INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
ATTENTION MODULATED
ASSOCIATIVE COMPUTING AND
CONTENT-ASSOCIATIVE SEARCH IN IMAGE ARCHIVE

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ELECTRICAL ENGINEERING

AUGUST 1995

By

Muhammad Javed Iqbal Khan

Dissertation Committee:

David Y.Y. Yun, Chairperson
Robert Cole
Shu Lin
W.W. Peterson
Tong-Jye Lee
In memory of my beloved father Fazlur Rahman Khan, who would be the happiest person to see me graduating, and to my beloved mother Jahan Ara Begum, who inspired me to seek.
ACKNOWLEDGEMENTS

I would like to thank my advisor David Y. Y. Yun for his encouragement and support throughout this research. He gave me the freedom to explore and roam in the realm of ideas and just the right amount of encouragement, advice and inspiration to guide me towards completion. I am also grateful to the members of my committee, W. Wesley Peterson, Shu Lin, Robert Cole and Tong Jye Lee, for their patient and careful evaluation of the dissertation. I would also like to mention my teachers Dr. Mahbubur Rahman, Dr. Tim Brown and Dr. Bharat Kinariwala whose inspiration and teaching prepared me for this work.

I have also benefitted from the countless discussions with members of our lab who have provided me encouragement and feedback, especially, Tse Cheng, Hui Liu, Xiangming Feng, Yiqun Ge and Zong Ling. I also appreciate the help from Glenn Yamashita, Linda Hamada, Earl Bethke, Silvia Chan, Joann Skudlarick, and Berney Trodd.

I acknowledge the support of the East West Center (EWC), and the United States Information Agency (USIA) for the scholarship that supported my Ph.D. program. Also, a part of this research has been funded by Advanced Communications Technology Satellite (ACTS) and Supercomputing in Remote, Cooperative Medical Triage Support and Radiation Treatment Planning Project, Advanced Research Projects Agency (ARPA), Department of Defence, under contract DABT 63-93-C-0056.

I am thankful to Amma, Taskina Apa and Asif for the moral and emotional support and good wishes that made it possible for me to continue through all these years so far away from home. The most special thanks goes to Farial Sabrina for unfailing love and tireless help in proofreading this dissertation for a ’grammatically challenged’ husband who was mostly married to machines during these years in Hawaii.
ABSTRACT

Principal Result (AAM): This dissertation presents an Artificial Associative Memory (AAM) which can learn pattern-associations and retrieve them with changeable attention. The attention refers to the fact that the user can specify any subset of the elements in the example query pattern and expect the memory to confine its match within the specified field of attention. Existing AAMs lack such flexibility and perform match based on unchangeable unary attention over all the elements of the query pattern.

Corollary 1 (dynamic attention): Unlike the existing AAMs, in this new associative memory user can dynamically vary the field of attention during each query. Within the specified field of attention, it performs statistically robust matching.

Corollary 2 (robustness): This attentive AAM can retrieve information from very small cue. Conventional AAMs require the effective cue in the query pattern to be large (approximately 50% of the query frame) and statistically dominant.

Corollary 3 (MNC feedback): It also has the unique ability to generate a feedback (called MNC) on the quality of match corresponding to the retrieved pattern.

Approach: In contrast to the conventional AAM, the proposed approach is based on (i) a new representation of information, which includes confidence, instead of only measurement (ii) mapping of measurements on the surface of a hypersphere, instead of on a linear interval, and (iii) a new rule of synaptic efficacy based on trigonometric averaging, rather than a statistical sum.

Application: As an application of this attentive memory an automatic direct content-based search mechanism has been developed for querying into image database (IDB). For an IDB of p images of n pixels, it improves the cost of direct content based search from O(np) to O(n.logp). It can also use a fuzzy logic based query and inference formalism. This mechanism avoids the problem of representation insufficiency and inaccuracy of intermediate model encoding and can help in supplementing existing content based IDB query mechanisms.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** ................................................................. v
**ABSTRACT** .............................................................................. vi
**TABLE OF CONTENTS** ............................................................... vii
**LIST OF TABLES** .................................................................... ix
**LIST OF FIGURES** ................................................................. x
**LIST OF ABBREVIATIONS** ....................................................... xii

**CHAPTER 1** **INTRODUCTION** .................................................. 1
  1.1 Introduction ......................................................................... 1
  1.2 Summary and Contribution .................................................. 11
  1.3 Dissertation Overview ....................................................... 13

**CHAPTER 2** **ARTIFICIAL ASSOCIATIVE COMPUTING** .................. 15
  2.1 Introduction ......................................................................... 15
  2.2 Development of Artificial Associative Computing ................. 15
  2.3 Learning and Retrieval In Artificial Associative Computing ..... 28
  2.4 Retrieval with Changeable Attention (RCA) ......................... 34
  2.5 Role of Synaptic Transmission Rule .................................... 49
  2.6 Solution Approach ............................................................ 53
  2.7 Chapter Summary .............................................................. 54

**CHAPTER 3** **CONTENT BASED SEARCH IN IMAGE ARCHIVE** ........ 55
  3.1 Introduction ......................................................................... 55
  3.2 Image Information ............................................................. 56
  3.3 Review of Current Database Systems ................................... 60
  3.4 Difficulty of Coherent Modeling ......................................... 67
  3.5 Holographic Approach ....................................................... 69
  3.6 Other Mechanisms for Direct Search .................................... 71
  3.7 Application of MHAC in Direct Image Database Search .......... 79
  3.8 Chapter Summary .............................................................. 80

**CHAPTER 4** **THEORY OF HOLOGRAPHIC COMPUTING** .............. 81
  4.1 Introduction ......................................................................... 81
  4.2 Information with Modifiable Significance ............................ 82
  4.3 Computational Representation ............................................ 84
  4.4 Proposed Holographic Computational Model ....................... 88
  4.5 Confirmation of Basic Associative Recovery (type-U/BM) ....... 90
  4.6 Confirmation of Expectation-1 (type-A/BM) ......................... 94
  4.7 Confirmation of Expectation-2 (MNC) .................................. 96
  4.8 Significance of Holographic Computation ........................... 97
  4.9 Other Learning Models ...................................................... 100
  4.10 Chapter Summary ............................................................ 101

**CHAPTER 5** **CHARACTERISTICS OF HOLOGRAPHIC COMPUTING** .... 102
  5.1 Introduction ......................................................................... 102
  5.2 Analysis of Characteristics ................................................ 103
  5.3 Simulation Results on MHAC Characteristics ....................... 115

vii
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4 Operational Range Space</td>
<td>133</td>
</tr>
<tr>
<td>5.5 Chapter Summary</td>
<td>149</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>150</td>
</tr>
<tr>
<td>6.2 System Overview</td>
<td>151</td>
</tr>
<tr>
<td>6.3 Encoding Process</td>
<td>153</td>
</tr>
<tr>
<td>6.4 Query Process</td>
<td>164</td>
</tr>
<tr>
<td>6.5 Search Complexities</td>
<td>169</td>
</tr>
<tr>
<td>6.6 Chapter Summary</td>
<td>170</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>171</td>
</tr>
<tr>
<td>6.2 Compound Query Language</td>
<td>173</td>
</tr>
<tr>
<td>6.3 Inference Mechanism for Compound Queries</td>
<td>178</td>
</tr>
<tr>
<td>6.4 Modeling of Subjective Reasoning</td>
<td>181</td>
</tr>
<tr>
<td>6.5 Examples of Inferences</td>
<td>184</td>
</tr>
<tr>
<td>6.6 Chapter Summary</td>
<td>188</td>
</tr>
<tr>
<td>8.1 Summary: Contributions and Limitations</td>
<td>189</td>
</tr>
<tr>
<td>8.2 Future Research Directions</td>
<td>192</td>
</tr>
<tr>
<td>APPENDIX-A SURVEY OF ARTIFICIAL NEURAL NETWORKS</td>
<td>196</td>
</tr>
<tr>
<td>APPENDIX-B SURVEY OF IDB SEARCH TECHNIQUES</td>
<td>198</td>
</tr>
<tr>
<td>APPENDIX-C AUXILIARY DERIVATIONS</td>
<td>200</td>
</tr>
<tr>
<td>APPENDIX-D SCENE INTERPRETATION</td>
<td>202</td>
</tr>
<tr>
<td>APPENDIX-E QUERY IN SEQ001 ARCHIVE</td>
<td>205</td>
</tr>
<tr>
<td>APPENDIX-F QUERY IN MEDIA ARCHIVE</td>
<td>211</td>
</tr>
<tr>
<td>APPENDIX-G PERFORMANCE PROJECTIONS</td>
<td>215</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>220</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Complexities of Procedural Search Methods</td>
<td>76</td>
</tr>
<tr>
<td>4.1 Expectations</td>
<td>83</td>
</tr>
<tr>
<td>5.1 Operational Range Space</td>
<td>147</td>
</tr>
<tr>
<td>5.2 Typical Assymetry</td>
<td>148</td>
</tr>
<tr>
<td>5.3 Typical Memory Loading</td>
<td>148</td>
</tr>
<tr>
<td>6.1 A RWLP Coding Example</td>
<td>158</td>
</tr>
<tr>
<td>6.2 Elementary Object Table</td>
<td>165</td>
</tr>
<tr>
<td>6.3 Retrieval Performance</td>
<td>169</td>
</tr>
<tr>
<td>A.1 Neural Network Models (1)</td>
<td>196</td>
</tr>
<tr>
<td>B.1 Image Database Systems</td>
<td>198</td>
</tr>
<tr>
<td>I.1 Trademark Image Archive</td>
<td>216</td>
</tr>
<tr>
<td>I.2 Server Performance for Trademark Image Archive</td>
<td>216</td>
</tr>
<tr>
<td>I.3 Planetary Navigation Archive</td>
<td>217</td>
</tr>
<tr>
<td>I.4 Server Performance for PN Archive</td>
<td>217</td>
</tr>
<tr>
<td>I.5 A Typical Image Archive</td>
<td>218</td>
</tr>
<tr>
<td>I.6 Server Performance of Single Site</td>
<td>219</td>
</tr>
<tr>
<td>I.7 Server Performance for Distributed Image Archive</td>
<td>219</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Robustness of Hopfield Models</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Attention Modulated Retrieval</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Contribution</td>
<td>12</td>
</tr>
<tr>
<td>1.1 Logic Operations in MP-neuron</td>
<td>17</td>
</tr>
<tr>
<td>1.2 Basic Neuron Cell</td>
<td>29</td>
</tr>
<tr>
<td>1.3 An AM Model</td>
<td>31</td>
</tr>
<tr>
<td>1.4 Robustness of Basic Neuron</td>
<td>51</td>
</tr>
<tr>
<td>3.1 Physical Image Information</td>
<td>58</td>
</tr>
<tr>
<td>3.2 Model Based Approach</td>
<td>60</td>
</tr>
<tr>
<td>3.3 User Profile Based Approach</td>
<td>66</td>
</tr>
<tr>
<td>3.4 Direct Content Search Approach</td>
<td>69</td>
</tr>
<tr>
<td>3.5 A Model of Conventional Memory</td>
<td>72</td>
</tr>
<tr>
<td>4.1 Points on Hyperspherical Surface</td>
<td>85</td>
</tr>
<tr>
<td>5.1 Unit Vectors</td>
<td>105</td>
</tr>
<tr>
<td>5.2 Geometry of Phase Error</td>
<td>107</td>
</tr>
<tr>
<td>5.3 Convergence of Recall Accuracy</td>
<td>117</td>
</tr>
<tr>
<td>5.4 Convergence of Recalled MNC</td>
<td>118</td>
</tr>
<tr>
<td>5.5 Recall Accuracy and Number of Patterns</td>
<td>119</td>
</tr>
<tr>
<td>5.6 Recall Accuracy and Load Factor</td>
<td>120</td>
</tr>
<tr>
<td>5.7 Effect of Attention Strength</td>
<td>122</td>
</tr>
<tr>
<td>5.8 Meta-Knowledge and Noise</td>
<td>124</td>
</tr>
<tr>
<td>5.9 Meta-Knowledge and Attention</td>
<td>125</td>
</tr>
<tr>
<td>5.10 SNR for Stored/Unstored Patterns</td>
<td>126</td>
</tr>
<tr>
<td>5.11 MNC for Stored/Unstored Patterns</td>
<td>127</td>
</tr>
<tr>
<td>5.12 Focus and SNR</td>
<td>128</td>
</tr>
<tr>
<td>5.13 Focus and MNC</td>
<td>129</td>
</tr>
<tr>
<td>5.14 Effect of Representation Dimension (I)</td>
<td>130</td>
</tr>
<tr>
<td>5.15 Effect of Representation Dimension (II)</td>
<td>131</td>
</tr>
<tr>
<td>5.16 SNR and Dimension</td>
<td>132</td>
</tr>
<tr>
<td>5.17 Crosstalk and Dimension</td>
<td>132</td>
</tr>
<tr>
<td>5.18 ORS (I)</td>
<td>136</td>
</tr>
<tr>
<td>5.19 ORS (II)</td>
<td>137</td>
</tr>
<tr>
<td>5.20 ORS (III)</td>
<td>138</td>
</tr>
<tr>
<td>5.21 ORS (IV)</td>
<td>139</td>
</tr>
<tr>
<td>5.22 ORS (V)</td>
<td>140</td>
</tr>
<tr>
<td>5.23 ORS (VI)</td>
<td>141</td>
</tr>
<tr>
<td>5.24 ORS (VII)</td>
<td>142</td>
</tr>
<tr>
<td>5.25 ORS (VIII)</td>
<td>144</td>
</tr>
<tr>
<td>5.26 ORS (IX)</td>
<td>144</td>
</tr>
<tr>
<td>5.27 ORS (X)</td>
<td>145</td>
</tr>
<tr>
<td>5.28 ORS (XI)</td>
<td>146</td>
</tr>
<tr>
<td>5.29 ORS (XII)</td>
<td>146</td>
</tr>
<tr>
<td>6.1 CAIDSS Overview</td>
<td>152</td>
</tr>
<tr>
<td>6.2 Spiral Spreading Function</td>
<td>155</td>
</tr>
<tr>
<td>6.3 Spiral mapping with k=7.0</td>
<td>156</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAM</td>
<td>Artificial Associative Memory</td>
</tr>
<tr>
<td>AES</td>
<td>Associative Encoding Subsystem</td>
</tr>
<tr>
<td>AM</td>
<td>Associative Memory</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>AQS</td>
<td>Associative Query Subsystem</td>
</tr>
<tr>
<td>CAIDSS</td>
<td>Content-Associative Image Database Search System</td>
</tr>
<tr>
<td>CQF</td>
<td>Compound Query Formalism</td>
</tr>
<tr>
<td>CQL</td>
<td>Compound Query Language</td>
</tr>
<tr>
<td>DDC</td>
<td>Dimension Dispersion Component</td>
</tr>
<tr>
<td>LMS</td>
<td>Least Mean Square</td>
</tr>
<tr>
<td>MCN</td>
<td>Multidimensional Complex Numeric</td>
</tr>
<tr>
<td>MHAC</td>
<td>Multidimensional Holographic Associative Memory</td>
</tr>
<tr>
<td>MNC</td>
<td>Mean Normalized Confidence</td>
</tr>
<tr>
<td>ORS</td>
<td>Operational Range Space</td>
</tr>
<tr>
<td>RCA</td>
<td>Retrieval with Changeable Attention</td>
</tr>
<tr>
<td>RLP</td>
<td>Response Label Pattern</td>
</tr>
<tr>
<td>RMWC</td>
<td>Reverse Modulo Weighted Code</td>
</tr>
<tr>
<td>SAS</td>
<td>Storage and Access Subsystem</td>
</tr>
<tr>
<td>SST</td>
<td>Search Summary Table</td>
</tr>
</tbody>
</table>
"For example, there is '111'; that's a 'Nelson' because that's what you call a score of 111 in cricket. I see a '312' in there: area code of Chicago. The '1745' I remember as 39, because Ben Franklin was 39 in 1745..."

On associations:
Rajan Mahadevan,
Prodigious numerical memorists who among other things recited 31,811 digits of pi from memory.
CHAPTER 1

INTRODUCTION

1.1 Introduction

Associative computing has long been thought of as an integral part of human cognition. Aristotle (387-322 B.C.) observed that mental items (ideas, perceptions, sensations, or feeling) are connected in memory if they occur simultaneously, occur in close succession, are similar, or are contrary. The modern advancements in artificial associative memory (AAM) research came after the rapid development of neurophysiology and computational science, beginning in the late nineteenth century. Since then, the observations on the behavioral and physical structure of natural memory obtained from psychological and neurobiological sciences always fascinated and consequently propelled the researchers in this field.

Flexible learning with rigid retrieval: One of the first such observations that fascinated researchers of the time [Sher06, Jaco93] was the physically distributed construct of biological memories (Ramon y Cajal 1888-1911, Sherrington 1885-1951) In 1943, McCulloch and Pitts’ invention of formal neuron for the first time provided a formal architecture for a brain like distributed processing of information. It was extraordinary, because reinforced by the theory of symbolic logic (Russell & Whitehead, 1910, 1912, 1913) it promised universal computability and artificial realizability of almost unlimited complex systems [McPi43, RuWh10].

1 Among the many pioneers of the time, even Sigmund Freud attempted to construct a distributed model of "ego" with interconnected neurons, complete with a sketch. He wrote "Let us imagine the ego as a network of connected neurons..." [Jaco93,p28].
The optimism it sparked has been enormous. The principal focus of the research immediately shifted to the pursuit of the other fascinating aspect of biological memory: it’s adaptability or ability to learn. Beginning with Rosenblatt’s Perceptron learning, till today researchers have focused, and in many ways confined themselves, to perfecting the learning behavior of artificial associative memory systems.

During these years, increasingly more intricate and complex properties of learning phenomena have been pursued in great depth. Versatility (how arbitrary complex associations can be learned), efficiency (how more patterns can be learned), learnability of causality and temporal relations (Grossberg 1967, Klopf 1987), self-organization (Kohonen 1987, Oja 1982), autonomous unsupervised adaptation (Grossberg 1976, Carpenter & Grossberg, 1987) are just few examples of the intricacies through which research in artificial learning matured [Gros67, Klop87, Koho87, Oja82, Gros76, CGMR92]. Surprisingly, during this period of vigorous emphasis on the learning aspect of associative memory systems, very few attempts had been made to examine it’s recollection aspect, other than assuming a very simple model of retrieval.

Almost all the proposed learning models since McCulloch and Pitts have been constructed on the assumption of a simple and restricted retrieval scenario², where the sample of the content that is used during query is a close replica of the target. However more complex and versatile retrieval formalism is not only conceivable but also seems to be an integral part of natural associative memories.

**Limitation of the current models:** Interestingly, experiments performed by previous researchers on current models of AAM contain empirical evidence of severe retrieval inadequacy of the existing models.

---

² Consequentially, most of these learning methods break down when the test of learning is based on the generalized retrieval scenario.
The simulation result shown in Fig-1.1 is one such demonstration\(^3\). It plots the probability of correct recall as compared to the number of bits in error for several learning algorithms. It may be noted that as the number of error bits approach approximately just 40%, the probability of correct recall vanishes sharply irrespective of the sophistication of the learning algorithm. Experiments published by many other researchers also contain the fingerprint of such pure statistical nature of neural network convergence\(^4\) [Hopf82, MiFa90] (although, it apparently always managed to escape attention). Simple investigations on other artificial neural networks (ANN) also reveal the same limitation. As soon as the amount of

\(^{3}\) The demonstration is inadvertent. The experiment is originally intended to compare the performance of several high performance neural models based on higher order encoding including Hopfield model. HNN, CHNNet etc. are various models compared. The simulation demonstrates that high-order asymmetric learning algorithms dramatically enlarge the basin of attraction and increase probability of correct recall.

\(^{4}\) Hopfield in his famous 1982 paper reported "...memories too close to each other are confused and tend to merge... For \(N=100\) [number of elements in a pattern], a pair of random memories should be separated by [at least] 50±5 Hamming units".
"incorrect" part becomes statistically dominant over the "correct" part (close to around 50%) in the entire pattern element space of input information the network ceases to operate. The reason is the indiscriminate statistical matching used by the networks.

This is a profound limitation. Any effective memory (like the biological memories) must be able to retrieve information with a far smaller cue. However this limitation has never been questioned, and the myth about the robustness of neural networks continued without any clarification. Consequently, the principal motivation behind this dissertation was:

Motivation: to investigate the limitation of conventional associative memories due to hard statistical nature of internal matching and to extend its retrieval capability so that cues can be used based on their cognitive importance, irrespective of their statistical dominance.

As it will be shown, the failure of conventional networks to work with statistically weak sensory stimulus cue is a manifestation of a deeper limitation, namely; their inability to perform attention modulated retrieval.

Fig-1.2 illustrates the problem. Let an associative memory be allowed to learn the image frames A, B and C. If during the retrieval, template-D is used as a sensory input, then it is natural to expect that the memory should retrieve frame-A based on the roller, that appears to be the most cognitively significant index in the template. However it can be demonstrated that most of the conventional AAMs instead will retrieve frame-C as the closest match (indeed B and C are closer to D than A; both in LMS, and dot-product sense). The reason for such an unexpected result is the statistical weakness of the cognitively important roller pixels compared to the statistical strength of cognitively less important background pixels. In contrast, a natural memory seems to be immune to such statistical weakness and can retrieve information based on localized focus on cognitively important zones.
A more intriguing aspect of natural associative memory is that it can change the distribution of attention over element space dynamically during query. Consider template-E. Now, there are two objects of focus, and two possible answers. If desired, a natural memory can shift its attention to any other object (for example Plant) in the template and retrieve totally different match (frame-B) apparently without any significant internal reorganization. In contrast, a conventional AAM lacks such flexibility. For a given state of learning, it acts as a deterministic machine where each initial state flows into a pre-determined single attractor. Conventional AAMs have no mechanism to accommodate change in the precoded distribution of attention over its element space.
Clearly, the inability to retrieve with statistically weak cue is just a manifestation of the deeper problem of attention deficiency of conventional AAM. A conventional AAM requires any effective cue to be statistically significant compared to the overall pattern size because it cannot focus in any subspace of the template pattern as a basis for match.

The principal objective of this dissertation is to develop an associative computing mechanism that can support intelligent retrieval (even when the cue is statistically weak) based on dynamically modulated selective attention on the sensory stimulus field.

The hypothesis that was never challenged: It is surprising that, despite the invention of so many artificial neural network (ANN) models over the enormously productive fifty years following McCulloch and Pitts, the rule specifying the transformation of signal of a neuron has remained unaltered. It is so prevalent that hardly anyone bothered to name it. In an invited paper, Gail Carpenter has pointed it out in the last World Congress on Neural Network [Carp94]:

“What, if anything, do all the neural networks of the past fifty years have in common? ... from McCulloch and Pitts neuron to the many biological and engineering models at this year’s conferences, do have one element in common, namely the rule setting the net signal from a source node to a target node equal to a path signal times a synaptic weight. This product rule of synaptic transmission is in such universal use that it is almost treated as a nameless fact rather than a hypothesis, although neurophysiology has so far neither confirmed nor refuted the rule.”

The easy acceptance of the product rule of synaptic efficacy is probably partly due to its computational convenience and partly due to historic inertia. This dissertation demon-

5 Why in the first place, McCulloch and Pitts decided to use a function (transfer function) of the form of scalar weighted sum as a basic building block of neural computation? Possibly, because of its resemblance to the conjunctive normal forms of bi-variate symbolic logic.
strates that the reason behind the inability of conventional neural computing to perform attention modulated retrieval in fact, lies right at the non-optimality of this rule of synaptic efficacy.

Intuitively, it is difficult to imagine how a summation process can be robust against loss of its components, specially when their numbers are not so large. Nevertheless, it has been accepted by almost all researchers in this field without question as a foundation of artificial neuro computing. Almost all of the current neural memories are based on this elementary neuron model. They vary mostly in the way these neurons are interconnected. Consequently, almost all of today’s ANN models bear this fundamental limitation.

**Holographic attentive computing:** This dissertation demonstrates a new associative computing paradigm which overcomes the above limitations of the conventional associative computing. The solution is based on two fundamental features. First, it embraces a new representation of information that includes the notion of attention at the very core. Secondly, it employs an analogous but computationally different synaptic transmission rule as a basic building block for computation. In this dissertation, this memory is referred as holographic because of the adopted complex numeric representation and mode of computation.

The memory retains both the desirable features of conventional neural computing, adaptability as well as cellular distributed architecture, and at the same time it provides the capability to perform attention modulated search. As a natural part of this new representation it also has another unique property (compared to existing neural computing), it not only accepts but also communicates back a *significance feedback* associated with the information retrieved from the memory.

