

1J-02 単純な音楽を効率的に検索するためのニューラルネットワーク[†]

荒木智行 堀 良幸 山本富士男

神奈川工科大学 工学部 情報工学科

1 はじめに

コンピュータ技術の発展に伴い、従来から行われていた事務処理や計算などの実務的な処理だけではなく、趣味、ゲーム、情報家電などに対するコンピュータの需要が高まりつつあり、コンピュータを用いた音楽に対する需要も増大しつつある。こうした分野で使われる音楽は、その作成の過程からコンピュータが用いられることが多い。音楽に対するコンピュータの応用の一分野として、音楽の検索等があげられる。

従来、音楽データベースの検索は、曲名、歌詞、演奏者や作曲家などの文字情報からの検索が行われてきた。しかしながら音楽の検索においては、メロディのような音情報としてのコンテンツによる検索が必要となる場合が多い。しかしながら検索キーとなる音楽情報は、曲の一部分の特徴的なメロディなどである場合が多く、曲全体や小節全体である場合は希で、多くの場合、不完全な情報である場合がほとんどであると考えられる。

本報告ではこれらの点を補う手法として木内、萩原²⁾によって提案されているニューラルネットワークを用いたメロディの記憶と想起の手法を基にして、メロディ、特にコンピュータゲームや携帯電話の着信メロディーなどのような単純なメロディの記憶と想起を行う手法について考察する。ここで、本研究が単純なメロディを対象とした理由は、上述したようにコンピュータを内蔵したゲームなどで使われる音楽では、音楽的に優れたものというより、画面上でのキャラクターの動きや、プレイヤーがゲームをするリズムを促すような単純なメロディが用いられることが多く、今後、ゲームなどで使用される音楽の数は、爆発的に増加するものと考えられ、「純粋な音楽としての音楽」の数を上回るものと考えられる。したがって、このような音楽をメロディにより検索するようなシステムは、今後益々必要となるものと考えられる。

ニューラルネットワークによる想起は逐次的に行われるため、データベースとのパターンマッチングを用いる文字情報による検索方法と比べて、入力に対するインタラクティブな処理を行うことが可能であるという特性を持っている。しかし連想記憶モデルは、時間的なパターンを扱う際には一括入力によって静的パター

ンへの変換を行う必要がある。したがって、本質的な時系列処理にはならない。一方、ニューラルネットワーク自身に、時間的な処理能力を持たせる方法として、時間遅延を持つ結合を付加する方法がある。その一つにネットワーク内に遅延を伴うフィードバック結合を付加する方法がある。これをリカレントニューラルネットワーク (Recurrent Neural Network: RNN) と呼ぶ。そこで本論文では RNN の一つである、中間層に自己フィードバックを持つネットワーク (Local Feedback Multilayered Network: LFMN) を用いてメロディの記憶と想起を行うシステムについて考察した。LFMN を用いたメロディの記憶は、階層型ニューラルネットワークの代表的な学習アルゴリズムであるバックプロパゲーション法を用いて、入力したメロディの次の入力メロディが出力される様に学習を行う。また、メロディの想起は、ある一定の長さのメロディを入力した後、その出力値を LFMN の外部でのフィードバックを行うことにより、次の入力値とする。その実験過程でメロディの1回の記憶における誤差の推移に着目し、想起の正解率及び、記憶に掛かる時間に関して、木ノ内、萩原により提案されている手法より優れた手法を提案し、実験的に検証を行った。実験により、提案したシステムでは、学習回数を可変長にし、“学習曲線の平行な場所を求める”と“テスト想起”という二つのアルゴリズムをメロディの記憶に組み込むことにより、不安定なニューラルネットワークでも確実な想起が可能であることを示した。

2 本研究で用いたニューラルネットワーク

本研究で用いたニューラルネットワークは、図1のような自己フィードバックを持つニューロンで構成される図2のような階層化ニューラルネットワークを用いた。

3 木内、萩原の手法²⁾の改良

木内、萩原の手法を追試した結果、ニューラルネットワークの学習曲線 (誤差曲線) は、学習回数に必ずしも依存してないと考えられる。学習曲線は、大別すると図3(1), (2), (3)にあるような三つの特徴的な部

[†]A Neural Network for Retrieving Simple Music Effectively, Tomoyuki Araki, Yoshiyuki Hori, Fujio Yamamoto
Dept. of Information and Computer Sciences,
Kanagawa Institute of Technology 243-0292 Japan

