QUANTUM-INSPIRED ASSOCIATIVE MEMORIES FOR
INCORPORATING EMOTION IN A HUMANOID

NAOKI MASUYAMA

FACULTY OF COMPUTER SCIENCE AND INFORMATION
TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016

QUANTUM-INSPIRED ASSOCIATIVE MEMORIES FOR INCORPORATING EMOTION IN A HUMANOID

NAOKI MASUYAMA

THESIS SUBMITTED IN FULFILMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF COMPUTER SCIENCE AND INFORMATION
TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Naoki Masuyama

Registration/Matrix No.: WHA120044

Name of Degree: Doctor of Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Field of Study: Artificial Intelligence (Robotics)

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature   Date

Subscribed and solemnly declared before,

Witness's Signature   Date

Name:
Designation:

ABSTRACT

Associative memory is essential to human activity. In the past decades, several artificial neural associative memory models have been developed and were expected to provide a new perspective for the modeling of the human brain, and these models were expected to form the basis for a robot to exhibit human-like behavior. However, the conventional models suffered from limited abilities. In 2006, Rigatos and Tzafestas applied the concept of Quantum Mechanics to associative memory, and introduced Quantum-Inspired Hopfield Associative Memory (QHAM). Although QHAM showed outstanding potential and superiority, it is limited as an auto-association model with binary state information. With regards to events in the real world, information representation by a binary or bipolar state is insufficient. Therefore, the ability and functional improvements of associative memory based on quantum-inspired model are defined as the main objectives in this thesis. In regards to the ability improvements in terms of memory capacity and noise tolerance, the quantum-inspired hetero-association models with batch/incremental learning algorithm are developed based on QHAM. Furthermore, the quantum-inspired complex-valued hetero-association models are considered to accommodate the high dimensional problems. Based on the results of experiments, it is shown that the quantum-inspired hetero-association models have outstanding abilities. In regards to the functional improvements, the emotion affected association model is developed in an interactive robot system based on the relationship between memory and emotion from the viewpoint of psychology and neuroscience, which is called the mood-congruency effect. The experimental results show that the emotion affected association model is able to associate the emotion dependent information to the robot similar with the mood-congruency effect.

ABSTRAK

Associative Memory adalah penting untuk aktiviti seharian manusia. Beberapa dekad yang lalu, model Associative Memory telah dibangunkan dan dapat memberi perspektif baru bagi pemodelan otak manusia bagi membentuk tingkahlaku asas bagi robot untuk meniru tingkah laku manusia. Walau bagaimanapun, model konvensional mempuyai kebolehan yang terhad. Pada tahun 2006, Rigatos dan Tzafestas menggunakan konsep Quantum Mechanics kepada Associative Memory dan memperkenalkan Quantum-Inspired Hopfield Associative Memory (QHAM). Walaupun QHAM menunjukkan potensi yang cemerlang, ia terhad kepada model Auto-Association dengan maklumat didalam bentuk binari. Di dunia nyata, maklumat yang mewakili daripada bentuk binari ataupun bipolar adalah tidak mencukupi. Oleh itu, keupayaan dan peningkatan fungsi Associative Memory adalah berdasarkan model Quantum-Inspired yang digunapakai untuk dijadikan sebagai objektif utama dalam thesis ini. Peningkatan upaya dari segi kapasiti memori dan toleransi bunyi, model Quantum-Inspired Hetero-Association dibangunkan bedasarkan model QHAM. Selain itu, model Quantum-Inspired Complex-Valued Hetero-Association dianggap dapat menampung masalah dimensi yang tinggi. Berdasarkan keputusan eksperimen, model Quantum-Inspired Hetero-Association mempunyai kebolehan yang luar biasa. Berhubung dengan peningkatan fungsi, model Emotion Affected Association dibangunkan melalui sistem robot interaktif berdasarkan kepada hubungan antara memori dan emosi daripada pandangan psikologi dan neurosains, dimana ia dipanggil sebagai kesan keselarasan emosi. Keputusan eksperimen menunjukkan model Emotion Affected Association mampu untuk mempengaruhi maklumat pergantungan emosi kepada robot yang sama dengan kesan keselarasan emosi.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

E	:	an energy of particle.
H	:	Hamiltonian operator.
$I_{(t)}^{CA}$	:	the state of core affect.
$I_{(t)}^{E}$	:	the state of emotion.
$I_{(t)}^{MI}$	:	the multi-modal input.
$I_{(t)}^{M}$	:	the state of mood.
L	:	a number of layers in the association model.
M	:	a number of neurons in Y-layer.
N	:	a number of neurons in X-layer.
P_β	:	the effect of personality factors to Pleasure axis.
P_β	:	the effect of personality factors to Arousal axis.
P_γ	:	the effect of personality factors to Dominance axis.
Q	:	the state of an isolated quantum system.
Ve	:	an external potential.
$\Delta\alpha$	:	the effect of emotional effect to Pleasant axis.
$\Delta\beta$	:	the effect of emotional effect to Arousal axis.
Ω	:	the emotion appraisal model.
$\boldsymbol{S}$	:	a temporal states of associated pattern in X-layer.
$\boldsymbol{U}$	:	a temporal states of associated pattern in Y-layer.
$\boldsymbol{W}^T$	:	a transpose weight matrix of $\boldsymbol{W}$.
$\boldsymbol{W}$	:	a weight matrix.
$\boldsymbol{X}$	:	the original memory vectors in X-layer.
$\boldsymbol{Y}$	:	the original memory vectors in Y-layer.
δ	:	a general output parameter.
η	:	a learning parameter of incremental learning.
γ^M	:	a suppression ratio in emotional model.
$\hbar$	:	Plank's constant.
λ	:	relaxation factor.
$\left|\psi_{(x,t)}\right|^2$	:	a probability that the particle is at position x at time t.
μ, ν	:	a fuzzy membership grade.
ψ	:	a memory vector in quantum state.
ψ^*	:	a conjugate memory vector of ψ.
ξ	:	normalizing constant.
k	:	a number of original memory vectors.
p	:	a momentum of particle.

r	:	a quantization value on the complex unit circle.
$s_{(\alpha)}$	:	the orthonormalized memory vectors in α-th layer.
$s_{(\beta)}$	:	the orthonormalized memory vectors in β-th layer.
x	:	an element of original memory in X-layer.
y	:	an element of original memory in Y-layer.
AM	:	Associative Memory.
ANN	:	Artificial Neural Network.
BAM	:	Bidirectional Associative Memory.
CBAM	:	Complex-Valued Bidirectional Associative Memory.
CHAM	:	Complex-Valued Hopfield Associative Memory.
CI	:	Computational Intelligence.
CLM	:	Constrained Local Model.
CMAM	:	Complex-Valued Multi-directional Associative Memory.
DP	:	Dynamic Programming.
FAM	:	Fuzzy Associative Memory.
FS	:	Fuzzy System.
GA	:	Genetic Algorithm.
GBAM	:	General Bidirectional Associative Memory.
GNG	:	Growing Neural Gas.
HAM	:	Hopfield Associative Memory.
IBAM	:	Incremental Bidirectional Associative Memory.
IMAM	:	Incremental Multi-directional Associative Memory.
IQBAM	:	Incremental Quantum-Inspired Bidirectional Associative Memory.
IQMAM	:	Incremental Quantum-Inspired Multi-directional Associative Memory.
LBP	:	Local Binary Pattern.
MAM	:	Multi-directional Associative Memory.
MLNN	:	Multi-layered Neural Network.
NN	:	Neural Network.
PC	:	Personal Computer.
QBAM	:	Quantum-Inspired Bidirectional Associative Memory.
QCBAM	:	Quantum-Inspired Complex-Valued Bidirectional Associative Memory.
QCI	:	Quantum-Inspired Computational Intelligence.
QCMAM	:	Quantum-Inspired Complex-Valued Multi-directional Associative Memory.
QFS	:	Quantum-Inspired Fuzzy System.

QGA	:	Quantum-Inspired Genetic Algorithm.
QHAM	:	Quantum-Inspired Hopfield Associative Memory.
QL-BAM	:	Quick Learning Bidirectional Associative Memory.
QL-MAM	:	Quick Learning Multi-directional Associative Memory.
QM	:	Quantum Mechanics.
QMAM	:	Quantum-Inspired Multi-directional Associative Memory.
QNN	:	Quantum-Inspired Neural Network.
RHAM	:	Rotor Hopfield Associative Memory.
RNN	:	Recurrent Neural Network.
SNN	:	Spiking Neural Network.
SSGA	:	Steady-State Genetic Algorithm.
SVM	:	Support Vector Machine.
TGART	:	Topological Gaussian Adaptive Resonance Theory Algorithm.

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Associative Memory (AM) is regarded as a basis for human activity. In other words, human thought is always associative. Thus, the establishment and development of artificial neural associative memory is expected to provide a new perspective for the modeling of the human brain. Furthermore, it would be a basis for the robot to exhibit human-like behavior. In this chapter, AM as a human brain function and its computational models are introduced. Furthermore, the improvements in computational models of AM from an algorithm and the structural perspective are reviewed. Finally, based on the discussions and reviews, the motivations and approaches of this thesis are identified.

1.1 Associative Memory System in Brain

The brain is considered as a huge, complex, and sensitive system that is based on numerous synaptic connections called Neural Networks (NNs). Even today, the complete elucidation of the entire systems of the brain have not been explained. In past decades, neurologists have revealed that the brain has different functions such as abilities of vision, language, and memory depending on the specific regions of brain. With regard to memory, the memory system is fundamentally associative. In other words, when one acquire new knowledge, the brain will attempt to construct links with knowledge that has been previously acquired. Therefore, humans are able to associate the knowledge that is already firmly anchored in their brain with a piece of new information. This brain function is called AM. AM is defined as the ability to learn and recall the relationships between related/unrelated information. In particular, the association process can be classified into two types. If the input information and associated information are the same, it is regarded

as the auto-association. However, if the input information and associated information are different, it is considered as hetero-association. Presently, it has become clear that between the temporal lobe, the hippocampus and its neuronal connections, these hold the key factors for AM (Hirabayashi et al., 2013). Figure 1.1 shows roughly the positions of hippocampus and temporal lobe in brain.
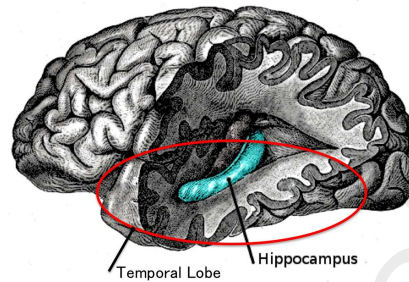
Figure 1.1: Hippocampus; located in the medial temporal lobe of the brain.

According to previous brain studies, declarative memory, a form of memory that is based on objective facts and personal experiences, is handled in the temporal lobe of the cerebrum. (Furthermore, the neuron group for long-term memory that stores the declarative memory was discovered in the cerebral temporal lobe of primates (Squire & Zola Morgan, 1991).)

Hippocampus, which is phylogenetically old cortex, belongs to the limbic system. It is not only related to instinctive behaviors, but also deeply involved with the functions of memory. In particular, recursive synaptic connections were found in an area of hippocampus called CA3. If CA3 is not functioning properly, a person cannot recall memory correctly (Scoville & Milner, 1957). This condition is called pattern completion. In addition, an interesting function of CA3 in the enhancement of long-term memory in collaboration with the temporal lobe has been reported (Nakazawa et al., 2002).

In the latest research results of the experiments conducted on brain activity, only the rudimentary and underlying understandings of the simple pattern of synaptic activities for memory systems have been obtained. The recursive synaptic connections of hippocampus

and the temporal lobe as a storage of long-term memory, however, are involved in the associative memory system.

1.2 Computational Approach of Brain Function

In the recent advances in computer technology, replicating the outstanding abilities of the brain in a computer has become an active research field. In particular, attempting to realize higher level brain functions in a computer, based on the analysis and modeling of the essential functions of a biological neuron and its complicated networks, is termed neurocomputing. In 1943, McCulloch and Pitts (1943) introduced a mathematical model for artificial neurons based on the signal transmission mechanism of neuronal connections in brain. Hebb (1949) introduced the hypothesis of the learning process for connections between artificial neurons based on the mechanism of neural plasticity, known as Hebb learning. Since then, several types of fundamental Artificial Neural Networks (ANNs), with connective strength between artificial neurons defined by Hebb learning, have been developed. In general, the structures of ANN are mainly divided into two types of models; one is the multi-layered neural network model which is represented by Perceptron, the other Recurrent Neural Network (RNN) model is represented by Association (Nakano, 1972). In general, artificial neural associative memory model is represented in the latter network model.

1.2.1 Artificial Neural Associative Memory Model

Hopfield (1982) introduced a fundamental recurrent network model called Hopfield Associative Memory (HAM). This model works as a content-addressable memory system with bipolar threshold nodes, and its learning process is defined by Hebb learning. HAM is key to understanding the human memory system. HAM belongs to the artificial neural associative memory models. In this model, it has been pointed out that a relationship with CA3 of hippocampus is involved in the associative memory and cerebral cortex of brain.

Furthermore, mathematically, it is guaranteed to converge to a local minimum. Although HAM is an interesting and significant model in terms of human memory function, the model has several disadvantages due to the structural design of the model. In general, the associative functions of the human brain are classified into two types of operations, and they are referred to as auto-association and hetero-association. Auto-association is able to associate the stored information from its incomplete information. Hetero-association is able to associate several types of stored information from its paired information. Because HAM is composed of a single directional layer, it is limited to auto-association. In order to overcome the above mentioned problem, Kosko and Hagiwara developed Bidirectional Associative Memory (BAM) (Kosko, 1987b) and Multi-directional Associative Memory (MAM) (Hagiwara, 1990), respectively. BAM consists of two layers of neurons, which are fully connected to each other, whereas MAM consists of fully connected multi layers of neurons.

The connections of neuron in the fundamental models of AM, namely HAM, BAM and MAM, are defined by correlation learning as Hebb learning. Thus, these models suffer from low memory capacity and recall reliability. In the past, several types of improvements have been introduced to increase the capabilities of the model. As examples of the structural improvements of model, dummy neurons (Y. F. Wang et al., 1991) and hidden layers (Kang, 1994) have been added to the model. These methods are simple but effective improvements. From an algorithmic improvement point of view, a projection matrix based on the least mean squared error minimization (Kohonen, 1972; Personnaz et al., 1985; C. Leung, 1993) is an effective method. However, this approach is not a local process. With regards to other approaches, several models have shown superior abilities by integration with effective learning algorithms; for example, Hattori et al. (1994) applied Hebb and pseudo relaxation learning in two stages learning to MAM. Zhuang et al. (1993) integrated a householder encoding. To obtain the optimal weight matrices for

efficient recall of any prototype input pattern, the population generation technique, the crossover operator, and the fitness evaluation function as genetic algorithm are applied to calculate the weight matrices (Kumar & Singh, 2010). Other models also use the non-binary neurons in association model for improving memory capacity and noise tolerance (Salavati et al., 2014). In addition, the effectiveness of sparseness, which is based on human brain modeling in the neural associative memory model, has been recently discussed and summarized (Palm, 2013). Furthermore, based on various physiological experiments, chaotic behavior of the real neurons has been observed. In particular, it is considered that chaos plays an important role in memory and learning of human brain. Several researches have introduced a chaotic associative memory to imitate human brain functions (Adachi & Aihara, 1997; He et al., 2008). These models show more superior abilities than fundamental models in terms of memory capacity, noise tolerance, and reduction of spurious attractors, however, the complexity of the model structure is greatly increased.

From the point of view of learning algorithm of AM, typical models use an offline, one-shot learning rule in Hebb learning with a non-linear output function (Hopfield, 1982; Kosko, 1987b; Hagiwara, 1990; Rigatos & Tzafestas, 2006). Conventionally, several types of iterative learning algorithms have been introduced (Xu et al., 1994; Oh & Kothari, 1994; Shi et al., 1998; Storkey & Valabregue, 1999; Eom et al., 2002). Learning algorithms based on time difference Hebb association (Storkey & Valabregue, 1999), and iterative relaxation learning (Oh & Kothari, 1994) are simple, but effective approaches. However, these models are not online. Chartier and Boukadoum proposed a self-convergent iterative learning algorithm with a nonlinear output function applied to BAM and MAM, called Incremental Bidirectional Associative Memory (IBAM) (Chartier & Boukadoum, 2006a, 2006b; Chartier et al., 2008) and Incremental Multi-directional Associative Memory (IMAM) (Chartier & Boukadoum, 2011), respectively, which can learn online without being subject to overlearning. In addition, it does not require a batch

process, or a reference to the previous state. Self-convergent iterative learning causes the network to progressively develop a resonance state between input and output. As a result, iterative online learning is performed by incorporating explicit learning in the output. These algorithmic variations are fundamentally important in modern systems such as big data processing.

1.2.2 Approach for Higher Dimension Problem by Complex Domain

As mentioned in the previous section, numerous improvements have been introduced, and the capabilities of model have improved tremendously. However, these models are only able to handle binary or bipolar states because the models are based on the McCulloch and Pitts neuron model. With regards to the real world events, information representation using binary or bipolar state is insufficient. For example, by comparing Figures 1.2b and 1.2c, it is apparent that Figure 1.2b has more information loss than Figure 1.2c as compared to the original image (Figure 1.2a).

(a) original image

(b) binary image

(c) grayscale image

Figure 1.2: Information expressive power of images.

In the 1973, the concept of complex-valued neuron model was first introduced by Aizenberg (1973) based on the McCulloch and Pitts neuron model. One of the signifi-

cant advantages in a complex-valued neuron is that it can handle higher dimensions as compared to a real-valued neuron. The fundamentals of the multi-state neural networks was introduced by Noest (1988b). Jankowski et al. (1996) introduced a Complex-Valued Hopfield Associative Memory (CHAM) wherein neurons are processed by a complex-valued discrete activation function. Subsequently, Lee et al. (1998) introduced a Complex-Valued Bidirectional Associative Memory (CBAM), and Kobayashi et al. (2005) implemented a Complex-Valued Multi-directional Associative Memory (CMAM). Similar to the features in the real-valued model, several improvements in the form of extension to the complex domains have been introduced to the complex-valued model. Lee (2006) argued that the real-valued projection rule can be generalized to the complex domain such that the weight matrix can be designed by using a simple and effective method. Hattori et al. (1994) introduced the two-stage learning algorithm called Quick Learning where Hebb learning and pseudo relaxation learning are applied to the learning process. Based on real-valued Quick Learning, Kobayashi (2008) developed a complex-valued Quick Learning algorithm. Furthermore, chaotic behavior is also integrated into the complex-valued neuron model while maintaining the properties of chaos. Several researches introduced the Complex-Valued Chaotic Associative Memories, and reported its superior abilities (R. S. Lee, 2004; Chakravarthy et al., 2008; Shimizu & Osana, 2010).

1.3   Quantum-Inspired Computing

In recent years, an interesting research area, which links the concept of quantum information and computer science, called Quantum-Inspired Computing, was introduced. In the past decades, Quantum Mechanics (QM) has been developed as a theory to explain the fundamental principles of substance. QM provides several mathematical concepts, such as duality of waves and particles, complementarity and non-locality, to improve the comprehension of the micro world. These new findings have affected the field of sci-

ence and technology. Quantum concepts have been applied to typical intelligence such as ANNs, Genetic Algorithms (GAs) and Fuzzy Systems (FSs) in the field of Computational Intelligence (CI), namely, Quantum-Inspired Neural Networks (QNNs), Quantum-Inspired Genetic Algorithms (QGAs) and Quantum-Inspired Fuzzy Systems (QFSs) in Quantum-Inspired Computational Intelligence (QCI). In fact, several studies have reported that the performances of QNNs, QGAs and QFSs have improved dramatically compared to conventional models because of the application of quantum mathematics (Li et al., 2013; Zhang, 2011; Rigatos & Tzafestas, 2002; Manju & Nigam, 2014).

In terms of ANNs, Perus (1996) hypothesized that QM is based on mesoscopic features in the physical and biological or physiological processes of the brain. Furthermore, the unification of the neuronal process and quantum process have been attempted based on the above assumption. In the internal structure of neuron, the presence of the two quantum states in tubulin, which are proteins of the size 4nm×8nm and having a 20nm gap between the synapse, suggest that ANNs would be handled as a descriptive subject of QM. In general, however, QNN is based on the linear similarity within the descriptions of the neuronal and quantum states. Often, the discussion of QNN is based on the similarity of the concept of QM. Thus, in the current situation, it should be noted that QNN is not derived from the QM in the physics field.

Conventionally, several studies have shown the potentiality and superiority of QNNs (Levy & McGill, 1993; Gandhi et al., 2014; Chen et al., 2014). It can be noted that the application of the concept of QM is a successful approach of algorithmic improvement in ANNs. With regard to AM, Quantum-Inspired Hopfield Associative Memory (QHAM) proposed by Rigatos and Tzafestas (2006) is considered as one of the superior models. This model demonstrates that quantum information processing in neural structures results in an exponential increase in the storage capacity with simple algorithms and ability to explain the extensive memory and inference capabilities of humans. This model applies

fuzzy inference to weight matrix to satisfy the features of QM, referred to as parallelism and unitarity. However, QHAM is limited to auto-association and binary/bipolar processing because of its neuronal structure.

As mentioned in this section, QCI has significantly improved the capabilities of conventional algorithms. Interestingly, it has the potential to illustrate the dynamics of neurons in the human brain when viewed from a QM perspective. It is expected that the hetero-association model and complex domain model in AM based on concept of QM will exhibit superior abilities in terms of memory capacity and noise tolerance. Therefore, this thesis will focus on the improvements of quantum-inspired hetero-association models based on QHAM, which are regarded as improvements in ability aspects.

1.4   Role of Associative Memory in Communication

Communication is a fundamental basis of human activity. In the past decades, psychologists have defined and discussed human mental activities in the field of communication, such as in sociology, developmental psychology, relevance theory, and embodied cognitive science (Pfeifer & Scheier, 1999; Eysenck, 1998; Gregory, 1998; Minsky, 1986; Sperber & Wilson, 1995; Asada et al., 2009). In the "society of mind" theory, Minsky (1986) argues that human thoughts are not simply transmitted, but it is something shared between humans. In other words, when we communicate with other humans, we try to create and share our personal cognition, called cognitive environment (Yorita & Kubota, 2011). Multi-modal information, such as gesture, utterance, and facial expression, play an important role in focusing the attention to the subjects and sharing of our cognitive environment in human–human communication. Smooth communication enables us to share our cognitive environments with each other, and is therefore an important ability. In addition, relevance and continuity of the subjects are important in enlarging a cognitive environment (Sperber & Wilson, 1995). Therefore, in various communication scenes,

AM is a useful and significant brain function for handling relevance and placing the subject in a continuous state (Foss & Harwood, 1975; J. R. Anderson & Bower, 2014).

Meanwhile, in communication between humans and robot, the above discussion can be applied to human–robot communication given that the robot is characterized with human-like cognitive abilities. Conventionally, in the researches on human-robot communication, the robot provides the conversational subjects based on a history of conversations with humans, or human behaviors. As an alternative approach, the robot can also select general topics such as the weather and news that are available from the Internet (K. Zheng et al., 2013; Toris et al., 2014). In these approaches, there are possibilities that the relevance and continuity of the subject cannot be maintained. Thus, if the robot is able to perform multi-modal communication with AM under a shared cognitive environment, then it is able to provide suitable information in any situation. As a result, communication between the robot and the human would be more active.

1.5 Emotional Effects in Associative Memory

In artificial neural associative memories, numerous improvements have been introduced with reference to abilities, such as memory capacity and recall reliability. However, these models are independent from the other functions of the human brain. Incidentally, during a decision making process in human-human communication, the brain will not only be affected by the logical thinking factors, but also by emotional factors (Bechara et al., 2000). The importance of interaction between memory and emotion has been presented in (Reisberg & Hertel, 2004; LeDoux, 1994; Smith & Petty, 1995), with interesting mutual relationships. For instance, memory can be recalled by emotional information. Similarly, emotional information can be recalled from memory. In addition, it is possible to recall more clearly information that is memorized with emotional factors. In general, during positive feelings, we have the tendency to recall positive things, based on our personal

experiences and knowledge. Similarly, during negative feelings, negative things are likely to be recalled. Thus, as we feel both positive and negative feelings, it becomes easier to obtain positive and negative information. This psychological phenomena is known as the mood-congruency effect (Ekkekakis, 2013).

With regards to associative memory, the association relationships between stored memories can be represented by nodes which represent the information, and its connective relationships. As an example, Figure 1.3a shows the association map of Red Ball which is defined by several nodes and its connections. In general, the connections between nodes are different depending on individuals, and it will be changed based on personal experience and feelings. In Figure 1.3a, all nodes which are connected with Red Ball have the possibility to associate with the equal probability. Meanwhile, the association relationships between stored memories that are affected with the mood-congruency effect can be considered as Figure 1.3b. Here, each node is labeled to the positive, negative or neutral attribution based on experience, memory and feelings of individuals. Therefore, due to the mood-congruency effect, the node that has same emotional attribution with individual emotional information during association has the tendency to be associated. For instance, the person who has the knowledge as Figure 1.3b, and he/she has the positive emotional information, the nodes of Apple and Fruits are more likely to be associated than the nodes of Blood and Knife.

(a) without emotional factors   (b) with emotional factors
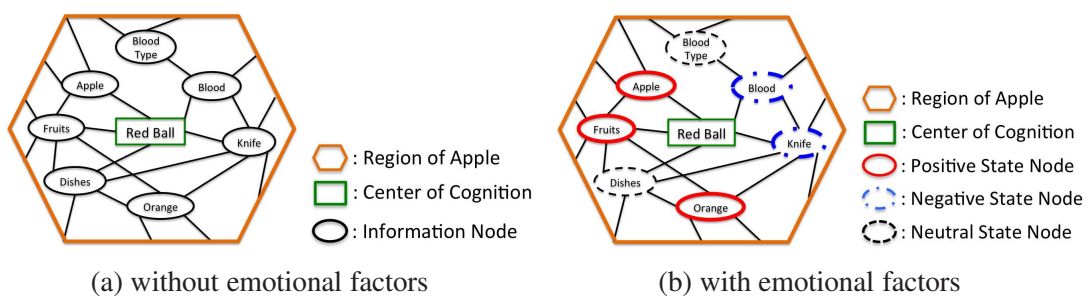
Figure 1.3: Region for Apple in association map; (a) without emotional factors, (b) with emotional factors.

This thesis intends to bring the above significant features between emotional effects and AM together into an integrated model. Furthermore, to evaluate the effectiveness of emotion affected AM, an interactive robot system with the integrated model will be developed. This integrated model is regarded as one of the functional improvements of AM.

1.6 Research Problems

From the discussions of biological and computational viewpoints, associative memory is regarded as a significant function of the human brain, and its ability is a fundamental activity of human beings. Therefore, the realization of artificial associative memory might bring new perspective to human brain modeling, and the emerging improvements of artificial intelligence in neurocomputing. Although several improvements have been introduced, conventional artificial associative memory models have several problems. Thus, the problems of this thesis are defined as follows;

Problem 1 :

The real-/complex-valued artificial associative memory models suffer from low memory capacity and noise tolerance.

Problem 2 :

Most associative memory models employ batch learning algorithm.

Problem 3 :

The real-valued model is not efficient for high dimensional problems.

1.7 Research Objectives

Several conventional studies consider QCI as an effective performance improvement approach of CIs. However, the discussion and studies pertaining to artificial associative memory models in QCI are insufficient when compared with the conventional association models in CI. The main objective of this thesis is to improve the artificial associative memory in terms of model abilities by developing QHAM (Rigatos & Tzafestas, 2006). In particular, the objectives are defined as follows;

Objective 1 :

> To improve the abilities of AM in terms of memory capacity and noise tolerance in keeping with stability, and to confirm it by simulation experiment.

Objective 2 :

> To develop the hetero-association models with the batch and incremental learning algorithms.

Objective 3 :

> To develop the complex-valued association model to deal with high dimensional problems.

The research objectives 1, 2 and 3 correspond to research problems 1, 2 and 3, respectively. The achievement of these research objectives will be confirmed by comparison experiments with conventional models under several conditions. Furthermore, the implementation of the emotion affected associative memory model based on the mood-congruency effect is considered for the functional improvement of AM.

1.8 Research Contributions

In this thesis, based on QHAM which is one of the most successful approaches of AM in QCI, real-valued quantum-inspired bi/multi-directional associative memory models with batch learning are presented as Quantum-Inspired Bidirectional Associative Memory (QBAM) and Quantum-Inspired Multi-directional Associative Memory (QMAM). Furthermore, based on QBAM and QMAM, self-convergent iterative incremental learning models are introduced as Incremental Quantum-Inspired Bidirectional Associative Memory (IQBAM) and Incremental Quantum-Inspired Multi-directional Associative Memory (IQMAM), respectively. The mathematical proofs and numerical examples confirm that the features of QM, namely superposition and unitarity, are maintained in the models. The details of the above topics will be presented in Chapter 3. In summary, the contributions of real-valued model are characterized as follows;

- QBAM and QMAM

    - Improvement of their memory capacity and noise tolerance; and
    - Realization of hetero-association.

- IQBAM and IQMAM

    - Improvement of their memory capacity and noise tolerance;
    - Realization of hetero-association; and
    - Implementation of incremental learning.

With regards to the corresponding model in the high dimensional problems, in relation with real-valued models as QBAM and QMAM, quantum-inspired complex-valued bi/multi-directional associative memory models called Quantum-Inspired Complex-Valued Bidirectional Associative Memory (QCBAM) and QCMAM are introduced. The details of the complex-valued models will be presented in Chapter 4. In summary, the contributions of complex-valued model are characterized as follows;

- QCBAM and QCMAM

  - Improvement of their memory capacity and noise tolerance;

  - Realization of hetero-association; and

  - Ability to handle higher dimensional problems.

The interest models of this thesis are shown in Table 1.1.

In the critical discussion involving the relationships between memory and emotion in the brain from a psychology and a neuroscience perspective, the significance of emotional effects on memory in communication, which is referred to as the mood-congruency effect, is revealed. In this thesis, the emotion affected associative memory model consisting of QCMAM and emotional model is introduced. The realization of emotion affected associative memory will enhance the capability of associative memory model for humanoid.

Table 1.1: Models of interest in thesis.

Model		Auto-Association	Hetero-Association	
			Bidirectional	Multi-directional
Basic Model	One-Shot Learning	HAM (Hopfield, 1982)	BAM (Kosko, 1987a)	MAM (Hagiwara, 1990)
		CHAM (Jankowski et al., 1996)	CBAM (Donq & Wen, 1998)	CMAM (Kobayashi & Yamazaki, 2005)
	Iterative Online Learning	—	IBAM (Chartier & Boukadoum, 2006a)	IMAM (Chartier & Boukadoum, 2011)
Quantum-Inspired Model	One-Shot Learning	QHAM (Rigatos & Tzafestas, 2006)	Quantum-Inspired Bidirectional Associative Memory (QBAM)	Quantum-Inspired Multidirectional Associative Memory (QMAM)
		—	Quantum-Inspired Complex-Valued Bidirectional Associative Memory (QCBAM)	Quantum-Inspired Complex-Valued Multidirectional Associative Memory (QCMAM)
	Iterative Online Learning	—	Incremental Quantum-Inspired Bidirectional Associative Memory (IQBAM)	Incremental Quantum-Inspired Multidirectional Associative Memory (IQMAM)

1.9 Outline

This thesis is organized into four parts; Introduction and Literature Review (Chapters 1 and 2), Models (Chapters 3 and 4), Interactive Robot System (Chapter 5), and Conclusion and Future Works (Chapter 6).

Chapter 2 reviews the fundamentals of artificial neuron model and structure of ANNs, including HAM, BAM, and MAM as a basis of AM. Furthermore, studies of the interactive robot system with AM that takes into account the external stimulus for association are also provided. Chapter 2 presents the reader with useful background knowledge about improvements in AM.

Chapter 3 formally defines the structures of QBAM, QMAM, IQBAM, and IQMAM. Furthermore, detailed mathematical proofs for satisfying the features of QM and its numerical example are presented.

Chapter 4 introduces the structure of QCBAM and QCMAM using the concept of complex domain. Additionally, Chapter 4 presents the mathematical proofs and a numerical example for the complex domain model that satisfies with the rationale of QM.

Chapter 5 describes the configurations of an interactive robot system with the emotion affected associative memory. In addition, simulation results for evaluating the capability of the emotion affected associative memory are presented.

Chapter 6 presents the conclusions, contributions and future works.

CHAPTER 2

LITERATURE REVIEW

In the past decades, numerous types of "Artificial Intelligence" models have been proposed in the field of computer science in order to realize the human-like abilities in a humanoid. In particular, attempting to realize the higher level brain functions in a computer, based on the analysis and modeling of essential functions of a biological neuron and its complicated networks, is called neurocomputing. It has been noted that one of the interesting and challenging subjects is the imitation of memory function of the human brain. In general, the above function is handled as RNNs, which is a part of ANNs in the computer science field. This chapter presents the fundamentals of neurocomputing, namely, the artificial neuron model and its network models such as Perceptron and RNNs. Studies of RNN are specifically described in detail. Useful background knowledge about improvements of AM is also presented. Furthermore, studies on the interactive robot system with AM will be reviewed.

2.1 Neuron

2.1.1 Biological Perspective

The biological nervous system is composed of numerous neurons, which perform parallel information processing using a complex combination of neurons. A neuron comprises a dendrite as an input terminal, axon as an output terminal, and soma as illustrated in Figure 2.1.

The dendrite of the neuron receives an input from neurotransmitters via synapse from the axon of other neurons. This input is a weak electrical signal that is classified as excitatory, which increases voltage and inhibitory, which decreases voltage. The receiver
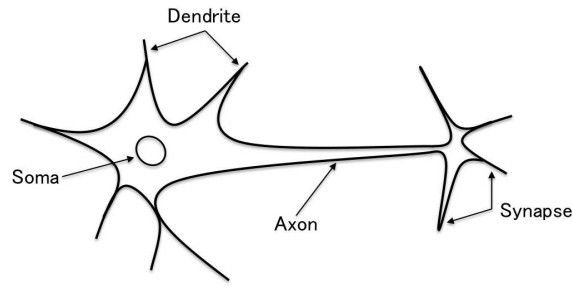
Figure 2.1: Neuron – an electrically excitable cell.

side of the neuron accepts the positive or negative weighted signal from the connected neurons. Once the sum of the weighted signal has reached a certain threshold, the receiver side of neuron is fired, thereby influencing the other neurons that are connected by the synapse of fired neuron.

2.1.2 Artificial Model

In 1943, W.S. McCulloch and W. Pitts (1943) introduced the artificial neuron model based on the biological perspective in Figure 2.2.
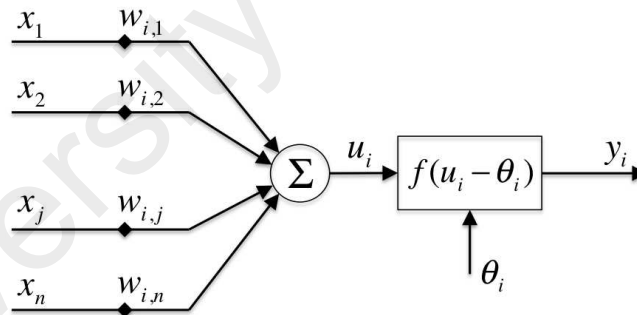
Figure 2.2: Artificial neuron model.

Let the input signal from neuron j as x_j, the connective weight between neurons j and i as $w_{i,j}$, and threshold for firing as θ_i, the sum of the inputs u_i to neuron i is as following;

$$u_i = \sum_{j=1}^{n} w_{i,j} x_j. \tag{2.1}$$

$$y_i = f(u_i - \theta_i) = f\left(\sum_{j=1}^{n} w_{i,j} x_j - \theta_i\right). \tag{2.2}$$

19

where, $f(\cdot)$ denotes an activation function for firing the neuron. In general, a step function or a sigmoid function is applied depending on network systems.

(a) Step Function   (b) Sigmoid Function
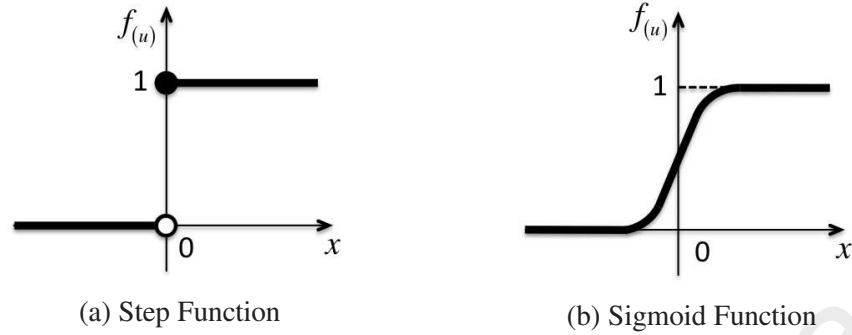
Figure 2.3: Activation function.

A step function is a classic mathematical model shown as Figure 2.3a. The output y_i based on the sum of the inputs u_i, and threshold θ_i is given as;

$$y_i = f(u_i - \theta_i) = \begin{cases} 1, & \text{If } \left(\sum_{j=1}^{n} w_{i,j} x_j - \theta_i \right) \geq 0 \\ 0, & \text{If } \left(\sum_{j=1}^{n} w_{i,j} x_j - \theta_i \right) < 0 \end{cases}. \tag{2.3}$$

It is clear that a step function is applied to the discrete time problem. In RNNs, a step function with bipolar state is often used as an activation function, whereas, the sigmoid function in Figure 2.3b is applied as a differentiable continuous function to the continuous time problem. The output y_i is given as;

$$y_i = f(u_i - \theta_i) = \frac{1}{1 + \exp\{-(u_i - \theta_i)\}}. \tag{2.4}$$

2.1.3 Hebb Learning

Hebb (1949) proposed the neurobiological hypothesis of signal transmission in neurons. The hypothesis can be stated as, "*When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing*

B, is increased." In short, if neurons are fired simultaneously, the synaptic connections between neurons will intensify. This is known as Hebb learning which is the basis of learning theory in NN. Here, Figure 2.2 assumes the presence or absence of fired neurons expressed as $x_j, y_i \in \{0, 1\}$, and the connection coefficient from neuron y_i to neuron x_j as $w_{i,j}$. The Hebb learning can be formulated as follows;

$$\Delta w_{i,j} = \lambda x_j y_i. \tag{2.5}$$

where, λ $(0 < \lambda \leq 1)$ denotes learning coefficient. Thus, when the neuron y_i is fired, the connection coefficient is updated as following;

$$w_{i,j} \leftarrow w_{i,j} + \lambda x_j. \tag{2.6}$$

In Hebb learning, the connection coefficient is updated by proportion to the firing frequency of the neurons.

2.2 Perceptron

Rosenblatt (1958) introduced an Multi-layered Neural Networks (MLNNs) called Perceptron, based on McCulloch-Pitts neuron model and Hebb rule as a learning function, and applied it to the pattern recognition problem. Perceptron is a 3-layer hierarchical NN that consists of a sensory unit (*S*-unit), associative unit (*A*-unit), and response unit (*R*-unit). In particular, it is a 3-layer hierarchical feed-forward network that satisfies only one neuron in *R*-unit, and there are no connections between neurons in the same unit. This network is called a Simple Perceptron, as shown in Figure 2.4.

The network will perform the following procedures; when the pattern to be classified is inputted into the network, *S*-unit will activated. Then, *A*-unit is activated as output from *S*-unit, and outputs the signals to *R*-unit. Finally, *R*-unit is activated by the output from
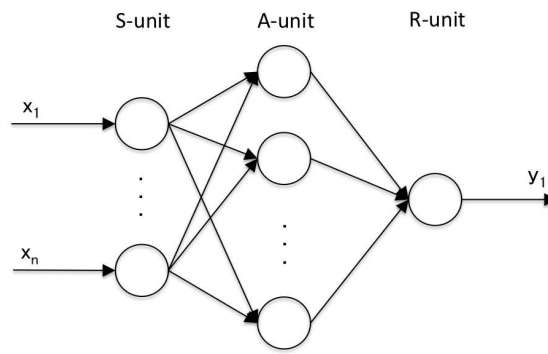
Figure 2.4: Simple Perceptron.

