

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4246516号
(P4246516)

(45) 発行日 平成21年4月2日(2009.4.2)

(24) 登録日 平成21年1月16日(2009.1.16)

(51) Int. Cl. F 1
G06T 17/40 (2006.01) G06T 17/40 A
G06T 1/00 (2006.01) G06T 1/00 200A

請求項の数 8 (全 18 頁)

<p>(21) 出願番号 特願2003-37381 (P2003-37381) (22) 出願日 平成15年2月14日 (2003.2.14) (65) 公開番号 特開2004-246729 (P2004-246729A) (43) 公開日 平成16年9月2日 (2004.9.2) 審査請求日 平成18年2月13日 (2006.2.13)</p>	<p>(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4丁目1番8号 (74) 代理人 100105371 弁理士 加古 進 (72) 発明者 星野 准一 茨城県つくば市春日1丁目6-2-501 審査官 相澤 祐介</p>
--	---

最終頁に続く

(54) 【発明の名称】 人物動画画像生成システム

(57) 【特許請求の範囲】

【請求項1】

人物動画画像生成システムであって、
 衣服画像を3次元姿勢の姿勢パラメータとともに格納した仮想衣服データベースを有し

、
ビデオ映像中の人物から、背景部分、肌色部分及び前記ビデオ映像中の人物の3次元姿勢を推定して人体頭部部分を削除することにより、衣服画像を抽出する衣服画像抽出手段と、

前記衣服画像抽出手段により抽出した衣服画像を、前記推定した3次元姿勢の姿勢パラメータとともに、前記仮想衣服データベースに格納する仮想衣服データベース格納手段と

10

、
合成対象のビデオ映像中の人物の3次元姿勢の姿勢パラメータを推定する3次元姿勢推定手段と、

前記3次元姿勢推定手段により推定した3次元姿勢の姿勢パラメータに対応する衣服画像を前記仮想衣服データベースから選択する衣服画像選択手段と、

前記合成対象のビデオ映像中の人物に前記衣服画像選択手段により選択した衣服画像を合成する衣服画像合成手段と

を備えることを特徴とする人物動画画像生成システム。

【請求項2】

請求項1に記載の人物動画画像生成システムにおいて、さらに、

20

前記衣服画像選択手段は、前記3次元姿勢推定手段により推定した3次元姿勢の姿勢パラメータと一致する姿勢パラメータをもつ衣服画像を選択し、一致する衣服画像がない場合には類似する姿勢パラメータをもつ衣服画像を選択し、

前記衣服画像合成手段は、前記衣服画像選択手段で類似する姿勢パラメータをもつ衣服画像を選択した場合に、選択した衣服画像を合成対象のビデオ映像中の人物の姿勢に合うように調整すること

を特徴とする人物動画像生成システム。

【請求項3】

請求項2に記載の人物動画像生成システムにおいて、

前記衣服画像合成手段における前記調整は、衣服画像から衣服の境界を検出し、関節部分を中心に身体部位の回転軸を回転させることにより合成対象のビデオ映像中の人物の姿勢に合わせること

を特徴とする人物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人物動画像生成システムにおいて、さらに、

ヘアスタイル画像を3次元姿勢の姿勢パラメータとともに格納したヘアスタイルデータベースと、

合成対象のビデオ映像中の人物の頭部の3次元姿勢の姿勢パラメータを推定する頭部3次元姿勢推定手段と、

前記頭部3次元姿勢推定手段により推定した頭部の3次元姿勢の姿勢パラメータに対応するヘアスタイル画像を前記ヘアスタイルデータベースから選択するヘアスタイル画像選択手段と、

前記合成対象のビデオ映像中の人物に前記ヘアスタイル画像選択手段により選択したヘアスタイル画像を合成するヘアスタイル画像合成手段と

を備えることを特徴とする人物動画像生成システム。

【請求項5】

請求項4に記載の人物動画像生成システムにおいて、さらに、

ビデオ映像中の人物からヘアスタイル画像を抽出するヘアスタイル画像抽出手段と、

前記ヘアスタイル画像抽出手段により抽出したヘアスタイル画像を前記ヘアスタイルデータベースに格納するヘアスタイルデータベース格納手段とを備え、

前記ヘアスタイル画像抽出手段は、前記頭部3次元姿勢推定手段を用いて前記ビデオ映像中の人物の頭部の3次元姿勢の姿勢パラメータを推定し、顔領域を削除することによりヘアスタイル画像を抽出し、

前記ヘアスタイルデータベース格納手段は、ヘアスタイル画像とともに、前記推定した頭部の3次元姿勢の姿勢パラメータを格納し、

前記ヘアスタイル画像選択手段は、前記頭部3次元姿勢推定手段により推定した3次元姿勢の姿勢パラメータと一致する姿勢パラメータをもつヘアスタイル画像を選択し、一致するヘアスタイル画像がない場合には類似する姿勢パラメータをもつヘアスタイル画像を選択すること

を特徴とする人物動画像生成システム。

【請求項6】

請求項4又は5のいずれかに記載の人物動画像生成システムにおいて、

前記頭部3次元姿勢推定手段は、ビデオ映像中の人物の画像と、頭部の3次元モデルを2次元射影変換したテンプレート画像とをマッチングして顔領域を検出することにより、ビデオ映像中の人物の頭部の3次元姿勢の姿勢パラメータを推定すること

を特徴とする人物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人物動画像生成システムにおいて、

前記衣服画像合成手段および前記ヘアスタイル画像合成手段は、

アルファブレンドにより衣服画像およびヘアスタイル画像の輪郭部分をぼかして自然な

10

20

30

40

50

輪郭を得た後に合成を行なうこと

を特徴とする人物動画像生成システム。

【請求項 8】

請求項 1 ~ 7 のいずれかに記載された人物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プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ビデオ映像から生成した人物の 3 次元モデルを用いた、人物動画像生成システムに関する。

【0002】

【技術的背景】

近年、実写映像と CG ベースの仮想衣服を合成した映像を与える、仮想ファッションのシステムが提案されている（例えば、非特許文献 1 ~ 3）。

仮想衣服の生成方法として、さまざまな手法が提案されている。従来手法の例を図 1 に示す。図 1 において、(b) に示すような衣服の CG データを作成し、(a) に示す人物の実写映像の 3 次元動作（3 次元姿勢）を推定（例えば、特許文献 1 の手法による）して、その動作にあわせて、クロスシミュレーションを行なっている。

従来手法により合成された結果の画像は、例えば図 1 (c) のようになる。

しかし、これらの手法では、利用者の動作に連動した仮想衣服の生成に重点がおかれているため、合成結果がフォトリアル、ビデオリアルに見えるかどうかについてはあまり検討されていなかった。また、CG ベースの仮想衣服を生成するために多くの時間や手間を必要とする。そのため、複雑なしわの生成や、複雑な衣服の合成をすることができないという問題や、合成結果が不自然に見えるという問題があった。

また、近年のインターネットの普及と、パーソナルコンピュータの性能の向上により、利用者の体型に合わせた 3 次元モデルに実写の仮想衣服を着用させ、インターネット経由で商品イメージを提示するサービスなども見られるようになってきた（例えば、非特許文献 4 ~ 5）。しかし、一定の姿勢の 3 次元モデルに仮想衣服を合成するために、実際に利用者が着用した時のイメージがつかみにくいという問題がある。

