

# 敵対的学習を用いた触覚提示向け振動の自動生成

宇治土公 雄介<sup>1,3,a)</sup> 伴 祐樹<sup>2</sup> 広田 光一<sup>3</sup>

**概要:** 本研究では木や金属などの素材の属性を入力として、触覚提示用の振動を出力する生成モデルを構築した。本モデルを使用することによって、開発者が提示したい表面特性をもつ素材を手元に保持しない場合や、振動データセットに所望の素材に対応する振動が存在しない場合にも、触覚設計者は触覚提示用振動の収集を可能とすることで、振動設計の生産性向上をめざす。実験により、本モデルを用いて生成した振動と実際の素材表面から記録した振動とを区別することができないこと、および生成振動は実際の振動と同程度の現実感を感じさせられることを示した。

## Vibration Generation for Tactile Rendering using Generative Adversarial Network

### 1. はじめに

人は木材や岩などの素材表面をペンでなぞった際に、その素材表面の微細なテクスチャを区別することができる。このとき、ペンと素材の界面で生じる振動の情報を頼りに細かいテクスチャを知覚している。このことから逆に、現実にはその場にはないバーチャルな素材表面のテクスチャ感を、ペンを通じてその振動を提示することにより、ユーザに伝達できるはずである。実際、研究としては粗滑等のテクスチャ知覚について、振動を使って提示可能であることが示されている [3]。

最近では振動アクチュエータが組み込まれたスマートフォン等のタッチ端末の登場がしつつある。現状アクチュエータ付きの製品は限定的であるが、将来的には普及が見込まれ、振動による素材のテクスチャ感提示を表現として取り入れたアプリケーションの登場が期待される。

しかしながら現在のところ、そのようなタッチ端末のアプリケーションにおける振動の活用は、振動の単純なオンオフを切り替えてタッチの際のレスポンスに用いるなどの、入力操作に対する記号的なフィードバックとしての利

用に留まっている。素材のテクスチャ感提示のために振動を有効に活用したアプリケーション開発の問題の1つとして、提示に用いる振動を収集・選定するのに手間がかかることが挙げられる。

振動の収集・選定の一般的な手段としては、記録対象となる素材を手元に複数取り寄せた上で、振動記録ツール [11] を用いて素材をなぞる・叩く等する際の振動を記録することで収集し、その記録振動を提示したときのテクスチャ感を評価し、それが所望の感覚でない場合には別の素材で試す、という一連の作業の反復である。物理的な素材を手元に用意しなければならないために、振動の収集・選定に時間がかかってしまう。

近年では、自前での振動収集の代替手段として、整備されつつある触覚用途向け振動データベース [18] が利用できるが、現状高々 100 種類の素材に対応する振動が提供されている段階であり、現実世界には無数の素材が存在することを踏まえると、アプリケーションでの用途に必要な素材に対応する振動が見つからない場合が多い。

そこで本稿では、様々な素材種類に対応する振動を生成可能なモデルを構築する。具体的には、素材に対応した振動をアプリケーション開発者が直感的に収集・選定可能な、素材の属性や画像を入力とする振動生成モデルを構築することをめざす。様々な種類の素材の見た目をしたコンテンツや UI 部品 (ボタンやスライダー等) に触れる際のフィードバックとして振動を提示するために、アプリケーション開発者がこのモデルを使って生成した振動を活用すること

<sup>1</sup> 株式会社日立製作所  
Hitachi, Ltd., Yokohama, Japan

<sup>2</sup> 東京大学  
The University of Tokyo

<sup>3</sup> 電気通信大学  
The University of Electro-Communications

a) yusuke.ujitoko.uz@hitachi.com

を想定している。例えば、スマートフォン画面上でのボタンアイコンとして金属や木材などの見た目をしたものがあるが、これらをタップする際に素材に対応した振動を提示できれば、ボタンにより現実感をもたせることができる。

近年、高次元のデータ分布からデータをサンプルできる深層生成モデルが注目を集めている。特に、Generative Adversarial Networks(本稿では GAN と呼称) という敵対的学習フレームワーク [4] は画像生成や画像変換などのタスクにおいて従来の機械学習モデルを上回る成果を残している。GAN は属性 [12], [14] など条件付けて、画像を生成するように学習することができる。すなわち属性を入力して、その入力属性に適した画像を生成することができる。このときに属性を連続的に変化させて入力すると、生成画像も離散的でなく連続的に変化することが知られている。以上を踏まえると、GAN のフレームワークを用いれば、属性や画像を条件として所望の振動を生成することができるはずである。その際には、学習に用いた属性や画像だけでなく、それらの属性や画像から合成した属性や画像を入力した場合にもねらった振動を生成できる見込みがあると言える。しかしこれらの実績にも関わらず、GAN を用いて時系列データを生成する試みは多くない。特に、GAN による振動の生成はこれまで実現されていない。GAN はフレームワークとして学習が不安定で難しいことが知られている。それを緩和するための学習テクニックが多く開発されてきたが、それらはこれまで GAN で学習対象とされてきたデータ構造が主に画像等の 2 次元構造に依存しているため、1 次元データである時系列データの学習には直接活かすことができないため学習が著しく難化してしまう。これが、GAN による振動生成がこれまで実現されてこなかった理由の 1 つと考えられる。

