

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-174690

(P2007-174690A)

(43) 公開日 平成19年7月5日(2007.7.5)

(51) Int. Cl.	F I			テーマコード (参考)	
HO4N 7/26 (2006.01)	HO4N	7/13	Z	5C059	
HO4N 7/01 (2006.01)	HO4N	7/01	Z	5C063	

審査請求 有 請求項の数 26 O L (全 30 頁)

(21) 出願番号	特願2007-17136 (P2007-17136)	(71) 出願人	396001980
(22) 出願日	平成19年1月26日 (2007.1.26)		株式会社モノリス
(62) 分割の表示	特願2001-331544 (P2001-331544) の分割	(74) 代理人	100105924 弁理士 森下 賢樹
原出願日	平成13年10月29日 (2001.10.29)	(72) 発明者	秋吉 仰三 東京都港区麻布十番1-7-3 株式会社モノリス内
		(72) 発明者	秋吉 信雄 東京都港区麻布十番1-7-3 株式会社モノリス内
		Fターム(参考)	5C059 KK02 LB16 MA00 UA02 UA05 5C063 BA08 BA12 BA20 CA05 CA07

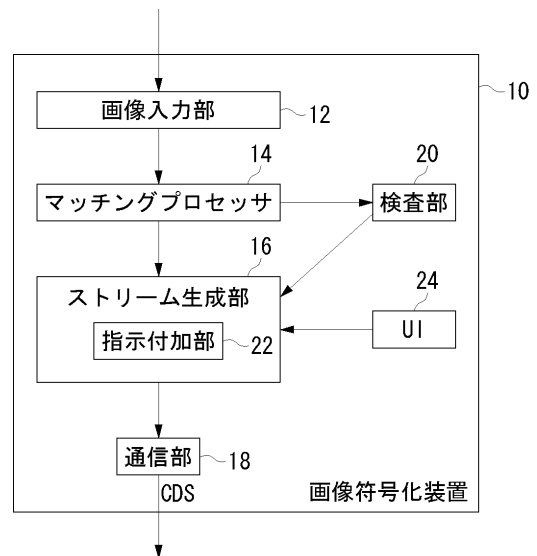
(54) 【発明の名称】 画像符号化方法、画像復号方法、画像符号化装置および画像復号装置

(57) 【要約】

【課題】 補間によって動画を生成する際、中間フレームの画質が悪い場合がある。

【解決手段】 画像入力部12はキーフレームのデータ入力し、マッチングプロセッサ14がそれらの間の対応点情報を計算する。検査部20はキーフレーム間の変化量が大きいときこれを検出し、ストリーム生成部16へ伝える。ストリーム生成部16がキーフレームと対応点情報をもとに符号化データストリームを生成する際、検査部20から前述の通知があれば、復号時に中間フレームを生成する際、故意的にその画質を低下させる指示を付加する。

【選択図】 図18



**【特許請求の範囲】****【請求項 1】**

キーフレームのデータを取得する工程と、  
キーフレーム間の対応点情報を取得する工程と、  
前記対応点情報とキーフレームのデータをもとに、別のフレームを生成する工程と、  
を含み、  
別のフレームを生成する工程において、故意的な画質低下処理を施しつつ別のフレームを生成することを特徴とする画像復号方法。

**【請求項 2】**

前記画質低下処理に関する指示を検出する工程をさらに含み、その指示が検出されたときに限り、前記画質低下処理を施して別のフレームが生成される請求項 1 に記載の方法。 10

**【請求項 3】**

前記対応点情報を検査する工程をさらに含み、キーフレーム間の対応点の移動量が大い場合に限り、前記画質低下処理を施して別のフレームが生成される請求項 1 に記載の方法。

**【請求項 4】**

ユーザの要求を受け付ける工程をさらに含み、要求が出された場合に限り、前記画質低下処理を施して別のフレームが生成される請求項 1 に記載の方法。

**【請求項 5】**

前記画質低下処理は、別のフレームにモーションブラー効果を与える処理である請求項 1 ~ 4 のいずれかに記載の方法。 20

**【請求項 6】**

前記モーションブラー効果は、キーフレーム間に生成される複数の別のフレームをブレンドして得られる請求項 5 に記載の方法。

**【請求項 7】**

前記画質低下処理は、別のフレームの解像度を下げる処理である請求項 1 ~ 4 のいずれかに記載の方法。

**【請求項 8】**

前記画質低下処理は、別のフレームに画像フィルタをかける処理である請求項 1 ~ 4 のいずれかに記載の方法。 30

**【請求項 9】**

キーフレームのデータを取得する工程と、  
キーフレーム間の対応点情報を取得する工程と、  
取得されたキーフレームおよび対応点情報に、キーフレーム間の別のフレームに対する故意的な画質低下処理に関する指示を付加して符号化データを生成する工程と、  
を含むことを特徴とする画像符号化方法。

**【請求項 10】**

キーフレーム間の対応点の移動量が大い場合に限り、前記指示を付加する請求項 9 に記載の方法。

**【請求項 11】**

ユーザの要求を受け付ける工程をさらに含み、要求が出されたとき前記指示を付加する請求項 9 に記載の方法。 40

**【請求項 12】**

キーフレームのデータとキーフレーム間の対応点情報を取得する画像取得部と、  
前記対応点情報とキーフレームのデータをもとに、別のフレームを生成する生成部と、  
を含み、  
前記生成部は、故意的な画質低下処理を施しつつ別のフレームを生成することを特徴とする画像復号装置。

**【請求項 13】**

ユーザの要求を受け付ける UI 部をさらに含み、ユーザの要求が出されたとき、前記画 50

質低下処理が施される請求項 1 2 に記載の装置。

【請求項 1 4】

生成される複数の別のフレームのブレンド処理部をさらに含み、別のフレームのブレンドによって前記画質低下処理が実現される請求項 1 2 に記載の装置。

【請求項 1 5】

生成される別のフレームの解像度低減部をさらに含み、別のフレームの解像度を下げて前記画質低下処理が実現される請求項 1 2 に記載の装置。

【請求項 1 6】

前記解像度低減部は、別のフレームの解像度を下げたのち、その表示画素数がもとの別のフレームと等しくなるよう画像拡大処理を施す請求項 1 5 に記載の装置。

10

【請求項 1 7】

生成される別のフレームのフィルタをさらに含み、別のフレームのフィルタリングによって前記画質低下処理が実現される請求項 1 2 に記載の装置。

【請求項 1 8】

キーフレームのデータとキーフレーム間の対応点情報を取得する画像取得部と、  
取得されたキーフレームおよび対応点情報に、キーフレーム間の別のフレームに対する故意的な画質低下処理に関する指示を付加して符号化データを生成するストリーム生成部と、  
を含むことを特徴とする画像符号化装置。

【請求項 1 9】

前記対応点情報を生成するマッチングプロセッサをさらに含む請求項 1 8 に記載の装置。

20

【請求項 2 0】

前記対応点情報を検査する検査部をさらに含み、キーフレーム間の対応点の移動量が大い場合に限り、前記指示が付加される請求項 1 8 に記載の装置。

【請求項 2 1】

ユーザの要求を受け付ける UI 部をさらに含み、ユーザの要求が出されたとき、前記指示が付加される請求項 1 8 に記載の装置。

【請求項 2 2】

キーフレームのデータを取得する工程と、  
キーフレーム間の対応点情報を取得する工程と、  
前記対応点情報とキーフレームのデータをもとに、別のフレームを生成する工程と、  
をコンピュータに実行せしめるコンピュータプログラムであって、  
別のフレームを生成する工程において、故意的な画質低下処理を施しつつ別のフレームを生成することを特徴とするコンピュータプログラム。

30

【請求項 2 3】

キーフレームのデータを取得する工程と、  
キーフレーム間の対応点情報を取得する工程と、  
取得されたキーフレームおよび対応点情報に、キーフレーム間の別のフレームに対する故意的な画質低下処理に関する指示を付加して符号化データを生成する工程と、  
をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

40

【請求項 2 4】

前記キーフレーム間の変化の程度を検査する工程をさらに含み、その変化が大い場合に限り、前記画質低下処理を施して別のフレームが生成される請求項 1 に記載の方法。

【請求項 2 5】

前記キーフレーム間の変化が大い場合に限り、前記指示を付加する請求項 9 に記載の方法。

【請求項 2 6】

前記キーフレーム間の変化の程度を検査する検査部をさらに含み、キーフレーム間の変化が大い場合に限り、前記指示が付加される請求項 1 8 に記載の装置。

50

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

この発明は、画像処理技術に関し、とくにふたつの画像をマッチング技術をもとに補間する符号化と復号技術に関する。

## 【背景技術】

## 【0002】

動画圧縮の事実上の世界標準であるMPEG(Motion Picture Expert Group)は、CDなどストレージメディアからネットワークや放送などの伝送メディアへとそのターゲットエリアが広がった。放送のデジタル化はMPEGを中心とする圧縮符号化技術なしに考

10

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0003】

そうした混沌の中でも、動画の圧縮技術に関する方向性について確かなことがある。すなわち、より高い圧縮率と画質の両立である。MPEGは周知のごとく高圧縮化をブロック歪みが阻むことがある。また、動画の配信が現実的になるにしたがい、リアルタイム処理に対するいままで以上の配慮が必要になる。

20

## 【0004】

本発明はこうした現状に鑑みてなされたものであり、その目的は、画像データの効率的圧縮を実現する符号化および復号技術の提供にある。別の目的は、画質の維持と圧縮率の向上という相反する要望に適う画像符号化および復号技術の提供にある。さらに別の目的は、キーフレームから補間によって中間フレームを生成するアプリケーションにおいて、中間フレームの見た目の画質改善を図ることにある。

## 【課題を解決するための手段】

## 【0005】

本発明は画像符号化および復号技術に関する。この技術は本出願人が先に特許第2927350号にて提案した画像マッチング技術(以下「前提技術」という)を利用すること

30

## 【0006】

本発明のある態様は、画像復号方法に関する。この方法は、キーフレームのデータを取得する工程と、キーフレーム間の対応点情報を取得する工程と、前記対応点情報とキーフレームのデータをもとに、故意的な画質低下処理を施しつつ中間フレームを生成する工程とを含む。

## 【0007】

キーフレームはネットワーク経由、記憶装置からの読出などのほか、自らカメラ撮影して取得してもよい。対応点情報も外部から取得するほか、自らキーフレーム間のマッチングを計算して取得してもよい。

40

## 【0008】

「画質低下処理」は、中間フレームの解像度を下げたり、中間フレームを所定のフィルタ、たとえローパスフィルタに通したり、生成された複数の中間フレームにモーションブラー効果を加えることなどで実現できる。モーションブラーとは、被写体の動きによって発生する画像上の「ぶれ」をいう。モーションブラー効果自体は画像処理上既知であり、例えば特開平11-283041号にも記載されている。