A-unit, and subsequently outputs the classified signal corresponding to the input pattern. The learning of Perceptron converges the weight connections by calculating iteratively until the output of *A*-unit reaches the condition of linear separation.

However, Perceptron is unable to solve the linear inseparable classification problem. For example, "exclusive-or" (or "XOR"), that is demonstrated by Minsky and Papert (1969), is one of the typical problems in Perceptron. Due to this problem, the studies on NN subsided. In 1986, Rumelhart et al. (1986) introduced a back propagation algorithm to solve the XOR problem. As a result, numerous types of MLNN have been studied and have led to the formation of a basis for development into Deep Learning today (Hinton et al., 2006; Schmidhuber, 2015). In general, MLNN is applied to pattern recognition, data compression, noise reduction, and many more areas. Though the MLNN is an improved model, it can be regarded as a category of Perceptron.

2.3 Recurrent Neural Network

RNN is one of the categories of ANN. RNN has connections between units from a directed cycle. Thus, the network is able to handle a comprehensive information analysis of the past and current states. It means that the network exhibits dynamic temporal behavior. RNN is classified into three categories from the viewpoint of connection structures of neurons as shown in Table 2.1. In addition, the typical structures of RNN are shown in Figure 2.5.

Table 2.1: Classification of recurrent neural networks.

	Classification		Typical Model
Recurrent Neural Networks	Fully Connection		Boltzmann Machine (Ackley et al., 1985)
	Symmetric Connection	Synchronism	Associatron (Nakano, 1972)
		Asynchronism	**Hopfield Associative Memory (Hopfield, 1982)**
	Asymmetric Connection	Hierarchical	Jordan Networks (Jordan, 1997) Elman Networks (Elman, 1990)
		Module	Long Short Term Memory (Hochreiter & Schmidhuber, 1997)
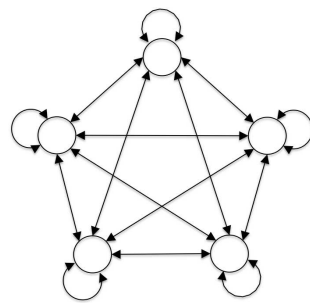
(a) Fully connection
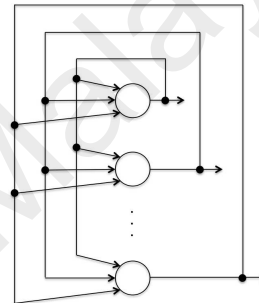
(b) Symmetric connection
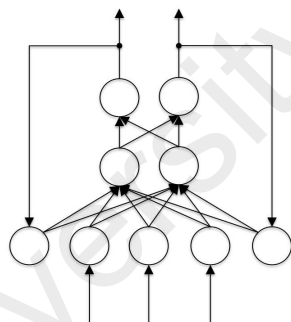
(c) Hierarchical connection
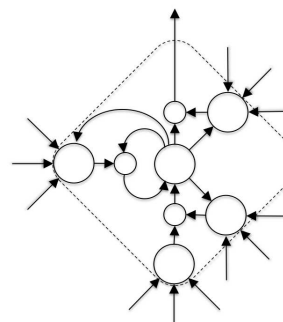
(d) Module type

Figure 2.5: Typical models of recurrent neural network.

Asymmetric connection model is available as internal memory to process arbitrary sequences of input information. It would provide the ability to recognize the unsegmented continuous information, such as handwriting recognition, which is the best known result. However, it requires the fixed input information in place of the sequence pattern information for the symmetric connection model that is typified by Associatron and HAM. The inner state of neurons in Associatron will be updated synchronously. On the other hand,

the asymmetric connection model will be updated asynchronously in HAM. In general, symmetric connection model has the function of AM. A full connection model that is typified by the Boltzmann Machine (Ackley et al., 1985) is also made available for the sequence pattern information and fixed input information by changing the network design in accordance to the problem.

In this thesis, models with association function, namely the types of HAM in Table 2.1, are primarily discussed and reviewed. Thus, the following sections mainly focus on the symmetric connection models of RNN.

2.3.1 Auto-Association Model

HAM is regarded as a typical auto-association model that is able to associate stored information from incomplete data. HAM works as content-addressable memory system with bipolar threshold nodes, and with a learning process defined by Hebb learning. Because HAM consists of a single directional layer, as shown in Figure 2.6, its ability is limited to auto-association only.
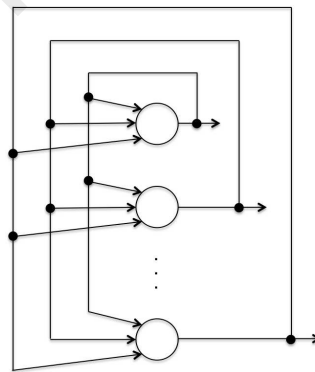
Figure 2.6: Structure of Hopfield associative memory.

HAM is composed of N neurons (a number of neurons in X-layer (N)), and a number of original memory vectors (k) of the memory pattern $\left\{ \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots, \boldsymbol{X}^{(k)} \right\}$ ($\boldsymbol{X} = [x_1, x_2, \ldots, x_N]$, $x \in \{-1, 1\}$) (an element of original memory in X-layer (x), and the original

memory vectors in X-layer ($\boldsymbol{X}$)) to be stored. In general, HAM is formalized as follows;

$$
\begin{cases}
\boldsymbol{S}^{(k)} = \sum_{j=1}^{N} \boldsymbol{W}\boldsymbol{X}^{(k)} & \text{(2.7a)} \\
\boldsymbol{X}^{(k)} = \phi\left(\boldsymbol{S}^{(k)}\right) & \text{(2.7b)}
\end{cases}
$$

where, a temporal states of associated pattern in X-layer ($\boldsymbol{S}$). $\phi(\cdot)$ denotes the activation function. The activation function of HAM is usually applied as a step function with bipolar state. According to Hebb learning, a weight matrix ($\boldsymbol{W}$) is defined as the following;

$$
\boldsymbol{W} =
\begin{cases}
\dfrac{1}{k}\sum_{p=1}^{k} \boldsymbol{X}^{(p)}\boldsymbol{X}^{(p)T}, & (i \neq j) \\
0, & (i = j)
\end{cases}
. \qquad (2.8)
$$

where, exponential T denotes transpose operation, and k denotes the number of memory pairs.

HAM is guaranteed to converge to a local minimum, however, convergence to a false pattern (wrong local minimum) rather than the stored memory pattern (expected local minimum) could occur.

2.3.2 Hetero-Association Model

Hetero-association model can recall the complete information from a piece of linked information. Furthermore, the part of information that is input to the network is restored to the complete information, similar to an auto-association model. The fundamental models of hetero-associative memory, known as BAM and MAM, were introduced by Kosko (1987b) and Hagiwara et al. (1990), respectively. In this section, fundamentals of BAM, MAM, and its improvement models are presented.

2.3.2.1 Bidirectional Model

BAM consists of two layers of neurons, which are fully connected to each other as shown in Figure 2.7.
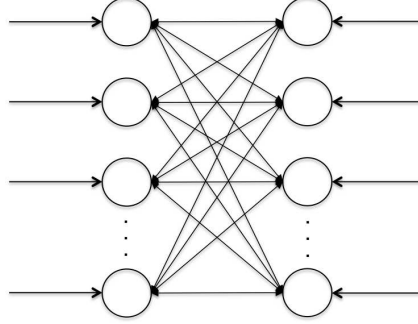
Figure 2.7: Structure of Bidirectional Associative Memory.

Here, N neurons in X-layer and a number of neurons in Y-layer (M) with k of the memory pairs $\left\{ \left(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)} \right), \left(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)} \right), \ldots, \left(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)} \right) \right\}$, where $\mathbf{X} = [x_1, x_2, \ldots, x_N]$ ($x \in \{-1, 1\}$) (an element of original memory in X-layer (x), and the original memory vectors in X-layer ($\mathbf{X}$)) and $\mathbf{Y} = [y_1, y_2, \ldots, y_M]$ ($y \in \{-1, 1\}$) (an element of original memory in Y-layer (y), and the original memory vectors in Y-layer ($\mathbf{Y}$)), are stored in BAM. Thus, BAM is formalized as;

- X-layer to Y-layer

$$
\begin{cases}
\mathbf{S}^{(k)} = \sum_{j=1}^{M} \mathbf{W}^T \mathbf{X}^{(k)} & \text{(2.9a)} \\
\mathbf{Y}^{(k)} = \phi \left(\mathbf{S}^{(k)} \right) & \text{(2.9b)}
\end{cases}
$$

- Y-layer to X-layer

$$
\begin{cases}
\mathbf{U}^{(k)} = \sum_{i=1}^{N} \mathbf{W} \mathbf{Y}^{(k)} & \text{(2.10a)} \\
\mathbf{X}^{(k)} = \phi \left(\mathbf{U}^{(k)} \right) & \text{(2.10b)}
\end{cases}
$$

where, are the temporal states of the associated pattern in X-layer ($\mathbf{S}$) and a temporal states of associated pattern in Y-layer ($\mathbf{U}$). N and M denote the number of neurons in X-

layer and Y-layer, respectively. The $\phi(\cdot)$ denotes an activation function as a step function with bipolar state. The weight connections $\boldsymbol{W}$ and a transpose weight matrix of $\boldsymbol{W}$ ($\boldsymbol{W}^T$) are defined by Hebb learning as;

- X-layer to Y-layer

$$\boldsymbol{W}^T = \frac{1}{k}\sum_{p=1}^{k}\boldsymbol{Y}^{(p)}\boldsymbol{X}^{(p)T}. \tag{2.11}$$

- Y-layer to X-layer

$$\boldsymbol{W} = \frac{1}{k}\sum_{p=1}^{k}\boldsymbol{X}^{(p)}\boldsymbol{Y}^{(p)T}. \tag{2.12}$$

where, exponential T denotes transpose operation. k denotes the number of memory pairs.

2.3.2.2 Multi-directional Model

The multi-directional model can be constructed from the L ($L > 2$) of the layers (a number of layers in the association model (L)). Thus, the focusing layer is defined as α-th layer, and the other layers are referred to as β-th layers. MAM consists of fully connected multiple layers of neurons as shown in Figure 2.8.
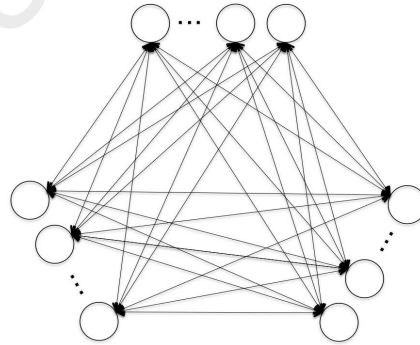
Figure 2.8: Structure of multi-directional associative memory (3-layers model).

In general, MAM can be considered as multiple BAM. The state of α-th associated layer is determined by the sum of inner state from β-th layers through an activation function. Here, let k of the memory pairs $\left\{\boldsymbol{X}_{(1)}^{(k)},\boldsymbol{X}_{(2)}^{(k)},\ldots,\boldsymbol{X}_{(L)}^{(k)}\right\}$ be stored, where $\boldsymbol{X}$ denotes the memory vectors as $\boldsymbol{X}_{(l)}^{(k)} = \left[x_1,x_2,\ldots,x_{N_{(l)}}\right]$ ($x \in \{-1,1\}, l = 1,2,\ldots,L$), L denotes

the number of layers, and $N_{(l)}$ denotes the number of neurons in l-th layer. Thus, MAM is formalized as follows;

- α-th layer to β-th layers

$$
\begin{cases}
\mathbf{S}_{(\beta)}^{(k)} = \sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^{L} \mathbf{W}_{(\alpha\beta)}^{T} \mathbf{X}_{(\alpha)}^{(k)}, & (\beta = 1, 2, \ldots, L; \ \beta \neq \alpha) \quad (2.13\text{a}) \\
\mathbf{X}_{(\beta)}^{(k)} = \phi\left(\mathbf{S}_{(\beta)}^{(k)}\right) & (2.13\text{b})
\end{cases}
$$

- β-th layers to α-th layer

$$
\begin{cases}
\mathbf{S}_{(\alpha)}^{(k)} = \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{L} \mathbf{W}_{(\alpha\beta)} \mathbf{X}_{(\beta)}^{(k)}, & (\alpha = 1, 2, \ldots, L; \ \alpha \neq \beta) \quad (2.14\text{a}) \\
\mathbf{X}_{(\alpha)}^{(k)} = \phi\left(\mathbf{S}_{(\alpha)}^{(k)}\right) & (2.14\text{b})
\end{cases}
$$

where, α and β denote the layer number. $\mathbf{S}$ denotes the temporal states of associated patterns. L denotes the number of layers, and N denotes the number of neurons. The $\phi(\cdot)$ denotes an activation function as a step function with bipolar state. The weight connections $\mathbf{W}_{(\alpha\beta)}^{T}$ and $\mathbf{W}_{(\alpha\beta)}$ are defined by Hebb learning as;

- α-th layer to β-th layers

$$
\mathbf{W}_{(\alpha\beta)}^{T} = \frac{1}{k} \sum_{p=1}^{k} \mathbf{X}_{(\beta)}^{(p)} \mathbf{X}_{(\alpha)}^{(p)T}. \quad (2.15)
$$

- β-th layers to α-th layer

$$
\mathbf{W}_{(\alpha\beta)} = \frac{1}{k} \sum_{p=1}^{k} \mathbf{X}_{(\alpha)}^{(p)} \mathbf{X}_{(\beta)}^{(p)T}. \quad (2.16)
$$

where, exponential T denotes transpose operation and k denotes the number of memory pairs. The neurons in each layer continue to be subjected to cyclic updates until the layer reaches an equilibrium.

2.4 Improvements of Associative Memory Models

In general, the improvements in capabilities are achieved by the changes made to the model structure and learning algorithm, or by applying relevant knowledge from other research fields.

2.4.1 Structural and Algorithmic Improvements

Structural improvements are carried out by network expansion such as adding a dummy neuron and a layer, or adopting a hidden layer, whereas, algorithmic improvements are made by learning algorithms of the weight matrix of memory vectors.

2.4.1.1 Structural Improvement

Wang et al. (1990) introduced two coding strategies for BAM consisting of multiple training that can guarantee the recall of a single trained pair under suitable initial conditions, and a dummy augmentation that can guarantee the recall of all trained pairs, if dummy data is attached to the training pairs. The simulation experiment demonstrated that an improvement over the original BAM can be achieved for recall of multiple pairs as well. Furthermore, the sufficient condition for weight matrix of BAM to take a local minimum was also discussed by (Y. F. Wang et al., 1991). The study showed a linear programming/multiple training method that determines the weights that satisfy the conditions when a solution is feasible.

Wu and Pados (2000) applied a dummy layer to represent feed-forward connections in BAM. The feed-forward BAM is a three-layer network of McCulloch–Pitts neurons with guaranteed perfect bidirectional recall. Based on their theoretical analysis, they showed that the model guarantees perfect bidirectional recall for arbitrarily correlated patterns. The overall network design procedure is fully scalable in the sense that any number of bidirectional associations can be implemented.

Kang (1994) proposed a triple-layered hybrid neural network that consists of the following layers: the first synapse is a one-shot associative memory using the modified Kohonen's adaptive learning algorithm with arbitrary input patterns; the second is a fundamental BAM consisting of orthogonal input/output basis vectors that satisfy the strict continuity condition; the third is a simple one-shot AM. In addition, the mathematical proofs based on relationships between local minima and noise-free recall are provided, and the conditions of robust capacity of a triple-layered hybrid neural network are derived.

In general, it is regarded that the increasing model scale is a simple and effective improvement for NNs in terms of memory capacity and noise tolerance. However, the learning process in above models might be slower than fundamental models because of the network size.

2.4.1.2 Algorithmic Improvement

The original models of AM, namely HAM, BAM and MAM, have limited memory capacity, recall reliability, and spurious attractors, because of correlation learning. In an earlier study based on the spin glass formulation, Personnaz et al. (1985; 1986) proposed and analyzed the well-known rule called projection matrix based on least mean squared error minimization. The learning process based on an inverse matrix principle is not a local process. The spin glass formulation of the neural network problem leads to particularly simple results which allow an analytical evaluation of the attractivity of the memorized states. It is shown that, for any preassigned set of states to be memorized, the parameters of the network can be completely calculated in most cases in order to guarantee the stability of these states.

Zhuang et al. (1993) integrated a Householder transformation (Householder, 1958; Press et al., 1996) and BAM. Because of Hebb-type correlation learning, the traditional BAM cannot take fixed points of all training pattern pairs, despite the small number of

training pairs. The Householder transformation is a linear transformation that describes a reflection about a plane or hyperplane containing the origin. Simulation results show that the BAM with a Householder transformation has superior memory capacity as compared to original BAM, particularly when the input dimensions are large. Leung (1993) introduced an enhanced Householder encoding algorithm, which is based on the Householder encoding algorithm and projection matrix, and applied it to BAM.

Several studies focused on the asymmetric connections between neurons and their weight matrix (Oh & Kothari, 1991, 1994; Xu et al., 1994; Shi et al., 1998). The concept of asymmetric connection contributed to the performance improvements of AM, while reducing the computational cost.

Wang (1996) introduced the BAM that have inter-connected neurons inside each layer. The model uses an optimal AM matrix instead of the standard correlation matrix. This study presents the designs of a linear BAM and a nonlinear BAM. Furthermore, the analysis of the stability and other performance criteria of the models is provided. Based on the simulation experiments, the capacities of the inter-connective BAM are far higher than the original BAM. In a similar study, Eom et al. (2002) utilized a Hamming distance in recall procedure of usual asymmetrical BAM that is replaced with modified Hamming distance to define the weight matrix. Consequently, the model is able to store highly correlated information by adjusting weighting factors appropriately. This property is proven to be essential in implementing multiple associations with an elaborated dimension augmentation scheme. This property models the imperfectness in training data during the learning process.

Shi et al. (1998) developed a theory for General Bidirectional Associative Memory (GBAM). The paper discussed the associative recalling process as a dynamic system under stability and asymptotic stability conditions, and developed an algorithm for learning the asymptotic stability conditions of GBAM. GBAM does not require the usual

assumption that the interconnection weight from a neuron in the X-layer to a neuron in the Y-layer is the same as the one from the Y-layer to the X-layer. The effectiveness of GBAM in terms of storage capacity, attraction, and avoiding spurious memories are demonstrated by some outstanding experimental results. Oh and Kothari (1991, 1994) introduced a pseudo relaxation learning algorithm to HAM and BAM. This iterative learning algorithm is adapted from the relaxation method, which replaces the gradient descent technique used in existing iterative learning algorithms for solving systems with linear inequalities. It is well suited for neural network implementation, guarantees the recall of all training patterns, is highly insensitive to learning parameters, and offers high scalability for large applications. As an extended algorithm, Quick Learning, which applies pseudo relaxation learning algorithm and Hebb learning, which defines initial learning states, have been introduced and are applied to BAM and MAM called Quick Learning Bidirectional Associative Memory (QL-BAM) (Hattori et al., 1994) and Quick Learning Multi-directional Associative Memory (QL-MAM) (Hattori & Hagiwara, 1995), respectively.

Gripon and Berrou (2011) introduced an alternative construction to define the weight matrix, which builds on ideas from the theory of error correcting codes; the alternative construction greatly outperforms HAM in capacity, diversity, and efficiency based on sparse connection. In this model, three levels of sparsity are provided. The first level is related to the size of information that are smaller than the number of available neurons. The second is provided by a particular coding rule, acting as a local constraint in the neural activity. The third is a characteristic of the low final connection density of the network after the learning phase. Yao et al. (2013) implemented a variation of the Gripon-Berrou AM that is processed on a graphical processing unit. Further discussion about effectiveness of sparseness, which is based on human brain modeling in AM, is well summarized by Palm (2013).

In the past decades, numerous fusion studies of computational intelligences, such as NNs, FS and GA, were carried out, and their advantages were shown (Kar et al., 2014; Pham & Karaboga, 2012; Ding et al., 2013; Sakawa, 2012; Fazzolari et al., 2013).

In terms of AM, several types of fusion models are introduced. In the fusion models with GA, Imada and Araki (1995) proposed a simple technique of genetic algorithm to modify the weight matrix to enlarge the HAM, which keeps the original weight matrix unchanged during the evolution. The individuals of each generation are slightly different from the original weight matrix, keeping the diagonal elements zero. Kumar and Singh (2010) applied the population generation technique, the crossover operator and the fitness evaluation function to obtain the optimal weight matrices in AM for efficient recall of any prototype input pattern. From the perspectives of various studies of GA, the optimization of weight connection in AM is applied to evaluate the proposed GA (F. H. Leung et al., 2003; Tsai et al., 2006). This indicates that GA is applicable for the improvement of AM.

Furthermore, in the fusion models with FS, the earliest attempt to utilize the fuzzy theory to describe an associative memory called Fuzzy Associative Memory (FAM) was introduced by Kosko (1991b). This model was described in terms of a nonlinear matrix product called max-min (or max-product) composition, and the weight matrix of AM is given by fuzzy Hebb learning. Chung and Lee (1996) generalized the Kosko's FAM model using max-t compositions. Liu (1999) extended and generalized the max-min FAM by introducing a threshold at each unit of max-min FAMs. Sussner and Valle (2006) proposed an implicative FAM, which is described by OR/AND neurons (Hirota & Pedrycz, 1994) with threshold. The weight matrix is determined by the minimum implications of presynaptic and postsynaptic activations. The model exhibits unlimited storage capacity, one-pass convergence, and tolerance with respect to erosive noise. Additionally, Sussner and Valle (2008) provided the general framework for FAM.

Based on various physiological experiments, chaotic behavior in neurons of human brain has been observed. In particular, it is considered that chaos plays an important role in memory and learning of human brain. Aihara et al. (1990) first introduced the chaotic responses of a biological neuron, and proposed the concept of chaotic neural network. Applying the concept of chaotic neurons, Osana et al. (1996) proposed BAM with chaotic neurons. In this model, each training pair requires contextual information, and chaotic neurons are used in a part of the network corresponding to the contextual information. Adachi and Aihara (1997) analyzed the chaotic associative dynamics in terms of spatio-temporal output patterns, quasi-energy function, distances between internal state vectors and orbital instability. In a recent study, a chaotic hetero-associative memory with dynamic behavior is proposed by Aghajari et al. (2015). This model can store twice as much as a regular hetero-associative memory using an extension of sparse learning algorithm that gives the network ability for successive learning. The storage capacity and recall reliability show superior performances, however, the system is complicated because of chaotic neuron dynamics.

In general, most of existing systems are required to periodically update in order to correspond to new information. However, the learning algorithms in HAM, BAM, MAM, and the majority of models that are reviewed above are limited to offline and one-shot learning rules because of the Hebb-type learning algorithms called batch learning. In batch learning, when the system learns new information, the existing learning result is discarded, and the learning process is performed considering the whole information once again. Incremental learning is able to integrate new information to existing learned result incrementally. It provides the reduction of computational cost of the learning process, and the ability to perform big data learning (Read et al., 2012).

Storkey and Valabregue (1999) proposed an incremental learning algorithm for HAM based on Hebb/anti-Hebb learning. It retains the important functionalities, such as incre-

mentality and locality, which are found lacking in projection matrix. Salavati et al. (2014) discussed the problems of AM in terms of learning and pattern retrieval parts from the perspective of conventional batch learning algorithms. Furthermore, they proposed a simple iterative algorithm that utilizes the neural graph theory with sparsity to improve memory capacity and recall reliability. Chartier and Boukadoum proposed and analyzed the model with a self-convergent iterative learning rule based on Hebb/anti-Hebb approach, and a nonlinear output function for BAM and MAM, called IBAM (Chartier & Boukadoum, 2006a; Chartier et al., 2008) and IMAM (Chartier & Boukadoum, 2011), respectively. IBAM and IMAM are able to process the online learning without being subject to over-learning. Simulation experiments show that IBAM IMAM have fewer spurious attractors as compared to BAM and other learning algorithms.

Though these models show more superior abilities than fundamental models in terms of memory capacity, recall reliability, and reduction of spurious attractors, the complexity of the model has the tendency to increase significantly.

2.4.2 Complex Domain Model

In 1973, the concept of complex-valued artificial neuron model was first introduced by Aizenberg (1973). The structure of complex-valued neuron is the same as shown in Figure 2.2, except that it works with complex numbers. A significant advantage of complex-valued neural network is that it can handle higher dimensions as compared to real-valued neural networks like Perceptron. In terms of activation function, a real-valued model utilizes a step function and sigmoid function as shown in Figure 2.3, whereas, a complex-valued model applies the discrete multivalued threshold logic for an activation function as in Figure 2.9.
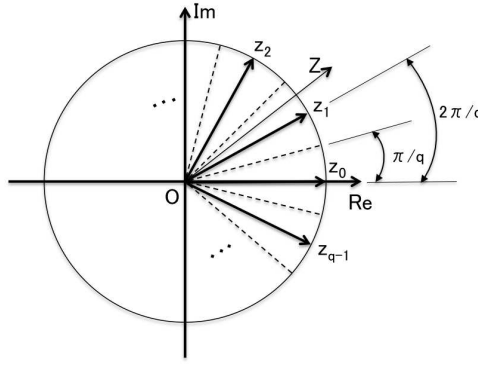
Figure 2.9: Complex-valued activation function.

In general, the discrete complex-valued activation function $\phi(\cdot)$ is summarized as follows;

$$\phi(Z) = \begin{cases} \exp(j2\pi n/r), & \text{If } \left| \text{Arg}\left\{ \frac{Z}{\exp(j2\pi n/r)} \right\} \right| < \pi/r \text{ and } Z \neq 0 \\ \text{previous state}, & \text{If } Z = 0 \end{cases} . \tag{2.17}$$

where, a quantization value on the complex unit circle (r). $\text{Arg}(\cdot)$ denotes the phase angle which is taken to range over $(-\pi, \pi)$, and n takes an integer. Assume $z_0, z_1, \ldots, z_{r-1}$, are r quantization values. Here, z_1 is considered as an output that is closest to Z.

The early model of multi-states associative memory was introduced by Noest (1988a, 1988b). Moreover, based on real-valued HAM, BAM, and MAM, Jankowski et al. (1996) introduced CHAM, Lee et al. (1998) introduced CBAM, and CMAM was implemented by Kobayashi et al. (2005), respectively. It is known that a complex number is equal to a real number if the phase information of complex number is zero. Conventionally, numerous studies that improve complex-valued models are introduced, based on the improvements for real-valued models.

Lee et al. (2006) showed that the projection matrix proposed by Personnaz can be generalized to a complex domain. Furthermore, they applied a complex-valued projection matrix to CHAM and analyzed the stability of the model by using energy function approach, which shows that in synchronous update mode the model is guaranteed to con-

verge to a fixed point from any given initial state. Kobayashi (2008) applied a pseudo-relaxation learning algorithm to CHAM. Furthermore, the complex-valued chaotic behavior models are also considered based on real-valued models (R. S. Lee, 2004; Chakravarthy et al., 2008; Shimizu & Osana, 2010; Yoshida & Osana, 2011). Gislén et al. (1992) proposed Rotor Hopfield Associative Memory (RHAM) as an extension of CHAM. The most significant attribute of RHAM is that it does not store the rotated patterns of training patterns, which have less noise robustness because they store rotated patterns. The learning algorithms of fundamental RHAM utilize Hebb learning and gradient descent learning. Kitahara and Kobayashi (2014) applied a projection rule for RHAM to increase the noise robustness of RHAM.

2.4.3 Quantum-Inspired Model

Conventionally, several types of QNN have been introduced based on fundamental NNs, and shown their potential and superior abilities to solve problems pertaining recognition, classification, and optimization (Levy & McGill, 1993; Purushothaman & Karayiannis, 1997; Lu et al., 2013; Gandhi et al., 2014; Chen et al., 2014). It can be regarded that applying the concept of QM is a successful approach to algorithmic improvement in ANNs.

With regard to AM, Perus et al. (2007) provided a comprehensive introduction into associative processing and memory-storage in quantum-physical framework, particularly, it introduced a neural-network-like quantum information dynamics based on HAM (or spin glass model) and holographic NN. They revealed the relationship between the conventional neural dynamics and quantum processing. Furthermore, algorithms of HAM viewed from quantum processing perspective are discussed. Zheng and Zhou (2008) proposed a novel quantum associative memory using the quantum holography scheme. In this model, the memory vectors are encoded onto a set of independent quantum states using an iterative learning algorithm. Because of quantum implementation, the model is not

only able to improve the performances, but it may also be a suitable model to explain the biological associative phenomena because the inherent phase of oscillation in the model. These models perform excellently in terms of mathematical properties of quantum processing. At the same time, however, the models are limited to auto-association model, due to the constraint of quantum theory based on the probability interpretation, namely $\left\langle \psi^i | \psi^i \right\rangle = 1$ (a memory vector in quantum state (ψ)).

One of the successful approaches is QHAM and was proposed by Rigatos and Tzafestas (2006). This model demonstrates that quantum information processing in neural structures results in an exponential increase in storage capacity with simple algorithms, and is able to explain the extensive memory and inferencing capabilities of humans. This model is applied as fuzzy inference to weight matrix to satisfy the features of QM, namely parallelism and unitarity. Although QHAM is limited to auto-association and bipolar processing because of its neuronal structure, the model has the potential to be extended as a hetero-association model.

2.4.4 Interactive Robot System with Emotion Affected Associative Memory

As mentioned in Chapter 1, associative memory function in the human brain plays a significant role in human–human communication. In addition, human beings have the tendency to recall emotions from information based on their knowledge and experience (Smith & Petty, 1995). Consequently, several interactive robot system with AM have been introduced to improve the human-robot interaction. It can be considered that system implementation is a functional improvement of AM.

Conventionally, a number of computational emotion models with several factors such as personality and biometric information have been introduced. In terms of emotion, personality is one of the significant factors. In the past, several psychologists have discussed relationships between the human emotional factor and personality factor (Gray,

1987; Russell & Barrett, 1999). From a behavioral perspective (Johns & Silverman, 2001), various rule-based models (André et al., 2000) and probabilistic models (Ball & Breese, 2000) have been introduced. Costa et al. (Costa & McCrae, 1992) introduced the OCEAN model based on five factors, i.e., openness, conscientiousness, extraversion, agreeableness, and neuroticism. Mehrabian (1996) utilized the five factors of personality to represent the Pleasant-Arousal-Dominance (PAD) temperament model. The relationship between the five factors of personality and PAD model is derived through linear regression analysis (Goldberg, 1992). Based on these psychological backgrounds, several types of computational based robotic emotional models have been introduced (Schneider & Adamy, 2014). Han et al. (2013) employed the five factors of personality to a 2D (pleasure-arousal) scaling model that was introduced by Russell and Bullock (1985) to represent a robotic emotional model. This model generates a robot mood state based on the human facial expression information.

Applying such computational emotion models, several types of interactive robot systems, which utilize the emotion affected associative memory model, have been introduced. Hiolle et al. (2012) developed a system to elicit care-giving behavior in a human-robot interaction with adolescents/adults by using associative memory model. The robot collects the multi-modal information to associate the suitable behavior for the situations. Rumbell et al. (2012) discussed and reviewed emotional mechanisms that are often used in artificial agents as a method of improving action selection. Of utmost importance is the emotional factors' impact on the cognitive process, which is focused on attention and knowledge retrieval from memory. Conventionally, several types of robot systems that apply the emotion affected associative memory to select suitable robot behaviors have been introduced. Itoh et al. (2005) developed an emotion expression humanoid robot and its interactive system. The system is able to associate the emotional expression for robot from human behavior using a chaotic complex-valued AM whose output is controlled

by the mental model comprising the mental space, mood, equations of emotion, robot personality, need model, consciousness model, and behavior model. Yi et al. (2014) argued that AM is essential to realize man–machine cooperation in the natural interaction between humans and robots. They developed an emotional robot platform with associative memory model, which is applied in emotional robots. The system calculates the emotional state of the robot's dynamic change of mood state while concurrently considering their own needs based on the external information. Valverde et al. (2014) studied and proposed the computation model for associative memory and emotional information using ideas inspired by neuroscience research into neural-endocrine systems interaction. The model is able to create the emotional memory of the robot. The emotional memory is utilized to predict future emotional states based on past experiences, and it would be affected by selecting the emotional response actions of the robot. Furthermore, the model is able to monitor the emotional state of a communication partner, to identify if that individual is in a state of flow, and providing positive support that might be needed.

2.5 Summary

In this chapter, the fundamentals of artificial neuron and its networks are introduced, and the comprehensive review of recurrent networks models are presented. Furthermore, the conventional studies for improving the performance of associative memory models are reviewed. Finally, the robot systems which have the emotion affected associative memory models are introduced. The summary of the models and systems are listed in Tables 2.2, 2.2 and 2.4.

As summarized in this chapter, the several studies have introduced to improve the artificial associative memory models. Although these improvements have certain advantages, the other problems are raised such as increasing the size and complexity of networks, the learning algorithm is still batch learning and so on. In terms of QCI, the models show the superior abilities with the simple network architecture in several fields. In this thesis, carrying out the solution of the research problems by developing QHAM.

Table 2.2: Summary for improvements of associative memory model (I)

Model	Aspect		Learning Type		Association Type		Neuron State		Fusion with
	Structural	Algorithmic	Batch	Incremental	Auto-	Hetero-	Real-valued	Complex-valued	
Wang et al. (1990)	√		–	–	–	–	–	–	–
Wu and Pados (2000)	√		–	–	–	–	–	–	–
Kang (1994)	√		–	–	–	–	–	–	–
Personnaz et al. (1985, 1986)		√	√		√		√		
Zhuang et al. (1993)		√	√			√	√		
Leung (1993)		√	√			√	√		
Zhuang et al. (1993)		√	√			√	√		
Shi et al. (1998)		√	√			√	√		
Oh and Kothari (1991)		√	√		√		√		
Oh and Kothari (1994)		√	√			√	√		
Hattori et al. (1994)		√	√			√	√		
Hattori and Hagiwara (1995)		√	√			√	√		
Gripon and Berrou (2011)		√	√		√		√		
Imada and Araki (1995)		√	√		√		√		GA
Kumar and Singh (2010)		√	√		√		√		GA
Kosko (1991a)		√	√			√	√		FS
Chung and Lee (1996)		√	√			√	√		FS
Liu (1999)		√	√			√	√		FS
Sussner and Valle (2006)		√	√			√	√		FS
Osana et al. (1996)		√	√			√	√		Chaos

Table 2.3: Summary for improvements of associative memory model (II)

Model	Aspect		Learning Type		Association Type		Neuron State		Fusion with
	Structural	Algorithmic	Batch	Incremental	Auto-	Hetero-	Real-valued	Complex-valued	
Adachi and Aihara (1997)		√	√			√	√		Chaos
Storkey and Valabregue (1999)		√		√	√		√		
Salavati et al. (2014)		√		√	√		√		
Chartier and Boukadoum (2008; 2011)		√		√	√		√		
Lee et al. (2006)		√	√		√			√	
Kobayashi (2008)		√	√		√			√	
Lee (2004)		√	√		√			√	Chaos
Chakravarthy et al. (2008)		√	√		√			√	Chaos
Shimizu and Osana (2010)		√	√			√		√	Chaos
Yoshida and Osana (2011)		√	√			√		√	Chaos
Gislén et al. (1992)		√	√		√			√	
Kitahara and Kobayashi (2014)		√	√		√			√	
Zheng and Zhou (2008)		√	√		√		√		
Rigatos and Tzafestas (2006)		√	√		√		√		
QBAM		√	√			√	√		
QMAM		√	√			√	√		
IQBAM		√		√		√	√		
IQMAM		√		√		√	√		
QCBAM		√	√			√		√	
QCMAM		√	√			√		√	

Table 2.4: Summary for discussed interactive robot system incorporaing emotion information

System	Emotion Model	Mood State	Personality	Input Modality	Biometric Information	Association Model	Communication Partner
Hiolle et al. (2012)	Arousal model (self-defined)	None	None	Touch, Gesture	None	Multidirectional (many-to-one)	Animal robot
Itoh et al. (2005)	Pleasant-Arousal -Certainty model (self-defined)	Positive and Negative	None	Voice	None	Bidirectional (one-to-one)	Upper body of humanoid (WE-4RII)
Yi et al. (2014)	Emotional energy model (self-defined)	Positive and Negative	None	Object, Voice	None	Bidirectional	Facial robot
Valverde et al. (2014)	Plutchik emotion model (universal emotion)	None	None	Hormonal state (Testosterone, Cortisol, Serotonin)	None	Bidirectional	Autonomous robot
Robot system in this thesis	Pleasant-Arousal model (universal emotion)	Positive and Negative	OCEAN model	Gesture, Object, Voice, Facial Exp.	Age, Gender	Multidirectional (one-to-many)	Facial display and arms robot

CHAPTER 3

QUANTUM-INSPIRED ASSOCIATIVE MEMORIES

3.1 Introduction

Applying the concepts of QM is a successful approach for algorithmic improvement in ANNs. In terms of AM, several types of quantum-inspired auto-association models have been introduced, and their superior abilities have been exhibited. However, the improvement of a hetero-association model from these auto-association models is difficult, because of the constraint of quantum theory, which is based on the probability interpretation and is described as $\langle \psi^i | \psi^i \rangle = 1$. Rigatos and Tzafestas (2006) proposed QHAM that shows the potential to extend to hetero-association model. QHAM is applied to a fuzzy inference to weight matrix to satisfy the features of QM, that are parallelism and unitarity. QHAM demonstrates quantum information processing in neural structures that results in an exponential increase in the storage capacity using simple algorithms, and be able to explain the extensive memory and inferencing capabilities of humans. Therefore, it can be assumed that quantum-inspired hetero-association model based on QHAM also shows outstanding abilities.

In this chapter, the mathematical models of quantum-inspired bidirectional and multi-directional association models, namely QBAM and QMAM, are presented. In addition, based on IBAM (Chartier & Boukadoum, 2006a, 2006b; Chartier et al., 2008) and IMAM (Chartier & Boukadoum, 2011), which are applied to a self-convergent iterative learning algorithm that is based on Hebb/anti-Hebb approach with a nonlinear output function, QBAM and QMAM with an incremental learning algorithm called IQBAM and IQMAM, respectively, are also presented. Furthermore, the mathematical proofs of the features of

QM, such as superposition and unitarity, and its numerical example are provided. Finally, the simulation experiments to evaluate the memory capacity and noise tolerance of the model compared with those of conventional models are presented.

3.2 Fundamental Structures of Bidirectional Association Model

This section presents the mathematical models of QBAMand IQBAM. Because the models only have two layers, the layers are defined as X-layer and Y-layer. Here, let k of the original memory pairs $\left\{ \left(\boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(1)} \right), \left(\boldsymbol{X}^{(2)}, \boldsymbol{Y}^{(2)} \right), \ldots, \left(\boldsymbol{X}^{(k)}, \boldsymbol{Y}^{(k)} \right) \right\}$, where $\boldsymbol{X} = [x_1, x_2, \ldots, x_N]$ $(x \in \{-1, 1\})$ and $\boldsymbol{Y} = [y_1, y_2, \ldots, y_M]$ $(y \in [-1, 1])$, are stored , in both QBAM and IQBAM. N and M denote the number of neurons in X-layer and Y-layer, respectively.