そこで本研究では GAN を活用して時間-周波数領域の 2 次元パターンを画像として出力するモデルを構築し、生成パターンに対して変換処理を施して時系列の振動データを取得。具体的なモデルへの入力としては、アプリケーション開発者が直感的な触覚デザインを行えるよう、素材の属性とした。開発者が所望の素材の属性値に見当がついている場合には、素材の属性をモデルに入力して生成振動を取得してもらうことを想定している。モデルをアプリケーション開発時の触覚デザインのこのようなケースにおいて活用することで、触覚提示に用いる振動の収集・選定のコストを低減し、触覚デザインを促進することをねらう。本研究の最終的なねらいは、学習に用いた素材以外の素材の属性をモデルへ入力する場合にも、それらに即した振動をモデルが生成することである。ただし本稿ではその第一歩として、学習に用いた素材属性に対して、適切な振動データを生成できるかをユーザスタディにより検証した。本稿の貢献は次の 2 点である。

- 素材に対応した触覚提示用振動の生成を問題として提

起し、GAN に基づく学習方法で当問題を解決するためのモデルを構築したこと

- 構築したモデルが、学習に用いた素材属性に対して、適切な振動データを生成できることを実験により検証したこと

## 2. 関連研究

### 2.1 従来の振動収集・選定方法

振動の収集・選定の一般的手段として、振動記録ツール [11] を用いて素材から直接振動を記録する方法や、触覚提示用の振動データベースから取得する方法がある。しかしながら、素材を手元に用意する労力がかかることや、データベースが整備途上であることから、現状ではそれらの方法を利用するのは困難である。

振動が必要となった際に都度収集するという上記のアプローチではなく、振動をデータドリブンにモデル化しておくというアプローチもある。例えば、従来研究ではペンで素材表面をなぞる際の、ペンの速度やペンにかかる力に基づいた振動のモデル化を試みている [3][17]。このような従来研究のモデルは、ペイントアプリなど、ペンの速度や印加される力に大きな変動があるインタラクションに適している。しかし、このモデルだと単一の素材に対応した振動しか出力できないという制約があり、様々な素材に対応する振動を収集・選定するという用途にはそのまま適用できない。もし仮に複数の種類の素材に対応する振動を生成したい場合には、その分のモデルを用意する必要がある。しかしながら実世界に無数に存在する素材の種類すべてに対して、専用の生成モデルを用意するのは現実的には不可能である。一方で、我々は素材の種類に応じて適切な振動を生成するような単一のモデルを構築する。アプリケーション開発者が直感的な触覚デザインを容易に行えるように、モデルの入力として素材の属性をとるモデルを構築する。

### 2.2 GAN による学習フレームワーク

近年、画像などの高次元のデータ分布からサンプル可能な生成モデルが注目を集めている。画像変換 [6]、画像の高解像度化 [10] 等、その用途も幅広い。近年注目を集めている代表的な生成モデルとしては、Autoregressive Model(AR モデル)[19], [20]、Variational Auto Encoder(VAE)[8]、Generative Adversarial Network(GAN)[4] の 3 つがある。このうち、AR モデルは推論に長時間かかり、VAE は鮮明なデータを生成できないという、触覚デザイン用途としては回避できない欠点が存在するため、本研究では GAN をベースにして生成モデルを構築する。

具体的には、GAN の学習フレームワークを用いて時間-周波数表現のパターンを画像として生成し、振動に変換して最終的にユーザに提示する。GAN は Goodfellow ら [4] によって提案された学習フレームワークであり、内部の構