**Applications of AM models:** Today’s associative memory systems based on artificial neural computing can be considered as memory systems that offer ample flexibility in
learning of information but for the purpose of a rigid retrieval. Consequently, most of the successes in their application have been gained mainly in the area of adaptive classification or filtering\(^6\) [WiLe90, Carp89]. Adaptive applications capitalize on their ability to learn. But because of the rigid nature of their recollection, hardly any application flourished which can take advantage of its associative memory characteristics. This dissertation shows how the added retrieval flexibility of the proposed holographic memory may change this scenario. It is demonstrated that the flexibility of associative computing, coupled with the ability to perform attention-modulated retrieval, can be used to develop an effective *content based search and retrieval* mechanism for image information, which itself is an outstanding and difficult problem in intelligent image management.

**Difficulty of current image database technology:** The transcendental nature of image information makes it very difficult to construct a content based search strategy for image databases through conventional symbolic methods. Conventional technology requires strict structure. Its representation is almost tabular. Only recently it is shifting towards object oriented approaches, which allows at least some flexibility. However when it comes to the management of image information, even object oriented approaches fall far short of the flexibility required to cope with its formlessness.

A second level of difficulty is associated with the inference mechanism used by the current image database technology. Almost all existing approaches use strict bi-valued symbolic logic. However it is very difficult to reason with image objects using absolute certainty of bi-valued symbolic logic.

\(^6\) The potential application areas of ANN are frequently classified in literature into three major categories, associative memory, adaptive filtering, and optimization.
**Amorphousness of image information:** The current approach of image database management is to model the 'meaning' of an image as intermediate objects and to apply conventional database technology to perform query into this structured annotation of content [Niba93, ChYD88, JoCa88, TuPr91, ChFu80, YaSa93]. For almost all of these proposed methods the translation of 'meaning' of an image into a structured model itself has proved to be a formidable task. Neither is there a coherent finite formalism (or language) that can contain the possible ranges of 'meaning', nor is there any method that can consistently extract such 'meaning', because of the inherent amorphousness of image information. As a result, almost all such intermediate model-based methods currently face the difficulty of obtaining consistent and coherent models.

This research explores an alternate approach. It demonstrates that holographic associative computing with the capability to perform attention modulated query can bypass the need for structured intermediate modeling of meaning.

**Contemporary realization:** It is interesting to note that, when this new approach to avoid the intermediate model (and thus the need to impose 'meaning' at the very encoding stage) was being developed, some of the pioneers who lead the model based research trend, began voicing their own frustration over the fundamental steps involved in these techniques. S. K. Chang, in a recent evaluation report [Chan92] titled "Where Do We Go From Here?" writes:

"Spatial entities (object) and relationships (image features) in image do not carry any semantic meaning by themselves. Associating semantic meanings by naming cause some problems with image information. First, the same image could be interpreted in different ways. Second, the same image could be used in different ways during different time periods. Third, since the image interpretation is an approximation in many cases, it may be changed due to better recognition techniques. In any of the above cases, directly associating semantic meanings to image entities and relationships will severely limit the usage of image information"
Similar realizations have been expressed in the recent writings of several others leading researchers [Jain93, GrMe89]. However any solution is yet to be proposed.

**Proposed solution:** Holographic content based search with modulated attention, provides a different approach to perform content based search. It removes the need for attaching adhoc 'meaning' to the content just for the sake of intermediate modeling. The approach proposed in this research is to store and represent information as a physical entity as it is provided, and allow the inquirer to attach any 'meaning' to it just prior to the query.

The existing procedural search techniques that can support such direct search are paralyzingly inefficient. On the other hand, existing distributed associative search techniques (AAMs) are efficient but they lack the vital capability to focus. The holographic associative computing model with the ability to perform attention actuated search fills this critical void. It offers an efficient mechanism to perform the required search.

The inference mechanisms of traditional database systems are based on the classical bi-variate symbolic logic. The brittleness of bi-variate logic in the face of inherent structurelessness of image information also makes the traditional database technology ill-suited for managing images. The recent breakthrough in multivariate (fuzzy) logic, invented by Lotfi Zadeh [Zade65, Kosk87, Kosk88] has the ingredients to overcome the brittleness of bi-variate logic.

However coupling a high level fuzzy inference engine with a low level search engine will require some mechanism for bi-directional multivariate feedback to and from the search engine. Conventional AAMs do not have any such mechanism. In contrast, the *significance feedback* from holographic memory can be used perfectly for this coupling. This research also shows how the attention modulated search capability of holographic associative computing can be combined with the flexibility of fuzzy logic to obtain an intelligent content based search approach for image archives.
1.2 Summary and Contribution

The contributions of this dissertation are twofold and are summarized below.

**Principal Result 1 (Attention & AAM):** *This dissertation presents a new artificial associative memory which can learn pattern-associations and retrieve information with changeable attention on the stimulus field.*

**Corollary 1 (robustness):** The ability to focus makes the proposed memory one of the few AMs which can retrieve information from statistically weak cue. Within the specified field of attention it performs statistically robust matching.

**Corollary 2 (dynamic attention):** The attention can be changed dynamically during query without reorganization of the memory.

**Corollary 3 (MNC feedback):** It also has the unique ability to generate a feedback (called MNC) on the quality of match corresponding to the retrieved pattern.

Existing AAMs perform searches based on the entire sensory stimulus field. Thus, they lack the capability to perform effective retrieval, when the useful cue in the query pattern is statistically weak. Also, existing associative memories do not have any mechanism to convey the confidence measure on the search (as a consequence it cannot distinguish between a stored and unstored pattern). This new model brings artificial associative memories one step closer to the biological memories from the perspective of retrieval behavior.

As an application of this new associative computing mechanism it has been shown that how this memory system can be used for content-based search in image archive. Thus, **Principal Result 2 (IDB):** *This dissertation also presents one of the first associative computing based direct content-based search mechanism for query into image archive with automatic encoding.***

This new content-based search approach does not require the meaning of the images to be pre-extracted and modeled at the encoding stage, rather it attaches the meaning
dynamically during query in the form of attention. Thus, it can avoid the representational
insufficiency and encoding inaccuracy faced by the model based methods. Nor, does it
require intensive human involvement in encoding.

Conventional procedural search methods (such as hashing, table-lookup, binary sort,
etc.) fail to provide any efficient solution because multidimensional patterns can not be
pre-ordered to satisfy dynamic attention modulated search. Therefore, when compared to
the procedural search methods, the associative search of holographic system can be con­sidered almost exponentially (with respect to the number of images) faster.

Notably, the two principal research problems addressed in this dissertation (specially
the limitation of the existing AAMs) are yet not well identified. Perhaps, the identification
of these problems can be considered as part of the contribution made by this dissertation.
In the future, there may be other and better solutions available to these problems than the
ones proposed in this dissertation.
Fig-1.3 outlines the relationship between the two principal results. Holographic approach improves over the conventional search techniques. From the perspective of the existing multidimensional procedural search methods it adds an associative order to improve efficiency. From the perspective of distributed associative search methods it adds the capability to retrieve with changeable attention (RCA). The resulting search technique enables efficient direct content based search into image archive.

1.3 Dissertation Overview

The core of this dissertation has been organized in three successive themes with two chapters allocated for each. The first two chapters are dedicated to the identification of the problems (and review of the current methods) in artificial associative memory and image database technology. The next two chapters present the proposed paradigm of holographic associative computing. And the final two chapters show how this new computing technology can be used to perform intelligent associative search into image archive. Below the content of each of these chapters is described briefly.

Chapter 2 formally demonstrates the inadequacy of the existing associative computing to support attention modulated search. It also identifies the principal cause of such inability. The chapter also provides a review of the existing associative memories and formally characterizes various retrieval techniques beyond the traditional form supported by the existing AAMs.

Chapter 3 reviews the state-of-the-art in the search techniques used in image database. It reviews the current methodologies and their limitations. It also reviews the search complexities of existing procedural search methods, and demonstrates how these methods fail to provide any efficient support for direct content based search into large image database.
Chapter 4 presents the theory of this new computational method and explains how it works as an associative memory. It also demonstrates how this new computing paradigm can support retrieval based on changeable attention.

Chapter 5 presents the important characteristics of this memory. It first provides analytical investigation into the critical characteristics of this method. Finally, it provides the results of extensive computer simulation that has been performed to validate and further investigate the characteristics of this new paradigm.

Chapter 6 presents the proposed content associative image database search system (CAIDSS), its architecture, basic encoding and decoding formalism. It discusses the key design issues in the process of interpreting the associative retrieval capabilities of holographic computing into a database search system. It presents the design of the implemented CAIDSS prototype. It also presents analysis of its performance.

Chapter 7 presents the query formalism CQF. It explains both the query language and inference mechanism of the formalism. It shows how the unique output confidence measure of MHAC can be used to build a fuzzy logic based reasoning capability to CAIDSS. Finally, Chapter 8 summarizes the research by addressing the potential implications of this dissertation.

In each of these chapters, the introduction briefly explains the objective of the corresponding chapter and its logical organization. Also, each chapter is summarized at the end. When appropriate, a question-answer format has been used to visit important relevant but diversionary issues which deserve specific attention.
CHAPTER 2

ARTIFICIAL ASSOCIATIVE COMPUTING

2.1 Introduction

The objective of this chapter is to review the learning and retrieval mechanisms of existing parallel and distributed models of artificial associative memories (AAM) and to explain one of the fundamental limitations of these memories: their inability to retrieve information based on changeable attention.

This chapter is organized as follows. First, section 2.2 briefly reviews the current models of AAMs. A table has been compiled in appendix-A, which summarizes the principal characteristics of some important AAM models available today. Section 2.3 summarizes the characteristics of the two principal learning paradigms that have been used by current AAMs to store and retrieve information. Section 2.4 formally defines the generalized framework of attention modulated retrieval. Finally, section 2.5 identifies the cause underlying the failure of current learning and retrieval models to support such selective attention modulated search. It examines the role of classical scalar product rule of synaptic efficacy, and demonstrate that any network, which uses such a transmission rule, can not realize associative retrieval with changeable attention.

2.2 Development of Artificial Associative Computing

A parallel and distributed artificial associative memory (AAM) or artificial neural network (ANN) can be considered as a memory, which attempts to learn efficient representations of complex patterns, and is able to retrieve the stored information from a part of its content.
Associative memory differs from conventional computer memory in the sense that information is retrieved based on association. Conventional memory retrieves content based on address. In contrast, associative memory uses a part of content and retrieves the address. The current distributed models of artificial associative memory (AAM) are based on artificial neural computing to such an extent that these are synonymous to ANN.

2.2.1 Invention of AAM Architecture

2.2.1.1 Warren McCulloch & Walter Pitts [1943]

The first computational framework to construct an artificial associative memory was invented by McCulloch and Pitts in 1943 [McPi43]. They proposed the use of 'formal' neuron as a basis for general purpose computation (it will be referred to as MP-neuron). MP-neuron performs two step computation. In the first step it computes a weighted sum of its inputs and a bias. In the second step, it computes a nonlinearity over the sum. It can be any non-linear function such as binary, sigmoid, or threshold-linear. The second step is known as activation function.

\[
y_i = \sum_i w_{ij} s_j + b_i \quad \ldots(2.1a)
\]

\[
z_i = f(y_i) \quad \ldots(2.1b)
\]

However the overall transfer function does not have any name. It has remained so integral and inalienable part of neuro-computing, that hardly any one felt the necessity of naming it\(^7\). However for the purpose of the discussion presented in this dissertation, one is needed now. Since, the MP-neurons computes a scalar product of the inputs, It will be referred as the scalar product rule of synaptic efficacy (SPRSE) in the rest of this dissertation.

\(^7\)Except only recently, G. Carpenter, while referring to another non-optimality of this transfer function, called it as product rule of synaptic efficacy [Carp94]. I shall follow her.
McCulloch and Pitts showed that this simple MP-neuron can implement three simple operations of formal logic: AND, OR, and NOT. Fig-1 shows the weight and bias settings used for these operations. The neurons fire to +1 when the net excitation equals or exceeds zero, otherwise remain -1. The states are bipolar +1 or -1.

John Von Neumann [Neum58] demonstrated that redundant MP-nets can be designed to carry out arithmetical calculations with high reliability.

2.2.1.2 Aristotle to Symbolic Logic

The implication of the MP-neuron has been enormous and can be best understood in the light of the long human quest to find the basics of human intelligence, reasoning and thought process and the endeavor to recreate them artificially.

The formal methods for dialogue compiled by Aristotle can be considered as one of the earliest recorded attempts to find 'algorithm' of human reasoning. Although these rules for argument were mistaken for a tool for sound thinking, and were repeatedly used and criticized in the next two thousand years, it was just a hundred years ago, when a brilliant

---

8 Von-Neumann eventually decided not to lean towards parallel and distributed model of computation. He suggested the following changes to the serial architecture of ENIAC developed by Presper Eckert and John Mauchly of U. of Pennsylvania: (a) utilize binary numbers instead of analog (disbelieve in analog precision of MP-neurons?), and (b) store the program in the memory, instead of in the circuitry (doubt in easy analog universality of MP-nets?). The result was EDVAC, the first modern digital computer.

9 Aristotle laid down the foundation of formal logic by writing up a handbook with the rules for playing a kind of "yes-no" game. This "yes-no" game was prevalent in Plato's school as a formal method of debate. Plato's 'Dialogues' contains innumerable play-by-play descriptions of the game. The game was prevalent since the time of Soocrates. However by the time Aristotle arrived at the scene, as a young teacher in Plato's school, all sorts of variations were introduced, some of it were appalling. Hardly anyone was playing it in classical way. Aristotle decided to write up the rules, most probably to save the original beautiful game [FlesSI].

---
English mathematician named George Boole realized that Aristotelian method of logic was nothing but a 'mathematics' for thinking. Boole translated the complex system of formal logic into streamlined Boolean algebra, and hence sparked the science of symbolic logic. Subsequent development finally climaxed into the monumental development of three volumes of "Principia Mathematica" by B. Russell and A. N. Whitehead published in 1910, 1912, 1913 [RuWh10]. The achievement, described in the words of Susanne K. Lagler [Lang37]:

"Thousands of men, through thousands of years, have had millions of headaches over Aristotle’s system of logic.. with symbolic logic it can be reduced to three sets of equations..... For a true Aristotelian this exhausts the abstract system of logic."

Symbolic logic demonstrates how the knowledge pertaining to any complex reasoning mechanism can be represented through three basic logical operations. McCulloch and Pitts have demonstrated that MP-neuron can perform exactly these three basic operations. In the light of symbolic logic, the profound implication of McCulloch and Pitts work is that, any system which can be described by an expression of symbolic logic can be implemented by a network of MP-neurons\textsuperscript{10}.

\textbf{2.2.1.3 Hilberts Thirteenth Problem & Komogorov's Proof}

In the mean time mathematicians had already begun a more ambitious investigation into the decomposability of arbitrary analog functions into similar repetitive structures. In 1900, David Hilbert poised a fundamental question which is known as the thirteenth problem

\textsuperscript{10} In a parallel development, Alan Turing [1950] also looked into the same problem of universal approximation of Boolean expressions from mechanistic viewpoint and proved the existence of his Universal Turing Machine, that can simulate any mechanism describable in its formalism. Equivalence at the level of input and output is considers as the basis of formal equivalence for both systems.
[Lore76]. It can be stated as whether any continuous multivariate function can be approxi-
imated by a finite number of single variate continuous functions and a finite number of
additions and a finite number of compositions [Vemu92].

Although Hilbert himself conjectured that such decomposition can not be achieved,
in 1957 Kolmogorov and Arnold [Kolm57] independently confirmed the existence of a
negative solution to this problem. Kolmogorov demonstrated that:

**Theorem (Kolmogorov's):** Any continuous function $f(x_1, ..., x_n)$ of several variables
defined on $I^n (n \geq 2)$, where $I = [0,1]$ can be represented in the form

$$f(x_1, ..., x_n) = \sum_{j=1}^{2^n+1} \chi_j \left( \sum_{i=1}^{n} \psi_{ij}(x_i) \right)$$

Where $\chi_j$ and $\psi_{ij}$ are continuous functions of one variable and $\psi_{ij}$ are monotonic
functions that do not depend on $f$.

Kolmogorov's proof has immense implication. The gems of the proof are the facts
that component functions are independent of the universal function and the decomposition
is finite. Only $n(2n+1)$ pre-defined functions are enough to represent an arbitrary analog
function.

The first implication of Kolmogorov's theorem is of representability and computability
of universal systems. If the same set of component functions can be used to construct any
function just by adjusting some simple parameters, then, it means those parameters will be
a representation of the function. Thus, it will be possible to represent any function through
a finite set of parameters. It also implies that a finite architecture will be able to imitate the
input output behavior of any system.

The second implication of Kolmogorov's theorem is of learnability of universal sys-
tems. The problem of devising a machine that will be able to automatically learn any system
or mechanism reduces to the autonomous finding of the appropriate parameters. If, the number of parameters is finite, then a finite set of learning rules should be able to learn and mimic any function.

The possibility of using only a few functions as the basis set is even more exciting. It means that one has to worry about constructing only a few basic machine types. In such a case, a universal machine, no matter how complex is its target function, can be constructed by interconnecting these basis ones. In addition if one is looking for a mechanism to automatically learn how to mimic a function, only a few learning rules should be enough.

### 2.2.1.4 MP-neuron in the Light of Hilbert Program

McCulloch & Pitts showed in their paper that the formal MP-neuron can perform simple arithmetic, classification, and storage and retrieval of finite set of data etc. Donald M. Mackey [Mack54] expressed the capacity as follows:

"If you assert that there are certain process that a computer cannot go through, and if you can describe in words exactly what constitutes such a process, then at least one McCulloch and Pitt's net that can embody and carry out the process exists."

Of course, the above optimism hinges on the subtle imposition of the condition 'describable'. The universality of the MP-neuron requires the process to be describable by symbolic logic. The apparent neurobiological similarities, as well as the implications of Kolmogorov's theorem created enormous optimism about the capabilities of MP-neurons. Over the next 50 years, researchers begun extensive investigations on the capabilities of interconnected networks of MP-neurons. However it is interesting, that almost none of the researchers questioned the basic scalar product rule of synaptic efficacy, proposed by McCulloch and Pitts (the issue will be revisited in detail chapter-2.4).
However it should be pointed out that, there is still an important gap between Kolmogorov's proof and that of today's networks based on MP-neurons. It is still unanswered whether, the basis function (transfer function) used by MP-neurons is the one that will satisfy Kolmogorov's existential proof.

A series of works on functional approximation has been carried out subsequently with great optimism to prove specifically the universality of MP-neuron that uses this rule as the neuronal transfer function. Researchers such as Sprecher 1965 [Spre65], Ueseka 1971, Lorentz 1976 [Lore76], Poggio 1983 [Pogg83], Irie and Miyaka 1988 [IrMi88], Le Cun 1987 [leCu87], Lapeds and Farber 1988 [LaFa88], Gallant and White 1988, Kurt Hornik, et. al. 1989 [HoSW89], Funahashi 1989 focused on the universality of known transfer functions and some of its variants [Funa89]. However all these proofs specifically aimed at the MP-neuron based networks compromise either the finiteness or the accuracy conditions. As it is pointed out in previous section, the whole issue depends on the simultaneous satisfaction of both the conditions. It is still an open question, what exactly those basis functions of Kolmogorov look like. It would be appropriate to quote Hornik [HOSW89]:

"Are the successes observed to date reflective of some deep and fundamental approximation capability, or are they merely flukes, resulting from selective reporting and a fortuitous choice of problems? Are multi layer feedforward networks in fact inherently limited to approximate only some fairly special class of functions?"

The reason to bring out the above discussion is two fold. First, to trace how the very idea of scalar product rule of synaptic efficacy evolved and thrived, and second, to pin point the subtle gap left abridged on the road, as the march continued. This is just to suggest a cautious optimism in evaluating capabilities of conventional artificial neural computing, as this dissertation leads to the identification of a fundamental limitation of this very rule, that
has become so unalienable to artificial neural computing. However before introducing this limitation in section-2.4 in detail, the subsequent developments in artificial neurocomputing that followed the invention of MP-neuron will be discussed.

2.2.2 Learning

McCulloch and Pitts invented the architecture to represent and compute arbitrary functions (patterns). Immediately after, the attention shifted towards the autonomous learnability of these functions by learning the weights from input output instances (examples).

The principal research issue became the capability, efficiency, and capacity of MP-nets to learn, store, and retrieve arbitrary patterns efficiently with minimum error. Since then, countless models have been proposed to date. Appendix-A provides a brief survey of the characteristics of some of the important models of today. These models differ primarily on the way these MP-neurons are interconnected (network type), the type of activation linearity, or the environment (supervised, unsupervised, autonomous, etc.) under which these learn.

Current learning algorithms can be traced back to two independent lineages. The first one, now known as delta learning, evolved directly from the MP-neuron. The other lineage evolved through the study of learning itself as a psychological and biological phenomenon, now known as Hebbian learning. There also existed a third independent line that started when in 1948 Gabor [Gab048] invented the first artificial associative phenomena, optical Holography. Biological and physical groups converged shortly thereafter and emerged as correlation learning (also known as Hebbian learning). Next two subsections briefly describe these two principal paradigms.
2.2.3 Delta Learning Models

2.2.3.1 Rosenblatt's Perceptron

In 1958 Frank Rosenblatt proposed the first learning algorithm called "back-coupled learning rule" to make MP-neurons adaptive [Rose58]. Rosenblatt demonstrated that a single layer network of MP-neurons called Perceptron, can automatically learn a set of associations between stimuli and responses directly from examples without explicitly specifying their logical relations. This was the first encoding or learning algorithm for AAMs.

2.2.3.2 ADALINE/PADALINE and LMS Learning

Shortly after, in 1960, Bernard Widrow and M. E. Hoff presented an improved version of learning rule now known as Widrow Hoff Learning, along with their version of MP-nets (ADALINE/MADALINE). They demonstrated that ADALINE and MADALINE can retrieve a pattern which is provably nearest to the query example in least mean square (LMS) error sense [WiHo60].

2.2.3.3 Learning Ability of Single Layer MP-net

In 1969, Minsky and Papert [MiPa62] demonstrated that a single layer Perceptron, which depends on linear sum cannot learn arbitrary functions (for example XOR problem). However at that time, there was no known learning algorithm which could be applied to multilevel Perceptrons. In addition, Minsky and Papert speculated:

"It [Perceptron] has many features to attract attention; its linearity; its intriguing learning theorem... There is no reason to suppose that any of these virtues carry over to the many layered version... Perhaps some powerful convergence theorem will be discovered, or some profound reason for the failure to produce an interesting "learning theorem" for multi-layered machine will be found."
2.2.3.4 Multi-layer Feedforward Net and Back-propagation Learning

In contrary to Minsky and Papert’s speculation, Paul Webros [Webr74], David Parker [Park85] and others independently invented the Back-propagation learning algorithm for multi-layer feed forward nets.

2.2.4 Hebbian Learning Models

An independent research was already under way to understand the mechanism of learning based on neurobiological theories and psychological experiments. Donald Hebb [Hebb49] in his book "Organization of Behavior" published in 1943 provided a clue about how learning is performed in biological neurons.

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes place 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."

The idea was not new. The investigation of brain structure and neurons by Santiago Ramon y Cajal during 1852-1934 [Jaco93] and discovery of classical conditioning by psychologist Ivan P. Pavlov [Pavl27, CaBu90] leads to Hebb's rule. However the timing was crucial. As G. Carpenter [Carp89] writes it:

"But in the decade of McCulloch and Pitts, the formulation of the idea in the above sentence crystalized the notion in such a way that it become widely influential in the emerging neural network field."

As we will see, later Stephan Gorssberg [Gros69], A. Henry Klopf [Klop87] and others meticulously refined the idea and translated it into a series of learning algorithms, which were able to mimic biological learning phenomenon to a greater detail.
2.2.4.1 Matrix Memory and Correlation Learning

In 1956 Wilfrid Taylor constructed the first net from the direct application of Hebb's law. The emphasis was to directly address associative encoding as seen in biological memory. Taylor's net looked like a three layer Perceptron, however the individual neurons (cells) were analog instead of binary devices.

Shortly after this, in 1961, a very similar net was introduced by Karl Steinbuch called the "learning matrix". Following Steinbuch but mostly independently many others devised similar nets- such as James Anderson, Peter Buneman, Christopher Longuet Higgins, David Marr and Teuvo Kohonen [HiAn85, CoSh88, Koho72].

2.2.4.2 Neohebbian and Differential Hebbian Learning

In the mid-1950s, when researchers began to write computer simulations of Hebbian systems, they quickly realized that there are some practical problems. First of all, Hebbian learning only suggests weight increase. Such a rule is bound to lead to unlimited growth of the weights vectors.

In mid 1960s Grossberg presented explicit mathematical framework of Hebbian learning and introduced 'forgetting' [Gros67, Gros69] to solve the problem of unbounded growth. His expanded rule is known as Neohebbian learning. Neohebbian learning accounts for the fact that biological systems not only learn but also forget.

