

(51)Int.Cl.

G 0 6 T 13/40

(2011.01)

F I

G 0 6 T 15/70

B

請求項の数9 (全18頁)

(21)出願番号 特願2001-250583(P2001-250583)
 (22)出願日 平成13年8月21日(2001.8.21)
 (65)公開番号 特開2003-058906(P2003-58906A)
 (43)公開日 平成15年2月28日(2003.2.28)
 審査請求日 平成20年8月8日(2008.8.8)

(73)特許権者 503360115
 独立行政法人科学技術振興機構
 埼玉県川口市本町四丁目1番8号
 (74)代理人 100110191
 弁理士 中村 和男
 (72)発明者 星野 准一
 茨城県つくば市春日一丁目6-2-501
 審査官 西出 隆二

最終頁に続く

(54)【発明の名称】人物動画作成システム

(57)【特許請求の範囲】

【請求項1】

2次元的な人物の線画から動画を生成することができる動画像生成システムであって、
 該システムは、人物の一連の動作に対する3次元の姿勢パラメータである動作セグメント
 を複数蓄積するデータ蓄積部を有しており、
 動画を作成するために、2次元的な人物の線画により記述するストーリーボード記述部と

前記ストーリーボード記述部に入力した2次元的な線画から、3次元の姿勢パラメータを
 推定する姿勢推定部と、

前記姿勢推定部からの複数の姿勢パラメータに基づいて、姿勢間の補間を行うために、前
 記データ蓄積部から、複数の姿勢パラメータに当てはまる動作セグメントを用いて補間す
 ることで、動作を生成する動作生成部と、
 前記生成した動作を表示する画像表示部と
 を備えることを特徴とする動画像生成システム。

【請求項2】

請求項1に記載の動画像生成システムにおいて、
 前記動作生成部は、前記姿勢推定部で得られた複数の姿勢パラメータを固有空間に投影し
 て、投影した固有空間内の各点との距離がもっとも小さい動作セグメントを選択すること
 を特徴とする動画像生成システム。

【請求項3】

10

20

請求項 1 又は 2 に記載の動画像生成システムにおいて、
前記ストーリーボード記述部は、時間軸上の関係を記述することができることを特徴とする動画像生成システム。

【請求項 4】

請求項 1 ~ 3 のいずれかに記載の動画像生成システムにおいて、
ストーリーボード記述部を複数備え、各ストーリーボード記述部で記述されたストーリーボードを共通の蓄積部に蓄積することができることを特徴とする動画像生成システム。

【請求項 5】

請求項 4 に記載の動画像生成システムにおいて、
前記蓄積するストーリーボードに対して、進行状態を示すように、表示状態を変化させることができることを特徴とする動画像生成システム。 10

【請求項 6】

請求項 1 ~ 5 のいずれかに記載の動画像生成システムにおいて、
人物が動作している 3 次元の姿勢パラメータを入力し、該姿勢パラメータを分割して動作セグメントとして、前記データ蓄積部に蓄積する動作入力部をさらに備えることを特徴とする動画像生成システム。

【請求項 7】

請求項 1 ~ 6 のいずれかに記載の動画像生成システムにおいて、
前記動作入力部における 3 次元の姿勢パラメータの分割は、前記姿勢パラメータを固有空間に投影して動作曲線を生成し、該動作曲線を最小記述規範により分割して、動作セグメントを得ることにより行うことを特徴とする動画像生成システム。 20

【請求項 8】

請求項 1 ~ 7 のいずれかに記載の動画像生成システムをコンピュータ・システムに構成させるためのプログラム。

【請求項 9】

請求項 8 に記載の動画像生成システムをコンピュータ・システムに構成させるためのプログラムを格納した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、2 次元的な線画から人物動画を生成することができる動画像生成システムであり、特に時間軸上の前後関係を記述することができる動画像生成システムに関する。 30

【0002】

【背景技術】

近年のエンターテインメントや教育などのマルチメディア・コンテンツの制作現場では、CG (コンピュータ・グラフィクス) による複数の仮想人物の相互作用や、仮想人物と人間の対話が含まれるコンテンツが増加している。

ところが、従来の CG ソフトウェアで利用されているキーフレーム・アニメーションのような動作生成法や、CG プログラミング言語による動作の記述で、多様な相互作用を記述して生成することは極めて困難である。 40

CG アニメーションを作成するときには、キーフレームにより間欠的な動作を指定して、フレーム間の動作をスプライン曲線などの単純な手法で補間する方法が一般的である。このような手法で自然な動作を生成するためには、多数のキーフレームを設定することが必要となり、熟練も必要となる。また、フレーム単位での動きを直接扱う必要があるため、ストーリーに沿って長い動作を組み立てるための手法はまた別途考える必要がある。

【0003】

CG アニメーションの制作現場では、作品に対する漠然としたイメージを定着させて、様々なアイデアを統合化するための道具として、紙によるストーリーボードが利用されている。ストーリーボードには、台詞に加えて、ト書きによる状況説明や、アニメーション作成時の様々な指定事項 (色や動き方の指定など) を記述する。また、複数のアニメーター 50

がイメージを共有化するための場としても活用されている。

ところが、紙によるストーリーボードでは、アニメーションを作成するプロセスと分離されているという問題がある。そのため、アニメーションを作成するときには、最初から作り直すことが必要となる。また、作業の途中で書き換えたり編集するなどのインタラクティブな操作ができない。

仮想人物の対話を記述する方法としては、T V M L (テレビ番組制作言語 : TV program Making Language) のような特殊なスクリプト言語により動作の記述を行うことが一般的である。しかし、言語による抽象化された記述のみで、多様な動作を表現することは困難である。抽象化された言語で、細かい動作を記述すると、結局はプログラミング言語のように数値レベルでの記述が必要となる。

10

【 0 0 0 4 】

【 発明が解決しようとする課題 】

本発明は、ストーリーボードに描いた2次元的な線画から仮想人物の3次元姿勢を推定し、3次元アニメーションを生成することを目的とする。

【 0 0 0 5 】

【 課題を解決するための手段 】

上記目的を達成するために、本発明は、2次元的な人物の線画から動画を生成することができる動画像生成システムであって、該システムは、人物の一連の動作に対する3次元の姿勢パラメータである動作セグメントを複数蓄積するデータ蓄積部を有しており、動画を作成するために、2次元的な人物の線画により記述するストーリーボード記述部と、前記ストーリーボード記述部に入力した2次元的な線画から、3次元の姿勢パラメータを推定する姿勢推定部と、前記姿勢推定部からの複数の姿勢パラメータに基づいて、姿勢間の補間を行うために、前記データ蓄積部から、複数の姿勢パラメータに当てはまる動作セグメントを用いて補間することで、動作を生成する動作生成部と、前記生成した動作を表示する画像表示部とを備える。

