

# 手描きスケッチ入力による衣服3Dモデルの生成手法

何 毅<sup>1,a)</sup> 謝 浩然<sup>1,b)</sup> 宮田 一乗<sup>1,c)</sup>

**概要:** 単視点画像からの3Dモデル再構成はディープラーニングの発展により大きな進歩を遂げたが、主流のテンプレートメッシュの変形による再構成手法は衣服などの非水密3Dモデルの再構成には不向きとされている。また、手書きスケッチは手軽にデザイン案を出せるため、画像入力による手法より使いやすい。本研究では、手描きスケッチ入力から3D衣服モデルを生成できる衣服デザイン支援システムを提案する。既存研究と異なり、提案手法は入力スケッチ画像から目標となる3Dモデルの符号なし距離関数(UDF)を推測し、算出したUDFからメッシュの抽出を行う。さらに、Marching Cubesにより生成されたメッシュの編集も可能となった。提案システムで生成した衣服モデルに対し、定量評価と主観評価両方で本研究の有効性を確認した。

## 1. はじめに

衣服の3DモデリングはAR・VR技術やメタバースが発展している今、バーチャル試着、オンラインショッピング、キャラクター3D化などの場面で広く利用されている。これらの需要が爆発的に増える一方、デザイナーによる衣服のデザインプロセスは複雑で、アイデアを検証するにも時間がかかる。一方、手描きスケッチはストロークで構成され、目標オブジェクトに対して様々な抽象化が施されているにもかかわらず、人間はそのスケッチから対応する3Dオブジェクトを想像することができる。そのため、スケッチがデザインの初期段階で使われていることが多い。デザイナーにとって、デザインをスケッチから実物に落とすよりも、仮想空間でデザインを再現し、アイデアの検証を行う方が手っ取り早く、かつコストが低い。そのため、スケッチを利用した3Dモデルの作成ツールが必要となる。

スケッチにおける3Dオブジェクトの生成はここ数年で広く研究されてきているが、衣服の生成問題を解決するための研究は管見の限りまだ発表されていない。既存研究のうち、スケッチを用いて3D形状データベースから目標形状を探索する手法[24]では、入力スケッチを予め処理した輪郭データとの比較で目標形状を検索する。また、特定スケッチのストロークを三次元空間へマッピングし、3Dモデルを作る手法[11]も代表的なものである。これらの手法は三次元形状の取得において有効ではあるが、前者におい

て、既に存在している形状の取得しか対応していない、後者の場合、ユーザに一定水準の3Dモデリングスキルを要求しているため、素人には難しい。

ディープラーニングの発展により、簡単なスケッチ画像一枚でユーザが思い描く目標物の生成がより簡単に、より精密になった[8], [27]。Zhangら[27]が提案した手法は、テンプレートメッシュを目標形状に変形させることによる生成法で、近年の単視点画像からの3Dモデル再構成では主流の手法である。Guillardら[8]は陰関数(Implicit Function)学習に基づいた手法を提案しており、符号付き距離関数(SDF, Signed Distance Function)の学習により目標形状を生成する。しかしながら、Zhangらの手法はテンプレートのトポロジー構造によって結果が制限され、Guillardらの手法の生成結果も水密(Watertight)モデルに限る。そのため、既存手法では、衣服のような非水密(Non-Watertight)モデルの生成には対応できない。ここ数年、符号なし距離関数(UDF, Unsigned Distance Function)による三次元オブジェクトの表現[4]は非水密モデルや複雑表面およびシーンも対応でき、応用範囲が広いことがわかった。

この論文では、UDFを用いて、スケッチによる衣服3Dモデル再構成手法を提案する。研究概要を図1に示す。本研究はAuto-Encoder構造を利用し、UDFs表現の学習を行う。Auto-Encoderの出力はUDFデータとなり、またMarching Cubes手法[9]によりUDFからメッシュ表面を抽出する。また、素人にも手軽にスケッチの作成および生成された3D衣服モデルの編集ができるユーザインターフェイスを開発した。生成されたモデルの輪郭を編集する

<sup>1</sup> 北陸先端科学技術大学院大学  
JAIST, Nomi, Ishikawa 923-1292, Japan

a) s2010035@jaist.ac.jp

b) xie@jaist.ac.jp

c) miyata@jaist.ac.jp

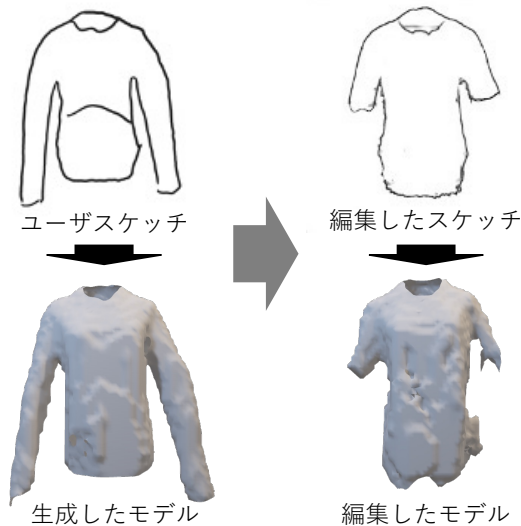


図 1: 研究概要図

ことで、潜在ベクトルの最適化が行える。最適化された潜在ベクトルを Decoder によって、修正したモデルとして出力する。最後に、提案システムを使用し、スケッチからの衣服 3D モデルの生成効果を実験で検証した。

