

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3679426号

(P3679426)

(45) 発行日 平成17年8月3日(2005.8.3)

(24) 登録日 平成17年5月20日(2005.5.20)

(51) Int.Cl.⁷

F I

H04N 7/32

H04N 7/137

Z

G06T 9/00

G06T 9/00

H03M 7/30

H03M 7/30

Z

請求項の数 7 (全 22 頁)

(21) 出願番号 特願平6-44055
 (22) 出願日 平成6年3月15日(1994.3.15)
 (65) 公開番号 特開平7-95592
 (43) 公開日 平成7年4月7日(1995.4.7)
 審査請求日 平成13年3月14日(2001.3.14)
 (31) 優先権主張番号 031642
 (32) 優先日 平成5年3月15日(1993.3.15)
 (33) 優先権主張国 米国(US)

(73) 特許権者 591013573
 マサチューセッツ・インスティテュート・
 オブ・テクノロジー
 MASSACHUSETTS INSTI
 TUTE OF TECHNOLOGY
 アメリカ合衆国マサチューセッツ州021
 39, ケンブリッジ, マサチューセッツ
 ・アベニュー 77

(74) 代理人 100089705
 弁理士 社本 一夫
 (74) 代理人 100071124
 弁理士 今井 庄亮
 (74) 代理人 100076691
 弁理士 増井 忠武

最終頁に続く

(54) 【発明の名称】 画像データを符号化して夫々がコヒーレントな動きの領域を表わす複数の層とそれら層に付随する動きパラメータとにするシステム

(57) 【特許請求の範囲】

【請求項1】

一連の画像フレームから成るフレーム・シーケンスに関するデジタル画像データを符号化するためのシステムにおいて、

A. 連続したフレームにおける夫々のピクセル近傍における局所的動きを推定して該推定した局所的動きを表す局所的動きベクトルを生成し、そしてそれに関連した稠密な動きモデルを生成する局所的動き推定部であって、生成した前記局所的動きベクトルは、各前記ピクセル近傍に含まれる1つのピクセルに割り当てられ、前記稠密動きモデルは、フレームを構成するピクセルに割り当てられた前記局所的動きベクトルのアレイから成る、前記の局所的動き推定部と、

B. 前記稠密動きモデルから

i. 前記フレーム・シーケンスの全体の中のコヒーレントな動きの領域と、

ii. フレームからフレームへかけての前記コヒーレント動き領域の各々の動きを表わす夫々の領域動きモデルと、

を決定する動きセグメンテーション処理部と、

C. 前記コヒーレント動き領域の各々について層を形成する層抽出処理部であって、前記層が、

i. 当該コヒーレント動き領域に含まれる前記ピクセルの各々について、前記フレーム・シーケンス内のフレームから得たピクセル輝度情報を表すデータと、

ii. 当該層の他の層に対する順序付けを、前記画像における深さにより表すデータで

あって、該層抽出処理部が、当該層を形成する前記コヒーレント動き領域内のピクセル位置におけるピクセルが、前記シーケンス中における前記フレームの任意のものにおいて覆い隠されるかどうかを判定し、前記シーケンス中において前記ピクセル位置にて覆い隠されないピクセルの個数を算出し、該算出したピクセル個数を当該層の前記深さとして使用することにより、当該層の前記深さを判定するようになった、前記のデータと、

iii. 前記動きセグメンテーション処理部による当該コヒーレント動き領域に関連した前記領域動きモデルに係するパラメータを含む動き情報を表すデータと、
を含む、前記の層抽出処理部と、
を備えたことを特徴とするシステム。

【請求項 2】

前記動きセグメンテーション処理部が、

a. 前記コヒーレント動き領域の各々について前記領域動きモデルを生成する動き推定部と、

b. 互いに類似した領域動きモデルどうしをグループ化してクラスタを形成し、そして各クラスタについてクラスタ動きモデルを生成するクラスタ化処理部と、

c. 個々のピクセルを、各該ピクセルに関連した前記局所的動きベクトルに最も良く対応する動きを定める領域動きモデルまたはクラスタ動きモデルと関連付ける、ピクセル割当て処理部と、

を含んでいることを特徴とする請求項 1 記載のシステム。

【請求項 3】

前記局所的動き推定部が、

a. 前記画像データにフィルタ処理を施して画像にブレないしボケを加えるフィルタであって、フィルタ処理後データを発生するようにしたフィルタと、

b. 前記フィルタ処理後データのサブサンプリングを行なって、低解像度版の画像を生成する手段と、

c. 前記画像の前記低解像度版の選択したピクセル近傍どうしの間の小さな動きを推定する手段と、

d. 前記画像の前記低解像度版における推定した前記小さな動きを用いてその画像の関連したピクセル近傍における動きを推定する手段と、

を含んでいることを特徴とする請求項 1 記載のシステム。

【請求項 4】

複数のフレームから成る画像シーケンスの中の、複数のコヒーレント動き領域と、それら領域に関連する動きとを判定する方法において、

A. 連続したフレームにおける夫々のピクセル近傍における局所的動きを推定して該推定した局所的動きを表す局所的動きベクトルを生成し、そしてそれに関連した稠密動きモデルを生成するステップであって、生成した前記局所的動きベクトルは、各前記ピクセル近傍に含まれる 1 つのピクセルに割り当てられ、前記稠密動きモデルは、フレームを構成するピクセルに割り当てられた前記局所的動きベクトルのアレイから成る、推定ステップと、

B. 前記稠密動きモデルを用いて、一連の画像フレームから成る画像シーケンスの全体の中の互いに類似の局所的動きベクトルを有する夫々のコヒーレント動き領域と、フレームからフレームへかけての前記コヒーレント動き領域の各々の動きを表わす夫々の領域動きモデルとを判定するステップと、

C. 前記コヒーレント動き領域の各々について層を形成するステップであって、前記層が、

i. 当該コヒーレント動き領域に含まれる前記ピクセルの各々について、前記フレーム・シーケンス内のフレームから得たピクセル輝度情報を表すデータと、

ii. 当該層の他の層に対する順序付けを、前記画像における深さにより表すデータであって、当該層の前記深さの判定は、当該層を形成する前記コヒーレント動き領域内のピクセル位置におけるピクセルが、前記シーケンス中における前記フレームの任意のものに

10

20

30

40

50

において覆い隠されるかどうかを判定し、前記シーケンス中において前記ピクセル位置にて覆い隠されないピクセルの個数を算出し、該算出したピクセル個数を当該層の前記深さとして使用することにより行う、前記のデータと、

iii. 当該コヒーレント動き領域に関連する前記領域動きモデルに係したパラメータを含む動き情報を表すデータと、

を含む、前記のステップと、

D. 前記画像シーケンスを、各層が関係した動き情報を含んでいる一連の順序付けした複数の層として格納するステップと、
を含んでいることを特徴とする方法。

【請求項 5】

10

前記ステップ A が、

a. 前記画像データにフィルタ処理を施して画像にブレないしボケを加えるステップであって、フィルタ処理後データを発生するようにしたステップと、

b. 前記フィルタ処理後データのサブサンプリングを行なって、前記画像の低解像度版を生成するステップと、

c. 前記画像の前記低解像度版の選択したピクセル近傍どうしの間の小さな動きを推定するステップと、

d. 前記画像の前記低解像度版における推定した前記小さな動きを用いてその画像の関連するピクセル近傍における動きを推定するステップと、
を含んでいることを特徴とする請求項 4 記載の方法。

20

【請求項 6】

前記ステップ B が、

a. 前記コヒーレント動き領域の各々について前記領域動きモデルを推定するステップと、

b. 互いに類似した領域動きモデルどうしをグループ化してクラスタを形成し、それらクラスタの各々についてクラスタ動きモデルを生成するステップと、

c. 個々のピクセルを、このピクセルに関連した前記局所的動きベクトルを最も良く近似する動きを定める、領域動きモデルまたは前記ステップ b で生成したクラスタ動きモデルに関連付けるステップと、
を含んでいることを特徴とする請求項 4 記載の方法。

30

【請求項 7】

前記ステップ B が、更に、

d. 前記ステップ c において個々のピクセルに関連付けられた前記領域動きモデルまたはクラスタ動きモデルを用いて、推定された前記領域動きモデルを反復して更新するステップと、

e. 前記ステップ d において更新された前記領域動きモデルを用いて、前記クラスタ動きモデルを反復して更新するステップと、

f. 前記ステップ c ~ e を、その反復実行の回数が最大回数に達するか、或いは、連続する反復実行の間で得られる前記クラスタ動きモデルに大きな変化が生じなくなるまで、反復実行するステップと、
を含んでいることを特徴とする請求項 6 記載の方法。

40

【発明の詳細な説明】

【0001】

【産業上の利用分野】

本発明は、広くは画像の符号化に関するものであり、より詳しくは、画像データに対して、その格納、伝送、ないし復号化が容易なように、圧縮処理を施す機構に関するものである。

【0002】

【従来の技術及び発明が解決しようとする課題】

本願の関連出願として、その発明の名称を「画像の符号化のための層構造表現 (Layered

50

d Representation For Image Coding)」とした、Edward H. Adelson による米国特許出願第 07 / 8 8 8 8 0 1 号があり、同米国出願は、この言及をもって本開示に組み込むものとする。同米国出願には、1つの画像を、その画像中の「深さ」によって順序付けた一連の複数の層で表わすようにした、画像圧縮システムが記載されている。このシステムは、それら複数の層の各々に、画像シーケンスの継続時間に互って、その層にどのような操作を加えて行けば良いか、あるいは、その層をどのように変形させて行けば良いかということに関する情報を包含させるようにしており、それによって、画像シーケンスを表わせるようにしている。

