DEVELOPING HOPFIELD NEURAL NETWORKS USING GAUSSIAN DISTRIBUTED SMALL WORLD TOPOLOGY FOR VISUAL OBJECT TRACKING

SUN JUN

UNIVERSITI SAINS MALAYSIA

2024

DEVELOPING HOPFIELD NEURAL NETWORKS USING GAUSSIAN DISTRIBUTED SMALL WORLD TOPOLOGY FOR VISUAL OBJECT TRACKING

by

SUN JUN

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

June 2024

ACKNOWLEDGEMENT

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully at this great University, the Unviersiti Sains Malaysia (USM).

I would like to express my sincere gratitude to my supervisor, Dr. Saratha Sathasivam., Associate Professor in the School of Mathematical Science, for giving me the opportunity to conduct research and providing invaluable guidance throughout this research. Her dynamism, vision, sincerity, and motivation have deeply inspired me. She has taught me the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honor to work and study under her guidance. I am extremely grateful for what she has offered me. I would also like to thank her for her friendship, empathy, and great sense of humor. I am extending my heartfelt thanks to her family for their acceptance and patience during the discussion I had with her on research work and thesis preparation.

I am extremely grateful to my parents for their love, prayers, care, and sacrifices in raising me and preparing me for my future. I am very much grateful to my wife and my daughter for their love, understanding, prayers, and continued support in completing this research work.

I would like to thank all the students of the School of Mathematical Science, USM, for their support during my research work. I would also like to thank all the staff of the Mathematical Science School for their kindness. My Special thanks goes to Assoc Prof. Dr. Farah Aini Binti Abdullah for the keen interest shown to complete this thesis successfully.

Finally, my thanks go to all the people who have supported me to complete the research work directly or indirectly.

Sun Jun

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

r	The radius from the centroid to the current data layer
L	Total number of nodes in the network
d	Dimension of the underlying lattice structure of the small-world network
P(r)	Wiring probability distribution function
N(r)	Total number of nodes at radius r.
$N_b(r)$	Number of nodes connected to the centroid at radius r
k	The initial range of the small-world network
δ	The parameter of the Gaussian distributed wiring function
ξ	The length scale of the Gaussian distributed small-world network
L_b	The total number of shortcuts in the network
L_c	The number of centroids in the network
$w_{i,j}$	Weight matrix of Gaussian distributed small-world Hopfield neural network.
S_i	Signum activation function of Gaussian distributed small-world Hopfield neural network.
V_i	State of neuron node i
T_i	The threshold of neuron node i
s	Number of patterns (s = 1,2,3…. m)
E_i'	Checking metrics for change detection of prior frame
E_i''	Checking metrics for change detection of second frame
I (x, y, t)	The value of the feature point at the position (*x*, *y*) at time *t*

LIST OF ABBREVIATIONS

USM	Unviersiti Sains Malaysia
VOT	Visual object tracking
ANN	Artificial neural network
HNN	Hopfield neural network
CHNN	Continuous Hopfield neural network
DHNN	Discrete Hopfield neural network
2-SAT	2-Satisfiability
GDSWHNN	Gaussian distributed small-world Hopfield neural network
WSSW	Watts and Strogatz small-world
NWSW	Newman and Watts small-world
CNN	Cellular neural network
FC	Fully connected
APL	Average path length
fMRI	Functional magnetic resonance imaging
MEG	Magnetoencephalography
EEG	Electroencephalogram
HOG	Histogram of oriented gradient
SVM	Support vector machine
TLD	Tracking learning detection
KCF	Kernelized correlation filter
ECO	Efficient convolution operators
MDNet	Multi-domain network
IoU	Intersection over union
AOS	Average overlap score
OPE	One pass evaluation
RNG	Random number generation
AUC	Area Under Curve
AO	Average overlap
SR	Success Rate
AOI	Area of Interest

MEMBANGUNKAN RANGKAIAN NEURAL HOPFIELD MENGGUNAKAN TOPOLOGI DUNIA KECIL TERAGIH GAUSSIAN UNTUK PENGESANAN OBJEK VISUAL

ABSTRAK

Pengecaman objek visual (VOT) dilihat sebagai sebuah topik penyelidikan yang mencabar dalam bidang kecerdasan buatan. Pada masa kini, kebanyakan industri bergantung kepada teknologi pengecaman objek bagi mengenalpasti kesilapan, memantau persekitaran, dan membuat keputusan tepat berdasarkan hasil pengecaman tersebut. Pengecaman objek visual telah mewujudkan pelbagai inovasi seperti kenderaan autonomi, sistem pemantauan trafik, sistem diagnostik perubatan jarak jauh, dan banyak lagi aplikasi canggih dalam masa terdekat. Namun begitu, dalam kemajuan yang ketara ini, tidak dapat dinafikan bahawa otak manusia lebih efisien dalam mengesan objek dan memerlukan sumber yang sedikit berbanding teknologi pengecaman objek tersebut. Kajian neurosains yang terkini telah membuktikan bahawa rangkaian neural buatan yang dicipta seperti sistem saraf sebenar manusia boleh berfungsi dengan lebih efisien dalam tugas pengecaman yang kompleks. Oleh itu, kaedah baru pengecaman objek visual berdasarkan rangkaian neural buatan oleh Hopfield merupakan objektif utama dalam kajian ini. Sebuah rangkaian dunia kecil digunakan sebagai topologi untuk model rangkaian neural. Walau bagaimanapun, ciri-ciri biologi tersebut telah digabungkan ke dalam rangkaian dunia kecil seperti algoritma kerosakan eksponen, yang boleh menyerupai beberapa ciri struktur korteks serebrum. Pada rangkaian neural tersebut, setiap piksel bingkai video dipadankan dengan neuron pada posisi yang selari. Keupayaan piksel dikategorikan sebagai keadaan neuron. Bingkai video itu dihafal setelah semua saraf dalam rangkaian neural telah dilatih untuk berada dalam keadaan stabil. Sebuah mekanisme bionik yang menggunakan sifat memori berkaitan sistem neural bionik Hopfield telah dicadangkan untuk mengecam objek dalam bingkai video. Maklumat pergerakan piksel objek yang bergerak seperti orientasi dan halaju, diperolehi daripada anggaran peralihan kedudukan neuron dalam rangkaian neural. Selain itu, maklumat gerakan subset neuron digabungkan dengan topologi dunia kecil oleh Gaussian untuk mengundurkan semula maklumat pergerakan objek sasaran, dengan mengambil kira sifat perilaku kolektif neuron berdekatan. Kaedah yang dipilih telah dilaksanakan menggunakan alat perkembangan dalam MATLAB R2016a. Simulasi berangka turut dilakukan pada platform perisian dengan sistem operasi Windows 11. Prestasi kaedah yang dicadangkan telah dinilai berdasarkan IoU (Intersection of Union), ketepatan, kekukuhan, masa pengiraan, dan sumber pengiraan. Ujian perbandingan telah dijalankan terhadap lima set data yang memang wujud. Keputusan ujian ini telah membuktikan kelebihan luar biasa bagi kaedah yang dicadangkan dari segi ketepatan dan penggunaan sumber.