【0003】

【特許文献 1】

特開 2002 - 269580 号公報

【非特許文献 1】

中野敦，星野准一："利用者の動作と連動する仮想ファッション"，インタラクシオン 2002，pp.202-208

【非特許文献 2】

中野敦，星野准一："対話型仮想ファッションシステム"，第 17 回 NICOGRAPH 論文集，pp.155-160，2001

【非特許文献 3】

星野准一，齊藤啓史："ビデオ映像と CG の合成によるヴァーチャルファッションの実現"，情報処理学会論文誌，Vol.42，No.5，pp.1182-1193，2001

【非特許文献 4】

日本ランズエンド（マイ・バーチャル・モデル）<http://www.landsend.co.jp/>

【非特許文献 5】

デジタルファッション株式会社（HAOREBA）<http://www.dressingsim.com/>

【0004】

【発明が解決しようとする課題】

本発明は、上記の問題を解決するために、従来のような CG ベースの仮想衣服・ヘアスタイルを合成するのではなく、実写ベースの仮想衣服・ヘアスタイルを実写映像中の人物の動作に連動して合成し、リアルな映像を生成する手法を提案するものである。

10

20

30

40

50

【 0 0 0 5 】

【課題を解決するための手段】

上記の課題を解決するために、本発明は、人物動画像生成システムであって、衣服画像を3次元姿勢の姿勢パラメータとともに格納した仮想衣服データベースを有し、ビデオ映像中の人物から、背景部分、肌色部分及び前記ビデオ映像中の人物の3次元姿勢を推定して人体頭部部分を削除することにより、衣服画像を抽出する衣服画像抽出手段と、前記衣服画像抽出手段により抽出した衣服画像を、前記推定した3次元姿勢の姿勢パラメータとともに、前記仮想衣服データベースに格納する仮想衣服データベース格納手段と、合成対象のビデオ映像中の人物の3次元姿勢の姿勢パラメータを推定する3次元姿勢推定手段と、前記3次元姿勢推定手段により推定した3次元姿勢の姿勢パラメータに対応する衣服画像を前記仮想衣服データベースから選択する衣服画像選択手段と、前記合成対象のビデオ映像中の人物に前記衣服画像選択手段により選択した衣服画像を合成する衣服画像合成手段とを備えることを特徴とする人物動画像生成システムである。

10

また、前記衣服画像選択手段は、前記3次元姿勢推定手段により推定した3次元姿勢の姿勢パラメータと一致する姿勢パラメータをもつ衣服画像を選択し、一致する衣服画像がない場合には類似する姿勢パラメータをもつ衣服画像を選択し、前記衣服画像合成手段は、前記衣服画像選択手段で類似する姿勢パラメータをもつ衣服画像を選択した場合に、選択した衣服画像を合成対象のビデオ映像中の人物の姿勢に合うように調整することを特徴としていてもよい。

前記衣服画像合成手段における前記調整は、衣服画像から衣服の境界を検出し、関節部分を中心に身体部位の回転軸を回転させることにより合成対象のビデオ映像中の人物の姿勢に合わせることを特徴としていてもよい。

20

【 0 0 0 6 】

上記の人物動画像生成システムは、さらに、ヘアスタイル画像を3次元姿勢の姿勢パラメータとともに格納したヘアスタイルデータベースと、合成対象のビデオ映像中の人物の頭部の3次元姿勢の姿勢パラメータを推定する頭部3次元姿勢推定手段と、前記頭部3次元姿勢推定手段により推定した頭部の3次元姿勢の姿勢パラメータに対応するヘアスタイル画像を前記ヘアスタイルデータベースから選択するヘアスタイル画像選択手段と、前記合成対象のビデオ映像中の人物に前記ヘアスタイル画像選択手段により選択したヘアスタイル画像を合成するヘアスタイル画像合成手段とを備えることを特徴とする人物動画像生成システムであっててもよい。

30

この人物動画像生成システムは、さらに、ビデオ映像中の人物からヘアスタイル画像を抽出するヘアスタイル画像抽出手段と、前記ヘアスタイル画像抽出手段により抽出したヘアスタイル画像を前記ヘアスタイルデータベースに格納するヘアスタイルデータベース格納手段とを備え、前記ヘアスタイル画像抽出手段は、前記頭部3次元姿勢推定手段を用いて前記ビデオ映像中の人物の頭部の3次元姿勢の姿勢パラメータを推定し、顔領域を削除することによりヘアスタイル画像を抽出し、前記ヘアスタイルデータベース格納手段は、ヘアスタイル画像とともに、前記推定した頭部の3次元姿勢の姿勢パラメータを格納し、前記ヘアスタイル画像選択手段は、前記頭部3次元姿勢推定手段により推定した3次元姿勢の姿勢パラメータと一致する姿勢パラメータをもつヘアスタイル画像を選択し、一致するヘアスタイル画像がない場合には類似する姿勢パラメータをもつヘアスタイル画像を選択することを特徴とする人物動画像生成システムであっててもよい。

40

また、前記頭部3次元姿勢推定手段は、ビデオ映像中の人物の画像と、頭部の3次元モデルを2次元射影変換したテンプレート画像とをマッチングして顔領域を検出することにより、ビデオ映像中の人物の頭部の3次元姿勢の姿勢パラメータを推定することを特徴としていてもよい。

また、前記衣服画像合成手段および前記ヘアスタイル画像合成手段は、アルファブレンドにより衣服画像およびヘアスタイル画像の輪郭部分をぼかして自然な輪郭を得た後に合成を行なうことを特徴としていてもよい。

また、上記の人物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プログラム

50

も、本発明である。

【0007】

【発明の実施の形態】

本発明の人物動画像生成システムでは、上述した従来技術の問題点を解決するために、CG衣服を作成してクロスシミュレーションを行なうのではなく、次のような手法を用いる。すなわち、ビデオ映像から様々な姿勢の衣服画像を切り出してデータベース化し、入力された人物動画像にデータベースの衣服画像を合成して、様々な服装の人物動画像を生成する。

また、本発明の人物動画像生成システムでは、ヘアスタイルの合成についても同様の手法を用いる。すなわち、ビデオ映像から様々な姿勢のヘアスタイル画像を切り出してデータベース化し、入力された人物動画像にデータベースのヘアスタイル画像を合成して、様々なヘアスタイルの人物動画像を生成する。

10

【0008】

以降で説明する本発明の実施形態においては、あらかじめテレビカメラにより撮影した人物の映像を入力映像として、人物の動作（各フレームにおける人物の姿勢）に連動した仮想衣服・ヘアスタイルを合成し、表示する人物動画像生成システムを構築する。これにより、映像中の人物の3次元動作（3次元姿勢）に連動して、仮想衣服・ヘアスタイルが動くことにより、合成結果がビデオリアリスティックに見えることを目的としている。

以降、図を参照しながら、本発明の人物動画像生成システムの実施形態を詳細に説明する。

20

まず、ビデオ映像から様々な姿勢の衣服画像を切り出してデータベース化する処理（仮想衣服データベースの作成）について説明し、次に、入力された人物動画像にデータベースの衣服画像を合成して、様々な服装の人物動画像を生成する（実写映像と衣服画像の合成）について説明する。