【0003】

1つの層は、一連の複数のデータ・マップによって構成されており、それらマップの各々は、画像中の、他のどの物体の動きとも明確に異なった動きをしている1つの物体ないし物体の1つの部分に関連付けられている。各々のマップは、離散した2次元位置に関するデータの集合として構成されており、それらマップのうちには、第3次元として時間の次元を含んでいるものもある。それらマップは、各々の位置ごとに、(i)例えばシーケンスの開始点等の一定の瞬間における輝度値と、(ii)時間の経過に伴う減衰量と、(iii)時間の経過に伴う速度変化とを示している。更に任意に含めることのできるマップとして、コントラスト変化マップ、ブレノボケ・マップ、等々のマップがあり、コントラスト変化マップは、それに対応する層の輝度値マップにどのような数に乗じて時間の経過に伴うコントラストの変遷を作り出すべきかを記述したマップであり、また、ブレノボケ・マップは、1つないし複数の位置に、動きブレや焦点ボケを追加するものである。これら任意マップを、必要に応じて含めるようにして、対応した物体の時間の経過に伴う変化を記述するようにしても良い。それら複数の層から画像シーケンスを再生あるいは復号化するには、それら複数の層を然るべき順序で組合せた上で、それら層に時間の経過に従って然るべき操作を加えて行き、即ち「ワープ(warp)」させて行くようにする。

【0004】

物体を規定するための方法、即ち物体の境界を判別するための方法には、様々なものがある。例えば、動き解析を反復実行するという方法があり、その一例は主要動き解析という方法であって、この方法は、画像シーケンスの中にただ1つの主要動きが存在しているということを繰り返して仮定し、その仮定した動きに基づいて1つの物体(主要物体)を規定する。続いて、この主要物体をマスクした上で、あるいは画像から除外した上で、その画像の残りの部分について第2の主要動きが存在するものと仮定して再度解析を行なう。第2の物体が識別されたならば、その物体をも画像から消去した上で、その画像に再びワープ処理を施して更なる動き解析を行なうようにすれば良い。この種の主要動き解析は「反復的動き解析を利用した画像シーケンス・エンハンスメント(Image Sequence Enhancement Using Multiple Motions Analysis)」という題名の Iraniと Pelegによる論文に記載されている。この方法は、画像中にただ1つの、そして明確に主要動きであるといえる動きが存在している場合には、その主要物体を正確に識別することができる。しかしながら、画像中に複数の主要動きが存在している場合には、それら複数の物体を正確に識別することができない。

【0005】

画像中の物体を判別するための更に別の方法に、ブロック・マッチング法がある。この方法では、1つのフレームからその次のフレームへかけての、複数の矩形のピクセル・ブロックの動きを判定し、それらブロックに対して夫々に動きベクトルを対応付ける。そして、互いに類似した動きをしているブロックどうしは同一の物体の夫々の部分であると見なす。不整な形状の物体をも含めて、全ての物体がそれらブロックの組合せで表わされる。「低ビット・レート映像信号の圧縮のための、セグメンテーションを利用した、動きの差分及び動きフィールド画像の符号化(Segmentation-Based Coding of Motion Difference and Motion Field Images For Low Bit-Rate Video Compression)」という題名の、Liuと Hayesによる論文に記載されているように、1つの物体を表わすのに様々な大きさのブロックを使用することができるが、不整な形状の物体は、それらブロックによって

完全に正確に表わすことはできない。

【 0 0 0 6 】

【課題を解決するための手段】

本発明は、複数の物体を包含している画像について、(i) 画像シーケンス又はフレーム・シーケンスの全体の中における複数のコヒーレント動き領域の夫々の境界と、(ii) 1つのフレームからその次のフレームへかけてのそれら領域の変形を記述した、それら領域の各々に対応した動き式の動きパラメータ又は係数を同時に判別するためのシステム、並びにそのシステムの動作方法である。その動き式が例えばアフィン変換式である場合には、1つのフレームからその次のフレームへかけてのその画像の領域の全体の動きは、6個のアフィン・パラメータから成るアフィン・パラメータ集合で記述される。このシステムは、この情報を用いて画像シーケンスを符号化することにより、その画像シーケンスを各々が1つずつの領域に対応した一連の複数の層にすると共に、速度マップを圧縮して動きパラメータの集合にする。

10

【 0 0 0 7 】

本発明のシステムは、画像中に主要動きが存在していることを必要とせず、画像中の動きが、選択した動き式の集合で記述し得るような動きでありさえすれば良い。例えばアフィン変換式は、滑らかな動きを記述するものであるため、非常に複雑な動きはエラー状態として取り扱われることになる。このシステムは更に、物体が、例えば複数の矩形ブロックで表現し得るような、整った形状であることを必要としてない。このシステムは、個々のピクセルをコヒーレント動き領域に割当てるようにしており、そのため、基本的に、いかなる形状の物体でも表現することができる。

20

【 0 0 0 8 】

手短に要約して述べると、このシステムは、先ず最初に、動きの局所的な推定を行ない、それには、1つの画像フレーム i からその次の画像フレーム $i + 1$ へかけての、幾つかのピクセルから成る小近傍の動きを判定して、画像の光学流れモデル、又は稠密な動きモデルすなわち稠密動きモデルを生成する。このシステムは続いて、アフィン変換式等の低次の滑らかな変換式を用いて、このシステムが先にコヒーレント動きをしているものと識別したところの、即ちフレーム $i - 1$ からフレーム i へかけての動きを解析した際に識別したところの、複数の領域から成る1つの領域集合の中の動きの推定を行なう。このシステムは続いて、それら領域の各々ごとに動きモデルすなわち領域動きモデルを作成する。

30

【 0 0 0 9 】

続いてこのシステムは、互いに類似している動きモデルどうしをグループ化、又はクラスタ化し、そして、その画像についての、更新した動きモデルの集合を反復して生成する。このシステムは続いて、局所的動き推定結果に基づいて、画像中の個々のピクセルを、そのピクセルの動きを最も良く近似している動きモデルに割当てることによって、コヒーレント動き領域の更新を行なう。このシステムは、こうして更新した領域に基づいて、再びその動きモデルを更新し、そして更なる更新をすべきである場合には、コヒーレント動き領域を更に更新し、以下同様にして、反復実行してもピクセルの割当てが殆ど変化しなくなったならば、反復実行を終了する。

【 0 0 1 0 】

40

続いてこのシステムは、フレーム $i + 1$ 及び $i + 2$ についての動き解析を行ない、その際に、このフレーム・ペアにおけるコヒーレント動き領域の初期推定結果として、その直前のフレーム i 及び $i + 1$ の解析において識別した更新した領域の集合を用いる。このシステムは続いて、局所的動きの推定処理と、上述の領域の動きを推定する処理と、ピクセル割当て処理とを反復実行し、そして、今回のフレーム・ペアに対応した、反復実行してもピクセルの割当てが殆ど変化しない更新領域が識別された時点で、その反復実行を終了する。続いてこのシステムはその更新領域の集合を、その次のフレーム・ペアの解析において使用し、以下同様にして、フレーム・シーケンスの中の全てのフレーム・ペアの解析を完了するまで以上の処理を繰り返す。

【 0 0 1 1 】

50

このシステムは、全てのフレーム・ペアの解析を完了した時点において、既に画像のセグメンテーションを完了して、その画像を複数のコヒーレント動き領域に分け終っている。従ってこのシステムは、それら複数の領域の夫々に対応する層を「抽出」する必要があり、即ち、それら夫々の層を、(1)ピクセル輝度値と、(2)その層に対応した動きモデル・パラメータと、(3)画像中における「深さ」の順序とによって規定する必要がある。それら層の順序付けは、覆い覆われている関係と不透明性の関係とを保存するためのものであり、なぜならば、この順序付けによって、前景の物体に対応している層が、背景中の物体に対応している層の手前に位置付けられるからである。

【0012】

このシステムは、層抽出を行なうためには、層の中の各々のピクセル位置ごとにピクセル輝度情報を決定する。従って、このシステムは、各フレームにおいて処理対象の領域に関係している個々のピクセルの動きを判定する。このシステムは1つのフレームを選択し、例えばフレーム・シーケンスの中央のフレーム等を選択する。そして、その被選択フレーム以外の各々のフレームの中の、処理対象の領域内に存在するピクセルを、被選択フレームの中の対応する領域に対して位置揃えし、それを、その領域に付随している動きパラメータを用いて、1つのフレームからその次のフレームへのピクセルの動きをたどることによって行なっている。またそのときに、このシステムは、それら動きパラメータに修正処理を施し、即ち、それら動きパラメータどうしを適当に結合させて、それら動きパラメータが、所与のフレームの中の領域と被選択フレームの中の同じ領域との間での動きを記述できるようにしており、これについては後に詳しく説明する。続いてこのシステムは、その層の中の各ピクセル位置ごとの輝度値を決定し、この輝度値は、互いに位置揃えした夫々のピクセルの輝度値の組合せに基づいた値である。そのためにこのシステムはそれらピクセルの輝度値を積算し、即ち、それら輝度値を時間の流れの中で結合し、そして、その結合における中央値を算出する。続いてこのシステムは、算出したその中央値を、その層の中の対応するピクセル位置に割当てて、これによって層の輝度マップが得られる。続いてこのシステムは、残りの各層について位置揃えとピクセル輝度値の算出とを反復して実行し、それによって、各層ごとに、その層に関する輝度値マップを作成する。

【0013】

あるフレームにおいて他のものに覆い隠されているピクセルの輝度値は、層の輝度値の算出に使用しないようにしている。このシステムは、各々の層の輝度値の算出に使用したピクセルの個数をカウントしており、輝度値の算出に使用された寄与ピクセルの個数が多い層を、輝度値の算出に使用された寄与ピクセルの個数が少ない層よりも手前に置くようにして、層どうしの間の順序付けを行なう。続いてこのシステムは、輝度値マップと、修正後動きパラメータと、順序付け情報とを、1つの層として格納する。

