

(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2009-237747

(P2009-237747A)

(43) 公開日 平成21年10月15日(2009.10.15)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 13/00 (2006.01)	G06T 13/00 B	5B050
G10L 21/04 (2006.01)	G10L 21/04 I2OD	

審査請求 未請求 請求項の数 7 O L (全 17 頁)

(21) 出願番号 特願2008-80930 (P2008-80930)
 (22) 出願日 平成20年3月26日 (2008. 3. 26)

(71) 出願人 000004260
 株式会社デンソー
 愛知県刈谷市昭和町1丁目1番地
 (74) 代理人 100095751
 弁理士 菅原 正倫
 (72) 発明者 立石 雅彦
 愛知県刈谷市昭和町1丁目1番地 株式会
 社デンソー内
 Fターム(参考) 5B050 BA07 BA08 BA12 CA07 EA05
 EA13 EA18 EA19 EA24 FA02

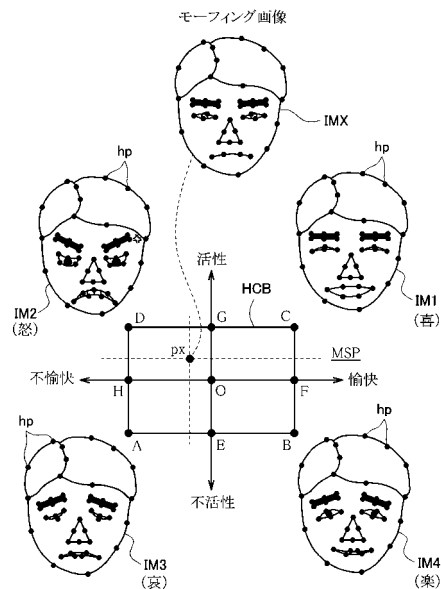
(54) 【発明の名称】 データポリモーフィング方法及びデータポリモーフィング装置

(57) 【要約】 (修正有)

【課題】 画像合成演算の回数を大幅に削減することができるデータモーフィング方法を提供する。

【解決手段】 M次元(2 M)のモデルデータ配置空間MSP上に少なくとも4個のモーフィング元座標点を定義し、各モーフィング元座標点に対応付けた形でモデルデータ列IM1~IM4を用意する。モデルデータ配置空間MSPは、モーフィング元座標点を頂点とする超立方体からなる単位セルHCBに区画される。モデルデータ配置空間MSPにおいて単位セルHCBの各頂点をモーフィング元座標点として特定するとともに、単位セルHCB内の任意の座標点をモーフィング先座標点pxとして選定する。そして、モーフィング元座標点に対応する4組のモデルデータ列を、各モーフィング元座標点のモーフィング先座標点pxまでの距離に応じた重みにてポリモーフィングする。

【選択図】 図7



【特許請求の範囲】

【請求項 1】

M次元(2 M)のモデルデータ配置空間上に少なくとも2 M個のモーフィング元座標点を定義し、各モーフィング元座標点に対応付けた形でモーフィング処理対象となるモデルデータ列を用意するとともに、

前記モデルデータ配置空間を、前記モーフィング元座標点を頂点とする頂点数2 M個の超直方体からなる単位セルに区画し、前記モデルデータ配置空間において前記単位セルの各頂点をモーフィング元座標点として特定するとともに、前記単位セル内の任意の座標点をモーフィング先座標点として選定し、

各前記モーフィング元座標点に対応する2 M組の前記モデルデータ列を、前記モデルデータ配置空間における各前記モーフィング元座標点の前記モーフィング先座標点までの距離に応じた重みにてポリモーフィングすることにより、前記モーフィング先座標点に対応する合成データ列を作成し、出力することを特徴とするデータポリモーフィング方法。

10

【請求項 2】

前記超直方体を、前記モーフィング先座標点を通して各面と平行なM個の平面で切断することにより、それぞれ前記モーフィング先座標点を共有し、かつ前記超直方体の頂点をなす前記モーフィング元座標点を排他的に1個ずつ取り合う2 M個の部分直方体に区切り、各部分直方体の前記超直方体に対する相対体積を、当該部分直方体に含まれるモーフィング元座標点の前記超直方体の対角線方向反対側に位置するモーフィング元座標点への重みとする形でポリモーフィングを行なう請求項1記載のデータポリモーフィング方法。

20

【請求項 3】

前記モデルデータ配置空間が2次元であり、前記単位セルが長方形である請求項1又は請求項2に記載のデータポリモーフィング方法。

【請求項 4】

前記モデルデータ配置空間が二次元であり、長方形をなす前記単位セルの互いに平行対向する1対の各辺への前記モーフィング先座標点の正射影点を分点とする形で、各辺の両端の前記モーフィング元座標点に対応する前記モデルデータ列同士を梃子の関係に従う重みにて一次モーフィングすることにより1対の中間データ列を得、各前記正射影点が張る線分上の前記モーフィング先座標点を新たな分点とみなし、その分点に関して前記1対の中間データ列同士を梃子の関係に従う重みにて二次モーフィングすることにより前記合成データ列を得る請求項3記載のデータポリモーフィング方法。

30

【請求項 5】

前記モデルデータ列が画像データを構成するものである請求項1ないし請求項4のいずれか1項に記載のデータポリモーフィング方法。

【請求項 6】

前記モデルデータ列が音声波形データを構成するものである請求項1ないし請求項5のいずれか1項に記載のデータポリモーフィング方法。

【請求項 7】

M次元(2 M)のモデルデータ配置空間上に定義された少なくとも2 M個のモーフィング元座標点に個別に対応付けられた、モーフィング処理対象となるモデルデータ列を取得するモデルデータ列取得手段と、

40

前記モデルデータ配置空間を、前記モーフィング元座標点を頂点とする頂点数2 M個の超直方体からなる単位セルに区画し、前記モデルデータ配置空間において前記単位セルの各頂点をモーフィング元座標点として特定するとともに、前記単位セル内の任意の座標点をモーフィング先座標点として選定するモーフィング先座標点選定手段と、

各前記モーフィング元座標点に対応する2 M組の前記モデルデータ列を、前記モデルデータ配置空間における各前記モーフィング元座標点の前記モーフィング先座標点までの距離に応じた重みにてポリモーフィング処理することにより、前記モーフィング先座標点に対応する合成データ列を作成するポリモーフィング処理手段と、

前記ポリモーフィング処理により得られた合成データ列を出力する合成データ列出力手

50

段と、

を備えたことを特徴とするデータポリモーフィング装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データポリモーフィング方法及びデータポリモーフィング装置に関する。

【背景技術】

【0002】

【特許文献1】特開2000-354517号公報

【特許文献2】特開2002-229579号公報

10

【特許文献3】特開2002-229579号公報

【非特許文献1】IEEE Computer Graphics and Applications, January/February 1998, 60-73

【非特許文献2】Kawahara, H., Katayose, H., Cheveign'e, de A., and Patterson, R. D.: Fixed Point Analysis of Frequency to Instantaneous Frequency Mapping for Accurate Estimation of F0 and Periodicity, Eurospeech '99, Vol. 6, pp.2781-2784

【非特許文献3】「STRAIGHTを用いた音声モーフィングの事例に基づくデザイン支援への応用」The 20th Annual Conference of the Japanese Society for Artificial Intelligence, 2006, 1D1-5