成要素のネットワークとして Generator と Discriminator を持つ。Generator と Discriminator の学習は互いに競い合うように進められるという特徴がある。訓練データの分布が与えられると、Generator は乱数シード  $z$  をもとにデータを生成する。この生成データが訓練データを模倣するよう Generator は学習していく。一方で Discriminator は訓練データと生成データを区別できるように学習していく。GAN は学習が終了した段階で、Generator が乱数シード  $z$  に基づいて生成したデータは、Discriminator が区別不可能な程度に訓練データを模倣しているという仕掛けとなっている。Generator と Discriminator がともに誤差を最小化するように学習しようとするが、お互いが相手の誤差を大きくすることになるため、ナッシュ均衡を安定的に見つけるのが難しい。それを緩和するための学習テクニックが多く開発されてきたが、それらはこれまで GAN で学習対象とされてきたデータ構造が主に画像等の 2 次元構造に依存しているため、1 次元データである時系列データの学習には直接活かすことが困難なものが多い。そのため本研究では GAN の生成対象として、1 次元の振動の時系列データではなく、時間-周波数表現の 2 次元のデータ構造を対象とすることで、既存の GAN の学習テクニックを活かすねらいがある。

GAN の生成データを制御する条件付き GAN[12] として幾つかの変種が考案されている。その中でも AC-GAN[13] が条件付き GAN の代表例である。AC-GAN では、属性ラベル  $c$  と乱数シード  $z$  に基づいて generator がデータを生成し、discriminator が生成されたデータの属性ラベル  $c$  を区別するタスクも解くというフレームワークとなっている。学習後の AC-GAN は、属性ラベル  $c$  を指定するとその属性に対応したデータを生成することができる。

### 2.3 GAN による感覚情報のモデル化

条件付き GAN を使ってクロスモーダルなデータの変換をめざす研究も行われている。テキストから画像への変換は [14] によって行われている。この研究では、鳥や花を表現するテキストを入力すると、それに応じた画像が生成される様子が示されている。楽器演奏時の音響と演奏画像を相互に変換するモデルも提案されている [1]。我々の研究と似たアプローチをとっており、音響に変換可能なスペクトログラムを GAN で生成する試みを行っているが、生成されたスペクトログラムは音響に変換すると粗く、鑑賞可能な音響の質には至っていない。その理由の 1 つとして、音響向けにカバーすべき 22.05kHz の周波数帯域までカバーした画像を生成することは困難であることが挙げられる。単純に考えると 22.05k pixel のサイズをもつ画像を生成しなければならないが、GAN の最先端モデルであっても 1k pixel オーダの画像の生成が限界であり、音響向けデータを GAN で生成するのはまだまだ難しいといえる。一方で、

触覚受容器を刺激するのに必要な振動の周波数帯域は音響刺激よりも小さくてよく、GAN で十分カバー可能な範囲にあるといえる。

## 3. 生成モデルの構築

### 3.1 モデルの全体像

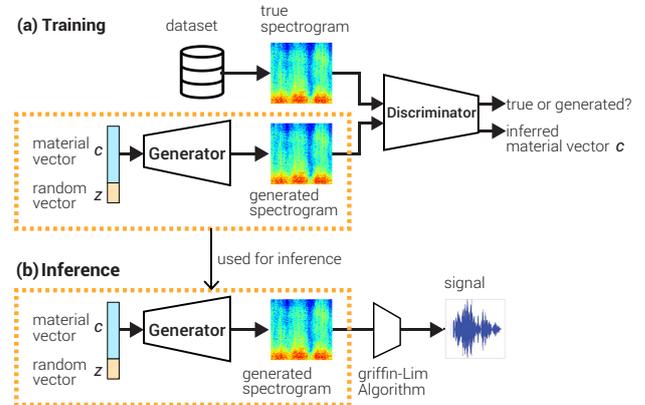


図 1 モデルのアーキテクチャ。(a) 学習時のモデル構成。(b) 推論時のモデル構成。

Fig. 1 Architecture of the model.

我々の提案する振動生成モデルのアーキテクチャの全体像を図 1 に示す。学習時のモデルは Generator と Discriminator の 2 つの構成要素からなる。Generator と Discriminator は GAN の学習フレームワークで同時に学習させる。学習の詳細は次節で述べる。学習時と反して、推論時には Discriminator は取り除き、Generator のみ利用する。

Generator の出力は振動である。Generator に対する開発者による入力素材の属性を示す素材ベクトル  $c$  である。素材の属性を示す素材ベクトルとは例えば、 $(c_{\text{木材 A}}, c_{\text{金属 B}}, c_{\text{布 C}}, \dots) = (0.3, 0, 0.1, \dots)$  として指定する。上記の例では素材ベクトルの基底を木材、金属、布等と設定したが、ある大まかな属性の素材 (例: 木材) にスコープを限定して、その中で細かく基底 (木材 A, 木材 B, 木材 C 等) を設定してもよい。また、素材そのものをベクトルの基底に設定せずに、素材を別の観点で表現した基底としてよい。例えば粗滑や硬軟等のテクスチャの知覚次元で表現した基底としてもよい。本稿で構築したモデルでは、使用したデータセット [18] の素材のクラスを素材ベクトルの基底としてそのまま利用した。Generator は素材ベクトル  $c$  と乱数ベクトル  $z$  を結合し、その結合ベクトルをスペクトログラムに変換する。生成されたスペクトログラムは Griffin-Lim アルゴリズムで位相復元し加速度へと変換される。そしてその加速度が振動アクチュエータを介してユーザの皮膚に伝達される。このモデル全体を使うことで、アプリケーション開発者は素材の属性を入力として、振動情報を得ることができる。