【0014】

格納されている複数の層に復号処理を施して、それらを1つの画像シーケンスにするためには、復号器が、先ず、それら層を順序付け情報に従って重ね合わせて1つの画像を生成し、続いてその画像に対し、格納されている動きパラメータに従ってワーブ処理を施すことによって、画像フレーム・シーケンスを生成する。層抽出を実行した際に、層の中の全てのピクセル位置の各々に1つずつ輝度値を付与してあるため、フレーム・シーケンスの一部分においてある領域のある部分が他のものに覆い隠されている場合でも、その覆い隠されている部分は、輝度値マップの中では、覆い隠された状態から回復されて表現されている。従って、覆い隠している側の物体に対応した層を排除して、その覆い隠している側の物体をフレーム・シーケンスから消去することができ、そうすることによって、本来は部分的に覆い隠されていた物体を完全な姿で生成することができる。このシステムは、それら層を利用して、「超高解像度」の画像シーケンス、あるいは、フレーム・レートを異ならせた画像シーケンスを生成することもでき、それが可能であるのは、このシステムは、画像シーケンスの全体を通しての夫々のコヒーレント動き領域及びそれらコヒーレント動き領域の中の個々のピクセルの算出された動きに基づき、補間法を用いて、中間ピクセル及び/または中間フレームを正確に生成することができるからである。

10

20

30

40

50

【 0 0 1 5 】

【実施例】

本発明の以上の利点、並びにその他の利点は、添付図面と共に以下の説明を参照することによって更に明瞭に理解することができる。図 1 は、Motion Picture Expert Group (M P E G) が製作した「Flower Garden (花園) 」というタイトルの作品の、一連のフレームから成るフレーム・シーケンスのうちの 3 枚のフレームの画像を図示している。以下に、これらフレームを参照しつつ、「画像の層」の概念について更に詳細に説明して行く。

【 0 0 1 6 】

それらのフレームは、1つの撮影シーンを横切って移動するカメラから発生された一連のフレームから成るフレーム・シーケンスの一部であり、前景の立ち木 1 と、中景の地面の花園 2 と、背景の家並み 3 とを含んでいる。立ち木 1、花園 2、それに家並み 3 という夫々の物体は、このフレーム・シーケンスの中で時間的コヒーレンシーを有しており、時間的コヒーレンシーを有しているというのは、それら物体がその形状ないしテクスチャのパターンを時間の経過と共にゆっくりと変化させているという意味である。それら物体は更に、このフレーム・シーケンスの中で空間的コヒーレンシーを有しており、空間的コヒーレンシーを有しているというのは、それら物体が、異なったフレームでは異なった部分が覆い隠されているため、その部分に関する情報が後のフレームでは含まれたり含まれなかったりしているにもかかわらず、このフレーム・シーケンスの全体に亘ってその形状ないしテクスチャのパターンを滑らかに変形させつつ変化させているという意味である。立ち木 1、花園 2、それに家並み 3 は、カメラに対する相対位置が互いに異なっているため、このフレーム・シーケンスの全体に亘って夫々に異なった動きをしている。

【 0 0 1 7 】

図示した画像は、それにセグメンテーション処理を施すことによって、各層が夫々、立ち木 1 と、家並み 2 と、花園 3 と、実質的に動きのない背景即ち空 4 との各々に対応した、合計 4 つの層 1 a ~ 4 a に分けることができる。そして、そのように 4 つの層に分けたならば、それら 4 つの層 1 a ~ 4 a の各々をその層に付随する動きパラメータに従って変形させつつ、それら 4 つの層を合成することによって、このフレーム・シーケンスを表わすということが可能になる。従ってそれら層と、それら層の変形量と、時間の経過に従ったそれら層の変形の仕方を記述した数式とで、この画像シーケンスを表わすことによって、この画像シーケンスを圧縮形で表わすことができる。ここでいう層とは、アニメーション用のセルのようなものであり、アニメーション用のセルは、例えば、静止或いは動いている背景に対してそのセルを相対的にずらしたり平行移動させたりすることによって、そのセルに描き込んである人物や物体の動きを表現することができる。ただしここでいう層は、アニメーション用のセルとは異なり、それら層を利用することによって、アニメーション用のセルの平行移動よりはるかに複雑な動きを表現することができる。

【 0 0 1 8 】

画像にセグメンテーション処理を施して複数の層に分けるには、複数の物体の境界を判定することが、即ち、複数のコヒーレント動き領域を判定することが必要であり、また更にそれと同時に、それら領域の夫々の動きを判定することも必要である。本発明者らは、アフィン変換式を用いてそれらの動きを表わすことにしたが、ただし、アフィン変換式に限らず、その他の種類の、滑らかな低次の関数の集合を使用することも可能である。アフィン動きモデルでは、ある動きを規定するのに、平行移動、剪断変形、縮倍変形、ないしはそれらの組合せとして規定するため、使用するパラメータの個数は 6 個になる。例えば回転は、垂直剪断変形と水平剪断変形との組合せである。

【 0 0 1 9 】

次に図 2 について説明すると、システム 10 は、データを符号化し、従って圧縮し、そしてその圧縮したデータを復号化して画像を再生するシステムであり、符号器 12 を含んでいる。符号器 12 は、一連の画像フレームから成る画像フレーム・シーケンスに対応したデジタル画像データを符号化して圧縮し、それによって、(i) 各層がその画像の中

10

20

30

40

50

のコヒーレント動き領域の１つずつを表わしている一連の順序付けされた複数の層と、(ii) それら層の各々に付随する、１つのフレームからその次のフレームへかけてのその層の変形を記述した動きパラメータとを生成する。データ格納／伝送装置１４は、様々な層とそれら層に付随する動きパラメータとを表わしているデータを格納すると共に、その情報を復号器１６へ伝送するものであり、復号器１６は、それら様々な層を組み立てて１つの画像を生成した上、更に、格納されていた動きパラメータに従って様々な層をワープさせて行くことによって画像シーケンスを再生する。続いて復号器１４は、このようにして復号化したフレーム・シーケンスを、ビデオ・ディスプレイ・ユニット（不図示）へ送出してディスプレイさせる。

【００２０】

次に図３について説明すると、符号器１２は、局所的動き推定部２２を含んでおり、この局所的動き推定部２２は、前後に連続した２枚のフレーム i と $i+1$ との間における、ピクセルの小さなアレイ又はピクセル近傍の中の動きを推定するものである。この近傍は、例えば３ピクセル×３ピクセルのアレイとすることができる。推定部２２は、それら近傍の各々について、その平均的動きを表わす速度ベクトルを発生し、そしてその速度ベクトルを、その近傍の中央に位置しているピクセルに対応付ける。続いてこの推定部２２は、その次のピクセルに対応した、先の近傍との間に重なり部を有する新たな近傍を選択し、その平均的動きベクトルを決定し、そしてその速度ベクトルを、その新たな近傍の中央に位置するピクセルに対応付ける。推定部２２は、以下同様に反復実行して、画像中の全てのピクセルの各々に動きベクトルを対応付ける。これによって例えば図４に示したような、その画像の光学流れモデル又は稠密動きモデルが得られる。以下に、図５並びに図６および図７を参照しつつ、この局所的動き推定部の動作について更に詳細に説明して行く。

【００２１】

動きセグメンテーション処理部２４は、コヒーレント動き領域推定部２４ａと動き推定部２４ｂとで構成されており、それらは協働して、画像中の複数のコヒーレント動き領域と、それら領域に付随する夫々の動きモデルの集合とを同時に判定する。動き推定部２４ｂは、コヒーレント動き領域推定部２４ａが識別した複数のコヒーレント動き領域における夫々の動きを推定して、それら領域の各々に対し１つずつの推定した動きモデル（例えばアフィン動きモデル）を生成する。続いて、動き推定部２４ｂは、互いに類似した動きモデルどうしをまとめてグループ化し、各グループごとに１つずつの合成動きモデルを生成する。続いてコヒーレント動き領域推定部２４ａが、個々の画像ピクセルをそれら合成動きモデルに対応付けて行き、その際に、個々のピクセルを、それら動きモデルのうち、そのピクセルの局所的動きを最も良く近似して表わしている動きモデルに対応付けるようにする。この処理部は、これを行なうことによって、更新したコヒーレント動き領域を生成する。この処理部は、それら更新した領域を用いて更に新たな動きモデルを生成し、それによって更に領域を更新するというようにして、反復実行により更新を繰り返す。

【００２２】

その反復実行の回数が所定回数に達した時点で、或いは、ある反復実行によって得られたピクセルの対応付けの結果が、その直前の反復実行によって得られたピクセルの対応付けの結果と殆ど変わらないということが最初に起きた時点で、処理部２４は、動きモデルのパラメータと、夫々のコヒーレント動き領域を特定している情報とを、そのときの処理対象のフレーム・ペアのうちの後の方のフレーム、即ちフレーム $i+1$ に対応した格納位置に格納する。処理部２４は更に、こうして特定されたコヒーレント動き領域を、続く次のフレーム・ペア、即ちフレーム $i+1$ 及び $i+2$ を対象とした解析における、コヒーレント動き領域の推定の初期値として使用する。以下にこの動きセグメンテーション処理部２４の動作について、図８を参照しつつ更に詳細に説明して行く。