## 2. 関連研究

### 2.1 単視点画像による衣服 3D モデル再構成

近年、深層学習の爆発的発展により、一枚の単視点画像からの 3D モデル再構成手法も続々と発表されている [1], [13], [19], [28]。それらの大多数はテンプレートベース手法であり、予め用意されたテンプレートメッシュを目標形状に変形させることが手法の肝である。この手法では、生成結果はテンプレートメッシュのトポロジー構造に制限され、大まかな形状しか生成できない。そのため、衣服のような複雑形状を表現するには、他の追加情報を提供する必要がある。例えば、Pavlakos ら [19] はパラメトリック 3D モデルを利用し、画像から抽出した 2 次元の特徴に適合するようにモデルのパラメータを最適化することで人体モデルを生成する。DeepWrinkles[13] は、用意した衣服のテンプレートメッシュの法線マップに微細な衣服のしわを追加することで、グローバルな形状変形と表面の詳細を併合して表現している。Tex2shape[1] は形状表現問題を画像の style transfer 問題と考え、法線マップなどモデル表面情報を予測し、用意された人体モデルに追加することで着衣した人のモデルを生成した。より複雑な表面情報を提供し、複雑な表面の学習を目指したが、これらの手法は用意されたテンプレートメッシュのトポロジー構造の制限を突破できない。

テンプレートベース手法とは別に、PIFu[22] は陰関数学習に基づき、着衣した身体 3D 再構成のための深層学習フ

レームワークを提案した。より多くの表面形状特徴を保持するために、三次元空間上の全ての点  $(x, y, z)$  情報と 2D 画像画素を局所的に整列させ、より複雑な表面構造を再現することに成功した。Zhao ら [28] は陰関数学習ベースの手法に基づいて、RGB 画像から表面周辺に配置された微細な表面に特徴を代表する Anchor Points と呼ばれるキーポイント群を予測して、より精度の高い 3D 表面表現ができる深層学習フレームワークを提案した。これらの手法は陰関数の学習に基づいたもので、テンプレートベース手法の制限がない。しかし、これらの手法は、衣服写真を入力としており、スケッチベースの生成は難しい。

### 2.2 陰関数学習に基づいた 3D モデル再構成

陰関数学習に基づいた 3D 再構成手法は多数発表されている。このような手法は Implicit Field による 3D オブジェクトの表現によって成立し、点群と同じような利点を持っている。代表的なものに、Occupancy Field[15] と SDF[17], [21] や UDF[4], [28] がある。これらの手法は連続的な形状表現を学習でき、再構成された 3D モデルに解像度の制限もない。また、Occupancy Field や SDF による陰関数学習には、水密 (Watertight) モデルの生成しか対応できないが、UDF による手法は水密モデルも含め、非水密モデルまたは複雑表面の生成も対応できる。陰関数によるモデル表現はテンプレートベース手法よりも、テンプレートのトポロジー構造に制限されない、自由なサイズのモデルの表現ができる、また非水密モデル生成 (UDF による) も対応できるなどの点において有利と言われている。衣服データは基本非水密なため、内外判定はできなく、UDF を用いた陰関数学習は最適手法である。また、B.Guillard ら [9] は UDF から高速にメッシュを生成する手法を提案したことで、UDF から効率的にメッシュを抽出することも可能になった。

本研究はこれらの手法を踏まえて、UDF による陰関数学習手法を用いた、衣服の非水密 3D モデル手法を提案する。

### 2.3 スケッチベース形状生成及び衣服デザイン支援

スケッチによる形状生成手法はこの数年広く研究されてきている。初期の試みとして、五十嵐ら [11] は手描き線分を特定形状にマッピングすることで、スケッチだけで三次元モデルの作成が可能になった。また、深層学習の爆発的発展により数々の学習ベース手法も提案されている [6], [14]。これらの手法は複数視点情報を必要とする。単視点スケッチによる 3D モデルの生成例として、Zhang ら [27] はスケッチから直接にメッシュモデルを生成する手法を提案したが、生成した形状の ill-pose 問題 (例えば入力スケッチはどの方向から見たものが曖昧など) 解決を中心としたため、複雑な形状の生成とリアルタイムのモデル生成と編集は対応できていない。B.Guillard ら [8] は SDF