【非特許文献4】<http://www.wakayama-u.ac.jp/~kawahara/Miraikandemo/straightMorph.swf>

20

【0003】

画像加工技術の一つにモーフィングが知られている（例えば、特許文献1）。具体的にはモーフィング元となる2画像間で複数個の対応点（基準点）を指定し、モーフィング元画像の各対応点を合成比の上下限值に対応させる。そして、任意の中間の合成比を指定し、モーフィング元画像の対応点を上記合成比に応じた重みにて補間することにより合成後対応点を求め、対応点近傍の各ピクセル濃度を同様の補間演算によりブレンドすることにより合成画像（以下、「モーフィング先画像」ともいう）が得られる。図6は、人物の顔画像を第一モーフィング元画像とし、犬の顔画像を第二モーフィング元画像として、上記モーフィングの原理に従い画像合成した実例を示すものである。また、第一モーフィング元画像から第二モーフィング元画像に向けて、合成比を段階的に変化させ、補間の重み付けが異なる複数のモーフィング先画像を作成し、これをコマとして動画再生すれば、第一モーフィング元画像（人物）が徐々に変形して第二モーフィング元画像（犬）へ移り変わってゆく独特の遷移アニメーションが得られる。この遷移アニメーションを作成する技術を狭義に「モーフィング」と呼ぶこともある。

30

【0004】

そして、近年、2つのモーフィング元画像の合成処理を対象とする従来のモーフィングに加え、3枚以上の複数のモーフィング元画像を合成するポリモーフィング技術に関する研究が進められている（非特許文献1）。

【0005】

一方、モーフィング技術は音声合成にも応用されている（例えば、特許文献2、非特許文献2, 3）。すなわち、音声入力波形はフーリエ変換処理によりパワースペクトラムの形で二次元マッピングでき、画像の場合と同様にモーフィングによる合成処理が可能となるのである。例えば、同一人物が「あいしてる」と発音する場合でも、当該の人物の感情により異なる波形が得られる。そこで、「喜び」「怒り」「悲しみ」といった各感情を典型的に反映した音声波形のパワースペクトラムをモーフィング元データとして用意し、画像の場合の対応点に相当する特徴点をスペクトル上に定め、その特徴点を用いて同様の補間処理を行なうことにより合成パワースペクトラムを得ることができる。これを音声波形に逆変換すれば中間の感情状態を示す音声波形を合成出力できる（非特許文献4）。非特許文献4の音声モーフィングでは、3つの音声波形によるポリモーフィング技術が採用さ

40

50

れている。

【0006】

ところで、非特許文献1に開示されている画像のポリモーフィングは、2個の画像に適用される通常のモーフィングのパラダイムを、任意数(M)個の画像のモーフィングに拡張したものである。M個の画像をモーフィングする場合、各画像の規格化された合成比率は加算すれば1になる制約があるので、独立変数はM-1個であり、M-1次元空間上の座標により表現することができる。非特許文献1では、各画像を1個の(M-1)次元シンプレックスの1個の頂点で定式化するとともに、ポリモーフィングの各画像の合成比を、そのシンプレックス内の1点により表現している。例えば、モーフィング元画像が3個であれば合成比率は2次元座標点として表現可能であり、この場合のシンプレックスは三角形である(参考のために付言すれば、これと類似のシンプレックスの概念が、多成分系の状態図(相図)における組成表現にも採用されており、例えば3成分系状態図において個々の組成比は、組成三角形と称されるシンプレックス内の1点により表わすことができる)。

10

【0007】

そして、非特許文献1に開示されているポリモーフィングの概念では、このシンプレックスを用いた合成比表現に厳密に従いつつ、2画像の補間合成に分解した形で実施される。図10は、三角形シンプレックスの頂点に対応する3つの画像P0, P1, P2を合成する場合を示している。WijはPiからPjへのワーブ関数で、Pi上の各点に対応するPj上の点を特定する。最終的には合成画像Pを生成するには、まずWijをPjの重心座標gjに適用してPi毎にWijを線形内挿し、中間ワーブ関数Wiバーを導く。各Piは、隣接する2つのものが、Wiバーによりモーフィング先座標点pxの重心座標G*に応じた重みで中間合成され、中間画像Piバーを生成する。合成画像Pxは、Piバーの各点を重心座標gjが示す重みにて線形結合して得られる。

20

【0008】

図11は、これをさらに具体的に展開して示すものであり、モーフィング元座標点pa, pb, pcをそれぞれ点A, B, Cとし、また、モーフィング先座標点pxを点Xとする。三角形ABCの各頂点A, B, Cから、点Xを通して各辺と交差する直線を考え、各辺との交点をD, E, Fとすると、pxの重心座標G*の各成分は、図中のga, gb, gcとして式(1)により表わされる。図中の各点の座標値及び各線分の長さは周知の解析幾何学の手法により計算できるが、いずれも初等的であるため詳細な説明は略する。すると、3つの中間画像Piバーは、図中のPd, Pe, Pfとして式(2)により計算される。その結果、Pxはga, gb, gcを重みとするPd, Pe, Pfの線形結合として計算される。また、非特許文献4も3つの音声波形を任意比率で合成するポリモーフィングの事例を示すものであるが、非特許文献1と同様のアルゴリズムを採用しているものと推測される。

30

【発明の開示】

【発明が解決しようとする課題】

【0009】

上記従来のポリモーフィングの概念では、3画像を合成する場合は、図10からも明らかごとく、第一段階としてシンプレックスの辺数に対応した3回の2画像の補間合成処理を順次実行してそれぞれの中間合成画像を得、さらにそれらの中間合成画像を重心座標G*について線形結合処理する処理(以下、重心結合処理という)が必要となり、都合4回もの画像合成処理が必要となる。また、4画像を合成する場合は、シンプレックスは三角錐(辺数6)となり、処理は一層複雑となる。この場合、各頂点からモーフィング先座標点pxを通してシンプレックスの対向面(三角形)と交わる直線を考え、各三角形にできる交点を中間合成比点とみなすことで、4枚の三角形についての上記の3画像合成処理に分割して考えることができる。そして、その4つの合成結果を上記各直線のモーフィング先座標点pxによる分割比に応じて重心結合処理することにより最終的な合成画像が得られるが、6回の補間合成処理と、4+1=5回の重心結合処理との都合11回の画像合

40

50

成処理が必要となる。また、5画像合成では、シンプレックスは5個の三角錐に取り囲まれる辺数10、面数5の四次元超立体となり、10回の補間合成処理と、 $5 + 5 + 1 = 11$ 回の重心結合処理との都合21回の画像合成処理が必要となる。このように、画像数増加に伴い、画像合成回数ひいては演算負荷が急激に増大する欠点があることは明らかである。

【0010】

本発明の課題は、4以上の画像データや音声データ等をモーフィングするに際しても、画像合成演算の回数を大幅に削減することができ、ひいてはその技術的な汎用性を高めることができるデータモーフィング方法及び装置を提供することにある。

【課題を解決するための手段及び発明の効果】

10

【0011】

上記の課題を解決するために、本発明のデータポリモーフィング方法は、

M次元($2 \leq M$)のモデルデータ配置空間上に少なくとも2M個のモーフィング元座標点を定義し、各モーフィング元座標点に対応付けた形でモーフィング処理対象となるモデルデータ列を用意するとともに、

モデルデータ配置空間を、モーフィング元座標点を頂点とする頂点数2M個の超直方体からなる単位セルに区画し、モデルデータ配置空間において単位セルの各頂点をモーフィング元座標点として特定するとともに、単位セル内の任意の座標点をモーフィング先座標点として選定し、

各モーフィング元座標点に対応する2M組のモデルデータ列を、モデルデータ配置空間における各モーフィング元座標点のモーフィング先座標点までの距離に応じた重みにてポリモーフィングすることにより、モーフィング先座標点に対応する合成データ列を作成し、出力することを特徴とする。

