

MCNN と SOM を用いた動画像の記録と動的想起

渡邊 駿*

Dynamical Recollection and Storage of Video Images via MCNN and SOM

Shun WATANABE

Abstract: 本稿では時系列情報の記憶と再生が可能な連想記憶モデルである多層カオスニューラルネットワーク (MCNN) に関する自己組織化マップ (SOM) を用いた動画像への応用を紹介する。

Keywords: association memory, chaotic neural network, self-organizing map

1. はじめに

脳のように物事の記憶と再生を行う神経細胞網の数理モデルは連想記憶モデルと呼ばれ、数多くのモデルが研究されている。その中でも単純な神経細胞モデルを用いた連想記憶モデルとして Hopfield が提案したネットワークモデル (Hopfield network: HN) ^[1] が有名である。このモデルの特徴は Hebb が提唱した神経細胞の学習仮説^[2]を用いて神経細胞モデルを相互に結合したネットワーク構造である。この構造により HN は個々の細胞が独立しながらも規則的な情報を表現でき、脳のような情報の記憶と再生を行う並列分散処理的モデルとしても注目されている。(以降、情報の記憶を“記録”, 再生を“想起”と呼ぶ。) しかし, HN の挙動はある時点で収束してしまう。この収束性は、実際の神経細胞が行うような常に動的で様々な情報を表現する挙動を表現していないことを意味している。

実際の神経細胞メカニズムを基に提案された神経細胞の簡易モデルを用いた連想記憶モデルとして, Adachi らのカオスニューラルネットワーク (chaotic neural network: CNN) ^[3] が知られている。このモデルは HN と異なり, 常に動的な挙動を行い, 記録した複数の情報を不規則に想起することができる。この動的挙動と情報の想起はカオスニューロンモデルによるカオスの挙動と HN と同様の相互結合ネットワーク構造による情報表現によって実現している。しかし, Adachi らの CNN は不連続な情報しか取り扱えず, 脳のように

に時系列情報の記録と再生が行えなかった。

複数の時系列情報の記録と動的想起を実現する連想記憶モデルとして, Kuremoto らは CNN を改良した多層カオスニューラルネットワーク (multi-layer chaotic neural network: MCNN) を提案した^[4]。このモデルは, 複数の CNN を相互に結合させたネットワーク構造であり, 情報を想起する CNN 層を制御し, 情報を想起する CNN 層を切り替えることによって時系列情報の表現を可能にしている^[5]。MCNN はこれまでに制御手法の改良や層数による性能への影響調査などが行われている。

MCNN が取り扱ってきた情報はこれまで単純なものが多く, 実際の環境を基にした時系列情報は取り扱われていなかった。そこで, 著者らは自己組織化マップ (self-organizing map: SOM) ^[6] と MCNN を組み合わせることにより, 実際の動画像を用いた時系列情報の記録と動的想起を試みた^[7]。その結果, SOM による動画像と時系列情報の相互変換により MCNN への記録と動的想起を実現した。

本稿では, 2015 年 4 月に電気学会に掲載された提案モデル^[7]について紹介し, 実動画像の記録と動的連想のプロセスについて説明する。

2. 連想システム

本章では, 著者らが提案した実動画像の連想システム^[7]について紹介する。

図 1 は連想システム全体のネットワーク構造であり, 中央に多層カオスニューラルネットワーク (MCNN 部), 前後に自己組織化マップ (SOM 部)