【００２３】

局所的動き推定部２２は、各々のフレーム・ペアごとに、局所的動きの推定の処理と各ピクセルに１つずつの動きベクトルを対応付ける処理とを行なって行く。続いて、動きセ

10

20

30

40

50

グメンテーション処理部 24 が、その処理対象の画像のセグメンテーション処理を実行して、その画像を複数のコヒーレント動き領域に分ける。この処理は、上述の処理手順に従って、先行フレーム・ペアに関して識別された夫々のコヒーレント動き領域を、現在フレーム・ペアにおける夫々のコヒーレント動き領域の推定の初期値として使用して行なうものであり、これを次々と反復実行して、そのフレーム・シーケンスの中の全てのフレーム・ペアについての解析を完了する。従ってこの処理部は、そのフレーム・シーケンスの中にあつてコヒーレント動きを行なっている夫々の領域を特定すると共に、各々のフレーム・ペアごとに、そのフレーム・ペアの第 1 のフレームから第 2 のフレームへかけてのそれら夫々の領域の動きを記述した動きモデルの集合を決定する。

【0024】

層抽出処理部 26 は、フレーム・シーケンスの中の全てのフレームの夫々に対応したコヒーレント動き領域の情報とそのコヒーレント動き領域に関連した動きモデルの情報とを結合して、各コヒーレント動き領域ごとに 1 つずつの層を生成する。この処理部 26 は、先ず最初に、フレーム・シーケンスの中央に位置するフレームを選択し、その被選択フレームより先行しているフレーム及びその被選択フレームより後のフレームの各々について、そのフレームの中の様々なコヒーレント動き領域をどのように変形させれば、それらコヒーレント動き領域を、被選択フレームの中のそれらに対応するコヒーレント動き領域に位置揃えすることができるかを判定する。そのためにこの処理部 26 は、ある 1 つのフレームの中のある 1 つのコヒーレント動き領域に付随している動きパラメータを、そのフレームと被選択フレームとの間に存在している中間フレームの中の同じ領域に付随している動きパラメータと結合して、修正した動きパラメータ集合を生成するようにしている。

【0025】

処理部 26 は、こうして得られた修正後動きパラメータを用いて、被選択フレームの中の領域に、各々のフレームの中のそれらに対応する夫々の領域を位置揃えし、そして、その領域内の全てのピクセル位置の各々に対する夫々の合成ピクセル輝度値を算出する。続いてこのシステムは、決定したそれら輝度値を用いて、層の輝度値マップを作成する。

【0026】

処理部 26 はそれら合成輝度値を算出する際に、他のものに覆い隠されているピクセルの輝度値は計算に用いないようにしている。従って、部分的に覆い隠されている物体に対応している層では、合成輝度値を算出するために結合するピクセル輝度値の個数が、そうでない層と比べて少なくなっている。処理部 26 は、合成輝度値の算出に寄与しているピクセルの個数を数えることによって、層どうしの間の順序付けを行なうようにしている。従って各々の層は、(i) ピクセル輝度値のマップと、(ii) 被選択フレームからフレーム・シーケンスの中のその他の各々のフレームへかけてのその層の動きを規定した修正後動きパラメータの集合と、(iii) 画像を復号化即ち再生する際に、その層をその他の層とどのようにして合成すべきかを指示した層順序情報(層深さ情報)とで構成されている。また、処理部 24 は、必要に応じて、各層の中に更に、減衰、動きブレ、等々に関する情報を含ませるようにもしている。続いて処理部 26 は、それらの層情報をデータ格納/伝送装置 14 (図 1) へ送出して、格納及び/または伝送を行なわせる。以下に、この層抽出処理部 26 の動作について、図 11 を参照しつつ更に詳細に説明して行く。

【0027】

システム 10 は、局所的動きの推定から出発して再び局所的動きの推定へ帰るようにして、それら局所的動きに関係したコヒーレント動き領域の動き及び境界の決定を反復して行なうことにより、画像の中の複数のコヒーレント動き領域の内部における単一の動きの推定を複数回に亘って行なう、その画像全体の中に存在している複数の動きを解析するという課題を達成している。このようにしているため、システム 10 は、僅かな数の層と、それら層に付随する動きパラメータとで、画像データを表わすことが可能になっている。従って、このシステム 10 では、画像データを格納する際に、従来公知のシステムにおいて必要とされていた記憶容量よりも小さな記憶容量の中に格納することと、画像データを伝送する際に、従来公知のシステムにおいて必要とされていた帯域幅よりも狭い帯域幅

10

20

30

40

50

で伝送するということが、潜在的に可能になっている。例えば、30枚のフレームから成るフレーム・シーケンスを、各々が複数の層から成る数枚の静止画像と、1つの層の1枚のフレームについて6個ずつのパラメータとで表わすことができる。

【0028】

次に図3～図7を参照しつつ説明すると、局所的動き推定部22は、局所的動きを推定するために、ある1枚のフレーム20_iからその次のフレーム20_{i+1}へかけての、複数のピクセル31から成る小さなアレイ又は近傍30の内部での動きを判定する。画像の動きの大きさに関してはいかなる前提条件も存在していないため、即ち、その動きが小さい場合も大きい場合もあり得るため、各々の近傍の内部の動きを判定するために、「初めは粗く次第に細かく」という方式の動き推定方法を用いるようにしており、それによって、
10
大きな動きも小さな動きもいずれも追跡できるようにしている。好適実施例においては、この推定部22は、マルチ・スケール最小二乗法を用いたものにしている。基本的に、この局所的動き推定部22は、画像から「ズーム」アウトするために、その画像に対応したガウス・ピラミッド(図6)を形成し、即ち、大きな動きであっても小さな動きとして検出されるようになるまで、その画像の解像度を次第に低下させた表示を次々と生成して行く。この推定部22は、エリアシングの発生を防止するために、2枚のフレームの画像に対してガウス・フィルタを用いたフィルタ処理を施すことによって、それら画像にボケを導入するようにしている。続いてこの推定部22は、そのフィルタ処理後データに対してサブサンプリング処理を施し、即ち、そのフィルタ処理後データの、例えばその他の全てのピクセル位置のデータ等の部分集合を抽出し、それによって、新たな画像21_i及び21_{i+1}の集合を形成する。この推定部22は、それら画像に対して更にフィルタ処理を施し、そのフィルタ処理後データのサブサンプリングを行なうことによって、更に低い解像度の画像を生成して、適度に低い解像度の画像I_i及びI_{i+1}が生成されるまで、この処理を反復実行する。
20

【0029】

続いてこの推定部22は、こうして得られた新たな画像I_i及びI_{i+1}の互いに対応する近傍どうしの間の小さな動きを、公知の勾配方式の最小二乗推定法を用いて推定する。従ってこの推定部22は、その低解像度の近傍に関して、画像I_iの中のその近傍30の内部の位置の関数としてピクセル輝度値を表わしている曲線F(x)と、画像I_{i+1}の中のその近傍30の内部の位置の関数としてピクセル輝度値を表わしている曲線G(x)との間の「隔たり」を表わす、動き速度ベクトルhを決定する。2次元最小二乗法という解析法を用いて、以下の式を反復して解くことによって、F(x)からのG(x)の変位の大きさを表わすこのベクトルhを求めることができる。
30

【数1】

$$h_{k+1} = h_k + \sum w(x) \frac{\partial [F(x-h_k) + G(x)]}{\partial x} \frac{[G(x) - F(x-h_k)]}{2} C^{-1}$$

$$\text{ここで、 } C = \sum w(x) \frac{\partial [F(x-h_k) + G(x)]}{\partial x} \frac{\partial [F(x-h_k) + G(x)]}{\partial x}$$

ここで、 ∇_x は勾配演算子であり、 $w(x)$ は、その動きを推定しようとしているピクセルから遠く離れたピクセルほど寄与分が減少するようにするための重み付け係数である。小さな動きを検出するための同様の方法に、ときに「画像見当合わせ(image registration)と呼ばれる方法があり、これについてはLucasとKanadeによる次の論文の中に詳細に記載されている：「An Iterative Image Registration Technique with an Application to Stereo Vision by Lucas and Kanade, Image Understanding Workshop, April, 1981, pp.121-130」。推定部22は、動きを解析する際に、G(x)が、その新たな位置を占めるためにはどこから移動したはずであるかという、元の位置を判定するようにしており、即ち、「ワーブ元」を解析するようにしている。従ってこの推定部22は、上掲のLucasとKanadeの論文に記載されている方法の変形方法を用いているということが
40
50

でき、Lucasと Kanade の論文に記載されている方法では、動きを解析するのに、曲線がどこへ向かって動いているのか、即ち「ワープ先」を解析するようにしている。推定部 2 2 は、「ワープ元」という方式を用いて、動きの解析を行なうことによって、ピクセル 1 個分の大きさ以下の小さな動きを、より正確に判定し得るものとなっている。

【 0 0 3 0 】

画像 I_i 及び I_{i+1} についての小型動き解析によって判定された動き情報は、ガウス・ピラミッドにおける 1 つ上の階層へ伝播され、この階層において、より高い解像度を有するそれら画像を用いて動き解析が行なわれる。この階層の第 1 番目の画像は、基本的に、それより 1 段階低い解像度の階層で推定された動きに対応した位置へ変位しており、即ちワープされており、この画像について小型動き解析を再び実行する。それによって新たな推定が得られたならば、その推定を、ガウス・ピラミッドの更に 1 つ上の階層へ伝播させ、そして、その階層における動き解析を実行し、最終的には、解像度を低下させる前の元の解像度を有する画像の、被選択近傍の中での動きが判定されることになる。図 7 には、階層数が 3 層のピラミッドを示した。

【 0 0 3 1 】

以上の局所的動き解析を、画像中の全てのピクセルの各々について行なうことによって、その画像の光学流れ動きモデル又は稠密動きモデル（図 4）が作成される。上述の近傍の寸法は、この「初めは粗く次第に細かく」という方式の動き検出法を用いたときに、小さな動きのディテールが失われることもなく、また、大きな動きの検出も正確に行なえるような寸法に選定する。