Neohebbian learning cannot explain causality of biological learning. In Pavlov's experiment it was observed that if both conditional and unconditional stimuli are applied simultaneously then it does not induce learning\textsuperscript{11}. The bell is required to ring slightly before the food. In 1987 Henry Klopf proposed the drive-reinforcement theory [Klop87] and

\textsuperscript{11} Such condition of causality is perfectly sensible from survival perspective. If there is no time to react, then there is no survival benefit from learning the association between the bell and food.
explained causality. The model required both inhibitory and excitatory synaptic connections, however, the same connection does not need to change modality. Interestingly, biological neurons are now thought to have such property [Sher06].

2.2.4.3 Grossberg’s Instar/ Outstar

Since the late 60’s Grossberg began to investigate the more complex behavioral aspects of biological learning; the capability of natural systems to learn autonomously. For autonomy he introduced a process of continuous learning through simultaneous classification and recollection. Grossberg hypothesized the existence of two distinct learning notions called Instar and Outstar [Gros69, Carp89]. Instar primarily focuses on classifying concepts and forming internal clusters. On the other hand, Outstar type focuses on the accurate regeneration of the patterns from their internal cluster representation. In his model, each neuron inside a network plays dual role in a continuum of time. Grossberg experimented with various learning rules (both Hebbian and non-hebbian) for Instar and Outstar learning.

Such an autonomous system that does not distinguish between learning and the recollection phase suffers from stability-plasticity dilemma [Carp89]. This is a problem that originates from the indiscriminate continuous forgetting introduced earlier by Grossberg himself. In the unsupervised mode it allows a network to forget important past learning, while it keeps on acquiring new ones even when these are not important.

In 1976 Grossberg [CaGr87a] presented, the adaptive resonance theory (ART) to resolve this dilemma. The networks managed to be stable enough to preserve significant past learning, as well as adaptable enough to incorporate new information. The key was a vigilance mechanism, which allowed the important past learning to resonate and lock itself
out in case the new pattern is not relevant. In 1987, he and Gail Carpenter proposed ART2 [CaGr87b], an analog version of ART1. However ART2 seems to suffer from stability problems.

2.2.4.4 Hopfield Network and Convergence of Networks

In 1982, physicist Hopfield provided the spin-glass analogy and demonstrated that a network of MP-neurons stores information through dynamically stable configurations. Hopfield showed how the collective convergence dynamics of ANN systems can be understood by the Lyapunov function.

Hopfield and Tank [HoTa85] also demonstrated how MP-neurons can solve optimization and constraint satisfaction problems. Optimization problems were considered difficult for ANN to learn because the optimization constraints are generally difficult to communicate just by sets of input and output. They demonstrated how the appropriate weights for solving Travelling Salesman problem (TSP) can be directly learned (or rather computed) using Lyapunov’s functions (also known as energy function).

2.2.4.5 Holographic Associative Memory

In the same decade as of McCulloch, Pitts and Hebb, in 1948, D. Gabor [Gabo48] who was working on a very different field, invented the first artificial system to demonstrate associative memory phenomena, optical holography. A number of authors, probably Van Heerden was the first among them [vaHe63], immediately drew the analogy between holograms and the brain. Their basis of the analogy was the distributed nature of both the

---

12 Like many other seminal works, Hopfield’s principle is also implicit in other people’s work [list in Vemu92, pp38]. For example, W. Ross. Ashby [Vemu92,pp23], in early days of cybernetics pointed out that brain activity patterns always tend towards dynamically stable configuration.
memories and their ability to withstand partial damage. Gabor studied the phenomena closely and noted correlation as the key to such memory [Gabo69], which was effectively a restatement of Hebb's learning rule.

Following the work of Gabor, in late 60's Willshaw started investigating the design of a distributed content-addressable memory on holographic principles [Will85]. In 1971 he proposed the correlograph model, using a simplification of optical holographic principles. Interestingly, one of these simplifications was the use of one-dimensional scalar representation for numbers instead of two-dimensional phasor optical waves. The methodology pursued in this dissertation can be considered as an investigation into this lost dimensionality. However, it not only investigates 2-D phasor representation, but also looks into a generalized higher dimensional representations.

2.3 Learning and Retrieval In Artificial Associative Computing

The above discussion presents briefly (but not completely) the highlights of the extensive research in artificial associative computing that has been carried out since its invention. The variety of architectural configuration and learning techniques that can be interpreted as artificial associative memory is staggering. Almost any book on neural networks contains excellent description, as well as classification schemes for existing AAMs [Vemu92]. Now, a brief technical description of the key aspects of current generation AAMs will be provided.

Almost all the models of ANNs and AAMs can be characterized to be built upon the computational cell (neuron) shown above. Fig-2.2 and equation (2.3) together show what can be almost unquestionably considered as an essential constituent of an AAM. Here, \( s \) represents the inputs to this neuron. Generally these assume values between 0 to 1. In a bipolar system they assume values from -1 to 1. The collection of \( w_i \) represents the memory.
It is surprising that, since the proposal of McCulloch and Pitts, the original model of artificial neuron cell (Fig-2.1) has remained virtually unaltered. The dynamics of this model have also remained almost unaltered as a two-step transformation. Virtually every model assumes this transfer function as a universal nameless fact of neurocomputing rather than a hypothesis. However neurophysiology has so far neither confirmed nor refuted the rule [Sher06, Carp94]. Also, as discussed in section-2.2.1.4 computational theory has not so far proven its universality as a finite real approximator.

Section-2.3.1 will summarize the principal learning and retrieval capabilities of existing associative memories. However before that notations of a formal memory system will be defined in section-2.3.1.
2.3.1 Memory Model

Fig-2.3 outlines schematic of an associative memory system. An associative memory stores a set in pattern associations. Let us consider a set of such stimulus patterns and a corresponding set of response patterns, which are to be stored in a memory:

\[ S = \{ S^1, S^2, \ldots, S^p \} \]
\[ R = \{ R^1, R^2, \ldots, R^p \} \]

Each of these patterns consists of elements, such that the stimulus and response patterns are vectors of form:

\[ S^\mu = [s^\mu_1 \ s^\mu_2 \ \ldots \ s^\mu_p] \]
\[ R^\mu = [r^\mu_1 \ r^\mu_2 \ \ldots \ r^\mu_p] \]

The superscript refers to index of the pattern vector, and the subscript refers to the particular element in the pattern vector. The individual vector elements are obtained from some external measurements. These measurements are mapped in a finite and real interval \( I = [a, b] \). The term content of a pattern generally refers to these measurements. Thus, an associative memory can be defined as:

**Definition (Associative memory):** Given, a set of stimulus pattern vectors \( S = \{ S^\mu \mid 1 \leq \mu \leq P \} \) and a set of response pattern vectors \( R = \{ R^\mu \mid 1 \leq \mu \leq P \} \) an associative memory is capable of learning the associations between a stimulus member \( S^\mu \in S \) and the corresponding response member \( R^\mu \in R \) in such a way that, given a query pattern \( S^Q \), it can retrieve a pattern \( R^R = R^T \) such that \( R^T \in R \), and \( S^Q \) is closest to \( S^T \in S \) according to a matching criterion \( D \).

An associative memory system needs (i) a learning algorithm \( A_{learn} \) which converts all the \( \{ S^\mu, R^\mu \} \) associations into some internal representation, (ii) a physical storage medium and representation formalism \( A_M \) to store the associations, (iii) a decoding algorithm \( A_{retrieve} \) to recollect stored information \( R^R \) from a given query stimulus \( S^Q \), and (iv) a matching criterion \( D \) to measure the closeness of stimulus patterns to the query pattern.
Principal Learning Paradigms: As stated earlier (section-2.2.2, the learning algorithms that have evolved over the years can be classified into two main categories, (a) Delta learning, and (b) Hebbian learning. However it is possible to demonstrate that the matching criterion used in both of these learning paradigms are related [Koho87]. In the following two subsections, closer observation about both of these paradigms will be made.

2.3.2 Hebbian Learning

As noted, this is one of the earliest and the best known learning rules based on Hebb's observation of biological learning. N e o h e b b i a n learning [Gros69], Differential Hebbian learning, Drive reinforcement learning [Klop87] Hopfield rule [Hopf82], Energy Function minimization rule [HoTa84], Kohonen's competitive learning [Koho89], Principal component learning [Sang89] Taylor, Steinbuch and Willshaw's correlation learning are also variant of this learning rule. Below the description and principal characteristics of the Hebbian class of learning are given.

2.3.2.1 Algorithm

According to Hebb's rule the weights are computed by:
\[ w_q^{(\text{new})} = w_q^{(\text{old})} + s_i r_j \] ...(2.4a.1)

However generally, the simple Hebbian rule does not require the weights to be computed iteratively. It can be computed directly by outer product of the stimulus and response pattern:

\[
\begin{bmatrix}
S^1 \\
S^2 \\
\vdots \\
S^n \\
\end{bmatrix}
\begin{bmatrix}
r_1^\mu \\
r_2^\mu \\
\vdots \\
r_m^\mu \\
\end{bmatrix}
= 
\begin{bmatrix}
S^1 r_1^\mu & S^1 r_2^\mu & \ldots & S^1 r_m^\mu \\
S^2 r_1^\mu & S^2 r_2^\mu & \ldots & S^2 r_m^\mu \\
\vdots & \vdots & \ddots & \vdots \\
S^n r_1^\mu & S^n r_2^\mu & \ldots & S^n r_m^\mu \\
\end{bmatrix} 
\] ...(2.4a.2)

To store, the association set \( S \) and \( R \), these product matrices are superimposed. The resultant weight matrix is given by:

\[
W = \sum_{\mu} [S^\mu]^T R^\mu 
\] ...(2.4a.3)

2.3.2.2 Constraints

Hebb's rule requires two conditions to operate:

1. The first issue is whether such a weight-set exists which can represent the inputs to desired output. If the input vectors are linearly independent, then such a weight exists. This is in fact the extension of the linear separability problem for vector output.

2. The second constraint is perfect recollection. If the input vectors are orthogonal (uncorrelated), then the weight matrix given by this rule will be able to recall perfectly. If, they are not, there will cross-contribution from all other patterns.
2.3.2.3 Matching Criterion

The weight-set learned by Hebbian rule is always attracted by one of the attractors from the encoded set, which has the maximum projection along the query vector. The retrieval equation provided by inner product is actually the estimation of the projection. The matching criterion is thus:

\[
\max_{\mu}^{p} \left[ \| S^\mu \| \cdot \| S^Q \| \cos \theta^{(\mu, Q)} \right] = \min_{\mu}^{p} \left[ -\| S^\mu \| \cdot \| S^Q \| \cos \theta^{(\mu, Q)} \right] \tag{2.4b}
\]

2.3.3 LMS Learning

LMS learning too represents a large class of learning algorithms. Its origin is in the Perceptron learning algorithm introduced by Rosenblatt [Rose58, Rose62]. Widrow and Hoff [WiHo60] provided the LMS rule with the proof of optimality. Subsequently, it was extended for multi-layer feedforward networks by Parker, Webrros and several other researchers, which is now known as Backpropagation learning. Among the various influential versions are generalized Backpropagation [Faus94], recurrent Backpropagation [RuHW86], Quickprop [FahI88, Faus94]. Below the description and principal characteristics of the LMS class of learning are given.

2.3.3.1 Algorithm

The weights are computed by the following iterative equations in two steps. \( z = f(y) \) represents an arbitrary activation function that is differentiable. First, each of the MP-neurons computes the weighted sum for the given inputs:

\[
y_j = \sum_{i}^{n} w_i s_i, \text{ and } z_j = f(y_j) \tag{2.5a.1}
\]

The change of weight is computed by the following:

\[
\Delta w_j(t + 1) = \alpha(r_j - z_j)s_i \cdot f'(y_j) \tag{2.5a.2}
\]
Here $\alpha$ represents the rate of learning. It guides the fraction of error that is learned in each iteration. It is generally less than 2.

### 2.3.3.2 Constraints

1. As before, the first issue is whether such a weight set exists that can map the inputs to desired output. If the input vectors are linearly independent, then such a weight exists. In other words this is the extension of the linear separability problem for vector output. Linearly separable mappings can be encoded by using a single layer of neurons using delta rule.

2. However it does not require orthogonality among the patterns.

### 2.3.3.3 Matching Criterion

Delta rule produces the least square solution, even when input patterns are not linearly separable [RuMc86,Faus94]. Therefore, the weight-set learned by this rule always is attracted by one of the attractors from the encoded set, which has the minimum mean squared error from the query vector.

$$\min_{\mu} \left[ \sum_{i} (r_i^\mu - z_i^\mu)^2 \right] \quad \ldots(2.5b)$$

### 2.4 Retrieval with Changeable Attention (RCA)

Since the invention of the MP-neuron by McCulloch and Pitts, the principal research impetus on artificial models of associative memory has been concentrated on perfecting learning aspects of memories. Research has been vigorously pursued on issues such as how arbitrary associations can be learned, how more patterns can be learned, how patterns can be stored more accurately, how learning can be accrued autonomously in the continuum of time, etc. However the problems addressed in all these years can be considered as principally learning centric. Although, recollection of information is the ultimate objective of learning, surprisingly, very few attempts have been made to examine the recollection aspect of these
memories, other than assuming a very simple model of retrieval. Almost all the proposed learning models have been constructed on the assumption of a simple and restricted retrieval scenario, where the sample of the content that is used during query is a close replica of the target. However it is possible to show that almost all these models of AAMs are unable to support attention modulated retrieval. On the other hand, such retrieval seems to be a natural capability of biological memories.

**Changeable attention:** The phenomena have been already illustrated in section-1 (Fig-1.1). Below we formally define the problem in the context of associative memory model presented earlier:

**Definition (attention):** Attention refers to the fact that any subset\(^{13}\) \(F^Q \in N\) of the elements in the example query pattern \(S^Q\) can be dynamically specified as a field of attention and the memory can confine its associative match only within \(F^Q\).

The most important aspect of attention based retrieval is the ability to change the distribution of attention over element space dynamically during post-learning query. This ability is equivalent to the capability of changing perspective over the same sensory input. If a specific distribution of attention is given during encoding before learning, a conventional AAM can hard-encode it in the synaptic weights. However, once the learning is over, it does not allow the flexibility to recast the distribution of attention any more during query. For a given learning, it acts as a deterministic machine where each initial state flows into a predetermined single attractor. Conventional AAMs have no mechanism to accommodate post-learning change in the distribution of attention over its element space.

---

\(^{13}\) The membership in the subset comprising the field of attention can be bivalued or analog (also known as fuzzy set). In analog set a particular element can be a member of this subset partially, instead of being fully included or being fully excluded.
Statistically weak cue: One of the repercussions of such limitation is the inability to deal with a statistically weak cue. Intuitively, a mechanism without the ability to focus will always have to consider all elements of information even if they are knowingly wrong. On the other hand, the mechanism that has the ability to ignore, should be able to cope with small cue by ignoring the rest. The empirical evidence that demonstrates the inability of conventional AMs to cope with small cue has been shown earlier in section-1.1.

Next two sections formally investigates the theoretical cause of this fundamental limitation of current models. Section-2.4.1 defines what is meant by retrieval with changeable attention (RCA), and how it relates to the simple retrieval mechanism assumed by existing models. Then, section-2.4.2 demonstrates the inability of current models to perform RCA in the form of a simple proof. Section 2.5 identifies the roots of this problem. The problem originates from the very representation, and subsequently from the basic scalar product rule of synaptic efficacy (SPRSE).

In this dissertation, a method will be demonstrated which can overcome this fundamental limitation of existing ANN models and is capable of retrieving information even when the useful cue becomes very small by selective attention over the effective bits.

2.4.1 Definition of RCA queries

Concept of pattern distance is essential to any match. The overall distance between two patterns is compositional and is estimated as a function of individual distance measures between the corresponding elements of the two patterns. In this section, an attempt will be made to construct a measure of such compositional distance with minimum possible unnecessary specifications. Below, such a generalized compositional distance measure is given.

\[ D(S^a, S^b) = \left[ \sum_i^N \delta(\text{dist}(s^a_i, s^b_i)) \right] \]  \hspace{1cm} \text{(2.6)}
Where, \( \text{dist}() \) is the \textit{distance measure function} (DMF) between any two pattern elements. It can be any arbitrary function with the constraint that it is a monotonically increasing or decreasing function with respect to \( |s_i^a - s_i^b| \), and symmetric for all \((s_i^a, s_i^b)\) pairs.

\( \delta() \) is the \textit{distance composition function} (DCF). It also can be any arbitrary function with the constraint that it too is monotonic with \( \text{dist}() \), and the direction of its monotonicity is same as that of \( \text{dist}() \).

Finally, the overall function is required to have the property, that \( D(S^a, S^a) = c \), where \( c \) is a constant independent of specific pattern index \( a \).

The above function, only assumes that the distance between two patterns \( S^a \) and \( S^b \) is measured by comparing the distances between the individual elements and the overall distance measure is monotonic.

From, this generalized measure of distance, it is possible to define the following three matching criteria and corresponding query types for the memory system model defined in section-2.3.1. As it will be evident, the first of these cases directly corresponds to the query supported by existing conventional AAMs.

2.4.1.1 Retrieval with Unchangeable Attention (unary)

The problem that an associative memory can retrieve based on global match over the element space performs, can be stated in the following form:

**Definition (type-U query):** Can an AM be constructed that can retrieved \( R^R \equiv R^{T^a} \), when:

\[
D\left( S^O, S^T \right) = \min_{\mu} \left[ \min_{i} \text{dist}(s_i^O, s_i^\mu) \right] \quad \text{...(2.7)}
\]
2.4.1.2 Retrieval with Changeable Attention (binary)

If $F^Q \subseteq N$ represents a subspace of the total element space $N$, then problem that an associative memory can retrieval with changeable attention can be stated in the following form:

**Definition (type-B query):** Can an AM be constructed that can retrieved $R^R = R^T_s$, where the set of elements in attention vector $F^Q$ is dynamically specified during query, and such that:

$$D(S^Q, S^T_s, F^Q) = \min_{\mu} \left[ \text{dist}(s_i^Q, s_i^T) \right]$$  \hspace{1cm} \text{(2.8)}

2.4.1.3 Retrieval with Changeable Attention (analog)

The above retrieval can be further generalized when, the attention on a specific element can be distributed instead of being completely included or excluded. The generalized form of the problem of retrieval with changeable analog attention can be stated in the following form:

**Definition (type-A query):** Can an AM be constructed that can retrieved $R^R = R^T_s$, where the analog attention on the stimulus elements is represented by the dynamically specifiable query vector, $A^Q = [\lambda_1^Q \lambda_2^Q ... \lambda_n^Q]$, and $\lambda_i^Q : [0, 1]$.

$$D(S^Q, S^T_s, A^Q) = \min_{\mu} \left[ \text{dist}(s_i^Q, s_i^T) \right]$$  \hspace{1cm} \text{(2.9)}

Where, the analog attention on the stimulus elements is represented by an additional query vector of length $n$, $A^Q = [\lambda_1^Q \lambda_2^Q ... \lambda_n^Q]$ with $\lambda_i^Q : [0, 1]$, and such that:
Query type-A is a generalized form of both other types of query. If \( \Sigma^Q = [1.0 \ 1.0 \ 1.0 \ \ldots \ 1.0] \), then criterion (2.9) becomes equivalent to criterion (2.7). On the other hand, if \( \lambda_i^Q = 1.0 \), when \( i \in F^Q \), or \( \lambda_i^Q = 0.0 \) otherwise, then search criterion (2.9) becomes equivalent to search criterion (2.8).

2.4.2 Existing Neural Computation and Search Types

The above three types of retrieval vary based on their scope of the matching functions over the element space. If the matching criteria used by the existing neural computational models are compared to the above categories, it becomes evident that invariably these belong to the type-U category. In this category the distance composition function \( \delta() \) is single-variate linear function. As shown in (2.4b) and (2.5b), the distance measure in Hebbian class of learning is dot-product of the patterns, and that for the LMS class of learning is mean square error. There are few other variants of measure (for example entropy) used in conventional neuro-computing. However the key feature to note is that for all of these the distance composition functions are a single-variate function of distance. Whereas, both type-A and type-B distance measures have additional query variable(s) (either \( F^Q \), or \( \Lambda^Q \)) representing the attention distribution.

2.4.3 Non-Optimality of Conventional AAMs in RCA Search

Now the inability of the all inclusive optimization based AAMs in performing RCA type-B and type-A search will be demonstrated formally.

**Theorem:** An associative memory constructed by interconnecting cells with the scalar product rule of synaptic transmission specified by the equation below can not realize the retrieval of type-B, or type-A. Where, \( f() \) is any single variate function, and \( s_i \) is a real valued number in the range \( I = [0,1] \), and the weights \( w_i \) contains the learned pattern.

\[
r = f\left(\sum_{i}^{n} w_{qi} s_i + b_i \right)
\]
Proof: The proof has been constructed in two parts. In the first part, it is shown that a network realizing all element inclusive scope of optimization cannot converge to a correct result with respect to the RCA type-B and type-A search criterion. In the second part, it is shown that the scope of the optimization can not be modified during query for any network which is based on MP-neurons with a SPRSE transfer function.

Part 1 (Problem of All-inclusive Optimization): Let us consider a trained network, which has memorized two patterns \( S^1 = [s^1_1 \ s^1_2 \ \ldots \ s^1_n] \) and \( S^2 = [s^2_1 \ s^2_2 \ \ldots \ s^2_n] \). Let us divide the set of total element space \( N \) into two arbitrary subsets \( A \) and \( B \), such that \( A \cup B = N \).

Without loss of generality let us also assume:

\[
\left[ \sum_{i}^{A} \text{dist}(s^1_i, s^2_i) \right] < \left[ \sum_{i}^{B} \text{dist}(s^1_i, s^2_i) \right]
\]

\( \ldots (2.10) \)

Now, let us consider a query stimulus, \( S^Q = [s^Q_1 \ s^Q_2 \ \ldots \ s^Q_n] \) carefully constructed in such a way that:

\( s^Q_i = s^1_i, \text{ when } i \in A \)

\( = s^2_i, \text{ when } i \in B \)

Let us also consider an attention distribution vector, \( A^Q = [\lambda^Q_1 \ \lambda^Q_2 \ \ldots \ \lambda^Q_n] \) where:

\( \lambda^Q_i = 0.0, \text{ when } i \in A \)

\( = 1.0, \text{ when } i \in B \)

Now, first investigate the optimum result expected from the query of type -A. From the initial state \( S^Q \), the distance measure according to a search of type-A, between \( S^Q \) and \( S^1 \) is:
\[ D(S^Q, S^1, \Lambda^Q) = \left[ \sum_{i}^{N} \lambda_i^Q \text{dist}(s_i^Q, s_i^1) \right] \]

\[ = \left[ \sum_{i}^{B} \text{dist}(s_i^Q, s_i^1) \right] \] \hspace{1cm} \text{...(2.11a)}

And, the distance from the second stimulus is:

\[ D(S^Q, S^2, \Lambda^Q) = \left[ \sum_{i}^{N} \lambda_i^Q \text{dist}(s_i^Q, s_i^2) \right] \]

\[ = \left[ \sum_{i}^{B} \text{dist}(s_i^Q, s_i^2) \right] \]

\[ = 0 \] \hspace{1cm} \text{...(2.11b)}

Thus, from (2.11a) and 2.11b):

\[ D(S^Q, S^1, \Lambda^Q) > D(S^Q, S^1, \Lambda^Q) \] \hspace{1cm} \text{...(2.11c)}

Which implies the expected result is \( R^R = R^2 \).

Now let us see the actual output of a cell with above definition. The distance between the patterns according to estimate A:

\[ D(S^Q, S^1) = \sum_{i}^{N} \text{dist}(s_i^Q, s_i^1) \]

\[ = \sum_{i}^{A} \text{dist}(s_i^Q, s_i^1) + \sum_{i}^{B} \text{dist}(s_i^Q, s_i^1) \]

\[ = 0 + \sum_{i}^{B} \text{dist}(s_i^Q, s_i^1) \] \hspace{1cm} \text{...(2.12a)}

Similarly, the distance measure between \( S^Q \) and \( S^2 \) is given by:
\[ D(S^Q, S^2) = \sum_{i}^{N} \text{dist}(s_i^Q, s_i^2) \]

\[ = \sum_{i}^{A} \text{dist}(s_i^Q, s_i^2) + \sum_{i}^{B} \text{dist}(s_i^Q, s_i^2) \]

\[ = \sum_{i}^{A} \text{dist}(s_i^Q, s_i^2) + 0 \]

\[ \cdots (2.12b) \]

An optimally trained network of such cells will converge to a pattern which is at minimum distance from the query.