* 釧路工業高等専門学校電子工学科

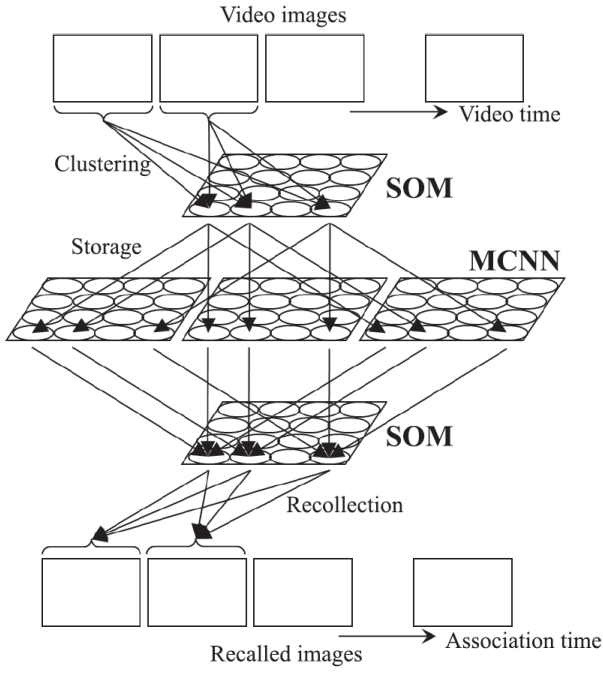


図1 MCNNとSOMを用いた連想システム

が接続されている。以下に図1に示す連想システムの処理の流れを示す。

1. 動画画像から特徴マップを生成（前SOM部）
2. 特徴マップを用いて動画画像を時系列2値パターンへ変換（前SOM部）
3. 時系列2値パターンをMCNNへ記名（MCNN部）
4. MCNNによる記録パターンの動的想起（MCNN部）
5. MCNNの想起パターンを動画画像へ変換（後SOM部）

以降において、SOMと特徴マップの生成方法と動画画像の時系列2値化、MCNNへの時系列2値パターンの記録方法と動的想起、想起パターンの動画画像への変化方法について説明する。

2.1. SOMと特徴マップの生成

SOMとは、Kohonenによって提案された、神経細胞の学習機構を用いた教師なし学習モデルであり、学習データの集合から特徴的なデータ集団を特徴ベクトルとして抽出できるモデルである。動画画像は複数の画像が順番に切替わることで、画像が動いているように表現しており、 τ 番目画像は $\psi(\tau)$ と表現できる。そこで、提案された連想システムでは、2次元トラス型のSOMと複数の画像 $\psi(\tau)$ を用いて以下の式を $t_1 = T_1$ まで繰り返すことによりSOMへ特徴マップを生成している。

$$i^{(win)}, j^{(win)} \leftarrow \arg \min_{i,j} \|v_{ij} - \psi(\tau)\| \quad (1)$$

$$\rho = \rho^{(max)} - (\rho^{(max)} - \rho^{(min)}) \frac{t_1 - 1}{T_1 - 1} \quad (2)$$

$$\sigma = \sigma^{(max)} - (\sigma^{(max)} - \sigma^{(min)}) \frac{t_1 - 1}{T_1 - 1} \quad (3)$$

$$d(i, k, L) = \min\{|i - k|, L - |i - k|\} \quad (4)$$

$$\kappa_{ij} \leftarrow \exp\left(\frac{d(i, i^{(win)}, L_i)^2 + d(j, j^{(win)}, L_j)^2}{-2\sigma}\right) \quad (5)$$

$$v_{ij} \leftarrow v_{ij} + \rho \kappa_{ij} (\psi(\tau) - v_{ij}) \quad (6)$$

2.2. 動画画像の時系列2値パターン化

連想システムでは以下の式により各 τ 番目画像 $\psi(\tau)$ に対して類似するF個の特徴ベクトル v_{ij} から時系列2値パターンの τ 番目2値パターン $\chi_k(\tau)$ へ変換を行っている。

$$\chi_k(\tau) = \begin{cases} 1 & \text{if } k \in \{(i-1)L_j + j \mid i, j \in P(\psi(\tau))\}, \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

2.3. MCNNへのパターン記録と動的想起

MCNNとは、Kuremotoらによって提案されたAdachiらが提案したCNNを複数用いて構成される時系列2値パターンの記録と動的想起が可能な連想記憶モデルである。以下の式はM層のCNNで構成されたMCNNのN個のニューロンで構成されるn層目CNNである。また、n'層は「想起層」と呼び、出力 $x_i^{(n)}(t+1)$ により2値パターンを動的に想起する。一方 $n \neq n'$ 層は「提示層」と呼び想起したパターンを想起層へ提示する。