## 【0009】

本発明では画質低下を中間フレームの生成に当たって故意的に発生せしめるものである。たとえば、キーフレーム間にオクルージョン、すなわち被写体の隠蔽がある場合やキーフレーム間の差が大きい場合、対応点情報の誤差が大きくなり、または対応点情報では表

50

現の限界があり、中間フレームの画質が低下することがある。そうしたとき、画像のシャープさを捨てて、いくぶん曖昧さのある画像を生成したほうが、むしろ人間の目に自然なことがある。本発明はそうした本発明者の知見に基づく。したがって、本発明でいう「画質低下」は、画像処理上客観的な画質低下をいい、それが主観画質としてはむしろ画質改善に結びつく点に本発明の特徴がある。

#### 【0010】

画像低下処理に関する指示を検出する工程をさらに含み、その指示が検出されたときに限り、当該処理を加えて中間フレームが生成されてもよい。対応点情報を検査する工程をさらに含み、キーフレーム間の対応点の移動量が大きい場合に限り、画像低下処理を施して中間フレームが生成されてもよい。また、ユーザの要求を受け付ける工程をさらに含み、要求が出された場合に限り、画像低下処理を施して中間フレームが生成されてもよい。

10

#### 【0011】

本発明の別の態様は、画像符号化方法である。この方法は、キーフレームのデータを取得する工程と、キーフレーム間の対応点情報を取得する工程と、取得されたキーフレームおよび対応点情報に、故意的な画質低下処理に関する指示を付加して符号化データを生成する工程とを含む。符号化に当たり、キーフレーム間の対応点の移動量が大きい場合に限り、指示を付加してもよいし、ユーザの要求が出されたとき付加してもよい。

#### 【0012】

以上の態様のうち、キーフレーム間の対応点情報を生成する符号化方法、およびそれを利用して中間画像を生成する復号方法は、前提技術を利用することができる。ただし、本発明は前提技術を必須とはしない。また、以上の各構成、工程を任意に入れ替えたり、方法と装置の間で表現を一部または全部入れ替え、または追加したり、表現をコンピュータプログラム、記録媒体等に変更したのもまた、本発明として有効である。

20

#### 【発明を実施するための最良の形態】

#### 【0013】

はじめに、実施の形態で利用する多重解像度特異点フィルタ技術とそれを用いた画像マッチング処理を「前提技術」として詳述する。これらの技術は本出願人がすでに特許第2927350号を得ている技術であり、本発明との組合せに最適である。ただし、実施の形態で採用可能な画像マッチング技術はこれに限られない。図18以降、前提技術を利用した画像処理技術を具体的に説明する。

30

#### 【0014】

##### [前提技術の背景]

最初に[1]で前提技術の要素技術の詳述し、[2]で処理手順を具体的に説明する。さらに[3]で実験の結果を報告する。

##### [1]要素技術の詳細

##### [1.1]イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから精細なレベルへと順に解像度の階層を辿っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

40

#### 【0015】

本前提技術は、例えば完全に自動的なモーフィング、物体認識、立体写真測量、ボリュームレンダリング、少ないフレームからの滑らかな動画像の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ボリュームレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

#### 【0016】

##### [1.2]特異点フィルタの階層

50

前提技術に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の輝度及び位置を保存することができる。ここで画像の幅をN、高さをMとする。以下簡単のため、 $N = M = 2^n$  (nは自然数)と仮定する。また、区間 $[0, N]$  RをIと記述する。(i, j)における画像の画素を $p(i, j)$ と記述する( $i, j \in I$ )。

【0017】

ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第mレベルにおける各画像のサイズは $2^m \times 2^m$  ( $0 \leq m \leq n$ )とする。特異点フィルタは次の4種類の新たな階層画像をnから下がる方向で再帰的に構築する。

【数1】

$$p_{(i,j)}^{(m,0)} = \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)}))$$

$$p_{(i,j)}^{(m,1)} = \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)}))$$

$$p_{(i,j)}^{(m,2)} = \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)}))$$

$$p_{(i,j)}^{(m,3)} = \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)}))$$

(式1)

ただしここで、

【数2】

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式2})$$

とする。以降これら4つの画像を副画像(サブイメージ)と呼ぶ。 $\min_{x, y} p_{(x, y)}$ 、 $\max_{x, y} p_{(x, y)}$ をそれぞれ及びと記述すると、副画像はそれぞれ以下のように記述できる。

【0018】

$$p_{(x, y)}^{(m, 0)} = \min_{x, y} p_{(x, y)}^{(m+1, 0)}$$

$$p_{(x, y)}^{(m, 1)} = \max_{x, y} p_{(x, y)}^{(m+1, 1)}$$

$$p_{(x, y)}^{(m, 2)} = \min_{x, y} p_{(x, y)}^{(m+1, 2)}$$

$$p_{(x, y)}^{(m, 3)} = \max_{x, y} p_{(x, y)}^{(m+1, 3)}$$

すなわち、これらはとのテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明らかなように、特異点フィルタはもとの画像について $2 \times 2$ 画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、前提技術では輝度を採用するが、画像に関するいろいろな数値を採用することができる。ふたつの方向の両方について最大画素値となる画素は極大点、ふたつの方向の両方について最小画素値となる画素は極小点、ふたつの方向の一方について最大画素値となるとともに、他方について最小画素値となる画素は鞍点として検出される。

【0019】

特異点フィルタは、各ブロックの内部で検出された特異点の画像(ここでは1画素)でそのブロックの画像(ここでは4画素)を代表させることにより、画像の解像度を落とす。特異点の理論的な観点からすれば、 $\min_{x, y} p_{(x, y)}$ は極小点を保存し、 $\max_{x, y} p_{(x, y)}$ は極大点を保存し、 $\min_{x, y} p_{(x, y)}$ 及び $\max_{x, y} p_{(x, y)}$ は鞍点を保存する。

【0020】

はじめに、マッチングをとるべき始点(ソース)画像と終点(デスティネーション)画

10

20

30

40

50

像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

#### 【0021】

この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとれられていく。まず  $p^{(m, 0)}$  を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p^{(m, 1)}$  を用いて鞍点のマッチングがとられ、 $p^{(m, 2)}$  を用いて他の鞍点のマッチングがとられる。そして最後に  $p^{(m, 3)}$  を用いて極大点のマッチングがとられる。

#### 【0022】

図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像  $p^{(5, 0)}$  を示している。同様に、図1(e)と図1(f)は  $p^{(5, 1)}$ 、図1(g)と図1(h)は  $p^{(5, 2)}$ 、図1(i)と図1(j)は  $p^{(5, 3)}$  をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず  $p^{(5, 0)}$  によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p^{(5, 1)}$  によれば口が明確になる。口は横方向で輝度が低いためである。 $p^{(5, 2)}$  によれば首の両側の縦線が明確になる。最後に、 $p^{(5, 3)}$  によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

#### 【0023】

特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影された画像の特徴と、予め記録しておいたいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

#### 【0024】

##### [1.3] 画像間の写像の計算

始点画像の位置  $(i, j)$  の画素を  $p^{(n)}(i, j)$  と書き、同じく終点画像の位置  $(k, l)$  の画素を  $q^{(n)}(k, l)$  で記述する。 $i, j, k, l \in I$  とする。画像間の写像のエネルギー(後述)を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ  $p^{(m, 0)}$  と  $q^{(m, 0)}$  間の写像  $f^{(m, 0)} : p^{(m, 0)} \rightarrow q^{(m, 0)}$  が計算される。 $f^{(m, 0)}$  に基づき、最小エネルギーを持つ  $p^{(m, 1)}$ 、 $q^{(m, 1)}$  間の写像  $f^{(m, 1)}$  が計算される。この手続は、 $p^{(m, 3)}$  と  $q^{(m, 3)}$  の間の写像  $f^{(m, 3)}$  の計算が終了するまで続く。各写像  $f^{(m, i)}$  ( $i = 0, 1, 2, \dots$ ) を副写像と呼ぶことにする。 $f^{(m, i)}$  の計算の都合のために、 $i$  の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

#### 【0025】

##### 【数3】

$$f^{(m, i)} : p^{(m, \sigma(i))} \rightarrow q^{(m, \sigma(i))} \quad (\text{式3})$$

ここで  $(i) \in \{0, 1, 2, 3\}$  である。

#### 【0026】

##### [1.3.1] 全単射

始点画像と終点画像間のマッピングを写像で表現する場合、その写像は両画像間で全単射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射かつ単射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全単射のデジタル版である。前提技術では、画素は格子点によって特定される。

#### 【0027】

始点副画像(始点画像について設けられた副画像)から終点副画像(終点画像について設けられた副画像)への写像は、 $f^{(m, s)} : I/2^{n-m} \times I/2^{n-m} \rightarrow I/2^{n-m} \times I/2^{n-m}$  ( $s = 0, 1, \dots$ ) によって表される。ここで、 $f^{(m, s)}(i, j) = (k, l)$  は、始点画像の  $p^{(m, s)}(i, j)$  が終点画像の  $q^{(m, s)}(k, l)$

10

20

30

40

50

$(i, j)$  に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$  が成り立つとき画素  $q(k, l)$  を  $q_{f(i, j)}$  と記述する。

【0028】

前提技術で扱う画素（格子点）のようにデータが離散的な場合、全単射の定義は重要である。ここでは以下のように定義する（ $i, i', j, j', k, l$  は全て整数とする）。まず始めに、始点画像の平面において  $R$  によって表記される各正方形領域、

【数4】

$$P_{(i,j)}^{(m,s)} P_{(i+1,j)}^{(m,s)} P_{(i+1,j+1)}^{(m,s)} P_{(i,j+1)}^{(m,s)}$$

(式4)

を考える（ $i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$ ）。ここで  $R$  の各辺（エッジ）の方向を以下のように定める。

【数5】

$$\overrightarrow{P_{(i,j)}^{(m,s)} P_{(i+1,j)}^{(m,s)}}, \overrightarrow{P_{(i+1,j)}^{(m,s)} P_{(i+1,j+1)}^{(m,s)}}, \overrightarrow{P_{(i+1,j+1)}^{(m,s)} P_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{P_{(i,j+1)}^{(m,s)} P_{(i,j)}^{(m,s)}}$$

(式5)

この正方形は写像  $f$  によって終点画像平面における四辺形に写像されなければならない。 $f^{(m,s)}(R)$  によって示される四辺形、

【数6】

$$Q_{f(i,j)}^{(m,s)} Q_{f(i+1,j)}^{(m,s)} Q_{f(i+1,j+1)}^{(m,s)} Q_{f(i,j+1)}^{(m,s)}$$