【 0 0 3 2 】

次に図 8 について説明すると、動きモデル推定部 2 4 b（図 3）の一部をなしているアフィン動きモデル推定部 3 2 は、局所的動き推定部 2 4（図 3）が生成した稠密動きモデルの解釈を行なって、フレーム・ペア（ i 、 $i+1$ ）に関するコヒーレント動き領域の夫々のアフィン動きモデルを表わす推定を生成する。物体の形状に関する初期制約条件は存在しておらず、従ってコヒーレント動き領域の推定の初期値は存在しないため、上述の処理部は、フレーム・シーケンスの中のフレーム 0 とフレーム 1 とに関連しているコヒーレント動き領域の推定の初期値としては、各々が n ピクセル \times m ピクセルの寸法の r 個の矩形ブロックから成るグリッドを用いる。続いてこの処理部は、コヒーレント動き領域の各々の内部の動きを、次の形のアフィン変形式としてモデル化する。

【 数 2 】

$$V_x = a_{x0} + a_{x1}x + a_{x2}y$$

$$V_y = a_{y0} + a_{y1}x + a_{y2}y$$

これらの式において、 V_x と V_y とは夫々、位置（ x ， y ）に存在しているピクセルの動き速度ベクトルの x 成分と y 成分とであり、また、 a_k は、 x 方向ないし y 方向における、夫々、平行移動、剪断変形、及び傾斜変形を表わすアフィン・パラメータである。

【 0 0 3 3 】

推定部 3 2 は、線形回帰法を用いることによって、アフィン動きパラメータの推定結果の値を、それに関係している局所的動きベクトルから求めるようにしている。また、計算量を減らすために、各々の速度成分ごとに個別にこの線形回帰法を適用するようにしており、これが可能であるのは各々の速度成分が互いに独立だからである。ここで、 H_i を、アフィン・パラメータ空間内のアフィン・パラメータの第 i 番目の推定の値とし、その x 成分と y 成分とを、 H_{xi} と H_{yi} とし、更に を回帰ベクトルとするならば、以下の式が得られる。

【 数 3 】

10

20

30

40

$$H_i = [H_{yi} \ H_{xi}]$$

$$H_{xi}^T = [a_{xi} \ a_{di} \ a_{zi}]$$

$$H_{yi}^T = [a_{yi} \ a_{li} \ a_{zi}]$$

$$\phi^T = [1 \ x \ y]$$

そして、

$$[H_{yi} \ H_{xi}] = [\sum_n \phi \phi^T]^{-1} \sum_n (\phi^T [V_r(x,y) \ V_d(x,y)])$$

10

以上の式において、 P_i はその画像中の第 i 番目の領域である。これは、基本的に、速度空間内における平面の当てはめに他ならない。

【0034】

アフィン・パラメータ推定部32は、ある領域についてのアフィン動きパラメータを推定したならば、従って、その領域に対する推定した動きモデルを得たならば、それに続いて、その動きモデルを、その動きに対応している複数の局所的動きの推定の夫々と比較することによって、その領域の全域に互る分散の値を算出する。続いて、この推定部32は、その分散の値の逆数である信頼性等級をその動きモデルに付与する。信頼性等級が低い推定は「悪い」推定であると見なされ、無視されることになる。「悪い」推定は、典型的な例としては、1つないし2つ以上の物体境界を含んでいる領域に付随するものであり、なぜならば、そのような領域の中に存在している2つの動きが単一のアフィン動きモデルには適合するということは、通常あり得ないからである。

20

【0035】

初期値として用いる領域の寸法は、ある1つの領域が物体境界を包含する確率を低く抑えることができる十分に小さな寸法とする一方で、その領域に付随するアフィン動きパラメータの合理的な推定値を得るための基礎となるデータが十分に含まれるような十分に大きな寸法となるように選定している。

【0036】

アフィン・パラメータ推定部32は、設定した上述のグリッドをなしている複数の矩形ブロックの各々に対して1つずつ、合計 r 個の動きモデルを生成する。それら動きモデルのうちには、互いに同じ物体に対応した動きモデルであるものがあり、そのような動きモデルどうしは互いに類似しているはずである。従ってそのような動きモデルを1つにまとめるようにすれば、画像のセグメンテーション処理の精度を更に高めることができる。

30

【0037】

再び図6および図7について説明すると、クラスタ化処理部34は、 r 個のアフィン動きモデルの夫々に付随するパラメータを受け取り、それらのうちで最も信頼性の高い方から C 個の動きモデルを、クラスタ化「中心」の初期値として選択する。選択するそれら中心の各々は、その他のいずれの中心からも、少なくとも、アフィン・パラメータ空間における正規化距離の最小値として定めた距離だけ離れていなければならない。

40

【0038】

クラスタ化処理部34は、 k 平均クラスタ化法を用いて、選択した各々の中心から所定半径以内に存在している動きモデルをまとめてグループ化し、それによって C 個の初期クラスタを形成する。続いてこの処理部34は、所定個数未満の動きモデルしか含んでいないクラスタを無視して、それ以外の夫々のクラスタについて、新たなクラスタ中心の推定値を生成する。即ち、この処理部34は、任意の1つのクラスタに関して、そのクラスタの中に含まれている複数の動きモデルの各々に付随している夫々のアフィン・パラメータの平均値を求め、それら平均値を新たな中心のパラメータとして使用する。それら平均値に対して、動きモデルに付随している信頼性等級値で重み付けをすることが適当である場合には、そのようにしても良い。

50

【 0 0 3 9 】

もし、新たな中心のうちの、ある 2 つの中心の間の距離が、アフィン・パラメータ空間内の所定の離隔距離だけ離れていなかったならば、それら 2 つの中心に対応しているクラスタどうしを併合した上、その併合クラスタに対応した新たな 1 つの中心を算出する。クラスタ化処理部 3 4 は、このクラスタ化の処理を行なっている間に、もはや処理結果に変化が生じなくなったならば、その処理結果として得られた q 個 ($q \leq r$) のアフィン動きモデルを、仮説検定部 3 6 へ受け渡し、この仮説検定部 3 6 が、夫々のピクセルを個別に、それら動きモデルの各々と突き合わせて、最良のピクセル - 動きモデル間の一致を判定し、即ち、それら動きモデルのうちのどの動きモデルが、そのピクセルの推定した局所的動きを最も良く近似して表わしているかを判定する。

10

【 0 0 4 0 】

検定部 3 6 は、この仮説検定を実行する間に、各ピクセルをそのピクセルの位置において誤差を最小にする動きモデルに割当てて行くが、それには、次の式で示す値を最小にする動きモデルを選択するようにする。

【 数 4 】

$$(V_{\text{actual}}(x,y) - V_{\text{Hi}}(x,y))^2$$

ここで、 $V_{\text{actual}}(x,y)$ は、局所的動きベクトルのベクトル場であり、また、 V_{Hi} は第 i 番目の動きモデル又は仮説に対応したアフィン動きの場であって、動きモデル $i = 0, 1, \dots, q$ の検定を行う。位置 (x, y) にあるピクセルは、その位置においてそのピクセルの局所的動きを最も良く近似して表わしている動きモデルに割当てるようにする。いずれの動きモデルとの間で算出した誤差も所定の最小値を超えているピクセルが存在していた場合には、検定部 3 6 は、そのピクセルを割当てないままにしておく。続いて検定部 3 6 は、画像のセグメンテーション処理を実行し、それには、1 つ 1 つのピクセルごとに仕分けを行なって、 q 個の領域と 1 つないし複数の未割当てピクセルとに分ける。検定部 3 6 は、未割当てのピクセルの割当てを次のようにして行なっても良く、即ち、フレームとフレームとの間におけるそのピクセルの輝度値の変化量を、夫々の動きモデルによって予測される夫々の輝度値の変化量とを比較した上で、そのピクセルの輝度値を所定の誤差の範囲内で最も良く近似して予測している動きモデルに、そのピクセルを割当てるというものである。従ってこのシステムは、あるピクセルの局所的動きの推定が悪い推定である場合でも、そのピクセルをいずれかの領域に割り振ることができる。

20

30

【 0 0 4 1 】

こうして得られた q 個の領域に基づいて、アフィン・パラメータ推定部 3 2 がアフィン・パラメータの値を更新する。そして、その更新されたアフィン・パラメータの値に基づいて、クラスタ化処理部 3 4 がクラスタ化処理を、仮説検定部 3 6 が仮説検定処理を実行することにより、今度はそれら領域が更新される。更に、その更新された領域に基づいてアフィン動きモデルの生成が再び行なわれ、こうして新たに生成されたアフィン動きモデルを対象として再び上述のクラスタ化処理と仮説検定処理とが実行される。この反復実行は、反復実行して得られた領域が、その直前の実行によって得られた領域と殆ど変化がなくなった時点か、或いは、反復実行の回数が適当に定めた最大回数に達した時点で終了する。以上の反復実行処理によって生成された、あるフレーム・ペアに関する最終的な更新された領域は、フレーム・シーケンスの中のそれに続く次のフレーム・ペアの画像の動きセグメンテーション処理のための領域の初期値として使用される。

40

【 0 0 4 2 】

符号器 1 0 (図 3) が、連続するフレーム・ペアの夫々に対応した動きモデルを次々と推定して行く処理において実行する動作を、図 9 および図 1 0 のフローチャートに示した。局所的動き推定部 2 2 は、フレーム i からフレーム $i + 1$ へかけての個々のピクセルの夫々の局所的動きを、マルチ・スケール勾配方式の方法を用いて推定して行き、それによって画像の稠密動きモデルを生成する (ステップ 4 0)。これに続いて、フレーム $i - 1$ からフレーム i へかけての動きの解析によって得られた複数のコヒーレント動き領域に対

50

応した、夫々のアフィン動きパラメータを算出する（ステップ42）。なお、そのとき解析しているのが、フレーム・シーケンスの中の第1番目のフレーム・ペアである場合には、互いに重なり部分を持たない任意の複数の領域を用いるようにする。

