Introduction

Hopfield Network
Introduction

Hopfield Network
Low storage capacity!
Hopfield Network
Low storage capacity!

(Chiueh and Goodman, 1991)

Recurrent Correlation Networks
High storage capacity!
Introduction

Quaternionic Hopfield Networks
Low storage capacity!

(Kobayashi, 2016; Valle, 2014)

Hopfield Network
Low storage capacity!

(Chiueh and Goodman, 1991)

Recurrent Correlation Networks
High storage capacity!
Introduction

Quaternionic Hopfield Networks
Low storage capacity!

(Hobayashi, 2016; Valle, 2014)

Hopfield Network
Low storage capacity!

(Chiueh and Goodman, 1991)

Recurrent Correlation Networks
High storage capacity!

Quaternionic Recurrent Correlation Neural Network
High storage capacity!
Outline

Introduction

Review the Bipolar Recurrent Correlation Neural Networks

Basic Concepts on Quaternions

Quaternionic Recurrent Correlation Neural Networks (QRCNNs)

QRCNNs as Associative Memories

Concluding Remarks
From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{u^1, \ldots, u^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$ 

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t + 1) = \begin{cases} 
\frac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\
x_i(t), & v_i(t) = 0,
\end{cases}$$

where

$$v_i(t) = \sum_{j=1}^{N} m_{ij} x_j(t), \quad \forall i = 1, \ldots, N,$$
From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{ \mathbf{u}^1, \ldots, \mathbf{u}^K \} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$ 

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \frac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{j=1}^{N} \left( \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi \right) x_j(t), \quad \forall i = 1, \ldots, N,$$
From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u^\xi_i u^\xi_j.$$ 

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t + 1) = \begin{cases} \frac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{j=1}^{N} \sum_{\xi=1}^{K} \left( \frac{1}{N} \sum_{j=1}^{N} x_j(t) u^\xi_j \right) u^\xi_i, \quad \forall i = 1, \ldots, N.$$
From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{u^1, \ldots, u^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u^\xi_i u^\xi_j.$$ 

Given an input $x(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \frac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{\xi=1}^{K} \left( \frac{1}{N} \left\langle u^\xi, x(t) \right\rangle \right) u^\xi_i, \quad \forall i = 1, \ldots, N.$$
From Hopfield Network to RCNNs

Given a set of vectors \( \mathcal{U} = \{u^1, \ldots, u^K\} \subseteq \{-1, +1\}^N \), compute

\[
m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^{\xi} u_j^{\xi}.
\]

Given an input \( x(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N \), define

\[
x_i(t+1) = \begin{cases} 
\frac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\
x_i(t), & v_i(t) = 0,
\end{cases}
\]

where

\[
v_i(t) = \sum_{\xi=1}^{K} f \left( \frac{1}{N} \langle u^{\xi}, x(t) \rangle \right) u_i^{\xi}, \quad \forall i = 1, \ldots, N,
\]

and \( f : [-1, 1] \rightarrow \mathbb{R} \) is a continuous nondecreasing function.
From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{u^1, \ldots, u^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$ 

Given an input $x(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \frac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{\xi=1}^{K} w_\xi u_i^\xi$$

and

$$w_\xi = f \left( \frac{1}{N} \langle u^\xi, x(t) \rangle \right).$$

Here, $f : [-1, 1] \rightarrow \mathbb{R}$ is a continuous nondecreasing function.
Quaternions

A quaternion is a four-dimensional hyper-complex number

\[ q = q_0 + q_1 i + q_2 j + q_3 k, \]

where \( i, j, \) and \( k \) are imaginary numbers such that

\[ i^2 = j^2 = k^2 = ijk = -1. \]

The real part and the vector part of \( q \) are

\[ \text{Re} \{ q \} \equiv q_0 \quad \text{and} \quad \text{Ve} \{ q \} \equiv \tilde{q} = q_1 i + q_2 j + q_3 k. \]

The sum and the product of quaternions are

\[ p + q = (p_0 + q_0) + (p_1 + q_1)i + (p_2 + q_2)j + (p_3 + q_3)k. \]

and

\[ pq = p_0 q_0 - \tilde{p} \cdot \tilde{q} + p_0 \tilde{q} + q_0 \tilde{p} + \tilde{p} \times \tilde{q}. \]
The conjugate and the norm are defined by

\[ \bar{q} \equiv q_0 - \bar{q} \quad \text{and} \quad |q| = \sqrt{\bar{q}q} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}. \]

We say that \( q \) is a unit quaternion if \( |q| = 1 \).

