同様に構成され同様に動作する。

図１２において、７４ａはＢオブジェクト参照判定部、７５ａはＢオブジェクト参照指示フラグである。

ここで動きパラメータ５２ａは検出に用いたメモリ情報も含むものとする。パラメータの値から実施の形態１と同様に、予測画像メモリアド

レス算出部２２ａ、メモリ読み出し部２４ａを用いて予測画像候補を生成する（ステップＳ４４からステップＳ４９）。また、Ｂオブジェクト参照判定部７４ａは動きパラメータ５２ａの中の参照するメモリ情報から、当該マクロブロックの予測にＢオブジェクト用のメモリを使用しているかどうかを判定する（ステップＳ４３）。

Ｂオブジェクト参照判定部７４ａは、判定結果をＢオブジェクト参照指示フラグ７５ａとして出力し、復号側で当該オブジェクトが単独で、自身のメモリａ，ｂ，ｃのみを用いて再生できるかどうかの判定に用いるため、ビットストリーム７３に多重化されて復号側に伝送される。符号化時には、外部から与える信号８５ａにより、常に復号側で当該オブジェクトの単独再生が可能になるように、動きパラメータの検出に際して、自身の予測用のメモリ（ａ，ｂ，ｃ，のみ）しか使用しないように制御することも可能とする。

以上述べた動き補償部５３ａ、５３ｂの構成を持つ符号化装置により、フレームの局所的な動きの性質に応じて複数のメモリを効率的に使い分ける仕組みが供給され、予測効率を向上させることができる。

また、この実施の形態ではオブジェクトをマクロブロック単位に符号化する装置について述べたが、任意の形状を持つ部分画像または複数個の固定サイズブロックの組み合わせによって構成される可変形状のブロックなどの単位で符号化する装置についても、同様の効果が期待できる。

また、この実施の形態で述べた動き補償部５３ａ、５３ｂと同一の部材を用いた復号装置を構成することもできる。復号装置に用いられる場合は、動き補償部５３は復号された動きパラメータに関わる予測画像だけを生成すればよい。また、ビットストリームから他オブジェクト参照指示フラグ７５ａ、７５ｂ相当のビットを復号して、復号中のオブジェ

クトが単独で再生できるかどうかを認識できる構成とすれば、オブジェクトデータを誤りなく復号再生できるようになる。

実施の形態５．

次に、動画像オブジェクトの時間的変化に対応して、メモリの数または容量を自在に変化可能な構成の符号化装置の実施形態を説明する。実施の形態５においては、図１に示した構成の符号化装置におけるメモリ更新部１５の構成を置きかえた符号化装置を考える。

図１４にこの実施の形態におけるメモリ更新部１５ａの内部構成を示す。同図において、７６はメモリ拡張判定部、７７はメモリ拡張指示信号、７８はメモリ縮退判定部、７９はメモリ縮退指示信号である。図１５にメモリ更新部１５ａの動作フロー（ステップＳ５１からステップＳ６３）を示す。

シーンチェンジなどによりメモリに蓄積されていた動画像シーケンスの履歴とは著しく異なる画像が現れた場合、シーンチェンジ後の予測においては、既存メモリに含まれる参照画像だけでは予測効率が劣化することがある。そのような場合、シーンチェンジ検出部８０によりシーンチェンジを検出し、シーンチェンジ直後のフレームはイントラフレーム符号化するなどして、それを新たな参照画像としてメモリに追加蓄積しておくことができれば、それ以降の予測効率を向上できる。

また、追加できるメモリの容量には物理的限界があるため、メモリ上の参照画像のうち、ほとんど予測に使われない部分を逐次縮退できる仕組みも持たせる。メモリａ，ｂ，ｃの各メモリ領域予測に使われる頻度を予測モード８に基づいてメモリ更新部１５ａで計測しておき、メモリ更新部に頻度の小さいメモリ領域を使用領域から開放する仕組みをもたせる。この実施の形態によれば、例えばソフトウエアで符号化装置を構

成するような場合、限られたRAM（ランダムアクセスメモリ）資源を有効に活用することができる。

以上の観点から、この実施の形態におけるメモリ更新部１５ａは、動画像シーケンスの時間的変化の状況に応じてメモリ領域を増やしたり、予測にあまり使われない参照画像を含むメモリ領域を縮退したりする機能を備える。

メモリａについては、実施の形態１と同様、メモリａ更新判定部２７において更新の可否が決定され（ステップＳ５０）、更新する場合は局所復号画像１４をメモリａに書き込む（ステップＳ５６、ステップＳ５７）。また、予測モード８にしたがって各メモリへ局所復号画像１４が書き込まれる（ステップＳ５１からステップＳ５５）。

これらメモリ内容更新は、メモリ拡張／縮退の判定をともなって行われる。メモリ拡張判定部７６においては、メモリａの更新に用いるアクティビティ２６をもとにメモリａ（またはメモリｂ，ｃ）の容量を増やすかを判定する（ステップＳ５８からステップＳ６０）。シーンチェンジなどにより容量を増やした方がよいと判定されれば、メモリ拡張指示信号７７によってメモリの拡張が指示される。また、メモリ縮退判定部７８においては、予測モード８をもとに予測に用いられるメモリ領域のカウントをおこなっておき、所定の回数以下しか予測に用いられないメモリ領域については、メモリ縮退指示信号７９によってメモリの縮退が指示される（ステップＳ６１からステップＳ６３）。

以上述べたメモリ更新部１５ａの構成を持つ符号化装置により、動画像シーケンスの時間的変化に追随して効率のよい予測が行えるとともに、予測に必要なメモリ領域をダイナミックに割り当てる仕組みが供給され、予測効率を向上させるとともにメモリ資源の有効活用が可能になる。

この実施の形態ではフレームごとに符号化する装置について述べたが、任意の形状を持つ動画像オブジェクト（VOP）を符号化する装置についても同様の効果が期待できる。

また、この実施の形態ではフレームをマクロブロック単位に符号化する装置について述べたが、任意の形状を持つ部分画像などの画像単位または複数個の固定サイズブロックの組み合わせによって構成される可変形状のブロック領域などの画像単位で画像を符号化する装置についても、同様の効果が期待できる。

また、この実施の形態で述べたメモリ更新部１５ａとまったく同一の部材を用いた復号装置を構成することもできる。

実施の形態６．

上記の各実施の形態においてはマクロブロックで予測に使用するメモリを変更する場合を示したが、フレームもしくは動画像オブジェクトの単位で、予測に使用するメモリを変更して予測を行う構成をとることもできる。これにより、フレームもしくは動画像オブジェクトの単位で符号化すべきメモリ関連の情報、およびマクロブロック単位で符号化すべきメモリ選択情報（予測モード８に含まれる）を符号化する必要がなくなり、効率的な符号化を行うことができる。

例えば実施の形態１における図１の符号化装置においては、マクロブロック単位で予測に使用するメモリを切替えることが可能になっているため、マクロブロック単位でどのメモリを予測に使用したかを示す付加情報を伝送する必要がある。これに対して、本実施の形態ではこの使用メモリの切替の単位をフレームまたは動画像オブジェクトに制限して予測を行うことにより、マクロブロック単位に伝送すべき付加情報を効果的に削減することを可能とする。図２３に、図２０に示した実施の形態

１の伝送ビットストリーム２１と本実施の形態の伝送ビットストリーム２１との違いを示す。図２３に示す例は、フレーム単位に予測に使用するメモリを変更する場合を示しており、予測メモリ指示情報８００がフレーム単位のヘッダ情報に含まれている。図２３に示す例は、例えば、動画像シーケンスの映像の性質が定常的に変化せず、マクロブロックレベルの局所的な変化が少ないような場合に有効である。また、以上のようにして符号化されたビットストリームを復号してフレームもしくは動画像オブジェクトを再生する復号装置を構成することもできる。

実施の形態７.

上記の各実施の形態において、複数のメモリのうち、任意の複数、例えば２つ、のメモリ（例えばメモリａとｂ）から取り出された２つの予測画像候補を加算平均した画像を予測画像候補６の１つとする構成または予測画像９として用いる構成にすることができる。また、以上のようにして符号化されたビットストリームを復号してフレームもしくは動画像オブジェクトを再生する復号装置を構成することもできる。

実施の形態８.

上記の各実施の形態に示した符号化装置において、予測を行う単位となる画像領域の空間的な複雑さ、注視度、などを規定する特徴量パラメータをあらかじめ検出しておき、これらを予測モード決定、メモリ更新の際の判定尺度として利用する構成をとることができる。

例えば、複雑な動きを含み、与えられた符号量内で許容品質を達成する符号化を行うことが困難な動画像を想定する。このような場合は、被予測画像領域（マクロブロック、任意形状画像領域、任意形状ブロックなど）ごとに重要度を見極め、重要度の低い領域はある程度品質を落し

て符号量を減少させ、その分を重要度の高い領域に割り当てて全体の品質を向上させる。本発明のように複数のメモリを任意のタイミングで切り替えて予測を行う装置では、被予測画像領域の重要度を表す特徴量パラメータを検出して、これに基づいて動的にメモリ使用方法を決定することにより、より画像の性質に適応した予測を行うことができる。例えば、図１６に示すように、領域ごとに特徴量パラメータを検出して重要度を決定する領域重要度検出部９５を設ける。領域重要度検出部９５は領域重要度を予測モード選択部７ａへ伝えるとともに領域重要度に基づく量子化パラメータをテクスチャ符号化部１１ａへ伝える。領域重要度検出部９５で重要度が高いと判定された領域は、複数用意された予測モードのうち、もっとも複雑なモードを用いて予測を行う。例えば、各メモリａ，ｂ，ｃからの参照画像を用いてそれぞれ複雑な動きモデルに基づいた動きパラメータと予測画像を求め、これら予測画像の任意の組み合せ（加算平均など）も含めて最も予測効率の高い予測モードを予測モード選択部７ａにおいて選択する。この際、予測に使用したすべてのメモリに対して参照画像を更新する。さらに、テクスチャ符号化部１１ａでは、量子化ステップ幅が細かい量子化パラメータにより符号化を行う。一方、重要度の低い領域では、簡易な予測モード（１つのメモリだけを用いた平行移動量検出）によって予測を行い、得られた予測誤差信号の振幅に関わらず少ない符号量になるように、量子化ステップの粗い量子化パラメータを用いて符号化する。以上の制御を行うことによって、重要度の低い領域はある程度品質を落しても、重要度の高い領域は高度な予測によって品質が保たれ、与えられた符号量で全体の品質を向上することができる。

実施の形態９．

複数のメモリを用いて動画像シーケンスを予測、符号化する装置において、動画像シーケンスの各時刻において使用可能な符号量、あるいは当該時刻でのシーンの変化量（シーンチェンジ検出など）、実施の形態８で述べたような被予測画像領域の特徴量や重要度などのパラメータを検出しておき、これらの値を、当該時刻における画像の予測に用いることのできる予測方式、あるいは参照メモリ領域の選択の際の判定尺度として利用する構成をとることができる。例えば、図１７のように、フレーム単位で重要度を決定するフレーム重要度検出部９６を設け、前フレームに対する当該フレームの変化量（シーンチェンジ検出部８０からのシンーンチェンジを検出）、新しいオブジェクトの出現・隠蔽の有無などを検出し、送信バッファ２０から通知される当該フレームで使用できる符号量を考慮して最終的なフレーム重要度を決定する。これに基づいて、重要度の高いフレームでは用意されたすべての予測方式と参照メモリ領域を用いて最大限予測効率を向上し、重要度の低いフレームでは予測方式や参照メモリ領域を限定して用い、符号化処理を簡素化して処理量を低減する、という装置構成が考えられる。また、シーンチェンジ時に予測を行わずにすべてイントラフレーム符号化する装置も考えられる。さらに、実施の形態８で述べた領域重要度検出部９５と併用すれば、よりきめ細かい品質制御を行うことができる。以上の制御を行うことによって、重要度の低いフレームはある程度品質を落しても、重要度の高いフレームは高度な予測によって品質が保たれ、与えられた符号量で全体の品質を向上することができる。

本実施の形態の考え方は、処理プロセスや使用可能なメモリ量が流動的なソフトウエアによる符号化の際に、利用できる資源を最大限に活用して効率よく符号化処理を行わせる場合にも適用可能である。重要度の低いフレームでは処理量を低減でき、全体の処理速度を向上させること

ができる。

実施の形態１０.

図１１に示したような、複数のメモリを用いて複数の動画像オブジェクトから構成される動画像シーケンスを予測、符号化する装置において、シーケンスとして使用できる総符号量、各動画像オブジェクトの各時刻において使用可能な符号量、あるいは当該時刻での動画像オブジェクトの変化量（オブジェクトの出現、隠蔽など）、各動画像オブジェクトのシーンの中での重要度／注視度のレベル、実施の形態８や９で述べたような被予測画像領域の特徴量や重要度などのパラメータを検出しておき、これらの値を、当該時刻における動画像オブジェクトの予測に用いることのできる予測方式、あるいは参照メモリ領域の選択の際の判定尺度として利用する構成をとることができる。

例えば、図１８のように、各オブジェクト１～ｎに対応する重要度検出部９７ａ～９７ｎを設け、オブジェクトの各時刻における変化量やオブジェクトの出現・隠蔽の有無などを表すパラメータを検出するとともに、すべてのオブジェクトの符号化データが蓄積される全体バッファ７２ｘの占有率、各オブジェクトの仮想バッファ７２ａ～７２ｎの占有率とを考慮して、各時刻においてオブジェクトの重要度を決定する。例えば、オブジェクト内に他のオブジェクトの一部が重なるなどの結果として新しいタイプの領域が出現した場合、これは以降の予測効率に大きく影響するため、ある程度オブジェクトの仮想バッファに余裕のない場合であっても重要度を高くし、符号化画像をきれいにしておく、などの制御が考えられる。重要度検出部９７ａ～９７ｎで検出された重要度はオブジェクト１～N符号化部９８ａ～９８ｎに渡され、重要度の高いオブジェクトでは用意されたすべての予測方式と参照メモリ領域を用いて最

大限予測効率を向上し、重要度の低いオブジェクトでは予測方式や参照メモリ領域を限定して用い、符号化処理を簡素化して処理量を低減する、という装置構成が考えられる。また、フレームからリアルタイムにオブジェクト分離を行って符号化する装置では、新しいオブジェクトの出現や、既存オブジェクトの隠蔽などによって当該オブジェクトの内容が著しく変化した場合に、予測を行わずにすべてイントラフレーム符号化する装置も考えられる。さらに、オブジェクト１～N符号化部９８ａ～９８ｎにおいて実施の形態８で述べた領域重要度検出部９５も併用すれば、オブジェクト内の各被予測対象領域単位によりきめ細かい品質制御を行うことができる。以上の制御を行うことによって、重要度の低いオブジェクトはある程度品質を落しても、重要度の高いオブジェクトは高度な予測によって品質が保たれ、与えられた符号量で全体の品質を向上することができる。

実施の形態１１.

また、図１９に示すように、予測に関する符号化情報（参照メモリ番号など）の符号割り当て（符号化）を行う予測情報符号化部９１を設ける場合がある。

複数のメモリａ，ｂ，ｃを用いて動画像シーケンスまたは動画像オブジェクトを予測、符号化する装置において、予測に用いられるメモリの使用頻度に応じて複数のメモリのランク付けを行い、かつこのランク付けを符号化中にダイナミックに変更することができるようにして、上記予測情報符号化部９１における予測に関する符号化情報（参照メモリ番号など）の符号割り当てを、予測に用いる複数のメモリのランクに応じて行うようにする構成をとることができる。

例えば、図１９の符号化装置において、メモリ更新部１５ｂに、メモ

リａ，ｂ，ｃそれぞれの予測に用いられる頻度をカウントするカウンタ９２を設けておき、そのカウント値にしたがってメモリａ，ｂ，ｃをランク付けし、ランク情報９０を出力する。このランク付けは、例えば1フレームもしくは動画像オブジェクトのある時刻における画像（VOP）を単位に行ってもよいし、より細かい単位である被予測画像領域（マクロブロック、任意形状領域、任意形状ブロックなど）毎に行うこともできる。

これにより、どのメモリがどれだけ頻繁に予測に使われるかを認識できる。頻繁に予測に用いられるメモリは、予測を行うにあたって重要度の高いメモリであり、参照頻度が高いほどランクを高くする。

このようにして、各被予測画像領域単位に予測に用いたメモリの頻度情報を符号化する場合、頻繁に参照するメモリ（＝ランクが高いメモリ）には短い符号長を割り当てるほうが符号化効率が高まる。

また、各被予測画像領域単位に検出した動きパラメータも、参照したメモリのランクに対応して符号長を割り当てることができれば、頻繁に発生する動きパラメータ値に対して短い符号長を割り当てることができ、効率的な予測情報の符号化が可能になる。これらの仕組みは、可変長符号化･多重化部１９の予測情報符号化部９１がメモリ更新部１５ｂのカウンタ９２から各メモリのランクを受け取るようにしておき、このランク情報９０に基づいて符号長を可変にして予測情報の符号化を行う構成で実現できる。

実施の形態１２．

図２４は、本実施の形態における、圧縮符号化されたディジタル画像を伸長再生する画像復号装置の構成を示したものである。同図において、

２１は符号化されたビットストリーム、１１９は可変長復号部、１２は量子化DCT係数、１２ａは量子化直交変換係数、１２ｂは量子化ステップ、１１１はテクスチャ復号部、１１１ａは逆量子化部、１１１ｂは逆直交変換部、１９０は復号加算部、１０１は復号画像、１９１は表示制御部、８は予測モード、１００１はメモリｂ更新情報、１００２はメモリｃ更新情報、４は動きベクトル（動きパラメータ）、８００は予測メモリ指示情報、１９５は被予測画像領域の画面内位置、１０５は動き補償部、１１６はメモリａ、１１７はメモリｂ、１１８はメモリｃ、１１５はメモリ更新部、１０６は予測画像である。動き補償部１０５とメモリ更新部１１５により予測画像生成部１００ａを構成している。また、メモリａ，ｂ，ｃによりメモリ領域２００ａを構成している。

この実施の形態では、メモリａは１フレーム分の画像データを蓄積するフレームメモリであり、メモリｂはスタティックスプライトメモリであり、メモリｃはダイナミックスプライトメモリであるものとする。また、この実施の形態の復号装置は、図２２に示したビットストリーム２１を入力するものとする。また、図２２には示していないが、メモリｂ更新情報１００１及びメモリｃ更新情報１００２がビットストリームの中で送られてくるものとする。メモリｂ更新情報１００１には、スタティックスプライトメモリを全面更新するための更新指示と全面更新するための画像データが含まれているものとする。同様に、メモリｃ更新情報１００２には、ダイナミックスプライトメモリを全面更新する更新指示と全面更新するための画像データが含まれているものとする。

以下、上記構成の装置の動作を説明する。まず、可変長復号部１１９においてビットストリーム２１が解析され、個々の符号化データに切り分けられる。量子化直交変換係数１２ａは逆量子化部１１９ａに送られ、量子化ステップ１２ｂを用いて逆量子化される。この結果が逆直交変

換部１１１ｂにおいて逆直交変換されテクスチャが復号され、復号加算部１９０に送られる。直交変換はDCT（離散コサイン変換）など、符号化装置側で用いるものと同じものを用いる。

動き補償部１０５には、可変長復号部１１９でビットストリーム２１から復号された動きベクトル４、予測メモリ指示情報８００と、被予測画像領域の画面内位置１９５を示す情報が入力される。動き補償部１０５は、これらの３種類の情報にしたがって複数のメモリａ，ｂ，ｃに格納されている参照画像から所望の予測画像を取り出す。被予測画像領域の画面内位置１９５はビットストリームに含まれている情報ではなくマクロブロックの数をカウントすることにより計算できる。予測画像生成の処理については以下の動き補償部１０５の動作説明の箇所で詳述する。

復号加算部１９０は、予測モード８の情報に基づいて、イントラフレーム符号化されたブロックならば逆直交変換部１１１ｂの出力をそのまま復号画像１０１として出力し、インターフレーム符号化されたブロックなら逆直交変換部１１１ｂの出力に予測画像１０６を加算して復号画像１０１として出力する。復号画像１０１は表示制御部１９１に送られ、表示デバイスに出力されるとともに、以降の復号処理において参照画像として用いるためにメモリａ～ｃに書き込まれる。メモリへの書き込みは予測モード８に基づいて、メモリ更新部１１５で制御される。

次に、予測画像生成部１００ａの動き補償部１０５における予測画像生成処理について説明する。本実施の形態では、画像の予測方式は予測メモリ指示情報８００に基づいて決定される。本実施の形態による復号装置では、予測画像を生成するにあたって、動きベクトル４と予測メモリ指示情報８００とに基づいて、所定の座標変換処理および内挿処理により参照画像から予測画像を生成する。座標変換処理の方法はあらかじ

め予測に使用するメモリに対応づけられているものとする。例えば、実施の形態３で述べた画像変形方式と同じ以下のような変形手法の例を考えることができる。

（１）メモリａを予測に用いる場合（＝予測メモリ指示情報８００がメモリａの使用を指示している場合）

動きベクトルによって被予測対象領域の各画素の座標を平行移動させ、メモリａ中の対応する位置の画像データを予測画像として取り出す。

（２）メモリｂを予測に用いる場合（＝予測メモリ指示情報がメモリｂの使用を指示している場合）

動きベクトルに基づいてアフィン変換式を求め、同変換式によって被予測対象領域の各画素の座標を変位させ、メモリｃ中の対応する位置の画像データを予測画像として取り出す。

（３）メモリｃを予測に用いる場合（＝使用メモリ指示情報がメモリｃの使用を指示している場合）

動きベクトルに基づいて遠近法変換式を求め、同変換式によって被予測対象領域の各画素の座標を変位させ、メモリｂ中の対応する位置の画像データを予測画像として取り出す。

動き補償部１０５の内部構成を図２５に示す。同図において、１６１はスイッチ、１６２はメモリａ用対応点決定部、１６３はメモリｂ用対応点決定部、１６４はメモリｃ用対応点決定部、１６５はメモリ読み出しアドレス生成部、１６６はスイッチ、１６７は内挿処理部である。また、図２６はその動作の様子を示すフローチャートである。

以下、図２５および図２６をもとに、本実施形態における動き補償部１０５の動作を説明する。

１）対応点の決定

まず、予測メモリ指示情報８００に基づき、スイッチ１６１によって対

応するメモリ用の対応点決定部を選択し、選択された対応点決定部に動きベクトル４を入力する。ここでは、各メモリに対応した予測画像位置の算出を行う。以下、メモリ対応で説明する。

１－１）予測メモリ指示情報８００がメモリａを指示している場合（ステップＳ１００）

動きベクトルによる平行移動によって、予測画像位置を算出する（ステップＳ１０１）。具体的には、動きベクトル(a, b)によって、被予測画像領域の位置(x, y)の画素に対する予測画像位置(x', y')を下式によって決定する。

$$x' = x + a$$

$$y' = y + b$$

決定された予測画像位置をメモリ読み出しアドレス生成部１６５に出力する。

１－２）予測メモリ指示情報８００がメモリｂを指示している場合（ステップＳ１０３）

動きベクトル４に基づいてアフィン変換式を決定する。具体的には、被予測画像領域を取り囲む矩形領域の頂点の動きベクトルを用いて、下式のアフィンパラメータ(a, b, c, θ)を決定する。

$$x' = a(\cos\theta)x + a(\sin\theta)y + b$$

$$y' = a(-\sin\theta)x + a(\cos\theta)y + c$$

これにより、被予測画像領域の位置(x, y)の画素に対する予測画像位置(x', y')を求め、メモリ読み出しアドレス生成部１６５に出力する（ステップＳ１０４）。

１－３）予測メモリ指示情報８００がメモリｃを指示している場合（ステップＳ１０６）

動きベクトルに基づいて遠近法変換式を決定する。具体的には、被予

測画像領域を取り囲む矩形領域の頂点の動きベクトルを用いて、下式の遠近法パラメータ(a, b, c, d, e, f)を決定する。

$$x' = (a x + b y + c)/(g x + h y + 1)$$

$$y' = (d x + e y + f)/(g x + h y + 1)$$

これにより、被予測画像領域の位置(x, y)の画素に対する予測画像位置(x', y')を求め、メモリ読み出しアドレス生成部に出力する（ステップＳ１０７）。

２）予測画像生成用データの読み出し

選択された対応点決定部から出力される予測画像位置(x' y')を基に、メモリ読み出しアドレス生成部１６５がメモリに蓄積されている参照画像中の予測画像生成に必要な画像データの位置を特定するメモリアドレスを生成し、予測画像生成用データを読み出す（ステップＳ１０２、Ｓ１０５、Ｓ１０８）。

３）予測画像の生成

予測画像を構成する画素のうち、整数画素位置の画素ならば、予測画像生成用データがそのまま予測画像構成画素となる。一方、実数精度の画素位置の画素の場合、内挿処理部１６７によって予測画像生成用データの内挿処理によって内挿画素値が生成される（ステップＳ１０９，Ｓ１１０，Ｓ１１１）。内挿画素値の生成は図２６による。図２６において、$(i_r, j_r)$は整数画素位置を示し、(i', j')は実数精度の画素位置を示し、wは重みを示す。

４）メモリ（参照画像）の更新

図２８に、メモリ更新部１１５の制御のフローチャートを示す。メモリ更新部１１５では、予測モード８（又は、予測メモリ指示情報８００）に基づいて予測画像取り出しの単位（例えばマクロブロック単位）で各メモリの更新を制御する。メモリａを用いて予測した場合（ステップ

Ｓ１１２）は、復号画像１０１によって逐次、メモリａおよびｃの内容を更新する（ステップＳ１１３）。メモリｂを用いて予測した場合（ステップＳ１１４）は、メモリｂはスタティックスプライトメモリなので、メモリｂの参照画像は予測画像取り出しの単位では更新しない。しかし、復号画像１０１によって逐次、メモリａおよびｃの内容を更新する（ステップＳ１１５）。一方、メモリｂ更新情報１００１による更新指示を受けた場合、メモリｂ更新情報１００１に含まれる画像データを受け取りメモリｂの内容を全面更新する（ステップＳ１１６）。また、メモリｃを用いて予測した場合（ステップＳ１１７）は、復号画像１０１によって逐次、メモリａおよびｃの内容を更新する（ステップＳ１１８）。また、メモリ更新情報により更新指示を受けた場合に、メモリｃ更新情報１００２に含まれる画像データを受け取りメモリｃの内容を更新する（ステップＳ１１９）。

この実施の形態においては、メモリａ，ｂ，ｃの３つのメモリを用いる場合を示したが、２つのメモリを用いる場合であっても構わない。例えば、メモリａとｂ、即ち、フレームメモリとスタティックスプライトメモリを用いる場合でも構わない。或いは、メモリａとｃ、即ち、フレームメモリとダイナミックスプライトメモリを用いる場合でも構わない。

以上のように、本実施の形態の復号装置によれば、画像の動きに対応して様々な種類の動きパラメータを用いて効率よく予測を行った符号化されたビットストリーム２１を復号できるとともに、符号化側で定められるタイミングで参照画像の内容を任意に更新する仕組みに対応できるので、より画像の性質に適応した復号処理を行うことが可能である。

本実施の形態では、直交変換符号化以外の別の符号化方式によって予測誤差信号を符号化したビットストリームであっても、動き補償部・メ

モリ更新部以外の、予測誤差信号復号処理のための部材を変更することで同様の効果を得ることができる。

また、本実施の形態は、固定サイズブロックを単位として復号処理を行うような通常のテレビ信号のフレームを単位とする復号装置に適用できるだけでなく、被予測対象領域を固定サイズブロックに限定せずに、任意形状画像オブジェクト（例：ISO/IEC JTC1/SC29/WG11/N1902で開示されるVideo Object Planeなど）を単位とする復号装置にも適用可能である。

産業上の利用可能性

以上説明したように、この発明によれば、参照用画像を蓄積するメモリ領域を複数個持つので、動画像シーケンスの性質に応じてメモリに振り分けて蓄積することができ、また、複数のメモリ領域のうち１つ以上のメモリ領域の内容を任意のタイミングで更新可能なので、背景画像などの時不変な画像内容は長時間内容の更新を行わず、局所的に変化する画像領域は頻繁にメモリ内容を更新するなどの制御が可能となり、動画像シーケンスの履歴を生かした効率的な予測が行える。

また、複数のメモリ領域それぞれに、各メモリ領域が有効となる変形パラメータ値域を設定し、被予測画像領域の変形パラメータの値に応じてメモリ領域を切り替えて予測に用いるので、動画像シーケンスの局所的／大域的な動きの大きさに応じて効率的な予測が行える。同時に、被予測画像領域ごとに符号化すべき動きパラメータは参照メモリ領域の有効動きパラメータ値域の範囲内で効率的に符号化することができる。

また、複数のメモリ領域それぞれに、各メモリで有効となる変形手法を設定し、被予測画像領域の変形パラメータの種類に応じてメモリを切り替えて予測できるので、動画像シーケンスの局所的／大域的な動きの

複雑さに応じて効率的な予測が行える。同時に、被予測画像領域の動きの性質にあわせて変形手法を適宜選択できることで、動きパラメータを効率的に符号化することができる。

* C L M S *

# 請 求 の 範 囲

1．予測に際して参照する画像データを蓄積するとともに、それぞれ個別の変形手法が対応づけられている複数のメモリと、

被予測画像領域の動きを表現するパラメータを入力し、該被予測画像領域の予測に用いたメモリに蓄積された画像データを用いて、上記パラメータと、該メモリに対応づけられた変形手法に基づき予測画像を生成する予測画像生成部とを備え、

少なくとも符号化装置および復号装置のいずれかに用いられることを特徴とする動画像を予測する動画像予測方式。

2．上記符号化装置は、上記予測画像の生成に用いた該メモリを示す予測メモリ指示情報信号を生成し、復号側でも該メモリに蓄積された画像データを用いて、該メモリに対応づけられた変更手法に基づき上記予測画像が生成できるように、上記予測メモリ指示情報信号と上記パラメータを復号側へ送信することを特徴とする請求項1記載の動画像予測方式。

3．上記復号装置は、上記パラメータと上記予測画像の生成に用いたメモリを示す予測メモリ指示情報を符号化側から受け取り、

上記予測画像生成部は、上記パラメータと該メモリに対応づけられた変形手法に基づき、該メモリに蓄積された画像データを用いて上記予測画像を生成することを特徴とする請求項1記載の動画像予測方式。

4．予測に際して参照する画像データを蓄積するとともに、それぞれ個別のパラメータ有効値域が対応づけられている複数のメモリと、

被予測画像領域の動きを表現するパラメータを入力し、該パラメータの値を含む上記パラメータ有効値域が対応づけられたメモリを選択し、選択されたメモリに蓄積された画像データを用いて予測画像を生成する予測画像生成部とを備え、

少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する動画像予測方式。

5．予測に際して参照する画像データを蓄積する複数のメモリと、

被予測画像領域の動きを表現するパラメータを入力し、該パラメータに基づいて上記複数のメモリに蓄積された画像データを用いて予測画像を生成する動き補償部と、上記複数のメモリの少なくとも１つのメモリに蓄積される画像データを任意のタイミングで更新するメモリ更新部とを有する予測画像生成部とを備え、

少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する動画像予測方式。

6．上記動画像予測方式は、第１と第２の動画像オブジェクトを有する動画像シーケンス上の動画像を予測し、

上記複数のメモリは、上記第１と第２の動画像オブジェクトに対応する第１と第２の個別の複数のメモリを備え、上記予測画像生成部は、上記第１と第２の動画像オブジェクトに対応する第１と第２の個別の生成部を備えており、

上記第１の生成部は、上記第１のオブジェクトを予測する場合に、少なくとも上記第１と第２の複数のメモリのいずれかに蓄積されている画像データを用いて上記予測画像を生成するとともに、第２の複数のメモリを第１のオブジェクトの予測に使用しているか否かを示す情報を予測

画像に付加するために生成することを特徴とする請求項１、４、５いずれかに記載の動画像予測方式。

７．上記予測画像生成部は、動画像の各時刻における変化に応じて、上記複数のメモリの数とサイズのいずれかを増減することにより予測画像を生成することを特徴とする請求項１、４、５いずれかに記載の動画像予測方式。

８．上記予測画像生成部は、動画像の各時刻における変化に応じて、予測に用いるメモリを限定することにより予測画像を生成することを特徴とする請求項１、４、５いずれかに記載の動画像予測方式。

９．上記予測画像生成部は、上記複数のメモリに蓄積された上記画像データを用いてそれぞれに生成された複数の予測画像を演算することにより予測画像を生成することを特徴とする請求項１記載の動画像予測方式。

１０．上記動画像予測方式は、被予測画像領域の重要度を示す特徴量パラメータを検出する重要度検出部を備え、

上記予測画像生成部は、上記特徴量パラメータに基づいて、少なくとも複数の予測方式と、複数のメモリと、複数のメモリ更新方法のいずれかについて、複数の選択肢の中から少なくとも１つを選択することにより予測画像を生成することを特徴とする請求項１記載の動画像予測方式。

１１．上記動画像予測方式は、少なくとも、被予測画像領域が利用可能

な符号量と、各時刻における該画像領域の変化量と、該画像領域の重要度のいずれかを表すパラメータを検出する重要度検出部を備え、

上記予測画像生成部は、上記パラメータに基づいて、少なくとも複数の予測方式と、複数のメモリと、複数のメモリ更新方法のいづれかについて、複数の選択肢の少なくとも１つを選択することにより予測画像を生成することを特徴とする請求項１又は５記載の動画像予測方式。

１２．上記動画像予測方式は、動画像を動画像オブジェクト単位で予測し、

上記動画像予測方式は、少なくとも、被予測動画像オブジェクトが利用可能な符号量と、各時刻における該動画像オブジェクトの変化量と、該動画像オブジェクトの重要度のいずれかを表すパラメータを検出する重要度検出部を備え、

上記予測画像生成部は、上記パラメータに基づいて、少なくとも複数の予測方式と、複数のメモリと、複数のメモリ更新方法のいづれかについて、複数の選択肢の少なくとも１つを選択することにより予測画像を生成することを特徴とする請求項１又は５記載の動画像予測方式。

１３．上記動画像予測方式は、動画像の符号化において、予測に係る情報を符号化する予測情報符号化部を備え、

上記予測画像生成部は、予測に使用されるメモリの使用頻度数をカウントし、カウントした使用頻度数に基づいて上記複数のメモリのランクを決定し、

上記予測情報符号化部は、予測に用いられるメモリのランクに応じて、予測に係わる情報に符号長を割り当てることにより予測に係わる情報を符号化することを特徴とする請求項１又は５記載の動画像予測方式。

１４．上記複数のメモリは、少なくとも、画像データをフレーム単位で蓄積するフレームメモリと、スプライト画像を蓄積するスプライトメモリとを有することを特徴とする請求項１記載の動画像予測方式。

１５．上記スプライトメモリは、少なくとも、逐次更新を伴うダイナミックスプライトメモリと逐次更新を伴わないスタティックスプライトメモリのいずれかを有することを特徴とする請求項第１４記載の動画像予測方式。

１６．上記複数のメモリは、少なくとも、平行移動、アフィン変換、遠近法変換の内のいずれかの変形手法が変更可能に対応づけられることを特徴とする請求項第１４記載の動画像予測方式。

１７．予測に際して参照する画像データを複数のメモリに蓄積する工程と、

上記複数のメモリにそれぞれ個別の変形手法を対応づける工程と、

被予測画像領域の動きを表現するパラメータを入力する工程と、

上記被予測画像領域の予測に用いたメモリに蓄積された画像データを用いて、上記パラメータと該メモリに対応づけられた変形手法に基づき予測画像を生成する工程を備え、

少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する方法。

１８．上記動画像を予測する方法は、さらに、上記予測画像生成に用いた該メモリを示す予測メモリ指示情報信号を生成する工程と、

上記予測メモリ情報信号と上記パラメータを復号側へ送信する工程とを備えたことを特徴とする請求項１７記載の動画像を予測する方法。

１９．上記動画像を予測する方法は、さらに、予測画像の生成に用いたメモリを示す予測メモリ指示情報信号と被予測画像領域の動きを表現するパラメータを符号化側から受け取る工程を備え、復号に用いられることを特徴とする請求項１７記載の動画像を予測する方法。

２０．予測に際して参照する画像データを複数のメモリに蓄積する工程と、

上記複数のメモリにそれぞれ個別のパラメータ有効値域を対応づける工程と、

被予測画像領域の動きを表現するパラメータを入力する工程と、

上記パラメータの値を含む上記パラメータ有効値域が対応づけられたメモリを選択する工程と、

選択されたメモリに格納された画像データを用いて予測画像を生成する工程を備え、少なくとも符号化および復号のいずれかに用いられることを特徴とする動画像を予測する方法。

２１．予測に際して参照する画像データを複数のメモリに蓄積する工程と、

被予測画像領域の動きを表現するパラメータを入力する工程と、

上記パラメータに基づいて、上記複数のメモリに蓄積された画像データを用いて予測画像を生成する工程と、

上記複数のメモリの少なくとも１つのメモリに格納される画像データを任意のタイミングで更新する工程を備え、少なくとも符号化および復

号のいずれかに用いられることを特徴とする動画像を予測する方法。

図1

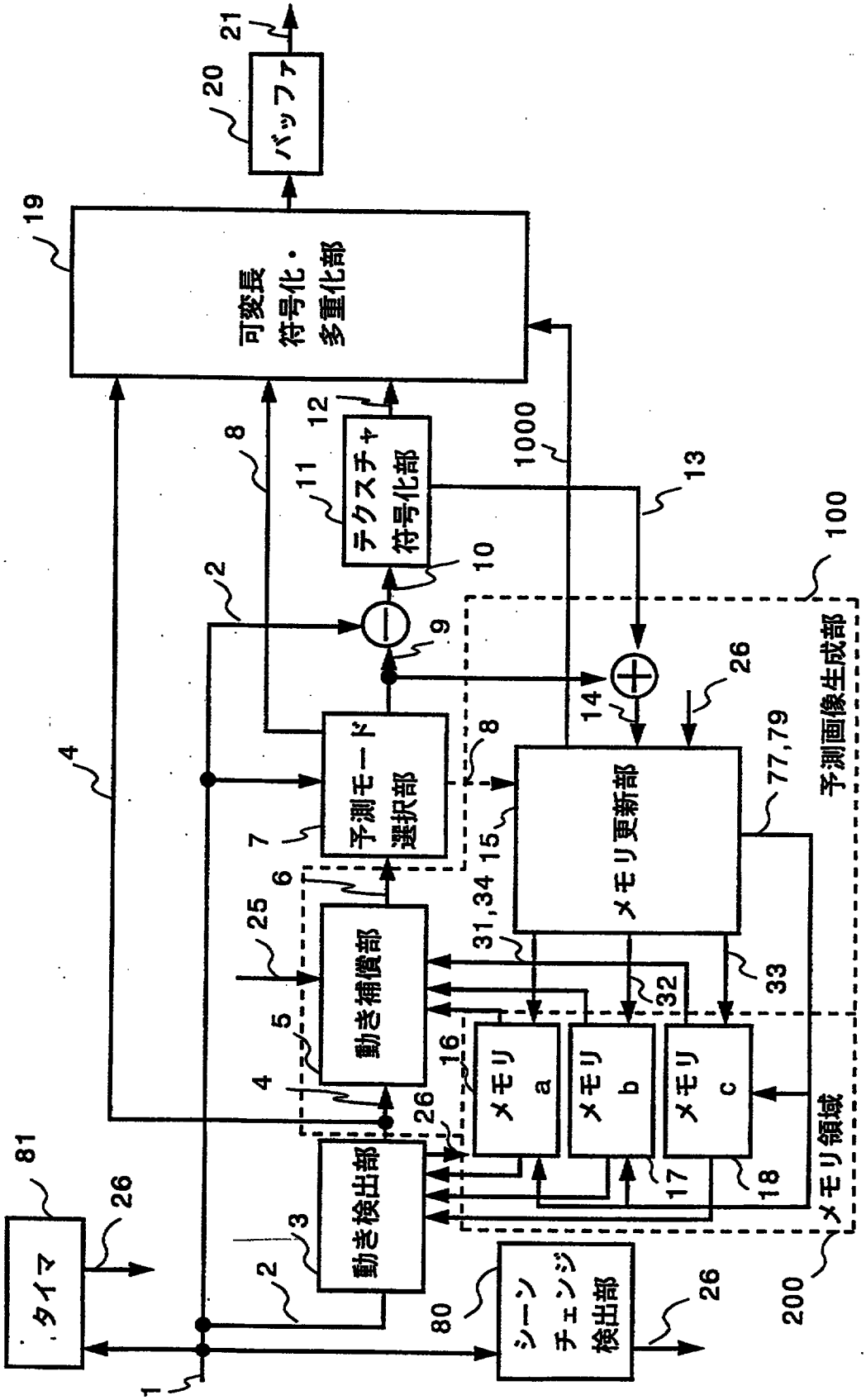

21
20
バッファ
19
可変長
符号化・
多重化部
12
11
テクスチャ
符号化部
1000
13
8
2
10
9
100
予測画像生成部
26
14
4
予測モード
選択部
8
7
77,79
メモリ更新部
15
6
25
動き補償部
31,34
33
32
5
4
16
メモリ
a
メモリ
b
メモリ
c
26
81
26
動き検出部
17
18
メモリ領域
3
2
80
シーン
チェンジ
検出部
26
200
タイマ
1

図2

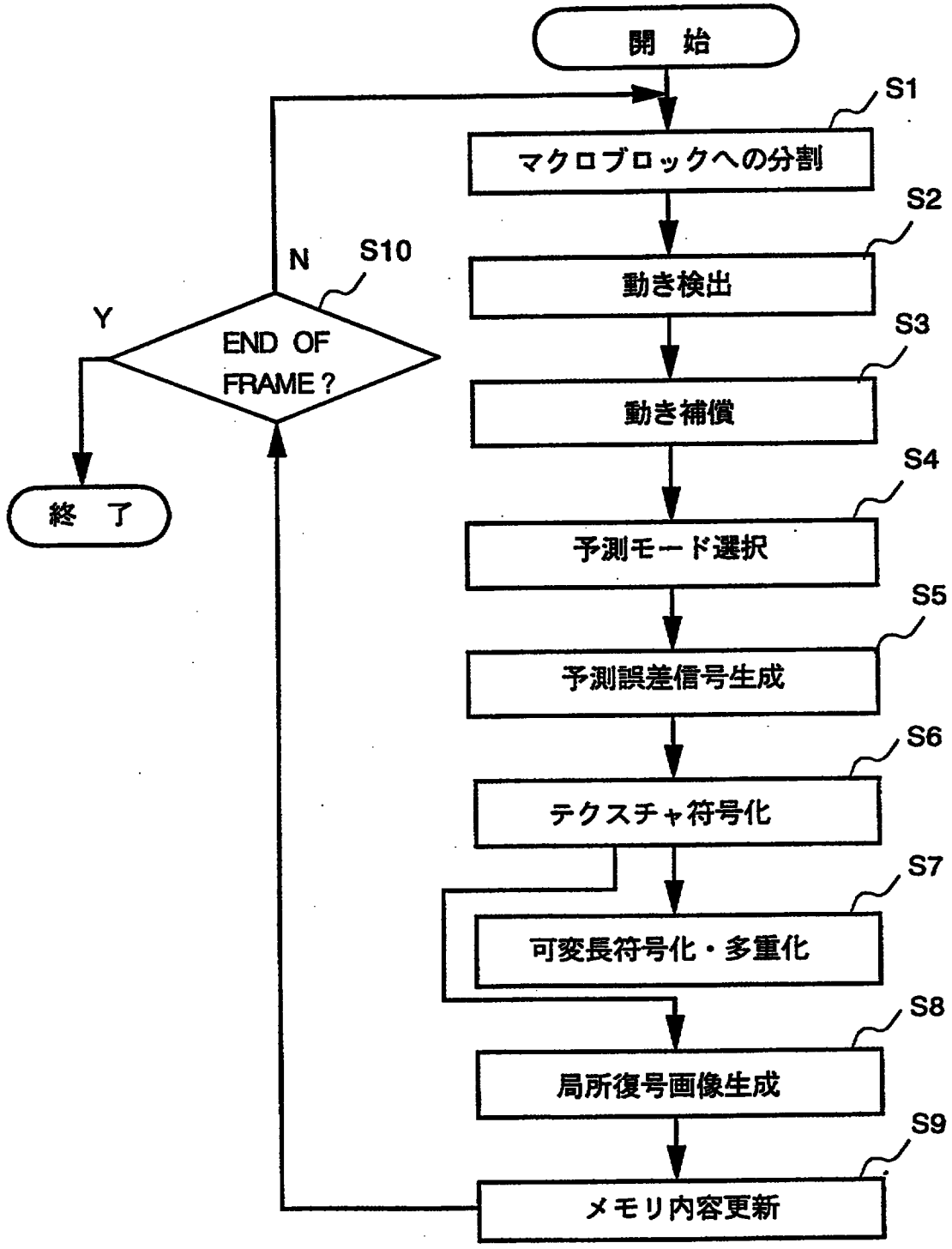

開 始
S1
マクロブロックへの分割
S2
動き検出
S3
動き補償
S4
予測モード選択
S5
予測誤差信号生成
S6
テクスチャ符号化
S7
可変長符号化・多重化
S8
局所復号画像生成
S9
メモリ内容更新
S10
END OF FRAME ?
N
Y
終 了

図3

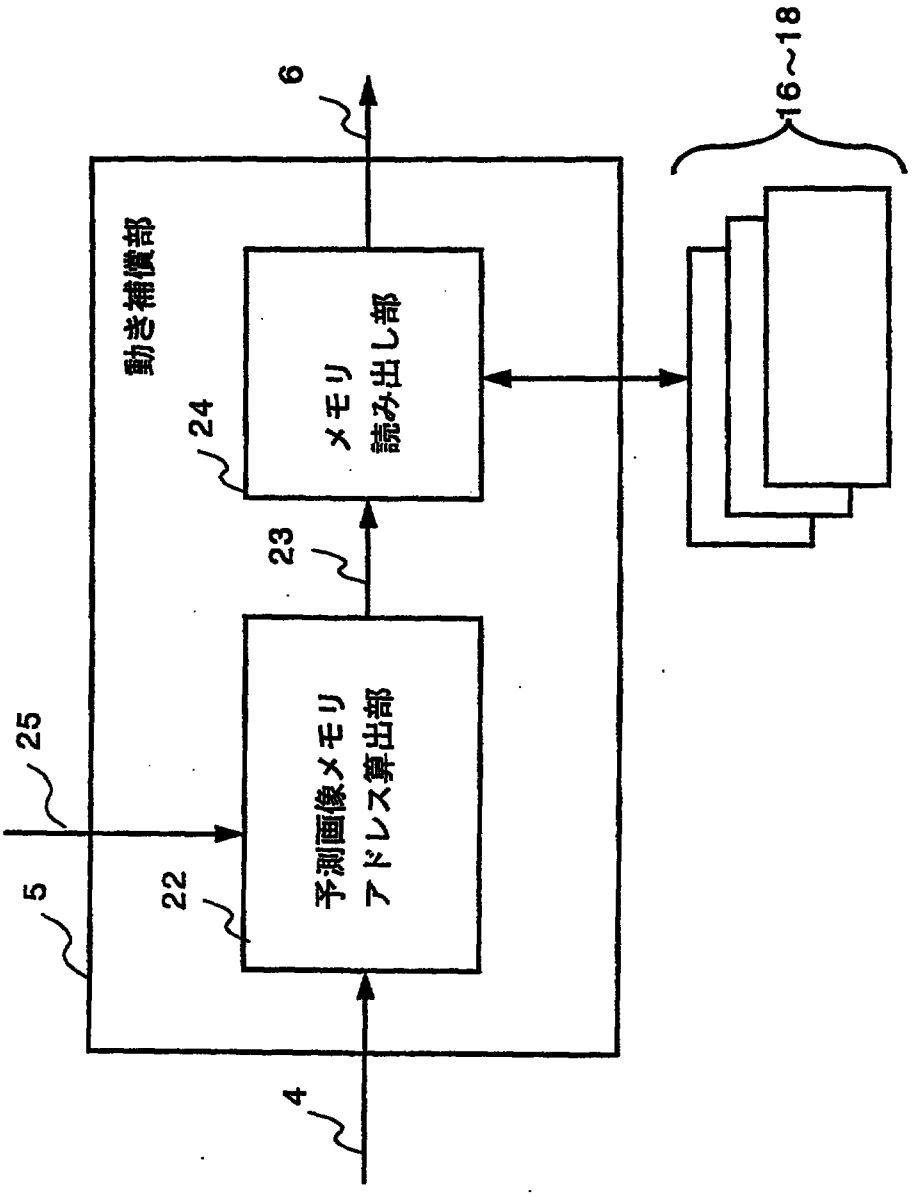


図4

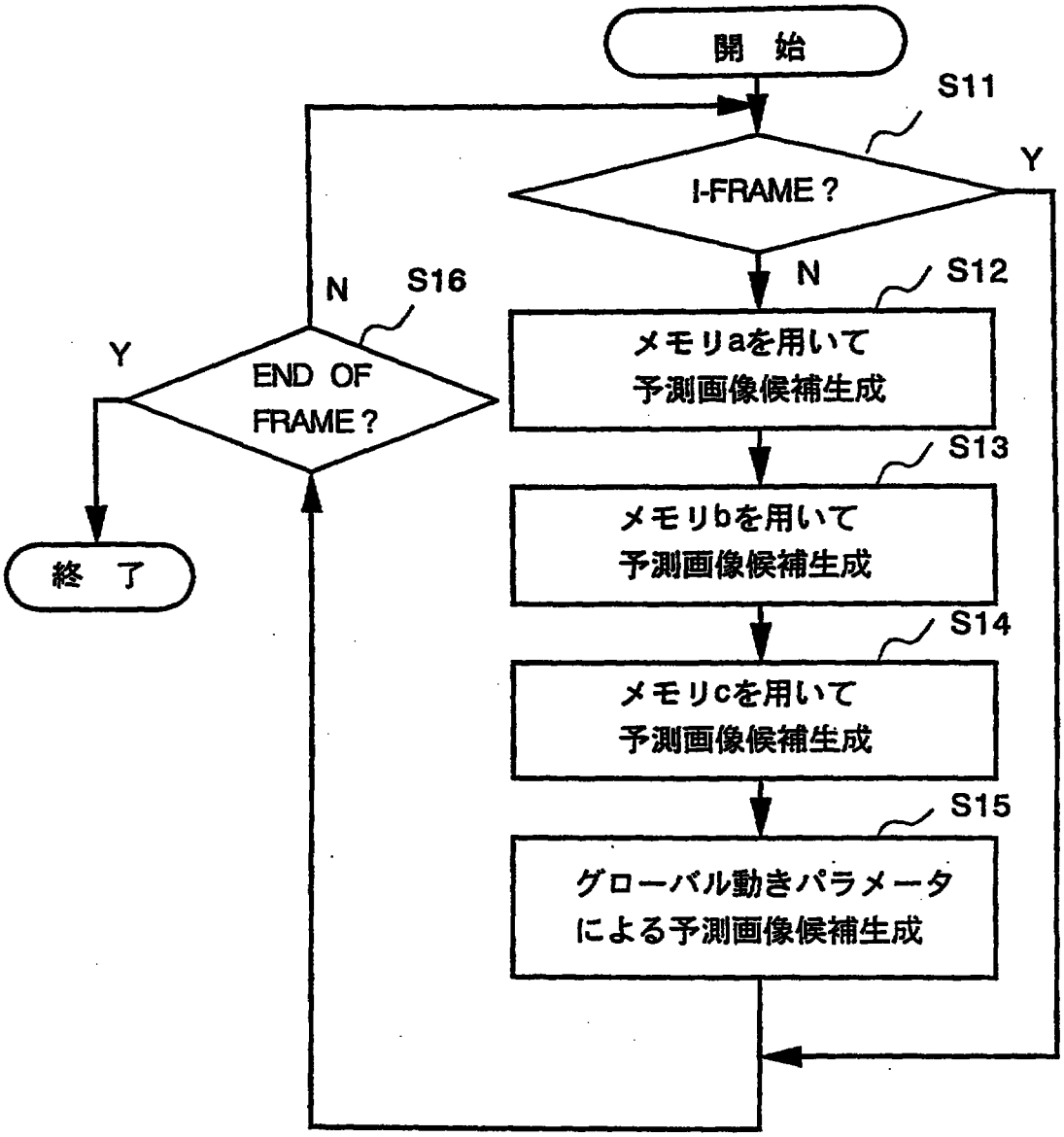

開 始
S11
I-FRAME ?
Y
N
S12
メモリaを用いて
予測画像候補生成
S13
メモリbを用いて
予測画像候補生成
S14
メモリcを用いて
予測画像候補生成
S15
グローバル動きパラメータ
による予測画像候補生成
N
S16
END OF
FRAME ?
Y
終 了

