(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2005 年12 月22 日 (22.12.2005)

PCT

(10) 国際公開番号
WO 2005/122593 A1

(51) 国際特許分類[7]: H04N 7/36

(21) 国際出願番号: PCT/JP2005/005941

(22) 国際出願日: 2005 年3 月29 日 (29.03.2005)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願2004-175979 2004 年6 月14 日 (14.06.2004) JP

(71) 出願人 (米国を除く全ての指定国について): 株式会社モノリス (MONOLITH CO., LTD.) [JP/JP]; 〒1060045 東京都港区麻布十番１－７－３ Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 山下 伸一 (YAMASHITA, Shinichi) [JP/JP]; 〒1060045 東京都港区麻布十番１－７－３ 株式会社モノリス内 Tokyo (JP). 遠藤 増春 (ENDO, Masuharu) [JP/JP]; 〒1060045 東京都港区麻布十番１－７－３ 株式会社モノリス内 Tokyo (JP). 秋吉 仰三 (AKIYOSHI, Kozo) [JP/JP]; 〒1060045 東京都港区麻布十番１－７－３ 株式会社モノリス内 Tokyo (JP). 秋吉 信雄 (AKIYOSHI, Nobuo) [JP/JP]; 〒1060045 東京都港区麻布十番１－７－３ 株式会社モノリス内 Tokyo (JP).

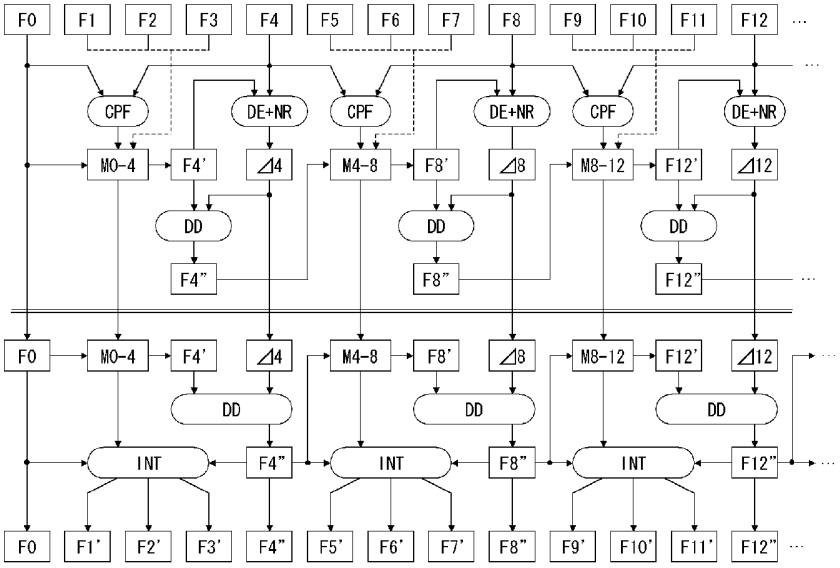
(74) 代理人: 森下 賢樹 (MORISHITA, Sakaki); 〒1500021 東京都渋谷区恵比寿西２－１１－１２ Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,

[続葉有]

(54) Title: MOVING PICTURE ENCODING METHOD AND MOVING PICTURE DECODING METHOD

(54) 発明の名称: 動画符号化方法および動画復号方法



(57) Abstract: When a moving picture is encoded with a high compression ratio by the MPEG, the block noise becomes remarkable. Matching is calculated between the first and the second key frame (F0, F4) and corresponding point information (M0-4) is generated. According to this corresponding point information, a virtual second key frame (F4') is generated. A difference between the actual second key frame (F4) and the virtual second key frame (F4') is compressed/encoded by a difference encoder DE. The F0, M0-4, and the compressed/encoded difference Δ4 are outputted as encoded data between the key frame F0 and F4.

(57) 要約: ＭＰＥＧで動画を高い圧縮率で符号化するとブロックノイズが顕著になる。第１、第２キーフレーム（Ｆ０、Ｆ４）間でマッチングを計算し、対応点情報（Ｍ０－４）を生成する。この対応点情報をもとに、仮想の第２キーフレーム（Ｆ４'）を生成する。現実の第２キーフレーム（Ｆ４）と仮想の第２キーフレーム（Ｆ４'）との差分を差分符号化器ＤＥで圧縮符号化する。Ｆ０、Ｍ

[続葉有]

WO 2005/122593 A1

BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) 指定国 *(表示のない限り、全ての種類の広域保護が可能)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

２文字コード及び他の略語については、定期発行される各*PCT*ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

０－４、および、圧縮符号化された差分Δ４をキーフレームＦ０、Ｆ４間の符号化データとして出力する。

明 細 書

動画符号化方法および動画復号方法

技術分野

[0001] この発明は、画像処理技術に関し、とくにマッチングを用いた動画符号化技術および動画復号技術に関する。

背景技術

[0002] MPEG(Motion Picture Experts Group)は動画圧縮のひとつの標準技術である。MPEGでは、ブロックマッチングが利用される。このマッチングは、ブロック間の差分が最小になるようブロック探索を行う。そのため、差分は確かに小さくなるが、必ずしもフレーム間で本来対応しあう領域どおしが対応づけられるわけではない。

発明の開示

発明が解決しようとする課題

[0003] MPEGでは、圧縮率を上げようとすると、いわゆるブロックノイズが問題になる。このノイズの発生を抑え、フレーム間コヒーレンシに注目した圧縮率をさらに上げるためには、現状のブロックマッチングベースの技術を改める必要である。求める技術は、本来対応しあう領域なり画素なりが正しく対応するよう符号化すべきであり、また、単純なブロックマッチングは避けることが望ましい。

本発明の目的は、この課題を解決する動画符号化および復号技術を提供することにある。

課題を解決するための手段

[0004] 本発明は画像マッチング技術を利用する。この技術は本出願人が先に特許第2927350号にて提案した技術(以下「前提技術」という)を利用することができる。本発明の動画符号化方法は、以下の処理を実行する。

[0005] a) 1以上の画像フレームを間に挟む第1、第2キーフレーム間でマッチングを計算し、第1、第2キーフレーム間の対応点情報を生成するステップと、

b) 第1、第2キーフレーム間の対応点情報をもとに第1キーフレームに含まれる画素を移動させることによって、仮想の第2キーフレームを生成するステップと、

c） 現実の第2キーフレームと仮想の第2キーフレームとの差分を圧縮符号化するステップと、

d） 第1キーフレーム、第1、第2キーフレーム間の対応点情報、および、現実の第2キーフレームと仮想の第2キーフレーム間で圧縮符号化された差分をこれらのキーフレーム間の符号化データとして出力するステップ。

[0006] 一方、本発明の動画復号方法は、以下の処理を実行する。

k） 1以上の画像フレームを間に挟む第1、第2キーフレーム間の対応点情報、および第1キーフレームを取得するステップと、

l） 第1、第2キーフレーム間の対応点情報をもとに、第1キーフレームに含まれる画素を移動させることによって、仮想の第2キーフレームを生成するステップと、

m） 予め符号化側にて求められた現実の第2キーフレームと仮想の第2キーフレームとの差分の圧縮符号化データを取得するステップと、

o） 取得された差分の圧縮符号化データと前記仮想の第2キーフレームとから、改良された仮想の第2キーフレームを生成するステップと、

p） 第1、第2キーフレーム間の対応点情報をもとに、第1キーフレームと改良された仮想の第2キーフレーム間で補間計算をすることにより、これらのキーフレームの間に存在すべき中間フレームを生成するステップと、

q） 第1キーフレーム、生成された中間フレーム、改良された仮想の第2キーフレームをこれらのキーフレーム間の復号データとして出力するステップ。

[0007] 本発明の動画符号化方法は、さらに前記a）のステップにおけるマッチングの正確さを評価するステップを備え、評価の結果に依存して、前記c）のステップにおける圧縮スキームを切り替えてもよい。評価するステップでは、キーフレーム間のマッチングエネルギー値に着目してマッチングの正確さを評価してもよい。マッチングエネルギーとは、例えば後述する前提技術を利用した画像マッチングにおいて算出される、対応点どうしの距離と画素値の違いに基づく値である。

[0008] 本発明の別の態様は、動画符号化方法である。この方法は、第1、第2画像フレーム間で、領域ベースの画像マッチングを計算し、そのマッチング結果を利用して、少なくとも第3画像フレームを符号化する方法であって、画像マッチングの結果の良否

を領域ごとに判定するステップと、第3画像フレームの符号化プロセスにおいて、判定するステップにおいて判定された良否に基づき、領域ごとに量子化スキームを選択するステップとを備える。

[0009] 以上の各ステップを入れ替えたり、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ものもまた、本発明として有効である。

発明の効果

[0010] 本発明によれば、比較的高い圧縮率を実現しつつ、動画の画質が高まる。

図面の簡単な説明

[0011] [図1]図1(a)と図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η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λを変えながら求められた$f^{(m,s)}(\lambda=i\Delta\lambda)$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]ηを変えながら求められた$f^{(n)}(\eta=i\Delta\eta)$(i=0, 1, …)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]改良後の前提技術において第mレベルにおける副写像を求めるフローチャートである。

[図19]実施の形態に係る動画の符号化装置および復号装置の構成および処理を示す図である。

[図20]実施の形態に係る差分符号化器およびノイズリデューサの構成を示す図である。

符号の説明

[0012] Fx 現実の画像フレーム、 CPF 画像マッチングプロセッサ、 DE 差分符号化器、 NR ノイズリデューサ、 DD 差分復号器、 INT 補間プロセッサ、 Fx' 仮想の画像フレーム、 Fx" 改良された仮想の画像フレーム、 Mx－y 対応点情報ファイル

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19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。

[前提技術の実施の形態]

[0014] 最初に[1]で前提技術の要素技術を詳述し、[2]で処理手順を具体的に説明する

。さらに[3]で前提技術に基づき改良を施した点について述べる。

[1]要素技術の詳細

[1．1]イントロダクション

[0015] 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。

[0016] 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。

[1．2]特異点フィルタの階層

[0017] 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∈I)。

[0018] 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$2^m\times 2^m$($0\leqq m\leqq n$)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。

[0019] [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]

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

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。$\min_{x \leqq t \leqq x+1}$、$\max_{x \leqq t \leqq x+1}$をそれぞれα及びβ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。

$$P^{(m,0)} = \alpha(x)\,\alpha(y)\,p^{(m+1,0)}$$

$$P^{(m,1)} = \alpha(x)\,\beta(y)\,p^{(m+1,1)}$$

$$P^{(m,2)} = \beta(x)\,\alpha(y)\,p^{(m+1,2)}$$

$$P^{(m,3)} = \beta(x)\,\beta(y)\,p^{(m+1,3)}$$

[0020] すなわち、これらはαとβ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2×2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。

[0021] 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、α(x)α(y)は極小点を保存し、β(x)β(y)は極大点を保存し、α(x)β(y)及びβ(x)α(y)は鞍点を保存する。

[0022] 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。

[0023] 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。

[0024] 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。

[0025] 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。

[1.3]画像間の写像の計算

[0026] 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∈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, …)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。

[0027] [数3]

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

ここでσ(i)∈{0, 1, 2, 3}である。

[1. 3. 1]全単射

[0028] 始点画像と終点画像の間のマッチ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。

[0029] 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, …)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$q^{(m,s)}_{(k,l)}$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。

[0030] 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]

$$p^{(m,s)}_{(i,j)}\,p^{(m,s)}_{(i+1,j)}\,p^{(m,s)}_{(i+1,j+1)}\,p^{(m,s)}_{(i,j+1)} \qquad (式4)$$

を考える(i=0, …, 2^m-1、j=0, …, 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。

[0031] [数5]

$$\overrightarrow{p^{(m,s)}_{(i,j)}p^{(m,s)}_{(i+1,j)}},\ \overrightarrow{p^{(m,s)}_{(i+1,j)}p^{(m,s)}_{(i+1,j+1)}},\ \overrightarrow{p^{(m,s)}_{(i+1,j+1)}p^{(m,s)}_{(i,j+1)}}\ \text{and}\ \overrightarrow{p^{(m,s)}_{(i,j+1)}p^{(m,s)}_{(i,j)}} \qquad (式5)$$

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、

[0032] [数6]

$$q^{(m,s)}_{f(i,j)}q^{(m,s)}_{f(i+1,j)}q^{(m,s)}_{f(i+1,j+1)}q^{(m,s)}_{f(i+1,j+1)} \qquad (式6)$$

