

ケーブルの三次元形状の復元

范 哲銘^{†1} 林 豊洋^{‡2} 大橋 健^{‡2}

概要：現代社会は電化が進んでおり、身の回りには大量の機器が存在する。これらの機器に電力や信号を送る為に大量のケーブルが使用されており、問題になっている。ロボットを利用してケーブルを整理できるようになれば大幅に便利になるが、現在のところロボットでケーブルを操作するシステムはない。この原因は、ケーブルは変形するため状態の認識が難しいからである。本研究では、ケーブルを対象として、ロボットでケーブルを操作する為に、2Dのケーブル画像から3Dケーブルモデルを作るシステムを開発した。

キーワード：ケーブル、画像処理、画像認識、深層学習、3Dモデル

Make 3D cable model from RGB image

ZHEMING FAN^{†1} TOYOHITO HAYASHI^{‡2} TAKESHI OHASHI^{‡2}

Abstract: Along with the electrification, there are so many electrical equipment are used in our daily life. In using those equipment we need cables to charge them or transmit signal and so on. Therefore there is a requirement on cables in our life. But, too many cables will become problems. We think if we can resolve those problems with robot it will save much time for people. But, unfortunately, from now, there is no system which can manipulate cables with robot. The reason is cable has a deforming nature that its shape is irregular. When you want to use robot to manipulate something you need to recognize it from picture. However, we can not recognize cables state and we can not record cables position correctly from picture. But, we think recognize cables from 3D model is possible. Therefore, in this paper, we made a system which can build a 3D voxel model base on 2D image as a beforehand for use robot to manipulate cables.

Keywords: Cable, Image Processing, Image Recognition, Deep Learning, 3Dmodel

1. はじめに

現代社会は電化が進んでおり、家、学校、会社などの場所には大量の機器が存在する。これらの機器では電力や信号を送る為にケーブルが使用されている。このように身近にケーブルは大量に存在しており、これが問題になることがある。例えば掃除や保守には時間がかかり、これらの操作は人に頼らざるを得ない。人間にとってこのような仕事は難しくはなく、幾つの簡単な動作の繰り返しで実現できる。しかし、この仕事に人手を無駄に使うのは勿体無いので、ロボットを利用してケーブルを整理できるようになれば大幅な人手を削減するなどが期待できる。しかし、ロボットがケーブルを扱うためには、ケーブルの位置と形を正確に把握できなければならない。残念ながら、現在のところケーブルの状態を認識できるシステムはない。本研究では、ロボットを利用してケーブルを整理することを想定して、ケーブルの状態を認識するために、深層学習技術を用いて2Dのケーブル画像から対応する3Dケーブルモデルを作る方法を提案する。

2. 従来方法の紹介

従来方法[1][2]はRGB画像と距離画像を入力として、画像処理技術を利用してケーブルの認識を行っている。我々は、これらの研究を基礎として新たな方法[3]を提案した。この方法では入力データを減少し、RGB画像一枚でケーブルを認識する。画像中のピクセルの連結を調べ、エッジなどの情報を抽出し、統合して画像中のケーブルの状態を把握する。しかし、この方法にはロボットで実用するためにはまだ欠点がある。例えば、認識の精度は解像度に依存しており、画像の背景色とケーブル色は単色しか対応できない。また、二次元情報しかないのも、ロボットで操作するには情報が足りない。これらの問題を解決する為に、深層学習技術を利用してシステムを大幅に変更した。

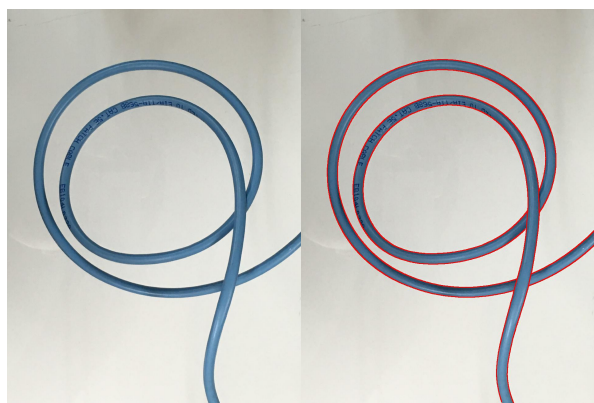
図1に論文[3]の結果を示す。図1の中の上半部分は成功した例で、下半部分は失敗した例である。成功した例にはケーブルの輪郭線を赤い曲線で重ねて表示している。このようにケーブルの輪郭位置と交差部分の上下関係が正しく示されている。