図5

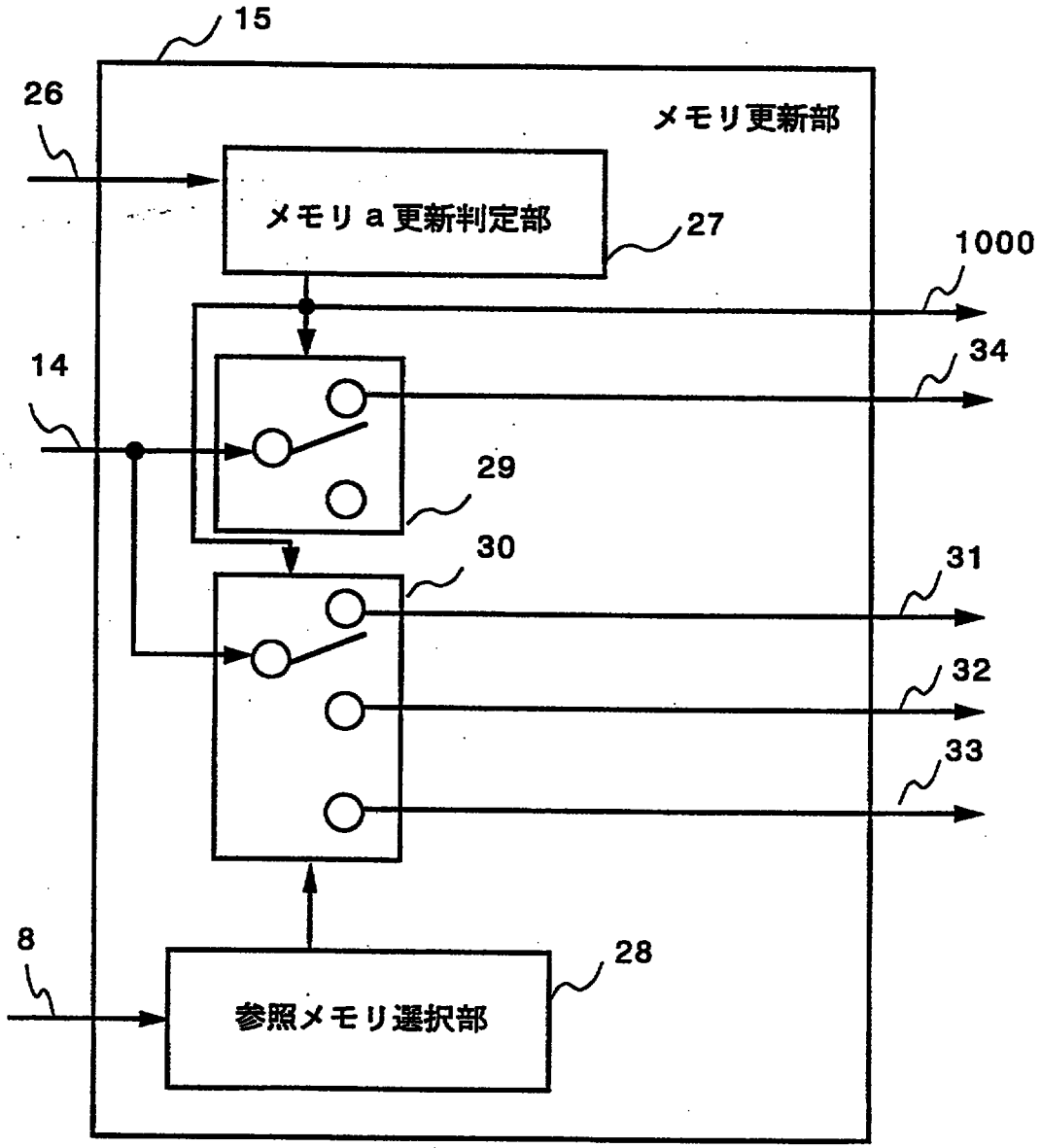

15
26
メモリ更新部
メモリａ更新判定部
27
1000
14
34
29
30
31
32
33
8
28
参照メモリ選択部

図6

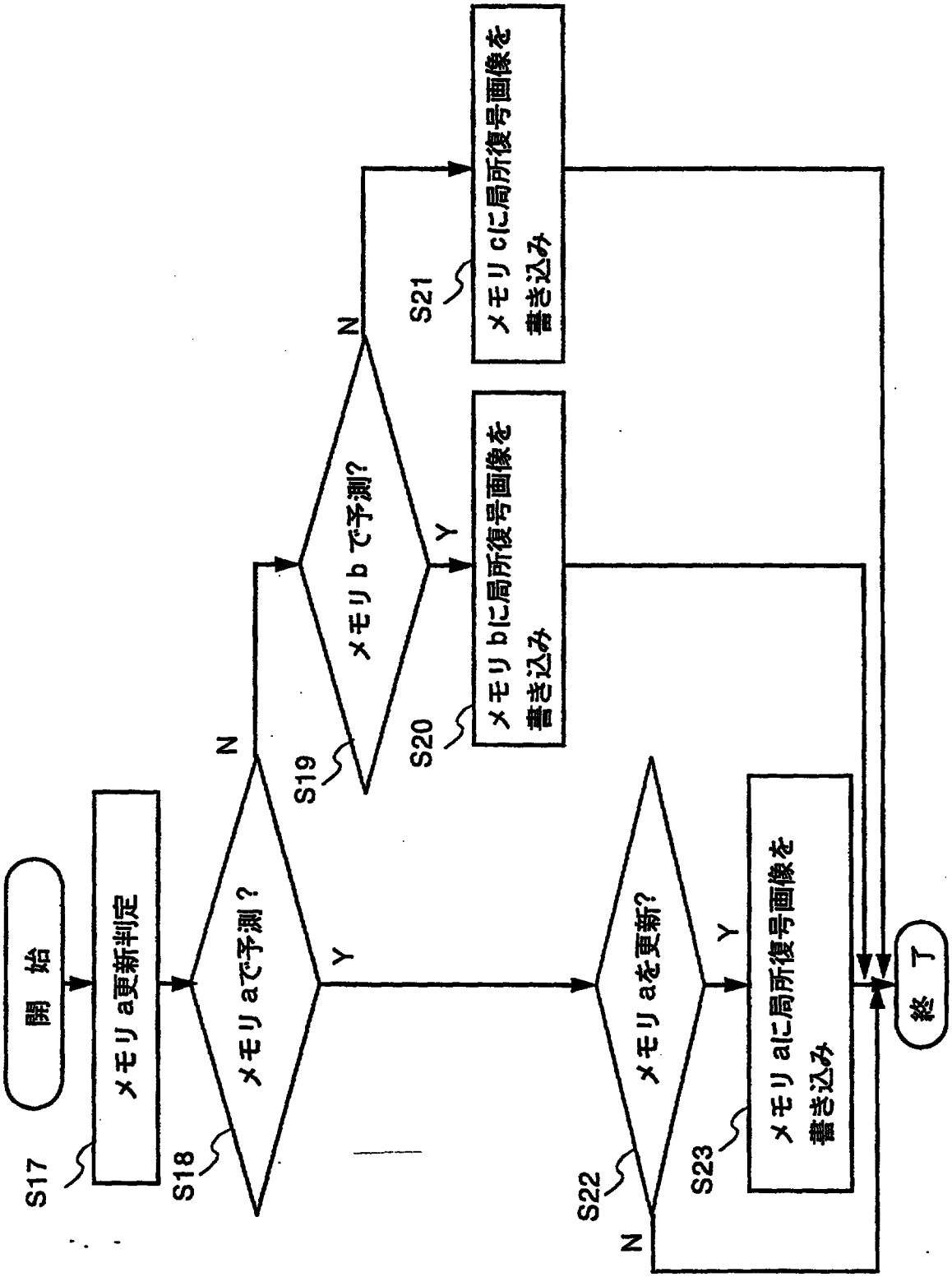

開 始
メモリ a更新判定
S17
メモリ aで予測？
S18
N
Y
メモリ b で予測?
S19
N
Y
メモリ bに局所復号画像を書き込み
S20
メモリ cに局所復号画像を書き込み
S21
メモリ aを更新?
S22
N
Y
メモリ aに局所復号画像を書き込み
S23
終 了

図7

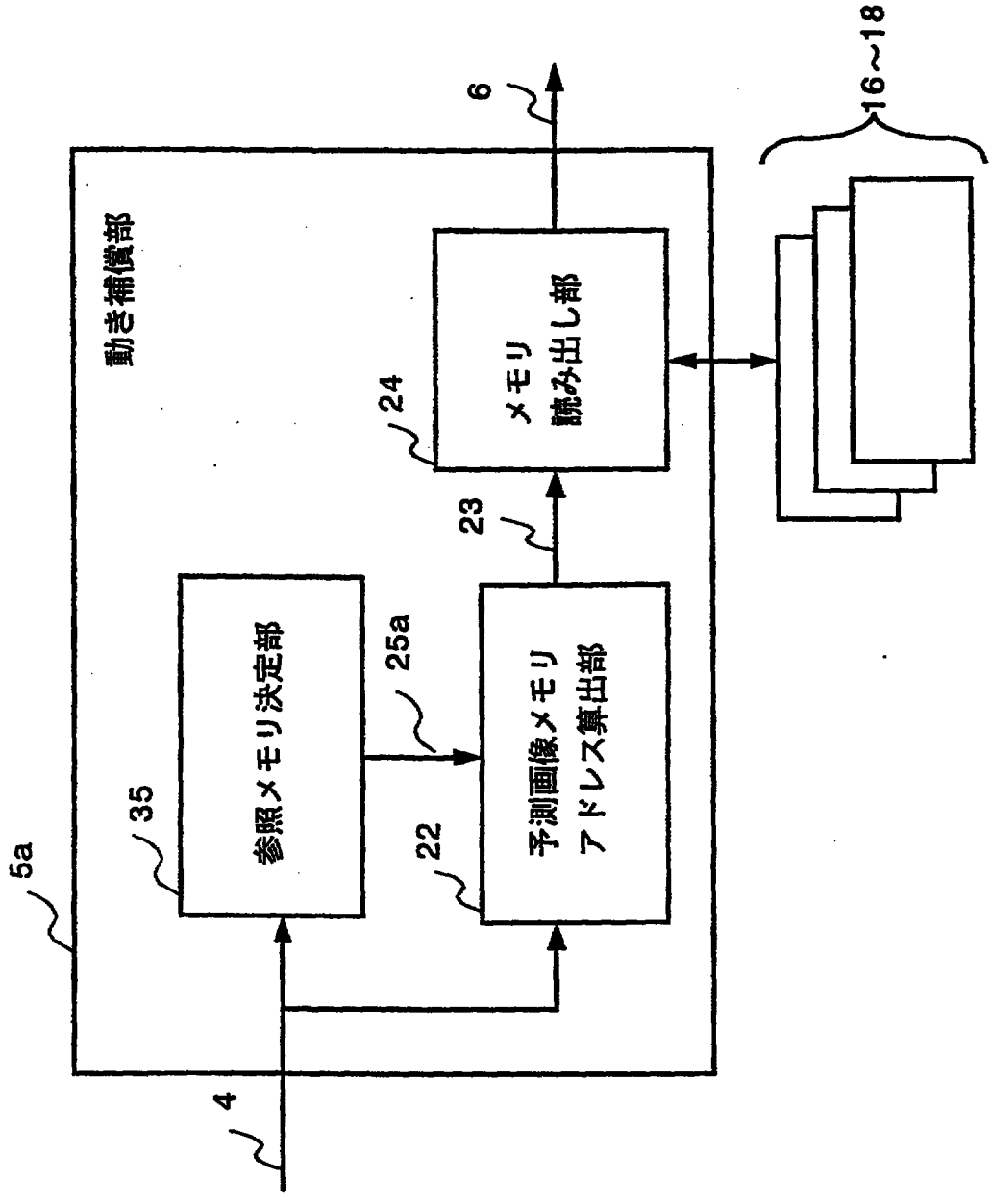

16~18
6
動き補償部
メモリ
読み出し部
24
23
参照メモリ決定部
25a
予測画像メモリ
アドレス算出部
35
22
5a
4

図8

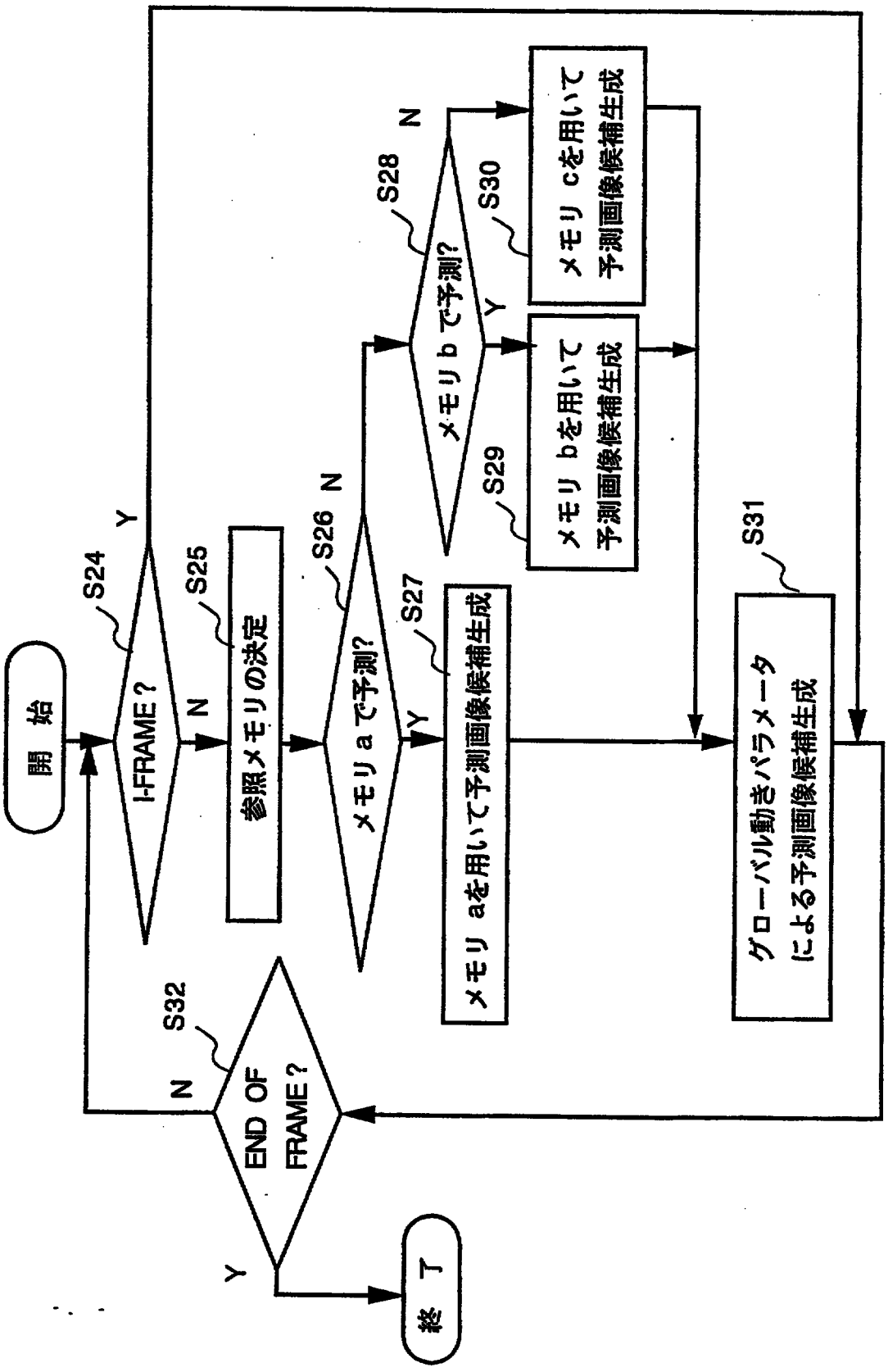

開 始
S24
I-FRAME?
Y
N
S25
参照メモリの決定
S26
メモリ a で予測?
N
Y
S27
メモリ aを用いて予測画像候補生成
S28
メモリ b で予測?
N
Y
S29
メモリ bを用いて予測画像候補生成
S30
メモリ cを用いて予測画像候補生成
S31
グローバル動きパラメータによる予測画像候補生成
S32
END OF FRAME?
N
Y
終 了

図9

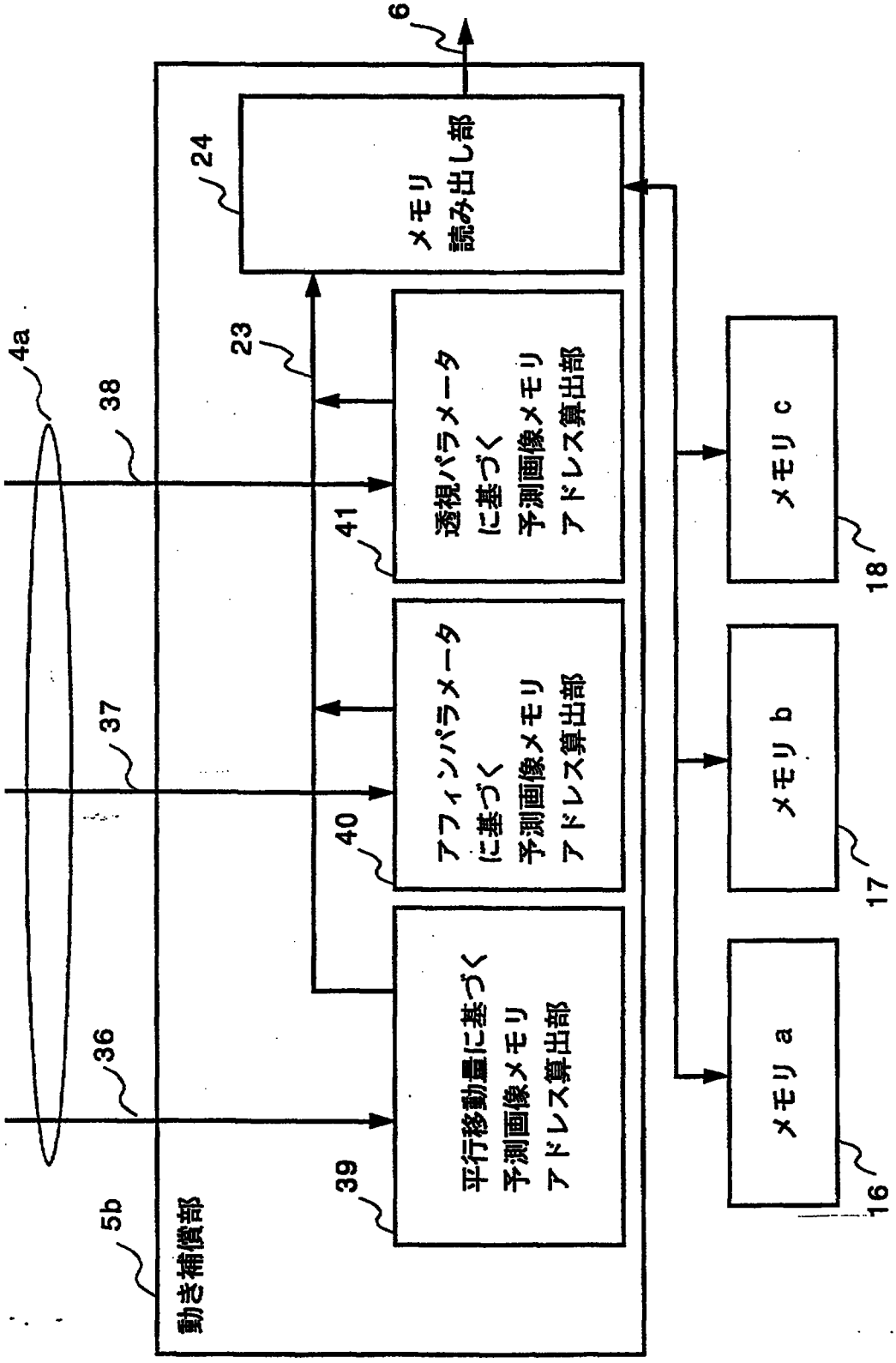

6
24
メモリ
読み出し部
23
4a
38
透視パラメータ
に基づく
予測画像メモリ
アドレス算出部
41
メモリ c
18
37
アフィンパラメータ
に基づく
予測画像メモリ
アドレス算出部
40
メモリ b
17
36
平行移動量に基づく
予測画像メモリ
アドレス算出部
39
メモリ a
16
5b
動き補償部

図10

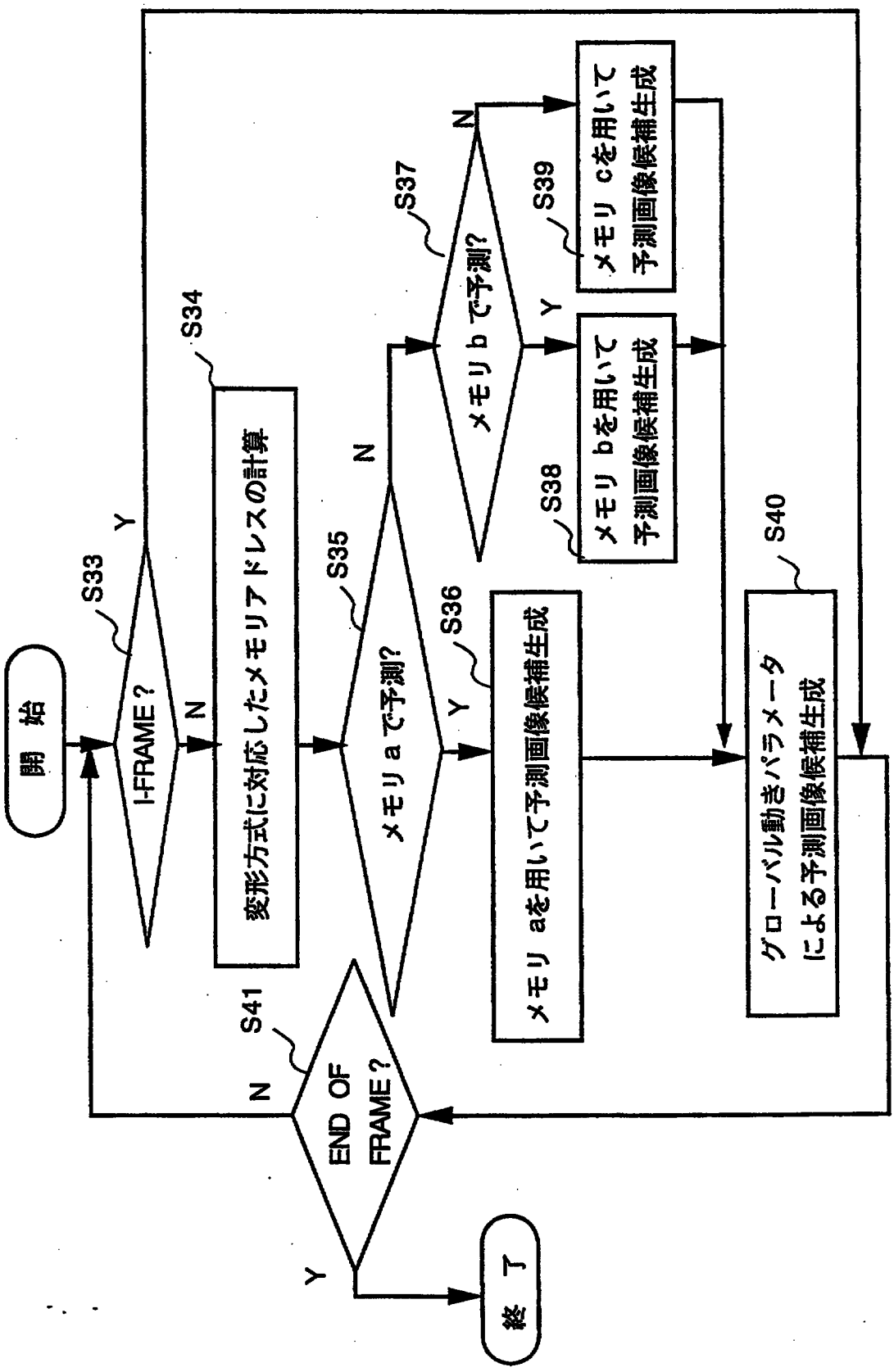

開始
S33
I-FRAME?
Y
N
S34
変形方式に対応したメモリアドレスの計算
S35
メモリaで予測?
N
Y
S36
メモリ aを用いて予測画像候補生成
S37
メモリbで予測?
N
Y
S38
メモリ bを用いて予測画像候補生成
S39
メモリ cを用いて予測画像候補生成
S40
グローバル動きパラメータによる予測画像候補生成
S41
END OF FRAME?
N
Y
終了

図11

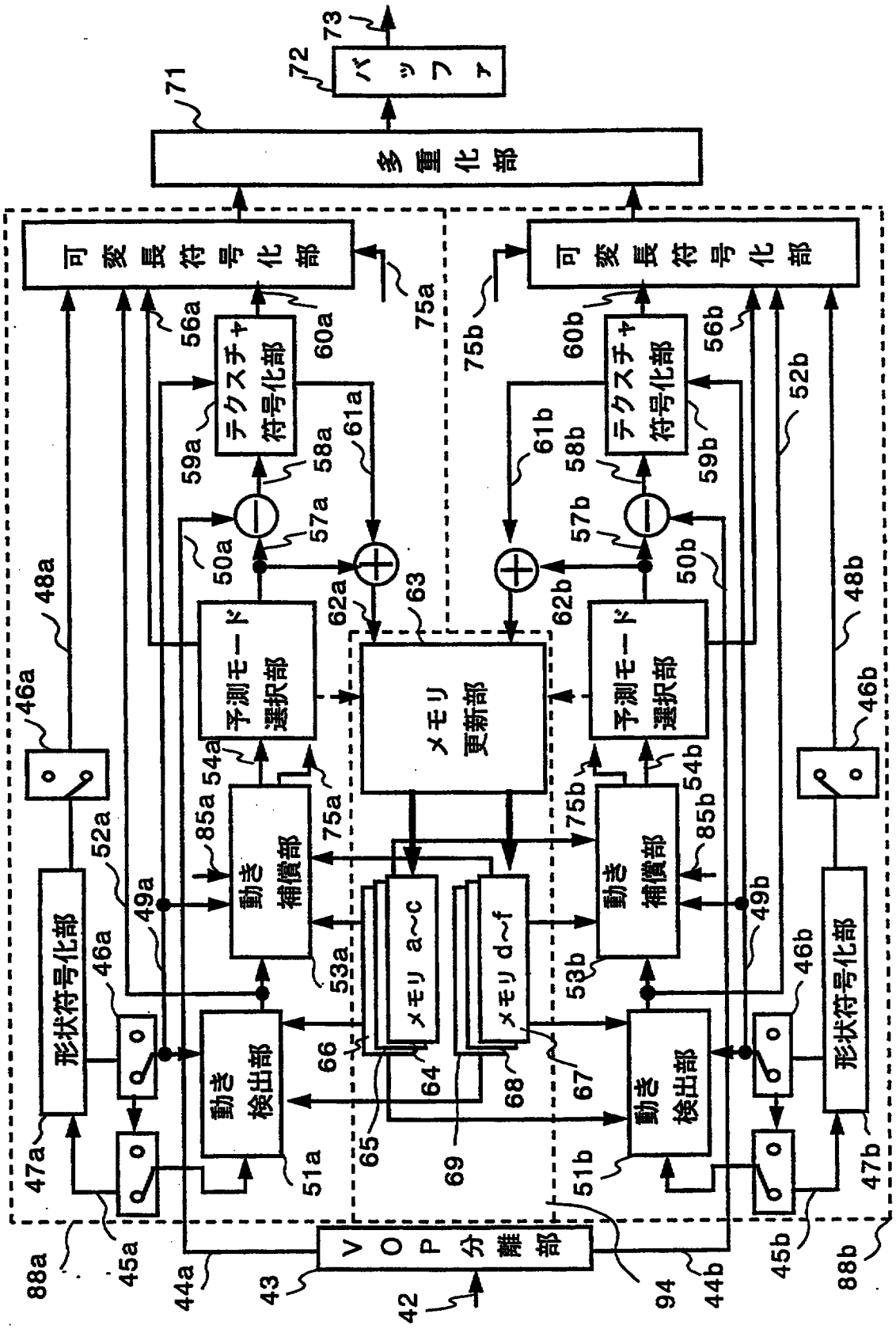

多重化部
71
バッファ
72
73
可変長符号化部
テクスチャ符号化部
予測モード選択部
動き補償部
メモリ更新部
メモリ a~c
メモリ d~f
動き検出部
形状符号化部
VOP分離部
42
43
44a
44b
45a
45b
46a
46b
47a
47b
48a
48b
49a
49b
50a
50b
51a
51b
52a
52b
53a
53b
54a
54b
56a
56b
57a
57b
58a
58b
59a
59b
60a
60b
61a
61b
62a
62b
63
64
65
66
67
68
69
75a
75b
85a
85b
88a
88b
94

図12

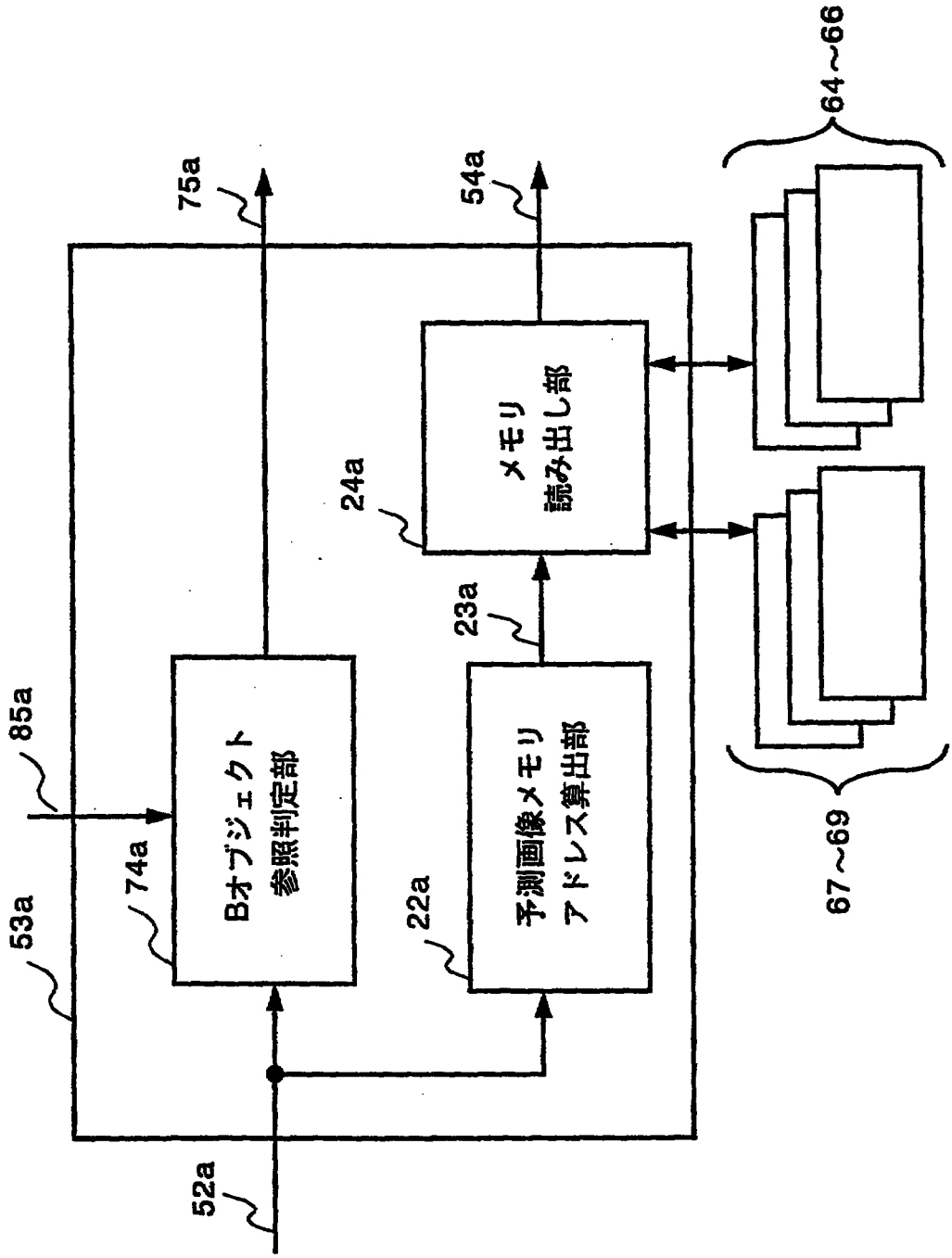

64～66
75a
54a
メモリ
読み出し部
24a
23a
85a
Bオブジェクト
参照判定部
予測画像メモリ
アドレス算出部
74a
22a
53a
67～69
52a

図13

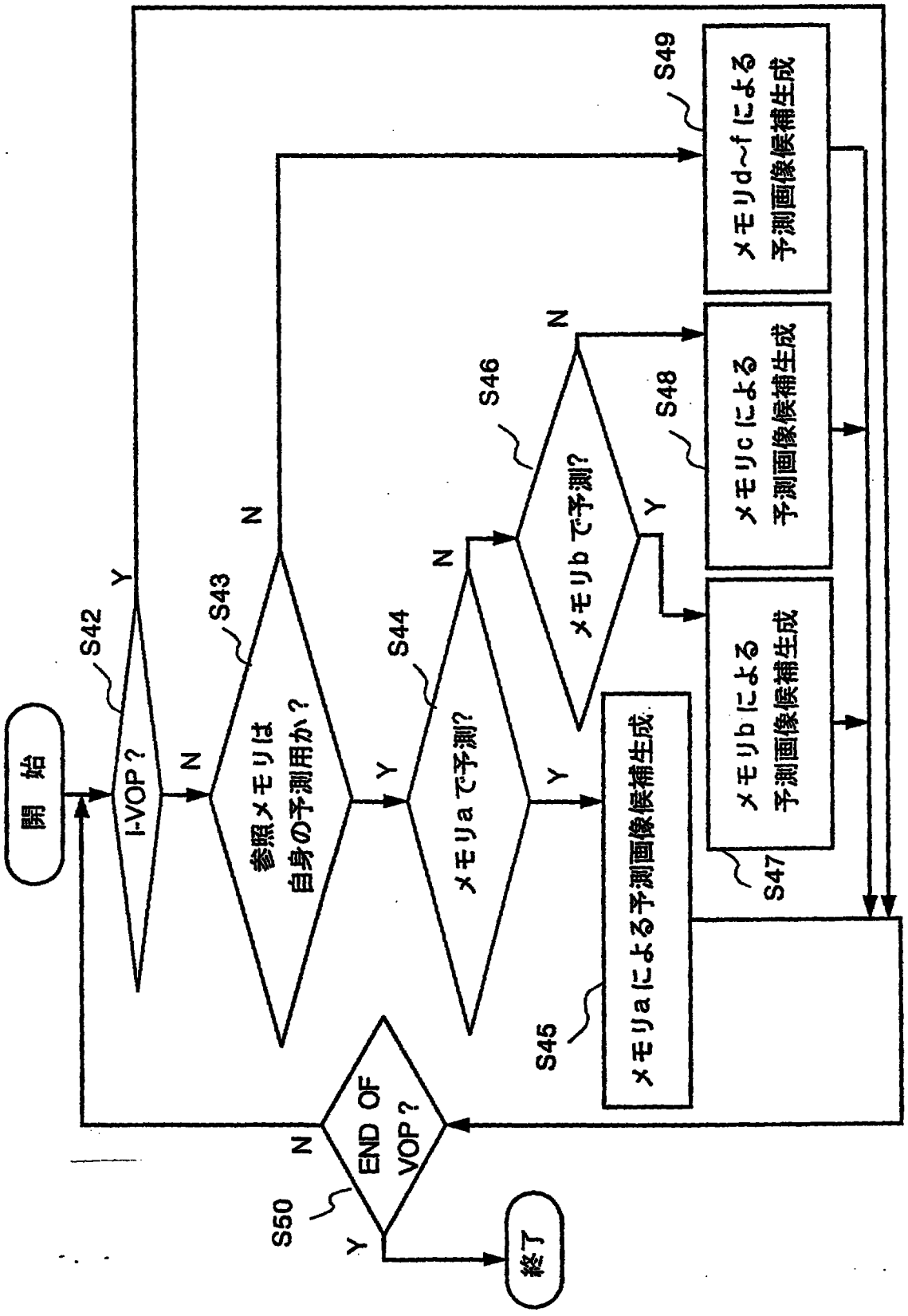

開始
I-VOP?
S42
Y
N
参照メモリは
自身の予測用か?
S43
N
Y
メモリaで予測?
S44
N
Y
メモリbで予測?
S46
N
Y
メモリaによる予測画像候補生成
S45
メモリbによる
予測画像候補生成
S47
メモリcによる
予測画像候補生成
S48
メモリd~fによる
予測画像候補生成
S49
END OF
VOP?
S50
N
Y
終了

図14

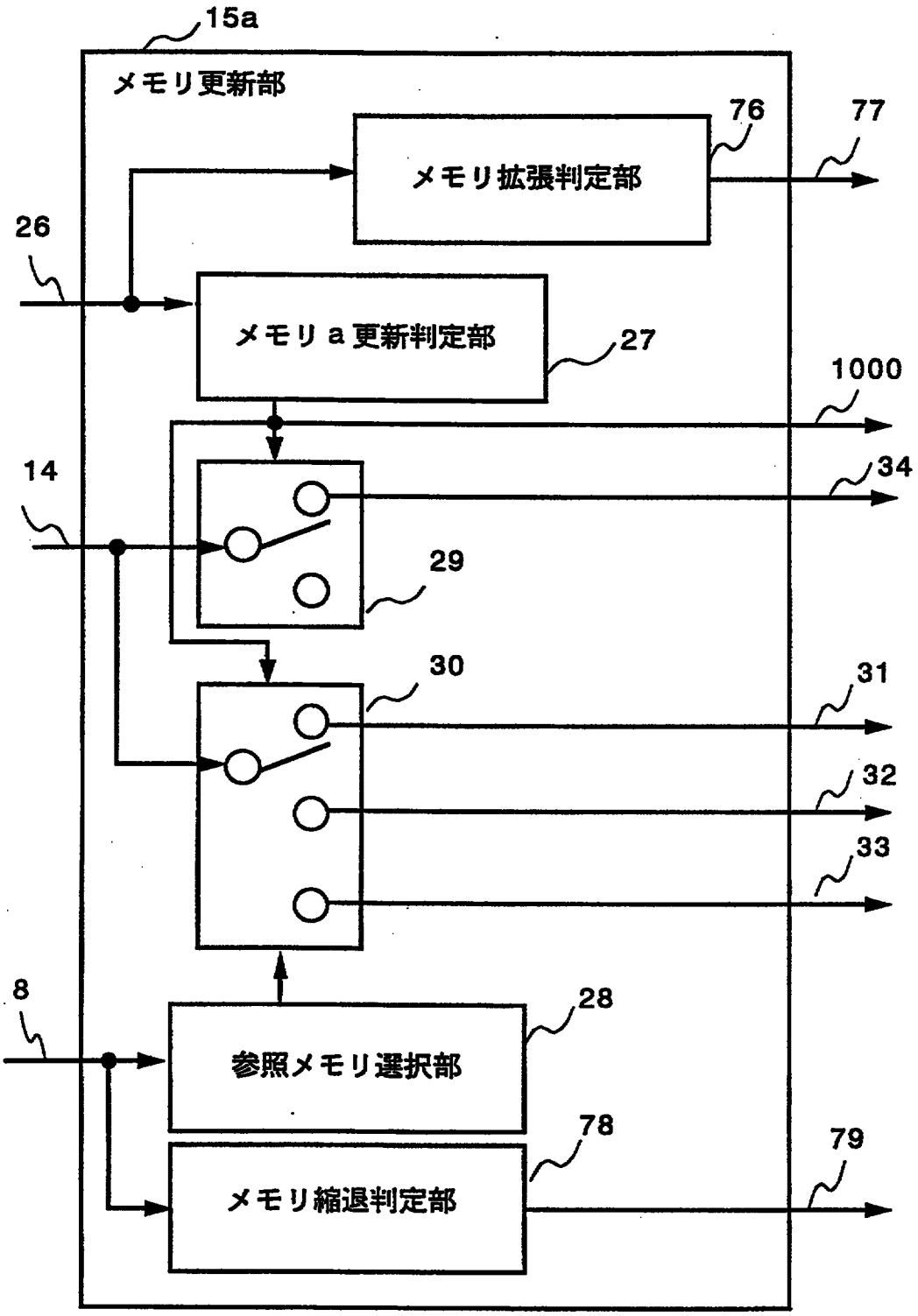

15a
メモリ更新部
76
77
メモリ拡張判定部
26
メモリａ更新判定部
27
1000
14
34
29
30
31
32
33
8
28
参照メモリ選択部
78
79
メモリ縮退判定部

図15

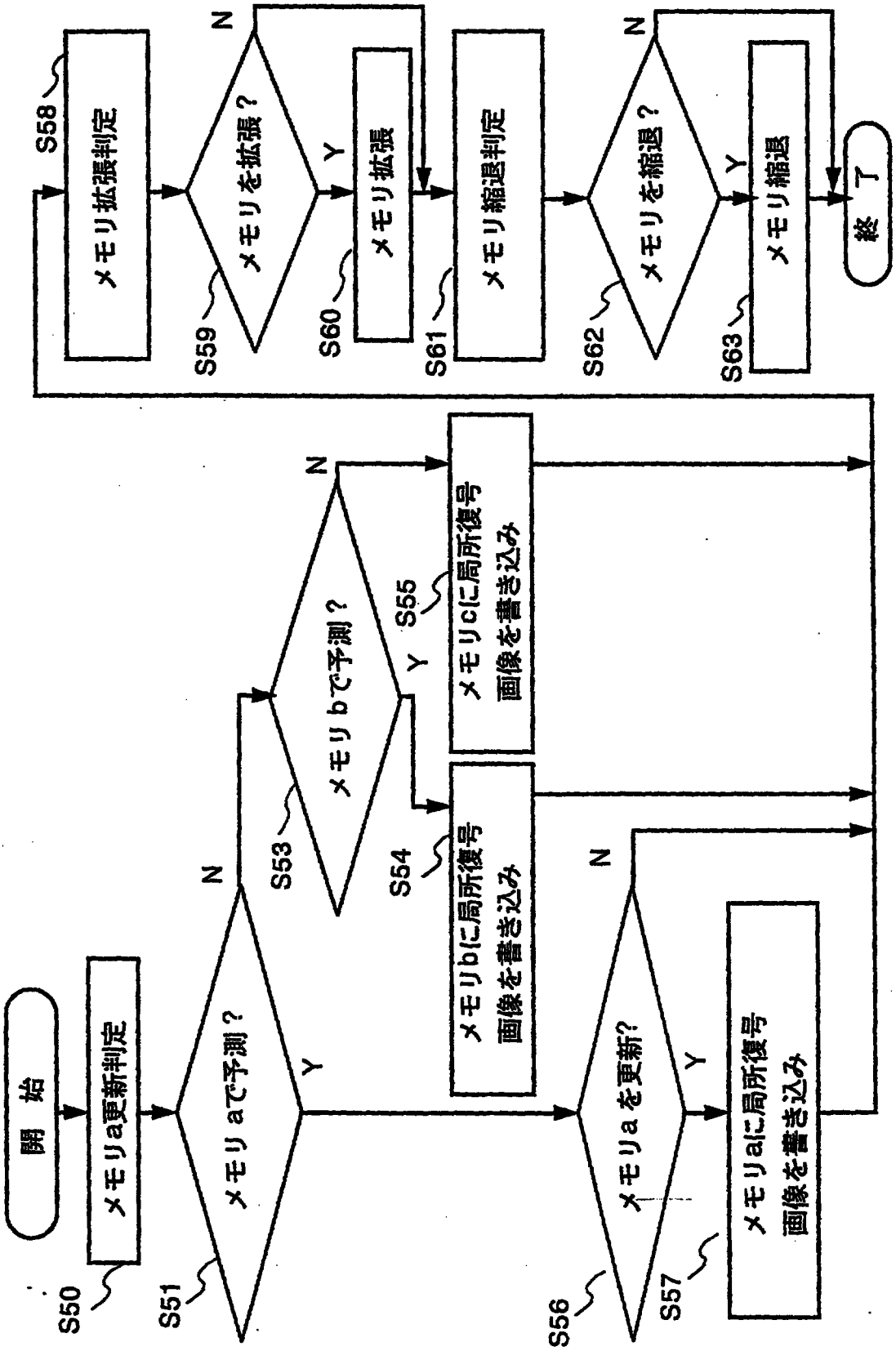

開 始
メモリa更新判定
S50
メモリ a で予測？
S51
N
Y
メモリ b で予測？
S53
N
Y
メモリbに局所復号
画像を書き込み
S54
メモリcに局所復号
画像を書き込み
S55
メモリ a を更新?
S56
N
Y
メモリaに局所復号
画像を書き込み
S57
メモリ拡張判定
S58
メモリを拡張？
S59
N
Y
メモリ拡張
S60
メモリ縮退判定
S61
メモリを縮退？
S62
N
Y
メモリ縮退
S63
終 了

図16

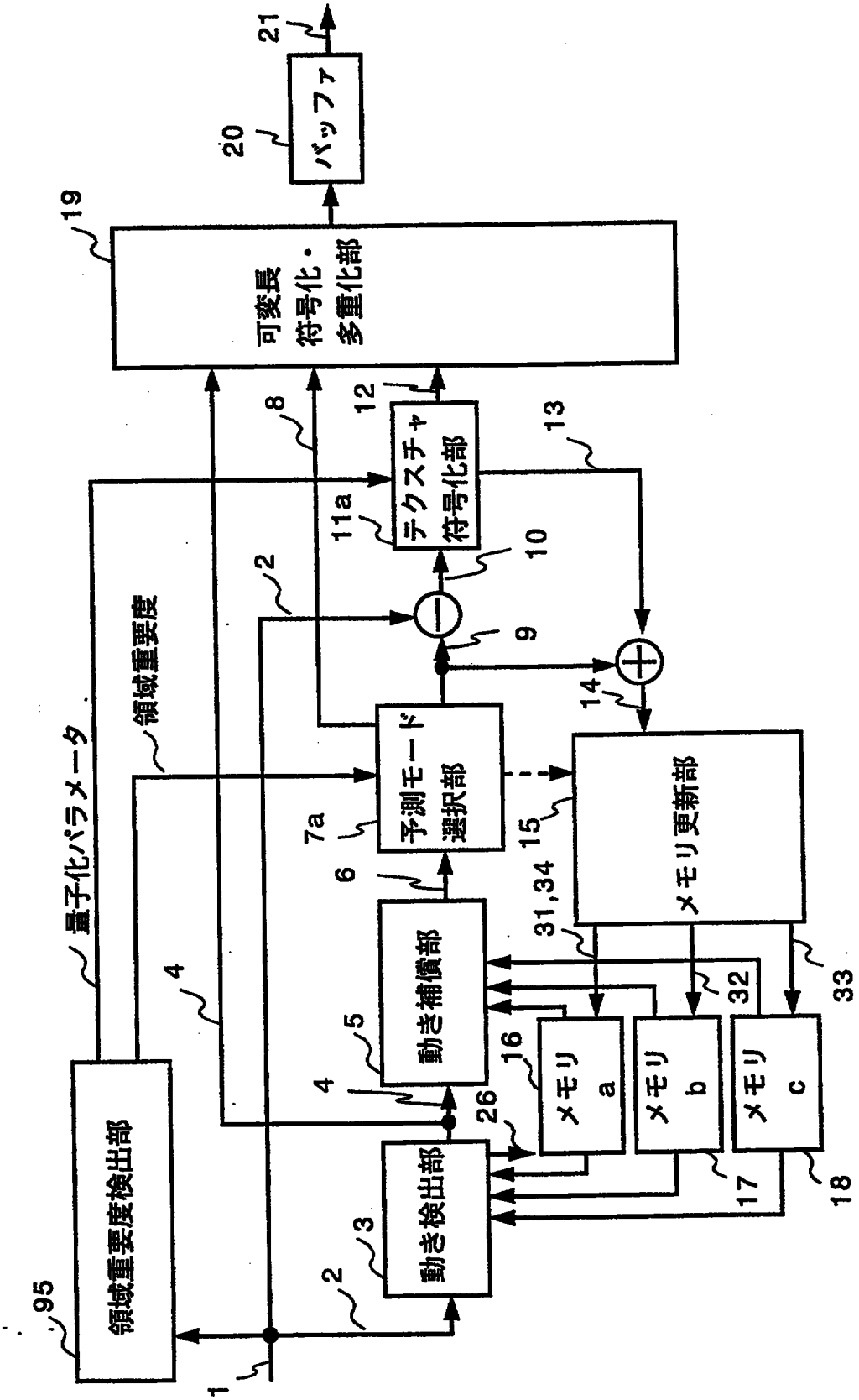

領域重要度検出部
動き検出部
動き補償部
予測モード選択部
テクスチャ符号化部
可変長符号化・多重化部
バッファ
メモリ更新部
メモリ a
メモリ b
メモリ c
量子化パラメータ
領域重要度
1
2
3
4
5
6
7a
8
9
10
11a
12
13
14
15
16
17
18
19
20
21
26
31,34
32
33
95

図17

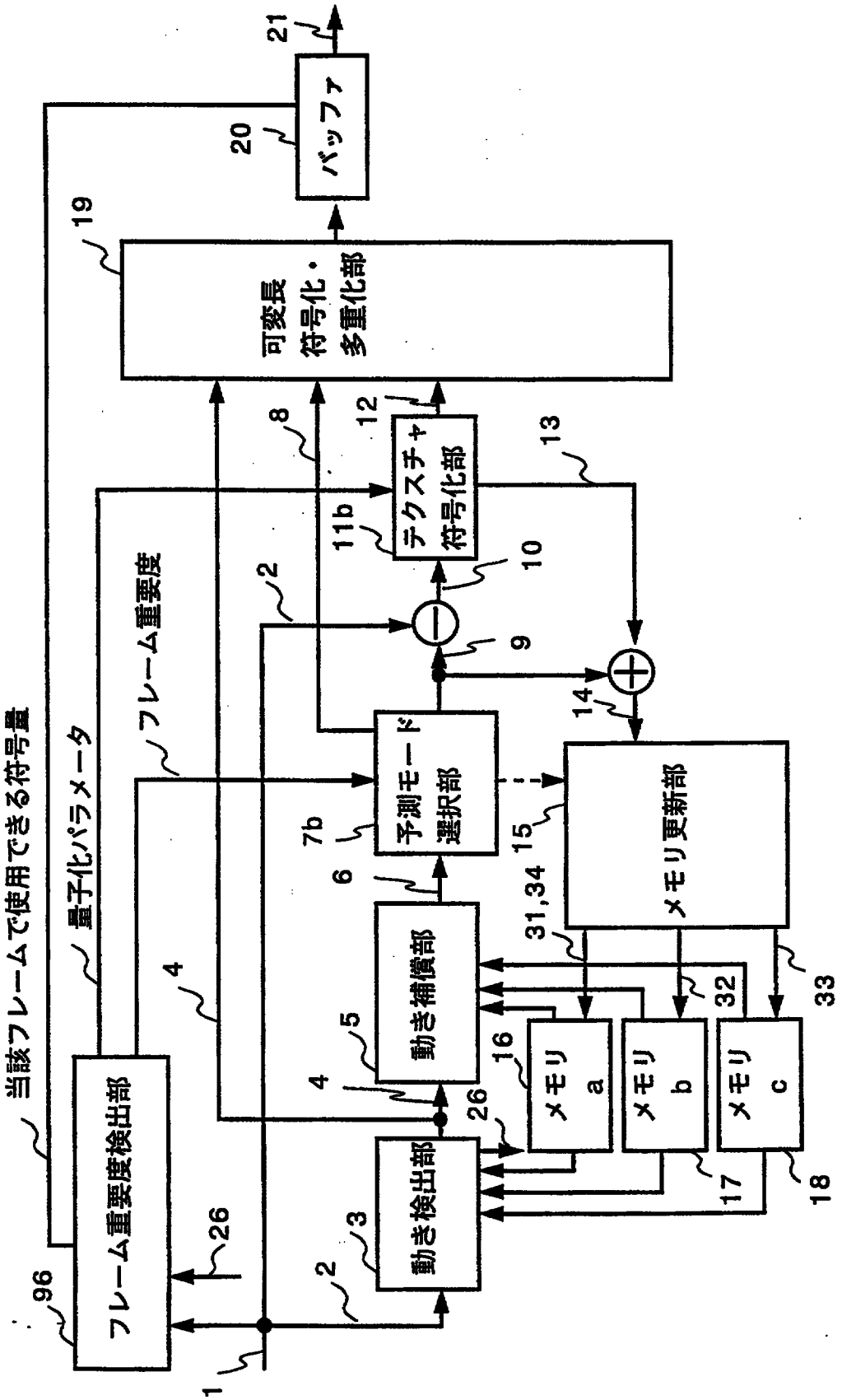

当該フレームで使用できる符号量
量子化パラメータ
フレーム重要度
フレーム重要度検出部
動き検出部
動き補償部
予測モード選択部
テクスチャ符号化部
可変長符号化・多重化部
バッファ
メモリ更新部
メモリ a
メモリ b
メモリ c
96
1
2
3
4
5
6
7b
8
9
10
11b
12
13
14
15
16
17
18
19
20
21
26
31,34
32
33