【0043】

このシステムは、互いに類似している動きモデルどうし、即ち、対応したパラメータ空間内において所定の最小離隔距離よりも更に互いに近接して存在している動きモデルどうしを、まとめてグループ化し、幾つかのクラスタを形成する。最初のクラスタ中心は予め決めてあり、中心間の距離がアフィン・パラメータ空間内における所定離隔距離より小さいクラスタどうしは併合し、一方、ある1つのクラスタに包含されている動きモデルどうしが所定の最大正規化距離以上に互いから離隔している場合には、そのクラスタを2つのクラスタに分割する（ステップ44）。続いてこのシステム10は、同じ1つのクラスタの中に包含されている複数の動きモデルに付随している夫々のアフィン動きパラメータの平均値を求め、そして、各クラスタごとに、新たなアフィン・パラメータ集合が付随した新たなクラスタ中心を生成する（ステップ46）。以上を更に反復実行することによって、クラスタを繰り返し更新し、そして、ある反復実行をしたときに、それによって割当てを変更された動きモデルの個数が、所定個数に達しなくなった時点で、或いは、反復実行の回数が適当に定めた最大回数に達した時点で、反復実行を終了する（ステップ46～48）。

【0044】

続いてこのシステムは、以上のようにして得られた複数の動きモデルの各々に個々のピクセルを突き合わせ、そして個々のピクセルを、それら動きモデルのうちで、そのピクセルに対応した位置においてそのピクセルの局所的動きを最も良く近似して表わしている動きモデルに割当てする（ステップ52）。これに関して、もしあるピクセルpを2つ以上の動きモデルが完全に同一程度に近似して表わしていた場合には、そのピクセルpを、そのピクセルpに最も近い幾つかの画像ピクセルに関連して推定された動きに基づいて生成された動きモデルに割当てするようにしている。また、これと別の方法として、そのピクセルpを、割当てられたピクセルの個数が最大の動きモデルに割当てするようにしても良い。あるピクセルの局所的動きを、どの動きモデルも、所定の最小誤差（例えば±1ピクセル）以内の精度で近似していないという場合には、そのようなピクセルは、未割当てピクセルとして取扱うようにし、この種のピクセルは、輝度値の変化量の解析によってその割当てを決定するようにしても良い。

【0045】

領域を更新したならばそのたびに、それら領域のアフィン・パラメータの値を推定する処理手順と、クラスタ化する処理手順と、ピクセル割当ての処理手順とを反復実行し、それらによって、それら更新した領域に更に更新を施し、これを反復実行して更新を繰り返す。そして、新たに反復実行して得た領域が、その直前の反復実行によって得た領域と殆ど変わらないものとなったか、或いは、反復実行の回数が、適当に定めた最大回数に達したならば、その時点で反復実行を終了する（ステップ54～58）。続いてこのシステムは、更新して得たそれら領域を、続く次のフレーム・ペアにおける複数の動きとそれら動きに付随する夫々のコヒーレント動き領域とを解析するための初期条件として使用し、これを次々と繰り返して全てのフレーム・ペアの解析を完了する（ステップ60～62）。解析が完了したフレームの数が増えるにつれて、コヒーレント動き領域の推定をより高精度で行なえるようになり、また、それら領域に付随するアフィン・パラメータが安定して行く。従って、新たなフレームにおける動きを解析するために必要な時間は次第に短縮されて行く。

【0046】

フレームi及びi+1に関して判定された夫々の領域に付随するアフィン・パラメータは、更に、そのフレーム・シーケンスの中の続く次のフレームi+2における動きモデルとして使用するようにしても良く、そうすることによってシステムを安定化させることができる。更に、それらアフィン・パラメータは、仮説検定処理における、動きの予測値と

10

20

30

40

50

して使用することができる。例えば、続く次のフレームにまで同じ動きが継続しているような場合には、ピクセルをその動きモデルに容易に割当てることができ、そのため、収束させるための反復実行の回数を減らすことができる。

【 0 0 4 7 】

以上のようにしてコヒーレント動き領域の識別を完了したならば、このシステムは、続いて、それら領域に対応した夫々の層を抽出する処理を実行し、そのために、動き補償処理（ステップ 6 4）を実行する。あるコヒーレント動き領域の境界とその領域に付随するアフィン動きパラメータとが共に正確に判定されているならば、適当なアフィン変換式を用いて画像にワープ処理を施すことによって、各フレームの中の互いに対応する領域どうしを位置揃えすることができる。従って、例えば図 1 の立ち木等の物体を、適当なワープ処理を施したフレーム・シーケンスの中に静止させたままで、花壇と家並みとを動かすようにすることも可能である。また、互いに異なったフレームの中の互いに同一の領域の中に存在する対応するピクセルどうしを結合させて、その領域に対応した層における輝度値を、その層の中の各々のピクセル位置ごとに決定することができるようにしている。

【 0 0 4 8 】

いずれのフレームであれ、1 枚のフレームからだけでは得られないような情報を層に包含させることも可能であり、例えば、他のものによって覆い隠されている位置にあるピクセルの輝度値のような情報も、そのピクセルがフレーム・シーケンスの中のその他全てのフレームにおいても同様に覆い隠されているのでない限り、層に包含させることができる。尚、あるピクセルが全てのフレームにおいて覆い隠されているような場合には、そのフレーム・シーケンスを正確に符号化することは要求されない。

【 0 0 4 9 】

次に図 1 1 について説明すると、層抽出処理部 2 6（図 3）は、層の輝度値マップを作成するために、フレーム・シーケンスの中の中央に位置するフレーム f_s を選択する。続いて層抽出処理部 2 6 は、その他のフレームのコヒーレント動き領域をこのフレームの対応する領域に対して位置揃えし、それによって層抽出処理部 2 6 は、様々なフレームから得たピクセル情報を組合せて、領域の輝度値マップ、又は層の輝度値マップを生成することができるようになる。これを行なうために、このシステムは、フレーム・シーケンスの中の夫々のフレームに付随している動きパラメータに修正を加えて、そのフレームから被選択フレームへの動きを記述するパラメータを生成する（ステップ 7 2）。被選択フレームよりも先行しているフレーム i については、このシステムは、そのフレーム i に付随しているアフィン・パラメータ P_i と、そのフレーム i から被選択フレームまでの間に存在している中間フレームに付随しているアフィン・パラメータ P_j （ここで $j = i + 1, i + 2, \dots, s$ である）とを組合せて、フレーム i からフレーム s へかけての動きを表わす修正後動きパラメータ P_i を生成する。説明を分かり易くするために、 $i = 0$ 、且つ、 $s = 2$ であるものと仮定し、従って、中間フレームが 1 枚だけであるものとし、また更に、 I_2 が画像データを表わし、 x_2 がピクセル位置を表わすものとすれば、修正後アフィン・パラメータは、以下に示すように定義される。

【 数 5 】

$$I_2(x_2) = I_1(x_2 - V_2)$$

この式は「ワープ元」を記述している式であり、この式において、 $V_2 = A_1 + B_1 x_2$ はアフィン変換式であり、更にここで、 A_1 はフレーム 1 からフレーム 2 へかけての定数アフィン・パラメータを表わし、 B_1 はフレーム 1 からフレーム 2 へかけての変数アフィン・パラメータの係数を表わしている。従って次のようになる。

【 数 6 】

10

20

30

40

$$I_2(x_2) = I_1(x_2 - A_1 - B_1(x_2));$$

また更に、

$$I_1(x_1) = I_0(x_1 - A_0 - B_0(x_1));$$

従って、

$$I_2(x_2) = I_0(x_2 - A_{02} - B_{02}(x_2));$$

10

ここで、

$$A_{02} = A_0 + A_1 - B_0 A_1$$

$$B_{02} = B_0 + B_1 - B_0 B_1$$

以上の式において、 $A_{0,2}$ は、フレーム 0 からフレーム 2 へかけての定数アフィン・パラメータを表わし、 $B_{0,2}$ は、フレーム 0 からフレーム 2 へかけての変数アフィン・パラメータの係数を表わしている。

被選択フレームよりも後のフレームに付随するパラメータを算出するには、以上に説明したのと同様にして、まず、被選択フレーム s からそのフレーム j へかけてのパラメータを算出し、その後、フレーム j からフレーム s へ戻るためのパラメータを以下の式によって求める。

20

【数 7】

$$A_{js} = [B_{sj} - I]^{-1} A_{sj}$$

$$B_{js} = [B_{sj} - I]^{-1} B_{sj}$$

これらの式において、 I は単位行列である。

30

【0050】

このシステムは、このようにして求めた修正後動きパラメータを用いて、各々のフレームの中の処理対象のコヒーレント動き領域を、被選択フレーム f_s の中の同じ領域に対して位置揃えする。続いてこのシステムは、それら各々のフレームの中の互いに対応するピクセル位置の輝度値を組合せるようにし、従って、ある長さの時間に亘って輝度値を組合せ、それによって、ピクセル位置に各々ごとに輝度値の中央値を求める（ステップ 74）。続いてこのシステムは、このようにして求めた輝度値の中央値を、層の中の夫々のピクセル位置に割当て、それによって、層の中の全てのピクセル位置の各々に 1 つずつのピクセル輝度値を当てはめる（ステップ 76）。

【0051】

40

処理対象の領域におけるある特定のピクセル位置が、あるフレームにおいては他のものに覆い隠されているという場合には、その領域に関する輝度値中央値の算出を行なう際に、その覆い隠されているピクセル位置の輝度値は考慮しない。このシステムは、各々のピクセル位置に対応した輝度値中央値を算出する際にその計算に使用したピクセルの総数を常に追跡している（ステップ 78）。そしてこのシステムは、その値の算出に寄与したピクセルの個数に基づいて、複数の層を、画像中の深さとして順序付けるようにしている。前景の物体は決して他のものによって覆い隠されることがないため、その全てのピクセル位置において、算出に寄与したピクセルの個数が大きな数になっている。一方、他のものに覆い隠されている物体は、そのうちの幾つかのピクセル位置において、寄与ピクセルの個数が小さな数になっている。従ってこのシステムは、覆い隠されている部分を有する層