20

【0012】

また、本発明のデータポリモーフィング装置は、

M次元($2 \leq M$)のモデルデータ配置空間上に定義された少なくとも2M個のモーフィング元座標点に個別に対応付けられた、モーフィング処理対象となるモデルデータ列を取得するモデルデータ列取得手段と、

モデルデータ配置空間を、モーフィング元座標点を頂点とする頂点数2M個の超直方体からなる単位セルに区画し、

30

モデルデータ配置空間において単位セルの各頂点をモーフィング元座標点として特定するとともに、単位セル内の任意の座標点をモーフィング先座標点として選定するモーフィング先座標点選定手段と、

各モーフィング元座標点に対応する2M組のモデルデータ列を、モデルデータ配置空間における各モーフィング元座標点のモーフィング先座標点までの距離に応じた重みにてポリモーフィング処理することにより、モーフィング先座標点に対応する合成データ列を作成するポリモーフィング処理手段と、

ポリモーフィング処理により得られた合成データ列を出力する合成データ列出力手段と、を備えたことを特徴とする。

【0013】

40

この発明では、ポリモーフィング対象となるモデルデータ列をモデルデータ配置空間上にモーフィング元座標点と対応付けてマッピングするとともに、モーフィング元座標点を配置する単位セルの頂点数は $2M$ ($M \geq 2$)であり、従来採用していたM次元シンプレックス(頂点数 $M + 1$)よりも頂点数の多い単位セルが採用されている。そして、該単位セルが、具体的には頂点数2M個の超直方体として選ばれている。ここでいう超直方体は、モデルデータ配置空間を直交座標系としたとき、次元数Mが3のときは直方体(立方体を概念として含む)、次元数Mが2のときは長方形(正方形を概念として含む)となる。

【0014】

単位セルの頂点、すなわちモーフィング元座標点の全てをランダムに設定した場合は、モーフィング元座標点1つにつきM個の座標成分が存在することから、ポリモーフィング

50

演算には $M \times$ (頂点数) の座標値を独立変数として考慮しなければならない。しかし、上記のような超直方体を採用すれば、超直方体の各辺の長さ (M 通り) が与えられれば、超直方体の頂点をなす1つのモーフィング元座標点の座標値から、他のモーフィング元座標点の座標値も決定される。その結果、シンプレックスを用いる場合と比較してポリモーフィング処理のための補間演算を大幅に簡略化できる。

【0015】

上記超直方体の各頂点をなすモーフィング元座標点とモーフィング先座標点との幾何学的な関係に基づき、各モーフィング元座標点に対応するモデルデータ列を線形補間合成して合成画像を得る場合は、次の手法を採用することにより、ポリモーフィングアルゴリズムの大幅な簡略化を図ることができる。すなわち、超直方体のモーフィング先座標点を通じて各面と平行な M 個の平面で切断する。これにより、超直方体は、それぞれモーフィング先座標点を共有し、かつ超直方体の頂点をなすモーフィング元座標点を排他的に1個ずつ取り合う $2M$ 個の部分直方体に区切られる。モデルデータ配置空間を直交座標系としたとき、部分直方体は、次元数 M が3のときは直方体 (立方体を概念として含む) であり、超直方体に対する分割数は8、次元数 M が2のときは長方形 (正方形を概念として含む) であり、超直方体に対する分割数は4となる。 M 次元に一般化した場合、超直方体の部分直方体による分割数は 2^M である。

10

【0016】

そして、各部分直方体の超直方体に対する相対体積を、当該部分直方体に含まれるモーフィング元座標点の超直方体の対角線方向反対側に位置するモーフィング元座標点への重みとする形でポリモーフィングを行なう。この方法によれば、ポリモーフィングの重み演算を各部分直方体の体積演算に転換することができ、例えば2点間線形補間によるモデルデータ列合成を比較的少数回繰り返すだけで最終的な合成画像を簡単に得ることができる。

20

【0017】

部分直方体の相対体積を用いた上記のポリモーフィング演算方法は、これと数学的に等価な結果が得られるアルゴリズムであれば、その演算手順は特に限定されない。例えば、2組のモデルデータ列をモーフィングする処理を、空間の次元数だけ段階的に繰り返す方式を例示できる。例えば、モデルデータ配置空間が二次元であれば、長方形をなす単位セルの互いに平行対向する1対の各辺へのモーフィング先座標点の正射影点を分点とする形で、各辺の両端のモーフィング元座標点に対応するモデルデータ列同士を梃子の関係に従う重みにて一次モーフィングすることにより1対の中間データ列を得る。そして、上記一次モーフィングで使用した各正射影点が張る線分上のモーフィング先座標点を新たな分点とみなし、その分点に関して1対の中間データ列同士を梃子の関係に従う重みにて二次モーフィングすることにより合成データ列を得る。長方形をなす単位セルの平行対辺は2組存在するが、どちらの平行対辺を採用しても、得られる結果は数学的に等価である。

30

【0018】

本発明のポリモーフィング方法及び装置の適用対象は特に限定されないが、周知のモーフィング手法と同様に、画像データないし音声データをその典型的な適用対象として例示することができる。画像モーフィングの場合は、モデルデータ列は画像データを構成する。具体的には、各モーフィング元画像上に設定された、互いに一対一の対応関係を充足する対応点や、その対応点周辺のピクセルがデータ列を構成する。また、音声モーフィングの場合は、モデルデータ列が音声波形データを構成する。具体的には、各モーフィング元音声波形のパワースペクトラムから周知のケプストラム解析により分離した、スペクトラム包絡 (主に声道の共振特性を反映した情報となる) とスペクトラム微細構造 (主に、声帯での音源的な特徴を反映した情報となる) の各特徴が反映されるように、パワースペクトラムないしケプストラムプロファイル上に定義された対応点等がデータ列を構成する。

40

【発明を実施するための最良の形態】

【0019】

以下、本発明の実施の形態を、図面を用いて説明する。

50

図1は、本発明のデータポリモーフィング装置の電気的構成の一例を示すブロック図である。データポリモーフィング装置1の制御主体は、CPU51、RAM52、ROM53及び入出力部54をバス接続した構造の周知のマイコン50により構成され、ROM53にはポリモーフィング処理の全体動作を管理するための制御ソフトウェアと、ポリモーフィング結果等を入力するための出力ソフトウェアがROM53内に格納されている。CPU51はRAM51をワークエリアとしてこれらソフトウェアを実行することにより、本発明の方法に従う後述のポリモーフィング処理と、ポリモーフィング結果(画像ないし音声)の出力処理を行なう。

【0020】

入出力部54にはマウス、キーボード、音声認識入力部等からなる入力部57、CDやDVDなどの外部メディアの記録内容を読み取るメディアドライブ58、モニタ59、プリンタ60、音声データに基づいて音声波形信号を合成する音声合成部61及び音声出力のためのスピーカ62が接続されている。

10

【0021】

マイコン50のバスにはモーフィング処理LSI55及びグラフィックメモリ56が接続されている。モーフィング処理LSI55はCPU51からの指示に従い、ポリモーフィング対象となるモデルデータ列(画像データないし音声データ(パワースペクトラムやケプストラムのプロファイルであってもよい))を用いたポリモーフィング演算及びデータ列合成処理を実行する。グラフィックメモリ56内にはデータ列合成処理を行なうためのモーフィング処理エリアと、ポリモーフィング対象となるモデルデータ列の格納メモリエリアが形成されている。