図18

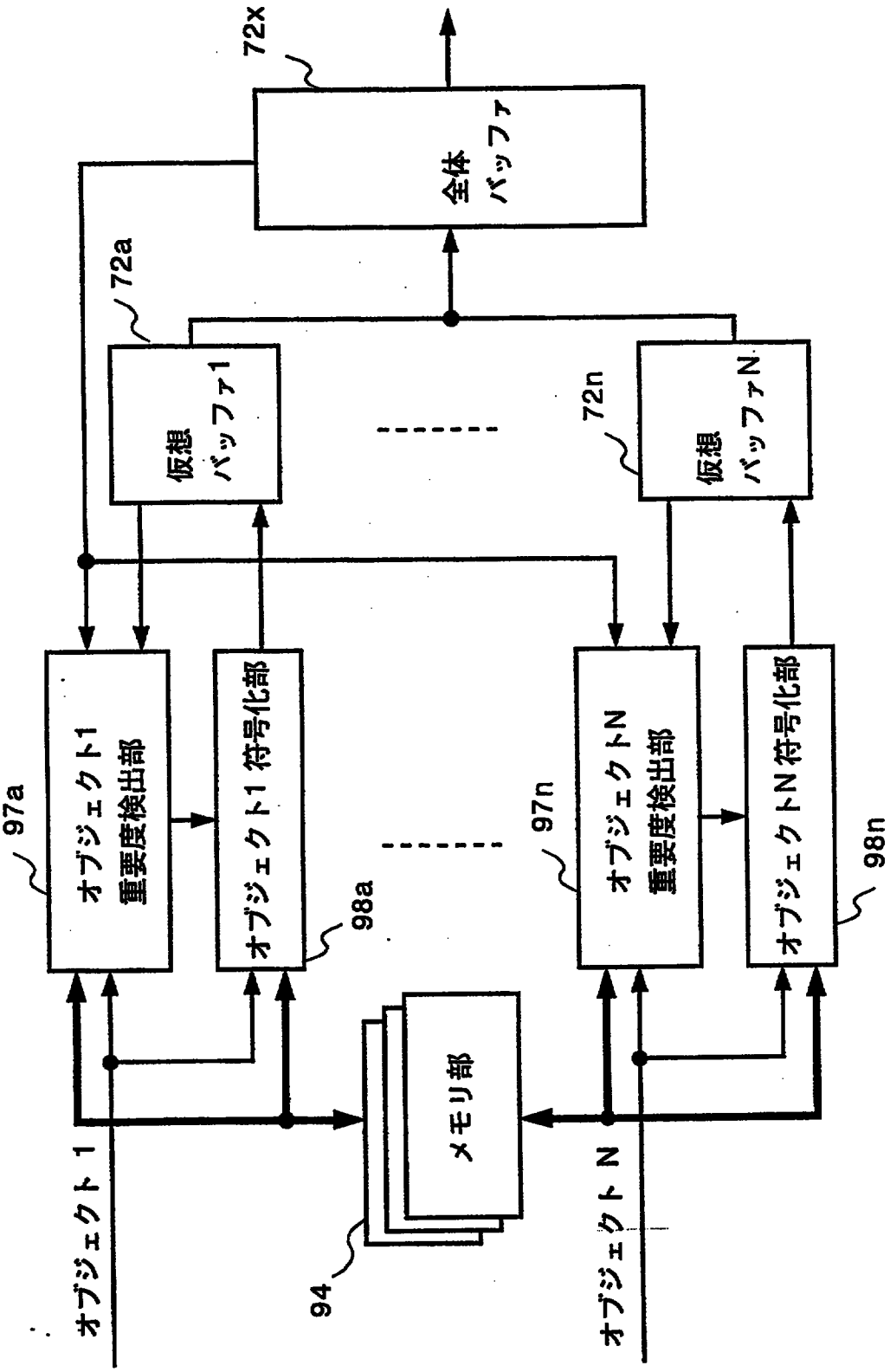

72x
全体
バッファ
72a
仮想
バッファ1
72n
仮想
バッファN
97a
オブジェクト1
重要度検出部
オブジェクト1 符号化部
98a
97n
オブジェクトN
重要度検出部
オブジェクトN 符号化部
98n
メモリ部
94
オブジェクト 1
オブジェクト N

図19

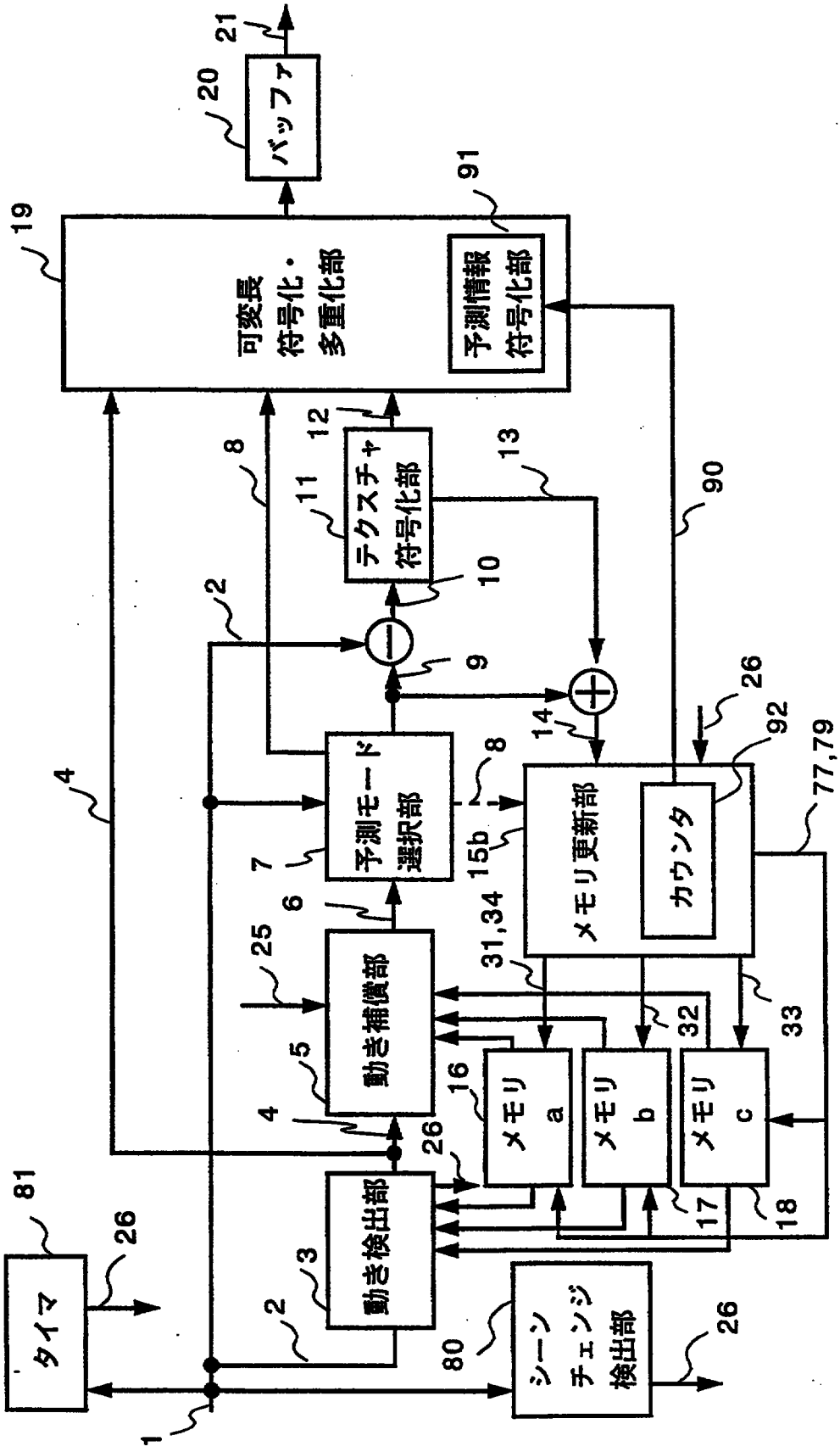

タイマ
81
26
1
2
3
動き検出部
4
5
動き補償部
6
7
予測モード選択部
8
9
10
11
テクスチャ符号化部
12
13
14
15b
メモリ更新部
カウンタ
77,79
92
16
メモリ a
メモリ b
メモリ c
17
18
31,34
32
33
25
26
4
2
80
シーンチェンジ検出部
26
19
可変長符号化・多重化部
91
予測情報符号化部
90
20
バッファ
21

図２０

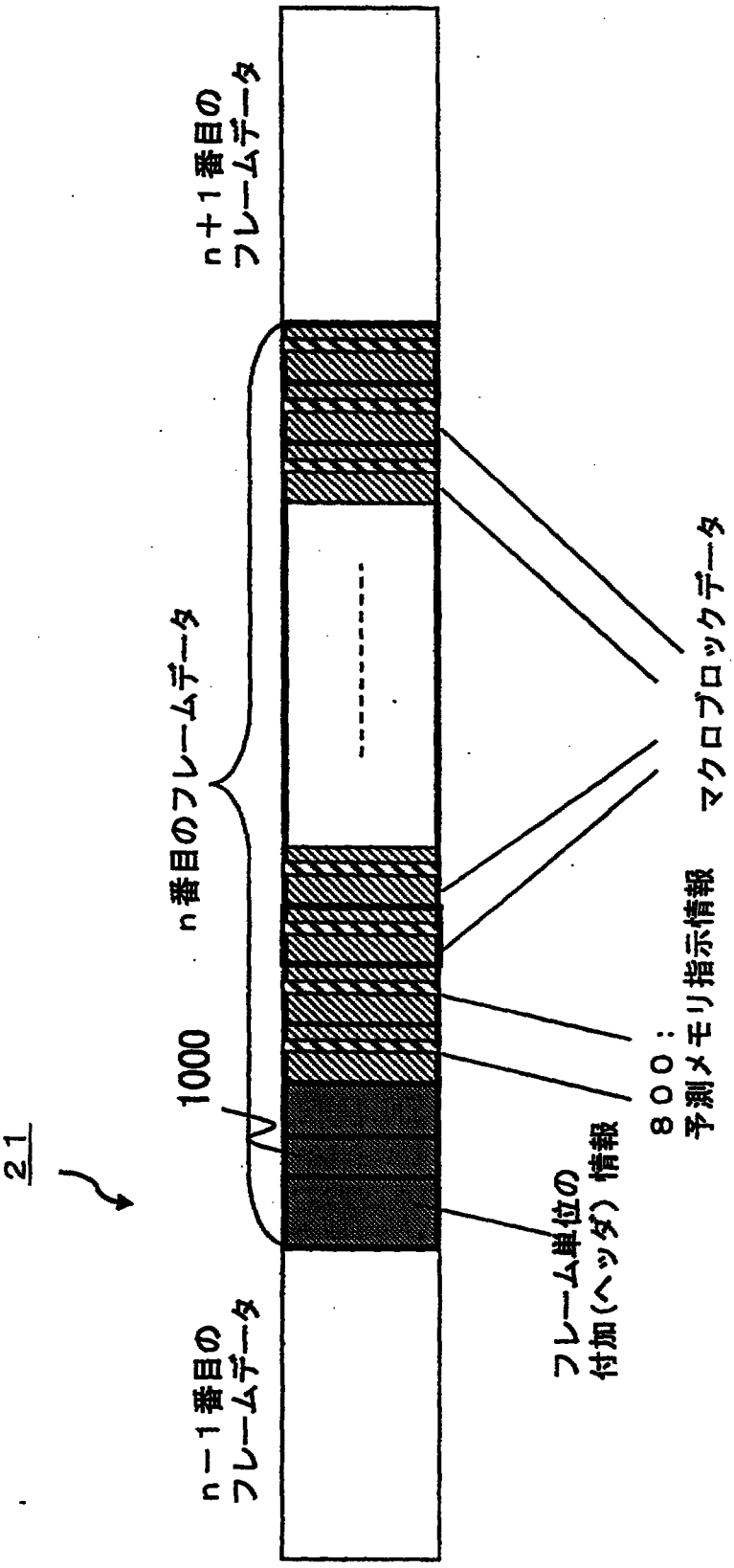

21
1000
ｎ＋１番目の
フレームデータ
ｎ番目のフレームデータ
ｎ－１番目の
フレームデータ
フレーム単位の
付加（ヘッダ）情報
８００：
予測メモリ指示情報
マクロブロックデータ

図２１

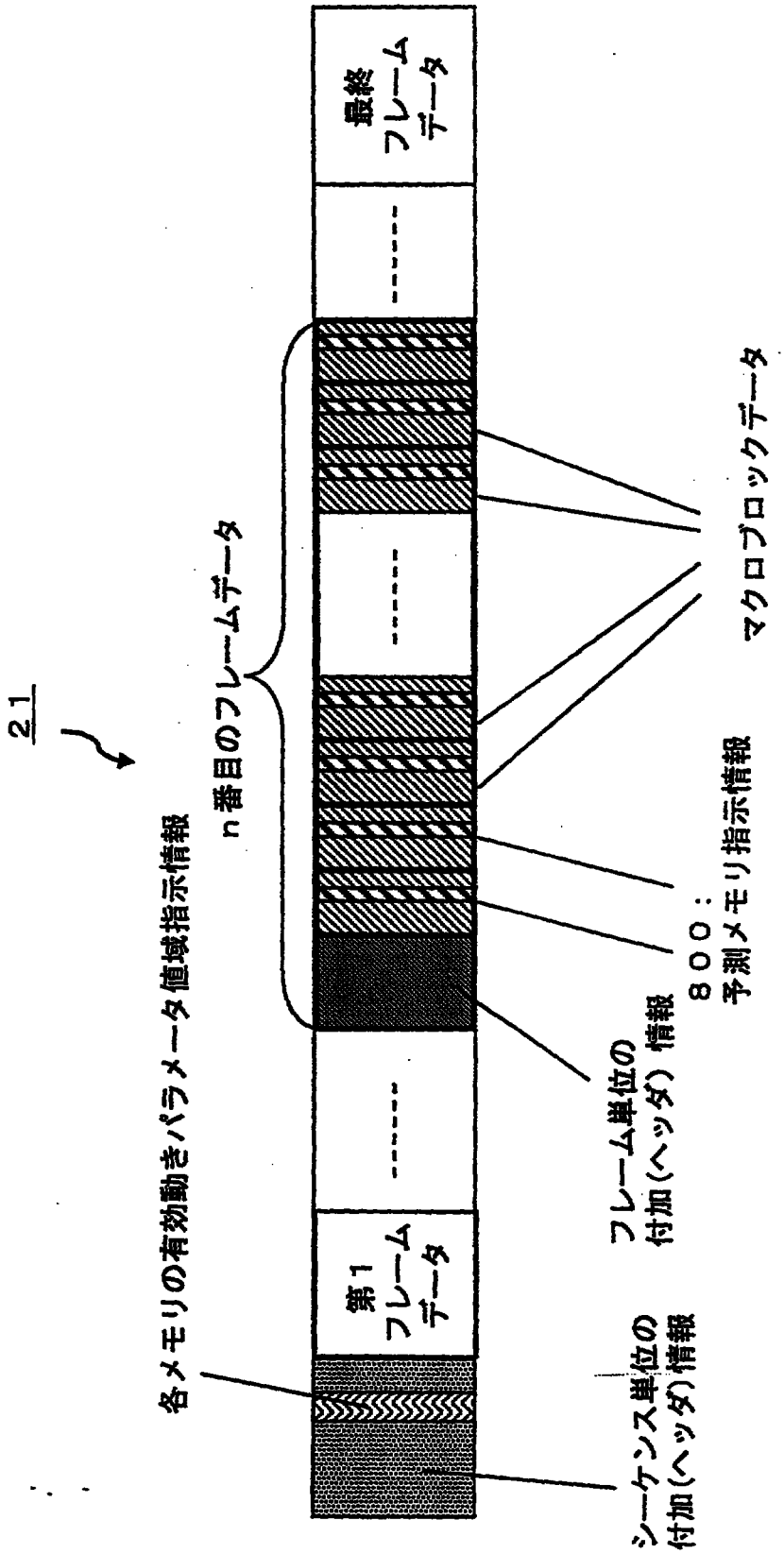

21
各メモリの有効動きパラメータ値域指示情報
n番目のフレームデータ
第1
フレーム
データ
最終
フレーム
データ
シーケンス単位の
付加(ヘッダ)情報
フレーム単位の
付加(ヘッダ)情報
800：
予測メモリ指示情報
マクロブロックデータ

図２２

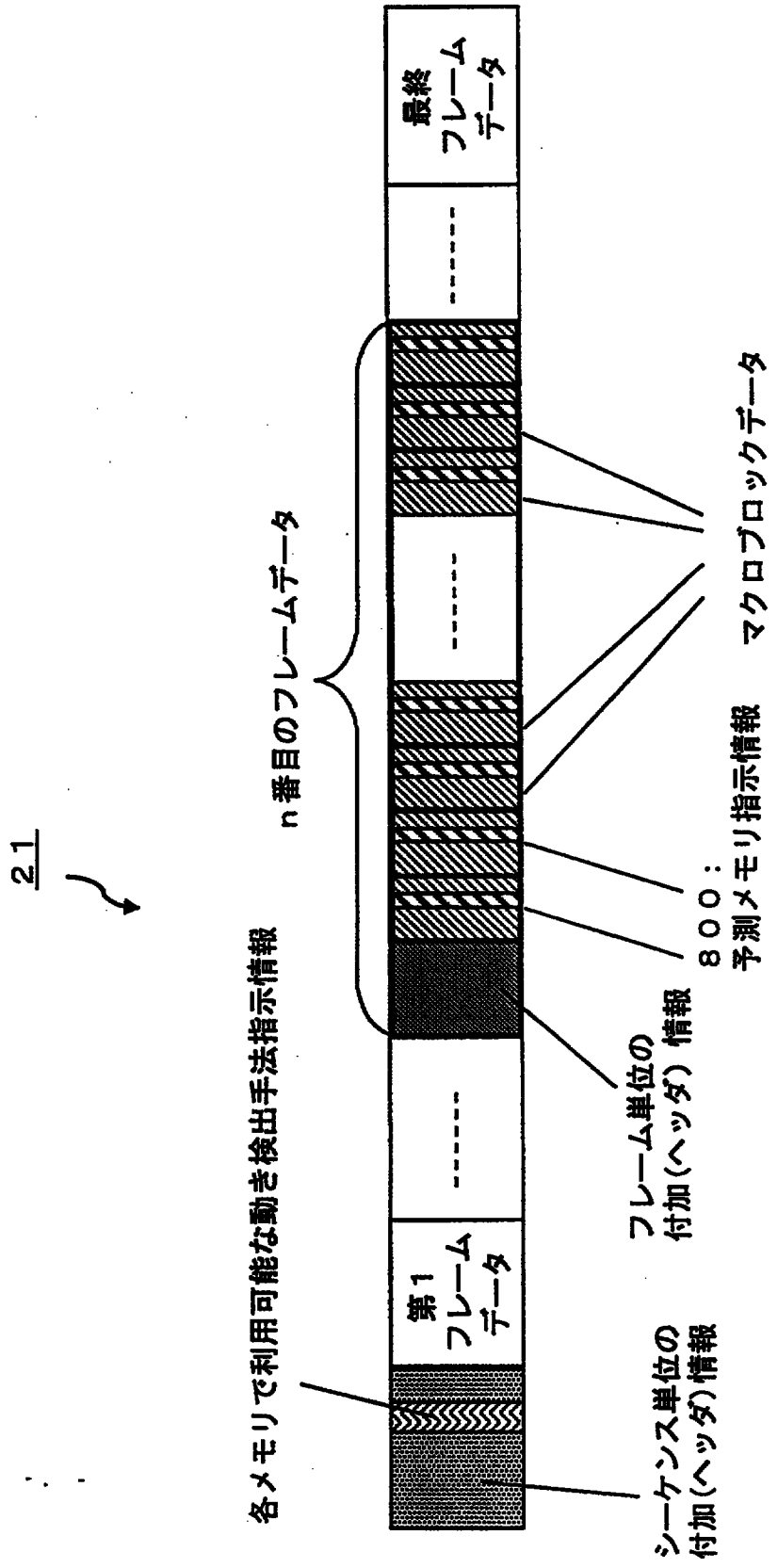

21
各メモリで利用可能な動き検出手法指示情報
n番目のフレームデータ
第1
フレーム
データ
最終
フレーム
データ
シーケンス単位の
付加(ヘッダ)情報
フレーム単位の
付加(ヘッダ)情報
800：
予測メモリ指示情報
マクロブロックデータ

図２３

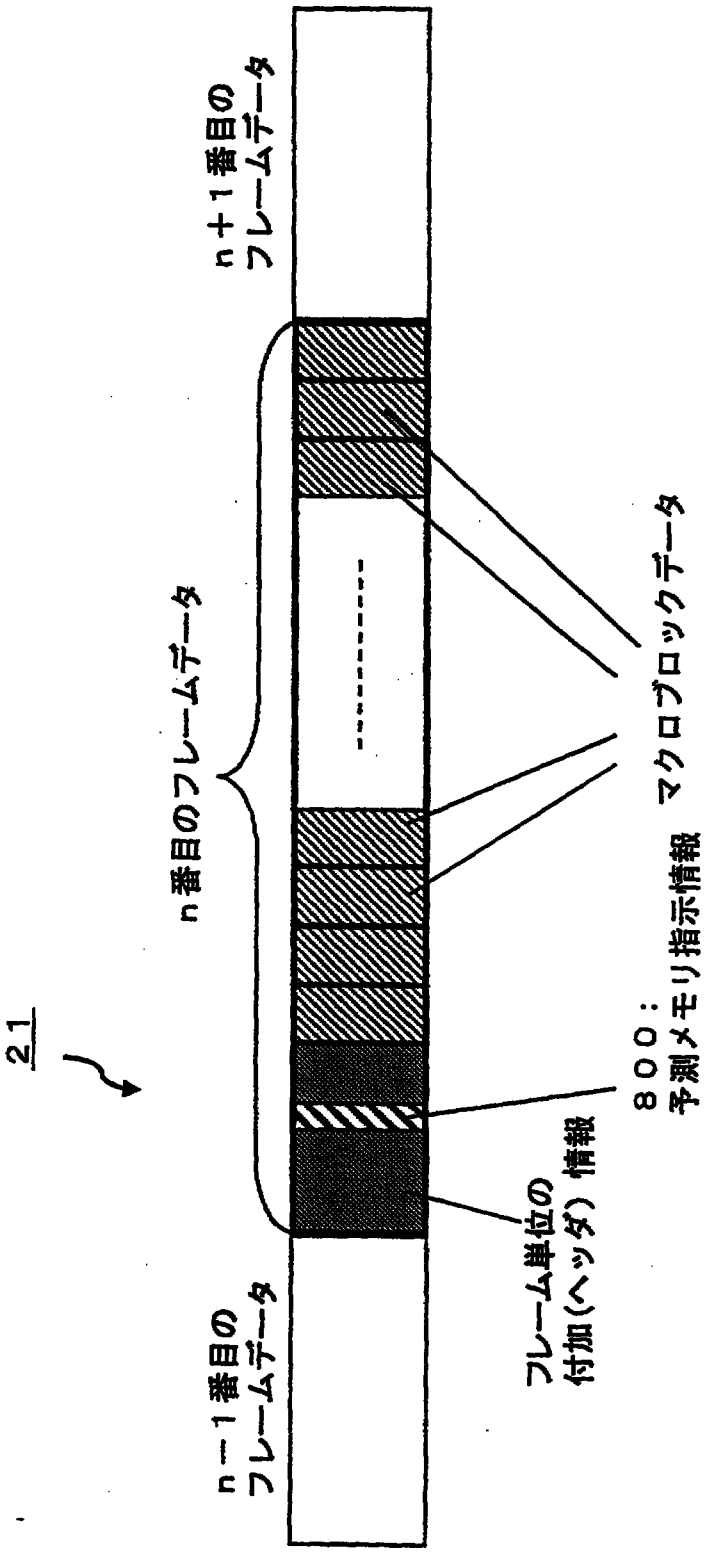

21
n＋1番目の
フレームデータ
n番目のフレームデータ
n－1番目の
フレームデータ
マクロブロックデータ
８００：
予測メモリ指示情報
フレーム単位の
付加（ヘッダ）情報

図24

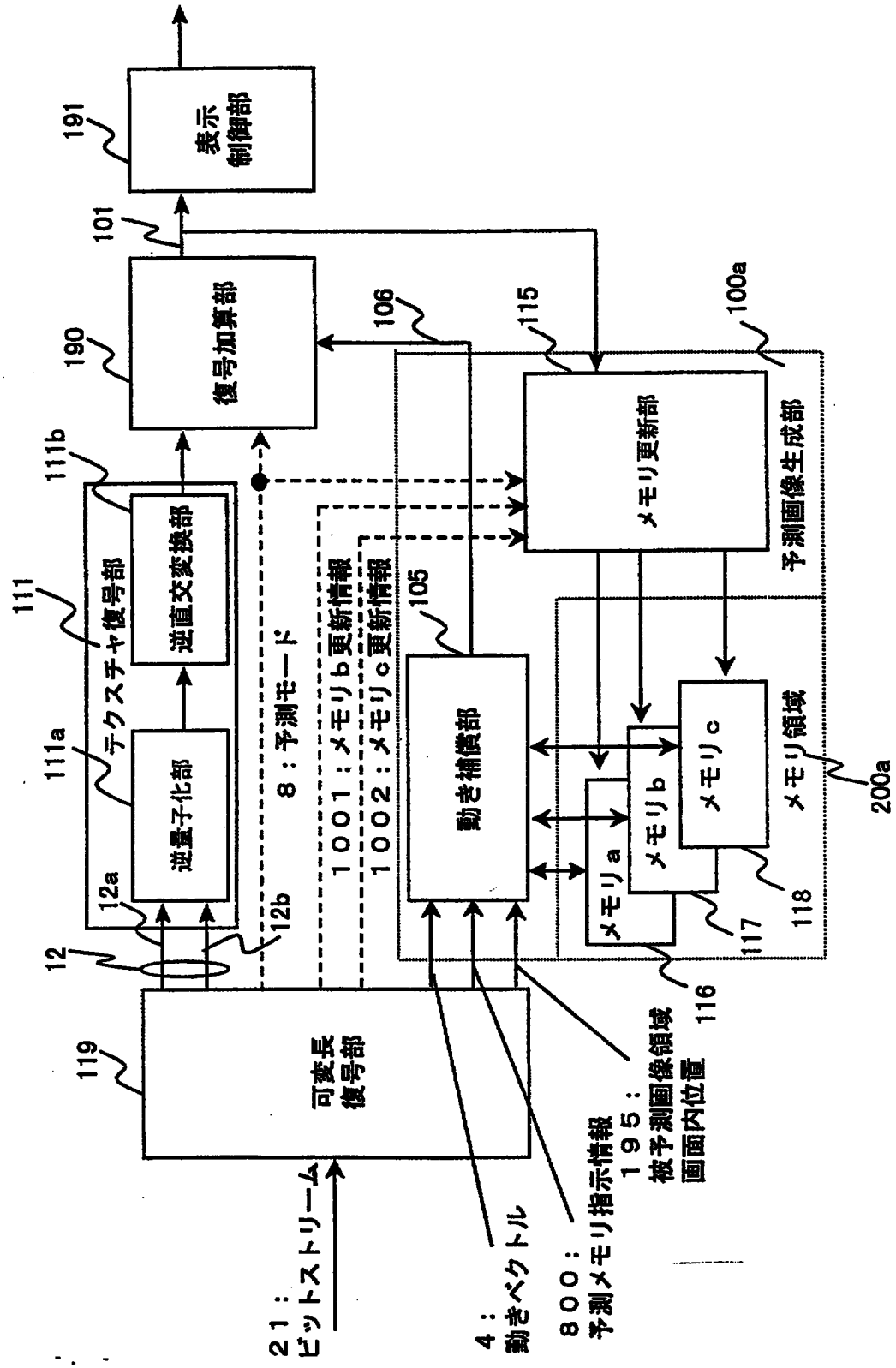

191
表示制御部
101
190
復号加算部
106
115
100a
111b
逆直交変換部
予測画像生成部
111
テクスチャ復号部
メモリ更新部
111a
逆量子化部
8：予測モード
1001：メモリb更新情報
1002：メモリc更新情報
105
動き補償部
メモリc
メモリ領域
200a
メモリb
メモリa
12a
12b
12
118
117
116
119
可変長復号部
21：ビットストリーム
4：動きベクトル
800：予測メモリ指示情報
195：被予測画像領域画面内位置

図25

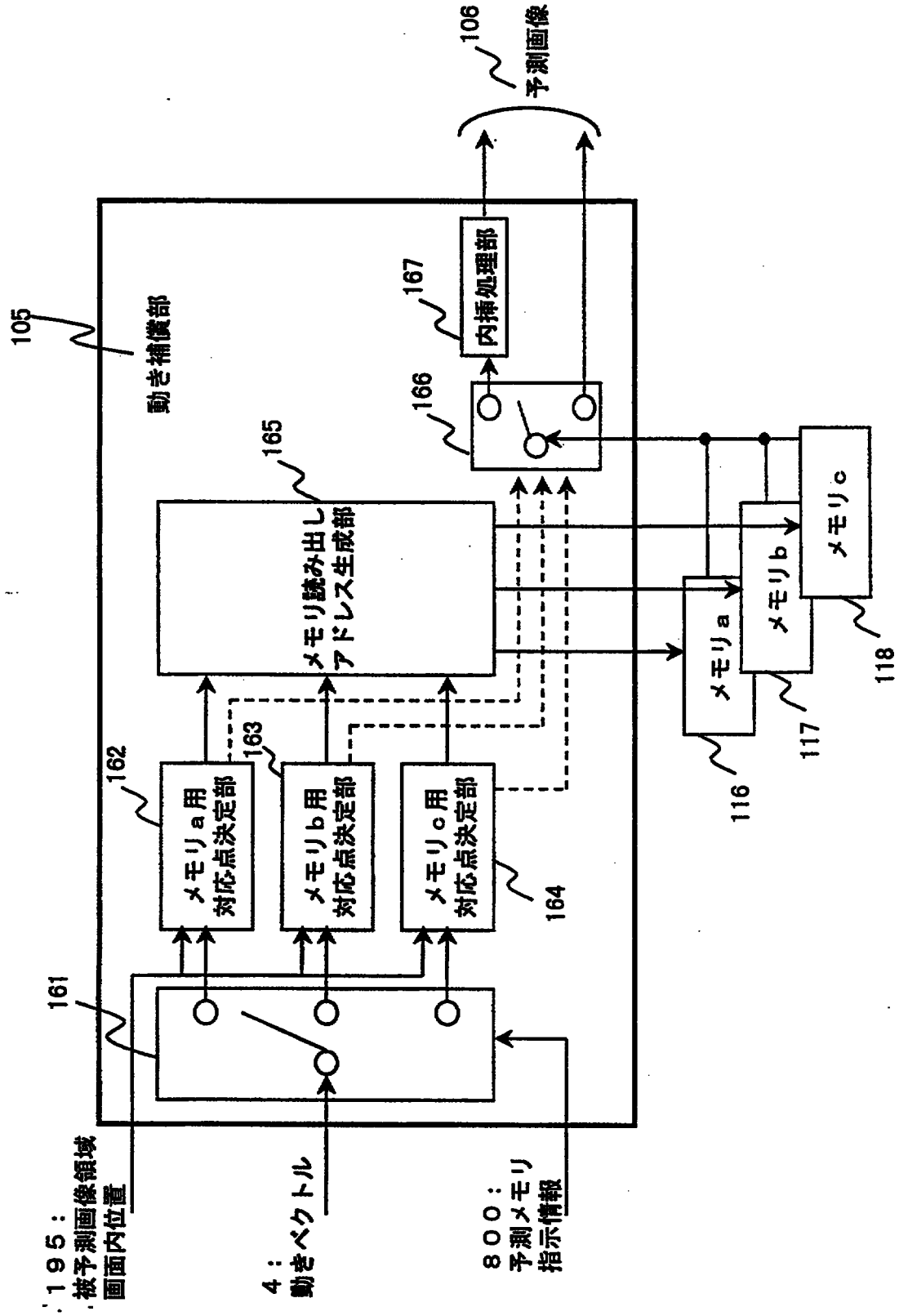

106
予測画像
105
動き補償部
167
内挿処理部
166
165
メモリ読み出しアドレス生成部
メモリa
メモリb
メモリc
118
117
116
162
163
メモリa用対応点決定部
メモリb用対応点決定部
メモリc用対応点決定部
164
161
195：被予測画像領域画面内位置
4：動きベクトル
800：予測メモリ指示情報

図26

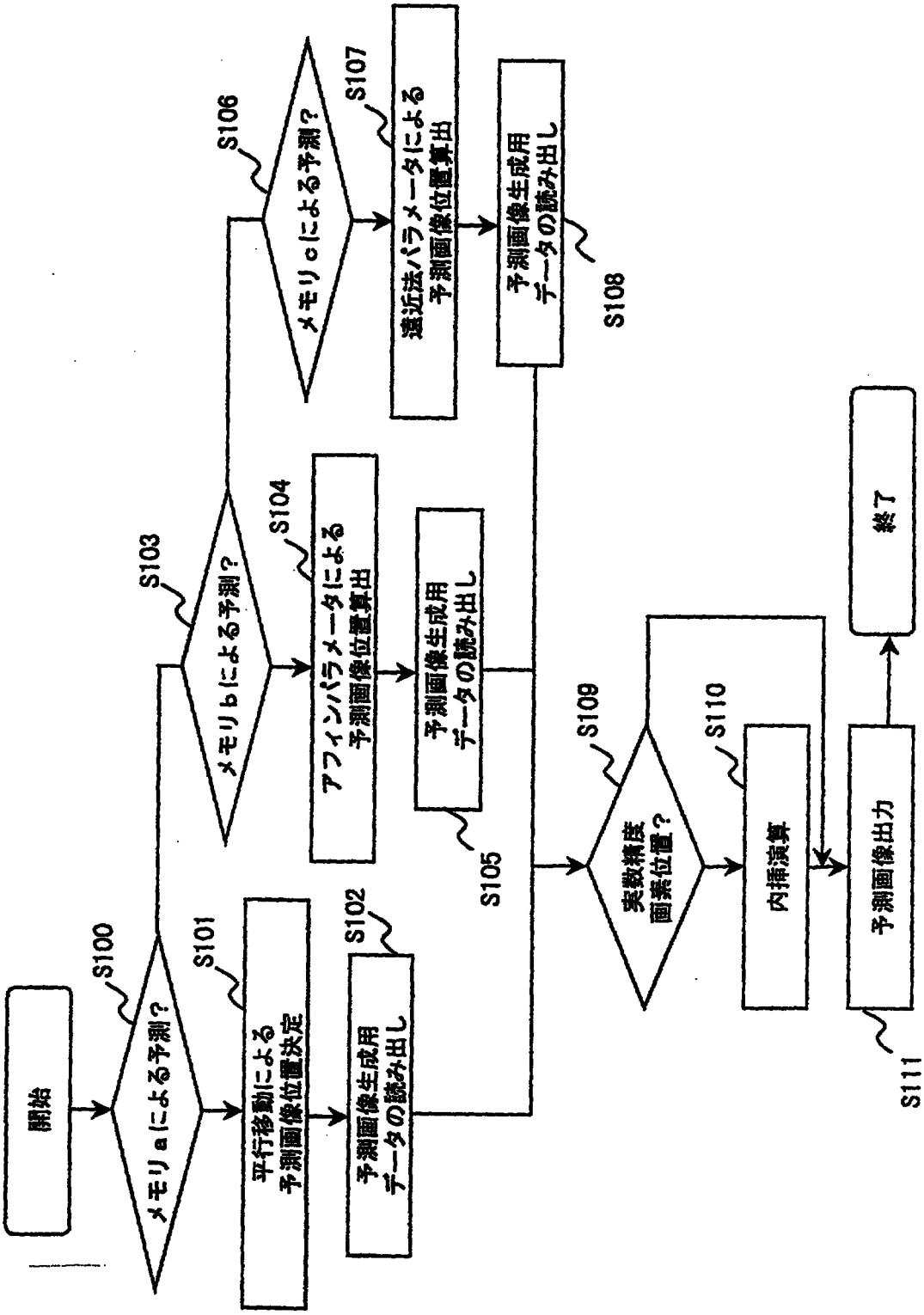

開始
S100
メモリaによる予測?
S101
平行移動による予測画像位置決定
S102
予測画像生成用データの読み出し
S103
メモリbによる予測?
S104
アフィンパラメータによる予測画像位置算出
S105
予測画像生成用データの読み出し
S106
メモリcによる予測?
S107
遠近法パラメータによる予測画像位置算出
S108
予測画像生成用データの読み出し
S109
実数精度画素位置?
S110
内挿演算
S111
予測画像出力
終了

図27

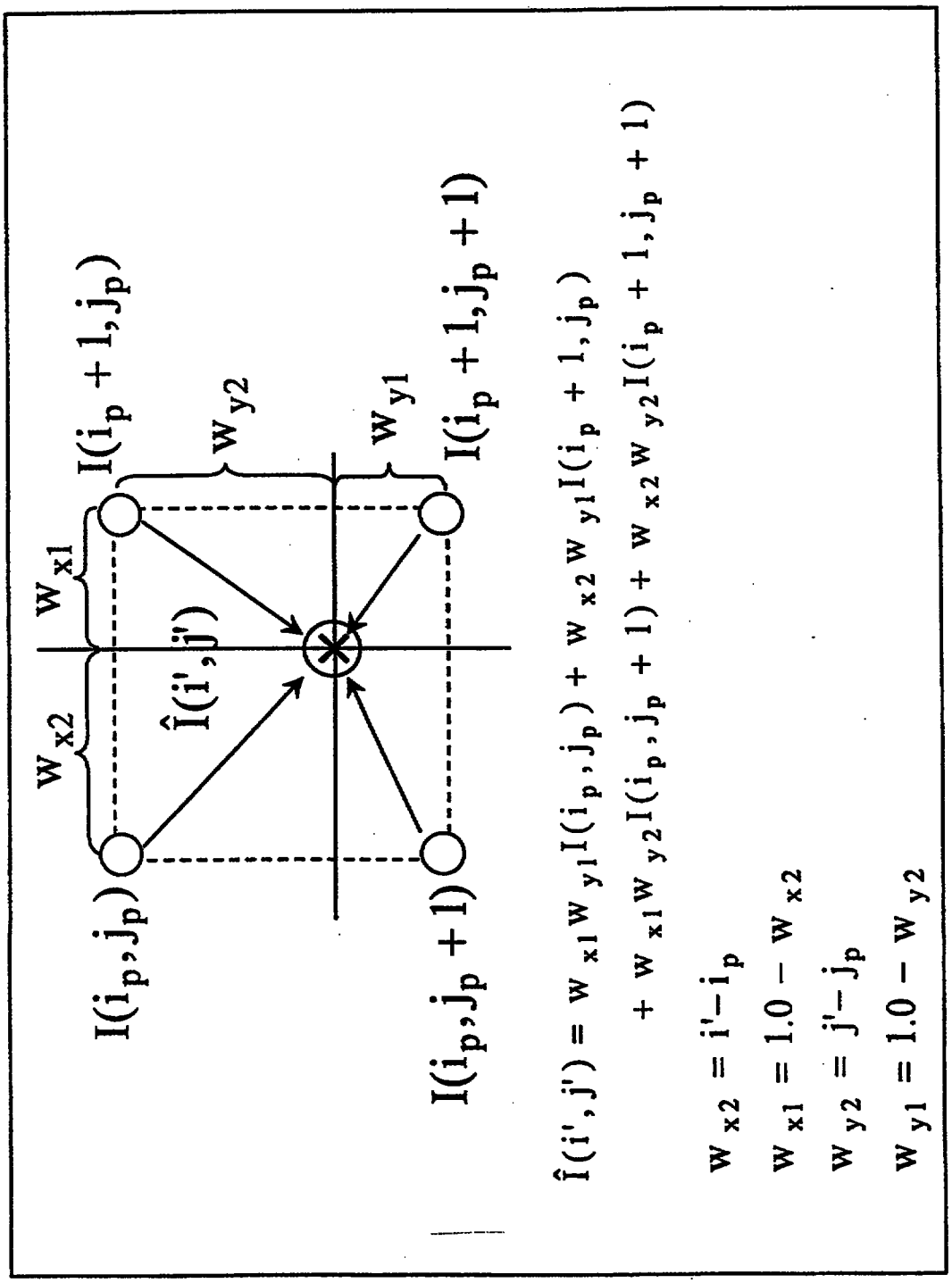

$I(i_p, j_p)$
$I(i_p + 1, j_p)$
$I(i_p, j_p + 1)$
$I(i_p + 1, j_p + 1)$
$\hat{I}(i', j')$
$w_{x1}$
$w_{x2}$
$w_{y1}$
$w_{y2}$
$\hat{I}(i', j') = w_{x1} w_{y1} I(i_p, j_p) + w_{x2} w_{y1} I(i_p + 1, j_p)$
$+ w_{x1} w_{y2} I(i_p, j_p + 1) + w_{x2} w_{y2} I(i_p + 1, j_p + 1)$
$w_{x2} = i' - i_p$
$w_{x1} = 1.0 - w_{x2}$
$w_{y2} = j' - j_p$
$w_{y1} = 1.0 - w_{y2}$

図２８

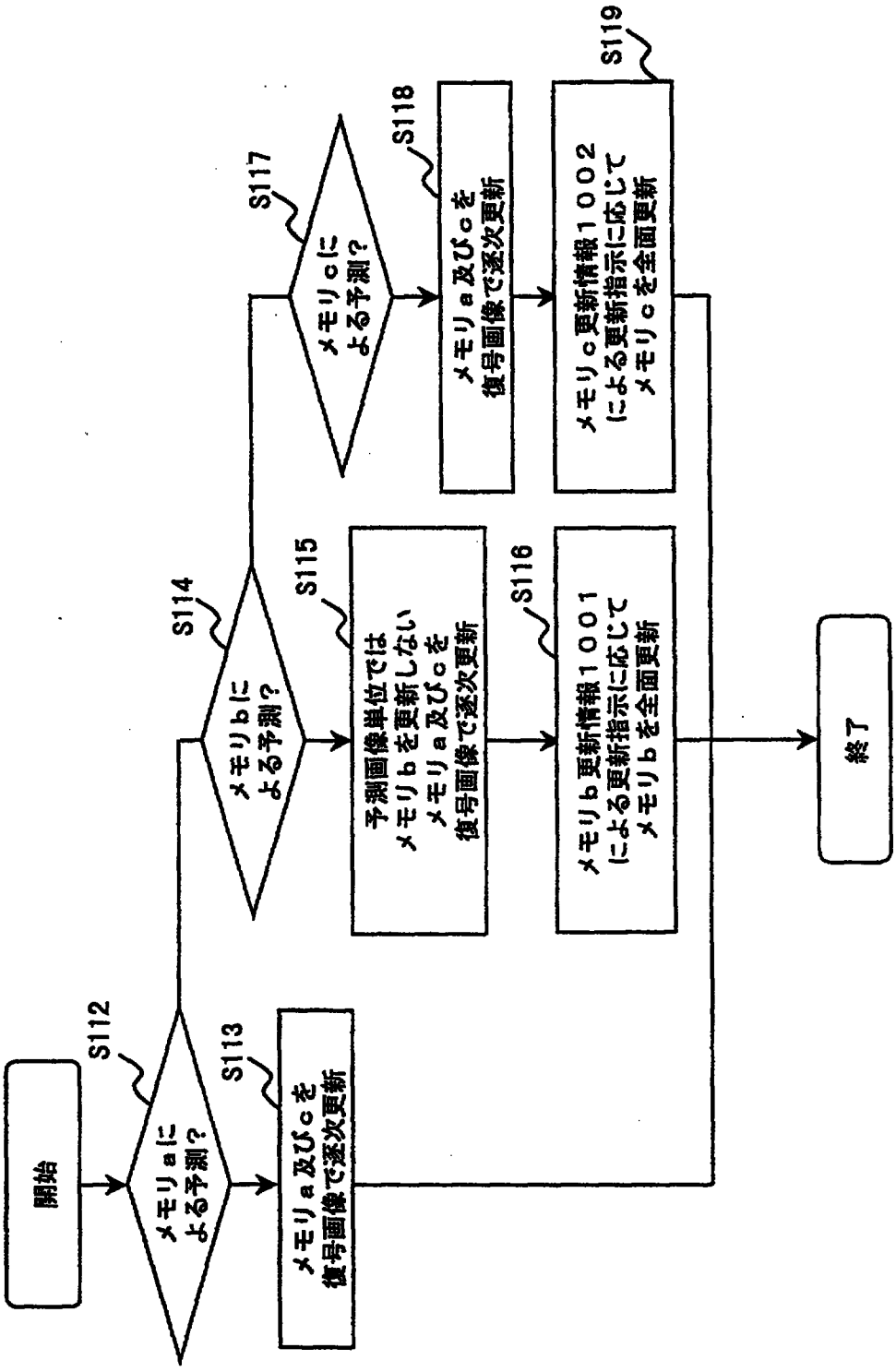

開始
S112
メモリaによる予測？
S113
メモリa及びcを復号画像で逐次更新
S114
メモリbによる予測？
S115
予測画像単位ではメモリbを更新しないメモリa及びcを復号画像で逐次更新
S116
メモリb更新情報１００１による更新指示に応じてメモリbを全面更新
S117
メモリcによる予測？
S118
メモリa及びcを復号画像で逐次更新
S119
メモリc更新情報１００２による更新指示に応じてメモリcを全面更新
終了

図29

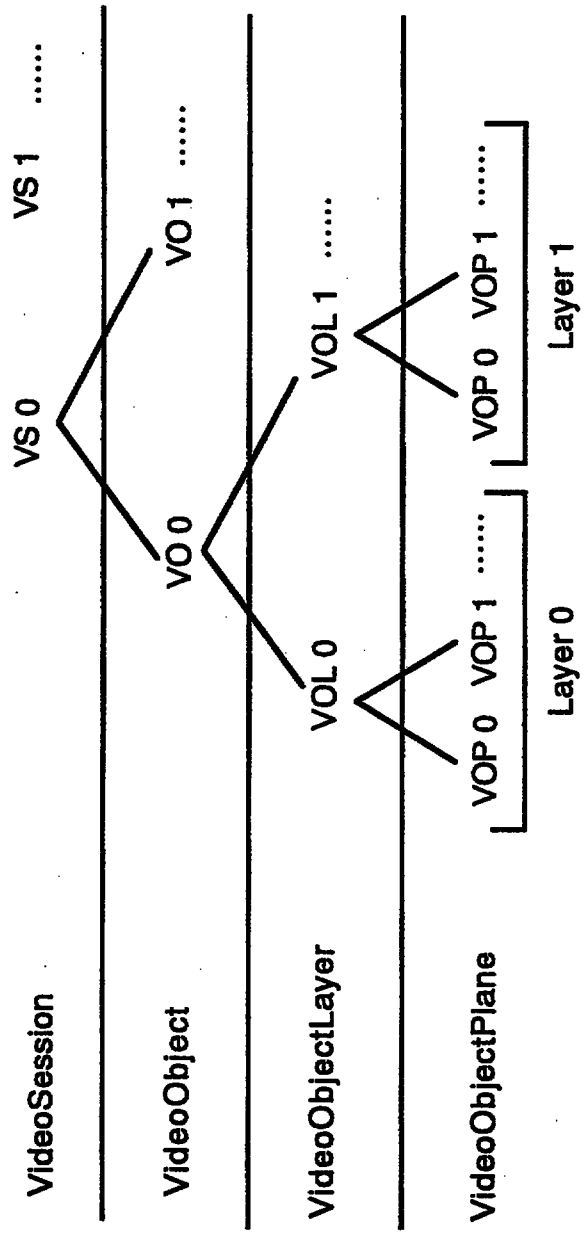

VideoSession
VideoObject
VideoObjectLayer
VideoObjectPlane
VS 0
VS 1 .....
VO 0
VO 1 .....
VOL 0
VOL 1 .....
VOP 0 VOP 1 .....
VOP 0 VOP 1 .....
Layer 0
Layer 1

図30

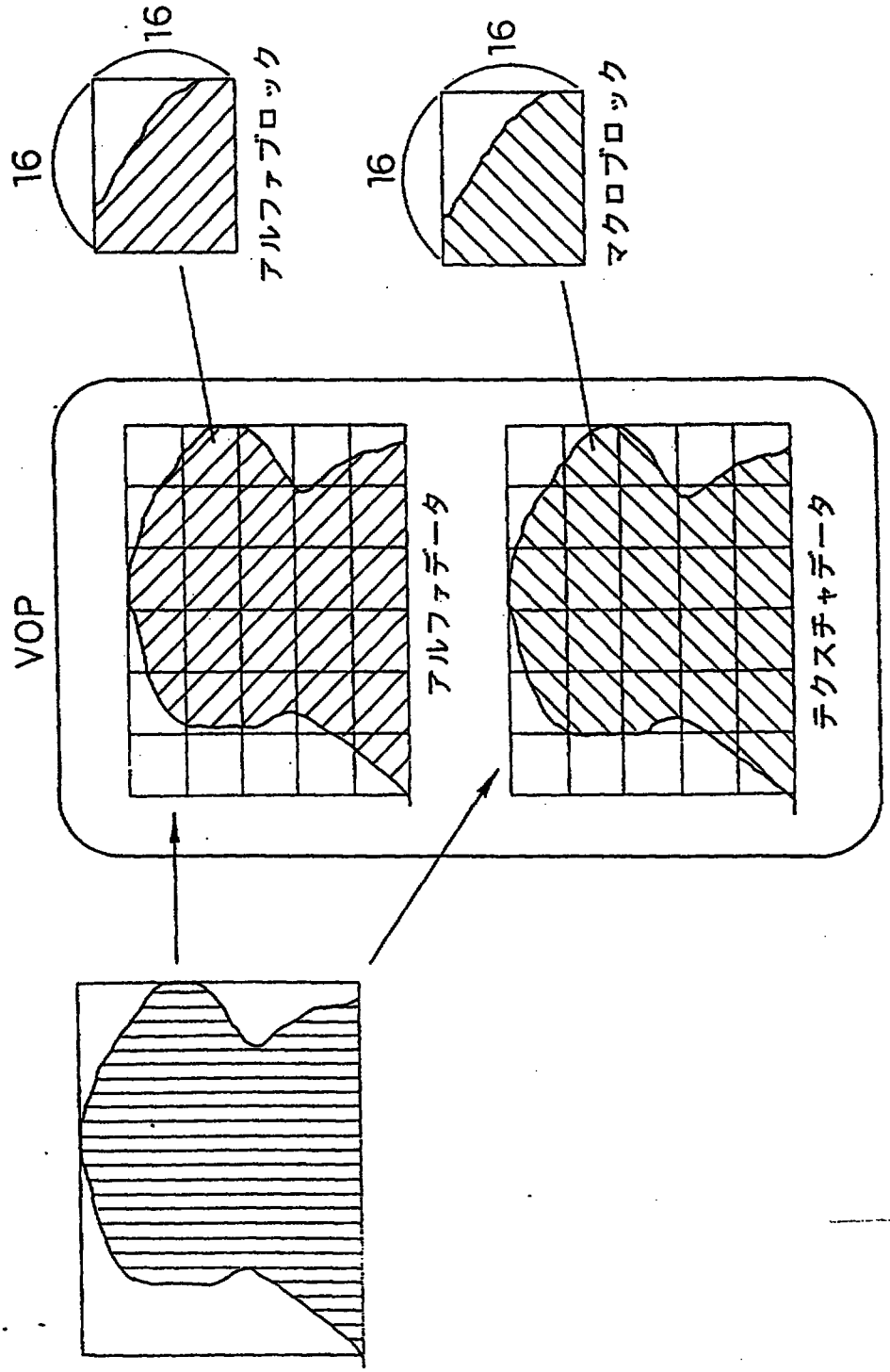


図31

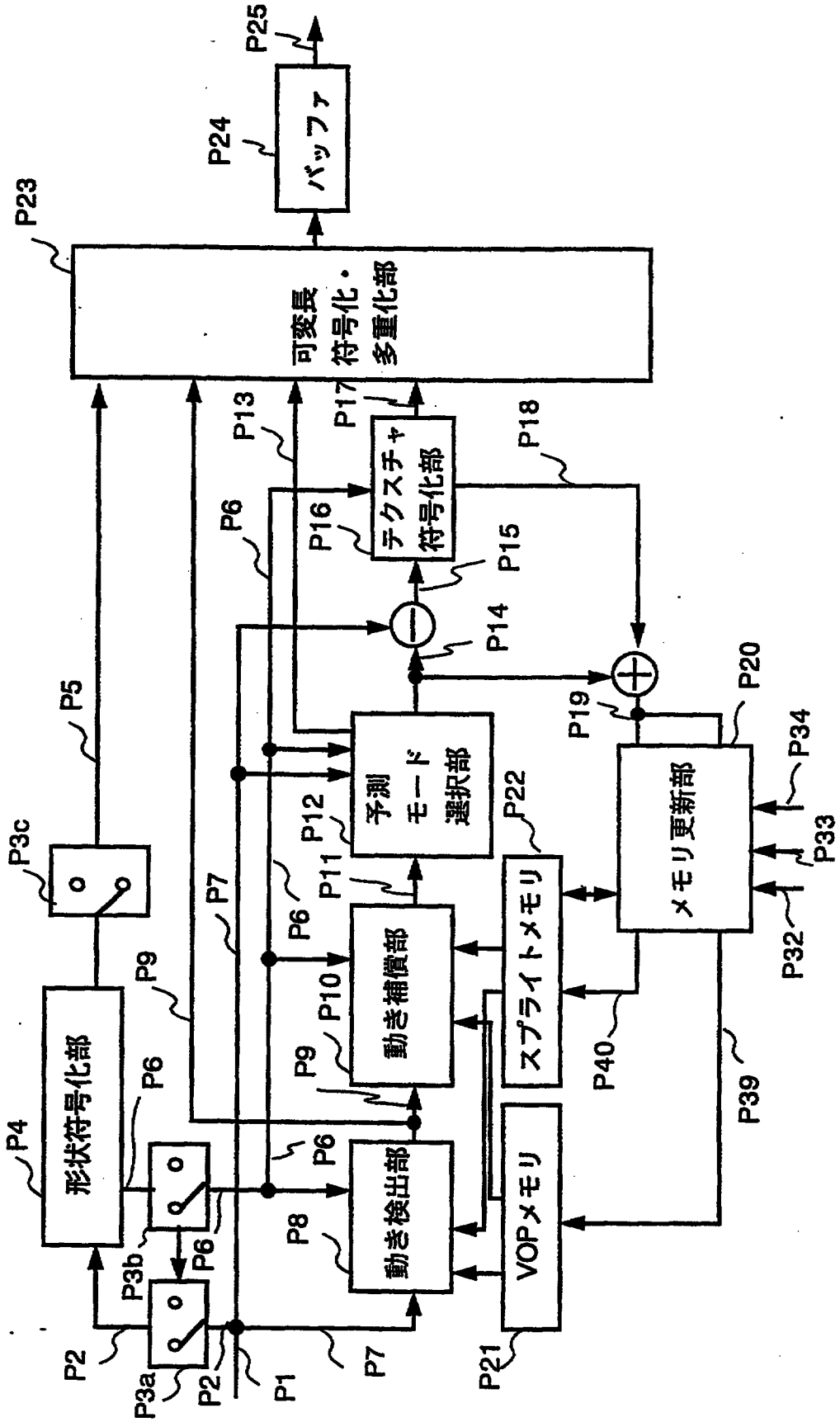

P25
P24
バッファ
P23
可変長符号化・多重化部
P13
P17
テクスチャ符号化部
P18
P6
P16
P15
P14
P5
予測モード選択部
P19
P20
P12
P22
P34
P3c
P11
メモリ更新部
P33
P7
P6
P32
P9
P10
動き補償部
スプライトメモリ
P4
形状符号化部
P6
P9
P40
P6
P6
P3b
P8
動き検出部
VOPメモリ
P39
P2
P7
P21
P3a
P2
P1