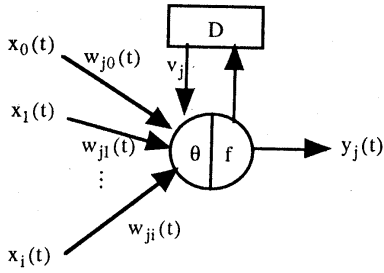


図 1: 自己フィードバックを持つニューロン

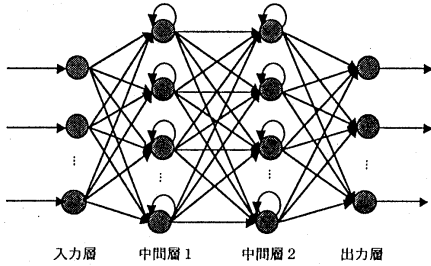
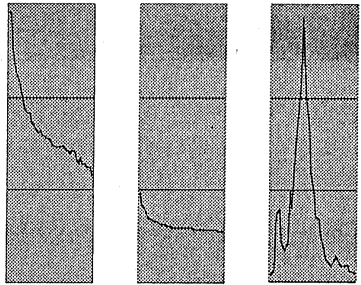


図 2: LFMN

分に分けられる。図 3(1) の状態が続けば、1 回毎の学習によって誤差は少なくなってゆく。図 3(2) が続く場合は、学習曲線は平衡状態にあると言える。この状態では、ネットワークの重みの変化量も小さくなり、出力が安定してくる。図 3(3) は学習の振動の場合である。本研究では、改良法として、“平衡状態を求める”



(1)下り坂 (2)平衡状態 (3)跳ね上がり
図 3: 学習曲線の三つの特徴

と“テスト想起”という二つの評価法をメロディの記憶アルゴリズムに組み込む。具体的には、

- (1) 前の誤差 2 乗和と今の誤差 2 乗和を比較して、その差が 0.1 以下の状態が 5 回続けば平衡状態とする。
- (2) 想起を行う。もし全て望ましい出力が得られれば学習を終了する。もし得られなければ今の誤差 2 乗和が大きすぎるので、現在の誤差 2 乗和 - 0.5 という次の目標値を保持し、学習を進める。これはむやみにテスト想起を行わない為の処置である。その後誤差 2 乗和が目標値よりも低ければもう一度想起を行う。また、誤差 2 乗和が十分低い想起に失敗してしまう場合や目標値が最小値よりも

低い場合があるので、目標値が更新されてから 10 回ごとに保持している値に 0.1 を加算し、数居値の調整を行う。

これらの改良法を取り入れて、表 1 のデータに対して評価を行った結果が表 2 である。改良型では、想起正解率が確実に 100% 付近となる。また、文献 [3] に収録されているメロディに対しても同様な結果が得られ、提案手法の有効性が確認された。

表 1: メロディデータ [2]

#1	cde-fedrefgfedc-
#2	cdedefg-fedcd-e-fgfedcd-efed----
#3	edcdcbabcded-c-babcded-cbaba---
#4	a-g-fefga---g-f-gfede---f-g-g-e-fga-g---feded---
#5	bcd-cde-dcdef-e-fed-e-dcbcd-c---
#6	abc-d-cba---bcd-c-bab---cbaga-b-cdcbag-abc-b-aga---
#7	gab-cba-babcd---cdc-bab-rabcb-a-g---

表 2: 学習回数と想起正解率の比較

学習方法	300回	500回	700回	改良型	改良型 (学習回数1000回)
平均学習時間 (秒)	346.1	631	804.5	551.9433	435.0633
最大学習時間 (秒)	—	—	—	4689.1	1252.3
最小学習時間 (秒)	—	—	—	204.9	204.9
平均学習回数	300	500	700	473.4	374.0677
最大学習回数	—	—	—	3994	1083
最小学習回数	—	—	—	177	177
想起	7	13	12	16	30
正解	6	5	7	4	0
数	5	1	0	3	0
[回]	4	1	1	1	0
	3	2	0	1	0
	2	4	2	1	0
	1	1	4	0	0
	0	3	4	4	0
想起正解率 (%)	69	65.7	76.2	100	99.5

4 むすび

リカレントニューラルネットワークの一つである LFMN を用いてメロディの記憶と想起を行うシステムについて考察した。実験により、提案したシステムでは、学習回数を可変長にし、“学習曲線の平行な場所を求める”と“テスト想起”という二つのアルゴリズムをメロディの記憶に組み込むことにより、ニューラルネットワークにおける学習回数に依存した学習の不安定さを回避でき、かつ、学習の効率を高くできることを示した。

参考文献

- [1] 木ノ内, 萩原. 複素ニューロンによる時系列の学習. 電気学会論文誌, Vol.116-C, No.7, pp.748-754 (1996).
- [2] 木ノ内, 萩原. 複素リカレントニューラルネットワークを用いたメロディの記憶と想起. 情報処理学会論文誌, Vol.39, No.5, pp.1232-1239 (1998).
- [3] 株式会社ほにーてる. ケータイ着メロ♪ドレミ BOOK 2. 双葉社, 1998.