

音楽解析用回路ニューラルネットワークの

活性パルス強度分布とメロディ人気度との関連

Correlations between melody popularity and active pulse strength distributions of circuit neural network for use in music analysis

水村駿太[†] 田村 大河[†] 竹内 幹[†]
Shunta Mizumura Taiga Tamura Kan Takeuchi

1. はじめに

音楽には多くの人に受け入れられる曲、いわゆるヒット曲がある。メロディ、歌詞、ビート、アレンジなど多くの要素が関与しているが、たとえば、ビートルズの Let It Be や Hey Jude など世界で幅広く支持される曲は、そのメロディラインが人間の感性を心地よく刺激する。本研究は、ヒット曲に共通するメロディの特徴をニューラルネットワークで理解しようとする試みである。人間の感性をできるだけ直接的に計算機上に再現することを目指し、デジタル AI 解析ではなく、生体系を模した電気回路ニューラルネットワークを構築して解析する[1]。実際の神経網をそのまま電子回路として再現するアプローチはニューロモーフィックとして知られている[2]。本研究では、デジタル AI や多くニューロモーフィックで採用されてきたパーセプトロンを多段連結する回路構成ではなく、時系列を作り込んだ別の回路構成を基本とする。メロディラインは時系列そのものであるし、人間の感性は予測(期待、不安)と実際の結果との相互作用が基本であると考えからである。本研究では、この回路型ニューラルネットワークに、ビートルズの 200 を超える楽曲のうち、音楽配信サイト(Apple Music)の

ダウンロード上位曲及び下位曲それぞれ 20 曲のメロディを入力した。その結果、内部ニューロンの活性パルス強度分布に有意な差を見出したので報告する。

2. シミュレーション方法

2.1 電気回路ニューラルネットワーク

図 1 に回路ニューラルネットワークの構成を示す。図 1 の右下に示す通り、 V_0, V_1 を入力とし、 V_{OUT} を出力とする単位ニューロン回路が約 5000 個連結している。電気回路シミュレータ LTspice を使い、ピアノの鍵盤を叩くように、音程 G3 から A5 の範囲の電圧パルスを曲のメロディに沿って次々に入力する。正確には、楽器が発生する倍音[3]を模した D5 から C#8 と命名された倍音端子のうち、5 個(基本波から 5 倍波相当)の端子から回路ニューラルネットワークに電圧パルスを入力する。入力電圧波形の作成は、music21[4]を使った Python コードで行った。曲の midi データ[5,6]からメロディのみを抽出し、LTspice の PWL 電圧波形に変換した。