3.2.1 Quantum-Inspired Bidirectional Associative Memory

Based on above conditions, QBAM is formalized as followings;

- X-layer to Y-layer

$$\begin{cases} \boldsymbol{S}^{(k)} = \sum_{j=1}^{N} \sum_{i=1}^{M} W_{ij}^T x_i^{(k)} & \text{(3.1a)} \\ \boldsymbol{Y}^{(k)} = \phi\left(\boldsymbol{S}^{(k)} \right) & \text{(3.1b)} \end{cases}$$

- Y-layer to X-layer

$$\begin{cases} \boldsymbol{U}^{(k)} = \sum_{i=1}^{M} \sum_{j=1}^{N} W_{ij} y_j^{(k)} & \text{(3.2a)} \\ \boldsymbol{X}^{(k)} = \phi\left(\boldsymbol{U}^{(k)} \right) & \text{(3.2b)} \end{cases}$$

where, $\boldsymbol{S}$ and $\boldsymbol{U}$ denote the temporal states of associated patterns in X-layer and Y-layer, respectively. The $\phi(\cdot)$ denotes an activation function as a step function with bipolar state. The weight connections $\boldsymbol{W}^T$ and $\boldsymbol{W}$ are defined by Hebb learning as follows;

- X-layer to Y-layer

$$W^T = \frac{1}{k} \sum_{p=1}^{k} \sum_{j=1}^{N} \sum_{i=1}^{M} v_j^{(p)} u_i^{(p)T}. \tag{3.3}$$

- Y-layer to X-layer

$$W = \frac{1}{k} \sum_{p=1}^{k} \sum_{i=1}^{M} \sum_{j=1}^{N} u_i^{(p)} v_j^{(p)T}. \tag{3.4}$$

where, exponential T denotes transpose operation, and k denotes the number of memory pairs. Here, orthonormalized vectors u and v are calculated by Gram-Schmidt orthogonalization as follows; $\boldsymbol{a}_1 = \boldsymbol{A}_1 / \|\boldsymbol{A}_1\| (p = 1)$, $\boldsymbol{b}_p = \boldsymbol{A}_p - \sum_{i=p-1}^{k-1} (\boldsymbol{a}_i, \boldsymbol{A}_i) \boldsymbol{a}_i$ and $\boldsymbol{a}_p = \boldsymbol{b}_p / \|\boldsymbol{b}_p\|$ $(2 \leq p \leq k)$, where $\boldsymbol{A}$ denotes the memory vector, $\boldsymbol{a}$ and $\boldsymbol{b}$ denote the orthonormalized vector and the orthogonalized vector, respectively.

Algorithm 1 shows the pseudo code of above weight learning process for a weight matrix W in a batch learning model.

Algorithm 1 An algorithm for a weight matrix W in a batch learning model

---

Require: weight matrix W, fundamental memory vectors $X^{(k)}$ and $Y^{(k)}$, number of pairs k
Ensure: weight matrix W
  Initialize W as a zero matrix
  Calculate orthogonalized vectors u and v by Gram-Schmidt orthogonalization from X and Y
  Set $d = 1$
  while $k \geq d$ **do**
    Calculate $W_{ij} \leftarrow W_{ij} + \frac{1}{k}(u_i v_j^T)$
    Set $d \leftarrow d + 1$
  end while

3.2.2 Quantum-Inspired Incremental Bidirectional Associative Memory

IQBAM is formalized as followings;

- X-layer to Y-layer

$$\begin{cases} \boldsymbol{S}^{(k)} = \sum_{j=1}^{N} \sum_{i=1}^{M} W_{ji} x_i^{(k)} & \text{(3.5a)} \\ \boldsymbol{Y}^{(k)} = (\delta + 1)\boldsymbol{S}_{(t)}^{(k)} - \delta \boldsymbol{S}_{(t)}^{(k)3} & \text{(3.5b)} \end{cases}$$

47

- Y-layer to X-layer

$$\begin{cases} \boldsymbol{U}^{(k)} = \displaystyle\sum_{i=1}^{M}\sum_{j=1}^{N} V_{ij} y_j^{(k)} & \text{(3.6a)} \\[4mm] \boldsymbol{X}^{(k)} = (\delta+1)\boldsymbol{U}_{(t)}^{(k)} - \delta \boldsymbol{U}_{(t)}^{(k)^3} & \text{(3.6b)} \end{cases}$$

where, δ is a general output parameter, t denotes the number of iterations over the network.

The weight matrices $\boldsymbol{W}$ and $\boldsymbol{V}$ are defined as follows;

$$\boldsymbol{W}_{(t+1)} = \boldsymbol{W}_{(t)} + \eta \left(v_{(0)} - v_{(t)} \right) \left(u_{(0)} - u_{(t)} \right)^T \tag{3.7}$$

$$\boldsymbol{V}_{(t+1)} = \boldsymbol{V}_{(t)} + \eta \left(u_{(0)} - u_{(t)} \right) \left(v_{(0)} - v_{(t)} \right)^T \tag{3.8}$$

where, a learning parameter of incremental learning (η). $u_{(0)}$ and $v_{(0)}$ denote the initial inputs, which are selected randomly based on a uniform distribution, subsequently, those inputs are iterated t times through the network. u and v are the orthonormalized vectors of x and y, respectively, that are calculated by Gram-Schmidt orthogonalization.

The learning convergence is guaranteed with the condition of the learning parameter η (Chartier & Boukadoum, 2006b);

$$\eta < \frac{1}{2(1-2\delta)\text{Max}[M,N]}, \quad \delta \neq 1/2. \tag{3.9}$$

where, a general output parameter (δ), M and N denote the number of neurons in X-layer and Y-layer, respectively.

Algorithm 2 shows the pseudo code of above weight learning process for a weight matrix W in an incremental learning model. Similarly, a weight matrix V is calculated.

Algorithm 2 An algorithm for a weight matrix W in an incremental learning model

---

Require: weight matrix W, fundamental memory vectors $X^{(k)}$ and $Y^{(k)}$, number of pairs k,
    number of iterations d, general output parameter δ, learning parameter η
Ensure: weight matrix W
  if $d = 1$ **then**
    Initialize W as a zero matrix
  end if
  while $k \geq d$ **do**
    Set $r = \text{uniform\_rand}[1, k]$
    Calculate $S^{(k)}$ using Equation (3.5a)
    Calculate $Y^{(k)}$ using Equation (3.5b)
    Calculate orthogonalized vectors u and v by Gram-Schmidt orthogonalization from X and Y
    Calculate weight matrix W using Equation (3.7) with $t = r$
    Set $d \leftarrow d + 1$
  end while

3.3 Fundamental Structures of Multi-directional Association Model

QMAM and IQMAM might consist of L $(L > 2)$ of the layers. Thus, the associated layer is defined as α-th layer, and the other layers are referred to as β-th layers ($\beta = 1, 2, \ldots, L$; $\alpha \neq \beta$). The neurons in each layer continue to be subject to cyclic updating until the layers reach an equilibrium state. Here, let k of the original memory pairs $\left\{ \boldsymbol{X}_{(1)}^{(k)}, \boldsymbol{X}_{(2)}^{(k)}, \ldots, \boldsymbol{X}_{(L)}^{(k)} \right\}$ are stored in QMAM and IQMAM, where $\boldsymbol{X}_{(l)}^{(k)} = \left[x_1, x_2, \ldots, x_{N_{(l)}} \right]$ ($x \in [-1, 1], l = 1, 2, \ldots, L$). N denotes the number of neurons in l-th layer.

3.3.1 Quantum-Inspired Multi-directional Associative Memory

QMAM is formalized as followings;

- α-th layer to β-th layers

$$
\begin{cases}
\boldsymbol{S}_{(\beta)}^{(k)} = \sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^{L} \sum_{j=1}^{N_{(\beta)}} \sum_{i=1}^{M_{(\alpha)}} W_{ij(\alpha\beta)}^{T} x_{i(\alpha)}^{(k)}, & (\beta = 1, 2, \ldots, L; \ \beta \neq \alpha) \quad (3.10\text{a}) \\[2em]
\boldsymbol{X}_{(\beta)}^{(k)} = \phi\left(\boldsymbol{S}_{(\beta)}^{(k)} \right) & (3.10\text{b})
\end{cases}
$$

- β-th layers to α-th layer

$$
\begin{cases}
\boldsymbol{U}_{(\alpha)}^{(k)} = \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{L} \sum_{i=1}^{M_{(\alpha)}} \sum_{j=1}^{N_{(\beta)}} W_{ij(\alpha\beta)} x_{j(\beta)}^{(k)}, & (\alpha = 1, 2, \ldots, L; \ \alpha \neq \beta) \quad (3.11\text{a}) \\[2em]
\boldsymbol{X}_{(\alpha)}^{(k)} = \phi\left(\boldsymbol{U}_{(\alpha)}^{(k)} \right) & (3.11\text{b})
\end{cases}
$$

where, α and β denote the layer numbers ($\alpha \neq \beta$). S and U denote the temporal states of associated patterns. L denotes the number of layers. M and N represent the number of neurons in α-th layer and β-th layer, respectively. Here, $\phi(\cdot)$ denotes an activation function as a step function with bipolar state. The weight connections $W_{ij(\alpha\beta)}^T$ and $W_{ij(\alpha\beta)}$ are defined by Hebb learning as follows;

- α-th layer to β-th layers

$$\boldsymbol{W}_{(\alpha\beta)}^T = \frac{1}{k} \sum_{p=1}^{k} \sum_{j=1}^{N_{(\beta)}} \sum_{i=1}^{M_{(\alpha)}} s_{j(\beta)}^{(p)} s_{i(\alpha)}^{(p)T}. \tag{3.12}$$

- β-th layers to α-th layer

$$\boldsymbol{W}_{(\alpha\beta)} = \frac{1}{k} \sum_{p=1}^{k} \sum_{i=1}^{M_{(\alpha)}} \sum_{j=1}^{N_{(\beta)}} s_{i(\alpha)}^{(p)} s_{j(\beta)}^{(p)T}. \tag{3.13}$$

where, exponential T denotes transpose operation. k denotes the number of memory pairs, and the orthonormalized memory vectors in α-th layer ($s_{(\alpha)}$) and the orthonormalized memory vectors in β-th layer ($s_{(\beta)}$), respectively, that are calculated by Gram-Schmidt orthogonalization as follows; $\boldsymbol{a}_1 = \boldsymbol{A}_1 / \|\boldsymbol{A}_1\| (p=1)$, $\boldsymbol{b}_p = \boldsymbol{A}_p - \sum_{i=p-1}^{k-1} (\boldsymbol{a}_i, \boldsymbol{A}_i) \boldsymbol{a}_i$ and $\boldsymbol{a}_p = \boldsymbol{b}_p / \|\boldsymbol{b}_p\| (2 \leq p \leq k)$, where $\boldsymbol{A}$ denotes memory vector, $\boldsymbol{a}$ and $\boldsymbol{b}$ denote the orthonormalized vector and the orthogonalized vector, respectively.

Similar with the bidirectional model with a batch learning, the above batch learning is summarized as Algorithm 1 in Section 3.2.1.

3.3.2 Quantum-Inspired Incremental Multi-directional Associative Memory

QMAM is formalized as followings;

- α-th layer to β-th layers

$$
\begin{cases}
\boldsymbol{S}_{(\beta)}^{(k)} = \displaystyle\sum_{\substack{\alpha=1 \\ \alpha\neq\beta}}^{L} \sum_{j=1}^{N_{(\beta)}} \sum_{i=1}^{M_{(\alpha)}} W_{ji(\alpha\beta)} x_{i(\alpha)}^{(k)}, \quad (\beta = 1, 2, \ldots, L; \ \beta \neq \alpha) & \text{(3.14a)} \\[2em]
\boldsymbol{Y}_{(\beta)}^{(k)} = (\delta+1)\boldsymbol{S}_{(\beta)(t)}^{(k)} - \delta\boldsymbol{S}_{(\beta)(t)}^{(k)3} & \text{(3.14b)}
\end{cases}
$$

- β-th layers to α-th layer

$$
\begin{cases}
\boldsymbol{U}_{(\alpha)}^{(k)} = \displaystyle\sum_{\substack{\beta=1 \\ \beta\neq\alpha}}^{L} \sum_{i=1}^{M_{(\alpha)}} \sum_{j=1}^{N_{(\beta)}} V_{ij(\alpha\beta)} x_{j(\beta)}^{(k)}, \quad (\alpha = 1, 2, \ldots, L; \ \alpha \neq \beta) & \text{(3.15a)} \\[2em]
\boldsymbol{X}_{(\alpha)}^{(k)} = (\delta+1)\boldsymbol{U}_{(\alpha)(t)}^{(k)} - \delta\boldsymbol{U}_{(\alpha)(t)}^{(k)3} & \text{(3.15b)}
\end{cases}
$$

where, α and β denote the layer numbers ($\alpha \neq \beta$). $\boldsymbol{S}$ and $\boldsymbol{U}$ denote the temporal states of associated patterns. L denotes the number of layers. M and N represent the number of neurons in α-th layer and β-th layer, respectively. Here, δ is a general output parameter, t denotes the number of iterations over the network. The weight matrices $\boldsymbol{W}$ and $\boldsymbol{V}$ are defined as follows;

$$
\boldsymbol{W}_{(t+1)} = \boldsymbol{W}_{(t)} + \eta \left(v_{(\beta)(0)}^{(k)} - v_{(\beta)(t)}^{(k)} \right) \left(u_{(\alpha)(0)}^{(k)} - u_{(\alpha)(t)}^{(k)} \right)^T \tag{3.16}
$$

$$
\boldsymbol{V}_{(t+1)} = \boldsymbol{V}_{(t)} + \eta \left(u_{(\alpha)(0)}^{(k)} - u_{(\alpha)(t)}^{(k)} \right) \left(v_{(\beta)(0)}^{(k)} - v_{(\beta)(t)}^{(k)} \right)^T \tag{3.17}
$$

where, η represents the learning parameter. $u_{(\alpha)(0)}^{(k)}$ and $v_{(\beta)(0)}^{(k)}$ denote the initial inputs which are selected randomly based on a uniform distribution, then, those inputs are iterated t times through the network. $u_{(\alpha)}^{(k)}$ and $v_{(\beta)}^{(k)}$ are orthonormalized vectors of $x_{(\alpha)}^{(k)}$ and $x_{(\beta)}^{(k)}$, respectively, that are calculated by Gram-Schmidt orthogonalization. The learning convergence is guaranteed with the condition of the learning parameter η (Chartier & Boukadoum, 2006b);

$$
\eta < \frac{1}{2(1-2\delta)\text{Max}[M,N]}, \quad \delta \neq 1/2. \tag{3.18}
$$

51

where, δ is a general output parameter, M and N denote the number of neurons in α-th layer and β-th layer, respectively.

Similar with the bidirectional model with an incremental learning, the above incremental learning is summarized as Algorithm 2 in Section 3.2.2.

3.4 Quantum Mechanics for Associative Memory

Superposition and unitarity are key features of quantum mechanics. Superposition can be explained as "multiple states" that exist simultaneously in the quantum system. The evolution of a closed quantum system is described by a unitary transformation. In quantum-inspired associative memory, superposition is satisfied by the decomposed weight matrices based on fuzzy inference. Unitarity is satisfied by rotations between spaces that are spanned by the eigenvectors of decomposed weight matrices.

In general, the multi-directional model is considered as a combination of the multiple bidirectional connections that are depicted in Figure 3.1. In other words, it can be considered that the weight connections between 2 layers in a multi-directional model are equivalent to weight connections in a bidirectional model. Thus, the mathematical proofs for the superposition and unitarity are forced on between α-th layer and β-th layer in the bidirectional/multi-directional model.

(a) Bidirectional Model

(b) Multi-directional Model (3-layers)
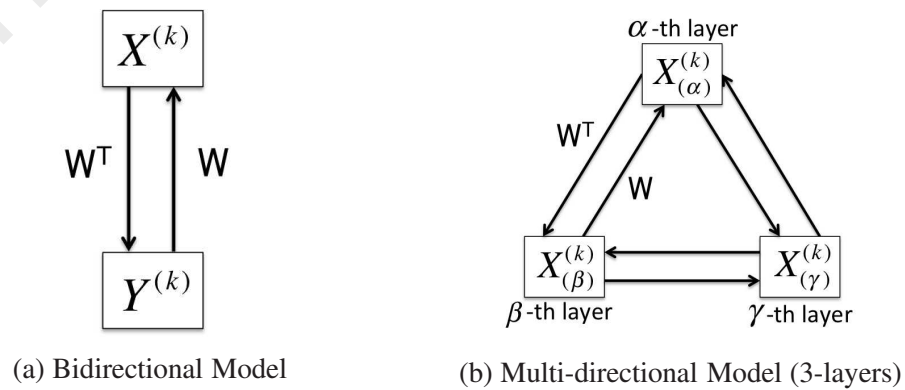
Figure 3.1: Conceptual structures of the hetero-association model. Here, the model (a) can be considered as a part of the model (b).

3.4.1 Fundamentals of Quantum Mechanics

In QM, the state of an isolated quantum system (Q) is represented by a vector $|\psi_{(t)}\rangle$ in a Hilbert space. This vector satisfies Schrödinger's diffusion equation (Griffiths, 1995).

$$i\hbar\frac{d}{dt}|\psi_{(t)}\rangle = H\psi_{(t)}. \tag{3.19}$$

where, Hamiltonian operator (H) gives the total energy of a particle $H = \left(p^2/2m\right) + Ve_{(x)}$. Equation (3.19) shows that a momentum of particle (p) and an energy of particle (E) are diffused in the wave using a probability density, that is proportional to $\left|\psi_{(x,t)}\right|^2$ (a probability that the particle is at position x at time t ($\left|\psi_{(x,t)}\right|^2$)), and an external potential (Ve) is defined as $Ve = -\left(p^2/2m\right) + E$. For $Ve = 0$ or constant, the solution of Equation (3.19) is a superposition of plane waves of the form as;

$$|\psi_{(x,t)}\rangle = e^{i(px-Et)/\hbar}. \tag{3.20}$$

where, subscript i denotes the imaginary unit, x denotes the position of the particle, and Plank's constant ($\hbar$). These results can be applied to harmonic potential when only the basic mode is taken into account. The probability of finding the particle between x and $x + dx$ at time t is given by $P_{(x)}dx = \left|\psi_{(x,t)}\right|^2$. The total probability should equal unity, i.e., $\int_{-\infty}^{\infty}\left|\psi_{(x,t)}\right|^2 dx = 1$. The average position x of the particle is given by;

$$\langle x\rangle = \int_{-\infty}^{\infty} P_{(x)}dx = \int_{-\infty}^{\infty}\left(\psi^* x\psi\right)dx. \tag{3.21}$$

where, a conjugate memory vector of ψ (ψ^*). The wave function $\psi_{(x,t)}$ can be analyzed as a set of orthonormal eigenfunctions in a Hilbert space;

$$\psi_{(x,t)} = \sum_{m=1}^{\infty} c_m\psi_m. \tag{3.22}$$

here, the coefficient c_m indicates the probability of describing the particle's position x at time t using the eigenfunction ψ_m due to the orthonormality of ψ_m, and c_m is given by $c_m = \int_{-\infty}^{\infty} \psi_m \psi^* dx$. Moreover, the eigenvalues and eigenvectors of the quantum operator of position x can be defined as $\psi_m = a_m \psi_m$, where ψ_m is the eigenvector, and a_m is the associated eigenvalue. Using Equations (3.19) and (3.20), the average position of the particle is found using;

$$\langle x \rangle = \sum_{m=1}^{\infty} \|c_m\|^2 a_m. \tag{3.23}$$

where, $\|c_m\|^2$ denotes the probability that the particle's position can be described by the eigenfunction ψ_m, when the position of x is given by the associated eigenvalue a_m. The eigenvalue a_m is chosen with probability $P \propto \|c_m\|^2$.

In the same way, the probability of finding the particle between y and $y + dy$ at time t is given by $P_{(y)} dy = |\psi_{(y,t)}|^2$, from which the following can be derived;

$$\langle y \rangle = \sum_{n=1}^{\infty} \|c_n\|^2 b_n. \tag{3.24}$$

where, $\|c_n\|^2$ denotes the probability of describing the particle's position using the eigenfunction ψ_n.

3.4.2 Hebb-like Learning based on Fuzzy Inference

The elements of the weight matrix w_{ij} are considered as fuzzy variables, and the increase/decrease operation for the element of weight w_{ij} is performed based on the following rules (Tzafestas & Rigatos, 2000);

● Increase

IF $w_{ij(k)}$ is A_1 THEN $w_{ij(k+1)}$ is A_2

IF $w_{ij(k)}$ is A_2 THEN $w_{ij(k+1)}$ is A_3

$\vdots$

IF $w_{ij(k)}$ is A_{n-1} THEN $w_{ij(k+1)}$ is A_n

● Decrease

IF $w_{ij(k)}$ is A_2 THEN $w_{ij(k+1)}$ is A_2

IF $w_{ij(k)}$ is A_3 THEN $w_{ij(k+1)}$ is A_3

$\vdots$

IF $w_{ij(k)}$ is A_n THEN $w_{ij(k+1)}$ is A_{n-1}

Here, A_i ($i=1,\ldots,n$) denotes the fuzzy subsets in which the universe of discourse of the variable w_{ij} is partitioned. The fuzzy sets A_i based on a triangular fuzzifier are depicted as Figure 3.2. The derivation of the fuzzy relational matrices R_n^i and R_n^d apply the min t-norm, which are mentioned in Section 3.4.5. The max-min inference is used, while the defuzzifier is a center of average one.

The above fuzzy rules can be considered as a Hebb-like learning in the case of fundamental memory vectors x_k and y_k can be stated as follows;

$$\text{IF sgn}(x_k^i y_k^j) > 0 \text{ THEN increase } w_{ij}$$

$$\text{IF sgn}(x_k^i y_k^j) < 0 \text{ THEN decrease } w_{ij}$$

The weight learning process with the above fuzzy learning rule is a consequence of a fuzzy weight matrix. Thanks to the fuzzy learning, not a Hebb learning, the weight matrix W of associative memory can be decomposed and satisfied unitarity in the decomposed weight matrices $\bar{W}$, which are detailed in Sections 3.4.4 and 3.4.5.
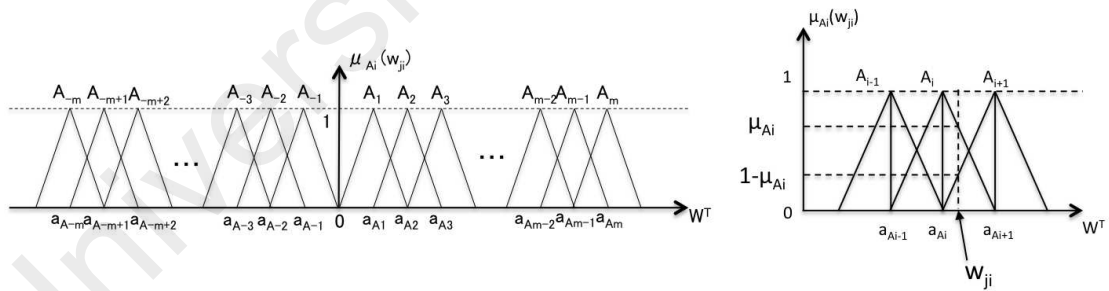
Figure 3.2: Fuzzy subsets for weight matrix in α-th layer.

3.4.3 Similarity between Quantum Features and Fuzzy Inference

Assume that the fuzzy variables x and y belong to a universe of discourse that can be quantized into an infinite number of fuzzy subsets A_i, A_{-i} ($i = 1, 2, \ldots, \infty$) and B_j, B_{-j} ($j = 1, 2, \ldots, \infty$), e.g., the axis of real number R is partitioned into an infinite number of fuzzy subsets with the same space and width. For example, Figure 3.2 shows the fuzzy subsets in α-th layer. The fuzzy subsets have the following properties;

(a) Summation of a fuzzy membership grade (μ, ν) is unity as followings;

- α-th layer to β-th layer

$$\sum_{m=1}^{\infty} \mu_{A_m(x)} + \sum_{m=1}^{\infty} \mu_{A_{-m}(x)} = 1. \tag{3.25}$$

- β-th layer to α-th layer

$$\sum_{n=1}^{\infty} \nu_{B_n(y)} + \sum_{n=1}^{\infty} \nu_{B_{-n}(y)} = 1. \tag{3.26}$$

(b) The center of fuzzy subsets a_m and b_n are belong to A_m and B_n, respectively.

(c) The average values of variables x and y are given by;

- α-th layer to β-th layer

$$\begin{cases} \langle x \rangle = \sum_{m=1}^{\infty} \mu_{A_m(x)} a_m & \left(W_{ij}^T \geq 0 \right) & \text{(3.27a)} \\ \langle x \rangle = \sum_{m=1}^{\infty} \mu_{A_{-m}(x)} a_{-m} & \left(W_{ij}^T < 0 \right) & \text{(3.27b)} \end{cases}$$

- β-th layer to α-th layer

$$\begin{cases} \langle y \rangle = \sum_{n=1}^{\infty} \nu_{B_n(y)} b_n & \left(W_{ij} \geq 0 \right) & \text{(3.28a)} \\ \langle y \rangle = \sum_{n=1}^{\infty} \nu_{B_{-n}(y)} b_n & \left(W_{ij} < 0 \right) & \text{(3.28b)} \end{cases}$$

Rigatos and Tzafestas summarized the details of relationships between fuzzy inference and QM from the above equations as following (Rigatos & Tzafestas, 2006; Rigatos, 2010);

(i) The membership grades $\mu_{A_m(x)}$, $\mu_{A_{-m}(x)}$ correspond to $\|c_m\|^2$, $\|c_{-m}\|^2$.

(ii) The centers a_m and a_{-m} of the fuzzy subsets A_m and A_{-m} represent the eigenvalues a_m and a_{-m}, respectively, of the position operator x.

(iii) The fuzzy subsets A_m and A_{-m} that correspond to the probability of the particle position are given by $|\psi_x|^2$.

(iv) The membership grades $\mu_{A_m(x)}$ and $\mu_{A_{-m}(x)}$ of fuzzy subsets A_m and A_{-m} satisfy the condition $\sum_{m=1}^{\infty} \mu_{A_m(x)} + \sum_{m=1}^{\infty} \mu_{A_{-m}(x)} = 1$. This condition is equivalent to total probability of particle position that equals unity in a quantum system, i.e., $\sum_{-\infty}^{\infty} |\psi_{(x)}|^2 \, dx = 1$.

(v) The outputs of the quantum system are the eigenvalue A_m and A_{-m} that are associated with the eigenvectors ψ_m and ψ_{-m}, with probabilities $P \propto \|c_m\|^2$ and $P \propto \|c_{-m}\|^2$, respectively. In the fuzzy system, the centers of the fuzzy sets A_m and A_{-m}, which become the output of the fuzzy system with probability $P \propto \mu_{A_m(x)}$ and $P \propto \mu_{A_{-m}(x)}$, respectively.

(vi) The particle wave equation can be represented as a vector in a Hilbert space, with orthonormalized vectors ψ_m and ψ_{-m}, while the fuzzy variable can be represented as a vector in a space which is transitioned on the fuzzy subsets A_m and A_{-m}, respectively.

In case of the centers b_n and b_{-n} of the fuzzy subsets B_n and B_{-n}, respectively, satisfy the above specified relationships.

3.4.4 Existence of Superposition in the Weight Matrix

The decomposition process of weight matrices between α-th layer and β-th layer is shown.

- α-th layer to β-th layer

Consider the element w_{ji} of weight matrix W^T. Because of the strong fuzzy partition, this weight element belongs to two adjacent fuzzy subsets A_i and A_{i+1} (Figure 3.2). The corresponding centers of the fuzzy subsets are a_{ji}^i and a_{ji}^{i+1}, and the associated membership grades are $\mu_{ji} = \mu_{A_i}$ and $1 - \mu_{ji} = \mu_{A_{i+1}}$, respectively. Therefore, w_{ji} can be represented by the set of $\left\{\mu_{ji}, a_{ji}^{A_i}\right\}$ and $\left\{1 - \mu_{ji}, a_{ji}^{A_{i+1}}\right\}$. The matrices that consist of the membership grades μ_{ji} and $1 - \mu_{ji}$ are generated from the possible combinations of membership grades of each element of weight matrix. Therefore, the weight matrix W^T can be decomposed into the set of superposition matrices $\bar{W}_i^T$ $\left(i = 1, 2, \ldots, 2^{NM}\right)$ as Table 3.1.

Table 3.1: Possible combinations of membership grades and corresponding centers of fuzzy subset in W^T.

	(Membership grade, Center of the fuzzy subset)			
$\bar{W}_1^T$:	$(\mu_{11},\ a_{11}^{A_i})$	$\cdots$	$(\mu_{n(m-1)},\ a_{n(m-1)}^{A_i})$	$(\mu_{nm},\ a_{nm}^{A_i})$
$\bar{W}_2^T$:	$(\mu_{11},\ a_{11}^{A_i})$	$\cdots$	$(\mu_{n(m-1)},\ a_{n(m-1)}^{A_i})$	$(1 - \mu_{nm},\ a_{nm}^{A_{i+1}})$
$\bar{W}_3^T$:	$(\mu_{11},\ a_{11}^{A_i})$	$\cdots$	$(1 - \mu_{n(m-1)},\ a_{n(m-1)}^{A_{i+1}})$	$(\mu_{nm},\ a_{nm}^{A_i})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{W}_{2^{NM}-1}^T$:	$(1 - \mu_{11},\ a_{11}^{A_{i+1}})$	$\cdots$	$(1 - \mu_{n(m-1)},\ a_{n(m-1)}^{A_{i+1}})$	$(\mu_{nm},\ a_{nm}^{A_i})$
$\bar{W}_{2^{NM}}^T$:	$(1 - \mu_{11},\ a_{11}^{A_{i+1}})$	$\cdots$	$(1 - \mu_{n(m-1)},\ a_{n(m-1)}^{A_{i+1}})$	$(1 - \mu_{nm},\ a_{nm}^{A_{i+1}})$

Therefore, the weight matrix W^T can be decomposed as;

$$W^T = \sum_{i=1}^{2^{MN}} \mu_i \bar{W}_i^T. \tag{3.29}$$

It can be regarded as superposition.

In Figure 3.2, $A_1, A_2, \ldots, A_{m-1}, A_m$ and $A_{-1}, A_{-2}, \ldots, A_{-m+1}, A_{-m}$ are fuzzy subsets in the universe of discourse of the variable w_{ji}. The sets A_i and A_{-i} are selected in order

to have the same spread and satisfy the strong fuzzy partition equalities $\sum_{i=1}^{m} \mu_{A_m(x)} + \sum_{i=1}^{m} \mu_{A_{-m}(x)} = 1$.

- β-th layer to α-th layer

As in the case of α-th layer to β-th layer, consider the element w_{ij} of weight matrix W. The elements of w_{ij} can be represented by the set of $\left\{ v_{ij}, b_{ij}^{B_j} \right\}$ and $\left\{ 1 - v_{ij}, v_{ij}^{B_{j+1}} \right\}$. Thus, the weight matrix W can be decomposed into the set of superposition matrices $\bar{W}_j$ $\left(j = 1, 2, \ldots, 2^{MN} \right)$ as Table 3.2.

Table 3.2: Possible combinations of membership grades and corresponding centers of fuzzy subset in W.

	(Membership grade, Center of the fuzzy subset)			
$\bar{W}_1:$	$(v_{11}, b_{11}^{B_j})$	$\cdots$	$(v_{m(n-1)}, b_{m(n-1)}^{B_j})$	$(v_{mn}, b_{mn}^{B_j})$
$\bar{W}_2:$	$(v_{11}, b_{11}^{B_j})$	$\cdots$	$(v_{m(n-1)}, b_{m(n-1)}^{B_j})$	$(1 - v_{mn}, b_{mn}^{B_{j+1}})$
$\bar{W}_3:$	$(v_{11}, b_{11}^{B_j})$	$\cdots$	$(1 - v_{m(n-1)}, b_{m(n-1)}^{B_{j+1}})$	$(v_{mn}, b_{mn}^{B_j})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{W}_{2^{NM}-1}:$	$(1 - v_{11}, b_{11}^{B_{j+1}})$	$\cdots$	$(1 - v_{m(n-1)}, b_{m(n-1)}^{B_{j+1}})$	$(v_{mn}, b_{mn}^{B_j})$
$\bar{W}_{2^{MN}}:$	$(1 - v_{11}, b_{11}^{B_{j+1}})$	$\cdots$	$(1 - v_{m(n-1)}, b_{m(n-1)}^{B_{j+1}})$	$(1 - v_{mn}, b_{mn}^{B_{j+1}})$

Therefore, as in the case of α-th layer to β-th layer, the weight matrix W can be decomposed as;

$$W = \sum_{j=1}^{2^{NM}} v_j \bar{W}_j. \tag{3.30}$$

For example, consider the m by n weight matrix W. Here, the elements of weight w_{mn} are considered fuzzy variables. The possible combinations of membership grades v_{mn} and center of fuzzy subsets $b_{mn}^{B_i}$ are represented by the decomposed weight matrices $\bar{W}_k$ $(k = 1, 2, \ldots, 2^{mn})$ as Equation (3.31);

$$
\begin{aligned}
W \quad &= \frac{v_{11} + v_{12} + \cdots + v_{mn}}{NM \cdot 2^{NM-1}}
\begin{bmatrix}
b_{11}^{B_i} & b_{12}^{B_i} & \dots & b_{1n}^{B_i} \\[4pt]
b_{21}^{B_i} & b_{22}^{B_i} & \dots & b_{2n}^{B_i} \\[4pt]
\vdots & \vdots & \ddots & \vdots \\[4pt]
b_{m1}^{B_i} & b_{m2}^{B_i} & \dots & b_{mn}^{B_i}
\end{bmatrix} \\[10pt]
&+ \frac{v_{12} + \cdots + v_{mn} - v_{11} + 1}{NM \cdot 2^{NM-1}}
\begin{bmatrix}
b_{11}^{B_{i+1}} & b_{12}^{B_i} & \dots & b_{1n}^{B_i} \\[4pt]
b_{21}^{B_i} & b_{22}^{B_i} & \dots & b_{2n}^{B_i} \\[4pt]
\vdots & \vdots & \ddots & \vdots \\[4pt]
b_{m1}^{B_i} & b_{m2}^{B_i} & \dots & b_{mn}^{B_i}
\end{bmatrix} \\[10pt]
&+ \frac{v_{11} + v_{13} + \cdots + v_{mn} - v_{12} + 1}{NM \cdot 2^{NM-1}}
\begin{bmatrix}
b_{11}^{B_i} & b_{12}^{B_{i+1}} & \dots & b_{1n}^{B_i} \\[4pt]
b_{21}^{B_i} & b_{22}^{B_i} & \dots & b_{2n}^{B_i} \\[4pt]
\vdots & \vdots & \ddots & \vdots \\[4pt]
b_{m1}^{B_i} & b_{m2}^{B_i} & \dots & b_{mn}^{B_i}
\end{bmatrix} \\[10pt]
&+ \cdots + \frac{-v_{11} - v_{12} \cdots - v_{mn} + NM}{NM \cdot 2^{NM-1}}
\begin{bmatrix}
b_{11}^{B_{i+1}} & b_{12}^{B_{i+1}} & \dots & b_{1n}^{B_{i+1}} \\[4pt]
b_{21}^{B_{i+1}} & b_{22}^{B_{i+1}} & \dots & b_{2n}^{B_{i+1}} \\[4pt]
\vdots & \vdots & \ddots & \vdots \\[4pt]
b_{m1}^{B_{i+1}} & b_{m2}^{B_{i+1}} & \dots & b_{mn}^{B_{i+1}}
\end{bmatrix}.
\end{aligned}
\tag{3.31}
$$

where, the elements in matrices that are denoted by N_i and associated $\|L_1\|$ are calculated. Each L_1 norm is divided by the number of elements in matrices N_i. It can be regarded as superposition.

Lemma 1. *The $\|L_1\|$ of the matrices N_i as $\sum_{i=1}^{M}\sum_{j=1}^{N}|v_{ij}|$ divided by the number of elements in matrices, i.e., NM and by $2^{(NM-1)}$, where M and N are the number of neurons in α-th layer and β-th layer, respectively, equals unity.*

$$\frac{1}{NM \cdot 2^{(NM-1)}} \sum_{i=1}^{M}\sum_{j=1}^{N}|v_{ij}| = 1. \tag{3.32}$$

Proof. There are 2^{NM} matrices N_i. Because of the same width triangular fuzzy subsets, two matrices N_i and N_j are obtained as $v(w_{ij})$ and $1 - v(w_{ij})$, respectively. Thus, the sum of the corresponding L_1 norms $\|N_i\| + \|N_j\|$ normalized by the number of elements equals unity. This normalization procedure can be used to derive the membership grade of the weight matrices $\bar{W}_i$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

3.4.5 Unitary Operators in Fuzzy Inference

The following theorems interpret the function of "increase" and "decrease" operators from a quantum mechanics point of view (Rigatos & Tzafestas, 2002). Here, the t-norm is used to derive the fuzzy relational matrices R_m^i and R_m^d (α-th layer), and R_n^i and R_n^d (β-th layer), where the exponentials i and d denote "*increase*" and "*decrease*", respectively. These matrices have the following properties;

- α-th layer to β-th layer

$$\begin{cases} A_m = R_m^i \circ A_{m-1}, \ A_{m-1} = R_m^d \circ A_m & \left(W_{ij}^T \geq 0 \right) & \text{(3.33a)} \\[2mm] A_{-m} = R_{-m}^i \circ A_{-m+1}, \ A_{-m+1} = R_{-m}^d \circ A_{-m} & \left(W_{ij}^T < 0 \right) & \text{(3.33b)} \end{cases}$$

- β-th layer to α-th layer

$$\begin{cases} B_n = R_n^i \circ B_{n-1}, \ B_{n-1} = R_n^d \circ B_n & \left(W_{ij}^T \geq 0 \right) & \text{(3.34a)} \\[2mm] B_{-n} = R_{-n}^i \circ B_{-n+1}, \ B_{-n+1} = R_{-n}^d \circ B_{-n} & \left(W_{ij}^T < 0 \right) & \text{(3.34b)} \end{cases}$$

Theorem 1. *The increase and decrease fuzzy relational operators described in the rule base are unitary.*

Proof.

- α-th layer to β-th layer

The fuzzy relational matrices $R_i \left(W_{ij}^T \geq 0 \right)$ and $R_{-i} \left(W_{ij}^T < 0 \right)$ used by the increase and decrease operators, respectively, satisfy the following fuzzy relational equations;

(1) Increase mode:

$$A_2 = R_1^i \circ A_1, \ A_3 = R_2^i \circ A_2, \ldots, \ A_m = R_{m-1}^i \circ A_{m-1} \qquad \left(W_{ij}^T \geq 0 \right) \text{(3.35a)}$$

$$A_{-2} = R_{-1}^i \circ A_{-1}, \ A_{-3} = R_{-2}^i \circ A_{-2}, \ldots, \ A_{-m} = R_{-m+1}^i \circ A_{-m+1} \quad \left(W_{ij}^T < 0 \right) \text{(3.35b)}$$

(2) Decrease mode:

$$A_1 = R_1^d \circ A_2, \ A_2 = R_2^d \circ A_3, \ldots, \ A_{m-1} = R_{m-1}^d \circ A_m \qquad \left(W_{ij}^T \geq 0 \right) \text{(3.36a)}$$

$$A_{-1} = R_{-1}^d \circ A_{-2}, \ A_{-2} = R_{-2}^d \circ A_{-3}, \ldots, \ A_{-m+1} = R_{-m+1}^d \circ A_{-m} \quad \left(W_{ij}^T < 0 \right) \text{(3.36b)}$$

In both cases, "$\circ$" denotes the max-min composition. Here, substituting $A_{m-1} = R_{m-1}^d \circ A_m$ for $A_m = R_{m-1}^i \circ A_{m-1}$, $A_m = R_{m-1}^i \circ \left(R_{m-1}^d \circ A_m \right)$ is obtained. Similarly, $A_{-m} = R_{-m+1}^i \circ \left(R_{-m+1}^d \circ A_{-m} \right)$ is obtained. Because of the associativity of the max-min operator, it derives $A_m = \left(R_{m-1}^i \circ R_{m-1}^d \right) \circ A_m$, and $A_{-m} = \left(R_{-m+1}^i \circ R_{-m+1}^d \right) \circ A_{-m}$, respectively. Therefore, it is maintained as following;

$$\begin{cases} \left(R_{m-1}^i \circ R_{m-1}^d \right) = I & \text{(3.37a)} \\ \\ \left(R_{-m+1}^i \circ R_{-m+1}^d \right) = I & \text{(3.37b)} \end{cases}$$

Similarly, substituting $A_m = R_{m-1}^i \circ A_{m-1}$ in $A_{m-1} = R_{m-1}^d \circ A_m$, and $A_{-m} = R_{-m+1}^i \circ A_{-m+1}$ in $A_{-m+1} = R_{-m+1}^d \circ A_{-m}$, $A_{m-1} = \left(R_{m-1}^d \circ R_{m-1}^i \right) \circ A_{m-1}$ and $A_{-m+1} = \left(R_{-m+1}^d \right)$

$\circ R^i_{-m+1}$ $\circ A_{-m+1}$, respectively., will be obtained. Therefore, it is maintained as follows;

$$\begin{cases} \left(R^d_{m-1} \circ R^i_{m-1} \right) = I & \text{(3.38a)} \\ \left(R^d_{-m+1} \circ R^i_{-m+1} \right) = I & \text{(3.38b)} \end{cases}$$

Furthermore, here, Mandani's inference system (Kosko, 1991b) is applied to generate matrices R^i_{m-1} and R^d_{m-1}, R^i_{-m+1} and R^d_{-m+1}. Thus, the following relationships can be established;

$$\begin{cases} R^d_{m-1} = \left(R^i_{m-1} \right)^T & \text{(3.39a)} \\ R^d_{-m+1} = \left(R^i_{-m+1} \right)^T & \text{(3.39b)} \end{cases}$$

The following can be derived from Equations (3.37a) and (3.38a), Equations (3.37b) and (3.38b), respectively;

$$\begin{cases} \left(R^d_{m-1} \right)^{-1} = \left(R^i_{m-1} \right)^T & \text{(3.40a)} \\ \left(R^d_{-m+1} \right)^{-1} = \left(R^i_{-m+1} \right)^T & \text{(3.40b)} \end{cases}$$

Therefore, the fuzzy relational matrices R_{m-1} and R_{-m+1}, with increase and decrease operators for α-th layer to β-th layer are unitary.