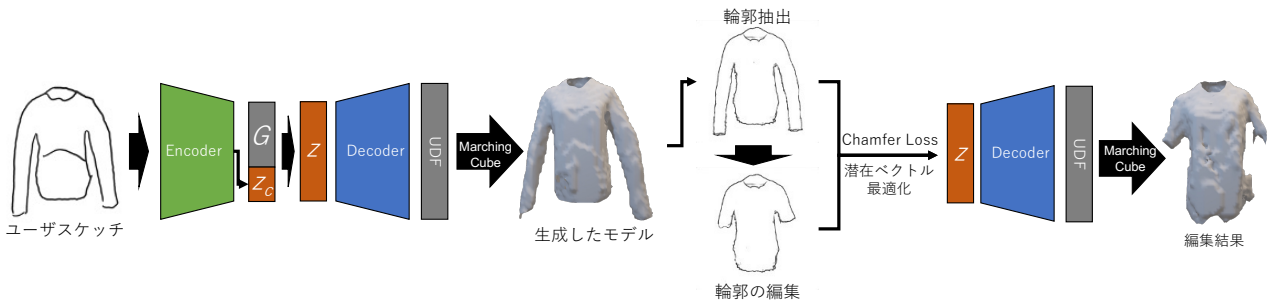


図 2: 提案システムのフレームワーク

を利用し、ハイクオリティメッシュモデルの生成および編集手法を提案した。しかしこれらの手法は水密モデルしか生成できなく、服装など非水密モデルの生成に対応できない。それ以外に、スケッチによる衣服のデザイン手法も発表されている。Wang ら [25] はスケッチ、メッシュデータおよびパラメータが共有する潜在空間の学習を行い、衣服形状を人体モデルにマッピングすることで着衣した人体モデルの生成を提案した。Chen ら [3] は 2D 衣服スケッチを入力として、人体形状パラメータと組み合わせ、特定の人体形状にフィットする 3D 衣服メッシュを生成する手法を提案した。

これらの手法は着衣した人体の生成がベースで、衣服だけの生成は対応できていない。またユーザに 3D モデリングやデザインの専門知識を要求しているため、素人にとってハードルが高い。これらの問題点を解決するため、本研究はスケッチ作成および生成したモデルとのインタラクションができる使いやすいユーザインターフェースの開発を行う。

### 3. 提案手法

本研究は Auto-Encoder 構造 [17] ニューラルネットワークを実装して、スケッチから衣服の 3D モデル再構成を行った。そして、UDF データを衣服のモデルからサンプリングし、データセットの構築を行った。また、スケッチの作成及び生成したモデルを閲覧・編集するためのユーザインターフェースを開発した。

#### 3.1 フレームワーク

本研究の概要を図 2 に示す。本研究の目的は入力スケッチから 3D モデルを再構成することである。3D モデルの再構成は符号なし距離場 UDF に対して等値面抽出と三角メッシュ生成によって行う。システムはスケッチ画像を潜在コード  $z_c$  にエンコードするエンコーダーと潜在ベクター  $z$  から UDF を生成するデコーダーの二つのネットワークにより構成される。 $z$  は三次元空間のランダムサンプリング点群 Grid  $G$  と  $z_c$  の結合で、デコーダーは  $z_c$  を基づいて全ての  $p_i \in G$  に対して符号なし距離  $d_i^u$  を予測し、UDF を

出力する。そして Marching Cubes 手法 [9] を利用し、3D モデルを出力する。生成したモデルは、提案されたユーザインターフェイスによって編集操作もできる。

#### 3.2 スケッチ画像のデータ収集

本研究が使用する Deep Fashion3D データセットはペアになるスケッチ画像を提供していない。スケッチを入手するために、まずモデルごとにレンダリング画像を生成する必要がある。生成されたレンダリング画像を使用し、スケッチ画像を生成する。RGB 画像からのスケッチ画像の生成には輪郭抽出手法を使用する。画像内の輪郭抽出手法は数多く存在しているが、代表的な手法は Suggestive Contours[5] や Holistically-nested Edge Detection (HED) [26] がある。本研究では、最低限のストロークだけを確保するため、Suggestive Contours ではなく、HED 手法を適用してスケッチ画像を入手する。本手法で入手したスケッチデータとペアとなるメッシュデータを図 3 に示す。



図 3: 生成したスケッチとペアになるメッシュ

#### 3.3 ネットワークの学習

本研究では UDF 学習に図 4 に示す Auto Encoder 構造 [17] を使う。入力スケッチを潜在ベクターにエンコードするエンコーダーを  $E$ 、UDF データ生成のデコーダーを  $D$ 、 $S$  を入力スケッチとする。入力スケッチ  $S$  はエンコーダー  $E$  によって潜在コード  $z_c$  にエンコードされる。また、Grid sampling は 3D 形状が存在する  $[-1, 1]^3$  の三次元空

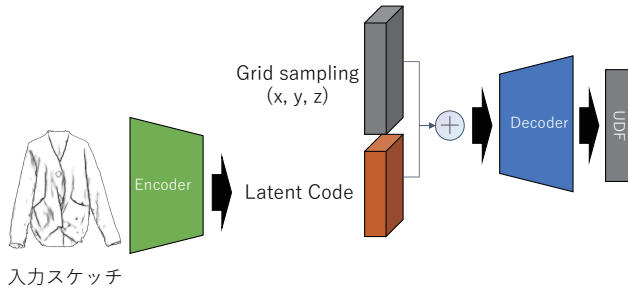


図 4: 提案システムのネット構造

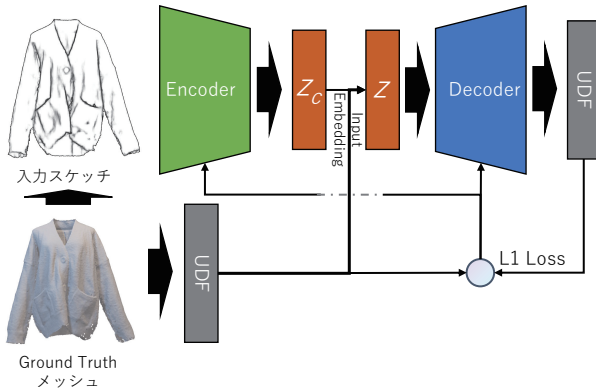


図 5: ネットワークの学習フロー

間内に離散正規 Grid をサンプリングした結果であり、 $G$  で表す。潜在コード  $z_c$  は  $G$  と接続して潜在ベクトル  $z$  になり、デコーダー  $D$  によって UDF に生成する。この過程を式 (1) に示す。デコーダーは入力全ての座標  $p_i \in G$  に対して、符号なし距離  $d_i^u$  を予測する。ネットワークの学習フローを図 5 に示す。