DEVELOPING HOPFIELD NEURAL NETWORKS USING GAUSSIAN DISTRIBUTED SMALL WORLD TOPOLOGY FOR VISUAL OBJECT TRACKING

ABSTRACT

Visual object tracking (VOT) is considered a challenging research topic in artificial intelligence. Today, many industries rely on object tracking technologies to identify errors, monitor environments, and make timely decisions based on tracking results. Visual object tracking has enabled many innovations, such as autonomous vehicles, traffic monitoring systems, remote medical diagnostic systems, and more cutting-edge applications are on the horizon. However, among these notable achievements, it is worth noting that, unlike these object-tracking techniques, a human brain is more efficient for object tracking tasks and requires fewer resources. Recent neuroscience studies have shown that artificial neural networks organized as real cortical connectivity may perform more efficiently in complex recognition tasks. Therefore, a novel visual object tracking method based on Hopfield neural networks is proposed in this study. A small-world network is employed as the topology of the neural network model. However, a biological feature is integrated into the small-world network model: the exponential decay rule, which may mimic some characteristics of the structure of the cerebral cortex. In the neural network, each pixel of video frames is assigned to a neuron at the corresponding position. Pixel strength is characterized as the state of a neuron. The video frame is memorized after all neurons in the neural network have been trained to a stable state. A bionic mechanism utilizing the associative memory property of a bionic Hopfield neural network is proposed to track objects in video frames. The movement information of the pixel of the tracking object,

e.g., the orientation and the velocity, are obtained from the estimation of the state transition of neurons in the neural network. Additionally, the motion information of a subset of neurons is merged on the Gaussian distributed small-world topology to regress the motion information of local trunks of the target object, taking into account the collective behavioural properties of neighbouring neurons. The proposed method was implemented using the development tool in MATLAB R2016a. Numerical simulations are performed on a hardware platform with the Windows 11 operating system. The performance of the proposed method is examined based on its IoU (Intersection of Union), accuracy, robustness, computation time, and computational resources. Comparative tests were performed on five real-life data sets. These test results confirm the outstanding advantages of the proposed method in terms of accuracy and resource consumption.

CHAPTER 1

INTRODUCTION

This thesis presents a novel visual object tracking method based on a Gaussian distributed small-world Hopfield neural network. This chapter begins with an overview of artificial neural networks. In the following sections, the problem statement, research objectives, methodology, and organization of this work are presented separately.

1.1 Introduction

An artificial neural network is a nonlinear information system consisting of numerous interconnected processing units that, inspired by neuroscience, may process their information by mimicking the mechanism of the cerebral cortex. Each processing unit analogizes a neuron with an activation function for its output. The connection strength between each pair of processing units in the neural network that constitutes the synapse of neurons is called the weight. A learning mechanism that mimics the brain's memory function shapes the weight between two neurons. From a mathematical point of view, the results of the neural network trained by the bionic learning mechanism are commonly viewed as an approximation of a function. In 1982, John Hopfield introduced a neural network model in his research on content addressable memories. The Hopfield neural network uses a fully connected topology that organizes each of its neurons to connect all other neurons. Its activation function is utilized to determine the state of neurons according to a weighted sum of all connected neuron states. Since their invention, Hopfield neural networks have been widely used in many industries, such as image recognition (Nasrabadi et al., 1991), image enhancement (Paik et al., 1992), image change detection (Pajares, 2006), and combination optimization (Hopfield et al., 1985). In addition, in-depth studies were conducted by some front-line scientists, and remarkable improvements were made to Hopfield's neural networks. Abdullah (1992) proposed a new learning method by mining logic for a Hopfield neural network. Sathasivam and Abdullah (2011) extended this method with the reverse analysis method to induce logical rules in Hopfield neural networks. Inspired by the presidential election model, Sathasivam et al. (2020) further proposed a metaheuristic method to improve the learning performance of Hopfield neural networks. However, previous studies have shown that the fully connected topology used in Hopfield neural networks is biologically unrealistic (Braitenberg et al., 1998). Biological studies have observed that the real brain has a small-world topology.