図32

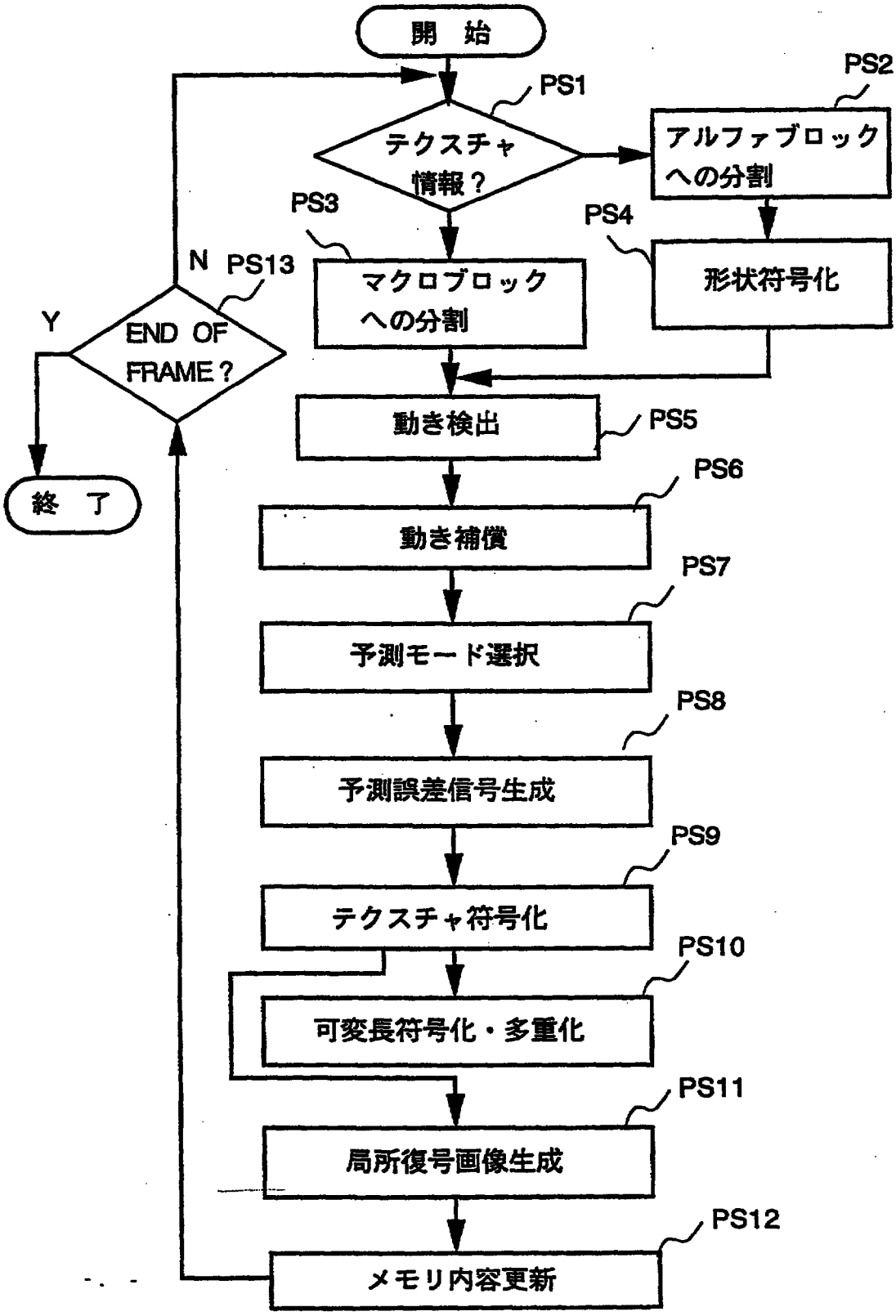

開 始
PS1
テクスチャ情報？
PS2
アルファブロックへの分割
PS3
マクロブロックへの分割
PS4
形状符号化
N
PS13
END OF FRAME？
Y
終 了
動き検出
PS5
PS6
動き補償
PS7
予測モード選択
PS8
予測誤差信号生成
PS9
テクスチャ符号化
PS10
可変長符号化・多重化
PS11
局所復号画像生成
PS12
メモリ内容更新

図33

| VOPタイプ | 予測なし | 過去のVOPからの予測 | 過去および未来のVOPからの予測 | スプライトから予測 |
|---|---|---|---|---|
| I-VOP | ○ | × | × | × |
| P-VOP | ○ | ○ | × | × |
| B-VOP | ○ | ○ | ○ | × |
| SPRITE-VOP | ○ | ○ | × | ○ |

図34

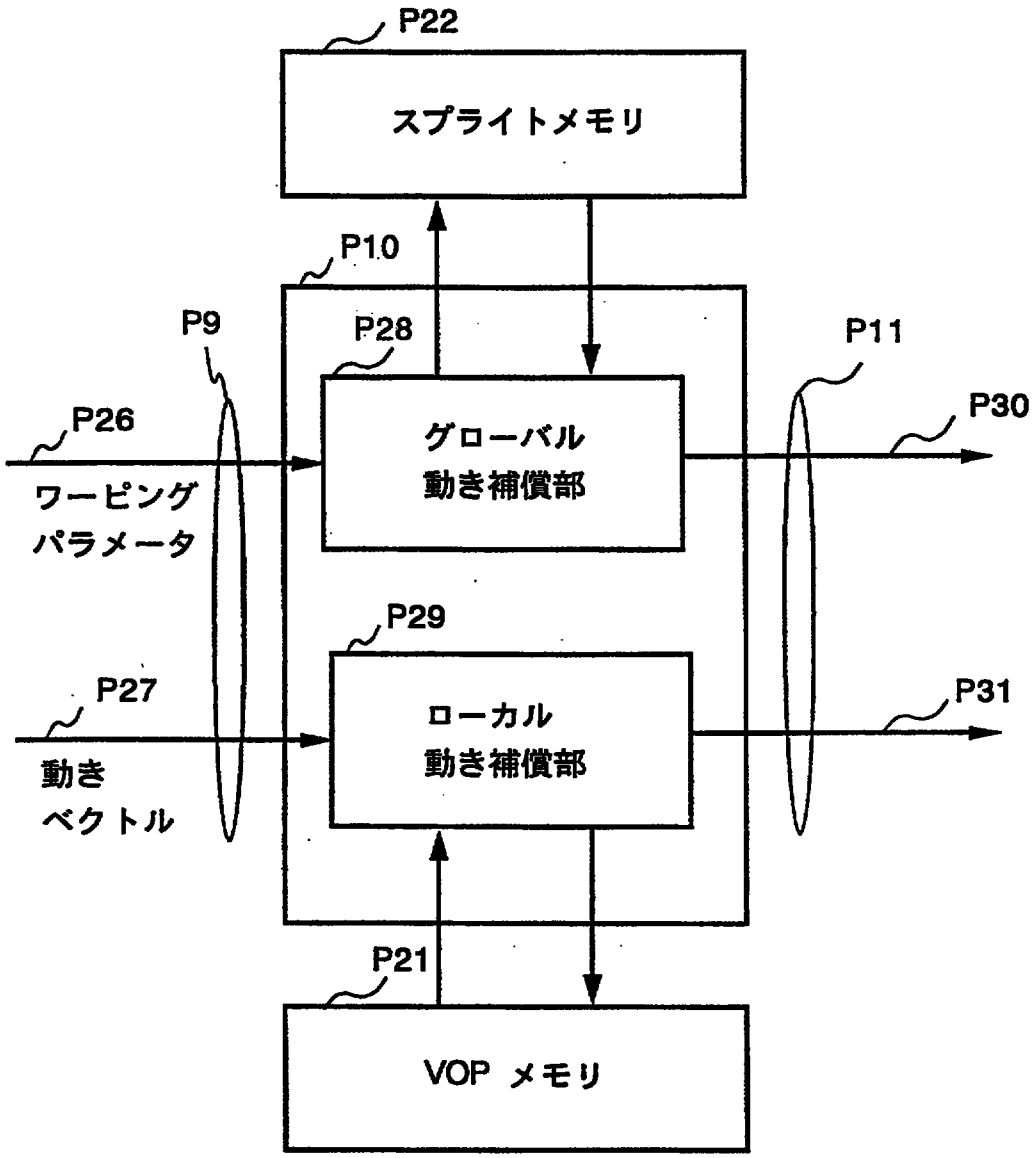

P22
スプライトメモリ
P10
P9
P28
P11
P26
グローバル
動き補償部
P30
ワーピング
パラメータ
P29
P27
ローカル
動き補償部
P31
動き
ベクトル
P21
VOP メモリ

図35

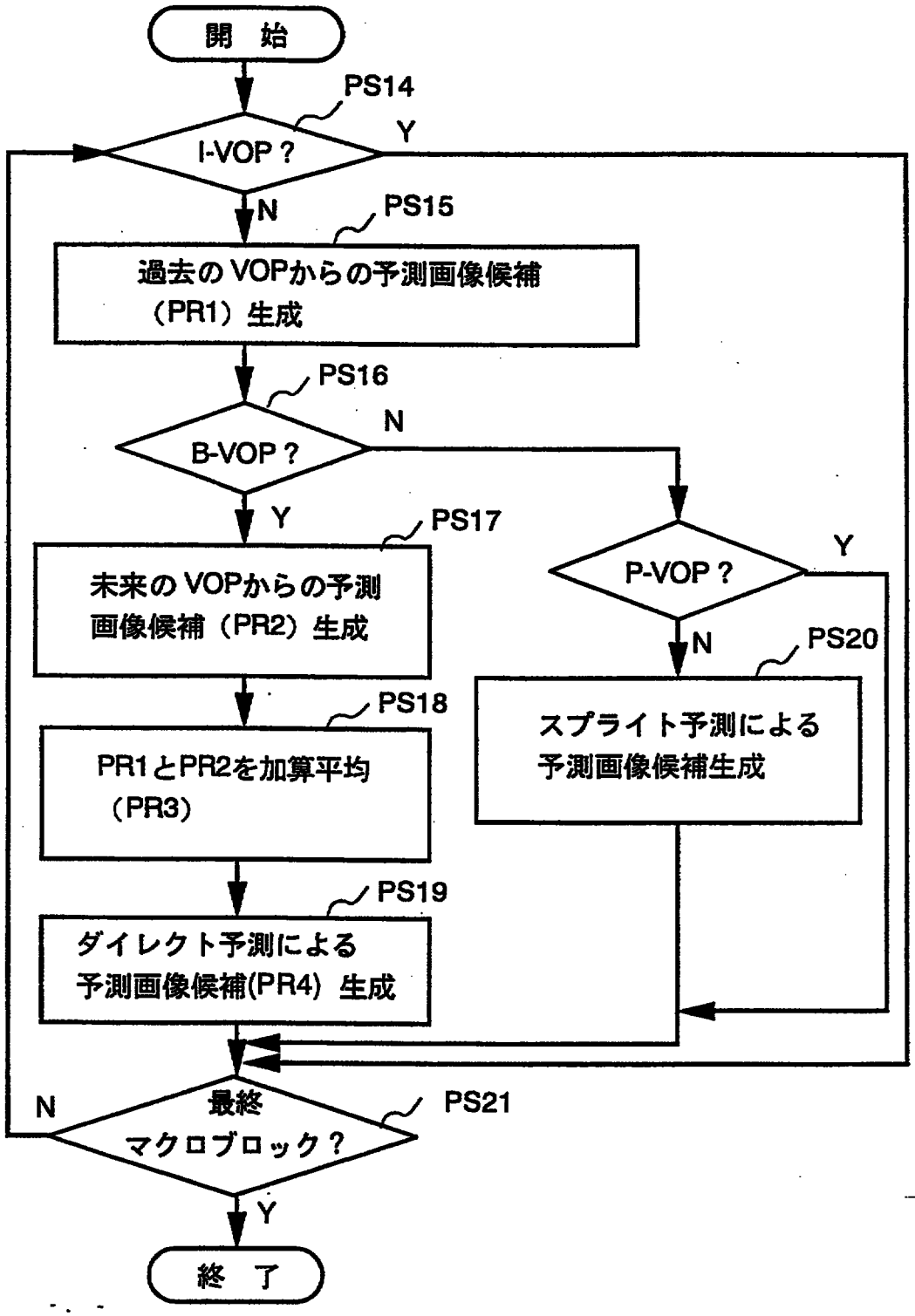

開 始
PS14
I-VOP ?
Y
N
PS15
過去のVOPからの予測画像候補（PR1）生成
PS16
B-VOP ?
N
Y
PS17
未来のVOPからの予測画像候補（PR2）生成
P-VOP ?
Y
N
PS20
スプライト予測による予測画像候補生成
PS18
PR1とPR2を加算平均（PR3）
PS19
ダイレクト予測による予測画像候補(PR4) 生成
N
最終マクロブロック ?
PS21
Y
終 了

図36

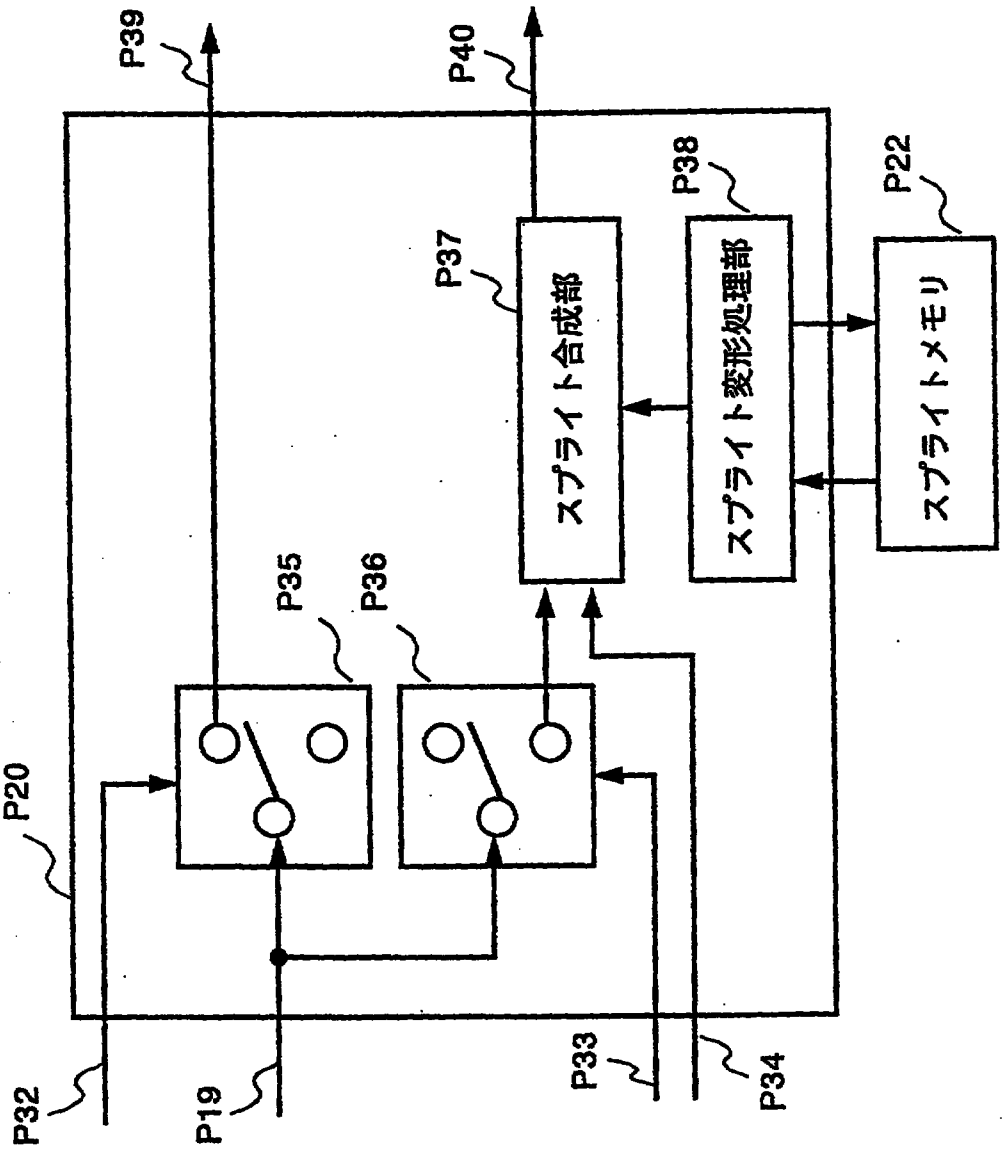
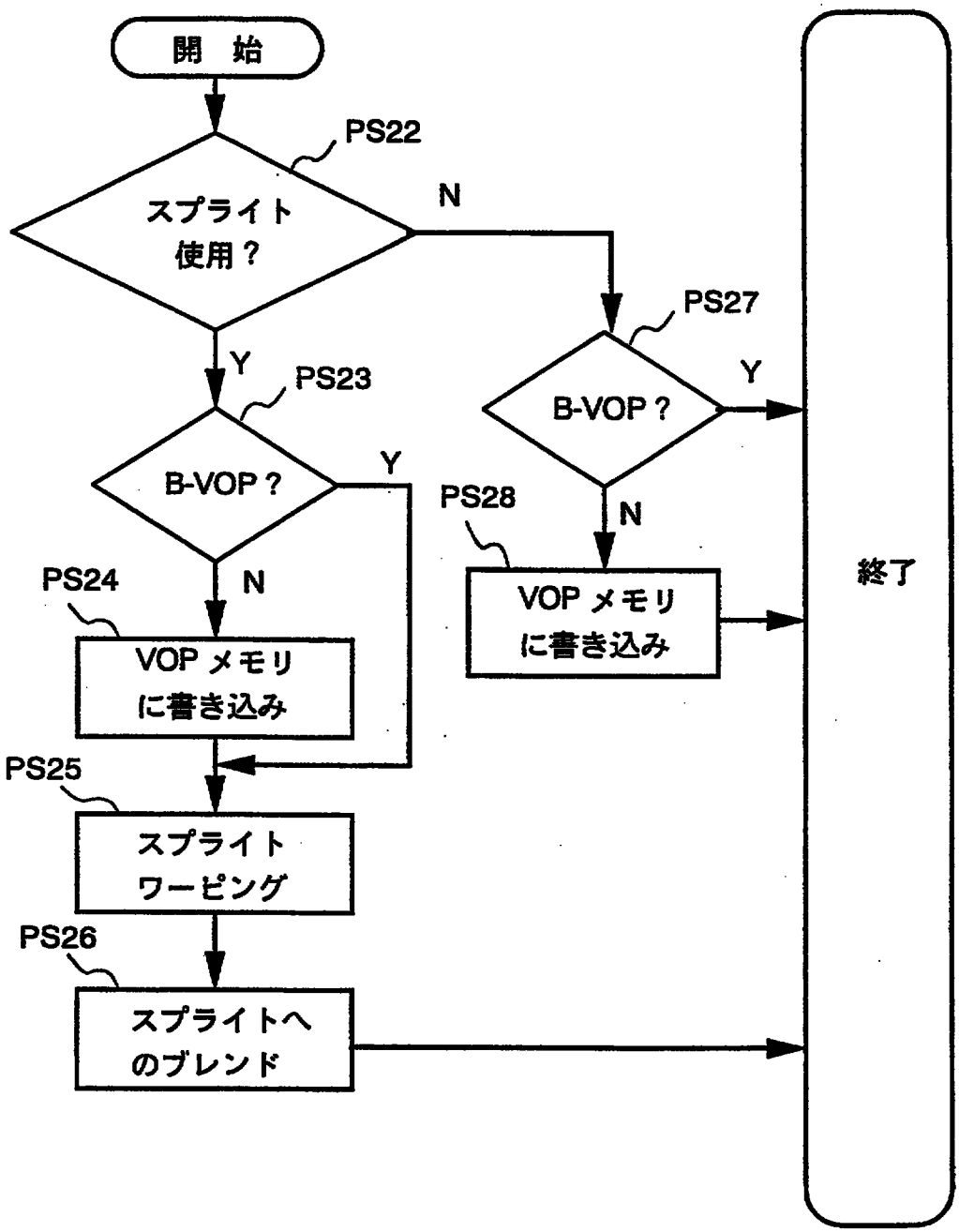

P39
P40
P38
P22
P37
スプライト合成部
スプライト変形処理部
スプライトメモリ
P35
P36
P20
P32
P19
P33
P34

図37


開 始
PS22
スプライト
使用？
N
Y
PS23
B-VOP？
Y
N
PS24
VOP メモリ
に書き込み
PS25
スプライト
ワーピング
PS26
スプライトへ
のブレンド
PS27
B-VOP？
Y
N
PS28
VOP メモリ
に書き込み
終了

# MOVING PICTURE PREDICTION SYSTEM

## Technical Field

The present invention relates to the prediction of a moving picture implemented, for example, in

a moving picture encoder/decoder used in a portable/stationary video communication device and the like for visual communications in a video telephone system, a video conference system or the like,

a moving picture encoder/decoder used in a picture storage/recording apparatus such as a digital VTR and a video server, and

a moving picture encoding/decoding program implemented in the form of a single software or a firmware as a Digital Signal Processor (DSP).

## Background Art

MPEG-4 (Moving Picture Experts Group Phase-4) Video Encoding/Decoding Verification Model (hereinafter referred to by the initials VM) whose standardization is in progress by ISO/IEC JTC1/SC29/WG11 may be introduced as a conventional type of predictive encoding/decoding in an encoding/decoding system of moving pictures. The VM continues to revise its contents according to the progress being made in standardization of MPEG-4. Here,

Version 5.0 of the VM is designated to represent the VM and will be simply referred to as VM hereinafter.

The VM is a system for encoding/decoding each video object as one unit in view of a moving picture sequence being an aggregate of video objects changing their shapes time-/space-wise arbitrarily. Fig. 29 shows a VM video data structure. According to the VM, a time-based moving picture object is called a Video Object (VO), and picture data representing each time instance of the VO, as an encoding unit, is called a Video Object Plane (VOP). If the VO is layered in time/space, a special unit called a Video Object Layer (VOL) is provided between the VO and the VOP for representing a layered VO structure. Each VOP includes shape information and texture information to be separated. If the moving picture sequence includes a single VO, then the VOP is equated to a frame. There is no shape information included, in this case, and the texture information alone is then to be encoded/decoded.

The VOP includes alpha data representing the shape information and texture data representing the texture information, as illustrated in Fig. 30. Each data are defined as an aggregate of blocks (alphablocks/macroblocks), and each block in the aggregate is composed of 16x16 samples. Each alphablock sample is represented in eight bits. A macroblock

THIS SHEET HAS BEEN REPLACED



includes accompanied chrominance signals being associated with 16x16 sample luminance signals. VOP data are obtained from a moving picture sequence externally processed outside of an encoder.

Fig. 31 is a diagram showing the configuration of a VOP encoder according to the VM encoding system. The diagram includes original VOP data P1 to be inputted, an alphablock P2 representing the shape information of the VOP, a switch P3a for passing the shape information, if there is any, of the inputted original VOP data, a shape encoder P4 for compressing and encoding the alphablock, compressed alphablock data P5, a locally decoded alphablock P6, texture data (a macroblock) P7, a motion detector P8, a motion parameter P9, a motion compensator P10, a predicted picture candidate P11, a prediction mode selector P12, a prediction mode P13, a predicted picture P14, a prediction error signal P15, a texture encoder P16, texture encoding information P17, a locally decoded prediction error signal P18, a locally decoded macroblock P19, a sprite memory update unit P20, a VOP memory P21, a sprite memory P22, a variable-length encoder/multiplexer P23, a buffer P24, and an encoded bitstream P25.

Fig. 32 shows a flowchart outlining an operation of the encoder.

Referring to the encoder of Fig. 31, the original

This Sheet Has Been Replaced

VOP data P1 are decomposed into the alphablocks P2 and the macroblocks P7 (Steps PS2 and PS3). The alphablocks P2 and the macroblocks P7 are transferred to the shape encoder P4 and the motion detector P8, respectively. The shape encoder P4 is a processing block for data compression of the alphablock P2 (step PS4), the process of which is not discussed here further in detail because the compression method of shape information is not particularly relevant to the present invention.

The shape encoder P4 outputs the compressed alphablock data P5 which is transferred to the variable-length encoder/multiplexer P23, and the locally decoded alpha data P6 which is transferred sequentially to the motion detector P8, the motion compensator P10, the prediction mode selector P12, and the texture encoder P16.

The motion detector P8, upon reception of the macroblock P7, detects a local-motion vector on a macroblock basis using reference picture data stored in the VOP memory P21 and the locally decoded alphablock P6 (step PS5). Here, the motion vector is one example of a motion parameter. The VOP memory P21 stores the locally decoded picture of a previously encoded VOP. The content of the VOP memory P21 is sequentially updated with the locally decoded picture of a macroblock whenever the

macroblock is encoded. In addition, the motion detector P8 detects a global warping parameter, upon reception of the full texture data of the original VOP, by using reference picture data stored in the sprite memory P22 and locally decoded alpha data. The sprite memory P22 will be discussed later in detail.

The motion compensator P10 generates the predicted picture candidate P11 by using the motion parameter P9, which is detected in the motion detector P8, and the locally decoded alphablock P6 (step PS6). Then, the prediction mode selector P12 determines the final of the predicted picture P14 and corresponding prediction mode P13 of the macroblock by using a prediction error signal power and an original signal power (step PS7). In addition, the prediction mode selector P12 judges the coding type of the data either intra-frame coding or inter-frame coding.

The texture encoder P16 processes the prediction error signal P15 or the original macroblock through Discrete Cosine Transformation (DCT) and quantization to obtain a quantized DCT coefficient based upon the prediction mode P13. An obtained quantized DCT coefficient is transferred, directly or after prediction, to the variable-length encoder/multiplexer P23 to be encoded (steps PS8 and PS9). The variable-length

encoder/multiplexer P23 converts the received data into a bitstream and multiplexes the data based upon predetermined syntaxes and variable-length codes (step PS10). The quantized DCT coefficient is subject to dequantization and inverse DCT to obtain the locally decoded prediction error signal P18, which is added to the predicted picture P14, and the locally decoded macroblock P19 (step PS11) is obtained. The locally decoded macroblock P19 is written into the VOP memory P21 and the sprite memory P22 to be used for a later VOP prediction (step PS12).

Dominant portions of prediction including a prediction method, a motion compensation, and the update control of the sprite memory P22 and the VOP memory P21 will be discussed below in detail.

(1) Prediction method in the VM

Normally, four different types of VOP encoding shown in Fig. 33 are processed in the VM. Each encoding type is associated with a prediction type or method marked by a circle on a macroblock basis. With an I-VOP, intra-frame coding is used singly involving no prediction. With a P-VOP, past VOP data can be used for prediction. With a B-VOP, both past and future VOP data can be used for prediction.

All the aforementioned prediction types are motion

vector based. On the other hand, with a Sprite-VOP, a sprite memory can be used for prediction. The sprite is a picture space generated through a step-by-step mixing process of VOPs based upon a warping parameter set

$$\vec{\alpha}=(a, b, c, d, e, f, g, h)$$

detected on a VOP basis (The mark → denotes a vector hereinafter). The warping parameter set is determined by the following parametric equations.

$$x'=(ax+by+c)/(gx+hy+1)$$

$$y'=(dx+ey+f)/(gx+hy+1)$$

The sprite is stored in the sprite memory P22.

Referring to the parametric equations, (x, y) represents the pixel position of an original VOP in a two-dimensional coordinate system. (x', y') represents a pixel position in the sprite memory corresponding to (x, y,) based upon a warping parameter. With the Sprite-VOP, the warping parameter set can be used uniformly with each macroblock to determine (x', y') in the sprite memory for prediction to generate a predicted picture. In a strict sense, the sprite includes "Dynamic Sprite" used for prediction and "Statistic Sprite" used for prediction as well as for another purpose of an approximate representation of VOP at a decoding station. In Figs. 34 through 37 below, "sprite" stands for Dynamic Sprite.

The motion detector P8 detects the motion vector and

the warping parameter to be used for the aforementioned prediction types. The motion vectors and the warping parameters are generically called the motion parameter P9 hereinafter.

(2) Motion compensation

Fig. 34 is a diagram showing the configuration of the motion compensator P10 in detail. In the figure, a warping parameter P26, a motion vector P27, a global-motion compensator P28, a local-motion compensator P29, a warping-parameter based predicted picture candidate P30, and a motion-vector based predicted picture candidate P31 are shown. The warping-parameter and motion-vector based predicted picture candidates 30, 31 are generically called the predicted picture candidates P11 hereinafter.

Fig. 35 shows a flowchart outlining the operation of the motion compensator P10 including steps PS14 through PS21.

The motion compensator P10 generates the predicted picture candidate P11 using the warping parameter P26 of a full VOP detected on a macroblock P7 basis in the motion detector P8 or a macroblock based motion vector P27. The global-motion compensator P28 performs a motion compensation using the warping parameter P26, and the local-motion compensator P29 performs a motion

compensation using the motion vector P27.

With the I-VOP, the motion compensator P10 does not operate. (The operating step proceeds to step PS21 from step PS14.) With a VOP other than the I-VOP, the local-motion compensator P29 reads out a predicted picture candidate PR1 from the locally decoded picture of a past VOP stored in the VOP memory P21 by using the motion vector P27 (step PS15). With the P-VOP, the predicted picture candidate PR1 is only available to be used.

When the B-VOP is identified in step PS16, the local-motion compensator P29 further reads out a predicted picture candidate PR2 from the locally decoded picture of a future VOP stored in the VOP memory P21 by using the motion vector P27 (step PS17). In addition, an arithmetic mean of the predicted picture candidates PR1, PR2 obtained from the past and future VOP locally decoded pictures to obtain a predicted picture candidate PR3 (step PS18).

A predicted picture candidate PR4 is generated also through Direct Prediction (step PS19). (Direct Prediction is based upon a prediction method corresponding to B-Frame in an encoding method H.263, Recommendation ITU-T. A vector for B-Frame is produced based upon a group of P-VOP vectors, which is not discussed further here in detail.) In Fig. 34, the

motion-vector based predicted picture candidates P31 is a generic term for all or part of the predicted picture candidates PR1 through PR4.

If a VOP is of neither I-VOP nor B-VOP, then the VOP is of Sprite-VOP. With the Sprite-VOP, the predicted picture candidate PR1 is read out from the VOP memory based upon the motion vector. In addition, the global-motion compensator P28 reads out the predicted picture candidate P30 from the sprite memory P22 based upon the warping parameter P26 in step PS20.

The global-motion compensator P28 calculates the address of a predicted picture candidate in the sprite memory P22 based upon the warping parameter P26, and reads out the predicted picture candidate P30 from the sprite memory P22 to be outputted based upon a resultant address. The local-motion compensator P29 calculates the address of a predicted picture candidate in the VOP memory P21 based upon the motion vector P27 and reads out the predicted picture candidate P31 to be outputted based upon a resultant address.

These predicted picture candidates P11 are evaluated along with an intra-frame coding signal of the texture data P7 in the prediction mode selector P12, which selects a predicted picture candidate having the least power of a prediction error signal along with a

THIS SHEET HAS BEEN REPLACED



prediction mode.

(3) Updating of memories

The memory update unit P20 controls the VOP memory P21 and sprite memory P22 to be updated (step PS12). The contents of these memories are updated regardless of the prediction mode P13 selected on a macroblock basis.

Fig. 36 is a diagram showing the configuration of the memory update unit P20. Fig. 37 shows a flowchart including steps PS22 through PS28 illustrating the operation of the memory update unit P20.

In Fig. 36, an externally supplied VOP encoding type P32, an externally supplied sprite prediction identification flag P33 for indicating the use of the sprite memory for prediction, an externally supplied blend factor P34 used for prediction with the sprite memory, switches P35, P36, a sprite blender P37, a sprite transformer P38, a VOP memory update signal P39, and a sprite update signal P40 are shown.

Firstly, the use of the sprite with the current VO or VOL is examined if being designated by the sprite prediction identification flag P33 (step PS22). With no use of the sprite designated, the data are examined if being the B-VOP (step PS27). With the B-VOP, then no updating is performed with the VOP memory P21. With either the I-VOP or the P-VOP, then the VOP memory P21 is

This Sheet Has Been Replaced



written over with the locally decoded macroblock P19 on a macroblock basis (step PS28).

With the use of the sprite designated in step PS22, then the VOP memory P21 is updated in the same manner as above (steps PS23, PS24), and in addition, the sprite memory PS22 is updated through the following procedure.

a) Sprite warping (step PS25)

In the sprite transformer P38, an area

$M(\vec{R},t-1)$

in the sprite memory P22 ($M(\vec{R},t-1)$ is an area having the same size as that of a VOP having the origin of the coordinates at a position in the sprite memory P22 with the VOP at a time t) is subject to warping (transformation) based upon a warping parameter

$\vec{\alpha}=(a,b,c,d,e,f,g,h)$.

b) Sprite blending (step PS26)

By using a resultant warped picture from a) above, a new sprite memory area is calculated in the sprite blender P37 according to the following expression,

$$M(\vec{R},t)=(1-\alpha)\cdot W_b[M(\vec{R},t-1),\vec{\alpha}]+\alpha\cdot VO(\vec{r},t),$$

where $\alpha$ is the blend factor P34, $W_b[M,\vec{\alpha}]$ is the resultant warped picture, and $VO(\vec{r},t)$ is a pixel value of a locally decoded VOP with a location $\vec{r}$ and a time t.

With a non-VOP area in a locally decoded macroblock, it is assumed that

$VO(\vec{r},t)=0$.

As the blend factor $\alpha$ is assigned on a VOP basis, a locally decoded VOP is collectively blended into the sprite memory P22 based upon a weight $\alpha$, regardless of the contents of a VOP area.

According to the aforementioned prediction system in the conventional encoding system, the video object is predicted by using the memory designed to be used for detecting the motion vector alone and the memory designed to be used for detecting the warping parameter alone, both of which are structurally allowed the maximum use of a single screen alone each. Thus, the limited use of reference pictures is only available for prediction, thereby hindering a sufficient improvement in prediction efficiency.

Further, in such a system where two or more video objects are encoded concurrently, these memories only include a reference picture representing the past record of a video object to be predicted alone, which limits the variation of a reference picture and precludes the utilization of a correlation among video objects for prediction.

Further, the memories are updated regardless of such items as the internal structure, a characteristic, and the past record of the video object. This results in the insufficient storage of information lacking significant

THIS SHEET HAS BEEN REPLACED



data for predicting a video object, thereby posing a problem of failing to enhance prediction efficiency.

The present invention is directed to solving the aforementioned problems. An objective of this invention is to provide the prediction system for encoding/decoding of picture data where two or more memories are provided to store the past record of the moving picture sequence effectively in consideration of the internal structure and characteristic of the moving picture sequence, thereby achieving a highly efficient prediction as well as encoding/decoding. In addition, the prediction system provides a sophisticated inter-video object prediction performing among two or more video objects.

## Disclosure of the Invention

According to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoder and a decoder, includes a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being corresponding to different transform methods, respectively, and a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED



stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories.

The encoder generates a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and transmits the prediction memory indication information signal and the parameter to a decoding station so as to generate the predicted picture using the picture data stored in the one of the plurality of memories based upon the one of the transform methods corresponding to the one of the plurality of memories in the decoding station.

The decoder receives the parameter and a prediction memory indication information sinal indicating the one of the plurality of memories used for generating the predicted picture from an encoding station, wherein the prediction picture generation section generates the predicted picture using the picture data stored in the one of the plurality of memories based upon the parameter and the one of the transform methods corresponding to the one of the plurality of memories.

Further, according to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoding

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED



and a decoding, includes a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being assigned to different parameter effective value ranges, respectively, and a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, for selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and for generating a predicted picture using the picture data stored in a selected memory.

Still further, according to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes a plurality of memories for storing picture data for reference to be used for prediction and a prediction picture generation section including a motion compensator for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture by using the picture data stored in the plurality of memories based upon the parameter, and a memory update unit for updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

The moving picture prediction system predicts the

This Sheet Has Been Replaced



moving picture in a moving picture sequence having first and second video objects, wherein the plurality of memories includes separate first and second pluralities of memories corresponding to the first and second video objects, respectively, and the prediction picture generation section includes separate first and second generators, respectively, corresponding to the first and second video objects, wherein the first generator uses the picture data stored in at least one of the first and second pluralities of memories to generate the predicted picture when predicting the first object, and generates information indicating a use of the second plurality of memories for predicting the first object, the information being added to the predicted picture.

The prediction picture generation section generates the predicted picture through a change of either one of a number and a size of the plurality of memories in response to a change in the moving picture at each time instance.

The prediction picture generation section generates the predicted picture in a limited use of memories for prediction in response to a change in the moving picture at each time instance.

The prediction picture generation section generates the predicted picture by calculating a plurality of the

predicted pictures generated by using the respective picture data stored in the plurality of memories.

The moving picture prediction system further includes a significance detector for detecting a feature parameter representing a significance of the picture segment to be predicted, wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

The moving picture prediction system further includes a significance detector for detecting a parameter representing at least one of an amount of bits available for coding the picture segment to be predicted, an amount of change of the picture segment at each time instance, and a significance of the picture segment, wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, a plurality of memory update methods.

The moving picture prediction system predicts the moving picture on a video object basis, wherein the moving picture prediction system further includes a significance detector for detecting a parameter

representing at least one of an amount of bits available for coding a video object to be predicted, an amount of change in the video object at each time instance, and a significance of the video object, wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

The moving picture prediction system further includes a prediction information encoder for encoding prediction relating information of the moving picture, wherein the prediction picture generation section counts times of a memory used for prediction and determines a rank of the plurality of memories based upon a counted number of the times, wherein the prediction information encoder allocates a code length to the prediction relating information to be encoded based upon the rank of a memory used for prediction.

The plurality of memories includes at least a frame memory for storing the picture data on a frame basis and a sprite memory for storing a sprite picture.

The sprite memory includes at least one of a dynamic sprite memory involving a regular updating, and a static sprite memory not involving the regular updating.

The one of the transform methods corresponding to

THIS SHEET HAS BEEN REPLACED



the one of the plurality of memories is at least one of a parallel translation, an affine transformation, and a perspective transformation in an interchangeable manner.

Still further, according to the present invention, a method for predicting a moving picture to be implemented in at least one of an encoding or a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, corresponding different transform methods with the plurality of memories, respectively, receiving a parameter representing a motion of a picture segment to be predicted, and generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods being corresponding to the one of the plurality of memories.

The method for predicting a moving picture further includes the steps of generating a prediction memory indication information signal indicating the one of the plurality of memories used for the picture segment to be predicted, and transmitting the prediction memory indication information signal and the parameter to a decoding station.

The method for predicting a moving picture is implemented in the decoding, and further includes the

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED



step of receiving a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and the parameter representing a motion of the picture segment to be predicted from an encoding station.

Still further, according to the present invention, a method, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, assigning separate parameter effective value ranges to the plurality of memories, respectively, receiving a parameter representing a motion of a picture segment to be predicted, selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and generating a predicted picture using the picture data stored in a selected memory.

Still further, according to the present invention, a method, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, receiving a parameter representing a motion of a picture segment to be predicted, generating a predicted picture

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED



using the picture data stored in the plurality of memories based upon the parameter, and updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

Brief description of drawings

Fig. 1 is a structural diagram of a moving picture encoder according to an embodiment of this invention.

Fig. 2 is a flowchart illustrating an operation of the moving picture encoder according to the embodiment of this invention.

Fig. 3 is a structural diagram illustrating the configuration of a motion compensator of the moving picture encoder of the embodiment of this invention.

Fig. 4 is a flowchart illustrating an operation of the motion compensator.

Fig. 5 is a structural diagram illustrating the structure of a memory update unit of the moving picture encoder of the embodiment of this invention.

Fig. 6 is a flowchart illustrating an operation of the memory update unit.

Fig. 7 is a structural diagram illustrating the configuration of a motion compensator of a moving picture encoder according to another embodiment of this invention.

This Sheet Has Been Replaced

Fig. 8 is a flowchart illustrating an operation of the motion compensator of Fig. 7.

Fig. 9 is a structural diagram illustrating the configuration of a motion compensator of a moving picture encoder according to another embodiment of this invention.

Fig. 10 is a flowchart illustrating an operation of the motion compensator of Fig. 9.

Fig. 11 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 12 is a structural diagram showing the configuration of a motion compensator of the moving picture encoder according to the embodiment of this invention.

Fig. 13 is a flowchart illustrating an operation of the motion compensator of Fig. 12;

Fig. 14 is a structural diagram illustrating the configuration of a memory update unit of a moving picture encoder according to another embodiment of this invention.

Fig. 15 is a flowchart illustrating an operation of the memory update unit of Fig. 14.

Fig. 16 is a structural diagram of a moving picture encoder according to another embodiment of this

invention.

Fig. 17 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 18 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 19 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 20 is a diagram showing bit stream 21 according to the first embodiment of this invention.

Fig. 21 is a diagram showing bit stream 21 according to the second embodiment of this invention.

Fig. 22 is a diagram showing bit stream 21 according to the third embodiment of this invention.

Fig. 23 is a diagram showing bit stream 21 according to the sixth embodiment of this invention.

Fig. 24 is a structural diagram of a moving picture decoder according to another embodiment of this invention.

Fig. 25 is a structural diagram illustrating the configuration of a motion compensator of the moving picture decoder according to the embodiment of this invention.

Fig. 26 is a flowchart illustrating an operation of the motion compensator.

Fig. 27 is an exemplary diagram of interpolation.

Fig. 28 is a flowchart illustrating an operation of a memory update unit of the moving picture decoder according to the embodiment of this invention.

Fig. 29 is an exemplary diagram of a video data configuration according to the VM encoding system.

Fig. 30 is an exemplary diagram of a VOP data structure.

Fig. 31 is a structural diagram illustrating the configuration of a VM encoder.

Fig. 32 is a flowchart illustrating an operation of the encoder of Fig. 31.

Fig. 33 is an exemplary diagram of VOP encoded types and corresponding prediction types.

Fig. 34 is a structural diagram illustrating the configuration of a motion compensator of the encoder of Fig. 31.

Fig. 35 is a flowchart illustrating an operation of the motion compensator of Fig. 34.

Fig. 36 is a structural diagram illustrating the configuration of a memory update unit of the encoder of Fig. 31.

Fig. 37 is a flowchart illustrating an operation of the memory update unit of Fig. 36.

Best mode for carrying out the invention

Embodiment 1.

Fig. 1 is a block diagram showing the configuration of an encoder according to a first embodiment and the following embodiments. The diagram illustrates an input moving picture signal 1, texture data 2, a motion detector 3, a motion parameter 4, a motion compensator 5, a predicted picture candidate 6, a prediction mode selector 7, a prediction mode 8, a predicted picture 9, a prediction error picture 10, a texture encoder 11, a quantized DCT coefficient 12, a locally decoded prediction error picture 13, a locally decoded picture 14, a memory update unit 15, a memory-a 16, a memory-b 17, a memory-c 18, a variable-length encoder/multiplexer 19, a transmission buffer 20, a bitstream 21, a scene-change detector 80, and a timer 81. Particularly, the motion compensator 5 and the memory update unit 15 forms a predicted picture generation section 100 which implemented a prediction system. Memories a, b, c forms a memory area 200. In the figure, portions not mentioned in this embodiment will be discussed in the following embodiments. Fig. 2 shows a flowchart illustrating an

operating flow of the encoder.

This embodiment is based upon the assumption that a plural number, e.g. three, of memories, are used adaptively according to the significance of an input moving picture based upon such characteristics of motion as an amount of motion and an intensity of color. It is also assumed that the content of an arbitrary memory (area), the memory-a for example, are updated at an arbitrary period of time and a moving picture sequence is received on a frame basis.

(1) Input signal

As aforementioned, the encoder inputs a frame representing a picture at each time instance of the moving picture sequence and decomposes the frame into the encoding units of macroblocks which are one example of picture segments subject to prediction (step S1).

(2) Adaptive use of memories

The memories store previously decoded pictures or previously provided fixed pictures. In this embodiment, the three memories are used adaptively according to the significance of a picture segment in a frame as follows.

The memory-a stores a least significant picture segment (i.e., a background-like picture segment whose motion is static or flat and the texture is flat.)

The memory-b stores a less significant picture

segment (i.e., a picture segment of an object whose motion is relatively small.)

The memory-c stores a most significant picture segment (i.e., a picture segment of an object whose motion is complicated or drastic.)

The least significant picture segment to be stored in the memory-a may be a background picture segment in a video conference scene or the like. The least significant picture segment also corresponds to a background segment in a camera-work relating flat motion scene of a full screen including slightly moving objects. With this type of motion, it is efficient to obtain a frame based amount of a motion to substitute for a macroblock based motion, rather than to obtain a macroblock based amount of a motion. Specifically, a transform parameter corresponding to the sprite warping parameter discussed in the conventional art is obtained and the transform parameter of a full frame is then used as the motion parameter of a macroblock in the frame. The motion parameter may be selected from among a simple parallel translation parameter (=a motion vector), an affine motion parameter involving transformation, and a perspective motion parameter involving transformation. Here, a motion vector is one example of the motion parameter.

The less significant picture segment to be stored in the memory-b may be a picture segment of a moving figure who is not a speaker in a video conference scene or the like. This segment type of the object may be considered less attracting in the scene. The most significant picture segment to be stored in the memory-c may be a segment of an object attracting most attention in the video conference scene such as a speaker.

A picture segment stored in the memory-b or the memory-c representing a unique type of motion of an object should have a macroblock based unique motion parameter. The motion parameter of this case may be selected from among the simple parallel translation parameter (=a motion vector), the affine motion parameter involving transformation, the perspective motion parameter involving transformation, etc.

(3) Motion detection (step S2)

The motion detector 3 of this embodiment is designed to detect an arbitrary transform parameter of the respective three memories on a macroblock basis, involving no distinction between a motion vector and a warping parameter in the conventional art. The motion detector 3 is provided with additional functions of a global-motion parameter detection for detecting a frame based transform parameter using the memory-a and a local-

motion parameter detection for detecting a macroblock based transform parameter using the memories a through c.

(4) Motion compensation (step S3)

Fig. 3 shows the configuration of the motion compensator 5 of this embodiment in detail. In the figure, a prediction picture memory address calculator 22, a prediction picture memory address 23, a memory reader 24, and a reference memory indicator signal 25 which is supplied externally are shown. In this embodiment, the reference memory indicator signal 25 indicates the use of the memory a, b, c. Fig. 4 shows a flowchart including steps S11 through S16 illustrating an operation of the motion compensator 5.

Initially, with an I(Intra)-frame, no motion compensation is performed (step S11). With a frame other than the I-frame, predicted picture candidates are generated based upon the global-motion and local-motion parameters corresponding to the respective memories detected in the motion detector 3 (steps S12 through S15). Specifically, the prediction picture memory address calculator 22 calculates the prediction picture memory address 23 of a predicted picture candidate in a memory identified by the reference memory indicator signal 25 based upon the motion parameter 4. Upon reception of the prediction picture memory address 23,

the memory reader 24 reads out the predicted picture candidate 6 from a corresponding memory to be outputted.

In this embodiment, the global-motion and local-motion parameters are obtained through the same transform method, thereby allowing the motion compensator 5 of Fig. 3 to be shared by both global-motion and local-motion parameter based approaches of generating a predicted picture. When generating the predicted picture candidate 6 through the global-motion parameter, the memory-a is always used as a reference memory (step S15).

(5) Prediction mode selection (step S4)

The prediction mode of this embodiment is assumed to include the following.

(a) a mode for using the memory-a,

(b) a mode for using the memory-b,

(c) a mode for using the memory-c, and

(d) a mode for using an intra-frame coding signal.

Similarly to the discussion in the conventional art, the prediction mode selector 7 selects the predicted picture candidate 6 having the least power (amplitude) of a prediction error signal, for example, from among all the predicted picture candidates 6 generated in the motion compensator 5 along with an intra-frame coding signal, and outputs a selected one of the predicted picture candidates 6 as the predicted picture 9 and an

corresponding one of the prediction mode 8. The prediction mode 8 includes memory selection information indicating a memory used for predicting the selected predicted picture 9. The prediction mode 8 is transferred to the variable-length encoder/multiplexer 19 to be encoded with an allocated length of code in the bitstream 21 as prediction memory indication information 800.

(6) Memory updating

The memory update unit 15 controls the memories to be updated. Fig. 5 shows the configuration of the memory update unit 15 of this embodiment in detail. In the figure, an activity 26 used for updating the memory-a (which will be discussed later), a memory-a update judger 27, a reference memory selector 28, switches 29, 30, picture data 31 for updating the memory-a, picture data 32 for updating the memory-b, picture data 33 for updating the memory-c, and global prediction picture data 34 for updating the memory-a are shown. Fig. 6 shows a flow of a memory updating operation.

The memory updating operation of this embodiment has the following procedure. Upon reception of the locally decoded picture 14, the memory update unit 15 judges the necessity of updating the memory-a with a frame including the locally decoded picture 14 in the memory-a update

judger 27 (step S17). The reference memory selector 28 selects a memory used for predicting the locally decoded picture based upon the prediction mode 8 (steps S18, S19). Then, a reference picture stored in a selected memory is updated with one of the picture data 31, 32, 33 for updating the memory-a, the memory-b, the memory-c, respectively, and the global prediction picture data 34 for updating the memory-a of the locally decoded picture 14 based upon the following rule. A memory is assumed to be updated each frame on a prediction unit (macroblock) basis.

(1) Frame based regular updating of the memory-b and the memory-c (steps S20, S21):

The locally decoded picture 14 is written into either the memory-b or the memory-c used for predicting the picture.

(2) Frame based adaptive updating of the memory-a (steps S22, S23):

The locally decoded picture 14 is written into the memory-a used for predicting the picture for an arbitrary frame only or at an arbitrary period of time based upon a memory-a update judgement 1000 obtained in step S17.

The content of memory-a is the past record of a time-unvarying picture segment such as a background picture. This removes the necessity of the regular

updating of the content of memory unless a full-screen involved complicated or drastic movement such as a scene change occurs to cause a drastic change in the content of a picture segment.

As aforementioned, a frame based regular updating is performed with a complicated or drastic area of an object, whereas a longer-term based updating is performed with the content of the memory-a, thereby achieving an effective prediction with a background picture half visible among moving objects.

Viewed in this light, the memory-a is updated in an arbitrary period of time in this embodiment. Specifically, possible arbitrary update criteria are as follows.

a. A full-screen content is updated all at once with a global-motion parameter indicating a complicated or drastic motion, whereas no updating operation is performed with the parameter indicating a rather static motion.

b. A full-screen content is updated all at once on a predetermined period basis, regardless of a frame based period of time.

c. A full-screen content is updated all at once only with a frame immediately after a scene change detected.

In this embodiment, data as the arbitrary update

criteria are generically called as the activity 26 used for updating memory-a. Initially, the memory-a update judger 27 judges whether to update the content of memory-a based upon the activity 26 (step S17). Specifically, the activity 26 corresponds to a value of the global-motion parameter detected in the motion detector 3 with the arbitrary update criterion-a, a time stamp of the current frame from the timer 81 with the arbitrary update criterion-b, and a flag indicating a scene change detection outputted from the scene-change detector 80 with the arbitrary update criterion-c.

When the content of memory-a is judged to be updated, the content of the locally decoded picture 14 is outputted as the global prediction picture data 34 for updating the content of the memory-a (step S23). When no updating is judged with the content of the memory-a, then no updating is performed with the memory-a.

The memory-a update judgement 1000 of a frame is multiplexed in the bitstream 21 to be transmitted to a decoding station so that the same updating of the memory-a can be performed with the frame in the decoding station.

Fig. 20 is a diagram illustrating the bitstream 21 of this embodiment.