$$UDF = D(E(S), G) \quad (1)$$

### 3.3.1 ネットワーク最適化目標

ネットワークの学習目的は、与えられた  $S$  に対し、最も適合する UDF を生成できることである。学習における最適化目標は  $L1$  距離を利用して式 (2) で示す。ここで、 $df_{gt}^u$  をグラウンドトゥルースとする。ネットワークに散乱しているサンプル座標の影響を軽減するため、 $df_{gt}^u$  及び  $D(z)$  は  $\delta$  の値を閾値に離散値の除去を行う。また、入力の潜在ベクトル  $z$  に対して正規化を行い、正規化結果に対しても最適化を行う。正規化損失を式 (3) に示す。  $\lambda$  は正規化強度を指す。以上を踏まえ、ネットワークの損失関数を式 (4) で定義する。

$$L(D(z), df_{gt}^u) = |clamp(D(z), \delta) - clamp(df_{gt}^u, \delta)| \quad (2)$$

$$L(z_i) = \lambda \|z_i\|_2 \quad (3)$$

$$L = L(D(z), df_{gt}^u) + L(z_i) \quad (4)$$

### 3.3.2 生成した 3D モデルの最適化

生成結果の最適化及び編集において、本研究は B.Guillard ら [8] の手法を参考し、生成された 3D モデルにレンダリングを行い、HED で輪郭を抽出する。その結果とスケッチの Chamfer 距離を最小化することで、デコーダー入力としての潜在ベクトルの最適化を行う。最適化の制限条件には Chamfer 距離を用いる。本研究では、Chamfer 距離を 2D に実装する。2D Chamfer Loss  $L_{CD}$  を式 (5) で定義する。ここで、 $S_p$  は 3D メッシュを投影した輪郭を指し、 $F_s$  はスケッチ領域を表す。

$$L_{CD} = \sum_{p_p \in S_p} \min_{p_s | F_s(p_s)=0} \|p_p - p_s\|^2 + \sum_{p_s | F_s(p_s)=0} \min_{p_p \in S_p} \|p_p - p_s\|^2 \quad (5)$$

生成結果の最適化の本質は生成結果と入力間の損失によって勾配情報の Backpropagate が、UDF データも 3D モデルも微分可能性がないため、最適化の目標はデコーダー入力の潜在ベクトルとなる。潜在ベクトルはネットワークの学習過程の一部として、微分可能性を持ち合わせているため、最適化が可能である。その最適化によってネットワークの出力結果が調整できる。具体的に、生成した 3D メッシュを  $M$  とし、その表面頂点のサンプリングを  $V$  とする場合、潜在ベクトル  $z$  に対し、勾配  $\frac{\partial L_{CD}}{\partial z}$  を式 (6) で算出する。算出した勾配を Backpropagate 過程を通して潜在ベクトルの最適化を行う。

$$\frac{\partial L_{CD}}{\partial z} = \sum_{v \in V} -\frac{\partial L_{CD}}{\partial v} \nabla D(v, z) \frac{\partial D}{\partial z}(v, z) \quad (6)$$

そして、最適化した潜在ベクトルを用いて、デコーダーによる UDF データの生成と 3D モデルの再構成を行う。モデルの編集過程を図 6 に示す。

### 3.4 ユーザインターフェイス

提案システムは、スケッチ入力による非水密衣服 3D モデルの生成システムとスケッチ作成及び編集用ユーザインターフェイスによって構成する。提案するユーザインターフェイス (以下 UI) を図 7 に示す。本 UI は二つの領域で構成する。左の領域はスケッチの作成及び編集の機能を備えて、右の領域は生成したオブジェクトの閲覧機能が搭載されている。本 UI のボタンは上部に配置されている。「Sketch Pad」領域において、ボタンアイコンの左から右それぞれは「ブラシ」、「消しゴム」、「クリア」と「保存」であり、これらの機能をユーザが利用してスケッチの作成を行う、完成したら「保存」を押し、作成したスケッチからモデルの生成が始まる。また、右側の上部に Model Viewer の操作ボタンが配置されている。左から右へそれぞれ「キャ