は、以下の全単射条件を満たす必要がある。

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。

[0033] 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。

[0034] 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。

[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

[0035] 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。

[0036] [数7]

$$C^{(m,s)}_{(i,j)} = |V(p^{(m,s)}_{(i,j)}) - V(q^{(m,s)}_{f(i,j)})|^2 \qquad (式7)$$

ここで、$V(p^{(m,s)}_{(i,j)})$及び$V(q^{(m,s)}_{f(i,j)})$はそれぞれ画素$p^{(m,s)}_{(i,j)}$及び$q^{(m,s)}_{f(i,j)}$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$C^{(m,s)}_{(i,j)}$の合計で定義できる。

[数8]

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

[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

[0037] 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Dfを導入する。このエネルギーは画素の輝度とは関係なく、$p^{(m,s)}_{(i,j)}$および$q^{(m,s)}_{f(i,j)}$の位置によって決まる(i=0, …, 2^m-1, j=0, …, 2^m-1)。点(i, j)における写像$f^{(m,s)}$のエネルギー$D^{(m,s)}_{(i,j)}$は次式で定義される。

[0038] [数9]

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

ただし、係数パラメータηは0以上の実数であり、また、

[数10]

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

[数11]

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

とする。ここで、

[数12]

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

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の間の隔たりを表す。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。

[0039] [数13]

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

[1．3．2．3]写像の総エネルギー

[0040] 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$\lambda C_f^{(m,s)} + D_f^{(m,s)}$で定義される。ここで係数パラメータλ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。

[0041] [数14]

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \qquad (式１４)$$

[0042] λ＝0及びη＝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, …, 2^m-1及びj＝0, …, 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λ＝0及びη＝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λの位置を変えて$C_f^{(m,s)} + \lambda D_f^{(m,s)}$と定義したとすれば、λ＝0及びη＝0の場合に総合評価式が$C_f^{(m,s)}$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っ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。

[0043] オプ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、オプ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。

[1．3．3]多重解像度の導入による写像の決定

[0044] 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。

まず、

[数15]

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \qquad (式15)$$

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16]

$$parent(i, j) = ([\frac{i}{2}], [\frac{j}{2}]) \qquad (式16)$$

[0045] $p^{(m, s)}_{(i, j)}$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, …, 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。

[0046] [数17]

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

（式17）

ただしここで、

[数18]

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

（式18）

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。

[0047] 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$q^{(m,s)}_{f(m)(i,j)}$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。

[0048] 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]

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

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]

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

[0049] 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。

[0050] 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。

[0051] 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にくくしている。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。

[1. 4]最適なパラメータ値の自動決定

[0052] 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。

[0053] 前提技術に係るシステムはふたつのパラメータ、λ及びηを含む。端的にいえば、λは画素の輝度の差の重みであり、η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η＝0に固定してλを0から徐々に増加させる。λ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$C_f^{(m,s)}$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、λが最適値を超えると以下の現象が発生する。

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。

3. その結果、式14において$D_f^{(m,s)}$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、$D_f^{(m,s)}$の急激な増加を抑制するよう$f^{(m,s)}$が変化し、その結果$C_f^{(m,s)}$が増加する。

[0054] したがって、λ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$C_f^{(m,s)}$が減少から増加に転じる閾値を検出し、そのλをη＝0における最適値とする。つぎ

にη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ηを自動決定する。そのηに対応してλも決まる。

[0055] 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。

[1．4．1]λの動的決定

[0056] λは0から所定の刻み幅で増加されていき、λ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$\lambda C^{(m,s)}_f + D^{(m,s)}_f$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$\lambda C^{(m,s)}_{(i,j)}$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、$\lambda C^{(m,s)}_{(i,j)}$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。

[0057] この条件のもと、λ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。$\lambda l^2 \geqq 1$が成り立つために、例えば$l^2 = 1/\lambda$の場合を考える。λがλ_1からλ_2まで微小量変化するとき、

[数21]

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

(式２１)

で示されるA個の画素が、

[数22]

$$C_f^{(m,s)} - l^2 = C_f^{(m,s)} - \frac{1}{\lambda} \qquad \text{(式２２)}$$

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、

[数23]

$$\partial C_f^{(m,s)} = -\frac{A}{\lambda} \qquad (式２３)$$

だけ変化することを示し、その結果、

[数24]

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \qquad (式２４)$$

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

なお、H(h＞0)及びkを定数とするとき、

[0058] [数25]

$$h(l) = Hl^k = \frac{H}{\lambda^{k/2}} \qquad (式２５)$$

と仮定すれば、

[0059] [数26]

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \qquad (式２６)$$

が成り立つ。このときk≠－3であれば、

[数27]

$$C_f^{(m,s)} = C + \frac{H}{(3/2+k/2)\lambda^{3/2+k/2}} \qquad (式２７)$$

となる。これが$C_f^{(m,s)}$の一般式である(Cは定数)。

[0060] λ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、

[0061] [数28]

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \qquad (式２８)$$

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。

[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)＝Hl^kを仮定するとき、例えば、

[0062] [数31]

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

（式３１）

は定数になる。しかしλ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_{0thres}を越えるかどうかを検査し、λ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_{1thres}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。

[0063] なお実験では、副写像$f^{(m,s)}$を計算する際、もしλ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。λ＞0．1のとき、画素の輝度255レベル中のわずか「3」の違いが副写像の計算に影響したためであり、λ＞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。

[1．4．2]ヒストグラムh(l)

[0064] $C^{(m,s)}_f$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(λ，$C^{(m,s)}_f$)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$\lambda^{(1-k)/2}$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$\lambda^{1/2}$である。しかし、こうした差は閾値B_0

を正しく設定することによって吸収することができる。
thres

[0065] 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]

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

(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。

[0066] [数33]

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

(式33)

[0067] 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。

[数34]

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

(式34)

[0068] 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中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27が一般に減少関数であることが保障される。

[0069] なお、式34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2m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2mを導入した。

[1．4．3] ηの動的決定

[0070] パラメータηも同様の方法で自動決定できる。はじめにη＝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、ηをある値$\Delta\eta$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(

$^{n)}_{f}$を計算し直す。この過程を最適値が求まるまで続ける。ηは写像の剛性を示す。次式の重みだからである。

[0071] [数35]

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

（式３５）

[0072] ηが0のとき、$D^{(n)}_{f}$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、ηが非常に大きな値のとき、$D^{(n)}_{f}$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。ηの値が0から次第に増えるとき、後述のごとく$C^{(n)}_{f}$は徐々に減少する。しかしη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η、Y軸はC_fである。

[0073] この方法で$C^{(n)}_{f}$を最小にする最適なηの値を得ることができる。しかし、λの場合に比べていろいろな要素が計算に影響する結果、$C^{(n)}_{f}$は小さく揺らぎながら変化する。λ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、η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$C^{(n)}_{f}$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かれば、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。

[1．5]スーパーサンプリング

[0074] 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×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、

[0075] [数36]

$$V\left(q^{(m,s)}_{f^{(m,s)}(i,j)}\right)$$

（式３６）

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、

[数37]

$$V\left(q^{(m,s)}_{(i,j)+(0.5,0.5)}\right) \qquad (式37)$$

は、

[数38]

$$\left(V\left(q^{(m,s)}_{(i,j)}\right)+V\left(q^{(m,s)}_{(i,j)+(1,1)}\right)\right)/2 \qquad (式38)$$

によって与えられた。

[1．6]各画像の画素の輝度の正規化

[0076] 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。

[0077] 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。

[1．7]インプリメンテーション

[0078] 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ずつ増やしながら各$f^{(m,s)}$(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、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。

[0079] まず(s mod 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決め

ていく。(s mod 4)が1の場合、最上行の右端点を開始点とし、iを減少、jを増加させながら決めていく。(s mod 4)が2のとき、最下行の右端点を開始点とし、i及びjを減少させながら決めていく。(s mod 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。

[0080] 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, …, 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ϕを掛け、一方、第1または第2の条件を破る候補にはϕを掛ける。今回はϕ＝2、ϕ＝100000を用いた。

[0081] 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を確かめる。

[数39]

$$W = \vec{A} \times \vec{B} \qquad (式39)$$

ただしここで、

[数40]

$$\vec{A} = \overrightarrow{q^{(m,s)}_{f(m,s)(i,j-1)} q^{(m,s)}_{f(m,s)(i+1,j-1)}} \qquad (式40)$$

[数41]

$$\vec{B} = \overrightarrow{q^{(m,s)}_{f(m,s)(i,j-1)} q^{(m,s)}_{(k,l)}} \qquad (式41)$$

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ϕ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。

[0082] 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。

[1.7.1]副写像の順序

[0083] 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$\sigma(0)=0$、$\sigma(1)=1$、$\sigma(2)=2$、$\sigma(3)=3$、$\sigma(4)=0$を用い、奇数のときは$\sigma(0)=3$、$\sigma(1)=2$、$\sigma(2)=1$、$\sigma(3)=0$、$\sigma(4)=3$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。

[1.8]補間計算

[0084] 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($0 \leqq t \leqq 1$)である中間画像の画素r(x, y, t) ($0 \leqq x \leqq N-1$, $0 \leqq y \leqq M-1$)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。

[0085] [数42]

$$
\begin{aligned}
(x,y) \;=\; & (1-dx)(1-dy)(1-t)(i,j)+(1-dx)(1-dy)tf(i,j) \\
+\; & dx(1-dy)(1-t)(i+1,j)+dx(1-dy)tf(i+1,j) \\
+\; & (1-dx)dy(1-t)(i,j+1)+(1-dx)dytf(i,j+1) \\
+\; & dxdy(1-t)(i+1,j+1)+dxdytf(i+1,j+1)
\end{aligned}
$$

(式42)

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。

[0086] [数43]

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

（式４３）

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。

[1．9]拘束条件を課したときの写像

[0087] 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。

[0088] 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。

[0089] 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。

[0090] 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44]

$$p_{(i_0,j_0)}, p_{(i_1,j_1)}, \cdots, \qquad p_{(i_{n_s-1},j_{n_s-1})}$$

（式４４）

を特定するとき、以下の値を決める。

[数45]

$$F^{(n)}(i_0, j_0) = (k_0, l_0),$$
$$F^{(n)}(i_1, j_0) = (k_1, l_1), \ldots,$$

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

（式４５）

[0091] 始点画像の他の画素の変位量は、$p_{(ih, jh)}$ (h＝0, …, 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。

[0092] [数46]

$$F^{(m)}(i,j) = \frac{(i,j) + \sum_{h=0}^{h=n_s-1}(k_h - i_h, l_h - j_h) weight_h(i,j)}{2^{n-m}}$$

（式４６）

ただしここで、

[数47]

$$weight_h(i,j) = \frac{1/||(i_h - i, j_h - j)||^2}{total\ weight(i,j)}$$

（式４７）

[数48]

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

（式４８）

とする。

[0093] 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49]

$$D^{(m,s)}_{(i,j)} = E^{(m,s)}_{0_{(i,j)}} + \eta E^{(m,s)}_{1_{(i,j)}} + \kappa E^{(m,s)}_{2_{(i,j)}}$$

（式４９）

である。ただし、

[数50]

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

（式５０）

であり、κ，$\rho \geqq 0$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。

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、

[0094] [数51]

$$\left[\frac{\rho^2}{2^{2(n-m)}}\right]$$

（式５１）

以内であるとき、$E_{2_{(i,j)}}^{(m,s)}$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。

[0095] 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(S1)、つづいて始点画像と終点画像のマッチングをとる(S2)。ただし、S2は必須ではなく、S1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。

[0096] 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。

[0097] 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。

[0098] 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$p^{(m,0)}$～$p^{(m,3)}$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$p^{(m,0)}$であると考える。[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。

[0099] 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1×1である。

[0100] 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。

[0101] 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$C_f^{(m,s)}$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$D_f^{(m,s)}$がそれである。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$\lambda C_f^{(m,s)}+D_f^{(m,s)}$がそれであり、[1．3．2．2]で導入した$\eta$を用いれば、

[0102] [数52]

$$\Sigma\Sigma\left(\lambda C^{(m,s)}_{(i,j)}+\eta E_0^{(m,s)}_{(i,j)}+E_1^{(m,s)}_{(i,j)}\right) \qquad (式52)$$

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

[0103] 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とられる。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。

[0104] 図12のごとく、まず係数パラメータη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も、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。

[0105] 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$f^{(m,3)}$まで求まった後、これを拘束条件として$f^{(m,0)}$を一回更新するという手続をとった。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。

[0106] 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。

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。

[0107] 2．点a～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～dはそれぞれ画素A～Dに属する。ただし、画素A～Cは本来存在しない仮想的な画素である。