$$x_i^{(n)}(t+1) = \begin{cases} f(y_i^{(n)}(t+1)) & \text{if } n = n' \\ x_i^{(n)}(t) & \text{otherwise} \end{cases} \quad (8)$$

$$y_i^{(n)}(t+1) = \eta_i^{(n)}(t+1) + \zeta_i^{(n)}(t+1) + \xi_i^{(n)}(t+1) \quad (9)$$

$$\eta_i^{(n)}(t+1) = k_1^{(n)} \eta_i^{(n)}(t) + \sum_{j=1}^N w_{ij}^{(n,n)} x_j^{(n)}(t) \quad (10)$$

$$\zeta_i^{(n)}(t+1) = k_r \zeta_i^{(n)}(t) - \alpha x_i^{(n)}(t) + a_i^{(n)} \quad (11)$$

$$\xi_i^{(n)}(t+1) = k_2^{(n)} \xi_i^{(n)}(t) + \sum_{m \neq n} \sum_{j=1}^N w_{ij}^{(n,m)} x_j^{(m)}(t) \quad (12)$$

$$f(z) = \frac{1}{1 + e^{-z/\epsilon}} \quad (13)$$

MCNNへの時系列2値パターン $\chi_k(\tau)$ の記録にはHebb学習則を基にした以下の式を用いる。また、時系列2値パターンはMCNNの各CNN層へ交互に記録される。

$$g = (\tau - 1) \bmod M + 1 \quad (14)$$

$$\tau' = M \left\lfloor \frac{\tau - m}{M} \right\rfloor + m \quad (15)$$

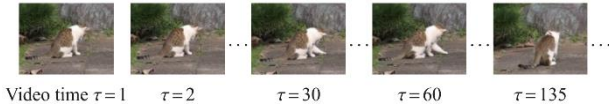


図2 実験で用いた動画の一部

$$\Delta w_{ij}^{(n,m)}(\tau) = \begin{cases} 0 & \text{if } n \neq g \vee \tau' \leq 0, \\ 0 & \text{if } i = j \wedge n = m, \\ \beta \chi_i(\tau) \chi_i(\tau') & \text{otherwise} \end{cases} \quad (16)$$

$$w_{ij}^{(n,m)} \leftarrow w_{ij}^{(n,m)} + \Delta w_{ij}^{(n,m)}(\tau) \quad (17)$$

MCNN は想起層の内部パラメータ $(k_1^{(n)}, k_2^{(n)})$ を制御することにより記録時系列パターンをより適切に想起することができる。本稿ではパラメータ制御手法として最適パラメータ探索手法である粒子群最適化 (particle swarm optimization: PSO) を用いる。また、探索したパラメータの最適性は以下のエネルギー関数が最小値となるパラメータを最適として評価した。

$$E(t) = -\frac{1}{2} \sum_{n,m} \sum_{i,j} x_i^{(n)}(t) w_{ij}^{(n,m)} x_j^{(m)}(t) \quad (18)$$

2.4. 想起パターンの動画化

連想システムにおいて、MCNN で想起された2値パターン $x_k^{(n)}(t)$ は、下記の式により SOM の特徴マップ v_{ij} を用いることで再構成される。また、再構成には中央値を用いることで、外れ値の影響を抑制している。

$$A(t) = \{i, j | (i-1)L_j + j = k \wedge x_k^{(n)}(t) \geq 0.5\} \quad (19)$$

$$\phi(t) = \text{median}_{i,j \in A(t)} \{v_{ij}\} \quad (20)$$

3. 実験結果

本章では、連想システムによる実動画の記録と動的想起のそれぞれに関する結果を示す。本実験で用いた動画は、図2に示す14秒間15fpsの動画から抽出した210枚の120×90ピクセル24