20

【0022】

この実施形態では、二次元のモデルデータ配置空間を採用して4つのモデルデータ列を合成する場合を例にとる。従って、モデルデータ列の数は $2^2 = 4$ であり、モデルデータ列の格納メモリエリアも4つ形成されているが、モデルデータ配置空間の次元数は3ないしそれ以上の値であってもよいことはもちろんである。例えばモデルデータ配置空間が三次元の場合はモデルデータ列の数は $2^3 = 8$ であり、格納メモリエリアは8つ形成される。

【0023】

画像モーフィングの場合、モデルデータ列は、それぞれモーフィング元となる画像データであり、この実施形態では、二次元のモデルデータ配置空間MSP上に長方形の単位セルHCBが定義され、該単位セルHCBのつの頂点をモーフィング元座標として、それら各モーフィング元座標と一義的に対応付けた形で4つのモーフィング元画像データ(I(0,0)、I(1,0)、I(1,1)、I(0,1))が用意されている。これらのモーフィング元画像データは、例えば外部メディアに書き込まれた状態でメディアドライブ58にて読み取られ、モーフィング処理LSI55を経由してグラフィックメモリの各格納メモリエリアに転送される。なお、これらのモーフィング元画像データは、データポリモーフィング装置1に接続された通信網を介して外部の配信元からダウンロードして取得するようにしてもよい。

30

【0024】

モーフィング元画像データの内容、つまりポリモーフィングの対象とする画像の内容は特に限定されない。例えば、図6に示すように、モーフィング元画像データの1つを人物の顔画像500Aとし、1つを動物の顔画像500Bとすれば、人物と動物が混合された架空の合成顔画像500Mがポリモーフィング結果として得られ、奇抜な効果を上げることができる。

40

【0025】

各モーフィング元画像データ500A, 500B上には、顔上の特徴部を表す複数個の対応点hpが設定されており、モーフィング合成比率を指定することにより、両モーフィング元画像500A, 500Bの対応点hpの座標を該合成比率により補間して、合成顔画像500M上の対応点位置として演算する。また、各対応点hp近傍の所定位置関係を

50

なす複数個のピクセルの出力設定値も、ポリモーフィング処理に際して上記の合成比率により補間演算され、合成顔画像500M上の対応点近傍のピクセルが当該補間演算された出力値に設定される。対応点hpの一部は、顔の輪郭や、目、眉、口、鼻といった部位の輪郭を示す経路に沿って複数配置することができ、その経路内のピクセル同士を、出力設定値の補間演算を行なうための対応ピクセル群として指定することができる。

【0026】

以下、本実施形態では、人物の顔画像同士をポリモーフィングする場合を例にとる。例えば、父母の幼少時の顔画像を合成すれば、その両親に生まれてくると推定される子供の顔画像をモーフィング結果として得ることができる。しかし、ここでは、同一人物の異なる表情の顔画像をポリモーフィングする場合を考える。図7に示すように、モデルデータ配置空間MSPは、精神活性度（覚醒度）を縦軸に、横軸に愉快度を定めた2次元感情平面として規定されており、単位セルHCBは長方形である。そして、該長方形の頂点をなすモーフィング元座標点C, D, A, Bに対応する4つの顔画像データIM1, IM2, IM3及びIM4は、該2次元感情平面MSP上にて対応する各モーフィング元座標点C, D, A, Bの示す精神活性度及び愉快度に対応した、同一人物の4つの表情の顔画像データとしている。

10

【0027】

該2次元感情平面は、いわゆるラッセル・メーラビアン感情平面の概念に基づくものであり、個々の象限は喜怒哀楽に対応した4つの精神状態、すなわち、盛り上がり状態（喜：精神活性度大/愉快）、怒り・興奮状態（怒：精神活性度大/不愉快）、落胆・倦怠状態（精神活性度小/不愉快）を表す。そして、原点O（感情的特徴の少ない中立的な精神状態を示す）からの距離が大きくなるほど、個々の象限に特徴付けられた喜怒哀楽の精神状態が強く反映される形になる。

20

【0028】

例えば、図7に示すように、単位セルHCBをこの4つの象限にまたがるように設定すると、各頂点をなすモーフィング元座標点C, D, A, Bには、同一人物の典型的な喜怒哀楽の4つの表情を顔画像データIM1, IM2, IM3及びIM4として配置できる。そして、記単位セルHCB（2次元感情平面MSP）上にて所望の感情状態に対応するモーフィング先座標点pxを入力部57からの入力情報に基づき設定し、そのモーフィング先座標点pxが規定する重み付けにて後述の方法により顔画像データIM1, IM2, IM3及びIM4をポリモーフィングすれば、所望の感情状態に対応する顔画像を自由に合成することができる。

30

【0029】

顔画像データIM1, IM2, IM3及びIM4は、同一人物の顔画像であり、顔の輪郭や、目、眉、口、鼻、頭髪などの部位の輪郭も、感情に応じた変形を除けばほぼ近接した形状となるから、対応点hpの設定も容易である。その結果、4枚の顔画像データを元に、その人物のあらゆる感情状態に対応した顔画像を自然に表現することが可能となる。また、従来のポリモーフィング技術では単位セルとして三角形を用いていたので、1つの三角形では喜怒哀楽の4つの感情をカバーすることができなかったが、本発明では長方形単位セルを用いることで4つの感情にまたがってより自由で自然な感情表現が可能となる。なお、図7では顔画像データは二次元顔画像データであったが、三次元顔画像データを用いることももちろん可能であり、また、人物の顔の実写画像データに代えて絵画やイラストによる顔画像データを採用することも可能である。

40

【0030】

以下、長方形単位セルを、各辺の長さを1とした正方形とし、入力されるモーフィング先座標点px（二次元ベクトル）を (x_1, x_2) （ $0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1$ ）と表記して、本発明に従うポリモーフィング方法のアルゴリズムについて説明する。図3Aに示すように、長方形単位セルHCBは2本の軸を有する。それらを X_1, X_2 と表記する。モーフィング先座標点pxは $X_1 = x_1, X_2 = x_2$ となる長方形単位セルHCBの内点をなす。ここでは、 $x_1 = 0.3, x_2 = 0.8$ の例を示している。最終的に得たいの

50

は、4個の画像 $I(b_1, b_2)$ から、モーフィング先座標点 p_x に対応する合成画像 $I(x_1, x_2)$ を求めることである。ここでは、モデルデータ配置空間 MSP の次元数 M が2なので、具体的な計算は2ステップで完了する(前述の通り、モデルデータ配置空間 MSP の次元数が M であれば、計算は M ステップ必要となる)。以下、各ステップについて説明する。

【0031】

第一ステップでは、図3Bに示すように、長方形単位セル HCB の互いに平行対向する1対の各辺(図3Aでは $pa - pb$ と $pd - pc$)へのモーフィング先座標点 p_x の正射影点 pe, pf を分点として求める。そして、各辺の両端のモーフィング元座標点に対応する顔画像データ(モデルデータ列)同士(つまり、 $pa: I(0, 0)$ と $pb: I(1, 0)$ 、及び $pd: I(0, 1)$ と $pc: I(1, 1)$)とを梹子の関係に従う重みにて補間演算する形で、図6を用いて既に説明した周知の方法にて一次モーフィングすることにより、1対の中間画像データ($I(0.3, 0)$, $I(0.3, 1)$)を得る。

