

リカレントニューラルネットワークモデルを用いた 図形の描画系列学習

高田 優子[†][†] 徳島大学 工学部西出 俊[‡][‡] 徳島大学 大学院社会産業理工学研究部康 鑫[‡]任 福継[‡]

1. はじめに

近年、深層学習の様々なモデルが提案される中でリカレントニューラルネットワークを用いた時系列の学習が映像・音声など様々なメディアの学習において有効性が確認されている [1][2]。本研究では人間のスキルの一つとして描画を対象とし、リカレントニューラルネットワークを用いて学習することで人間の描画技術をモデルで表現することを目的とする。同一の図形を様々な方法で描画した際の描画系列を学習し、きれいに描画する描き順をモデル内で表現することが可能であることを確認した。

2. Multiple Timescales Recurrent Neural Network (MTRNN)

本研究ではリカレントニューラルネットワークの一種である Multiple Timescales Recurrent Neural Network (MTRNN) を学習モデルとして用いる [3]。MTRNN の構成を図 1 に示す。MTRNN は従来のリカレントニューラルネットワークモデルと同様、現在の状態 $IO(t)$ を入力することで次の状態 $IO(t+1)$ を出力する予測器として構成されている。通常のリカレントニューラルネットワークモデルとは異なり、コンテキスト層は複数存在する。本研究では 2 種類のコンテキスト層 (Fast Context(C_f) 層と Slow Context(C_s) 層) を持つ MTRNN を用いる。MTRNN の各ニューロン群には時定数と呼ばれるパラメータが設定されており、その値によって各ニューロンの時間的変化が決定される。時定数が大きいニューロン群の時間的変化は小さく、時定数が小さいニューロン群の時間的変化は大きくなる。本モデルでは入出力ニューロン群 (IO) の時定数が最も小さく、 C_f ニューロン群、 C_s ニューロン群の順に時定数が大きくなっていくように設定した。これにより、 C_f と C_s には異なるレベルの文脈情報を自己組織的に学習することができる。

MTRNN の学習は Back Propagation Through Time (BPTT) を用いる [4]。BPTT では各学習データに対して、Forward 計算によって各時刻の入力に対する出力値を計算し、最終時刻の出力誤差を初期時刻まで逆伝播する。

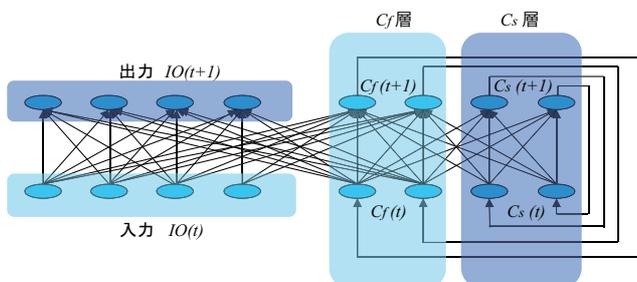


図 1: MTRNN の構成

Learning Drawing Sequences Using Recurrent Neural Network Model
Yuko Takata (Tokushima Univ.), Shun Nishide (Tokushima Univ.), Xin Kang (Tokushima Univ.), and Fuji Ren (Tokushima Univ.)

各重みにかかる誤差値積算し、初期時刻における誤差の積算値に基づいて重み値を更新する。この際、 C_f の初期値 $C_f(0)$ は一定値 0.5 とし、 C_s の初期値 $C_s(0)$ は重みの更新と同様の計算で各データに対して更新する。これにより、 $C_s(0)$ は各学習データに対して固有の値として自己組織化される。

MTRNN の学習後、任意の時系列に対して、BPTT と同様の手法で $C_s(0)$ 値のみを更新 (重みは固定) することで時系列に対応した $C_s(0)$ が計算できる。この計算過程を時系列の認識と呼ぶ。また、与えられた $C_s(0)$ 値と時系列の初期値 ($IO(0)$) を MTRNN に入力することで逐次計算によって各時刻における時系列の出力値が計算でき、時系列全体を復元できる。この計算過程を時系列の生成と呼ぶ。