図 1 では、F6⇒D6⇒A6 の倍音端子から次々と電圧パルスが入力された場合に、F6D6A6 の 3 連音活性パルスが発

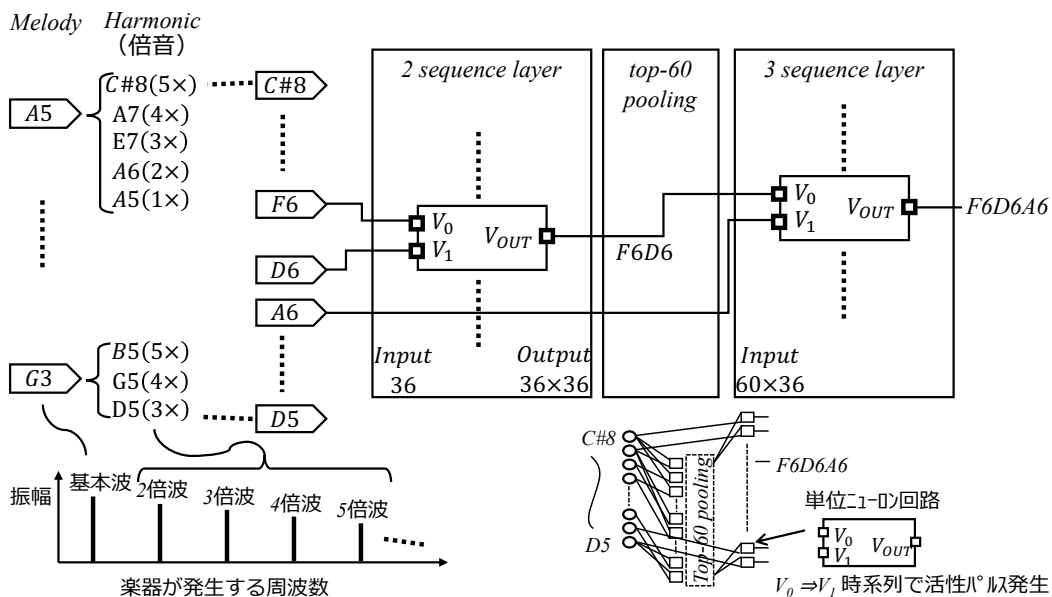


図 1 電気回路ニューラルネットワーク構成

[†] 明星大学理工学部総合理工学科
Faculty of Science and Engineering, Meisei University

生ずる様子を示している。2 sequence layer のうち、 V_0, V_1 がそれぞれ F6, D6 に接続された単位ニューロンは、 $F6 \Rightarrow D6$ の時系列入力に反応して 2 連音 F6D6 活性パルスが発生する。次の 3 sequence layer のうち、 V_0, V_1 がそれぞれ F6D6, A6 に接続された単位ニューロンは、 $F6 \Rightarrow D6 \Rightarrow A6$ の時系列入力に反応して 3 連音 F6D6A6 活性パルスが発生する。このようにして、メロディに対応した 3 連音活性パルスが次々と発生する。なお、回路規模縮小のため、2 sequence layer と 3 sequence layer との間には top-60 pooling 層が設けられている。曲ごとに、2 sequency layer のシミュレーション後、2 連音のうち活性パルスが大きかった上位 60 のみの活性パルス波形を、次の 3 sequence layer の入力としている。

2.2 単位ニューロン

図 2 に単位ニューロンの構成[1]を示す。 V_0 入力に引き続き V_1 入力があった場合に活性パルス V_{OUT} を出力する。時系列を作り込む観点で、文献[7]の第 6 章のアメフランのニューロン結合を参考にした。 V_0, V_1 は、たとえば、それぞれ倍音端子入力 F6, D6 であるが、以下の動作説明 (図 3) で明らかとなる通り、F6, D6 との間に別の音、たとえば、C6 が入っても活性パルス F6D6 は発生する。また、F6 と D6 との時間間隔が短いほど、大きな活性パルス F6D6 が発生する。

