

# 仮想試着システムにおける 体型推定を用いた衣服画像合成技術

Clothes Image Synthesis Technologies for Virtual Fitting System Using Body Shape Estimation

関根 真弘      杉田 馨      フランク ペルベール

■ SEKINE Masahiro      ■ SUGITA Kaoru      ■ Frank PERBET

東芝は、ユーザーの全身画像に衣服画像を合成することで、仮想的な試着体験を提供する仮想試着システムを開発している。対象の衣服をあたかもほんとうに着ているかのように見せるためには、モニタに映し出されたユーザーの体型に応じて衣服画像を精度よく合成する必要がある。

そこで、距離センサから得られた距離データに基づいてユーザーの体型を約1秒で自動推定する技術と、体型の推定結果や肩の輪郭線に基づいて衣服画像の合成時の大きさや位置を算出する技術を開発した。これらの技術により、フィット感の高い画像合成が可能となり、システムの実用性を高めることができた。

Toshiba has been developing a virtual fitting system that provides users with a virtual experience of trial fitting of clothes, by synthesizing images of the clothes on whole-body images of the user to simulate the wearing of those clothes. In order to enhance the reality of the experience in such virtual fitting systems, it is necessary to accurately transform the images of the clothes on the monitor display according to the body shape of each user.

We developed two technologies to address this issue: a technology that automatically estimates the user's body shape in approximately one second based on distance data captured by a depth sensor, and a technology that determines the size and position of the clothes images based on the estimated body shape and shoulder line. Our newly developed virtual fitting system incorporating these technologies realizes accurate synthesis of fitting images, thereby offering enhanced practical applicability.

## 1 まえがき

近年、一般消費者向け小売業界では、新たなユーザー価値を創造することでショッピングスタイルの差異化を図る動きが活発になっている。そのような背景の下、数多くの衣服から自分に合うものを選ぶ行動を支援するための仮想試着システムが注目を集めている。

仮想試着システムとは、大型モニタに映し出されたユーザーの等身大の全身画像に対して衣服を合成することで、その衣服を試着しているようすを姿見の鏡で見ているかのような体験を提供するものである。このシステムにおいて、ユーザーに本物の衣服を着ているように感じさせるためには、主に二つの点を考慮する必要がある。

一つ目は、衣服の色や柄、質感などを忠実に再現することである。合成する衣服コンテンツとして、2次元画像を用いる手法と3次元CG(コンピュータグラフィックス)を用いる手法がある。2次元画像は、本物の衣服を撮影したものであるため見た目の再現性が高い。一方、3次元CGは、実時間で高品位に描画することが難しい。小売業界で商品の魅力をアピールすることを目的とする場合、2次元画像を用いる手法が適していると考えられる。