The small-world network is a network named by analogy to the small-world phenomena in nature. The small-world network model was first proposed in 1998 by Duncan J. Watts and Steven H. Strogatz. Their small-world network model consists of two parts: the ring lattice and random rewiring. The ring lattice part forms a highly clustered structure in the network, and the random rewiring structure is formed by randomly rewiring the nodes with a probability P in the network. This new network model has attracted much attention from scientists. Sporns et al. (2004) and Bassett et al. (2006) showed in these studies that brain networks have a small-world network structure. Thereafter, this finding inspired extensive reports at the scientific frontiers. However, the current small-world network model has yet to consider some biological constraints, such as the exponential decay rule. In real biological situations, Ercsey-Ravasz et al., 2013 and Oldham et al., 2022 discovered a rule that the probability of two neurons being connected decreases exponentially with their distance. These biological rules have been shown to improve the performance of neural networks in practical applications (Pajares, 2006). In addition, scientific communities have extensively elucidated the stability of small-world Hopfield neural networks. Li and Chen (2003) examined the stability of their neural network model with small-world connection. Their results highlight the impact of the number of neurons and wiring probabilities on the stability of small-world neural networks. Sinha (2005) examined how small-world topological structures in networks affect their stability. The results prove the affection between increasing network structural complexity and the stability-instability transition. Yu et al., 2017 examined the neural network model with small-world connection. Their study discovered that a small wiring probability of inhibitory shortcuts helps maintain stability. Rüdiger et al. (2020) and Arvin et al. (2022) pointed out in their studies the stable properties of the small-world topology network: that long-range wirings dominate the state of the small-world topology, while short-range wirings shape the transmission dynamics for the small-world network. These results both suggest the potential contribution of long- and short-range wring distributions to the stability of small-world neural network models.

On the other hand, object tracking technology has been recognized as a successful practical application of artificial neural networks and has become an attractive research area. The main principle of the object tracking method is to build a prediction model for detecting the object in the subsequent video frames according to its data in the initial video frames. Object tracking methods are divided into two categories: generative methods and discriminative methods. In generative methods, the tracked object is determined by comparing the similarity between its candidates in the subsequent video frame and its predictions in the initial video frame. In discriminative methods, a decision function or classifier is designed to recognize and position the tracked object in the subsequent video frame using its information in the initial video frame. Both of these methods can be embodied in their respective neural networks. In practical applications, however, these methods still face challenges in terms of computational efficiency and resource requirements, which suggests that these methods may differ from how a human brain recognizes an object. Recent neuroscience studies have shown that artificial neural networks organized as real brain connectivity can perform more efficiently on complex recognition tasks (Suárez, 2021). Therefore, we proposed a simple Hopfield neural network with bionic brain topology to improve computational efficiency and resource requirements for current object tracking methods.

This study aims to develop a high-performance object tracking method by applying a brain-like topological neural network model. We complement a biological constraint to the small-world network to better mimic the characteristics of the cerebral cortex. As the main contribution of this work, a novel Gaussian distributed small-world Hopfield neural network model for visual object tracking is proposed. The proposed method stores the initial position of the tracked object in the small-world neural network model, and its motion path is predicted according to its position in subsequent frames.

The proposed neural network model was tested against five popular object tracking benchmarks. These datasets include OTB, VOT, UAV123, GOT-10K, and TC128. The OTB dataset is considered the most commonly used dataset in visual object tracking (Muller et al., 2018). The OTB dataset comprised 100 video sequences divided into 11 test scenarios. Li et al. (2019) tested their Target-Aware Deep Tracking based tracking method using the OTB dataset in the CVPR 2019 conference. Kiran et al. (2022) used the OTB dataset to test their proposed adaptive Siamese tracking method. The VOT dataset, also known as the VOT Challenge dataset, is utilized by the annual VOT Challenge contest at the International Conference on Computer Vision (ICCV) to compare the performance of various trackers. VOT dataset contains 90 video sequences divided into 6 test scenarios. In the ICCV 2018 VOT Challenge contest, George et al. (2018) presented the performance of their KCF tracker using the VOT dataset. Sun et al. (2020) used the VOT dataset as the benchmark for their tracker based the proposed fast template matching and updating techniques. Luo et al. (2018) use the VOT dataset for measure their reinforcement learning based tracker. The UAV123 dataset is professionally suited for visual object tracking in the aerial domain. It offers 123 aerial video sequences captured from the aerial at low altitudes. Bi et al. (2019) have used the UAV aerial dataset to tested their proposed enhanced MDNet tracking method. The SiamRpn++ method proposed by Li et al. in 2019 was also tested on the UAV123 dataset. Huang et al. (2019) tested their correlation filter methods on the UAV dataset for real-time tracking. The GOT-10K dataset is a collection of over 10,000 real-world video sequences, featuring 563 object categories. The dataset offers a significant amount of manually labelled ground truth data to aid in performance evaluation. Gou et al. (2021) tested their SiamGAT tracking algorithm on the GOT-10K dataset. Lan et al. extended performance testing for their proposed ProContEXT tracker on the GOT-10K dataset in 2023. Yu et al. (2021) used the GOT-10K dataset to test the proposed discriminative tracker, based on an encoder-decoder Transformer. The TC128 dataset contains 128 video sequences, especially for the color-enhanced tracker. Li et al. (2019) used theTC128 dataset for the Gradient-guided network method (GradNet). Kiani Galoogahi et al. (2017) proposed the Learning background-aware correlation filters method and tested it on the TC128 dataset. Zhou et al. tested their SiamCAN method on the TC128 dataset. Through the comparison test on the above five datasets, the test results showed that the proposed neural network algorithm has an advantage over existing object tracking algorithms in many performance metrics.

1.2 Research Gaps and Problem Statement

Neural networks that simulate brain mechanisms have been applied to visual object tracking, and some advances have been made. However, its computational efficiency and resource requirements still lag behind those of a real brain. Although little is known about the truth of object recognition in the natural brain, some biologists have observed in previous studies that the biological brain is a small-world network that is more sparsely structured than the fully connected structure in the current artificial neural network (Braitenberg et al., 1998). Some frontline studies have shown that using small-world networks as the topology can lead to a neural network model with more biological reality, higher computational efficiency, and reduced computational resource requirements (Calcraft, 2006). Recent neuroscience studies have also shown that artificial neural networks organized as real brain connectivity can perform more efficiently on complex recognition tasks (Suárez, 2021). However, the current small-world network has yet to include the biological constraints in the brain connectomes: the growth-cost rule and the exponential decay rule.