[0108] 3．第0レベルですでに求まっている画素A～Dの対応点A'～D'を$q^{(1,s)}$の中にプロットする。画素A'～C'は仮想的な画素であり、それぞれ画素A～Cと同じ位置にあるものとする。

[0109] 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'～d'をプロットし、点a'～d'で相続四辺形を作る。

[0110] 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。

[0111] 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'～D'の間隔が空いていく状況が発生する。

[0112] 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$\eta=0$に関して定まったものであるから、$f^{(n)}(\eta=0)$と書く。

[0113] つぎに異なるηに関する写像も求めるべく、ηを$\Delta\eta$だけシフトし、mをゼロクリアする(S24)。新たなηが所定の探索打切り値η_{max}を超えていないことを確認し(S25)、S21に戻り、今回のηに関して写像$f^{(n)}(\eta=\Delta\eta)$を求める。この処理を繰り返し、S21で$f^{(n)}(\eta=i\Delta\eta)$(i=0, 1, …)を求めていく。$\eta$が$\eta_{max}$を超えたときS26に進み、後述の方法で最適な$\eta=\eta_{opt}$を決定し、$f^{(n)}(\eta=\eta_{opt})$を最終的に写像$f^{(n)}$とする。

[0114] 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η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λを独立して決める。

[0115] 同図のごとく、まずsとλをゼロクリアする(S210)。つぎに、そのときのλについて(および暗にη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$f^{(m,s)}(\lambda=0)$と書く。異なるλに関する写像も求めるべく、λを$\Delta\lambda$だけシフトし、新たなλが所定の探索打切り値λ_{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$f^{(m,s)}(\lambda=i\Delta\lambda)$(i=0, 1, …)を求める。$\lambda$が$\lambda_{max}$を超えたときS214に進み、最適な$\lambda=\lambda_{opt}$を決定し、$f^{(m,s)}(\lambda=\lambda_{opt})$を最終的に写像$f^{(m,s)}$とする(S214)。

[0116] つぎに、同一レベルにおける他の副写像を求めるべく、λ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。

[0117] 図16は、あるmとsについてλを変えながら求められた$f^{(m,s)}(\lambda=i\Delta\lambda)$(i=0, 1, …)に対応するエネルギー$C_f^{(m,s)}$の挙動を示す図である。[1. 4]で述べたとおり、λが増加すると通常$C_f^{(m,s)}$は減少する。しかし、λが最適値を超えると$C_f^{(m,s)}$は増加に

転じる。そこで本前提技術では$C_f^{(m,s)}$が極小値をとるときのλをλ_{opt}と決める。同図のように$\lambda > \lambda_{opt}$の範囲で再度$C_f^{(m,s)}$が小さくなっていっても、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。λ_{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。

[0118] 一方、図17は、ηを変えながら求められた$f^{(n)}(\eta = i\Delta\eta)$(i=0, 1, …)に対応するエネルギー$C_f^{(n)}$の挙動を示す図である。ここでも$\eta$が増加すると通常$C_f^{(n)}$は減少するが、$\eta$が最適値を超えると$C_f^{(n)}$は増加に転じる。そこで$C_f^{(n)}$が極小値をとるときの$\eta$を$\eta_{opt}$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。η_{opt}が決まれば$f^{(n)}$を最終決定することができる。

[0119] 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。

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。

[0120] 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$E_{tot} = \alpha E_0 + E_1$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αを自動決定する。つまり、いろいろなαについてE_{tot}が最小になるような写像を求める。それらの写像のうち、αに関してE_1が極小値をとるときのα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。

[0121] 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。

[0122] パラメータも、αのみ、前提技術のごとくηとλ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。

(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$C_f^{(m,s)}$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$\alpha E_0 + \beta E_1$を総合評価式とし、$\alpha + \beta = 1$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。

(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3×3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。

[3]前提技術の改良点

[0123] 以上の前提技術を基本とし、マッチング精度を向上させるためのいくつかの改良がなされている。ここではその改良点を述べる。

[3. 1]色情報を考慮に入れた特異点フィルタおよび副画像

[0124] 画像の色情報を有効に用いるために、特異点フィルタを以下のように変更した。まず色空間としては、人間の直感に最も合致するといわれているHISを用いた。但し色

を輝度に変換する際は、輝度Iに代わり人間の目の感度に最も近いといわれている輝度Yを選択した。

[0125] [数53]

$$H=\frac{\frac{\pi}{2}-\tan^{-1}\left(\frac{2R-G-R}{\sqrt{3(G-B)}}\right)}{2\pi}$$
$$I=\frac{R+G+B}{3}$$
$$S=1-\frac{\min(R,G,B)}{3}$$
$$Y=0.299\times R+0.587\times G+0.114\times B \qquad \text{（式５３）}$$

[0126] ここで画素aにおけるY（輝度）をY（a）、S（彩度）をS（a）として、次のような記号を定義する。

[数54]

$$\alpha_Y(a,b)=\begin{cases}a\cdots(Y(a)\le Y(b))\\ b\cdots(Y(a)>Y(b))\end{cases}$$
$$\beta_Y(a,b)=\begin{cases}a\cdots(Y(a)\ge Y(b))\\ b\cdots(Y(a)<Y(b))\end{cases}$$
$$\beta_S(a,b)=\begin{cases}a\cdots(S(a)\ge S(b))\\ b\cdots(S(a)<S(b))\end{cases} \qquad \text{（式５４）}$$

上の定義を用いて以下のような5つのフィルタを用意する。

[0127] [数55]

$$p_{(i,j)}^{(m,0)}=\beta_Y(\beta_Y(p_{(2i,2j)}^{(m+1,0)},p_{(2i,2j+1)}^{(m+1,0)}),\beta_Y(p_{(2i+1,2j)}^{(m+1,0)},p_{(2i+1,2j+1)}^{(m+1,0)}))$$
$$p_{(i,j)}^{(m,1)}=\alpha_Y(\beta_Y(p_{(2i,2j)}^{(m+1,1)},p_{(2i,2j+1)}^{(m+1,1)}),\beta_Y(p_{(2i+1,2j)}^{(m+1,1)},p_{(2i+1,2j+1)}^{(m+1,1)}))$$
$$p_{(i,j)}^{(m,2)}=\beta_Y(\alpha_Y(p_{(2i,2j)}^{(m+1,2)},p_{(2i,2j+1)}^{(m+1,2)}),\alpha_Y(p_{(2i+1,2j)}^{(m+1,2)},p_{(2i+1,2j+1)}^{(m+1,2)}))$$
$$p_{(i,j)}^{(m,3)}=\alpha_Y(\alpha_Y(p_{(2i,2j)}^{(m+1,3)},p_{(2i,2j+1)}^{(m+1,3)}),\alpha_Y(p_{(2i+1,2j)}^{(m+1,3)},p_{(2i+1,2j+1)}^{(m+1,3)}))$$
$$p_{(i,j)}^{(m,4)}=\beta_S(\beta_S(p_{(2i,2j)}^{(m+1,4)},p_{(2i,2j+1)}^{(m+1,4)}),\beta_S(p_{(2i+1,2j)}^{(m+1,4)},p_{(2i+1,2j+1)}^{(m+1,4)}))$$

（式５５）

[0128] このうち上から4つのフィルタは改良前の前提技術におけるフィルタとほぼ同じで、輝度の特異点を色情報も残しながら保存する。最後のフィルタは色の彩度の特異点

をこちらも色情報を残しながら保存する。

[0129] これらのフィルタによって、各レベルにつき5種類の副画像（サブイメージ）が生成される。なお、最も高いレベルの副画像は元画像に一致する。

[0130] [数56]

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)}^{(n,4)} = p_{(i,j)}$$

（式５６）

[3. 2]エッジ画像およびその副画像

[0131] 輝度微分（エッジ）の情報をマッチングに利用するため、さらに一次微分エッジ検出フィルタを用いる。このフィルタはあるオペレータGとの畳み込み積分で実現できる。第nレベルの画像の、水平方向、垂直方向の微分に対応した2種類のフィルタをそれぞれ以下のように表す。

[数57]

$$p_{(i,j)}^{(n,h)} = Y(p_{(i,j)}) \otimes G_h$$
$$p_{(i,j)}^{(n,v)} = Y(p_{(i,j)}) \otimes G_v$$

（式５７）

[0132] ここでGは画像解析においてエッジ検出に用いられる一般的なオペレータを適用することが可能であるが、演算スピードなども考慮して以下のようなオペレータを選択した。

[0133] [数58]

$$G_h = \frac{1}{4}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$
$$G_v = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

（式５８）

[0134] 次にこの画像を多重解像度化する。フィルタにより0を中心とした輝度をもつ画像が生成されるため、次のような平均値画像が副画像としては最も適切である。

[数59]

$$p_{(i,j)}^{(m,h)} = \frac{1}{4}(p_{(2i,2j)}^{(m+1,h)} + p_{(2i,2j+1)}^{(m+1,h)} + p_{(2i+1,2j)}^{(m+1,h)} + p_{(2i+1,2j+1)}^{(m+1,h)})$$

$$p_{(i,j)}^{(m,v)} = \frac{1}{4}(p_{(2i,2j)}^{(m+1,v)} + p_{(2i,2j+1)}^{(m+1,v)} + p_{(2i+1,2j)}^{(m+1,v)} + p_{(2i+1,2j+1)}^{(m+1,v)})$$

（式５９）

[0135] 式59の画像は後述するForward Stage、すなわち初回副写像導出ステージの計算の際、エネルギー関数のうち新たに導入された輝度微分(エッジ)の差によるエネルギーに用いられる。

エッジの大きさ、すなわち絶対値も計算に必要なため、以下のように表す。

[0136] [数60]

$$p_{(i,j)}^{(n,e)} = \sqrt{(p_{(i,j)}^{(n,h)})^2 + (p_{(i,j)}^{(n,v)})^2}$$

（式６０）

この値は常に正であるため、多重解像度化には最大値フィルタを用いる。

[0137] [数61]

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

（式６１）

式61の画像は後述するForward Stageの計算の際、計算する順序を決定するのに用いられる。

[3．3]計算処理手順

[0138] 計算は最も粗い解像度の副画像から順に行う。副画像は5つあるため、各レベルの解像度において計算は複数回行われる。これをターンと呼び、最大計算回数をtで表すことにする。各ターンは前記Forward Stageと、副写像再計算ステージであるRefinement Stageという二つのエネルギー最小化計算から構成される。図18は第mレベルにおける副写像を決める計算のうち改良点に係るフローチャートである。

[0139] 同図のごとく、sをゼロクリアする（S40）。つぎにForward Stage（S41）において始点画像pから終点画像qへの写像$f^{(m,s)}$および、終点画像qから始点画像pへの写像$g^{(m,}$

$^{s)}$を順次、エネルギー最小化によって求める。以下、写像$f^{(m,s)}$の導出について記述する。ここで最小化するエネルギーは、改良後の前提技術においては、対応する画素値によるエネルギーCと、写像の滑らかさによるエネルギーDの和である。

[0140] [数62]

$$\min_f \left(C^f(i,j)+D^f(i,j)\right) \qquad \text{（式６２）}$$

[0141] エネルギーCは、輝度の差によるエネルギーC_I（前記改良前の前提技術におけるエネルギーCと等価）と、色相、彩度によるエネルギーC_C、輝度微分（エッジ）の差によるエネルギーC_Eで構成され、以下のように表される。

[0142] [数63]

$$C_I^f(i,j)=\left|Y(p_{(i,j)}^{(m,s)})-Y(q_{f(i,j)}^{(m,s)})\right|^2$$

$$C_C^f(i,j)=\left|S(p_{(i,j)}^{(m,s)})\cos(2\pi H(p_{(i,j)}^{(m,s)}))-S(q_{f(i,j)}^{(m,s)})\cos(2\pi H(q_{f(i,j)}^{(m,s)}))\right|^2$$
$$+\left|S(p_{(i,j)}^{(m,s)})\sin(2\pi H(p_{(i,j)}^{(m,s)}))-S(q_{f(i,j)}^{(m,s)})\sin(2\pi H(q_{f(i,j)}^{(m,s)}))\right|^2$$

$$C_E^f(i,j)=\left|p_{(i,j)}^{(m,h)}-q_{f(i,j)}^{(m,h)}\right|^2+\left|p_{(i,j)}^{(m,v)}-q_{f(i,j)}^{(m,v)}\right|^2$$

$$C^f(i,j)=\lambda C_I^f(i,j)+\psi C_C^f(i,j)+\theta C_E^f(i,j) \qquad \text{（式６３）}$$