二つ目は、ユーザーの全身画像に対して衣服を、大きさや位置がずれることなく合成することである。ユーザーが立って

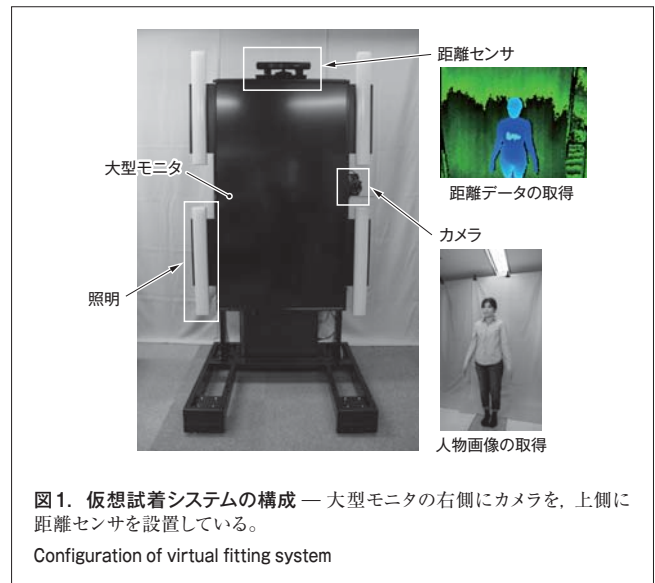


図1. 仮想試着システムの構成 — 大型モニタの右側にカメラを、上側に距離センサを設置している。

Configuration of virtual fitting system

る空間をセンシングする手法として、カメラ画像を用いる手法と距離センサから得られる距離データを用いる手法がある。距離データを用いる手法は、カメラ画像の色や輝度を用いる手法に比べて、背景が著しく変化するシーンでのセンシングに向いている。人が行き交う店頭や通路に設置する場合、距離データを用いる手法が適していると考えられる。

東芝は、距離データを用いてユーザーの体型を推定し<sup>(1)</sup>、推

定した体型モデルに基づいて、人物画像に精度よく2次元の衣服画像を合成する技術を開発した<sup>(2),(3)</sup>。この技術により、ユーザーの体型に応じたフィット感の高い画像合成が可能となり、仮想試着システムの実用性を高めることができた。

ここでは、仮想試着システム向けに開発した当社の衣服画像合成技術とその評価結果について述べる。

## 2 仮想試着システムの概要

仮想試着システムの構成を図1に示す。大型モニタの右側にカメラを、上側に距離センサを設置する構成にしている。カメラでは人物画像を取得し、距離センサでは人物が立っている空間の距離データを取得する。

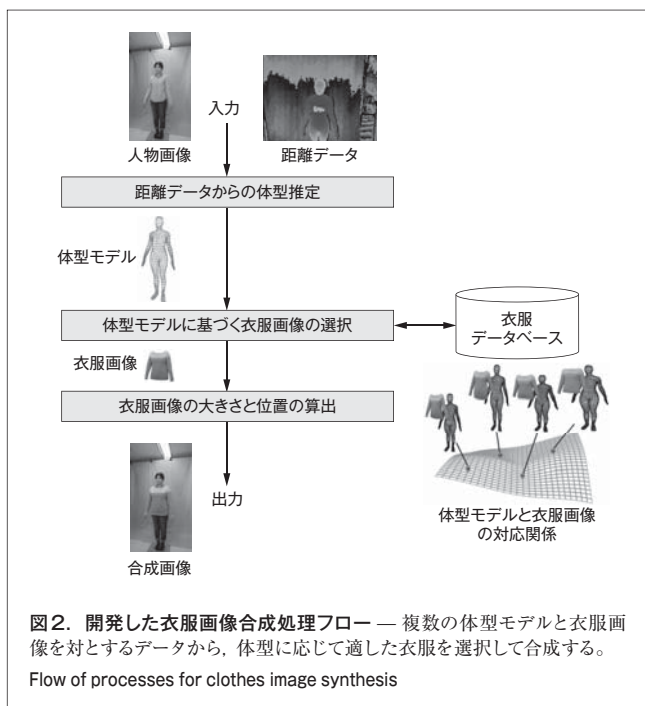
距離センサには、距離データに加えて人物の骨格中の関節位置を実時間で取得できるものがある。既存技術<sup>(4)</sup>では、この骨格に合わせて衣服画像を合成している。しかし、人物の骨格を基に合成すると体型の違いによって、衣服画像の大きさや位置がずれてしまうという問題がある。

そこで当社は、距離データからユーザーの体型を推定し、その体型に応じて衣服画像の大きさや位置を算出する技術を開発した。この技術について以下に説明する。

## 3 体型推定による衣服画像合成

### 3.1 全体の処理フロー

開発した衣服画像合成の処理フローを図2に示す。事前に衣服データベースを構築する際、従来は、ある衣服に対して



1枚の衣服画像だけを準備するが、ここでは、衣服を着た人物若しくはマネキンの体型モデルと衣服画像とを対応付け、ある衣服に対して複数の体型モデルと衣服画像を対するデータを準備しておく。