20

上述の構成により、ストーリーボードに記述した2次元的な人物の線画から動画を生成することが可能となる。

【 0 0 0 6 】

前記動作生成部は、前記姿勢推定部で得られた複数の姿勢パラメータを固有空間に投影して、投影した固有空間内の各点との距離がもっとも小さい動作セグメントを選択する。また、前記ストーリーボード記述部は、時間軸上の関係を記述することができ、さらに、ストーリーボード記述部を複数備え、各ストーリーボード記述部で記述されたストーリーボードを共通の蓄積部に蓄積することができる。また、前記蓄積するストーリーボードに対して、進行状態を示すように、表示状態を変化させることもできる。

30

【 0 0 0 7 】

人物が動作している3次元の姿勢パラメータを入力し、該姿勢パラメータを分割して動作セグメントとして、前記データ蓄積部に蓄積する動作入力部をさらに備えることができる。

。

前記動作入力部における3次元の姿勢パラメータの分割は、前記姿勢パラメータを固有空間に投影して動作曲線を生成し、該動作曲線を最小記述規範により分割して、動作セグメントを得ることにより行う。

40

【 0 0 0 8 】

なお、動画像生成システムをコンピュータ・システムに構成させるコンピュータ・プログラムおよびコンピュータ・プログラムを記録した記録媒体も本発明である。

【 0 0 0 9 】

【 発明の実施の形態 】

< 概要 >

本発明のシステムは、2次元的な絵コンテから3次元CGアニメーションを生成するシステムであり、ストーリーボードに記述した2次元的な線画から3次元的な姿勢を自動的に推定する。そして、あらかじめサンプリングして蓄積してあるサンプル動作のデータより

50

、ストーリーボードで記述した間欠的な動作から推定して連続的な3次元動作を生成することができる。

また、ストーリーボードに時間軸上の前後関係を記述する機能を付加することで、複数の仮想人物間の対話のタイミング等を表現できる。更に、複数のCGアニメーターの間でストーリーボードを共有することで、CG作成時の共同作業を支援することも可能となる。以下、上述のシステムの実施形態を図面を用いて詳細に説明する。

【 0 0 1 0 】

< 構成 >

図1は本発明の人物動画像生成システムの機能構成を示す図である。動作信号入力部12と動作分離部14を備えた動作入力部10、ストーリーボード22, 24, 26を備えたストーリーボード記述部20、3次元姿勢推定部32と3次元動作補間部34を備えた動作生成部30、データ蓄積部40、画像表示部50で構成されている。

動作入力部10は、動作信号入力部12を用いてサンプル動作の3次元姿勢を入力し、動作分離部14を用いて、入力されたサンプル動作の自動分割を行う。入力データ蓄積部40は、仮想人物の各関節のパラメータを蓄積する。ストーリーボード記述部20は、利用者がCGを作成したいシーンの2次元的なスケッチをストーリーボード22, 24, 26に記述するのに用いられる。これは、例えば、マウス等のポインティング・デバイスや、タブレットから、ストーリーボード欄に入力することで行われる。動作生成部30では、3次元姿勢推定部32を用いて2次元姿勢からの3次元姿勢の推定を行い、サンプル動作に基づく補間を3次元動作補間部34にて行う。そして、生成したCGアニメーションを画像表示部50に表示する。

次に、本発明のシステムの動作処理を、処理ごとに説明する。

【 0 0 1 1 】

< 2次元姿勢から3次元姿勢の算出 >

ストーリーボードからアニメーションを生成するための最初のステップは、2次元の絵コンテから、3次元モデルの姿勢を推定することである。ストーリーボード上に記述した2次元的な姿勢の絵コンテから、3次元姿勢を推定する手法については、本発明の発明者による特許出願である「動画像生成システム(特願2001-63854)」の明細書中の段落[0010]以降に記述した<モデルに基づく人物動作の推定>に用いた手法を利用する。以下に詳細に説明する。

【 0 0 1 2 】

動作生成部30においては、人体の3次元モデルを当てはめて、姿勢を推定している。図2に、動作生成部30で用いている仮想人物モデル100を示す。図2のように、仮想人物モデル100は、各パーツをオブジェクトで表し、それらを接続することで全体を表している。この仮想人物モデル100は、頭部101、胴体102、右上腕103、左上腕104、右前腕105、左前腕106、右腿107、左腿108、右膝下109、左膝下110の各パーツ(各オブジェクト)で構成している。動作生成部30では、このような仮想人物モデル100を利用して姿勢を推定する。仮想人物モデル100の各パーツ(各オブジェクト)の接続関係は、胴体102を頂点として、腕や頭部といった末端のパーツに向かって親子関係の接続として定めることができる。親パーツが動くと、その子に当たるパーツは親パーツの座標系で規定される接続関係で一緒に動くことになる。この仮想人物モデル100を用いて、以下説明を行う。

図3は、カメラ座標系およびワールド座標系におけるストーリーボード上の線画と3次元人物モデルの関係を示す図である。この図において、動画像はカメラ座標系で記述される画像面150に表示される。仮想人物モデル100上の任意の点を、ワールド座標系を用いて、

【 数 1 】

$$p = (x, y, z)^T$$

とする(上式で括弧の肩に付けられているTは転置行列を表す)と、カメラ座標系に投影した画像面150上の点(X, Y)は、

10

20

30

40

50

【数 2】

$$X = f x / z$$

$$Y = f y / z \quad (f \text{ は画像面 } 150 \text{ を生成するカメラの焦点距離})$$

となる。

【0013】

(姿勢の推定)

以下では、拡張カルマン・フィルタにより、シルエット画像から抽出した中心線と、人物モデルの体軸を一致させる手法を述べる。

ワールド座標系での仮想人物モデル100の各パーツの姿勢 a を表した式を以下に示す。

【数 3】

$$a = [r , d]^T$$

d はワールド座標の原点からの移動成分を表すベクトルである。また、 r はワールド座標の各座標軸である回転軸に沿ったベクトルであり、ベクトルの大きさは回転角に等しい。

【0014】

図4は、画面上の線分の関係と、それに対応する仮想人物モデル100中のパーツとの関係を示す図である。3次元(3D)線分Pの投影された画像面150上の2次元(2D)予測線分がQである。図4における2次元(2D)観測線分Q'と2D予測線分Qを、2次元の画像面上150での始点および終点を表わすワールド座標系のベクトル q_1' , q_2' , q_1 および q_2 で表わした式を以下に示す。

【数 4】

$$Q' = (q_1' , q_2')$$

$$Q = (q_1 , q_2)$$

また、3D線分Pをワールド座標系での線分の始点および終点を表わすベクトル p_1 および p_2 で表わした式を以下に示す。

【数 5】

$$P = (p_1 , p_2)$$

そして、初期位置で観測された2D観測線分Q'と3D線分Pの組を l として、以下の式に示す。

【数 6】

$$l = (P , Q')$$

この3D線分Pと、それに対応する画像面150上での2D観測線分Q'の組 l を求める手法を図5を参照しながら以下に説明する。図5はストーリーボードに記述した絵コンテの線画より人物モデルの初期位置を推定するまでの処理の過程を示す図である。