50

を、覆い隠されている部分のない層の背後に置くことによって、層の順序付けを行なうことができる。

【0052】

このシステムは、符号化処理によって複数の層を生成するものであり、それら複数の層は、その各々が、ピクセル輝度値マップと、修正後動きパラメータの集合と、順序付け情報とを含んでいる。このシステムはそれらの情報を、後に復号化処理を実行したり伝送したりするために、格納しておくようにしている。

【0053】

それら複数の層は、このシステムに、フレーム・シーケンスの中の1枚のフレームからだけでは得られない情報を提供する。例えば、再び図1について説明すると、視線を遮る立ち木が存在しているため、フレーム・シーケンスの中に花園の全体が見えるフレームは1枚も存在していないにもかかわらず、花園に対応した層2aは、花園の全てのピクセルを含んでいる。同様に、家並みに対応した層3aは、図示のフレーム・シーケンスの最後の部分に至らねば見えない家に関するピクセル情報を含んでいる。このシステムは、これらピクセル情報の全てを保持しているため、動きパラメータに補間法を適用することによって、単位時間あたりのフレーム数を変化させる場合等に必要な中間フレームの生成を行なうことができる。更にこのシステムは、認識システムへデータを供給する際には、覆い隠されている部分を解消した後にそのデータを供給することによって、その認識システムが物体をより容易に認識できるようにすることができる。更にこのシステムは、物体の動きに基づいてその物体を認識するようにした認識システムへデータを提供するためのシステムとすることもでき、それには、その種の認識システムへ、処理対象の物体だけに関係したデータを供給するようにする。

【0054】

物体が不透明なものであれば、全ての画像情報が層及び動きパラメータの中に包含されることになる。一方、画像のある1つの部分ないしは幾つかの部分が透明である場合や、焦点ボケや動きブレ等の「特殊効果」を呈している場合には、そのような効果に関係した追加情報を、別に格納しておくようにする。更には、そのような領域の縁部における透明度の変化をより正確に表わすためには、更なる修正計算処理が必要になることもあり得る。

【0055】

ある領域が静止している場合、或いは、ある領域がテクスチャを持たない場合には、そのような領域は、1つのフレームから別のフレームへかけて変形されることのない単一の層に対応付けるようにしている。そのような領域へ動きを対応付けることにおけるいかなる誤りも、検出することは不可能であり、従って無視することになる。

【0056】

ある特定の動きが、選択した種類の動き式（本実施例ではアフィン変換式）では正確に記述することができないような動きである場合には、その動きによって影響を受けるピクセルの輝度値を制御するための誤差修正情報を層に含めておく必要がある。例えば、回転することによってその姿が著しく変化する物体を記述している層には、誤差マップを含めておくようにする。

【0057】

簡潔に要約するならば、このシステムは、コヒーレント動き領域を判定するための解析の一部として局所的動きの推定を行なって、画像を符号化して一連の複数の層にするものである。このシステムは、その解析の対象を主要な動きや小さな動きに限ってはいない。更にこのシステムは、物体の境界の割当てを行なったり、物体に境界を任意に制限したりしてはいない。そうではなくて、このシステムは、個々のピクセルを、そのピクセルの局所的動きを最も良く近似して予測している動きモデルに割当てるようにしているのである。

【0058】

画像にセグメンテーション処理を施して、その画像を複数のコヒーレント動き領域に分

10

20

30

40

50

け、その後に、それら領域の各々を個別の層に対応付けるようにしているため、画像シーケンスを、一連の順序付けした複数の層のマップとそれに付随する動きパラメータ集合という形で格納しておけるようになっている。そのため非常に僅かな記憶容量の中に画像を格納することができ、また、非常に狭い帯域幅で画像を伝送することができる。

【 0 0 5 9 】

このシステムは、各層に、その層の中の全てのピクセル位置ごとに1つずつの輝度値を含ませるようにしている。従って、このシステムにおいては、1枚だけのフレームの中に存在している画像情報以上の量の画像情報が入手可能となっている。このシステムは更に、その情報を利用して、例えば特定の物体を消去した画像シーケンスを生成することや、単位時間あたりのフレーム枚数が異なる「超高解像度」のフレーム・シーケンスを生成することもできる。このシステムは更に、その情報を利用して、ブレがなくなるように動き補償する低ノイズ化処理を行なうことも可能にしている。

10

【図面の簡単な説明】

【図1】 「花園 (Flower Garden) 」という題名の画像シーケンスの3枚のフレームとその画像シーケンスに対応した、画像の複数の層を示した図である。

【図2】 1つの画像を一連の複数の層として符号化及び復号化するためのシステムの機能を示したブロック図である。

【図3】 図2のシステムに含まれている符号器の機能を示したブロック図である。

【図4】 図1に示したうちの1枚のフレームの稠密動きモデルを示した図である。

【図5】 局所的動きの判定に用いるピクセル近傍を示した図である。

20

【図6】 ガウス・ピラミッドの中での画像の変化を示した図である。

【図7】 ガウス・ピラミッドの中での画像の変化を示した図である。

【図8】 図2のシステムに含まれている動きセグメンテーション処理部の機能を示したブロック図である。

【図9】 図2のシステムに含まれている局所的動き推定部及び動きセグメンテーション処理部の動作を示したフローチャートであり、図10へ続く図である。

【図10】 図2のシステムに含まれている局所的動き推定部及び動きセグメンテーション処理部の動作を示したフローチャートであり、図9から続く図である。

【図11】 図2のシステムに含まれている層抽出処理部の動作を示したフローチャートである。

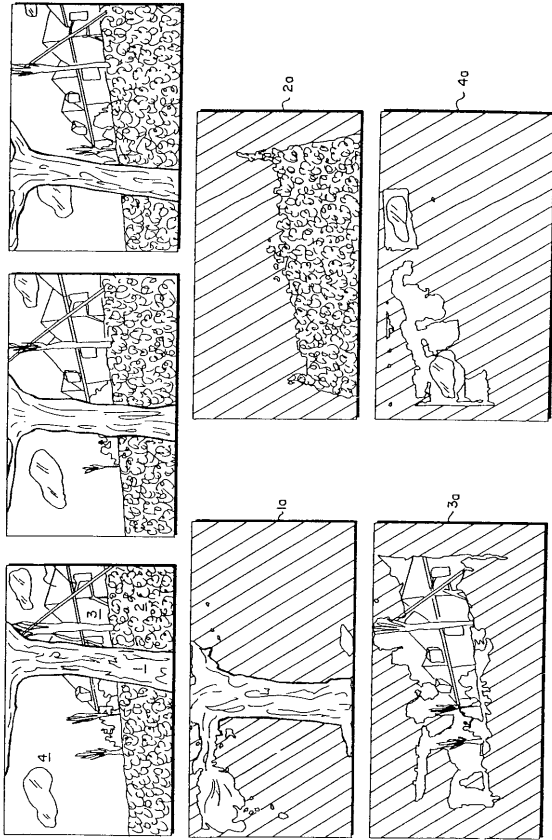
30

【符号の説明】

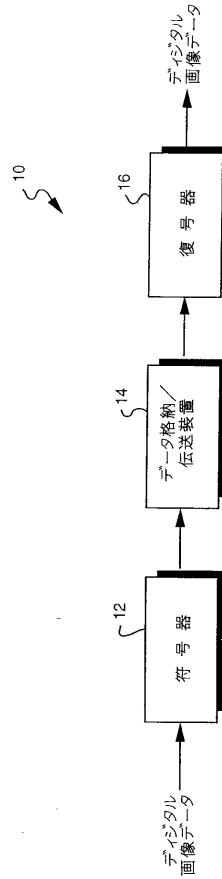
- 1 0 符号化 / 復号化システム
- 1 2 符号器
- 1 4 データ格納 / 伝送装置
- 1 6 復号器
- 2 2 局所的動き推定部
- 2 4 動きセグメンテーション処理部
- 2 4 a コヒーレント動き領域推定部
- 2 4 b 動きモデル推定部
- 2 6 層抽出処理部
- 3 2 アフィン動きモデル推定部
- 3 4 クラスタ化処理部
- 3 6 仮説検定部