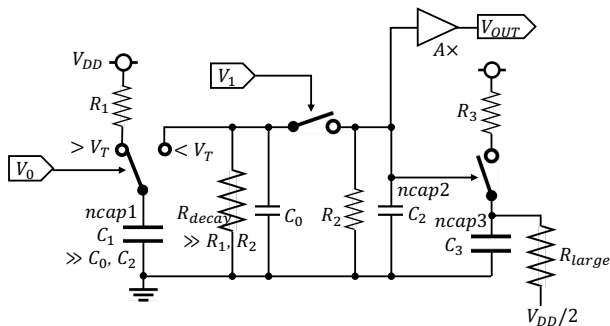


図 2 単位ニューロン

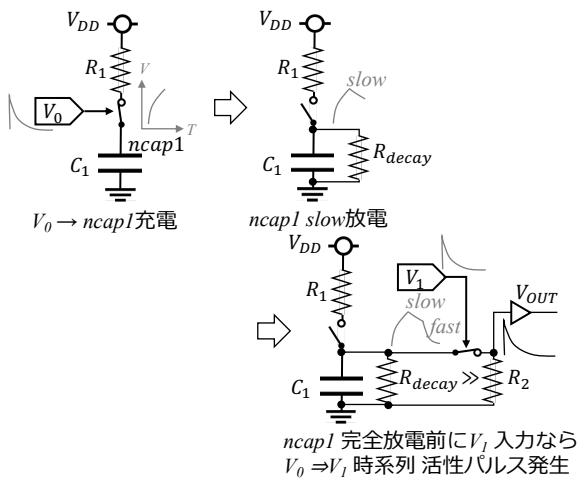


図 3 単位ニューロンの動作

図 3 に単位ニューロンの動作原理を示す。まず、 V_0 入力パルス幅に依存して容量 C_1 が充電される。その後、大きな

抵抗 R_{decay} により C_1 電荷はゆっくり放電される。すべて放電されるより前に、 V_1 パルスが入力されると、残っていた C_1 電荷は抵抗 R_2 により急速に放電され、活性パルス V_{OUT} を発生する。

2.3 活性パルス波形とその解析方法

1 つの曲のメロディ入力に対して、たとえば、図 4 に示すような 3 連音活性パルス波形が得られる。図 4 では 3 つの 3 連音 F6D6A6, F6D6D7, G5F6A#5 の波形を示しているが、他の 3 連音についても同様な波形が得られる。横軸はメロディが流れる時間である。

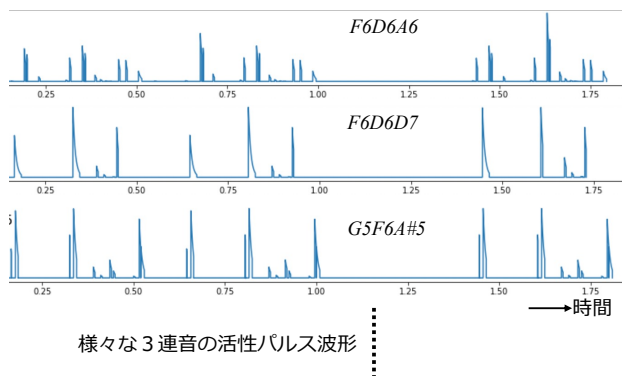


図 4 活性パルス (3 連音) 波形 (曲 A の例)

活性パルス波形のひとつの解析方法は、3 連音ごとの活性パルス最大ピーク電圧を、たとえば、表 1 のように集計し、この表の曲間の類似度を t-SNE などの手法で可視化することである[1]。すべての曲は C 調に転調し、対応する 3 連音を曲間で比較できるようにした。

表 1 活性パルス最大値 (曲 A の例)

3 連音	活性パルス最大値
F6D6A6	0.786258
F6D6C6	0.786143
F6D6F5	0.786105
F6D6F6	0.761677
G5F6D7	0.623658
G5F6A#6	0.622809

この最大値評価に加えて、本研究では、波形全体の特徴を抽出するため、1 つの曲を通しての活性パルス強度の頻度分布に着目した。まず、ある曲に対して、活性パルス最大ピーク電圧が大きい Top 50 の 3 連音を選び、それらの活性パルス強度の最大ピーク電圧に対する比の頻度分布を累積する。

図 5 に、例として、6 つの曲について、この方法で得られた頻度分布を比較する。左側はダウンロード数が多く、右側は少ない曲のヒストグラムである。ヒストグラムの横軸は活性パルス電圧強度、縦軸はそれぞれの曲で発生した頻度である。この 6 曲に限れば、人気上位曲では強度が小さい活性パルスの発生頻度が滑らかに上昇する特徴がある。

ヒストグラム形状の特徴を図 6 に示す方法で数値化した。すなわち、図 5 のグラフを両対数化して、その近似線の傾きと相関係数で「強度が小さい活性パルスの発生頻度が滑らかに上昇する」傾向を数値化した。「滑らかに」は相関