【0015】

(中心線の推定)

ストーリーボードより入力した2次元的な線画から中心線を推定する際には、手や足などの部位が胴体に隠れた場合、姿勢を推定できないため、絵コンテに記述した仮想人物のシルエット画像を用いて行う。以下の手順で中心線を推定する。

【0016】

1 中心線が抽出しやすい姿勢のシルエット画像をあらかじめ選択してデータ蓄積部40に記憶しておく。このときに、シルエット画像に対応した人物モデルの姿勢パラメータも与えておく。姿勢パラメータは後述する手法により姿勢を微調整することができるので、大まかな値で構わない。

2 図5(a)(i)および(ii)のような、ストーリーボード記述部20からの入力画像系列で、あらかじめデータ蓄積部40に記憶してあるシルエット画像とのマッチングを行い、姿勢が類似しているフレームを選択する。

3 選択したシルエット画像では、人物モデルの大まかな姿勢が分かっているため、画像面150上に各パーツの領域を投影することができる。それぞれの領域内のシルエットを切り出して、主軸を求めて中心線とする。図5(b)(i)および(ii)は入力画像に人物モデルを当てはめ、各パーツの中心線を抽出し、表示した画像を示す。

10

20

30

40

50

【 0 0 1 7 】

上記の手法より、抽出した中心線を利用して、3D線分Pと、それに対応する画像面150上での2D観測線分Q'の組lを求めることができ、後述する手法により人物モデルの正確な姿勢の推定を行うことができる。

【 0 0 1 8 】

(姿勢の調整)

各パーツの姿勢aが正確な場合は、観測画像から得られた画像面150上の2D観測線分Q'とカメラの原点Oとが作る平面Mの上に3D線分Pが含まれる。ところが各パーツの姿勢aに誤差がある場合は、図4(a)のように、画像面150上に表示される線分Q'とカメラの原点Oとが作る平面Mと3D線分Pの間に距離が生じる。この距離の最短距離を、予測誤差 $h(a, l)$ とし、以下の式のように表す。

10

【 数 7 】

$$h(a, l) = \left[\frac{(Rp_1 + d) \cdot n}{(Rp_2 + d) \cdot n} \right]$$

$h(a, l)$ は 2×1 のベクトルである。Rはrから導出される 3×3 の回転行列である。また、nは面Mの単位法線ベクトル

【 数 8 】

$$n = \frac{q'_1 \times q'_2}{\|q'_1 \times q'_2\|}$$

20

である。aに誤差がない場合は、 $h(a, l) = 0$ となり、図4(b)のように画像面150上の2D予測線分Qが2D観測線分Q'と重なる。そして2D観測線分Q'とカメラの原点が作る平面の上に3D線分Pが含まれるようになる。このときの3D線分Pの姿勢aが画像面150上に表示されている2D観測線分Q'に対応しているということになる。 $h(a, l)$ は非線形なので、観測値 $l = \hat{l}_i$ と推定値 $a = \hat{a}_i$ のまわりに1次テーラー展開をして線形近似する。

【 数 9 】

30

$$h(a_i, l_i) = \frac{\partial h}{\partial a} (a_i, l_i)(a - a_i) + \frac{\partial h}{\partial l} (a_i, l_i)(l - l_i)$$

ただし h / a , h / l は偏微分を表す。これらの偏微分係数は、以下のカルマン・フィルタを適用して人物モデルの各パーツの姿勢 $a = [r, d]^T$ の最適値を求める際に利用する。

【 0 0 1 9 】

対応付けがついた2D観測線分Q'と3D線分Pの誤差 $h(a, l)$ が最小になるように、人体姿勢aを補正し、位置の誤差分布 x_i を減少させる。直前に示した $h(a, l_i)$ の線形近似式をゼロとおいて書き直すと、次のような観測方程式が得られる。

40

【 数 1 0 】

$$z_i = H_i a + \epsilon_i$$

ただし、

【 数 1 1 】

$$z_i = \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial a} \hat{a}_{i-1} - h(\hat{a}_{i-1}, \hat{l}_i)$$

$$H_i = \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial a}$$

$$v_i = \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial l} (l_i - \hat{l}_i)$$

10

である。ここで σ_i の分散 B_i は次のように求める。

【数12】

$$E[v_i v_i^T] = B_i = \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial l} L_i \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial l}^T$$

【0020】

以上の観測方程式より、人体パーツの姿勢 x_i の2乗誤差を最小にする推定値は、カルマン・フィルタで与えられる。

【数13】

20

$$\hat{a}_i = \hat{a}_{i-1} + K_i [z_i - H_i \hat{a}_{i-1}]$$

$$K_i = \sigma_{i-1} H_i (H_i \sigma_{i-1} H_i^T + B_i)^{-1}$$

$$\sigma_i = \sigma_{i-1} - K_i H_i \sigma_{i-1}$$

人体姿勢 a_i と誤差分布 σ_i は、新しい線分の対応付けが行われるたびに更新される。以上の姿勢推定を人体の各部位に適用することで、ストーリーボード上の2次元の線画の3次元モデルにおける姿勢を推定する。本発明のシステムでは、胴体の姿勢を最初に求めて、上腕、上脚、下腕、下脚と胴体に近い順番に中心線とのマッチングを行い、姿勢の推定を行う。

30

このようにして、図5(a)(i)および(ii)の入力画像に初期姿勢を推定し、図5(c)(i)および(ii)のように人物モデルを重ね合わせた画像を得ることができる。

【0021】

(2D、3D線分の対応付け)

上述の手法では、シルエット画像のマッチングにより、抽出した中心線を利用して、3D線分Pと対応する2D観測線分Q'の組1が求めたとして説明した。しかし、撮影環境内に中心線として抽出可能な複数の線分がある場合、1つの3D線分に対して複数の2D線分の候補があることとなる。そこで以下に示す2つの手法を併用して対応付けを行う。まず、2D画像面150上で見たときに、2D観測線分Q'に射影したときに射影像が2D観測線分Q'と重なる線分を候補として選択する。次に、推定値の誤差分布と、2D、3D線分の誤差分布を考慮に入れて、統計的に一致すると判定できる線分を選択する。

40

【0022】

1 2次元画面上での対応付け

図4を用いて2次元画像面150上で対応付けを行う手法を示す。まず、図5(a)のように、3D線分Pを画面上に投影して、2D予測線分Q = (q₁, q₂)を求める。また、観測画像から2D観測線分Q' = (q₁', q₂')を得る。2D観測線分Q'に対して射影した2D予測線分Qの射影像が重複領域をもつ場合は、対応の候補として選択する。重複領域の判定方法としては、2D予測線分Qを2D観測線分Q'に射影したときに、