- β-th layer to α-th layer

As in the case of α-th layer to β-th layer, the fuzzy relational matrices $R_j \left(W_{ij} \geq 0 \right)$, $R_{-j} \left(W_{ij} < 0 \right)$, used by the increase and decrease fuzzy operators, satisfy fuzzy relational equations. Similar to the case of α-th layer to β-th layer, the following conditions will be satisfied;

$$\begin{cases} \left(R^d_{n-1} \right)^{-1} = \left(R^i_{n-1} \right)^T & \text{(3.41a)} \\ \left(R^d_{-n+1} \right)^{-1} = \left(R^i_{-n+1} \right)^T & \text{(3.41b)} \end{cases}$$

Therefore, the fuzzy relational matrices R_{n-1} and R_{-n+1}, with increase and decrease operators for β-th layer to α-th layer are unitary. $\qquad\square$

3.4.6 Evolution of Eigenvector Spaces via Unitary Rotations

Here, it will be shown that the transition between the vector spaces that are associated with the decomposed matrices $\bar{W}^T$ and $\bar{W}$ are described by unitary matrix.

Theorem 2. *The transition between the spaces that are spanned by the eigenvectors of the weight matrices $\bar{W}^T$, and $\bar{W}$ are unitary operators.*

Proof.

- α-th layer to β-th layer

Let w_i, x_i, y_i, z_i and w_j, x_j, y_j, z_j be the elements of the vector which span the spaces associated with the matrices $\bar{W}_i^T$ and $\bar{W}_j^T$, respectively. Here, the memory vector p can be described in each space as $p = [p_{w_i}, p_{x_i}, p_{y_i}, p_{z_i}]^T$ and $p = [p_{w_j}, p_{x_j}, p_{y_j}, p_{z_j}]^T$. The transition from the space $\bar{W}_i^T \rightarrow \{w_i, x_i, y_i, z_i\}$ to the space $\bar{W}_j^T \rightarrow \{w_j, x_j, y_j, z_j\}$ is represented by the matrix R. Here, $p_{\bar{W}_i^T} = p_{w_i} w_i + p_{x_i} x_i + p_{y_i} y_i + p_{z_i} z_i$ and $p_{\bar{W}_j^T} = p_{w_j} w_j + p_{x_j} x_j + p_{y_j} y_j + p_{z_j} z_j$ are given by elements of the vector and memory vectors. In addition, the following equations can be defined;

$$p_{\bar{W}_i^T} = R \cdot p_{\bar{W}_j^T}$$

$$\Rightarrow \begin{bmatrix} p_{w_i} \\ p_{x_i} \\ p_{y_i} \\ p_{z_i} \end{bmatrix} = \begin{bmatrix} w_i w_j & w_i x_j & w_i y_j & w_i z_j \\ x_i w_j & x_i x_j & x_i y_j & x_i z_j \\ y_i w_j & y_i x_j & y_i y_j & y_i z_j \\ z_i w_j & z_i x_j & z_i y_j & z_i z_j \end{bmatrix} \begin{bmatrix} p_{w_j} \\ p_{x_j} \\ p_{y_j} \\ p_{z_j} \end{bmatrix}. \tag{3.42}$$

Similarly, considering the transition from $p_{\bar{W}_j^T}$ to $p_{\bar{W}_i^T}$ as described $p_{\bar{W}_i^T} = Q \cdot p_{\bar{W}_j^T}$. Since "dot products" are commutative, one obtains $Q = R^{-1} = R^T$. Therefore, the transition from the space $\bar{W}_j^T$ to the space $\bar{W}_i^T$ is described by unitary operators;

$$QR = R^T R = R^{-1} R = I. \tag{3.43}$$

• β-th layer to α-th layer

Let w_j, x_j, y_j, z_j and w_i, x_i, y_i, z_i be the elements of the vector which span the spaces associated with the matrices $\bar{W}_j$ and $\bar{W}_i$, respectively. Here, the memory vector s can be described in each space as $s = \begin{bmatrix} s_{w_j}, s_{x_j}, s_{y_j}, s_{z_j} \end{bmatrix}^T$ and $s = \begin{bmatrix} s_{w_i}, s_{x_i}, s_{y_i}, s_{z_i} \end{bmatrix}^T$. The transition from the space $\bar{W}_j \to \{w_j, x_j, y_j, z_j\}$ to the space $\bar{W}_i \to \{w_i, x_i, y_i, z_i\}$ is represented by the matrix V. Here, $s_{\bar{W}_j} = s_{w_j} w_j + s_{x_j} x_j + s_{y_j} y_j + s_{z_j} z_j$ and $s_{\bar{W}_i} = s_{w_i} w_i + s_{x_i} x_i + s_{y_i} y_i + s_{z_i} z_i$ are given by elements of the vector and memory vectors. In addition, the following equations can be defined;

$$s_{\bar{W}_i} = V \cdot s_{\bar{W}_j}$$

$$\Rightarrow \begin{bmatrix} s_{w_j} \\ s_{x_j} \\ s_{y_j} \\ s_{z_j} \end{bmatrix} = \begin{bmatrix} w_j w_i & w_j x_i & w_j y_i & w_j z_i \\ x_j w_i & x_j x_i & x_j y_i & x_j z_i \\ y_j w_i & y_j x_i & y_j y_i & y_j z_i \\ z_j w_i & z_j x_i & z_j y_i & z_j z_i \end{bmatrix} \begin{bmatrix} s_{w_i} \\ s_{x_i} \\ s_{y_i} \\ s_{z_i} \end{bmatrix}. \tag{3.44}$$

Similarly, considering the transition from $s_{\bar{W}_i}$ to $s_{\bar{W}_j}$ as described $s_{\bar{W}_j} = U \cdot s_{\bar{W}_i}$. Since "dot products" are commutative, one obtains $V = U^{-1} = U^T$. Therefore, the transition from the space $\bar{W}_i$ to the space $\bar{W}_j$ is described by unitary operators;

$$VU = U^T U = U^{-1} U = I. \tag{3.45}$$

$\square$

Theoretically, the weight matrix in the incremental model satisfies the features of QM, because of the weight is defined in the basis of Hebb/anti-Hebb learning.

3.5 Numerical Example

The numerical examples of superposition and unitarity based on Section 3.4 is presented. Through the examples in this section, the fundamental condition is defined. Here, the following binary vectors x_i and y_j $(i, j = 1, 2, 3, 4)$ will be set as fundamental memory vectors for quantum-inspired hetero-association model. Even if the model has two fundamental weight matrices W^T in α-th layer to β-th layer and W in β-th layer to α-th layer, the weight matrices W^T and W are linked with a transposition relationship. Therefore, if the weight W is satisfied with superposition and unitarity, the weight W^T also satisfies these properties. Thus, the following examples are performed with the weight W^T.

Firstly, the fundamental memory vectors x_i and y_j $(i, j = 1, 2, 3, 4)$ are described.

$$
x_1 = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \quad x_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}. \tag{3.46}
$$

$$
y_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad y_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \quad y_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \tag{3.47}
$$

In general, if a high dimensional vector space is chosen randomly, the vector space will exhibit orthogonality. However, the dimensions of memory vectors are low, therefore, to have orthogonality, Gram-Schmidt orthogonalization is applied to fundamental memory vectors x_i and y_i. Here, an orthonormalized vector will be utilized to calculate the weight matrix W, and the weight matrix is calculated by Hebb-like learning.

The Gram-Schmidt orthogonalization gives the orthonormalized vectors $u_{(i)}$ and $v_{(i)}$ from memory vectors x_i and y_i, respectively, as follows;

$$
u_{x_1} = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}, \quad
u_{x_2} = \begin{bmatrix} -0.5 \\ -0.5 \\ 0.5 \\ 0.5 \end{bmatrix}, \quad
u_{x_3} = \begin{bmatrix} -0.5 \\ 0.5 \\ 0.5 \\ -0.5 \end{bmatrix}. \tag{3.48}
$$

$$
v_{y_1} = \begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{bmatrix}, \quad
v_{y_2} = \begin{bmatrix} -0.5 \\ 0.5 \\ -0.5 \\ 0.5 \end{bmatrix}, \quad
v_{y_3} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}. \tag{3.49}
$$

3.5.1 Superposition of the Weight Matrix

The weight matrix W is defined by Hebb learning with orthonormalized memory vectors u_x and u_y;

$$
W = \begin{bmatrix}
-0.0833 & -0.2500 & 0.0833 & -0.0833 \\
0.0833 & -0.0833 & 0.2500 & 0.0833 \\
-0.0833 & 0.0833 & 0.0833 & 0.2500 \\
-0.2500 & -0.0833 & -0.0833 & 0.0833
\end{bmatrix}. \tag{3.50}
$$

According to Equation (3.30), Equation (3.50) is regarded as a composition weight matrix. Here, it can be considered that the number of elements in weight matrix decreases.

Because of the duplication of the value in the weight matrix, it can be regarded as;

$$
W = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \\ w_5 & w_6 & w_7 & w_8 \\ w_9 & w_{10} & w_{11} & w_{12} \\ w_{13} & w_{14} & w_{15} & w_{16} \end{bmatrix}
$$

$$
\Rightarrow W = \begin{bmatrix} w_1 & w_2 & w_3 & w_1 \\ w_3 & w_1 & w_4 & w_3 \\ w_1 & w_3 & w_3 & w_4 \\ w_2 & w_1 & w_1 & w_3 \end{bmatrix} \tag{3.51}
$$

Therefore, the number of possible combinations are reduced from $2^{NM} = 2^{16}$ to 2^4 (non duplication). Here, elements in Equation (3.50) are considered as the universe of discourse of the fuzzy variables depicted in Figure 3.3.
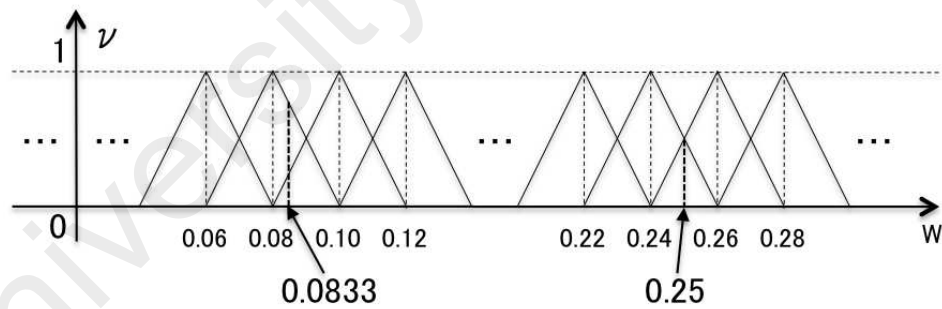
Figure 3.3: Derivation of membership grade from element of weight matrix.

Based on Equation (3.31), Equation (3.50) can be described as Equation (3.52);

$$
W = \frac{0.165 + 0.165 + \cdots + 0.165}{16 \cdot 2^3}
\begin{bmatrix}
-0.10 & 0.10 & -0.10 & -0.24 \\
-0.24 & -0.10 & 0.10 & -0.10 \\
0.10 & 0.26 & 0.10 & -0.10 \\
-0.10 & 0.10 & 0.26 & 0.10
\end{bmatrix}
$$

$$
+ \frac{0.165 + \cdots + 0.165 - 0.835 + 1}{16 \cdot 2^3}
\begin{bmatrix}
-0.08 & 0.10 & -0.08 & -0.24 \\
-0.24 & -0.08 & 0.10 & -0.08 \\
0.10 & 0.26 & 0.10 & -0.08 \\
-0.08 & 0.10 & 0.26 & 0.10
\end{bmatrix}
$$

$$
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.52)
$$

$$
+ \frac{0.165 + 0.165 + \cdots - 0.835 + 1}{16 \cdot 2^3}
\begin{bmatrix}
-0.10 & 0.08 & -0.10 & -0.24 \\
-0.24 & -0.10 & 0.08 & -0.10 \\
0.08 & 0.26 & 0.08 & -0.10 \\
-0.10 & 0.08 & 0.26 & 0.08
\end{bmatrix}
$$

$$
+ \cdots + \frac{-0.835 - 0.835 \cdots - 0.835 + 16}{16 \cdot 2^3}
\begin{bmatrix}
-0.08 & 0.08 & -0.08 & -0.26 \\
-0.26 & -0.08 & 0.08 & -0.08 \\
0.08 & 0.24 & 0.08 & -0.08 \\
-0.08 & 0.08 & 0.24 & 0.08
\end{bmatrix}.
$$

where, each membership grade v is calculated as Equation (3.31), that are;

$$
v_1 = 0.0311, \quad v_2 = 0.0625, \quad v_3 = 0.0625, \quad v_4 = 0.0311,
$$

$$
v_5 = 0.0311, \quad v_6 = 0.0939, \quad v_7 = 0.0625, \quad v_8 = 0.0625,
$$

$$
v_9 = 0.0625, \quad v_{10} = 0.0625, \quad v_{11} = 0.0311, \quad v_{12} = 0.0939,
$$

$$
v_{13} = 0.0939, \quad v_{14} = 0.0625, \quad v_{15} = 0.0625, \quad v_{16} = 0.0939.
$$

The above membership grades v of each matrix $\bar{W}$ are satisfied with $\sum_{j=1}^{2^4} v_j = 1$, which is provided by Lemma. 1 in Section 3.4.4.

3.5.2 Unitarity in the Decomposed Weight Matrix

Unitarity is maintained as the fuzzy relational matrix R that is defined by eigenvectors of decomposed weight matrices $\bar{W}_i$. Here, as examples, unitarity between vector spaces $\bar{W}_1$, $\bar{W}_8$ and $\bar{W}_{16}$ will be shown. Tables 3.3 shows eigenvalues λ and eigenvectors ε of decomposed weight matrices $\bar{W}_i$.

Table 3.3: Eigenvalues and eigenvectors of decomposed weights $\bar{W}_1$, $\bar{W}_8$ and $\bar{W}_{16}$.

	Eigenvalue		Eigenvector			
$\bar{W}_1$:	$\lambda_1 =$	-0.0101	$\varepsilon_1 = [$ -0.765	0.300	-0.355	$-0.445]^T$
	$\lambda_2 =$	-0.0046	$\varepsilon_2 = [$ -0.141	-0.877	0.118	$-0.443]^T$
	$\lambda_3 =$	0.0046	$\varepsilon_3 = [$ 0.443	-0.118	-0.877	$-0.141]^T$
	$\lambda_4 =$	0.0101	$\varepsilon_4 = [$ -0.445	-0.355	-0.300	$0.765]^T$
$\bar{W}_8$:	$\lambda_1 =$	-0.0188	$\varepsilon_1 = [$ -0.762	0.327	-0.325	$-0.456]^T$
	$\lambda_2 =$	-0.0075	$\varepsilon_2 = [$ -0.176	-0.863	0.167	$-0.444]^T$
	$\lambda_3 =$	0.0089	$\varepsilon_3 = [$ 0.383	-0.183	-0.896	$-0.133]^T$
	$\lambda_4 =$	0.0199	$\varepsilon_4 = [$ -0.492	-0.340	-0.254	$0.760]^T$
$\bar{W}_{16}$:	$\lambda_1 =$	-0.0293	$\varepsilon_1 = [$ 0.767	-0.186	0.303	$0.533]^T$
	$\lambda_2 =$	-0.0116	$\varepsilon_2 = [$ 0.056	0.913	-0.199	$0.352]^T$
	$\lambda_3 =$	0.0116	$\varepsilon_3 = [$ 0.352	-0.199	-0.913	$-0.056]^T$
	$\lambda_4 =$	0.0293	$\varepsilon_4 = [$ -0.533	-0.304	-0.186	$0.767]^T$

According to Theorem 2, unitarity is maintained in the transition matrix between the spaces that are spanned by the eigenvectors of decomposed weight matrices $\bar{W}_i$ and $\bar{W}_j$. For example, it will be shown that the transition matrix R which performs from $\bar{W}_1$ to $\bar{W}_{16}$ is calculated by Equation (3.44) as a following transition matrix $R_{\bar{W}_1 \bar{W}_{16}}$;

$$R_{\bar{W}_1 \bar{W}_{16}} = \begin{bmatrix} \varepsilon_1^{(\bar{W}_1)}\varepsilon_1^{(\bar{W}_{16})} & \varepsilon_1^{(\bar{W}_1)}\varepsilon_2^{(\bar{W}_{16})} & \varepsilon_1^{(\bar{W}_1)}\varepsilon_3^{(\bar{W}_{16})} & \varepsilon_1^{(\bar{W}_1)}\varepsilon_4^{(\bar{W}_{16})} \\ \varepsilon_2^{(\bar{W}_1)}\varepsilon_1^{(\bar{W}_{16})} & \varepsilon_2^{(\bar{W}_1)}\varepsilon_2^{(\bar{W}_{16})} & \varepsilon_2^{(\bar{W}_1)}\varepsilon_3^{(\bar{W}_{16})} & \varepsilon_2^{(\bar{W}_1)}\varepsilon_4^{(\bar{W}_{16})} \\ \varepsilon_3^{(\bar{W}_1)}\varepsilon_1^{(\bar{W}_{16})} & \varepsilon_3^{(\bar{W}_1)}\varepsilon_2^{(\bar{W}_{16})} & \varepsilon_3^{(\bar{W}_1)}\varepsilon_3^{(\bar{W}_{16})} & \varepsilon_3^{(\bar{W}_1)}\varepsilon_4^{(\bar{W}_{16})} \\ \varepsilon_4^{(\bar{W}_1)}\varepsilon_1^{(\bar{W}_{16})} & \varepsilon_4^{(\bar{W}_1)}\varepsilon_2^{(\bar{W}_{16})} & \varepsilon_4^{(\bar{W}_1)}\varepsilon_3^{(\bar{W}_{16})} & \varepsilon_4^{(\bar{W}_1)}\varepsilon_4^{(\bar{W}_{16})} \end{bmatrix} \tag{3.53}$$

Therefore, from the eigenvectors in Table 3.3, the following matrix is calculated;

$$R_{\bar{W}_1 \bar{W}_{16}} = \begin{bmatrix} -0.202 & 0.060 & -0.526 & -0.824 \\ 0.330 & -0.725 & 0.440 & -0.414 \\ -0.414 & 0.440 & 0.725 & -0.330 \\ -0.824 & -0.526 & -0.060 & 0.202 \end{bmatrix}. \tag{3.54}$$

It holds;

$$R_{\bar{W}_1 \bar{W}_{16}} R_{\bar{W}_1 \bar{W}_{16}}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.55}$$

Similarly, the transition matrix R which performs from $\bar{W}_8$ to $\bar{W}_{16}$ is presented;

$$R_{\bar{W}_8 \bar{W}_{16}} = \begin{bmatrix} -0.197 & 0.054 & -0.455 & -0.867 \\ 0.321 & -0.709 & 0.501 & -0.380 \\ -0.419 & 0.468 & 0.733 & -0.260 \\ -0.826 & -0.525 & -0.070 & 0.191 \end{bmatrix}. \tag{3.56}$$

It holds;

$$R_{\bar{W}_8 \bar{W}_{16}} R_{\bar{W}_8 \bar{W}_{16}}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.57}$$

In principle, unitarity is satisfied with any combinations of eigenvectors space.

3.6 Simulation Experiments

This section presents the evaluation of quantum-inspired bidirectional models (QBAM, IQBAM) and multi-directional models (QMAM, IQMAM) in terms of memory capacity and recall reliability.

Memory capacity is an important element for the performance of associative memory. In general, memory capacity is sensitive to the number of neurons in layer, therefore, if the number of neurons in layer increase, memory capacity will also increase correspondingly. In addition, the differences of the number of neurons in each layer also affect to the performance of memory capacity.

Noise tolerance is another significant property in associative memory. In general, "noise" is roughly defined into two types in the association models. One is because the similarity of the stored patterns. The other is due to the noise being contained in the stored patterns. In this section, the noise contained in the stored patterns will be discussed. Here, the initial setting of each layer is as follows: the number of stored pairs is 30, the layer 1 is set the information that contain a certain ratio $(0 - 100[\%])$ of salt and pepper noise, and others are assigned the random bipolar patterns.

Here, several representative models in AM are compared as shown in Table 3.4.

Table 3.4: Conventional models for simulation experiment.

	Learning Type	Comparison Model
Bidirectional Model	Batch	BAM (Kosko, 1987a), GBAM (Shi et al., 1998), QL-BAM (Hattori et al., 1994)
	Incremental	IBAM (Chartier & Boukadoum, 2006a)
Multi-directional Model	Batch	MAM (Hagiwara, 1990), QL-MAM (Hattori & Hagiwara, 1995)
	Incremental	IMAM (Chartier & Boukadoum, 2011)

In Table 3.4, the learning algorithms in QL-BAM, QL-MAM, IBAM, IMAM, IQBAM and IQMAM have the parameters. These are relating the learning ratio, convergence ratio

and network stability. Through the simulation experiment, the parameters in each model are determined as shown in Table 3.5.

Table 3.5: Learning parameters for the simulation experiment.

Model	Learning parameter	
QL-BAM,	relaxation factor (λ) :	1.9
QL-MAM	normalizing constant (ξ) :	0.1
IBAM, **IMAM**,	general output parameter (δ) :	0.4
IQBAM, **IQMAM**	learning parameter (η) :	1.5

The details of the experimental setting are presented in the following subsections.

3.6.1 Bidirectional Association Model

Through the simulation experiments, the association will be performed from the layer 1 to 2. The initial state of each layer differs based on the experiments of memory capacity and noise tolerance.

3.6.1.1 Memory Capacity

The conditions for memory capacity are shown as Table 3.6.

Table 3.6: Conditions for memory capacity with the bidirectional model.

	Condition	Layer No.	
		[1]	**[2]**
Number of Neurons	**(A-1)**	100	100
	(A-2)	200	200
	(B-1)	150	250
	(B-2)	250	150
Neuron Representation	Bipolar		
Data Set Configuration	Random Pattern		

The conditions are roughly divided into two types; (i) all layers having a same number of neurons, and (ii) each layer having a different number of neurons. Evaluation of the sensitivity of the models to the differences in number of neurons for memory capacity is carried out. Here, the initial setting of each layer is as follows: the layer 1 is set the desired information, and another is set the random bipolar patterns.

73

In general, the number of neurons in layers affects the memory capacity directory. Thus, the recall rate in Figure 3.4c is better than in Figure 3.4a. In Figures 3.4a and 3.4c, the recall rate of conventional models (BAM, GBAM, and QL-BAM) reduces quickly when the number of stored pairs is increased. However, QBAM maintains a high recall rate even if the number of stored pairs is increased. In addition, IQBAM also shows similar properties with QBAM comparing with IBAM.
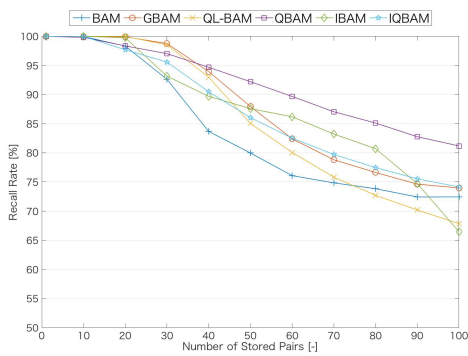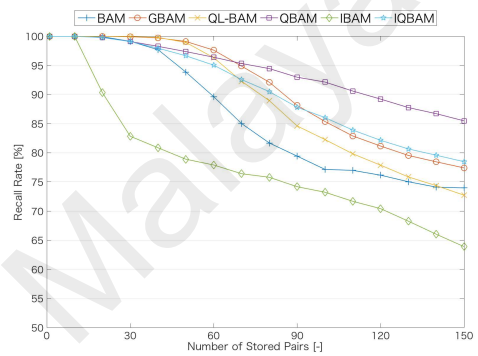
| (a) condition (A-1) | (b) condition (B-1) |

| (c) condition (A-2) | (d) condition (B-2) |
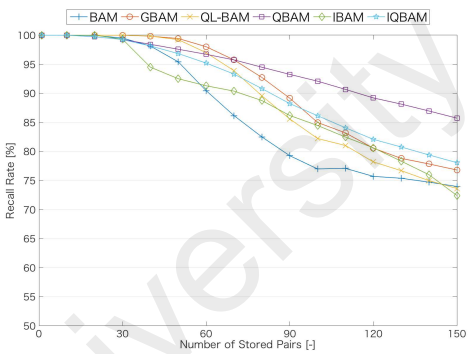
Figure 3.4: Results of memory capacity with conditions A and B.

Based on Figures 3.4b and 3.4d, the sensitivity of models to the differences in the number of neurons for memory capacity can be evaluated. In general, the association from the layer that has a large number of neurons to a layer that has small number of neurons has a higher accuracy than association in opposite direction. Comparing the Figures 3.4b and 3.4d, the recall rate of conventional models is different between conditions (B-1) and (B-2). In contrast, QBAM and IQBAM show a similar recall rate. This implies that

QBAM and IQBAM exhibit stability irrespective of the differences in the number of neurons in bidirectional structure. In terms of IBAM and IQBAM, there is iterative process for the weight learning. Here, the number of iterations is defined to obtain the enough convergence of IQBAM. Under this condition, IBAM shows much less performance than IQBAM.

3.6.1.2   Noise Tolerance

The conditions for noise tolerance are presented in Table 3.7.

Table 3.7: Conditions for noise tolerance with the bidirectional model.

	Condition	Layer No.	
		[1]	[2]
Number of Neurons	**(C-1)**	100	100
	(C-2)	200	200
	(D-1)	150	250
	(D-2)	250	150
Neuron Representation	Bipolar		
Data Set Configuration	Random Pattern		
Number of Pairs	30		

Here, the settings of the number of neurons in layers are same as memory capacity. However, the number of stored pairs is fixed as 30. Thus, the recall rates in Figure 3.5 with $0[\%]$ noise rate are equal to the recall rate of Figure 3.4 in the point of the 30 stored pairs. In Figure 3.5, QBAM and IQBAM show the better recall rate with any conditions. In Figures 3.5b and 3.5d, QBAM and IQBAM show stability against the effect from the difference in the number of neurons in bidirectional structure. In addition, similar with the results of memory capacity, comparing IBAM and IQBAM, IMAM shows much less performance than IQMAM due to the number of learning iterations condition.

From the results of memory capacity and noise tolerance, it can be considered that QBAM and IQBAM have the superior abilities than conventional models.
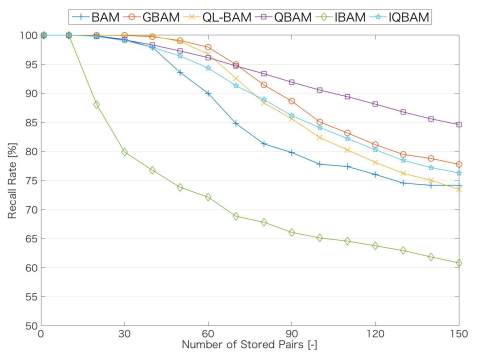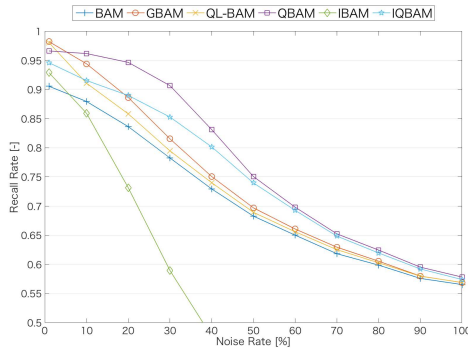
(a) condition (C-1)
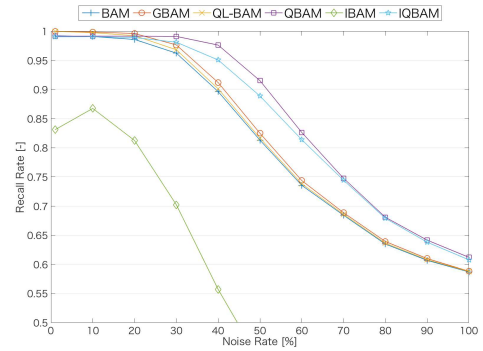
(b) condition (D-1)
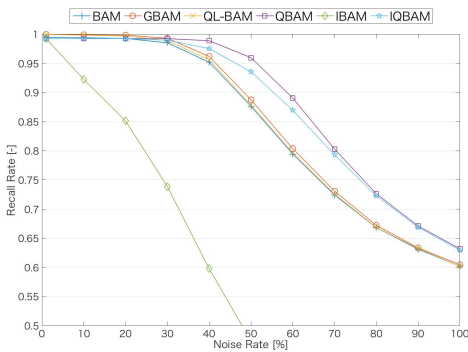
(c) condition (C-2)

(d) condition (D-2)

Figure 3.5: Results of noise tolerance with conditions C and D.

3.6.2 Multi-directional Association Model

Here, through the simulation experiments, the association from the layer 1 to 3 (3-layers model) or 5 (5-layers model) will be performed cyclically until the equilibrium state is attained. The initial state of each layer differs based on experiments of memory capacity and noise tolerance. As mentioned earlier, the multi-directional model consists of multiple bidirectional models. Therefore, it is assumed that the results and properties of the multi-directional model are similar to those of the bidirectional model.

3.6.2.1 Memory Capacity

In terms of memory capacity, two types of conditions are defined. The first is when layers have the same number of neurons, and the second is when the layers have the different number of neurons. As in the case with the bidirectional model, the sensitivity of models to the difference in number of neurons for memory capacity is evaluated. Here,

the initial setting of each layer is as follows; layer 1 is set the desired information, and others layers are set the random bipolar patterns. Initially, the models with a constant number of neurons in layers are compared. The conditions are summarized in Table 3.8.

Table 3.8: Conditions for memory capacity with constant number of neurons in the multi-directional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	3-layers	(E-1)	100	100	100	–	–
		(E-2)	200	200	200	–	–
	5-layers	(F-1)	100	100	100	100	100
		(F-2)	200	200	200	200	200
Neuron Representation	Bipolar						
Data Set Configuration	Random Pattern						

In Figure 3.6, as in the case with the bidirectional model, since the number of neurons in each layer is large, the conditions (E-2) and (F-2) show better recall rate than the conditions (E-1) and (F-1), respectively. Here, similar to the case of the bidirectional model, as the number of layers increase, the memory capacity will correspondingly decrease. In the multi-directional model, the initial conditions, which are random bipolar state, in layers are regarded as noise information. Comparison between the 3-layers model and the 5-layers model shows that the 3-layers model has better recall rate in the multi-directional model. Thus, comparing Figures 3.6a and 3.6b, and Figures 3.6c and 3.6d, the recall rate in Figures 3.6a and 3.6c show better results, respectively. In Figure 3.6, the recall rate of conventional models will decrease quickly with the increase in the number of stored pairs. In contrast, the recall rate of QMAM gradually decreases with the increase in the number of stored pairs. This feature can be regarded as an advantage of QMAM. Focusing the IMAM and IQMAM, IMAM shows the superior recall rate in conditions (E-1) and (E-2). However, when the number of neurons is increased, the recall rate is decreased as conditions (F-1) and (F-2). This is because the number of iterations of weight learning is insufficient for IMAM.
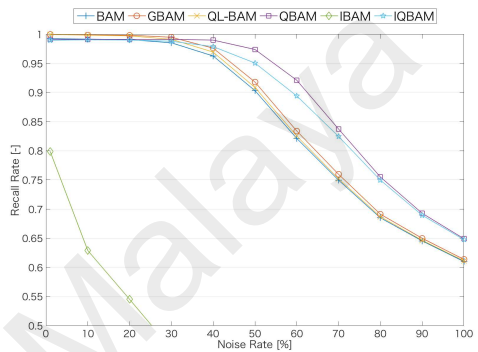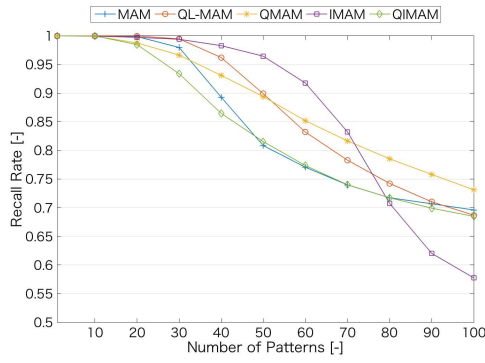
(a) condition (E-1)
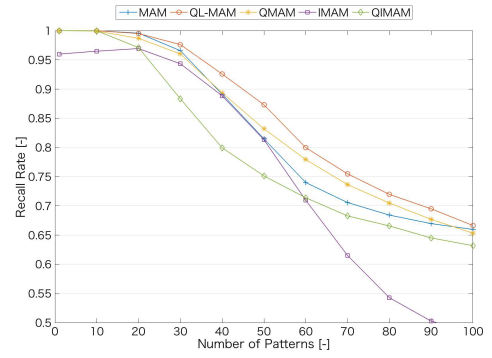
(b) condition (F-1)
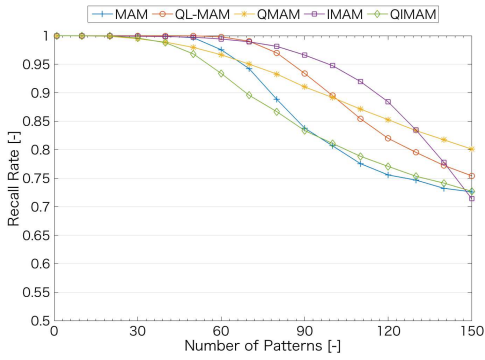
(c) condition (E-2)

(d) condition (F-2)

Figure 3.6: Results of memory capacity with conditions E and F.

Next, the models with the different number of neurons in layers are compared. The conditions are summarized in Table 3.9.

Table 3.9: Conditions for memory capacity with different number of neurons in the multidirectional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	**3-layers**	**(G-1)**	150	200	250	–	–
		(G-2)	250	200	150	–	–
	5-layers	**(H-1)**	160	180	200	220	240
		(H-2)	240	220	200	180	160
Neuron Representation	Bipolar						
Data Set Configuration	Random Pattern						

The recall process from the layer that has a small number of neurons to the layer that has a large number of neurons is harder than the recall process in the opposite direction, due to the differences of the number of neurons in each layer. In other words, the conditions (G-2) and (H-2) show superior recall rate than the conditions (G-1) and

(H-1), respectively. In Figure 3.7, the conventional models show the much difference between conditions G and H in terms of recall rate. However, QMAM and IQMAM show the similar recall rate between conditions (G-1) and (G-2), (H-1) and (H-2), respectively. It can be considered that QMAM and IQMAM have the stability against the different number of neurons in layers. In a real environment, information association is performed between multi-modal information. Typically, each modal is composed of different dimensions. Thus, the resistance to dimensional difference is a significant factor for associative memory. This can also be regarded as an advantage of QMAM and IQMAM. Comparing with IMAM and IQMAM in Figure 3.7, IMAM shows the large difference of recall rate between conditions.
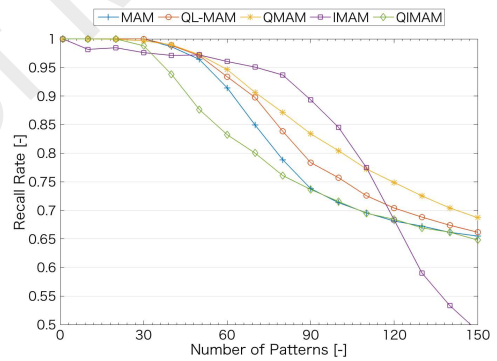
(a) condition (G-1)   (b) condition (H-1)
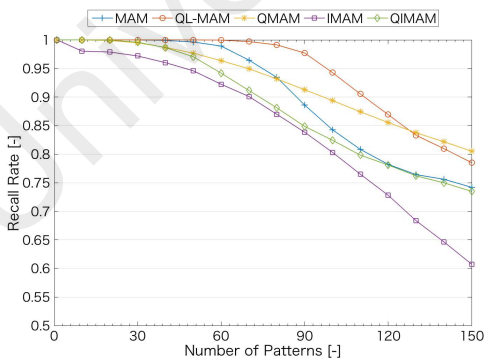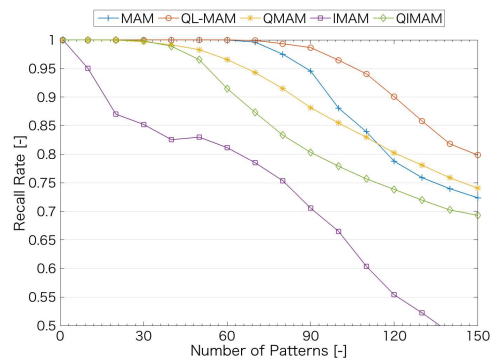
(c) condition (G-2)   (d) condition (H-2)

Figure 3.7: Results of memory capacity with conditions G and H.

Table 3.10 shows the experimental conditions for noise tolerance.