^{†1} 九州工業大学情報工学府

Kyushu Institute of Technology, Graduate School of Computer Science and System Engineering

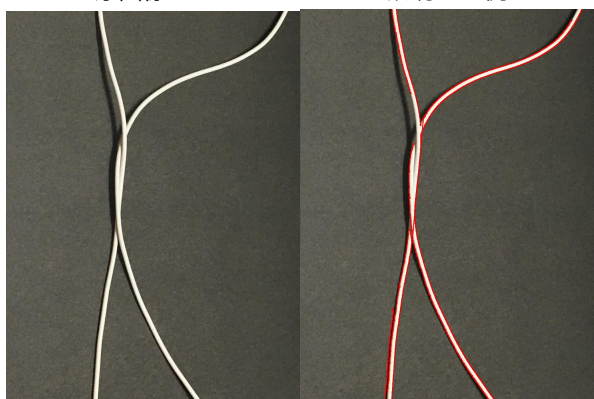
^{‡2} 九州工業大学情報科学センター

Kyushu Institute of Technology, Information Science Center



原画像

成功した例



原画像

失敗した例

図1 従来方法[3]の結果

3. システムの流れ

本研究ではケーブルの状態を認識するシステムを作成し、ロボットでケーブルを操作する為に必要な三次元形状を含む情報を集める。このシステムには四つの処理があり、その流れを図2に示す。

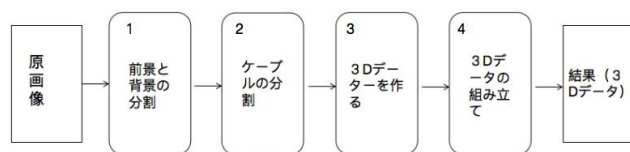


図2 システムの流れ図

システムの入力（図2の原画像）は通常のRGB画像であり、ケーブルはテーブルなどの台の上に置いて、上から撮影する。最終結果は画像中のケーブルに対応する3Dのボクセルデータである。四つの処理内容は以下の通りである：

- ① 画像内のケーブル部分を探す。
- ② ケーブル部分を分割し、原画像から一連の小画像を抽

出する。

- ③ 各小画像に対応する3Dデータを作る。

- ④ 一連の3Dデータを連結し、原画像に対応している3Dデータ作る。

この中で①、③部分には深層学習技術を利用した。②、④部分はOpenCVを利用してC++とpython言語で実装した。

4. 実現方法

4.1 前景と背景の分割

画像中のケーブル部分は前景として他の部分は背景に分割する。これは画像を図3のようにいくつかの種類に分割する処理である。



図3 前景と背景分割の効果

本研究では画像中のケーブル位置を認識する為にSegNet[4]を利用する。SegNetはオープンソースで提供されており、深層学習技術を利用する画像分割技術である。

SegNetは畳み込み式13回とmaxpooling5回（前半）プラス畳み込み式13回とupsampling5回（後半）を行うネット構造である。図4にSegNetの構造を示す。システムの入力は360*480のRGB画像で、出力は同じサイズの分割画像である。

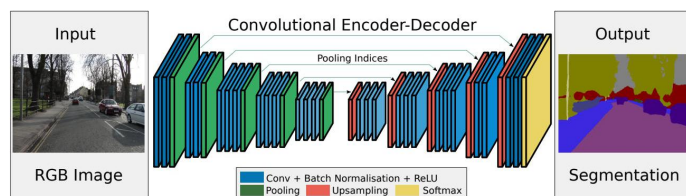


図4 SegNetのネット構造([4]より)

SegNetは11種類の目標に分割、認識する為に設計された。しかし、本研究の認識目標の種類は3種類だけなので、SegNetの構造を簡略化して利用する。

本研究では畳み込み式5回とmaxpooling5回（前半）プラス畳み込み式5回とupsampling5回（後半）のネット構造を使う（図4の構造の青い層を全部一つにしたものに相当