50

q_1 あるいは q_2 のどちらかの点が Q' の内部にあれば、2D 予測線分 Q の射影像が 2D 観測線分 Q' に重なっていると考えることができる。まず、 q_1 を Q' に射影した場合の座標 c は、 Q' に沿う単位方向ベクトルを

【数 14】

$m = N [q_2' - q_1']$ ($N [\]$ はベクトルの正規化を表わす) とすると、

【数 15】

$$c = q_1' + m \cdot (q_1 - q_1')$$

である。そして、 c が Q' の内部に含まれる条件は、

【数 16】

$$(c - q_1') \cdot (c - q_2') < 0$$

となる。この操作を q_1, q_2, q_1', q_2' に関して行い、上記の条件を満たさない線分は棄却する。

【0023】

2 誤差分布を考慮に入れた対応付け

人体姿勢の誤差分布 x_i と、2D, 3D 線分の誤差分布 l_i を使って χ^2 検定を行い、統計的に一致しないと考えられる線分を棄却する。ここでは、2D-3D 線分が一致していると仮定して、仮説検定を行う。人体姿勢 $a = [r, d]^T$ と観測値 $l = (P, Q')$ の誤差分布を、正規分布 $a_i \sim N(0, P_0)$, $l_i \sim N(0, L_0)$ と仮定する。 $h(\hat{a}_{i-1}, \hat{l}_i)$ の 1 次、2 次統計量は、

【数 17】

$$E [h(\hat{a}_{i-1}, \hat{l}_i)] = 0$$

【数 18】

$$E [h(\hat{a}_{i-1}, \hat{l}_i)] = Z_i = \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial a} P_{i-1} \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)^T}{\partial a} + \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)}{\partial l} L_{i-1} \frac{\partial h(\hat{a}_{i-1}, \hat{l}_i)^T}{\partial l}$$

となる。これよりマハラノビス距離 k を次のように求めることができる。

【数 19】

$$k(\hat{a}_{i-1}, \hat{l}_i) = h(\hat{a}_{i-1}, \hat{l}_i)^T Z_i^{-1} h(\hat{a}_{i-1}, \hat{l}_i)$$

$h(\hat{a}_{i-1}, \hat{l}_i)$ の誤差分布は $h(\hat{a}_{i-1}, \hat{l}_i) \sim N(0, Z_i)$ であるため、 $k(\hat{a}_{i-1}, \hat{l}_i) \sim \chi^2(\)$ となる (ただし $\ = \text{Rank}(Z_i)$)。 $\chi^2(\)$ 分布から、確率に基づくしきい値を設置することができる。例えば $k(\hat{a}_{i-1}, \hat{l}_i) < 9.21$ は、 Z_i がランク 2 ($\ = 2$) のときに 99% の線分が受け付けられることを意味している。しきい値以上の線分は、統計的に一致すると考えられないため、棄却することができる。複数の線分が候補として残った場合は、 $k(\hat{a}_{i-1}, \hat{l}_i)$ の距離値が最小のものを選択する。

以上の 2 つの手法を併用して、2D 予測線分 Q が 2D 観測線分 Q' に一番近い 3D 線分 P と画像面 150 上で対応する 2D 観測線分 Q' の組 $l = (P, Q')$ を求めることができる。

なお、上述の人物モデルの姿勢は、利用者が人物モデルの姿勢を推定して画像面に投影し、入力画像に重ね合わせて当てはめることにより求めてもよい。

【0024】

< サンプル動作に基づく補間 >

ストーリーボードから連続的な人物動作を生成する際には、フレーム間の動作の間隔が広い場合、単純に動作を補間するだけでは自然な動作とならないという問題がある。そのため、データ蓄積部 40 に蓄積したサンプル動作を当てはめることで、欠落している動作情報を補う。

10

20

30

40

50

サンプル動作を取得する手法としては、ビデオ映像で人物動作を記録して、動画像処理によって、手動もしくは自動にて3次元姿勢を求めることができる。また、磁気センサや光学式センサによるモーション・キャプチャ装置を利用することもできる。

ここで、マッチングの際の計算量を低減するとともに、サンプル動作の変形により多様な動作を生成するために、固有空間へ投影することでパラメータ数を減少させる。これについて、以下説明する。

まず、時刻 t における人体の姿勢パラメータを、

【数 2 0】

$$x_t = [x_t^1, x_t^2, \dots, x_t^m]^T$$

10

と表す。 m は姿勢パラメータの総数である。本明細書では、等間隔に撮影した動画像から姿勢パラメータを推定するため、 t はフレーム番号に対応する。 N を教示動作の総サンプル数とすると、人体の姿勢の平均 c は、

【数 2 1】

$$c = \frac{1}{N} \sum_{t=1}^N x_t$$

である。教示動作の各姿勢から、姿勢の平均 c を差し引いた行列を X とすると、姿勢データ列の共分散行列 Q は、

20

【数 2 2】

$$Q = X X^T$$

である。次の固有方程式を解いて固有値と固有ベクトルを求める。

【数 2 3】

$$\lambda_i e_i = Q e_i$$

k 次元の固有空間は、固有値の中から大きい順に k 個選んだ固有ベクトルを、基底ベクトルとすることにより得られる。固有値 m 個の中の大きい方から n 個の固有値が占める割合を累積寄与率

【数 2 4】

$$S_n = \frac{\sum_{j=1}^n \lambda_j}{\sum_{j=1}^m \lambda_j}$$

30

と呼ぶ。なお、本明細書では多項式近似を容易にするために、上位 2 つの固有ベクトルを利用した。

固有空間内の動作曲線は、姿勢パラメータを固有空間に射影することで得られる。

【数 2 5】

$$g_t = [e_1, e_2]^T (x_t - c)$$

図 6 は以上の手法により求めた固有空間における人体の動作曲線と、各座標値に対応する 3 次元姿勢の例を示す図である。

40

【0 0 2 5】

< 動作曲線の自動分割 >

図 7 は固有空間の動作曲線の分割手法を示す図であり、図 7 (a) は分割前の動作曲線、図 7 (b) は分割後の動作曲線を示す。入力画像から人物の姿勢パラメータを推定した後、以下の 2 つの手順で連続動作の分割と記述を行う。

1 上記のサンプル動作に基づく補間を求める手法を利用して、姿勢パラメータを固有空間に投影して次元数を低減し、動作曲線 (action trajectory) を生成する。