本研究では加速度を振動データとして用いている。提案



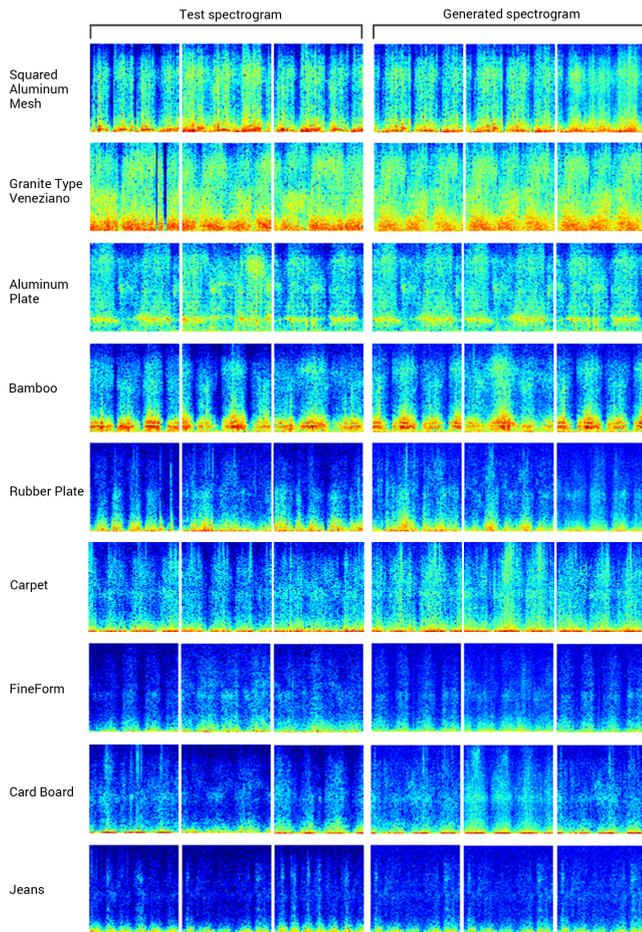


図 4 各クラスのテストデータセットのスペクトログラムと generator の生成したスペクトログラム  
Fig. 4 The spectrogram for each class label in test dataset and the one generated by generator.

「リアリティ」と定義し呼称する。

一般的に GAN の生成物を定量的に評価することは難しいとされている。GAN の定量評価は通常は Inception score[15] やその改良版の Fréchet Inception Distance (FID) [5] が用いられるが、今回のケースにはこれらは適用できない。なぜなら Inception score や FID を適用できるのは、CIFAR-10 などの標準データセットを用いた場合に限られ、今回のようなカスタムデータセットに対しては適切な評価を得ることができないためである。代わりに、t-SNE[21] を用いて生成された 16384 次元 (=128x128) のデータを 2 次元に圧縮し、その圧縮空間上でデータの分布を観察する方法をとる。t-SNE を用いて GAN による生成データのクラスターを観察することは、カスタムデータセットを GAN に適用する際に標準的に用いられる方法である [7], [16]。

t-SNE による可視化結果を図 5 に示す。可視化結果を見ると、各クラスごとにテストデータも生成データもまとまっていることがわかる。

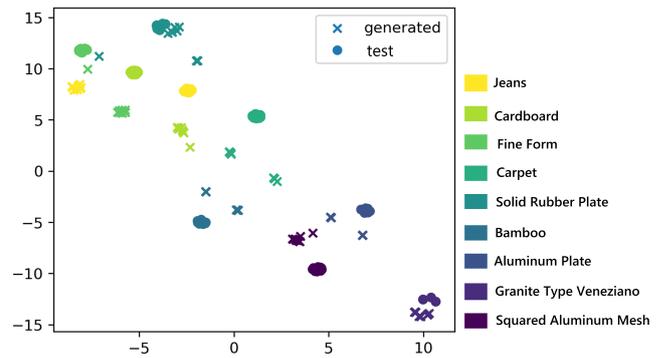


図 5 各クラスのテストデータセットのスペクトログラムと generator の生成したスペクトログラムの t-SNE による可視化  
Fig. 5 t-SNE visualization of test dataset and the one generated by generator.