する)。入力は 360×480 の RGB 画像で、出力は同じサイズの分割画像である。

訓練ラベルは背景、単一ケーブル、複雑ケーブル（ケーブルの交差、結び、捻りの所）の三種類である。訓練データとラベルを図5に示す。左の図は原画像で、右の図は訓練画像（黒の部分は背景、赤い部分は単一ケーブル、緑の部分は複雑な部分のラベルを示す）である。

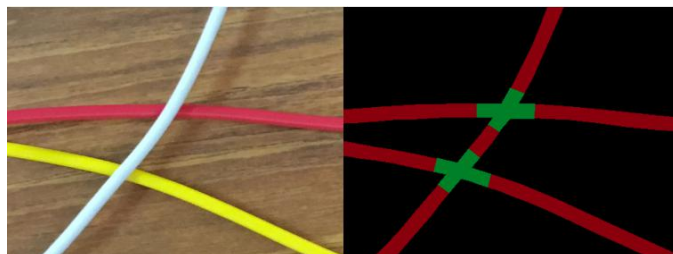


図5 訓練データとラベル

ネットワークの訓練後のテスト結果を図6に示す。



図6 訓練後のテスト結果

図6を見ると画像中のケーブル部分と背景部分はよく分割できている。この結果は次の処理の入力として利用する。従来方法と比べて、制限が少なくなり、性能も向上している。

4.2 ケーブル部分の分割

ケーブル分割は原画像のケーブル部分から一連の小画像に分割する。三次元形状の復元には深層学習技術を利用するが、大きい画像をそのまま復元することには二つの問題点がある。

- ① 画像内の目標は小さいので画像の情報が損失が多く含まれる。
- ② 復元部分のネットの訓練後のテスト結果は良くない

そこで原画像を多数の小画像に分割し、各小画像に対して3Dデータを復元する。そして、得られた小画像に対する3Dデータを連結してケーブル全体を表現することにする。図7に小画像の例を示す。

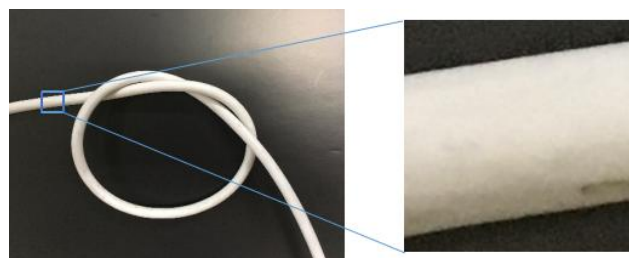


図7 小画像

小画像は以下のルールで切り出すことをする：

- 正方形の画像。
- 画像中のケーブル面積は小画像面積の3分の2程度。
- ケーブル部分の中心と小画像の中心は重なっている。

分割方法は sliding window 技術を使って実現する。図8にウィンドウボックスのサイズとパスの決める方法を示す。

図8の上の画像は背景分離した画像であり、この画像からケーブルの縦と横の幅（赤い線分）を求める。ウィンドウボックスは正方形なので、幅の中で短い方を選択して、ウィンドウボックスのサイズはこの幅の1.5倍に設定する。図8の例では横の幅を選択する。

図8の下画像は上の画像の細線化画像であり、細線化画像はケーブルの中心エッジ線を表示している。このエッジ線はウィンドウボックスの経路になる。図中の赤いボックスたちは移動経路に沿っており、ボックスの中心はエッジ線と重なっている。

