

再帰型ニューラルネットワークを用いた GAN による触覚振動生成

我妻正太郎*, 高橋伸 (筑波大学), 黒木詢也, 嵯峨智 (熊本大学)

Vibrotactile Signal Generation by Recurrent Neural Network with Generative Adversarial Network

Shotaro Agatsuma*, Shin Takahashi (University of Tsukuba)

Jyunya Kurogi, Satoshi Saga (Kumamoto University)

Abstract

To create intuitive tactile display, collecting vibrotactile information is important, though, the collection procedure requires manual scanning of textures. Thus, collection of vast information is difficult. However, by employing machine learning technology, there is a possibility to generate further virtual data from existing collected data. In this paper, we made a generation model of vibrotactile signal from the collected acceleration data by using Recurrent Neural Network with GAN. We generated several vibrotactile signals by the model and compared the generated data with the collected one.

キーワード: 機械学習, 振動生成, 加速度データ

(Machine Learning, Vibrotactile signal generation, Acceleration)

1. はじめに

振動を用いた触覚ディスプレイにおいては, 記録された振動をディスプレイの出力として用いる手法が数多くの研究においてリアリティの高い触覚提示を実現している. これらの触覚ディスプレイの出力として用いられる振動は, 人が実際に物体を触察する際の振動を元に作成されている場合が多い. これに伴い, 触覚ディスプレイへの応用のため, 触察時の情報を収集する研究が多く行われている.

例えば, Burka ら⁽¹⁾ は物体表面の視覚情報と振動情報を収集するデバイスを提案した. Abdulali ら⁽²⁾ は, 圧力センサ, 加速度センサを用いて物体表面の情報を収集することにより, 異方性素材における物体表面の触覚のモデリングの手法を提案した. これらの他にも, 様々な触覚情報収集手法が提案されている. しかし, 先行研究では, なぞり方などの収集の条件を限定しているものがほとんどである. 例えば, Stresse ら⁽³⁾ は板状の様々な素材の物体 (テクスチャ) をペン型機器で左から右になぞることでその際の振動などの情報収集を行ったが, なぞり方の条件を変化させた場合, 収集できるデータに違いがある可能性がある. ペン型機器でテクスチャをなぞる動作の中には機器をテクスチャに押し付ける力やなぞる速度, 機器のなぞる角度, なぞるテクスチャの種類など様々な要素が存在している. これら全ての条件を考慮して網羅的に収集を行うのは非現実的である.

しかし, 収集済みのデータを元に収集しきれない条件のデータに類似したデータを生成することができれば, 様々な条件のデータを実際に収集せずとも取得することができる可能性がある. 例えば, 一定の速さで収集したデータを用いて, また別の速度で収集したデータと類似したデータを生成できれば, その分収集するコストを削減することができる.

これらの背景から, 我々は, 触覚情報生成の第一歩とし

て, Generative Adversarial Network (GAN) による信号生成の手法を提案している⁽⁴⁾. GAN は主に画像生成で用いられる機械学習手法であるが, GAN の入力と出力を調整することで, 画像だけでなく, 加速度などの時系列データを高精度に生成できる可能性がある. 本稿では, 以前提案した機械学習モデル⁽⁴⁾を改良し, 再帰型ニューラルネットワーク (RNN) を用いた GAN を構築した. さらに 3 軸加速度データを訓練データとして機械学習し, 実際に生成を行った. 訓練データには, 加速度センサを指に取り付け, テクスチャをなぞった際のデータを用いた. また, 実際に生成されたデータと訓練データの簡単な比較を行い, 考察を行った.

2. GAN の構成について

ここでは, 振動データ生成に用いる GAN について説明する. GAN では, 実際にデータ生成を行う生成器である Generator と, 生成データと本物のデータを分類する Discriminator の 2 種類の機械学習モデルを用いる. Discriminator の分類結果から, 本物のデータに生成データが近づくように Generator を学習することで, 学習が進むごとに Generator がより高精度なデータを生成できるようになる. 本稿では, 振動データ生成には Generator 側と Discriminator 側それぞれ RNN の一種である, Long Short Term Memory (LSTM) を用いたモデルを使用した. また, 得られた加速度データのなかから, ランダムに抽出した連続 1000 点のデータを学習に利用した.