実行時の処理としては、まず、入力したワンショットの距離データからユーザーの体型を推定する。次に、推定した体型モデルに基づいて適切な衣服画像を衣服データベースから選択する。最後に、ユーザーの立ち位置に応じ、選択されたその衣服画像の合成時の大きさや位置を算出し、人物画像に合成する。

このような処理によって、ユーザーが仮想的に衣服を試着しているような合成画像が出力される。

### 3.2 距離データからの体型推定

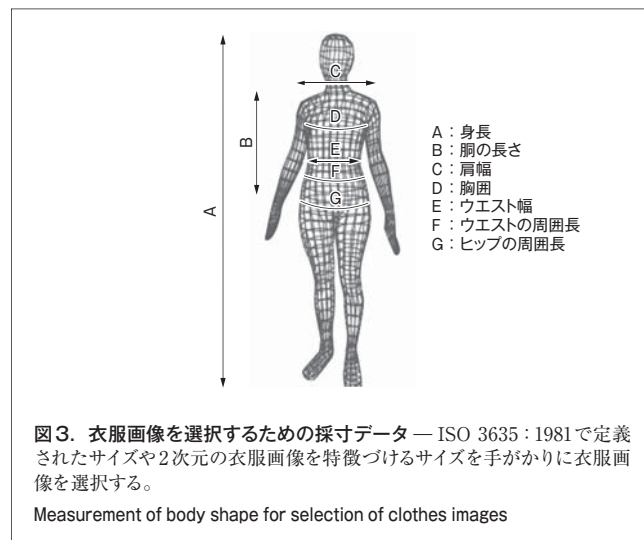
距離センサから得られるワンショットの距離データを用いて、事前学習した様々な人物の3次元体型データベースからもっとも類似する体型を高速に検索する技術を開発した。

数百万パターンの体型モデルを学習し、おおまかな体型の特徴と詳細な体型の特徴とを分離して検索できるようにすることで高速化した。また、抽出された体型モデルをユーザーに合わせて更に変形することで、衣服内部のユーザー本来の体型を推定する。これらの処理を約1秒で行うことができる。

### 3.3 体型モデルに基づく衣服画像の選択

ある体型で撮影した衣服画像を異なる体型の人物に合成すると、衣服の袖や裾の丈がずれてしまう問題があるため、より類似した体型の衣服画像を用いることで、合成時のずれを解消できると考えられる。

そこで、推定した人物の体型モデルに基づき、衣服データベースの中から、もっとも類似した体型モデルと類似した衣服画像を選択する。ここでは、図3に示すような体型を表現する様々な採寸データを用いる。身長、胸囲、ウエストの周囲長、及びヒップの周囲長はISO 3635 (国際標準化機構規格



3635) : 1981で定義されたサイズから引用した。また、胴の長さ、肩幅、及びウエスト幅は、2次元の衣服画像を特徴づける独自のサイズを用いた。このような採寸データを基に、体型の違いを式(1)の評価値  $r$  によって評価する。

$$r = \sum_i k_i |p_i - q_i| \quad (1)$$

ここで、 $p_i$  は推定した人物の体型モデル上の各採寸データ、 $q_i$  は衣服データベースに含まれる体型モデル上の各採寸データ、 $k_i$  は各採寸データに対する重み係数である。