図3 10×10 トーラス型 SOM の特徴マップ

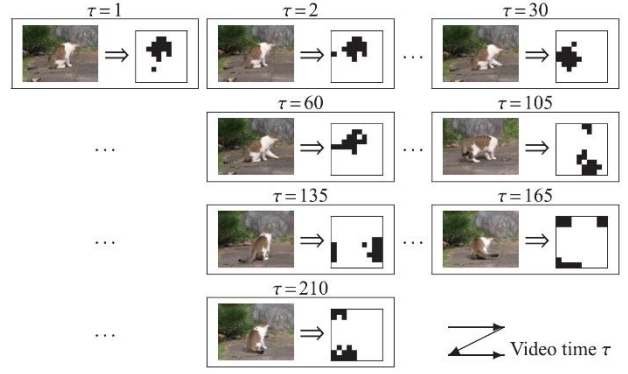


図4 動画から生成された2値パターンの一部

ビット RGB カラーの画像 $\psi(\tau), \tau = 1, \dots, 200 (= \tau^{(end)})$ である。また、この実験では画素値を0~255の値範囲から0~1の値範囲へ正規化している。

まず、図3にSOMにより動画から生成された10×10の特徴マップと参照ベクトルの画像を示す。この特徴マップは動画画像中の特徴的な画像(場面)が参照ベクトルとして記憶されていることがわかる。例として、丸印○で示した図3の左から2、上から2の画像は猫がしゃがんでいる様子である。一方で、右から1、上から4の画像の猫は立ち上がっている様子が見て取れる。

次に、特徴マップを用いて動画から生成された時系列2値パターンを図4に示す。図4は類似する画像に対応する2値パターンも同様に類似し、類似していない画像に対する2値パターンも互いに類似していないことを表している。例として、 $\tau = 1$ と $\tau = 2$ の画像と2値パターンは互いに類似しているが、 $\tau = 1$ と $\tau = 30, \tau = 135$ は画像も2値パターンも異なっていることが判断できる。

そして、図5にMCNNにより想起された2値パターンから特徴マップを用いて再構成された画像の一部を示す。この結果は、記録された動画画像の特徴的な画像を順番に想起していることを表している。例として、図5の $t = 5, 7, 77$ で示された画像の変化は図4の $\tau = 30 \sim 60$ で示された画像の変化(猫が手を出している様子)に類似している。さらに、図5の $t = 81, 82$ で示された画像の変化は図4の $\tau = 165 \sim 210$ で示された画像の変化

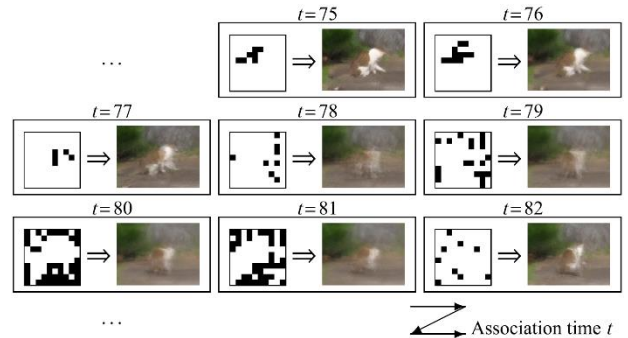


図5 MCNN の動的想起から再構成された動画画像

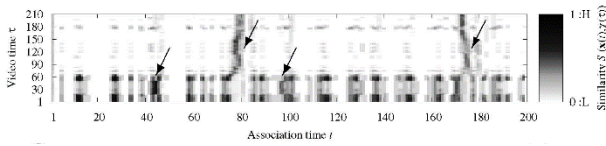


図6 MCNNの想起と時系列2値パターンとの類似性

(猫が頭を下げた状態から、後ろ向きになった様子)に類似している。

最後に、図5で示した時系列2値パターンの動的想起過程について評価したものを図6に示す。図6は動画像の時系列2値パターンとMCNNの想起パターンを以下の類似性を示す式で表現したものである。縦軸は時系列2値パターンの時刻 τ 、横軸はMCNNの想起時刻 t であり、ある時刻 τ の時系列2値パターンとある時刻 t のMCNNの想起パターン間の類似性が高い場合は色を黒く、低い場合では色を白く表現している。

$$S(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\dagger \mathbf{b} / (\|\mathbf{a}\| \|\mathbf{b}\|) \quad (21)$$