#### 4. 実験:学習に用いた素材属性入力時の生成振動の評価

本実験では、我々の提案モデルの使い方を想定し、Generator を用いて素材ベクトルから生成した振動をユーザを対象としてテストした。具体的には、ユーザがテストデータと生成データを触覚として区別できるかどうかと、生成データに対してユーザがリアリティを感じるかを評価項目とした。ただし入力とする素材は学習に用いたものに限定した。

実験参加者は 22 歳から 25 歳までの男女 10 名 (男性:8 名, 女性:2 名) で、全参加者は右利きであった。なお、本論文における全実験は東京大学倫理委員会による承認 (承認番号: KE17-63) を受けて実施されたものである。

##### 4.1 実験システム

本実験での参加者のタスクは、タブレット端末の表示画面上で、ペン型デバイスから振動提示を受けつつ表面をなぞることである。実験システムは、タブレット端末 (Apple Inc., iPad Pro 9.7inch), アンプ (Lepai Inc., LP-2020A+), および振動子付きペン型デバイスの 3 つから構成される (図 6)。

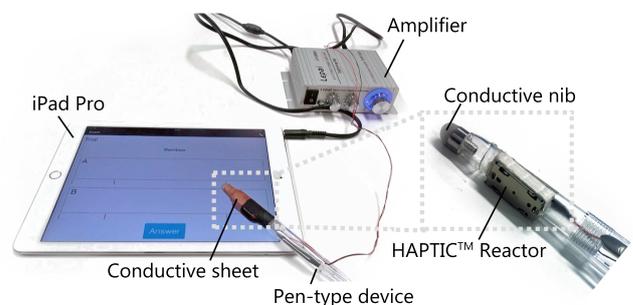


図 6 実験システム  
Fig. 6 Experimental system.

ペン型デバイスは重量約 20g, 長さ約 140mm, 把持部の

直径は約 10mm である。ペンの先端には導電性材料を装着した。先端と把持部を接続するように、導電性シートを巻いて静電容量式タッチスクリーンと反応するようにした。ペン型デバイスの把持部（ペンの先端から 2cm の位置）に振動子（ALPS 社、HAPTIC Reactor）を組み込んだ。振動子は小さく（35.0mm × 5.0mm × 7.5mm）、軽い（約 5g）ため、参加者がペンを動かすのを阻害するものではない。

参加者がタブレット上にペンを接触させてなぞると、振動用信号がタブレットのイヤホンジャックから出力され、アンプを経由して振動として参加者の指に伝わる。

#### 4.2 タスク設計

本実験は参加者内計画で行った。ある素材クラスのクラス名が画面上部に表示されておりそれを見ながら（図 7 左）、参加者は、テストデータ内の振動またはモデルが生成した振動を提示されつつ、ペン型デバイスを画面上の所定の経路上を動かした。次に同じくテスト用の振動もしくは生成振動を受けながら、所定の経路上でペン型デバイスを動かした。2つの経路上で異なる振動刺激を受けた後、参加者はどちらの振動刺激がモデルで生成されたものであるかを区別し、さらに画面上部に表示されている素材クラス（クラス名または素材画像）に属する振動としてのリアリティについても評価した。参加者に提示する振動は図 3 に示した 9 つの素材のうちの 1 つに対応する素材ベクトルを Generator に入力することで生成したものと、同一の素材に対応するテストデータとしてデータセットに備わる振動を用いた。Generator 実験で視覚提示する素材の情報はその素材のクラス名である。なお、参加者からするとテストデータおよび生成データのいずれも初見である。参加者は視覚提示される素材のクラス名または素材画像から想起されるものに基づいて、実際に提示された振動に対して判断を下す。

参加者の 1 試行での実験手順を述べる。以下の手順を繰り返し行っていく。参加者は、ペンを左から右へ約 100mm の距離を一定速度で動かす。移動速度を一定にするために画面上に延伸するバーが表示されるので、このバーの延伸スピードに合わせて参加者はペンを動かす。バーの延伸スピードは 1.6 秒で約 100mm を移動し終える設定となっている。参加者の指に対して振動を直接提示するために、参加者は振動子が組み込まれたペンの把持部を持つよう指示されている。画面上の 2 つの経路上（図 7 左）で順にペンを動かした後、どちらの振動が生成されたものと感じたかを画面上のボタンを押して答える。加えて、2 つの振動刺激のリアリティについてもそれぞれを VAS（Visual Analogue Scale）評定法に従って評価する（図 7 右）。100mm の長さの線分が画面上に表示されており、この線分のどこかを参加者は感じたリアリティの程度を評価するためチェックをつける。線分の左端は”全くリアリティを感じない”ことを