10

【0032】

続いて、第二ステップでは、図3Cに示すように、上記の正射影点 pe, pf が張る線分上のモーフィング先座標点 p_x を新たな分点とみなし、その分点に関して1対の中間データ列($pe: I(0.3, 0)$, $pf: I(0.3, 1)$)を梹子の関係に従う重みにて二次モーフィングすることにより、最終的な合成画像データ(合成データ列)を得る。合成画像データは図1のモニタ59やプリンタ60から出力でき、通信やネットワークにて接続された外部のシステムに向けて転送出力することも可能である。また、モーフィング先座標点 p_x の入力値も、入力部57からの入力情報として取得する方法以外に、通信やネットワークにて外部システムから取得することが可能である。

20

【0033】

図5に、次元数 M が2の場合に係るポリモーフィングのアルゴリズムを概念的に示している。モーフィング元座標点 pa, pb, pc, pd をそれぞれ A, B, C, D として、長方形 HCB (超直方体)のモーフィング先座標点 p_x を通って各辺(CA, DB 及び CD, AB)と平行な2本の直線(2個の平面)で切断する。これにより、長方形 HCB は、それぞれモーフィング先座標点 $X(p_x)$ を共有し、かつ長方形 HCB の頂点をなすモーフィング元座標点を排他的に1個ずつ取り合う4個(2M個)の部分長方形 SCB 、具体的には長方形 $CKXN$ (面積: S_b)、 $NXL D$ (面積: S_a)、 $KAM X$ (面積: S

30

【0034】

そして、各部分長方形(部分直方体) SCB の長方形(超直方体) HCB に対する相対面積(相対体積)を、当該部分長方形 SCB に含まれるモーフィング元座標点の長方形 HCB の対角線方向反対側に位置するモーフィング元座標点(すなわち、 pa に対しては pd 、 pb に対しては pc 、 pd に対しては pa 、 pc に対しては pb)への重みとする形でポリモーフィングを行なう。すなわち、長方形 HCB の面積を S_0 とすれば、合成画像 P_x は、

$$P_x = (1/S_0) \times (S_a \cdot P_a + S_b \cdot P_b + S_c \cdot P_c + S_d \cdot P_d) \quad (13)$$

40

にて合成することができる。

【0035】

図3A~図3Cでは、辺 $pa - pb$ 及び辺 $pd - pc$ から演算を開始する形で段階的にポリモーフィング演算を行なったが、辺 $pa - pd$ 及び辺 $pb - pc$ から演算を開始しても、最終的に得られる結果はいずれも(13)式の結果と等価である。図5にその確認のための計算例を示している。すなわち、線分 DB へのモーフィング先座標点 p_x の正射影点を L とし、線分 CA へのモーフィング先座標点 p_x の正射影点を KL とすれば、線分 DB 側の中間画像データ PL が図中の式(11)により、線分 CA 側の中間画像データ PK が図中の式(12)により補間計算される。そして、線分 KL 上にはモーフィング先座標点 X が存在するので、これを分点として一次中間画像 PL 及び PK を用いて同様の補間計

50

算を行なうと、(13)式通りの合成画像 P_x が得られることは、幾何学的に容易に理解できる。なお、 P_x を a 及び a を用いて表した結果を(17)式に示している。

【0036】

なお、図7に示すように、4つの象限にまたがる長方形セルに対し、原点 O と各座標軸との交点 $G H E F$ とを新たなモーフィング元座標点として追加し、対応する5つの顔画像データ(モデルデータ列)を用意することができる。この場合、原点 O を中心として象限毎に長方形単位セル($O F C G$, $O G D H$, $O H A E$, $O E B F$)が都合4つ隣接形成される。この場合、モーフィング先座標点 p_x が決定された後、そのモーフィング先座標点 p_x がどの長方形単位セルに属するかを判定した後、関係する長方形単位セルの顔画像データ(モデルデータ列)を用いて同様のポリモーフィング処理を実施することができる。

10

【0037】

また、本発明において、モデルデータ配置空間 $M S P$ の次元数 M は2以上の値であれば制限はない。例えば、モデルデータ配置空間が三次元であれば単位セルは直方体となるが、この場合は、次のような3段階のモーフィングを実施すればよい。すなわち、3組ある平行対向面(長方形)のいずれを選んで、各平行対向面へのモーフィング先座標点の正射影点を求め、それら正射影点について各面をなす長方形毎に、二次元の場合と全く同様の2段階のモーフィングを行ない、一次中間データ列を得る。そして、各面上の正射影点同士が張る線分上にモーフィング先座標点の正射影点を新たに分点として設け、その分点に関して1対の一次中間データ列同士を梃子の関係に従う重みにて三次モーフィングすることにより最終的な合成データ列を得る。図4は、次元数 M を n に一般化した場合のアルゴリズムを示すフローチャートである。

20

【0038】

すなわち、上記ポリモーフィング演算のアルゴリズムは、実は、次のような補間合成演算を逐次的に実行して合成画像 P_x を得ると数学的に全く等価である。すなわち、超直方体 $H C B$ の各座標軸方向に隣接する2つのモーフィング元座標点間にて、それらモーフィング元座標点が張る線分へのモーフィング先座標点 p_x の正射影点を分点とする形で、梃子の原理により一次中間画像を合成する。次いで、超直方体 $H C B$ の各面の対向する2辺について得られた一次中間画像に対し、対応する正射影点が張る線分についてモーフィング先座標点 p_x の正射影点を新たに分点として設け、その分点に関してそれら一次中間画像同士を梃子の原理により合成し、二次中間画像とする(以上、S4)。この一連の処理(S4)を、分点がモーフィング先座標点 X にたどり着くまで繰り返す(S5 S3)

30

【0039】

なお、上記本発明のポリモーフィング方法は、モデルデータ列を画像データから音声データに置き換えても矛盾なく実施することができる。音声モーフィングの場合は、モデルデータ列は音声波形データ(ないし、そのパワースペクトラムプロファイルあるいはケプストラムプロファイル)を構成する。波形データやスペクトラムないしケプストラムのプロファイルは、二次元平面上に描画が可能であるから、原理的にはこれを画像とみなすことで画像モーフィングと同様の取り扱いは可能である。

【0040】

40

しかし、例えば特許文献3や非特許文献2, 3に開示されているような、すでに周知となっている音声工学特有のテクニックを導入すれば、より簡便で妥当性の高い音声モーフィングが可能となる。例えば、モーフィング元音声波形のパワースペクトラム(図9上)からは、周知のケプストラム解析により、スペクトラム包絡(図9中:主に声道の共振特性を反映した情報となる)とスペクトラム微細構造(図9下:主に、声帯での音源的な特徴を反映した情報となる)を分離できる。この場合、モーフィング元音声波形のスペクトラム包絡とスペクトラム微細構造と個別にポリモーフィングすればよいことになる。

【0041】

スペクトラム包絡については、ピーク点(図9中に丸印で示している)などの特徴点のみを補間演算してもよいし、さらに妥当性の高い方法としては積分スペクトル逆関数を用

50

いる方法もあるが、公知であるため詳細な説明は略する。一方、スペクトラム微細構造は声帯音源からの基本波のピッチを支配する要素であり、基本波を構成する倍音列に対応した多数のピークを有している。このピーク点をそれぞれ周波数方向に補間してピッチを伸縮させる手法が、音声モーフィングでは一般的である。

【0042】

また、2つの音声データをモーフィングする周知のエンジンとしては、非特許文献2, 3に開示されているS T R A I G H Tを本発明においても活用することが可能である。S T R A I G H Tはいわゆるチャンネル・ボコーダのアーキテクチャに基づき、音声からフィルタ情報(スペクトル包絡)と音源情報とを分離し抽出するようになっている。S T R A I G H Tでは、音源の基本周波数に適應する相補的な時間窓と周波数領域でのスプライン関数理論に基づく適應平滑化により、調波位置での値を保存することと音源の周期性により生ずるスペクトル包絡への干渉の完全な除去を両立させている(例えば、Speech Communication, Vol. 27, No. 3-4, pp. 187-207 (1999)を参照)。