【0009】

<仮想衣服データベースの作成>

仮想衣服データベースの作成は、以下に示す処理によって行なう。なお、この処理を図2に示す。各々の処理については後で詳しく説明する。

(1) 入力された人物動画像（仮想衣服生成用映像）(S201)を用いて、人物の3次元姿勢を推定する(S202)。

30

(2) 推定した3次元姿勢を利用して、画像中から人体部分を削除し、衣服部分を切り出す(S203)。

(3) 映像の各フレームで上記(1)(2)を繰り返すことで様々な姿勢の衣服のデータを取得し、仮想衣服データベースを構築する(S204)。

これにより、入力された人物動画像の3次元姿勢に対応する衣服データを取得できるようになる。

【0010】

まず、上記(1)の3次元姿勢の推定について説明する。

実写映像中の人物動作に連動した合成映像を生成するためには、実写映像中の人物の3次元姿勢（3次元動作）を推定する必要がある。本実施形態では、人物の人体幾何モデルに合わせた動力学モデルから得られる関節駆動力を運動情報とし、ビデオ映像上における人体領域のフレーム間輝度値2乗誤差を併用することで、人物の3次元姿勢を推定する。

40

なお、上記の手法については、例えば上述の特許文献1（特開2002-269580号公報）、「中野敦，星野准一："動力学モデルに基づくビデオモーションキャプチャ"，TVRSJ Vol.7, No.4.2002」等を参照されたい。

【0011】

次に、上記(2)の衣服部分を切り出す処理について説明する。

本実施形態では、求められた3次元人体モデルと姿勢パラメータを利用して、実写画像中から衣服部分を切り出し、仮想衣服データベースを作成する。

効率よく衣服画像部分を抽出するために、本実施形態では、例えば背景にブルースクリー

50

ンを用い、抽出する衣服以外は青色の布で隠し、背景差分処理を行なう。また、閾値処理には比較的輝度変化にロバストなH V S表色系のH（色相）とS（彩度）を利用する。しかし、背景差分処理だけでは肌色領域や、頭部の領域を削除することができない。本実施形態では、頭部は3次元人体モデルの情報を利用し、首の部分より上部の人体領域を頭部と判断し、削除する。また、肌色領域は色相がほぼ一定であるため、H V S基底のH（色相）を利用することによって削除する。R G B基底からH V S基底への変換は以下の式を用いる。

【0012】

【数1】

$$H = \cos^{-1} \left\{ \frac{(G-B) + (G-R)}{2\sqrt{(G-B)^2 + (G-R)(B-R)}} \right\}$$

10

【数2】

$$V = \text{MAX}(R, G, B)$$

【数3】

$$S = \sqrt{\left(\frac{(R-G) + (R-B)}{2} \right)^2 + \left(\frac{\sqrt{3}(G-B)}{2} \right)^2}$$

20

【0013】

図4に、衣服画像の抽出結果を示す。図4において、(a)は入力画像であり、(b)は(a)から上記の手法を用いて抽出した衣服画像である。

なお、上記の肌色領域の削除処理は従来技術を利用したものである。詳しくは、例えば「松橋 聡，藤本研司，中村 納，南 敏，“顔領域抽出に有効な修正HSV表色系の提案”，テレビジョン学会誌，vol.49,no.6,pp.787-797,1995」などを参照されたい。

30

【0014】

最後に、上記(3)の仮想衣服データベースの構築について説明する。

上記の(1)(2)の処理を実写映像のフレームごとに行ない、画像中から衣服画像を抽出する。この衣服画像と、3次元人体モデルの姿勢パラメータとを対応付けて、仮想衣服データベースを構築する。

【0015】

<実写映像と衣服画像の合成>

実写映像と衣服画像の合成は、以下に示す処理によって行なう。なお、この処理を図3に示す。各々の処理については後で詳しく説明する。

(1)入力された人物動画像(合成対象の人物の映像)(S301)を用いて、人物の3次元姿勢を推定する(S302)。

40

(2)推定した3次元姿勢を利用して、仮想衣服データベースから一致する姿勢、または類似する姿勢の衣服データを参照する(S303)。

(3)入力画像の人物の姿勢に一致する姿勢の衣服データがない場合、類似する姿勢の衣服データを、入力映像の姿勢に合うように調整する(S304)。

(4)調整した衣服画像と実写映像中の人物を合成し(S305)、生成した画像を出力する(S306)。

(5)入力映像の各フレームで上記(1)~(4)を繰り返し行なうことで、実写映像と衣服画像の合成を行なう。

【0016】

50

まず、上記(1)の3次元位置・姿勢の推定については、上述で説明した<仮想衣服データベースの作成>と同様の手法で推定を行なう。本実施形態ではこの手法により推定した3次元姿勢に合わせて、実写ベースの仮想衣服を合成する。

【0017】

次に、上記(2)の衣服データの参照について説明する。

上述の<仮想衣服データベースの作成>で説明したように、画像中から抽出した衣服画像は、3次元人体モデルの姿勢パラメータと対応付けて、仮想衣服データベースに格納されている。このため、上記(1)で推定した3次元姿勢を利用して、仮想衣服データベースから一致する姿勢、または類似する姿勢の衣服データを参照することができる。

【0018】

次に、上記(3)の衣服データの調整、および(4)の衣服画像と実写映像中の人物との合成について説明する。

本実施形態では、入力画像中の人物の姿勢と同じ姿勢の衣服データがない場合、類似した姿勢の衣服データを調整し、姿勢を一致させる必要がある。本実施形態では、次の手順で類似した姿勢の衣服データから袖部分を切り出し、2次元的に変形させることで、入力画像中の人物の姿勢と一致した衣服データを生成する。これは、次のような処理により行なわれる。

1. 衣服の袖を切り出すために、袖の境界を求める。
2. 入力画像中の人物の姿勢と一致するように、切り出した袖の部分を、腕の回転軸を中心に回転させる。
3. 回転によって生じるオクルージョン領域の補間を行なう。
4. アルファマップを用いて境界部分をぼかし、合成する。

なお、本実施形態では袖(腕)の部分の調整を例として説明しているが、他の部分についても同様の手法で衣服データを調整することができる。この場合、腕以外の身体部位(例えば脚など)に対応する衣服の境界を検出し、上記の腕の回転と同様に、関節部分を中心に身体部位の関節回転軸を回転させることにより合成対象の映像中の人物の姿勢に合わせることができる。

【0019】

(1. 袖の境界の検出)

画像中から袖の領域を検出するためには、衣服のモデルを作成し対応する領域を求めることが考えられる。しかし、衣服の種類ごとにモデルを作成しなければならず、CGの仮想衣服と同様に、複雑な形状のモデルを作ることが困難である。

代表的な領域分割手法には、クラスタリングによる手法、ヒストグラム解析を用いた閾値領域による手法、領域成長法などがある。しかし、本実施形態で扱う衣服画像の場合、境界が明確にわからない場合が多いため、推定した3次元姿勢(動作)の結果から得られたおおよその両腕の領域を考慮に入れ、袖の境界を検出する。

袖の境界は、多くの場合明度が低くなる傾向がある。得られた両腕の領域付近で、境界線の滑らかさを考慮に入れ、明度が低い(暗い)画素を探索する。探索には、従来技術であるSnakeのアルゴリズムを取り入れ、下記の式によりエネルギーを計算し、最小になる画素を境界とする。