(式6)

は、以下の全単射条件を満たす必要がある。

【0029】

1. 四辺形  $f^{(m,s)}(R)$  のエッジは互いに交差しない。
2.  $f^{(m,s)}(R)$  のエッジの方向は  $R$  のそれらに等しい（図2の場合、時計回り）。
3. 緩和条件として収縮写像（リトラクション：retractions）を許す。

【0030】

何らかの緩和条件を設けないかぎり、全単射条件を完全に満たす写像は単位写像しかないためである。ここでは  $f^{(m,s)}(R)$  のひとつのエッジの長さが0、すなわち  $f^{(m,s)}(R)$  は三角形になってもよい。しかし、面積が0となるような図形、すなわち1点または1本の線分になってはならない。図2(R)がもとの四辺形の場合、図2(A)と図2(D)は全単射条件を満たすが、図2(B)、図2(C)、図2(E)は満たさない。

【0031】

実際のインプリメンテーションでは、写像が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというものである。すなわち、 $f(i, j) = (i, j)$ （ただし  $i = 0, i = 2^m - 1, j = 0, j = 2^m - 1$  の4本の線上）である。この条件を以下「付加条件」とも呼ぶ。

【0032】

[1.3.2] 写像のエネルギー

[1.3.2.1] 画素の輝度に関するコスト

写像  $f$  のエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像  $f^{(m,s)}$  の点  $(i, j)$  におけるエネルギー  $C^{(m,s)}(i, j)$  は次式によって定まる。

10

20

30

40

【数 7】

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_{f(i,j)}^{(m,s)})|^2 \quad (\text{式 7})$$

ここで、 $V(p_{(i,j)}^{(m,s)})$  及び  $V(q_{f(i,j)}^{(m,s)})$  はそれぞれ画素  $p_{(i,j)}^{(m,s)}$  及び  $q_{f(i,j)}^{(m,s)}$  の輝度である。 $f$  のトータルのエネルギー  $C^{(m,s)}$  は、マッチングを評価するひとつの評価式であり、つぎに示す  $C_{(i,j)}^{(m,s)}$  の合計で定義できる。

【数 8】

$$C_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} C_{(i,j)}^{(m,s)} \quad (\text{式 8})$$

10

[ 1 . 3 . 2 . 2 ] 滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギー  $D_f$  を導入する。このエネルギーは画素の輝度とは関係なく、 $p_{(i,j)}^{(m,s)}$  および  $q_{f(i,j)}^{(m,s)}$  の位置によって決まる ( $i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$ )。点  $(i, j)$  における写像  $f^{(m,s)}$  のエネルギー  $D_{(i,j)}^{(m,s)}$  は次式で定義される。

【数 9】

$$D_{(i,j)}^{(m,s)} = \eta E_{0(i,j)}^{(m,s)} + E_{1(i,j)}^{(m,s)} \quad (\text{式 9})$$

20

ただし、係数パラメータ  $\eta$  は 0 以上の実数であり、また、

【数 10】

$$E_{0(i,j)}^{(m,s)} = \|(i, j) - f^{(m,s)}(i, j)\|^2 \quad (\text{式 10})$$

【数 11】

$$E_{1(i,j)}^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i, j) - (i, j)) - (f^{(m,s)}(i', j') - (i', j'))\|^2 / 4 \quad (\text{式 11})$$

30

とする。ここで、

【数 12】

$$\|(x, y)\| = \sqrt{x^2 + y^2} \quad (\text{式 12})$$

であり、 $i' < 0$  および  $j' < 0$  に対して  $f(i', j')$  は 0 と決める。 $E_0$  は  $(i, j)$  及び  $f(i, j)$  の距離で決まる。 $E_0$  は画素があまりにも離れた画素へ写影されることを防ぐ。ただし  $E_0$  は、後に別のエネルギー関数で置き換える。 $E_1$  は写像の滑らかさを保証する。 $E_1$  は、 $p_{(i,j)}$  の変位とその隣接点の変位間の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー  $D_f$  は次式で定まる。

40

【数 13】

$$D_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} D_{(i,j)}^{(m,s)} \quad (\text{式 13})$$

[ 1 . 3 . 2 . 3 ] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は  $C_f^{(m,s)} + D_f^{(m,s)}$  で定義される。ここで係数パラメータ  $\eta$  は 0 以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次式で示す最小エネルギーを与える写像を見いだすことである。

【数 1 4】

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \tag{式 1 4}$$

= 0 及び = 0 の場合、写像は単位写像になることに注意すべきである（すなわち、全ての  $i = 0, \dots, 2^m - 1$  及び  $j = 0, \dots, 2^m - 1$  に対して  $f^{(m,s)}(i, j) = (i, j)$  となる）。後述のごとく、本前提技術では最初に = 0 及び = 0 の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の位置を変えて  $C^{(m,s)}_f + D^{(m,s)}_f$  と定義したとすれば、 = 0 及び = 0 の場合に総合評価式が  $C^{(m,s)}_f$  だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけで対応づけられ、写像が無意味なものになる。そうした無意味な写像をもとに写像を変形していてもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

10

【0033】

オブティカルフローもこの前提技術同様、画素の輝度の差と滑らかさを考慮する。しかし、オブティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。前提技術に係る特異点フィルタを用いることによって大域的な対応関係を検出することができる。

20

【0034】

[ 1 . 3 . 3 ] 多重解像度の導入による写像の決定

最小エネルギーを与え、全単射条件を満足する写像  $f_{min}$  を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0035】

まず、

30

【数 1 5】

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \tag{式 1 5}$$

が成り立つとき、  $p^{(m-1,s)}(i', j')$ 、  $q^{(m-1,s)}(i', j')$  をそれぞれ  $p^{(m,s)}(i, j)$ 、  $q^{(m,s)}(i, j)$  の parent と呼ぶことにする。  $[x]$  は  $x$  を越えない最大整数である。また  $p^{(m,s)}(i, j)$ 、  $q^{(m,s)}(i, j)$  をそれぞれ  $p^{(m-1,s)}(i', j')$ 、  $q^{(m-1,s)}(i', j')$  の child と呼ぶ。関数  $parent(i, j)$  は次式で定義される。

【数 1 6】

$$parent(i, j) = ([\frac{i}{2}], [\frac{j}{2}]) \tag{式 1 6}$$

40

$p^{(m,s)}(i, j)$  と  $q^{(m,s)}(k, l)$  の間の写像  $f^{(m,s)}$  は、エネルギー計算を行って最小になったものを見つけることで決定される。  $f^{(m,s)}(i, j) = (k, l)$  の値は  $f^{(m-1,s)}(m = 1, 2, \dots, n)$  を用いることによって、以下のように決定される。まず、  $q^{(m,s)}(k, l)$  は次の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

。

【数 17】

$$q_{g^{(m,s)}(i-1,j-1)}^{(m,s)} q_{g^{(m,s)}(i-1,j+1)}^{(m,s)} q_{g^{(m,s)}(i+1,j+1)}^{(m,s)} q_{g^{(m,s)}(i+1,j-1)}^{(m,s)}$$

(式 17)

ただしここで、

【数 18】

$$g^{(m,s)}(i,j) = f^{(m-1,s)}(\text{parent}(i,j)) + f^{(m-1,s)}(\text{parent}(i,j) + (1,1))$$

(式 18)

10

である。こうして定めた四辺形を、以下  $p^{(m,s)}(i,j)$  の相続 (inherited) 四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0036】

図 3 は以上の手順を示している。同図において、始点画像の A, B, C, D の画素は、第  $m-1$  レベルにおいてそれぞれ終点画像の A', B', C', D' へ写影される。画素  $p^{(m,s)}(i,j)$  は、相続四辺形 A' B' C' D' の内部に存在する画素  $q^{(m,s)}(i,j)$  へ写影されなければならない。以上の配慮により、第  $m-1$  レベルの写像から第  $m$  レベルの写像への橋渡しがなされる。

20

【0037】

先に定義したエネルギー  $E_0$  は、第  $m$  レベルにおける副写像  $f^{(m,0)}$  を計算するために、次式に置き換える。

【数 19】

$$E_{0(i,j)} = \|f^{(m,0)}(i,j) - g^{(m)}(i,j)\|^2 \quad (\text{式 19})$$

また、副写像  $f^{(m,s)}$  を計算するためには次式を用いる。

【数 20】

$$E_{0(i,j)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2 \quad (1 \leq i) \quad (\text{式 20})$$

30

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式 20 により、異なる特異点に対応する副写像が、副写像どうしの類似度が高くなるように同一レベル内で関連づけられる。式 19 は、 $f^{(m,s)}(i,j)$  と、第  $m-1$  レベルの画素の一部と考えた場合の  $(i,j)$  が射影されるべき点の位置との距離を示している。

【0038】

仮に、相続四辺形 A' B' C' D' の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、A' B' C' D' の境界線からの距離が  $L$  (始めは  $L=1$ ) である画素を調べる。それらのうち、エネルギーが最小になるものが全単射条件を満たせば、これを  $f^{(m,s)}(i,j)$  の値として選択する。そのような点が発見されるか、または  $L$  がその上限の  $L^{(m)}_{\max}$  に到達するまで、 $L$  を大きくしていく。 $L^{(m)}_{\max}$  は各レベル  $m$  に対して固定である。そのような点が全く発見されない場合、全単射の第 3 の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f^{(m,s)}(i,j)$  を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の第 1 及び第 2 条件を外す。

40

【0039】

多重解像度を用いる近似法は、写像が画像の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことがで

50

きない。さらに、通常写像に滑らかさを要求するため、そうした画素間の対応関係を見つけていくとしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、そうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル（粗いレベル）において小さいためである。

#### 【0040】

##### [1.4] 最適なパラメータ値の自動決定

既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。前提技術に係る方法によれば、最適なパラメータ値を完全に自動決定することができる。

10

#### 【0041】

前提技術に係るシステムはふたつのパラメータ、 $\alpha$  及び  $\beta$  を含む。端的にいえば、 $\alpha$  は画素の輝度の差の重みであり、 $\beta$  は写像の剛性を示している。これらのパラメータの値は初期値が0であり、まず  $\beta = 0$  に固定して  $\alpha$  を0から徐々に増加させる。 $\alpha$  の値を大きくしながら、しかも総合評価式（式14）の値を最小にする場合、各副写像に関する  $C^{(m, s)}_f$  の値は一般に小さくなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、 $\alpha$  が最適値を超えると以下の現象が発生する。

#### 【0042】

20

1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。

2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる。

#### 【0043】

3. その結果、式14において  $D^{(m, s)}_f$  が急激に増加しようとする。

4. その結果、式14の値が急激に増加しようとするため、 $D^{(m, s)}_f$  の急激な増加を抑制するよう  $f^{(m, s)}$  が変化し、その結果  $C^{(m, s)}_f$  が増加する。