サイズと移動経路を組み合わせ、スライディングボックスにより原画像を分割する。

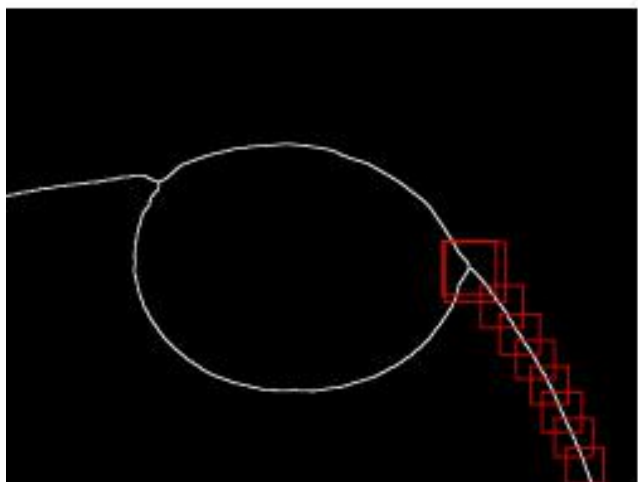
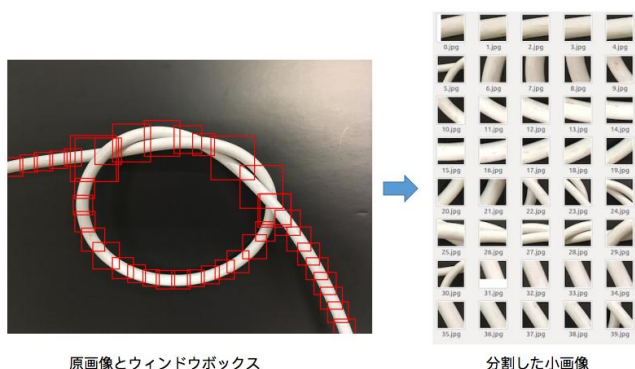


図8 sliding window の説明図

ケーブルの分割結果を図9に示す。



原画像とウィンドウボックス

分割した小画像

図9 ケーブル分割の結果

図9の左は原画像の上にウィンドウボックスを表示したものである。原画像のケーブルの状態に合わせてボックスのサイズは変化する。右の画像は分割した40枚の小画像（ウィンドウボックスの区域）である。

4.3 3Dデータを作る

3D-GAN[5]はディープラーニングを用いるオープンソースの2D画像から3Dボクセルデータを作る技術である。図10は3D-GANの流れ図である。

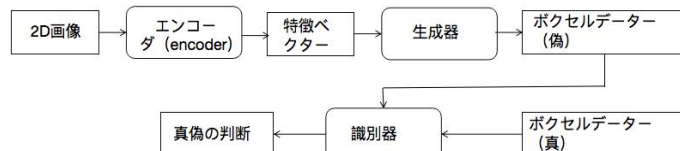


図10 3D-GANの流れ

GANはGenerative-Adversarial Networkの略名であり、GANには生成器と識別器二つの部分がある。この二つの部分を対抗して各自のパラメタを更新する。3D-GANはGANの枠組上でエンコーダを追加している。エンコーダの入力は2D画像、出力は特徴ベクターである。また、生成器の入力は特徴ベクター、出力は3Dボクセルデータである。そして、識別器の入力はボクセルデータ（真偽二種）、出力はボクセルデータの真偽を判断する。生成器は識別器が本物と区別ができない偽物データを作ること为目标である。識別器は生成器が作ったボクセルデータ（偽）と本物のボクセルデータの真偽を識別する为目标とする。3D-GANはエンコーダ、生成器、識別器に対応して三つのLOSS関数を設定する必要がある。識別器の認識成功率が高い時エンコーダと生成器のパラメタを更新し、識別器の認識成功率が低い時識別器のパラメタを更新する。

各部分のネット構造を図11に表す。

エンコーダ:		生成器:		識別器:	
conv2d	5	conv3d	5	conv3d	5
fc	1				
(batch,256,256,3)		(batch,1,1,1,200)		(batch,64,64,64,1)	
(batch,200)		(batch,64,64,64,1)		(batch,1,1,1,1)	

図11 3D-GANのネット構造

図11の上半部分はネットワークの構造を示しており、エンコーダは2Dの畳み込み式計算、生成器と識別器は3D畳み込み式計算を行う。下半部分は訓練データの変化の流れを示しており、特にエンコーダから生成器に入力時の(batch, 200)から(batch, 1, 1, 1, 200)になるのは2Dから3Dを復元するための重要部分である。

本研究の入力は小画像であり、本来の3D-GANの入力と比べかなり簡単な対象である。このため本研究では3

D-GAN の構造を図 12 の 3D ネットワークのように簡略化して利用する。

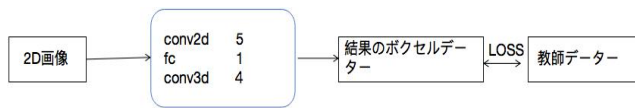


図 12 3D ネットワークの構造

大きな変更点は識別器を削除した点とエンコーダと生成器を一つにまとめた点である。本来は三つの LOSS 関数 (LOSS エンコーダ, LOSS 生成器, LOSS 識別器) が必要であったが、一つ (LOSS :生成データと教師データの比べ) のみになった。このネットはもはや GAN の形ではないので、3D ネットワークと呼ぶことにする。