図6の矢印 \downarrow はMCNNの動的想起が特徴的な変化を行った場所を表している。特に、左から2つ目の矢印では右斜め上へ黒さが移動している。図5に示したこの時刻領域でのMCNNの動的想起の変化は、右斜め上へ濃さが移動していることから、動画像のある場面から他の場面を順番に想起していることが判断できる。

図7AにMCNNへ記録した時系列2値パターンの相互類似性を表したものを示す。縦軸と横軸は共に時系列2値パターンの時刻 τ であり、パターン間の類似性が高いほど色を黒く、低いほど白く表現している。図7Aは黒く表現された類似性が高いパターンが時間を飛んで出現していることから、似た画像(場面)が時間をおいて再び出現していることを表している。図7Bは図7Aで類似している場面領域を任意に区画分け、単純化したものであり、類似画像の再出現とその変化順を表している。また、図7Bの Φ_A は猫が下を向いている画像(図4 $\tau=1$)、 Φ_B は猫が手を出している画像(図4 $\tau=30$)、 Φ_D は立ちながら後ろ向きの画像(図4 $\tau=135$)、 Φ_E は横向きの画像(図4 $\tau=105$)、 Φ_F は座りながら後ろ向きの画像(図4 $\tau=210$)である。

以上の実験結果から、連想システムは画像の記録だけでなく、その時間的な順序も記録し、想起していることがわかる。

4. まとめ

本稿では、自己組織化マップ(SOM)と多層カオスニューラルネットワーク(MCNN)を用いた連想システムを紹介し、実動画像への適応例を紹介した。本研究の大きな特徴として、静的な情

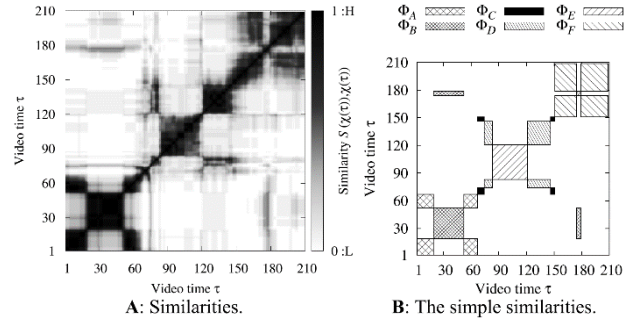


図7 時系列2値パターンの相互類似性

報の記憶と再生だけでなく、動的な情報の記憶と再生が行える点である。例えば、自然界における事象は、本実験で用いた実画像のような非線形的な時間変化を行う。私たちの脳はその非線形的な変化に適応し、正しくその順序や動きを記憶し、思い出すことができる。本稿で紹介した連想システムを発展させることで、脳のような記憶システムを表現することが期待できる。

今後の研究として、脳における時系列情報の取り扱いメカニズムと連想システムのメカニズムとの比較や、時系列情報の最大記憶容量や取り扱い可能な最長時系列などの調査が挙げられる。

参考文献

- [1] J.J. Hopfield: "Neurons with graded response have collective computational properties like those of two-state neurons", Proceedings of the National Academy of Sciences, Vol.81, No.10, pp.3088-3092 (1984)
- [2] D.O. Hebb: The organization of behavior, New York, Wiley (1949)
- [3] M. Adachi and K. Aihara: "Associative dynamics in a chaotic neural network", Neural Networks, Vol.10, No.1, pp.83-98 (1997)
- [4] T. Kuremoto, T. Eto, M. Obayashi, and K. Kobayashi: "A multilayered chaotic neural network for associative memory", Proceedings of SICE Annual Conference, pp.1020-1023 (2005)
- [5] S. Watanabe, T. Kuremoto, K. Kobayashi, S. Mabu, and M. Obayashi: "The recollection characteristics of a generalized MCNN", Proceeding of the SICE Annual Conference, pp.1375-1380 (2013)
- [6] T. Kohonen: Self-organizing maps, Vol.30, Springer (2001)
- [7] 渡邊駿, 吳本堯, 小林邦和, 間普真吾, 大林正直: "MCNNとSOMを用いた動画像の記録と動的想起", 電気学会論文誌C(電子・情報・システム部門誌), Vol.135, No.4, pp.414-422 (2015)