したがって、 $\alpha$  を増加させながら式14が最小値をとるという状態を維持しつつ  $C^{(m, s)}_f$  が減少から増加に転じる閾値を検出し、その  $\alpha$  を  $\beta = 0$  における最適値とする。つぎに  $\beta$  を少しずつ増やして  $C^{(m, s)}_f$  の挙動を検査し、後述の方法で  $\alpha$  を自動決定する。その  $\beta$  に対応して  $\alpha$  も決まる。

30

#### 【0044】

この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できるとき、その目が固定される。

#### 【0045】

##### [1.4.1] $\alpha$ の動的決定

$\alpha$  は0から所定の刻み幅で増加されていき、 $\alpha$  の値が変わる度に副写像が評価される。式14のごとく、総エネルギーは  $C^{(m, s)}_f + D^{(m, s)}_f$  によって定義される。式9の  $D^{(m, s)}_f$  は滑らかさを表すもので、理論的には単位写像の場合に最小になり、写像が歪むほど  $E_0$  も  $E_1$  も増加していく。 $E_1$  は整数であるから、 $D^{(m, s)}_f$  の最小刻み幅は1である。このため、現在の  $C^{(m, s)}_{(i, j)}$  の変化（減少量）が1以上でなければ、写像を変化させることによって総エネルギーを減らすことはできない。なぜなら、写像の変化に伴って  $D^{(m, s)}_f$  は1以上増加するため、 $C^{(m, s)}_{(i, j)}$  が1以上減少しない限り総エネルギーは減らないためである。

40

#### 【0046】

この条件のもと、 $\alpha$  の増加に伴い、正常な場合に  $C^{(m, s)}_{(i, j)}$  が減少することを示す。 $C^{(m, s)}_{(i, j)}$  のヒストグラムを  $h(l)$  と記述する。 $h(l)$  はエネルギー  $C^{(m, s)}_{(i, j)}$  が  $l^2$  である画素の数である。 $l^2 - 1$  が成り立つために、例えば  $l^2 = 1$  の場合を考える。 $\alpha$  が  $\alpha_1$  から  $\alpha_2$  まで微小量変化するとき、

50

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式 2 1)

で示される A 個の画素が、

$$C_f^{(m,s)} - l^2 = C_f^{(m,s)} - \frac{1}{\lambda} \quad (\text{式 2 2})$$

10

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素のエネルギーがすべてゼロになると近似している。この式は  $C^{(m,s)}_f$  の値が、

$$\partial C_f^{(m,s)} = -\frac{A}{\lambda} \quad (\text{式 2 3})$$

だけ変化することを示し、その結果、

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (\text{式 2 4})$$

20

が成立する。  $h(l) > 0$  であるから、通常  $C^{(m,s)}_f$  は減少する。しかし、  $C_f$  が最適値を越えようとするとき、上述の現象、つまり  $C^{(m,s)}_f$  の増加が発生する。この現象を検出することにより、  $C_f$  の最適値を決定する。

【 0 0 4 7 】

なお、  $H (h > 0)$  及び  $k$  を定数とすると、

【 数 2 5 】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (\text{式 2 5})$$

30

と仮定すれば、

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (\text{式 2 6})$$

が成り立つ。このとき  $k > 3$  であれば、

【 数 2 7 】

$$C_f^{(m,s)} = C + \frac{H}{(3/2 + k/2)\lambda^{3/2+k/2}} \quad (\text{式 2 7})$$

40

となる。これが  $C^{(m,s)}_f$  の一般式である ( $C$  は定数)。

【 0 0 4 8 】

の最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率を  $p_0$  と仮定する。この場合、

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式 2 8})$$

が成立しているため、全単射条件を破る画素の数は次式の率で増加する。

50

【数 29】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}}$$

(式 29)

従って、

【数 30】

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1$$

(式 30)

は定数である。仮に  $h(l) = H l^k$  を仮定するとき、例えば、

【数 31】

$$B_0 \lambda^{3/2+k/2} = p_0 H$$

(式 31)

は定数になる。しかし  $k$  が最適値を越えると、上の値は急速に増加する。この現象を検出し、 $B_0 \lambda^{3/2+k/2} / 2^m$  の値が異常値  $B_{0 \text{ t h r e s}}$  を越えるかどうかを検査し、 $k$  の最適値を決定することができる。同様に、 $B_1 \lambda^{3/2+k/2} / 2^m$  の値が異常値  $B_{1 \text{ t h r e s}}$  を越えるかどうかを検査することにより、全単射の第3の条件を破る画素の増加率  $B_1$  を確認する。ファクター  $2^m$  を導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギー  $C^{(m, s)}_f$  の観察では検出し損なった写像の過度の歪みを検出するために用いることができる。

【0049】

なお実験では、副写像  $f^{(m, s)}$  を計算する際、もし  $k$  が 0.1 を越えたら  $f^{(m, s)}$  の計算は止めて  $f^{(m, s+1)}$  の計算に移行した。  $k > 0.1$  のとき、画素の輝度 255 レベル中のわずかに「3」の違いが副写像の計算に影響したためであり、  $k > 0.1$  のとき正しい結果を得ることは困難だったためである。

【0050】

[1.4.2] ヒストグラム  $h(l)$

$C^{(m, s)}_f$  の検査はヒストグラム  $h(l)$  に依存しない。全単射及びその第3の条件の検査の際、 $h(l)$  に影響を受けうる。実際に  $(k, C^{(m, s)}_f)$  をプロットすると、 $k$  は通常 1 付近にある。実験では  $k = 1$  を用い、 $B_0 \lambda^2$  と  $B_1 \lambda^2$  を検査した。仮に  $k$  の本当の値が 1 未満であれば、 $B_0 \lambda^2$  と  $B_1 \lambda^2$  は定数にならず、ファクター  $(1-k)^{1/2}$  に従って徐々に増加する。 $h(l)$  が定数であれば、例えばファクターは  $1/2$  である。しかし、こうした差は閾値  $B_{0 \text{ t h r e s}}$  を正しく設定することによって吸収することができる。

【0051】

ここで次式のごとく始点画像を中心が  $(x_0, y_0)$ 、半径  $r$  の円形のオブジェクトであると仮定する。

【数 32】

$$p(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_0)^2 + (j-y_0)^2}) & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases}$$

(式 32)

一方、終点画像は、次式のごとく中心  $(x_1, y_1)$ 、半径が  $r$  のオブジェクトであるとする。

【数 33】

$$q(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases}$$

(式 33)

ここで  $c(x)$  は  $c(x) = x^k$  の形であるとする。中心  $(x_0, y_0)$  及び  $(x_1, y_1)$  が十分遠い場合、ヒストグラム  $h(l)$  は次式の形となる。

10

20

30

40

【数 3 4】  

$$h(l) \propto r l^k \quad (k \neq 0)$$

(式 3 4)

k = 1 のとき、画像は背景に埋め込まれた鮮明な境界線を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。k = - 1 のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従って暗くなる。一般のオブジェクトはこれらふたつのタイプのオブジェクトの間にあると考えてもさして一般性を失わない。したがって、k は - 1 < k < 1 として大抵の場合をカバーでき、式 2 7 が一般に減少関数であることが保障される。

【 0 0 5 2 】

なお、式 3 4 からわかるように、r は画像の解像度に影響されること、すなわち r は 2 m に比例することに注意すべきである。このために [ 1 . 4 . 1 ] においてファクター 2 m を導入した。

【 0 0 5 3 】

[ 1 . 4 . 3 ] の動的決定

パラメータ も同様の方法で自動決定できる。はじめに  $\alpha = 0$  とし、最も細かい解像度における最終的な写像  $f^{(n)}$  及びエネルギー  $C_f^{(n)}$  を計算する。つづいて、 $\alpha$  をある値だけ増加させ、再び最も細かい解像度における最終写像  $f^{(n)}$  及びエネルギー  $C_f^{(n)}$  を計算し直す。この過程を最適値が求まるまで続ける。 $\alpha$  は写像の剛性を示す。次式の重みだからである。

【数 3 5】

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式 3 5)

$\alpha$  が 0 のとき、 $D_f^{(n)}$  は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 $\alpha$  が非常に大きな値のとき、 $D_f^{(n)}$  は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 $\alpha$  の値が 0 から次第に増えるとき、後述のごとく  $C_f^{(n)}$  は徐々に減少する。しかし  $\alpha$  の値が最適値を越えると、図 4 に示すとおり、エネルギーは増加し始める。同図の X 軸は  $\alpha$ 、Y 軸は  $C_f$  である。

【 0 0 5 4 】

この方法で  $C_f^{(n)}$  を最小にする最適な  $\alpha$  の値を得ることができる。しかし、 $\alpha$  の場合に比べていろいろな要素が計算に影響する結果、 $C_f^{(n)}$  は小さく揺らぎながら変化する。 $\alpha$  の場合は、入力が微小量変化するたびに副写像を 1 回計算しなおすだけだが、 $\alpha$  の場合はすべての副写像が計算しなおされるためである。このため、得られた  $C_f^{(n)}$  の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つければ、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【 0 0 5 5 】

[ 1 . 5 ] スーパーサンプリング

画素間の対応関係を決定する際、自由度を増やすために、 $f^{(m,s)}$  の値域を  $R \times R$  に拡張することができる ( $R$  は実数の集合)。この場合、終点画像の画素の輝度が補間され、非整数点、

【数 3 6】

$$V(q_{f^{(m,s)}(i,j)}) \quad (式 3 6)$$

における輝度を持つ  $f^{(m,s)}$  が提供される。つまりスーパーサンプリングが行われる。実験では、 $f^{(m,s)}$  は整数及び半整数値をとることが許され、

10

20

30

40

【数 37】  

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)}) \quad (式 37)$$

は、

【数 38】  

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2 \quad (式 38)$$

によって与えられた。

【0056】

[1.6] 各画像の画素の輝度の正規化

始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー  $C^{(m,s)}_f$  が大きくなりすぎ、正しい評価がしづらいためである。

【0057】

例えば、人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふたつの顔の間の副写像を計算するために、まず副画像を正規化する。すなわち、最も暗い画素の輝度を0、最も明るいそれを255に設定し、他の画素の輝度は線形補間によって求めておく。

【0058】

[1.7] インプリメンテーション

始点画像のスキャンに従って計算がリニアに進行する帰納的な方法を用いる。始めに、1番上の左端の画素  $(i, j) = (0, 0)$  について  $f^{(m,s)}$  の値を決定する。次に  $i$  を1ずつ増やしながらか  $f^{(m,s)}(i, j)$  の値を決定する。 $i$  の値が画像の幅に到達したとき、 $j$  の値を1増やし、 $i$  を0に戻す。以降、始点画像のスキャンに伴い  $f^{(m,s)}(i, j)$  を決定していく。すべての点について画素の対応が決まれば、ひとつの写像  $f^{(m,s)}$  が決まる。