教師データの作成は人為的に作った入力 of 2D 小画像に対応している 3D ボクセルデータである。図 13 はケーブルの 3D の表示方法を示す。

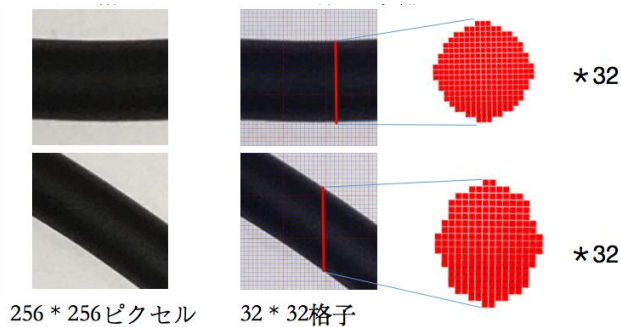


図 13 3D データの表示方法

本研究ではケーブルを円柱体で表現する。この円柱体は複数の断面データを繋げて表現する。このため、水平と垂直方向のケーブルの断面は円形で表現され、斜め方向のケーブルの断面は楕円形で表現される。図 13 の中の赤い線分は断面を表し、右のボクセルは断面を表す。断面の円形、楕円形を 32 枚を繋げてケーブルを表現する。

3D ネットワークの入力画像サイズは 256 * 256、出力のサイズは 32 * 32 * 32 とする。教師データを作る時、まずは、小画像のサイズを正規化して 256 * 256 の画像にする。次に小画像を 32 * 32 の格子で分割し、各格子の中にケーブルがあるかどうかを判断してボクセルデータに充填する。