A learning based on LMS rule or its variant characterized by equation (2.4b) will converge in a least mean square error solution, on the other hand a network, with Hebbian learning or its variant characterized by equation (2.5b) will converge to maximum dot product solution. Due to inequality (2.10), in both cases:

\[ D(S^Q, S^2) > D(S^Q, S^1) \]

\[ \cdots (2.12c) \]

Thus, from (2.12a) and (2.12b), the produced result will be \( R^B = R^1 \). Which is a clear contradiction to the expected result from search of type-B or type-A.

\( \text{(proved) \hspace{1cm} \text{Part 2 (Scope Inflexibility of SPRSE neuron): Let us consider, the role of any } i^{th} \text{ neuron in the network. Let us also consider that to reconstruct the expected pattern } R^2, \text{ its corresponding ideal output is } z_i^{\text{ideal}} = f(y_i^{\text{ideal}}). \text{ The corresponding ideal input vector is } S^{\text{ideal}} = [s_1, s_2, s_3, \ldots, s_n]. \text{ Let, the learned weight vector is } W = [w_1, w_2, w_3, \ldots, w_n]. \text{ Therefore, the ideal weighted summation output of the cell is:} \)

\[ y_i^{\text{ideal}} = \sum_{j}^{N} w_j s_j + b_i \]

Now, if the attention vector \( \Lambda^{Q} \) is imposed on it as a scope constraint, then the modified and constrained output of the cell becomes:
\[ y_i^{\text{const}} = \sum_j^B w_{ij} s_i + b_i \]

Thus, the error is:

\[ y_i^{\text{error}} = y_i^{\text{ideal}} - y_i^{\text{const}} = \sum_j^A w_{ij} s_i \]

Individual terms in this summation are in the order of \( O(w \cdot s) \). In addition, if:

\[ B \ll N, \text{ or } \sum_i^B \lambda_i \ll N \]

Then the overall summation itself will be in the order of \( y^{\text{error}} \equiv O(y^{\text{ideal}}) \). For large \( N \equiv A \gg 1 \) the sum will behave like a random walk and the expected value of its growth will be of the order of \( |w_s| \cdot \sqrt{N} \).

The exact error at the output will depend on the specific type of the activation function \( f(.) \). For any non-linearity with unimodal first derivative function (which includes all sigmoidal and step non-linearity used by conventional AAMs), small errors will be corrected but larger errors will be magnified. Thus, the actual output of this cell will also be numerically off-balanced from the ideal output in the order of:

\[ z^{\text{error}} = g(z_i^{\text{ideal}}) - g(z_i^{\text{const}}) = O(z_i^{\text{ideal}}) \quad (proved) \]

The above analysis is true for any cell in a network. For \( z^{\text{error}} \equiv O(z_i^{\text{ideal}}) \), a network of non-linear SPRSE neurons will run into avalanche magnification of error\(^{14}\). Therefore, collectively a network made of SPRSE rule will fail to converge when the scope is altered.

\[^{14}\text{However, a network of non-linear neurons will be error correcting if } z^{\text{error}} = 0.\]
2.4.4 Scalar-Product-Rule of Synaptic Transmission & RCA

Now the inability of the all inclusive optimization based AAMs in performing RCA type-B and type-A search will be demonstrated formally.

**Theorem:** An associative memory constructed by interconnecting cells with the scalar product rule of synaptic efficacy (SPRSE) specified by the equation below can not realize the retrieval of type-B, or type-A. Where, \( f() \) is any single variate function, and \( s_i \) is a real valued number in the range \( i=[0,1] \), and the weights \( w_{ij} \) contains the learned pattern.

\[
r = f\left( \sum_{i} w_{ij} s_i + b_i \right)
\]

Proof: The proof has been constructed in two parts. In the first part, it is shown that a network realizing all element inclusive scope of optimization cannot converge to a correct result with respect to the RCA type-B and type-A search criterion. In the second part, it is shown that the scope of the optimization can not be modified during query for any network which is based on MP-neurons with a SPRSE transfer function.

**Part 1 (Problem of All-inclusive Optimization):** Let us consider a trained network, which has memorized two patterns \( S^1 = [s^1_1 \ s^1_2 \ \ldots \ s^1_i] \) and \( S^2 = [s^2_1 \ s^2_2 \ \ldots \ s^2_i] \). Let us divide the set of total element space \( N \) into two arbitrary disjoint subsets \( A \) and \( B \), such that \( A \cup B = N \), and \( A \cap B = \emptyset \). Without loss of generality let us also assume:

\[
\left[ \sum_{i} \text{dist}(s^1_i, s^2_i) \right] < \left[ \sum_{i} \text{dist}(s^1_i, s^2_i) \right] \quad \ldots(2.10)
\]

Now, let us consider a query stimulus, \( S^Q = [s^Q_1 \ s^Q_2 \ \ldots \ s^Q_i] \) carefully constructed in such a way that:

\[s^Q_i = s^1_i, \text{ when } i \in A\]

\[= s^2_i, \text{ when } i \in B\]
Let us also consider an attention distribution vector, $\Lambda^O = [\lambda^O_1 \ \ \lambda^O_2 \ \ \ldots \ \ \lambda^O_n]$ where:

$\lambda^O_i = 0.0$, when $i \in A$

$= 1.0$, when $i \in B$

Now, first investigate the optimum result expected from the query of type -A. From the initial state $S^Q$, the distance measure according to a search of type-A, between $S^Q$ and $S^1$ is:

$$D(S^Q, S^1, \Lambda^O) = \left[ \sum_{i}^{N} \lambda^O_i \text{dist}(s^Q_i, s^1_i) \right]$$

$$= \left[ \sum_{i}^{B} \text{dist}(s^Q_i, s^1_i) \right] \quad \text{...(2.11a)}$$

And, the distance from the second stimulus is:

$$D(S^Q, S^2, \Lambda^O) = \left[ \sum_{i}^{N} \lambda^O_i \text{dist}(s^Q_i, s^2_i) \right]$$

$$= \left[ \sum_{i}^{B} \text{dist}(s^Q_i, s^2_i) \right]$$

$$= 0 \quad \text{...(2.11b)}$$

Thus, from (2.11a) and 2.11b):

$$D(S^Q, S^1, \Lambda^O) > D(S^Q, S^1, \Lambda^O) \quad \text{...(2.11c)}$$

Which implies the expected result is $R^R = R^2$.

Now let us see the actual output of a cell with above definition. The distance between the patterns according to estimate A:
\[ D(S^Q, S^1) = \sum_{i}^{N} \text{dist}(s_i^Q, s_i^1) \]
\[ = \sum_{i}^{A} \text{dist}(s_i^Q, s_i^1) + \sum_{i}^{B} \text{dist}(s_i^Q, s_i^1) \]
\[ = 0 + \sum_{i}^{B} \text{dist}(s_i^Q, s_i^1) \]
\[ \text{... (2.12a)} \]

Similarly, the distance measure between \( S^Q \) and \( S^2 \) is given by:
\[ D(S^Q, S^2) = \sum_{i}^{N} \text{dist}(s_i^Q, s_i^2) \]
\[ = \sum_{i}^{A} \text{dist}(s_i^Q, s_i^2) + \sum_{i}^{B} \text{dist}(s_i^Q, s_i^2) \]
\[ = \sum_{i}^{A} \text{dist}(s_i^Q, s_i^2) + 0 \]
\[ \text{... (2.12b)} \]

An optimally trained network of such cells will converge to a pattern which is at minimum distance from the query.

A learning based on LMS rule or its variant characterized by equation (2.4b) will converge in a least mean square error solution, on the other hand a network, with Hebbian learning or its variant characterized by equation (2.5b) will converge to maximum dot product solution. Due to inequality (2.10), in both cases:

\[ D(S^Q, S^2) > D(S^Q, S^1) \]
\[ \text{... (2.12c)} \]

Thus, from (2.12a) and (2.12b), the produced result will be \( R^R = R^1 \). Which is a clear contradiction to the expected result from search of type-B or type-A.

\( \text{proved} \)

\textbf{Part 2 (Scope Inflexibility of SPRSE neuron):} Let us consider, the role of any \( i^{th} \) neuron in the network. Let us also consider that to reconstruct the expected pattern \( R^2 \), its corre-
sponding ideal output is \( z_i^{\text{ideal}} = f(y_i^{\text{ideal}}) \). The corresponding ideal input vector is \( S^{\text{ideal}} = [s_1, s_2, s_3, \ldots, s_n] \). Let, the learned weight vector is \( W = [w_1, w_2, w_3, \ldots, w_n] \).

Therefore, the ideal weighted summation output of the cell is:

\[
y_i^{\text{ideal}} = \sum_{j}^{N} w_j s_i + b_i
\]

Now, if the attention vector \( \Lambda^Q \) is imposed on it as a scope constraint, then the modified and constrained output of the cell becomes:

\[
y_i^{\text{const}} = \sum_{j}^{B} w_j s_i + b_i
\]

Thus, the error is:

\[
y_i^{\text{error}} = y_i^{\text{ideal}} - y_i^{\text{const}} = \sum_{j}^{A} w_j s_i
\]

Individual terms in this summation is in the order of \( O(ws) \). In addition, if:

\[
B \ll N, \quad \text{or} \quad \sum_{i}^{B} \lambda_i \ll N
\]

Then the overall summation itself will be in the order of \( y^{\text{error}} \equiv O(y^{\text{ideal}}) \). For large \( N \equiv A \gg 1 \) the sum will behave like a random walk and the expected value of its growth will be of the order of \( |ws| \sqrt{N} \).
The exact error at the output will depend on the specific type of the activation function \( f(\cdot) \). For any non-linearity with unimodal first derivative function (which includes all sigmoidal and step non-linearity used by conventional AAMs), small errors will be corrected but larger errors will be magnified. Thus, the actual output of this cell will also be numerically off-balanced from the ideal output in the order of:

\[
z_{\text{error}} = g(z_{\text{ideal}}^i) - g(z_{\text{const}}^i) = O(z_{\text{ideal}}^i)
\]

(proved)

This analysis is true for any cell in a network. For \( z_{\text{error}} = O(z_{\text{ideal}}) \), a network of non-linear SPRSE neurons will run into avalanche magnification of error\(^\text{15}\). Therefore, collectively a network made of SPRSE based neurons will fail to converge when the scope varies during query.

**Rigidity of Conventional AAMs Retrieval:** As far as various learning algorithms are concerned, the objective of the learning rules is to select a set of weights \( w \), which allows the cell to have only a finite number of stable points which correspond to the learned responses. On the other hand, the objective of retrieval algorithms is to guarantee that a given query stimulus state converges to the nearest stable set. However the key point is that a cell of the above architecture after training with a fixed weight-set \( w \) can be considered as deterministic machine, in the sense, that for every initial state \( S^q \), there is a predetermined final state \( R^q \), where the network converges. But, both the queries of type B, and A, incorporates the possibility of convergebility of an initial pattern state \( S^q \), to multiple final states depending on \( \Lambda^q \).

\(^{15}\) However, a network of non-linear neurons will be error correcting if \( z_{\text{error}} = 0 \).
2.5 Role of Synaptic Transmission Rule

Previous section has analyzed the limitation of the collective characteristics of conventional neuro-computing. In this section, now a closer look at the functionality of the MP-neuron will be taken. How the cause of failure of conventional AAMs to support RCA originates right at this cell is shown here.

First, the nature of robustness that is generally attributed to neural computing without that much of clarification will be investigated. Then, it will be shown how actually this is related to RCA.

A natural question that follows is whether some form of architectural or algorithmic modification of the basic MP-neuron dynamics would be able to correct such limitation. Unfortunately, the solution is not so intuitive (or rather intuitive solutions are not the actual solution). This important issue is also illustrated in this section.

**Nature of Neural Robustness:** Fig-2.4 shows the operation of a typical McCulloch Pitts neuron. The computation involves a weighted summation set of inputs $S_j$. The weights are determined by learning and remain static for a given learning. The final output $z$ of the neuron is obtained by passing the weighted sum through the activation function.

Let us consider, that the network has been trained to produce a perfect output $z$ for the memorized input vector $S = [s_1, s_2, s_3, \ldots, s_n]$. Let, the corresponding weight vector is $W = [w_1, w_2, w_3, \ldots, w_n]$. Now let us consider the situation when a slightly noisy version of input, $S' = S + E$ is used as input.

The weighted sum $y'$ then shifts by a resultant error of amount $WE$ in the sum, which in terms shifts the final output to $z'$. The amount of error reflected in the final output ($z - z'$) is not necessarily proportional to $WE$, rather, it depends on the non-linearity. In fact, the
network may even rectify the error completely depending on the location of \( y \), and type of the non-linearity. For example, a step function as shown in the figure can correct the error when \( WE \) is less than \( y \).

The error correction capability of a network of such neurons critically depends on activation non-linearity. What is generally referred as the 'robustness' of an ANN originates from this effect of non-linearity. However the same non-linearity that helps in correcting error may also catastrophically amplify error. For example, if \( y \) lies at the vicinity of 0, then even a small resultant error \( WE \) can tilt the balance from completely correct recall to completely erroneous recall.

The size of resultant error depends on the statistical balance between the 'correct' versus 'incorrect' components of error. The exact weight of a particular erroneous element is decided by fixed vector \( W \), which is pre-decided during learning, long before the query, and can not be modified dynamically at query. This eliminates the possibility of multiplicative modification of synaptic inputs in proportion to their attention.

The above discussion shows that what is generally known as robustness of ANN, is actually its robustness against noisy input. For correct recall, the actual signal strength in the query-pattern must be statistically significantly dominant over the noise strength.

**Attention and Robustness:** Now, let us turn to the issue of attention. If, one considers the roller retrieval problem illustrated in fig-1.2, it will be evident that another way of describing attention is to ask whether it is possible to selectively ignore. To be specific, the pattern elements outside the focus of attention must be ignored from the query pattern set during query. Thus, RCA is equivalent to a fault situation where some of the pattern elements are missing. Are MP-neurons robust against such a fault?
To answer this question, let us investigate what happens if during the query, it is known that a subset of input elements are erroneous. Within the framework of the MP-neuron is it possible to obtain the optimum result with respect to this added knowledge?

Let us assume that it is a priori known that the first element is wrong. Can we utilize this information and stop them contributing to the resultant error \( WE \)? It may appear that, one obvious solution is to set element \( s_1 \) to zero (zero setting), or equivalently, not to include the inputs from \( s_1 \) in the summation by some hardware means (switch). But, a closer look will reveal, rather than solving it worsens the problem. Such attempt will still contribute an error of the order \( w_1(s_1 - 0) \). In addition, it increases the possibility of convergence towards some other encoded pattern that has \( s_1 = 0 \) in the first location. Thus zero setting or switch, none can help in barring \( s_1 \) from contributing the error. Thus, the scalar product rule of
synaptic efficacy is computational non-optimal when such meta information is available. MP-neurons, and thus the collective network built upon them, does not have any mechanism to be robust against missing elements.

The above discussion reveals the inability of the current McCulloch cells, and the networks built upon such cells, to cope with the query-modulated attention over the stimulus pattern space. It also provides strong indication that simple architectural modification cannot solve this non-optimality of the cell. In general, it exposes the limits and nature of "robustness" that is generally endowed to an ANN.

Finally, the above discussions will be summarized in the form of identifying the precise causes that make current neuro-computing incapable of supporting RCA search.

2.5.1 Source of Limitations

2.5.1.1 Analog Summation Process

The first problem arises from the algebraic summation process itself. An analog summation process requires all the components to be present for its reconstruction. Death of even a single member in such an operation is likely to critically hurt the outcome for finite size of the member population. Intuitively it becomes difficult to imagine how an analog process so fundamentally dependent on summation can be robust in the face of fragility of its components.

2.5.1.2 Linear Representation

The second, but more fundamental problem is the representation. Obviously, the difficulty is to convey to the network the notion of "don't-care". Can we have a representation for 'don't-care'? It is clear from the previous discussion that if s's are variable enumerated by real numbers in the finite real interval I=[1,0], the use of a real value $d$ in this interval to convey the notion of "don't-care" to the cell can not purport the intended action. Such a
strategy merely introduces uneven bias towards other attractors in proportion to the distance of their elements at that position from \( d \). Such a solution favors some attractors over others not because of pattern similarity but because of representational proximity. Thus, current neuro-computing does not have a representation for 'don’t-care'.

### 2.6 Solution Approach

**Question:** Is there any simple intuitive solution?

If one looks deeper into the analysis presented in the earlier section it can be shown, that under certain restriction on the transfer function, a bipolar (or binary) network, may be able to avoid being partial. If \( s \) is restricted to only two values \([a,b]\), only then the concept of 'don’t-care' can be represented by a real number placing it equidistant from both the possible values, and thus avoiding the risk of introducing unwanted bias. For \([1,-1]\) bipolar representation \( d=0 \), and for \([0,1]\) binary representation \( d=.5 \). However How can one construct a representation, without bias when \( s \) is continuous in \([1,0]\)?

**Approach:** This dissertation proposes a scheme that solves this riddle of representation and of weighted summation. A representation formalism is shown where indeed it is possible to find a point equidistant from any number of attractors. And demonstrates a new synaptic transmission rule based on trigonometric weighted average which can avoid the fragility of traditional weighted sum. In fact this new rule is a natural extension of this new representation.

In chapter 4, section-4.3 explains this new representation and Section-4.4 explains this new synaptic transmission rule. It shows how this new computing paradigm extends beyond the capability of current neuro-computing by supporting RCA and the overall process, but otherwise retains all the attractiveness of MP-neurons as a fascinating basis for parallel and distributed computing.
2.7 Chapter Summary

This chapter has reviewed the learning and retrieval mechanisms of existing associative memories. It has also identified three forms of retrieval based on distribution of attention namely type-U, type-B, and type-A. It has been demonstrated that it is not possible to perform either of the type-B or type-A retrieval which requires changeable attention with the general representational and computational framework of conventional artificial neurocomputing.

This chapter has also explained the non-optimality of conventional scalar product rule when added meta knowledge is available about the location of errors in the stimulus pattern. The linear scalar representation in this scalar product rule is unable to represent the variation of significance of the constituent measurements of a sensory stimulus.
CHAPTER 3

CONTENT BASED SEARCH IN IMAGE ARCHIVE

3.1 Introduction

The objective of this chapter is to briefly survey the existing approaches and current research trends for content-based search in image database (IDB). It also identifies the principal difficulties and limitations faced by these current approaches. This dissertation demonstrates how the combination of associative computing, and the ability to retrieve memory content with changeable attention can provide a new and direct approach for content based search into image databases (IDB).

The chapter is organized as follows. First, section-3.2 describes the conceptual and physical nature of image information. Then section-3.3 provides a brief overview of current image database search systems that has been proposed to manage image information. Section 3.4 identifies the problem of synergistic modeling [KhYu94d] between the encoder and user that limits the capabilities of current model approaches. Finally, section 3.5 introduces the holographic associative computing approach that can provide an efficient mechanism to perform direct content based search into image archive, and bypass the difficulty of finite synergistic and coherent modeling faced by current approaches.

In this chapter, also a brief review of various procedural methods to perform direct search has been presented in section-3.6. The objective is to show that there is no efficient method for performing direct search in large image databases by conventional procedural means.
3.2 Image Information

3.2.1 Conceptual Structure of Information

*Amorphousness of image information:* One of the principal reasons that images are difficult to manage is probably that the very reality of the world (manifested in an image) lacks any rigid structure in terms of well-defined concepts. Any image is merely a projection of the real world. Imposition of "man-made" structure tends to confine the expression of the image content.

The conceptual information contained in an image tends to surpass any attempt to contain itself in a rigid and finite object-oriented language. Transcendental nature of image information has always baffled researchers in informatics. Interestingly, even in the early days of modern computing, John Von Neumann contemplated the difficulty of representing visual objects completely and unambiguously in words (consequently in logical symbols). He remarked [Neum58]:

"It is not at all certain that ..[a visual object] might not constitute the simplest description of itself..[Further,] it is .. not at all unlikely that it is futile to look for a precise logical concept, that is, for a precise verbal description [say] of "visual analogy". It is possible that the connection pattern of the visual brain itself is the simplest logical expression or definition of this principle."

The process involved with managing and understanding formless natural images are substantially different from (and more difficult than) that involved with crisp symbolic information. Images also come with various flavors of abstractions. Graphic images are somewhat more crystalized. Concepts and objects become less well defined and less compact as it moves from graphics to natural images.
Difficult of Conventional Database Technology: The success of conventional database technology in managing symbolic information over the last 30 years has lead the researchers to apply it to visual information also.

However the problem is that current database technology is strictly structured. It's representation is almost tabular. Only recently symbolic databases are shifting towards object oriented approaches, which allows at least some flexibility. However when it comes to the management of image information, even object oriented approaches fall far short of the flexibility that is required. After two decades of research and development, some modest progress in structure oriented approaches for image management have been made [Kato92, Chan92, ChKu81, ChFu81]. However some innovative technique to deal with the formlessness of images and the amorphousness of pixels is much needed.

After decades of direct trials with object based approaches, the limitations inherent in them are becoming increasingly felt. The very recent writings of leading researchers in this field reflect growing realization about the fundamental limitations of the current model based techniques [Chan92, Jain93, GrMe89]. However any solution is yet to be proposed. The notable exceptions are the very recently proposed user profiling methods (based on user's subjective feedback), which in some respect attempt to overcome the problems of main-stream model based techniques. However as it will be shown that they also do not solve the problem.

Besides representation, a second, and probably more profound limitation of the current database technology arises from the fact that its entire reasoning process is based on classical bi-variate system of symbolic logic. But just like the structured representation, such crisp logic seems to be severely inadequate to deal with the inherent amorphousness of image information.
This research incidentally shows how holographic computing technique can come to the aid of image information management. On one hand, this new technique offers an alternate choice other than using such structured representation. On the other hand, it allows the smooth integration of multi-variate (fuzzy) logic based inference. It inherently copes better with the shapelessness of image information compared to conventional bi-valued logic.

**LEVELS OF IMAGE INFORMATION**

![Image](Fig-3.1 Physical Image Information)

### 3.2.2 Physical Structure of Information

Information about any image can be of three physical types (Fig-3.1): (a) content (raw image), (b) context (register information), and (c) condensed (symbolic model). Generally along with the raw images called content, come labels such as name, location, time, etc. These pieces of information will be referred to as context. Contextual information is symbolic. Current data base technology maintains a condensed and symbolic description of an image. It will be referred to as condensed model.
The image management systems, which are currently matured and available commercially, mostly deal with the symbolic and structured information of the context registrar (such as PACs used in hospitals, Multimedia systems, Macro Mind Director, Author Ware, etc.).

### 3.2.3 Principal Research Trends

Search based on only registered context information is severely inadequate for intelligent image database management. The principal focus of current research is to provide access based on the content of the image. However, mainstream effort for content-based search does not directly deal with the raw image. Rather, it has been directed towards searching into a condensed intermediate model (Fig-3.1) which contains an annotated 'meaning' of the image. One of the impending reasons perhaps is to take advantage of the readily available symbolic database technology that has matured over the last 30 years. Besides, currently there is no efficient mechanism to directly search large image databases.

In these approaches, various object-oriented symbolic data structures are used for filing the meaning and various conventional database techniques are used to search in these crisp intermediate descriptions of images. IMAID [ChFu80], QBIC [Niba93], IIDS [ChYD88], PICDMS [JoCa88], IDB [TuPr91], are just some of the bright examples from the long list of such approaches which evolved from this pursuit.

However, the very basic step of this approach, the extraction of crisp 'meaning' from raw image has remained a formidable task. Despite some moderate successes, mainly those of graphical nature; in general the pre-imposition of a subjective structure through such 'meaning' extraction process appears to be a formidable limitation.

Only a few of the relatively earlier researches have attempted to deal directly with the content of the raw image (IMAID-ARES and GRAIN) [ChFu80, ChFu81]. A direct search
is not susceptible to the problems associated with intermediate modelling, but requires massive computations. Due to the lack of efficient algorithm for direct search (either in software or in hardware), mainstream attention reverted back shortly thereafter.

This dissertation demonstrates how the new technique of holographic associative computing can be used to develop an efficient direct content based search mechanism into image archive.

3.3 Review of Current Database Systems

In this section, various techniques used by currently proposed research databases will be investigated.

**Basic model:** Fig-3.2 provides a generalized schematic of the existing approaches for the image information storage and query. All approaches have two distinct stages; encoding and decoding.

![Fig-3.2 Model Based Approach](image-url)

---

60
In the encoding stage, each of the physical images is interpreted into an intermediate representation as designed by the meta-model. The interpretation involves encoder's (also referred as interpreter) knowledge and the physical image itself. The objective of this stage is to develop a condensed, efficient and crisp description of the image content.

In the decoding stage, the mental expectation of the interrogator (also referred to as user) is transformed into a template through interrogator’s knowledge. Interrogator also generally blends a constraint language with representation of expectation to generate one or more templates. The search is performed by matching this template with the stored models.