Fig. 20 is a conceptual diagram showing how frame

data are encoded in sequence to be transmitted. Each frame data is provided with header information at the front as a frame based additional information. The header information has the memory-a update judgement 1000 multiplexed to be transmitted to a decoding station. The header information is followed by the component macroblock data of the frame. The macroblock data include the prediction memory indication information 800 indicating a memory used for predicting the macroblock data. In a counterpart decoder, a memory for predicting a predicted picture is specified based upon the prediction memory indication information 800 of macroblock data to generate the predicted picture.

Although not shown in the figures, the memory-b update information or/and the memory-c update information may be transmitted to the decoding station along with, or alternatively to, the memory-a update judgement 1000.

The aforementioned encoder thus provides the adaptive and efficient use of two or more memories in response to the content of a moving picture sequence, thereby enhancing prediction efficiency. Specifically, a moving picture sequence is predicted based upon an arbitrary transform parameter through an adaptive use of two or more memories in response to the content and characteristic of a moving picture sequence. This

enables an efficient prediction of a moving picture in response to a local characteristic of a picture by even covering complicated motion. With the enhanced prediction efficiency, the encoder is allowed to reduce an amount of encoded data without deteriorating encoded picture quality. The same prediction system may be employed by the counterpart decoder for decoding a bitstream encoded through the prediction system of this invention.

This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

Further, this embodiment has thus disclosed the encoder performing on a macroblock basis as the macroblock being a picture segment subjected to prediction. The same effect can be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-

motion parameter detection. With no global-motion detection, no necessity occurs for transmitting prediction information indicating a global/local prediction as the prediction mode.

Further, this embodiment may include a special memory for prediction which stores reference picture data previously generated based upon the content of the moving picture sequence. The special memory is not updated during an encoding operation.

Further, this embodiment has thus disclosed the case that memories a, b, c store a picture segment each time and the memory update unit 15 updates one of the memories a, b, c each time. If two or all of the memories a, b, c share to store a picture in part or fully, then the memory update unit 15 updates the two or all of the memories a, b, c. In the case of the memory-a being a frame memory for storing a frame of reference picture data, the memory-b being a static sprite memory involving the adaptive updating, and the memory-c being a dynamic sprite memory involving the regular updating, the memory update unit 15 does not update the memory-b as the static sprite memory for storing previously fixed reference picture data but updates the memory-a and the memory-c concurrently when the memories store the same reference picture segment. Thus, if a duplicated storage of the

reference picture data occurs with the memories a, b, c, then the memory update unit 15 updates a duplicated segment stored in each memory.

The aforementioned can also be applied to the following embodiments.

Further, this embodiment has thus disclosed the use of three memories a, b, c, but alternatively, two of the memories may be utilized.

Further, a counterpart decoder may be provided with the prediction picture generation section 100 including the same components as the motion compensator 5 and the memory update unit 15 discussed in this embodiment. A motion compensator provided in the decoder, having no necessity of generating all the three predicted picture candidates, generates a single predicted picture alone based upon a decoded motion parameter.

Embodiment 2.

A second embodiment shows an encoder with a single replacement of the motion compensator 5 of the encoder shown in Fig. 1. The configuration and operation of a motion compensator 5a of the second embodiment are now described.

Fig. 7 shows the configuration of the motion compensator 5a of this embodiment in detail. The figure

includes a reference memory determiner 35. Fig. 8 shows a flowchart illustrating a detailed operation of the motion compensator 5a.

Initially, with the I-frame, no compensation is performed (step S24). With a frame other than the I-frame, the reference memory determiner 35 determines a reference memory based upon a value of the motion parameter 4 (step S25). The reference memory determiner 35 holds effective motion parameter value ranges (which will be discussed later in detail) allocated, respectively, to the memories a, b, c. The reference memory determiner 35 compares the respective effective motion parameter value ranges with the value of the motion parameter 4 to judge which memory is designated by the motion parameter 4 and outputs a reference memory indicator signal 25a for identifying the respective memories a, b, c.

The effective motion parameter value ranges are effective search ranges allocated to the respective memories for detecting a motion vector, for example. Specifically, if ±15 pixels are assumed to be given for a total search value range, then the memory-a is chosen to be used for prediction in a range of ±0 to 3 pixels, the memory-b is used in a range of ±4 to 8 pixels, and the memory-c is used in a range of ±9 to 15 pixels, for

example. Here, the reference memory determiner 35 operates only when the local-motion parameter is used for prediction, because the memory-a is exclusively used as a reference memory when the predicted picture candidate is generated based upon the global-motion parameter. This motion-vector value based approach of identifying a memory to be used for prediction is based upon the assumption that a background picture should include a static motion and a most attractive picture should include a complicated or drastic motion. This motion-vector value based approach of identifying a memory for prediction involves no necessity of encoding the prediction mode to be transmitted.

Next, the predicted picture candidate 6 is generated based upon the reference memory indicator signal 25a of a selected memory (steps S26 through S30). Specifically, the prediction picture memory address calculator 22 calculates the prediction picture memory address 23 of the predicted picture candidate 6 in an identified memory by the reference memory indicator signal 25a based upon the motion parameter 4. Based upon the prediction picture memory address 23, the memory reader 24 reads out the predicted picture candidate 6 from the memory to be outputted.

As the global-motion and local-motion parameters of

this embodiment are based upon the same transform method, both parameter based approaches can share the motion compensator 5a of Fig. 7 for generating the predicted picture candidate. When reading out the predicted picture candidate 6 based upon the global-motion parameter (step S31), the memory-a is always used as a reference memory.

The effective motion parameter value ranges may be fixed on a moving picture sequence basis, and alternatively, changed on a frame basis, for example. With the frame based changes, the effective motion parameter value ranges assigned to the respective memories of the frame are multiplexed in a bitstream to be transmitted to a decoding station to perform the same memory selection.

Fig. 21 is a diagram showing the bitstream 21 of this embodiment.

The bitstream is provided with header information added at the front on a moving picture sequence basis. The header information includes effective motion parameter value range indication information of the respective memories. By thus designating the effective motion parameter value range indication information at the front of a moving picture sequence, the moving picture sequence is predicted with the fixed effective

THIS SHEET HAS BEEN REPLACED



motion parameter value ranges in a counterpart decoder.

When varying the effective motion parameter value ranges each frame, the effective motion parameter indication information is to be included in the header information added on a frame basis.

Thus, the efficient and adaptive use of the memories in response to the magnitude of a local motion of a frame can be provided by the encoder including the motion compensator 5a, thereby enhancing prediction efficiency.

This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

Further, this embodiment has thus disclosed the encoder performing on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. No global-motion detection involves no necessity for transmitting information

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED



indicating a global/local prediction as the prediction mode.

Further, a counterpart decoder may be provided with the prediction picture generation section 100 including the same component as the motion compensator 5 discussed in this embodiment. In the decoder, a motion compensator only generates a single predicted picture based upon a decoded motion parameter.

Embodiment 3.

Another embodiment shows an encoder having a single replacement of the motion compensator 5 of the encoder of Fig. 1. The configuration and operation of a motion compensator 5b are now described. A motion detector 3a employed in this embodiment is assumed to output an amount of the parallel translation, the affine parameter, and the perspective parameter as motion parameters 4a.

Further, the memory-a of this embodiment is assumed to be a frame memory for storing a reference picture frame, the memory-b is assumed to be a static sprite memory, and the memory-c is assumed to be a dynamic sprite memory.

Fig. 9 shows the configuration of the motion compensator 5b of this embodiment in detail. In the figure, a parallel translation amount 36 (i.e., a motion

This Sheet Has Been Replaced



vector), an affine parameter 37, a perspective parameter 38, a parallel-translation based prediction picture memory address calculator 39, an affine parameter based prediction picture memory address calculator 40, and a perspective parameter based prediction picture memory address calculator 41 are shown. Fig. 10 is a flowchart illustrating the operation of the motion compensator 5b in detail.

Initially, with the I-frame, no prediction is performed (step S33). With a frame other than the I-frame, the prediction picture memory address calculators 39 through 41 calculate the respective prediction picture memory addresses 23 based upon the respective values of the motion parameters 4a (step S34).

The memory address calculators 39, 40, 41 calculate addresses based upon picture transform methods assigned, respectively, to the corresponding memories. In this embodiment, the parallel translation is assigned to the memory-a, the affine parameter involving such a simple transformation as a rotation and an expansion/contraction is assigned to the memory-b, and the perspective parameter involving a three-dimensional complicated motion is assigned to the memory-c. These transform methods may be expressed by the following transform expressions.

[Parallel translation]

Amount of parallel translation (a, b):

$$x'=x+a$$

$$y'=y+b$$

[Affine transform]

Affine parameter (a,b,c,θ):

$$x'=a(\cos\theta)x+a(\sin\theta)y+b$$

$$y'=a(-\sin\theta)x+a(\cos\theta)y+c$$

[Perspective transform]

Perspective parameter (a,b,c,d,e,f):

$$x'=(ax+by+c)/(gx+hy+1)$$

$$y'=(dx+ey+f)/(gx+hy+1)$$

Here, (x, y) in a two-dimensional coordinate system represents a pixel location of an original macroblock. (x', y') represents a pixel location in a memory corresponding to (x, y,) based upon each of the parameters. That is a location in a memory (x', y') is calculated based upon these parameters. Through this mechanism, a memory most suitable for the characteristic of a motion can be chosen to be used for prediction on a macroblock basis. With calculated prediction picture memory addresses 23 based upon the respective motion parameters 36, 37, 38, the memory reader 24 reads out the predicted picture candidates 6 from corresponding memories to be outputted (steps S35 through S39).

The transform methods assigned to the respective memories of the frame are multiplexed in the bitstream 21 to be transmitted to a decoding station as a motion detection method indication information so that the same motion compensation can be performed in the decoding station.

Fig. 22 is a diagram showing the bitstream 21 of this embodiment.

Header information added at the front of a moving picture sequence includes the motion detection method indication information. In the encoder, the transformation types to be used in the respective memories are interchangeable, and thus the motion detection method indication information indicating a memory-transform method relation is to be transmitted to the counterpart decoder as the header information of the moving picture sequence. Thus, transformation types assigned to be used with the respective memories can be identified in the decoder.

In the decoder, the identified transformation types are dynamically assigned to the respective memories.

Thus, the efficient and adaptive use of the memories in response to the characteristic of a local motion of a frame is provided by the encoder including the motion compensator 5b, thereby enhancing prediction efficiency.

THIS SHEET HAS BEEN REPLACED



This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

Further, this embodiment has thus disclosed the encoder performing on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. No global-motion detection involves no necessity of transmitting information of a global/local prediction as the prediction mode.

Further, this embodiment has thus disclosed the use of the memories a, b, and c. Alternatively, the use of memories a and b alone, memories a and c alone, or memories b and c alone, is also applicable.

Further, a decoder may be provided with the prediction picture generating section 100 including the same component as the motion compensator 5b discussed in this embodiment. A motion compensator in a decoder only

This Sheet Has Been Replaced



generates a single predicted picture based upon a decoded motion parameter.

Embodiment 4.

Another embodiment shows an encoder which receives a plural number, two for example, of different video objects, having shape information, intermingled in a moving picture sequence for a collective encoding. Fig. 11 shows the configuration of the encoder of this embodiment.

In the figure, an input picture frame 42, an object separator 43, object data 44a, 44b, shape blocks 45a, 45b, switches 46a, 46b, shape encoders 47a, 47b, compressed shape block data 48a, 48b, locally decoded shape blocks 49a, 49b, texture data (macroblocks) 50a,50b, motion detectors 51a, 51b, motion parameters 52a, 52b, motion compensators 53a, 53b, predicted picture candidates 54a, 54b, prediction mode selectors 55a, 55b, prediction mode information 56a, 56b, predicted pictures 57a, 57b, prediction error signals 58a, 58b, texture encoders 59a, 59b, compressed texture data 60a, 60b, locally decoded prediction error signals 61a, 61b, locally decoded macroblocks 62a, 62b, a memory update unit 63, memory-a 64, memory-b 65, memory-c 66, memory-d 67, memory-e 68, memory-f 69, variable-length encoders

70a, 70b, a multiplexer 71, a buffer 72, a bitstream 73, a memory section 94, an object-A encoder 88a for encoding an object-A, and an object-B encoder 88b for encoding an object-B are shown. The object encoders 88a, 88b are structurally identical to each other with the identical components.

This encoder inputs the picture frame 42, which is decomposed into the encoding units of objects in the object separator 43. The object separator 43 is assumed to be assigned a processing method arbitrarily.

The shape information of an object is transferred to the shape encoder 47a, 47b in a form of the shape block 45a, 45b to be encoded, and then transferred to the variable-length encoder 70a, 70b as the compressed shape block data 48a, 48b.

The motion detector 51a, 51b detects a motion parameter based upon the locally decoded shape block 49a, 49b in the same manner as that of the VM encoding system. A motion parameter can be detected on a macroblock basis by using all the memories a through f.

As a rule, however, the memories a through c are designed to be used for an object-A to be encoded in the object-A encoder 88a, and the memories d through f are designed to be used for an object-B to be encoded in the object-B encoder 88b.

Also, as for a motion type, an arbitrary transform parameter is assumed to be detected on a macroblock basis with all the memories in the memory section 94, involving no distinction between the motion vector and the warping parameter.

The motion compensator 53a, 53b generates all the predicted picture candidates 54a, 54b based upon the respective motion parameters 52a, 52b. Then, in the prediction mode selector 55a, 55b, the predicted picture 57a, 57b is obtained along with the prediction mode information 56a, 56b. The predicted picture 57a, 57b is then differentiated from an original signal or the texture data 50a, 50b to obtain the prediction error signal 58a, 58b, which is encoded in the texture encoder 59a, 59b to be transmitted to the variable-length encoder 70a, 70b. The locally decoded prediction error signal 61a, 61b is added to the predicted picture 57a, 57b to obtain the locally decoded macroblock 62a, 62b to be stored into the memories a through f in accordance with an indication by the memory update unit.

Object A/B data when encoded in the object-A/B encoder 88a, 88b are multiplexed in the bitstream 73 at the multiplexer 71 to be transmitted via the buffer 72.

The prediction of this embodiment is discussed below focusing on the motion compensator 53a, 53b playing a

primary role in the prediction.

The motion compensator 53a, 53b of this embodiment generates a predicted picture candidate based upon the motion parameter 52a, 52b detected in the motion detector 51a, 51b. Fig. 12 shows the configuration of the motion compensator 53a in detail. Fig. 13 shows a flowchart illustrating the operation of the motion compensator 53a in the object-A encoder 88a.

In Fig. 12, an object-B reference judger 74a and an object-B reference indicator flag 75a are shown.

The motion parameter 52a includes memory information used for detection. A predicted picture candidate is generated based upon a parameter value through the prediction picture memory address calculator 22a and a memory reader 24a in the same manner as that stated in the first embodiment (step S44 through step S49). The object-B reference judger 74a judges if the memories assigned to object-B are used for predicting the current macroblock based upon the reference memory information included in the motion parameter 52a (step S43).

The object-B reference judger 74a outputs a judged result as the object-B reference indicator flag 75a, which is multiplexed in the bitstream 73 to be transmitted to a decoding station so as to be used for deciding whether the object can be reproduced in a single

use of the memories a, b, c of self in the decoding station. In order to secure the single use of the memories of self when reproducing the object in the decoding station, a limited use of the memories (a, b, c alone) of self for prediction can be controlled by an externally supplied signal 85a at the time of encoding the object.

Thus, the efficient and adaptive use of the memories in response to the characteristic of a local motion of a frame is provided by the encoder including the motion compensator 53a, 53b, thereby achieving an efficient prediction.

This embodiment has thus disclosed the encoder for encoding an object on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, a decoder may be provided with the same components as the motion compensator 53a, 53b of this embodiment. A motion compensator 53 of the decoder only generates a single predicted picture based upon a decoded motion parameter. Further, if the decoder is structured so as to acknowledge whether a decoding object can be reproduced by itself by way of decoding a bit

corresponding to object reference indicator flag 75a, 75b of the other object in a bitstream, then an error-free secured reproduction of decoded object data can be achieved.

Embodiment 5.

Another embodiment shows an encoder where the number of memories or the size of a memory can be varied flexibly in response to a change in a video object at each time instance. The encoder of a fifth embodiment modifies the encoder of Fig. 1 with a replacement of the memory update unit 15.

Fig. 14 shows the configuration of a memory update unit 15a of this embodiment in detail. In the figure, a memory expansion judger 76, a memory expansion indicator signal 77, and a memory contraction judger 78, and a memory contraction indicator signal 79 are shown. Fig. 15 shows an operating flow (step S51 through S63) of the memory update unit 15a.

A picture substantially different from the past record of a moving picture sequence stored in the memories may occur due to a scene change or the like. This may cause a deterioration of prediction efficiency after the scene change if reference pictures stored in the existing memories are the only available. For such

an occasion, the scene-change detector 80 detects a scene change, a frame appearing immediately after the detected scene change is subject to intra-frame coding or the like, and resultant intra-frame coded data are stored additionally in a memory as new reference data, thereby enhancing prediction efficiency thereafter.

Further, in consideration of the physical limitation of storage capacity, a flexible approach of contracting the portions which are rarely used for prediction of the reference pictures stored in the memories is introduced. Specifically, the frequency in use of memory areas for prediction of the respective memories a, b, c is examined in the memory update unit 15a based upon the prediction mode 8. Consequently, the memory update unit releases a memory area identified low in frequency from an area for use. For example, with a software based implementation of this encoder, limited RAM resources may be used effectively.

Viewed in this light, the memory update unit 15a of this embodiment is provided with a function of expanding a memory area in response to each time instance of a time-varying moving picture sequence and contracting a memory area including a reference picture rarely used for prediction.

The memory-a, similarly to the first embodiment, is

judged in the memory-a update judger 27 whether to be updated (step S50). When updating the memory-a, the locally decoded picture 14 is written into the memory-a (steps S56, S57). The locally decoded picture 14 is written into the other memories as well in accordance with the prediction mode 8 (step S51 through S55).

The updating of the contents of the memories involves the judgement of memory expansion/contraction. The memory expansion judger 76 judges whether to expand the size of the memory-a (or the memory-b, or the memory-c) based upon the activity 26 used for updating the memory-a (steps S58 through S60). When a positive judgement is made due to a scene change or the like, the expansion of the memory is indicated by the memory expansion indicator signal 77. The memory contraction judger 78 counts the times of a memory area used for prediction based upon the prediction mode 8. With a memory area counted less than a predetermined number in use for prediction, the contraction of the memory area is indicated by the memory contraction indicator signal 79 (steps S61 through S63).

Thus, a highly efficient prediction can be achieved in response to each time instance of a time-varying moving picture sequence by the encoder including the memory update unit 15a. In addition, the dynamic

allocation of memory areas required for prediction contributes to the enhancement of prediction efficiency and the effective use of memory resources.

This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

Further, this embodiment has thus disclosed the encoder for encoding a frame on a macroblock basis. The same effect can be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, a counterpart decoder may be provided with the same component as the memory update unit 15a discussed in this embodiment.

Embodiment 6.

With reference to the respective previous embodiments, the memories to be used for prediction are changed on a macroblock basis. Alternatively, the memories to be used for prediction can be changed on a frame or a video object basis. This eliminates the necessity of encoding memory relating information to be encoded on a frame or a video object basis and memory

selection information (which is included in the prediction mode 8) to be encoded on a macroblock basis, thereby achieving an efficient encoding.

With reference to the encoder of Fig. 1 of the first embodiment, for example, the macroblock based changes of the memories used for prediction create the necessity of transmitting additional information identifying a memory used for prediction on a macroblock basis. According to this embodiment, the changing unit of the memories to be used for prediction is limited to a frame or a video object, thereby eliminating the additional information to be transmitted on a macroblock basis effectively. Fig. 23 shows a difference of the transmission bitstream 21 of this embodiment from the transmission bitstream 21 of Fig. 20 of the first embodiment. The bitstream of Fig. 23 represents a frame based change of the memories to be used for prediction with the prediction memory indication information 800 included in frame based header information. The bitstream of Fig. 23 may be effective, for example, in the case that the picture characteristic of a moving picture sequence changes infrequently including little changes locally on a macroblock level. Further, a decoder may be provided so as to decode the thus encoded bitstream to reproduce a frame or a video object.

Embodiment 7.

With reference to the previous embodiments, two predicted picture candidates read out from an arbitrary plural number, two (e.g., memories a and b) for example, of the memories are subject to an arithmetic mean to obtain a picture as a member of the predicted picture candidates 6 or as the predicted picture 9. Further, a decoder may be provided so as to decode the thus encoded bitstream to reproduce a frame or a video object.

Embodiment 8.

With reference to the encoders of the previous embodiments, a previously detected feature parameter representing the spatial complexity, perceptual significance and the like of a picture segment as a prediction unit may be utilized as tools for deciding a prediction mode and for judging the updating of the memories.

For example, a moving picture is assumed to include a motion too complicated to encode data in an acceptable quality within a given amount of encoding. In this case, significance is examined on a prediction picture segment (e.g., a macroblock, an arbitrary shaped picture segment, an arbitrary shaped block) basis. Consequently, a low

quality encoding is assigned to some extent to a less significant segment in order to save some amount of encoding for a more significant segment, thereby improving an overall picture quality. With the encoders of this invention where two or more memories are switched interchangeably at an arbitrary timing to be used for prediction, a more adaptive prediction can be achieved in response to the characteristic of a picture, through detecting a feature parameter representing the significance of a prediction picture segment and then determining the use of the memories dynamically based upon a detected feature parameter. For example, as shown in Fig. 16, a segment-significance detector 95 is provided for detecting the feature parameter on a segment basis to determine the significance of the segment. The segment-significance detector 95 transfers a segment-significance to a prediction mode selector 7a and a quantization parameter based upon the segment-significance to a texture encoder 11a. With a segment judged more significant in the segment-significance detector 95, a most complex mode among two or more prediction modes available is used for prediction. Specifically, reference pictures from the respective memories a, b, c are used to obtain the motion parameters and the predicted pictures, respectively, based upon a

complex motion model. In the prediction mode selector 7a, a prediction mode having the highest prediction efficiency is selected from among modes including an arbitrary combination (e.g., an arithmetic mean) of the predicted pictures. At the same time, reference pictures of all the memories used for prediction are updated. The texture encoder 11a performs an encoding using a quantization parameter having a smaller quantization step size. With a less significant segment, a simplified prediction mode (i.e., a parallel translation amount detection using a single memory) is employed for prediction and a quantization parameter having a larger quantization step size is utilized for encoding, regardless of the amplitude of an obtained prediction error signal, so that an amount of encoding be reduced. Through this control, a less significant segment reduces its picture quality to some extent and a more significant segment maintains its quality through a high-quality prediction, thereby improving an overall quality within a given amount of encoding.

Embodiment 9.

In an encoder where a moving picture sequence is predicted and encoded by using two or more memories, a parameter representing an amount of encoding available

for the moving picture sequence at each time instance, an amount of a change in a scene at a certain time instance (e.g., a scene change detection), or the feature parameter or significance of a prediction picture segment described in the eighth embodiment may be detected previously. The values of these parameters may be used for predicting a picture at a particular time instance in a prediction system. Alternatively, these values may be utilized as judging tools for selecting a reference memory area in a prediction system. A frame-significance detector 96 may be provided for determining the significance on a frame basis as shown in Fig. 17. The frame-significance detector 96 detects, for example, an amount of a change in a motion between the current and the previous frame (e.g., a scene change detection by the scene change detector 80), the appearance of a new object or the disappearance of an object or the like. A final significance of the current frame is determined in consideration of an amount of encoding available for the current frame informed by the transmission buffer 20. Based upon the final significance, a more significant frame may be predicted by using all the prediction methods and reference memory areas available for the maximum possible improvement of prediction efficiency, whereas a less significant frame is predicted in a

limited use of the prediction methods and the reference memory areas for a simplified encoding so as to reduce the throughput. An alternative encoder for performing intra-frame coding alone involving no prediction at a scene change may be possible. In addition, a more sophisticated quality control may be achieved with a joint use of the segment-significance detector 95 discussed in the eighth embodiment. Through this control, a less significant frame reduces its quality to some extent and a more significant frame maintains its quality by a high-quality prediction, thereby improving an overall picture quality within a given amount of encoding.

The idea of this embodiment is also applicable to a software based encoding being associated with unsteady transaction processes and unsteady size of available storage in order to achieve an efficient encoding in the maximum use of available resources. This reduces a throughput with a less significant frame, thereby accelerating a general processing speed.

Embodiment 10.

In an encoder where a moving picture sequence including two or more video objects is predicted and encoded by using two or more memories, as shown in Fig.

11, a parameter representing a gross amount of encoding available for the sequence, an available amount of encoding of a video object at each time instance, an amount of a change in a video object at each particular time instance (e.g., the appearance/disappearance of an object), a level of significance/attention of a video object in a particular scene, or the feature parameter or significance of a prediction picture segment discussed in the eighth and ninth embodiments may be detected previously. The values of these parameters may be utilized for predicting a video object at each particular time instance. Alternatively, these values may be utilized as judging tools for selecting a reference memory area.

For example, as shown in Fig. 18, significance detectors 97a through 97n responsive, respectively, to objects 1 through n may be provided for detecting a parameter representing an amount of a change in an object at each time instance, or the appearance/disappearance of an object. In addition, the significance of an object at each time instance is determined in consideration of the occupational proportion of a buffer 72x for storing the encoded data of all the objects and the occupational proportion of virtual buffers 72a through 72n for the respective objects. When a new type of segment appears

as a result of an object having another object overlapped in part, for example, then the following control may be applied to this type of segment because this type of segment has a great influence on prediction efficiency thereafter. A higher significance may be assigned to this type of segment to obtain an encoded picture in high quality even without enough space available for storage in the corresponding virtual buffer of the object. Significance detected in the significance detectors 97a through 97n is transferred to object 1 through N encoders 98a through 98n, where a full use of the prediction methods and the reference memory areas available is allowed to a more significant object so as to improve prediction efficiency to the maximum, whereas a limited use of the prediction methods and the reference memory areas is assigned to a less significant object so as to simplify the encoding, thereby reducing the throughput. Further, with an encoder for encoding objects decomposed from a frame through a real time separation, when a considerable amount of a change occurs in the contents of the object due to the appearance of a new object or the disappearance of an existing object, the object may be subject to intra-frame coding alone with no prediction involved. A more sophisticated quality control may be achieved on a prediction segment basis of an object in a

joint use of the object 1 through N encoders 98a through 98n and the segment-significance detector 95 discussed in the eighth embodiment. Through this control, a less significant object is reduced in quality to some extent and a more significant object manages to maintain its quality through a sophisticated prediction, thereby improving an overall quality within a given amount of encoding.

Embodiment 11.

An alternative encoder may be provided with a prediction information encoder 91 for allocating a code (encoding) to prediction relating encoding information (e.g., a reference memory number) as shown in Fig, 19.

In the encoder where a moving picture sequence or a video object is predicted and encoded by using the memories a, b, c, the memories may be ranked based upon frequency in use for prediction with ranks being updated dynamically during an encoding operation. Consequently, a code allocation is performed to the prediction relating encoding information (e.g., a reference memory number) based upon the ranks of the respective memories used for prediction in the prediction information encoder 91.

For example, in the encoder of Fig. 19, the memory update unit 15b may be provided with a counter 92 which

counts the times of the respective memories a, b, c to be used for prediction, ranks the memories a, b, c based upon counted values, and outputs resultant ranking information 90. This ranking may be performed on a picture (VOP) basis at a particular time instance of a frame or a video object, and alternatively, on a smaller unit basis of a prediction picture segment (e.g., a macroblcok, an arbitrary shaped segment, and an arbitrary shaped block).

This shows how often the respective memories are used for prediction. A memory in a frequent use for prediction is the most significant for prediction, and thus, a high frequency in use for reference corresponding to a high rank.

When encoding information on the frequency in use for prediction of the memories on a prediction picture segment basis, a memory in a frequent use for reference (i.e., a high-rank memory) is allocated a short code to enhance encoding efficiency.

In addition, if the motion parameter detected on a prediction picture segment basis is allocated a code length in response to the rank of a memory used for reference, then a shorter code may be assigned to a motion parameter value generated frequently, thereby achieving an efficient encoding of the prediction

information. This may be materialized with an alternative encoder where the prediction information encoder 91 in the variable-length encoder/multiplexer 19 receives the ranks of the respective memories from the counter 92 in the memory update unit 15b and encodes the prediction information using a variable-length code based upon the ranking information 90.

Embodiment 12.

Fig. 24 shows the configuration of a picture decoder where an encoded digital picture through compression is reproduced through expansion according to another embodiment. In the figure, the encoded bitstream 21, a variable-length decoder 119, the quantized DCT coefficient 12, a quantization orthogonal transform coefficient 12a, a quantization step 12b, a texture decoder 111, a dequantizer 111a, an inverse orthogonal transformer 111b, a decoding adder 190, a decoded picture 101, a display controller 191, the prediction mode 8, memory-b update information 1001, memory-c update information 1002, the motion vector 4 (a motion parameter), the prediction memory indication information 800, an in-screen location 195 of a prediction picture segment, a motion compensator 105, a memory-a 116, a memory-b 117, a memory-c 118, a memory update unit 115,

and a predicted picture 106 are shown. The motion compensator 105 and the memory update unit 115 form a prediction picture generation section 100a. The memories a, b, c form a memory area 200a.

According to this embodiment, the memory-a is assumed to be a frame memory designed to store a frame of picture data, the memory-b is assumed to be a static sprite memory, and the memory-c is assumed to be a dynamic sprite memory. The decoder of this embodiment is assumed to receive the bitstream 21 of Fig. 22. Although not shown in Fig. 22, the memory-b update information 1001 and the memory-c update information 1002 are assumed to be transmitted in the bitstream. The memory-b update information 1001 is assumed to include an update indication for a full updating of the static sprite memory and picture data for the full updating. Similarly, the memory-c update information 1002 is assumed to include an update indication for a full updating of the dynamic sprite memory and picture data for the full updating.

The operation of the thus configured decoder is described below. The variable-length decoder 119 analyzes the bitstream 21 and decomposes it into separate encoded data. The quantization orthogonal transform coefficient 12a is transferred to the dequantizer 119a to

be dequantized by using the quantization step 12b. A dequantized result is subject to inverse orthogonal transformation in the inverse orthogonal transformer 111b to obtain a decoded texture, which is transferred to the decoding adder 190. Orthogonal transformation employed here is the same as that employed in an encoding station such as Discrete Cosine Transformation (DCT).

The motion compensator 105 inputs the motion vector 4, the prediction memory indication information 800, and information indicating the in-screen location 195 of a prediction picture segment included in the bitstream 21, all of which are decoded in the variable-length decoder 119. The motion compensator 105 reads out a right predicted picture from reference pictures stored in the memories a, b, c based upon the three kinds of information. The in-screen location 195 of a prediction picture segment can be obtained by counting the number of macroblocks, other than from the information included in the bitstream. The process of generating a predicted picture will be discussed in a later section for describing the operation of the motion compensator 105 in detail.

The decoding adder 190, based upon the information of the prediction mode 8, outputs an output from the inverse orthogonal transformer 111b directly as the

decoded picture 101 with a block through intra-frame coding, and, with a block through inter-frame coding, adds an output from the inverse orthogonal transformer 111b to the predicted picture 106 to be outputted as the decoded picture 101. The decoded picture 101 is transferred to the display controller 191 to be outputted to a display device and also transferred to the memories a, b, c to be stored as a reference picture for a later use in decoding. A memory writing operation is controlled by the memory update unit 115 based upon the prediction mode 8.

A predicted picture generation performed in the motion compensator 105 in the prediction picture generation section 100a is now discussed. According to this embodiment, the prediction method of a picture is determined based upon the prediction memory indication information 800. The decoder of this embodiment generates a predicted picture using a reference picture through predetermined coordinate transformation and interpolation based upon the motion vector 4 and the prediction memory indication information 800. Coordinate transform methods are assigned previously to the respective memories to be used for prediction. For example, the following approaches are possible similar to the picture transform methods described in the third

THIS SHEET HAS BEEN REPLACED



embodiment.

(1) the memory-a used for prediction (with the prediction memory identification information 800 indicating the use of the memory-a)

The coordinates of each pixel of a prediction segment are translated based upon the motion vector and picture data at a corresponding location in the memory-a is read out as the predicted picture.

(2) the memory-b used for prediction (with the prediction memory identification information 800 indicating the use of the memory-b)

An affine transform expression is found based upon the motion vector, the coordinates of each pixel of a prediction segment are displaced based upon the transform expression, and picture data at a corresponding location in the memory-c is read out as a predicted picture.

(3) the memory-c used for prediction (with the prediction memory identification information 800 indicating the use of the memory-c)

A perspective transform expression is found based upon the motion vector, the coordinates of each pixel of a prediction segment are displaced based upon the transform expression, and picture data at a corresponding location in the memory-b is read out as a predicted picture.

Fig. 25 shows the configuration of the motion compensator 105 in detail. In the figure, a switch 161, a corresponding point determiner 162 for the memory-a, a corresponding point determiner 163 for the memory-b, a corresponding point determiner 164 for the memory-c, a memory read address generator 165, a switch 166, and an interpolator 167 are shown. Fig. 26 is a flowchart illustrating the optation of the motion compensator 105.

The operation of the motion compensator 105 of this embodiment is described below with reference to Figs. 25 and 26.

1) Determining a corresponding point

Initially, the corresponding point determiner of a corresponding memory is selected by the switch 161 based upon the prediction memory indication information 800. The vector 4 is then inputted to a selected corresponding point determiner. In this section, a predicted picture location corresponding to each memory is calculated, which is explained below with each memory.

1-1) the memory-a indicated by the prediction memory indication information 800 (step S100)

A predicted picture location is calculated through parallel translation based upon a motion vector (step S101). Specifically, a predicted picture location (x',y') corresponding to a pixel at a prediction picture

segment location (x,y) is determined based upon a motion vector (a, b) according to the following expression.

$$x' = x + a$$

$$y' = y + b$$

A determined predicted picture location is outputted to the memory read address generator 165.

1-2) the memory-b indicated by the prediction memory indication information 800 (step S103)

An affine transform expression is determined based upon the motion vector 4. Specifically, an affine parameter (a, b, c, θ) of the following expression is determined by using the motion vector of a vertex of a rectangular area enclosing a prediction picture segment,

$$x' = a(\cos\theta)x + a(\sin\theta)y + b$$

$$y' = a(-\sin\theta)x + a(\cos\theta)y + c$$

thereby obtaining the predicted picture location (x',y') corresponding to a pixel at the location (x,y) of a prediction picture segment to be outputted to the memory read address generator 165 (step S104).

1-3) the memory-c indicated by the prediction memory indication information 800 (step S106)

A perspective transform expression is determined based upon a motion vector. Specifically, a perspective parameter (a, b, c, d, e, f) of the following expression is determined by using the motion vector of a vertex of a

rectangular area enclosing a prediction picture segment,

$$x' = (ax + by + c) / (gx + hy + 1)$$

$$y' = (dx + ey + f) / (gx + hy + 1)$$

thereby obtaining the predicted picture location (x',y') corresponding to a pixel at the location (x,y) of a prediction picture segment to be outputted to the memory read address generator (step S107).

2) Reading out data for generating a predicted picture

Based upon the predicted picture location (x', y') outputted from a selected corresponding point determiner, the memory read address generator 165 generates a memory address for specifying the location of picture data required for generating a predicted picture in a reference picture stored in a memory, and reads out the data for generating a predicted picture (steps S102, 105, 108).

3) Generating a predicted picture

Among the component pixels of a predicted picture, with a pixel at an integer pixel location, the data for generating a predicted picture is used directly as a component pixel of a predicted picture, and with a pixel at a real number precision pixel location, the data for generating a predicted picture is subject to interpolation in the interpolator 167 to generate an interpolated pixel value (steps S109, S110, S111). Fig.

THIS SHEET HAS BEEN REPLACED



26 illustrates an interpolated pixel value generation. In Fig. 26, $(i_0, j_p)$ denotes an integer pixel location, $(j^p, j^p)$ denotes a real number precision pixel location, and w denotes a weight.

4) Updating a memory (a reference picture)

Fig. 28 shows a flowchart illustrating a control operation of the memory update unit 115. The memory update unit 115 controls an updating of the respective memories on a readout unit basis (e.g., a macroblock) of a predicted picture based upon the prediction mode 8 (or the prediction memory indication information 800). With the memory-a used for prediction (step S112), the contents of the memory-a and the memory-c are updated regularly with the decoded picture 101 (step S113). With the memory-b used for prediction (step S114), the reference picture of the memory-b is not updated on a readout unit basis of a predicted picture because of the memory-b being a static sprite memory, whereas the contents of the memory-a and the memory-c are updated regularly with the decoded picture 101 (step S115). When receiving the update indication by the memory-b update information 1001, then the memory update unit updates a full content of the memory-b with received picture data included in the memory-b update information 1001 (step S116). With the memory-c used for prediction (step

S117), the contents of the memory-a and the memory-c are updated regularly by using the decoded picture 101 (step S118). When receiving the update indication by the memory update information, the memory update unit updates the content of the memory-c with received picture data included in the memory-c update information 1002 (step S119).

The use of the three memories a, b, c of this embodiment may be replaced by the use of two memories thereof, for example, with the memories a and b, that is, a frame memory and a static sprite memory. Alternatively, the memories a and c, that is, a frame memory and a dynamic memory, may be used.

As aforementioned, according to the decoder of this embodiment, the bitstream 21 encoded through an efficient prediction using the various kinds of motion parameters in response to the motion of a picture can be decoded. In addition, the decoder is applicable to the arbitrary updating approach of the contents of a reference picture at a timing determined in the encoding station, thereby achieving a more adaptive decoding in response to the characteristic of a picture.

According to this embodiment, if the bitstream includes a prediction error signal encoded through encoding other than orthogonal transform encoding, the

same effect may be obtained by replacing a component for decoding a prediction error signal, other than the motion compensator and memory update unit.

Further, this embodiment may be applied not only to a decoder for decoding data on a fixed-size block basis, e.g., for decoding a normal television signal on a frame basis, but also to a decoder for decoding an arbitrary shaped video object (e.g., a Video Object Plane disclosed in ISO/IEC JTC1/SC29/WG11/N1902) as a unit without limiting a prediction segment to a fixed-size block.

Industrial Feasibility

As discussed above, the memory areas provided for storing reference pictures according to this embodiment enables the adaptive use of the memories for storing data based upon the characteristic of the moving picture sequence. In addition, the contents of one or more of the memory areas can be updated at an arbitrary timing, so that the content of a time-unvarying picture, such as a background picture, is controlled to be updated on a longer-term basis, and the contents of a locally changing picture segment is controlled to be updated on a regular or sequential basis. This achieves an efficient prediction by reflecting the past record of the moving picture sequence.

Further, the transform parameter value ranges are assigned to the memory areas for making the respective memory areas effective, and the memory areas are switched to be used for prediction among them based upon the value of the transform parameter of a prediction picture segment, thereby achieving an efficient prediction in response to the magnitude of a local/global motion of the moving picture sequence. At the same time, the motion parameters to be encoded on a prediction picture segment basis can be encoded efficiently within the effective motion parameter value ranges of the reference memory areas.

Further, to the respective memory areas, the transform methods becoming effective in the respective memories are assigned, and the memories are switched to be used for prediction among them in response to the type of the transform parameter of a prediction picture segment, thereby achieving an efficient prediction in response to the complexity of a local/global motion of the moving picture sequence. At the same time, the transform method can be selected adaptively in response to the characteristic of a prediction picture segment, thereby achieving an efficient encoding of the motion parameter.

* C L M S *

THIS SHEET HAS BEEN REPLACED



What is claimed is:

1. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoder and a decoder, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories corresponding to different transform methods, respectively; and

a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories.

2. The moving picture prediction system of claim 1, wherein the encoder generates a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and transmits the prediction memory indication information signal and the parameter to a decoding station so as to generate the predicted picture using the

This Sheet Has Been Replaced

* C L M S *



picture data stored in the one of the plurality of memories based upon the one of the transform methods corresponding to the one of the plurality of memories in the decoding station.

3. The moving picture prediction system of claim 1, wherein the decoder receives the parameter and a prediction memory indication information sinal indicating the one of the plurality of memories used for generating the predicted picture from an encoding station,

wherein the prediction picture generation section generates the predicted picture using the picture data stored in the one of the plurality of memories based upon the parameter and the one of the transform methods corresponding to the one of the plurality of memories.

4. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being assigned to different parameter effective value ranges, respectively; and

a prediction picture generation section for receiving a

* C L M S *

THIS SHEET HAS BEEN REPLACED



parameter representing a motion of a picture segment to be predicted, for selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and for generating a predicted picture using the picture data stored in a selected memory.

5. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction; and

a prediction picture generation section including

a motion compensator for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture by using the picture data stored in the plurality of memories based upon the parameter, and

a memory update unit for updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

6. The moving picture prediction system of claim 1, 4,

This Sheet Has Been Replaced

* C L M S *

THIS SHEET HAS BEEN REPLACED



or 5, wherein the moving picture is predicted in a moving picture sequence having first and second video objects,

wherein the plurality of memories includes separate first and second pluralities of memories corresponding to the first and second video objects, respectively, and the prediction picture generation section includes separate first and second generators, respectively, corresponding to the first and second video objects,

wherein the first generator uses the picture data stored in at least one of the first and second pluralities of memories to generate the predicted picture when predicting the first object, and generates information indicating a use of the second plurality of memories for predicting the first object, the information being added to the predicted picture.

7. The moving picture prediction system of claim 1, 4 or 5, wherein the prediction picture generation section generates the predicted picture through a change of either one of a number and a size of the plurality of memories in response to a change in the moving picture at each time instance.

8. The moving picture prediction system of claim 1, 4 or 5, wherein the prediction picture generation section

This Sheet Has Been Replaced

generates the predicted picture in a limited use of memories for prediction in response to a change in the moving picture at each time instance.

9. The moving picture prediction system of claim 1, wherein the prediction picture generation section generates the predicted picture by calculating a plurality of the predicted pictures generated by using the respective picture data stored in the plurality of memories.

10. The moving picture prediction system of claim 1 further comprising:

a significance detector for detecting a feature parameter representing a significance of the picture segment to be predicted,

wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

11. The moving picture prediction system of claim 1 or 5 further comprising:

a significance detector for detecting a parameter

* C L M S *



representing at least one of an amount of bits available for coding the picture segment to be predicted, an amount of change of the picture segment at each time instance, and a significance of the picture segment;

wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, a plurality of memory update methods.

12. The moving picture prediction system of claim 1 or 5, wherein the moving picture is predicted on a video object basis,

wherein the moving picture prediction system further comprising:

a significance detector for detecting a parameter representing at least one of an amount of bits available for coding a video object to be predicted, an amount of change in the video object at each time instance, and a significance of the video object;

wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of

memory update methods.

13. The moving picture prediction system of claim 1 or 5 further comprising:

a prediction information encoder for encoding prediction relating information of the moving picture;

wherein the prediction picture generation section counts a number of times of a memory used for prediction and determines a rank of the plurality of memories based upon a counted number of the times,

wherein the prediction information encoder allocates a code length to the prediction relating information to be encoded based upon the rank of a memory used for prediction.

14. The moving picture prediction system of claim 1, wherein the plurality of memories includes at least a frame memory for storing the picture data on a frame basis and a sprite memory for storing a sprite picture.

15. The moving picture prediction system of claim 14, wherein the sprite memory includes at least one of a dynamic sprite memory involving a regular updating, and a static sprite memory not involving the regular updating.

* C L M S *

THIS SHEET HAS BEEN REPLACED



16. The moving picture prediction system of claim 14, wherein the one of the transform methods corresponding to the one of the plurality of memories is at least one of a parallel translation, an affine transformation, and a perspective transformation in an interchangeable manner.

17. A method for predicting a moving picture to be implemented in at least one of an encoding or a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;

corresponding different transform methods with the plurality of memories, respectively;

receiving a parameter representing a motion of a picture segment to be predicted; and

generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods being corresponding to the one of the plurality of memories.

18. The method for predicting a moving picture of claim 17, further comprising the steps of:

generating a prediction memory indication information

This Sheet Has Been Replaced

signal indicating the one of the plurality of memories used for the picture segment to be predicted; and

transmitting the prediction memory indication information signal and the parameter to a decoding station.

19. The method for predicting a moving picture of claim 17 to be implemented in the decoding, the method further comprising the step of:

receiving a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and the parameter representing a motion of the picture segment to be predicted from an encoding station.

20. A method for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;

assigning separate parameter effective value ranges to the plurality of memories, respectively;

receiving a parameter representing a motion of a picture segment to be predicted;

* C L M S *

THIS SHEET HAS BEEN REPLACED



selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter; and

generating a predicted picture using the picture data stored in a selected memory.

21. A method for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;

receiving a parameter representing a motion of a picture segment to be predicted;

generating a predicted picture using the picture data stored in the plurality of memories based upon the parameter; and

updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

This Sheet Has Been Replaced

ABSTRACT

To achieve an encoding system including a highly efficient prediction performed in response to the content of a scene, a significance, and a motion characteristic of a moving picture and the like, memories a, b, c, motion compensator 5 responsive to an arbitrary transform parameter representing the motion of a prediction picture segment for generating a predicted picture by using arbitrary data stored in the memories a, b, c based upon the transform parameter, and memory update unit 15 for allowing the content of one or more of the memories to be updated at an arbitrary period of time, are provided.

Fig.1

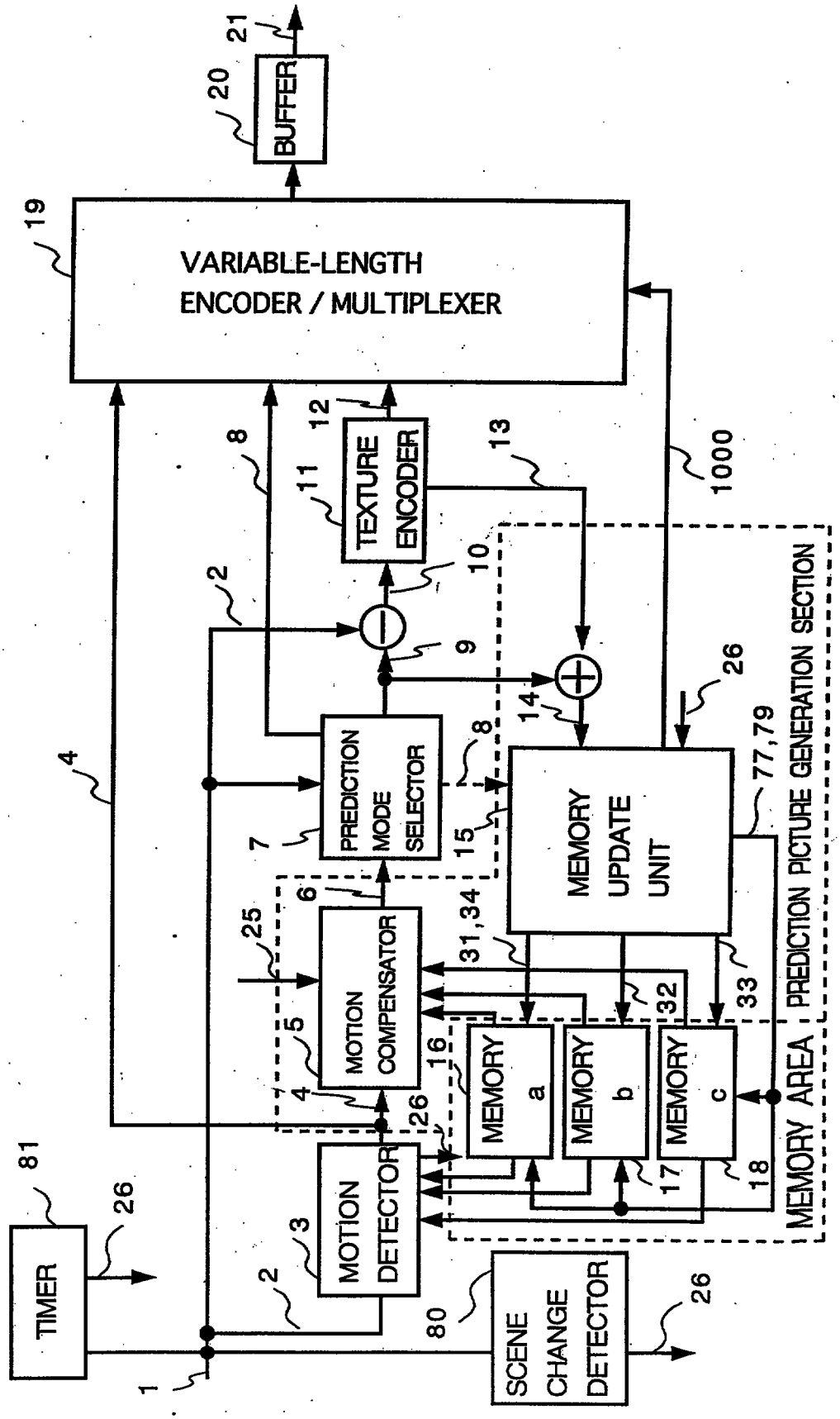

BUFFER
VARIABLE-LENGTH ENCODER / MULTIPLEXER
TEXTURE ENCODER
PREDICTION MODE SELECTOR
MOTION COMPENSATOR
MOTION DETECTOR
MEMORY UPDATE UNIT
MEMORY a
MEMORY b
MEMORY c
MEMORY AREA
PREDICTION PICTURE GENERATION SECTION
TIMER
SCENE CHANGE DETECTOR

Fig.2

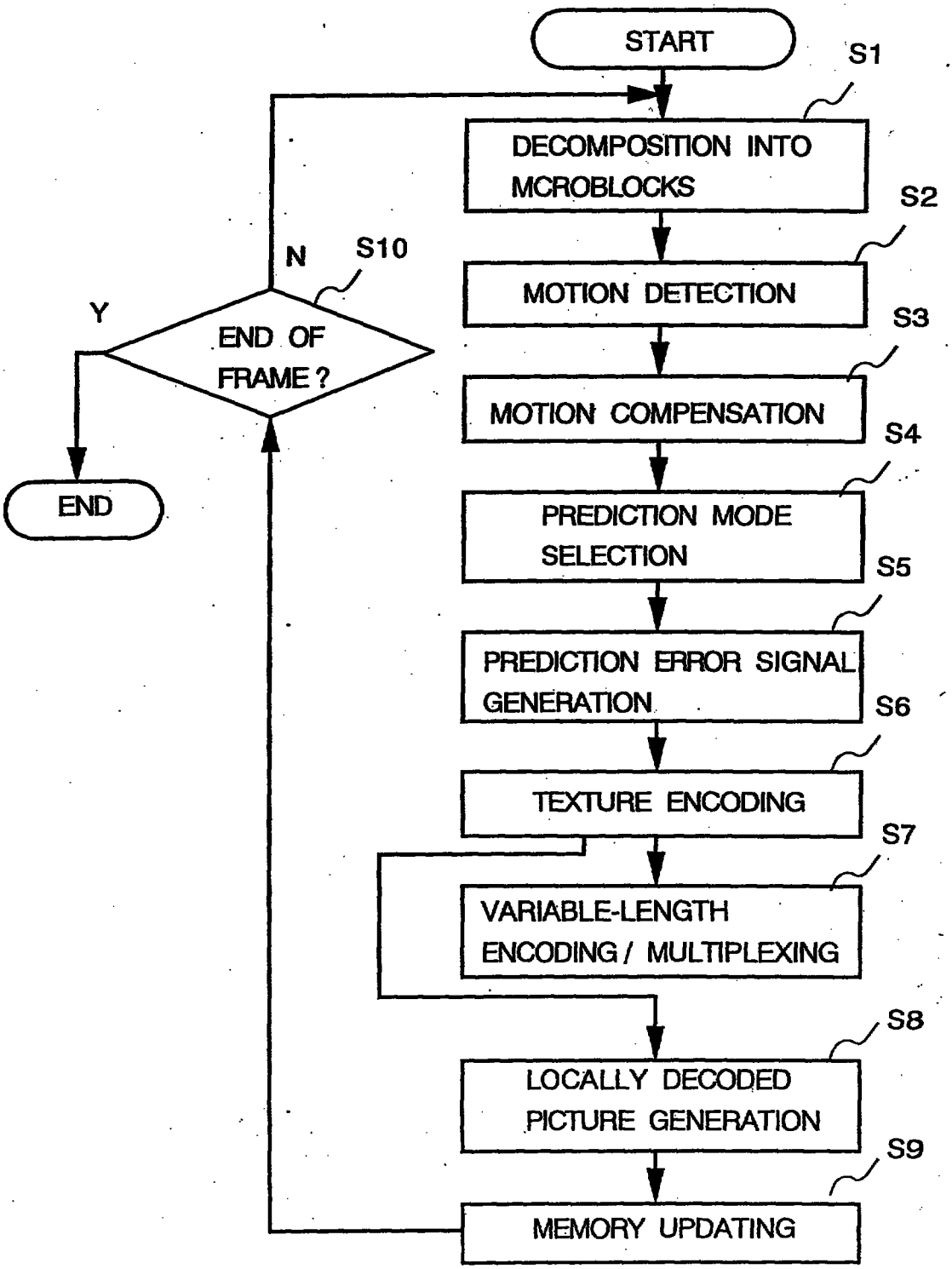

START
S1
DECOMPOSITION INTO MCROBLOCKS
S2
MOTION DETECTION
S3
MOTION COMPENSATION
S4
PREDICTION MODE SELECTION
S5
PREDICTION ERROR SIGNAL GENERATION
S6
TEXTURE ENCODING
S7
VARIABLE-LENGTH ENCODING / MULTIPLEXING
S8
LOCALLY DECODED PICTURE GENERATION
S9
MEMORY UPDATING
S10
END OF FRAME ?
N
Y
END