図 14 に四つのパターンのケーブル小画像とそれに対応する 3D ボクセルデータの形を示す。

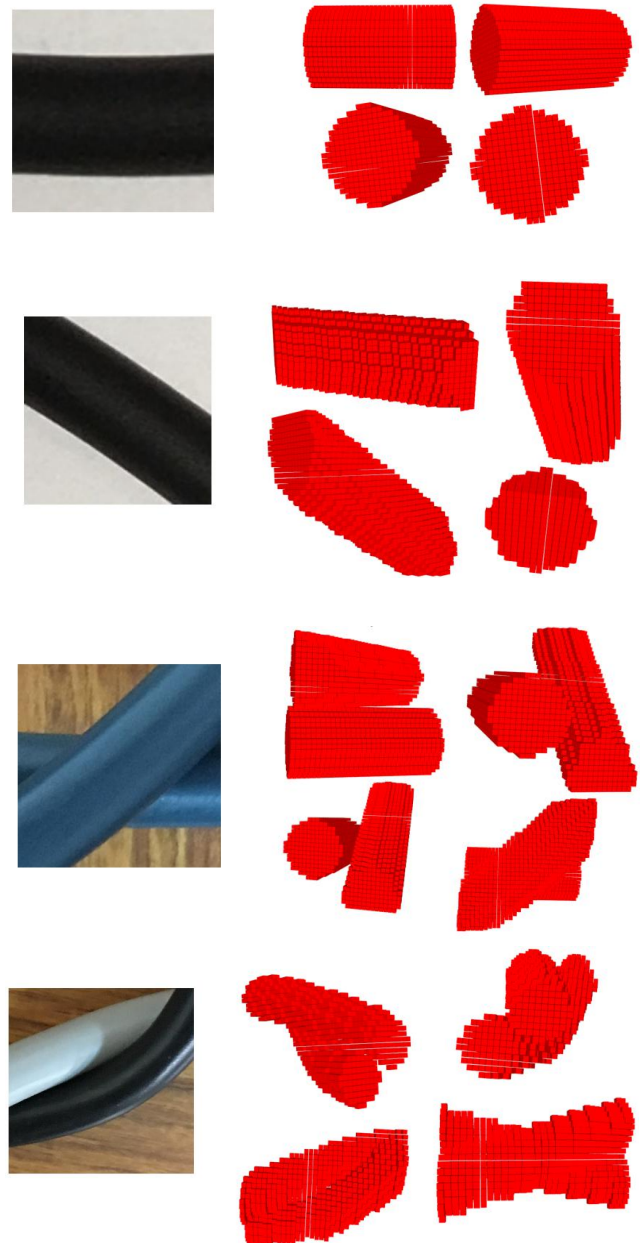


図 14 教師データ

図 14 の左部分は小画像、右の部分は対応した教師データであり、上から順に、水平（垂直）、斜め、交差、捻り四つのパターンを表す。

訓練データを用いてシステムを訓練した後にテストした結果を図 15 に示す。

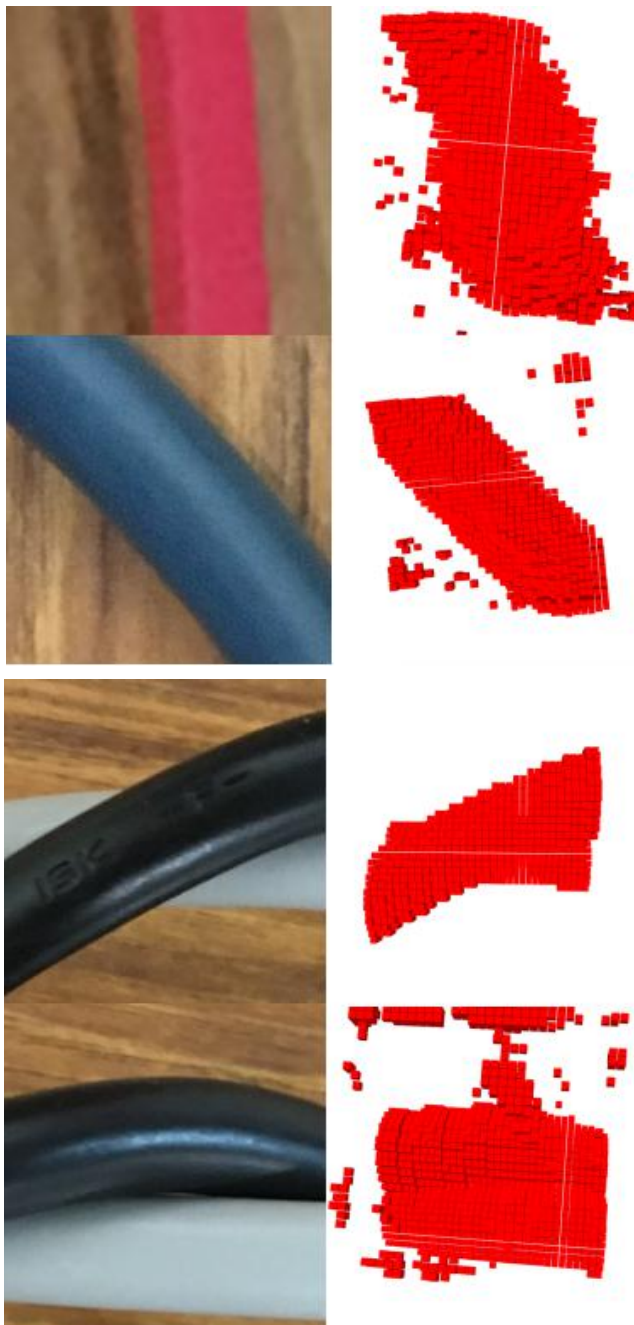


図 15 訓練結果

この結果を見ると孤立したボクセルが存在するがケーブルの概形を表示できている。

4.4 3Dデータの連結

一連の小画像に対する一連の3Dボクセルデータが得られたので、これらのボクセルデータを一つのボクセルデータに連結する。この連結したボクセルデータは原画像中のケーブルを表示している。