ここでパラメータλ、φおよびθは0以上の実数であり、本改良後の技術においては定数である。ここでこれらのパラメータを定数とできるのは、新たに導入されたRefinement Stageにより、パラメータに対する結果の安定性が向上したためである。また、エネルギーC_Eは副写像$f^{(m,s)}$の種類sに関わらず、座標と解像度のレベルによって決定する値である。

[0143] エネルギーDは前記改良前の前提技術と同じものを用いる。ただし前記改良前の前提技術において、写像の滑らかさを保証するエネルギーE_1を導出する際、隣接する画素のみを考慮していたが、周囲の何画素を考慮するかをパラメータdで指定できるように改良した。

[0144] [数64]

$$E_0^f(i,j) = \left\| f(i,j) - (i,j) \right\|^2$$
$$E_1^f(i,j) = \sum_{i'=i-d}^{i+d} \sum_{j'=j-d}^{j+d} \left\| \left(f(i,j) - (i,j) \right) - \left(f(i',j') - (i',j') \right) \right\|^2 \quad \text{（式６４）}$$

[0145] 次のRefinement Stageに備えて、このステージでは終点画像qから始点画像pへの写像$g^{(m,s)}$も同様に計算する。

[0146] Refinement Stage（S42）ではForward Stageにおいて求めた双方向の写像$f^{(m,s)}$および$g^{(m,s)}$を基に、より妥当な写像$f'^{(m,s)}$を求める。ここでは新たに定義されるエネルギーMについてエネルギー最小化計算を行う。エネルギーMは終点画像から始点画像への写像gとの整合度M_0と、もとの写像との差M_1より構成され、Mを最小とするような$f'^{(m,s)}$が求められる。

[0147] [数65]

$$M_0^{f'}(i,j) = \left\| g(f'(i,j)) - (i,j) \right\|^2$$
$$M_1^{f'}(i,j) = \left\| f'(i,j) - f(i,j) \right\|^2$$
$$M^{f'}(i,j) = M_0^{f'}(i,j) + M_1^{f'}(i,j) \quad \text{（式６５）}$$

[0148] 対称性を損なわないように、終点画像qから始点画像pへの写像$g'^{(m,s)}$も同様の方法で求めておく。

その後、sをインクリメントし（S43）、sがtを超えていないことを確認し（S44）、次のターンのForward Stage（S41）に進む。その際前記E_0を次のように置き換えてエネルギー最小化計算を行う。

[0149] [数66]

$$E_0^f(i,j) = \left\| f(i,j) - f'(i,j) \right\|^2 \quad \text{（式６６）}$$

[3. 4]写像の計算順序

[0150] 写像の滑らかさを表すエネルギーE_1を計算する際、周囲の点の写像を用いるため、それらの点がすでに計算されているかどうかがエネルギーに影響を与える。すなわち、どの点から順番に計算するかによって、全体の写像の精度が大きく変化する。そこでエッジの絶対値画像を用いる。エッジの部分は情報量を多く含むため、エッジの

絶対値が大きいところから先に写像計算を行う。このことによって、特に二値画像のような画像に対して非常に精度の高い写像を求めることができるようになった。

[動画符号化と復号に関する実施の形態]

以上の前提技術を一部利用した動画処理の具体例を述べる。

[0151] 図19は、動画の符号化装置と復号装置の構成および処理を示す。同図上段が符号化装置、下段が復号装置に関する。

[1]符号化装置の構成

[0152] CPF: 前提技術のCritical Point Filter、すなわち特異点フィルタを用いる画像マッチングプロセッサ。キーフレーム間のマッチングを画素単位で計算し、対応点情報を出力する。この情報はファイルとして出力される。このファイルは、ソース側のキーフレームの各画素がデスティネーション側のキーフレームのいずれの画素に対応するかを記述する。したがって、このファイルをもとに、これらのキーフレーム間で対応しあう画素の位置と画素値を内挿計算すれば、ふたつのキーフレーム間のモーフィング画像が得られる。なお、このファイルをソース側のキーフレームだけに作用させて内挿計算をすれば、単にソース側のキーフレームの各画素をこのファイルに記述した対応画素の位置へ徐々に移動させるモーフィング画像が得られる。この場合、対応画素間で位置だけが内挿されたことになる。

[0153] なお、CPFの代わりに、広く画像マッチングプロセッサを利用することができるが、本実施の形態の趣旨からいえば、精度が高い画素マッチングが理想的であり、前提技術はその条件を満たす。

[0154] DE: Differential Encoder 差分(誤差)符号化器。ふたつの画像フレーム間の差分をハフマン符号化その他の統計手法に基づき可変長符号化する。

[0155] NR: maskable Noise Reducer ノイズリデューサ。人間の視覚では微細な変化を認識できないことが多い。たとえば輝度の変化の激しい部分、つまり輝度の空間周波数成分が高い成分が強い領域では、輝度変化の誤差は視覚的には把握されない。動画情報にはさまざまな形でノイズが重畳しており、そのようなデータは視覚的には単にノイズとして認識されるだけで画像としての意味を持たない。そのような視覚的無意味なランダム情報、すなわち「視覚的マスク情報」を無視することが、より高い圧縮率

を達成するために重要である。

[0156] 現在のブロックマッチングにおける量子化は、輝度値に関する視覚的マスク情報を利用したものであるが、輝度値以外にもいくつかの視覚的マスク情報が存在する。NRは、空間位置情報ならびに時間位置情報に関する視覚的マスクを利用する。空間位置情報の視覚的マスクは、位置情報に関して、輝度変化が複雑な画像の場合は空間周波数の位相成分が視覚的に認識されにくいという事実を利用する。時間位置情報の視覚的マスクは、時間方向での変化が激しい部分では時間方向にデータの変化がずれたとしても、視覚的にはその差が認識されにくい事実を利用する。これらはいずれも所定のしきい値との比較して検出する。

[0157] 少なくともブロックマッチングと差分符号化という現在のMPEGのスキームでは、これらのマスクを積極的に利用することは困難である。これに対し、前提技術における復号処理は、視覚的な不自然さをもたらすような不連続性を回避するために、動画上の変化をトリリニアその他の補間で生成するものであり、それは誤差を輝度方向だけでなく、空間方向や時間方向に散らして視覚的に目立たなくする働きを持つ。NRは前提技術との組合せにおいて有用である。

[0158] DD: Differential Decoder 差分(誤差)復号器。DEで符号化された差分を復号し、その差分が生じた画像フレームに加算することで、その画像フレームの精度を高める。

[0159] なお、これらのほかに、ある単一のキーフレームに対応点情報を作用させ、そのキーフレームの画素移動だけから仮想的に別のキーフレームを生成する機能が存在する。以下、この機能を実現する機能ブロックを画素シフタとよぶ。

[2]符号化処理

[0160] 図19において、「F0」等は処理の対象となる動画の各フレーム、「M0－4」はCPFによって生成されたF0とF4間の対応点情報を示す。符号化は以下の手順で進む。

[0161] a) 1以上の画像フレーム(F1～F3)を間に挟む第1、第2キーフレーム(F0、F4)間でCPFによってマッチングを計算し、第1、第2キーフレーム間の対応点情報(M0－4)を生成するステップ。

b) 第1、第2キーフレーム間の対応点情報(M0－4)をもとに、画素シフタによって

第1キーフレーム(F0)に含まれる画素を移動させて仮想の第2キーフレーム(F4')を生成するステップ。

c) 現実の第2キーフレーム(F4)と仮想の第2キーフレーム(F4')との差分をNR機能付きDE(DE+NRと表記)で圧縮符号化するステップ。

d) 第1キーフレーム(F0)、第1、第2キーフレーム間の対応点情報(M0-4)、および、現実の第2キーフレームと仮想の第2キーフレーム間で圧縮符号化された差分(Δ4)をこれらのキーフレーム(F0、F4)間の符号化データとして出力するステップ。出力先は記録媒体、伝送媒体を問わない。実際には後述のj)で出力される情報と一体となり、動画符号化データとして記録媒体等に出力される。

[0162] つづいて、第2キーフレーム(F4)以降について以下の処理を行う。

e) 現実の第2キーフレーム(F4)と仮想の第2キーフレーム(F4')間で圧縮符号化された差分(Δ4)をDDで復号するステップ。

f) 復号された差分と前記仮想の第2キーフレーム(F4')とから、改良された仮想の第2キーフレーム(F4")をDDで生成するステップ。

g) 1以上の画像フレーム(F5～F7)を間に挟む第2、第3キーフレーム(F4、F8)間でCPFによってマッチングを計算し、第2、第3キーフレーム間の対応点情報(M4-8)を生成するステップ。

h) 第2、第3キーフレーム間の対応点情報(M4-8)をもとに、画素シフタによって、改良された仮想の第2キーフレーム(F4")に含まれる画素を移動させることによって、仮想の第3キーフレーム(F8')を生成するステップ。

i) 現実の第3キーフレーム(F8)と仮想の第3キーフレーム(F8')との差分をDE+NRで圧縮符号化するステップ。

j) 第2、第3キーフレーム間の対応点情報(M4-8)、および現実の第3キーフレームと仮想の第3キーフレーム間で圧縮符号化された差分(Δ8)をこれらのキーフレーム(F4、F8)間の符号化データとして出力するステップ。出力先は一般にd)の出力先と同じである。

[0163] 以下、さらに後続のキーフレームについて、図19のフレームF9以下に示すごとく、順次前記のe)からj)のステップを繰り返し、所定のグループ終了キーフレームに到達

したときに繰り返し処理を終了する。グループ終了キーフレームは、MPEGでいう1GOPの終了フレームに相当する。したがって、このフレームの次のフレームが新たなグループの先頭フレームとして新たに第1キーフレームと見なされ、a)以下の処理が繰り返される。以上の処理により、MPEGでいうGOPに相当するグループ(以下、単にグループとよぶ)について、キーフレーム(MPEGでいうIピクチャ)に相当する画像は1枚のみ符号化および伝送すればよい。

[3]復号装置の構成

符号化側にもましてシンプルな構成である。

DD: 符号化装置のDDと同じ。

INT: INTerpolator 補間プロセッサ。

[0164] これらの他に符号化側同様の画素シフタが存在する。ふたつの画像フレームと対応点情報から内挿処理による中間フレームを生成する。

[4]復号処理

復号は以下の手順で進む。

[0165] k) 1以上の画像フレーム(F1～F3)を間に挟む第1、第2キーフレーム(F0、F4)間の対応点情報(M0－4)、および第1キーフレーム(F0)を取得するステップ。取得は伝送媒体、記録媒体のいずれからでもよい。

l) 第1、第2キーフレーム間の対応点情報(M0－4)をもとに、画像シフタによって第1キーフレーム(F0)に含まれる画素を移動させることによって、仮想の第2キーフレーム(F4')を生成するステップ。

m) 予め符号化側にてl)同様の処理により、仮想の第2キーフレーム(F4')が生成され、符号化側でこれと現実の第2キーフレーム(F4)との差分の圧縮符号化データ(Δ4)を生成しているため、これを取得するステップ。

o) 取得された差分の圧縮符号化データ(Δ4)をDDで復号し、仮想の第2キーフレーム(F4')と加算して、改良された仮想の第2キーフレーム(F4'')を生成するステップ。

p) 第1、第2キーフレーム間の対応点情報(M0－4)をもとに、INTによって、第1キーフレーム(F0)と改良された仮想の第2キーフレーム(F4")間で補間計算をする

ことにより、これらのキーフレーム(F0、F4")の間に存在すべき中間フレーム(F1"～F3")を生成するステップ。

q) 第1キーフレーム(F0)、生成された中間フレーム(F1"～F3")、改良された仮想の第2キーフレーム(F4")をこれらのキーフレーム間の復号データとして表示装置等へ出力するステップ。

[0166] つづいて、第2キーフレーム(F4)以降について以下の処理を行う。

r) 1以上の画像フレーム(F5～F7)を間に挟む第2、第3キーフレーム(F4、F8)間の対応点情報(M4－8)を取得するステップ。

s) 第2、第3キーフレーム間の対応点情報(M4－8)をもとに、画素シフタによって、改良された仮想の第2キーフレーム(F4")に含まれる画素を移動させることによって、仮想の第3キーフレーム(F8')を生成するステップ。

t) 予め符号化側にてs)同様の処理により、符号化側でも仮想の第3キーフレーム(F8')が生成され、符号化側でこれと現実の第3キーフレーム(F8)との差分の圧縮符号化データ(Δ8)を生成しており、これを取得するステップ。

u) 取得された差分の圧縮符号化データ(Δ8)と仮想の第3キーフレーム(F8')とから、DDによって、改良された仮想の第3キーフレーム(F8")を生成するステップ。

v) 第2、第3キーフレーム間の対応点情報(M4－8)をもとに、INTによって、改良された仮想の第2キーフレーム(F4")と改良された仮想の第3キーフレーム(F8")間で補間計算をすることにより、これらのキーフレームの間に存在すべき中間フレーム(F5'～F7')を生成するステップ。

w) 改良された仮想の第2キーフレーム(F4")、生成された中間フレーム(F5'～F7')、改良された仮想の第3キーフレーム(F8")をこれらのキーフレーム(F4"、F8")間の復号データとして表示装置などへ出力するステップ。