係数が-1に近いことで評価し、「上昇する」は近似線の傾きが負に大きいことで評価する。対数化にあたり、頻度 0 は最小値 1 に置き換えた。

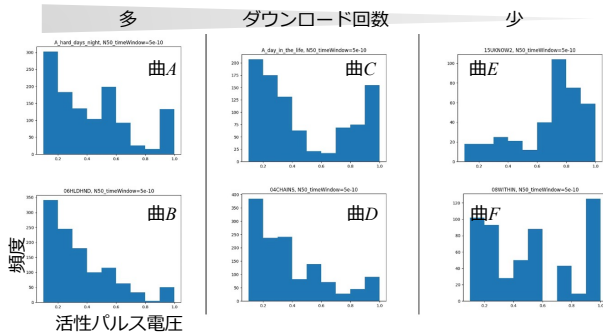


図 5 活性パルス電圧強度 (3 連音) vs. 頻度

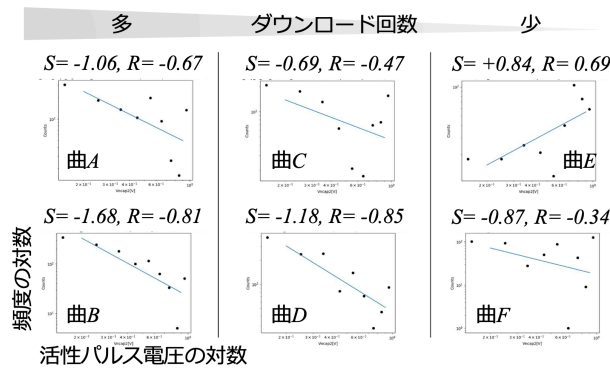


図 6 活性パルス電圧強度 vs. 頻度の両対数グラフ (S:近似線傾き、R:相関係数)

3. シミュレーション結果

3.1 活性パルス (3 連音) 最大値の曲間類似性

表 1 に示した 3 連音ごとの活性パルス最大値のパターンを曲間で類似性比較し、Isomap[8]により 2 次元マッピング

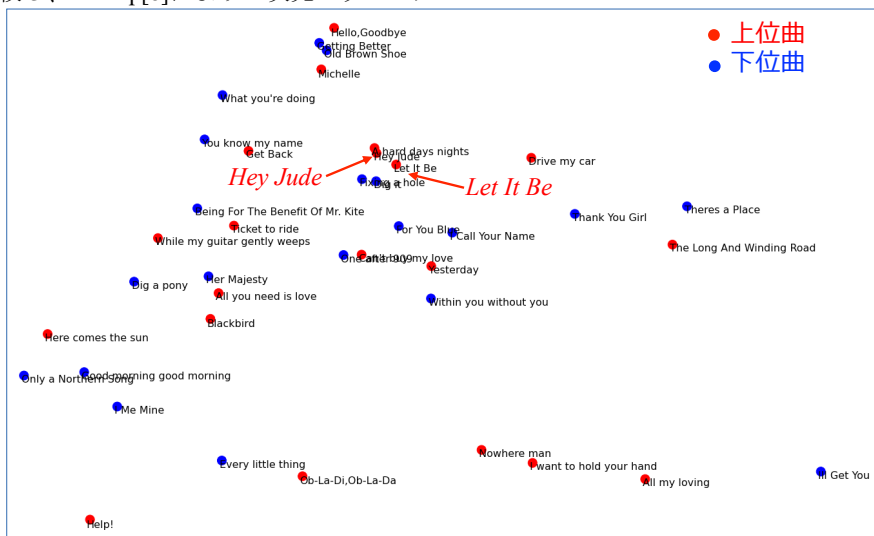


図 7 3 連音活性パルス最大値リスト (表 1) の曲間類似性 Isomap[8]

した結果を図 7 に示す。以前の結果[1]と同様に、良く知られた Let It Be と Hey Jude が近い位置になるなどの特徴は見られたものの、ダウンロード数上位曲 (赤) と下位曲 (青) との明確な分離は観測されなかった。