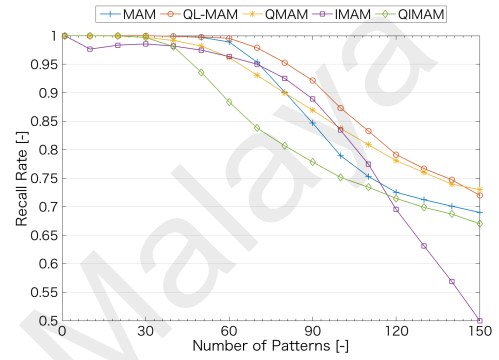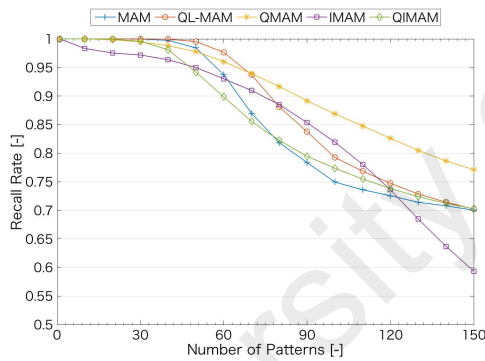Table 3.10: Conditions for noise tolerance in the multi-directional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	**3-layers**	**(I-1)**	200	200	200	–	–
		(I-2)	150	200	250	–	–
		(I-3)	250	200	150	–	–
	5-layers	**(J-1)**	200	200	200	200	200
		(J-2)	160	180	200	220	240
		(J-3)	240	220	200	180	160
Neuron Representation	Bipolar						
Data Set Configuration	Random Pattern						
Number of Pairs	30						

Here, the settings of the number of neurons in each layer are the same as the conditions in memory capacity tests. However, in this case, the number of stored pairs is fixed as 30. As described earlier, the number of layers is regarded as the source of noise. Thus, comparing Figures 3.8 and 3.9, Figure 3.8 shows better recall rate between the corresponding conditions. For any conditions in Figures 3.8 and 3.9, QMAM and IQMAM show the superior recall rate, even if the higher noise rate. From the results of the conditions (I-2) and (I-3) in Figure 3.8, and the conditions (J-2) and (J-3) in Figure 3.9, respectively, noise tolerance is affected by the influence of difference in the number of neurons in each layer. However, same case of memory capacity, QMAM and IQMAM show the stable recall rate even if the number of neurons in each layer is different. Comparing with IMAM and IQMAM in Figures 3.8 and 3.9, IQMAM shows the outstanding performance than IMAM. Although IMAM shows the superior memory capacity, IMAM has the disadvantage in noise tolerance.
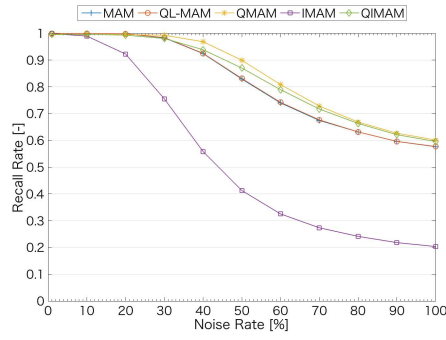
(a) condition (I-1)
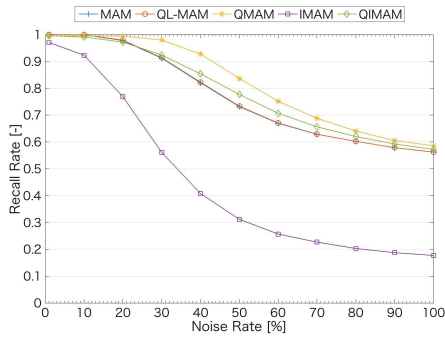
(b) condition (I-2)
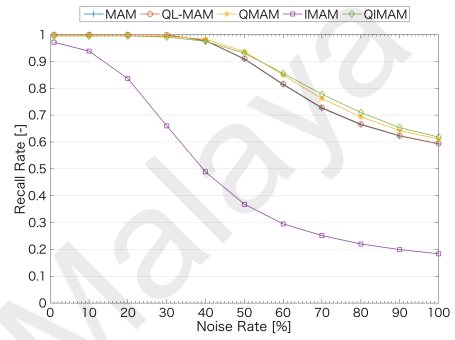
(c) condition (I-3)

Figure 3.8: Results of noise tolerance with condition I.
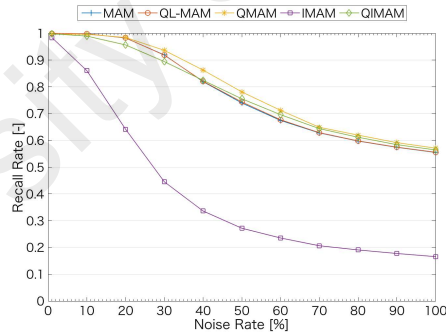
(a) condition (J-1)
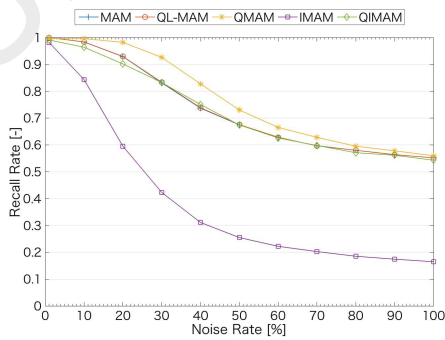
(b) condition (J-2)
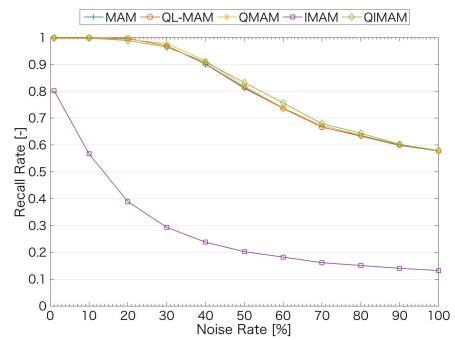
(c) condition (J-3)

Figure 3.9: Results of noise tolerance with condition J.

3.7 Summary

In this chapter, firstly, the mathematical model of quantum-inspired hetero-association models as QBAM and QMAM are presented based on QHAM that was proposed Rigatos and Tzafestas (2006). In addition, based on the incremental learning algorithm that was introduced by Chartier and Boukadoum (2006a, 2011), QBAM and QMAM with an incremental learning algorithm called IQBAM and IQMAM, respectively, are also presented.

In the quantum-inspired association models, superposition and unitarity are regarded as features of QM. In this chapter, the mathematical proofs for superposition and unitarity, which are maintained in hetero-association models, are presented. Furthermore, a numerical example is provided to confirm the satisfaction of superposition and unitarity in quantum-inspired hetero-association models.

From the results of simulation experiment, quantum-inspired hetero-association models show stable and improved abilities as compared to those of the conventional models. Noteworthy, quantum-inspired models show superior recall reliability with keeping stability, even though there is an influence of differing number of neurons in layers.

CHAPTER 4

QUANTUM-INSPIRED COMPLEX-VALUED ASSOCIATIVE MEMORIES

4.1 Introduction

The concept of complex-valued neuron model introduced by Aizenberg (Aizenberg et al., 1973) brought about significant and efficient abilities to NNs. A significant advantage of a complex-valued NN is that it is able to handle the higher dimension information better than the real-valued models. In the previous chapter, the fundamentals of quantum-inspired hetero-association models such as QBAM and QMAM, and the mathematical proofs of the features of QM were presented. Furthermore, the abilities of models pertaining memory capacity and noise tolerance were confirmed by simulation experiments. However, these models can only process real-valued (binary/bipolar) neurons.

In this chapter, the concept of complex-valued neuron model applied to QBAM and QMAM as an extension are called QCBAM and QCMAM, respectively. Similar to the scheme of the previous chapter, the mathematical proofs for features of QM for QCBAM and QCMAM are presented. Furthermore, a numerical example is also provided. Finally, the simulation experiments are conducted to evaluate the abilities of models in terms of memory capacity and noise tolerance.

4.2 Fundamental Structures of Complex-Valued Hetero-Association Model

This section presents the mathematical model of QCBAM and QCMAM. Essentially, the structures and dynamics of the complex-valued model are similar to the real-valued model.

4.2.1 Quantum-Inspired Complex-Valued Bidirectional Associative Memory

Here, the layers of QCBAM are defined as X-layer and Y-layer. Let k of the complex-valued memory pairs $\left\{ \left(\boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(1)} \right), \left(\boldsymbol{X}^{(2)}, \boldsymbol{Y}^{(2)} \right), \ldots, \left(\boldsymbol{X}^{(k)}, \boldsymbol{Y}^{(k)} \right) \right\}$, where $\boldsymbol{X} = [p_{x_1} + jq_{x_1}, p_{x_2} + jq_{x_2}, \ldots, p_{x_N} + jq_{x_N}]$ $(p_x, q_x \in \boldsymbol{R})$ and $\boldsymbol{Y} = [p_{y_1} + jq_{y_1}, p_{y_2} + jq_{y_2}, \ldots, p_{y_M} + jq_{y_M}]$ $(p_y, q_y \in \boldsymbol{R})$, are stored with N neurons in X-layer and M neurons in Y-layer. Based on above conditions, QCBAM is formalized as follows;

- X-layer to Y-layer

$$
\begin{cases}
\boldsymbol{S}^{(k)} = \sum_{j=1}^{N} \sum_{i=1}^{M} W_{ij}^{*} x_i^{(k)} & \text{(4.1a)} \\
\boldsymbol{Y}^{(k)} = \phi \left(\boldsymbol{S}^{(k)} \right) & \text{(4.1b)}
\end{cases}
$$

- Y-layer to X-layer

$$
\begin{cases}
\boldsymbol{U}^{(k)} = \sum_{i=1}^{M} \sum_{j=1}^{N} W_{ij} y_j^{(k)} & \text{(4.2a)} \\
\boldsymbol{X}^{(k)} = \phi \left(\boldsymbol{U}^{(k)} \right) & \text{(4.2b)}
\end{cases}
$$

where, $\boldsymbol{S}$ and $\boldsymbol{U}$ denote the temporal states of associated patterns in X-layer and Y-layer, respectively. The exponential asterisk denotes the conjugate transpose operation. The $\phi(\cdot)$ denotes an activation function based on a complex unit circle that is depicted as Figure 2.9. The activation function for the complex-valued model is formalized as follows;

$$
\phi(Z) = \begin{cases}
\exp(j2\pi n/r), & \text{If } \left| \text{Arg} \left\{ \frac{Z}{\exp(j2\pi n/r)} \right\} \right| < \pi/r \text{ and } Z \neq 0 \\
\text{previous state}, & \text{If } Z = 0
\end{cases} \tag{4.3}
$$

where, $\text{Arg}(\cdot)$ denotes the phase angle which is taken to range over $(-\pi, \pi)$. r denotes quantization value on the complex unit circle, n takes an integer. Assume that $z_0, z_1, \ldots, z_{r-1}$, are r quantized values. Here, the complex number Z will be defined as z_1 that is closest to Z.

The weight connections $\boldsymbol{W}^{*}$ and $\boldsymbol{W}$ are defined by Hebb-like learning, which is based

in Section 3.4.2 of Chapter 3, as follows;

- X-layer to Y-layer

$$W^* = \frac{1}{k} \sum_{p=1}^{k} \sum_{j=1}^{N} \sum_{i=1}^{M} v_j^{(p)} u_i^{(p)*}. \tag{4.4}$$

- Y-layer to X-layer

$$W = \frac{1}{k} \sum_{p=1}^{k} \sum_{i=1}^{M} \sum_{j=1}^{N} u_i^{(p)} v_j^{(p)*}. \tag{4.5}$$

where, an exponential asterisk denotes the conjugate transpose operation. k denotes the number of memory pairs. u and v are orthonormalized complex-valued vectors of x and y, respectively, that are calculated by complex-valued Gram-Schmidt orthogonalization as follows: $a_1 = A_1/\|A_1\| (p = 1)$, $b_p = A_p - \sum_{i=p-1}^{k-1} (a_i, A_i) a_i$ and $a_p = b_p/\|b_p\|$ $(2 \le p \le k)$, where A denotes complex-valued memory vector, a and b denote the orthonormalized and the orthogonalized complex-valued vectors, respectively.

Algorithm 3 shows the pseudo code of above weight learning process for a weight matrix W in a complex-valued model model.

Algorithm 3 An algorithm for a weight matrix W in a complex-valued model model

---

Require: weight matrix W, fundamental complex-valued memory vectors $X^{(k)}$ and $Y^{(k)}$, number of pairs k
Ensure: weight matrix W
  Initialize W as a zero matrix
  Calculate complex-valued orthogonalized vectors u and v by Gram-Schmidt orthogonalization from X and Y
  Set $d = 1$
  while $k \ge d$ **do**
    Calculate $W_{ij} \leftarrow W_{ij} + \frac{1}{k}(u_i v_j^*)$
    Set $d \leftarrow d + 1$
  end while

4.2.2 Quantum-Inspired Complex-Valued Multi-directional Associative Memory

As in the case of the real-valued model, the complex-valued multi-directional model with $L (L > 2)$ of the layers is comprised of multiple bidirectional models. Here, the associated layer is defined as α-th layer, and the other layers are referred to as β-th layers ($\beta = 1, 2, \ldots, L$; $\alpha \ne \beta$). The neurons in each layer continue to be subject to cyclic updates

until the layer reaches an equilibrium. Let k of the original complex-valued memory pairs $\left\{ \boldsymbol{X}^{(k)}_{(1)}, \boldsymbol{X}^{(k)}_{(2)}, \ldots, \boldsymbol{X}^{(k)}_{(L)} \right\}$ are stored in QCMAM, where $\boldsymbol{X} = [p_1 + jq_1, p_2 + jq_2, \ldots, p_{N_{(l)}} + jq_{(l)}]$ $(p, q \in \boldsymbol{R}, l = 1, 2, \ldots, L)$, L represents the number of layers, $N_{(l)}$ denotes the number of neurons in l-th layer, and subscript j denotes the imaginary unit. Based on above conditions, QCMAM is formalized as follows;

- α-th layer to β-th layers

$$
\begin{cases}
\boldsymbol{S}^{(k)}_{(\beta)} = \displaystyle\sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^{L} \sum_{j=1}^{N_{(\beta)}} \sum_{i=1}^{M_{(\alpha)}} W^{*}_{ij(\alpha\beta)} x^{(k)}_{i(\alpha)}, & (\beta = 1, 2, \ldots, L; \ \beta \neq \alpha) \quad (4.6a) \\[4mm]
\boldsymbol{X}^{(k)}_{(\beta)} = \phi\left(\boldsymbol{S}^{(k)}_{(\beta)} \right) & (4.6b)
\end{cases}
$$

- β-th layers to α-th layer

$$
\begin{cases}
\boldsymbol{U}^{(k)}_{(\alpha)} = \displaystyle\sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{L} \sum_{i=1}^{M_{(\alpha)}} \sum_{j=1}^{N_{(\beta)}} W_{ij(\alpha\beta)} x^{(k)}_{j(\beta)}, & (\alpha = 1, 2, \ldots, L; \ \alpha \neq \beta) \quad (4.7a) \\[4mm]
\boldsymbol{X}^{(k)}_{(\alpha)} = \phi\left(\boldsymbol{U}^{(k)}_{(\alpha)} \right) & (4.7b)
\end{cases}
$$

where, $\boldsymbol{S}$ and $\boldsymbol{U}$ denote the temporal states of associated patterns in α-th layer and β-th layer, respectively. The exponential asterisk denotes the conjugate transpose operation. The $\phi(\cdot)$ denotes an activation function based on a complex unit circle that is depicted in Figure 2.9. The activation function for the complex-valued model is formalized as follows;

$$
\phi(Z) = \begin{cases}
\exp(j2\pi n/r), & \text{If } \left| \text{Arg}\left\{ \frac{Z}{\exp(j2\pi n/r)} \right\} \right| < \pi/r \text{ and } Z \neq 0 \\[3mm]
\text{previous state}, & \text{If } Z = 0
\end{cases} \tag{4.8}
$$

where, $\text{Arg}(\cdot)$ denotes the phase angle which is taken to range over $(-\pi, \pi)$. r denotes quantization value on the complex unit circle, n takes an integer. Assume that $z_0, z_1, \ldots, z_{r-1}$, are r quantized values. Here, the complex number Z will be defined as z_1 that is closest to Z.

The weight connections $\boldsymbol{W}^*$ and $\boldsymbol{W}$ are defined by Hebb-like learning as follows;

• α-th layer to β-th layers

$$\boldsymbol{W}^T_{(\alpha\beta)} = \frac{1}{k} \sum_{p=1}^{k} \sum_{j=1}^{N_{(\beta)}} \sum_{i=1}^{M_{(\alpha)}} s_{j(\beta)}^{(p)} s_{i(\alpha)}^{(p)*}. \tag{4.9}$$

• β-th layers to α-th layer

$$\boldsymbol{W}_{(\alpha\beta)} = \frac{1}{k} \sum_{p=1}^{k} \sum_{i=1}^{M_{(\alpha)}} \sum_{j=1}^{N_{(\beta)}} s_{i(\alpha)}^{(p)} s_{j(\beta)}^{(p)*}. \tag{4.10}$$

where, the exponential asterisk denotes a transpose operation. k denotes the number of complex-valued memory pairs. $s_{(\alpha)}$ and $s_{(\beta)}$ represent orthonormalized complex-valued memory vectors $\boldsymbol{X}$ in α-th layer and β-th layer, respectively, that are calculated by complex-valued Gram-Schmidt orthogonalization as follows: $\boldsymbol{a}_1 = \boldsymbol{A}_1/\|\boldsymbol{A}_1\| (p = 1)$, $\boldsymbol{b}_p = \boldsymbol{A}_p - \sum_{i=p-1}^{k-1} (\boldsymbol{a}_i, \boldsymbol{A}_i) \boldsymbol{a}_i$ and $\boldsymbol{a}_p = \boldsymbol{b}_p/\|\boldsymbol{b}_p\|$ $(2 \leq p \leq k)$, where $\boldsymbol{A}$ denotes complex-valued memory vector, $\boldsymbol{a}$ and $\boldsymbol{b}$ denote the orthonormalized complex-valued vector and the orthogonalized complex-valued vector, respectively.

Similar with the real-valued multidirectional model, the above learning algorithm is summarized as Algorithm 3 in Section 4.2.1.

4.3 Features of Quantum Mechanics in Complex-Valued Associative Memory

In general, the real-valued model is considered as a special case of the complex-valued model, e.g., the phase equals to zero. Therefore, as in the case of the real-valued model, the features of QM are realized in the weight matrices of the quantum-inspired complex-valued model as the basis of similarity between QM and fuzzy inference, which is described in Section 3.4.3 of Chapter 3.

Here, the decomposition strategy for complex-valued weight matrix is considered. According to the mathematical definition, the weight matrix W of complex-valued model

is represented as follows;

$$
W = \begin{bmatrix}
p_{11} + jq_{11} & p_{12} + jq_{12} & \cdots & p_{1n} + jq_{1n} \\
p_{21} + jq_{21} & p_{22} + jq_{22} & \cdots & p_{2n} + jq_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
p_{m1} + jq_{m1} & p_{m2} + jq_{m2} & \cdots & p_{mn} + jq_{mn}
\end{bmatrix}.
\tag{4.11}
$$

where, subscript j represents the imaginary part. As in the case of the real-valued model, the elements of complex-valued weight matrix are regarded as the input for fuzzy inference with triangular membership functions that are applied to define the membership grades. Basically, the both elements of weight matrix p and q can be regarded as key factors for the decomposition process as shown in Figure 4.1.
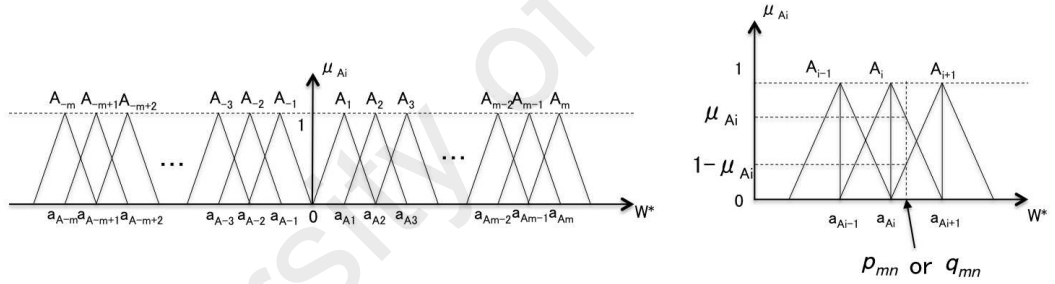
Figure 4.1: Fuzzy subsets for complex-valued weight matrix.

From the membership grade, the elements of complex-valued weight matrix are decomposed as following;

$$
p_{mn} + jq_{mn} = \mu \left(p_{mn} + jq_{mn} \right) + \left(1 - \mu \right) \left(p_{mn} + jq_{mn} \right).
\tag{4.12}
$$

here, the membership grade μ is determined by the element p or q. Thus, in the complex-valued model, the decomposition process can be accomplished by both the real and imaginary part.

88

4.3.1 Existence of Superposition in Complex-Valued Weight Matrix

In the following parts, the variable w_{ij} of weight matrix denotes the real or imaginary parts of Equation (4.11). As in the case of the real-valued model, an infinite number of triangular fuzzy subsets A_i, A_{-i} $(i = 1, 2, \ldots, \infty)$ and B_j, B_{-j} $(j = 1, 2, \ldots, \infty)$ are defined. Figure 4.1 shows the fuzzy subsets in α-th layer as an example. Here, the decomposition process of the complex-valued weight matrices $\boldsymbol{W}^*$ (α-th layer) and $\boldsymbol{W}$ (β-th layer) are presented.

- α-th layer to β-th layer

Consider the element w_{ji} of weight matrix W^*. Due to the strong fuzzy partition, this weight element belongs to two adjacent fuzzy subsets A_i and A_{i+1} (Figure 4.1). The corresponding centers of the fuzzy subsets are a_{ji}^i and a_{ji}^{i+1}, and the associated membership grades are $\mu_{ji} = \mu_{A_i}$ and $1 - \mu_{ji} = \mu_{A_{i+1}}$. Therefore w_{ji} can be represented by the set of $\left\{\mu_{ji}, a_{ji}^{A_i}\right\}$ and $\left\{1 - \mu_{ji}, a_{ji}^{A_{i+1}}\right\}$. The matrices that consist of the membership grades μ_{ji} and $1 - \mu_{ji}$ are generated from the possible combinations of membership grades of each element of the weight matrix. Therefore, the weight matrix W^* can be decomposed into a set of superposition matrices $\bar{W}_i^*$ $\left(i = 1, 2, \ldots, 2^{NM}\right)$ as shown in Table 4.1. Here, the exponential asterisk denotes a conjugate transpose operator.

Table 4.1: Possible combinations of membership grades and corresponding centers of fuzzy subset in W^*.

	(Membership grade, Center of the fuzzy subset)			
$\bar{W}_1^*:$	$(\mu_{11}, a_{11}^{A_i})$	$\cdots$	$(\mu_{n(m-1)}, a_{n(m-1)}^{A_i})$	$(\mu_{nm}, a_{nm}^{A_i})$
$\bar{W}_2^*:$	$(\mu_{11}, a_{11}^{A_i})$	$\cdots$	$(\mu_{n(m-1)}, a_{n(m-1)}^{A_i})$	$(1 - \mu_{nm}, a_{nm}^{A_{i+1}})$
$\bar{W}_3^*:$	$(\mu_{11}, a_{11}^{A_i})$	$\cdots$	$(1 - \mu_{n(m-1)}, a_{n(m-1)}^{A_{i+1}})$	$(\mu_{nm}, a_{nm}^{A_i})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{W}_{2^{NM}-1}^*:$	$(1 - \mu_{11}, a_{11}^{A_{i+1}})$	$\cdots$	$(1 - \mu_{n(m-1)}, a_{n(m-1)}^{A_{i+1}})$	$(\mu_{nm}, a_{nm}^{A_i})$
$\bar{W}_{2^{NM}}^*:$	$(1 - \mu_{11}, a_{11}^{A_{i+1}})$	$\cdots$	$(1 - \mu_{n(m-1)}, a_{n(m-1)}^{A_{i+1}})$	$(1 - \mu_{nm}, a_{nm}^{A_{i+1}})$

As a result, the weight matrix W^* can be decomposed as following;

$$W^* = \sum_{i=1}^{2^{NM}} \mu_i \bar{W}_i^*. \tag{4.13}$$

In Figure 4.1, $A_1, A_2, \ldots, A_{m-1}, A_m$ and $A_{-1}, A_{-2}, \ldots, A_{-m+1}, A_{-m}$ are fuzzy subsets in the universe of discourse of the variable w_{ji}. The sets A_i and A_{-i} are selected such that they have the same spread and satisfy the strong fuzzy partition equalities $\sum_{i=1}^{m} \mu_{A_m(x)} + \sum_{i=1}^{m} \mu_{A_{-m}(x)} = 1$.

• β-th layer to α-th layer

As in the case with α-th layer to β-th layer, consider the element w_{ij} of weight matrix W. The elements of w_{ij} can be represented by the set of $\left\{ v_{ij}, b_{ij}^{B_j} \right\}$ and $\left\{ 1 - v_{ij}, v_{ij}^{B_{j+1}} \right\}$. Thus, the weight matrix W can be decomposed into a set of superposition matrices $\bar{W}_j$ $\left(j = 1, 2, \ldots, 2^{MN} \right)$ as shown in Table 4.2.

Table 4.2: Possible combinations of membership grades and corresponding centers of fuzzy subset in W.

	(Membership grade, Center of the fuzzy subset)			
$\bar{W}_1:$	$(v_{11}, b_{11}^{B_j})$	$\cdots$	$(v_{m(n-1)}, b_{m(n-1)}^{B_j})$	$(v_{mn}, b_{mn}^{B_j})$
$\bar{W}_2:$	$(v_{11}, b_{11}^{B_j})$	$\cdots$	$(v_{m(n-1)}, b_{m(n-1)}^{B_j})$	$(1 - v_{mn}, b_{mn}^{B_{j+1}})$
$\bar{W}_3:$	$(v_{11}, b_{11}^{B_j})$	$\cdots$	$(1 - v_{m(n-1)}, b_{m(n-1)}^{B_{j+1}})$	$(v_{mn}, b_{mn}^{B_j})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{W}_{2^{NM}-1}:$	$(1 - v_{11}, b_{11}^{B_{j+1}})$	$\cdots$	$(1 - v_{m(n-1)}, b_{m(n-1)}^{B_{j+1}})$	$(v_{mn}, b_{mn}^{B_j})$
$\bar{W}_{2^{MN}}:$	$(1 - v_{11}, b_{11}^{B_{j+1}})$	$\cdots$	$(1 - v_{m(n-1)}, b_{m(n-1)}^{B_{j+1}})$	$(1 - v_{mn}, b_{mn}^{B_{j+1}})$

Therefore, as in the case of α-th layer to β-th layer, the weight matrix W can be decomposed as following;

$$W = \sum_{j=1}^{2^{NM}} v_j \bar{W}_j. \tag{4.14}$$

For example, a $m \times n$ weight matrix W (β-th layer to α-th layer) of a complex-valued hetero-association model is considered. Here, the elements of weight w_{mn} are regarded as fuzzy variables. The weight matrix decomposition process is applied to the

real part in Equation (4.11). The possible combinations of membership grades v_{mn} and corresponding centers of fuzzy subsets $b_{mn}^{B_i}$ are represented as the decomposed weight matrices $\bar{W}_k$ $(k = 1, 2, \ldots, 2^{mn})$ that are shown in Equation (4.15);

$$
\begin{aligned}
W &= \frac{v_{11} + v_{12} + \cdots + v_{mn}}{MN \cdot 2^{MN-1}}
\begin{bmatrix}
b_{11}^{B_i} + jq_{11} & b_{12}^{B_i} + jq_{12} & \ldots & b_{1n}^{B_i} + jq_{1n} \\
b_{21}^{B_i} + jq_{21} & b_{22}^{B_i} + jq_{22} & \ldots & b_{2n}^{B_i} + jq_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
b_{m1}^{B_i} + jq_{m1} & b_{m2}^{B_i} + jq_{m2} & \ldots & b_{mn}^{B_i} + jq_{mn}
\end{bmatrix} \\
&+ \frac{v_{12} + \cdots + v_{mn} - v_{11} + 1}{MN \cdot 2^{MN-1}}
\begin{bmatrix}
b_{11}^{B_{i+1}} + jq_{11} & b_{12}^{B_i} + jq_{12} & \ldots & b_{1n}^{B_i} + jq_{1n} \\
b_{21}^{B_i} + jq_{21} & b_{22}^{B_i} + jq_{22} & \ldots & b_{2n}^{B_i} + jq_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
b_{m1}^{B_i} + jq_{m1} & b_{m2}^{B_i} + jq_{m2} & \ldots & b_{mn}^{B_i} + jq_{mn}
\end{bmatrix} \\
&+ \frac{v_{11} + v_{13} + \cdots + v_{mn} - v_{12} + 1}{MN \cdot 2^{MN-1}}
\begin{bmatrix}
b_{11}^{B_i} + jq_{11} & b_{12}^{B_{i+1}} + jq_{12} & \ldots & b_{1n}^{B_i} + jq_{1n} \\
b_{21}^{B_i} + jq_{21} & b_{22}^{B_i} + jq_{22} & \ldots & b_{2n}^{B_i} + jq_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
b_{m1}^{B_i} + jq_{m1} & b_{m2}^{B_i} + jq_{m2} & \ldots & b_{mn}^{B_i} + jq_{mn}
\end{bmatrix} \\
&+ \cdots + \frac{-v_{11} - v_{12} \cdots - v_{mn} + MN}{MN \cdot 2^{MN-1}}
\begin{bmatrix}
b_{11}^{B_{i+1}} + jq_{11} & b_{12}^{B_{i+1}} + jq_{12} & \ldots & b_{1n}^{B_{i+1}} + jq_{1n} \\
b_{21}^{B_{i+1}} + jq_{21} & b_{22}^{B_{i+1}} + jq_{22} & \ldots & b_{2n}^{B_{i+1}} + jq_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
b_{m1}^{B_{i+1}} + jq_{m1} & b_{m2}^{B_{i+1}} + jq_{m2} & \ldots & b_{mn}^{B_{i+1}} + jq_{mn}
\end{bmatrix}.
\end{aligned}
\tag{4.15}
$$

where, the elements in matrices, which are denoted by N_i, and the associated $\|L_1\|$ are calculated. Each L_1 norm is divided by the number of elements in matrices N_i, which satisfies Lemma 1 in Section 3.4.4 of Chapter 3 as follows;

$$
\frac{1}{MN \cdot 2^{(MN-1)}} \sum_{i=1}^{M} \sum_{j=1}^{N} |v_{ij}| = 1.
\tag{4.16}
$$

Similarly, this process can be applied to the imaginary part of Equation (4.11) and can be regarded as superposition.

4.3.2 Unitary Operation in Complex-Valued Model

As in the case of the real-valued model, the unitary operators in fuzzy inference, which are described by the fuzzy relational matrices R_+ ($\boldsymbol{W}^* \geq 0$) and R_- ($\boldsymbol{W}^* < 0$) as the increase and decrease operators, respectively, are guaranteed based on Theorem 1 in Section 3.4.5 of Chapter 3. For example;

- α-th layer to β-th layer

$$
\begin{cases}
\left(R_{m-1}^d\right)^{-1} = \left(R_{m-1}^i\right)^* & \text{(4.17a)} \\
\left(R_{-m+1}^d\right)^{-1} = \left(R_{-m+1}^i\right)^* & \text{(4.17b)}
\end{cases}
$$

- β-th layer to α-th layer

$$
\begin{cases}
\left(R_{n-1}^d\right)^{-1} = \left(R_{n-1}^i\right)^* & \text{(4.18a)} \\
\left(R_{-n+1}^d\right)^{-1} = \left(R_{-n+1}^i\right)^* & \text{(4.18b)}
\end{cases}
$$

Furthermore, unitarity in the complex-valued models is satisfied in the eigenvector spaces between decomposed weight matrices $\bar{W}^*$ and $\bar{W}$, and is guaranteed by Theorem 2 in Section 3.4.6 of Chapter 3.

4.4 Numerical Example

Based on Section 4.3, superposition and unitarity are demonstrated by numerical examples. Through the example in this section, the fundamental condition is defined as follows: Here, the following vectors x_i and y_j ($i, j = 1, 2, 3, 4$) are set as the fundamental complex-valued memory vectors for quantum-inspired complex-valued hetero-association model. Even if the model has two fundamental weight matrices W^* in α-th layer to β-th layer and W in β-th layer to α-th layer, the weight matrices W^* and W are linked with a transposition relationship. Therefore, if weight W satisfies superposition and unitarity, weight W^* also satisfies these properties. Thus, the examples are performed using weight W.

First, the fundamental complex-valued memory vectors x_i and y_j $(i, j = 1, 2, 3, 4)$ are described.

$$x_1 = \begin{bmatrix} 1 \\ -1 \\ -j \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ j \end{bmatrix}, \quad x_3 = \begin{bmatrix} j \\ j \\ 1 \\ -1 \end{bmatrix}. \tag{4.19}$$

$$y_1 = \begin{bmatrix} -j \\ 1 \\ -j \\ -j \end{bmatrix}, \quad y_2 = \begin{bmatrix} j \\ j \\ 1 \\ 1 \end{bmatrix}, \quad y_3 = \begin{bmatrix} -j \\ -j \\ -j \\ 1 \end{bmatrix}. \tag{4.20}$$

Generally, the high dimensional vector space that is chosen randomly, demonstrates orthogonality. In this section, however, the dimensions of complex-valued memory vectors are low, therefore, to have orthogonality, the complex-valued Gram-Schmidt orthogonalization is applied to fundamental complex-valued memory vectors x_i and y_i. Here, orthonormalized complex-valued memory vectors will be utilized to calculate the weight matrix W. The complex-valued Gram-Schmidt orthogonalization gives the orthonormalized complex-valued memory vectors v_{x_i} and v_{y_i} from memory vectors x_i and y_i, respectively, as follows;

$$v_{x_1} = \begin{bmatrix} 0.5000 \\ -0.5000 \\ j0.5000 \\ 0.5000 \end{bmatrix}, \quad v_{x_2} = \begin{bmatrix} -0.3536 - j0.3536 \\ 0.3536 + j0.3536 \\ -0.3536 + j0.3536 \\ 0.3536 + j0.3536 \end{bmatrix},$$

$$v_{x_3} = \begin{bmatrix} j0.5773 \\ j0.5773 \\ 0.2887 + j0.2887 \\ -0.2887 + j0.2887 \end{bmatrix}. \tag{4.21}$$

$$v_{y_1} = \begin{bmatrix} -j0.5000 \\ 0.5000 \\ -j0.5000 \\ -j0.5000 \end{bmatrix}, \quad v_{y_2} = \begin{bmatrix} -0.6123 + j0.6123 \\ 0.2041 + j0.2041 \\ 0.2041 - j0.2041 \\ 0.2041 - j0.2041 \end{bmatrix},$$

$$v_{y_3} = \begin{bmatrix} 0.0 \\ -j0.4082 \\ 0.2041 - j0.6124 \\ 0.2041 + j0.6124 \end{bmatrix}. \tag{4.22}$$

4.4.1 Superposition of the weight matrix

The complex-valued weight matrix W is defined by Hebb-like learning with orthonormalized complex-valued vectors v_x and v_y as given by Equation (4.23).

$$W = \begin{bmatrix} j0.228 & -0.043 & -0.118 + j0.075 & 0.118 + j0.075 \\ -j0.228 & -0.114 & -0.118 + j0.004 & 0.118 + j0.004 \\ 0.062 & -0.039 + j0.171 & -0.171 + j0.079 & -0.053 - j0.039 \\ -j0.061 & 0.092 - j0.039 & -0.079 + j0.092 & 0.039 + j0.210 \end{bmatrix}. \quad (4.23)$$

where, the number of possible combinations are $2^{NM} = 2^{16}$, and elements in Equation (3.50) are considered as the universe of discourse of the fuzzy variables that are depicted in Figure 4.2.
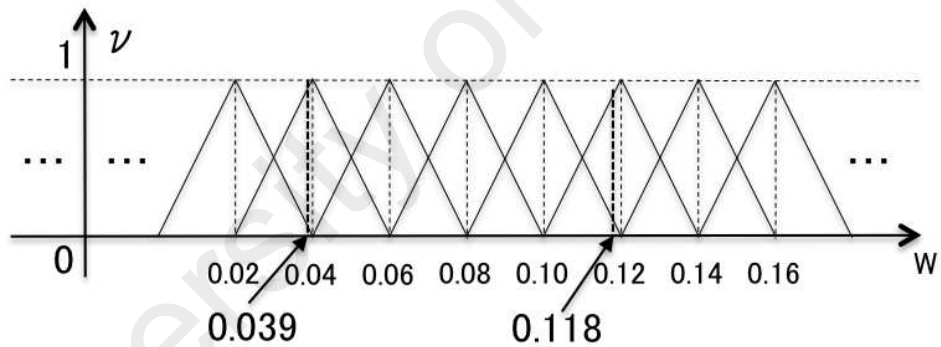
Figure 4.2: Example of derivation of the membership grade from element of weight matrix W.

Based on Equation (4.15), Equation (4.23) can be described as Equation (4.24) as following;

$$
W = \frac{0.85 + 0.10 + \cdots + 0.95}{16 \cdot 2^{15}}
\begin{bmatrix}
j0.228 & -0.04 & -0.10 + j0.075 & 0.12 + j0.075 \\
-j0.228 & -0.10 & -0.10 + j0.004 & 0.12 + j0.004 \\
0.08 & -0.02 + j0.171 & -0.16 + j0.079 & -0.04 - j0.039 \\
-j0.061 & 0.10 - j0.039 & -0.06 + j0.092 & 0.04 + j0.210
\end{bmatrix}
$$

$$
+ \frac{0.10 + 0.10 + \cdots + 0.95 - 0.85 + 1}{16 \cdot 2^{15}}
\begin{bmatrix}
j0.228 & -0.06 & -0.10 + j0.075 & 0.12 + j0.075 \\
-j0.228 & -0.10 & -0.10 + j0.004 & 0.12 + j0.004 \\
0.08 & -0.02 + j0.171 & -0.16 + j0.079 & -0.04 - j0.039 \\
-j0.061 & 0.10 - j0.039 & -0.06 + j0.092 & 0.04 + j0.210
\end{bmatrix}
\tag{4.24}
$$

$$
+ \cdots + \frac{-0.15 - 0.90 \cdots - 0.05 + 16}{16 \cdot 2^{15}}
\begin{bmatrix}
j0.228 & -0.06 & -0.12 + j0.075 & 0.10 + j0.075 \\
-j0.228 & -0.12 & -0.12 + j0.004 & 0.10 + j0.004 \\
0.06 & -0.04 + j0.171 & -0.18 + j0.079 & -0.06 - j0.039 \\
-j0.061 & 0.08 - j0.039 & -0.08 + j0.092 & 0.02 + j0.210
\end{bmatrix}.
$$

4.4.2 Unitarity in Decomposed Weight

Unitarity is maintained in the fuzzy relational matrix that is defined by complex-valued eigenvectors of decomposed weight matrices $\bar{W}_i$. Here, as examples, unitarity between the vector spaces $\bar{W}_1$, $\bar{W}_2$ and $\bar{W}_{2^{16}}$ is shown. Table 4.3 shows the eigenvalues λ and complex-valued eigenvectors τ of decomposed complex-valued weight matrices $\bar{W}_i$.

Table 4.3: Eigenvalues and complex-valued eigenvectors of decomposed weight matrices $\bar{W}_1, \bar{W}_2$ and $\bar{W}_{2^{16}}$.

	Eigenvalue		Eigenvector			
$\bar{W}_1:$	$\lambda_1 =$	$-2.841\mathrm{e}^{-6}$	$\tau_1 = [$	0.259	$-0.010 + j0.707$	$0.604 + j0.010$	$0.026 - j0.270]^T$
	$\lambda_2 =$	$-1.616\mathrm{e}^{-6}$	$\tau_2 = [$	0.672	$0.017 + j0.347$	$-0.571 + j0.081$	$-0.169 + j0.258]^T$
	$\lambda_3 =$	$5.190\mathrm{e}^{-7}$	$\tau_3 = [$	-0.303	$0.058 + j0.127$	$-0.125 + j0.426$	$-0.775 - j0.302]^T$
	$\lambda_4 =$	$2.238\mathrm{e}^{-6}$	$\tau_4 = [$	0.624	$0.014 - j0.604$	$0.303 + j0.116$	$-0.205 - j0.313]^T$
$\bar{W}_2:$	$\lambda_1 =$	$-8.088\mathrm{e}^{-5}$	$\tau_1 = [$	0.270	$-0.009 + j0.721$	$0.581 + j0.010$	$0.020 - j0.261]^T$
	$\lambda_2 =$	$-4.525\mathrm{e}^{-5}$	$\tau_2 = [$	0.658	$0.016 + j0.331$	$-0.592 + j0.082$	$-0.169 + j0.267]^T$
	$\lambda_3 =$	$1.429\mathrm{e}^{-5}$	$\tau_3 = [$	-0.311	$0.056 + j0.127$	$-0.133 + j0.425$	$-0.768 - j0.309]^T$
	$\lambda_4 =$	$6.048\mathrm{e}^{-5}$	$\tau_4 = [$	0.631	$0.014 - j0.591$	$0.304 + j0.119$	$-0.210 - j0.319]^T$
$\bar{W}_{2^{16}}:$	$\lambda_1 =$	$-1.179\mathrm{e}^{-4}$	$\tau_1 = [$	0.280	$-0.067 + j0.691$	$0.603 + j0.152$	$0.099 - j0.206]^T$
	$\lambda_2 =$	$-6.519\mathrm{e}^{-5}$	$\tau_2 = [$	0.625	$0.067 + j0.365$	$-0.567 + j0.020$	$-0.269 + j0.279]^T$
	$\lambda_3 =$	$1.588\mathrm{e}^{-5}$	$\tau_3 = [$	-0.368	$0.099 + j0.131$	$-0.009 + j0.426$	$-0.789 - j0.182]^T$
	$\lambda_4 =$	$8.215\mathrm{e}^{-5}$	$\tau_4 = [$	0.629	$0.021 - j0.594$	$0.290 + j0.161$	$-0.238 - j0.291]^T$

Based on Theorem 2 in Section 3.4.6 of Chapter 3, unitarity is maintained in the transition matrix between the spaces that are spanned by the complex-valued eigenvectors of decomposed weight matrices $\bar{W}_i$ and $\bar{W}_j$.