The design of an image database however requires another crucial step; meta-modeling (which is often trivial in symbolic databases). Meta-model dictates the boundaries within which the encoding and decoding process operates. The upper box in Fig-3.2 shows the design issues that must be resolved before any encoding or decoding process can be initiated. Various database systems which have been proposed, principally vary in the methods that are used to handle these issues.

This survey investigates specifically (i) meta-modeling, (ii) information representation, (iii) image encoding, and (iv) query process of the major image database systems that have been proposed over the last decade. Appendix-B contains case reviews of 11 representative image database systems. Table-3.1 presents a summary of the study. This section directly focuses on the specific techniques that have been proposed to manage these stages of image databases.

Classification of approaches: The principal objectives in the current research with image database are twofold.

**Question 1:** how completely information can be represented, reasoned and queried?

**Question 2:** how the human involvement can be reduced or eliminated from interpretation and reasoning stages?
Together they represent more of a dichotomy than objectives in cohesion. As a consequence current approaches are divided into two types according to their emphasis; mainly (a) meaning oriented, (b) automation oriented.

Meaning oriented approach concentrates on the representational depth and completeness of meta-model. If necessary it assumes substantial and sophisticated human involvement. In the contrary, automation oriented approach emphasizes the automation by substituting humans with computer programs, and if necessary compromises the quality of information.

Recently a third trend has emerged which tries to emphasize user's expectation over any other consideration, and can be called as user profiling approach. This new trend tries to use user's viewpoint as the feedback to obtain a connection between his query and expectation.

Algorithmically the archiving process requires (a) detection/annotation of concepts, (b) determination of class and structures in detected concepts, (c) construction of a knowledge structure to store a scene knowledge and, (d) archiving of the scene knowledge into database, and finally (e) query and search. Now, how these stages are handled in the three approaches will be examined.

3.3.1 Meaning Oriented Approach

Proposed systems in meaning oriented approaches can be divided into two main types; free-text based and semantic-model based.

Free-text: These approaches store a set of key words with quantum descriptions of the scene. A slight variant of this model stores full sentences. However the retrieval is based on keyword search. The positive aspects of this approach are; (a) the system is easier to use for someone who is familiar with the range and type of keywords used by the encoder, (b)
free-formatted text offers greater flexibility. There are no bounds of concepts (and their level of abstraction) that can be stored and retrieved. The drawbacks are; (a) needs agreement between the vocabulary of the encoder and user. Some system proposes judicious cross-indexing to alleviate some of the problem. (b) relations connecting two specific concepts can not be handled.

**Semantic network:** These approaches capture concepts, and also the specific relations connecting these concepts. As a first step, it divided concepts into four classes, (a) objects, (b) entities, (d) relations, and (d) attributes. It stores these concepts in various conceptual graphs (known as EAR models). Systems specialize on the nature of information it stores.

Such as, some systems encode the physical composition of the scene objects, with compositional relations such as such as "part-of" or "composed-of", etc. [HoHs92]. Some models encode abstraction relations such as 'generalization', 'instance-of' concepts through semantic network arcs [BeZi92, IrOX92, YaSa93], Some models store spatial relations. The need to improve geographical information systems, such as satellite imagery, map data, etc. has sparked considerable research in encoding spatial relations among the objects in a scene [Chan87, LeHS90]. Quad tree, B-tree, R-tree, 2D-string are some of the structures proposed for encoding spatial relations.

Another more ambitious and knowledge intensive trend targets managing event information. An interesting example of this group is the Birbeck system developed by Hibler et. al [HLMM92] developed in UK. From a given text sentence, it recognizes noun, adjective and verb respectively as the entity, attribute and relation in their EAR model. A human is needed to interpret a scene through English like sentences. A parser reads these specially formatted sentences and stored the concept types in a semantic net. EMIR [HaMo92] system is another example of a more complex and broader image representation formalism.
3.3.2 Automation Oriented Approach

The general approach is to decompose the overall concept space into fewer elementary concepts (also referred as features) for economy and emphasize the automatic detection of these elementary one. In contrast to knowledge based approach, it uses mathematically (or algorithmically) quantifiable features. These features are detected by pre-defined filters. Detected features are assimilated with the help of pre-encoded composite object models to recursively detect further higher level objects that ultimately bear some kind of semantic meaning. Automation oriented approaches can be classified according to the type of features they use.

**Global vs. local features:** Global features are derived from the entire object such as, shape, area, perimeter, a set of rectangular or triangular cover, moments, etc. [Jaga91, HoHs92]. Local features are composed of only some important segments of object, such as line, object contour, points of maximum curvature change, [GrJi92], dots on minimal rectangle [YaSa94], etc. Local features can withstand partial loss of object components, generally search is fast, but is susceptible to major errors in special situations. Moments can be used for shift, rotation or deformation, invariant compositional representation of objects [Hu1962, HoHs92].

**Meaningful vs. efficient features:** Features like moment, area, etc. have mathematical meaning attached to it. Instead of using such 'meaningful' features, some methods use 'efficient' features like Fourier coefficients, fractal coefficients, wavelet functions, or principal components. The objective is to optimize some performance measure, such as, to maximize distinguishability among the images, or to minimize the representation space through orthogonalization of the feature space. Generally these features adapt to the images.
Features are assimilated to detect complex objects and concepts. Computers require object models to determine what combination of which features will make an object. Obtaining such object-models as a necessary part of meta-modeling is usually difficult and tedious. If humans are to determine these models, then it becomes cognitively easier if the basic features have some conceptual 'meaning' (even when they are mathematical).

Detection techniques: Besides using specialized filters, the conventional neural computing has also been used for feature detection as well as for automatic object model determination from examples [CKLY93]. Generally rote learning networks have been used to adaptively construct filters for predefined fixed features. On the other hand, autonomous learning networks have been used to construct efficient features [RiSt93]. However such use of neural network are distinct from the proposed approach. Here neural networks are used only as trainable filters to model features or objects from examples, rather than from complicated analytic definitions.

3.3.3 User Profile Based Approach

Very recently Oommen & Fothergill from Canada[OoFo93] and Crehange et. al from France [CrSm92] have independently proposed methods which can be classified as a new and interesting approach for content based search. It eliminates the process of image annotation by a dedicated encoder, either human or machine. Instead it applies users’ own perspective to build a model of image content.

The approach is explained in Fig-3.3. The technique is to adaptively group/classify images to the objects in query by using user’s evaluation of success and failure of the responses. Initially images are randomly assigned to various clusters. During the first query, one image from each cluster (generally the one at the center of cluster) is presented to the user. The user then provides a reward/penalty response for each of these pictures. The
clustering algorithm then places the correct responses to the state of maximum certainty (generally in the same cluster) and moves the incorrect responses to a state of minimum certainty (generally to a different cluster). Gradually the system establishes connection between user’s query and expectation. The principal design issues are identified at the top of Fig-3.3.

As evident, although the approach eliminates and external encoding, or even metamodeling but in reality it shifts the task indirectly to its users. Another, interesting feature of this approach is that it does not require physical access to the image content. Thus, it is close to the meaning oriented annotation method. The connection it establishes between query keywords and the images, can be considered as the condensed description model of the images. However the model of content information built by this approach is entirely subjective because it never physically accesses the raw image information.
3.4 Difficulty of Coherent Modeling

The approaches based on intermediate representation of image information suffers from the difficulty of ensuring synergy between the encoder’s and the inquirer’s framework. The problem can be summarized as below:

(a) **The completeness and sufficiency of meta-model:** In this first level, the encoder has to correctly guess the bounds of users expectations and construct a sufficient framework (meta-model) from which he can satisfy users’ expectations. The difficulty is that there may not be any such framework of finite dimension to describe visual information. Consequently, any pre-modeling runs the risk of being incomplete as the subjective perspective of query changes. For example, to describe a historical scene the encoder may meticulously try to describe all the events portrayed. But, a future query may be on the spatial location of a character. In addition much of the visual information can not be represented using symbols, keywords, or even numerals. For example, there is no convenient language to describe shape, or texture.

(b) **Subjectivity and accuracy of image-model:** Even if we assume that such a framework of representation exists (or both the user and encoder comes to an agreement), in the second level, we run into the problem of subjective evaluation of the encoder during model extraction. A single situation can have large number of valid interpretations. Multiple interpretability is not only associated with the objects but also with relations that connects them. A scene can be analyzed from numerous equally valid viewpoints, resulting in numerous equally valid structured models, even within the bounds of a finite well-defined language. In general, judging the visual contents of an image itself is an imprecise task.

**Question 1:** Can the objectivity of mathematical features help in assuaging the problem of multiple interpretability?
Unfortunately no. In the automation oriented approach, it becomes the problem of synchronization between the user's subjective knowledge and the encoders mathematical model (i.e., the subjective model of the programmer embedded in the encoder).

The mathematical description of foundational features has the advantage that they can be precisely quantified. However these are very restrictive in their capability to model complex objects. For example, it is very difficult to quantify complex concepts like "hill", "grassy field", etc., from a finite set of pre-defined algorithmic features. The geometry of same object (even a strictly mathematical tetrahedron object) can be wildly different when viewed from different angles. On the other hand, the mathematical description of one object can easily match the mathematical description of the other. For example the circular geometry of "moon" is very close the circular geometry of "coin". Thus, systems that use mathematical features, suffer less in terms of accuracy of interpretation level, but suffer considerably more in terms of sufficiency of language.

**Question 2: Does the removal of encoder in user-profiling approach overcome this incoherence?**

Interestingly this new approach can be considered a reaction to the incoherence problem that is deeply associated with the first two approaches. It attempts to remove encoder and requires the interrogator to also perform the functions of the encoder. However it should be noted that it does not eliminate the role of encoder, rather it eliminates the formal encoder and places the user in that role.

This intermediate model representation is the subjective opinion of a collection of previous users. If there is consistency among the subjective knowledge of past users, then this approach can overcome the incoherence problem. However due to the fact that, usually, there are discrepancies among users' subjective interpretations and preferential vocabularies, this approach is also exposed to this fundamental problem of synchronization.
3.5 Holographic Approach

Removal of intermediate representation: In this research, an alternative content-based search approach based on holographic associative computing that avoids intermediate representation is described. It allows the interrogator to translate his/her subjective expectation directly into physical representation of image and search. As a result, the interrogator can directly query into the physical content of an image from his/her own framework of 'meaning', without going through the subjective interpretation of the intermediate agent. In contrast to user profile based approaches, it eliminates not only the formal encoder, but the role of encoder (in this case, played by collective users) altogether.

Fig-3.4 Direct Content Search Approach

Fig-3.4 shows the schematics of the approach. The user communicates with the search mechanism primarily through an example image, and performs a direct search into the database. However how this example image should be interpreted during search, depends
on the subjective 'meaning' structure perceived by the interrogator. The interrogator generally views a deep structure in terms of objects, concepts, and relations in a physical image, and expects match based on these perceived entities\textsuperscript{16}.

In this approach the interrogator uses an interpretive mask on the image to communicate this perceived 'meaning' structure and the resulting expectations to the search mechanism. The search mechanism then constructs physical templates from the example image and the sketched mask, and performs direct search into the image database.

**Isolation of physical representation from meaning:** In the traditional approach, an interpretation structure is imposed and subsequently hard encoded in the intermediate representation. In this approach such structure is used only at the query interface of the decoding stage. Thus, the difference of subjective interpretation between various users (or the very subjectivity of such structures) does not create problems usually associated with other methods. Here the 'physical' part of information is stored separately from the 'meaning'. Only, during query the 'meaning' is attached to a physical template in the form of mask, and finally the search mechanism performs a 'meaning' modulated search.

**Necessary components for this approach:** The approach can be summarized as (a) a flexible query interface that helps the user to translate his subjective expectation into masks, and (b) an efficient search mechanism to perform mask modulated direct search into image database. In this dissertation it will be demonstrated that how holographic associative computing can provide unique support to facilitate both of these stages.

**Representation for computational efficiency:** In the physical implementation level, the approach computes a holographic representation of the raw images. However this is

\textsuperscript{16} Traditional approaches demand a formal encoder to guess this 'meaning' structure before hand, and annotate the images accordingly and (b) store the result of annotation as intermediate model.
fundamentally different from the intermediate representation used by model based techniques. The objective of this intermediate holographic representation is to accomplish computational efficiency. This intermediate holographic representation (or any other representation used by associative memories), helps the system to 'emulate' a direct search (and thus gain efficiency) rather than really performing it like procedural methods. It does not impose any 'meaning' structure in this intermediate representation.

3.6 Other Mechanisms for Direct Search

As evident, direct content based search constitutes the key computational component of holographic approach. Before explaining how the proposed holographic associative computing paradigm provides an effective means to the requirement, in this section, the capabilities of current procedural search techniques will be reviewed.

**Question:** How good are the procedural search algorithms in performing direct search?

Conventional search requires time linearly in proportion to the size and number of images. In this section the complexities of few possible search methods will be briefly summarized. Previously, Minsky and Papert [MiPa72] have computed the capabilities of various search algorithms for exact and best match problem. Following analysis is an extension of their result for type-A and type-B searches.

---

17 The reorganization objectives can be stated as: (a) to maximize the distinguishability among the images with respect to any collection of pixels and pieces thereof; and (b) to speed up the search by arranging the images into a multidimensional ordered set.
3.6.1 A Simple Memory

Let us consider the conventional memory system of Fig-3.5, which is analogous to the associative memory system of (Fig-2.3). Here, the learning algorithm $A_{\text{learn}}$ is substituted by the filing algorithm $A_{\text{file}}$, and the retrieve algorithm $A_{\text{retrieve}}$ is substituted by the finding algorithm $A_{\text{find}}$.

![Model of Conventional Memory](image)

Let it be given a memory of $M$ bits. Before, seeing the data set the storage algorithm $A_{\text{file}}$ and the retrieval algorithm $A_{\text{find}}$ should be constructed with the following conditions:

1. $A_{\text{file}}$, given the binary data set with $p=2^n$ random words of length $n$, fills the memory bits, neither the data set nor $A_{\text{file}}$ is used again. Nor is $A_{\text{find}}$ allowed to get any information about what $A_{\text{file}}$ did except by inspecting the content of $M$.

2. $A_{\text{find}}$ is then given a random word $S^Q$ and asked to find the match, we would like to know how many bits in $M$ it has to access and compare before it can provide the desired answer.

3. The goal is to optimize the design of $A_{\text{file}}$ and $A_{\text{find}}$ to minimize the number of access and compare in the $A_{\text{find}}$ computation, averaged over all possible query words $SQ$.

3.6.2 Types of Searches

All three search types (type-U, type-B, and type-A) defined in previous section, have two forms; (i) exact (EM), or (ii) best (BM). These are:
Question 1 (exact match/EM): Given a random pattern $S^Q$, is there a word $S^T$ in the data set for which $D(S^Q, S^T) = 0$?

Question 2 (best match/BM): Given a random pattern $S^Q$, exhibit the word $S^T$ in the data set for which $D(S^Q, S^T) = \min_{\mu} \left[ \sum_{i}^{N} \text{dist}(s_i^T, s_i^\mu) \right]$.

There are various algorithms for procedural search [Sedg91]. It is well known that the cost of procedural search is dependent on the amount of memory. Below, strategies for five different memory categories\(^{18}\) are described.

3.6.2.1 CASE 1: Enormous Memory (table lookup)

Consider: $M \geq 2^n$

Let us consider an enormous amount of memory is available. In which case it is possible to construct a table that should have one bit for every conceivable value of word $S^\mu$, which has $n$ digits. Let $m_i$ be the $i^{th}$ bit in the memory, then there is a bit $m_\mu$ for every possible word $S^\mu$. It can be defined:

$A_{\text{file}}$: set $m_\mu$ to 1 if $S^\mu$ is in the data set.

$A_{\text{find}}$: $S^Q$ is in the data set if $m_Q$.

Thus, only one memory access is needed to answer question 1.

3.6.2.2 CASE 2: Sizable Memory (hashing)

Consider: $M \geq 2.n.2^n$

Still now a substantial margin of extra memory is available (about twice what is necessary to store the data set). In such a situation, a hash coding is one of the fastest strategy available. Let us consider the case of redundancy of 2 hash coding.

\(^{18}\) Minsky and Papert [MiPa72] also selected these algorithms as efficient ones in their respective memory categories.
Consider the hashing function $H(S,j)$ that can map any of the possible $2^b$ patterns $n$ bit patterns into an $(a+1)$ bit word. It is also assumed that this hashing function maps possible $2^b$ words more or less uniformly over $2^{a+1}$ output spaces. The corresponding memory procedures are:

$A_{\text{file}}$: Compute $H(S^a, 1)$ for each word $S^a$ in data set. If the register at that address is empty put $S^a$ in it, otherwise compute $H(S^a, 2), H(S^a, j)$, and if this is empty put the word in it, so on, until an unoccupied register is found.

$A_{\text{find}}$: Compute $H(S^O, 1)$. If this register contains $S^O$, then it is in data set, if not empty then the word is not in the data set. But, some other word is in it, then compute $H(S^O, 2), ..., H(S^O, j)$, and so on, until the word or an empty register is found.

It can be shown (the derivation will not be repeated, see [MiPa72]) that this method will on the average access 4 bits to answer question 1.

3.6.2.3 CASE 3: Ordered Search

Consider: $M = n.2^a$

Here, the memory size is just enough to store the ordered data set. Those will be stored according to a numerical order. Since, the original data does not come with any specific order between the $n$-digits, any order can be uses. Lets say, the $n$th digit is most significant digit and the rightmost one is the least significant. Then:

$A_{\text{file}}$: store the words of the data set in ascending numerical order.

$A_{\text{find}}$: perform a binary search to see which half might contain $S^O$, then which quartile, and so on.

This procedure requires at most $a = \log_2(2^a)$ inspections of $n$-bit words, that is in total $n.a$ bit inspection.
3.6.2.4 CASE 4: Exhaustive Search

Consider: \( M = (n-a).2^a \)

Here just memory is just enough memory to store an ordered data set. Thus:

\( A_{\text{me}}: \) First put the words of the data set in numerical order. Then compute their successive differences. These will take about \((n-a)\) bits each\(^{19}\). It will take about \((n-a).2^a\) bits total.

\( A_{\text{find}}: \) Begin from the top, add up successive differences in memory until sum equals or exceeds \( S^Q \). If equality occurs, then \( S^Q \) is in data set.

On the average it will search through half of the memory to answer question 1, which is \( \sim \frac{1}{2}(n-a)2^a \).

3.6.2.5 CASE 5: Data Table

Consider: \( M = n.2^a \)

This is almost like exhaustive search except, no coding is involved. Thus:

\( A_{\text{me}}: \) First put the words of the data set in numerical order.

\( A_{\text{find}}: \) Begin from the top, until sum equals or exceeds \( S^Q \). If equality occurs, then \( S^Q \) is in data set.

On the average it will search through half of the memory to answer question 1, which is \( \sim \frac{1}{2}n2^a \).

3.6.2.6 CASE 6: Inadequate Memory

Consider: \( M < (n-a).2^a \)

There is no conceivable algorithm to solve for question 1.

\(^{19}\) Any standard information theoretic encoding scheme can be adopted. Such as Huffman coding.
3.6.3 Performance for Type-U, Type-B, and Type-A search

Misnky & Papert investigated [MiPa72] the exact match as well as best match versions of the query problem. As it is evident by now, the matching criterion they have considered is of type-U. Table 3.1 provides the performance of the search strategies for type-B and type-A search. For, each types both exact match (EM) and best match (BM) versions have been analyzed.

Table-3.1 Complexities of Procedural Search Methods

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Memory</th>
<th>Type-U (exact)</th>
<th>Type-U (best)</th>
<th>Type-B &amp; A (exact)</th>
<th>Type-B &amp; A (best)</th>
<th>Type-B (exact)</th>
<th>Type-B (best)</th>
<th>Type-A (best)</th>
</tr>
</thead>
<tbody>
<tr>
<td>table</td>
<td>-</td>
<td>1</td>
<td>( n )</td>
<td>( \frac{1}{2}^{(n-c)} )</td>
<td>( n )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>lookup</td>
<td>( \geq 2^n )</td>
<td>( M = 2^n )</td>
<td>( M = n.2^n )</td>
<td>( M = 2^n )</td>
<td>( M = n \sum_{k=1}^{n!} \frac{1}{(n-k)!} )</td>
<td>( M = \infty )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hashing</td>
<td>( 2.n.2^n )</td>
<td>( 4 + \epsilon )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ordered set</td>
<td>( n.2^n )</td>
<td>( a.n )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>data-table</td>
<td>( n.2^n )</td>
<td>( \frac{1}{2}n.2^n )</td>
<td>( n.2^n )</td>
<td>( \frac{1}{2}c.2^n )</td>
<td>( n.2^n )</td>
<td>( c.2^n )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exhaustive</td>
<td>( (n-a).2^n )</td>
<td>( \frac{1}{2}(n-a).2^n )</td>
<td>( (n-a).2^n )</td>
<td>( \frac{1}{2}(n-a).2^n )</td>
<td>( (n-a).2^n )</td>
<td>( (n-a).2^n )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>impossible</td>
<td>( &lt;(b-a)2^n )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

The cost of type-U/EM is already shown with the description of the strategies. They can be viewed as a reference.

Now let us consider the type-U/BM problem first. As shown in the table, for this type of search, the table-look up method require about \( n \) times more memory to store the correct answer for a possible memory word. For best match, hashing strategy can not be applied at
all. In fact, hashing destroys the neighborhood or nearness information. More striking is the
failure of ordering method, which allows logarithmic time search for exact match, and seems
to be very efficient. However due to the lack of order among the $n$-digits of patterns, the
patterns themselves can not be sorted, thus the strategy fails. Relatively, the data-table and
exhaustive search suffer less. But each of these methods requires mandatory inspection of
the words all the way to the end (instead of expecting the match approximately halfway).

Now, let us look into type-B and type-A search. Let us consider that in type-B search,
the attention set includes only $c$-digits out of the $n$-digits of a word. The exact version of
these searches can be considered equivalent. In the table-lookup strategy, since, now $(n-c)$
digits are ignored, therefore it involves looking into $2^{(n-c)}$ locations in the table that shares
those $c$ bits. However if match exists, on the average it is expected to appear somewhere
halfway.

The hashing and ordering strategies fail both for type-A and type-B precisely for the
same reason as it fails for type-U/BM search. Only viable search strategy for this case seems
to be the data set search. It needs the inspection of the $c$ digits of each word.

For type-B/BM, the lookup-table becomes enormous, because of the possible ranges
of the query. When it comes to type-A/BM, it becomes impossible to construct any lookup
table that can store answers corresponding to all the possible queries.

The analysis summarized in Table-3.1 clearly shows the limitation of current strategies
in handling type-A (or type-B) search. At the end, exhaustive search, (either in raw or in
lossless compressed form) remains as only viable strategy.

---

20 For type-A search, $c$ should be considered as the number of digits which has attention $\lambda_i \neq 0$
What is the reason behind the brittleness of conventional logarithmic? The efficiency of these algorithms is derived from the fact that the words can be ordered according to some linear measure. However, words cannot be sorted in any meaningful way if the $n$-digits, those make the words, themselves are not ordered. In the type-U exact match case, an arbitrary discriminant order of the digits is assumed (or imposed). There is no problem so far if it is exact match. Any ordering converges to same solution for such exact match. However, once the problem becomes that of best match, all the digits require equal importance (errors at all digits bear equal weight). So the imposed ordering becomes meaningless. The problem grows more out of hand for type-A and type-B search. These searches not only defy any pre-imposed order but impose its own new order every time they perform a query. Thus, logarithmic search strategy falls apart for all these cases (even when they are only looking for exact match).

3.6.4 Associative Search vs. Conventional Search

The distributed associative search technique based on artificial neural computing, is fundamentally different from the conventional search. It will be relevant to look into two important distinctions:

**Retrieval Accuracy:** Artificial associative memories are inherently best match machines. Exact match is guaranteed only under restricted situation. In contrast, table-1 clearly shows that conventional algorithmic search techniques are primarily exact match machine.

**Retrieval Speed:** Retrieval in associative memories is generally very fast because, these memories do not actually search, rather compute the matching pattern. It will be demonstrated, that MHAC is logarithmic even with type-A or type-B search. This is roughly true for conventional AAMs except that these cannot support type-A or type-B search. The
efficiency of any search mechanism lies in its ability to obtain some form of order in the input data. As a search method, the efficiency of associative memories is derived by internal clustering, which can be considered as approximation of multidimensional partial orders. This adaptive organization characteristics can be called as an associative ordering process.

The information in an image is representationally large but sparse in content. Thus, from both of these points of differences, associative search better suits image application. However the inability of previous associative memories to perform type-A and B search becomes an insurmountable obstacle to their usefulness in IDB search.