Fig.3

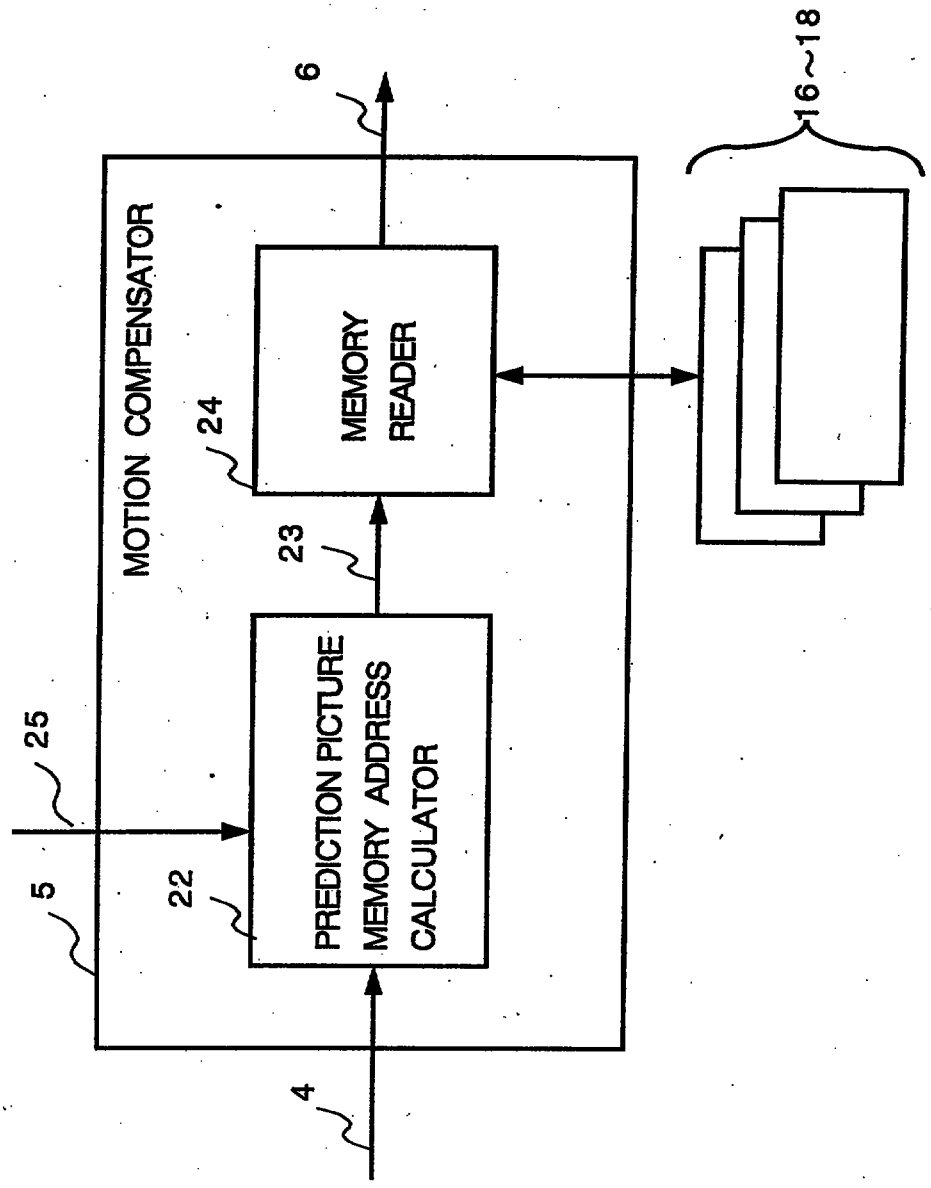

MOTION COMPENSATOR
MEMORY READER
PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
6
16~18
24
23
25
5
22
4

Fig.4

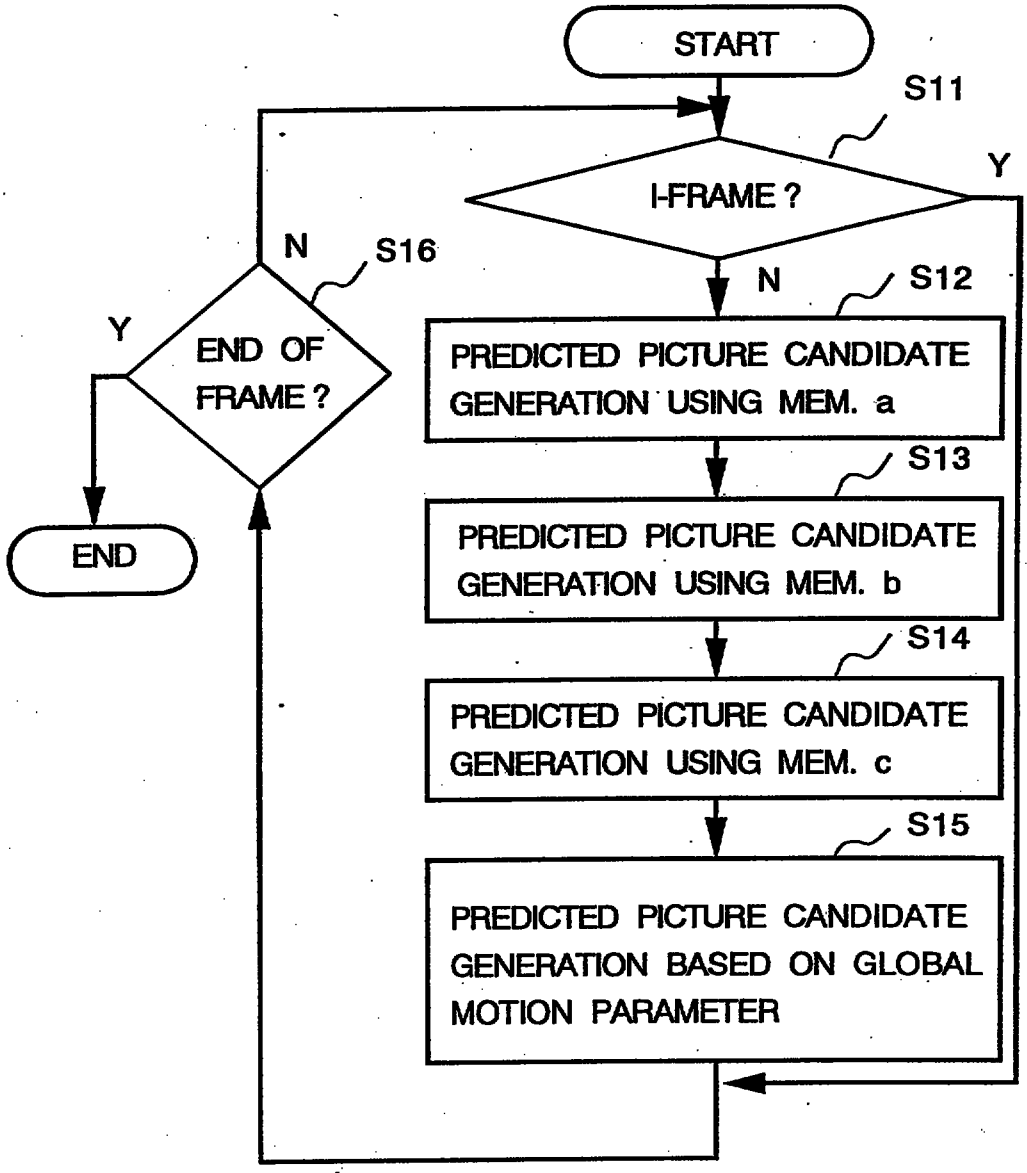

START
S11
I-FRAME ?
Y
N
S12
PREDICTED PICTURE CANDIDATE GENERATION USING MEM. a
S13
PREDICTED PICTURE CANDIDATE GENERATION USING MEM. b
S14
PREDICTED PICTURE CANDIDATE GENERATION USING MEM. c
S15
PREDICTED PICTURE CANDIDATE GENERATION BASED ON GLOBAL MOTION PARAMETER
N
S16
END OF FRAME ?
Y
END

Fig.5

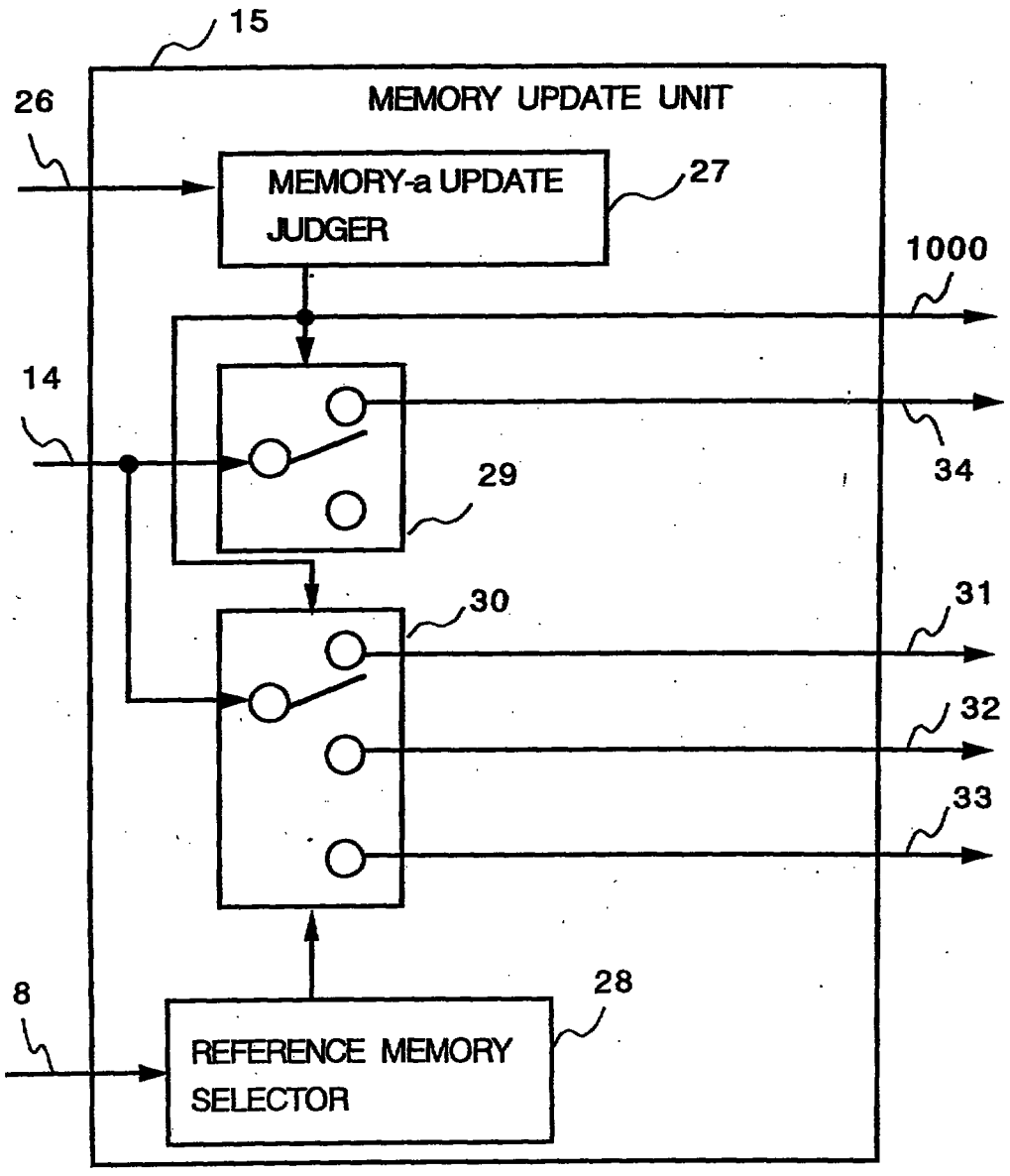

15
MEMORY UPDATE UNIT
26
MEMORY-a UPDATE JUDGER
27
1000
14
29
34
30
31
32
33
8
28
REFERENCE MEMORY SELECTOR

Fig.6

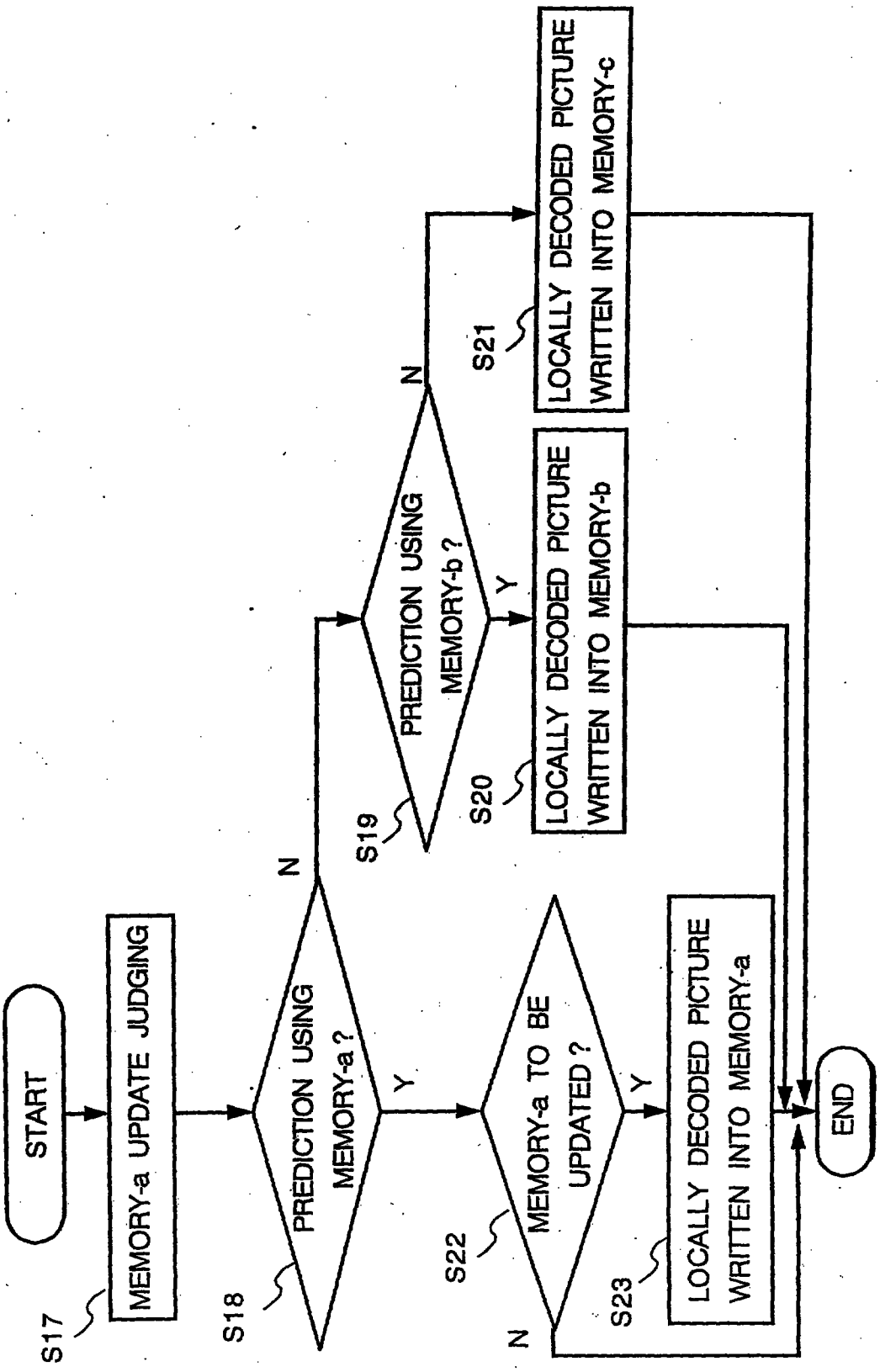

START
MEMORY-a UPDATE JUDGING
S17
PREDICTION USING MEMORY-a ?
S18
N
Y
PREDICTION USING MEMORY-b ?
S19
N
Y
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-c
S21
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-b
S20
MEMORY-a TO BE UPDATED ?
S22
N
Y
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-a
S23
END

Fig.7

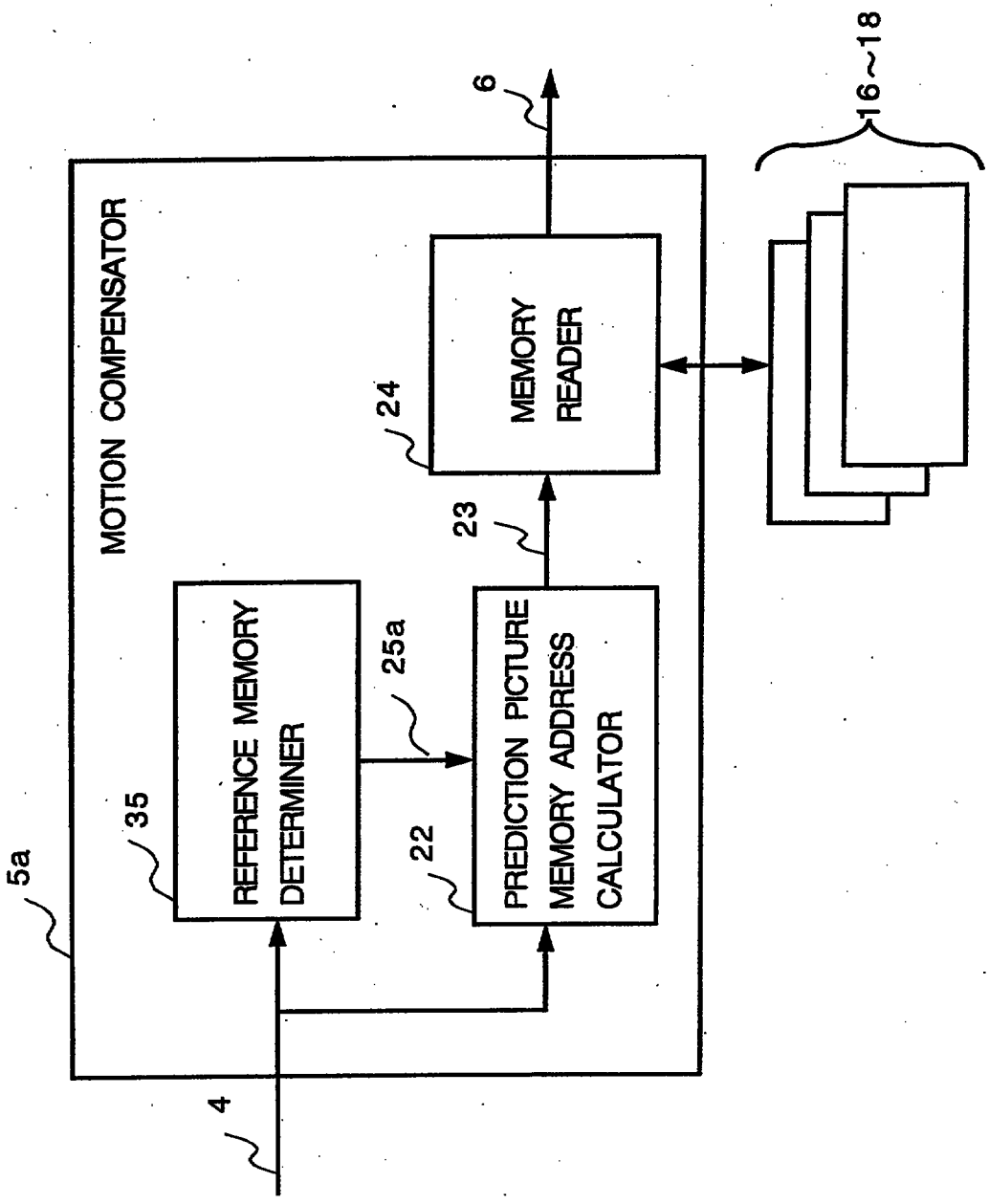

MOTION COMPENSATOR
REFERENCE MEMORY DETERMINER
PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
MEMORY READER
4
5a
6
16~18
22
23
24
25a
35

Fig.8

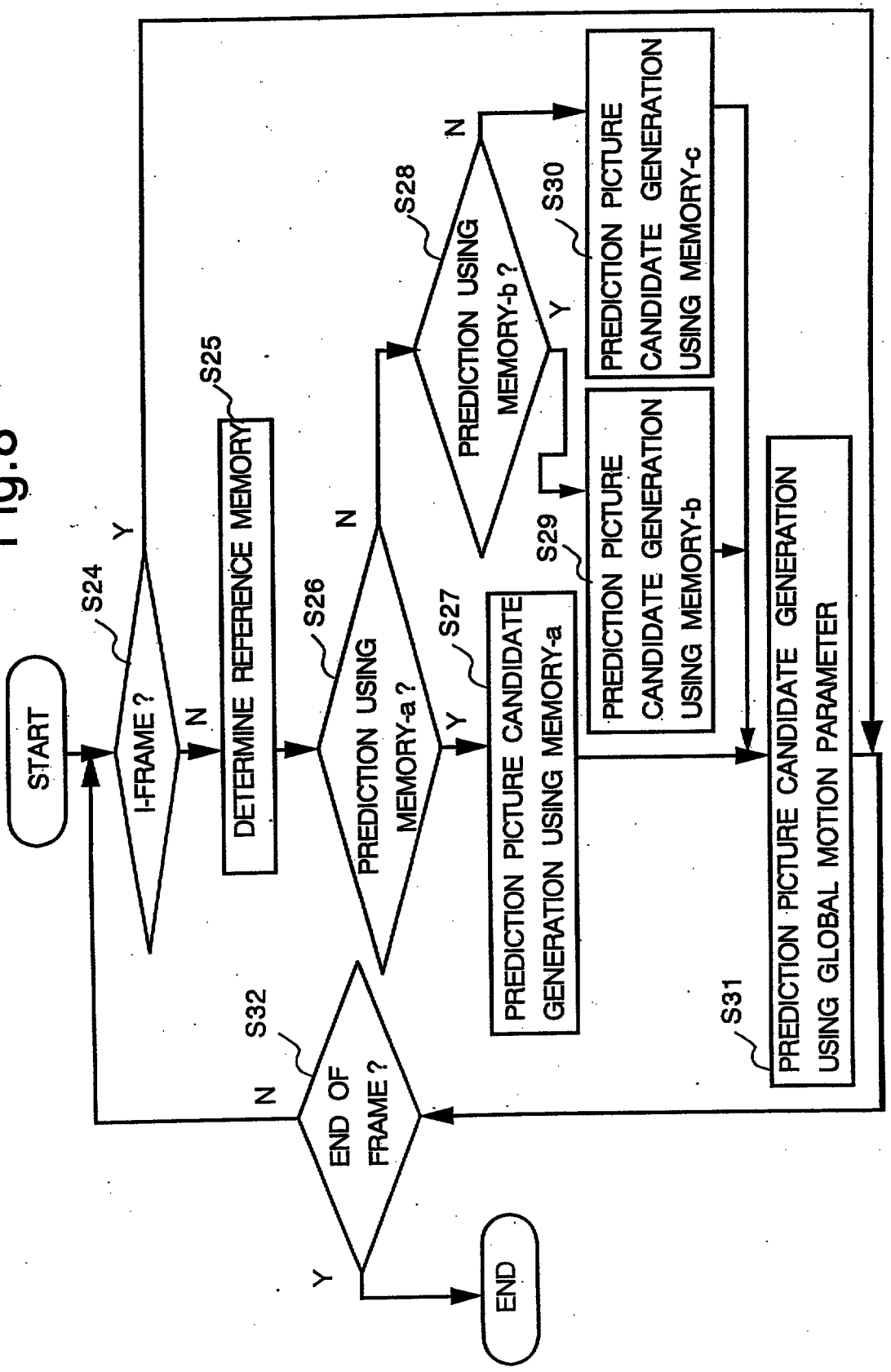

START
I-FRAME ?
Y
N
S24
DETERMINE REFERENCE MEMORY
S25
PREDICTION USING MEMORY-a ?
N
Y
S26
PREDICTION USING MEMORY-b ?
N
Y
S28
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-a
S27
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-b
S29
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-c
S30
PREDICTION PICTURE CANDIDATE GENERATION USING GLOBAL MOTION PARAMETER
S31
END OF FRAME ?
N
Y
S32
END

Fig.9

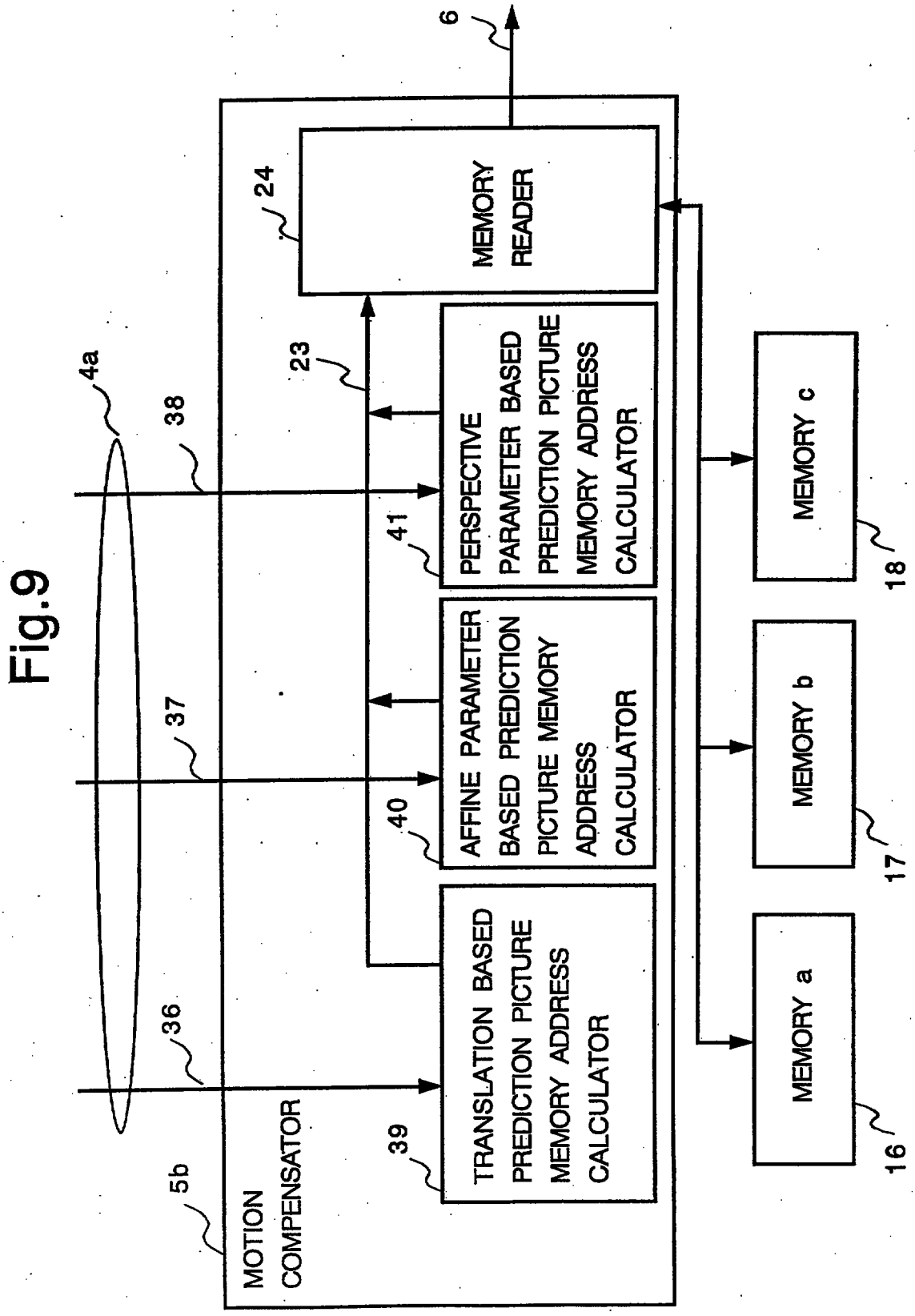

MOTION COMPENSATOR
5b
TRANSLATION BASED PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
39
AFFINE PARAMETER BASED PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
40
PERSPECTIVE PARAMETER BASED PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
41
MEMORY READER
24
6
4a
36
37
38
23
MEMORY a
16
MEMORY b
17
MEMORY c
18

Fig.10

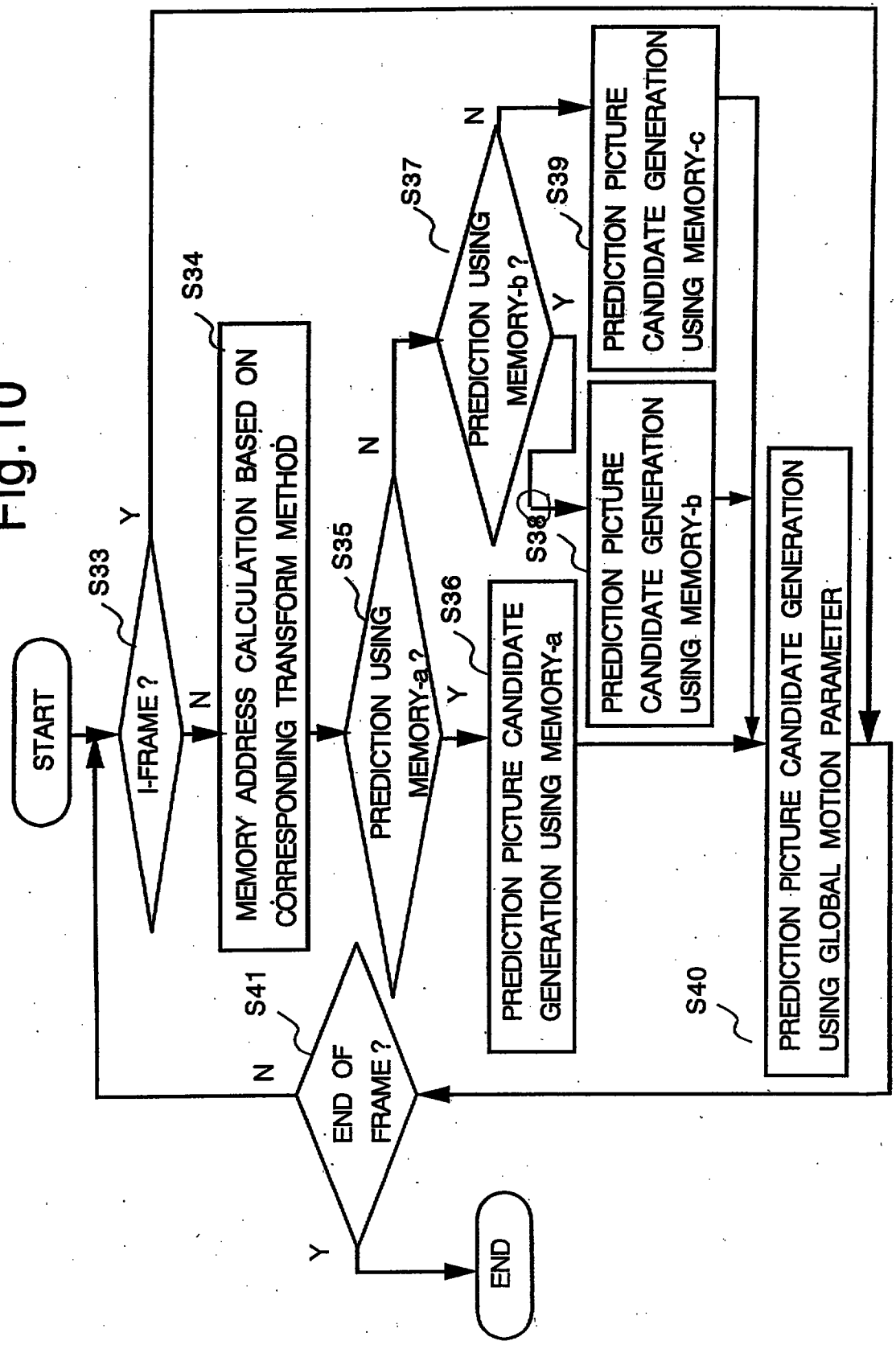

START
S33
I-FRAME ?
Y
N
S34
MEMORY ADDRESS CALCULATION BASED ON CORRESPONDING TRANSFORM METHOD
S35
PREDICTION USING MEMORY-a ?
N
Y
S36
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-a
S37
PREDICTION USING MEMORY-b ?
N
Y
S38
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-b
S39
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-c
S40
PREDICTION PICTURE CANDIDATE GENERATION USING GLOBAL MOTION PARAMETER
S41
END OF FRAME ?
N
Y
END

Fig.11

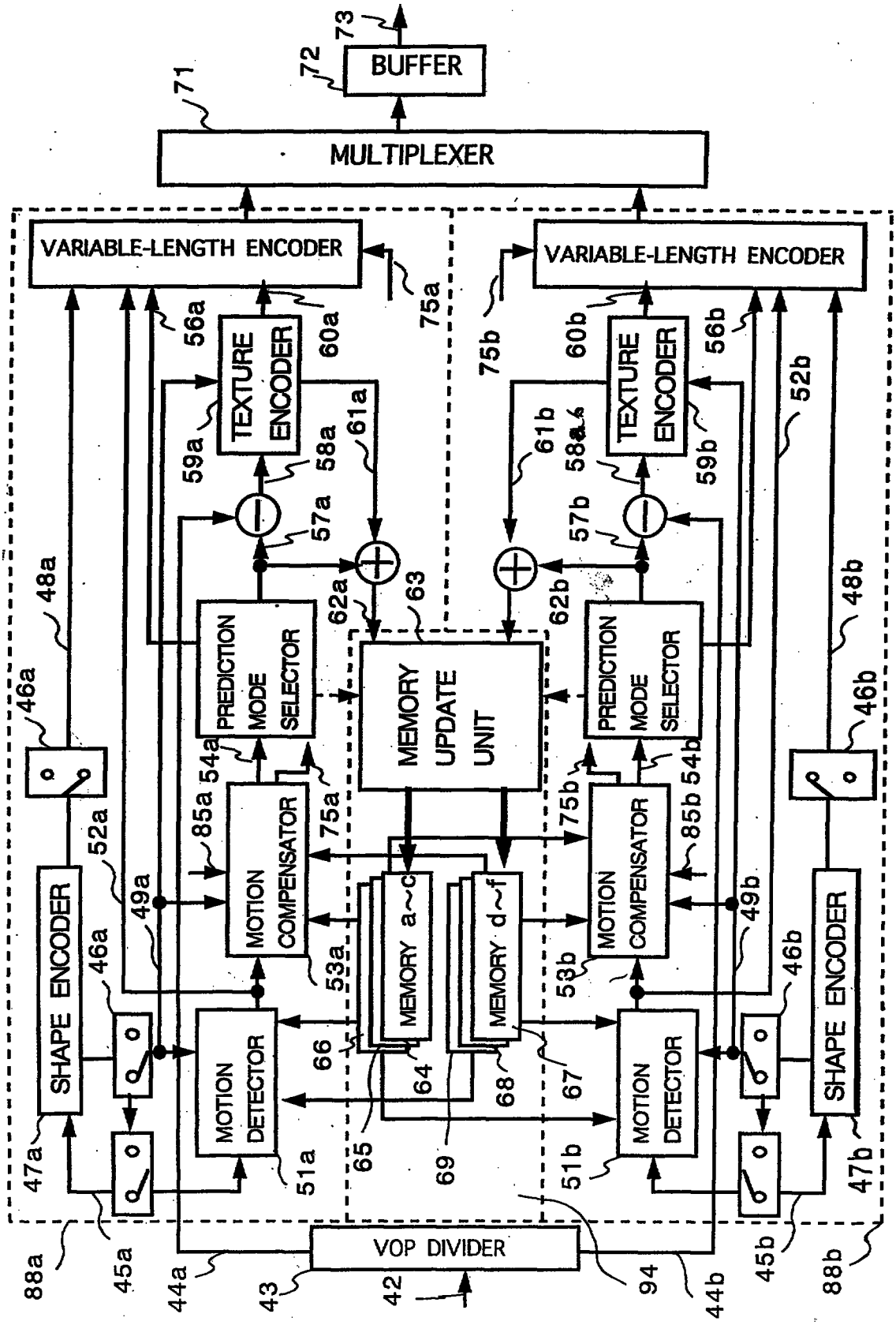

BUFFER
MULTIPLEXER
VARIABLE-LENGTH ENCODER
VARIABLE-LENGTH ENCODER
TEXTURE ENCODER
TEXTURE ENCODER
PREDICTION MODE SELECTOR
PREDICTION MODE SELECTOR
MEMORY UPDATE UNIT
MOTION COMPENSATOR
MOTION COMPENSATOR
MEMORY a~c
MEMORY d~f
MOTION DETECTOR
MOTION DETECTOR
SHAPE ENCODER
SHAPE ENCODER
VOP DIVIDER

Fig.12

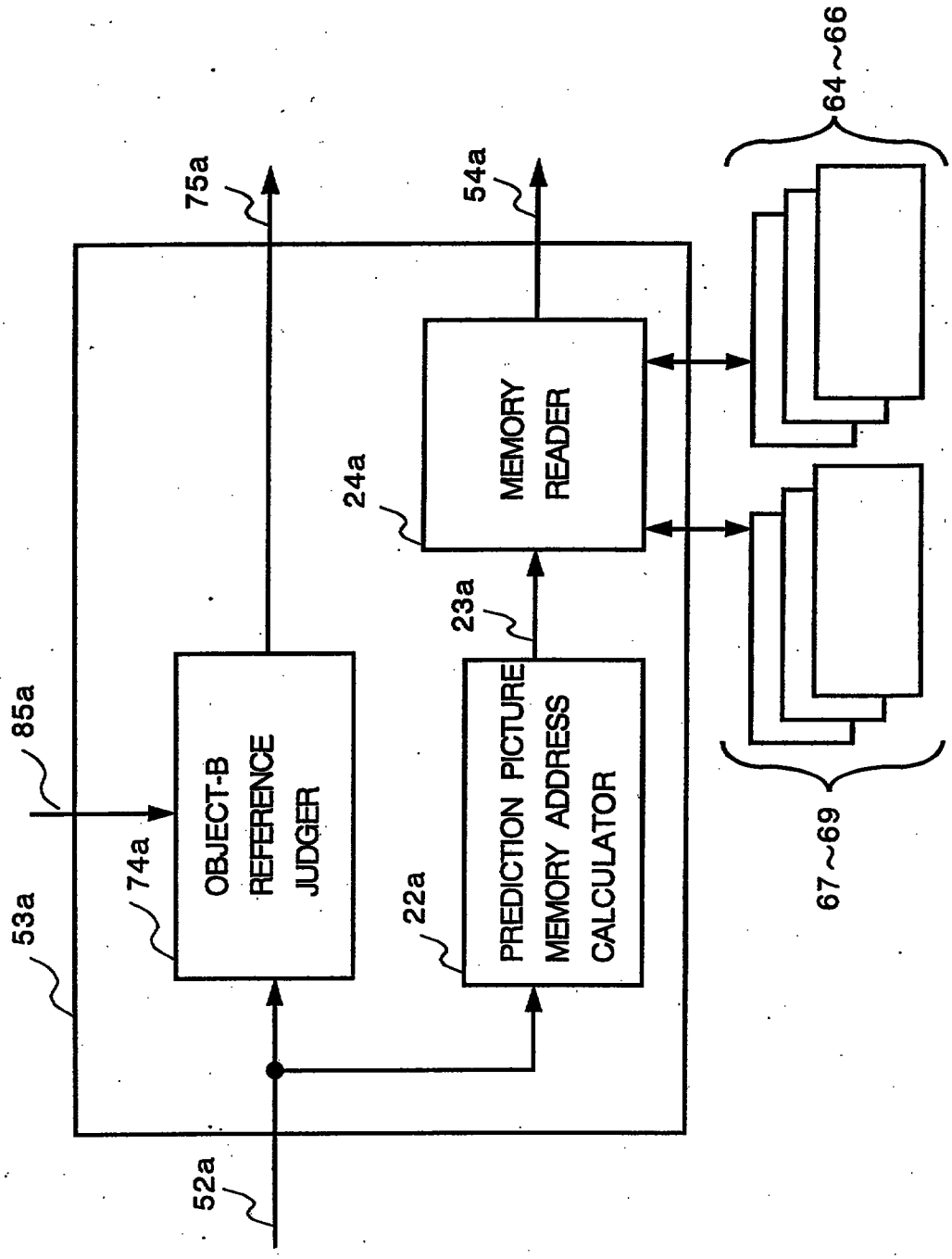

64~66
75a
54a
MEMORY READER
24a
23a
85a
OBJECT-B REFERENCE JUDGER
PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
74a
22a
53a
67~69
52a

Fig.13

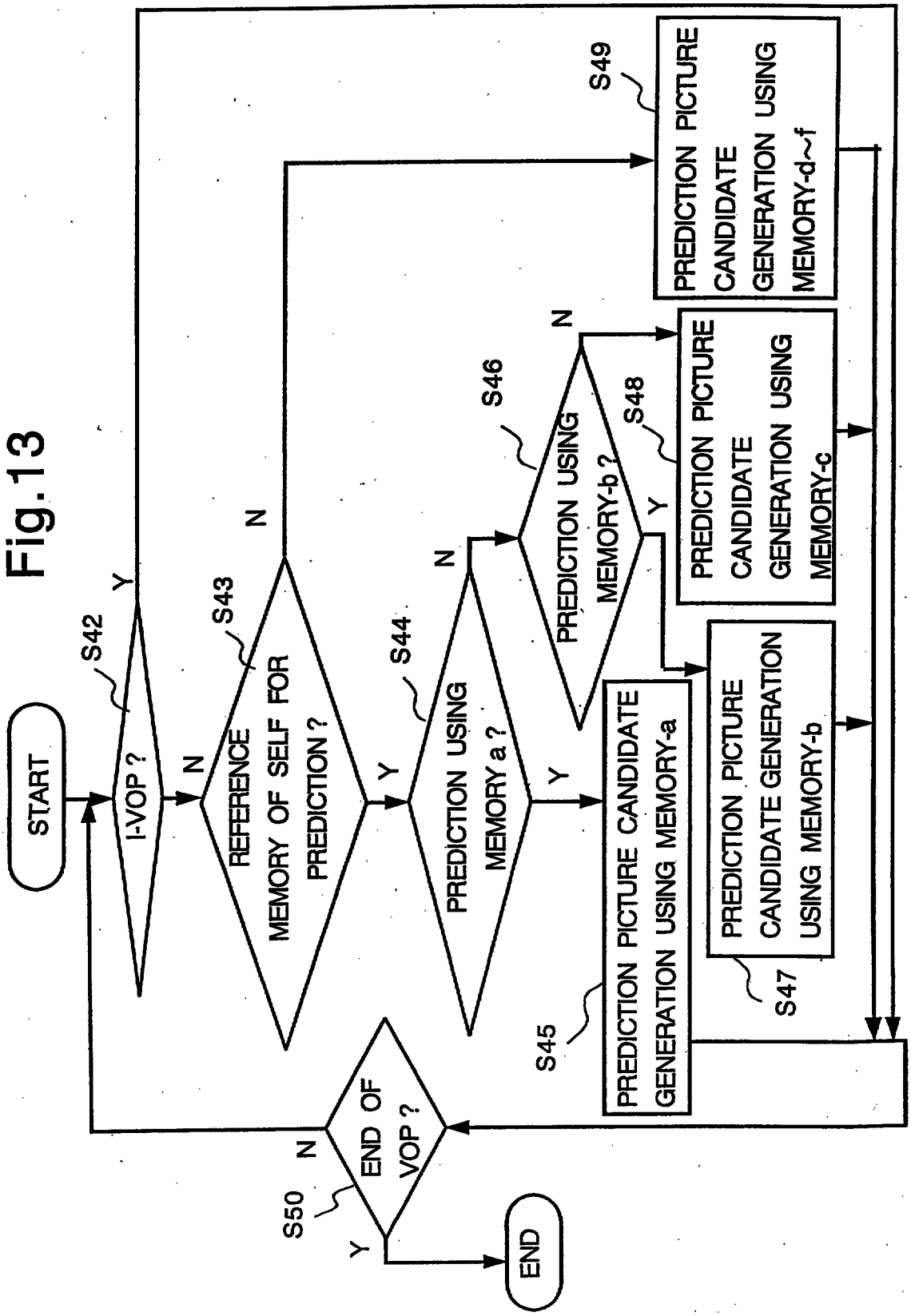

START
S42
I-VOP ?
Y
N
S43
REFERENCE MEMORY OF SELF FOR PREDICTION ?
N
Y
S44
PREDICTION USING MEMORY a ?
N
Y
S46
PREDICTION USING MEMORY-b ?
N
Y
S48
S49
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-d~f
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-c
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-b
S47
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-a
S45
S50
END OF VOP ?
N
Y
END

Fig.14

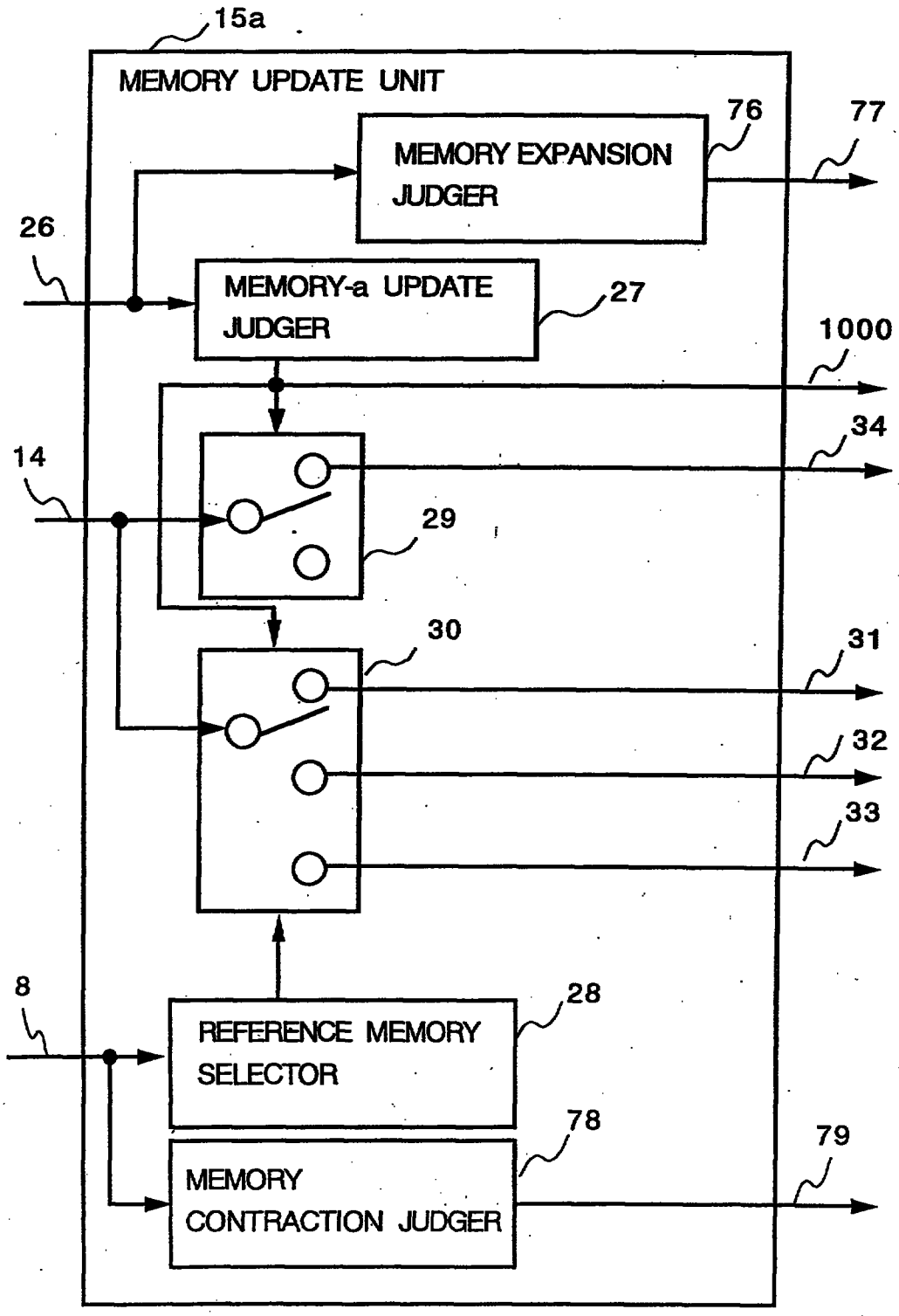

15a
MEMORY UPDATE UNIT
76
77
MEMORY EXPANSION JUDGER
26
MEMORY-a UPDATE JUDGER
27
1000
34
14
29
30
31
32
33
8
28
REFERENCE MEMORY SELECTOR
78
79
MEMORY CONTRACTION JUDGER

Fig.15

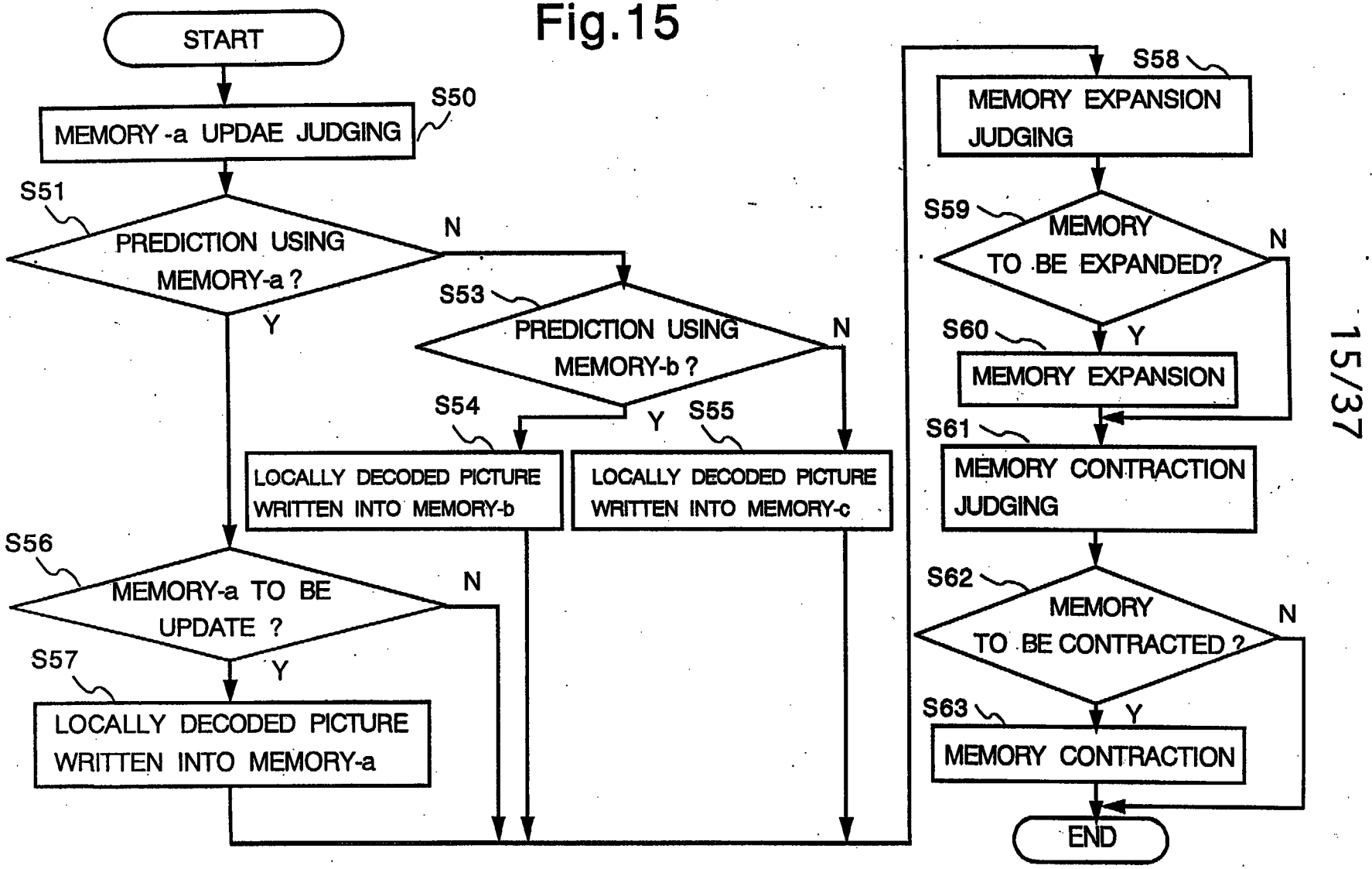

START
S50
MEMORY -a UPDAE JUDGING
S51
PREDICTION USING MEMORY-a ?
N
Y
S53
PREDICTION USING MEMORY-b ?
N
Y
S54
S55
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-b
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-c
S56
MEMORY-a TO BE UPDATE ?
N
Y
S57
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-a
S58
MEMORY EXPANSION JUDGING
S59
MEMORY TO BE EXPANDED?
N
Y
S60
MEMORY EXPANSION
S61
MEMORY CONTRACTION JUDGING
S62
MEMORY TO BE CONTRACTED ?
N
Y
S63
MEMORY CONTRACTION
END