算出した  $r$  がもっとも小さくなる、体型モデルに対応する衣服画像を選択する。

### 3.4 衣服画像の大きさと位置の算出

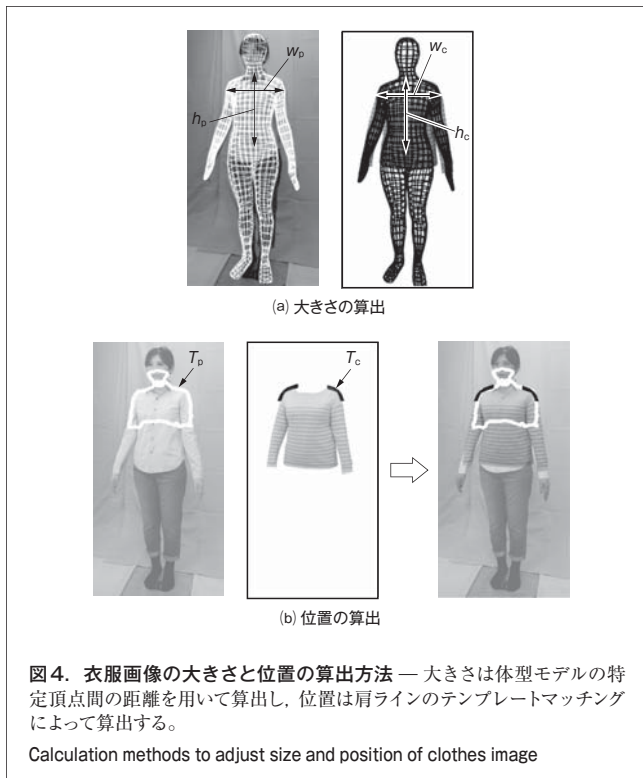
衣服画像の大きさと位置の算出方法を図4に示す。

大きさの算出は、体型モデルの特定頂点間の距離を用いて、横方向のスケール  $s_w$ 、縦方向のスケール  $s_h$  を式(2)で求める。

$$s_w = \frac{w_p}{w_c} \cdot s_h = \frac{h_p}{h_c} \quad (2)$$

ここで、 $w_p$ 、 $w_c$  は、人物と衣服それぞれの体型モデルにおける肩幅、 $h_p$ 、 $h_c$  は、人物と衣服それぞれの体型モデルにおける胴の長さを示す。

位置の算出は、人物と衣服の画像から抽出した肩ラインを用いる。肩ラインは、衣服を着装するうえで重要なポイントとなるという知見<sup>(5)</sup>に基づき、人物と衣服の肩ラインに対してテンプレートマッチング処理を行う。具体的には、図4(b)に示すような人物と衣服からそれぞれ抽出した肩ラインの領域  $T_p$ 、 $T_c$  を照



合し、二つの領域がもっとも合致する位置を算出する。

このような処理によって、フィット感の高い衣服画像の合成が可能になる。

## 4 評価

### 4.1 フィッティング精度の定義

ここでは、合成画像のフィッティング精度を評価する。フィッティング精度として、式(3)で定義するF値を用いる。

$$F = \frac{2R}{N+C} \quad (3)$$

ここで、 $R$  は図5に示すような理想画像と合成画像が合致する衣服領域のピクセル数、 $N$  は合成画像の衣服領域のピクセル数、 $C$  は理想画像の衣服領域のピクセル数である。

F値は0.0から1.0までの値となり1.0に近いほどフィッティング精度が高いことを示す。

### 4.2 衣服画像の選択に関する評価

まず、体型モデルに基づく衣服画像を選択した場合としない場合のフィッティング精度を比較した。52人の人物と10種類の衣服のデータセットを用いた評価を行った。衣服画像を選択しない場合は、評価対象としている人物以外の人物が着て撮影した衣服画像を合成した。