3.7 Application of MHAC in Direct Image Database Search

Type-A search: As evident, one of the key element of the proposed approach is the ability to perform 'meaning' modulated direct search very efficiently. Computationally, this requirement translates to a mechanism which can search with variable significance attached to the pixel elements [KhYu94c, KhYu94e]. The significance of the pixel clusters should be allowed to vary dynamically during query sessions. In fact, this is the type-A search as defined before.

Suitability of MHAC: What is needed is a mechanism, that can support type-A search as well as it is efficient. Current models of AAMs, although efficient, cannot support type-A search, and thus can not perform meaning modulated associative recall. On the other hand, traditional procedural approaches can support type-A retrieval (the implementation is fairly trivial), but are prohibitively expensive and thus are impractical.
The unique ability of holographic computing to perform both type-A search as well as retain the efficiency of conventional associative computing for the first time fills up this critical 'gap' of conventional search techniques. In this dissertation, it will be shown how a space and time efficient organization of the raw image pixels through Holographic computing can emulate the direct search.

3.8 Chapter Summary

Direct search into image databases can help eliminate the problem of incoherent modeling. A retrieval method based on direct content based search can compliment existing intermediate model based search techniques. However there is no conventional technique which can support such direct search efficiently.

The conventional search techniques have the ability to perform complex searches required for intelligent image management. However these techniques require massive amount of computation. Recently, a significant effort has been diverted to implement these conventional strategies into hardware [OgYN85, Robi89, Gros92]. Despite significant performance improvement these hardware systems are still limited in the problem size they can handle.

On the other hand, the emerging artificial associative computing demonstrates significant promise as a technique for extremely fast best match search machine. However till now, these models lacked the critical ability to perform type-A and B search, and thus failed to provide any significant help for intelligent image database systems.

The capability of holographic computing to perform attention modulated associative search can bridge this void and can be applied to develop an efficient direct content-based search mechanism for image information.
4.1 Introduction

This chapter presents the theory of the multidimensional complex associative computing (MHAC) paradigm, that introduces the ability to perform retrieval with changeable attention (RCA) in parallel and distributed associative computing. The proposed solution is founded on a new bi-modal definition of information that encompasses the notion of imperfection at the very representational level.

This chapter is organized as follows. Section 4.2 first explains the new and generalized notion of information in the context of changeable attention. This section also outlines the fundamental behavioral expectations from the new kind of generalized memory system that incorporates attention without being specific about any of computational embodiment. Then section 4.3 presents the proposed computational representation based on multidimensional complex numbers (MCN) in a hyperspherical space. It shows how MCN can be used to represent and process the bi-modal notion of information. Section 4.4 describes the learning and retrieval mechanisms. Then sections 4.5-4.7 confirm how the proposed mechanisms together can satisfy the behavioral expectations of associative recollection with changeable attention (RCA) outlined in section 4.2. Finally, section 4.8 explains the key distinguishing attribute of this new computing paradigm that enables it to overcome the limitations of existing associative computing stated in section 2.5.
4.2 Information with Modifiable Significance

A general stimulus pattern is a collection of elements, which can be considered as a symbolic vector:

\[ S = \{ s_1, s_2, \ldots, s_n \}. \]

Each of the individual elements represents a piece of information, or a feature, collectively which defines the pattern. Each of these elements bears implicit positional significance. The values of these elements correspond to a measurement obtained by some physical sensor. These measurements are generally enumerated on a real number interval. While constructing an instance of a pattern, the conventional representation approach is to fill these element slots directly with real numbers obtained from corresponding sensory measurements.

The proposed formalism of multidimensional holographic associative computing (MHAC) assumes the trust in information as inherently of undulating nature. Where the elements are susceptible to distortion error, loss, or even purposeful neglection. Therefore, in contrast to conventional approach, MHAC approach includes the meta-knowledge about the state of each given piece of information (measurement) as an integral part of its basic notion of information. Each piece of information is modeled as a bi-modal pair \((\alpha, \beta)\). Where, \(\alpha\) represents the measurement of the information elements and \(\beta\) represents the meta-knowledge associated with the measurement. Thus, an element of stimulus suit is defined as:

\[ s_k = \{ \alpha_k, \beta_k \}. \]

Here meta-knowledge \(\beta\) is interpreted as additional knowledge about the state of significance of the knowledge contained in associated element set \(\alpha\). This meta knowledge will also be referred as significance of information.
The above interpretation, in the context of a general memory (irrespective of its implementation mechanism) that computes on imperfect knowledge, generates some specific expectations about the operational behavior of this quantity. These are summarized below:

**Inflow Expectation (1):** The matching criterion should put more importance to a piece of information that is attributed with high degree of $\beta$ than to a piece attributed with low $\beta$ in the query. The expectation can be stated as a matching criterion:

$$D\left(s^q, s'^r, B\right) = \sum_{i}^{N} \beta_i \text{dist}(\alpha_i^q, \alpha_i^r)$$  \(\ldots(4.1)\)

In the simple memory framework stated earlier, this expectation corresponds to type-A RCA match.

**Outflow Expectation (2):** If $\alpha$-suit of query demonstrates high degree of resemblance to the $\alpha$-suit of a priory encoded stimulus pattern, then memory should retrieve the associated $\alpha$-suit with higher degree of accuracy and high degree of $\beta$. On the other hand, if it does not then it should generate a response with low degree of $\beta$, as detailed in table-1.

<table>
<thead>
<tr>
<th>Query</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{query}$</td>
<td>$\alpha_{query}$</td>
</tr>
<tr>
<td>HIGH</td>
<td>CLOSE</td>
</tr>
<tr>
<td>LOW</td>
<td>CLOSE</td>
</tr>
<tr>
<td>HIGH</td>
<td>NOT-CLOSE</td>
</tr>
<tr>
<td>LOW</td>
<td>NOT-CLOSE</td>
</tr>
</tbody>
</table>

Table-4.1 Expectations

Expectation (1) relates to the inward communication of the meta-knowledge into the memory system. An external querying system supplies the stimulus elements and the additional significance level of each stimulus element. The generalized memory is expected to utilize this additional meta-knowledge during a query to logically enhance its search.
Expectation (2) relates to the outward communication of the meta-knowledge from the memory system. Here, the external querying body is communicated back by the memory system not only the retrieved contents but also meta-knowledge about the status of the retrieved content. Both of the transfers are essential in the context of imperfect knowledge transaction.

The above expectations can also be considered as the behavioral definition of the memory system which incorporates possibility of imperfection in given measurements. It does not specify, how to achieve the expected relations between the two.

As demonstrated in the chapter-2, none of the existing analog models of parallel and distributed computation (which includes existing AAMs) can realize the above operational characteristics within the framework of their representation and transfer function formalisms. The question that now immediately follows is that:

**Question:** How one can obtain a computational representation and mechanism which can satisfy the above expectations?

This research sets out to realize the above operational expectations in an associative memory mechanism. A computational framework is demonstrated that realizes above operational expectations in the context of an associative memory scheme. Next few subsections provide the answer. First, the representation, and then the mechanism is explained.

### 4.3 Computational Representation

Traditional systems use only real numbers as computational representation of information elements. However in this approach because of the bi-variate nature of the input information, a new computational representation is required.
In this approach, each piece of information is mapped onto a multidimensional complex number (MCN). Equivalently, each $\alpha_k$ is mapped onto a set of phase elements $\theta_{j,k}$ in the range of $\pi \geq \theta_{j,k} \geq \pi$ through a mapping transformation $m^{+\alpha}(x)$, and corresponding meta information $\beta_k$ becomes its magnitude $\lambda_k$ through another mapping transform $m^{+\beta}(x)$.

$$s_k = (\alpha_k, \beta_k) \Rightarrow \lambda_k e^{\left[\sum_{j=0}^{d-1} i \theta_{i,j}\right]} \quad \ldots \quad (4.2)$$

where, each $s(\lambda_k, \theta_{1,k}, \theta_{2,k}, \ldots \theta_{d-1,k})$ is a vector that is expanded inside a unit sphere in a $d$-dimensional spherical space. Each $\theta_{j,k}$ is the spherical projection (or phase component) of the vector along the dimension $i_j$.

The magnitude of this multidimensional vector corresponds the meta-knowledge $\lambda_k$ of information. The set $\{\theta_{1,k}, \theta_{2,k}, \ldots, \theta_{d,k}\}$ represents the measurement of information $\alpha_k$. Symmetrically, a holographic representation can be reverted back through inverse transforms $m^{-\alpha}(x)$ and $m^{-\beta}(x)$.

**Fig-4.1 Points on Hyperspherical Surface**

**Mapping of measurements:** External stimulus measurements are obtained within a real number interval $I[a, b]$. Such a real linear interval of measurements $\alpha_k$ are mapped onto
a point in the surface of a hypersphere with a suitable mapping function \( m^\alpha_0 \). Mapping transform \( m^\alpha_0() \) converts the external real valued domain into internal phase domain of holographic computation.

A class of functions can be used as the mapping transform. The function should be single valued and continuous. However for discrete inputs, continuity is required only at the defined points. A second desirable characteristic of the mapping transform is that it should maximize the symmetry at the phase domain. The issue of symmetry will be further discussed in section-6.3.1.

**Mapping of significance:** Each piece or subset of information elements is assigned a positive number as an indication of its importance in the subsequent processing. As it will be evident later on, the contribution of each information element in both encoding and decoding operations are influenced by this assignment. For example, if two elements are assigned the same magnitude, in all subsequent processing they will contribute equally to the decision stages of this computing. On the other hand, an element with 0.0 magnitude will have no effect on the outcome of the computing. The upper bound of the magnitude can be normalized to 1.0 through the meta-knowledge mapping transform \( m^{\beta}() \). Such normalization establishes a probabilistic interpretation of certain aspects of this representation.

With the normalized representation, if all elements are assigned equal magnitude, this new paradigm will be functionally equivalent to the conventional AAMs. However the opportunity to modify these magnitude values dynamically during query, provides a new capability of selective focusing of this new paradigm, which is the principal theme of this dissertation.
**Combined representation:** Thus, each of the information elements is represented as a vector bounded in the unit multidimensional spherical space. The subsequent operations, such as addition, subtraction, multiplication, etc., are also performed in this multidimensional space. Although, the resulting computation appears analogous to conventional associative computing paradigm, but it demonstrates fundamentally different characteristics than conventional associative computing.

The multidimensional mapping of stimulus element from the external scalar field intensities is performed by some non-linear mapping. In many physical cases, the phase components can be directly obtained from the sensors. For example for color images, the three basic color intensities can be directly translated into three dimensional components resulting in a 4-D representation.

Thus, a stimulus pattern is represented as:

\[ S^u = \begin{bmatrix} \lambda_1^u e^{\left( \sum \frac{1}{j} \phi_{1j} \right)} \\ \lambda_2^u e^{\left( \sum \frac{1}{j} \phi_{2j} \right)} \\ \vdots \\ \lambda_n^u e^{\left( \sum \frac{1}{j} \phi_{nj} \right)} \end{bmatrix} \]

Similarly, the response patterns are also obtained. A similar mapping on the external scalar response field intensities provides the response representation:

\[ R^v = \begin{bmatrix} \gamma_1^v e^{\left( \sum \frac{1}{j} \phi_{1j}^{\prime} \right)} \\ \gamma_2^v e^{\left( \sum \frac{1}{j} \phi_{2j}^{\prime} \right)} \\ \vdots \\ \gamma_n^v e^{\left( \sum \frac{1}{j} \phi_{nj}^{\prime} \right)} \end{bmatrix} \]

Here, the phasor \( \phi \) represents the measurement of the expected associated response pattern elements from the memory system, and \( \gamma \) represents the expected system confidence of the response.

Use of complex number is not a totally new concept in artificial associative computing, at least in 2-dimension. In 1990, Sutherland [Suth90], in his pioneering work presented the first truly holographic associative memory with learning algorithm analogous to correlation...
4.4 Proposed Holographic Computational Model

4.4.1 Encoding

Information is stored in the form of associations. It is expected that a response can be retrieved from any part of the associated stimulus. In the encoding process, the association
between each individual stimulus and its corresponding response is defined in the form of a correlation matrix computed by the inner product of the conjugate transpose of the stimulus and the response vectors. Equation (1) shows the operation:

\[ [X^\mu] = [S^\mu]^T \cdot [R^\mu] \tag{4.3} \]

If the stimulus is a pattern with \( n \) elements and the response is a pattern with \( m \) elements, then \([X]\) is a \( n \times m \) matrix with \( d \)-dimensional complex elements.

The strength of this new paradigm is derived from the fact that an enormous number of such associations can be learned and stored on the same \( n \times m \) space by superimposing the individual correlation matrices onto the same storage elements.

A suit of associations derived from a set of stimulus and corresponding response is stored in the following correlation matrix \( X \). The resulting memory substrate containing the correlation matrix is referred as Holograph.

\[ [X] = \sum_\mu [X^\mu] = \sum_\mu [S^\mu]^T \cdot [R^\mu] \tag{4.4} \]

The above method, allows the associations to be learned simultaneously through a single iteration. Here the learning is not influenced by the previously learned state of the memory. However an enhanced encoding scheme can be used where the learning is directly influenced by priori learned information. Section 4.9 will describe the enhanced learning scheme.

4.4.2 Retrieval

During recall, an excitatory stimulus pattern \( [S^\ast] \) is obtained from the query pattern:

\[ [S^\ast] = \left[ \lambda_1 e^{d_1 i \phi_{s1}}, \lambda_2 e^{\frac{d_2 - 1}{d} i \phi_{s2}}, ..., \lambda_m e^\Sigma j i \phi_{s,n} \right] \]

89
In the event, that this new stimulus resembles closely to a priorly encoded stimulus, then the corresponding response pattern is generated with high magnitude. The decoding operation is performed by computing the inner product of the excitatory stimulus and the correlation matrix X:

\[ [R^e] = \frac{1}{c} [S^e] \cdot [X] \]

where, \[ c = \sum_k \lambda_k \]

In the following sections, first, it is demonstrated how equations (4.3), (4.4) and (4.5) together operates as an associative memory system, like any other type-U system. Next, section-4.6, and 4.7 shows how it can satisfy the behavioral expectations stated in section-4.2. Finally section-4.8 explains the fundamental distinctions of the stated technique that makes this new parallel and distributed computing paradigm capable of supporting type-A & B RCA search.

4.5 Confirmation of Basic Associative Recovery (type-U/BM)

In this section, it will be demonstrated that the above computational mechanism can store patterns and perform content based retrieval from query patterns similar to any of priori learned pattern. First, its recollection capability on the basis of type-U search like other conventional AMs will be demonstrated. Here the match is based on similarity over the entire element space.

By combining, the encoding and decoding operations expressed in (1) and (2), the retrieved association can be decomposed into principal and crosstalk components.

\[ [R^e] = \frac{1}{c} \cdot [S^e] [S^e]^T [R'] + \frac{1}{c} \cdot \sum_{\mu \neq i} ^p [S^e] [S^\mu]^T [R^\mu] \]

\[ = [R^e_{principal}] + [R^e_{crosstalk}] \]

\[ \ldots(4.6) \]
Where, $S'$ is considered the candidate match. Next, both the natures of crosstalk and principal components will be derived.

4.5.1 Principal Component

**Lemma 1:** If, the excitatory stimulus $[S^e]$, bears type-U similarity to any priory encoded stimulus $[S']$, in their α-suit then the principal component of generated response $[R^e]$ resembles its corresponding response pattern $[R']$.

Proof: First, the principal component given by (4.6) will be analyzed to estimate the $\lambda$ of the recalled response in terms of the closeness of the query stimulus $S^e$ and the candidate match with a priory encoded pattern $S'$. Let us consider the retrieval of the $j^{th}$ component of the response. Individual response components are retrieved independently in identical manner. It is also assumed that all the encoded stimulus patterns have $\lambda = 1$. Unity magnitude assignment sets a reference value for the initially encoded information.

$$r_{j,principal}^e = \frac{1}{c} [S^e] [S']^T r_j$$

$$= \frac{1}{c} \left[ \lambda_1 e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} , \lambda_2 e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} , \ldots , \lambda_n e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} \right] \left[ \begin{array}{c} 1.e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} \\ 1.e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} \\ 1.e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} \end{array} \right]$$

$$= \frac{1}{c} \sum_{k} \lambda_k e^{\left( \sum \lambda_i \delta_{i,j} \right) \delta_{i,j}} r_j^i$$

$\ldots (4.7)$
If, the query stimulus, and the target stimulus corresponds closely, then \(\theta_{j,k}^t \rightarrow \theta_{j,k}^t\).

Thus, all the exponent terms become unity with no phase disturbance. Which, reduces to,

\[
\ell_j^{*, principal} \equiv \frac{1}{c} \sum_k \lambda_k r_j^t
\]  

...(4.8)

The phase of the retrieved response corresponds to the retrieved information, and is equivalent to the phase of the encoded response:

\[
\text{argc}(\ell_j^{*, principal}) \equiv \text{argc}(r_j')
\]  

...(proved)

4.5.2 Crosstalk Component

**Lemma 2:** For sufficiently symmetrical distribution of the stimulus and response patterns, the magnitude of the crosstalk component tends to be zero.

Proof: The crosstalk component can be thought as a summation of a set of randomly oriented \(d\)-dimensional vectors. Let us, consider uni-normal projection of a set of such vectors \(A_{\mu}\),

\[
r_j^{*, \text{crosstalk}} = \frac{1}{c} \sum_{\mu \neq j} \mathcal{S}^{*} [T_\mu] [R^\mu] = \frac{1}{c} \left\{ \sum_{\mu \neq j} A_{\mu} \right\}
\]

Let the \(d\)-dimensional vector (left hand side components of (6)) is represented as:

\[
A_{\mu} \equiv \{ A_{\mu}^0, A_{\mu}^1, \ldots, A_{\mu}^d \} \equiv \{ x_{\mu}^0, x_{\mu}^1, \ldots, x_{\mu}^d \}
\]

Where, the middle representation is the representation of \(A\) in \(d\)-dimensional spherical co-ordinates and the right representation is its corresponding \(d\)-dimensional rectilinear representation. Let, the transformations between the two co-ordinate systems are given by:
\[ x_k^\mu = \prod_{j=1}^{k-1} \cos \theta_j^\mu \sin \theta_k^\mu \] when, \( k < d \)

\[ = \prod_{j=1}^{k} \cos \theta_j^\mu \] when, \( k = d \) ....(4.9a)

The expected value of the sum is given by:

\[ E\left[ \left( \sum_{\mu}^{N} A_{\mu,j} \right)^2 \right] = E\left[ \sum_{k}^{d} \left( \sum_{\mu}^{N} x_k^\mu \right)^2 \right] \]

\[ = \sum_{k}^{d} N E\left\{ x_k^\mu \right\}^2 \] ....(4.9b)

For mutually independent distribution of individual phases, for \( k<d \), (for \( k=d \) the derivation is also same):

\[ E\{ x_k^\mu \} = \prod_{j=1}^{k-1} E\{ \cos \theta_j^\mu \} E\{ \sin \theta_k^\mu \} \] ....(4.9c)

Let, \( f_\theta(x) \), the distribution functions of phases are iid (independent and identically distributed). If the phases are symmetrically distributed, i.e., \( f_\theta(x) = f_\theta(\pi-x) \) in the \( 0-2\pi \) range, then,

\[ E\{ \cos \theta \} = \int \cos x f_\theta(x) dx \to 0 \]

\[ E\{ \sin \theta \} = \int \sin x f_\theta(x) dx \to 0 \] ....(4.9d)

Which, directly implies that,

\[ r_{\text{crosstalk}}^e = \frac{1}{c} \left\{ \sum_{\mu \neq i}^{p} A_{\mu} \right\} \equiv 0 \] ....(proved)

4.5.3 Confirmation of Basic Associative Retrieval

The above two lemmas jointly confirm the associative characteristics of this memory.

It is stated in the form of the following two theorems:
Theorem 1: If, the excitatory stimulus \([S']\), bears similarity based on all of its elements (type-U) to any priory encoded stimulus \([S']\), in their \(\alpha\)-suit then the principal component of generated response \([R']\) resembles its corresponding response pattern \([R']\).

Theorem 2: If \([S']\) does not resemble based on all its elements (type-U) to any of the enfolded associations then the elements of \([R']\) demonstrate distinctive low magnitude indicating absence of the requested information in its enfolded memory.

The growth of the crosstalk decides the capacity of the network. For, acceptable recall performance, the magnitude of the crosstalk must remain well below unity. The magnitude of the crosstalk component indicates the saturation level of the holograph.

4.6 Confirmation of Expectation-1 (type-A/BM)

A query is performed with incomplete knowledge (about the expected pattern) and some associated meta-knowledge about the nature of imperfection. The operational memory returns a best possible but yet imperfect answer with some degree of associated meta-knowledge about the nature of its imperfection. This section investigates the relationship between the involved knowledge components.

This section demonstrates how the above defined computational mechanism satisfies the expectations about the input and output characteristics of meta-knowledge \(\beta\) stated in Table-1.

Lemma 3: Given a query stimulus \(S'\), with unequal analog distribution of attention distribution of its element field specified by \(\Lambda^e = [\lambda_1, \lambda_2, \ldots, \lambda_n]\), The memory dynamics specified by equation (4.3) and (4.5), satisfies expectation-1 and retrieves the pattern which best resembles \(S'\), where, the match is based on weighted importance specified in \(\Lambda^e\).
Proof: As before, the principal component given by (4.6) is analyzed to estimate the \( \lambda \) of the recalled response in terms of the closeness of the query stimulus \( S' \) and the candidate match with a priory encoded pattern \( S' \). Let us consider the retrieval of the \( j \)-component of the response. For the sake of notational simplicity let also assume \( d=2 \).

\[
\frac{r_j^{(\text{principal})}}{r_j} = \frac{1}{c} [S']^T r_j
\]

\[
= \frac{i}{c} \left[ \lambda_1 e^{i\theta_1}, \lambda_2 e^{i\theta_2}, \ldots, \lambda_n e^{i\theta_n} \right] \left[ \begin{array}{c} 1. e^{i\theta_1} \\ 1. e^{i\theta_2} \\ \vdots \\ 1. e^{i\theta_n} \end{array} \right] r_j
\]

\[
= \frac{1}{c} \sum_k \lambda_k e^{i(\theta_k - \theta_j)} r_j
\]

\[(4.10)\]

(4.10) shows that the reconstructed response term is in effect a weighted average. Each of the elements of the query stimulus attempts to reconstruct the desired response \( r_j \), modulated by the corresponding \( \lambda_k \). Thus, by adjusting individual \( \lambda_k \)'s the contribution of each query stimulus element can be controlled dynamically during query. Thus, it conforms to the expectation-1 stated at the beginning.

If, it is known during query that a particular \( \theta_k \), is corrupted, or should be ignored, then the matching can be obtained on the basis of other elements, by setting low value for corresponding \( \lambda_k \). This additional control opportunity lies at the heart of the unique capability of selective focusing of holographic computing paradigm.
4.7 Confirmation of Expectation-2 (MNC)

In this section, it will be shown that how the computational framework can itself reconstruct a measure of match, as a fundamental part of bi-modal representation. This measure of match is named mean normalized confidence (MNC).

**Lemma 4:** Given a query stimulus $S^q$, the memory dynamics specified by equation (4.3) and (4.5) satisfies the expected relation between the quality of match and the output distribution of confidence of generated response pattern $R^q$ specified as vector $\hat{\lambda}^q = [\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n]$ of corresponding element magnitudes stated in table-1.

Proof: The magnitude of only one of the response elements $\hat{\lambda}_j^q$ will be derived. With a set of some straightforward trigonometric and algebraic manipulations it is possible to deduce the following expression:

$$\hat{\lambda}_j^q = \frac{1}{c} \sqrt{\sum_k^m (\lambda_k^q)^2 + \sum_k^m \sum_{l \neq k} \lambda_k \lambda_l \cos(\theta_k^q - \theta_l^q - \phi_k^q + \phi_l^q)}$$

assuming $|r_j^q| = 1$

since, $c = \frac{1}{\sum_k^m \lambda_k^q}$, and let $(\theta_k^q - \theta_l^q - \phi_k^q + \phi_l^q) = \phi_k^q - \phi_l^q$

$$= \frac{1}{c} \sqrt{\left( \sum_k^m \lambda_k^q \right)^2 + \sum_k^m \sum_{l \neq k} \lambda_k \lambda_l (\cos \phi_k^q - \phi_l^q - 1)}$$

$$= \sqrt{1 + \left( \sum_k^m \lambda_k^q \cos \phi_k^q - 1 \right)}$$

$$\ldots (4.11)$$

Let us define a distance measure between two patterns $d$ such that, $\alpha$-suit elements of the stimulus $S^q$ and $S^l$ are bounded by the distance $d$ over the entire set, such that $|\theta_j^q - \theta_j^l| \leq d$, for all $j$ which implies, $0 \leq \phi_j^q - \phi_j^l \leq 2d$

96
If $d \to 0$, then the right term of (4.11) becomes zero, which indicates that irrespective of distribution of $\delta$ in the query pattern, $\lambda^*_j \to 1$. This conforms the first two dependencies of the expectation (2). On the other hand, if $d \gg 0$, indicating greater distance between the patterns, then right term of (4.11) becomes increasingly negative ($\cos \phi - 1 \ll 0$), which indicates the decay of $\lambda^*_j$. This satisfies other two parts of the expectation (2). The same results can be extended for other $\lambda^*_j$. (proved).