【数4】

$$ene = aene + bene + V + dif \times weight \rightarrow min$$

【0020】

ただし、3次元モデルから得られた対象領域(腕領域)をSnakeの初期輪郭とする。また、Vは対象画素の明度を表し、aene、beneはそれぞれ連続する点の座標の1次グラディエントの2乗、2次グラディエントの2乗を表す。difは3次元モデルから得られた対象領域の輪郭からの距離とする。

上記の手法により袖の境界を検出した様子を図5に示す。図5において、(a)は入力画像、(b)は検出した袖の境界を示す。510で示された線は3次元モデルから得られた

10

20

30

40

50

境界、520で示された線は検出された境界 (Actual Boundary) である。

【0021】

(2. 袖の回転)

入力画像中の人物の姿勢と一致するように、切り出した袖の部分を腕の回転軸を中心に回転させ、姿勢を変化させる。回転軸の画像中の座標は、3次元モデルを使った3次元姿勢(動作)の推定により求められる座標を利用した。

本実施形態では類似した姿勢からのみ姿勢の変形を行なうことを前提とし、変換に関し奥行き情報は無視できるくらい小さいと仮定する。回転後の各画素の x, y 座標 x', y' を、逆アフィン変換を用いて以下の式により計算する

【数5】

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} (x-x_0) \\ (y-y_0) \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

ただし、 x_0, y_0 は回転の中心の x, y 座標、 x, y は変換前の x, y 座標、 θ は半時計回りの回転角とする。

【0022】

(3. オクルージョン領域の補間)

袖の部分を切り出し、姿勢を変更すると、袖に隠れていた部分などのオクルージョン領域(隠蔽領域)の補間が必要となる場合がある。本実施形態では、衣服の3次元モデルがないため正確な補間領域を知ることはできない。また、画像中の隠蔽領域の補間方法としては様々な手法が提案されているが、あらかじめ画像の特徴をデータベース化するなど、前処理をする必要があるため、動画に適用する場合には多くの処理時間を必要とする。

本実施形態ではできるだけ簡便で汎用的な手法が望ましいため、例えば、3次元人体モデルの輪郭と、隠蔽領域の近傍の画素を利用して補間を行なうものとする。画像中を水平方向に走査し、補間すべき画素を調べ、補間領域の両側に衣服領域が存在する場合は、以下の式で線形補間を行なう。

【数6】

$$[x, y] = (1-p) \times [x_l, y] + p \times [x_r, y]$$

ただし、 $[x, y]$ は補間する対象画素のRGB値、 $[x_l, y]$ は補間領域の左端画素のRGB値、 $[x_r, y]$ は補間領域の右端画素のRGB値、 p は補間領域の水平方向の距離を1で正規化した時の、補間領域の左端からの距離とする。

また、片側だけに衣服領域が存在する場合は、補間領域の近傍の画素を、3次元人体モデルの輪郭に合わせるようにサイズ変換することで補間する。

【0023】

(4. アルファマップの作成)

単純に袖部分を回転した画像を合成する場合、輪郭部分の色の变化に連続性がないため不自然に見える。そこで本実施形態ではアルファブレンドの手法を利用して、境界線を自然に見えるようにする。この手法では、袖の境界 (Actual Boundary) を境にアルファ値(透明度)を変化させたアルファマップを作成し、衣服から袖の部分を切り出して合成する。

アルファマップを図6に示す。図6において、線610はActual Boundaryを、620は衣服領域を、630は衣服領域外を示している。衣服領域620はアルファ値を255(不透明)にし、衣服領域外630では、Actual Boundary 610から遠ざかるほど、アルファ値を0(透明)に近づける。

なお、アルファブレンドには以下の式を用いた。

【数7】

10

20

30

40

50

$$[x', y'] = [x_1, y_1] \times \alpha + [x_0, y_0] \times (1 - \alpha)$$

ただし、 $[x', y']$ はアルファブレンド後の対象画素の RGB 値、 $[x_1, y_1]$ は重ねる画素の RGB 値、 $[x_0, y_0]$ は元の画素の RGB 値、 α は対象画素のアルファ値とする。

上述した (1. 袖の境界の検出) ~ (4. アルファマップの作成) の手法により姿勢を調整した衣服画像を、図 7 に示す。図 7 において、(a) は入力画像を、(b) は入力画像に類似する姿勢の衣服画像、(c) は (a) の入力画像の人物に合わせて姿勢を調整した衣服画像である。

10

【0024】

(5. 3次元での仮想衣服の生成)

上述した 2次元での仮想衣服の生成では、衣服の 3次元モデルがないため、オクルージョン領域の正確な補間ができなかった。また、z軸方向への姿勢の変化を考慮に入れた、仮想衣服の姿勢の微調整を行うことができなかった。以降、この問題を解決するために、簡単な衣服の 3次元モデルを作り仮想衣服の姿勢を微調整する手法を説明する。

正確な衣服のモデルを作成するのは、多くの時間や手間がかかる。また、姿勢(動作)に連動した正確な衣服シミュレーションを行うことも非現実的である。本実施形態では、衣服領域を知ることを目的として、図 8 (b) のような簡単な衣服の 3次元モデルを手動で作成した。

20

本実施形態では以下の処理を行うことで、入力映像中の人物の姿勢に合った仮想衣服を生成する。

【0025】

まず、次式で類似した仮想衣服画像の各点の 3次元座標と入力画像の各点の 3次元座標との対応をとる。

【数 8】

$$P'_{3D} = M' M^{-1} (P_{3D} - R) + R'$$

ただし、

30

P_{3D} : 入力映像中の人物の、身体部分の腰からの 3次元ローカル座標

M : 入力映像中の人物姿勢の腰からの変換行列

R : 入力映像中の人物の、腰のグローバル座標

P'_{3D} : 類似姿勢仮想衣服の、身体部分の腰からの 3次元ローカル座標

M' : 類似姿勢仮想衣服の腰からの変換行列

R' : 類似姿勢仮想衣服の腰からのグローバル座標

とする。

【0026】

次に、次式で各点の 3次元座標を画面上の 2次元座標に変換し、各画素の対応をとる。

【数 9】

40

$$P'_{2D} = I' P'_{3D}$$

【数 10】

$$P_{2D} = I P_{3D}$$

ただし、

P_{2D} : P_{3D} を 2次元の画像平面に射影した座標

I : 3次元空間の座標を画像平面上の座標に変換するための行列

50

$P'_{2D} : P'_{3D}$ を 2次元の画像平面に射影した座標
 I' : 3次元空間の座標を画像平面上の座標に変換するための行列
 とする。

【0027】

これにより、類似姿勢の仮想衣服の各画素と、入力映像中の各画素との対応がとれ、衣服の3次元モデル上の領域に仮想衣服の画素をコピーすることで、入力映像中の人物の姿勢にあわせることができる。

図8に、上記の手法による、3次元モデルを用いた仮想衣服の調整を示す。図8において、(a)は入力画像を、(b)は衣服の3次元モデルを、(c)は上記の手法により入力画像の人物と衣服モデルを重ね合わせた画像である。

10

【0028】

次に、本実施形態におけるヘアスタイルの合成について説明する。

上述の従来技術による、3次元人体モデルを利用した動作推定の場合、正確な頭部の向きを得ることができない。そのため、ヘアスタイルを実写映像に正確に合成することが難しい。そこで、本実施形態では顔面の3次元モデルを用いて頭部の位置・姿勢を推定し、ヘアスタイルや帽子を合成する手法を提案する。