Fig.16

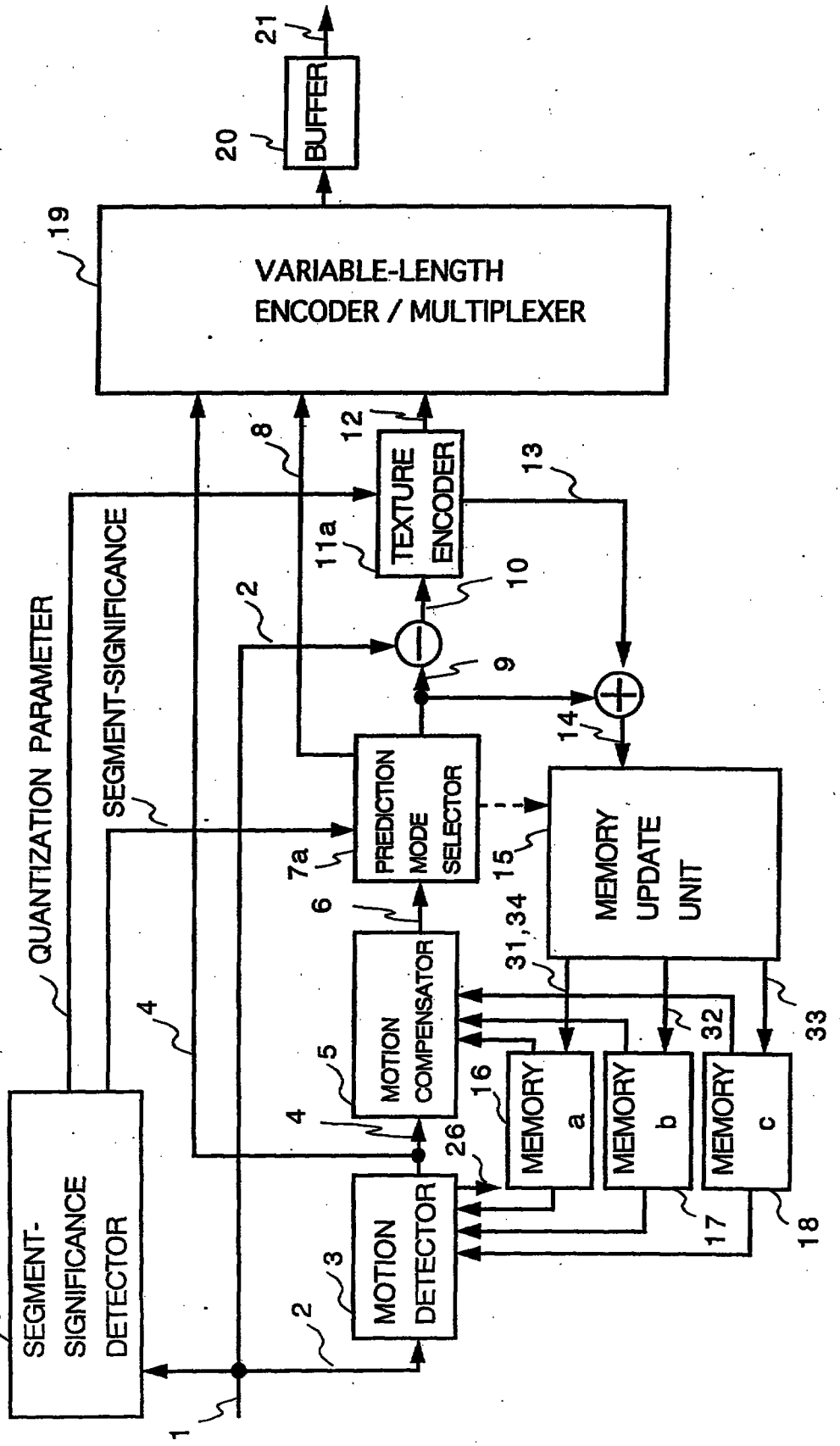

21
BUFFER
20
19
VARIABLE-LENGTH
ENCODER / MULTIPLEXER
8
12
TEXTURE
ENCODER
13
11a
10
2
9
14
QUANTIZATION PARAMETER
SEGMENT-SIGNIFICANCE
PREDICTION
MODE
SELECTOR
7a
15
MEMORY
UPDATE
UNIT
6
31,34
4
MOTION
COMPENSATOR
5
16
MEMORY
a
MEMORY
b
MEMORY
c
32
33
26
4
17
18
95
SEGMENT-
SIGNIFICANCE
DETECTOR
MOTION
DETECTOR
3
2
1

Fig.17

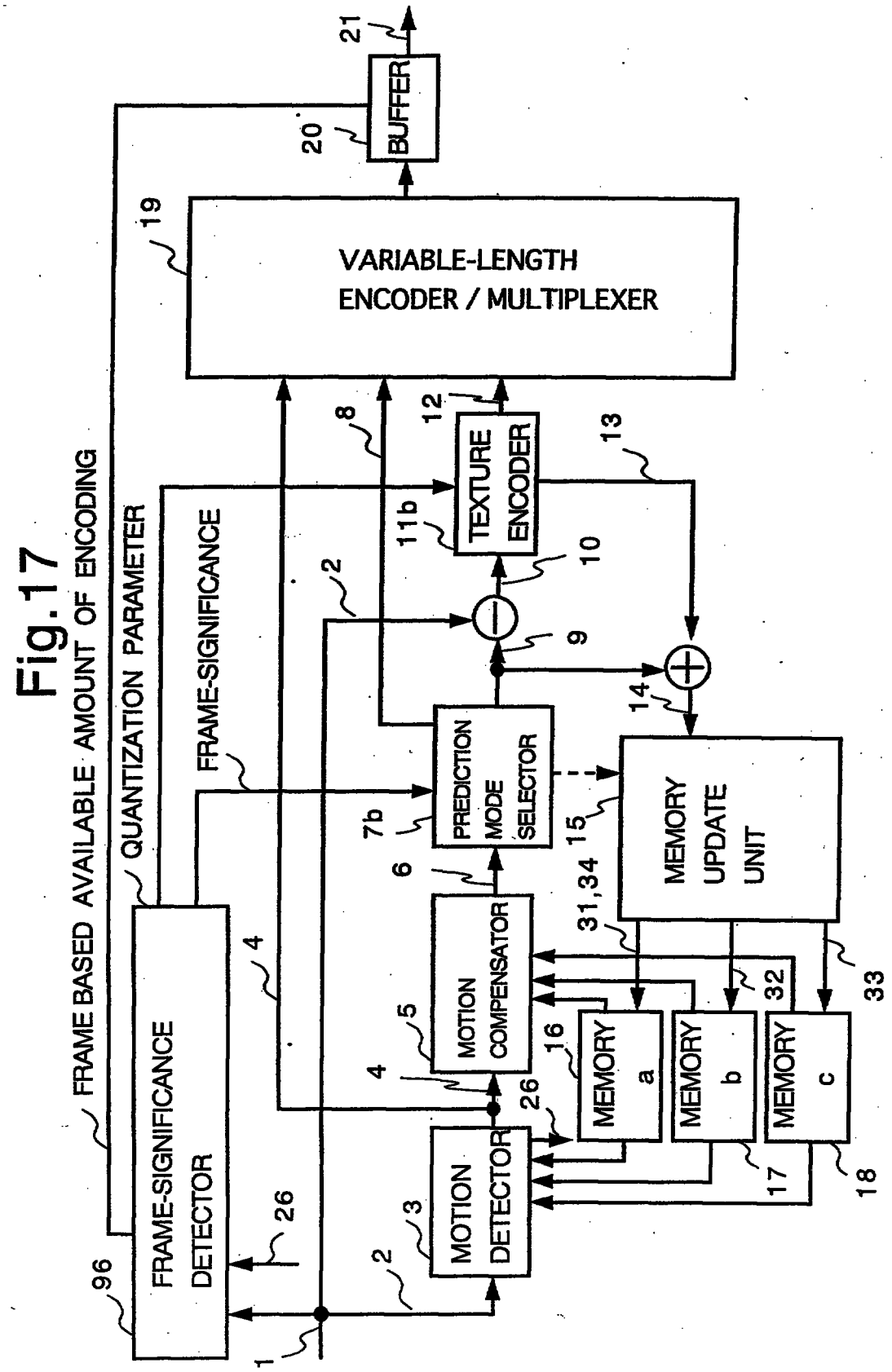

FRAME BASED AVAILABLE AMOUNT OF ENCODING
QUANTIZATION PARAMETER
FRAME-SIGNIFICANCE
FRAME-SIGNIFICANCE DETECTOR
MOTION DETECTOR
MOTION COMPENSATOR
PREDICTION MODE SELECTOR
TEXTURE ENCODER
VARIABLE-LENGTH ENCODER / MULTIPLEXER
BUFFER
MEMORY UPDATE UNIT
MEMORY a
MEMORY b
MEMORY c
96
1
2
3
4
5
6
7b
8
9
10
11b
12
13
14
15
16
17
18
19
20
21
26
31,34
32
33

Fig.18

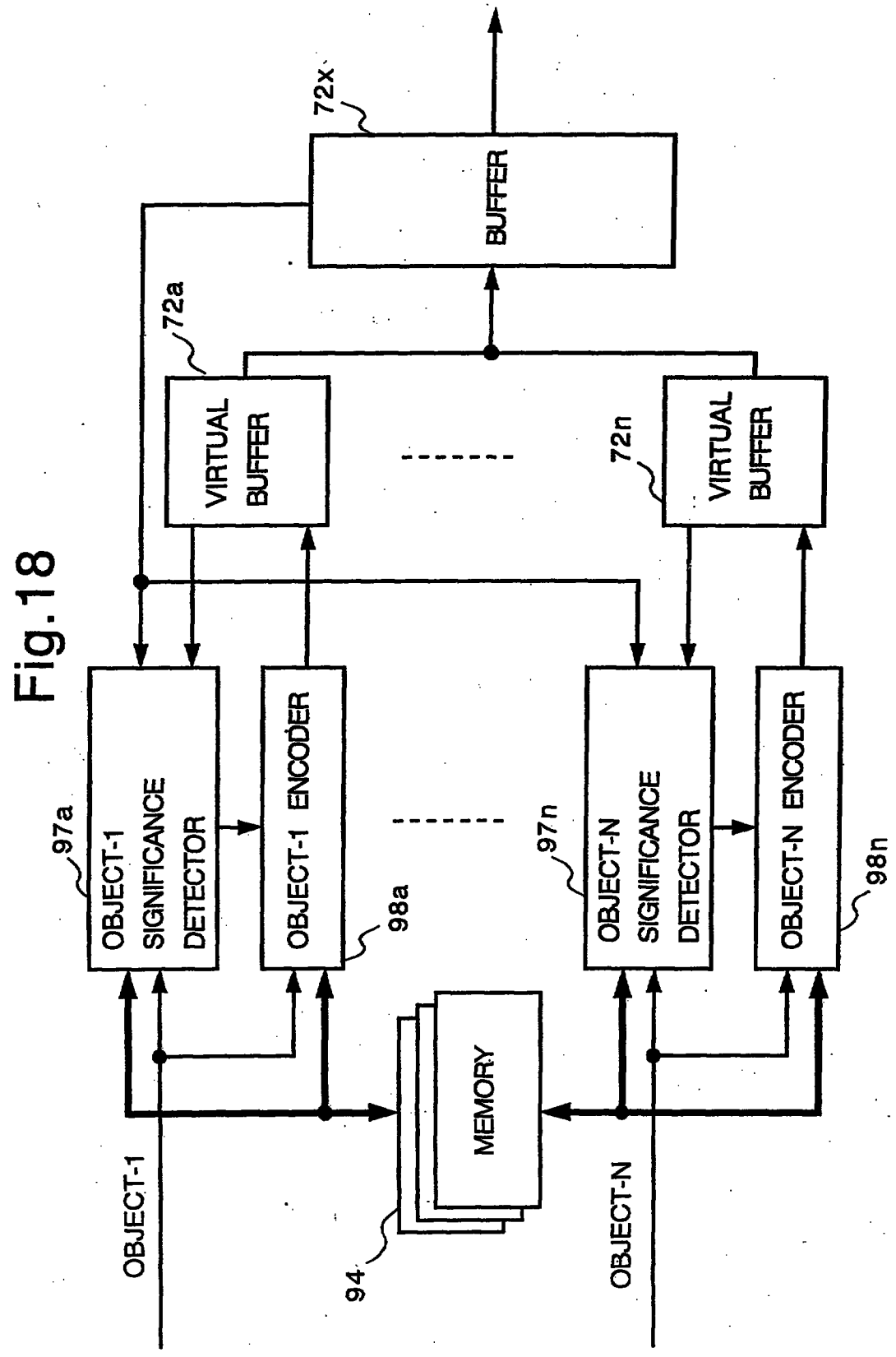

72x
BUFFER
72a
VIRTUAL BUFFER
72n
VIRTUAL BUFFER
97a
OBJECT-1 SIGNIFICANCE DETECTOR
OBJECT-1 ENCODER
98a
97n
OBJECT-N SIGNIFICANCE DETECTOR
OBJECT-N ENCODER
98n
MEMORY
94
OBJECT-1
OBJECT-N

Fig.19

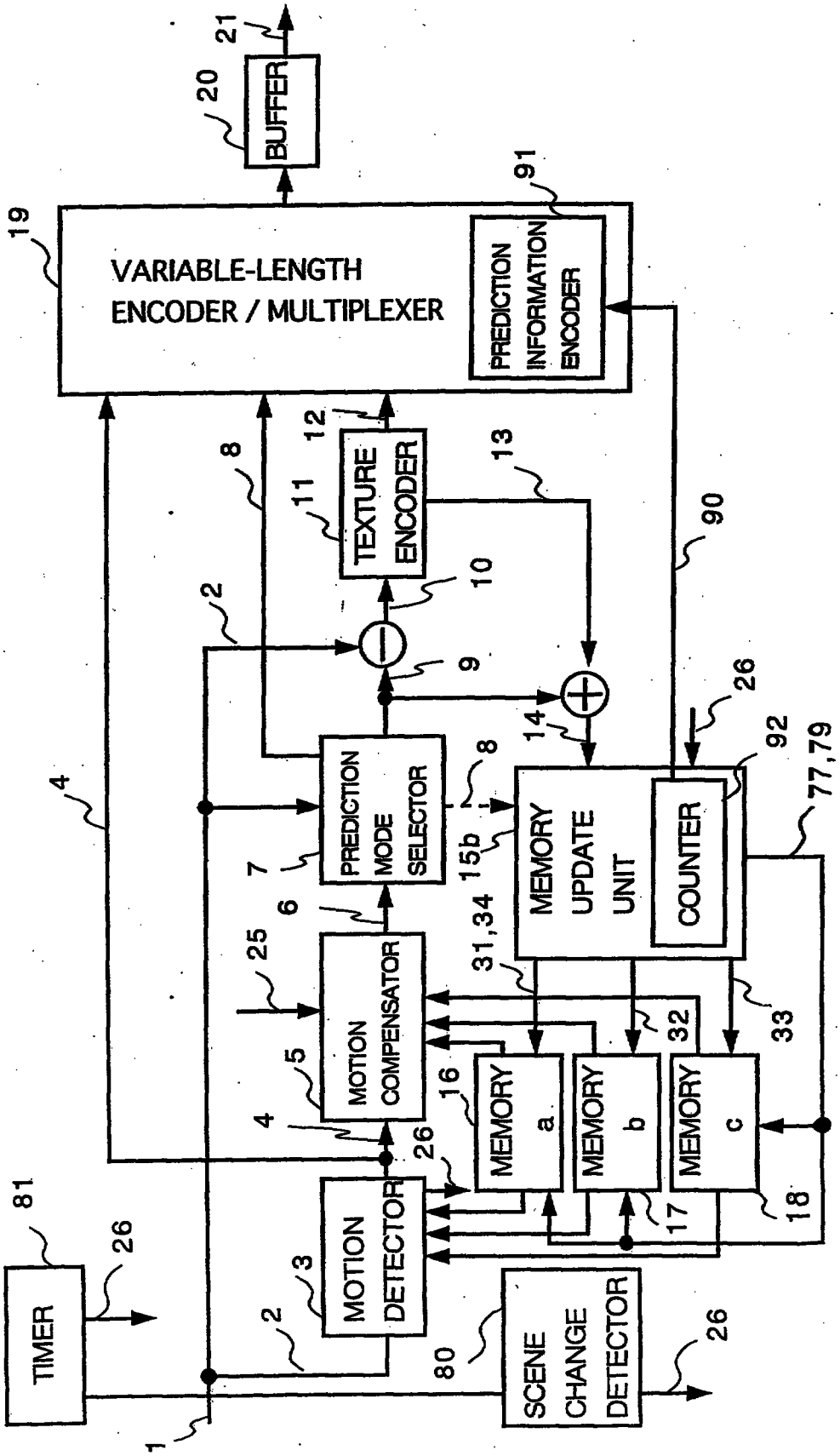

BUFFER
VARIABLE-LENGTH ENCODER / MULTIPLEXER
PREDICTION INFORMATION ENCODER
TEXTURE ENCODER
PREDICTION MODE SELECTOR
MOTION COMPENSATOR
MOTION DETECTOR
MEMORY UPDATE UNIT
COUNTER
MEMORY a
MEMORY b
MEMORY c
TIMER
SCENE CHANGE DETECTOR

Fig. 20

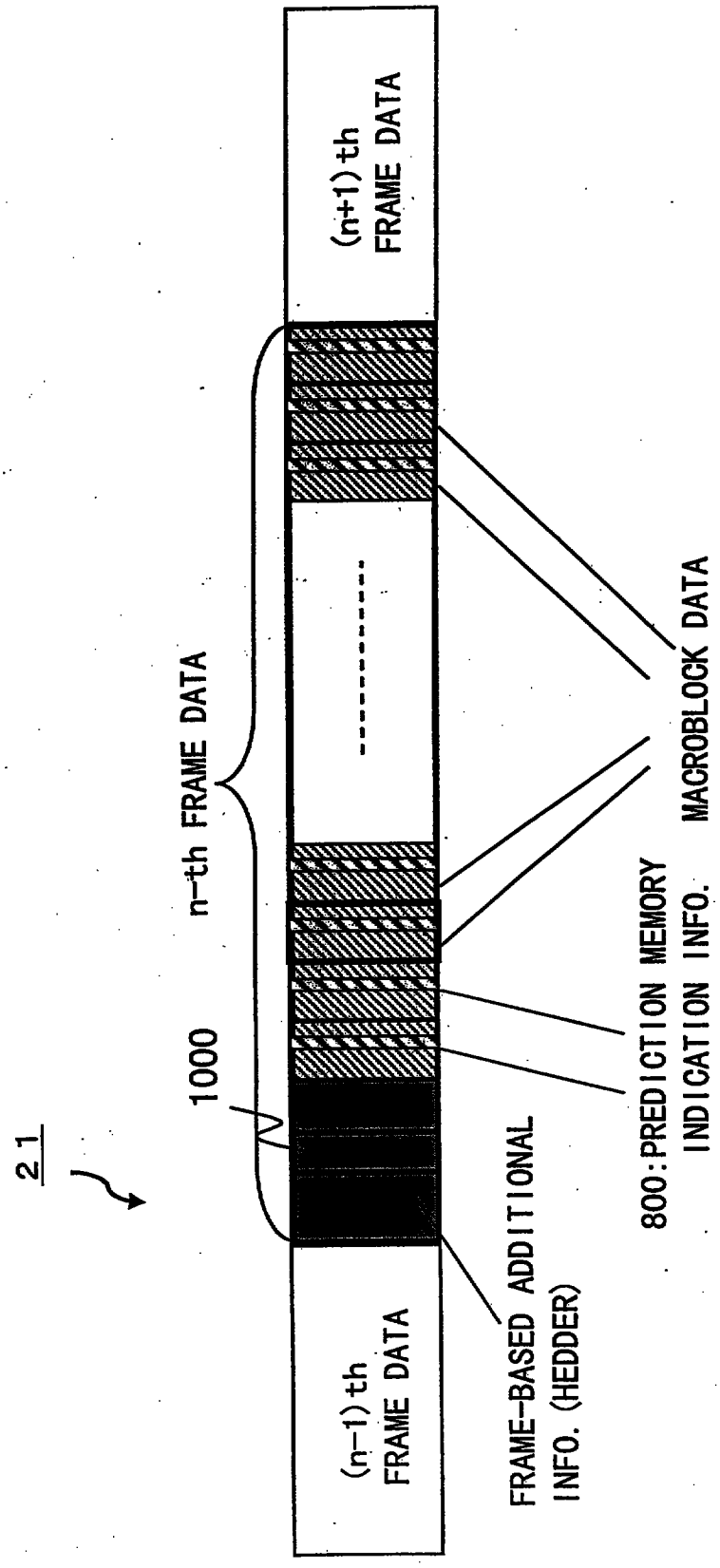

21
1000
n-th FRAME DATA
(n+1)th FRAME DATA
(n-1)th FRAME DATA
FRAME-BASED ADDITIONAL INFO. (HEDDER)
800:PREDICTION MEMORY INDICATION INFO.
MACROBLOCK DATA

Fig. 21

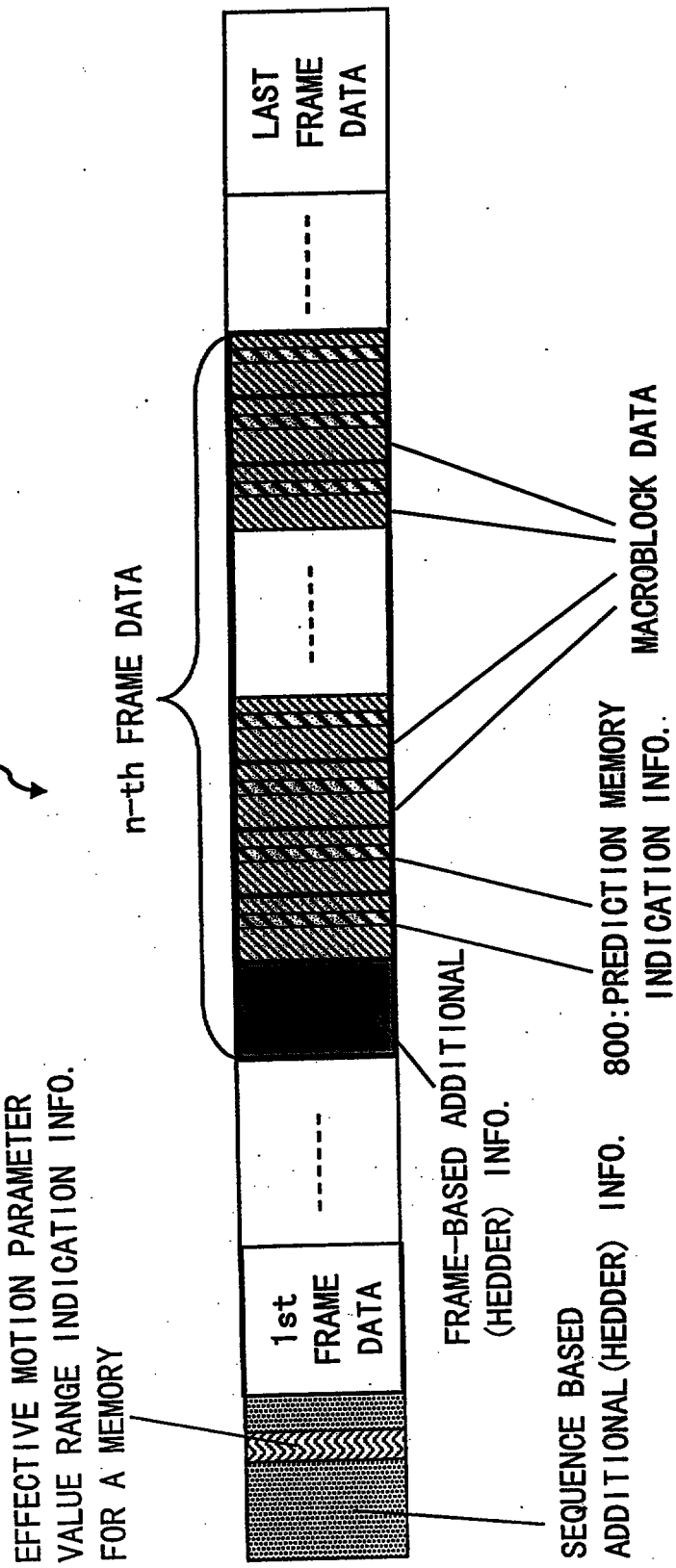

21
n-th FRAME DATA
LAST FRAME DATA
1st FRAME DATA
MACROBLOCK DATA
800:PREDICTION MEMORY INDICATION INFO..
FRAME-BASED ADDITIONAL (HEDDER) INFO.
SEQUENCE BASED ADDITIONAL (HEDDER) INFO.
EFFECTIVE MOTION PARAMETER VALUE RANGE INDICATION INFO. FOR A MEMORY

Fig. 22

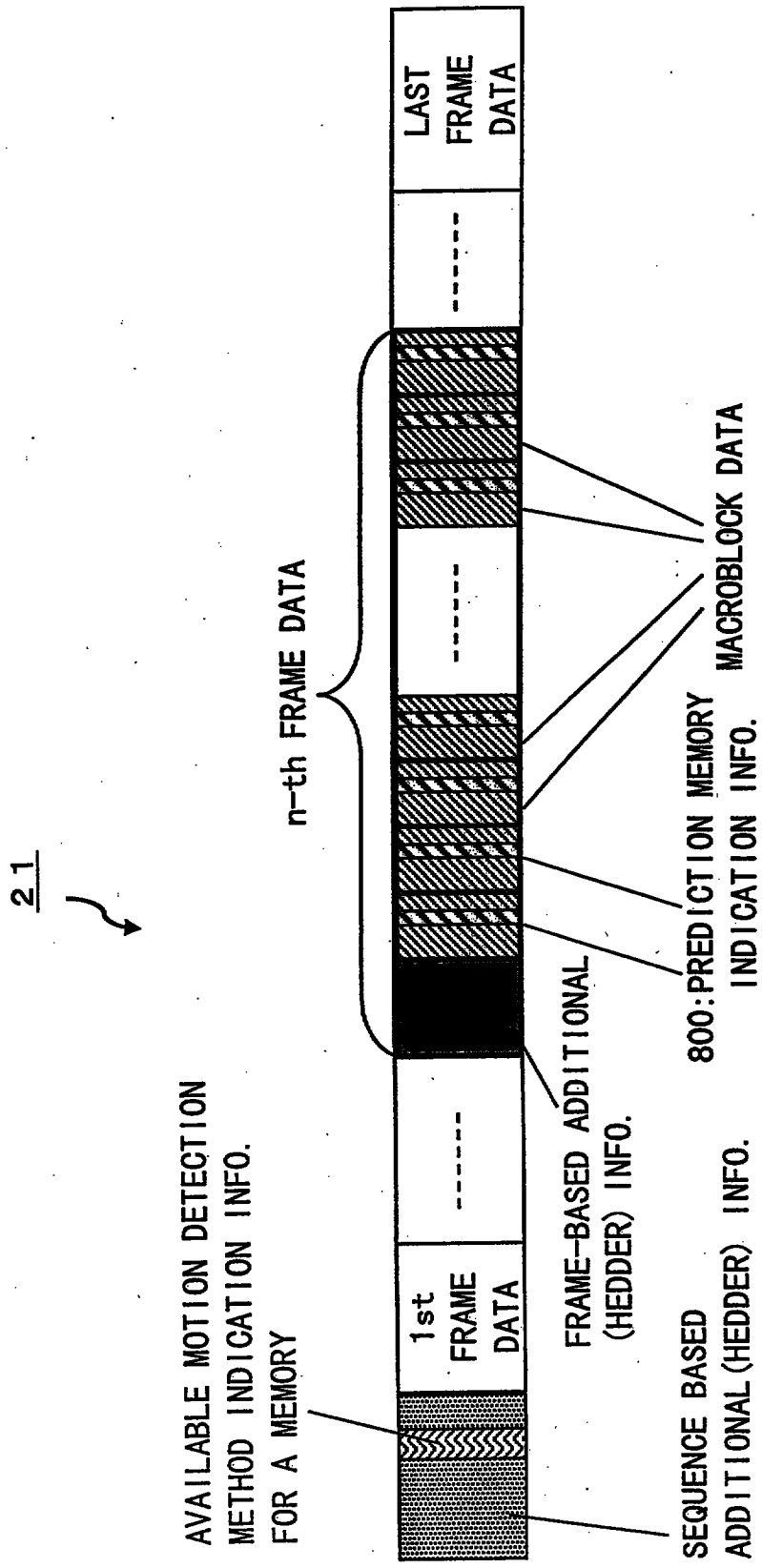

21
AVAILABLE MOTION DETECTION METHOD INDICATION INFO. FOR A MEMORY
SEQUENCE BASED ADDITIONAL (HEDDER) INFO.
1st FRAME DATA
FRAME-BASED ADDITIONAL (HEDDER) INFO.
n-th FRAME DATA
LAST FRAME DATA
800: PREDICTION MEMORY INDICATION INFO.
MACROBLOCK DATA

Fig. 23

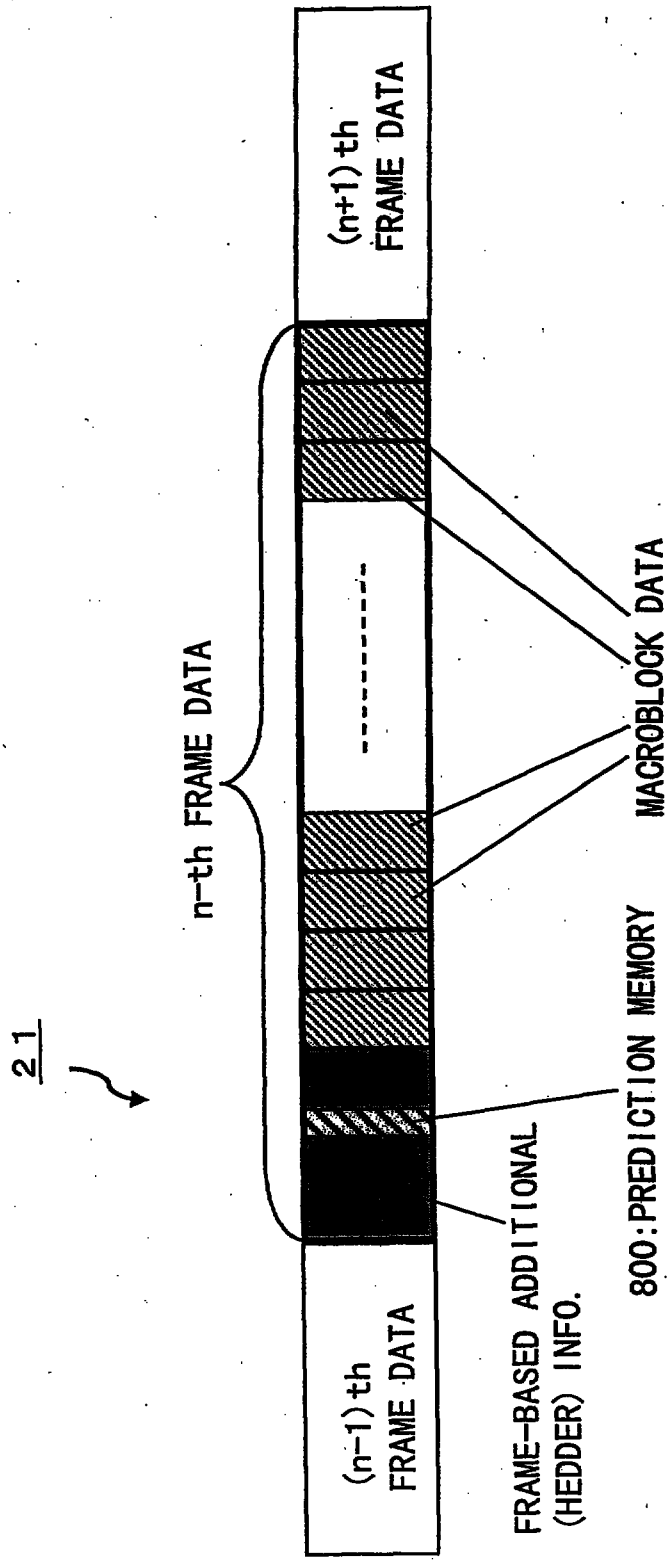

21
n-th FRAME DATA
(n+1)th FRAME DATA
(n-1)th FRAME DATA
MACROBLOCK DATA
800:PREDICTION MEMORY INDICATION INFO.
FRAME-BASED ADDITIONAL (HEDDER) INFO.

Fig. 24

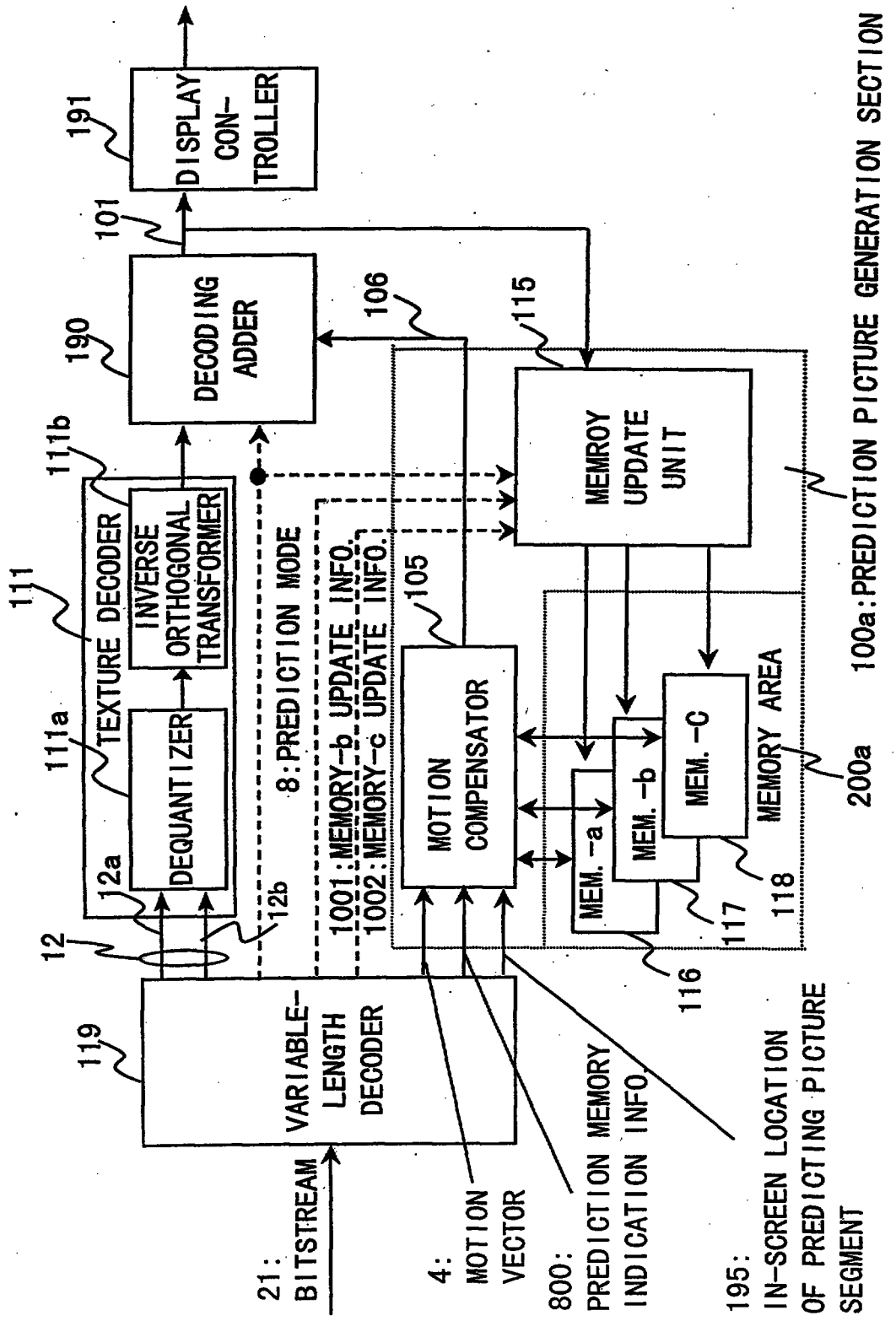

DISPLAY CON-TROLLER
191
DECODING ADDER
190
101
106
115
MEMROY UPDATE UNIT
111b
INVERSE ORTHOGONAL TRANSFORMER
111
TEXTURE DECODER
111a
DEQUANTIZER
12a
12b
12
8:PREDICTION MODE
1001:MEMORY-b UPDATE INFO.
1002:MEMORY-c UPDATE INFO.
105
MOTION COMPENSATOR
MEM. -a
MEM. -b
MEM. -C
MEMORY AREA
200a
118
117
116
119
VARIABLE-LENGTH DECODER
100a:PREDICTION PICTURE GENERATION SECTION
21: BITSTREAM
4: MOTION VECTOR
800: PREDICTION MEMORY INDICATION INFO
195: IN-SCREEN LOCATION OF PREDICTING PICTURE SEGMENT

Fig. 25

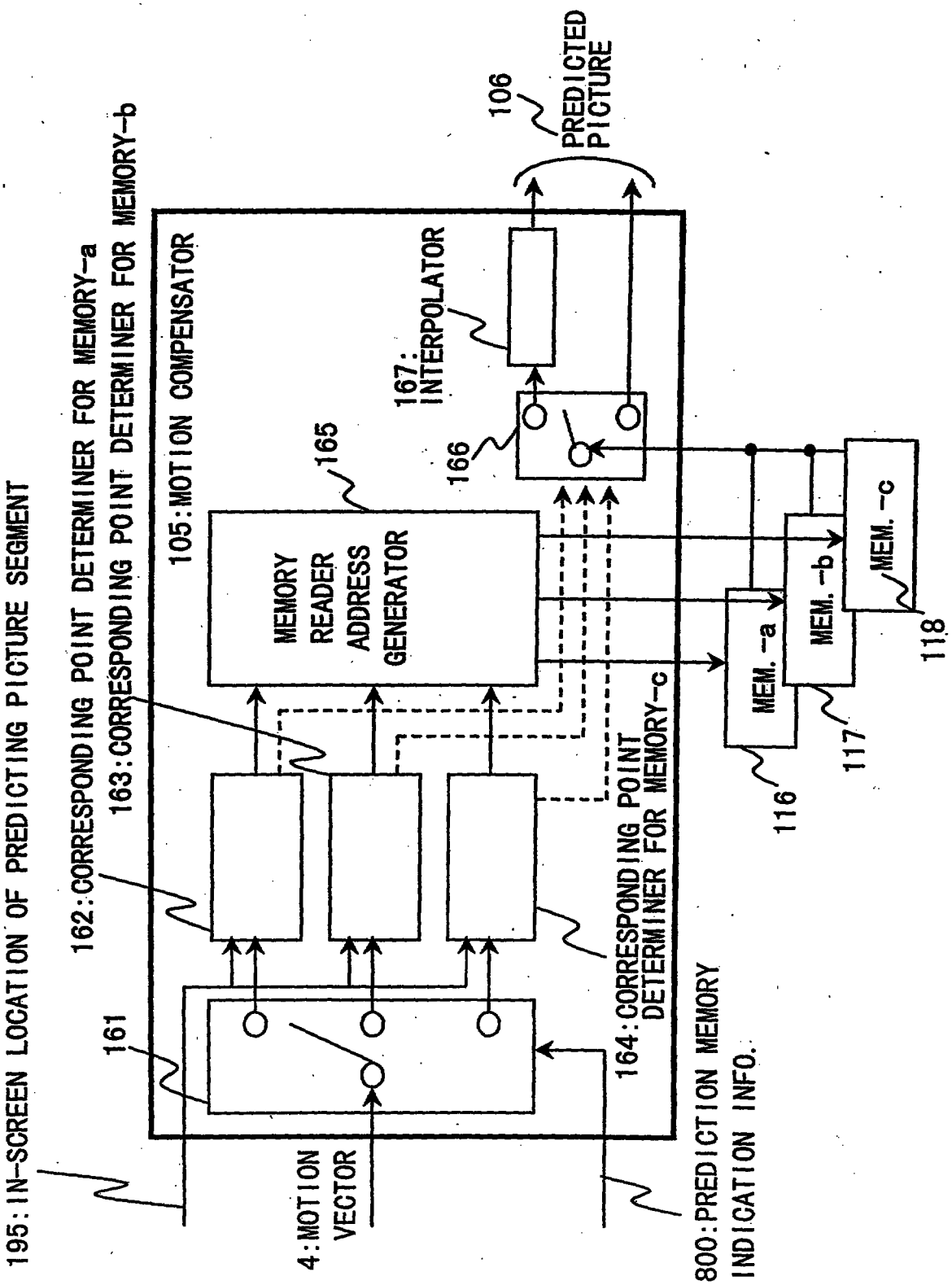

195: IN-SCREEN LOCATION OF PREDICTING PICTURE SEGMENT
162: CORRESPONDING POINT DETERMINER FOR MEMORY-a
163: CORRESPONDING POINT DETERMINER FOR MEMORY-b
105: MOTION COMPENSATOR
106
PREDICTED PICTURE
167: INTERPOLATOR
166
165
MEMORY READER ADDRESS GENERATOR
161
4: MOTION VECTOR
164: CORRESPONDING POINT DETERMINER FOR MEMORY-c
800: PREDICTION MEMORY INDICATION INFO.
MEM. -a
MEM. -b
MEM. -c
116
117
118

Fig. 26

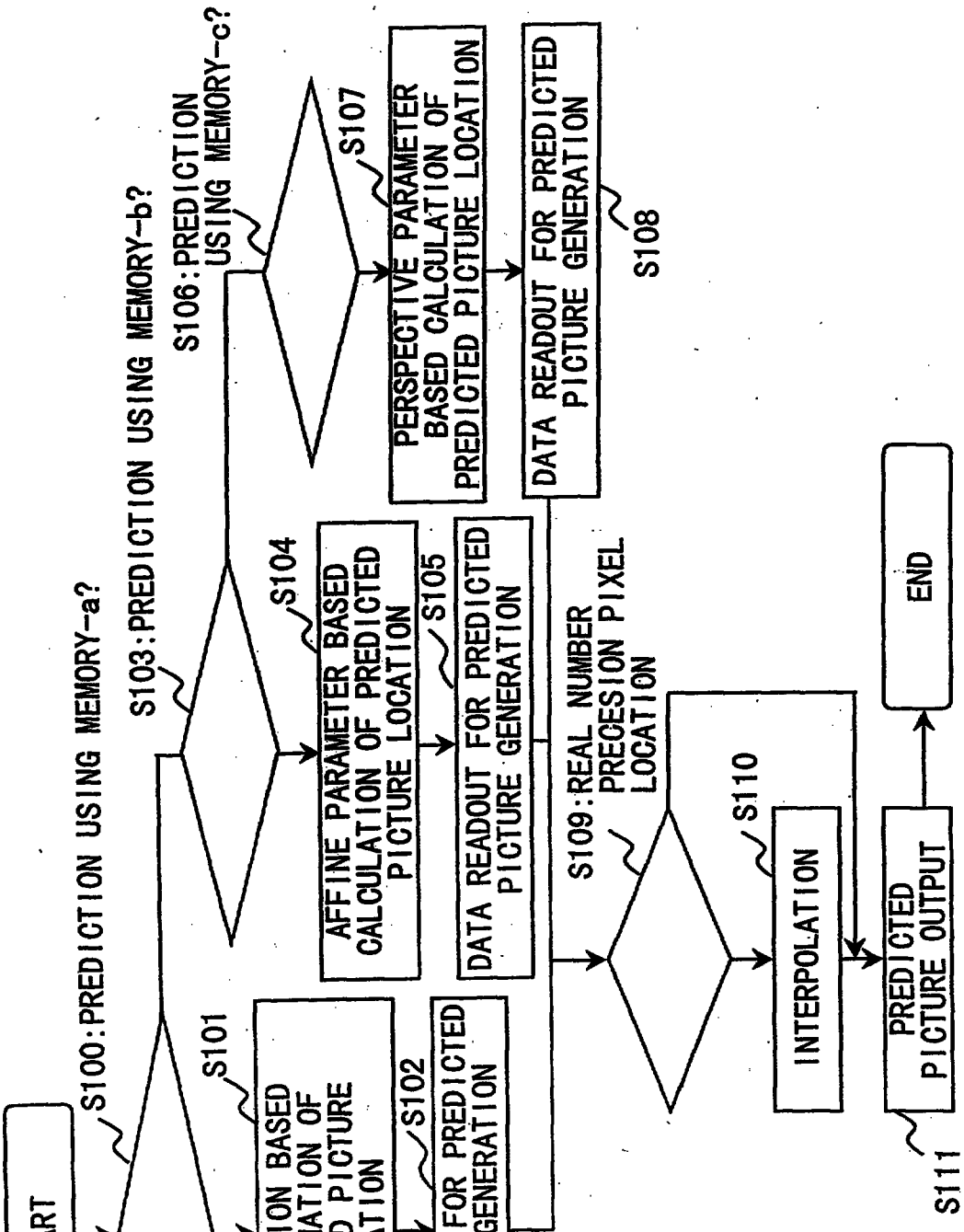

START
S100:PREDICTION USING MEMORY-a?
S101
TRANSLATION BASED DETERMINATION OF PREDICTED PICTURE LOCATION
S102
DATA READOUT FOR PREDICTED PICTURE GENERATION
S103:PREDICTION USING MEMORY-b?
S104
AFFINE PARAMETER BASED CALCULATION OF PREDICTED PICTURE LOCATION
S105
DATA READOUT FOR PREDICTED PICTURE GENERATION
S106:PREDICTION USING MEMORY-c?
S107
PERSPECTIVE PARAMETER BASED CALCULATION OF PREDICTED PICTURE LOCATION
S108
DATA READOUT FOR PREDICTED PICTURE GENERATION
S109:REAL NUMBER PRECESION PIXEL LOCATION
S110
INTERPOLATION
PREDICTED PICTURE OUTPUT
S111
END

Fig. 27

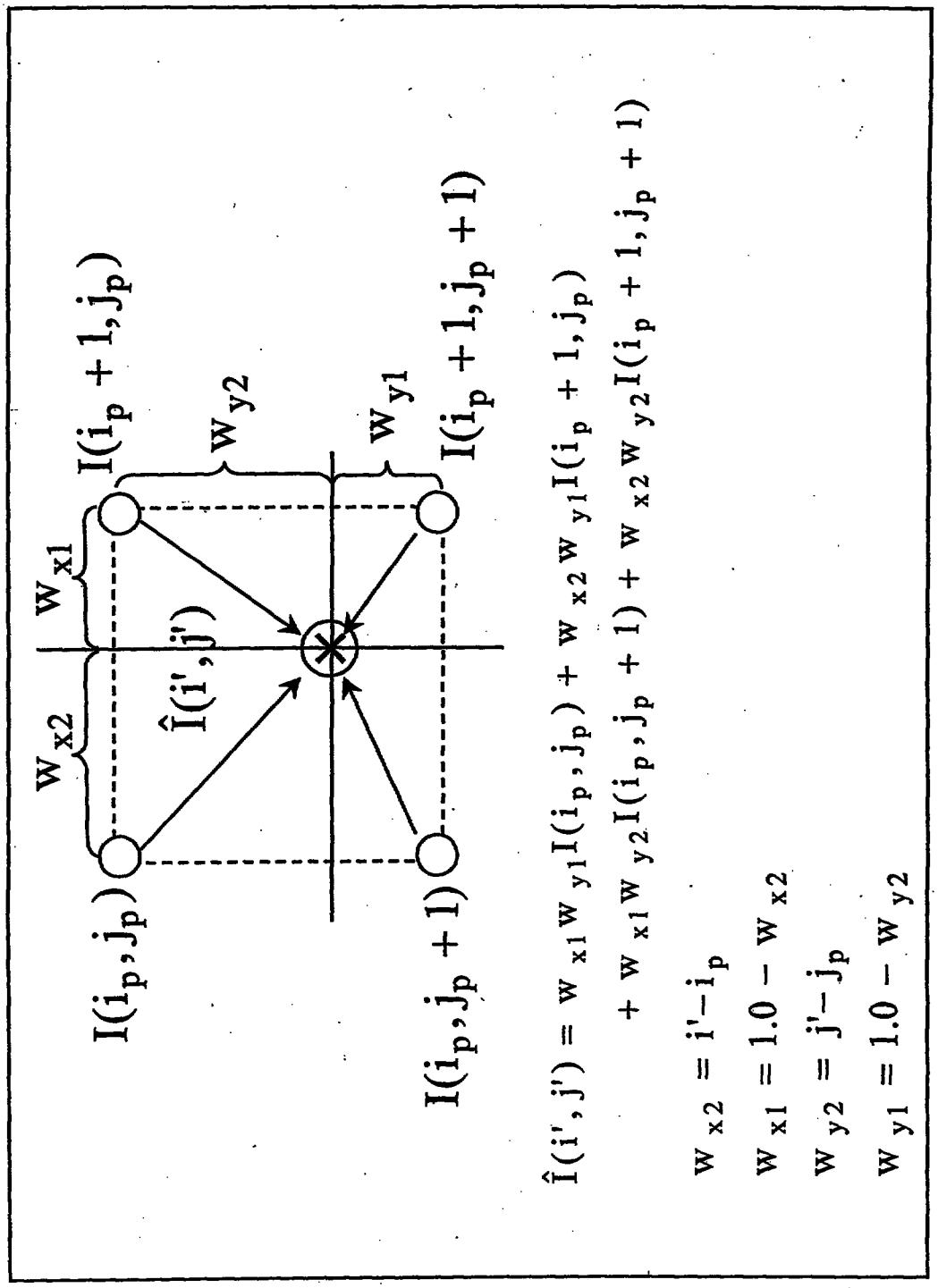

$I(i_p, j_p)$
$I(i_p + 1, j_p)$
$I(i_p, j_p + 1)$
$I(i_p + 1, j_p + 1)$
$\hat{I}(i', j')$
$w_{x1}$
$w_{x2}$
$w_{y1}$
$w_{y2}$
$\hat{I}(i', j') = w_{x1} w_{y1} I(i_p, j_p) + w_{x2} w_{y1} I(i_p + 1, j_p)$
$+ w_{x1} w_{y2} I(i_p, j_p + 1) + w_{x2} w_{y2} I(i_p + 1, j_p + 1)$
$w_{x2} = i' - i_p$
$w_{x1} = 1.0 - w_{x2}$
$w_{y2} = j' - j_p$
$w_{y1} = 1.0 - w_{y2}$

Fig. 28

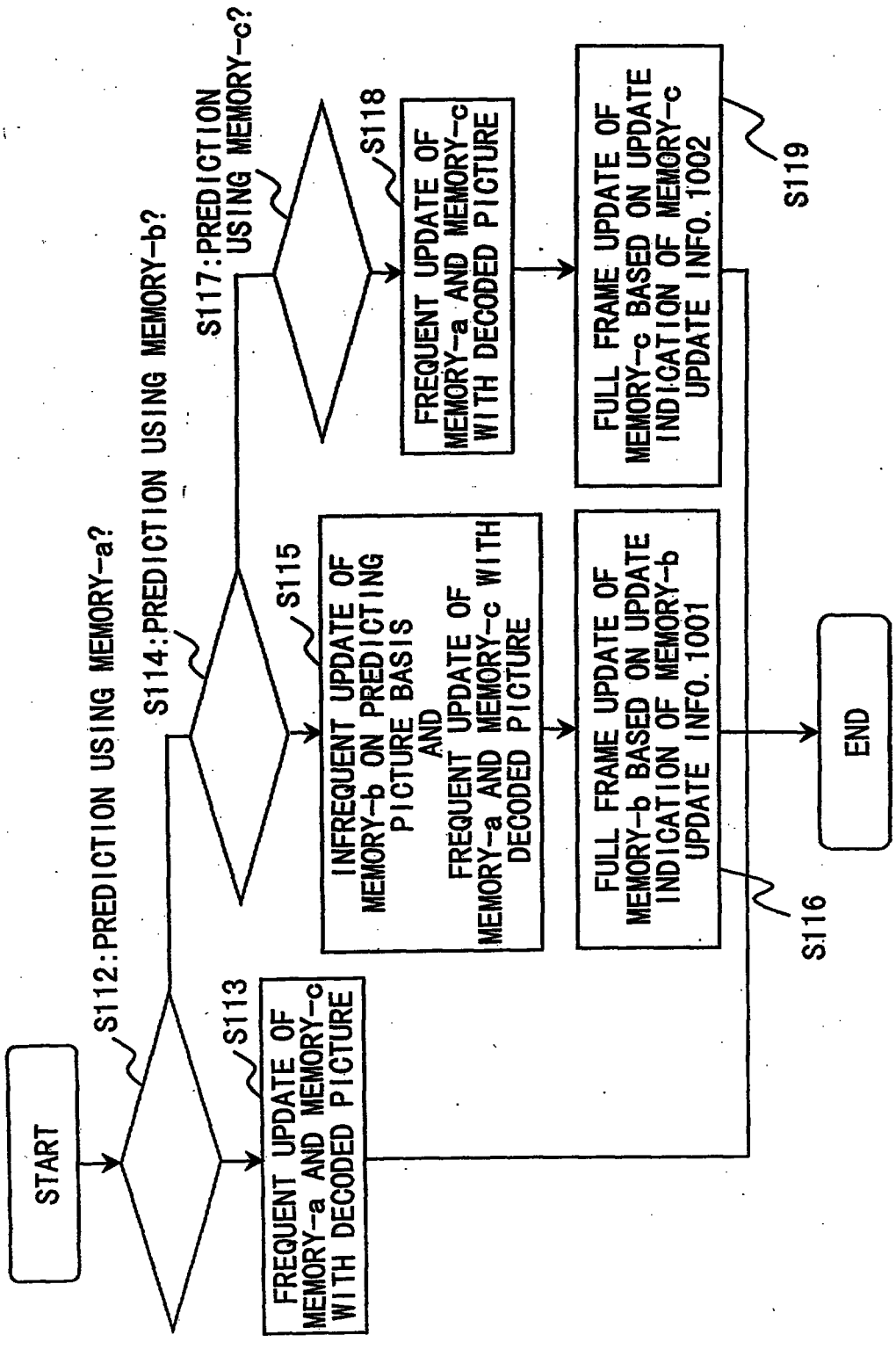
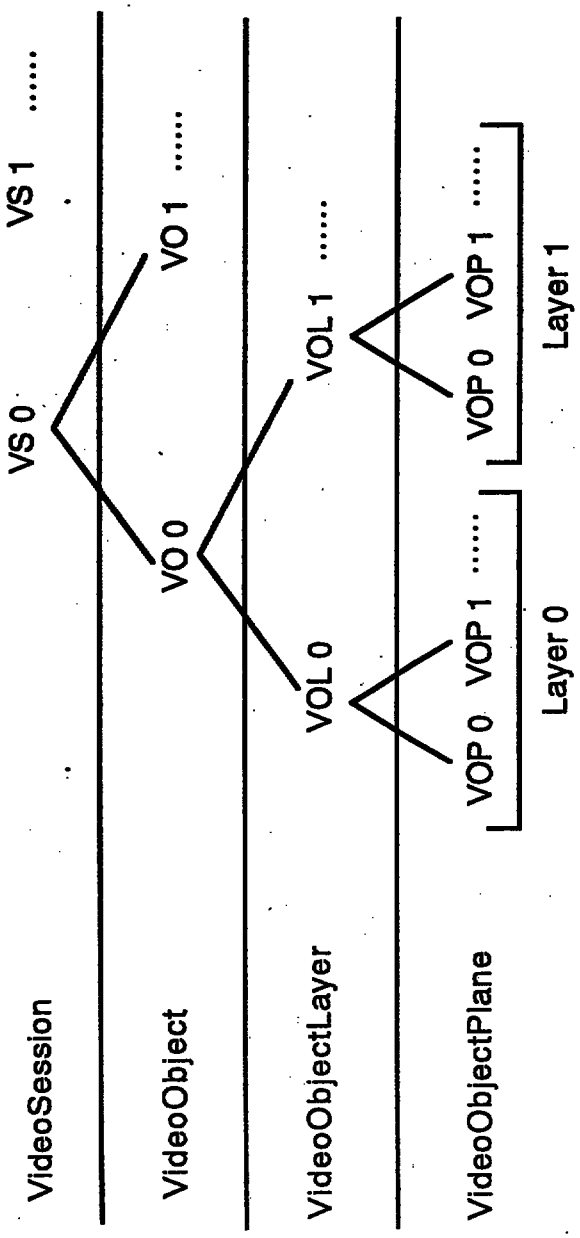