ある  $p(i, j)$  について対応点  $q_f(i, j)$  が決まれば、つぎに  $p(i, j+1)$  の対応点  $q_f(i, j+1)$  が決められる。この際、 $q_f(i, j+1)$  の位置は全単射条件を満たすために、 $q_f(i, j)$  の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに  $(0, 0)$  が最も優先される状態がつづくとき、求められる最終の写像に余計な偏向が加わる。本前提技術ではこの状態を回避するために、 $f^{(m,s)}$  を以下の方法で決めていく。

【0059】

まず  $(s \bmod 4)$  が0の場合、 $(0, 0)$  を開始点とし  $i$  及び  $j$  を徐々に増やしながらか決めていく。 $(s \bmod 4)$  が1の場合、最上行の右端点を開始点とし、 $i$  を減少、 $j$  を増加させながらか決めていく。 $(s \bmod 4)$  が2のとき、最下行の右端点を開始点とし、 $i$  及び  $j$  を減少させながらか決めていく。 $(s \bmod 4)$  が3の場合、最下行の左端点を開始点とし、 $i$  を増加、 $j$  を減少させながらか決めていく。解像度が最も細かい第  $n$  レベルには副写像という概念、すなわちパラメータ  $s$  が存在しないため、仮に  $s = 0$  及び  $s = 2$  であるとしてふたつの方向を連続的に計算した。

【0060】

実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補  $(k, l)$  の中からできる限り全単射条件を満たす  $f^{(m,s)}(i, j)$  ( $m = 0, \dots, n$ ) の値を選んだ。第3の条件を破る候補のエネルギー  $D(k, l)$  には  $\dots$  を掛け、一方、第1または第2の条件を破る候補には  $\dots$  を掛ける。今回は  $\dots = 2$ 、 $\dots = 1000000$  を用いた。

【0061】

前述の全単射条件のチェックのために、実際の手続として  $(k, l) = f^{(m,s)}(i, j)$  を決定する際に以下のテストを行った。すなわち  $f^{(m,s)}(i, j)$  の相続四辺形に含まれる各格子点  $(k, l)$  に対し、次式の外積の  $z$  成分が0以上になるかどうか

10

20

30

40

50

か確かめる。

【数 3 9】

$$W = \vec{A} \times \vec{B} \quad (\text{式 3 9})$$

ただしここで、

【数 4 0】

$$\vec{A} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{f(m,s)}^{(m,s)}(i+1,j-1)} \quad (\text{式 4 0})$$

【数 4 1】

$$\vec{B} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{(k,l)}^{(m,s)}} \quad (\text{式 4 1})$$

10

である（ここでベクトルは三次元ベクトルとし、z 軸は直交右手座標系において定義される）。もし W が負であれば、その候補については  $D^{(m,s)}(k,l)$  に を掛けることによってペナルティを与え、できるかぎり選択しないようにする。

【0 0 6 2】

図 5 ( a )、図 5 ( b ) はこの条件を検査する理由を示している。図 5 ( a ) はペナルティのない候補、図 5 ( b ) はペナルティがある候補をそれぞれ表す。隣接画素  $( i , j + 1 )$  に対する写像  $f^{(m,s)}(i,j+1)$  を決定する際、W の z 成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、 $q^{(m,s)}(k,l)$  は隣接する四辺形の境界線を越えるためである。

20

【0 0 6 3】

[ 1 . 7 . 1 ] 副写像の順序

インプリメンテーションでは、解像度レベルが偶数のときには  $( 0 ) = 0$ 、 $( 1 ) = 1$ 、 $( 2 ) = 2$ 、 $( 3 ) = 3$ 、 $( 4 ) = 0$  を用い、奇数のときは  $( 0 ) = 3$ 、 $( 1 ) = 2$ 、 $( 2 ) = 1$ 、 $( 3 ) = 0$ 、 $( 4 ) = 3$  を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は 4 種類であり、s は 0 ~ 3 のいずれかである。しかし、実際には s = 4 に相当する処理を行った。その理由は後述する。

30

【0 0 6 4】

[ 1 . 8 ] 補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面における正方形  $p(i,j) p(i+1,j) p(i,j+1) p(i+1,j+1)$  が終点画像平面上の四辺形  $q_f(i,j) q_f(i+1,j) q_f(i,j+1) q_f(i+1,j+1)$  に射影されると仮定する。簡単のため、画像間の距離を 1 とする。始点画像平面からの距離が t ( 0 ≤ t ≤ 1 ) である中間画像の画素  $r(x,y,t)$  ( 0 ≤ x ≤ N - 1, 0 ≤ y ≤ M - 1 ) は以下の要領で求められる。まず画素  $r(x,y,t)$  の位置 (ただし  $x, y, t \in R$ ) を次式で求める。

40

【数 4 2】

$$\begin{aligned} (x,y) = & (1-dx)(1-dy)(1-t)(i,j) + (1-dx)(1-dy)tf(i,j) \\ & + dx(1-dy)(1-t)(i+1,j) + dx(1-dy)tf(i+1,j) \\ & + (1-dx)dy(1-t)(i,j+1) + (1-dx)dyt f(i,j+1) \\ & + dx dy(1-t)(i+1,j+1) + dx dy t f(i+1,j+1) \end{aligned}$$

(式 4 2)

つづいて  $r(x,y,t)$  における画素の輝度が次の式を用いて決定される。

50

【数 4 3】

$$\begin{aligned}
 V(r(x, y, t)) = & (1 - dx)(1 - dy)(1 - t)V(p_{(i,j)}) + (1 - dx)(1 - dy)tV(q_{f(i,j)}) \\
 & + dx(1 - dy)(1 - t)V(p_{(i+1,j)}) + dx(1 - dy)tV(q_{f(i+1,j)}) \\
 & + (1 - dx)dy(1 - t)V(p_{(i,j+1)}) + (1 - dx)dytV(q_{f(i,j+1)}) \\
 & + dx dy(1 - t)V(p_{(i+1,j+1)}) + dx dy tV(q_{f(i+1,j+1)})
 \end{aligned}$$

(式 4 3)

ここで  $dx$  及び  $dy$  はパラメータであり、0 から 1 まで変化する。

10

【0 0 6 5】

[ 1 . 9 ] 拘束条件を課したときの写像

いままでは拘束条件がいっさい存在しない場合の写像の決定を述べた。しかし、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束条件としたうえで写像を決定することができる。

【0 0 6 6】

基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、しかる後、写像  $f$  を正確に計算する。

【0 0 6 7】

まず始めに、始点画像の特定の画素を終点画像の特定の画素に射影し、始点画像の他の画素を適当な位置に射影する大まかな写像を決める。すなわち、特定の画素に近い画素は、その特定の画素が射影される場所の近くに射影されるような写像である。ここで第  $m$  レベルの大まかな写像を  $F^{(m)}$  と記述する。

20

【0 0 6 8】

大まかな写像  $F$  は以下の要領で決める。まず、いくつかの画素について写像を特定する。始点画像について  $n_s$  個の画素、

【数 4 4】

$$P(i_0, j_0), P(i_1, j_1), \dots, P(i_{n_s-1}, j_{n_s-1})$$

(式 4 4)

30

を特定するとき、以下の値を決める。

【数 4 5】

$$\begin{aligned}
 F^{(n)}(i_0, j_0) &= (k_0, l_0), \\
 F^{(n)}(i_1, j_0) &= (k_1, l_1), \dots,
 \end{aligned}$$

$$F^{(n)}(i_{n_s-1}, j_{n_s-1}) = (k_{n_s-1}, l_{n_s-1})$$

40

(式 4 5)

始点画像の他の画素の変位量は、 $p_{(i_h, j_h)}$  ( $h = 0, \dots, n_s - 1$ ) の変位に重み付けをして求められる平均である。すなわち画素  $p_{(i, j)}$  は、終点画像の以下の画素に射影される。

【数 4 6】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{n_s-1} (k_h - i_h, l_h - j_h) \text{weight}_h(i, j)}{2^{n-m}} \quad (\text{式 4 6})$$

ただしここで、

【数 4 7】

$$\text{weight}_h(i, j) = \frac{1/\|(i_h - i, j_h - j)\|^2}{\text{total weight}(i, j)} \quad (\text{式 4 7}) \quad 10$$

【数 4 8】

$$\text{total weight}(i, j) = \sum_{h=0}^{n_s-1} 1/\|(i_h - i, j_h - j)\|^2 \quad (\text{式 4 8})$$

とする。

【0069】

つづいて、 $F^{(m)}$  に近い候補写像  $f$  がより少ないエネルギーを持つように、その写像  $f$  のエネルギー  $D^{(m, s)}(i, j)$  を変更する。正確には、 $D^{(m, s)}(i, j)$  は、

【数 4 9】

$$D_{(i,j)}^{(m,s)} = E_{0(i,j)}^{(m,s)} + \eta E_{1(i,j)}^{(m,s)} + \kappa E_{2(i,j)}^{(m,s)} \quad (\text{式 4 9}) \quad 20$$

である。ただし、

【数 5 0】

$$E_{2(i,j)}^{(m,s)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i, j) - f^{(m,s)}(i, j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i, j) - f^{(m,s)}(i, j)\|^2, & \text{otherwise} \end{cases} \quad (\text{式 5 0}) \quad 30$$

であり、 $\rho = 0$  とする。最後に、前述の写像の自動計算プロセスにより、 $f$  を完全に決定する。

【0070】

ここで、 $f^{(m, s)}(i, j)$  が  $F^{(m)}(i, j)$  に十分近いとき、つまりそれらの距離が、

【数 5 1】

$$\left\lfloor \frac{\rho^2}{2^{2(n-m)}} \right\rfloor \quad (\text{式 5 1})$$

以内であるとき、 $E_{2(i,j)}^{(m, s)}(i, j)$  が 0 になることに注意すべきである。そのように定義した理由は、各  $f^{(m, s)}(i, j)$  が  $F^{(m)}(i, j)$  に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

[ 2 ] 具体的な処理手順

[ 1 ] の各要素技術による処理の流れを説明する。

図 6 は前提技術の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い (S 1)、つづいて始点画像と終点画像のマッチングをとる (S 2)。ただし、S 2 は必須ではなく、S 1 で得られた画像の特徴をもとに画像認識などの処理を行ってもよい。

## 【0071】

図7は図6のS1の詳細を示すフローチャートである。ここではS2で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い(S10)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い(S11)、一連の終点階層画像を得る。ただし、S10とS11の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