本実施形態では、例えばミノルタ(株)のレーザーレンジファインダ(非接触3次元デジタルVIVID 300)により距離計測を行い、人物の正面を向いた顔の3次元モデルを作成し、同時に顔のテクスチャ画像を得る。レーザーレンジファインダを用いて、顔の3次元距離を測定し、シェーディングを行った。また、3次元モデルの傾きの正規化を行うことで、XYZ軸回転の角度を複数3次元モデル間で統一して表現する。レーザーレンジファインダにより計測した顔の3次元モデルを、鼻の下の位置を用いてXYZ軸周りの傾きを図9に示すように正規化した。

20

図9は顔の3次元モデルを示している。図9において、(a)は本実施形態で扱う座標系を示す。また、(b)~(d)は、それぞれX軸、Y軸、Z軸周りの傾きを正規化した顔の3次元モデルである。

以降、この3次元モデルを2次元射影変換して生成した画像を「テンプレート画像」と呼ぶ。

【0029】

上述したように、本実施形態におけるヘアスタイルの合成は、まず、ビデオ映像から様々な姿勢のヘアスタイル画像を切り出してデータベース化し(ヘアスタイルデータベースの作成)、入力された人物動画像にデータベースのヘアスタイル画像を合成して、様々なヘアスタイルの人物動画像を生成する(実写映像とヘアスタイル画像の合成)。以降、これらの処理を詳細に説明する。

30

【0030】

<ヘアスタイルデータベースの作成>

ヘアスタイルデータベースの作成は、以下に示す処理によって行なう。なお、この処理を図10に示す。各々の処理については後で詳しく説明する。

(1) 入力された人物動画像(仮想ヘアスタイル作成用映像)(S1001)を用いて、頭部の3次元位置・姿勢を推定する(S1002)。

40

(2) 推定した3次元姿勢を利用して、画像中から髪の部分を切り出す(S1003)。

(3) 映像の各フレームで上記(1)(2)を繰り返すことで様々な姿勢のヘアスタイルのデータを取得し、ヘアスタイルデータベースを構築する(S1004)。

これにより、入力された人物動画像の頭部の3次元姿勢に対応するヘアスタイルデータを得ることができるようになる。

【0031】

まず、上記(1)の3次元姿勢の推定について説明する。

頭部の3次元位置・姿勢の推定方法として、従来、目や口などの顔の幾何学的特徴点を利用したものや、対象物体のモデルと入力画像との照合による推定方法などが提案されている。しかし、幾何学的特徴を抽出する際には画像に含まれるノイズ等の影響が大きく、特

50

徴の対応づけを行う際に誤りを含んでしまう可能性がある。また、モデルを使った照合では、光源の位置や種類、物体の表面反射係数の影響が複雑となるために、モデルからの対象物体の濃淡画像を忠実に生成することが困難となるなどの問題がある。

そこで本実施形態では、正規化相関を用いたパターンマッチングを行い、入力画像とテンプレート画像との相関度を評価し、顔の位置・姿勢を推定する。正規化相関は、対応する画素値の平均と標準偏差によって正規化された絶対測度であり、照明変動などの外乱による影響を受けにくい評価尺度として知られている。

また、本実施形態では、テンプレート画像（3次元モデルから逐次生成される濃淡画像）を t_n ($n = 0, \dots, N - 1$)、探索対象画像（入力映像から得られる濃淡画像）を F_n ($n = 0, \dots, N - 1$)、探索対象画像における参照位置の、テンプレート画像と同サイズとなる部分画像を f_n ($n = 0, \dots, N - 1$) と表す。

【0032】

このとき、テンプレート画像 t_n ($n = 0, \dots, N - 1$) と部分画像 f_n ($n = 0, \dots, N - 1$) との正規化相関は次式で与えられる。なお、 R は相関値である。

【数11】

$$R = \frac{\sum_{n=0}^{N-1} (f_n - \bar{f})(t_n - \bar{t})}{\sqrt{\sum_{n=0}^{N-1} (f_n - \bar{f})^2} \sqrt{\sum_{n=0}^{N-1} (t_n - \bar{t})^2}}$$

10

20

ただし、

【数12】

$$\bar{f} = \frac{1}{N} \sum_{n=0}^{N-1} f_n$$

【数13】

$$\bar{t} = \frac{1}{N} \sum_{n=0}^{N-1} t_n$$

30

【0033】

複雑な条件下でマッチングを行う場合、顔パターンに類似した非顔領域の誤検出や、探索するパラメータの数が多くなるなどの問題が生じる可能性がある。本実施形態では処理を高速・安定化するため、撮影時の条件として、

- ・処理の始めは正面を向いている。
- ・シーン中には、背景を除き人物以外の物体は存在しない。
- ・顔の回転はx軸回転のみ、またはy軸回転のみとする。
- ・z軸方向（奥行き）の変化はほぼない。

40

という制約を加える。ここで用いた前提条件は常に満たされると限らないが、利用される状況の大部分はカバーしており実用上さほど問題にはならない。

【0034】

以下、本実施形態で行う頭部の位置・姿勢の推定手法を示す。図12は、本実施形態で行う頭部の位置・姿勢の推定処理の流れを示したフローチャートである。

(a) まず、初期化処理として、入力画像の顔領域の大きさとテンプレート画像の縮尺を合わせる(図12のS1202～S1207の処理)。

1201は、入力する顔の3次元モデルである。正面を向いたフレーム画像を入力し(S

50

1202)、これに対して、3次元モデルを2次元射影変換して生成したテンプレート画像を数段階にスケール変化させてマッチングを行う。

マッチングは、まず3次元モデル1201のz軸の奥行きを変化させ(S1203)、2次元射影変換を行なってテンプレート画像を生成し(S1204)、マッチングを行なう(S1205)。次に、上述した式により最大のR(相関値)を計算する。Rが最大値となった場合(S1207でYesの場合)に、上記S1203のz軸の奥行きを、最も相関が高いスケールであると判断して、このスケール(1208)を以後の処理に採用する。Rが最大でない場合(S1207でNoの場合)は、再びz軸の奥行きを変化させる処理(S1203)から繰り返す。

【0035】

(b)次に、3次元モデルの姿勢を逐次変化させて生成したテンプレート画像と入力画像とを比較して、入力画像(フレーム画像)の頭部の姿勢を求める(図12のS1209~S1214の処理)。

まず、3次元モデルのx軸、y軸回りの回転角度を変化させ(S1209)、2次元射影変換を行なってテンプレート画像を生成する(S1210)。生成したテンプレート画像を入力画像全体に走査してマッチングを行ない(S1211)、各場所での類似度を逐次求め、局所最大かつ、閾値以上の箇所を検出し相関値Rを求める(S1212)。これらの処理をRが最大値となるまで繰り返し(S1213)、様々な姿勢から得られるテンプレート画像の中で最も相関が高い姿勢を入力画像の頭部の姿勢(1214)とする。

(c)入力映像の各フレーム画像で(b)を繰り返し行う(S1215)。

なお、本実施形態では、入力画像とテンプレート画像を50%に縮小し処理を行った。また、テンプレート画像を作成する角度は前フレームで推定した顔の角度の、-20~+20度の範囲で刻み幅1度で行った。これにより、処理を大幅に高速化できるとともに、画像上の雑音成分を除去することができる。