意味し、線分の右端は”とてもリアリティを感じる”ことを意味している。テストデータの振動と生成振動の提示順序は 1 試行の中でランダムになっている。なお参加者はノイズキャンセリングヘッドフォンを装着し、音を遮断した状態で実験を行った。

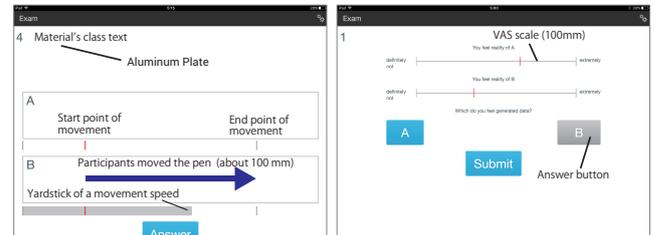


図 7 実験画面。左：画面上側に素材のクラス名が表示され、画面下側に参加者がなぞる 2 つの経路が存在する。右：VAS スケールを用いてリアリティを回答し、どちらが生成振動であったと感じたかをボタンで回答する

Fig. 7 Experimental window. Left: The material class name or is visualized on upper side of the screen. Right: Users evaluated the realism of the vibration and identify which is the generated one by pushing buttons.

9 種類の素材に対応する生成振動とテストデータの振動が本実験では用意された。生成振動は毎試行、モデルから新たに生成して用意した。テストデータの振動は、該当する 9 つの素材としてデータセットに予め用意されているテストデータの振動の中からランダムに抽出して使用した。参加者ごとに各素材 10 試行ずつ行った。1 人の参加者につき合計で 180 試行行ったこととなる。提示する素材種類の順序はカウンターバランスするよう割当を行った。

#### 4.3 結果

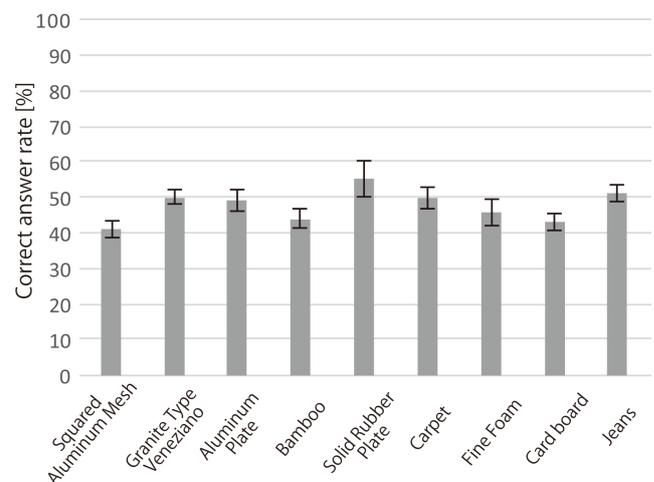


図 8 提示振動が生成されたものか、データセットから提供されたものか正しく回答した確率

Fig. 8 Percentage of identifying which stimuli were generated one out of two.

図 8 に、参加者に提示された振動が生成されたものか、データセットから提供されたテストデータかを参加者が正しく回答した確率（正答率）を素材種類ごとに示す。横軸が素材種類で、縦軸が素材ごとの正答率を示している。正答率が 50%に近い場合、参加者がテストデータと生成データを区別できなかったことを示している。全素材に対する正答率の平均は、 $47.7 \pm 1.49\%$ であった。各実験の各素材に関して、正答率が 50%から離れているかを検証するためカイ二乗適合度検定を行ったところ、いずれの素材も正答率が有意に 50%から離れていなかった。

一方、図 9 に、提示振動に対して感じるリアリティを素材種類ごとに示す。テストデータの振動に対するリアリティの評価値の平均は  $72.9 \pm 1.49$  であり、生成振動に対する評価値の平均は  $73.1 \pm 2.93$  であった。Student の t 検定をテストデータと生成データ間の平均値に対して行ったところ、Generator 実験における素材 Bamboo ( $p = 0.025$ ) に関して有意差があった。テスト振動全体と生成振動全体の間では有意差は見られなかった。

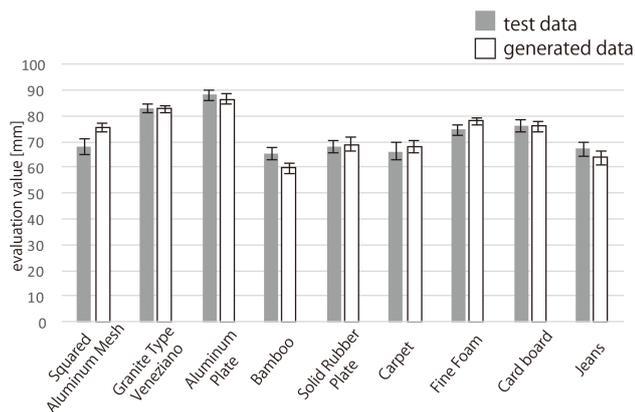


図 9 提示振動に対して感じたリアリティ。

Fig. 9 Felt realism for test signals and generated ones.