評価の結果、表1に示すように、衣服画像の選択の有無に

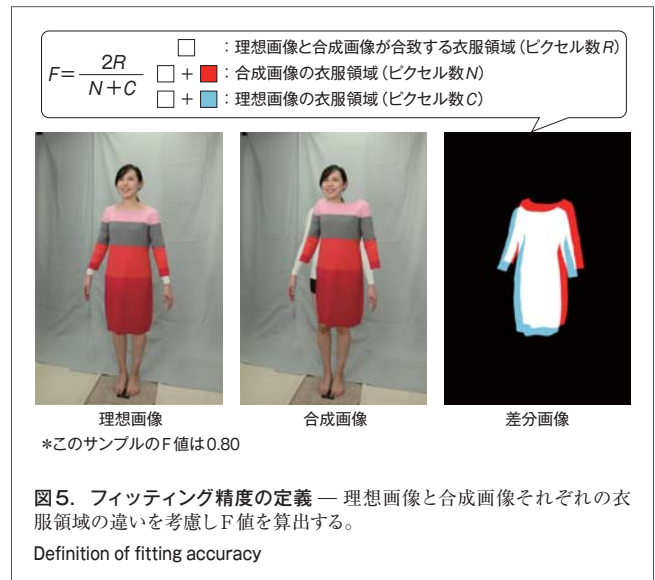


表1. 各評価におけるF値の結果

Results of F-measure values obtained in each clothes image synthesis evaluation case

評価項目	衣服画像の選択		衣服画像の大きさと位置の算出	
	衣服選択なし	衣服選択あり	既存技術	開発技術
平均値	0.916	0.921	0.919	0.932
最悪値	0.769	0.842	0.800	0.848

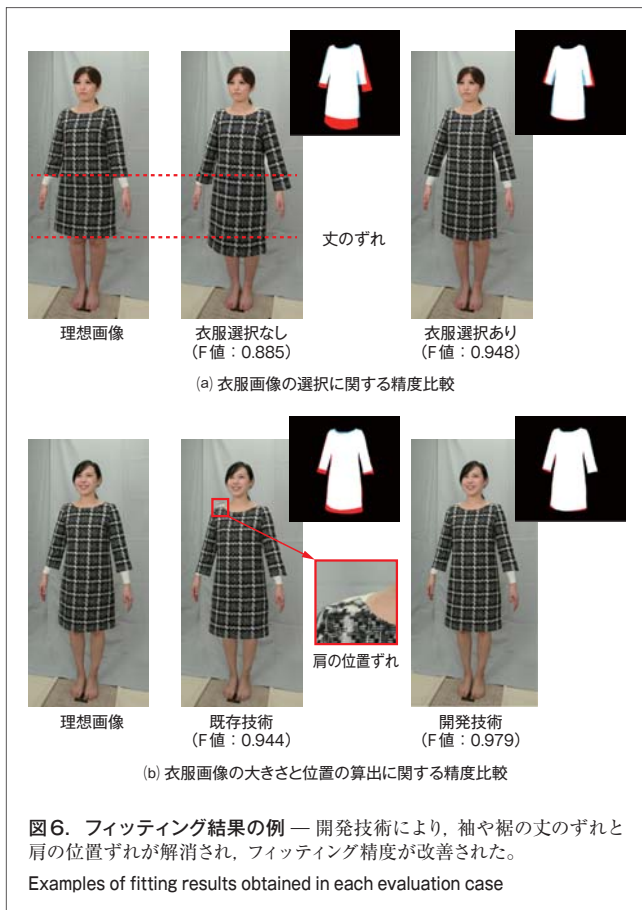


図6. フィッティング結果の例 — 開発技術により、袖や裾の丈のずれと肩の位置ずれが解消され、フィッティング精度が改善された。  
Examples of fitting results obtained in each evaluation case

よって、F値の平均値が0.916から0.921に、最悪値が0.769から0.842に改善されたことがわかり、衣服画像の選択の効果を確認することができた。外観的にも、図6(a)に示すように、袖や裾の丈のずれを解消できることを確認した。

### 4.3 衣服画像の大きさと位置の算出に関する評価

次に、衣服画像の大きさと位置の算出に関する評価を行った。骨格に合わせて衣服画像を合成する既存技術<sup>(5)</sup>を比較対象とし、それぞれのF値を比較した。データセットは前節で示したものと同一ものを用いた。