START
S112:PREDICTION USING MEMORY-a?
S113
FREQUENT UPDATE OF MEMORY-a AND MEMORY-c WITH DECODED PICTURE
S114:PREDICTION USING MEMORY-b?
S115
INFREQUENT UPDATE OF MEMORY-b ON PREDICTING PICTURE BASIS AND FREQUENT UPDATE OF MEMORY-a AND MEMORY-c WITH DECODED PICTURE
S116
FULL FRAME UPDATE OF MEMORY-b BASED ON UPDATE INDICATION OF MEMORY-b UPDATE INFO. 1001
S117:PREDICTION USING MEMORY-c?
S118
FREQUENT UPDATE OF MEMORY-a AND MEMORY-c WITH DECODED PICTURE
S119
FULL FRAME UPDATE OF MEMORY-c BASED ON UPDATE INDICATION OF MEMORY-c UPDATE INFO. 1002
END

Fig.29


VideoSession
VideoObject
VideoObjectLayer
VideoObjectPlane
VS 0
VS 1 .....
VO 0
VO 1 .....
VOL 0
VOL 1 .....
VOP 0 VOP 1 .....
VOP 0 VOP 1 .....
Layer 0
Layer 1

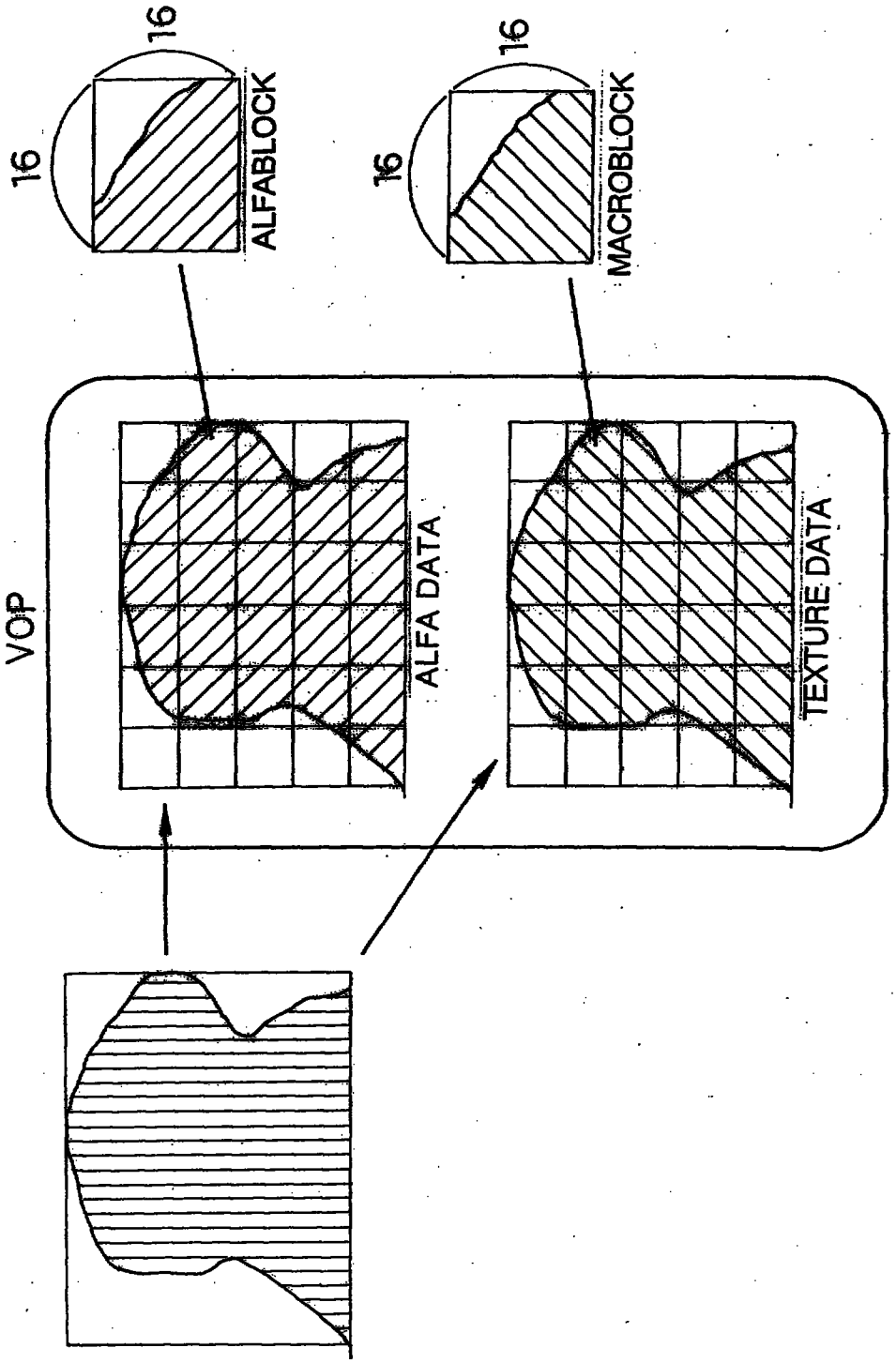



Fig.30

Fig.31

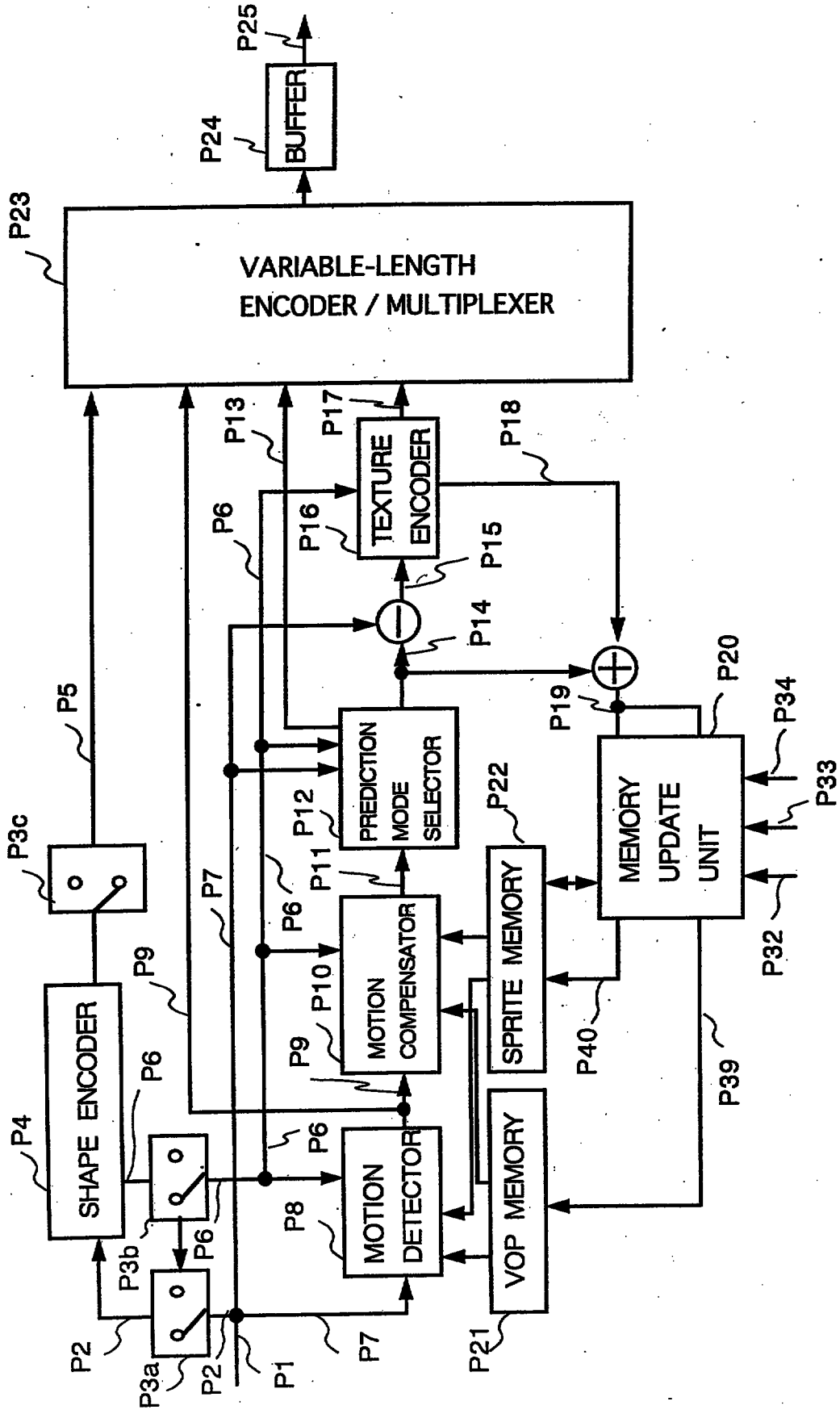

P25
BUFFER
P24
P23
VARIABLE-LENGTH
ENCODER / MULTIPLEXER
P13
P17
TEXTURE
ENCODER
P18
P6
P16
P15
P14
P20
P19
P34
P5
PREDICTION
MODE
SELECTOR
P12
P22
MEMORY
UPDATE
UNIT
P33
P3c
P11
P7
P6
P9
P10
MOTION
COMPENSATOR
SPRITE MEMORY
P32
P4
SHAPE ENCODER
P6
P9
P40
P6
P6
MOTION
DETECTOR
P8
VOP MEMORY
P39
P3b
P6
P2
P7
P21
P3a
P2
P1

Fig.32

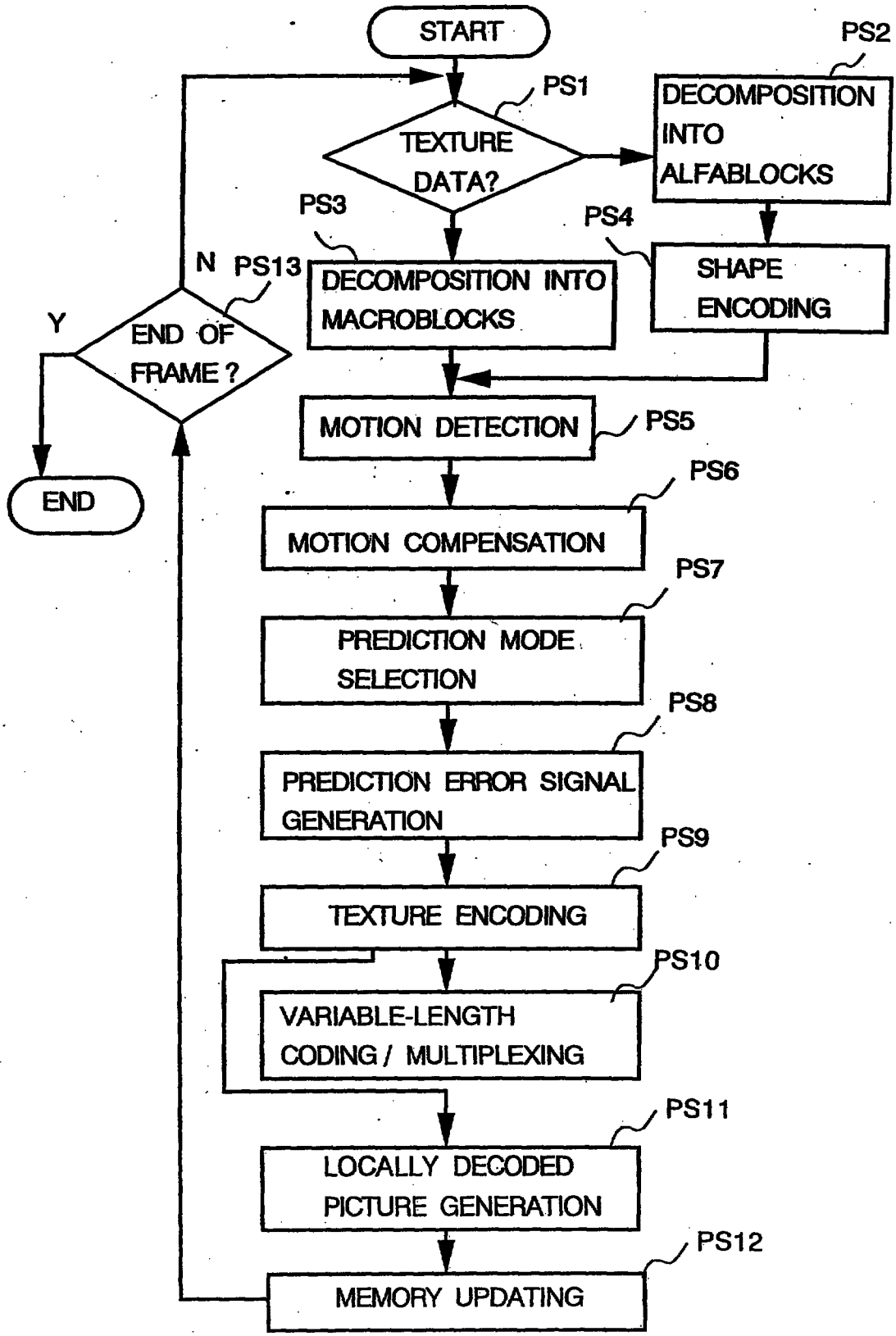

START
PS1
TEXTURE DATA?
PS2
DECOMPOSITION INTO ALFABLOCKS
PS3
DECOMPOSITION INTO MACROBLOCKS
PS4
SHAPE ENCODING
N
PS13
END OF FRAME ?
Y
END
MOTION DETECTION
PS5
MOTION COMPENSATION
PS6
PREDICTION MODE SELECTION
PS7
PREDICTION ERROR SIGNAL GENERATION
PS8
TEXTURE ENCODING
PS9
VARIABLE-LENGTH CODING / MULTIPLEXING
PS10
LOCALLY DECODED PICTURE GENERATION
PS11
MEMORY UPDATING
PS12

Fig.33

| VOP TYPE | NO PREDICTION | PAST VOP BASED PREDICTION | PAST AND FUTURE VOP BASED PREDICTION | SPRITE BASED PREDICTION |
|---|---|---|---|---|
| I - VOP | ○ | × | × | × |
| P - VOP | ○ | ○ | × | × |
| B - VOP | ○ | ○ | ○ | × |
| SPRITE - VOP | ○ | ○ | × | ○ |

Fig.34

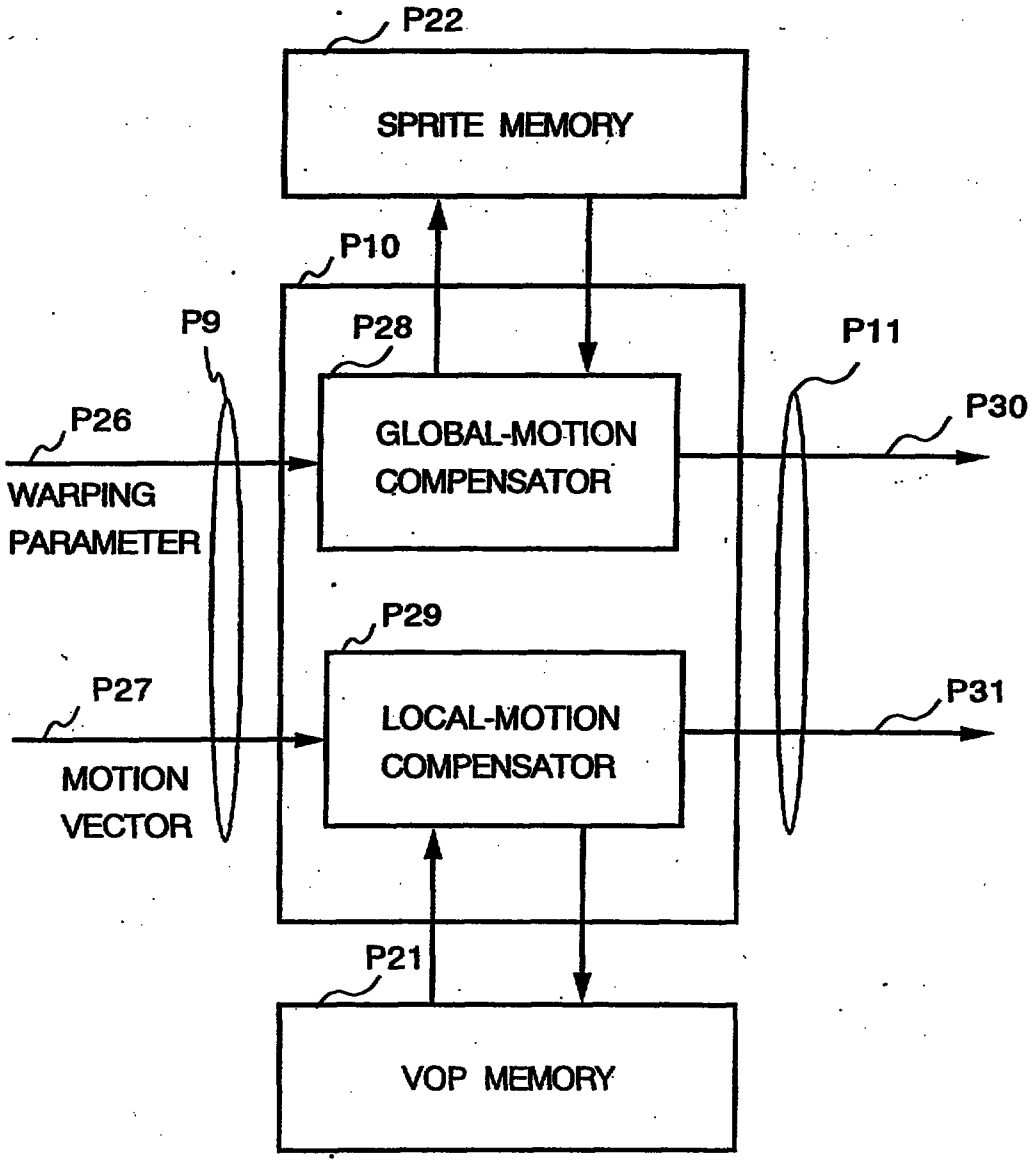

P22
SPRITE MEMORY
P10
P9
P28
P11
P26
GLOBAL-MOTION
COMPENSATOR
P30
WARPING
PARAMETER
P29
P27
LOCAL-MOTION
COMPENSATOR
P31
MOTION
VECTOR
P21
VOP MEMORY

Fig.35

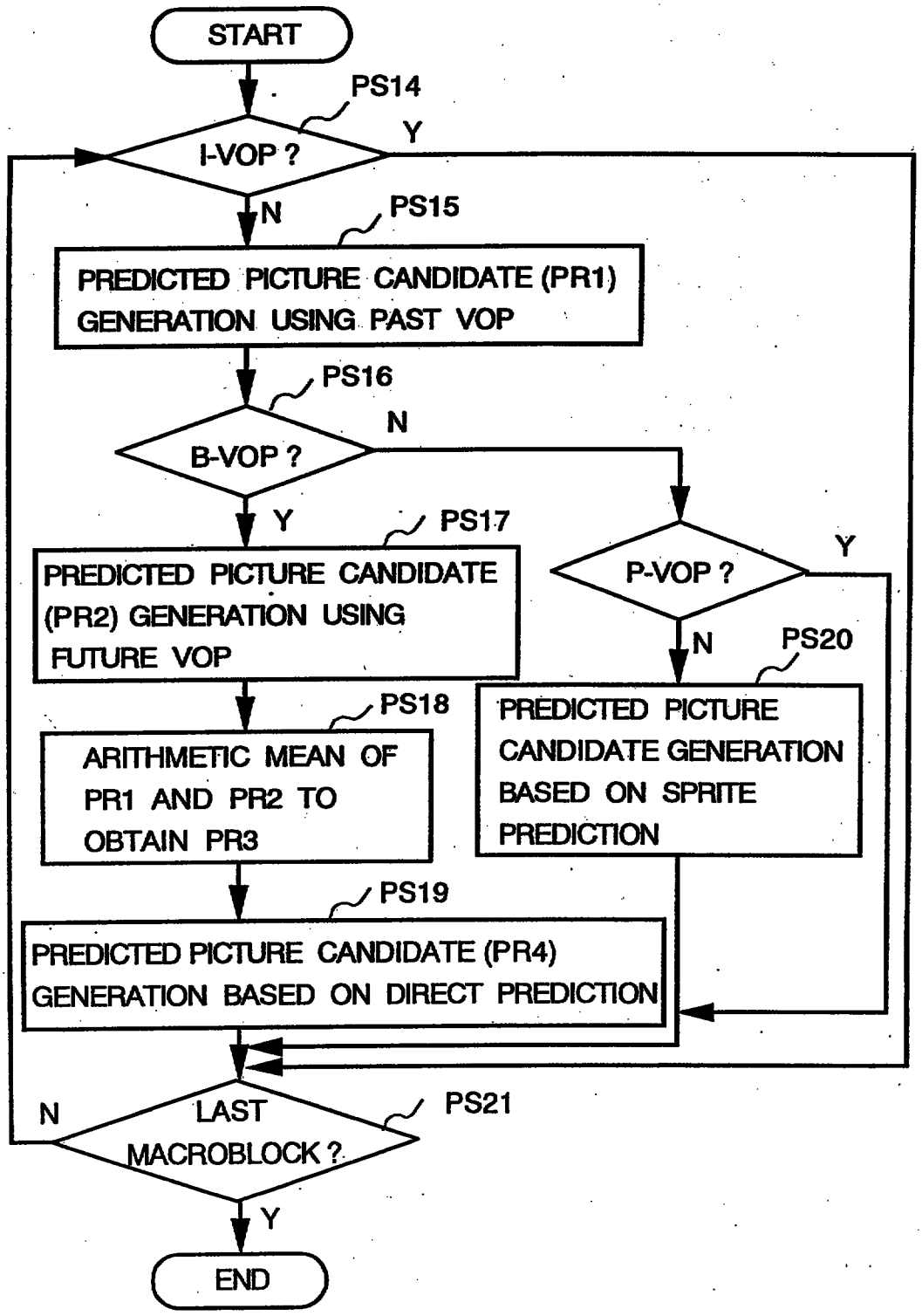

START
PS14
I-VOP ?
Y
N
PS15
PREDICTED PICTURE CANDIDATE (PR1) GENERATION USING PAST VOP
PS16
B-VOP ?
N
Y
PS17
PREDICTED PICTURE CANDIDATE (PR2) GENERATION USING FUTURE VOP
P-VOP ?
Y
N
PS20
PREDICTED PICTURE CANDIDATE GENERATION BASED ON SPRITE PREDICTION
PS18
ARITHMETIC MEAN OF PR1 AND PR2 TO OBTAIN PR3
PS19
PREDICTED PICTURE CANDIDATE (PR4) GENERATION BASED ON DIRECT PREDICTION
N
LAST MACROBLOCK ?
PS21
Y
END

Fig.36

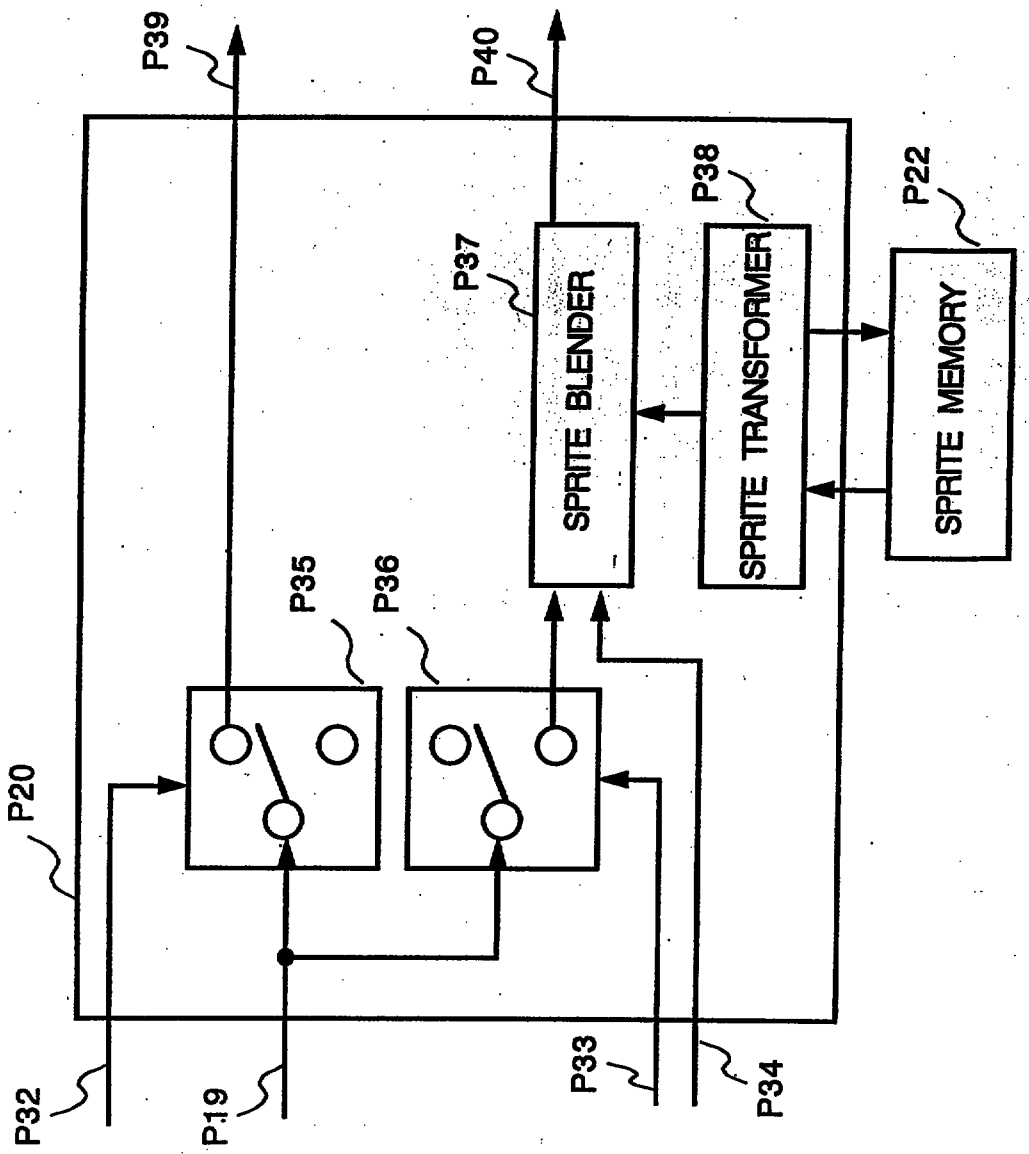
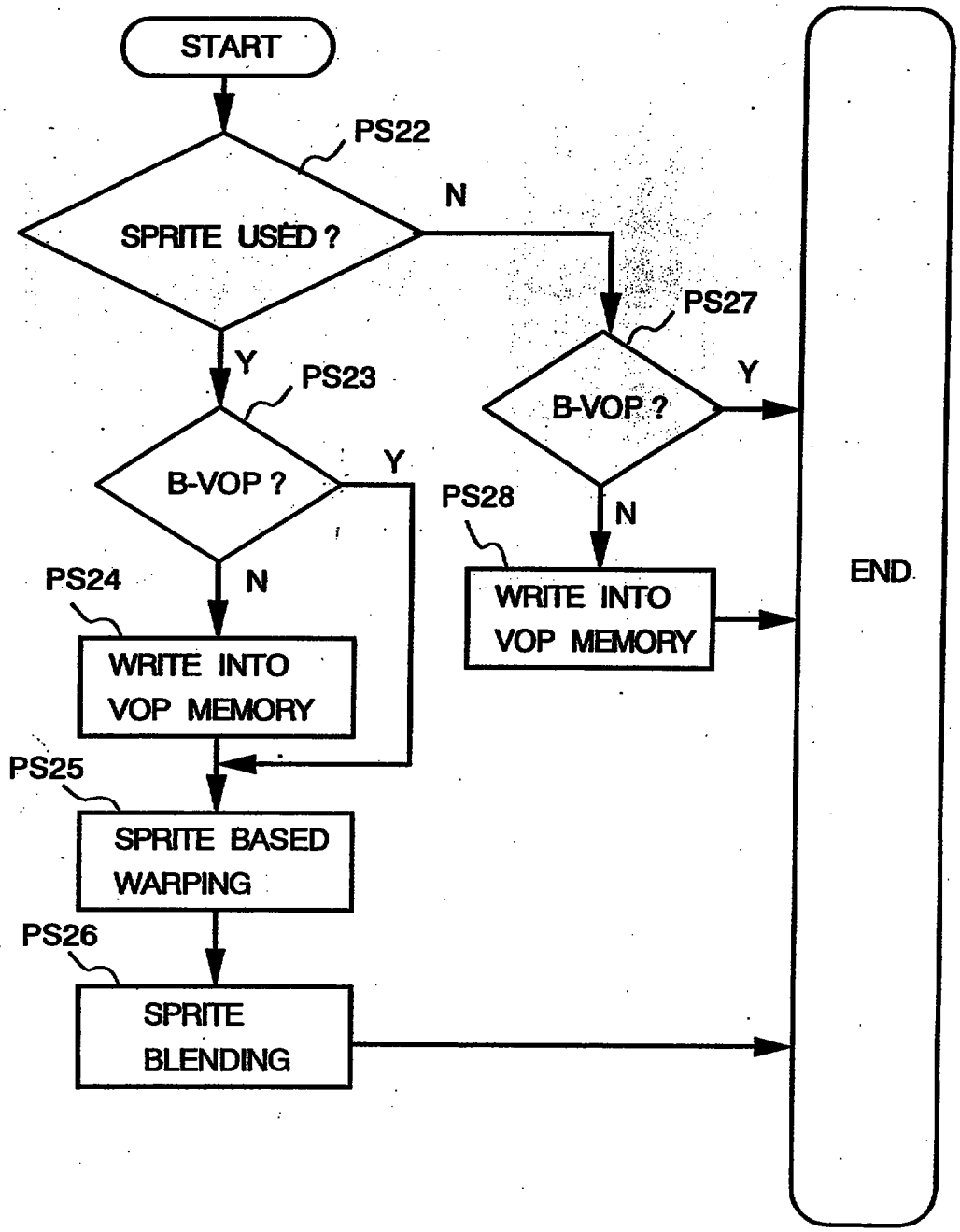

P39
P40
P38
P22
P37
SPRITE BLENDER
SPRITE TRANSFORMER
SPRITE MEMORY
P35
P36
P20
P32
P19
P33
P34

Fig.37


START
PS22
SPRITE USED ?
N
Y
PS23
B-VOP ?
Y
N
PS24
WRITE INTO VOP MEMORY
PS25
SPRITE BASED WARPING
PS26
SPRITE BLENDING
PS27
B-VOP ?
Y
N
PS28
WRITE INTO VOP MEMORY
END

includes accompanied chrominance signals being associated with 16x16 sample luminance signals. VOP data are obtained from a moving picture sequence externally processed outside of an encoder.

Fig. 31 is a diagram showing the configuration of a VOP encoder according to the VM encoding/decoding system. The diagram includes original VOP data P1 to be inputted, an alphablock P2 representing the shape information of the VOP, a switch P3a for passing the shape information, if there is any, of the inputted original VOP data, a shape encoder P4 for compressing and encoding the alphablock, compressed alphablock data P5, a locally decoded alphablock P6, texture data (a macroblock) P7, a motion detector P8, a motion parameter P9, a motion compensator P10, a predicted picture candidate P11, a prediction mode selector P12, a prediction mode P13, a predicted picture P14, a prediction error signal P15, a texture encoder P16, texture encoding information P17, a locally decoded prediction error signal P18, a locally decoded macroblock P19, a sprite memory update unit P20, a VOP memory P21, a sprite memory P22, a variable-length encoder/multiplexer P23, a buffer P24, and an encoded bitstream P25.

Fig. 32 shows a flowchart outlining an operation of the encoder.

Referring to the encoder of Fig. 31, the original

REPLACEMENT SHEET
09 OCT 1998

prediction mode.

(3) Updating of memories

The memory update unit P20 controls the VOP memory P21 and sprite memory P22 to be updated (step PS12). The contents of these memories are updated regardless of the prediction mode P13 selected on a macroblock basis.

Fig. 36 is a diagram showing the configuration of the memory update unit P20. Fig. 37 shows a flowchart including steps PS22 through PS28 illustrating the operation of the memory update unit P20.

In Fig. 36, an externally supplied VOP encoding type P32, an externally supplied sprite prediction identification flag P33 for indicating the use of the sprite memory for prediction, an externally supplied blend factor P34 used for prediction with the sprite memory, switches P35, P36, a sprite blender P37, a sprite transformer P38, a VOP memory update signal P39, and a sprite memory update signal P40 are shown.

Firstly, the use of the sprite with the current VO or VOL is examined if being designated by the sprite prediction identification flag P33 (step PS22). With no use of the sprite designated, the data are examined if being the B-VOP (step PS27). With the B-VOP, then no updating is performed with the VOP memory P21. With either the I-VOP or the P-VOP, then the VOP memory P21 is

REPLACEMENT SHEET
09 OCT 1998

motion parameter value ranges in a counterpart decoder.

When varying the effective motion parameter value ranges each frame, the effective motion parameter value range indication information is to be included in the header information added on a frame basis.

Thus, the efficient and adaptive use of the memories in response to the magnitude of a local motion of a frame can be provided by the encoder including the motion compensator 5a, thereby enhancing prediction efficiency.

This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

Further, this embodiment has thus disclosed the encoder performing on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. No global-motion detection involves no necessity for transmitting information

REPLACEMENT SHEET
09 OCT 1998

indicating a global/local prediction as the prediction mode.

Further, a counterpart decoder may be provided with the prediction picture generation section 100 including the same component as the motion compensator 5a discussed in this embodiment. In the decoder, a motion compensator only generates a single predicted picture based upon a decoded motion parameter.

Embodiment 3.

Another embodiment shows an encoder having a single replacement of the motion compensator 5 of the encoder of Fig. 1. The configuration and operation of a motion compensator 5b are now described. A motion detector 3a employed in this embodiment is assumed to output an amount of the parallel translation, the affine parameter, and the perspective parameter as motion parameters 4a.

Further, the memory-a of this embodiment is assumed to be a frame memory for storing a reference picture frame, the memory-b is assumed to be a static sprite memory, and the memory-c is assumed to be a dynamic sprite memory.

Fig. 9 shows the configuration of the motion compensator 5b of this embodiment in detail. In the figure, a parallel translation amount 36 (i.e., a motion

REPLACEMENT SHEET
09 OCT 1998

This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

Further, this embodiment has thus disclosed the encoder performing on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. No global-motion parameter detection involves no necessity of transmitting information of a global/local prediction as the prediction mode.

Further, this embodiment has thus disclosed the use of the memories a, b, and c. Alternatively, the use of memories a and b alone, memories a and c alone, or memories b and c alone, is also applicable.

Further, a decoder may be provided with the prediction picture generating section 100 including the same component as the motion compensator 5b discussed in this embodiment. A motion compensator in a decoder only

REPLACEMENT SHEET
09 OCT 1998

embodiment.

(1) the memory-a used for prediction (with the prediction memory identification information 800 indicating the use of the memory-a)

The coordinates of each pixel of a prediction segment are translated based upon the motion vector and picture data at a corresponding location in the memory-a is read out as the predicted picture.

(2) the memory-b used for prediction (with the prediction memory identification information 800 indicating the use of the memory-b)

An affine transform expression is found based upon the motion vector, the coordinates of each pixel of a prediction segment are displaced based upon the transform expression, and picture data at a corresponding location in the memory-b is read out as a predicted picture.

(3) the memory-c used for prediction (with the prediction memory identification information 800 indicating the use of the memory-c)

A perspective transform expression is found based upon the motion vector, the coordinates of each pixel of a prediction segment are displaced based upon the transform expression, and picture data at a corresponding location in the memory-c is read out as a predicted picture.

REPLACEMENT SHEET
09 OCT 1998

27 illustrates an interpolated pixel value generation. In Fig. 27, $(i_0, j_p)$ denotes an integer pixel location, $(j^p, j^p)$ denotes a real number precision pixel location, and w denotes a weight.

4) Updating a memory (a reference picture)

Fig. 28 shows a flowchart illustrating a control operation of the memory update unit 115. The memory update unit 115 controls an updating of the respective memories on a readout unit basis (e.g., a macroblock) of a predicted picture based upon the prediction mode 8 (or the prediction memory indication information 800). With the memory-a used for prediction (step S112), the contents of the memory-a and the memory-c are updated regularly with the decoded picture 101 (step S113). With the memory-b used for prediction (step S114), the reference picture of the memory-b is not updated on a readout unit basis of a predicted picture because of the memory-b being a static sprite memory, whereas the contents of the memory-a and the memory-c are updated regularly with the decoded picture 101 (step S115). When receiving the update indication by the memory-b update information 1001, then the memory update unit updates a full content of the memory-b with received picture data included in the memory-b update information 1001 (step S116). With the memory-c used for prediction (step

REPLACEMENT SHEET
09 OCT 1998

* C L M S *



or 5, wherein the moving picture is predicted in a moving picture sequence having first and second video objects,

wherein the plurality of memories includes separate first and second pluralities of memories corresponding to the first and second video objects, respectively, and the prediction picture generation section includes separate first and second generators, respectively, corresponding to the first and second video objects,

wherein the first generator uses the picture data stored in at least one of the first and second pluralities of memories to generate the predicted picture when predicting the first video object, and generates information indicating a use of the second plurality of memories for predicting the first video object, the information being added to the predicted picture.

7. The moving picture prediction system of claim 1, 4 or 5, wherein the prediction picture generation section generates the predicted picture through a change of either one of a number and a size of the plurality of memories in response to a change in the moving picture at each time instance.

8. The moving picture prediction system of claim 1, 4 or 5, wherein the prediction picture generation section

REPLACEMENT SHEET
09 OCT 1998

data for predicting a video object, thereby posing a problem of failing to enhance prediction efficiency.

The present invention is directed to solving the aforementioned problems. An objective of this invention is to provide the prediction system for encoding/decoding of picture data where two or more memories are provided to store the past record of the moving picture sequence effectively in consideration of the internal structure and characteristic of the moving picture sequence, thereby achieving a highly efficient prediction as well as encoding/decoding. In addition, the prediction system provides a sophisticated inter-video object prediction performing among two or more video objects.

REPLACEMENT SHEET
14 AUG 1999

98053 16-8
14/8/99

Disclosure of the invention

According to a first aspect of the present invention there is provided a moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoder and a decoder, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction, a plurality of transform methods each being assigned to a corresponding memory; and

a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories.

REPLACEMENT SHEET
14 AUG 1999

REPLACEMENT SHEET
14 AUG 1999

The encoder generates a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and transmits the prediction memory indication information signal and the parameter to a decoding station so as to generate the predicted picture using the picture data stored in the one of the plurality of memories based upon the one of the transform methods corresponding to the one of the plurality of memories in the decoding station.

The decoder receives the parameter and a prediction memory indication information sinal indicating the one of the plurality of memories used for generating the predicted picture from an encoding station, wherein the prediction picture generation section generates the predicted picture using the picture data stored in the one of the plurality of memories based upon the parameter and the one of the transform methods corresponding to the one of the plurality of memories.

Further, according to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoding

and a decoding, includes a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being assigned to different parameter effective value ranges, respectively, and a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, for selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and for generating a predicted picture using the picture data stored in a selected memory.

REPLACEMENT SHEET
14 AUG 1999

Still further, according to the present invention, there is provided a moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction, a plurality of transform methods each being assigned to a corresponding memory; and

a prediction picture generation section including a motion compensator for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segments to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories, and

a memory update unit for updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

The moving picture prediction system predicts the

REPLACEMENT SHEET
14 AUG 1999

REPLACEMENT SHEET
14 AUG 1999

the one of the plurality of memories is at least one of a parallel translation, an affine transformation, and a perspective transformation in an interchangeable manner.

Still further, according to the present invention there is provided, a method for predicting a moving picture to be implemented in at least one of an encoding or a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;

utilising a plurality of transform methods each being assigned to a corresponding memory;

receiving a parameter representing a motion of a picture segment to be predicted; and

generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods being corresponding to the one of the plurality of memories.

The method for predicting a moving picture further includes the steps of generating a prediction memory indication information signal indicating the one of the plurality of memories used for the picture segment to be predicted, and transmitting the prediction memory indication information signal and the parameter to a decoding station.

The method for predicting a moving picture is implemented in the decoding, and further includes the

REPLACEMENT SHEET
14 AUG 1999



step of receiving a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and the parameter representing a motion of the picture segment to be predicted from an encoding station.

Still further, according to the present invention, a method, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, assigning separate parameter effective value ranges to the plurality of memories, respectively, receiving a parameter representing a motion of a picture segment to be predicted, selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and generating a predicted picture using the picture data stored in a selected memory.

Still further, according to the present invention there is provided a method for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories, wherein one of a plurality of transform methods is assigned to a corresponding one of the plurality of memories;

receiving a parameter representing a motion of a picture segment to be predicted;

generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories; and

updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

REPLACEMENT SHEET
14 AUG 1999

REPLACEMENT SHEET
14 AUG 1999

Brief description of drawings

Fig. 1 is a structural diagram of a moving picture encoder according to an embodiment of this invention.

Fig. 2 is a flowchart illustrating an operation of the moving picture encoder according to the embodiment of this invention.

Fig. 3 is a structural diagram illustrating the configuration of a motion compensator of the moving picture encoder of the embodiment of this invention.

Fig. 4 is a flowchart illustrating an operation of the motion compensator.

Fig. 5 is a structural diagram illustrating the structure of a memory update unit of the moving picture encoder of the embodiment of this invention.

Fig. 6 is a flowchart illustrating an operation of the memory update unit.

Fig. 7 is a structural diagram illustrating the configuration of a motion compensator of a moving picture encoder according to another embodiment of this invention.

CLAIMS

1. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoder and a decoder, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction, a plurality of transform methods each being assigned to a corresponding memory: and

a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories.

2. The moving picture prediction system of claim 1, wherein the encoder generates a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and transmits the prediction memory indication information signal and the parameter to a decoding station so as to generate the predicted picture using the

REPLACEMENT SHEET
14 AUG 1999

* C L M S *



parameter representing a motion of a picture segment to be predicted, for selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and for generating a predicted picture using the picture data stored in a selected memory.

5. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the moving picture prediction system comprising:

a plurality of memories for storing picture data for reference to be used for prediction, a plurality of transform methods each being assigned to one corresponding memory; and

a prediction picture generation section including a motion compensator for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories, and

a memory update unit for updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

6. The moving picture prediction system of claim 1, 4,

REPLACEMENT SHEET
14 AUG 1999

* C L M S *

REPLACEMENT SHEET
14 AUG 1999



16. The moving picture prediction system of claim 14, wherein the one of the transform methods corresponding to the one of the plurality of memories is at least one of a parallel translation, an affine transformation, and a perspective transformation in an interchangeable manner.

17. A method for predicting a moving picture to be implemented in at least one of an encoding or a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;

utilising a plurality of transform methods each being assigned to a corresponding memory; and

receiving a parameter representing a motion of a picture segment to be predicted; and

generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods being corresponding to the one of the plurality of memories.

18. The method for predicting a moving picture of claim 17, further comprising the steps of:

generating a prediction memory indication information

* C L M S *



selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter; and

generating a predicted picture using the picture data stored in a selected memory.

21. A method for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories, wherein one of a plurality of transform methods is assigned to a corresponding one of the plurality of memories;

receiving a parameter representing a motion of a picture segment to be predicted;

generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and the one of a plurality of transform methods assigned to the one of the plurality of memories; and

updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

REPLACEMENT SHEET
14 AUG 1999

世界知的所有権機関
国際事務局

IPOS photocopied!

PCT

特許協力条約に基づいて公開された国際出願

9805316-8

| | | | |
|---|---|---|---|
| (51) 国際特許分類6<br>H04N 7/32 | A1 | (11) 国際公開番号 | WO98/36576 |
| | | (43) 国際公開日 | 1998年8月20日(20.08.98) |

(21) 国際出願番号 PCT/JP98/00232

(22) 国際出願日 1998年1月22日(22.01.98)

(30) 優先権データ
特願平9/28862 1997年2月13日(13.02.97) JP

(71) 出願人（米国を除くすべての指定国について）
三菱電機株式会社
(MITSUBISHI DENKI KABUSHIKI KAISHA)[JP/JP]
〒100 東京都千代田区丸の内二丁目2番3号 Tokyo, (JP)
(72) 発明者；および
(75) 発明者／出願人（米国についてのみ）
関口俊一(SEKIGUCHI, Shunichi)[JP/JP]
浅井光太郎(ASAI, Kohtaro)[JP/JP]
村上篤道(MURAKAMI, Tokumichi)[JP/JP]
西川博文(NISHIKAWA, Hirofumi)[JP/JP]
黒田慎一(KURODA, Shinichi)[JP/JP]
井須芳美(ISU, Yoshimi)[JP/JP]
長谷川由里(HASEGAWA, Yuri)[JP/JP]
〒100 東京都千代田区丸の内二丁目2番3号
三菱電機株式会社内 Tokyo, (JP)

(74) 代理人
弁理士 溝井章司(MIZOI, Shoji)
〒247 神奈川県鎌倉市大船二丁目17番10号 NTA大船ビル3F
Kanagawa, (JP)

(81) 指定国 CN, JP, KR, SG, US, 欧州特許 (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

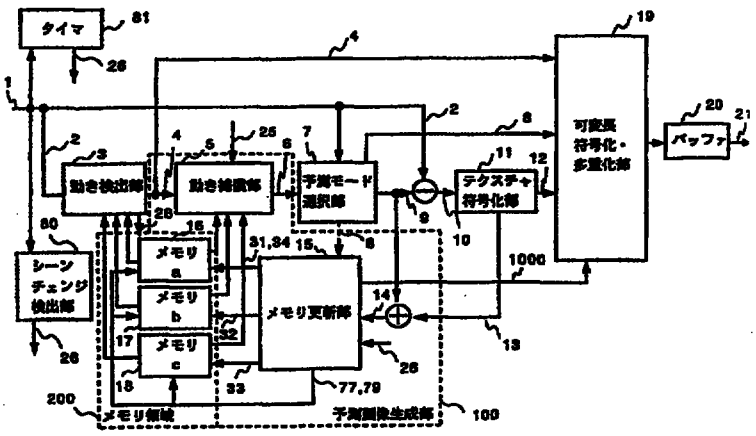
添付公開書類
国際調査報告書

(54)Title: MOVING IMAGE ESTIMATING SYSTEM

(54)発明の名称 動画像予測方式

(57) Abstract

To obtain an encoding system capable of estimating a moving image with a high accuracy in accordance with the details and degree of importance of a scene and nature of motion of the moving image, said system is provided with a plurality of memories a, b, c in which image data to be referred to are stored, a motion compensating element (5) capable of being operated correspondingly to an arbitrary deformation parameter representing a motion of an image region to be estimated, and adapted to generate an estimated image by using arbitrary data in these memories a, b, c on the basis of the deformation parameter, and a memory updating element (15) capable of updating the content of not less than one memory out of the memories a, b, c at arbitrary time intervals.




3...MOTION DETECTING ELEMENT
5...MOTION COMPENSATING ELEMENT
7...ESTIMATION MODE SELECTING ELEMENT
11...TEXTURE ENCODING ELEMENT
15...MEMORY UPDATING ELEMENT
16...MEMORY a
17...MEMORY b
18...MEMORY c
19...VARIABLE LENGTH ENCODING AND MULTIPLEXING ELEMENT
20...BUFFER
80...SCENE CHANGE DETECTING ELEMENT
81...TIMER
100...ESTIMATED IMAGE FORMATION UNIT
200...MEMORY REGION

（５７）要約

動画像のシーン内容や重要度、動きの性質などに応じて、高い精度の予測が行える符号化方式を得るために、参照する画像データが蓄積される複数のメモリａ、ｂ、ｃと、被予測画像領域の動きを表現する任意の変形パラメータに対応可能で該変形パラメータに基づいて前記複数のメモリａ、ｂ、ｃの任意のデータを用いて予測画像を生成する動き補償部５と、前記複数のメモリａ、ｂ、ｃのうち１つ以上のメモリの内容を任意の時間間隔で更新可能とするメモリ更新部１５とを備える。

ＰＣＴに基づいて公開される国際出願のパンフレット第一頁に掲載されたＰＣＴ加盟国を同定するために使用されるコード（参考情報）

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | アルバニア | FI | フィンランド | LT | リトアニア | SN | セネガル |
| AM | アルメニア | FR | フランス | LU | ルクセンブルグ | SZ | スワジランド |
| AT | オーストリア | GA | ガボン | LV | ラトヴィア | TD | チャード |
| AU | オーストラリア | GB | 英国 | MC | モナコ | TG | トーゴー |
| AZ | アゼルバイジャン | GE | グルジア | MD | モルドヴァ | TJ | タジキスタン |
| BA | ボスニア・ヘルツェゴビナ | GH | ガーナ | MG | マダガスカル | TM | トルクメニスタン |
| BB | バルバドス | GM | ガンビア | MK | マケドニア旧ユーゴスラヴィア共和国 | TR | トルコ |
| BE | ベルギー | GN | ギニア | ML | マリ | TT | トリニダッド・トバゴ |
| BF | ブルキナ・ファソ | GR | ギリシャ | MN | モンゴル | UA | ウクライナ |
| BG | ブルガリア | HU | ハンガリー | MR | モーリタニア | UG | ウガンダ |
| BJ | ベナン | IE | アイルランド | MW | マラウイ | US | 米国 |
| BR | ブラジル | IL | イスラエル | MX | メキシコ | UZ | ウズベキスタン |
| BY | ベラルーシ | IS | アイスランド | NE | ニジェール | VN | ヴィエトナム |
| CA | カナダ | IT | イタリア | NL | オランダ | YU | ユーゴースラヴィア |
| CF | 中央アフリカ | JP | 日本 | NO | ノルウェー | ZW | ジンバブエ |
| CG | コンゴー共和国 | KE | ケニア | NZ | ニュージーランド | | |
| CH | スイス | KG | キルギスタン | PL | ポーランド | | |
| CI | コートジボアール | KP | 北朝鮮 | PT | ポルトガル | | |
| CM | カメルーン | KR | 韓国 | RO | ルーマニア | | |
| CN | 中国 | KZ | カザフスタン | RU | ロシア連邦 | | |
| CU | キューバ | LC | セントルシア | SD | スーダン | | |
| CZ | チェコ共和国 | LI | リヒテンシュタイン | SE | スウェーデン | | |
| DE | ドイツ | LK | スリランカ | SG | シンガポール | | |
| DK | デンマーク | LR | リベリア | SI | スロヴェニア | | |
| EE | エストニア | LS | レソト | SK | スロヴァキア | | |
| ES | スペイン | | | SL | シエラ・レオーネ | | |