#### 4.4 考察

提示振動が生成されたものか、テストデータの振動か正しく回答した確率（正答率）を確認すると、殆どの条件においてほぼ 50%であり、参加者はテストデータの振動と生成振動を区別できていなかった。実験後の自由記述では、全参加者がテストデータの振動と生成振動に違いがなかったと回答していた。したがって、我々の提案モデルは、素材の属性や画像をもとに実際の振動を高い水準で模擬できたと言える。

リアリティに関しては、殆どの条件において 60 を超えており、かつ生成振動全体の平均とテストデータ全体の平均間で違いがなかった。このことから、提案モデルは概してテストデータ相当のリアリティのある振動を生成できていることがわかる。

実験後に行った実験参加者による自由記述にて、特定のクラスにおいて、クラス名から素材表面を想像するのが困難である旨の記述も見られた。これを踏まえて、今後は Generator に入力するベクトルを素材そのものの表現ではなく、人が解釈しやすい軸で表現する必要があることがわかった。例えば、テクスチャの知覚の軸である硬軟や粗滑などや、もしくは触感のオノマトペなどへ再設計するなどを検討する必要がある。

今回の実験では Generator への入力としては学習に用いた単一素材を示す素材ベクトルが入力した。しかしこのような、単一素材を示す素材ベクトルを入力する場合というのは、振動触感のデザインにおいては限定的なケースである。振動触感デザイン時の発展的なユースケースとして、より柔軟な素材ベクトルが Generator に入力される場合が考えられる。具体的には複数素材の中間にあたる素材の素材属性を入力して振動を生成することが想定される。このような場合の生成振動や Generator の応答性について今後検証を進めていく必要がある。

## 5. 本モデルの効果と制約

### 5.1 効果

本稿では、素材ベクトルを入力として振動を生成するモデルの全体像を提案した。加えて、モデルの学習パラメータやデータ処理方法等のモデル構築の際の詳細を定義し、それらを 2 つの実験を通じて検証した。実験 1 により、学習に用いた素材クラスに関して、高い水準でテストデータを模擬でき、リアリティのある振動を生成できる可能性が示唆された。この結果は、触覚デザインの際に幅広い範囲の素材に対して、本モデルが適用できることの可能性を示唆するものである。素材種類を増加させる場合など問題設定が複雑化した際にも、提案モデルの全体構成は活用できる。ただし今回用いたモデルの学習パラメータやデータ処理方法等が適用できるかについては別途検討を進める。

本稿の範囲では評価できていないが、本モデルは条件付き GAN の構成を採用しているため、学習に用いた属性に基づき合成した属性に対しても、適切な振動を生成できる可能性がある。例えば GAN の先行研究 [7] では、人の顔画像の生成問題において、Generator に入力する髪型に関する属性を「黒髪」「女性」から「金髪」「男性」に変化させて入力すると、出力画像としてもそのように変化することが知られている。加えて、入力属性の変化を連続的に行った際に、出力画像も離散的でなく連続的に変化することが示されている。このような連続的な出力振動の変化が本モデルにおいても可能であるとすると、現状高々 100 種類の素材に対応しているデータセットよりも、より多くの素材種類をカバーした振動を生成できる可能性がある。ただし、学習に用いた素材をもとに入力属性を合成するため、学習に用いた素材を基底とする素材空間に内包された素材

に対してのみ適合した振動を生成できるようになる見込みである。

## 5.2 制約

最後に提案モデルの制約としてペンの入力方法への対応である。提案モデルはペンの運動状態を入力としてとらないため、提案モデル単体ではペン入力に対応した振動を生成できない点が制約となる。しかし、ペンの入力は限定的であると仮定してもよいという見方もできる。実際 Apple はアプリケーション開発者に対しては、スワイプやタップなど 8 種類のシンプルな入力ジェスチャ以外を制限するガイドラインを設けている。これらのジェスチャは短時間で入力が終わるため、ペン速度やペンに印加される力は静的とみなすことができると考えられ、提案モデルを各ジェスチャに対して構築すれば十分であると考えられる。

一方、ペイントアプリなどペンの速度や力に変動があるインタラクションに対して、本モデルを用いたい場合がある。そのような場合においても、Culbertson et al.[2] により、振動がペンの速度に応じて変化しなければ、ユーザが振動に感じるリアリティは低減するが、ペンにかかる力には振動を対応させる必要がない可能性が指摘されている。そのため、本モデルで生成した振動をベースにペンの速度に応じて周波数変調を施すだけで十分である可能性がある。