40

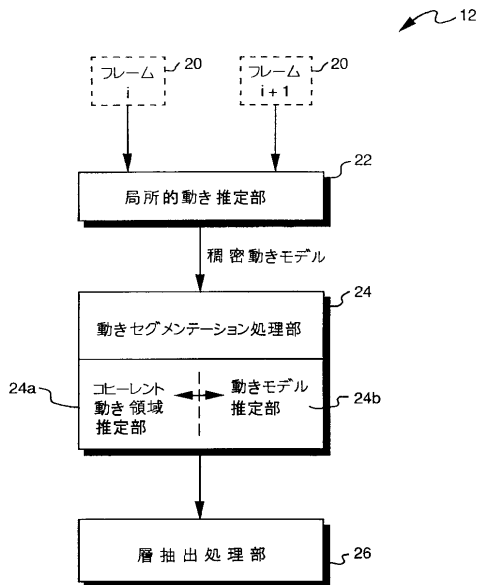
【図 1】



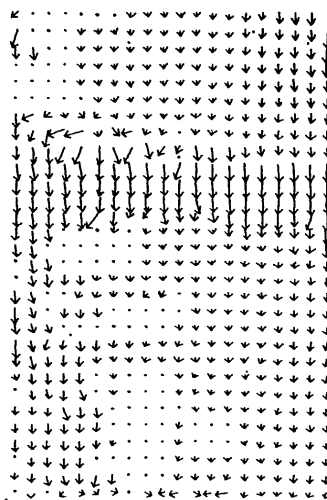
【図 2】



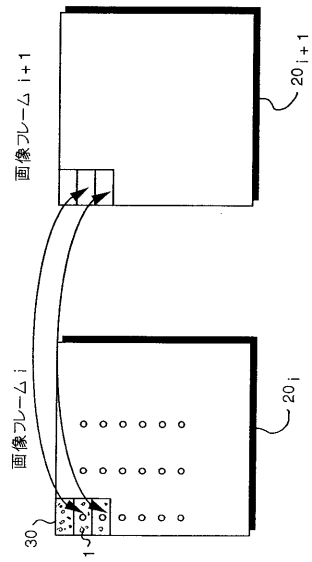
【図 3】



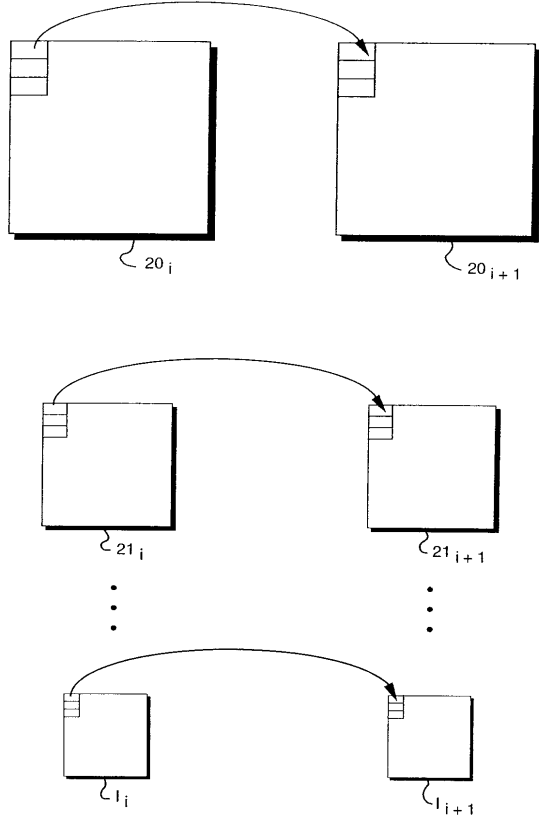
【図 4】



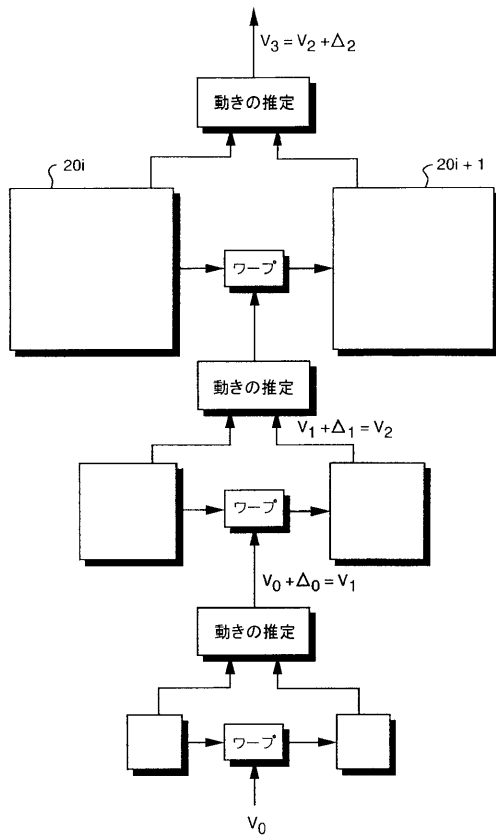
【図 5】



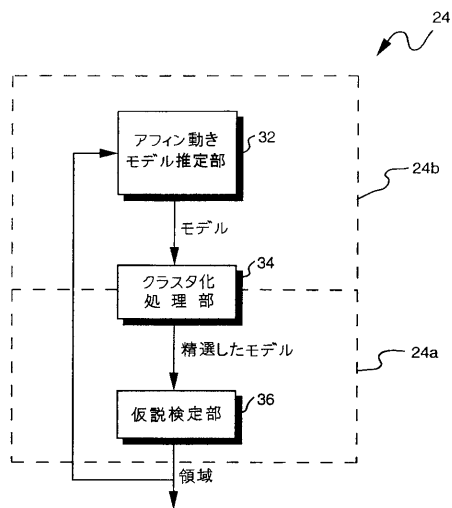
【図 6】



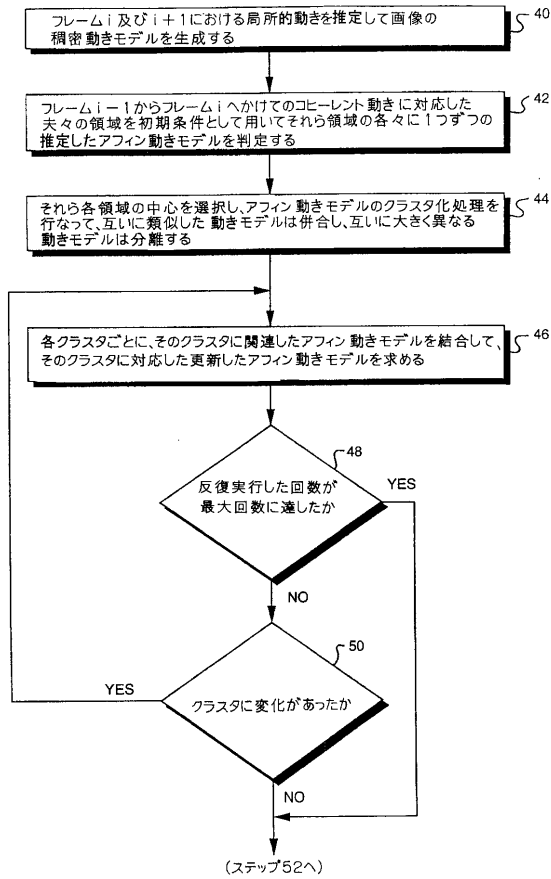
【図 7】



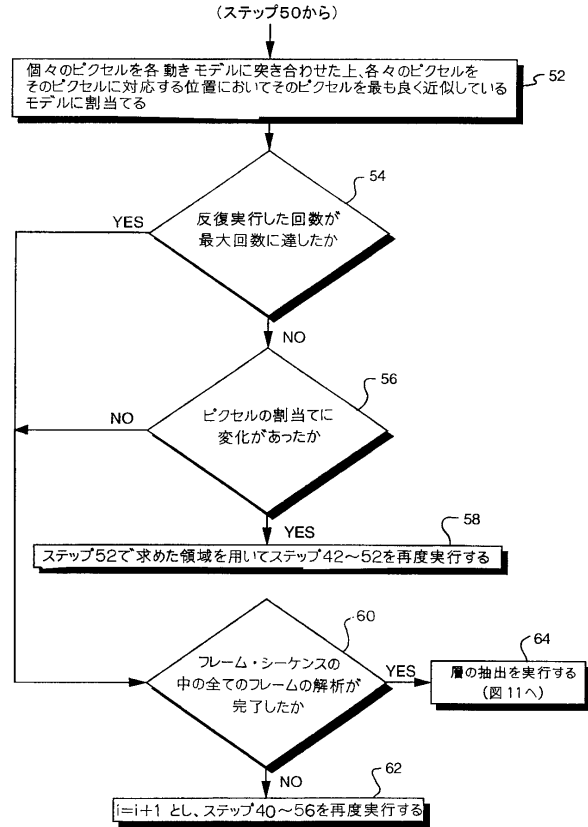
【図 8】



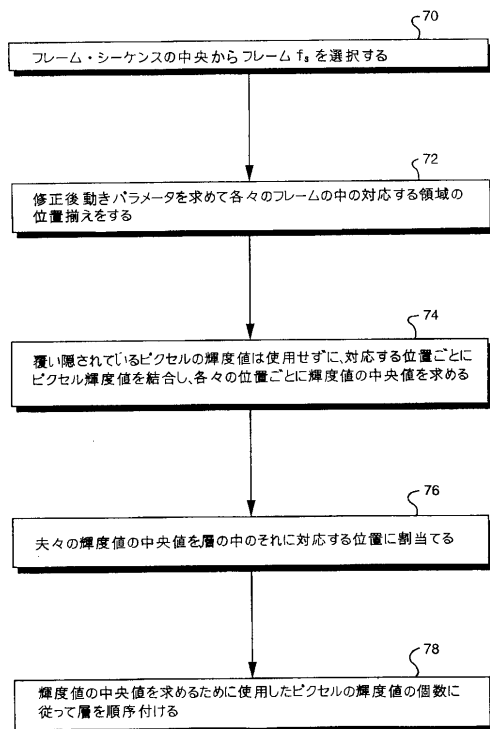
【図 9】



【図 10】



【図 11】



フロントページの続き

(74)代理人 100075236

弁理士 栗田 忠彦

(74)代理人 100075270

弁理士 小林 泰

(72)発明者 ジョン・ワイ・エイ・ウァン

アメリカ合衆国マサチューセッツ州02138, ケンブリッジ, マサチューセッツ・アベニュー
1010, アpartment 4

(72)発明者 エドワード・エイチ・アデルソン

アメリカ合衆国マサチューセッツ州02139, ケンブリッジ, フランクリン・ストリート 55
1, アpartment 2

審査官 菅原 道晴

(56)参考文献 特開平04-148283(JP, A)

(58)調査した分野(Int.Cl.⁷, DB名)

H04N 7/26-7/68

G06T 7/00-9/40