For example, the transition matrix R, which performs from $\bar{W}_1$ to $\bar{W}_{2^{16}}$, is calculated by a following transition matrix $R_{\bar{W}_1 \bar{W}_{2^{16}}}$ as given by Equation (4.25);

$$R_{\bar{W}_1 \bar{W}_{2^{16}}} = \begin{bmatrix} \tau_1^{(\bar{W}_1)} \tau_1^{(\bar{W}_2)} & \tau_1^{(\bar{W}_1)} \tau_2^{(\bar{W}_2)} & \tau_1^{(\bar{W}_1)} \tau_3^{(\bar{W}_2)} & \tau_1^{(\bar{W}_1)} \tau_4^{(\bar{W}_2)} \\ \tau_2^{(\bar{W}_1)} \tau_1^{(\bar{W}_2)} & \tau_2^{(\bar{W}_1)} \tau_2^{(\bar{W}_2)} & \tau_2^{(\bar{W}_1)} \tau_3^{(\bar{W}_2)} & \tau_2^{(\bar{W}_1)} \tau_4^{(\bar{W}_2)} \\ \tau_3^{(\bar{W}_1)} \tau_1^{(\bar{W}_2)} & \tau_3^{(\bar{W}_1)} \tau_2^{(\bar{W}_2)} & \tau_3^{(\bar{W}_1)} \tau_3^{(\bar{W}_2)} & \tau_3^{(\bar{W}_1)} \tau_4^{(\bar{W}_2)} \\ \tau_4^{(\bar{W}_1)} \tau_1^{(\bar{W}_2)} & \tau_4^{(\bar{W}_1)} \tau_2^{(\bar{W}_2)} & \tau_4^{(\bar{W}_1)} \tau_3^{(\bar{W}_2)} & \tau_4^{(\bar{W}_1)} \tau_4^{(\bar{W}_2)} \end{bmatrix}. \qquad (4.25)$$

Therefore, from the eigenvectors in Table 4.3, the following matrix is calculated;

$$
R_{\bar{W}_1 \bar{W}_{216}} = \begin{bmatrix} 0.997 & 0.011 - j0.014 & -0.041 - j0.024 & -0.064 - j0.007 \\ -0.004 - j0.013 & 0.996 - j0.022 & 0.058 + j0.025 & 0.050 + j0.028 \\ 0.049 - j0.031 & -0.061 + j0.037 & 0.980 - j0.086 & 0.150 - j0.036 \\ 0.057 & -0.040 + j0.022 & -0.156 - j0.034 & 0.982 + j0.068 \end{bmatrix} .
$$

(4.26)

Unitarity is maintained as given by Equation (4.27).

$$
\begin{aligned}
R_{\bar{W}_1 \bar{W}_{216}} R^*_{\bar{W}_1 \bar{W}_{216}} &= \begin{bmatrix} 0.998 + j0.003 & 0.001 - j0.032 & -0.004 - j0.048 & -j0.005 \\ 0.001 - j0.032 & 0.996 - j0.038 & 0.007 + j0.060 & -0.001 + j0.047 \\ -0.004 - j0.048 & 0.007 + j0.060 & 0.978 - j0.186 & -0.003 - j0.050 \\ -j0.005 & -0.001 + j0.047 & -0.003 - j0.050 & 0.987 + j0.143 \end{bmatrix} \\[2mm]
&\approx \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .
\end{aligned}
$$

(4.27)

Similarly, the transition matrix R, which performs from $\bar{W}_2$ to $\bar{W}_{216}$, is presented as given in Equation (4.28).

$$
R_{\bar{W}_2 \bar{W}_{216}} = \begin{bmatrix} 0.998 & 0.009 - j0.011 & -0.025 - j0.024 & -0.055 \\ -0.004 - j0.010 & 0.995 - j0.025 & 0.065 + j0.050 & 0.039 + j0.031 \\ 0.033 - j0.027 & -0.063 + j0.061 & 0.979 - j0.079 & 0.158 - j0.036 \\ 0.050 + j0.007 & -0.027 + j0.022 & -0.164 - j0.037 & 0.981 + j0.072 \end{bmatrix} . \quad (4.28)
$$

Equation (4.29) shows unitarity,

$$R_{\bar{W}_2\bar{W}_{216}}R^*_{\bar{W}_2\bar{W}_{216}} = \begin{bmatrix} 0.999+j0.001 & 0.002-j0.025 & -0.002-j0.05 & -0.001+j0.007 \\ 0.002-j0.026 & 0.992-j0.040 & 0.013+j0.109 & j0.044 \\ -0.002-j0.045 & 0.013+j0.109 & 0.976-j0.175 & -0.003-j0.051 \\ -0.001+0.007 & j0.044 & -0.003-j0.051 & 0.986+j0.153 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(4.29)

In principle, unitarity is satisfied with any combinations of eigenvectors space.

4.5 Simulation Experiments

This section presents the performance evaluation for QCBAM and QCMAM that are compared with the fundamental models as CBAM (Donq & Wen, 1998) and CMAM (Kobayashi & Yamazaki, 2005), respectively, in terms of abilities of memory capacity and noise tolerance. The abilities of complex-valued models are discussed from a real-valued model perspective. Therefore, as in the case of the real-valued model, memory capacity is sensitive to number of neurons in the layer. In addition, if the layers that have different number of neurons are set, memory capacity will be affected. Furthermore, "noise" in the association models is roughly classified into two types. First, the noise contained in the stored patterns, and second, the noise originating from the similarity in the stored patterns. In this section, the former type of noise pertaining to the bidirectional and multi-directional model is discussed. In addition, the latter type of noise is discussed with the multi-directional model. For noise tolerance, the initial setting of each layer is as follows; the number of stored pairs is 40, the layer 1 is set the information that contains a

certain ratio $(0 - 100[\%])$ of salt and pepper noise, and others are set randomly.

It is known that the abilities of complex-valued associative memory are dependent on the number of divisions in the complex-valued unit circle. In general, if the number of divisions is increased, the abilities of model will decrease correspondingly. In other words, the expressiveness of neurons is increased if the number of divisions of complex-valued unit circle is increased. However, at the same time, the recall rate decreases. In this section, the conditions of 4, 8, and 16 divisions in the complex-valued circle (Figure 4.3) are applied.
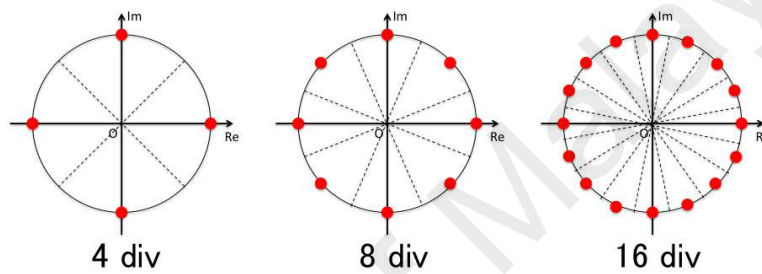
Figure 4.3: Complex unit circle with 4, 8 and 16 divisions.

The details of the experimental setting are presented in the following subsections.

4.5.1 Bidirectional Association Model

Through the simulation experiments, the association from layer 1 to 2 will be performed. The initial state of each layer differs based on the experiments on memory capacity and noise tolerance.

4.5.1.1 Memory Capacity

The conditions for memory capacity are shown in Table 4.4. Here, two types of conditions are defined; one is when the layers have a constant number of neurons, the other is when the layers have different number of neurons. It is feasible to evaluate the sensitivity of models to the difference in number of neurons for memory capacity. The initial inner state of each layer is defined as follows; layer 1 is set the desired information, and other layers are set the random phase information.

Table 4.4: Conditions for memory capacity with the bidirectional model.

	Condition	Layer No.	
		[1]	**[2]**
Number of Neurons	**(K-1)**	100	100
	(K-2)	200	200
	(L-1)	100	200
	(L-2)	200	100
Neuron Representation	Complex Number		
Data Set Configuration	Amplitude: 1.0 Phase: Random		

With regards to AM, the number of neurons in layers directly influences the memory capacity. Therefore, comparing the conditions (K-1) and (K-2) in Figure 4.4, the condition (K-2) shows better recall rate with both CBAM and QCBAM. Comparing the recall rates of CBAM and QCBAM, QCBAM maintains a high recall rate even when the number of pairs is increased. Furthermore, the reduction ratio of recall rate with QCBAM is smaller than CBAM. Considering the difference in the number of divisions, with the increase in the number of divisions, the recall rate of CBAM decreases correspondingly; whereas, QCBAM maintains a high recall rate even with a large number of stored pairs.

(a) CBAM with condition (K-1)

(b) QCBAM with condition (K-1)

(c) CBAM with condition (K-2)

(d) QCBAM with condition (K-2)
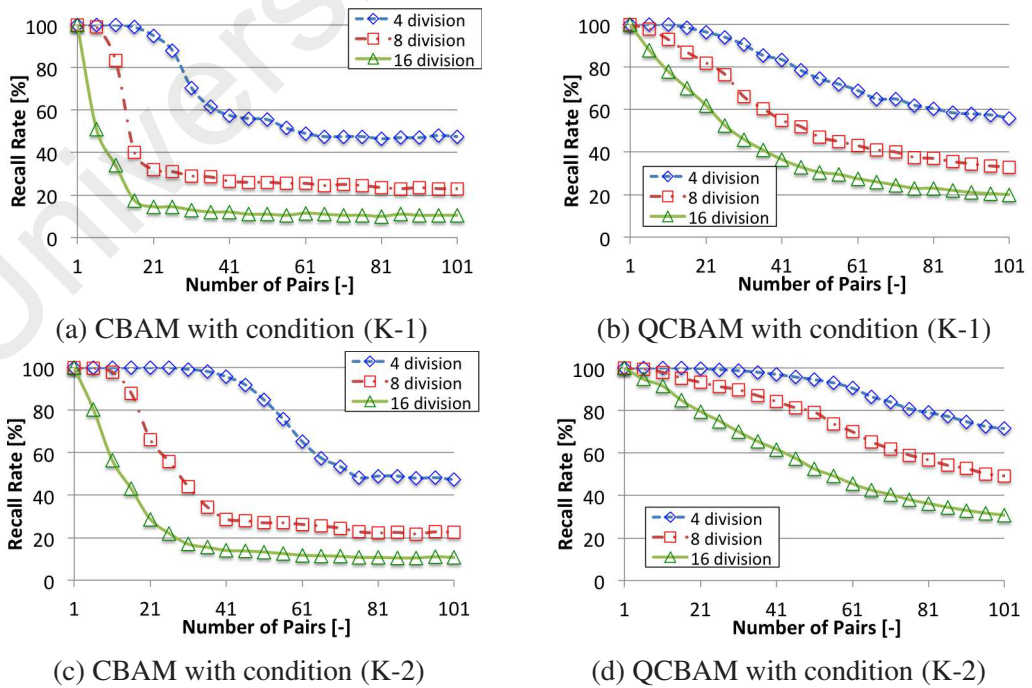
Figure 4.4: Results of memory capacity with random generated data on conditions K.

The difference in the number of neurons in layers also affects the memory capacity. As in the case of the real-valued model, the association from the layer that has a large number of neurons to the layer that has small number of neurons is easier as compared to the association in the opposite direction. Thus, pertaining the recall rate with CBAM in Figure 4.5, the condition (L-2) shows better recall rate than condition (L-1). However, QCBAM shows a similar result with the conditions (L-1) and (L-2).

| (a) CBAM with condition (L-1) | (b) QCBAM with condition (L-1) |
| (c) CBAM with condition (L-2) | (d) QCBAM with condition (L-2) |
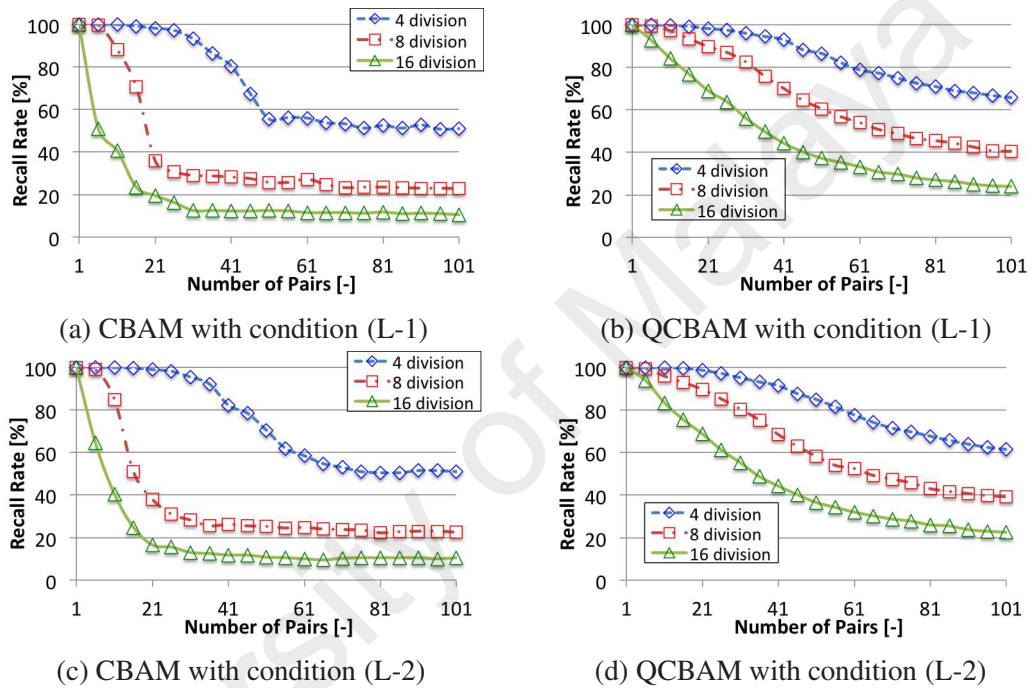
Figure 4.5: Results of memory capacity with random generated data on conditions L.

4.5.1.2 Noise Tolerance

The conditions for noise tolerance are presented in Table 4.5. Here, the settings of the number of neurons in layers are same as the conditions in memory capacity. However, the number of stored pairs is fixed as 40. Thus, the recall rates in Figures 4.6 and 4.7 with $0[\%]$ noise rate are equal to the recall rate of Figures 4.4 and 4.5, respectively, in the point of 40 stored pairs.

Figure 4.6 shows the result of noise tolerance with constant number of neurons. Because the number of neurons in the layers is large, the condition (M-2) shows better

Table 4.5: Conditions for noise tolerance with the bidirectional model.

	Condition	Layer No.	
		[1]	[2]
Number of Neurons	**(M-1)**	100	100
	(M-2)	200	200
	(N-1)	100	200
	(N-2)	200	100
Neuron Representation	Complex Number		
Data Set Configuration	Amplitude: 1.0 Phase: Random		
Number of Pairs	40		

results in CBAM and QCBAM. Comparing CBAM and QCBAM, QCBAM shows superior results under both conditions (M-1) and (M-2). Furthermore, QCBAM with a high number of divisions in the complex-valued unit circle shows outstanding performance as compared to CBAM.
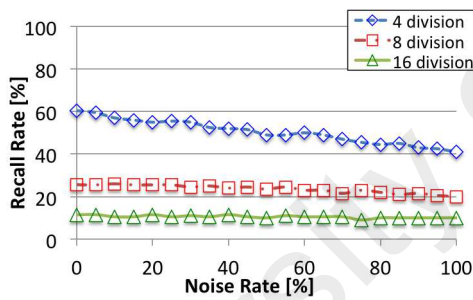
(a) CBAM with condition (M-1)
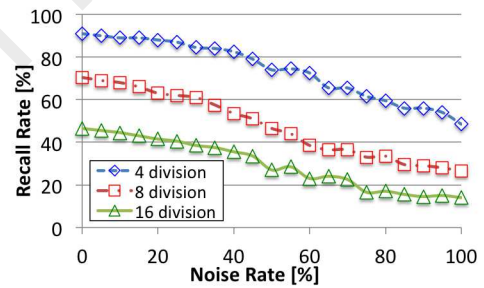
(b) QCBAM with condition (M-1)
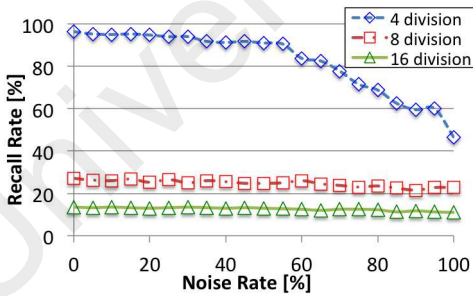
(c) CBAM with condition (M-2)

(d) QCBAM with condition (M-2)

Figure 4.6: Results of noise tolerance with random generated data on conditions M.

Figure 4.7 shows the result of noise tolerance with different number of neurons. Here, the features of results under conditions (N-1) and (N-2) are similar to the features of results in Figure 4.6. The results with conditions (N-1) and (N-2) in CBAM show a similar recall rate. It is considered that the association process of CBAM fails under these conditions; whereas, QCBAM is able to carry out the association process with superior performance.

(a) CBAM with condition (N-1)   (b) QCBAM with condition (N-1)

(c) CBAM with condition (N-2)   (d) QCBAM with condition (N-2)
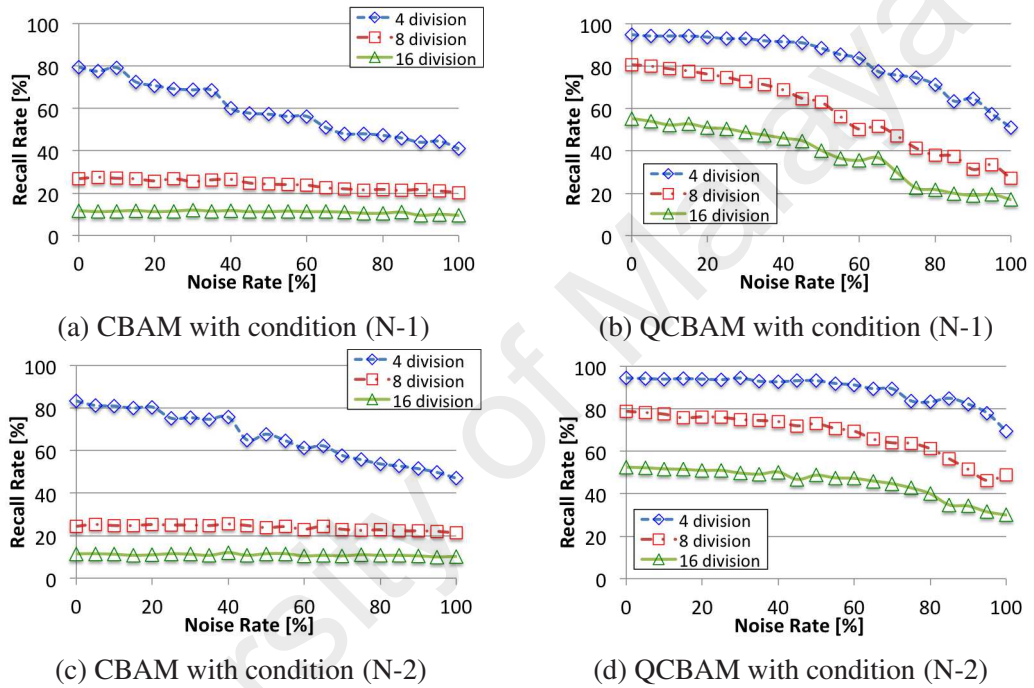
Figure 4.7: Results of noise tolerance with random generated data on conditions N.

Based on the results of memory capacity and noise tolerance, it is evident that QCBAM has superior abilities than CBAM.

4.5.2 Multi-directional Association Model with Random Bipolar Data

Through the simulation experiments, the association from the layer 1 to 3 (3-layers model) or 5 (5-layers model) is performed cyclically until an equilibrium state is attained. The initial state of each layer is different depending on experiments of memory capacity and noise tolerance. In addition, the initial settings of layers are roughly divided into two conditions; constant and different number of neurons in layers. As mentioned earlier, the

multi-directional model comprises multiple bidirectional models. Therefore, it is assumed that the results and properties of the multi-directional model are similar to those of the bidirectional model. In this section, random generated data is utilized for simulation experiments.

4.5.2.1 Memory Capacity

Firstly, the models that have a constant number of neurons are considered. The conditions for memory capacity with a constant number of neurons are in Table 4.6.

Table 4.6: Conditions for memory capacity with constant number of neurons in the multi-directional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	3-layers	(O-1)	100	100	100	–	–
		(O-2)	200	200	200	–	–
	5-layers	(P-1)	100	100	100	100	100
		(P-2)	200	200	200	200	200
Neuron Representation	Complex Number						
Data Set Configuration	Amplitude: 1.0, Phase: Random						

Here, the initial setting of each layer is as follows; layer 1 is set the desired information, and other layers are set the random phase information ($0 \leq 2\pi n/r < 2\pi$[rad]) depending on the number of divisions in the complex-valued unit circle.

As in the case of the bidirectional model, the number of neurons in layers affects the ability of memory capacity directly. Thus, comparing the conditions (O-1) and (O-2) in Figure 4.8, the condition (O-2) shows a better recall rate in both CMAM and QCMAM. Furthermore, in both conditions, QCMAM shows a better recall rate than CMAM, especially when the number of divisions is high. This implies that QCMAM has a superior ability to represent rich information. Comparing CMAM and QCMAM in Figure 4.8, we can infer that the decreasing ratio for the number of stored pairs of QCMAM is smaller than that for CMAM.
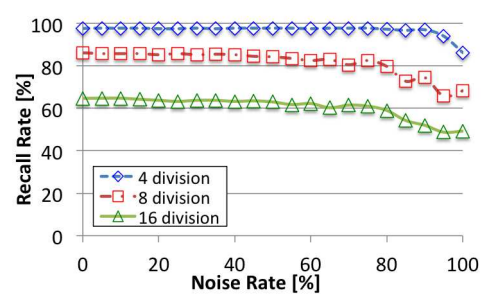
(a) CMAM with condition (O-1)
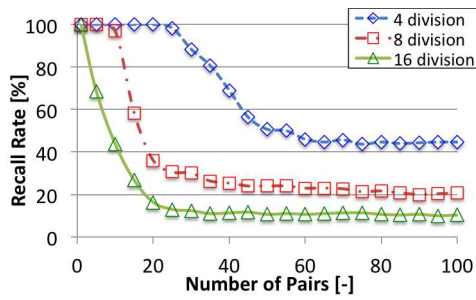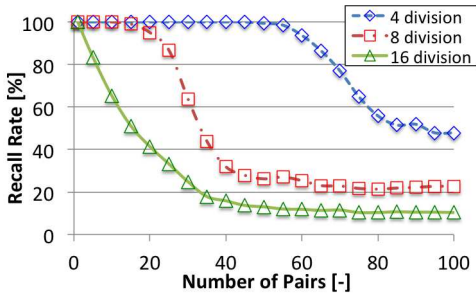
(b) QCMAM with condition (O-1)

(c) CMAM with condition (O-2)
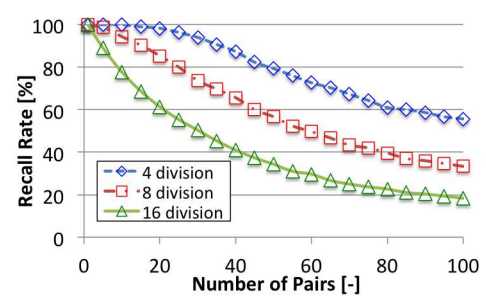
(d) QCMAM with condition (O-2)

Figure 4.8: Results of memory capacity with random generated data on conditions O.

Figure 4.9 shows the experimental results of 5-layer model. As in the case of the 3-layer model, by comparing the conditions (P-1) and (P-2) in Figure 4.9, it is evident that the condition (P-2) shows a better recall rate in both CMAM and QCMAM. This feature also is present in the real-valued model.

(a) CMAM with condition (P-1)

(b) QCMAM with condition (P-1)

(c) CMAM with condition (P-2)
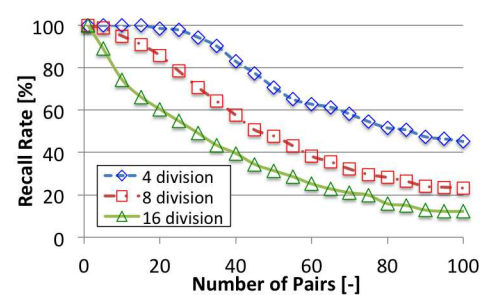
(d) QCMAM with condition (P-2)

Figure 4.9: Results of memory capacity with random generated data on conditions P.

Next, the models having different number of neurons are considered. The conditions of memory capacity with different number of neurons are shown in Table 4.7.

Table 4.7: Conditions for memory capacity with different number of neurons in the multi-directional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	3-layers	(Q-1)	150	200	250	–	–
		(Q-2)	250	200	150	–	–
	5-layers	(R-1)	160	180	200	220	240
		(R-2)	240	220	200	180	160
Neuron Representation	Complex Number						
Data Set Configuration	Amplitude: 1.0, Phase: Random						

Figure 4.10 shows the experimental results of 3-layer model with different number of neurons. Comparing with the conditions (Q-1) and (Q-2), as in the case of the bidirectional model, there are the differences in recall rates of CMAM and QCMAM, respectively. However, QCMAM shows the better recall rate under both conditions (Q-1) and (Q-2), especially when the number of divisions is high.

(a) CMAM with condition (Q-1)   (b) QCMAM with condition (Q-1)

(c) CMAM with condition (Q-2)   (d) QCMAM with condition (Q-2)
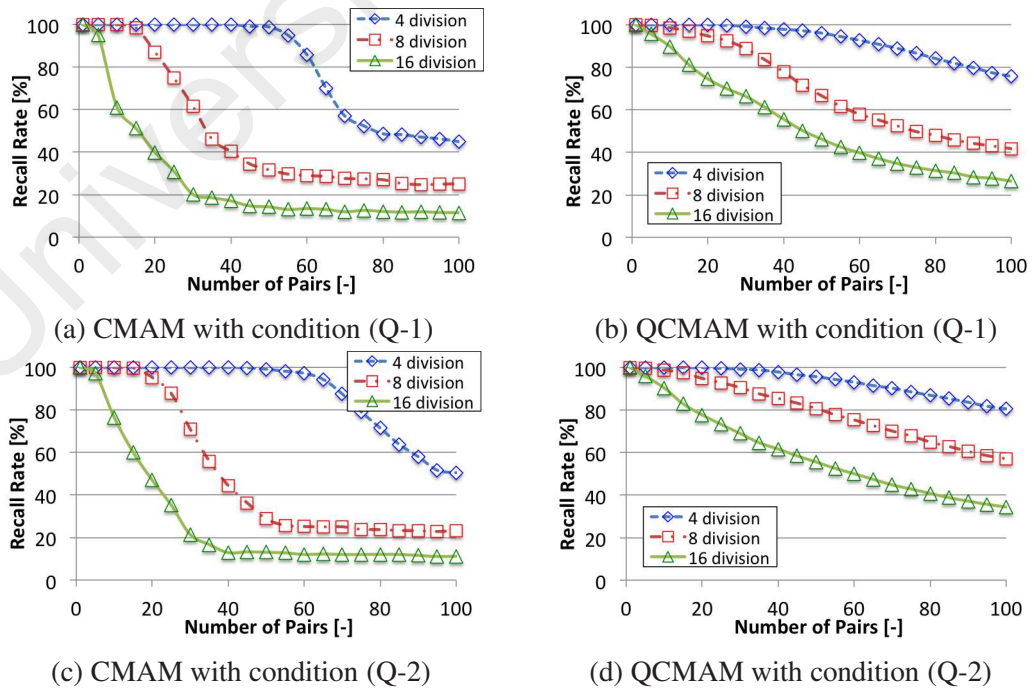
Figure 4.10: Results of memory capacity with random generated data on conditions Q.

Figure 4.11 shows the experimental results of 5-layer model with different number of neurons. Here, the features of experimental results of CMAM and QCMAM under conditions (R-1) and (R-2) are similar to the results in Figure 4.10. In addition, though 5-layers model, QCMAM exhibits superior stability against the effects of differences in the number of neurons in layers.
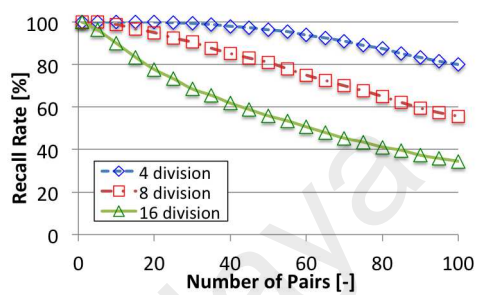

(a) CMAM with condition (R-1)   (b) QCMAM with condition (R-1)
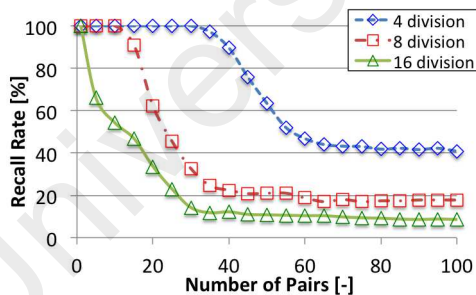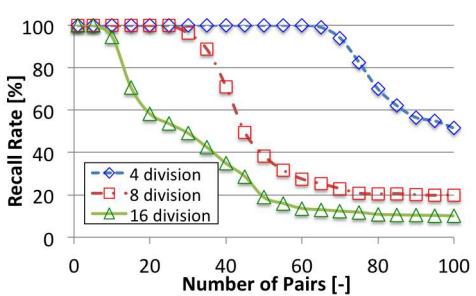
(c) CMAM with condition (R-2)   (d) QCMAM with condition (R-2)
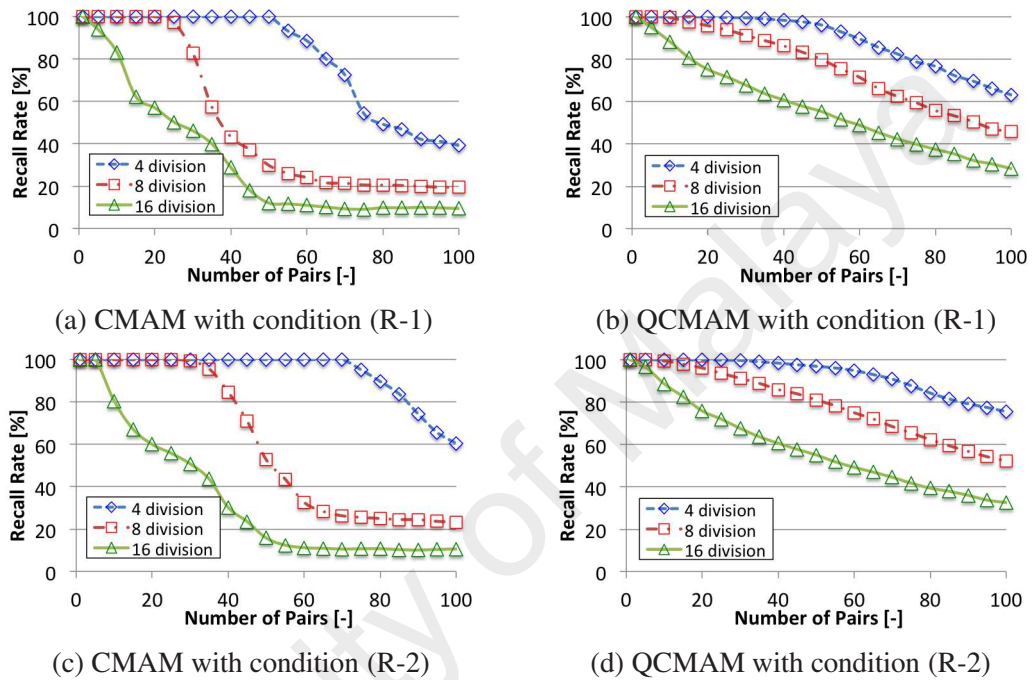
Figure 4.11: Results of memory capacity with random generated data on conditions R.

4.5.2.2 Noise Tolerance

The conditions for noise tolerance with random bipolar data are presented in Table 4.8. Here, the number of stored pairs is fixed as 40, and the noise ratio will be changed $(0 - 100[\%])$ in the layer 1. Similar to the conditions of memory capacity, the others layers are set the random phase information $(0 \leq 2\pi n/r < 2\pi[\text{rad}])$. The recall rates in Figures 4.12 and 4.13 with $0[\%]$ noise rate are equal to the recall rates of Figures 4.8 and 4.10, and Figures 4.9 and 4.11, respectively, in the point of 40 stored pairs.

Figure 4.12 shows the experimental results of noise rate in 3-layer model. Similar to the features of the bidirectional model, QCMAM shows better recall rate than CMAM,

Table 4.8: Conditions for noise tolerance in the multi-directional model.

| | Model | Condition | Layer No. | | | | |
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	**3-layers**	**(S-1)**	200	200	200	–	–
		(S-2)	150	200	250	–	–
		(S-3)	250	200	150	–	–
	5-layers	**(T-1)**	200	200	200	200	200
		(T-2)	160	180	200	220	240
		(T-3)	240	220	200	180	160
Neuron Representation	Complex Number						
Data Set Configuration	Amplitude: 1.0, Phase: Random						
Number of Pairs	40						

particularly, in presence of a high number of divisions in the complex unit circle. Furthermore, comparing the results with conditions (S-2) and (S-3), the stability against the effect of differences in the number of neurons in layers is also maintained in the case of noise tolerance.

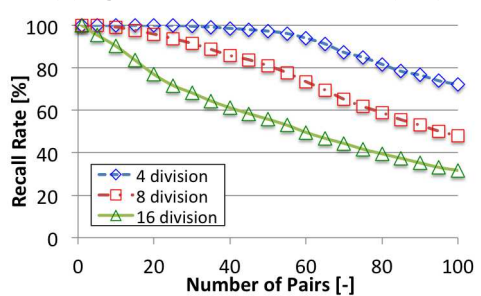(a) CMAM with condition (S-1)   (b) QCMAM with condition (S-1)

(c) CMAM with condition (S-2)   (d) QCMAM with condition (S-2)

(e) CMAM with condition (S-3)   (f) QCMAM with condition (S-3)
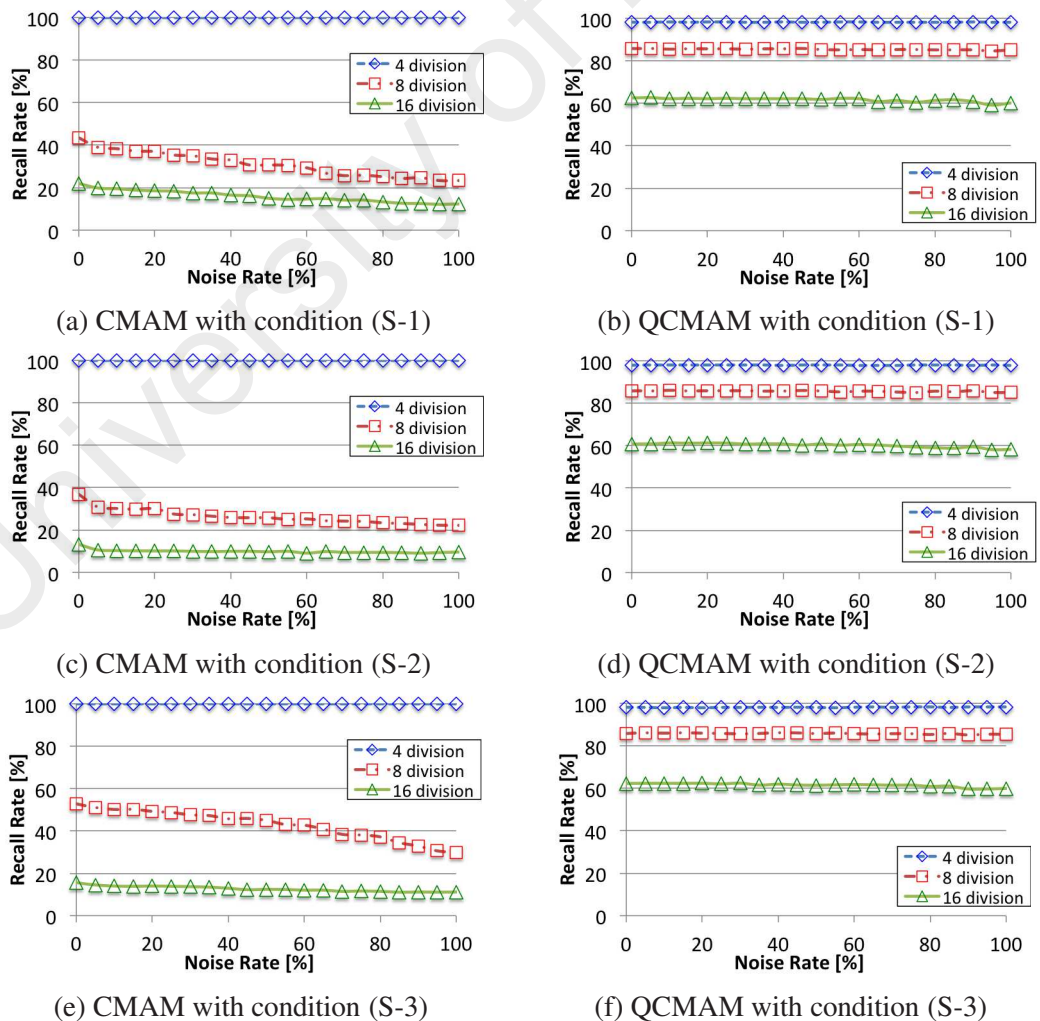
Figure 4.12: Results of noise tolerance with random generated data on conditions S.

Figure 4.13 shows the results of noise rate in 5-layers model. In general, if the number of layers in the model is increased, the noise tolerance ability also increased correspondingly. Therefore, comparing the conditions of the experimental results in 3-layer model (Figure 4.12) and 5-layer model (Figure 4.13), it is evident that the 5-layer model shows better recall rate in CMAM and QCMAM. In Figure 4.13, there is a huge difference between the results of CMAM with conditions (T-2) and (T-3). Whereas, similar to the 3-layer model, the results of QCMAM are quite stable and maintain high recall rate with varied conditions.
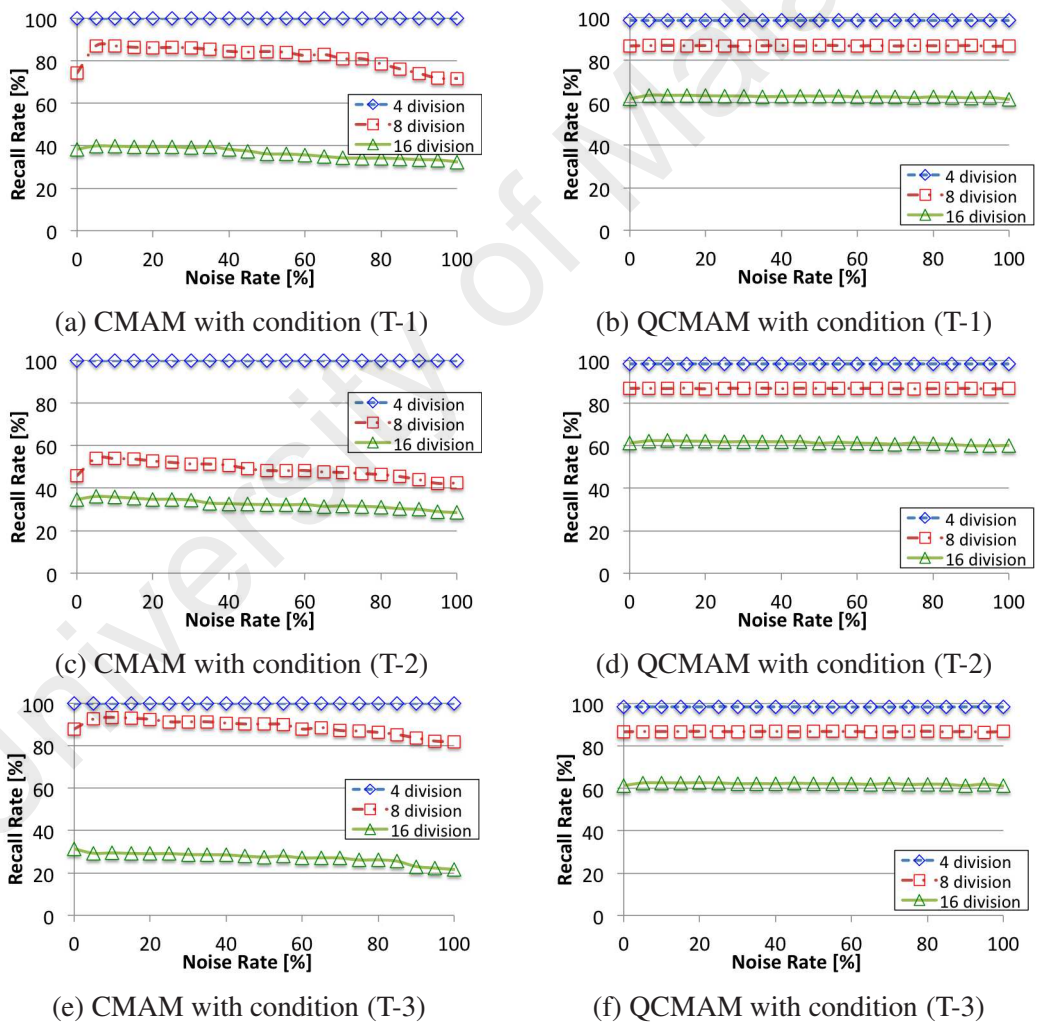
Figure 4.13: Results of noise tolerance with random generated data on conditions T.