## 【0072】

図8は図7のS10の詳細を示すフローチャートである。もとの始点画像のサイズは $2^n \times 2^n$ とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ $m$ を $n$ にセットする(S100)。つづいて第 $m$ レベルの画像 $p^{(m,0)}$ 、 $p^{(m,1)}$ 、 $p^{(m,2)}$ 、 $p^{(m,3)}$ から特異点フィルタを用いて特異点を検出し(S101)、それぞれ第 $m-1$ レベルの画像 $p^{(m-1,0)}$ 、 $p^{(m-1,1)}$ 、 $p^{(m-1,2)}$ 、 $p^{(m-1,3)}$ を生成する(S102)。ここでは $m=n$ であるため、 $p^{(m,0)} = p^{(m,1)} = p^{(m,2)} = p^{(m,3)} = p^{(n)}$ であり、ひとつの始点画像から4種類の副画像が生成される。

## 【0073】

図9は第 $m$ レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の数値は各画素の輝度を示す。同図の $p^{(m,s)}$ は $p^{(m,0)} \sim p^{(m,3)}$ の4つの画像を象徴するもので、 $p^{(m-1,0)}$ を生成する場合には、 $p^{(m,s)}$ は $p^{(m,0)}$ であると考え、[1.2]で示した規則により、 $p^{(m-1,0)}$ は例えば同図で輝度を記入したブロックについて、そこに含まれる4画素のうち「3」、 $p^{(m-1,1)}$ は「8」、 $p^{(m-1,2)}$ は「6」、 $p^{(m-1,3)}$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2^{m-1} \times 2^{m-1}$ になる。

## 【0074】

つづいて $m$ をデクリメントし(図8のS103)、 $m$ が負になっていないことを確認し(S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0レベルの副画像が生成された時点でS10が終了する。第0レベルの副画像のサイズは $1 \times 1$ である。

## 【0075】

図10はS10によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが4つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7のS11にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6のS1による処理が完了する。

## 【0076】

前提技術では、図6のS2に進むためにマッチング評価の準備をする。図11はその手順を示している。同図のごとく、まず複数の評価式が設定される(S30)。[1.3.2.1]で導入した画素に関するエネルギー $C^{(m,s)}_f$ と[1.3.2.2]で導入した写像の滑らかさに関するエネルギー $D^{(m,s)}_f$ がそれぞれである。つぎに、これらの評価式を統合して総合評価式を立てる(S31)。[1.3.2.3]で導入した総エネルギー $C^{(m,s)}_f + D^{(m,s)}_f$ がそれぞれであり、[1.3.2.2]で導入した

$$\left( C^{(m,s)}(i,j) + E_0^{(m,s)}(i,j) + E_1^{(m,s)}(i,j) \right) \quad (\text{式52})$$

となる。ただし、総和は $i, j$ についてそれぞれ $0, 1 \dots, 2^m - 1$ で計算する。以上でマッチング評価の準備が整う。

## 【0077】

図12は図6のS2の詳細を示すフローチャートである。[1]で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしで行われる

10

20

30

40

50

。画像間の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優れたものになる。

#### 【0078】

図12のごとく、まず係数パラメータを0、レベルパラメータ $m$ を0に設定する(S20)。つづいて、始点階層画像中の第 $m$ レベルの4つの副画像と終点階層画像中の第 $m$ レベルの4つの副画像のそれぞれの間でマッチングを計算し、それぞれ全単射条件を満たし、かつエネルギーを最小にするような4種類の副写像 $f^{(m, s)}$  ( $s = 0, 1, 2, 3$ )を求める(S21)。全単射条件は[1.3.3]で述べた相続四辺形を用いて検査される。この際、式17、18が示すように、第 $m$ レベルにおける副写像は第 $m-1$ レベルのそれらに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いま $m=0$ であってそれより粗いレベルはないが、この例外的な処理は図13で後述する。

一方、同一レベル内における水平的参照も行われる。[1.3.3]の式20のごとく、 $f^{(m, 3)}$ は $f^{(m, 2)}$ に、 $f^{(m, 2)}$ は $f^{(m, 1)}$ に、 $f^{(m, 1)}$ は $f^{(m, 0)}$ に、それぞれ類似するように決める。その理由は、特異点の種類が違って、それらがもともと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式20からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが良好とみなされる。

#### 【0079】

なお、最初に決めるべき $f^{(m, 0)}$ については同一のレベルで参照できる副写像がないため、式19に示すごとくひとつ粗いレベルを参照する。ただし、実験では $f^{(m, 3)}$ まで求めた後、これを拘束条件として $f^{(m, 0)}$ を一回更新するという手順をとった。これは式20に $s=4$ を代入し、 $f^{(m, 4)}$ を新たな $f^{(m, 0)}$ とすることに等しい。 $f^{(m, 0)}$ と $f^{(m, 3)}$ の関連度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では[1.7.1]に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの関連度を密接に保つ趣旨である。また、処理の開始点に依存する偏向を回避するために、 $s$ の値にしたがって開始点の位置を変える点は[1.7]で述べたとおりである。

#### 【0080】

図13は第0レベルにおいて副写像を決定する様子を示す図である。第0レベルでは各副画像がただひとつの画素で構成されるため、4つの副写像 $f^{(0, s)}$ はすべて自動的に単位写像に決まる。図14は第1レベルにおいて副写像を決定する様子を示す図である。第1レベルでは副画像がそれぞれ4画素で構成される。同図ではこれら4画素が実線で示されている。いま、 $p^{(1, s)}$ の点 $x$ の対応点を $q^{(1, s)}$ の中で探すとき、以下の手順を踏む。

#### 【0081】

1. 第1レベルの解像度で点 $x$ の左上点 $a$ 、右上点 $b$ 、左下点 $c$ 、右下点 $d$ を求める。
2. 点 $a \sim d$ がひとつ粗いレベル、つまり第0レベルにおいて属する画素を探す。図14の場合、点 $a \sim d$ はそれぞれ画素 $A \sim D$ に属する。ただし、画素 $A \sim C$ は本来存在しない仮想的な画素である。
3. 第0レベルですでに求まっている画素 $A \sim D$ の対応点 $A' \sim D'$ を $q^{(1, s)}$ の中にプロットする。画素 $A' \sim C'$ は仮想的な画素であり、それぞれ画素 $A \sim C$ と同じ位置にあるものとする。
4. 画素 $A$ の中の点 $a$ の対応点 $a'$ が画素 $A'$ の中にあるとみなし、点 $a'$ をプロットする。このとき、点 $a$ が画素 $A$ の中で占める位置(この場合、右下)と、点 $a'$ が画素 $A'$ の中で占める位置が同じであると仮定する。
5. 4と同様の方法で対応点 $b' \sim d'$ をプロットし、点 $a' \sim d'$ で相続四辺形を作る。

10

20

30

40

50

6. 相続四辺形の中でエネルギーが最小になるよう、点 $x$ の対応点 $x'$ を探す。対応点 $x'$ の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図14の場合、4つの画素がすべて候補になる。

【0082】

以上がある点 $x$ の対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第2レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図3に示すように画素 $A' \sim D'$ の間隔が空いていく状況が発生する。

【0083】

こうして、ある第 $m$ レベルの4つの副写像が決まれば、 $m$ をインクリメントし(図12のS22)、 $m$ が $n$ を超えていないことを確かめて(S23)、S21に戻る。以下、S21に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後にS21に戻ったときに第 $n$ レベルの写像 $f^{(n)}$ を決める。この写像は $\epsilon = 0$ に関して定まったものであるから、 $f^{(n)}(\epsilon = 0)$ と書く。

10

【0084】

つぎに異なる $\epsilon$ に関する写像も求めるべく、 $\epsilon$ を $\epsilon_{\text{step}}$ だけシフトし、 $m$ をゼロクリアする(S24)。新たな $\epsilon$ が所定の探索打切り値 $\epsilon_{\text{max}}$ を超えていないことを確認し(S25)、S21に戻り、今回の $\epsilon$ に関して写像 $f^{(n)}(\epsilon = \epsilon_{\text{step}})$ を求める。この処理を繰り返し、S21で $f^{(n)}(\epsilon = i \cdot \epsilon_{\text{step}})$ ( $i = 0, 1, \dots$ )を求めていく。 $\epsilon$ が $\epsilon_{\text{max}}$ を超えたときS26に進み、後述の方法で最適な $\epsilon = \epsilon_{\text{opt}}$ を決定し、 $f^{(n)}(\epsilon = \epsilon_{\text{opt}})$ を最終的に写像 $f^{(n)}$ とする。

20

【0085】

図15は図12のS21の詳細を示すフローチャートである。このフローチャートにより、ある定まった $\epsilon$ について、第 $m$ レベルにおける副写像が決まる。副写像を決める際、前提技術では副写像ごとに最適な $\epsilon$ を独立して決める。

【0086】

同図のごとく、まず $s$ と $\epsilon$ をゼロクリアする(S210)。つぎに、そのときの $\epsilon$ について(および暗に $\epsilon$ について)エネルギーを最小にする副写像 $f^{(m, s)}$ を求め(S211)、これを $f^{(m, s)}(\epsilon = 0)$ と書く。異なる $\epsilon$ に関する写像も求めるべく、 $\epsilon$ を $\epsilon_{\text{step}}$ だけシフトし、新たな $\epsilon$ が所定の探索打切り値 $\epsilon_{\text{max}}$ を超えていないことを確認し(S213)、S211に戻り、以降の繰り返し処理で $f^{(m, s)}(\epsilon = i \cdot \epsilon_{\text{step}})$ ( $i = 0, 1, \dots$ )を求める。 $\epsilon$ が $\epsilon_{\text{max}}$ を超えたときS214に進み、最適な $\epsilon = \epsilon_{\text{opt}}$ を決定し、 $f^{(m, s)}(\epsilon = \epsilon_{\text{opt}})$ を最終的に写像 $f^{(m, s)}$ とする(S214)。

30

【0087】

つぎに、同一レベルにおける他の副写像を求めるべく、 $\epsilon$ をゼロクリアし、 $s$ をインクリメントする(S215)。 $s$ が4を超えていないことを確認し(S216)、S211に戻る。 $s = 4$ になれば上述のごとく $f^{(m, 3)}$ を利用して $f^{(m, 0)}$ を更新し、そのレベルにおける副写像の決定を終了する。