2 最小記述規範により動作曲線を分割する。曲線を記述するモデルとしては、多項式モデルを利用する。

最小記述規範では、あるモデルを仮定したときに、モデルによる記述長が最短になる記述

50

を選択し、モデルの当てはめによる測定データの記述と、近似誤差のトレードオフを扱う。つまり、長時間の動作を簡単なモデルで記述した場合、モデルの記述長は減少するが、モデルから外れる値が増加する。一方で、動作を細かな時間間隔で区切り、モデルで近似をした場合、入力動作からの誤差は減少するが、モデルの記述に必要なビット数が増加する。そのため、モデルの記述長と、外れ値の記述長の両方が最小となるように最適化を行う。

【 0 0 2 6 】

動作曲線を近似するモデルとして、2次多項式(2nd order polynomial)を利用し、動作曲線の座標値をそれぞれ $G_1(t)$, $G_2(t)$ とすると、

【数 2 6】

$$G_1(t) = a_1 t^2 + a_2 t + a_3$$

$$G_2(t) = b_1 t^2 + b_2 t + b_3$$

である。これを最小2乗推定により上記の式の2次多項式の係数を推定する。

【数 2 7】

$$A = [a_1 , a_2 , a_3]^T$$

$$B = [b_1 , b_2 , b_3]^T$$

$$= [1 , t , t^2]^T$$

と置くと、次のように書くことができる。

【数 2 8】

$$G_1(t) = A^T A , G_2(t) = A^T B$$

また、ロバスト推定法により A_x , A_y を求めると、

【数 2 9】

$$A = [W^T]^{-1} [W G_1(t)]$$

$$B = [W^T]^{-1} [W G_2(t)]$$

となる。WはCauchyの重み関数を利用した。 は動作セグメントの始点と終点の間で和を計算する操作を表す。

【 0 0 2 7 】

人物の動作曲線は複雑な軌跡を描くことが多いため、全体を1つの多項式で近似することは困難である。そのため、多項式曲線から大きく外れる部分については、再帰的に分割を行い、個々の動作セグメントの近似精度を高める。動作曲線の記述長 $L_{s,k}$ は次の3つの項目から構成される。

【数 3 0】

$$L_{s,k} = L_{parameters} + L_{outliers} + L_{offset}$$

実数値のパラメータの記述長は、 をパラメータの個数、nをパラメータのしきい値とすると、 $(\log_2 n) / 2$ で求めることができる。よって、各パラメータの記述長は次の通りである。

1 モデルの記述長 ($L_{parameters}$)

多項式を表すためには、2つの終点と6つのパラメータを記述することが必要となる。 | を固有空間内の点のパラメータ幅、 | を多項式係数のパラメータ幅、 を解像度とすると、モデルの記述長は次式で表すことができる。

【数 3 1】

$$L_{parameters} = 6 \log_2 \frac{|\zeta|}{\varepsilon} + 4 \log_2 \frac{|\eta|}{\varepsilon}$$

2 外れ値の記述長 ($L_{outliers}$)

外れ値がある場合は、固有空間内のランダムな点として記述することができる。 個の外れ値があった場合の記述長は、次式で表すことができる。

【数 3 2】

10

20

30

40

$$L_{outliers} = 2\lambda \log_2 \frac{|\eta|}{\varepsilon}$$

3 モデルからの誤算の記述長 ($L_{model\ points}$)

モデルからの誤算の記述長は、ガウシアン分布 $\sim N(0, \sigma^2)$ を解像度 ε でサンプリングした場合の記述長に相当する。

【数 3 3】

$$L_{offset} = \frac{1}{2\log_2 e} \left(\frac{x-\mu}{\sigma} \right)^2 + \log_2 \left(\frac{\sigma}{\varepsilon} \right) + \frac{\log_2(2\pi)}{2}$$

分割候補点のところで分割するかどうかを決定するためには、分割前の動作セグメントと、分割後の動作セグメントの記述長を比較する。記述長が減少する場合は、分割点で動作セグメントを分割する。

【数 3 4】

$$L_{s\ k} > L_{s\ k\ 1} + L_{s\ k\ 2}$$

図 8 は一連の動作を固有空間に投影したところを示す図である。また、図 9 は上記の手法により、図 8 に示す動作を自動的に 0 ~ 249 のフレームに分割したところを示す図である。図 9 のグラフに \square 印を示し、数字を添えた部分は、その番号に対応したフレームに関するところである。

20

図 10 は 0 ~ 249 フレームに分割した動作セグメントに対応した 3 次元動作を示す図である。図 10 (a) は歩行動作に対応したフレーム (フレーム 0 ~ 59)、図 10 (b) は座る動作に対応したフレーム (フレーム 60 ~ 129)、図 10 (c) は立ち上がる動作に対応したフレーム (フレーム 198 ~ 249) である。各フレームの右下に記している番号が対応しているフレームを表す。

【0028】

< 人物動作の個性の分析と再構成 >

固有空間内の動作セグメントを変形することで、人物動作の個性の推定と再構成を行う。複数人で同じ動作を行った場合のサンプルを利用して、平均動作と、平均動作からの分散を求め、平均動作からの変動をランダム発生させて、動作セグメントを変形させることで、入力したサンプル動作とは異なる個性を持った動きを生成する。

30

図 11 は 3 名のモデルに同じ動作をさせた場合を平均化し、計算により新たな動作を生成したことを示す図である。図 11 (a) は 3 名が同じ動作をしているところを固有空間に投影したものを示す図である。そして、図 11 (b) は 3 名のモデルの平均動作を求めた後に、3 名の動作の分散を利用して新しい個性を持った動作を生成したところである。

【0029】

< 固有空間内の動作セグメントによる動作の補間 >

あらかじめデータ蓄積部 40 に蓄積してある動作セグメントを利用して、ストーリーボードによって記述された間欠的な姿勢を補間する手法を次に説明する。

40

1 上記のサンプル動作に基づく補間を求める手法を利用して、ストーリーボード上の 2 次元線分から 3 次元姿勢を推定する。

2 求めた 3 次元姿勢を固有空間に投影する。各姿勢に対応する固有空間内の点を p_1, p_2, \dots, p_n とする。

3 固有空間内の点 p_1, p_2, \dots, p_n からの距離が最も小さい動作セグメントを選択する。

図 12 はストーリーボードに入力した絵コンテと、上記の処理によって生成した CG アニメーションの例を示す図である。図 12 (a) はストーリーボードに記述された絵コンテであり、(i) ~ (iii) の順番でストーリーになっている。これを上記の間欠的な姿勢を補間する処理を施し、図 12 (b) (i) ~ (vi) に示すような連続した CG アニメー

50

ションを生成する。

【 0 0 3 0 】

< ストーリーボード・システムにおける対話の記述 >

図 1 3 はストーリーボードに、各キャラクタの動作の速さを制御する機能を追加した例を示す図である。上述した手法では、ストーリーボード間の動作の連結を表すことはできても、例えば同じストーリーボード内で複数のキャラクタが対話をしている場合の動作の前後関係や速さを表現することはできない。そのため、ストーリーボードに各キャラクタに対応したタイムラインの記述欄を追加することで、同一フレーム内の動作の因果関係を記述することができる。