GAN の具体的な構成について説明する. 用いた GAN の構成に関しては, Mogren⁽⁵⁾ が提案した音声生成 GAN を生成するモデルの構成を参考にしている. 構築したモデルの概略図を図 1 に示す. はじめに, Generator 側について説明する. Generator 側において, 入力 は平均 0, 標準偏差 1 の標準正規分布に基づき生成された 1×1000 の乱数行列

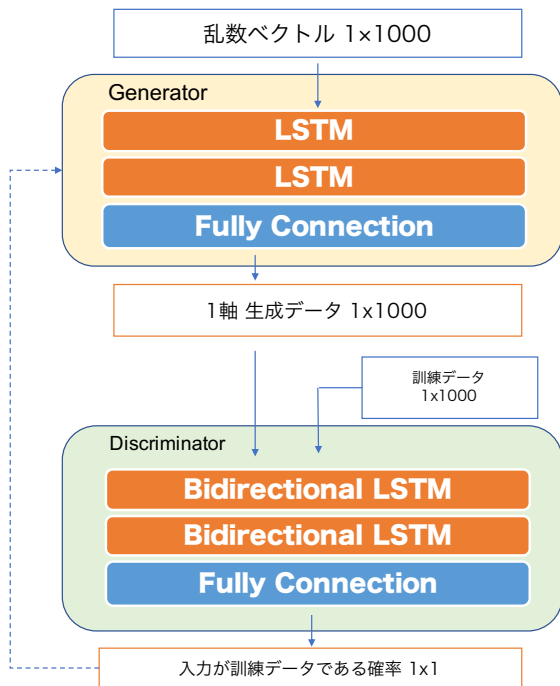


図1 構築した GAN の概略図。Generator は 1 軸振動データを生成する。Discriminator は生成データと訓練データを分類する。この分類結果を元に、Generator は Discriminator に見破られないようデータを生成できるように学習する。

とした。出力は 1×1000 の生成された 1 軸振動データである。なお、今回は 3 軸加速度の x, y, z 軸それぞれ 1 つずつ生成モデルを作成する。Generator は、LSTM 層 2 つと全結合層 1 つによって構成されている。LSTM 層内部の隠れ層の数については、Morgren⁽⁵⁾ と同じく 350 とした。隠れ層の活性化関数には Hyperbolic tangent (tanh) 関数を用いた。LSTM 内部の再帰計算時に使う活性化関数は Hard Sigmoid 関数を用いた。また、全結合層における活性化関数は tanh 関数を用いた。重みの最適化には Adam⁽⁶⁾ を用いた。Adam の学習率は 0.0001 とした。損失関数には、平均二乗誤差を用いた。

次に Discriminator 側については、入力は 1×1000 の振動データとし、出力は、入力データが生成データではなく、訓練データである確率とした。Discriminator 側は、Bidirectional LSTM 層 2 つと全結合層 1 つによって構成されている。LSTM 層内部の隠れ層の数と活性化関数、再帰計算時の活性化関数については Generator と同じ設定を使用した。全結合層での活性化関数は Sigmoid 関数を用いた。重みの最適化と損失関数についても、Generator と同じ設定を使用した。

3. データの生成実験

本稿で構築した GAN を用いて、振動データの生成実験を行う。加速度センサを装着した指でテクスチャをなぞり、その際の 3 軸加速度データを収集した。収集したデータを

訓練データとし、GAN によるデータ生成を行った。

〈3・1〉 訓練データの収集 訓練データの収集手法について示す。指に 3 軸加速度センサ (ADXL-335) をテープで固定し、その指で実際のテクスチャをなぞった際に発生する加速度データを収集する。収集者は一人 (20 代、男性) である。テクスチャを左から右に 5 秒間なぞる試行を 90 回行いデータを収集した。なぞる速さは約 5 cm/s で一定になるようにした。テクスチャをなぞる指の角度はテクスチャ表面から約 45 度とした。収集の様子を図 2 に示す。