[0167] 以下、さらに後続のキーフレームについて、図19のフレームF9以降に示すごとく、順次前記のr)からw)のステップを繰り返し、グループ終了キーフレームに到達したときに繰り返し処理を終了する。このフレームの次のフレームが新たなグループの先頭フレームとして新たに第1キーフレームと見なされ、k)以下の処理が繰り返される。

[5]本実施の形態によるメリット

[0168] 画像マッチングに前提技術のCPFを利用する場合、マッチング精度が高いため、本実施の形態で実現される圧縮率が高くなる。なぜなら、DE＋NRによって圧縮すべき差分が最初から小さく、かつ統計的な偏りが大きくなるためである。

[0169] 同様に、CPFを用いる場合、この符号化方法はブロックマッチングを用いないので、圧縮率を高めてもMPEGで問題となるブロックノイズがでない。もちろん、CPF以外の画像マッチングでも、ブロックノイズがでない処理方法を採用すればよい。

[0170] もともとMPEGは差分の最小化しか考慮しないが、CPFは本来対応すべき個所を検出するため、究極的にはMPEGよりも高い圧縮率が実現できる。

[0171] 符号化装置は画像マッチングプロセッサ、ノイズリダクション機能付き差分符号化器、差分復号器、画素シフタで構成でき、簡易である。また、ノイズリダクション機能はオプショナルな機能であり、これはなくともよい。同様に、復号装置も補間プロセッサ、差分復号器、画素シフタで構成でき、簡素である。とくに、復号装置は画像マッチングを行う必要もなく、処理量が軽い。

[0172] 仮想のキーフレームを生成するたびに、それと現実のキーフレームの差分をΔ4、Δ8などのように符号化データへ取り込むため、グループごとに1枚しか完全な形のキーフレームを符号化しないにもかかわらず、長い動画を再生しても誤差の蓄積がない。

[6]変形技術

[0173] 第1、第2キーフレーム(F0、F4)間のマッチング計算をして対応点情報ファイルを生成する際、それらキーフレーム間に存在する中間フレーム(F1～F3)も考慮してもよい(図19の破線矢印)。その場合、CPFはF0とF1、F1とF2、F2とF3、F3とF4のそれぞれの組についてマッチングを計算し、4個のファイル(仮に部分ファイルM0～M3とよぶ)を生成する。つづいて、これら4個のファイルを統合してひとつの対応点情報ファイルとして出力すればよい。

[0174] 統合のために、まず、F0の各画素がM0によってF1上のどこへ移動するかを特定する。つづいて、F1上で特定された画素がM1によってF2上のどこへ移動するかを特定する。これをF4まで行えば、4個の部分ファイルにより、F0とF4の対応がより正確になる。F0とF4は多少距離があり、それらの間よりも隣接する画像フレーム間のマ

ッチング精度のほうが一般に高いためである。

[0175] なお、この方法は最終的にF0とF4のマッチング精度を改善するものであるが、対応点情報ファイルを時間の関数として表現してもよい。その場合、部分ファイルを統合せず、4個の状態のまま、これらを対応点情報ファイルとみなして復号側へ提供すればよい。復号側はF0、F4、M0からF1を生成し、F0、F4、M0、M1からF2を生成し、という繰り返し処理でより正確な動画を復号できる。

[0176] (第2実施形態)

本発明の他の実施形態は、図19の符号化装置に関する。ここでは、画像マッチングの正確性を示す尺度として画像のマッチングエネルギーを導入し、これをDE+NRにおけるノイズリダクション等に利用する。以下、適宜図19を用いて説明するが、特に言及しない構成、機能については第1実施形態と同様である。

[0177] ここでいうマッチングエネルギーとは、対応点どうしの距離と画素値の違いで定まるものであり、例えば前提技術における式49に示されている。本実施形態では、CPFにおける画像マッチングの際得られるこのマッチングエネルギーをいわば副産物として利用する。前提技術の画像マッチングでは、キーフレーム間の各画素につき、写像のエネルギーが最小となるものを対応点として検出する。前提技術のこのような特徴に着目すれば、マッチングエネルギーの低い画素に関しては良好なマッチングがとれており、一方マッチングエネルギーの高い箇所については、当然キーフレーム間で位置や画素値の変化の大きい画素であったはずであるが、場合によってはマッチングエラーがあった可能性もあると評価できる。以下詳説するが、本実施形態ではマッチング精度の高い部分については差分の圧縮率を高める。また別の例では、マッチングエラーが推定される画素に関する差分情報を高く圧縮しても良い。

[0178] [1]符号化処理

本実施形態の符号化装置では、CPFが第1、第2のキーフレームのマッチングを計算する際に、併せて両フレーム間で対応しあう各画素のマッチングエネルギーを取得し、第1のキーフレーム(F0)上に各画素のマッチングエネルギーを記述したエネルギーマップを生成する。同様に、その他の隣接しあうキーフレーム間でもエネルギーマップを生成する。すなわち、エネルギーマップとは、キーフレーム間の対応点そ

れぞれのマッチングエネルギーを、基本的には前のキーフレームの各画素に関して記述したデータである。なお、エネルギーマップは前後のキーフレームのうち、後のキーフレーム上に表しても良い。エネルギーマップは不図示の経路によりCPFからDE＋NRに送られる。DE＋NRでは、このエネルギーマップを利用してキーフレーム間のマッチングの良否を評価し、それに基づいて、仮想のキーフレームと現実のキーフレームの差分を適応的に圧縮符号化する。なお、DE＋NRには、エネルギーマップの他、対応点情報ファイルも不図示の経路で送られている。

[0179] 図20は、本実施形態に係る図19のDE＋NRの構成を示す図である。図20のDE＋NRは差分計算器10と、差分圧縮部12と、エネルギー取得部14と、判定部16とを備える。このうち、前2者が専らDEに相当し、後2者が専らNRに相当する。以下第1のキーフレーム(F0)と第2のキーフレーム(F4)およびその中間の画像フレーム(F1～F3)を符号化する際のDE＋NRの動作を説明するが、後続の各キーフレーム、画像フレームの符号化においても、DE＋NRの動作は同様である。

[0180] 差分計算器10は、現実の第2キーフレーム(F4)と仮想の第2キーフレーム(F4')を取得して、位置的に対応しあう画素どうしの画素値の差分をとる。これにより、各画素が両キーフレーム間の画素値の差をもつ一種の画像が形成され、これを差分画像と呼ぶ。差分画像はエネルギー取得部14へと送られる。また、エネルギー取得部14には、現実の第1キーフレーム(F0)と現実の第2キーフレーム(F4)の間のエネルギーマップ及び対応点情報(M0－4)が、図19のCPFから入力される。エネルギー取得部14は、これらを利用して差分画像のマッチングエネルギーを取得する。

[0181] まず、取得部14は、第1、第2キーフレーム間の対応点情報(M0－4)をCPFから取得する。これを利用して、差分画像から仮想の第2キーフレーム(F4')、第1キーフレーム(F0)とたどっていくことで、差分画像のどの画素が第1キーフレーム(F0)のどの画素をシフトしたものに対応しているか、対応関係を取得する。その上で第1キーフレーム上に表されたエネルギーマップ上の各画素のエネルギーを参照し、差分画像の各画素に対応する第1キーフレーム(F0)上の画素のマッチングエネルギーを、差分画像の各画素のマッチングエネルギーとして取得する。差分画像のマッチングエネルギーはこうして求められる。

[0182] エネルギー取得部14は、差分画像のマッチングエネルギーを判定部16へと送る。判定部16は差分画像の各画素のマッチングエネルギーを利用して、差分画像のうち高圧縮対象領域を判定し、いずれの領域を高圧縮すべきかの情報を圧縮部12へと通知する。判定は例えば以下のように行われる。判定部16は、差分画像を16X16画素単位のブロックに分割し、各ブロックに含まれる画素の全てについてマッチングエネルギーを所定のしきい値と比較する。比較の結果、ブロック内の全ての画素のマッチングエネルギーがしきい値以下であった場合は、その領域を高圧縮対象ブロックと判定する。

[0183] 圧縮部12は、差分画像をJPEG形式にて圧縮する。この際、判定部16から通知された高圧縮対応領域の情報を利用し、圧縮率を通常の領域と高圧縮対応領域との間で適応的に変化させる。具体的には、高圧縮対象ブロックはDCT係数の量子化幅を通常のブロックに比べて大きくする処理などが利用できる。別の例では、差分画像では、高圧縮対象ブロックの画素値を0にしてしまってからJPEG圧縮をかける処理を行ってもよい。いずれにせよ、マッチングエネルギーが低い領域を高圧縮する理由は以下の考え方による。

[0184] すなわち、上述のごとくマッチングエネルギーの低い画素は、キーフレーム間のマッチング結果が良好であるとみなせる。従って、差分画像のうちマッチングエネルギーが低い部分に関しては、現実の第2のキーフレーム(F4)と仮想の第2のキーフレーム(F4')の間に差分は本来生じにくく、差分が生じているとすればそれはノイズであると考えてよい。よって、差分画像においてマッチングエネルギーが低い領域は、高圧縮による情報の欠落を気にすることなく、他の領域に比べて大幅に圧縮できる。一方、マッチングエネルギーの大きい領域については、マッチングにエラーが生じている可能性もあり、仮想の第2キーフレーム(F4')と現実の第2キーフレーム(F4)の差分は復号において重要な情報であるため、圧縮率を低くとどめ、復号時の正確性を優先する。

[0185] [2]第2実施形態によるメリット

以上の処理を経て、圧縮部18は、現実の第2キーフレーム(F4)と仮想の第2キーフレーム(F4')の圧縮符号化された差分(Δ4)を出力する。本実施形態による符号

化装置によれば、現実のキーフレームと仮想のキーフレームの差分情報を、符号化画像をより原画像に忠実に、正確な復号を行うための重要性に応じて適応的に圧縮可能であり、復号の正確性を保ちつつ高い符号化効率が実現できる。重要性とは、もちろん、本実施形態でも第1実施形態に係るメリットを享受できる。

[0186] [3]第2実施形態の変形技術

本実施形態の変形例として、マッチングエネルギーの大きい画素、中でも近傍の画素の対応傾向と著しく異なる対応傾向を示す画素はマッチングエラーを起こしている場合が多いと経験的に認められることから、マッチングエネルギーが周囲の画素と比べ大幅に異なる画素をマッチングエラーと評価し、これをノイズリダクションに導入することもできる。この場合、DE+NRは、第2キーフレーム(F4)の各画素のマッチングエネルギーを、例えば自身を中心とする9X9画素のブロック内の、他の画素のマッチングエネルギーの平均と比較する。比較の結果両者の差が所定のしきい値を超えている場合、そのような画素はマッチングエラーをおこしていると判定してもよい。

[0187] エラーを起こしている対応情報は復号側にとって無意味なデータであると考えることができ、現実の第2キーフレーム(F4)と仮想の第2キーフレーム(F4')間の差分情報中では、マッチングエラーを起こしている画素に関するデータはノイズといえる。よって、高圧縮による情報の欠落への配慮を不要とし、DE+NRは、現実のキーフレームと仮想のキーフレーム間の差分画像のうち、現実のキーフレーム間のマッチングエラーに対応する画素を他の画素に比べて高い率で圧縮する。なお、マッチングエラーの判定は、例えば、周囲の画素の動きベクトルの傾向と、注目する画素の動きベクトルの傾向を比較し、注目する画素の動きベクトルが周囲の傾向と著しく異なるか否かをもって行なっても良い。

[0188] 第2実施形態においても、第1実施形態と同様に、第1、第2キーフレーム(F0、F4)間の中間フレーム(F1～F3)を考慮し、これら全ての画像フレームの隣り合うそれぞれの組についてマッチングを計算して対応点情報ファイル(M0～M3)を生成し、それらを統合して第1、第2キーフレーム(F0、F1)間で一つの対応点情報ファイルを得る変形技術が考えられる。第1実施形態の変形技術同様、マッチング精度を向上し、正確な動画復号が実現できる。