【0036】

(推定姿勢の補正)

動画画像の場合には瞬きなどの動作があり、上述の3次元位置・姿勢の推定手法では、部分的に誤検出をする場合がある。その結果、頭部の推定した回転角が連続的に変化しないという問題が発生する。

そのような問題を解決するために、従来、更新型のテンプレートマッチングを併用した手法も提案されている。しかし、本実施形態でこの従来手法を用いると、あらかじめ計測した顔の3次元モデルからテンプレートを作成するため、いくつもの3次元モデルを用意し逐次レンダリングする必要があり非効率的である。

そこで本実施形態では、上述の3次元位置・姿勢の推定手法で推定した頭部の回転角度をSnakesのアルゴリズムと移動平均フィルタを併用することで、例外点を取り除き、推定した回転角度を連続的に変化させる。

図13に正面を向いている状態から、右を向き、さらに左を向いて、正面を向いた入力映像に対して、頭部の姿勢の推定を行った結果を破線で示す。また、補正後の回転角度の様子を実線で示す。本手法によって、より自然な頭部の姿勢の推定が可能となっていることがわかる。

【0037】

次に、上記(2)の髪の毛部分を切り出す処理について説明する。

本実施形態では、画像中から髪領域のみを抽出するために、ヘアスタイル抽出用のマスクを生成する。まず、髪領域の雑音成分を削減するために平滑化を行う。次に、背景差分法により人物領域と背景領域を分離し、頭部の位置・姿勢の推定で照合したテンプレート領域を削除する。最後に、髪領域の明度が肌色領域に比較して低いことを利用し、明度を用いて人物領域を肌領域と髪領域に分ける。本実施形態の手法により髪の毛部分を切り出した様子を、図14に示す。

図14において、(a)は入力画像である。(a)から平滑化を行い、背景領域とテンプレート画像が重なっている領域を削除した結果を(b)に、明度差による髪領域の抽出を

10

20

30

40

50

行った結果を(c)に示す。(d)は最終的に得られたヘアスタイル画像である。合成時には、上述したアルファマップの手法を用いて境界部分をぼかす。

【0038】

最後に、上記(3)のヘアスタイルデータベースの構築について説明する。

上記の(1)(2)の処理を実写映像のフレームごとに行ない、画像中からヘアスタイル画像を抽出する。このヘアスタイル画像と、3次元人体モデルの姿勢パラメータとを対応付けて、ヘアスタイルデータベースを構築する。

【0039】

<実写映像とヘアスタイル画像の合成>

実写映像とヘアスタイル画像の合成は、以下に示す処理によって行なう。なお、この処理を図11に示す。各々の処理については後で詳しく説明する。

(1)入力された人物動画像(合成対象の人物の入力映像)(S1101)を用いて、人物の頭部の3次元位置・姿勢を推定する(S1102)。

(2)推定した3次元姿勢を利用して、ヘアスタイルデータベースから一致する姿勢、または類似する姿勢のヘアスタイルデータを参照する(S1103)。

(3)検索したヘアスタイル画像を調整し(S1104)、実写映像中の人物と合成する(S1105)。最後に、生成された画像を出力する(S1106)。

(4)入力映像の各フレームで上記(1)~(3)を繰り返し行なうことで、実写映像とヘアスタイル画像の合成を行なう。

【0040】

まず、上記(1)の3次元位置・姿勢の推定については、上述で説明した<ヘアスタイルデータベースの作成>と同様の手法で推定を行なう。

【0041】

次に、上記(2)のヘアスタイルデータの参照について説明する。

上述の<ヘアスタイルデータベースの作成>で説明したように、画像中から抽出したヘアスタイル画像は、3次元人体モデルの姿勢パラメータと対応付けて、ヘアスタイルデータベースに格納されている。このため、上記(1)で推定した3次元姿勢を利用して、ヘアスタイルデータベースから一致する姿勢、または類似する姿勢のヘアスタイルデータを参照することができる。

【0042】

次に、上記(3)の実写映像とヘアスタイル画像との合成について説明する。

入力画像にヘアスタイルを合成する場合、合成する位置を入力画像中の人物の髪領域部に合わせる必要がある。本実施形態では、入力画像中のおおまかな髪領域をテンプレート画像がマッチングした座標を利用して指定し、初期条件として、入力映像中の頭部のサイズにあうように、手動で仮想ヘアスタイルのサイズ調整を行い、合成を行なっている。

【0043】

<実施例>

ここでは、ビデオ映像から人物の3次元姿勢(動作)の推定を行い、仮想衣服・ヘアスタイルの合成を行った例を説明する。

図15は、歩いている人物に仮想衣服を合成した結果を示す。図15において、(a)は入力画像、(b)は3次元姿勢の推定結果である。(c)は人物の姿勢に合わせて仮想衣服画像を生成したところである。(d)は(c)の仮想衣服画像を(a)の入力画像に合成した結果である。

同様に、人物の頭部の姿勢に合わせてヘアスタイルを合成した結果を図16に示す。図16において、(a)は入力画像、(b)は3次元姿勢の推定結果である。(c)は人物の姿勢に合わせてヘアスタイル画像を生成したところである。(d)は(c)のヘアスタイル画像を(a)の入力画像に合成した結果である。

本実施例によれば、図15、図16に示すように、人物の姿勢(動作)に連動した仮想衣服・ヘアスタイルの合成が可能になっている。

【0044】

【発明の効果】

本発明によれば、従来のCGベースの仮想衣服・ヘアスタイルではなく、人物の姿勢（動作）に連動した実写ベースの仮想衣服・ヘアスタイルを生成・合成する手法により、ビデオリアリスティックな合成結果を得ることができる。

【図面の簡単な説明】

【図1】従来技術を用いた仮想衣服と人物との合成を示した図である。

【図2】仮想衣服データベースの作成処理を示すフローチャートである。

【図3】実写映像と衣服画像の合成処理を示すフローチャートである。

【図4】本実施形態の手法による衣服画像の抽出結果を示す図である。

【図5】本実施形態の手法による袖の境界の検出結果を示す図である。

【図6】本実施形態で用いるアルファマップを示した図である。

【図7】本実施形態の手法により姿勢を調整した衣服画像を示す図である。

【図8】本実施形態の手法により3次元モデルを用いて調整した衣服画像を示す図である。

【図9】本実施形態の手法によりX軸、Y軸、Z軸周りの傾きを正規化した顔の3次元モデルを示す図である。

【図10】ヘアスタイルデータベースの作成処理を示すフローチャートである。

【図11】実写映像とヘアスタイル画像の合成処理を示すフローチャートである。

【図12】頭部の3次元位置・姿勢の推定処理の流れを示したフローチャートである。

【図13】本実施形態の手法による推定姿勢の補正結果を示す図である。

【図14】本実施形態の手法によるヘアスタイルの抽出結果を示す図である。

【図15】歩いている人物に仮想衣服を合成した実施例を示す図である。

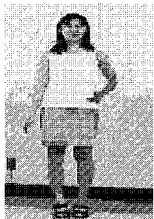
【図16】人物の頭部の動きに合わせてヘアスタイルを合成した実施例を示す図である。