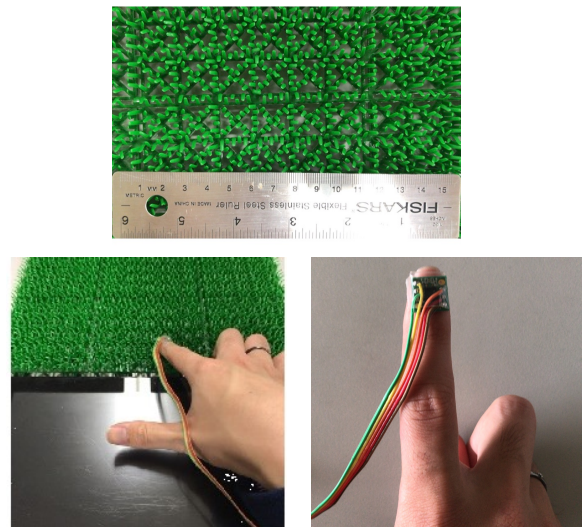


図2 訓練データの収集に用いたテクスチャと収集の様子。

本研究では、3次元のデータを正確に取得するため、加速度センサを接続した Arduino からシリアル通信を利用して加速度情報を PC へ送信し、PC 側では Processing により送信された加速度データを収集した。加速度は通信速度の関係上 1 kHz でサンプリングされる。データ収集に用いたテクスチャは図 2 に示すような人工芝のテクスチャである。本稿では試験的に 1 つのテクスチャのみでのデータ収集と生成を行う。収集の際のテクスチャの種類を増やし、様々なデータを生成できるようにするのは今後の課題の 1 つである。

〈3・2〉 データの生成と結果の考察 収集したデータを訓練データとし、構築した GAN による振動データ生成を行った。今回は全データの中から連続 1000 点のデータを 50 個抜き出して学習させる試行を 100 回繰り返して生成モデルを作成した。このモデルを元に生成されたデータを x, y, z 軸ごとにグラフ化した図を図 3 に示す。なお、Generator 側の出力層において、tanh 関数を活性化関数にすることで、-1 から 1 の範囲での出力データの正規化を行っている。これは Goodfellow ら⁽⁷⁾ の論文を参考に、より訓練データに近い波形のデータを生成するためである。また、訓練データの軸ごとの波形と生成データの軸ごとの波形を図 3 に示す。なお、図 3 ではデータの比較のため、訓練デー