From an application point of view, object tracking is an attractive area that has attracted much attention. Traditional object tracking methods can be divided into two categories: generative methods, which determine the object by comparing the similarity of its candidates in the subsequent video frame and its predictions in the initial video frame, and discriminative methods, which train a classifier through a neural network to detect and position the tracked object. In current visual object tracking technologies, generative methods based on various filters (Wang et al., 2018;

Zhu et al., 2020; Moorthy et al., 2020) and discriminative methods based on deep learning and machine learning technologies (Yun et al., 2018; Dong et al., 2019; Bae et al., 2017) dominate the directions. Although generative methods have also provided many insights for applications in this field, discriminative methods based on machine learning and deep learning have achieved remarkable results and have become popular in recent years. These methods and models are computationally and computer engineering oriented, which may deviate from the original intention of neural networks to mimic the working mechanism of the natural brain to solve complex recognition problems. In addition, all of these methods face the problem of large amounts of computation and large resource requirements.

Some neuroscience-oriented models and algorithms have emerged from frontline studies. These models and algorithms mimic the brain's visual mechanisms at a deep level and therefore have unique technical advantages in terms of computational efficiency and resource requirements. Notable examples of existing brain-like models include the lateral inhibition model (Collier et al., 1996; Gregor et al., 2011), the receptive field model (Schwartz et al., 2012; Araujo et al., 2019), the synchronous pulse emission model (Maass et al.,2001; Yang et al., 2019), the visual attention model (Itti et al., 2001; Mnih et al., 2014; Wang et al., 2017), and the memory recognition model (Shiffrin et al., 1997; Dörfel., 2009). These brain-like models are abstracted from the biological characteristics of the natural brain and have been used for visual object tracking or solving other practical problems. It is noteworthy that current research on brain-like models is still in the early exploratory phase, and small-world brain connectivity that conforms to biological constraints has not yet emerged in addition to these existing brain-like models.

However, given the current state of the art of the above research areas, such additions of small-world networks in terms of biological constraints to form a more realistic, brain-like model for visual object tracking or other practical problems remain a research gap. Therefore, the achievements of the proposed model obtained in this study can be viewed as a new, viable biological remedy to solve highly complex computationally and resource-intensive problems for which new insights from artificial intelligence studies are currently lacking.

1.3 Significance of Research

This study presents our work in developing a brain-like, high-performance Hopfield neural network model for tracking objects in visual data. In this model, we add biological features to the small-world network model to make the small-world network more biologically realistic. This model can mimic some characteristics of the cerebral cortex and thus achieve higher computational performance when applied to visual object tracking. We tested this proposed model against several popular object-tracking benchmarks. The results showed that the proposed brain-like, small-world topology Hopfield neural network outperforms most traditional object tracking methods. The significance of this study is twofold.

First, a biological feature is added to the small-world network model that allows it to satisfy two constraints, namely, the growth cost rule and the exponential decay rule. This addition improves the computational performance of small-world Hopfield neural networks. In addition, it provides a technical example and a theoretical basis for creating more realistic, brain-like models.

Second, in this study, a visual object-tracking solution based on this new brain-like neural network mode will be developed. Such a bionic computational model, based on brain-like neural networks, has unique advantages in terms of computational efficiency and resource requirements compared to traditional visual tracking methods. The existing and traditional approaches compensate for computing performance problems simply by changing or improving hardware devices. Therefore, this study also provides new insights to address the weaknesses of these modern approaches to artificial intelligence in terms of excessive computational complexity and high resource demands.

1.4 Research Objectives

The main objective of this study is to develop a high-performance Hopfield neural network model to reduce the highly complex computational and resource requirement problem in visual object tracking applications. This objective inspires the development of a brain-like Hopfield neural network model that complements biological features for the small-world network and integrates it as a topology of Hopfield neural networks.

The research objectives of this study are listed as follows:

1. To integrate the biological features in the small-world network model and develop the small-world network model with Gaussian distributed wiring.
2. To integrate the Gaussian distributed wiring small-world topology into Hopfield neural networks.
3. To develop the change detection algorithm based on the Gaussian distributed wiring small-world Hopfield neural network.
4. To develop the object motion prediction model based on the collective neuron algorithm.

5. To implement the comparison experiments and results in five mainstream visual object tracking benchmarks.

1.5 Methodology

The proposed brain-like Hopfield neural network model for visual object tracking applications is divided into three stages: the brain network formation stage, the neuron stimulation feature detection stage, and the object motion prediction stage. In the brain network formation stage, two biological features are added to the small-world networks to mimic the natural brain network, resulting in a more realistic Gaussian distributed small-world network. The formed Gaussian distributed small-world network is used as the topology of the Hopfield neural networks. Then, to mimic

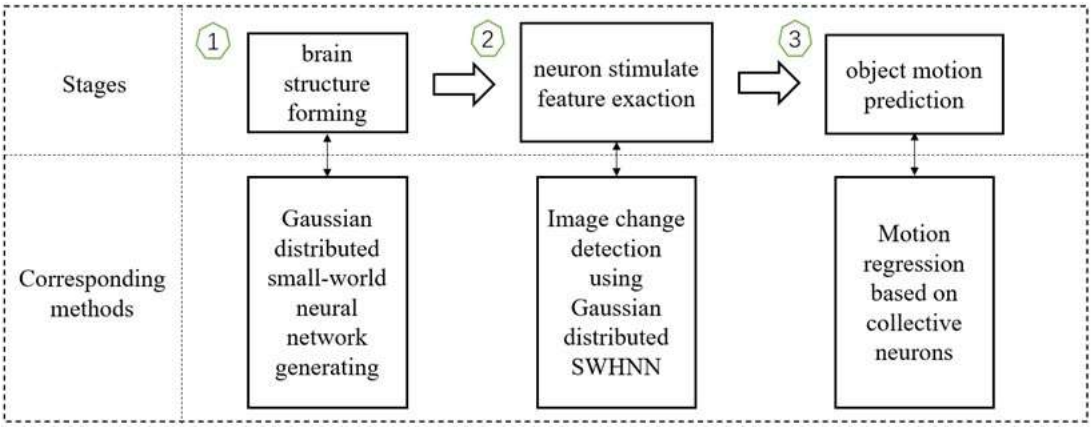


Figure 1.1 Stages of work of the Gaussian distributed small-world Hopfield neural network model for visual object tracking applications

the stimulation mechanism of the visual cortex in the brain, in the feature detection stage, neuron stimulation features are detected in the Gaussian distributed small-world Hopfield neural network. Based on the collective movement information contained in these stimulus features and their nearby neurons, in the object motion prediction stage, a regression model is established to predict the motions of the tracked object. The corresponding methodology of these stages is shown in Figure 1.1. Figure 1.1 shows a total of three methods for the relevant work stages.