【0088】

図16は、ある $m$ と $s$ について $\epsilon$ を変えながら求められた $f^{(m, s)}(\epsilon = i \cdot \epsilon_{\text{step}})$ ( $i = 0, 1, \dots$ )に対応するエネルギー $C^{(m, s)}_f$ の挙動を示す図である。[1.4]で述べたとおり、 $\epsilon$ が増加すると通常 $C^{(m, s)}_f$ は減少する。しかし、 $\epsilon$ が最適値を超えると $C^{(m, s)}_f$ は増加に転じる。そこで本前提技術では $C^{(m, s)}_f$ が極小値をとるときの $\epsilon = \epsilon_{\text{opt}}$ と決める。同図のように $\epsilon > \epsilon_{\text{opt}}$ の範囲で再度 $C^{(m, s)}_f$ が小さくなっていても、その時点ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 $\epsilon_{\text{opt}}$ は副写像ごとに独立して決めていき、最後に $f^{(n)}$ についてもひとつ定まる。

40

【0089】

一方、図17は、 $\epsilon$ を変えながら求められた $f^{(n)}(\epsilon = i \cdot \epsilon_{\text{step}})$ ( $i = 0, 1, \dots$ )に対応するエネルギー $C^{(n)}_f$ の挙動を示す図である。ここでも $\epsilon$ が増加すると通常

50

$C^{(n)}_f$  は減少するが、 $\theta$  が最適値を超えると  $C^{(n)}_f$  は増加に転じる。そこで  $C^{(n)}_f$  が極小値をとるときの  $\theta_{opt}$  と決める。図 17 は図 4 の横軸のゼロ付近を拡大した図と考えてよい。 $\theta_{opt}$  が決まれば  $f^{(n)}$  を最終決定することができる。

#### 【0090】

以上、本前提技術によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先験的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

10

#### 【0091】

なお、本前提技術について次のような変形技術も考えられる。

(1) 前提技術では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の 2 枚の画像間のマッチングをとる場合全般に利用できる。

#### 【0092】

たとえば 2 枚の画像間で、画素の輝度の差に関するエネルギー  $E_0$  と画素の位置的なずれに関するエネルギー  $E_1$  のふたつを評価式とし、これらの線形和  $E_{t.o.t} = E_0 + E_1$  を総合評価式とする。この総合評価式の極値付近に注目して  $\theta$  を自動決定する。つまり、いろいろな  $\theta$  について  $E_{t.o.t}$  が最小になるような写像を求める。それらの写像のうち、 $\theta$  に関して  $E_1$  が極小値をとるときの  $\theta$  を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

20

#### 【0093】

これ以外にも評価式の設定にはいろいろな方法があり、例えば  $1/E_1$  と  $1/E_2$  のように、評価結果が良好なほど大きな値をとるものを採用してもよい。総合評価式も必ずしも線形和である必要はなく、 $n$  乗和 ( $n = 2, 1/2, -1, -2$  など)、多項式、任意の関数などを適宜選択すればよい。

#### 【0094】

パラメータも、 $\theta$  のみ、前提技術のごとく  $\theta$  と  $\phi$  のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが 3 以上の場合はひとつずつ変化させて決めていく。

30

(2) 本前提技術では、総合評価式の値が最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である  $C^{(m,s)}_f$  が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によっては単に総合評価式の最小値が最小になるようにパラメータを決めても効果的である。その場合、例えば  $E_0 + E_1$  を総合評価式とし、 $\theta + \phi = 1$  なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてもよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

(3) 前提技術では各解像度レベルで 4 種類の特異点に関する 4 種類の副画像を生成した。しかし、当然 4 種類のうち 1、2、3 種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する  $f^{(m,3)}$  だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 $s$  に関する計算量が減る効果がある。

40

(4) 本前提技術では特異点フィルタによってレベルがひとつ進むと画素が  $1/4$  になった。例えば  $3 \times 3$  で 1 ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は  $1/9$  になる。

(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像を変換する。それ以外の方法として、RGB の各成分について副写像を計算してもよい。

#### 【0095】

[ 画像符号化と復号に関する実施の形態 ]

50

以上の前提技術を利用した画像の符号化および復号の具体例を述べる。

図18は画像符号化装置10の構成を示す。画像符号化装置10は、キーフレームを外部から取得する画像入力部12と、キーフレーム間で前提技術を用いた対応点計算をするマッチングプロセッサ14と、キーフレームおよび対応点情報を組み込んで符号化データストリーム（以下単にストリームともいう）を生成するストリーム生成部16と、そのストリームをネットワークを介して画像復号装置へ送信する通信部18を含む。

【0096】

画像符号化装置10はさらに、対応点情報からキーフレーム間の変化量が大きいかなを確認する検査部20と、画像処理に関するユーザの要求を受け付けるUI24を含む。また、ストリーム生成部16は指示付加部22を含み、これがストリームの中に画質低下処理の指示を埋め込む。

10

【0097】

検査部20は、例えばふたつのキーフレームKF1、KF2間の対応点位置のずれの総和を計算し、その平均値が所定のしきい値を超えたとき、変化量が大きいと判断する。そう判断したとき、その旨は指示付加部22へ伝えられ、少なくともキーフレームKF1とKF2の間の中間フレームについて画質低下処理を施すべく指示が組み込まれる。なお、検査部20による指示のほか、ユーザが編集過程にてUI24を介して明示的に画質低下処理の指示を入れてもよい。こうしたユーザインタラクションを実現することで、画像符号化装置10は画像編集ツールとしての性格をもつ。

【0098】

図19はストリームCDS（Coded Data Stream）の構造を示す。同図のごとく、まず2枚のキーフレームKF1、KF2のデータが置かれ、それらの対応点情報C<sub>12</sub>が続く、このあと、それらのキーフレーム間の中間フレームに画質低下処理を施すべきかなの指示Dが付加されている。この場合、さらに進んで、画質低下処理の内容を組み込んでよい。

20

【0099】

図20はストリームCDSを受けて実際に動画を再生する画像復号装置100の構成を示す。画像復号装置100は、任意のネットワークからデータを受信する通信部102と、受信データからストリームを入力する画像入力部104と、ストリームに含まれるキーフレームのデータおよび対応点情報をもとに補間計算で中間フレームを生成する中間画像生成部106と、その際にワークエリアおよび画像出力のタイミング調整のために利用されるバッファメモリ108と、最終的に得られた画像を表示する表示部110を含む。

30

【0100】

画像復号装置100はさらに、画像入力部104が入力したストリームから対応点情報を検出したうえでキーフレーム間の変化量の大きさを検証する検査部112と、ストリームに組み込まれた前述の指示があればこれを検出する指示検出部114と、画像再生に関するユーザの指示を受け付けるUI116とを含む。これらはすべて備える必要はなく、いずれかひとつでも十分である。

【0101】

中間画像生成部106は、ブレンド処理部120、解像度低減部122、LPF124を含む。これらはそれぞれ異なる画質低下処理を行うもので、ブレンド処理部120は複数の中間フレームにモーションブラー効果を施す。例えば連続する3枚の中間フレームをIF1、IF2、IF3とすると、モーションブラーは、これらのキーフレームを適当な値でブレンドして得られる。たとえば、IF1を表示すべきタイミングでは、IF1のブレンド比が大きくなり、同様にIF2やIF3の表示タイミングではそれぞれそれらのブレンド比が大きくなるように中間フレームを生成すればよい。

40

【0102】

解像度低減部122は、中間フレームの解像度をいったん落とし、その後、もとの画素数に戻して出力する。解像度低減部122は画素を単純に間引いてもよいし、4画素単位にアベレージフィルタを掛ける処理を数回繰り返してもよいし、本質的に画質が低下し、

50

または画像のシャープさが低減する処理を施せばよい。

【0103】

LPF124は、ローパスフィルタリングにより、画像をなまらせている。いずれかにしても、さして正確な中間フレームが得られないと判定されたときは、むしろ画質を落として対応点情報の誤差を見えにくくする点に特徴がある。なお、これら3つの処理部はいずれかひとつ存在すれば十分である。

【0104】

以上の画像復号装置100の動作は以下のとおりである。ストリームに指示が入っていることがわかっている場合は、指示検出部114によってその指示を探索し、あればその旨を中間画像生成部106へ伝える。中間画像生成部106はその指示のあるキーフレーム間の中間フレームを生成するとき、ブレンド処理部120、解像度低減部122、LPF124のいずれかを起動する。指示自体がいずれかを指定していればそれにしたがう。モーションブラー等の処理のために、バッファメモリ108をワークメモリとして利用する。また、最終的に得られる中間フレームをキーフレームとともに正しいタイミングで出力するためにも、バッファメモリ108がバッファとしての役割を果たす。

10

【0105】

ストリームに指示が入っているかどうか不明な場合や、仮に入っている場合、さらに独自に画質低下処理を施したい場合、検査部112またはUI116を利用する。検査部112は図18の検査部20同様の方法でキーフレーム間の変化の大きさを判定し、必要に応じて画質低下処理を中間画像生成部106へ指示する。

20

【0106】

例えば人がキーフレーム間で大きく移動するようになると、補間による中間フレームは一般に、人が関係のない背景の一部と誤対応を起こして不自然な画像になる。前提技術によれば、そうした場合でも相当に高い対応精度を得ることができるが、本実施の形態の効果を加えることにより、さらに自然な画像が得られることがわかった。一般論として、キーフレーム間の変化が大きいときは被写体が速く動いている場合が多く、したがって、モーションブラー効果は見た目にも自然であるところ、それが中間フレームの精度の問題を同時に解決しているためである。ローパスフィルタや解像度の低下も同様、画像のスピード感を損なうことなく、課題を解決している。

【0107】

以上、画像符号化と復号を説明した。この装置も符号化の場合と同様、ユーザインタラクションを確保することにより、画像編集ツールとして機能する。なお、これらの実施の形態は例示であり、いろいろな変形技術もある。以下、そうした例を挙げる。

30

【0108】

画像符号化装置10においてマッチングプロセッサ14は必須ではなく、外部で生成されたストリームに対して指示を付加するだけであってもよい。マッチングプロセッサ14がある場合も、それは前提技術を使うことは必須ではない。また、通信部18の代わりに図示しない記憶装置を備え、ストリームをそれに保存してもよい。

【0109】

画質低下処理には他にも例がある。例えば、誤差があるぐらいならその中間フレームを飛ばしてしまうという考え方もありうる。その場合、キーフレーム間がコマ送りになる。この場合も、ゼロ枚の中間フレームを生成するという意味で、中間フレームの生成における画質低下処理と把握できる。

40

【0110】

キーフレーム間の変化量の多寡は、フレーム全体ではなく、その一部のターゲット領域について検証してもよい。たとえば、一方のキーフレームにおいて領域を指定し、その領域内の各点が他方のキーフレームにおいていずれの点に対応するかを確認し、それらの点の平均のずれを所定のしきい値と比較してもよい。いずれの場合も、しきい値は実験的に定めることができる。