From the results of memory capacity and noise tolerance with random bipolar data, it can be regarded that QCMAM has the superior abilities than CMAM.

4.5.3 Multi-directional Association Model with Gray-scale Image Data

As mentioned in Section 4.5, "noise" in the association models is roughly classified into two types. In this section, the noise tolerance based on the similarity of the stored patterns is evaluated using the image based data. Here, through the simulation experiments, the association from the layer 1 to 3 (3-layers model) or 5 (5-layers model) will be performed cyclically until equilibrium state is attained. The initial state of each layer differs based on the experiments of memory capacity and noise tolerance.

The image based data is generated from tiny images of CIFER-10 image database (Krizhevsky & Hinton, 2009). This database is known as the large image database of 32×32 color images in 10 classes. In this section, the images are chosen randomly from 10 classes, converted from the color image to gray-scale image, and subsequently resized to 20×20 pixels. Thus, it can be considered that the selected images have much similarity as Figure 4.14. Furthermore, the images are posterized to 4, 8 and 16 divisions as Figure 4.15. In the noise tolerance simulation, the specific ratio of salt and pepper noise will be added to stored data as Figure 4.16.
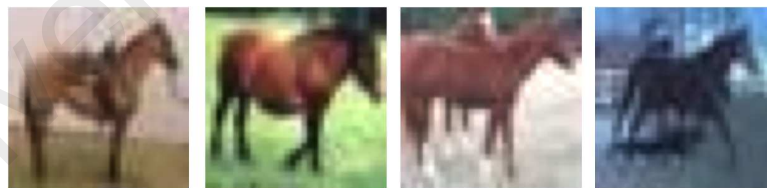
Figure 4.14: Images of horse class in CIFER-10 image database.
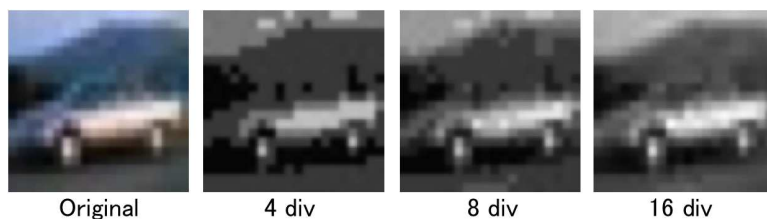
Figure 4.15: Posterized image based on each division.
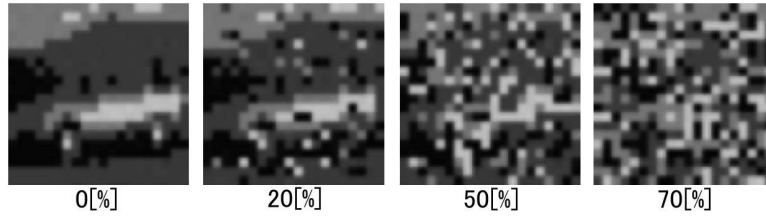
Figure 4.16: Noise examples on 4 division images.

4.5.3.1 Memory Capacity

The conditions of memory capacity with image based data are shown in Table 4.9. Same as Section 4.5.2.1, the initial setting of each layer is as follows; layer 1 is set the desired information, and other layers are set the random phase information ($0 \leq 2\pi n/r < 2\pi$[rad]).

Table 4.9: Conditions for memory capacity based on CIFER-10 image database in the multi-directional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	**3-layers**	**U**	400	400	400	–	–
	5-layers	**V**	400	400	400	400	400
Neuron Representation	Complex Number						
Data Set Configuration	CIFER-10 image database						

Figures 4.17a and 4.17b show the results of CMAM and QCMAM of 3-layers model, respectively. Here, the recall rate of CMAM decreases dramatically with the increase in the number of pairs, whereas, the recall rate of QCMAM maintains superior results even when a large number of pairs are stored.

(a) CMAM   (b) QCMAM
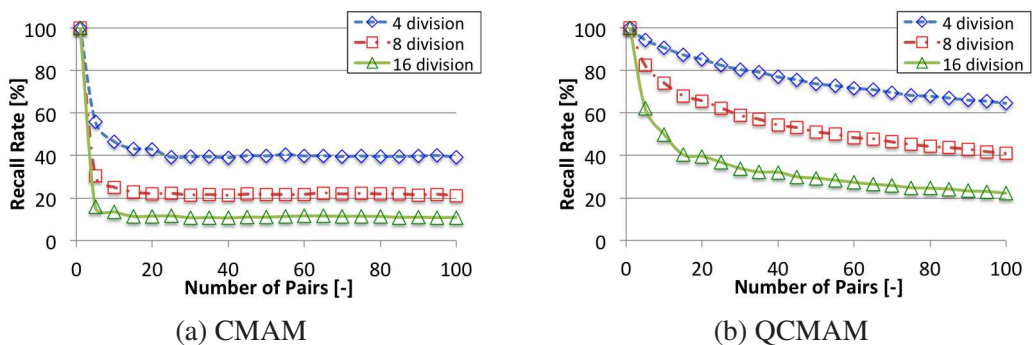
Figure 4.17: Results of memory capacity based on CIFER-10 image database with condition U.

Figures 4.18a and 4.18b show the results of CMAM and QCMAM of 5-layers model, respectively. As in the case of the 3-layer model, the recall rate of CMAM decreases quickly. On the other hand, the recall rate of QCMAM maintains outstanding results. Comparing the 3-layers model (Figure 4.17) with the 5-layers model (Figure 4.18), the 5-layers model shows a lower recall rate as compared to the 3-layers model. The same feature is shown in the results in Section 4.5.2.1 (Figures 4.8 and 4.9).

(a) CMAM
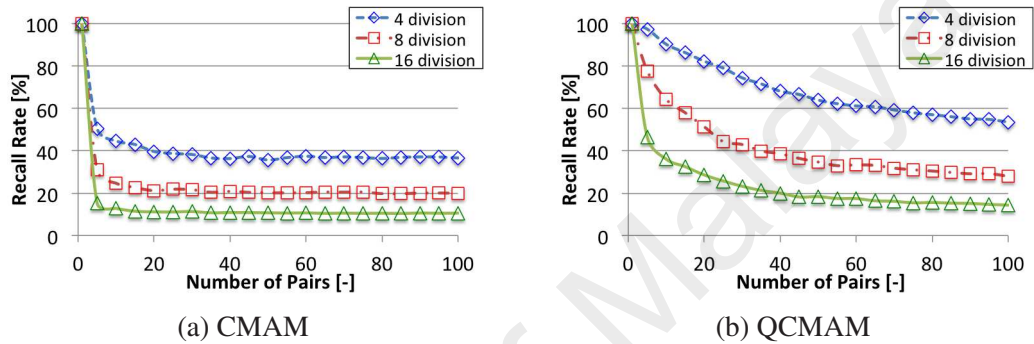
(b) QCMAM

Figure 4.18: Results of memory capacity based on CIFER-10 image database with condition V.

Here, due to the similarity of the stored data, the recall rate of CMAM and QCMAM (Figures 4.17 and 4.18) are lower than the results in Section 4.5.2.1 (Figures 4.8 and 4.9). This implies that not only is noise contained in the data, but also the similarity of the stored data influences the performance of the model. However, QCMAM has the superior resistance to the noise that is caused by the similarity of data.

4.5.3.2 *Noise Tolerance*

The conditions for noise tolerance with image based data are presented in Table 4.10. Here, the number of stored pairs is fixed as 5, and the noise ratio is changed $(0 - 100[\%])$ in the layer 1. As in the conditions of memory capacity, the others layers are set the random phase information $(0 \leq 2\pi n/r < 2\pi[\text{rad}])$. The recall rates in Figures 4.19 and 4.20 with $0[\%]$ noise rate are equal to the recall rate of Figures 4.17 and 4.18, respectively, in the point of 5 pairs stored.

Table 4.10: Conditions for noise tolerance based on CIFER-10 image database in the multi-directional model.

	Model	Condition	Layer No.				
			[1]	[2]	[3]	[4]	[5]
Number of Neurons	**3-layers**	**W**	400	400	400	–	–
	5-layers	**X**	400	400	400	400	400
Neuron Representation	Complex Number						
Data Set Configuration	CIFER-10 image database						
Number of Pairs	5						

Figures 4.19a and 4.19b show the results of noise tolerance of 3-layered CMAM and QCMAM with image based data. Here, as shown by the condition W in Table 4.10, the model stored only 5 stored pairs with 400 neurons. Comparing with the condition (S-1) in Table 4.8 (40 stored pairs with 200 neurons), it can be assumed that the results with the conditions W show better recall rate because the conditions are set with high number of neurons and low number of stored pairs. However, because the similarity of the stored data has great influence to the ability of model, the results of CMAM (Figure 4.19a) show poor performance than the results in Figure 4.12a. In contrast, the results of QCMAM (Figure 4.19b) show almost the same performance when compared to the results in Figure 4.12b. Thus, it can be considered that QCMAM has the superior tolerance to the impact of the similarity of data as compared to CMAM.

(a) CMAM   (b) QCMAM
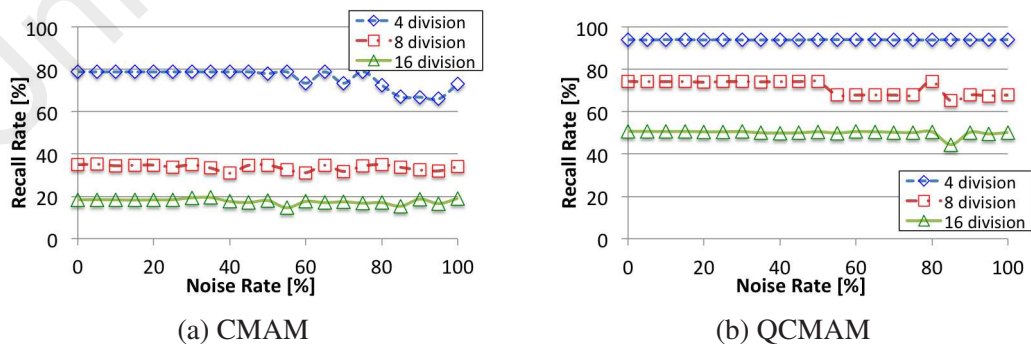
Figure 4.19: Results of noise tolerance based on CIFER-10 image database with condition W.

Figures 4.20a and 4.20b show the results of noise tolerance of 5-layers CMAM and QCMAM with image based data. Same as 3-layers model, comparing with the results with the conditions X in Table 4.10 and (T-1) in Table 4.8, the condition X with CMAM (Figure 4.20a) and QCMAM (Figure 4.20b) show the lower recall rate than the condition (T-1) with CMAM (Figure 4.13a) and QCMAM (Figure 4.13b), respectively. However, even in the case of 5-layers model, QCMAM shows better tolerance to the similarity of data than CMAM.

(a) CMAM   (b) QCMAM
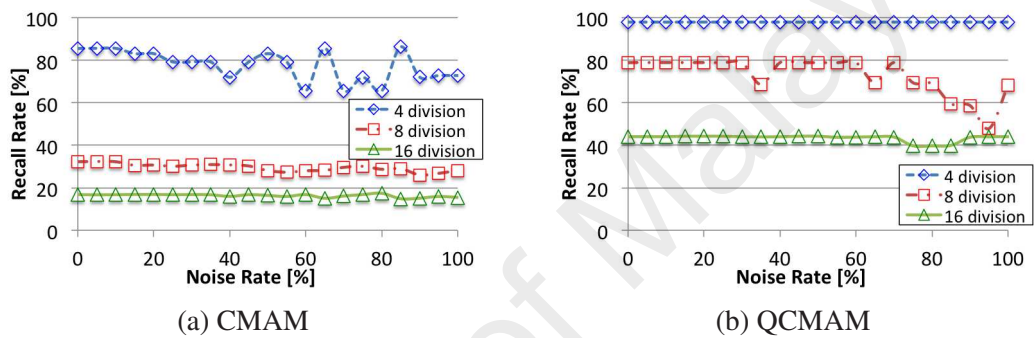
Figure 4.20: Results of noise tolerance based on CIFER-10 image database with condition X.

From the results of memory capacity and noise tolerance with image based data, it can be regarded that QCMAM has superior abilities than CMAM.

4.6 Summary

This chapter introduced the quantum-inspired complex-valued hetero-association models as QCBAM and QCMAM, which are extended models of QBAM and QMAM, respectively. Due to complex-valued neuron, the models are able to handle higher dimensional information better than the real-valued model. Based on real-valued model, the mathematical proofs of superposition and unitarity for the complex-valued model are presented. Furthermore, the numerical example is provided to confirm the satisfaction of superposition and unitarity in quantum-inspired complex-valued hetero-association models.

From the results of simulation experiment, quantum-inspired complex-valued association models with random phase information show the superior memory capacity and outstanding noise tolerance as compared to the conventional model. Furthermore, the results from the image based data show that QCMAM has high resistance to noise caused by the similarity of data. This feature is regarded to be advantageous in handling practical information such as image data.

CHAPTER 5

INTERACTIVE ROBOT SYSTEM WITH EMOTION AFFECTED ASSOCIATIVE MEMORY

5.1 Introduction

The decision making process in communication is affected by the internal and external factors from dynamic environments. Humans can exhibit a variety of behaviors in individuals, which is considered as humanity, and this has been found lacking in the robot. This chapter discusses the human psychological phenomenon during communication from the point of view of internal and external factors, such as perception, memory, and emotional information. Based on these, an interactive robot system is proposed. It can be expected that the capability of AM would enhance to simulate an actual function in human brain.

5.2 Memory and Emotion

The feature of research on emotion and memory in the psychological field does not only consider the emotional valence of the stimulus, but also consider the emotional state of the subject. The pioneering research done by Bower (1981) demonstrates the influence of emotional state on the human memory. In the experiment, subjects were artificially set to be in a Happy or Sad state. Subsequently, the subjects tried to memorize specific words, associate with past events, and remember sentences they wrote in a diary. Based on experiments by Bower, it is showed that while experiencing positive feelings, humans have the tendency to recall positive events based on personal experiences and knowledge. Similarly, while experiencing negative feelings, negative events are likely to be recalled.

Reisberg and Hertel (2004) discussed interesting mutual relationship between memory and emotion. For instance, a memory can be recalled using emotional information. Similarly, emotional information can be recalled based on a memory. Thus, since humans feel both positive and negative feelings, it becomes easier to obtain both positive and negative information. As mentioned in Section 1.5 of Chapter 1, this psychological phenomenon is called the mood-congruency effect (Ekkekakis, 2013).

In the past, several studies have shown the effects and properties of mood-congruency effects in human activities based on the analysis of brain signal processing and several human-based studies. Lewis et al. (2005) assumed that mood congruent facilitation is caused by mood-related reactivation and retrieval of emotional responses that are linked to the encoded valenced information in the associative memory. In the experiment, they presented the subjects with positive and negative words and manipulated their mood while monitoring brain activity by using fMRI. Pierce and Kensinger (2011) studied the effects of emotional valence and arousal on associative connection based on human study by using the negative, positive, and neutral word pairs. Murray and Kensinger (2013) considered the relationships of associative memory and emotion that takes into account the factors of age and gender by using the word pairs, in the case of emotional and non-emotional, respectively. Egidi and Nusbaum (2012) utilized visual and sound information to study the mood-congruency effect for a discourse comprehension based on analysis of EEG signal during experiment. Ravaja and Kätsyri (2014) revealed the relationships between facial expression and mood congruency effect based on facial electromyography analysis. From the above studies, the impact and effectiveness of mood-congruency effect on the association and recall processes with multi-modal information in the human brain can be confirmed.

5.3 Affinity of Emotional Information and Complex-Valued Associative Memory Model

In psychology, emotional information can be considered as waveform information (Rottenberg & Gross, 2003), for example, the emotion state is regarded as a short-term emotional state, and mood state is regarded as a long term emotional state, respectively. On the other hand, a number of studies of the human brain indicate that the neurons in the brain transmit electrical activity, which is the basis of neural oscillation (Başar, 2012). In addition, several studies have shown that the brain activity and emotional information are closely related each other (Hooker et al., 2012).

As shown in Section 5.2, the association process with the mood-congruency effect can be seen from the differences in brain activities through fMRI analysis. Thus, it is acceptable to represent the emotional information by the state of the neurons. In terms of the artificial neuron model, the complex-valued model is regarded as one of the oscillator models, due to the phase information. In general, the complex-valued model has the rich expressive power than the real-valued model. Thus, the complex-valued model has affinity to be applied to handle the internal state of neuron with emotion information.

Therefore, it is clear that the relationships between memory and emotion are interesting, and provide significant functions to humans. In this thesis, the complex-valued associative memory model and the emotional model are applied to simulate the mood-congruency effect for an interactive robot system, which aims to enhance the capability of associative memory model for the humanoid, is presented in Section 5.4.

5.4 Interactive Robot System

In this section, an interactive robot system that applies complex-valued associative memory and the emotional model is presented. Figure 5.1 shows the structure of the interactive robot system. The robot system is able to generate different emotional information from multi-modal information depending on personality factors. Moreover, recalling the

behaviors of the robot partner in the associative memory is affected by multi-modal information and the emotional factor.
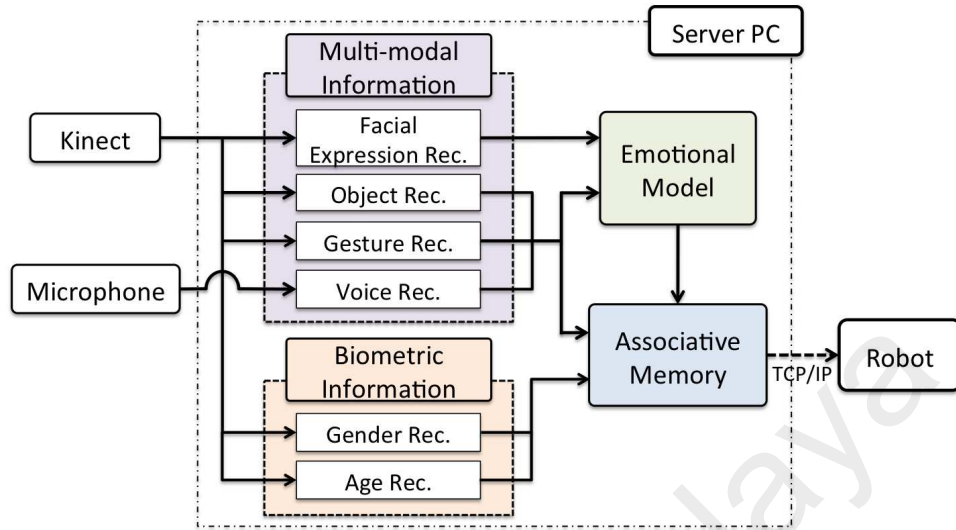
Figure 5.1: Configuration of interactive robot system

In this system, the QCMAM is applied as the complex-valued associative memory as described in Section 4.2.2. The three stage of the emotional model (i.e., core affect, emotion and mood) are based on the 2D (pleasure-arousal) scaling model with OCEAN personality factors. The structure of the emotional model is shown in Figure 5.2. Further details of model are described in APPENDIX A.

5.4.1 System configuration

The system consists of a robot, a Microsoft Kinect, a microphone, and a Personal Computer (PC). The system comprises cognitive intelligences, associative memory, and the emotional model.

The Kinect extracts red, green, and blue (RGB) data and distance data. Kubota and Toda (2012) implemented color based object recognition based on k-means algorithm and Steady-State Genetic Algorithm (SSGA), and hand gesture recognition based on Spiking Neural Networks (SNNs). The microphone records the human voice signals for Julius module (A. Lee & Kawahara, 2009) to process voice recognition. The human facial expressions are obtained using Constrained Local Model (CLM) (Cristinacce & Cootes,
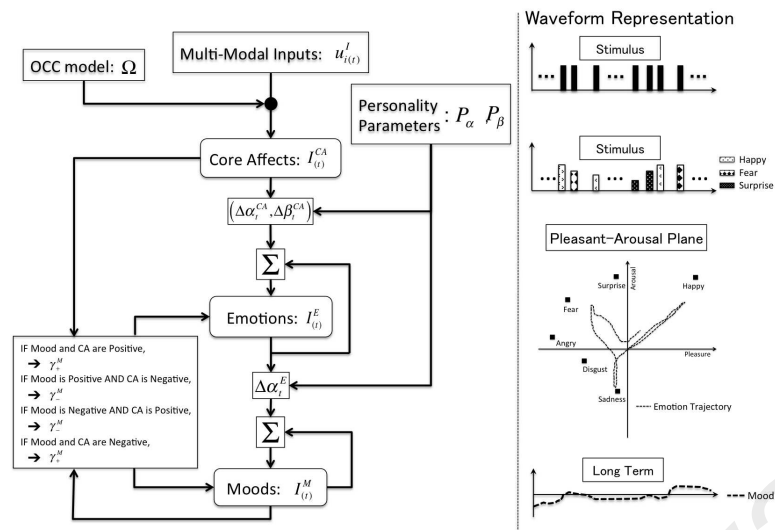
Figure 5.2: Structure of the emotional model and its waveform representation

2008) with two clustering algorithms. Furthermore, the age and gender information are estimated as the biometric information that is obtained from the facial information.

Based on multi-modal information, emotional states will be generated in the emotional model. The PC also calculates the relationships between information by using QCMAM depending on mood states. First, relationships between object, gesture, and voice information are associated. Subsequently, the robot actions are determined based on the result of association step and biometric information. The PC sends the behavior order to the robot by using TCP/IP. Here, a smart robot called iPhonoid, which was developed by Kubota and Toda (2012), is used as the robot for the system. The robot comprises an iPhone and four servo motors. It can perform arm motion, display representation, and audio output at simultaneously (Figure 5.3).
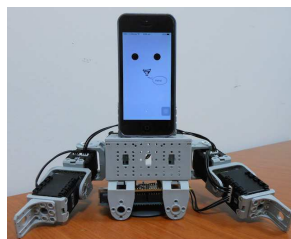
(a) iPhonoid
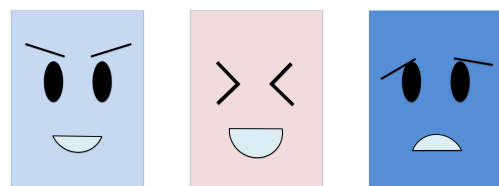
(b) Face variations: Neutral (left), Happy (center), Sad (right)

Figure 5.3: iPhonoid and example of the face templates

Majority of the conventional human–robot interaction systems focus on specific modalities, such as gesture-based, voice-based, facial expression based, or a combination of voice and facial-based systems. Therefore, these systems can handle limited communication situations. Furthermore, most of the systems focus on the detection of human emotional information to determine the actions of the robot, rather than the generation of emotional information for the robot. On the other hand, the interactive robot system in this thesis is able to generate the emotional information for the robot from the multi-modal and biometric information to determine the actions of the robot. Together with other sophisticated cognitive intelligences, the system is able to provide various types of expressions to communicate with a human.

5.4.2 Cognitive intelligence for robot

The robot system utilizes several existing studies to collect external information. In this section, the cognitive intelligences for multi-modal information (object, gesture, facial expression and voice) and the biometric information (age and gender) are introduced briefly.

5.4.2.1 Object recognition

Various types of algorithm have been proposed in regards to object recognition, such as the template matching methods by Dynamic Programming (DP) (Bertsekas, 1987), cellular neural network (J. A. Anderson, 1993), and GA (Back et al., 1997). While cellular neural network requires the exact templates for target, DP and GA detect targets based on similarity or distance based cost functions as an optimization problem. Basically, GA can be divided into a generational model (standard GA) and steady-state model (SSGA). The SSGA partially replaces a few individuals with offspring in a generation, not all individuals as standard GA. In particular, SSGA is suitable to solve optimization problems in the dynamic or changing environments (Kubota & Fukuda, 1999).

In general, the interactions are performed under dynamic environments with the object recognition processing only color information and not topological information. For this reason, the robot system utilizes the SSGA based object recognition (Kubota & Toda, 2012).

5.4.2.2 Gesture recognition

In general, the context of hand gesture is quite flexible depending on the hand speed and its movement. Therefore the significance of gesture recognition is to extract the spatio-temporal information from the hand movements. It is well known that SNN is capable of handling spatial and temporal context (Maass, 1997). In addition, because of the spike response model, SNN is able to reduce the computational cost as compared to integrate-and-fire model. Thus, the robot system applies SNN based hand gesture recognition method that is introduced by Kubota and Toda (2012).

5.4.2.3 Facial expression recognition

The facial feature tracking framework combines two incremental clustering algorithms with a classical method such as CLM (Cristinacce & Cootes, 2008), to allow the system to extract dynamic human facial features with higher accuracy and minimized tracking error while avoiding shape distortion. Nazrul et al. (2014) adopted two distinct incremental clustering algorithms and integrated them using the baseline method of CLM for improving feature tracking accuracy and robustness during the tracking of dynamic face appearance. Both algorithms, Patch Clustering (modified algorithm of LeaderP (Nuevo et al., 2010)) and Shape Clustering (modified algorithm of Topological Gaussian Adaptive Resonance Theory Algorithm (TGART) (Dawood et al., 2013)), are applied to cluster updating with incoming patches/shapes throughout tracking. In the robot system, facial expression recognition is processed by the CLM using two clustering algorithms, which was proposed by Nazrul et al. (2014)

5.4.2.4 Voice recognition

Verbal communication is essential to humans. In the past, various studies have been conducted on voice recognition. One of the established open source softwares is Julius (A. Lee & Kawahara, 2009). Julius works in real time, with a recognition accuracy that is over 90% in a 20,000-word reading test. Thus, Julius is applied to the Japanese language model for voice recognition in the robot system.

5.4.2.5 Biometric recognition

Biometric information is one of the significant factors in human emotional information and behaviors. Specifically, the human have tendency to change their attitudes and reactions based on the facial appearances of their communication partner. Thus, it can be considered that the above ability will enable the acquisition of unique characteristics for a robot. Eidinger et al. (2014) developed the age and gender estimation algorithm. The algorithm applied Local Binary Pattern (LBP) to extract the features, and classification of the features is performed using the standard linear Support Vector Machine (SVM) that is trained by feature vectors of LBP. Eidinger et al. (2014) report extensive tests that analyze the difficulty levels of contemporary benchmarks and the capabilities of their algorithm. These tests show that the algorithm outperforms the state-of-the-art algorithms by a wide margin. The system utilizes a part of above algorithm that is customized to only detect male/female as gender information and young/adult as age information for minimum configuration.

5.5 Experiments with Interactive Robot System

This section presents the experimental results of the robot system. The robot exhibits a number of behaviors based on the robot's personality and the biometric information of the communication partner. The following subsections describe experimental conditions that are related to the emotional model and associative memory model, and results of the sim-

ple interaction with the robot are presented. Due to the limitations of cognitive abilities of the robot, the practical relationships between individual information and movements are ignored. Thus, simple symbolic information is utilized for association. In practical terms, relationships are defined based on actual information, facts, common senses or personal knowledge. The main focus of this experiment is to consider the association result from AM that change based on the emotional factors and the biometric information of the communication partner.

The experiment is divided to two parts. First, the effect of emotional model is evaluated. Next, multi-modal communication is performed with robot to confirm the association results are changed based on emotional information as the mood-congruency effect. Based on the multi-modal information from the human, personality affected emotional information will be generated. Subsequently, association will be performed, and the robot actions will be determined.

5.5.1 Experimental Conditions of Personality Affected Emotional Model

As shown in Figure 5.1, the behavior of robot is controlled by the modules of emotional model and the associative memory model that are based on multi-modal biometric information. In each module, there are several predefined information. The emotional model is shown in Figure 5.2. The definitions of the variables in the emotional model are described as follows; each multi-modal input is assumed to be assigned a specific emotional intensity Ω by using the Ortony, Clore and Collins (OCC) model (Ortony, 1990), which is considered for appraisal of the emotional information of multi-modal information, as shown in Table 5.1. Table 5.2 shows the definitions of a suppression ratio in emotional model (γ^M). Four types of personalities are defined as given in Table 5.3. These personality factors are used to calculate Equations (A.1) and (A.2) in APPENDIX A. In the robot system, biometric information is utilized to determine the types of personality factors as

Table 5.4. In general, the parameters in Tables 5.1 and 5.2 can be considered as the personal knowledge and experience to generate the emotional information from multi-modal information. However, assumed experimental space in this thesis is small space with few simple symbolic information between the human and the robot. Therefore, the parameter settings are arranged empirically to obtain appropriate differences in results of emotional information.

Table 5.1: Emotional intensity Ω of Multi-modal Information

Attribution		Emotional Intensity Ω					
		Happy	Sad	Angry	Fear	Disgust	Surprise
Object:	Red Cir	0.7	0.0	0.0	0.0	0.0	0.1
	Blue Rec	0.0	0.6	0.0	0.1	0.1	0.0
Gesture:	Circle	0.6	0.0	0.0	0.0	0.0	0.1
	Bye-Bye	0.0	0.1	0.0	0.7	0.1	0.0
Voice:	Positive	0.7	0.0	0.0	0.0	0.1	0.0
	Negative	0.0	0.1	0.0	0.1	0.7	0.0
Facial Rep.:	Happy Face	0.7	0.0	0.0	0.0	0.0	0.1
	Sad Face	0.0	0.7	0.0	0.1	0.1	0.0

Table 5.2: Parameter settings of suppression ratio γ^M

	Input	Positive Mood	Negative Mood
γ^M	Positive Mood	0.95	0.60
	Negative Mood	0.60	0.95

Table 5.3: Four types of personality.

Factor	Personality			
	Type 1	Type 2	Type 3	Type 4
Openness	0.50	0.90	0.50	0.20
Extraversion	0.90	0.70	0.20	0.10
Agreeableness	0.50	0.50	0.20	0.30
Neuroticism	0.20	0.30	-0.40	-0.40

Table 5.4: Determination of personality type based on biometric information

Attribution		Gender:	
		Male	Female
Age:	Young	Type 1	Type 2
	Adult	Type 3	Type 4

5.5.2 Results of Personality Affected Emotional Model

Figure 5.4a shows the history of multi-modal inputs. Here, IN_ID represents the types of multi-modal inputs as given in Table 5.7. The multi-modal inputs, except for facial information, are used as inputs for QCMAM. Using the multi-modal inputs, the emotional model extracts the emotional information for core affect as shown in Figure 5.4b based on the predefined weight of OCC model that is shown in Table 5.1. According to emotional model as shown in Figure 5.2, core affect is not affected by the personality parameters. Thus, result of core affect shows same values even if the personality factors are different.

(a) Continuous inputs of multi-modal information (IN_ID)

(b) Intensity of core affect based on multi-modal information
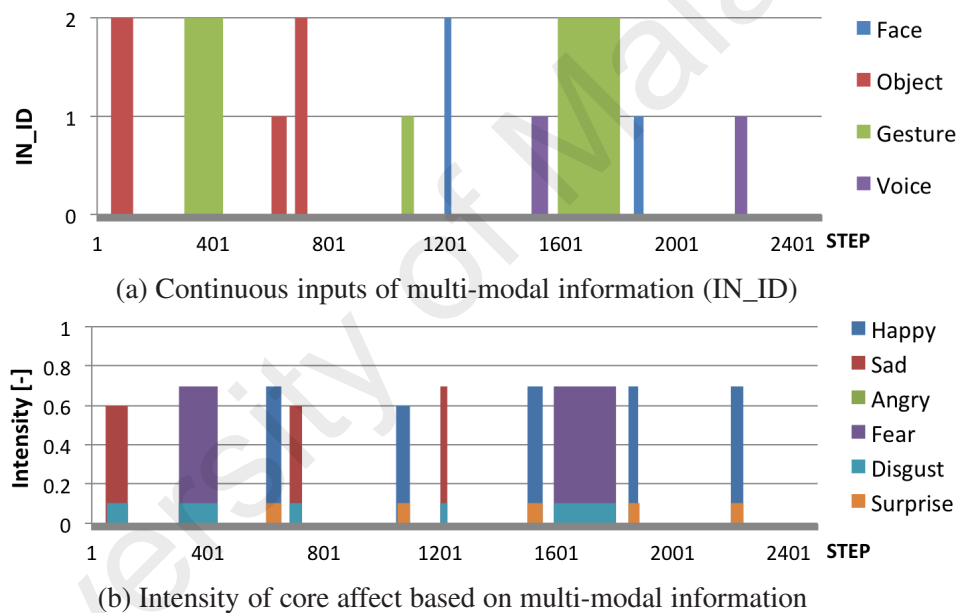
Figure 5.4: Internal states of emotional model. (a) Multi-modal inputs and (b) Core affect.

In contrast, due to the effect of personality factors given in Table 5.3, trajectory of emotion and mood state show different results as shown in Figures 5.5 and 5.6. Figure 5.5 shows the trajectory of emotion state on pleasure and arousal plane based on the outputs of core affects. For example, if Happy intensity is provided from core affects, the output of emotion will be closer to the coordinates of Happy ((pleasure, arousal)=(0.9, 0.9)) on the 2D plane.
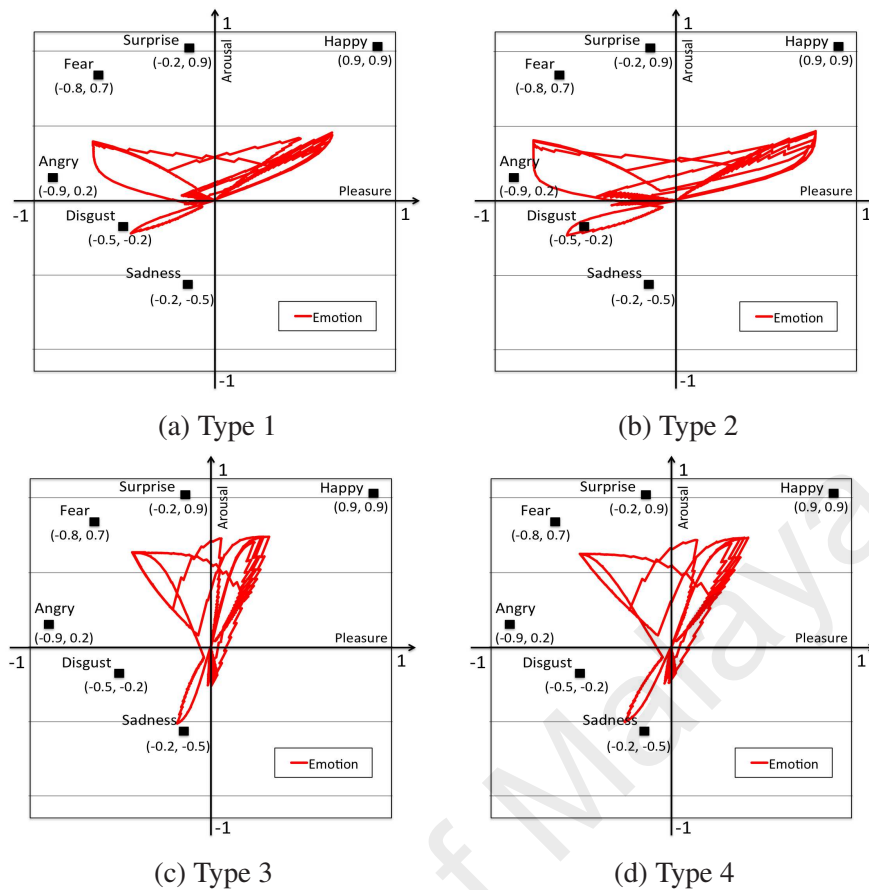
Figure 5.5: Trajectory of emotion states based on four types of personality

As defined in Tables 5.3 and 5.4, we can consider that a young woman has rich emotional sensitivity, high adaptability to the environment, whereas, an adult man has a calm character, and tolerance to the influence from the environment. From the definitions of personality factors given in Table 5.3, it can be regarded that Type 1 and 2 personalities indicate similar properties, as well as Type 3 and 4. Comparing Type 2 and 3 personalities given in Figures 5.5b and 5.5c, because Type 2 has high value of Openness and Extraversion as compared to Type 3, Type 2 personality shows a wide range of trajectories than Type 3, based on the same input information.

The mood state is calculated based on the emotion state by using Equation (A.14) in APPENDIX A. Figure 5.6 shows the trajectory of mood state. Because of the differences in emotion states and personality factors, the mood state in each personality type also generate different outputs. As mentioned, Type 2 personality is sensitive to stimulus.

Thus, the mood state of Type 2 personality changes quickly and violently. In contrast, Type 3 personality shows resistance to stimulus. Therefore, the mood state of Type 3 personality changes over a long term period as compared to the other types. Furthermore, the personality factors highlight a noteworthy feature of emotional model, which can be seen at ranges (i), (ii), and (iii) in Figure 5.6. At these ranges, the mood state of Type 1 and 2 indicate a positive state, while Type 3 and 4 indicate a negative state. Due to these differences in mood state, the association process is affected, which is shown in Section 5.5.4. The positions (a) to (f) are also mentioned in Section 5.5.4.
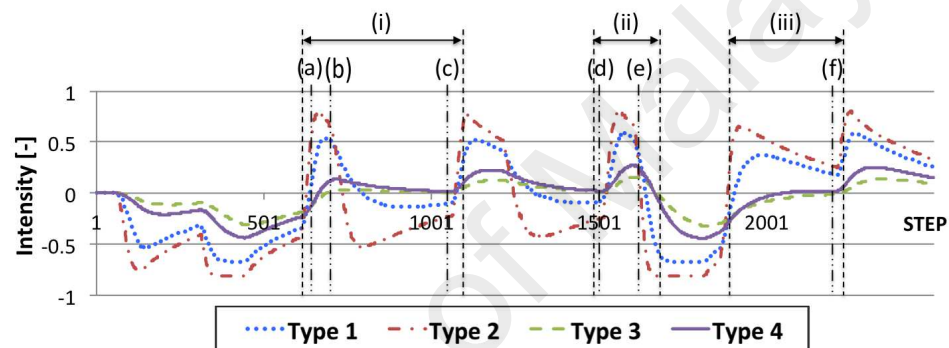
Figure 5.6: Trajectory of mood states based on four types of personality

From the results, it is clear that personality factors play an important role in introducing different emotional properties, and the functions of personality are successfully integrated in the emotional model. Thus, it can be considered that the robot is able to adapt its emotional reactions based on the appearance of the communicating partner; thus, providing appropriate responses.

5.5.3 Experimental Conditions of Interactive Communication with the Robot

In this section, all inputs of associative memory module are provided by the emotional model, namely, Figures 5.4a and 5.6. Furthermore, only 'male/female' as a gender factor and 'young/adult' as an age factor are utilized as biometric information for the simplicity of conditions. In the associative memory module, relationships between multi-modal information and robot action are defined as given in Table 5.5.