10

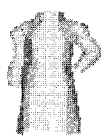
20

【図1】

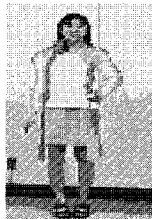
(a) オリジナル



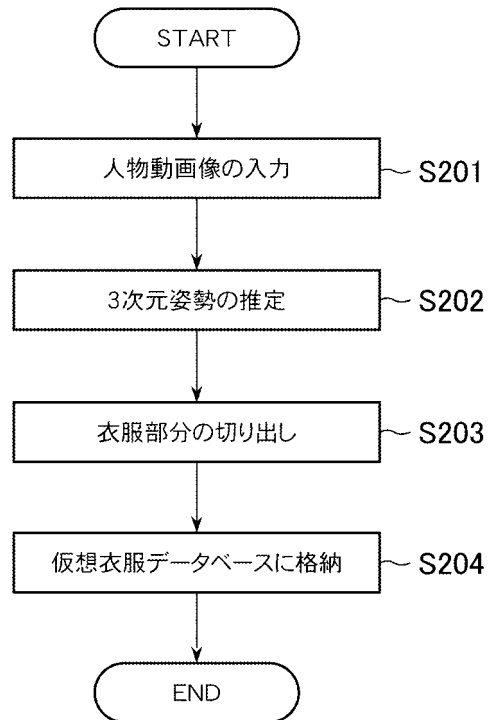
(b) CG衣服



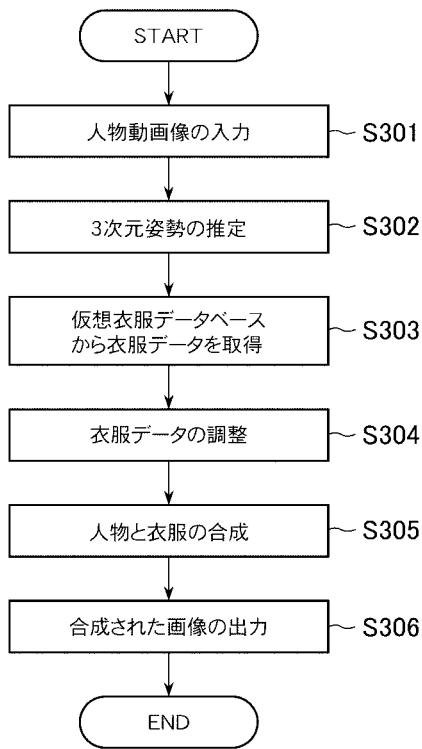
(c) 合成結果



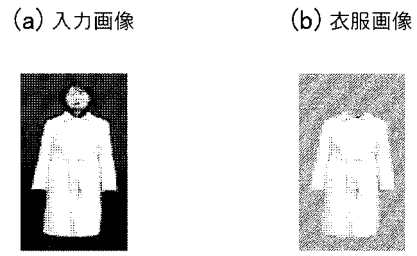
【図2】



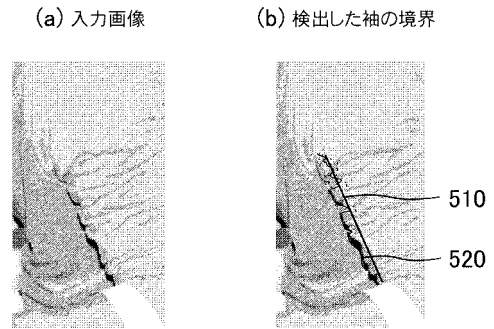
【図3】



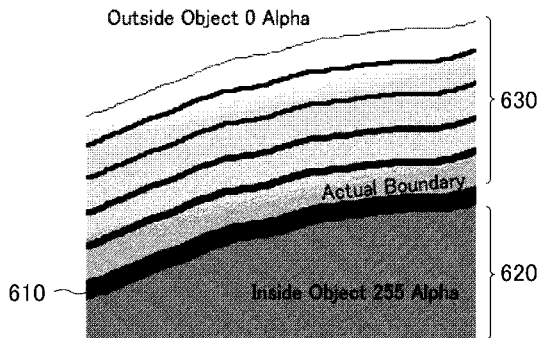
【図4】



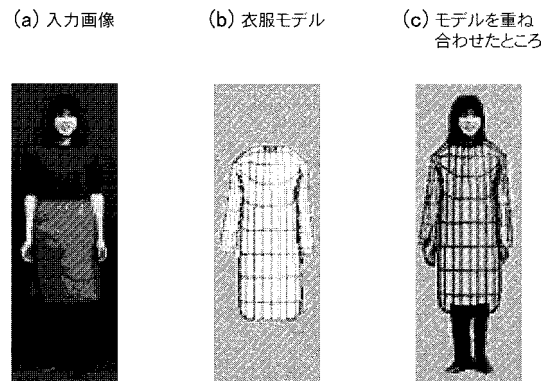
【図5】



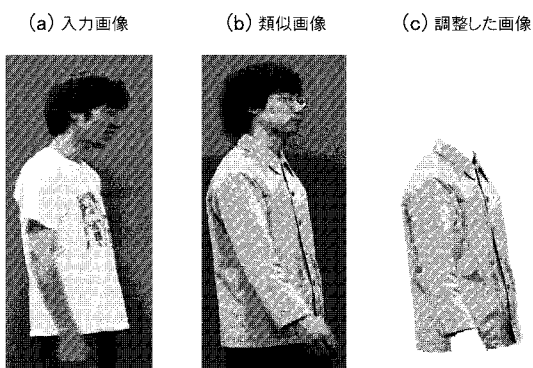
【図6】



【図8】

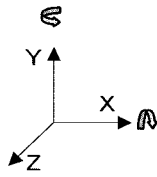


【図7】

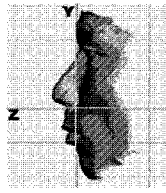


【図9】

(a) 本実施形態で扱う座標系



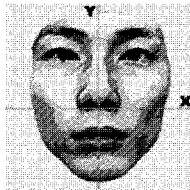
(b) X軸周りの正規化



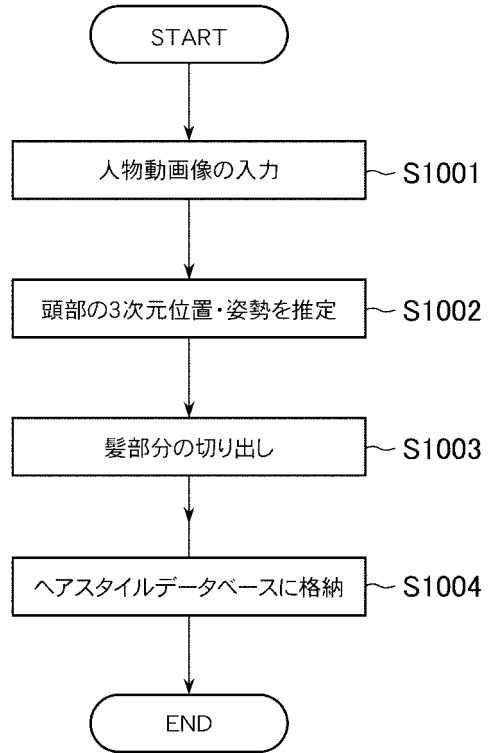