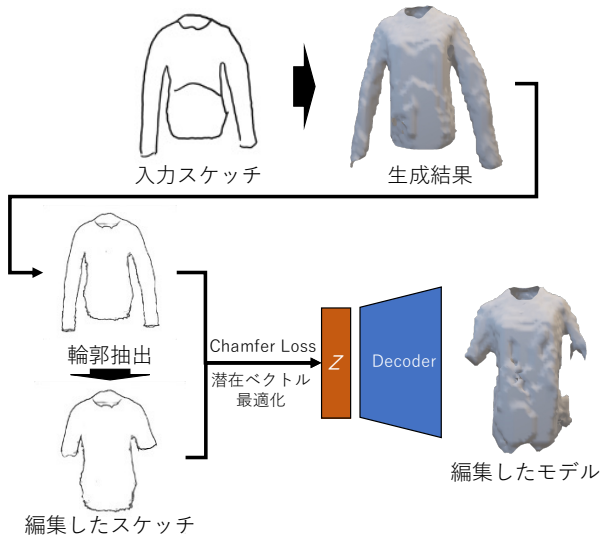


図 6: モデル編集のフレームワーク

プチャー」, 「リセット」と「モデル保存」になる。生成したオブジェクトは右側に表示され、マウススクロールで拡大、縮小、右クリック長押しでマウスを移動すると回転するなどの操作ができる。ユーザは「キャプチャー」ボタンで今の視点のスクリーンショットを入手し、システムはそれに基づきスケッチ画像を生成する。生成した画像は Sketch Pad に配置され、スケッチの作成機能を活用して編集を行う。編集が完了した後、もう一回「保存」を押すことで、システムは現在のモデルに対して修正したスケッチを基準に最適化を行い、モデルの編集ができる。また、「リセット」機能で、ユーザの編集内容が棄却され、最初の生成結果に戻ることができる。最後に、完成した 3D モデルに対し、「モデル保存」を押すことでダウンロードして保存することが可能である。

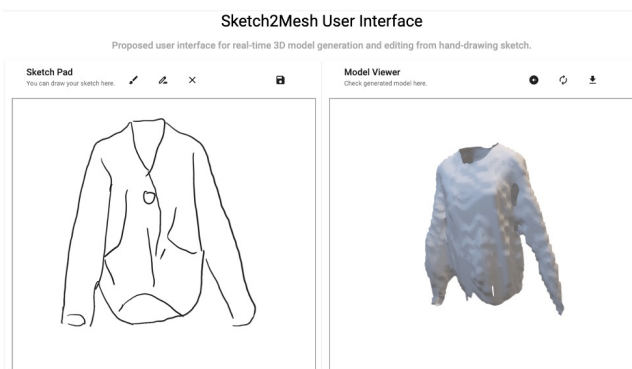


図 7: 提案ユーザインターフェース

## 4. 実装

### 4.1 ユーザインターフェイス

本研究において提案した UI は、利便性を目指して、ウェ

ブアプリの形で実装した。このような UI は計算処理を全部サーバ側が担当しているため、ユーザ側の端末に負荷をかけずにシステムを利用できる。本 UI は Angular[7] によって実装した。また、UI と 3D モデル生成システムを繋ぐために、サーバシステムを Flask[16] で実装した。ユーザがスケッチを作成してからモデルを取得するまでの所要時間は平均 1.13 秒であった。

### 4.2 ネットワーク

本システムの深層学習部分は Pytorch[18] 及び Pytorch3d[20] で実装した。UDF の生成にて使われている Auto Encoder は B.Guillard ら [9] の実装に従う。デコーダーは 9 層の MLP (Multilayer perceptron) 層で構成され、MLP 層の隠れ次元数は 512 とし、活性化関数は ReLU を利用する。また、入力の三次元座標に対し、フーリエ位置エンコーディング [23] を行うことで活性化関数によるディテール情報の損失を軽減する。スケッチ画像を潜在ベクトルにエンコードするエンコーダー部分は ResNet[10] を使用している。学習において、UDF による  $L1$  損失 (式 (2)) と入力正規化損失 (式 (3)) の和を損失関数とする。散乱値除去因数  $\delta$  を 0.1 にし、正規化強度  $\lambda$  を  $10^{-4}$  にする。

また、バッチサイズは 16、エポック数は 2000 で学習した。オプティマイザーには Adam[12] を使用し、学習率はエポック数に応じ、式 (7) に従って設定した。  $lr_{init}$  を初期学習率とし、初期値は 0.0005 に設定した。  $\omega$  は 0.5、  $\gamma$  は 500 である。エンコーダに対し、  $\alpha = 1$ 、デコーダーの場合、  $\alpha = 0.1$  に設定した。学習は二基の NVIDIA GeForce RTX 3090 を利用して行った。一つのエポックにつき学習時間は 150 秒前後、全エポック合わせておよそ 80 時間となる。

$$learning\_rate = \alpha \times lr_{init} \times (\omega^{\lfloor \frac{epoch}{\gamma} \rfloor}) \quad (7)$$

### 4.3 データセット

本研究は衣服の 3D モデル再構成として非水密モデルの生成を試みるため、Deep Fashion3D データセットを使用し実験を行う。本データセットは衣服メッシュ 2,075 個が含まれる。本研究において、学習セットと検証セットに分けられ、それぞれの数は 1,867 と 208 である。スケッチによる UDF を生成するには、スケッチと UDF のペアデータを用意する必要がある。

まずメッシュデータに対して、学習に適するために座標値を最大 1 に正規化する。そして、サンプリングを行う際は表面の外部と内部両方から行うため、モデル中心から距離が一番遠い表面の外からもサンプリングできるよう 1 より少し小さい任意半径の表面に納められるようにスケール調整し、調整されたメッシュに対しサンプリングを行う。