10

【0043】

S T R A I G H Tにおいて、音源情報は、基本周波数と、帯域毎の周期成分と非周期成分の割合を表す非周期性指標とから構成される。基本周波数の抽出には、フィルタ群の中心周波数から出力の瞬時周波数への写像における不動点を利用したアルゴリズムが用いられる。非周期性指標は、見かけ上の基本周波数が定数となるような時間軸の伸縮と楕円フィルタに基づき、シミュレーションによる補正を加えて計算される(例えば、Eurospeech '99, Vol. 6, pp. 2781-2784 (1999)参照)。S T R A I G H Tにおける合成音声は、こうして求められた音源情報とスペクトル包絡から計算される。スペクトル包絡は、最小位相のインパルス応答に変換され、群遅延を操作された混合音源(パルス+有色雑音)と畳み込まれてoverl ap and addにより合成音声波形が求められる。

20

【0044】

そして、S T R A I G H Tを用いたモーフィングでは、スペクトル包絡の系列を時間周波数表現として表示し、特徴的な位置に参照とするための点(基準点)を設定する。基準点は、時間方向では子音と母音からなる一つの音節について4~5個、周波数方向では、5000Hzまでに3~5個程度で十分であることが分かっている。モーフィングの第一段階では、それらの基準点が重なるように、一方の試料の時間周波数平面を変形する。こうして対応づけられた時間周波数平面の上で、それぞれの点においてモーフィング率に応じてパラメタを補間し、モーフィングされたパラメタ値を求める。最後に、モーフィング率に応じて、時間周波数平面を変形する。こうして求められたパラメタをS T R A I G H Tの合成部に渡すことにより、モーフィングされた音声合成される。

30

【0045】

図8には、同一人物による同一語彙(例えば「あいしてる」)の音声データをポリモーフィングする例を示している。モデルデータ配置空間M S Pは、図7と同様の2次元感情平面として規定されており、単位セルH C Bは長方形である。そして、該長方形の頂点をなすモーフィング元座標点C, D, A, Bに対応する4つの音声データW V 1, W V 2, W V 3及びW V 4は、該2次元感情平面M S P上にて対応する各モーフィング元座標点C, D, A, Bの示す精神活性度及び愉快度に対応した、同一人物の4つの感情を反映した音声データとしている。

40

【0046】

前述のごとく、2次元感情平面M S Pの個々の象限は喜怒哀楽に対応する4つの精神状態を表す。そして、原点O(感情的特徴の少ない中立的な精神状態を示す)からの距離が大きくなるほど、個々の象限に特徴付けられた喜怒哀楽の精神状態が、同じ語彙であってもアクセントや音高等に強く反映される形になる。

【0047】

そして、図7の場合と全く同様に、単位セルH C Bをこの4つの象限にまたがるように設定すると、各頂点をなすモーフィング元座標点C, D, A, Bには、同一人物の典型的な喜怒哀楽の4つの発声内容を音声データW V 1, W V 2, W V 3及びW V 4として配

50

置できる。そして、記単位セル H C B (2 次元感情平面 M S P) 上にて所望の感情状態に対応するモーフィング先座標点 p_x を入力部 5 7 からの入力情報に基づき設定し、既に説明した画像モーフィングの場合と同じ本発明のアルゴリズム (図 3 A ~ 図 5) に従い音声データ $W V 1$, $W V 2$, $W V 3$ 及び $W V 4$ をポリモーフィングすれば、所望の感情状態に対応する音声データを自由に合成することができる。この音声データは図 1 において、音声合成部 6 1 を介してスピーカ 6 2 から出力することができる。

【 0 0 4 8 】

例えば、共通のモーフィング先座標点 p_x の入力値を用いて、前述の顔画像のモーフィング処理と上記の音声モーフィング処理とを並列に実施し、各結果を同期させて出力させることにより、モーフィング先座標点 p_x の入力値に従い顔の表情と発話内容とを感情連動させた擬人化エージェントを実現可能である。

【 図面の簡単な説明 】

【 0 0 4 9 】

【 図 1 】 本発明のデータポリモーフィング装置の電気的構成の一例を示すブロック図。

【 図 2 】 モデルデータ配置空間と、モデルデータ列をなす画像データの配置関係を示す図

。

【 図 3 A 】 本発明のポリモーフィング方法の手順説明図。

【 図 3 B 】 図 3 A に続く説明図。

【 図 3 C 】 図 3 B に続く説明図。

【 図 4 】 画像のデータ例を示す図。

【 図 5 】 本発明のポリモーフィング方法のアルゴリズムの一例を示すフローチャート。

【 図 6 】 2 画像のモーフィング例を示す説明図。

【 図 7 】 画像ポリモーフィングに係る実施例を示す模式図。

【 図 8 】 音声ポリモーフィングに係る実施例を示す模式図。

【 図 9 】 音声パワースペクトラムの包絡 / 微細構造分離の概念を説明する図。

【 図 1 0 】 従来のポリモーフィング方法の概念説明図。

【 図 1 1 】 同じく計算方法の説明図。

【 符号の説明 】

【 0 0 5 0 】

1 データモーフィング装置

M S P モデルデータ配置空間

C P S 画像空間

p モーフィング元座標点

p_x モーフィング先座標点

p_a , p_b , p_c モーフィング元座標点

h_p 対応点

H C B 超直方体 (単位セル)

5 0 マイコン (ポリモーフィング処理手段)

5 5 モーフィング処理 L S I (ポリモーフィング処理手段)

5 7 入力部 (モーフィング先座標点選定手段)

5 8 メディアドライブ (モデルデータ列取得手段)

5 9 モニタ (合成データ列出力手段)

6 0 プリンタ (合成データ列出力手段)

6 2 スピーカ (合成データ列出力手段)