[0189] さらに、この変形技術では、各画像フレーム間のマッチングエネルギーを計算してそれをシーンチェンジ検出等に応用可能である。シーンチェンジ検出に係る構成は以下のとおりである。まず、CPFはF0とF1、F1とF2、F2とF3, F3とF4・・・それぞれの組について、マッチング計算をおこない、その副産物としてエネルギーマップ、E0、E1、E2、E3・・・を取得する。ここで、ある画像フレーム全体の画素に係るマッチングエネルギーの平均をとり、それを所定のシーンチェンジ検出用しきい値と比較し、その直後の画像を新たなグループとすればよい。例えばF5とF6の間のエネルギーマップE5に基づき、F5とF6のマッチングに係るF5の各画素のマッチングエネルギーを平均した結果、その値がキーフレーム追加用しきい値を越えたとする。この場合、直後のキーフレームすなわちF6以下を新たなグループとし、F6が次のグループの第1キーフレームとすればよい。マッチングエネルギーが大きい場合、画像間に大きな変化があったと考えることができるためである。これにより、自動的なシーンチェンジの検出ができ、シーンチェンジに対応してグループの選定が可能となる。

[0190] 各エネルギーマップに基づいて、各画像フレーム内画素の平均マッチングエネルギーを計算して、これを累積的に加算していき、その値が所定のしきい値を越えた時点でその画像フレームを新たにキーフレームとして登録しても良い。画像フレーム間の変化量の累積がある一定値を越えた時点でキーフレームを追加できれば、より復号時の画質の向上がはかれるためである。

産業上の利用可能性

[0191] 以上のように、本発明は動画の圧縮符号化および復号に利用することができる。

請求の範囲

[1] a) 1以上の画像フレームを間に挟む第1、第2キーフレーム間でマッチングを計算し、第1、第2キーフレーム間の対応点情報を生成するステップと、

b) 第1、第2キーフレーム間の対応点情報をもとに当該キーフレームに含まれる画素を移動させることによって、仮想の第2キーフレームを生成するステップと、

c) 現実の第2キーフレームと仮想の第2キーフレームとの差分を圧縮符号化するステップと、

d) 第1キーフレーム、第1、第2キーフレーム間の対応点情報、および、現実の第2キーフレームと仮想の第2キーフレーム間で圧縮符号化された差分をこれらのキーフレーム間の符号化データとして出力するステップと、

を備える動画符号化方法。

[2] 請求項1に記載の方法において、さらに、

e) 現実の第2キーフレームと仮想の第2キーフレーム間で圧縮符号化された差分を復号するステップと、

f) 復号された差分と前記仮想の第2キーフレームとから、改良された仮想の第2キーフレームを生成するステップと、

g) 1以上の画像フレームを間に挟む第2、第3キーフレーム間でマッチングを計算し、第2、第3キーフレーム間の対応点情報を生成するステップと、

h) 第2、第3キーフレーム間の対応点情報をもとに、改良された仮想の第2キーフレームに含まれる画素を移動させることによって、仮想の第3キーフレームを生成するステップと、

i) 現実の第3キーフレームと仮想の第3キーフレームとの差分を圧縮符号化するステップと、

j) 第2、第3キーフレーム間の対応点情報、および現実の第3キーフレームと仮想の第3キーフレーム間で圧縮符号化された差分をこれらのキーフレーム間の符号化データとして出力するステップと、

を備える動画符号化方法。

[3] 請求項2に記載の方法において、さらに後続のキーフレームについて、順次前記の

e)からj)のステップを繰り返す動画符号化方法。

[4] 請求項3に記載の方法において、前記のe)からj)のステップを所定のグループ終了キーフレームに到達したときに終了し、当該グループ終了キーフレームの次のキーフレームを新たに第1キーフレームと見なして前記のa)以下の処理を繰り返す動画符号化方法。

[5] k) 1以上の画像フレームを間に挟む第1、第2キーフレーム間の対応点情報、および第1キーフレームを取得するステップと、

l) 第1、第2キーフレーム間の対応点情報をもとに第1キーフレームに含まれる画素を移動させることによって、仮想の第2キーフレームを生成するステップと、

m) 予め符号化側にて求められた現実の第2キーフレームと仮想の第2キーフレームとの差分の圧縮符号化データを取得するステップと、

o) 取得された差分の圧縮符号化データと前記仮想の第2キーフレームとから、改良された仮想の第2キーフレームを生成するステップと、

p) 第1、第2キーフレーム間の対応点情報をもとに、第1キーフレームと改良された仮想の第2キーフレーム間で補間計算をすることにより、これらのキーフレームの間に存在すべき中間フレームを生成するステップと、

q) 第1キーフレーム、生成された中間フレーム、改良された仮想の第2キーフレームをこれらのキーフレーム間の復号データとして出力するステップと、

を備える動画復号方法。

[6] 請求項5に記載の方法において、さらに、

r) 1以上の画像フレームを間に挟む第2、第3キーフレーム間の対応点情報を取得するステップと、

s) 第2、第3キーフレーム間の対応点情報をもとに、改良された仮想の第2キーフレームに含まれる画素を移動させることによって、仮想の第3キーフレームを生成するステップと、

t) 予め符号化側にて求められた現実の第3キーフレームと仮想の第3キーフレームとの差分の圧縮符号化データを取得するステップと、

u) 取得された差分の圧縮符号化データと前記仮想の第3キーフレームとから、改

良された仮想の第3キーフレームを生成するステップと、

v) 第2、第3キーフレーム間の対応点情報をもとに、改良された仮想の第2キーフレームと改良された仮想の第3キーフレーム間で補間計算をすることにより、これらのキーフレームの間に存在すべき中間フレームを生成するステップと、

w) 改良された仮想の第2キーフレーム、生成された中間フレーム、改良された仮想の第3キーフレームをこれらのキーフレーム間の復号データとして出力するステップと、

を備える動画復号方法。

[7] 請求項6に記載の方法において、さらに後続のキーフレームについて、順次前記のr)からw)のステップを繰り返す動画復号方法。

[8] 請求項7に記載の方法において、前記のr)からw)のステップを所定のグループ終了キーフレームに到達したときに終了し、当該グループ終了キーフレームの次のキーフレームを新たに第1キーフレームと見なして前記のk)以下の処理を繰り返す動画復号方法。

[9] 請求項1に記載の方法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。

[10] 請求項5に記載の方法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。

[11] 前記a)のステップにおけるマッチングの正確さを評価するステップをさらに備え、

評価の結果に依存して、前記c)のステップにおける圧縮スキームを切り替えることを特徴とする請求項1から4のいずれかに記載の動画符号化方法。

[12] 前記現実の第2キーフレームと仮想の第2キーフレームとの差分のうち、前記評価するステップにおいて前記第1、第2キーフレーム間のマッチング精度が高いと判断された領域に対応する部分の圧縮率を高くすることを特徴とする請求項11に記載の動画符号化方法。

[13] 前記評価するステップでは、前記対応点情報を生成するステップにおいて相互に対応するとされた前記第1、第2キーフレーム間の各点のうち、対応傾向を示すパラメータが近接する点のものと比べて所定のしきい値以上に異なる点をマッチングエラー

として検出し、

前記現実の第2キーフレームと仮想の第2キーフレームとの差分のうち、マッチングエラーが検出された領域に対応する部分の圧縮率を高くすることを特徴とする請求項11に記載の動画符号化方法。

[14] 前記評価するステップでは、マッチングエネルギー値に着目してマッチングの正確さを評価することを特徴とする請求項11から13のいずれかに記載の動画符号化方法。

[15] 前記マッチングエネルギーは対応点の位置と画素値の双方に基づいて導出されることを特徴とする請求項14に記載の動画符号化方法。

[16] 前記第1、第2キーフレーム及び間に挟まれる画像フレームの全てについて隣り合う画像フレームどうしのマッチングをさらに計算するステップと、

前記計算するステップで得られる各画像間のマッチングエネルギーを計算して、各画像フレームのマッチングエネルギーに関するパラメータを取得するステップと、

前記パラメータを累積的に加算するステップと、

加算の結果が所定のしきい値を越える画像フレームを新たにキーフレームとして登録するステップと、

をさらに備えることを特徴とする請求項1から4、11から15のいずれかに記載の動画符号化方法。

[17] 前記第2キーフレーム以降の各画像フレームについても、隣り合う画像フレームどうしでマッチングを計算し、マッチングエネルギーに関するパラメータの累積的な加算値が前記しきい値を越えた場合はその直前の画像フレームを新たにキーフレームとして登録することを特徴とする請求項16に記載の動画符号化方法。

[18] 画像フレームの全てについて隣り合う画像どうしのマッチングをさらに計算するステップと、

前記計算するステップで得られる各画像フレーム間のマッチングエネルギーに基づいて各画像フレームのマッチングエネルギーに関するパラメータを取得するステップと、

前記パラメータが、所定のしきい値を越える画像フレームを前記グループ終了キー

フレームとするステップと、

をさらに備えることを特徴とする請求項4に記載の動画符号化方法。

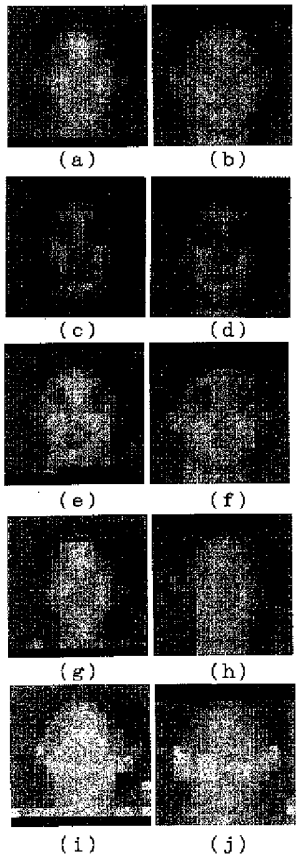
[19] 第1、第2画像フレーム間で、領域ベースの画像マッチングを計算し、そのマッチング結果を利用して、少なくとも第3画像フレームを符号化する方法であって、

画像マッチングの結果の良否を領域ごとに判定するステップと、

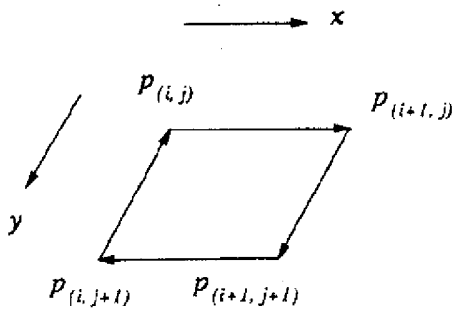
第3画像フレームの符号化プロセスにおいて、前記判定するステップにおいて判定された良否に基づき、領域ごとに量子化スキームを選択するステップと、

を備える動画符号化方法。

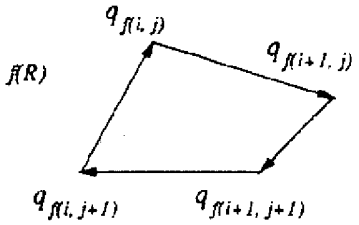
[図1]



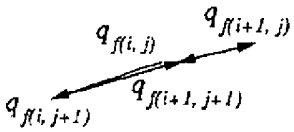
[図2]


x
y
$p_{(i,j)}$
$p_{(i+1,j)}$
$p_{(i,j+1)}$
$p_{(i+1,j+1)}$

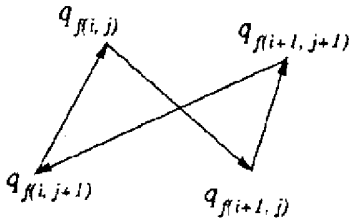
(R)


f(R)
$q_{f(i,j)}$
$q_{f(i+1,j)}$
$q_{f(i,j+1)}$
$q_{f(i+1,j+1)}$

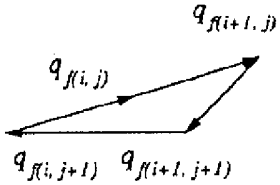
(A)


$q_{f(i,j)}$
$q_{f(i+1,j)}$
$q_{f(i,j+1)}$
$q_{f(i+1,j+1)}$

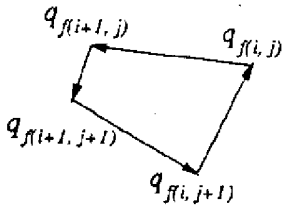
(E)


$q_{f(i,j)}$
$q_{f(i+1,j+1)}$
$q_{f(i,j+1)}$
$q_{f(i+1,j)}$

(B)