The first method adds two biological rules, the exponential distance rule and the growing cost rule, into the small-world network model to mimic brain structure. To forms the Gaussian distributed small-world network, a Gaussian function is proposed to describe the distribution of the long-short range connections of the neurons in the n-dimensional lattice network.

The second method integrates the proposed Gaussian distributed small-world network as a topology of the Hopfield neural network. A simple signum function is employed to determine the neuron state of the proposed small-world Hopfield neural network. And the Hebbian learning method is integrated to train the neural network model. Based on the properties of the proposed brain-like neural network, the Gaussian distributed small-world Hopfield neural network is designed to detect those stimulation features (changes between the two frames) according to its neuron energy to be minimized.

The third method is developed to predict the movement trends of the tracked object using the movement trend information of these brain stimulation features. The developed motion prediction model regresses relevant stimulus features' motion trends in the subsequent frames using position information from two temporally adjacent video frames. Furthermore, to overcome the accuracy drawback of the regression model, we construct the motion prediction model based on collective neuron behaviours.

1.6 Organization of Thesis

This chapter first introduces some basic concepts of artificial neural networks and then describes the background of the proposed method for visual object tracking based on the Gaussian distributed small-world Hopfield neural network. Subsequently, relevant content such as problem statements, research significance, research objectives, and methodology was presented one after the other. These introductions paint the overall picture of this study. The remaining chapters are structured as follows.

Chapter 2 provides a literature review on the Gaussian distributed small-world Hopfield neural network for visual object tracking purposes. Consequently, this chapter is partitioned into two primary sections: the bionic small-world neural network model and the visual object tracking techniques. The bionic small-world neural network model section includes the discrete Hopfield neural network model, the small-world network models, and the brain networks. The visual object tracking techniques section includes the general architecture of visual object tracking techniques, the performance and testing criteria of object tracking techniques, the classifications of conventional visual object tracking methods, and five standard datasets.

Chapter 3 presents our main development work of the proposed Gaussian distributed small-world Hopfield neural network for visual object tracking. In Chapter 4, comparison experiments and tests are introduced. The test results are presented and discussed in this chapter. Chapter 5 concludes this paper, and related future work is presented in this chapter.

CHAPTER 2
LITERATURE REVIEW

This chapter provides a literature review of Hopfield neural networks with Gaussian distributed small-world topology, particularly their potential application in the field of visual object tracking. It is structured into two primary parts: the bionic small-world Hopfield neural networks and the visual object tracking technique. However, the present chapter commences by providing a brief review of discrete Hopfield neural networks. In light of the biologically unrealistic drawback of the fully connected topology in the Hopfield neural network, an alternative is introduced: the small-world topology observed in the real brain. Furthermore, this part introduces the brain network theory and the biological constraints that are yet to be considered in the small-world network model. The part devoted to the visual object tracking technique provides an overview of a general architecture, including its motion model, feature model, observation model, and the classification of traditional visual object tracking techniques. In order to enhance comprehension of performance testing, this part introduces some performance definitions, as well as five data sets: OTB, VOT, GOT-10K, UAV123, and TC-128, which are widely used in the field of computer vision.

2.1 Introduction to Bionic Small-World Hopfield Neural Networks

Hopfield neural networks are often considered to be an excellent biological computational model for imitating the memory function of the natural brain. Hopfield neural networks can also be viewed as a spin dynamics system where each of its neurons is connected to all other neurons. Such a connection establishes an information transmission path between a pair of neurons. The weight set on the connection characterizes the intensity of the synapse. Each neuron collects the influences of all other neurons, performs a weighted summation, and processes its state through an activation function and its preset threshold. The Lyapunov function defines its network energy. By minimizing network energy, neurons are propelled toward the minimal solution. After its invention, the Hopfield neural network found wide application in various fields due to its newly emerging associative memory function. However, some studies have also noted the computational complexity drawbacks caused by the fully connected structure of the Hopfield neural network. Meanwhile, some biological studies have shown that fully connected networks are biologically unrealistic. Several frontline studies of the brain's anatomical and functional brain networks have observed that the actual brain structure exhibits a small-world network.

The small-world network is a complex network named after the small-world phenomenon that has its origins in social networking research. Small-world networks are networks with high clustering and short path lengths. In 1998, Watts and Strogatz proposed the first small-world network model in their study. This small-world model starts with a regular ring lattice, where each node is connected to a certain number of nodes from a nearby region. Then, rewire half of the nodes to other randomly chosen nodes according to a rewiring probability. Based on this mechanism, a small-world network is then formed that can switch between a random network and a regular network. In 1999, Newman and Watts further proposed a new small-world network model. On the one hand, the NW small-world network model replaces the rewiring mechanism of the WS small-world network model with the random addition of wiring. On the other hand, it integrates the standard two-dimensional lattices into the small-world network model, allowing for a better and more accurate definition of its length scale and network properties.