原画像中の小画像の位置情報と連結したボクセルデータを統合する。図16に原画像とボクセルデータを表示する。原画像の解像度は4K、ボクセルデータのサイズは252*

189*220である。

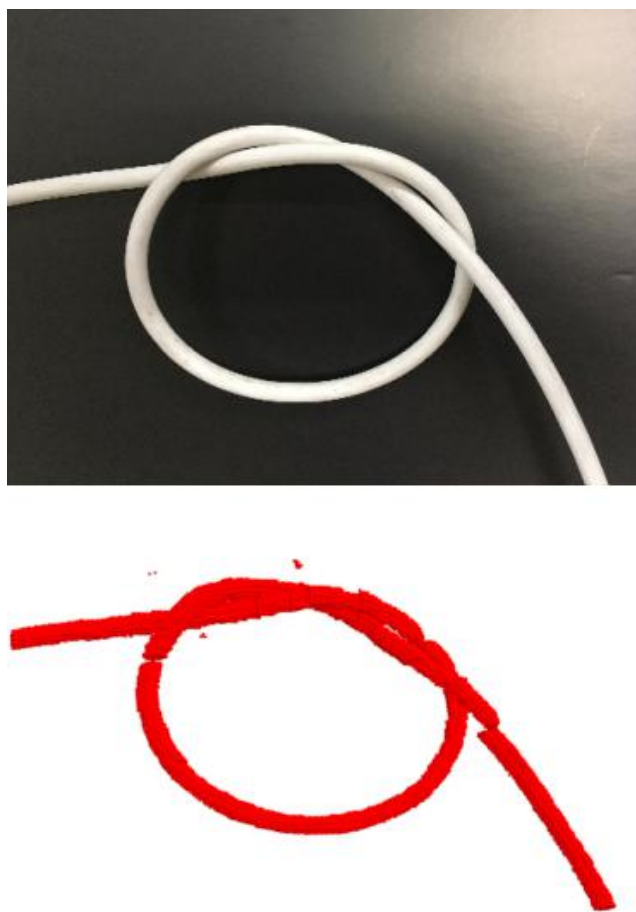


図 16 最後の結果

5. おわりに

本研究はケーブル画像を認識し三次元復元するシステムを提案した。そして、システムを実装し、ある程度三次元復元が実現できた。

本研究ではSegNetと3Dネットワーク二つの部分に深層学習を利用している。深層学習のモデルの性能に影響を与えるのは訓練データの数である。本研究のSegNet部分には500枚の訓練データを用意し、3Dネットワークの部分は250枚の訓練データを用意した。この数は十分であることは言えず、これが原因で過剰適合（overfitting）の状況である可能性が大きい。

また、図17のように3Dデータを連結した時にずれた部分が見られる。

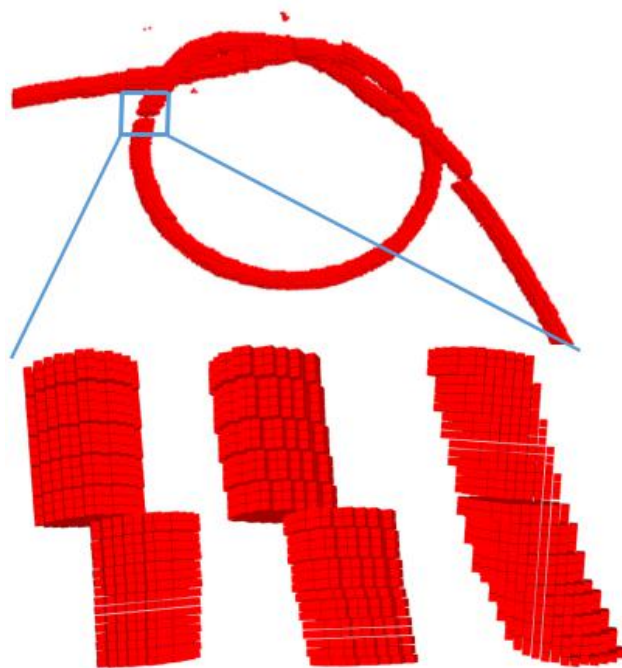


図 17 連結問題の可視化

今後は、訓練データを増やして深層学習部分の性能向上を目指し、連結時のずれた状況の補正を検討し、より精度の高い三次元復元を目指す。

参考文献

- [1] 佐藤 顕治, 岡田 慧, 稲葉 雅幸: 視覚による紐状物体認識に基づくヒューマノイドの生活支援, 2006
- [2] 近成 悠佑, 高林 悠, 浅井 康平, 中田 崇行: ひも認識を用いた3次元ジッとトコースターコースシミュレータ, 2016.
- [3] 范 哲銘, 大橋 健, 林 豊洋: 紐状物体認識システムの研究 2019
- [4] Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla, Senior Member: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, IEEE 2016
- [5] Jiajun Wu, Chengkai Zhang, Tianfan Xue, William t. Freeman, Joshua B. Tenenbaum: Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling, 2017.