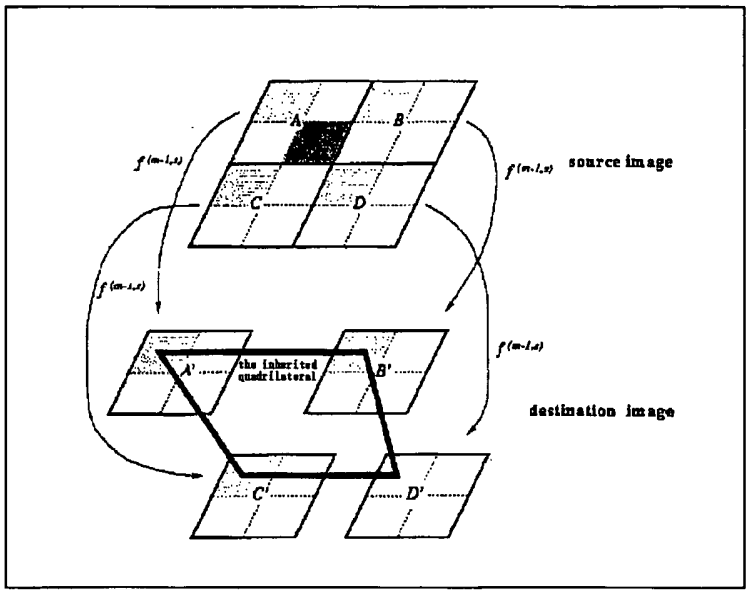

$q_{f(i+1,j)}$
$q_{f(i,j)}$
$q_{f(i,j+1)}$
$q_{f(i+1,j+1)}$

(D)

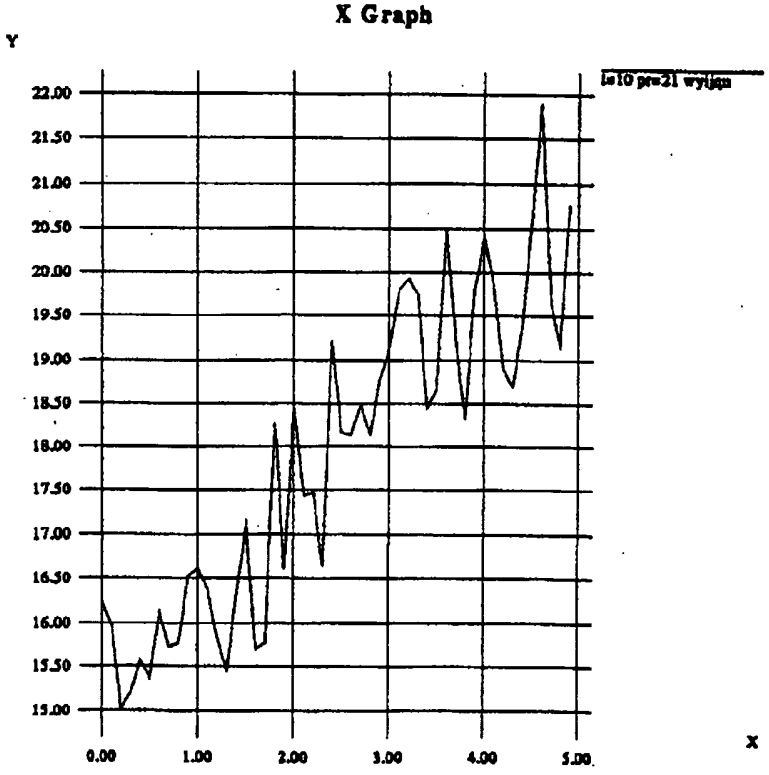

$q_{f(i+1,j)}$
$q_{f(i,j)}$
$q_{f(i+1,j+1)}$
$q_{f(i,j+1)}$

(C)

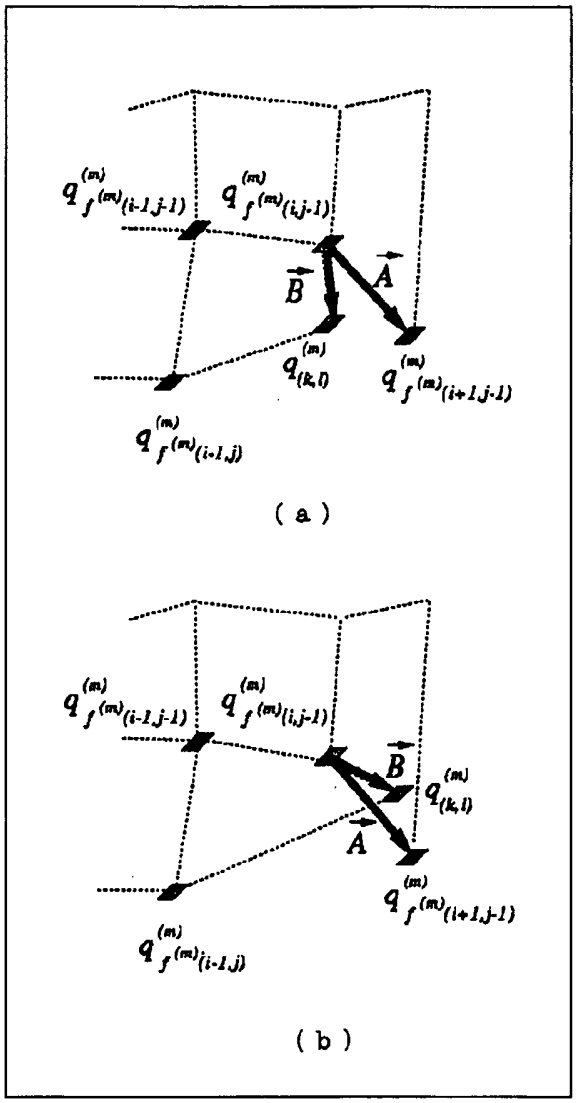
[図3]


A
B
C
D
source image
A'
the inherited quadrilateral
B'
destination image
C'
D'

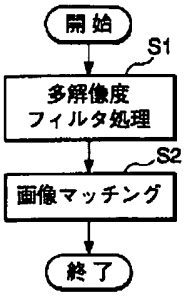
[図4]


X Graph
Y
22.00
21.50
21.00
20.50
20.00
19.50
19.00
18.50
18.00
17.50
17.00
16.50
16.00
15.50
15.00
0.00
1.00
2.00
3.00
4.00
5.00
X

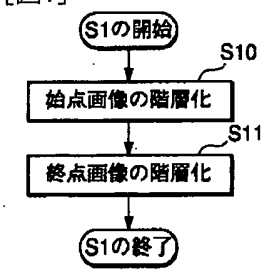
[図5]


$q^{(m)}_{f^{(m)}(i-1,j-1)}$
$q^{(m)}_{f^{(m)}(i,j-1)}$
$\vec{B}$
$\vec{A}$
$q^{(m)}_{(k,l)}$
$q^{(m)}_{f^{(m)}(i+1,j-1)}$
$q^{(m)}_{f^{(m)}(i-1,j)}$
(a)
$q^{(m)}_{f^{(m)}(i-1,j-1)}$
$q^{(m)}_{f^{(m)}(i,j-1)}$
$\vec{B}$
$q^{(m)}_{(k,l)}$
$\vec{A}$
$q^{(m)}_{f^{(m)}(i+1,j-1)}$
$q^{(m)}_{f^{(m)}(i-1,j)}$
(b)

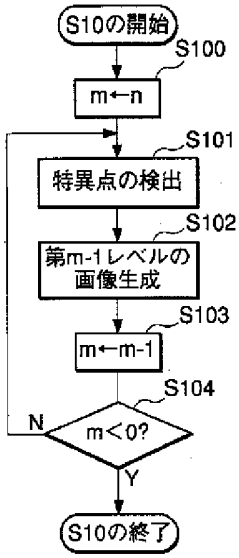
[図6]


開始
S1
多解像度
フィルタ処理
S2
画像マッチング
終了

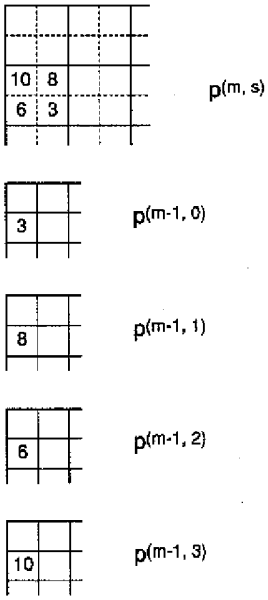
[図7]


S1の開始
S10
始点画像の階層化
S11
終点画像の階層化
S1の終了

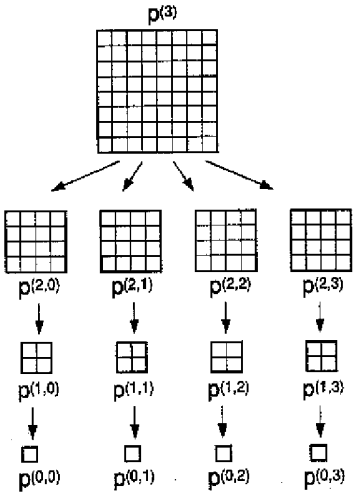
[図8]


S10の開始
S100
m←n
S101
特異点の検出
S102
第m-1レベルの画像生成
S103
m←m-1
S104
N
m<0?
Y
S10の終了

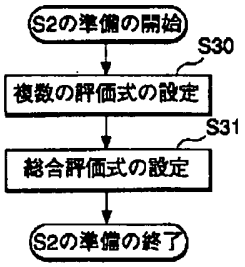
[図9]


10 8
6 3
p(m, s)
3
p(m-1, 0)
8
p(m-1, 1)
6
p(m-1, 2)
10
p(m-1, 3)

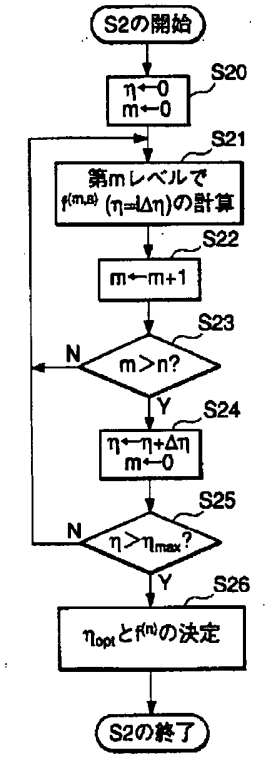
[図10]


p(3)
p(2,0)
p(2,1)
p(2,2)
p(2,3)
p(1,0)
p(1,1)
p(1,2)
p(1,3)
p(0,0)
p(0,1)
p(0,2)
p(0,3)

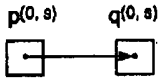
[図11]



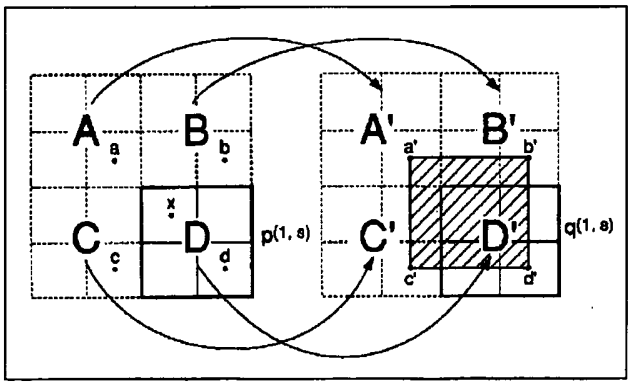
[図12]



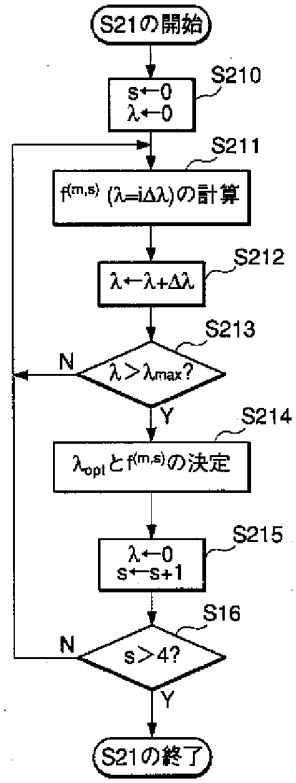
[図13]



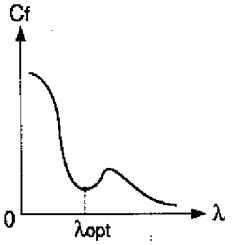
[図14]