3. 描画データの取得と実験設定

本実験では Wacom 社の液晶タブレット Cintiq 13HD (DTK-1301/KO) (図 2) を用いて実験データを取得する。本実験では被験者一人に対し、星型の図形を一筆描きで各頂点から時計回り、反時計回りに 3 回ずつ描いてもらった。描画してもらった系列の一例を図 3 に示す。星型には 5 つの頂点があるため、計 30 個のデータを取得した。描画の際、液晶タブレットの座標値を毎秒 10 フレームで取得した。

得られた 30 個の座標系列を学習データとして、8 つの MTRNN を用いて学習した。MTRNN のうち 4 つは IO ニューロンが 2 つ、 C_f ニューロンが 40 個、 C_s ニューロンが 2 つであり、残りの 4 つは IO ニューロン数と C_s ニューロン数が同じであり、 C_f ニューロンを 50 個とした。各 MTRNN の初期重み値はランダムに設定することで得られる学習結果も異なる。各ニューロン群の時定数は IO 、 C_f 、 C_s の順に 2、5、70 と設定した。時定数の値は実験的に定めた。

本実験では学習データである 30 個の手書き星型系列を各 MTRNN で学習後、星型の頂点が正五角形の頂点と一致するきれいな星型の系列 (各頂点から時計・反時計回りで 10 系列) を MTRNN で認識することできれいな星型の $C_s(0)$ を計算した。得られた $C_s(0)$ を MTRNN に再



図 2: ペンタブレット

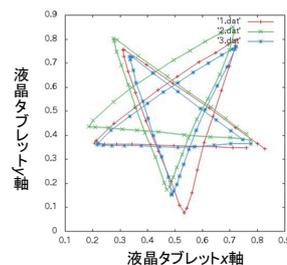


図 3: 被験者が描いた星型の描画系列の一例

入力することできれいな星型のモデル内表現系列を生成し、生成された星型系列ときれいな星型系列の誤差を計算した。誤差は各時刻における座標の絶対誤差を時系列全体で積算し、時系列のステップ数で割った値を用いた。それぞれのきれいな星型系列に対して、8つのMTRNNの誤差平均と標準偏差を評価した。

4. 実験結果

本実験の結果を図4に示す。図4では横軸方向に星型の各描画方法、縦軸方向に誤差平均としている。横軸の各数字は図4下部に示す描画方法の数字ラベルとそれぞれ対応している。誤差は液晶タブレットの幅を1に正規化した値を用いている。

図4より、左下の頂点から始め、反時計回りに星型を描画する系列が誤差平均・標準偏差共に最小になっていることが分かる。学習データ取得時の描画図形を定性的に評価したところ、図4で誤差平均・標準偏差が最も小さくなった描画方法7が最もきれいに描けていることが分かった。描画方法7で生成した星型の一例を図5に示す。図5より学習したMTRNNも描画方法7の星型はきれいに生成することが可能であることが示された。

5. 考察

本実験結果を以下の2点から考察する。

1. 学習系列とMTRNNの生成系列との関係性
2. 他の動作学習への応用

5.1 学習系列とMTRNNの生成系列との関係性

本実験では学習系列を定性的に評価した結果、きれいに描ける系列がMTRNNの生成系列においても誤差が小さくなることを確認した。きれいな系列は安定したダイナミクスを持っていると考えられ、MTRNNの学習によってより安定したアトラクタを形成すると考えられる。一般的にMTRNNの学習は初期状態依存であるが、より安定したダイナミクスを学習する場合は安定した学習も可能になるため、複数のMTRNNから生成される系列の誤差と標準偏差が小さくなると考えられる。

5.2 他の動作学習への応用

本稿では描画において手法の有効性を確認した。他の動作学習においても上記のように安定したダイナミクスを学習によって検出することについては本手法は有効であると考えられる。一方、学習データの次元数が増える場合はMTRNNの学習がより困難になるため、より効率的な学習手法を開発することも重要である。