3.2 活性パルス強度の頻度分布

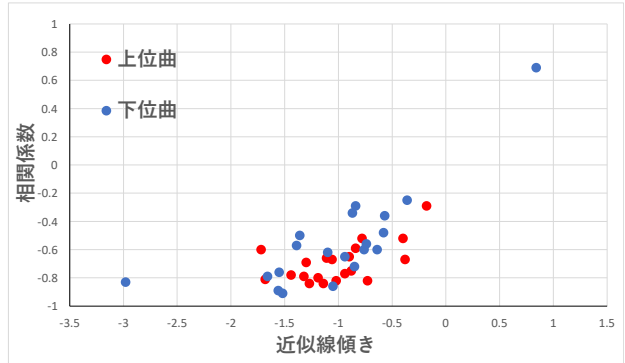


図 8 活性パルス強度 vs. 頻度分布の両対数グラフの近似線傾き及び相関係数 (曲間比較)

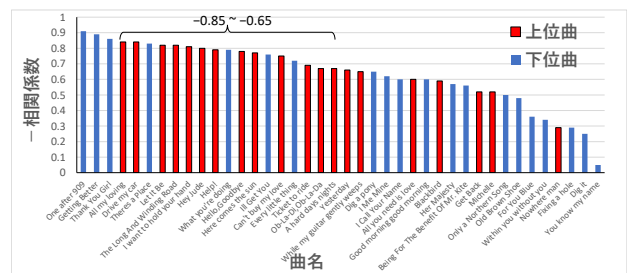


図 9 活性パルス強度 vs. 頻度 (対数) の相関係数 (負に大きい曲順)

ピーク電圧だけではなく、3 連音活性パルス強度の曲全体での挙動を定量化する方法として、2.3 で述べた解析方法、すなわち、活性パルス強度 vs. 頻度分布の両対数近似線傾き及び相関係数を曲間で比較した。結果を図 8

に示す。ダウンロード数上位曲 (赤) は下位曲 (青) に比べて、相関係数がよりマイナス側に多く分布している。これは、強度が小さい活性パルスの発生頻度が滑らかに上昇していることを意味する。

図 9 は、上位曲 20 曲、下位曲 20 曲を、相関係数が負に大きく -1 に近い順にならべたものである。グラフの左側ほど、すなわち相関係数が -1 に近いほど、強度の小さい活性パルスが多く発生し、かつ、その頻度の変化が滑らかであるといえる。相関係数が $-0.85 \sim -0.65$ 付近に、上位曲が有意に集まっている。すなわち、頻度の変化が、「適度に」滑らかである。これが、何らかの形で感性に心地よさを与えている可能性がある。

3.3 音の残響度の違いが及ぼす影響

単位ニューロン回路の放電抵抗 R_{decay} を 4 倍にして、音の残響度の違いが及ぼす影響を調査した。図 2 及び図 3 に示した通り、単位ニューロンに最初入力 V_0 が入った後、次の入力 V_1 が入るまでの間、 V_0 パルスで充電された C_1 電荷は大きな放電抵抗 R_{decay} でゆっくり放電される。次の入力 V_1 が入るまでの時間間隔が長いほど、 $V_0 \Rightarrow V_1$ 時系列に反応して発生する活性パルスは小さくなる。放電抵抗 R_{decay} を 4 倍にすると、 V_0 から V_1 までの時間間隔が 4 倍でも同じ強度の活性パルスを発生する。いわば、ニューラルネットワーク内部での音の残響度を大きくする。