Since the invention of small-world network models, small-world networks have attracted considerable attention in various scientific communities. Using advanced equipment and brain anatomy, some biologists have discovered that the actual structure of a brain is organized as a small-world network. In addition, top researchers have begun investigating details about the properties of small-world networks. Their results demonstrated the essential role of short- and long-range connections in enhancing the stability of the small-world network. Meanwhile, biological research has also taken note of the growth cost rule and the probability distribution of the short- and long-range connections between neurons in a natural brain environment, namely, the exponential distance rule. Unfortunately, however, the current small-world network has yet to address these biological limitations.

2.2 Discrete Hopfield Neural Networks

The Hopfield neural network is a feedback neural network invented by John Hopfield in 1982 as part of his research into content-addressable memory (Hopfield, 1982). Depending on the type of data to be processed, the Hopfield neural network can be divided into two categories. The continuous Hopfield neural network (CHNN) processes continuous data, and the discrete Hopfield neural network (DHNN) is a type of Hopfield neural network specialized in processing discrete data. Considering the binary data type used in the practical application of visual object tracking, we limit the scope of this study to DHNN. A detailed explanation of CHNN was not made. In a discrete Hopfield neural network, the state of the neuron is binary or bipolar and is usually denoted {0,1} or {-1,1}. Each neuron is connected to all other neurons and processes their influences in a weighted summation. An activation function, which is usually a signum function, and a preset threshold are used to determine the neuron state. Figure 2.1 shows a basic situation of a neuron in the discrete Hopfield neural network. In Figure 2.1, $x_{0,}x_1 \dots x_n$ represent the inputs of the neuron, and $w_{0,}w_1 \dots w_n$, represent

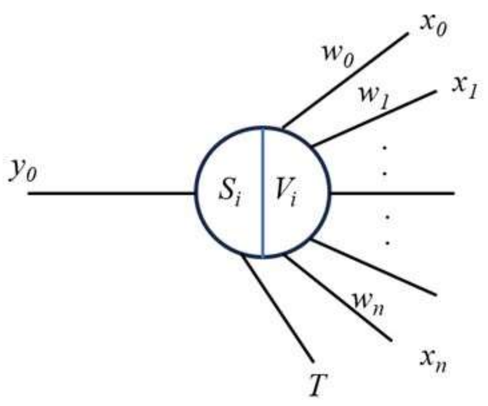


Figure 2.1 Basic situation of a neuron in the discrete Hopfield neural network

the weights of the relevant connections. y_0 is the output of the neuron. Each neuron processes a weighted summation for its inputs. Variable V_i is defined to store the intermediate result. V_i is calculated using Equation 2.1. Variable S_i denotes the state of the neuron. It is determined by its activation function, which in a discrete Hopfield

$$V_i = \sum_{i=1}^{n} x_i \cdot w_{i,j} \tag{2.1}$$

$$S_i = \begin{cases} 1, where\ V_i - T \geq 0 \\ 0, where\ V_i - T < 0 \end{cases} \tag{2.2}$$

neural network is usually a signum function, as defined in equation 2.2. In Equation 2.2, the variable T is the predefined threshold. If $V_i - T \geq 0$, then the neuron state is $S_i = 1$; otherwise, if $V_i - T < 0$, then the neuron state is $S_i = 0$. Then, an iteration is performed to update all neuron states in the network until all neurons remain in a stable state.

DHNN found wide application in practice after its invention. These include image recognition (Nasrabadi et al., 1991), image enhancement (Paik et al., 1992), image change detection (Pajares, 2006), and combination optimization (Hopfield et al., 1985). These applications have also enhanced DHNNs in various aspects. In particular, the emergence of new insights into the new learning method has significantly improved its performance.

2.2.1 Learning Methods of Discrete Hopfield Neural Networks

When DHNN was invented in 1982, John Hopfield employed the famous Hebbian learning method (Hebb, 1949) as the weight update rule in his neural network model. The learning mechanism of the discrete Hopfield neural network involves two phases, namely, the storage phase and the retrieval phase. In the storage phase, the weight updating rule is implemented in the following manner in equation 2.3 while learning m binary patterns. In equation 2.3, $w_{i,j}$ denotes the weight between neurons i

$$w_{i,j} = \sum_{s=1}^{m} (2V_i^s - 1)\,(2V_j^s - 1) \tag{2.3}$$

$$E(t) = -\frac{1}{2}\sum_{j}\sum_{i} w_{i,j}\, S_i(t)S_j(t) - \sum_{i} T_i S_i(t) \tag{2.4}$$

and j. V_i and V_j represent the states of neuron i and neuron j, respectively, and s is the number of patterns. It is worth noting that in Equation 2.3, $w_{i,j}$ will be positive if neuron i and neuron j have the same state, which would affect the value of neuron i, and neuron j tends to become equal. If, in turn, the state of neuron i differs from that of neuron j, the opposite occurs. During the storage phase, all network weights are updated until all network neurons remain in a stable state, which means that the state of the neurons remains unchanged after the previous round of weight updating. The pattern is then stored on the network. Its network energy function can be written as a Lyapunov energy function, as shown in Equation 2.4. In the retrieval phase, while inputting a pattern, the pattern stored in the neural network can be retrieved by computing Equation 2.1 and

Equation 2.2. Such a stable state of the neural network from the storage phase implements an associative memory function for the stored patterns. This newly emerged associative memory capability led to DHNNs becoming widely used in later decades. Some novel applications have inspired researchers to gain new insights to develop new learning methods for DHNNs.