6. おわりに

本稿では神経回路モデルMTRNNを用いた描画系列の学習について報告した。本手法では被験者が描画した際に取得した座標系列を複数のMTRNNによって学習し、きれいな描画系列をMTRNNで認識・生成した際の誤差平均・標準偏差を評価する。実験ではペンタブレットを用いて星型の各頂点から時計回り・反時計回りに3回ずつ描画した際に取得した座標系列を8種類のMTRNNで学習した。各MTRNNにきれいな星型を認識することで $C_s(0)$ 値を計算し、求めた $C_s(0)$ 値から生成した星型ときれいな星型の誤差を計算した。実験の結果、被験者の描画した星型が定性的にきれいであった描画方法が誤差平均・標準偏差共に小さくなることが確認された。

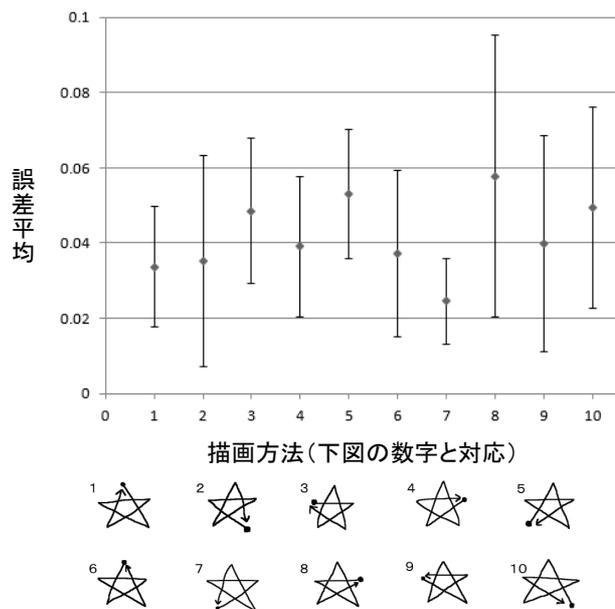


図4: 星型の各系列の誤差平均と標準偏差

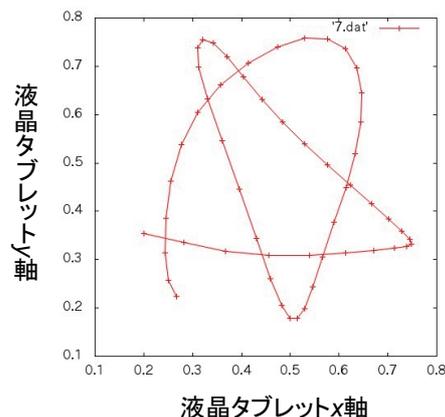


図5: 描画方法7で生成した星型の一例

本実験では定性的な評価のみであり、定量的な有効性は未確認である。今後は星型のきれいさを定量的に評価し、本実験と同様にMTRNNの生成系列の誤差と被験者が描画した図形のきれいさの間の相関関係について確認する。さらに、描画以外のスキルについても有効性を確認し、より汎用性の高い手法として確立する。

謝辞

本研究は科学研究費補助金、若手研究(A)(課題番号16H05877)の支援を受けた。

参考文献

- [1] Y. Yang, D. Krompass, and V. Tresp, "Tensor-Train Recurrent Neural Networks for Video Classification," in Proc. of Int. Conf. Machine Learning, Vol. 70, pp. 3891-3900, 2017.
- [2] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 6645-6649, 2013.
- [3] Y. Yamashita and J. Tani, "Emergence of Functional Hierarchy in a Multiple Timescale Recurrent Neural Network Model: a Humanoid Robot Experiment," PLoS Computational Biology, Vol. 4, No. 11, e1000220, 2008.
- [4] P. Werbos, "Backpropagation through time: What it does and how to do it," Proc. of the IEEE, Vol. 78, No. 10, pp. 1550-1560, 2002.