【0111】

50

キーフレームにもJPEG、GIF等のフレーム内圧縮を加えることにより、本実施の形態はアニメーションの生成に好適である。

【0112】

実施の形態では、検査部において、対応点情報の変化が大きいか否かを判定したが、それに限らず、キーフレーム間の変化の程度を検査してもよい。たとえば、シーンチェンジ検出技術を用いたり、キーフレーム間の差分画像を生成したりすることで、変化の程度が判明する。変化の程度が大きいと判定された場合にかぎり、指示を追加するようにしてもよい。

【0113】

実施の形態では、生成された中間フレームをぼかす等の処理によって目的を達成したが、これは中間フレームではなく、中間フレームそれを生成するもとになるキーフレームに対して行ってもよい。たとえば、問題となる中間フレームの両側にあるキーフレームに上述のいずれかまたはそれ以外の任意の画質低下処理を施した後、当該中間フレームを生成してもよい。

【図面の簡単な説明】

【0114】

【図1】図1(a)とは図1(b)は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図1(c)と図1(d)は、ふたりの人物の顔に関して前提技術で求められる $p^{(5,0)}$ の画像、図1(e)と図1(f)は、ふたりの人物の顔に関して前提技術で求められる $p^{(5,1)}$ の画像、図1(g)と図1(h)は、ふたりの人物の顔に関して前提技術で求められる $p^{(5,2)}$ の画像、図1(i)と図1(j)は、ふたりの人物の顔に関して前提技術で求められる $p^{(5,3)}$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図2】図2(R)はもとの四辺形を示す図、図2(A)、図2(B)、図2(C)、図2(D)、図2(E)はそれぞれ相続四辺形を示す図である。

【図3】始点画像と終点画像の関係、および第mレベルと第m-1レベルの関係を相続四辺形を用いて示す図である。

【図4】パラメータとエネルギー $C_f$ の関係を示す図である。

【図5】図5(a)、図5(b)は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図6】前提技術の全体手順を示すフローチャートである。

【図7】図6のS1の詳細を示すフローチャートである。

【図8】図7のS10の詳細を示すフローチャートである。

【図9】第mレベルの画像の一部と、第m-1レベルの画像の一部の対応関係を示す図である。

【図10】前提技術で生成された始点階層画像を示す図である。

【図11】図6のS2に進む前に、マッチング評価の準備の手順を示す図である。

【図12】図6のS2の詳細を示すフローチャートである。

【図13】第0レベルにおいて副写像を決定する様子を示す図である。

【図14】第1レベルにおいて副写像を決定する様子を示す図である。

【図15】図12のS21の詳細を示すフローチャートである。

【図16】ある $f^{(m,s)}$ についてを変えながら求められた $f^{(m,s)}$  ( $= i$ ) に対応するエネルギー $C^{(m,s)}_f$ の挙動を示す図である。

【図17】を変えながら求められた $f^{(n)}$  ( $= i$ ) ( $i = 0, 1, \dots$ ) に対応するエネルギー $C^{(n)}_f$ の挙動を示す図である。

【図18】実施の形態に係る画像符号化装置の構成図である。

【図19】図18の装置によって生成されるストリームの模式図である。

【図20】実施の形態に係る画像復号装置の構成図である。

【符号の説明】

【0115】

10

20

30

40

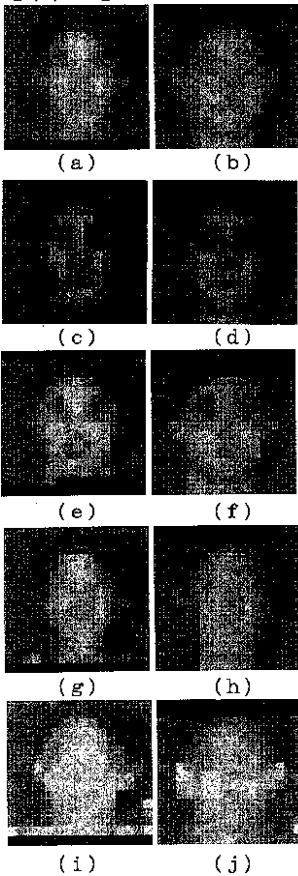
50

- 1 0 画像符号化装置
- 1 2 画像入力部
- 1 4 マッチングプロセッサ
- 1 6 ストリーム生成部
- 1 8 通信部
- 2 0 検査部
- 2 2 指示付加部
- 2 4 U I
- 1 0 0 画像復号装置
- 1 0 2 通信部
- 1 0 4 画像入力部
- 1 0 6 中間画像生成部
- 1 0 8 バッファメモリ
- 1 1 0 表示部
- 1 1 2 検査部
- 1 1 4 指示検出部
- 1 1 6 U I
- 1 2 0 ブレンド処理部
- 1 2 2 解像度低減部
- 1 2 4 L P F

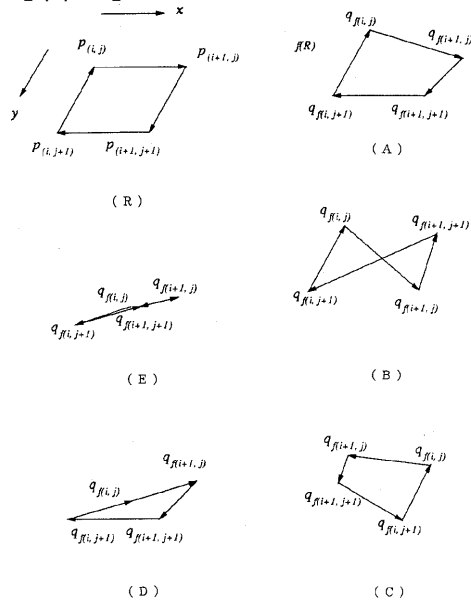
10

20

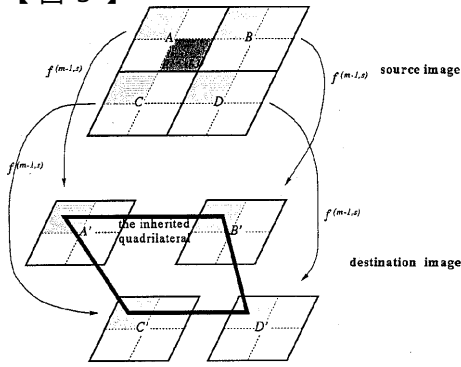
【 図 1 】



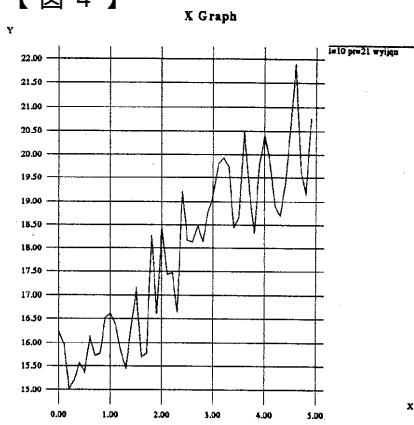
【 図 2 】



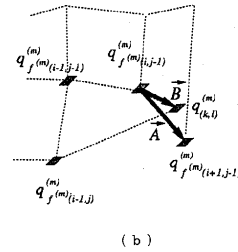
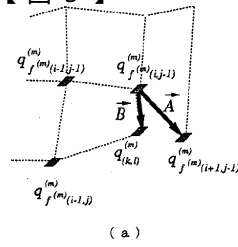
【図3】



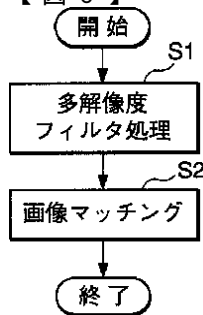
【図4】



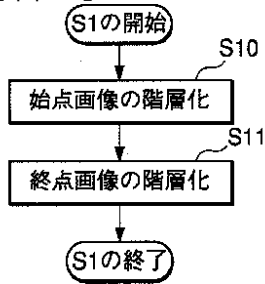
【図5】



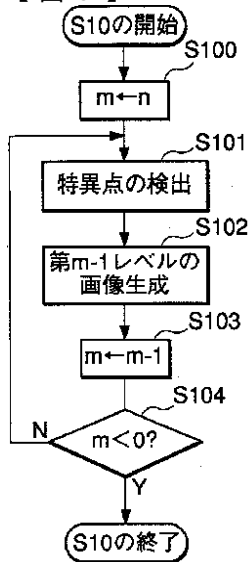
【図6】



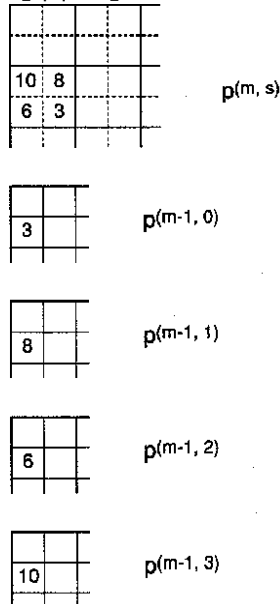
【図7】



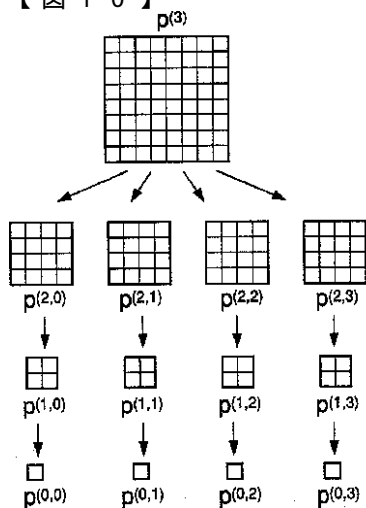
【図8】



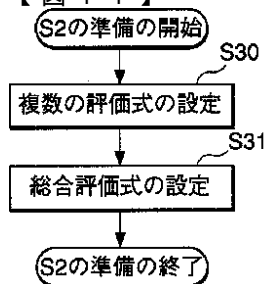
【図9】



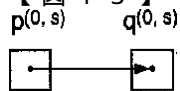
【図10】



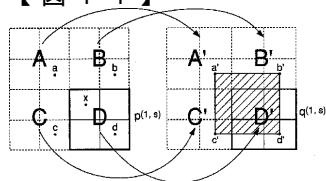
【図11】



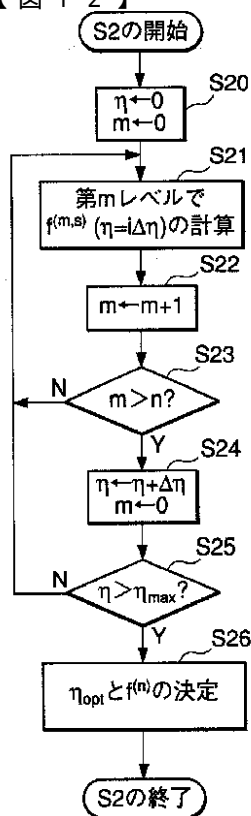
【図13】



【図14】



【図12】



【図15】