In 1992, Abdullah (1992) initiated a research program to discover the logic rule in discrete Hopfield neural networks. In this research, Abdullah proposed a new learning method that provides new insights into the discrete Hopfield neural network. In 2011, Sathasivam and Abdullah (2011) further expanded this method and officially named it the Wan Abdullah method. The Wan Abdullah method processes the relationship of neurons as clauses in a CNF (conjunctive normal form). Equation 2.5 shows an example of a two-literal CNF representing the binary relationship of neurons i and j. To identify inconsistencies in the CNF, Equation 2.5 is negated by applying the De Morgen rule; therefore, Equation 2.5 is converted to Equation 2.6. The cost function under the environment with minimized logic inconsistencies can therefore be written as Equation 2.7. Note that in Equation 2.5, β represents a two-literal CNF. If all clauses

$$\beta = (S_i \vee S_j) \wedge (S_i \vee \neg S_j) \wedge (\neg S_i \vee S_j) \wedge (\neg S_i \vee \neg S_j) \wedge (T_i \vee T_j) \tag{2.5}$$

$$\begin{aligned} \neg\beta = {} & (\neg S_i \wedge \neg S_j) \vee (\neg S_i \wedge S_j) \vee (S_i \wedge \neg S_j) \vee (S_i \wedge S_j) \\ & \vee (\neg T_i \wedge \neg T_j) \end{aligned} \tag{2.6}$$

$$\begin{aligned} E_\beta = {} & \frac{1}{2}(1 - S_i)\frac{1}{2}(1 - S_j) + \frac{1}{2}(1 - S_i)\frac{1}{2}(1 + S_j) \\ & + \frac{1}{2}(1 + S_i)\frac{1}{2}(1 - S_j) + \frac{1}{2}(1 + S_i)\frac{1}{2}(1 + S_j) \\ & + \frac{1}{2}(1 - T_i)\frac{1}{2}(1 - T_j) \end{aligned} \tag{2.7}$$

Table 2.1 Weight update rule for the 2-Satisfiability Wan Abdullah method

	$-\frac{1}{2}W_{i,j}S_iS_j - W_jS_j - W_iS_i - \frac{1}{2}W_{T_i,T_j}T_iT_j - W_{T_i}T_i - W_{T_j}T_j$				
	$=\frac{1}{4}(1 - S_j - S_i + S_iS_j)$	$=\frac{1}{4}(1 + S_j - S_i - S_iS_j)$	$=\frac{1}{4}(1 - S_j + S_i - S_iS_j)$	$=\frac{1}{4}(1 + S_j + S_i + S_iS_j)$	$=\frac{1}{4}(1 - T_j - T_i + T_iT_j)$
w_i	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	0
w_j	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	0
$w_{i,j}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	0
W_{T_i}	0	0	0	0	$\frac{1}{4}$
W_{T_j}	0	0	0	0	$\frac{1}{4}$
W_{T_i,T_j}	0	0	0	0	$-\frac{1}{2}$

of the CNF are true, then β is satisfied or 2-satisfied. The weight update rule for the 2-satisfiability Wan Abdullah method is given in Table 2.1.

Due to its excellent performance, the Wan Abdullah method later attracted a large number of researchers. Mansor et al. (2016) extended the Wan Abdullah method to 3-satisfiability. Kasihmuddin et al. (2018) introduced the maximum k-satisfiability for logic programming in DHNN. Sathasivam et al. (2020) presented a new method that hybridized the election algorithm with random k-satisfiability in the DHNN. The results of these studies showed that the performance of DHNN was significantly improved. Aside from these improvements in the learning method, researchers have also pointed out the computational complexity drawbacks introduced by the fully connected topology in the DHNN. In the meantime, the fully connected topology has also proven unrealistic in biological environments.

2.2.2 Fully Connected Topology

A typical discrete Hopfield neural network model uses a network system composed of a series of binary neurons. In this network system, each neuron is organized in connection with all other neurons but not with itself, thus forming a fully connected network. Figure 2.2 shows an example of a fully connected Hopfield neural network with four neurons. The links in Figure 2.2 represent connections between neurons, and the weights on these connections satisfy the conditions $w_{i,j} = w_{j,i}$ and $w_{i,i} = w_{j,j} = 0$. Therefore, these weights can be placed in a symmetric and zero-diagonal matrix. When this neural network is to be trained, each neuron receives the influences of all three other neurons and then performs a weighted summation. Each weight on

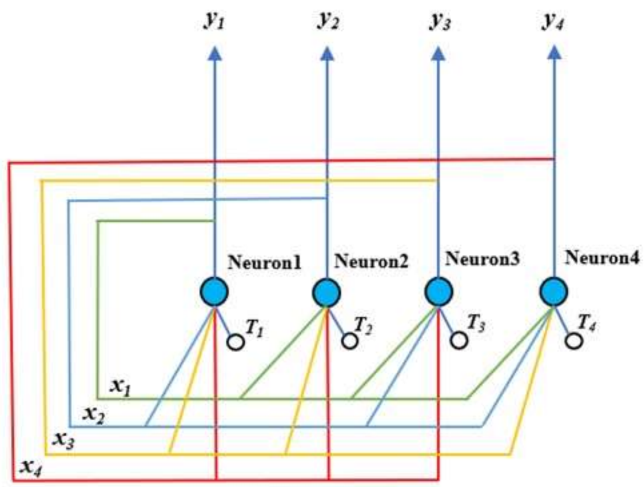


Figure 2.2 An example of a fully connected discrete Hopfield neural network with four neurons

these connections are updated until all neurons remain in a stable state. It is not difficult to observe that as the network size increases, the computational complexity of the network can lead to an insurmountable problem (Serpen, 2004). In practical applications, using visual data as an example, a 200×200 video frame would result in a neural network of 40,000 neurons generating ½ × 40,000 × (40,000 - 1) weights. As the network size increases, the resource and computational complexity of the fully connected topology leads to a complexity of $O(n^2)$ for the neural network model. This complexity issue has caused the fully connected topology to become a bottleneck that hinders DHNN in large applications. In addition, such a fully connected neural network is also biologically unrealistic.