The decay characteristic is cosine natured. For small perturbation, $\lambda^*_j$ decays very little from unity indicating statistical robustness. However the decay becomes sharper as the distance between the patterns increases. As shown by (4.11), for large $d$, the rate of decay is modulated by the asymmetry in the distribution of $\lambda$-suit in the query.

4.8 Significance of Holographic Computation

Now it will be explained exactly how this new computation formalism can solve the limitations of current AAMs identified in section-2.5.

4.8.1 Robustness of transfer function

The above, encoding and decoding algorithm can be realized in a distributed network of cells, where each cell will be responsible for a simple computation. The cell computation is equivalent to weighted sum of complex number. The transformation is given below:

The output of a cell is determined by analogous equation:

$$
\hat{z}_i = \sum_j \hat{w}_{ij} \hat{s}_j
$$

Where, all dot elements are complex numbers, instead of scalars. Despite the apparent analogous nature of computation, the transformation it realizes on the measurement com-
ponent of input information is fundamentally different from any existing transfer function.

Let, $\hat{w}_i = \|w_i\| e^{-i\omega_i}$. Then, the transformation between the input measurements and the output measurement is given by:

$$\phi_i = \cos^{-1}\left[ \frac{1}{c} \sum_{j} w_{ij} \cos(\theta_j - \omega_i) \right]$$

where, $c = \sum_{j} \|w_{ij}\|$ \hspace{1cm} \ldots(4.12)

In contrast, the scalar product rule of synaptic efficacy is given by (2.3). For the ease of comparison, it is stated below with equivalent notations.

$$\phi_i = f(y_i), \quad y_i = \sum_{j} w_{ij} \theta_j + b_i$$

\ldots(4.13)

This new transfer function has three characteristics that distinguishes itself from conventional transfer functions. The first is that the transfer function is a weighted trigonometric (cosine) mean function, in contrast to the conventional weighted sum. Secondly, that there is no explicit activation function. The third distinguishing feature is that instead of a cell wide single threshold, each of the synaptic inputs to the cell has its individual threshold.

The most important distinction is the first one. Intuitively, any mean function is more robust than any integral function in the face of missing inputs, specially when the number of inputs is limited\textsuperscript{21}. The property of mean function is that it is robust both in the sense of distortion in the measurements of the elements and in the sense of unpredictable loss of the

\textsuperscript{21} Sherrington's [Shen1906] observation on the existence of some form of integration process in the nervous sites is generally used to rationalize the use of linear weighted sum as something fundamental to conventional artificial neural computing. However still now the theory itself has not been decidedly validated or refuted. More importantly, our suggestion of weighted average does not imply the absence of integration. Sherrington's theory also suggests the existence of temporal summation [Jaco93]. Recent evidence suggests that in some cases two neurotransmitters can co-exist in axons [Bark91,p41]. It is also plausible that the pre-synaptic dendrites will also have their individual saturations like any other physical channels. All of which can potentially make the summation non-linear even at the channels.
elements. As it has been analyzed previously, a classical sum is robust only in the former sense. This is the key distinction that makes holographic cells robust against missing element, and thus also makes it possible to conduct RCA search.

The second and third distinctions are related and together these determine the mapping ability of the holographic memory. In a regular associative memory, the non-linear activation function is essential and it decides the nature of the discriminating hyperplane that maps stimulus patterns on to the response classes. It also determines the error tolerance of a network in a noisy environment. Although, there is no explicit non-linearity, the trigonometric transformation pairs serve as the implicit non-linearity in this cell. The computation in the complex domain is inherently non-linear in the phasor plane. Therefore, the cell can perform non-linear mapping. As a result, even a single layer interconnected network of such cell does not suffer from linear separability problem.

The only fundamental difference is that the non-linearity is local (like the existence of individual thresholds), as opposed to the global non-linearity used by conventional neurons. Such, localization of non-linearity is essential for attaining robustness against missing elements.

4.8.2 Role of Hyperspherical representational

The fundamental distinction of holographic cell can be visualized from a representational perspective also. It has been explained in section-5.2, that one of the basic limitation of the conventional network is that there is no representation of 'dont-care'. On a real line, it is not possible to obtain a point which is equidistant from all possible enumerations of an analog measurement. Any enumeration of 'dont-care' on a real line will always induce undue bias towards two of the enumerations than all others.
An obvious solution to this riddle is to place the enumerations on a plane. In fact, one way of explaining holographic representation is that, it generalizes the above solution a step further and puts the enumerations on the surface of a hypersphere (Fig-4.1). The center enumerates an unbiased 'dont-care'. Setting $\lambda = 0$ serves as the mechanism to arrive at such unbiased representation of 'dont-care'.

4.9 Other Learning Models

4.9.1 Differential Hebbian Model

The complex valued representation and the basic transfer function of the proposed holographic cell can be used for other learning models also. One such model is differential encoding scheme. In the previous learning scheme, all the associations are learned independently from one another. It is possible to obtain an enhance learning algorithm where new learning can be made a function of prior learning. In this mode, before any encoding, the stimulus is applied to obtain the unconditioned response of the holograph. Once, the unconditional response is obtained only the difference of it from the actual expected response is encoded using previous algorithm. The modified learning algorithm is given by equation (4.14).

$$X = X + L [S]^T \left( [R] - \frac{1}{c} [S] [X] \right)$$

Here, $L$ is the learning constant in the range of 0-1.0. In this differential mode when the holographic memory encounters a new pattern it learns only that part of an association which is new. Thus it significantly increases the encoding capacity.
4.9.2 Self Organizing Learning

Another learning model is competitive learning analogous to Kohonen’s feature map. In this model, the learning is based on the competition between an array of competing holographic cells. This learning model can be used to perform conceptual coding for image compression by attention weighted vector quantization.

4.9.3 Principal Component Learning

In this model an iterative method is used to find the most significant $m$ principal components of a given set of data. [KhYu94b] contains details of this technique. It has been used as an information compression mechanism for holographs trained with differential Hebbian learning.

However both of these later learning models are beyond the scope of current dissertation and will not be discussed further.

4.10 Chapter Summary

A model of associative memory based on multidimensional complex representation of information elements has been presented in this chapter. This chapter shows how this new representation provides the opportunity to process meta-information significance along with measurement within the framework of an associative memory system, and consequently can support associative search with changeable attention.

Conceptually, this new associative memory is also neural network like cellular. However the new transfer function used by its’ cell performs a weighted average over the measurements of information elements, instead of weighted sum by the conventional neural networks.
CHAPTER 5
CHARACTERISTICS OF
HOLOGRAPHIC COMPUTING

5.1 Introduction

The objective of this chapter is to present the principal characteristics of multidimensional holographic associative computing (MHAC). The characteristics have been investigated both analytically and experimentally.

The chapter is organized as follows. First, section 5.2 shows the analytical derivation of the following characteristics: (i) type-U capacity, (ii) type-U accuracy, (iii) type-A/B RCA capacity, (iv) type-A/B RCA accuracy and (v) effect of the representational dimensionality. For the ease of understanding, in each of these analyses the result has been stated before explaining the detail. Section 5.3, then experimentally shows the characteristics of MHAC. It presents experiment results on (i) the training characteristics, (ii) capacity characteristics, (iii) attention modulated retrieval characteristics, (iv) optimality of using meta-knowledge, (v) MNC characteristics, and (vi) representation dimensionality. It also provides empirical validation of the analyses performed in section 5.2.

This chapter also contains the results of a set of simulations on a complete parametric model space of MHAC to benefit the quick projection and estimation of new applications. These simulations together characterize an operational range-space (ORS). ORS is a sub-space of parameter space, spanned by a set of critical parameters those effect the performance of MHAC. Section-5.4 presents these simulation results and projected operational range-space for effective operation of the memory system.
5.2 Analysis of Characteristics

5.2.1 Capacity (type-U) of Holograph

Capacity refers to the number of patterns that can be effectively stored in a memory. For analog distributed associative systems it is dependant on the accuracy of retrieval. This section, first analyzes the capacity of MHAC when it functions as a conventional AM. A conventional AM uses uniform attention distribution on its element field for retrieval (type-U). The result of this analysis is stated below:

**Result (type-U/capacity):** For a MHAC whose learning is specified by equation (4.3) and (4.4), and retrieval is specified by (4.5), with \( p \) stored patterns each of length \( n \), the capacity is given by:

\[
\text{Capacity} = \sqrt{\frac{n}{p}} 
\]

...(5.1)

when (i) \( p \gg 1 \), and (ii) the input elements are uniformly distributed in phase space.

**Derivation:** Let us first define the saturation of a holograph in the context of previously derived (section-4.5) retrieval breakdown equation (4.7):

**Definition (Saturation):** Saturation of a holograph is defined as the ratio of the magnitude of principal component to the magnitude of the crosstalk component.

\[
SR = \frac{|r_{j}(\text{principal})|}{|r_{j}(\text{crosstalk})|} 
\]

...(5.2)

The lower is the saturation, the better will be the performance of the holographic memory. For effective operation, the saturation (or equivalently crosstalk) should be kept reasonably below unity. Saturation is also an indicator to the capacity of the holograph. It also limits the bounds of error in the retrieved measurements. In this analysis, saturation has been used as a measure of capacity.
If the phases are uniformly distributed, and if the attention is uniformly distributed then the crosstalk components can be considered as a collection of randomly oriented unit vectors in multidimensional space. The collective effect of the superimposed associations enfolded in a holograph during retrieval can be estimated as a sum of random walks in d-dimensional space. If there are \( p \) patterns stored, each with \( n \) elements, then \( (n-1)p \) is the total number of such vectors in the crosstalk. One may provide an upper limit to the expected growth of the crosstalk. Let us examine the average magnitude resulting from a sum of \( (n-1)p \) random walk assuming in the following way. Consider sum of \( P \) unit vectors in 2-dimension:

\[
N e^{i\theta_N} = \lambda_1 e^{i\theta_1} + \lambda_2 e^{i\theta_2} + \ldots + \lambda_P e^{i\theta_P}
\]

Where,

\[
N = \sqrt{\sum_i \lambda_i^2 + \sum_{i \neq j} \lambda_i \lambda_j \cos(\theta_i - \theta_j)}
\]

For \( \lambda_i = 1 \), the magnitude and the phase of the sum-vector are respectively given by:

\[
N = \sqrt{P + \sum_i \sum_{j \neq i} \cos(\theta_i - \theta_j)} \quad \ldots(5.3a)
\]

\[
\theta_N = \sin^{-1}\left[ \frac{1}{n} \sum_i \sin \theta_i \right] \quad \ldots(5.3b)
\]

For symmetric distribution of the phase components the cosine summation term tends to be zero \( \sum_i \sum_{j \neq i} \cos(\theta_i - \theta_j) \rightarrow 0 \), which implies:

\[
N = \sqrt{P}
\]

The above result can be expanded even for higher dimensional space. The similar trigonometric derivation for higher order representation results in higher ordered cosine and sin terms implying even faster decay of it. Fig-5.1 shows how the mean of a set of unit
vectors in a hyperspherical space grows for various dimensionality of the hypersphere. The phases were generated with uniform distribution in $I = [0, 2\pi]$. The growth is not only slow, but interestingly, it decreases even further with higher dimensionality.

This characteristic will be dealt in further depth later in section-5.2.5.2. However for now, returning to our previous discussion, the result of (5.3) implies:

$$\| r_{cross} \| \to \sqrt{n(p - 1)}$$  
\[ (5.4) \]

On the other hand, if the query stimulus $S^* \to S'$ element by element, the strength of the principal component is:

$$\| r_{principal} \| \to n$$  
\[ (5.5) \]

Thus, the ratio of the magnitude is given by:
\[
\frac{n}{\sqrt{n(p-1)}} = \sqrt{\frac{n}{p}} \quad \text{...(proved)}
\]

One way of estimating the capacity of a holographic memory is by the relative strength of the principal and the crosstalk components of the resultant response. If the crosstalk component is low, then only the principal component will contribute to the resultant response. Thus the above ratio, called as saturation, can also be used as a measure of capacity of a holographic memory.

5.2.2 Accuracy (type-U) of Holographic Retrieval

This section looks closely in the error in measurement introduced into the RCA type-U retrieved response from a perfect query pattern.

Result (type-U accuracy of retrieval): For a MHAC specified with equation (3), (4), and (5), with n stimulus elements and p stored patterns the maximum phase error in the retrieved response for uniform distribution of attention is given by

\[
|\Phi_{\text{error}}|_{\text{max}} = \sin^{-1}\left(\sqrt{\frac{p}{n}}\right) \quad \text{...(5.6)}
\]

when (i) \( p \gg 1 \), and (ii) the input elements are uniformly distributed in phase space.

Derivation: The resultant response is given by the sum of principal and crosstalk components. The accuracy of a recalled pattern depends on these two components. Let, us investigate the case, when the query stimulus \( S' \) resembles one of the stored pattern. Let, the crosstalk component is given by \( r_N e^{i\theta_N} \), and the principal component is given by \( r_S e^{i\theta_S} \), then the error in phase (which represents the measurement) of the resultant component is given by:

\[
\phi_e = \tan^{-1}\left[\frac{r_N \sin(\theta_N - \theta_S)}{r_S + r_N \cos(\theta_N - \theta_S)}\right] \quad \text{...(5.7)}
\]
Fig-5.2 illustrated the addition in hyperspherical space. The phase deviation, is maximum when, \( \Phi_{\text{max}} = (\theta_n - \theta_s) - 90^\circ \). Thus, for saturation given by equation (5.4) and (5.5), the maximum phase error is:

\[
|\Phi_{\text{error}}|_{\text{max}} = \sin^{-1}\left(\frac{\sqrt{p}}{n}\right)
\]

...(proved)

Fig-5.2 Geometry of Phase Error

**Error due to imperfect query pattern:** In fact, additional error may be present in the reconstructed output if the query pattern itself deviates from any encoded pattern. The total error in the resultant of the errors contributed by pattern deviation and that introduced by holograph loading (crosstalk). The error due to pattern deviation is the sum of the deviations of individual pattern elements. Thus it linearly moves away from target patterns, with mean of the shift in query \( \Phi_e = n \cdot \varepsilon_i \), when \( \theta'_i - \theta_f = \varepsilon_i \) when the error is small. It can be geometrically shown that the magnitude of the error due to pattern deviation grows in the order of \( \sqrt{2 \sin(\frac{\varepsilon_i}{2})} \).

5.2.3 Saturation in RCA Search

This section investigates the capabilities of this new memory to perform retrieval with changeable attention (RCA). First, the saturation with modulated attention is derived.
Result (type-A/Saturation): For a MHAC specified with equation (3), (4), and (5) with n stimulus elements and p stored patterns, and an unequal distribution of attention specified by the vector \( \Lambda^e = [\lambda_1, \lambda_2, \ldots, \lambda_n] \), then the saturation is given by:

\[
SR = \sqrt{\frac{n \cdot w}{(p - 1)}}
\]

...(5.8)

When (i) \( p \gg 1 \), and (ii) the elements are symmetrically distributed in phase space. Where \( w \) refers to the 'porosity' of the attention distribution.

**Derivation:** The saturation ratio is defined as the ratio of the signal-to-noise magnitude:\[^22\]

\[
SR = \frac{| r_{\text{principal}}^j |}{| r_{\text{cross talk}} |}
\]

From (4.6), it can be derived:

\[
SR = \sqrt{\frac{\sum_k \lambda_k^2 + \sum_{k \neq l} \lambda_k \lambda_l \cos \phi_{k-l}^e}{(p - 1) \sum_k \lambda_k^2 + \sum_{\mu \neq i} \sum_k \sum_l \lambda_k \lambda_l \cos \phi_{k-l}^\mu}}
\]

...(5.9a)

Assuming, independent, identical and symmetrical distribution of \( \alpha \)-suit \( (\theta^p) \), over all the element space of all the enfolded patterns:

\[
E[\cos \phi_{k-l}^e] = 0 \text{ or for sufficiently large } pn, \quad \left[ \sum_{\mu \neq i} \sum_k \sum_l \lambda_k \lambda_l \cos \phi_{k-l}^\mu \right] \rightarrow 0
\]

Thus,

\[^22\text{Note, that saturation ratio is not same as the signal-to-noise ratio (SNR). SNR is the ratio of the signal-to-noise measurements.}\]
The same distance measure as defined in section 4.6 is used. If, the distance between the candidate and query is large \( (d_{e.t} > 0) \), then according to (5.9b):

\[
SR = \sqrt{\frac{1}{(p-1)}} \left[ 1 + \frac{\sum_{i=1}^{n} \lambda_i \cos \phi_i}{\sum_{k} (\lambda_k)^2} \right] \quad \ldots (5.9b)
\]

On the other hand, for close match, \( (d_{e.t} \to 0) \):

\[
SR = \sqrt{\frac{n}{(p-1)w}}
\]

Where, \( w \) is attention strength.

**Definition (attention strength):** Attention strength \( w \) refers to relative strength of the attention distribution over the element space, and is defined by:

\[
w = \frac{\left[ \frac{\sum_{k} (\lambda_k)^2}{n \cdot \sum_{k} (\lambda_k)^2} \right]^{2}}{\frac{[E\{\lambda_i\}]^2}{E\{\lambda^2\}}} = \frac{[E\{\lambda_i\}]^2}{E\{\lambda^2\}} \quad \ldots (5.10)
\]

The attention strength \( w \) intuitively refers to the 'porosity' of the window frame. It varies from 0 to 1 and depends on the distribution of \( \lambda \) in the query field. 0 indicates absolute attention of the information in the query frame and 1 indicating absolute certainty. For type-U search \( w=1 \).
when, $\lambda_j \rightarrow 1$,

$$SR = \sqrt{\frac{n}{(p - 1)}}$$

\[\textbf{proved}\]

\textbf{5.2.4 Accuracy in RCA Search}

This section investigates the accuracy of type-A retrieval. In a similar manner of section-5.2.2 it can be derived:

\textbf{Result (type-A/accuracy of retrieval): For a MHAC specified with equation (3), (4), and (5) with n stimulus elements and p stored patterns the maximum distortion due to crosstalk is given by}

$$|\Phi_{error}|_{\text{max}} = \sin^{-1}\left(\sqrt{\frac{pw}{n}}\right)$$

\[\text{(5.11)}\]

When (i) $p \gg 1$, and (ii) the input elements are uniformly distributed in phase space.

The above analysis shows that the focus can be effectively (almost linearly) compensated with higher $n$ or lower $p$. This result is very significant, because, even for a fixed size problem it is possible to design a network with exponentially higher effective stimulus length ($n$), by techniques such as higher order encoding. The above analysis provides us the clue to select a suitable $n$ for a particular application. Section 5.4 investigates closely and discusses the effective range of loading for various attention ranges.

\textbf{5.2.5 Effect of Representational Dimension}

\textbf{5.2.5.1 Effect on Saturation Characteristics}

One of the fundamental aspect of this new memory system is the computational representation it has adopted to process individual elements of information. MHAC enumerates the measurements from sensory inputs on the surface of a unit hypersphere, whereas
conventional neuro computing enumerates them on a real linear interval. This section investigates how the dimensionality of this hypersphere effects the performance of this memory system.

Investigation shows that higher dimensional representation has the effect of significantly improving the recovery performance, specially in the case of reduced symmetry in the original input distribution.

The crosstalk component of holographic substrate as shown by equation-4.7 can be rewritten in the following form, where $A_{\mu}$ represents a resultant product vector:

$$[R_{\text{crosstalk}}] = \frac{1}{N} \cdot \left\{ \sum_{\mu=1}^{P} A_{\mu} \right\} \quad \ldots(5.12)$$

The normalized sum of the vectors is considered here, so that the crosstalk magnitude becomes equivalent to previously defined saturation. For practical purpose, the saturation can be thought as proportional to the sum of a set of randomly oriented vectors. Thus, the principal result can be stated as following:

**Principal result (dimensionality):** The holographic saturation due to the superimposition of a set of unit vectors tends to be smaller if the dimensionality of the vectors are larger. The tendency increases with the increase in the skew of phase distributions of the vectors.

However besides demonstrating the above result, the specific nature and characteristics of multidimensional representation are also investigated here.

This above result is demonstrated by comparing the growth of a set of vectors in respectively $m$ and $n$-dimensional spaces, where $m < n$. 
5.2.5.2 Vectors in Multidimensional Space

Let us consider a set of \( P \) vectors with uni-normal projection in \( n \) and \( m \) dimensional spaces \( \vec{A}_i \) and \( \vec{A}_i \) respectively, where \( 1 \leq i \leq P \). In this notation, the number with the vector arrow indicates the dimensionality of the vector. Now, the objective is to investigate the following inequality:

\[
| \sum_{j=1}^{P} [\vec{A}_i^m] | \geq | \sum_{j=1}^{P} [\vec{A}_i^n] | , \text{when} \ m < n 
\] ....(5.13)

Let the \( n \)-dimensional vector (left hand side components of (6)) is represented as:

\[
\vec{A}_i^n = A(\theta_1^i, \theta_2^i, \ldots, \theta_{n-1}^i) \equiv A(x_1^i, x_2^i, \ldots, x_n^i)
\]

Where, the middle expression is the representation of \( A \) in \( n \)-dimensional spherical co-ordinates and the right representation is its corresponding \( n \)-dimensional rectilinear representation. Let, the transformations between the two co-ordinate systems are given by:

\[
x_k^i = \prod_{j=1}^{k-2} \cos \theta_j \sin \theta_{k-1} \quad \text{when,} \ k < n
\]
\[
= \prod_{j=1}^{k-1} \cos \theta_j \quad \text{when,} \ k = n
\] ....(5.14)

Now, let us define an \( m \)-dimensional spherical sub-space of the above \( n \)-dimensional space using its first \( m-1 \) dimensions. Then the uni-normal projection vectors of \( A \) in \( m \)-dimensional space (\( \vec{A}_i^m \)), (right hand side components of (5.13)), become:

\[
\vec{A}_i^m = A(\theta_1^i, \theta_2^i, \ldots, \theta_{m-1}^i) \equiv A(y_1^i, y_2^i, \ldots, y_m^i)
\]

Where, its corresponding rectilinear co-ordinates are given by:

\[
y_k^i = \prod_{j=1}^{k-2} \cos \theta_j \sin \theta_{k-1} \quad \text{when,} \ k < m
\]
\[
= \prod_{j=1}^{k-1} \cos \theta_j \quad \text{when,} \ k = m
\] ....(5.15)
This section presents the notation and the co-ordinate transformation equations. Next, the effect of each individual dimensions is traced through a recursive formulation of (5.13). This effect is called dimension dispersion component.

5.2.5.3 Dimension Dispersion Component

Now the following sum of the vectors are computed:

\[
| \sum_i^n A_i |^2 = \sum_k^p \left( \sum_i x_i^k \right)^2 = \sum_k^p \left( \sum_i x_i^k \right)^2 + \left( \sum_i x_i^k \right)^2
\]

according to equation (5.14),

\[
= \sum_k^p \left( \sum_i x_i^k \right)^2 + \left( \sum_k^p \prod_k^i \cos \theta_k \right)^2 \quad \text{(step 1)}
\]

\[
= \sum_k^p \left( \sum_i x_i^k \right)^2 + \left( \sum_k^p \prod_k^i \cos \theta_k \right)^2 - B_{n-1} \quad \text{(step 2)}
\]

Using (5.14) and (5.15) and some trigonometric manipulation step 2 has been derived from step 1. Where,

\[
B_{n-1} = \sum_k^p \sum_{j \neq k} \left[ 1 - \cos(\theta_{n-1}^j - \theta_{n-1}^j) \prod_k^i \cos \theta_k \cos \theta_k \right]
\]

In a similar way it can be further derived:

\[
| \sum_i^n A_i |^2 = \sum_k^{n-1} \left( \sum_i x_i^k \right)^2 + \left( \sum_k^{n-1} \prod_k^i \cos \theta_k \right)^2 - B_{n-2} - B_{n-1}
\]

\[
= \sum_k^{n-1} \left( \sum_i x_i^k \right)^2 + \left( \sum_k^{n-1} \prod_k^i \cos \theta_k \right)^2 - B_{n-2} - B_{n-1}
\]

Using equation (5.14),

\[
= \sum_k^{n-1} \left( \sum_i \prod_j \cos \theta_j \sin \theta_{(k-1)} \right)^2 + \left( \sum_i \prod_k \cos \theta_k \right)^2 - B_{n-2} - B_{n-1}
\]
Using equation (5.15), therefore,

\[ | \sum_{k}^{m} \sum_{i}^{p} A_{i} |^2 = \sum_{k}^{m-1} \left( \sum_{i}^{p} y_{k} \right)^2 + \left( \sum_{i}^{p} y_{m} \right)^2 - B_{m} \ldots - B_{n-2} - B_{n-1} \]

\[ = \sum_{k}^{m} \left( \sum_{i}^{p} y_{k} \right)^2 - B_{m} \ldots - B_{n-2} - B_{n-1} \]

\[ = | \sum_{i}^{p} A_{i} |^2 - \sum_{k=m}^{n-1} B_{k} \]

\[ \ldots (5.16) \]

Each of the \( B_k \) terms accounts for the change in saturation for adding a dimension and it is called Dimension Dispersion Component (DDC) of the vector magnitude.