た、生成データともに平均 0、分散 1 になるようにデータを標準化している。

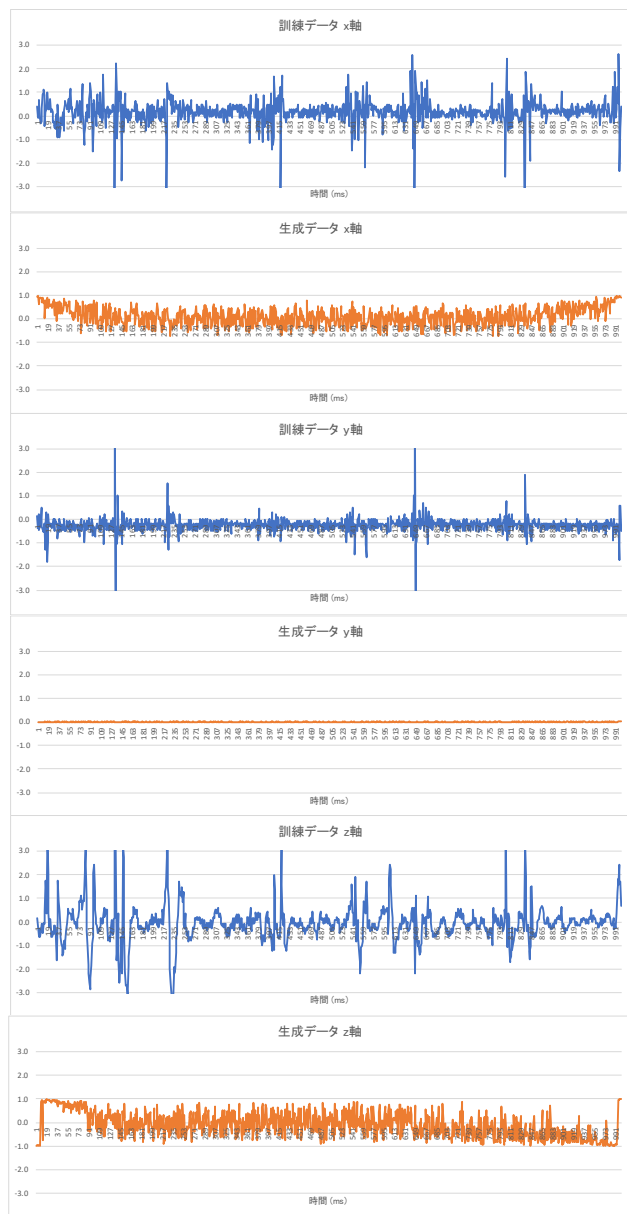


図 3 訓練データと生成データの軸ごとの比較。

図 3 から、まず x 軸について比較すると、値の変化の仕方は似ているものの、訓練データ中の値の変化の大きな箇所については、学習が不十分であることがわかる。また、訓練データの値がグラフ中の 0.4–0.6 の範囲に多く集まっているのに対し、生成データについては、各点が比較的バラバラな値を取ってしまっている。 y 軸については、生成データの値の変化の幅が訓練データに比べて非常に小さくなっていることがわかる。 z 軸についても x 軸と同じく、変化の仕方は似ている箇所があるものの、訓練データ中の

変化の大きな箇所については、学習が不十分であることがわかる。

訓練データの値の変化が大きい箇所について学習が不十分だった原因としては、訓練データ中の値の変化が小さい箇所が、変化が大きい箇所と同程度存在していることが関係していると考えられる。作成したモデルが、値の変化の小さい箇所を過剰に学習してしまい、値の変化の大きい箇所の学習が不十分になってしまった可能性がある。図 3 のグラフを見ても、特に y 軸のデータは値の変化の大きい箇所が少なく、モデルの学習に偏りが出てしまったと考えられる。この問題を解決するためには、データの値の変化が大きい箇所も学習できるよう、モデルの構成やパラメータを調整していく必要がある。今回モデル作成の際参考にした Morgen のモデル⁽⁵⁾ は音声生成を主としたものであるため、モデルの設定を見直すことでより加速度データの処理に特化したモデルを作成できる可能性がある。

4. おわりに

本稿では、収集しきれない条件においての振動データを収集済みの収集データから生成する手法の第一歩として、RNN の一種である LSTM をベースとした GAN による振動データ生成を行った。GAN については、Morgren のモデル⁽⁵⁾ を参考に学習モデルを構築した。収集したデータを訓練データとして 3 軸振動データの生成を行い、軸ごとに生成データと訓練データの比較を行った。結果、値の変化の仕方は似ているものの、訓練データ中の値が急激に変化する箇所については学習が不十分であることがわかった。

今後はより精度の高い生成を行うため、GAN の構成の見直しやパラメータチューニングを行う。データの前処理についても検討する。また、人工芝のテクスチャ以外のテクスチャについてもデータ収集と生成を行い、結果を考察する。さらに Conditional GAN⁽⁸⁾ を採用することで、モデルの入力データにラベル情報を追加できるようにし、どのようなデータを生成するか制御できるように実装する予定である。

参考文献

- (1) Alex Burka, Siyao Hu, Stuart Helgeson, Shweta Krishnan, Yang Gao, Lisa Anne Hendricks, Trevor Darrell, and Katherine J Kuchenbecker. Proton: A visuo-haptic data acquisition system for robotic learning of surface properties. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2016 IEEE International Conference on*, pp. 58–65. IEEE, 2016.
- (2) Arsen Abdulali and Seokhee Jeon. Data-Driven Modeling of Anisotropic Haptic Textures: Data Segmentation and Interpolation. In *Haptics: Perception, Devices, Control, and Applications: 10th International Conference, EuroHaptics 2016, London, UK*, pp. 228–239. Springer International Publishing, 2016.
- (3) Matti Strese, Yannik Boeck, and Eckehard Steinbach. Content-based surface material retrieval. In *World Haptics Conference (WHC)*, pp. 352–357. IEEE, 2017.

- (4) 我妻正太郎, 高橋伸, 嵯峨智. GAN による振動触覚ディスプレイのための信号生成. 第 23 回日本バーチャルリアリティ学会大会, pp. 31A–6. 日本バーチャルリアリティ学会, 2018.
- (5) Olof Mogren. C-RNN-GAN: Continuous recurrent neural networks with adversarial training. *Constructive Machine Learning Workshop (CML) at NIPS 2016 in Barcelona*, 2016.
- (6) Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *the 3rd International Conference for Learning Representations*, 2015.
- (7) Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- (8) Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *CoRR*, Vol. abs/1411.1784, , 2014.