経験的に、本研究ではスケールを 0.8 に設定する。

そしてメッシュ  $M_i$  毎に、頂点のサンプリング数を  $N$  とし、 $N$  個の三次元頂点座標  $p_i$  から  $M_i$  表面への距離  $d_i^u$  を求める。サンプリング数に制限はないが、データ処理速度を考慮し、なるべく時間がかからずに多くサンプリング数を取ることができるよう  $N$  を 120,000 と設定する。サンプリングデータは表面上、表面付近とメッシュを囲むバウンディングボックス内のランダム点によって構成する。サンプリングの重心をメッシュ表面に置いた。具体的には、経験に沿って、サンプリング数のうち、48,000 個は表面との距離が 0.05 以内の座標から、32,000 個は表面からの距離が 0.3 以内の座標からサンプリングする。また、24,000 個は  $M_i$  表面からランダムにサンプリングし、最後の 16,000 個はエッジ長さ 2 (範囲が  $[-1, 1]$ ) のバウンディングボックス内をサンプリングする。

$d_i^u$  は KDTree[2] を利用して算出する。メッシュ  $M_i$  自体に対して特に処理 (水密化など) をする必要がない。スケッチ画像は HED[26] によるレンダリング画像を用いる。メッシュ  $M_i$  毎に、36 方向からレンダリング画像を生成し、HED ネットワークに渡してスケッチ画像を入手する。

## 5. 評価

本システムの評価には、レンダリングによる生成されたスケッチと本研究が提案したユーザインターフェイス両方を使用した。また、提案手法で生成した衣服 3D モデルに対し、既存手法 (sketch2mesh[8]) との比較と合わせて、数値評価でも本手法の効果を検証する。

### 5.1 スケッチによる衣服 3D モデルの生成

スケッチによる衣服モデルの生成では、Deep Fashion3D[29] が提供する衣服データを使う。システムの生成効果を検証するため、レンダリング画像から生成したスケッチ画像とユーザインターフェイスを利用して作成したスケッチ画像の両方を使う。

本手法を利用して衣服モデルの生成を試みた結果を図 8 に示す。入力スケッチに対して、本システムは安定に非水密モデルを出力できる。また、図 9 に示す手描きスケッチに対する生成結果から、本システムの出力クオリティはあまり変わらなかったことがわかる。

### 5.2 既存スケッチベースメッシュ生成手法との比較

スケッチによる 3D モデルの生成手法の代表的なものとして、sketch2mesh[8] を比較対象として挙げる。SDF ベースの手法のため、本研究に使われた非水密モデルを水密化して学習を行う。比較結果を図 10 に示す。それぞれの正面図と上面図を示す。既存の SDF に基づいた手法はモデルの再構成ができるが、一つの連続表面としてしか再構成できない。したがって、現実的な再構成タスクには対応で

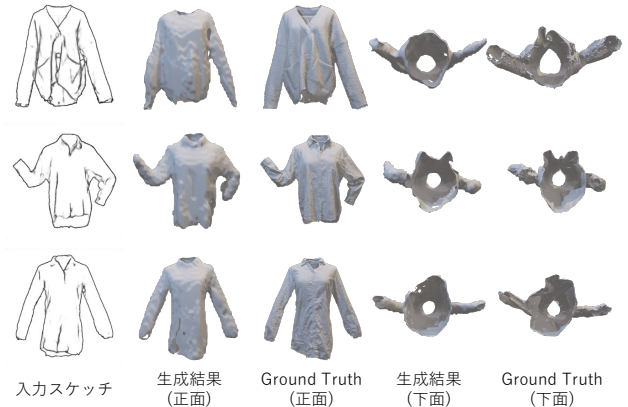


図 8: 生成されたスケッチにより再構成された衣服モデルの結果



図 9: 手描きスケッチにより再構成された衣服モデルの結果

きなく、役に立つことが難しい。一方、本手法では UDF を活用しているため、非水密モデルなど複雑表面の表現も対応できる。

数値評価の結果として、既存手法と本手法による生成モデルの Chamfer Distance と Earth Mover's Distance を表 1 に示す。本手法はより低い Chamfer 誤差と EMD 誤差を得られていることがわかる。

表 1: 生成したモデルの Chamfer Distance と Earth Mover's Distance.

Method	CD( $\times 10^{-3}$ )	EMD( $\times 10^{-2}$ )
sketch2mesh	4.59	7.65
ours	<b>4.12</b>	<b>6.84</b>

### 5.3 提案システムにより生成結果の編集

3.3 節に述べたように、本システムは出力モデルに対し、編集を行うこともできる。図 11 に示すように、ユーザは生成結果に不満を持つ部分に対して 3D モデルを直接編集するのではなく、操作が容易な 2D スケッチを介しての編