We denote by \( \mathbb{S} \) the set of all unit quaternions, i.e.,

\[ \mathbb{S} = \{ q \in \mathbb{H} : |q| = 1 \}. \]
A quaternionic RCNN (QRCNN) defines the sequence

\[ x_j(t + 1) = \begin{cases} 
\frac{v_j(t)}{|v_j(t)|}, & 0 < |v_j(t)| < +\infty, \\
x_j(t), & \text{otherwise},
\end{cases} \]

where

\[ v_j(t) = \sum_{\xi=1}^{K} w_\xi(t) u_\xi^j \quad \text{and} \quad w_\xi(t) = f \left( \frac{1}{N} \Re \left\{ \langle u_\xi^j, x(t) \rangle \right\} \right) , \]

and \( f : [-1, 1] \rightarrow \mathbb{R} \) is a continuous nondecreasing function.

Recall that \( \langle u_\xi^j, x(t) \rangle = \sum_{j=1}^{N} x_j(t) u_\xi^j \).
Examples of QRCNNs

- The correlation QRCNN is obtained by considering
  \[ f_c(x) = x. \]

- The exponential QRCNN is obtained considering
  \[ f_e(x) = e^{\alpha x}, \quad \alpha > 0. \]

- The high-order QRCNN is obtained by considering
  \[ f_h(x) = (1 + x)^q, \quad q > 1 \text{ is an integer.} \]

- The potential-function QRCNN is obtained by considering
  \[ f_p(x) = \frac{1}{(1 - x + \varepsilon_{mach})^L}, \quad L \geq 1. \]
Convergence

Theorem (Convergence of the QRCNNs)

The sequence produced by a QRCNN always converges.

The sequence \( \{x(t)\}_{t \geq 0} \) yields a local minima of the energy function

\[
E(x) = - \sum_{\xi=1}^{K} F \left( \frac{\text{Re} \{ \langle u^\xi, x \rangle \}}{N} \right), \quad \forall x \in \mathbb{S}^N,
\]

where \( F \) is a certain primitive of the real-valued function \( f \).

Thus, we can define the mapping \( \psi : \mathbb{S}^N \rightarrow \mathbb{S}^N \) by setting

\[
\psi(x) = \lim_{t \to \infty} x(t) \quad \text{with} \quad x(0) = x.
\]
We say that $\psi$ implements a $\tau$-associative memory designed for the storage of $u^1, \ldots, u^K$ if

$$\| \psi(u^\xi) - u^\xi \| \leq \tau, \quad \forall \xi = 1, \ldots, K,$$

where $\| \cdot \|$ denotes the Euclidean norm and $\tau > 0$ is a tolerance.

**Theorem**

Consider a fundamental memory set $\mathcal{U} = \{u^1, \ldots, u^K\} \subseteq S^N$ and a sufficiently small tolerance $\tau > 0$. If $f(x; \lambda) = [A(x)]^\lambda$, then there exists a sufficiently large parameter $\lambda$ such that

$$\| x(1) - u^\gamma \|_2 < \tau,$$

where $u^\gamma$ is the fundamental memory closest $x(0)$. 
Consequences of the Theorem

If the input $x(0)$ corresponds to a corrupted version of $u^\gamma$, then we can always fine tune the parameter $\lambda$ of a such that $x(1)$ sufficiently close to $u^\gamma$.

Any input in the open ball

$$\mathcal{B}(u^\gamma) = \left\{ x \in S^N : \|x - u^\gamma\|_2 < \min_{\xi \neq \gamma}\{\|x - u^\xi\|}\right\},$$

is attracted to $u^\gamma$.

The radius of attraction of $u^\gamma$ is, at least, half the Euclidean distance between the two closest fundamental memories.

Since $u^\gamma \in \mathcal{B}(u^\gamma)$, the inequality $\|\psi(u^\gamma) - u^\gamma\|_2 \leq \tau$ holds for all $\gamma$. 
Capability to Implement Associative Memory

Maximum of the Largest Output Error vs. $q, L, \alpha$

- High-order QRCNN
- Potential-function QRCNN
- Exponential QRCNN

Marcos Eduardo Valle (Unicamp)
Radius of the Basis of Attraction

High-order: $x(1)$
Potential-function: $x(1)$
Exponential: $x(1)$

High-order: $\psi(x)$
Potential-function: $\psi(x)$
Exponential: $\psi(x)$

$\delta_{\text{min}}$
Concluding Remarks

We generalized the bipolar RCNN using unit quaternions.

Quaternionic recurrent correlation neural networks (QRCNNs) can implement high-capacity associative memories.

- The sequence produced by an QRCNN always converge.
- But it may converge to an undesired spurious state.

If the activation function is an exponential in its parameter, the QRCNN can retrieve an item if the input is sufficiently close to it.

The radius of the basis of attraction of the resulting associative memory is at least half the distance between the two closest fundamental memories.

Thank you very much!
References (1)