図 1 3 に示すように、ストーリーボード記述部の下部に各キャラクタに対応したタイムラインを制御するための欄 2 2 0 を設けることができる。この欄 2 2 0 には、例えばビデオ映像から推定した動作のタイムライン（動作時間の表示）の各線分と、動作セグメントの開始点、終了点とを対応させる。そして、このタイムラインを動かす（スライダ機能）ことで、対応させた動作セグメントの動作を変更させる。即ち、このタイムラインを長くすると、これに対応したキャラクタの動作は遅くなる。また、このタイムラインの開始点や終了点を移動することで、動作の開始時点や終了時点を制御することができる。このように、この欄に表示されるタイムライン（線分）を変更させることを利用して、微妙な発話のタイミングを調整することができる。

10

【 0 0 3 1 】

< ストーリーボードの共有による共同作業支援 >

ストーリーボード・システムの応用として、複数の CG アニメーターの間でストーリーボードを共有することで、共同作業の支援を行う。また、複数のアニメーターがイメージを共有化するための場としても活用する。

20

図 1 4 は共同作業支援作業を行うためのシステムの構成例を示す図である。サーバ内のデータ蓄積部 4 0 に各種のデータを蓄積し、サーバにはストーリーボード記述部 3 0 2 , 3 0 4 , 3 0 6 を有するクライアントを接続する。各作業者はこれらクライアントからストーリーボードを参照することができる。

【 0 0 3 2 】

図 1 5 は複数制作者による共同作業を支援するための、ストーリーボードを利用した作業工程管理とメッセージの伝達手法を示す図である。

30

1 作業工程管理

複数人で CG アニメーションを作成するときには、どの部分がどの程度作業が進んでいるのかなどの作業工程の進捗を把握することが重要である。本システムでは、ストーリーボードを利用して作業工程の管理に必要な情報の記述を行う。例えば、モーションデータやモデルデータが得られていることを色などで示すことで、どこまでアニメーション作成が進んでいるかを表示している。

2 メッセージボード

複数の作業員間で情報交換を支援する方法として、ストーリーボード記述部上にメッセージボード 2 4 0 を付加することで、メッセージを添付する機能を追加することも可能である。例えば、各キャラクタの台詞、ト書きによる状況説明や、アニメーション作成時の様々な指定事項を記述する。これらのメッセージは、メール送信機能と連動することで、管理者等から特定の作業員に送ることもできる。

40

【 0 0 3 3 】

上述の動作入力部、ストーリーボード記述部、動作生成部、データ蓄積部、画像表示部のそれぞれの処理は、計算機上で実行することもできる。また、その処理を行うためにプログラムを格納した記録媒体から読み出したり、通信回線を介して受信等をしたプログラムを実行する等により、本発明の構成を実現することもできる。この記録媒体には、フロッピー・ディスク、CD-ROM、磁気テープ、ROMカセット等がある。また、通信回線としては、インターネット等がある。

【 0 0 3 4 】

50

【発明の効果】

2次元スケッチから3次元CGアニメーションを生成することができるため、アニメーション作成の生産性を大幅に向上することが可能となる。

そして、3次元のサンプル動作を自動的に分割して蓄積するとともに、蓄積したサンプル動作を利用して、ストーリーボードによる間欠的な動作から連続的な動作を生成することにも利用することができる。また、ストーリーボードに時間軸上の前後関係を記述する機能を付加することで、複数の仮想人物間の対話のタイミングを記述することが容易になる。

そのうえ、複数のCGアニメーターの間でストーリーボードを共有することができるため、共同作業の効率を大幅に向上させることが可能となる。

10

【図面の簡単な説明】

【図1】 本発明のシステムの機能構成を示す図である。

【図2】 仮想人物モデルを示す図である。

【図3】 ストーリーボード上の線画(画面)と3次元人物モデルの関係を示す図である。

【図4】 画面上の線分の関係と、それに対応する仮想人物モデルのパーツとの関係を示す図である。

【図5】 ストーリーボードに記述した絵コンテの線画より人物モデルの初期位置を推定するまでの処理の過程を示す図である。

【図6】 固有空間の動作曲線と、各座標値に対応する3次元姿勢の例を示す図である。

20

【図7】 固有空間の動作曲線の分割手法を示す図である。

【図8】 一連の動作を固有空間に投影したところを示す図である。

【図9】 図8に示す動作を分割したところを示す図である。

【図10】 分割した動作セグメントに対応した3次元動作を示す図である。

【図11】 3名のモデルに同じ動作をさせた場合を平均化し、計算により新たな動作を生成したものを示す図である。

【図12】 ストーリーボードに入力した絵コンテから生成したCGアニメーションの例を示す図である。

【図13】 ストーリーボードに時間フレーム(タイムライン)を追加した例を示す図である。

30

【図14】 共同作業支援作業を行うためのシステムの構成例を示す図である。

【図15】 ストーリーボードを利用した作業工程管理とメッセージの伝達手法を示す図である。

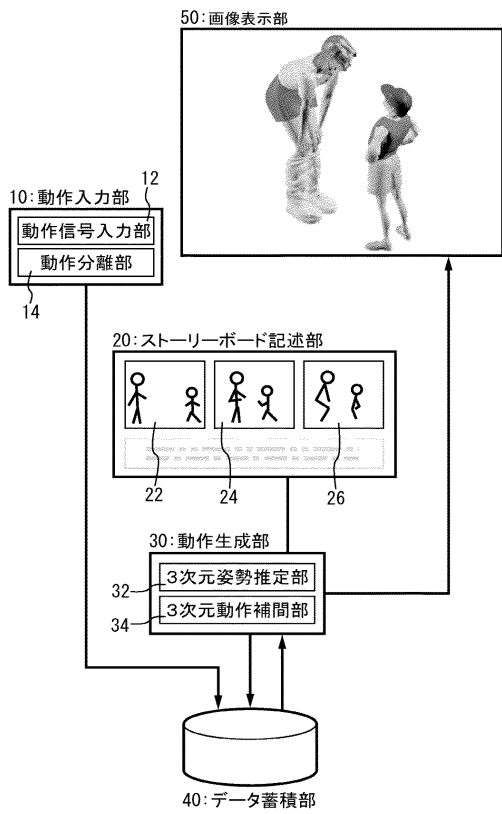
【符号の説明】

10	動作入力部
12	動作信号入力部
14	動作分離部
20	ストーリーボード記述部
22, 24, 26	ストーリーボード
30	動作生成部
32	3次元姿勢推定部
34	3次元動作補間部
40	データ蓄積部
50	画像表示部
100	仮想人物モデル
150	画像面
212, 214, 216	ストーリーボード
220	タイムライン記述欄
222, 224, 226	ストーリーボード
240	メッセージ欄