Table 5.5: Information Relationships for QCMAM (A.M._ID)

A.M._ID	Moods	Input Attribution	Input	Associated Information			Act._ID
				Object	Gesture	Voice	
0	P/N	-	No Input	-	-	-	0
1	Positive	Object	Red Circle	-	Circle	Happy	1
2	Negative	Object	Red Circle	-	Circle	Sad	3
3	Positive	Object	Blue Rectangle	-	Bye-Bye	Happy	2
4	Negative	Object	Blue Rectangle	-	Bye-Bye	Sad	4
5	Positive	Gesture	Circle	Red Circle	-	Happy	1
6	Negative	Gesture	Circle	Red Circle	-	Sad	3
7	Positive	Gesture	Bye-Bye	Blue Rectangle	-	Happy	2
8	Negative	Gesture	Bye-Bye	Blue Rectangle	-	Sad	4
9	Positive	Voice	Happy	Red Circle	Circle	-	1
10	Negative	Voice	Happy	Red Circle	Circle	-	3
11	Positive	Voice	Sad	Blue Rectangle	Bye-Bye	-	2
12	Negative	Voice	Sad	Blue Rectangle	Bye-Bye	-	4

These relationships are assigned specific ID, associative memory model is identified A.M._ID. Based on association results and mood states, the robot partner performs different actions, identified as Act._ID, as given in Table 5.6. For example, when the human shows a Red Circle as object information under positive moods to the robot, the robot will associate Circle and Happy as gestural and voice information, respectively. The robot will perform a circle gesture, and speak Happy with Happy face as Act._ID 1 in Table 5.6.

Table 5.6: Definitions of Robot Action ID (Act._ID)

Act._ID	Robot Action		
	Face	Gesture	Voice
0	Neutral	-	-
1	Happy	Circle	Happy
2	Sad	Up&Down	-
3	Happy	Up&Down	-
4	Sad	Bye-Bye	Sad

Figure 5.7 an example of gestural action by the robot. In this experiment, the specific ID (IN_ID) to multi-modal information is assigned as given in Table 5.7 for visualization. Here, information belonging to IN_ID 0 has a neutral attribution, IN_ID 1 has a positive emotional attribution, and IN_ID 2 has a negative emotional attribution.

(a) Circle   (b) Up&Down   (c) Bye-Bye
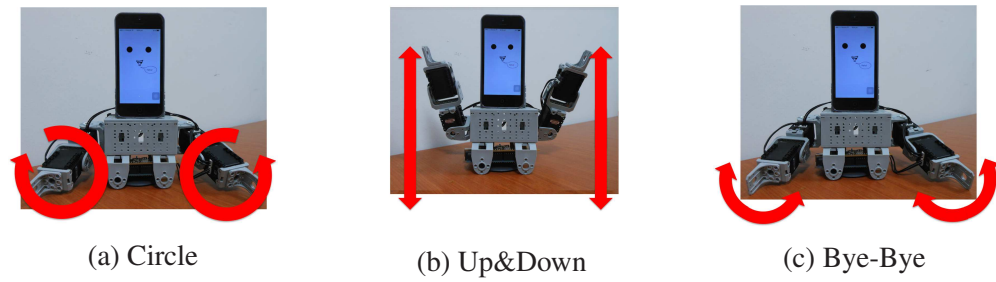
Figure 5.7: Example of robot actions with a neutral face

Table 5.7: Definitions of input information ID (IN_ID)

IN_ID	Information Attribution			
	Face	Object	Gesture	Voice
0	Neutral	No info.	No info.	No info.
1	Happy	Red Circle	Circle	Happy
2	Sad	Blue Rectangle	Bye-Bye	Sad

5.5.4 Results of Interactive Communication with the Robot

This section presents the association process of the robot action based on emotional factors. Figure 5.8a shows the association results (A.M._ID) in associative memory module, and Figure 5.8b shows the corresponding robot actions (Act._ID).
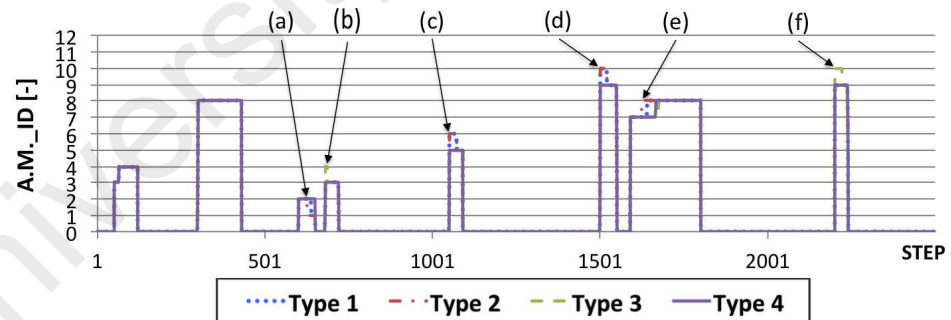
(a) Association results based on multi-modal inputs (A.M._ID)
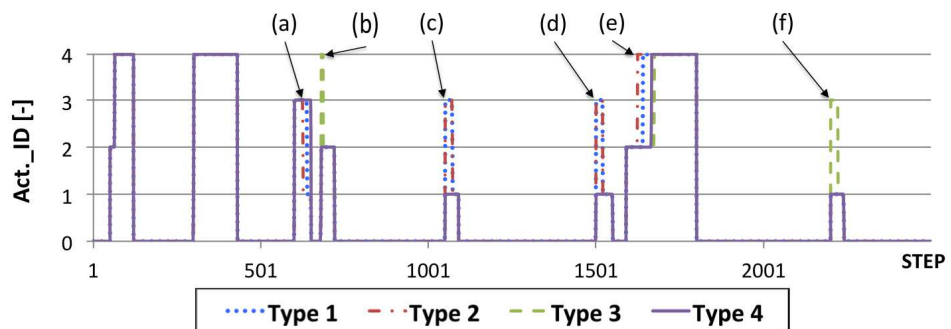
(b) Corresponding robot actions (Act._ID)

Figure 5.8: Association results of information relationships and corresponding robot actions

131

The associations are performed using multi-modal inputs and emotional information based on predefined relationships as given in Table 5.5. In Figures 5.8a and 5.8b, the positions of (a) to (f) are plotted in same step point in Figure 5.6, and these positions are included in ranges (i), (ii), and (iii) as shown in Figure 5.6.

Because of the differences in mood attribution, the association results are affected. In this thesis, from the definitions of personality factors given in Table 5.3, it is assumed that Type 1 and 2 personalities indicate similar properties, as well as Type 3 and 4. Thus, the positions (a), (c), (d), and (e) of Type 1 and 2 personalities show the same association results, which are different from Type 3 and 4. In particular, the positions (b) and (f) of Type 3 personality show unique results. Simultaneously, corresponding robot actions are also affected as shown in Figure 5.8b.

From these results, it can be considered that the differences in emotional information personality are able to manage different association results to the robot. It is assumed that suitable and actual information relationships with emotional factor can be prepared in AM, the emotion affected associative memory is able to perform the association corresponding to individual robots. It can be regarded that the capability enhancement of AM.

5.6 Summary

In this chapter, the relationships between associative memory and emotional information in the human brain are summarized from the perspective of psychology and brain science. Several studies have shown that during association tasks, the mood-congruency effect can be observed from the analysis of brain activity. Based on the above facts, the affinity between emotional information and complex-valued associative memory model is discussed. Furthermore, based on the above discussion, the emotion affected associative memory is implemented in the interactive robot system. In addition, cognitive intelligence in the robot system are briefly introduced.

From the experimental results, it is evident that the emotion affected associative memory is capable of associating different information from multi-modal inputs corresponding to the pre-defined relationships. It can be considered that the mood-congruency effect is integrated in AM.

In this thesis, the experimental simulations are not under practical conditions (applying symbolic information and simple information relationships). However, it is assumed that if the cognitive intelligence is well developed, there is a possibility to apply the system under complex relationships based on actual information, fact, and common sense. Therefore, the communication between human and robot would be more natural and active by the emotion affected associative memory.

CHAPTER 6

CONCLUSIONS, CONTRIBUTIONS AND FUTURE WORKS

In this chapter, the comprehensive summary of this thesis is presented, and the contributions are discussed. Furthermore, the future works are discussed from the viewpoint of limitations of the models in this thesis.

6.1   Conclusions

In this thesis, the artificial neural associative memory model, which is a recurrent neural network, has been improved from the perspective of ability and functionality. For the ability improvements, memory capacity and recall reliability of the real/complex-valued hetero-association models were evaluated by simulation experiments under several conditions. The experimental results showed that the quantum-inspired models have outstanding performance compared to the conventional models. In the case of functionality improvements, the relationships between memory and emotion in the human brain were discussed from a psychology and neuroscience point of view. As a result of the above discussions, the significance of the emotion affected associative memory, which is able to recall different information depending on the emotional information, was revealed and summarized based on neurocomputing. Furthermore, the emotion affected associative memory was implemented in the interactive robot system. The following descriptions summarize the knowledge obtained in each chapter:

In Chapter 1, the features and biological phenomena of the synaptic connections, and the associative memory system in the human brain were reviewed. An organized introduction to the approach to computational model and its improvement was also addressed. In addition, the relationships between memory and emotion were discussed from a psy-

chology and neuroscience perspective. The roles and significance of associative memory system in communication, based on the psychological theory of mood-congruency effect, were also revealed. Furthermore, based on discussions in the chapter, the need for associative memory system in neurocomputing was clearly described, and the research problems, objectives and contributions in this thesis were defined.

In Chapter 2, the fundamental knowledge pertaining the neuron, NNs, and associative memory model were presented. Furthermore, the conventional studies on associative memory models in terms of structural and algorithmic improvements, such as complex-valued model and quantum-inspired model, were reviewed. Additionally, conventional human-robot interaction systems with emotion affected associative memory model were introduced, and the significance of the system that integrated memory and emotional information for communication was pointed out.

In Chapter 3, the fundamentals of the real-valued quantum-inspired hetero-association memory models were presented. The mathematical proofs and numerical examples show that the features of QM, namely superposition and unitarity, are satisfied in the weight matrices of quantum-inspired hetero-association models. The simulation experiments showed that the quantum-inspired model shows superior memory capacity and stability against the effects of several types of noise tolerance as compared to conventional models.

In Chapter 4, based on the real-valued model, quantum-inspired complex-valued hetero-association models were introduced. Similar to Chapter 3, the mathematical proofs for the superposition and unitarity, which are satisfied in the weight matrices of quantum-inspired complex-valued hetero-association models model, were shown. Furthermore, numerical examples with complex numbers confirmed that superposition and unitarity are maintained in the weight matrices. The simulation experiments showed that the quantum-inspired models have superior performance as compared to the conventional model in terms of memory capacity and noise tolerance. In particular, quantum-inspired models

135

showed outstanding abilities even when the similarity of the stored information was high.

From the results of Chapters 3 and 4, it is evident that quantum-inspired models have superior memory capacity and stable noise tolerance under several conditions, when compared with the conventional models. Due to these advantages, it is possible to expect more stable operations than in the conventional models if the models are applied to the robot system under a practical environment.

Firstly, in Chapter 5, the conventional computational approaches of memory and emotion were reviewed, and the affinity of the complex-valued associative memory model and emotional information was discussed from a neurocomputing point of view. Based on the above discussions, the emotion affected associative memory was implemented in the interactive robot system as a functionality improvement. The system was able to associate the emotion dependent information corresponding to pre-defined relationships, which is characterized as the mood-congruency effect in the field of psychology, through the emotion affected associative memory model. With regards to interactive experiments for the robot, symbolic information and simple information relationships were utilized. From the experimental results, it was confirmed that the association results of emotion affected associative memory were changed based on emotional information.

6.2 Contributions

The research objectives of this thesis are achieved by corresponding research contributions, which are defined in Section 1.8 of Chapter 1, as the following viewpoints:

In Chapter 3, the features of QM were applied to real-valued hetero-association models with batch learning algorithm, namely QBAM and QMAM. It has succeeded in improving the memory capacity and recall reliability as compared to the conventional models. Furthermore, real-valued incremental learning algorithm models, namely IQBAM and IQMAM, were also introduced. The simulation experiments showed that IQBAM and IQMAM have succeeded in incorporating quantum features, and the models have performed comparably with the conventional models. Therefore, research objectives 1 and 2 have been addressed.

In Chapter 4, based on real-valued models, quantum-inspired complex-valued hetero-association models, namely QCBAM and QCMAM, were introduced. The simulation experiments showed that QCBAM and QCMAM have succeeded in improving the memory capacity and recall reliability while maintaining the features of QM. In particular, QCMAM showed an outstanding performance over the conventional models even with high similarity in the stored data. Thus, research objectives 1 and 3 have been achieved.

In Chapter 5, the emotion affected associative memory, which is focusing on the relationships between memory and emotion in the brain, was implemented in the interactive robot system as a functionality improvement. From the experimental results, it can be regarded that the new function has successfully been added into the associative memory model.

From the results and discussions in Chapters 3, 4, and 5, the research objectives, namely, improvements in artificial associative memory from the ability and functionality perspective, were addressed, thus the research problems in this thesis have been solved.

6.3 Future Works

QM have raised an interesting viewpoint to CI, and several studies of QCI have shown its superiorities and potentials. In this thesis, several types of artificial neural associative memory models have been improved from the Quantum information perspective. However, due to the definitions of network architecture and learning algorithms, there are several drawbacks.

With regards to the real world, information is mostly represented by continuous values. Furthermore, the quality, quantity and dimension of information are changed by various uncertainties. In this case, the models proposed in this thesis are not suitable, due to the prior requisition of fixing the number of neurons and layers in advance. In order to respond to this problem, the continuous-valued model with an incremental learning algorithm such as Growing Neural Gas (GNG) (Fritzke, 1995) can be suggested as one of the promising models. This model can classify information based on existing knowledge, and acquire a new category of information simultaneously. It could be expected that combining Quantum-Inspired association model and the growing neural architecture provides the possibility to represent properties closely related to human brain reasoning.

The ability for the robot to acquire new knowledge provides an emerging role to the interactive robot system. It allows the robot to acquire new knowledge, and is able to transfer the acquired knowledge to both robots and humans alike using suitable communication approaches. This kind of human-like behavior is considered as the basis of the adaptive robot, and it can initiate active and meaningful communication with humans.

REFERENCES

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines. *Cognitive science*, *9*(1), 147–169.

Adachi, M., & Aihara, K. (1997). Associative dynamics in a chaotic neural network. *Neural Networks*, *10*(1), 83–98.

Aghajari, Z., Teshnehlab, M., & Motlagh, M. J. (2015). A novel chaotic hetero-associative memory. *Neurocomputing*.

Aihara, K., Takabe, T., & Toyoda, M. (1990). Chaotic neural networks. *Physics letters A*, *144*(6), 333–340.

Aizenberg, N., Ivaskiv, Y. L., Pospelov, D., & Hudiakov, G. (1973). Multiple-valued threshold functions. ii. synthesis of the multi-valued threshold elements. *Kibernetika (Cybernetics)*, *1*, 53–66.

Anderson, J. A. (1993). *Neurocomputing* (Vol. 2). MIT press.

Anderson, J. R., & Bower, G. H. (2014). *Human associative memory*. Psychology press.

André, E., Klesen, M., Gebhard, P., Allen, S., & Rist, T. (2000). Integrating models of personality and emotions into lifelike characters. In *Affective interactions* (pp. 150–165). Springer.

Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., . . . Yoshida, C. (2009). Cognitive developmental robotics: A survey. *Autonomous Mental Development, IEEE Transactions on*, *1*(1), 12-34.

Back, T., Fogel, D. B., & Michalewicz, Z. (1997). *Handbook of evolutionary computation*. IOP Publishing Ltd.

Ball, G., & Breese, J. (2000). Emotion and personality in a conversational agent. *Embodied conversational agents*, 189–219.

Başar, E. (2012). *Brain function and oscillations: Volume ii: Integrative brain function. neurophysiology and cognitive processes*. Springer Science & Business Media.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, *10*(3), 295–307.

Bertsekas, D. (1987). *Dynamic programming: deterministic and stochastic models*. Prentice-Hall.

Bower, G. H. (1981). Mood and memory. *American psychologist*, *36*(2), 129.

Chakravarthy, V. S., Gupte, N., Yogesh, S., & Salhotra, A. (2008). Chaotic synchronization using a network of neural oscillators. *International journal of neural systems*,

18(02), 157–164.

Chartier, S., & Boukadoum, M. (2006a). A bidirectional heteroassociative memory for binary and grey-level patterns. *Neural Networks, IEEE Transactions on*, *17*(2), 385–396.

Chartier, S., & Boukadoum, M. (2006b). A sequential dynamic heteroassociative memory for multistep pattern recognition and one-to-many association. *Neural Networks, IEEE Transactions on*, *17*(1), 59–68.

Chartier, S., & Boukadoum, M. (2011). Encoding static and temporal patterns with a bidirectional heteroassociative memory. *Journal of Applied Mathematics*, *2011*.

Chartier, S., Renaud, P., & Boukadoum, M. (2008). A nonlinear dynamic artificial neural network model of memory. *New Ideas in Psychology*, *26*(2), 252–277.

Chen, C., Dong, D., Li, H., Chu, J., & Tarn, T. (2014, May). Fidelity-based probabilistic q-learning for control of quantum systems. *Neural Networks and Learning Systems, IEEE Transactions on*, *25*(5), 920-933.

Chung, F. L., & Lee, T. (1996). On fuzzy associative memory with multiple-rule storage capacity. *Fuzzy Systems, IEEE Transactions on*, *4*(3), 375–384.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The neo personality inventory. *Psychological assessment*, *4*(1), 5.

Cristinacce, D., & Cootes, T. (2008). Automatic feature localisation with constrained local models. *Pattern Recognition*, *41*(10), 3054 - 3067.

Dawood, F., Loo, C. K., & Chin, W. H. (2013). Incremental on-line learning of human motion using gaussian adaptive resonance hidden markov model. In *Neural networks (ijcnn), the 2013 international joint conference on* (pp. 1–7).

Ding, S., Li, H., Su, C., Yu, J., & Jin, F. (2013). Evolutionary artificial neural networks: a review. *Artificial Intelligence Review*, *39*(3), 251–260.

Donq, L., Liang, & Wen, W., June. (1998). A multivalued bidirectional associative memory operating on a complex domain. *Neural Networks*, *11*(9), 1623–1635.

Egidi, G., & Nusbaum, H. C. (2012). Emotional language processing: how mood affects integration processes during discourse comprehension. *Brain and language*, *122*(3), 199–210.

Eidinger, E., Enbar, R., & Hassner, T. (2014, Dec). Age and gender estimation of un-filtered faces. *Information Forensics and Security, IEEE Transactions on*, *9*(12), 2170-2179.

Ekkekakis, P. (2013). *The measurement of affect, mood, and emotion: A guide for health-behavioral research*. Cambridge University Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4),

169–200.

Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, *98*, 45–60.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179–211.

Eom, T. D., Choi, C., & Lee, J. J. (2002). Generalized asymmetrical bidirectional associative memory for multiple association. *Applied mathematics and computation*, *127*(2), 221–233.

Esmaileyan, Z., & Marvi, H. (2014). Recognition of emotion in speech using variogram based features. *Malaysian Journal of Computer Science*, *27*(3).

Eysenck, M. W. (1998). *Psychology: An integrated approach*. U.K.: Longman.

Fazzolari, M., Alcala, R., Nojima, Y., Ishibuchi, H., & Herrera, F. (2013). A review of the application of multiobjective evolutionary fuzzy systems: Current status and further directions. *Fuzzy Systems, IEEE Transactions on*, *21*(1), 45–65.

Foss, D. J., & Harwood, D. A. (1975). Memory for sentences: Implications for human associative memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(1), 1–16.

Fritzke, B. (1995). A growing neural gas network learns topologies. *Advances in neural information processing systems*, *7*, 625–632.

Gandhi, V., Prasad, G., Coyle, D., Behera, L., & McGinnity, T. (2014, Feb). Quantum neural network-based eeg filtering for a brain;computer interface. *Neural Networks and Learning Systems, IEEE Transactions on*, *25*(2), 278-288.

Gislén, L., Peterson, C., & Söderberg, B. (1992). Rotor neurons: Basic formalism and dynamics. *Neural Computation*, *4*(5), 737–745.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological assessment*, *4*(1), 26.

Gray, J. A. (1987). The neuropsychology of emotion and personality.

Gregory, R. L. (1998). *The mind*. Oxford Univ. Press.

Griffiths, D. J. (1995). *Introduction to quantum mechanics*. University Science Books.

Gripon, V., & Berrou, C. (2011). Sparse neural networks with large learning diversity. *Neural Networks, IEEE Transactions on*, *22*(7), 1087–1096.

Hagiwara, M. (1990). Multidirectional associative memory. In *International joint conference on neural networks* (Vol. 1, pp. 3–6).

Han, M. J., Lin, C. H., & Song, K. T. (2013). Robotic emotional expression generation based on mood transition and personality model. *Cybernetics, IEEE Transactions on*, *43*(4), 1290–1303.

Hattori, M., & Hagiwara, M. (1995). Quick learning for multidirectional associative memories. In *Neural networks, 1995. proceedings., ieee international conference on* (Vol. 4, pp. 1949–1954).

Hattori, M., Hagiwara, M., & Nakagawa, M. (1994). Quick learning for bidirectional associative memory. *IEICE TRANSACTIONS on Information and Systems*, *77*(4), 385–392.

He, G., Chen, L., & Aihara, K. (2008). Associative memory with a controlled chaotic neural network. *Neurocomputing*, *71*(13), 2794–2805.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New york: Wiley & Sons.

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*(7), 1527–1554.

Hiolle, A., Canamero, L., Davila Ross, M., & Bard, K. A. (2012). Eliciting caregiving behavior in dyadic human-robot attachment-like interactions. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *2*(1), 3.

Hirabayashi, T., Takeuchi, D., Tamura, K., & Miyashita, Y. (2013). Functional microcircuit recruited during retrieval of object association memory in monkey perirhinal cortex. *Neuron*, *77*(1), 192–203.

Hirota, K., & Pedrycz, W. (1994). Or/and neuron in modeling fuzzy set connectives. *Fuzzy Systems, IEEE Transactions on*, *2*(2), 151–161.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Hooker, C. I., Bruce, L., Fisher, M., Verosky, S. C., Miyakawa, A., & Vinogradov, S. (2012). Neural activity during emotion recognition after combined cognitive plus social cognitive training in schizophrenia. *Schizophrenia research*, *139*(1), 53–59.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554-2558.

Householder, A. S. (1958). Unitary triangularization of a nonsymmetric matrix. *Journal of the ACM (JACM)*, *5*(4), 339–342.

Imada, A., & Araki, K. (1995). Genetic algorithm enlarges the capacity of associative memory. In *Icga* (pp. 413–420).

Islam, M. N., & Loo, C. K. (2014). Geometric feature-based facial emotion recognition using two-stage fuzzy reasoning model. In *Neural information processing* (pp. 344–351).

Itoh, K., Miwa, H., Takanobu, H., & Takanishi, A. (2005). Application of neural network to humanoid robots—development of co-associative memory model. *Neural*

networks, *18*(5), 666–673.

Jankowski, S., Lozowski, A., & Zurada, J. (1996, Nov). Complex-valued multistate neural associative memory. *Neural Networks, IEEE Transactions on*, *7*(6), 1491-1496.

Johns, M., & Silverman, B. G. (2001). How emotions and personality effect the utility of alternative decisions: a terrorist target selection case study. *Center for Human Modeling and Simulation*, 10.

Jordan, M. I. (1997). Serial order: A parallel distributed processing approach. *Advances in psychology*, *121*, 471–495.

Kang, H. (1994). Multilayer associative neural networks (mann's): storage capacity versus perfect recall. *Neural Networks, IEEE Transactions on*, *5*(5), 812–822.

Kar, S., Das, S., & Ghosh, P. K. (2014). Applications of neuro fuzzy systems: A brief review and future outline. *Applied Soft Computing*, *15*, 243–259.

Kitahara, M., & Kobayashi, M. (2014). Projection rule for rotor hopfield neural networks. *Neural Networks and Learning Systems, IEEE Transactions on*, *25*(7), 1298–1307.

Kobayashi, M. (2008). Pseudo-relaxation learning algorithm for complex-valued associative memory. *International journal of neural systems*, *18*(02), 147–156.

Kobayashi, M., & Yamazaki, H. (2005). Complex-valued multidirectional associative memory. *IEEJ Transactions on Electronics, Information and Systems*, *125*, 1290–1295.

Kohonen, T. (1972). Correlation matrix memories. *Computers, IEEE Transactions on*, *100*(4), 353–359.

Kosko, B. (1987a). Adaptive bidirectional associative memories. *Applied optics*, *26*(23), 4947–4960.

Kosko, B. (1987b). Constructing an associative memory. *Byte*, *12*(10), 137–144.

Kosko, B. (1991a). Fuzzy associative memories. In *Nasa, lyndon b. johnson space center, proceedings of the 2 nd joint technology workshop on neural networks and fuzzy logic,* (Vol. 1).

Kosko, B. (1991b). *Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence*. Prentice-Hall, Inc.

Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images. *Computer Science Department, University of Toronto, Tech. Rep*, *1*(4), 7.

Kubota, N., & Fukuda, T. (1999). Ecological model of virus-evolutionary genetic algorithm. *Fundamenta Informaticae*, *37*(1, 2), 51–70.

Kubota, N., & Toda, Y. (2012). Multimodal communication for human-friendly robot

partners in informationally structured space. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, *42*(6), 1142-1151.

Kumar, S., & Singh, M. P. (2010). Pattern recall analysis of the hopfield neural network with a genetic algorithm. *Computers and Mathematics with Applications*, *60*(4), 1049–1057.

LeDoux, J. E. (1994). Emotion, memory and the brain. *Scientific American*, *270*(6), 50–57.

Lee, A., & Kawahara, T. (2009). Recent development of open-source speech recognition engine Julius. In *Proceedings: Apsipa asc 2009* (pp. 131–137).

Lee, D. L. (2006). Improvements of complex-valued hopfield associative memory by using generalized projection rules. *Neural Networks, IEEE Transactions on*, *17*(5), 1341–1347.

Lee, R. S. (2004). A transient-chaotic autoassociative network (tcan) based on lee oscillators. *Neural Networks, IEEE Transactions on*, *15*(5), 1228–1243.

Leung, C. (1993). Encoding method for bidirectional associative memory using projection on convex sets. *Neural Networks, IEEE Transactions on*, *4*(5), 879–881.

Leung, F. H., Lam, H. K., Ling, S. H., & Tam, P. K. (2003). Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *Neural Networks, IEEE Transactions on*, *14*(1), 79–88.

Levy, H., & McGill, T. (1993, May). A feedforward artificial neural network based on quantum effect vector-matrix multipliers. *Neural Networks, IEEE Transactions on*, *4*(3), 427-433.

Lewis, P., Critchley, H., Smith, A., & Dolan, R. (2005). Brain mechanisms for mood congruent memory facilitation. *Neuroimage*, *25*(4), 1214–1223.

Li, P., Xiao, H., Shang, F., Tong, X., Li, X., & Cao, M. (2013). A hybrid quantum-inspired neural networks with sequence inputs. *Neurocomputing*, *117*, 81–90.

Liu, P. (1999). The fuzzy associative memory of max-min fuzzy neural network with threshold. *Fuzzy Sets and Systems*, *107*(2), 147 - 157.

Lu, T. C., Yu, G. R., & Juang, J. C. (2013, Aug). Quantum-based algorithm for optimizing artificial neural networks. *Neural Networks and Learning Systems, IEEE Transactions on*, *24*(8), 1266-1278.

Maass, W. (1997). Networks of spiking neurons: the third generation of neural network models. *Neural networks*, *10*(9), 1659–1671.

Manju, A., & Nigam, M. J. (2014). Applications of quantum inspired computational intelligence: a survey. *Artificial Intelligence Review*, *42*(1), 79–156.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality

across instruments and observers. *Journal of personality and social psychology*, *52*(1), 81.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115–133.

Mehrabian, A. (1980). *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies*. Oelgeschlager, Gunn & Hain Cambridge, MA.

Mehrabian, A. (1996). Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology*, *48*(2), 86–92.

Minsky, M. (1986). *The society of mind*. New York: Simon and Schuster.

Minsky, M., & Seymour, P. (1969). *Perceptrons.* MIT press.

Murray, B. D., & Kensinger, E. A. (2013). Age-related changes in associative memory for emotional and nonemotional integrative representations. *Psychology and aging*, *28*(4), 969.

Nakano, K. (1972). Associatron-a model of associative memory. *Systems, Man and Cybernetics, IEEE Transactions on*(3), 380–388.

Nakazawa, K., Quirk, M. C., Chitwood, R. A., Watanabe, M., Yeckel, M. F., Sun, L. D., . . . Tonegawa, S. (2002). Requirement for hippocampal ca3 nmda receptors in associative memory recall. *Science*, *297*(5579), 211–218.

Noest, A. J. (1988a). Associative memory in sparse phasor neural networks. *EPL (Europhysics Letters)*, *6*(5), 469.

Noest, A. J. (1988b). Discrete-state phasor neural networks. *Physical Review A*, *38*(4), 2196–2199.

Nuevo, J., Bergasa, L. M., & Jiménez, P. (2010). Rsmat: Robust simultaneous modeling and tracking. *Pattern Recognition Letters*, *31*(16), 2455–2463.

Oh, H., & Kothari, S. (1991). A new learning approach to enhance the storage capacity of the hopfield model. In *Neural networks, 1991. 1991 ieee international joint conference on* (pp. 2056–2062).

Oh, H., & Kothari, S. C. (1994). Adaptation of the relaxation method for learning in bidirectional associative memory. *Neural Networks, IEEE Transactions on*, *5*(4), 576–583.

Ortony, A. (1990). *The cognitive structure of emotions*. Cambridge university press.

Osana, Y., Hattori, M., & Hagiwara, M. (1996). Chaotic bidirectional associative memory. In *Neural networks, 1996., ieee international conference on* (Vol. 2, pp. 816–821).

Palm, G. (2013). Neural associative memories and sparse coding. *Neural Networks*, *37*(0), 165 - 171.

Personnaz, L., Guyon, I., & Dreyfus, G. (1985). Information storage and retrieval in spin-glass like neural networks. *Journal de Physique Lettres*, *46*(8), 359–365.

Personnaz, L., Guyon, I., & Dreyfus, G. (1986). Collective computational properties of neural networks: New learning mechanisms. *Physical Review A*, *34*(5), 4217.

Perus, M. (1996). Neuro-quantum parallelism in brain-mind and computers. *Informatica(Ljubljana)*, *20*(2), 173–184.

Perus, M., Bischof, H., & Hadzibeganovic, T. (2007). A natural quantum neural-like network. *NeuroQuantology*, *3*(3).

Pfeifer, R., & Scheier, C. (1999). *Understanding intelligence*. The MIT Press.

Pham, D., & Karaboga, D. (2012). *Intelligent optimisation techniques: genetic algorithms, tabu search, simulated annealing and neural networks*. Springer Science & Business Media.

Pierce, B. H., & Kensinger, E. A. (2011). Effects of emotion on associative recognition: valence and retention interval matter. *Emotion*, *11*(1), 139.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1996). *Numerical recipes in c* (Vol. 2). Cambridge university press Cambridge.

Purushothaman, G., & Karayiannis, N. (1997, May). Quantum neural networks (qnns): inherently fuzzy feedforward neural networks. *Neural Networks, IEEE Transactions on*, *8*(3), 679-693.

Ravaja, N., & Kätsyri, J. (2014). Suboptimal facial expression primes in textual media messages: Evidence for the affective congruency effect. *Computers in Human Behavior*, *40*, 64–77.

Read, J., Bifet, A., Pfahringer, B., & Holmes, G. (2012). Batch-incremental versus instance-incremental learning in dynamic and evolving data. In *Advances in intelligent data analysis xi* (pp. 313–323). Springer.

Reisberg, D. E., & Hertel, P. E. (2004). *Memory and emotion.* Oxford University Press.

Rigatos, G. G. (2010). Stochastic processes and neuronal modelling: Quantum harmonic oscillator dynamics in neural structures. *Neural processing letters*, *32*(2), 167–199.

Rigatos, G. G., & Tzafestas, S. G. (2002). Parallelization of a fuzzy control algorithm using quantum computation. *Fuzzy Systems, IEEE Transactions on*, *10*(4), 451–460.

Rigatos, G. G., & Tzafestas, S. G. (2006). Quantum learning for neural associative memories. *Fuzzy Sets and Systems*, *157*(13), 1797–1813.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and

organization in the brain. *Psychological review*, *65*(6), 386.

Rottenberg, J., & Gross, J. J. (2003). When emotion goes wrong: Realizing the promise of affective science. *Clinical Psychology: Science and Practice*, *10*(2), 227–232.

Rumbell, T., Barnden, J., Denham, S., & Wennekers, T. (2012). Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Autonomous Agents and Multi-Agent Systems*, *25*(1), 1–45.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323–533.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of personality and social psychology*, *76*(5), 805.

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, *48*(5), 1290.

Sakawa, M. (2012). *Genetic algorithms and fuzzy multiobjective optimization* (Vol. 14). Springer Science & Business Media.

Salavati, A., Kumar, K., & Shokrollahi, A. (2014). Nonbinary associative memory with exponential pattern retrieval capacity and iterative learning. *Neural Networks and Learning Systems, IEEE Transactions on*, *25*(3), 557-570.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117.

Schneider, M., & Adamy, J. (2014). Towards modelling affect and emotions in autonomous agents with recurrent fuzzy systems. In *Systems, man and cybernetics (smc), 2014 ieee international conference on* (pp. 31–38).

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, *20*(1), 11.

Shi, H., Zhao, Y., & Zhuang, X. (1998). A general model for bidirectional associative memories. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *28*(4), 511–519.

Shimizu, Y., & Osana, Y. (2010). Chaotic complex-valued multidirectional associative memory. *IASTED Artificial Intelligence and Applications*.

Smith, S. M., & Petty, R. E. (1995). Personality moderators of mood congruency ef-

fects on cognition: the role of self-esteem and negative mood regulation. *Journal of personality and social psychology*, *68*(6), 1092.

Sperber, D., & Wilson, D. (1995). *Relevance – communication and cognition*. Oxford Univ. Press.

Squire, L. R., & Zola Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*(5026), 1380–1386.

Storkey, A. J., & Valabregue, R. (1999). The basins of attraction of a new hopfield learning rule. *Neural Networks*, *12*(6), 869–876.

Sussner, P., & Valle, M. E. (2006). Implicative fuzzy associative memories. *Fuzzy Systems, IEEE Transactions on*, *14*(6), 793–807.

Toris, R., Kent, D., & Chernova, S. (2014). The robot management system: A framework for conducting human-robot interaction studies through crowdsourcing. *Journal of Human-Robot Interaction*, *3*(2), 25–49.

Tsai, J. T., Chou, J. H., & Liu, T. K. (2006). Tuning the structure and parameters of a neural network by using hybrid taguchi-genetic algorithm. *Neural Networks, IEEE Transactions on*, *17*(1), 69–80.

Tzafestas, S., & Rigatos, G. (2000). Stability analysis of an adaptive fuzzy control system using petri nets and learning automata. *Mathematics and computers in Simulation*, *51*(3), 315–339.

Valle, M. E., & Sussner, P. (2008). A general framework for fuzzy morphological associative memories. *Fuzzy Sets and Systems*, *159*(7), 747 - 768.

Valverde Ibanez, R., Keysermann, M. U., & Vargas, P. (2014). Emotional memories in autonomous robots. In *Robot and human interactive communication, 2014 ro-man: The 23rd ieee international symposium on* (pp. 405–410).

Wang, Y. F., Cruz Jr, J. B., & Mulligan Jr, J. (1991). Guaranteed recall of all training pairs for bidirectional associative memory. *Neural Networks, IEEE Transactions on*, *2*(6), 559–567.

Wang, Y. F., Cruz Jr, J. B., & Mulligan Jr, J. H. (1990). Two coding strategies for bidirectional associative memory. *Neural Networks, IEEE Transactions on*, *1*(1), 81–92.

Wang, Z. O. (1996). A bidirectional associative memory based on optimal linear associative memory. *Computers, IEEE Transactions on*, *45*(10), 1171–1179.

Wu, Y., & Pados, D. A. (2000). A feedforward bidirectional associative memory. *Neural Networks, IEEE Transactions on*, *11*(4), 859–866.

Xu, Z. B., Leung, Y., & He, X. W. (1994). Asymmetric bidirectional associative memories. *Systems, Man and Cybernetics, IEEE Transactions on*, *24*(10), 1558–1564.

Yao, Z., Gripon, V., & Rabbat, M. G. (2013). A massively parallel associative memory based on sparse neural networks. *arXiv preprint arXiv:1303.7032*.

Yi, W., Zhi liang, W., & Wei, W. (2014). Research on associative memory models of emotional robots. *Advances in Mechanical Engineering*, *6*, 208153.

Yorita, A., & Kubota, N. (2011). Cognitive development in partner robots for information support to elderly people. *Autonomous Mental Development, IEEE Transactions on*, *3*(1), 64-73.

Yoshida, A., & Osana, Y. (2011). Chaotic complex-valued multidirectional associative memory with variable scaling factor. In *Artificial neural networks and machine learning–icann 2011* (pp. 266–274). Springer.

Zhang, G. (2011). Quantum-inspired evolutionary algorithms: a survey and empirical study. *Journal of Heuristics*, *17*(3), 303–351.

Zheng, K., Glas, D. F., Kanda, T., Ishiguro, H., & Hagita, N. (2013). Designing and implementing a human–robot team for social interactions. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, *43*(4), 843–859.

Zheng, L., & Zhou, C. (2008). Quantum associative memory that processes nonorthogonal patterns. In *Intelligent system and knowledge engineering, 2008. iske 2008. 3rd international conference on* (Vol. 1, pp. 578–583).

Zhuang, X., Huang, Y., & Chen, S. S. (1993). Better learning for bidirectional associative memory. *Neural networks*, *6*(8), 1131–1146.

APPENDIX B

PUBLICATIONS AND PAPERS PRESENTED

B.1 Journal Publication

[1] Naoki Masuyama, Chu Kiong Loo and Naoyuki Kubota,"Quantum-Inspired Bidirectional Associative Memory for Human-Robot Communication," International Journal of Humanoid Robotics, Vol. 11, No. 2, 1450006 (22 pages), 2014.

[2] Naoki Masuyama, Md. Nazrul Islam, Manjeevan Seera and Chu Kiong Loo, "Application of Emotion Affected Associative Memory based on Mood Congruency Effects for a Humanoid," Neural Computing and Applications, pp. 1-16, 2015.

[3] Md. Nazrul Islam, Naoki Masuyama, Manjeevan Seera and Chu Kiong Loo, "Simultaneous Facial Features Tracking and Emotion Recognition," Neurocomputing, 2015, *Accepted*.

B.2 Conference Publication

[1] Naoki Masuyama, Chee Seng Chan, Naoyuki Kubota and Jinseok Woo, "Computational Intelligence for Human Interactive Communication of Robot Partners," 12th Pacific Rim International Conference on Trends in Artificial Intelligence (PRICAI), pp. 771-776, 2012.

[2] Naoki Masuyama, Chu Kiong Loo and Naoyuki Kubota, "Quantum Mechanics Inspired Bidirectional Associative Memory for Human Robot Interaction," Advanced Computational Intelligence and Intelligent Informatics (IWACIII), 3rd International Workshop on, GS1-5, 2013.

[3] Naoki Masuyama, Chu Kiong Loo and Naoyuki Kubota, "Human-Robot Interac-

tion System with Quantum-Inspired Bidirectional Associative Memory," Robot, Vision and Signal Processing, Second International Conference on, pp.66-71, 2013.

[4] Naoki Masuyama and Chu Kiong Loo, "Quantum-Inspired Multidirectional Associative Memory for Human-Robot Interaction System," *IEEE* World Congress on Computational Intelligence (WCCI), pp. 1757-1764, 2014.

[5] Naoki Masuyama, Md. Nazrul Islam and Chu Kiong Loo, "Affective Communication Robot Partners using Associative Memory with Mood Congruency Effects," *IEEE* Symposium Series on Computational Intelligence (SSCI), #15057, 2014.

[6] Naoki Masuyama and Chu Kiong Loo, "Quantum-Inspired Complex-Valued Multidirectional Associative Memory," Neural Networks (IJCNN), *IEEE* International Joint Conference on, pp. 1-8, 2015.

[7] Naoki Masuyama and Chu Kiong Loo, "Robotic Emotional Model with Personality Factors base on Pleasant-Arousal Scaling Model," Robot and Human Interactive Communication (RO-MAN), *IEEE* 24th International Symposium on, 2015. (Best Paper Award Finalist)

[8] Zeeshan Rasool, Naoki Masuyama, Md. Nazrul Islam and Chu Kiong Loo, "Empathic interaction using the computational emotion model," *IEEE* Symposium Series on Computational Intelligence (SSCI), pp. 109–116, 2015.