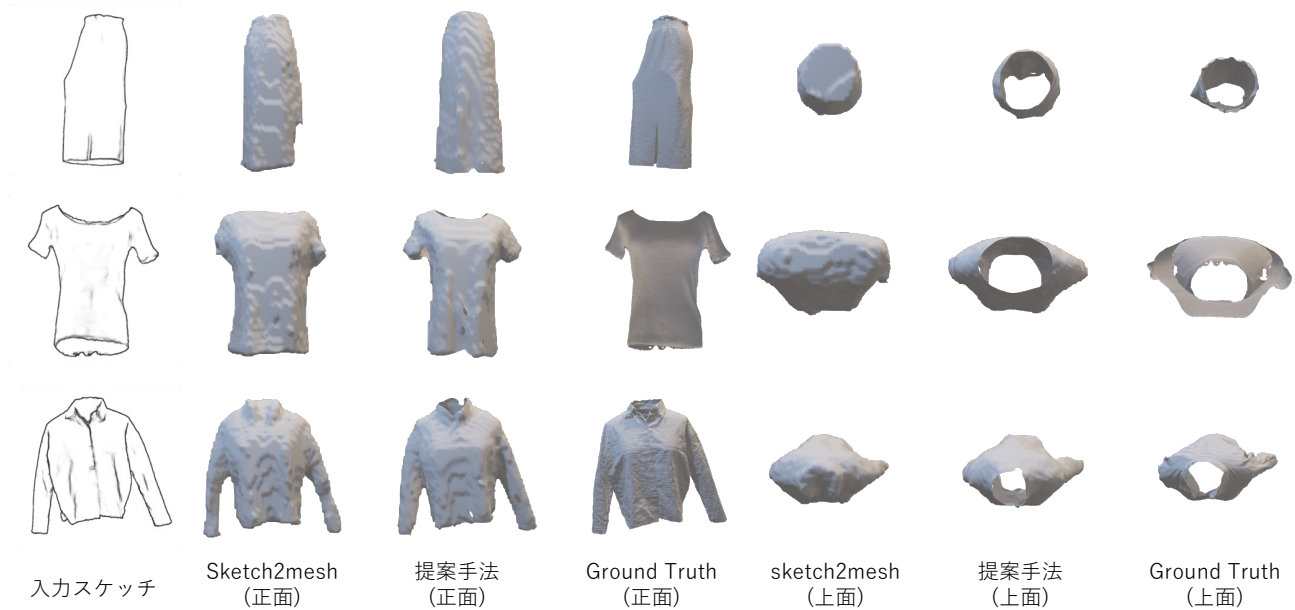


図 10: 既存手法との比較

集が可能となった。



図 11: スケッチから 3D 再構成したモデルの編集結果

## 6. 結論と今後の展望

本研究は、手描きスケッチ入力による衣服 3D モデル生成手法を提案した。UDF による陰関数学習によって、非水密的な衣服モデルの生成が可能となった。これにより、3D モデリングの専門知識を持ち合わせていないユーザでも、簡単に所望の衣服を 3D モデル化することが可能になった。

本研究において、UDF の学習をスケッチベース生成に適用して、衣服の 3D モデル生成に成功したが、衣服の表面特徴、ポケットやボタンなどの表現はまだ困難である。これは、スケッチ画像から画素レベルの特徴抽出を施したが、それらの特徴を 3D モデル上の目標構造と関連付けされていないことに起因する。この問題の解決を本研究の今後の課題とする。

また、本手法は 2D Chamfer Distance を利用し生成モデルの輪郭と入力スケッチを基づくモデルの最適化を行った

が、この手法で得られた結果は時に図 12 に示すような不安定な生成結果（予期しない空洞があくか編集前の衣服のトポロジー構造が残留するなど）が得られる場合がある。現在の最適化目標はある程度輪郭をスケッチにマッチングできるが、輪郭内の表面上に対して制限条件がないため、最適化の結果、表面にこのような破綻が出現するようになったと考えられる。より安定的にモデルの編集を行うため、編集過程の最適化にシルエット推測による入力と出力シルエットの IoU (Intersection over Union) 損失の導入などを検討する必要がある。

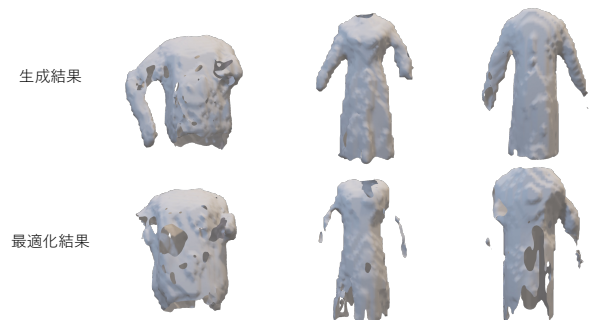


図 12: 編集過程において最適化失敗例

本手法により生成された 3D モデルは全てのパーツ（袖、ボタン、ポケットなど）は一体化されていて、実際のデザイン用途ではそれぞれのパーツの分割が必要と予想される。これに対する解決法も本研究の課題として検討している。また、本研究は UDF を使用し、スケッチから非水密モデルの代表例である衣服モデルの生成に成功したが、その他、スケッチ入力からの自動車の外観および内装や、