## 6. 結論

本研究では、素材の属性を入力として振動を出力するモデルを GAN により構築した。本モデルを用いることで、手元に素材そのものを持たない場合や、データセットに所望の素材の振動がない場合にも振動を収集可能である。実験によって、殆どの条件においてユーザがモデルが生成した振動を真の振動から区別できないことや生成振動に対して真の振動と同程度のリアリティを感じることを示した。今後は学習に用いてない複数素材の内挿ベクトルを入力した場合の生成振動についてより詳しく評価していく。

## 参考文献

- [1] Chen, L., Srivastava, S., Duan, Z. and Xu, C.: Deep Cross-Modal Audio-Visual Generation, *In Proceedings of the Thematic Workshops of ACM Multimedia 2017*, pp. 349–357 (2017).
- [2] Culbertson, H. and Kuchenbecker, K. J.: Should haptic texture vibrations respond to user force and speed?, *In Proceedings of IEEE World Haptics Conference*, pp. 106–112 (online), DOI: 10.1109/WHC.2015.7177699 (2015).
- [3] Culbertson, H., Unwin, J. and Kuchenbecker, K. J.: Modeling and rendering realistic textures from unconstrained tool-surface interactions, *IEEE Transactions on Haptics*, Vol. 7, No. 3, pp. 381–393 (2014).
- [4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative Adversarial Nets, *In Proceedings of the*

- Advances in Neural Information Processing Systems 27*, pp. 2672–2680 (2014).
- [5] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Klambauer, G. and Hochreiter, S.: GANs Trained by a Two Time-Scale Update Rule Converge to a Nash Equilibrium, *CoRR* (2017).
- [6] Isola, P., Zhu, J. Y., Zhou, T. and Efros, A. A.: Image-to-Image Translation with Conditional Adversarial Networks, *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976 (2017).
- [7] Jin, Y., Zhang, J., Li, M., Tian, Y., Zhu, H. and Fang, Z.: Towards the Automatic Anime Characters Creation with Generative Adversarial Networks, *In Proceedings of CoRR* (2017).
- [8] Kingma, D. P. and Welling, M.: Auto-Encoding Variational Bayes, *CoRR* (2013).
- [9] Kodali, N., Abernethy, J., Hays, J. and Kira, Z.: On Convergence and Stability of GANs, *In Proceedings of CoRR* (2017).
- [10] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and Shi, W.: Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 105–114 (2017).
- [11] Minamizawa, K., Kakehi, Y., Nakatani, M., Mihara, S. and Tachi, S.: TECHTILE toolkit, p. 1 (2012).
- [12] Mirza, M. and Osindero, S.: Conditional Generative Adversarial Nets, *In Proceedings of CoRR*, Vol. abs/1411.1 (2014).
- [13] Odena, A., Olah, C. and Shlens, J.: Conditional Image Synthesis with Auxiliary Classifier GANs, *In Proceedings of the 34th International Conference on Machine Learning*, Vol. 70, pp. 2642–2651 (2017).
- [14] Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B. and Lee, H.: Generative Adversarial Text to Image Synthesis, *In Proceedings of International Conference on Machine Learning*, pp. 1060–1069 (2016).
- [15] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. and Chen, X.: Improved Techniques for Training GANs, *Proceedings of Advances in Neural Information Processing Systems*, pp. 2234–2242 (2016).
- [16] Sankaranarayanan, S., Balaji, Y., Castillo, C. D. and Chellappa, R.: Generate To Adapt: Aligning Domains using Generative Adversarial Networks, *CoRR* (2017).
- [17] Shin, S., Osgouei, R. H., Kim, K. D. and Choi, S.: Data-driven modeling of isotropic haptic textures using frequency-decomposed neural networks, *In Proceedings of IEEE World Haptics Conference, WHC 2015*, pp. 131–138 (2015).
- [18] Strese, M., Schuwerk, C., Iepure, A. and Steinbach, E.: Multimodal Feature-Based Surface Material Classification, *IEEE Transactions on Haptics*, Vol. 10, No. 2, pp. 226–239 (2017).
- [19] van den Oord, A., Kalchbrenner, N. and Kavukcuoglu, K.: Pixel Recurrent Neural Networks, *CoRR* (2016).
- [20] van den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. and Kavukcuoglu, K.: WaveNet: A Generative Model for Raw Audio, *In Proceedings of CoRR* (2016).
- [21] van der Maaten, L. and Hinton, G.: Visualizing Data using t-SNE, *Journal of Machine Learning Research*, Vol. 9, pp. 2579–2605 (2008).