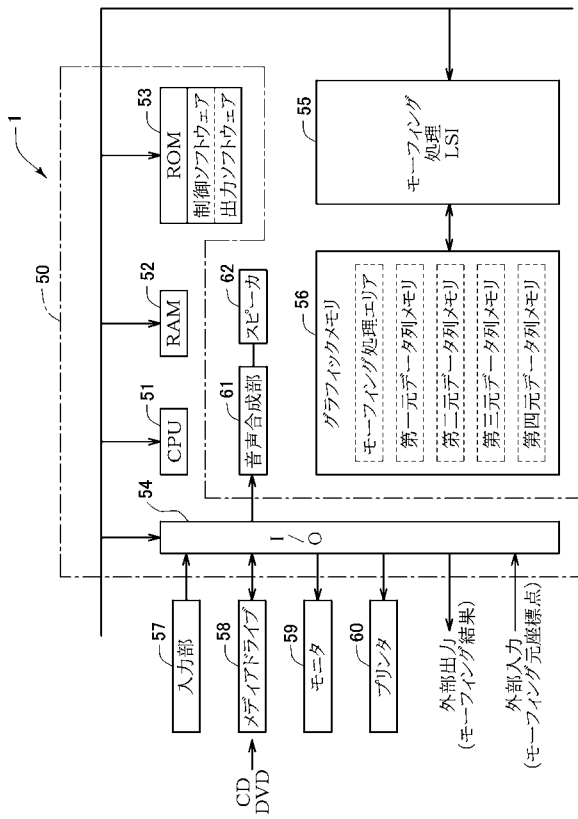
10

20

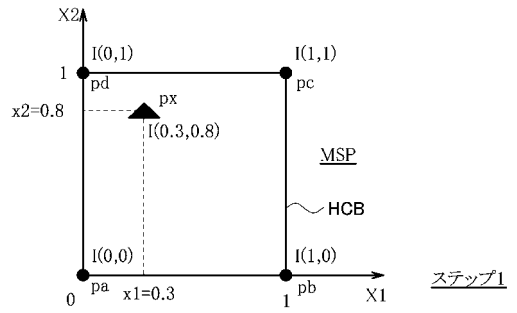
30

40

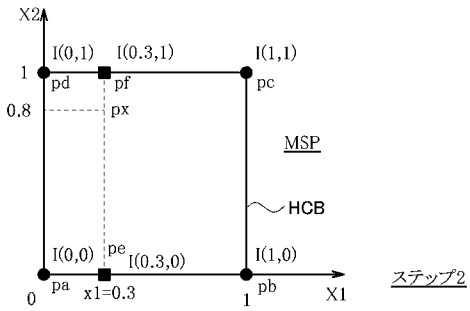
【 図 1 】



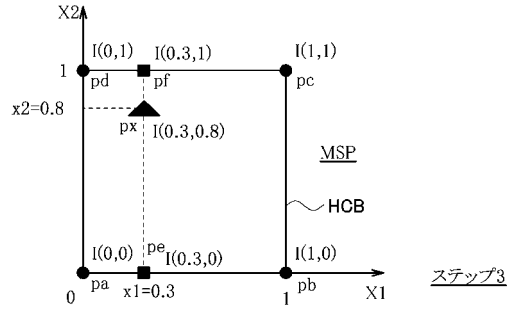
【 図 3 A 】



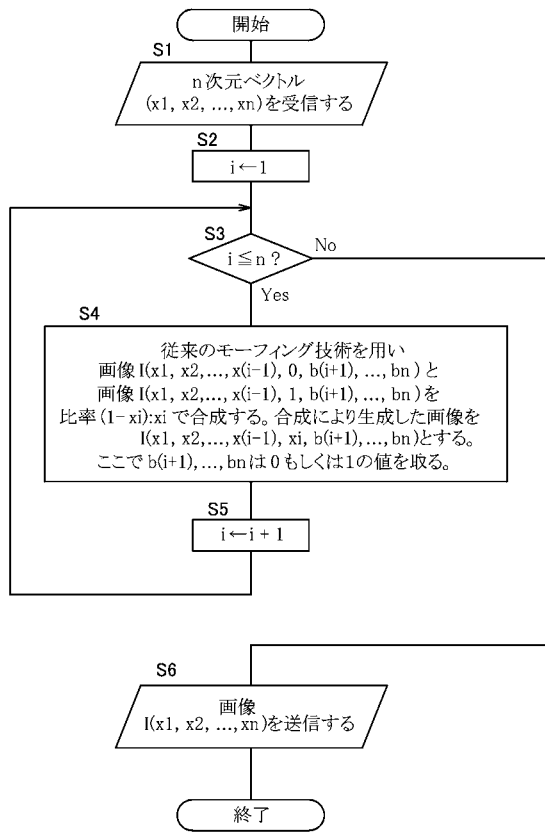
【 図 3 B 】



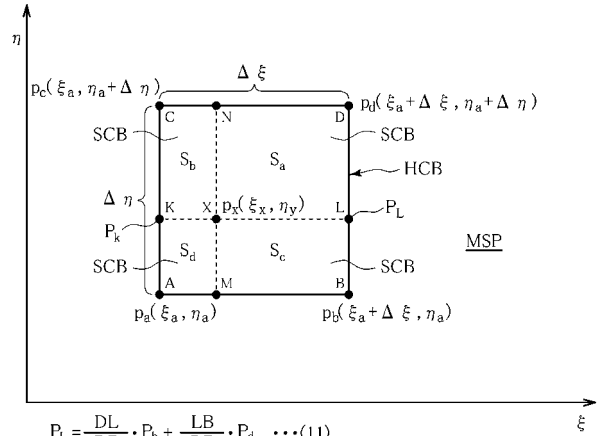
【 図 3 C 】



【 図 4 】



【 図 5 】



$$P_L = \frac{DL}{DB} \cdot P_b + \frac{LB}{DB} \cdot P_d \dots (11)$$

$$P_k = \frac{DL}{DB} \cdot P_a + \frac{LB}{DB} \cdot P_c \dots (12)$$

$$P_x = \frac{DN}{CD} \cdot P_k + \frac{NC}{CD} \cdot P_L$$

$$= \frac{DN}{CD} \left(\frac{DL}{DB} \cdot P_a + \frac{LB}{DB} \cdot P_c \right) + \frac{NC}{CD} \left(\frac{DL}{DB} \cdot P_b + \frac{LB}{DB} \cdot P_d \right)$$

$$= \frac{1}{S_0} (S_a \cdot P_a + S_b \cdot P_b + S_c \cdot P_c + S_d \cdot P_d) \dots (13)$$

ただし、 $S_0 \equiv CD \cdot DB$
 $S_a \equiv DN \cdot DL$ $S_b \equiv NC \cdot DL$ } (14) $\xi'_x \equiv \xi_x - \xi_a$ } (15) とおくと
 $S_c \equiv DN \cdot LB$ $S_d \equiv NC \cdot LB$ } $\eta'_y \equiv \eta_y - \eta_a$ }

$$S_0 = \Delta \xi \cdot \Delta \eta$$

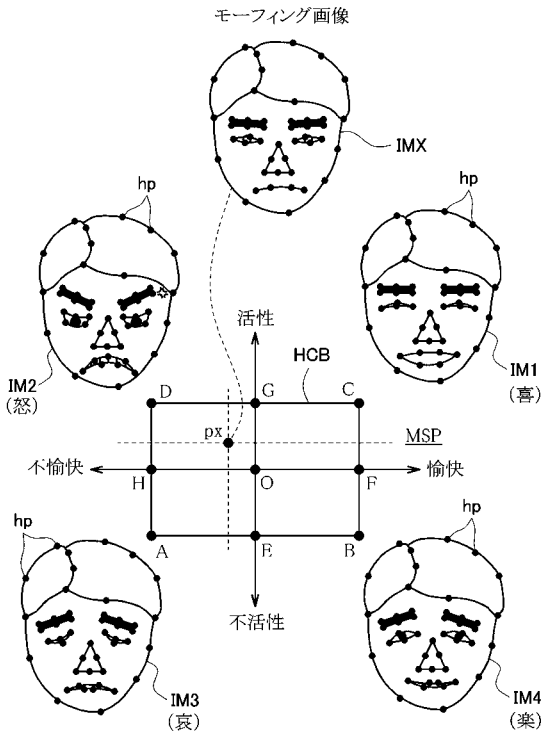
$$S_a = (\Delta \xi - \xi'_x) (\Delta \eta - \eta'_y) \quad S_b = \xi'_x (\Delta \eta - \eta'_y) \quad (16)$$