Biological studies have shown that the natural brain topology is sparser than a fully connected topology (Braitenberg et al., 1998). Unlike in fully connected networks, each neuron in the brain network has only a few connections to other neurons. In 2004, Sporns et al. (2004) reported their observation of small-world network attributes in the cerebral cortex. Achard et al. (2007) further confirmed the small-world topology from brain anatomy in their research on functional brain networks. After this finding, small-world networks aroused great interest in the scientific community. Scientists evaluated the influence of topology on network memoirist function and measured varieties of topologies in terms of the performance of storage and retrieval of patterns (Bohland, 2001; Davey et al., 2004). Their results (Calcraft, 2006) show that small-world networks with several shortcuts can have the same efficiency as a random network. This important result further motivates the emergence of neural network models using the small-world topology (Duan et al., 2016). In addition, small-world networks have also made outstanding contributions in the areas of brain functional networks and relevant medical research. Stam et al. (2007), Sanz-Arigita et al. (2010) and Zhao et al. (2012) showed that Alzheimer's disease is characterized by a loss of small-world network characteristics in the brain network. Wang et al. (2009) and Liu et al. (2015) observed altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. Berman et al. (2016) proposed a method to modulate the small-world architecture of functional brain networks in Parkinson's disease with levodopa. Based on the above facts, we confirm the existence of small-world networks in the biological brain environment. However, current small-world network models do not yet account for biological constraints.

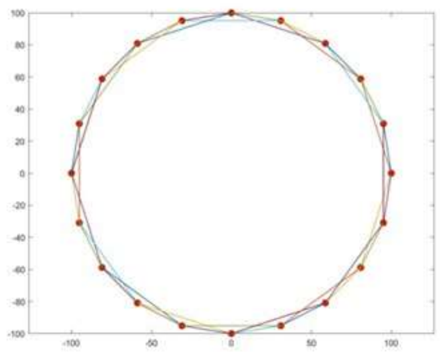
2.3 Small-World Network Models

A small-world network is a type of mathematical graph where most of the nodes are not connected to each other. However, each pair of nodes can reach each other via its neighbouring nodes in just a few steps. Such small-world networks explain a social phenomenon where any two people can be connected through only a few (six) acquaintances; if a node represents an individual, an edge between a pair of nodes denotes that they are acquainted with each other. In the social sciences, it is known as the “six-degree separation” theory (Milgram, 1967). After the advent of the six-degree separation theory, some scientists conducted a series of experiments to confirm the reality of small-world phenomena in real human society, such as the Erdős number experiment (Goffman, 1969) and the Bacon number experiment (Fass, 1996). Ultimately, however, all of these studies attest to the fact that despite the real-world human network having an enormous number of nodes, it takes only a few steps from one node to any other individual node. Examples of networks with such properties have appeared in other disciplines, such as biology, physics, and computer science, in addition to the social sciences. Recently, many empirical networks have been modelled by such networks, such as the World Wide Web, the transport network, the brain network, and the genetic network, all of which have the characteristics of such a small-world network.

2.3.1 WS Small-World Model

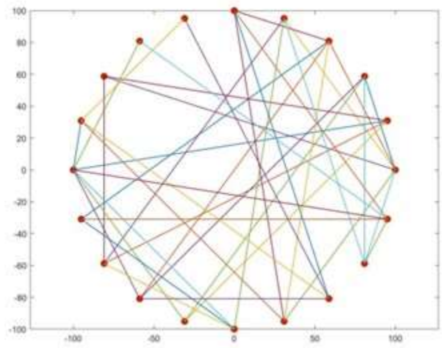
In 1998, Duncan J. Watts and Steven H. Strogatz proposed the small-world network model (Watts, 1998) based on a simple one-dimensional ring lattice. Watts and Strogatz describe their small-world network as a network model with a high cluster coefficient and a low average path length. Although the nodes in this network model are almost non-adjacent (connected) to each other, any one node in the network model can be accessed through only a few steps (nodes). Their small-world network is defined as a network where $L \infty \log N$, where L denotes the average path length of the network, and N denotes the number of nodes in the network. $L \infty \log N$ means L increases proportionally to $\log N$. Assume a node *i* that might connect with *k* nodes in the network and denoting *e* as the actual connections of node *i* and the maximum connection quantity can be written as $Max(e) = \frac{1}{2}\, k(k-1)$, representing node i being fully connected with the other nodes. Then, the clustering coefficient of node *i* is defined as $\frac{e}{\frac{1}{2}k(k-1)}$. The entire network clustering coefficient can be computed by averaging the total clustering coefficient of all nodes.

Later, Watts and Strogatz presented their method of forming a small-world network. Their method introduces two phases in building a small-world network: the initialization phase and the rewiring phase. The initialization phase suggests starting with a *n* nodes small-world network, and each node in the network is connected with *k* nodes from the near region to begin forming a high-centralization one-dimensional ring lattice. Denote N_i as each node of this network, and then in the rewiring phase, divide this node into left and right sides, with $\frac{k}{2}$ connections per side. *P* is the rewiring probability. Take its connections from the right side and rewire to a randomly chosen node by *P*, and stipulate that there is no self-loop. Perform this rewiring mechanism for each node in the network. Eventually, the WS small-world network is formed, and by adjusting the parameter P of the model, a small-world network can emerge between the regular network (P=0) and the random network (P=1).



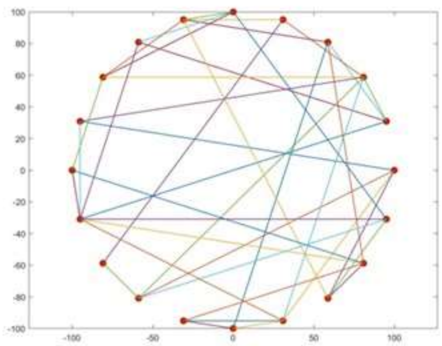
k =4, P = 0

(a)



k =4, P = 1

(b)



k =4, P = 0.5

(c)

Figure 2.3 An example of WS small-world networks

Figure 2.3 shows three networks that are formed with different values of the rewiring probability P. The case k = 4, P = 0 forms a regular ring lattice network, while k = 4, P = 1 generates a random network. The small-world network may be formed when P is in the interval from zero to one. In this case, a small-world network arises under k = 4, P = 0.5. This shows that by adjusting the wiring probability P, the small-world network can be obtained between the regular ring lattice network and random networks. However, such an underlying ring lattice-based network is relatively simple,