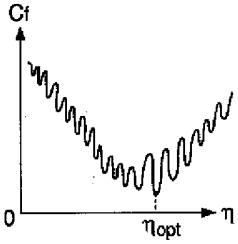
[図15]


S21の開始
S210
s←0
λ←0
S211
f(m,s) (λ=iΔλ)の計算
S212
λ←λ+Δλ
S213
N
λ>λmax?
Y
S214
λopt とf(m,s)の決定
S215
λ←0
s←s+1
S16
N
s>4?
Y
S21の終了

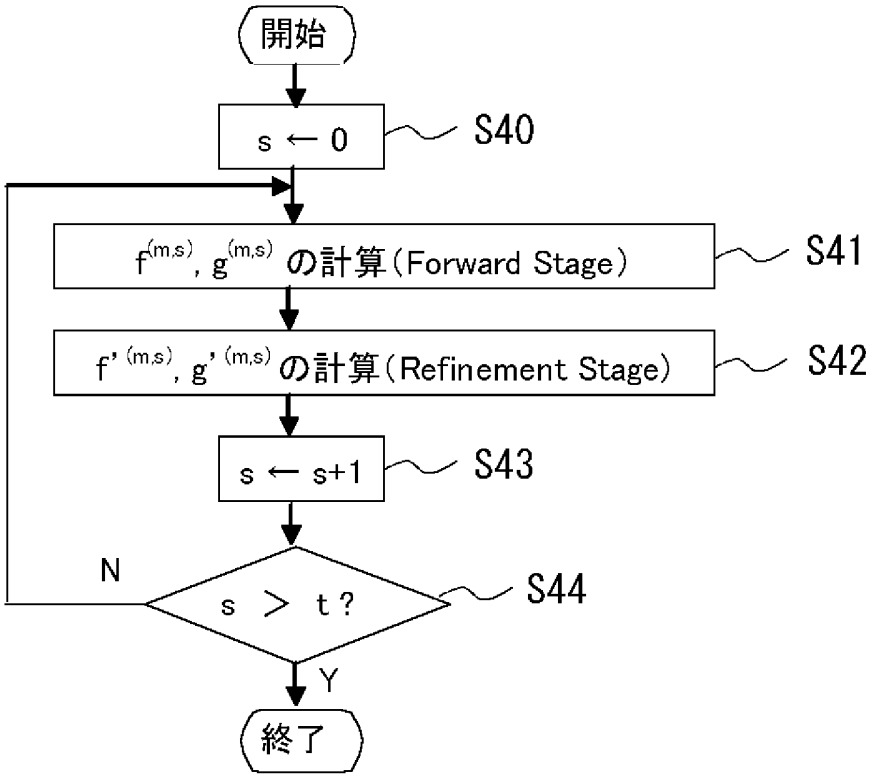
[図16]


Cf
0
λopt
λ

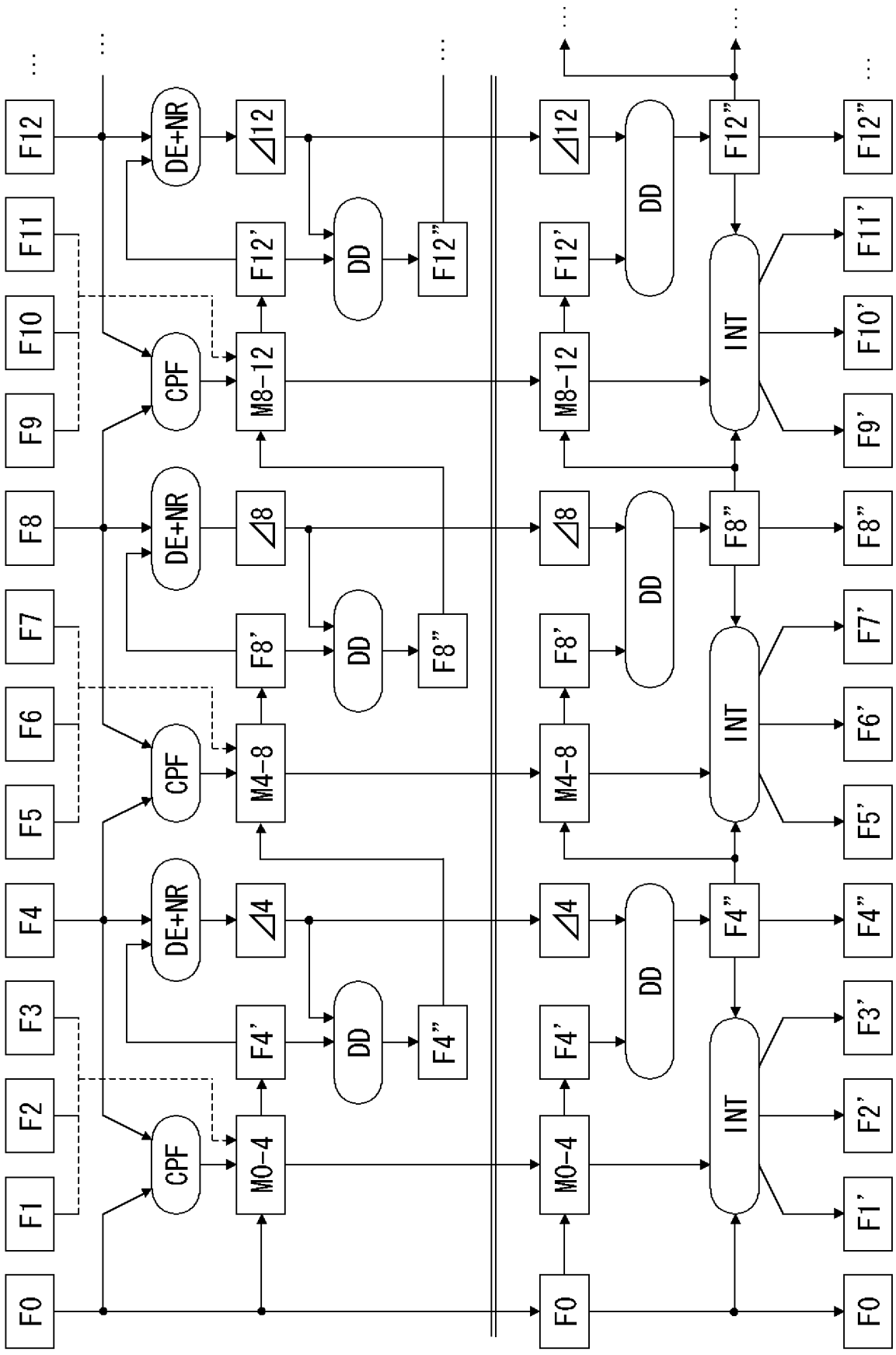
[図17]


Cf
0
ηopt
η

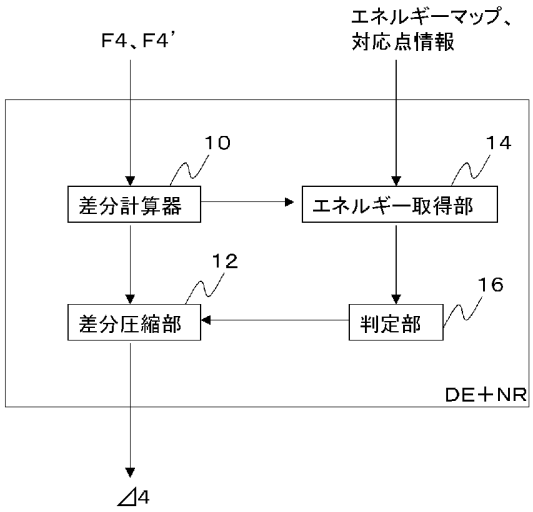
[図18]


開始
s ← 0
S40
f(m,s), g(m,s) の計算(Forward Stage)
S41
f'(m,s), g'(m,s) の計算(Refinement Stage)
S42
s ← s+1
S43
N
s > t ?
S44
Y
終了

[図19]


F0
F1
F2
F3
F4
F5
F6
F7
F8
F9
F10
F11
F12
CPF
DE+NR
M0-4
M4-8
M8-12
Δ4
Δ8
Δ12
F4'
F8'
F12'
DD
F4"
F8"
F12"
INT
F0
F1'
F2'
F3'
F4"
F5'
F6'
F7'
F8"
F9'
F10'
F11'
F12"

[図20]


F4、F4'
エネルギーマップ、
対応点情報
10
14
差分計算器
エネルギー取得部
12
16
差分圧縮部
判定部
DE+NR
⊿4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2005/005941

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[7] H04N7/36

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[7] H04N7/24-68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2005
Kokai Jitsuyo Shinan Koho 1971-2005 Toroku Jitsuyo Shinan Koho 1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 8-107558 A (NEC Corp.), 23 April, 1996 (23.04.96), Full text; all drawings (Family: none)	1-10 11-18
A	JP 10-304374 A (Sharp Corp.), 13 November, 1998 (13.11.98), Par. Nos. [0058] to [0061]; Fig. 1 & US 6052417 A & EP 874525 A	11-18
A	JP 2001-128179 A (NEC Corp.), 11 May, 2001 (11.05.01), Par. Nos. [0043] to [0047]; Fig. 3 & US 6611558 B1	11-18

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
09 September, 2005 (09.09.05)	27 September, 2005 (27.09.05)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2005/005941

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

The inventions of claims 1-18 relate to a method or a program for outputting a first key frame, the corresponding point information between the first and the second key frame, and a compressed/encoded difference between the actual second key frame and the virtual second key frame as encoded data between the key frames, or a method or a program for decoding the encoded data.

The invention of claim 19 relates to a method for encoding the third image frame.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-18

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２００５／００５９４

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.[7]　Ｈ０４Ｎ７／３６

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.[7]　Ｈ０４Ｎ７／２４－６８

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２００５年
日本国実用新案登録公報	１９９６－２００５年
日本国登録実用新案公報	１９９４－２００５年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Ｘ Ａ	ＪＰ　８－１０７５５８　Ａ（日本電気株式会社） １９９６．０４．２３，全文，全図（ファミリーなし）	１－１０ １１－１８
Ａ	ＪＰ　１０－３０４３７４　Ａ（シャープ株式会社） １９９８．１１．１３，【００５８】－【００６１】，図１ ＆　ＵＳ　６０５２４１７　Ａ　＆　ＥＰ　８７４５２５　Ａ	１１－１８

☑ Ｃ欄の続きにも文献が列挙されている。　☐ パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日 ０９．０９．２００５	国際調査報告の発送日 27.9.2005
国際調査機関の名称及びあて先 日本国特許庁（ＩＳＡ／ＪＰ） 郵便番号１００－８９１５ 東京都千代田区霞が関三丁目４番３号	特許庁審査官（権限のある職員）　５Ｃ　３６６０ 國分　直樹 電話番号　０３－３５８１－１１０１　内線　３５４１

様式ＰＣＴ／ＩＳＡ／２１０（第２ページ）（２００４年１月）

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Ａ	ＪＰ　２００１－１２８１７９　Ａ（日本電気株式会社） ２００１．０５．１１，【００４３】－【００４７】，図３ ＆　ＵＳ　６６１１５５８　Ｂ１	１１－１８

国際調査報告

国際出願番号 PCT／JP2005／005941

第Ⅱ欄 請求の範囲の一部の調査ができないときの意見（第１ページの２の続き）

法第8条第3項（PCT17条(2)(a)）の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. ☐ 請求の範囲＿＿＿＿＿＿＿＿ は、この国際調査機関が調査をすることを要しない対象に係るものである。
つまり、

2. ☐ 請求の範囲＿＿＿＿＿＿＿＿ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、

3. ☐ 請求の範囲＿＿＿＿＿＿＿＿ は、従属請求の範囲であってＰＣＴ規則6.4(a)の第２文及び第3文の規定に従って記載されていない。

第Ⅲ欄 発明の単一性が欠如しているときの意見（第１ページの３の続き）

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

請求の範囲１－１８に係る発明は、第１キーフレーム、第１、第２キーフレーム間の対応点情報、および、現実の第２キーフレームと仮想の第２キーフレーム間で圧縮符号化された差分をこれらのキーフレーム間の符号化データとして出力する方法もしくはプログラム、あるいは当該符号化データを復号する方法もしくはプログラムである。

請求の範囲１９に係る発明は、第３画像フレームを符号化する方法である。

1. ☐ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。

2. ☐ 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。

3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったので、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。

4. ☑ 出願人が必要な追加調査手数料を期間内に納付しなかったので、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

１－１８

追加調査手数料の異議の申立てに関する注意

☐ 追加調査手数料の納付と共に出願人から異議申立てがあった。

☐ 追加調査手数料の納付と共に出願人から異議申立てがなかった。