(c) Y軸周りの正規化



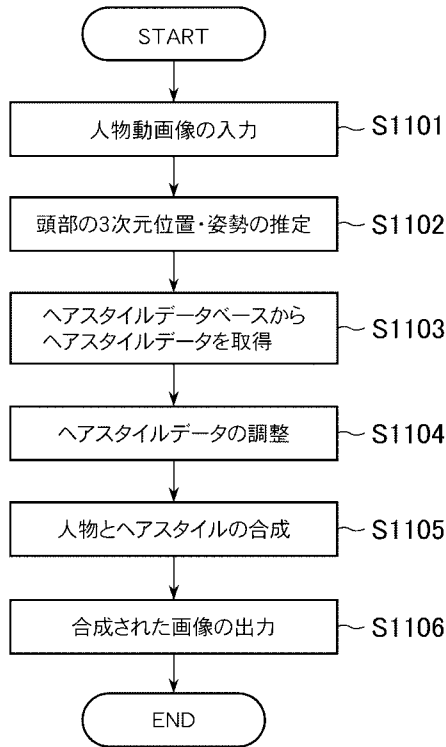
(d) Z軸周りの正規化



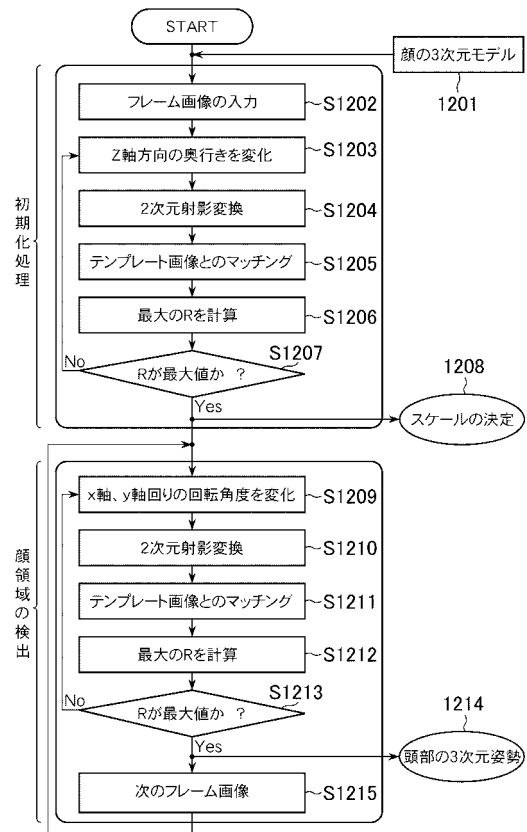
【図10】



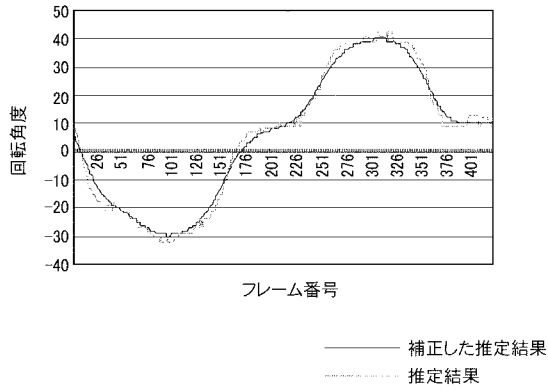
【図11】



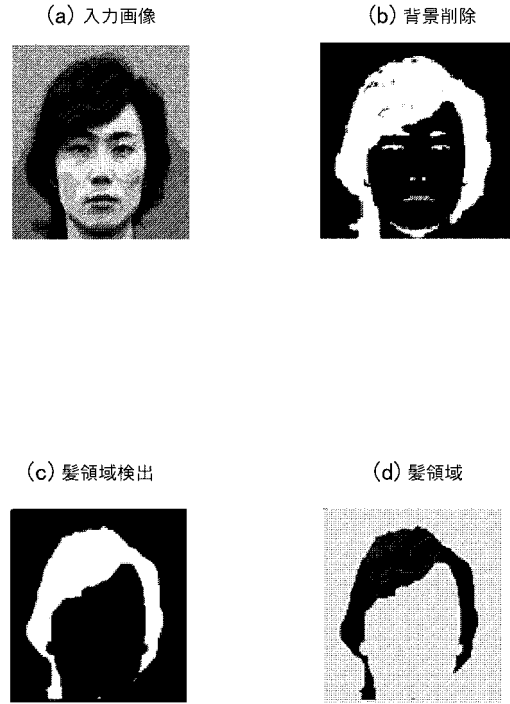
【図12】



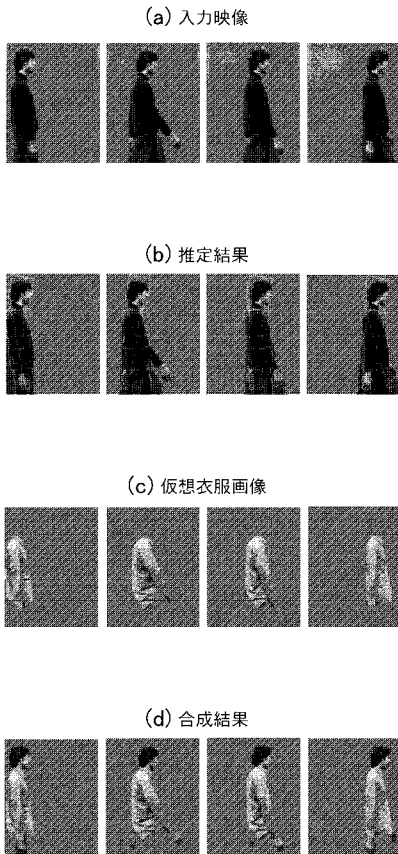
【図 13】



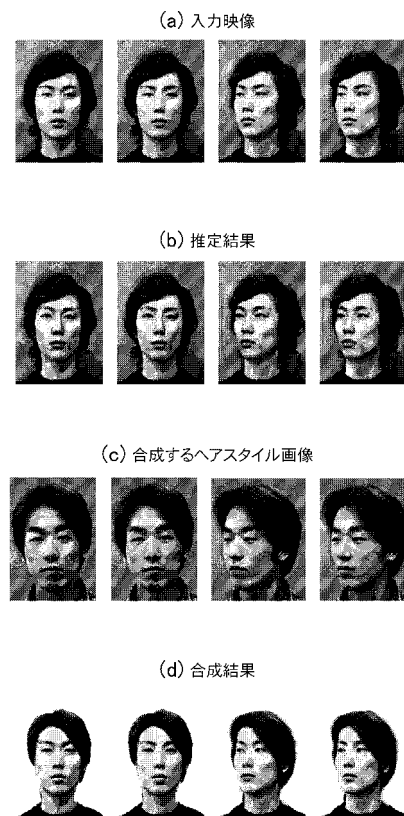
【図 14】



【図 15】



【図 16】



フロントページの続き

- (56)参考文献 特開平05 - 336445 (JP, A)
特開2001 - 084408 (JP, A)
特開2002 - 269580 (JP, A)
特開平08 - 131252 (JP, A)
特開平08 - 293038 (JP, A)
星野准一、外1名、ビデオ映像とCGの合成によるヴァーチャルファッションの実現、情報処理学会論文誌、日本、社団法人情報処理学会、2001年 5月15日、第42巻 第5号、P1182~P1193
星野准一、人物動作における個性の抽出と再構成、画像電子学会誌、日本、画像電子学会、2001年 9月25日、第30巻 第5号、P631~P640

(58)調査した分野(Int.Cl., DB名)

G06T 1/00-17/40