評価の結果、表1に示すように、開発技術によって、F値の平均値が0.919から0.932に、最悪値が0.800から0.848に改善されたことがわかり、肩ラインに注目した位置合わせの効果を確認することができた。外観的にも、図6(b)に示すように、肩ライン付近のずれを解消できることを確認した。

## 5 あとがき

フィット感の高い衣服画像合成を実現することを目的とし、距離データから推定したユーザーの体型モデルを用いた衣服画像合成技術を開発した。合成時のフィット感が高まることで、図7に示すような仮想試着システムの利用シーンに活用できることを確認した。

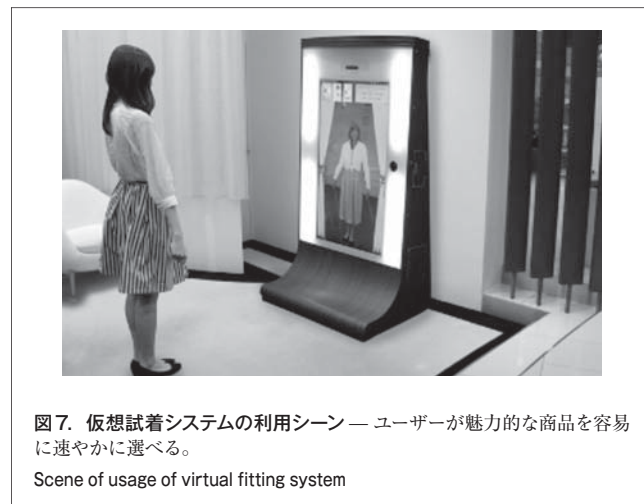


図7. 仮想試着システムの利用シーン — ユーザーが魅力的な商品を容易に速やかに選べる。  
Scene of usage of virtual fitting system

十分な精度を確保するために、どれくらいの体型のバリエーションで衣服データベースを構築すべきかを検討することが、課題として挙げられる。また、異なる姿勢の人物画像や、異なる照明環境で撮影した衣服画像に対する合成方法を検討する必要がある。今後も、フィッティング精度の更なる改善に向けて開発を進めていく。

## 文献

- (1) Perbet, F. et al. "Human body shape estimation using a multi-resolution manifold forest". Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA, 2014-06, IEEE Computer Society. 2014, p.668 - 675.
- (2) 杉田 馨 他. 人物シルエットへの時空間位置合わせによる衣服画像の合成. 電子情報通信学会論文誌 D. J97-D, 9, 2014, p.1519 - 1527.
- (3) Sekine, M. et al. "Virtual Fitting by Single-shot Body Shape Estimation". Proceedings of the 5th International Conference on 3D Body Scanning Technologies. Lugano, Switzerland, 2014-10, Hometrica Consulting. 2014, p.406 - 413.
- (4) Zhou, Z. et al. "Image-based clothes animation for virtual fitting". Proceedings of ACM SIGGRAPH ASIA 2012. Singapore, 2012-11, 12, ACM SIGGRAPH. 2012, Article No.33.
- (5) 大沼 淳 他 監修. ファッション辞典. 文化出版局, 1999, 767p.



関根 真弘 SEKINE Masahiro

研究開発統括部 研究開発センター インタラクティブメディアラボラトリー研究主務。コンピュータグラフィックス及びコンピュータビジョンに関する研究・開発に従事。映像情報メディア学会会員。Interactive Media Lab.



杉田 馨 SUGITA Kaoru, Ph.D.

研究開発統括部 研究開発センター インタラクティブメディアラボラトリー研究主務、博士(情報理工学)。コンピュータグラフィックス及びコンピュータビジョンに関する研究・開発に従事。Interactive Media Lab.



フランク ペルベール Frank PERBET, Ph.D.

東芝欧州研究所 ケンブリッジ研究所 コンピュータビジョングループ, Ph.D.。コンピュータビジョン及びコンピュータグラフィックスに関する研究・開発に従事。Toshiba Research Europe Ltd.