$$S_c = \eta'_y (\Delta \xi - \xi'_x) \quad S_d = \xi'_x \cdot \eta'_y$$

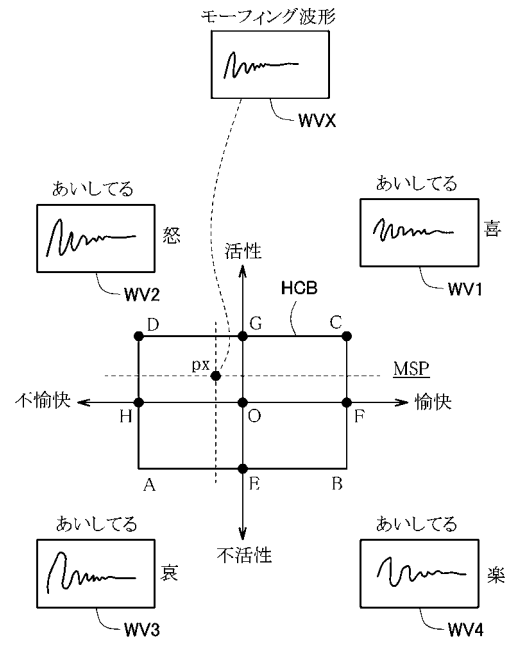
従って、(13)(16)から

$$P_x = \frac{1}{\Delta \xi \cdot \Delta \eta} \{ (\Delta \xi - \xi'_x) (\Delta \eta - \eta'_y) \cdot P_a + \xi'_x (\Delta \eta - \eta'_y) \cdot P_b + \eta'_y (\Delta \xi - \xi'_x) \cdot P_c + \xi'_x \cdot \eta'_y \cdot P_d \} \dots (17)$$

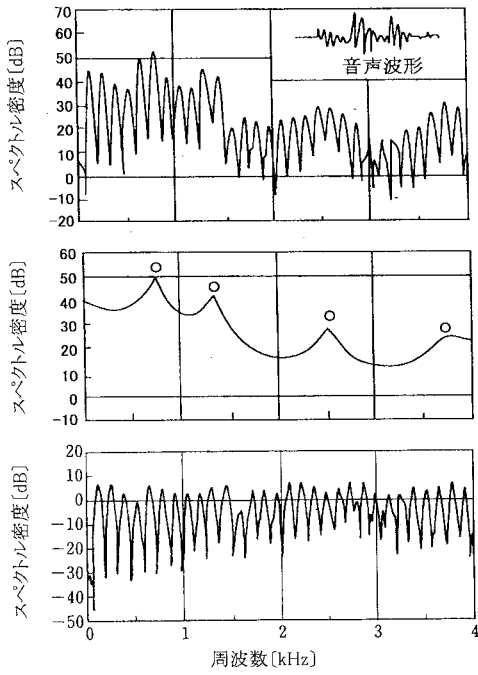
【 図 7 】



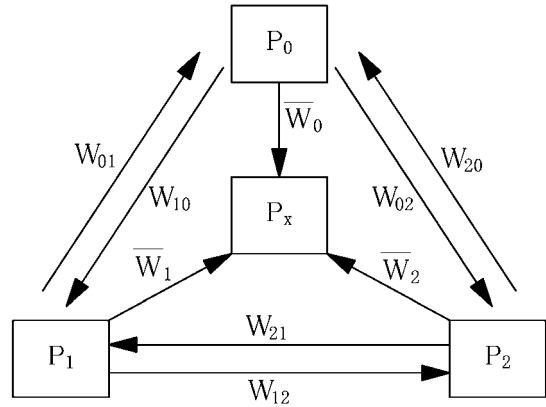
【 図 8 】



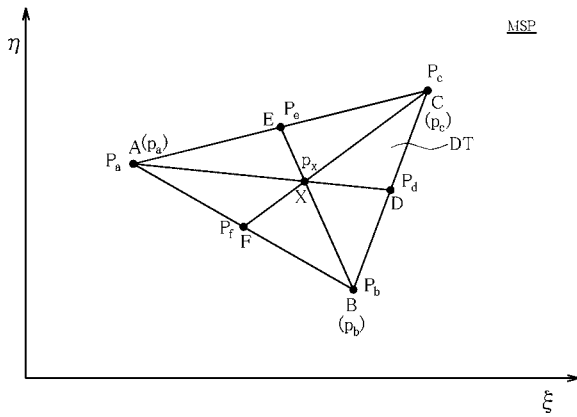
【 図 9 】



【 図 1 0 】



【 図 1 1 】



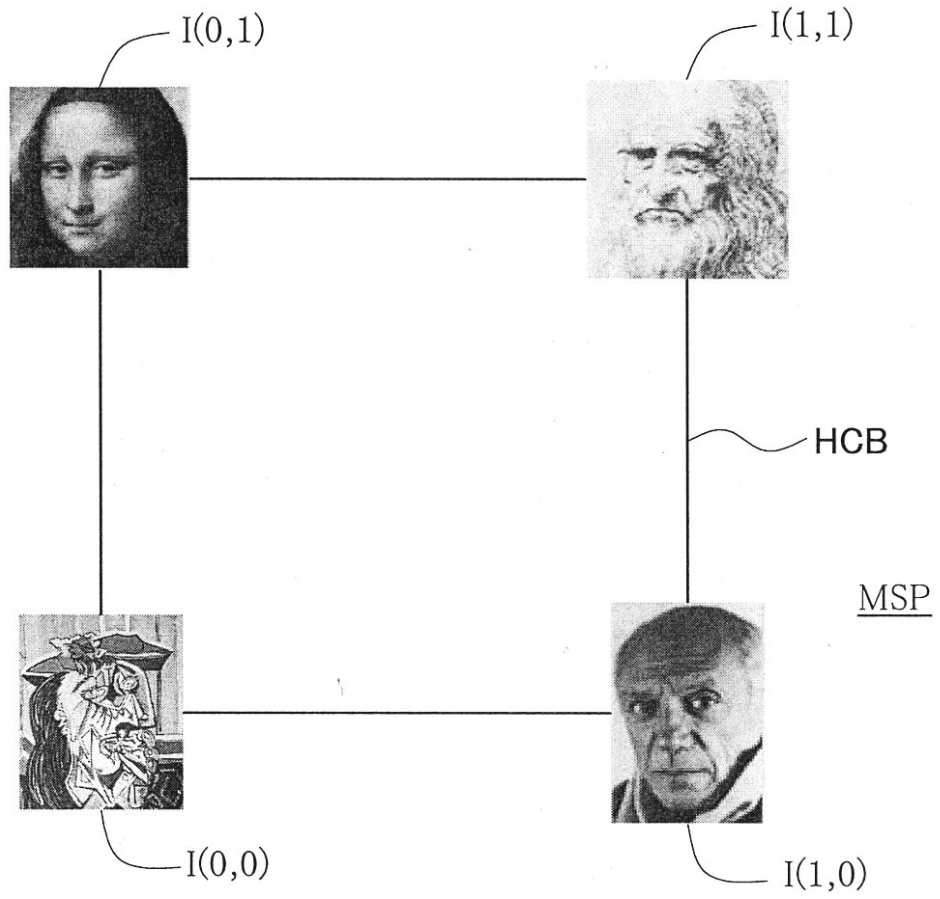
P_x の重心座標 $G^* \equiv (g_a, g_b, g_c)$

$$\left. \begin{aligned} g_a &= \frac{DX}{AD} \\ g_b &= \frac{EX}{BE} \\ g_c &= \frac{FX}{CF} \end{aligned} \right\} (1)$$

$$\left. \begin{aligned} P_d &= \frac{CD}{BC} \cdot P_c + \frac{BD}{BC} \cdot P_b \\ P_e &= \frac{AE}{CA} \cdot P_a + \frac{CE}{CA} \cdot P_c \\ P_f &= \frac{BF}{AB} \cdot P_b + \frac{AF}{AB} \cdot P_a \end{aligned} \right\} (2)$$

$$P_x = g_a \cdot P_d + g_b \cdot P_e + g_c \cdot P_f \quad (3)$$

【 図 2 】



【 図 6 】