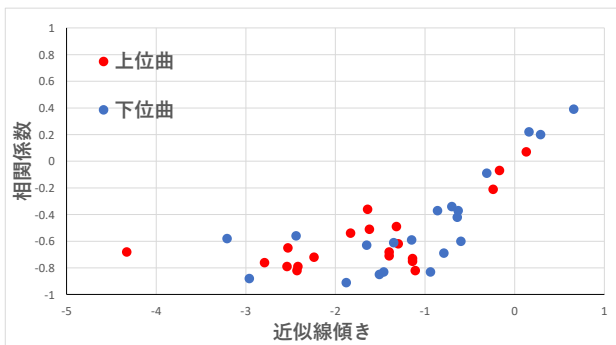


図 10 音の残響度 (R_{decay}) を図 8 に対して 4 倍にした場合

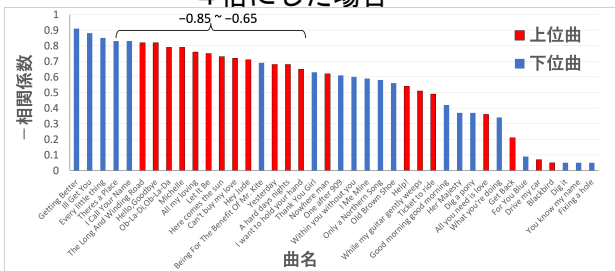


図 11 音の残響度 (R_{decay}) を図 9 に対して 4 倍にした場合

放電抵抗 R_{decay} を 4 倍にした場合の、図 8、図 9 と同様な解析結果を、図 10、図 11 にそれぞれ示す。残響度を大きくすると微小パルスが多くなり解析ノイズになると考え、活性パルス強度が最大値の 30% 以上のパルスを対象に、活性パルス強度 vs. 頻度分布の両対数近似線傾き及び相関係数を求めた。

結果を図 10 に示す。ダウンロード数下位曲 (青) は上位曲 (赤) に比べて、近似線傾き及び相関係数の絶対値が小さい領域に多く分布する形で差異が現れた。

また、図 11 に示す通り、相関係数 $-0.85 \sim -0.65$ 付近に上位曲が有意に集まる傾向が見られた。これは、図 9 と同様である。しかし、上位曲すべてではなく、いくつかの上位曲は相関係数の絶対値が小さい。これらの棒グラフの右側に位置する曲は、メロディの心地よさ以外の要因、たとえば、ビートなどで多くの人に支持されている可能性があると考えられる。

4. おわりに

生体系のアナログの側面を反映させた電気回路ニューラルネットワークを構築し、LTspice シミュレーションにより、人気のあるメロディの特徴を分析した。その単位ニューロンは入力 V_0 に引き続き入力 V_1 に応答して時系列活性パルスを発生し、その活性パルス強度は V_0 の長さや V_0 と V_1 との時間間隔に応じて変化する。曲の midi データのメロディのみを Python コードで LTspice の入力電圧波形に変換し、ニューラルネットワークの 1 層目では 2 連音、2 層目では 3 連音の特徴を活性パルス波形として抽出する。3 連音活性パルス強度の曲ごとの頻度分布を解析した結果、ビートルズ約 200 曲のうち、音楽配信サイトのダウンロード数上位 20 曲と下位 20 曲との間に、有意な差が見られた。すなわち、上位 20 曲では、強度が小さい 3 連音活性パルスの発生頻度が「適度に」滑らかに上昇する傾向が見られた。「適度に滑らかに」という特徴は活性パルス強度と頻度の対数の相関係数が、 $-0.85 \sim -0.65$ の間に偏在するという形で定量的に示された。ヒット曲はメロディ、歌詞、ビート、アレンジなど多くの要素が関与するものの、今回得られた結果は、人気の高いメロディの特徴の一つであると考えられる。

謝辞

一部のデータ取得に携わった、明星大学総合理工学科の小島太一郎氏、佐藤遼太郎氏に感謝いたします。

参考文献

- [1] 今本歩美, 小島太一郎, 本庄佑佳, 竹内幹, “時系列を作り込んだ SPICE 回路ニューラルネットワークを用いたメロディ人気要素の解析”, 信学技報, vol. 122, no. 223, PRMU2022-22, pp. 1-5, (2022).
- [2] M. Davies et. al., “Advancing Neuromorphic Computing with Loihi: A Survey of Results and Outlook,” Proc. of the IEEE, vol. 109, No. 5, pp. 911-934, 2021, DOI: 10.1109/JPROC.2021.3067593.
- [3] 小方厚, “音律と音階の科学”, 講談社
- [4] music21: a Toolkit for Computer-Aided Musicology
- [5] ザ・ビートルズ・コンプリート MIDI
- [6] SortMusic.com / The Beatles / Midis
- [7] E. R. Kandel(編), 金澤一郎(監訳), “カandel神経科学 5th edition”, MEDSi
- [8] sklearn.manifold. Isomap