室内シーンの3D再構成などもUDFに基づいた手法が適している[4]と考える。

スケッチはデザイナーにとってアイデアを一番早く表現できるツールの一つであるため、スケッチによる3Dモデル再構成は複雑なシーンの制作に活用できることが期待できる。

## 謝辞

本研究の一部は、北陸先端科学技術大学院大学研究拠点形成支援事業および科研費20K19845の支援により実施された。

## 参考文献

- [1] Alldieck, T., Pons-Moll, G., Theobalt, C. and Magnor, M. A.: Tex2Shape: Detailed Full Human Body Geometry From a Single Image, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2293–2303 (2019).
- [2] Bentley, J. L.: Multidimensional Binary Search Trees Used for Associative Searching, *Commun. ACM*, Vol. 18, No. 9, p. 509–517 (online), DOI: 10.1145/361002.361007 (1975).
- [3] Chen, Y., Xian, C., Jin, S. and Li, G.: 3D Shape-Adapted Garment Generation with Sketches, *CGI* (2021).
- [4] Chibane, J., Mir, A. and Pons-Moll, G.: Neural Unsigned Distance Fields for Implicit Function Learning, *ArXiv*, Vol. abs/2010.13938 (2020).
- [5] DeCarlo, D., Finkelstein, A., Rusinkiewicz, S. M. and Santella, A.: Suggestive contours for conveying shape, *ACM SIGGRAPH 2003 Papers* (2003).
- [6] Delanoy, J., Bousseau, A., Aubry, M., Isola, P. and Efros, A. A.: 3D Sketching using Multi-View Deep Volumetric Prediction, *Proc. ACM Comput. Graph. Interact. Tech.*, Vol. 1, pp. 21:1–21:22 (2018).
- [7] Google: Angular, the modern web developer’s platform. (2010).
- [8] Guillard, B., Remelli, E., Yvernay, P. and Fua, P.: Sketch2Mesh: Reconstructing and Editing 3D Shapes from Sketches, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13003–13012 (2021).
- [9] Guillard, B., Stella, F. and Fua, P.: MeshUDF: Fast and Differentiable Meshing of Unsigned Distance Field Networks, *ArXiv*, Vol. abs/2111.14549 (2021).
- [10] He, K., Zhang, X., Ren, S. and Sun, J.: Deep Residual Learning for Image Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778 (2016).
- [11] Igarashi, T., Matsuoka, S. and Tanaka, H.: Teddy: a sketching interface for 3D freeform design, *SIGGRAPH ’99* (1999).
- [12] Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, *CoRR*, Vol. abs/1412.6980 (2015).
- [13] Löhner, Z., Cremers, D. and Tung, T.: DeepWrinkles: Accurate and Realistic Clothing Modeling, *ECCV* (2018).
- [14] Lun, Z., Gadelha, M., Kalogerakis, E., Maji, S. and Wang, R.: 3D Shape Reconstruction from Sketches via Multi-view Convolutional Networks, *2017 International Conference on 3D Vision (3DV)*, pp. 67–77 (2017).
- [15] Mescheder, L. M., Oechsle, M., Niemeyer, M., Nowozin, S. and Geiger, A.: Occupancy Networks: Learning 3D Reconstruction in Function Space, *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4455–4465 (2019).
- [16] Pallets: Flask, web development, one drop at a time. (2010).
- [17] Park, J. J., Florence, P. R., Straub, J., Newcombe, R. A. and Lovegrove, S.: DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 165–174 (2019).
- [18] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J. and Chintala, S.: PyTorch: An Imperative Style, High-Performance Deep Learning Library, *Advances in Neural Information Processing Systems 32*.
- [19] Pavlakos, G., Choutas, V., Ghorbani, N., Bolkart, T., Osman, A. A. A., Tzionas, D. and Black, M. J.: Expressive Body Capture: 3D Hands, Face, and Body From a Single Image, *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10967–10977 (2019).
- [20] Ravi, N., Reizenstein, J., Novotny, D., Gordon, T., Lo, W.-Y., Johnson, J. and Gkioxari, G.: Accelerating 3D Deep Learning with PyTorch3D, *arXiv:2007.08501* (2020).
- [21] Remelli, E., Lukoianov, A., Richter, S., Guillard, B., Bagautdinov, T., Baque, P. and Fua, P.: *Advances in Neural Information Processing Systems*.
- [22] Saito, S., Huang, Z., Natsume, R., Morishima, S., Kanazawa, A. and Li, H.: PIFu: Pixel-Aligned Implicit Function for High-Resolution Clothed Human Digitization, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2304–2314 (2019).
- [23] Sitzmann, V., Martel, J. N. P., Bergman, A. W., Lindell, D. B. and Wetzstein, G.: Implicit Neural Representations with Periodic Activation Functions, *ArXiv*, Vol. abs/2006.09661 (2020).
- [24] Wang, F., Kang, L. and Li, Y.: Sketch-based 3d shape retrieval using convolutional neural networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1875–1883 (2015).
- [25] Wang, T. Y., Ceylan, D., Popović, J. and Mitra, N. J.: Learning a shared shape space for multimodal garment design, *ACM Transactions on Graphics (TOG)*, Vol. 37, pp. 1 – 13 (2018).
- [26] Xie, S. and Tu, Z.: Holistically-Nested Edge Detection, *International Journal of Computer Vision*, Vol. 125, pp. 3–18 (2015).
- [27] Zhang, S.-H., Guo, Y.-C. and Gu, Q.-W.: Sketch2Model: View-Aware 3D Modeling from Single Free-Hand Sketches, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6012–6021 (2021).
- [28] Zhao, F., Wang, W., Liao, S. and Shao, L.: Learning Anchored Unsigned Distance Functions with Gradient Direction Alignment for Single-view Garment Reconstruction, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12654–12663 (2021).
- [29] Zhu, H., Cao, Y., Jin, H., Chen, W., Du, D., Wang, Z., Cui, S. and Han, X.: Deep Fashion3D: A Dataset and



Benchmark for 3D Garment Reconstruction from Single  
Images (2020).