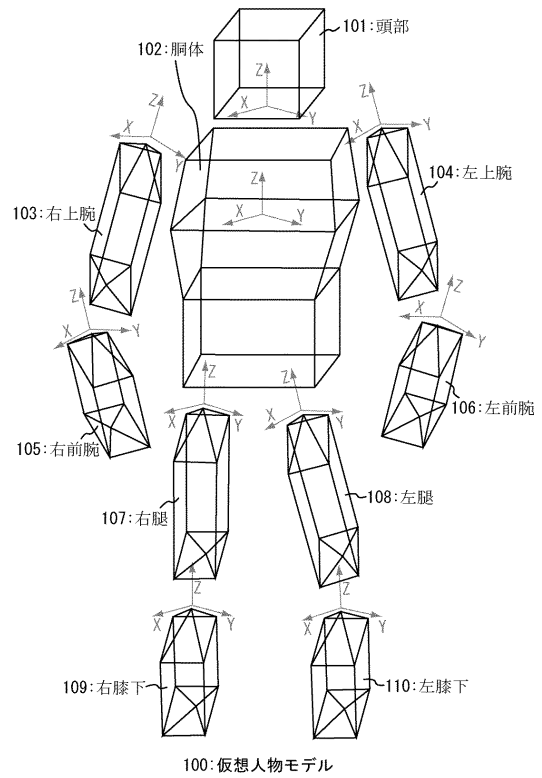
40

50

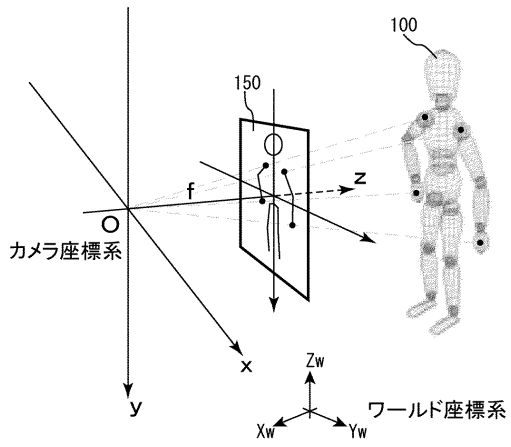
【 図 1 】



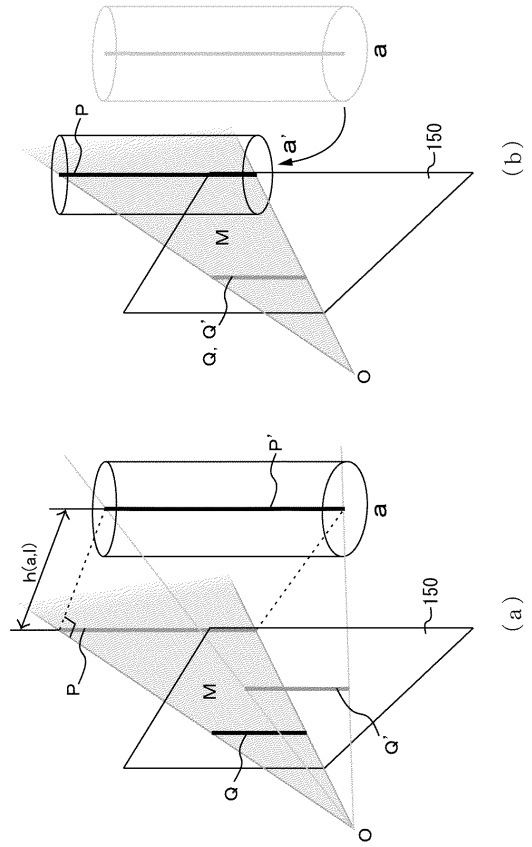
【 図 2 】



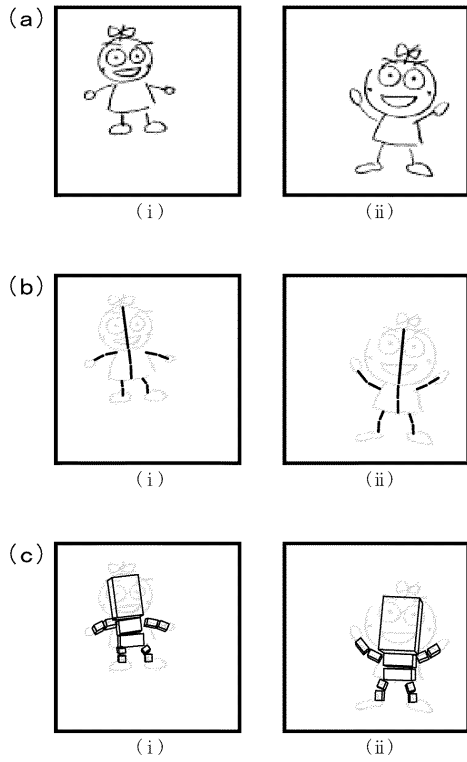
【 図 3 】



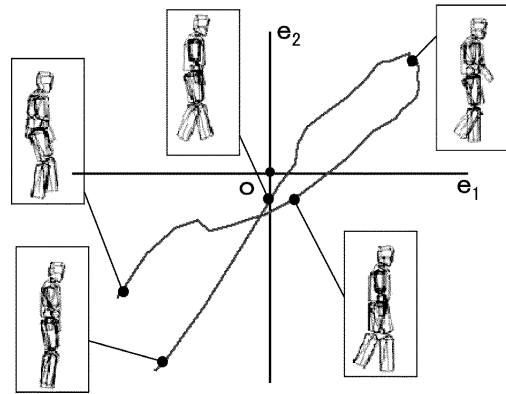
【 図 4 】



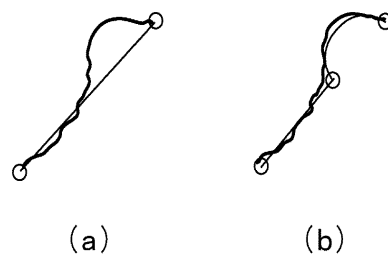
【 図 5 】



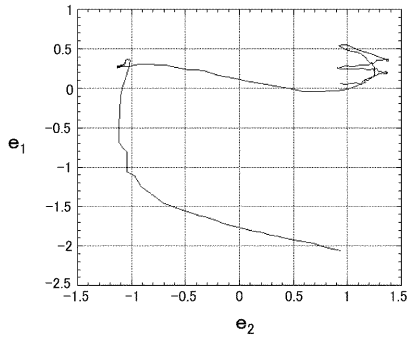
【 図 6 】



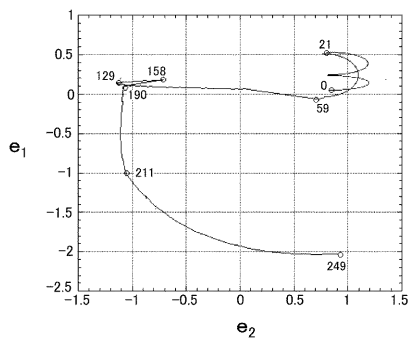
【 図 7 】



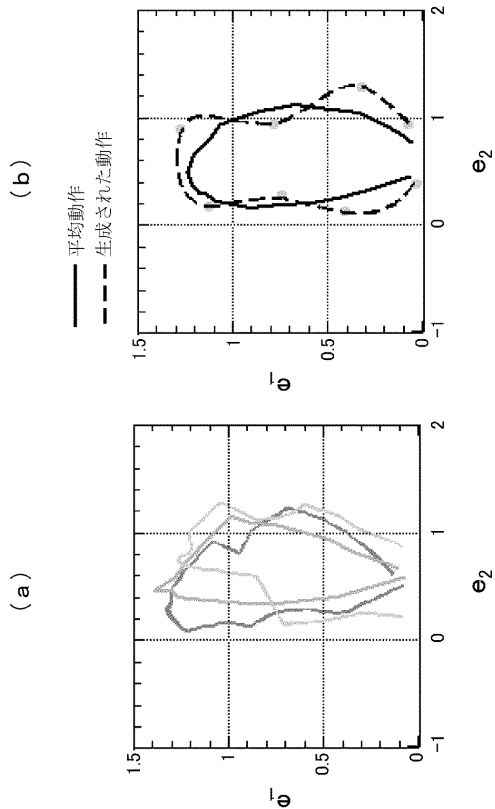
【 図 8 】



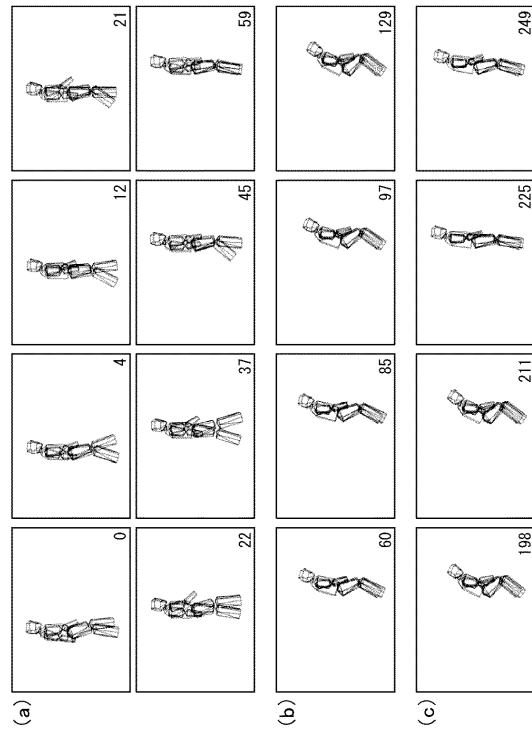
【 図 9 】



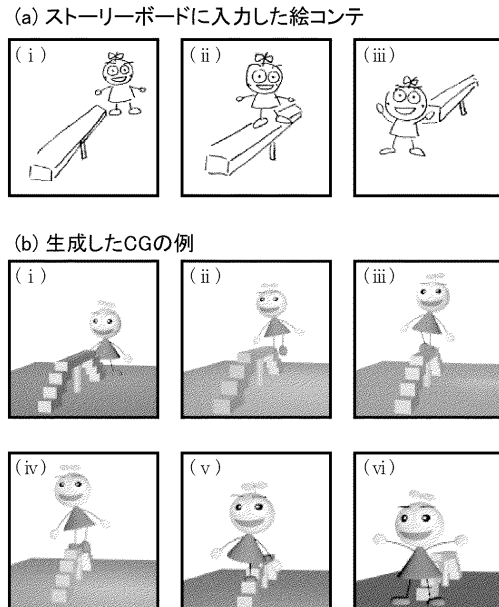
【 図 1 1 】



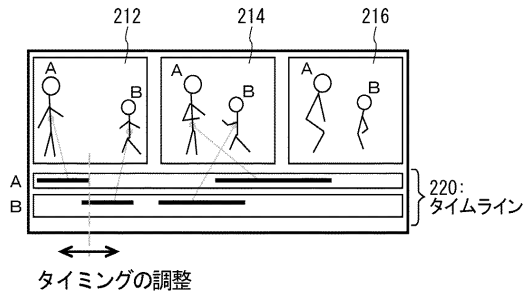
【 図 1 0 】



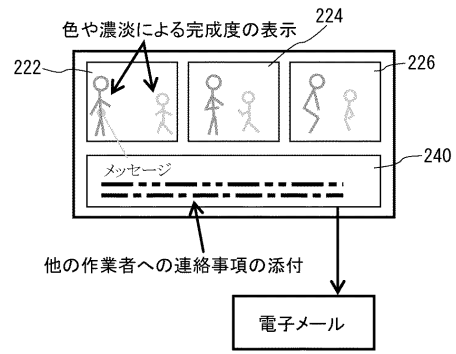
【 図 1 2 】



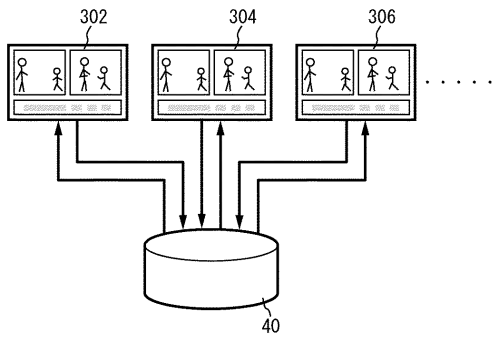
【 図 1 3 】



【 図 1 5 】



【 図 1 4 】



フロントページの続き

(56)参考文献 特開2001-155170(JP,A)

特開2000-251078(JP,A)

特開平11-345342(JP,A)

畑 幸一 Koichi HATA, 複雑な輪郭の自動抽出・追跡法 Automatic Extraction and Tracking of Complex Contours, 電子情報通信学会論文誌 (J81-D-II) 第4号 THE TRANSACTIONS OF THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS D-II, 日本, 社団法人電子情報通信学会 THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS, 1998年 4月25日, 第J81-D-II巻 第4号, 706~715頁

横井 茂樹 Shigeki YOKOI, コンピュータグラフィックス〔III・完〕 Computer Graphics [III Finish], 電子情報通信学会誌 第76巻 第6号 THE JOURNAL OF THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS, 日本, 社団法人電子情報通信学会 THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS, 1993年 6月25日, 第76巻 第6号, 627頁

(58)調査した分野(Int.Cl., DB名)

G06T 13/40