5.2.5.4 Holograph Saturation and Dimension Dispersion Component

DDC plays the key role in the saturation of the holograph. As demonstrated by equation-(5.16), positive DDC asserts the validity of the inequality stated in equation-(5.13). The greater is the value of DDC, the lower will be the saturation in the holographic substrate.

DDC is a function of the distribution of the individual phase components. Now, it will be investigated that how the distribution of phase may effect the magnitude and in particular the sign of DDC.

Stochastically, the expected value of DDC is:

\[ E\{B_n\} = \sum_{i}^{p} \sum_{j \neq i} \sum_{k}^{n-1} \left[ 1 - \cos(\theta_{n} - \theta_{j}) \right] \prod_{k \neq i} \cos \theta_{k} \cos \theta_{k} \]

Assuming statistical independence between the phase components,

\[ = \sum_{i}^{p} \sum_{j \neq i} \left[ E[1 - \cos(\theta_{n} - \theta_{j})] \prod_{k \neq i} E[\cos \theta_{k}] E[\cos \theta_{k}] \right] \]

the first term \( E[1 - \cos(\theta_{n} - \theta_{j})] = Q \) is always a positive quantity such that \( 0 \leq Q \leq 2 \).

On the other hand, further assuming identity of distribution among the phase components,

\[ = p_{c} Q \left. E[\cos \theta] \right|^{2(n-k)} \geq 0 \]
Thus, DDC is always stocastically positive. Thus, in equality (5.13) each additional dimension has the effect of reducing the saturation. The following important conclusions can be drawn from the above analysis.

5.2.5.5 Nature of DDC

**Result a:** Stochastically, the expected value of the DDC terms, in general, will always be positive, irrespective of their specific distribution with the assumptions of their independence and identity.

**Result b:** Deterministically, if the phase components are distributed in such a way that $-\pi/2 \leq \theta \leq \pi/2$, then also the DDC terms, in general will always be positive.

**Result c:** The expected value of DDC $E\{B_n\} \to 0$, as the distribution $f(\theta) \to C$ tends to uniformity. It implies, $\int_{-\pi}^{\pi} \cos \theta f(\theta) d\theta \to 0$. Thus, the lowering of holograph saturation due to dimension dispersion becomes increasingly pronounced as the distribution shifts towards non-symmetry.

**Result d:** $\cos \theta \leq 1$, hence, the DDC decrease with the increase in dimension. At higher dimension more and more $\cos \theta$ terms contribute to the product. Thus, the improvement due to DDC is relatively more effective at lower range of dimension.

Section 5.3.6 presents simulation results verifying each of these analytical results on DDC characteristics.

5.3 Simulation Results on MHAC Characteristics

In this section, results of extensive computer simulation on the characteristics of this model are presented. These experiments specifically explore the following issues: (i) the training characteristics, (ii) capacity characteristics, (iii) attention modulated retrieval characteristics, (iv) effectiveness of meta-knowledge (attention), (v) MNC characteristics, and (vi) dimensionality of representation.
The first two experiments investigate the speed of convergence and capacity of MHAC and validate the results of previous section. The next two experiments examine the RCA capability. The experiment on attention modulated retrieval shows how the quality of retrieval varies with the degree of attention. On the other hand, the experiment on metaknowledge shows the improvement in the accuracy of retrieval due to the presence of attention. Attention is a meta-knowledge. Conventional AM can not take advantage of this additional information. Experiment on mean normalized confidence (MNC) shows another unique feature of MHAC; its ability output a confidence measure (MNC) of the accuracy of retrieved information. Finally, the experiment on dimensionality of representation validates the analysis of section-5.2.5.

The following two measures have been used to quantify the accuracy of measurement and strength of significance components of patterns. Let, the expected response is $R_i^e$ and the retrieved response is $R_i^T(\omega)$ as in section-4.3.

**Definition (Accuracy of retrieval):** Accuracy of retrieval is measured as the peak signal to noise ratio in the measurement component of information over all the elements.

$$SNR = 20\log \frac{2\pi}{mse} \quad mse = \sqrt{\frac{1}{m} \sum [\phi_i^e - \phi_i^T(\omega)]^2}$$

...(5.17)

The peak signal is given by the dynamic phase range $2\pi$.

The performance of the memory system *Average SNR* is computed by averaging over all the pattern associations enfolded in the memory.

**Definition (Mean normalized confidence):** the confidence on the retrieved pattern is estimated as the attention strength of the retrieved pattern elements normalized by the attention strength of the expected pattern.
The performance of the memory system \( \text{Average MNC} \) is computed by averaging over all the pattern associations enfolded in the memory.

### 5.3.1 Training Characteristics

**TRAINING CHARACTERISTICS**

\[
MNC = \frac{\sum_i x_i^T}{\sum_i x_i^T}
\]  

\[\text{...(5.18)}\]

One of the features of holographic memory is its capability of fast learning. Fig-5.3 shows a simulation where \( p \) stimulus patterns with \( n=200 \) have been encoded in the memory. These patterns are randomly generated with phases uniformly distributed in \( 0 - 2\pi \) range. During each encoding iteration, the accuracy of the holograph is tracked by performing complete recall. This graph plots the average recall error of these iterations. It plots five
loading cases with total number of encoded patterns varying respectively p=20, 40, 60, 80, 100. As evident, the memory converges within 10-15 iterations. The average recall error drops to less than 2% of the dynamic range \((0 - 2\pi)\) within 6 iterations, in each of the cases.

**CONFIDENCE BUILDUP CHARACTERISTICS**

\[ n=100, p/n=0.1, 0.2, 0.3, 0.4, 0.5 \]

![Confidence Buildup Characteristics](image)

**Fig-5.4 Convergence of Recalled MNC**

Fig-5.4 plots the convergence of network confidence for these cases. The confidence is measured in terms of the average magnitude of the response elements, called *mean normalized confidence* (MNC). MNC=1 indicates the full confidence of the network in the retrieved response pattern. As evident, the MNC gradually converges to 1 as the retrieved patterns become more and more accurate.

MHAC characteristically demonstrates significantly faster convergence than that of conventional scalar ANN models. The structure of the hyperspherical representation space
allows MHAC to converge faster. In hyperspherical space, while moving from one point to another on the surface of the hypersphere, intermediate states can travel using a direct (and shorter) euclidian path through the hypersphere.

It has been observed that the convergence speed of MHAC remains equally faster even for extremely large $n$. Later, section-5.4 explains this issue in more details.

5.3.2 Capacity Characteristics

![Graph showing the response recall accuracy with the number of patterns stored.](image)

**Fig-5.5 Recall Accuracy and Number of Patterns**

Fig-5.5 plots the response recall accuracy with the number of pattern stored. A set of patterns is randomly generated with phase uniformly distributed in $0 - 2\pi$ range. For this experiment, three types of patterns are used with lengths respectively $n=256, 128, \text{ and } 64$. Each time the holograph has been trained with 30 iterations. Fig-5.5 shows how the recall error varies for holographs corresponding to each of these pattern classes as the number of
stored patterns increases. As evident, the average recall accuracy is well above 50 db (with respect to the dynamic range) even when the number of stored patterns reaches about 50% on the total elements.

For MHAC, the response recall error is related to the ratio of number of stimulus vector elements to the number of stored pattern (equation-5.6). Fig-5.6 plots the same three results, but x-axis representing the ratio L=p/n instead of p. It is evident, that all three curves almost merge on one another. Which confirms that the capacity is directly related to the ratio. Irrespective of n, the capacity (for a give accuracy) is dependent on the load factor L.

An enormous number of associations can be superimposed on a holograph, because of such scalability. In a typical situation, assuming a loading factor of .5 (corresponding to >75 db accuracy), almost 32,000 patterns, each of size 256x256 pixels, can be associatively memorized on a holograph. Corresponding associations can be recover with less than .017%
average error on its values. A regular database would require 2 Gigabyte memory to store such a volume of information. A holographic associative system with 8 element response would require about 1 Megabyte of holographic memory.

As can be seen, for MHAC a load factor of .3-.4 can easily be achieved. This far exceeds the capacity of conventional networks\textsuperscript{23}. Thus, even as a regular type-U memory, MHAC demonstrates significantly higher capacity.

Accuracy can also be further improved with higher order encoding as used in conjunction with other ANN models. In this technique, an input stimulus of size \( n \) is expanded to generate more physically effective stimuli elements by high order expansion [Vemu92, TaJo90]. The dimension of generated holograph depends on the effective \( n \) after higher order expansion. Almost unlimited number of patterns can be stored and accurately retrieved in a single holograph with combinatorial high order expansion of stimuli field. As expected, on one side, higher order encoding lowers saturation and increases distinguishability. On the other hand, it increases memory size and lowers generalization capability.

5.3.3 Attention Modulated Retrieval Characteristics

This experiment investigates the characteristics of holographic network for retrieving patterns with changeable attention. The stimulus patterns used for query were generated from the encoded stimulus patterns, but with partial attention. Fig-5.7 plots the average recall error with the focus density (x-axis). Focus density represents the percentage of stimulus elements which were used for query. It signifies the degree of focus or attention strength. During the experiments, focus had been varied from 100\% (all elements used) to 10\%.

\textsuperscript{23} For example, the maximum allowable load factor of Hopfield Model is 0.15 [VeMu92].
FOCUS CHARACTERISTICS

\[ n=200, \ p/n=.2 \]

Fig-5.7 plots the results in both average SNR (left y-axis) and average percentage error (right y-axis) scales. It plots the performance at three stages of training, after 3rd, 5th and 10th iteration of the reinforcement cycles for 20% loading (holograph encodes \( p=40 \) randomly generated patterns with \( n=200 \) elements each). This plots show that even with 10% focus strength, system can recover information with more than 15 db (within \( \pm 10\% \) of the expected analog value) accuracy, far exceeding the capability of conventional networks. A 10% focus corresponds to a situation where 90% of the stimulus elements are excluded from the query pattern. As evident, the accuracy steadily increases with the increase in attention strength. With full (100%) attention, the accuracy exceeds 70 db.

The capability of this new memory has two aspects. On one hand, the network, even when working as ordinary uniform attention (type-U) memory, can be considered extremely
accurate. On the other hand, while the attention falls below (rough) 50%, it can function with reasonable accuracy. Whereas, in this range of focus most other associative memories cease to operate.

5.3.4 Retrieval with Meta-knowledge

One of the unique characteristics of MHAC is its ability to utilize additional meta-knowledge about the state of basic information elements in the form of attention. It can accordingly modify the contribution of each element of the query pattern in the internal recollection process. In contrary, conventional ANNs can not take advantage of available additional meta-information. As explained before (section-2.4), this is because, these models do not have any representational framework to accommodate such additional information. The reason that conventional ANNs cannot recover from large error is that they cannot ignore error, even if some information is available about the region of error. On the other hand, in such a case (when ever such additional information about the state of basic elements is available) MHAC can converge to the best possible solution.

The experiment presented in section has been designed to investigate the difference that this meta-knowledge makes. This experiment uses a set of noisy versions of the originally stored patterns as query patterns. It then compares the performances when complete queries are performed with and without the meta-knowledge.

The experiment is based on the control of following three principal parameters: (a) number of distorted elements in the query stimulus, (b) the strength with which these elements were distorted, and finally, (c) the applied importance (meta-knowledge) on the

24 The type of distortion can effect the performance. For example, if the distortions are symmetrically distributed then there is a possibility that they might cancel each other. Under such rare circumstances even a network without attention, may be able to retrieve correctly. However in a general retrieval scenario, one can not expect such nice distortion.
distorted elements (DCF) during recollection. A simplified holograph with P=1, which associates a random stimulus with 100 with a random response with 100 elements, is considered here to observe the relationships among these three parameters.

Fig-5.8 plots the average SNR of the recollected responses (y-axis) against the noise strength of the distorted elements of the query pattern (y-axis). It plots the recall performance for both situations: (a) when no meta-knowledge is applied. This is done by assuming unity magnitude of all the query elements. And, (b) when the meta-knowledge is applied. This is done by setting the magnitude of the a-priori known distorted elements to 0.02. The three pairs of curves plot the performance for cases where respectively 30%, 50% and 80% of the query elements have been distorted.

**EFFECT OF META-KNOWLEDGE**

**SNR vs. Noise (H=100x100, P=1, DCF=.02)**

Fig-5.8 Meta-knowledge and Noise
EFFECT OF META-KNOWLEDGE
SNR vs. Distortion Window (H=100x100, P=1, DCF=.02)

The result clearly shows that application of meta-knowledge improves the recollection accuracy (almost by 30-40 db in all three pairs). The gentle slopes of the 'with meta-knowledge' curves show that there is some loss in ultimate performance as a general effect of the noise. However the 'without meta-knowledge' systems fails to achieve this level of accuracy.

The next figure plots the effect of number of distorted elements as the x-axis. The x-axis can also be thought as the degree of attention. It compares the recollection characteristics with and without meta-knowledge for three cases (0.2, 0.5 and 0.8) of applied noise magnitude on the distorted elements with original magnitude 1.0. The result is almost identical as before and demonstrates that meta-knowledge improves the recollection accuracy almost by 20-30 db in all three pairs.
This experiment demonstrates an important non-optimality of the conventional ANNs. For an application scenario, where the meta-knowledge about the spatial distribution of noise is available, a conventional ANN, handicapped by the inability to ignore, performs far less than optimum. Multidimensional holographic representation provides a way to overcome such non-optimality and to improve the retrieval performance under such circumstances.

5.3.5 MNC Characteristics

![Graph](image)

This simulation is directed for investigating, the other unique ability of this computing paradigm:- the ability to differentiate between stored and unstored patterns. As per expectation-2 (section-4.2) holographic memory not only retrieves a pattern, but also can convey its confidence in match. One of the implication of such capability is that it can clearly
Fig-5.11 MNC for Stored/Unstored Patterns

distinguish between the pattern which has been stored from those which has not been stored.

In this simulation a set of associations is randomly generated. Half of these are then encoded into a holograph. Then the holograph is tested with both of these sets. The following two figures show the result of the experiment for a holograph with 25% loading. In this experiment, patterns with indices 1-25 are encoded and with indices 26-50 are unencoded patterns. The first figure plots the SNR for each of the retrieved patterns and the second figure plots the MNC of the retrieved patterns. For both stored and unstored patterns, the recollection has been performed with full (100%) and half (50%) of the stimulus elements as focus.

As shown, for the unencoded patterns, the network produces distinctively low MNC in the range of 0.5-0.6. While, for the encoded patterns, the MNC remains close to 1.0. The
variation in MNC due to the variation of focus is almost negligible. The other figure plots the corresponding accuracy of the recollection. This also shows 60 to 10 db accuracy for the encoded patterns, and distinctively low accuracy for the non-encoded patterns.

The following two figures summarize the result for all the patterns and plot the average SNR and MNC over all the stored and unstored sets of patterns against the degree of focus. A set of recollections has been performed with degree of focus varying from 0% to 90%. The figures plot the response of the holograph both after 3rd and 5th iteration.

Almost irrespective of the amount of focus, the average MNC remained close to .5 to .6 range for unstored patterns and close to unity for stored pattern set. The results show that, although the quality of the retrieved signal is somewhat sensitive to the amount of stimulus elements used during recovery, the MNC still clearly can distinguish between the stored and unstored patterns.
This distinguishability is one of the unique and important capability of holographic computing. Conventional artificial neural networks do not have any such indicator. This important feedback mechanism can be used as a basis for fuzzy logical inference for subsequent high level reasoning, and has been discussed in section-CHAPTER 7.

5.3.6 Dimensionality of Representation

The objective of this simulation is to examine the effect of representation dimension on the performance of a holograph. It also verifies the analytical results obtained in section 5.2.5.

The first experiment observes the growth characteristics for vectors in hyperspherical space. In this experiment, a set of 100 vectors has been randomly generated. The phase elements of these vectors are uniformly distributed in $+R \geq \theta > -R$ range. A distribution
corresponding to \(| R | < 180^\circ\), corresponds to uniform symmetric distribution, whereas the first two ranges represent increasingly non-symmetric distribution. Two specific distributions with \(| R | < 90^\circ\) and \(| R | < 120^\circ\) has been considered.

Fig-5.14, shows the growth for \(| R | < 90^\circ\). The curves show the growth of the correlations for hyperspace dimensions \(D=2,3,4\) and 5. The curves clearly show that the correlation magnitude decreases for each added dimension (results a and b of section 5.2.5). The Saturation increases almost linearly with the increase of number of associations.

\[\text{MULTI-DIMENSIONAL ERROR GROWTH}\]

\[\text{Phase Distribution } |R|<90\]

![Graph showing growth for different hyperspace dimensions](image)

Fig-5.14 Effect of Representation Dimension (I)

Fig-5.15 shows the same plot for distribution \(| R | < 120^\circ\). The result clearly shows, that the rate of growth of the correlation magnitude increases as the distribution becomes narrower (result c of section 5.2.5). However the growth retards as the vectors spread to higher-dimensional space (result d of section 5.2.5).
Now, the effect of dimensionally of representation space on the characteristics of a complete holographic memory will be examined. The experiment has been performed with P (=16) vectors each with D dimensional S (=32) elements. Each of the phase components has been generated randomly with uniform step distribution within the range R.

In the encoding process, each of these vectors is associated with corresponding matching response vector. The associations are superimposed on multidimensional holograph. Each of these response vectors is then recalled. During the recall process the principal component and the cross component of the separately measured. The experiment is repeated for dimensions D=2 to 10.

Fig-5.16 plots the signal to noise ratio against the dimensionality. Fig-5.17 plots the crosstalk component against dimensionality.
MULTI-DIMENSIONAL ERROR GROWTH
Stimulus Length=32, Patterns=16

Fig-5.16 SNR and Dimension

MULTI-DIMENSIONAL ERROR GROWTH
Stimulus Length=32, Patterns=16

Fig-5.17 Crosstalk and Dimension
The results of this experiment thus clearly show that SNR improves (Fig-5.16) and crosstalk (Fig-5.17) decreases as one shifts to higher dimensional representation. The improvement is more drastic if the phase distribution window is narrow. This confirms the previous analysis.

5.4 Operational Range Space

The performance of MHAC is dependent on various factors such as nature of data, size of holograph, rigourness of learning, strength of focus, etc. The objective of the special set of experiments presented in this section is not only to investigate the abstract relationship among these factors, but also to determine the quantitative bounds of critical parameters for effective operation of the memory. The objective is to find the specific ranges of these critical parameters within which an effective and cost efficient memory with RCA retrieval capability can be constructed. The principal question around which it focuses the investigation is as following:

**Question:** How small focus can be used to perform a type-A RCA search, and what are the typical holographic performance parameter ranges for it?

To obtain answer to this central design question, a specially designed experimentation has been performed. In this approach first, a complete set of generic (not specific to any particular application) parameters have been identified that crucially effects the behavior of the target system (MHAC). Then, a set of measures has been developed to quantify each of these parameters.

These parameters together span a parameter space. The objective of subsequent experimentation is to determine a sub-space of this parameter space, called *operational range space* (ORS), within which it is possible to guarantee a target performance. The objective is to empirically determine the boundaries of this operational range space.
The availability of an ORS on any system model is advantageous from engineering point of view. Given an ORS, when a new application is taken under consideration, all that is required, is to measure application specific parameters and to verify whether it falls inside or outside the ORS. If, it is within, then the pre analyzed results available from ORS experimentation can be used to predict approximate performance, as well as to estimate necessary configuration of the system for that application. On the other hand, if it is outside, ORS experimentation results can still be readily used to identify the exact nature of special intervention that would be required to bring the application within ORS.

5.4.1 Parameters

Following are the five critical MHAC parameters on the basis of which this range-space simulation has been performed; (i) focus strength, (ii) load factor, (iii) distribution asymmetry and (iv) focus strength, (v) training iterations. These are explained below. Signal-to-noise ratio is the objective function and has been defined in section-5.3.

**Definition (focus strength):** The focus strength refers to the input strength of a query pattern compared to the overall strength of encoded pattern. Assuming uni-magnitude encoding of the pattern elements, it is defined as following, whose values varies from 0 to 1:

\[
\sum_{i} \lambda_i \\
\frac{1}{n}
\]

\[f = \frac{\sum_{i} \lambda_i}{n} \quad \text{...(5.19)}\]

**Definition (load factor):** A holograph loading factor (L) defined as the ratio of the total number of elements (n) in the patterns to the number of stimulus response associations (p) encoded.

\[L = \frac{P}{n} \quad \text{...(5.20)}\]

**Definition (Asymmetry):** The asymmetry (k) of a pattern refers to the circular distribution of the pattern elements around the center of the representation hypersphere. It is defined as following, whose value varies from 0 to 1:
\[
    k = \frac{\sum_{i} \lambda_i \exp\left(\sum_{j} d_j \theta_i^j\right)}{\sum_{i} \lambda_i}
\]  

\[\text{...}(5.21)\]

**Definition (Iterations):** Iteration (\(I\)) of learning refers to the number of times each training association is presented to the network.

The subsequent section presents experiments which establish the range space.

### 5.4.2 Analysis of Experiments

The range space is a multi-dimensional space. Therefore, to locate the boundaries of ORS, a set of experiments have been carefully organized which measures the RCA retrieval performance one by one with respect to the continuous variation of each of the four independent parameters (referred as principal parameter), and step variation of a second parameter (referred as secondary parameter), while keeping all other parameters (referred as tertiary parameters) in quasi critical (near boundary) states. The overall process of the empirical estimation of ORS boundaries is spiral because it requires determination of quasi critical states for secondary and tertiary parameters before the principal parameters can be varied continuously. However this section will not include the initial exploratory experimentation, rather will directly focus on the results of the advanced experiments.

In each of the following graphs, the objective parameter signal-to-noise ratio has been plotted along y-axis. The effect of continuous variation of the principal parameter is illustrated by plotting curves with x-axis representing the principal parameter. The effect of step variation of secondary parameter is illustrated by plotting several curves of the above type in each graph.
In all these experiments, pattern elements have been generated randomly with clipped Gaussian\textsuperscript{25} distribution to match natural distributions (such as image intensity). However the standard deviation has been varied to generate data with different asymmetry characteristics. Below, these experiments to determine each of the four parameter ranges namely (i) focus, (ii) training, (iii) loading, and (iv) data distribution ranges, and their corresponding results are explained.

**Focus Range:** For this experiment, first a set of holographs has been generated by encoding various numbers of patterns using different iterations. Then, by using a part of the originally stored patterns as query input to the holographs, the accuracy of retrieving the correct response pattern has been recorded. The focus strength has been varied by shutting off (setting $\lambda_i = 0$) specific number of random elements in the query pattern.

![FOCUS CHARACTERISTICS](image)

**FOCUS CHARACTERISTICS**

$n=1000, p/n=0.02, 0.04, 0.08, I=5$, $DV=3.0$

---

25 Because of the circular nature of phase space, only those random generations have been used which falls between 0 to 2 $\pi$. The phase values, otherwise has been generated with mean 0 and various standard deviations (sd).
Fig-5.18 shows a typical average recall error with smooth variation of focus strength of the query pattern (when other parameters are in quasi critical state). In this graph, the left y-scale shows the error as signal-to-noise ratio, and the right y-scale shows the performance in percentage of dynamic range. Fig-5.18 uses load factor as its secondary parameter. The three curves in this graph show the focus characteristics for three different load factors. Fig-5.19 and Fig-5.20 are almost similar plot except that these respectively use data distribution and training iteration as their secondary parameter.

A typical focus characteristics curve is monotonic and resembles a fat sigmoid. These curves generally demonstrate three distinct zones, (a) high-performance zone, (b) linear-zone, and (c) cut-off zone. The high performance zone corresponds to RCA type-U search performance. This zone is characterized with high accuracy above 30 db (or less than 2-3% magnitude error). After, the high-performance zone, the accuracy gracefully decays with
the focus strength. This is called as linear-zone. Finally, in the cut-off zone, the accuracy falls below 20 db (or more than 10% magnitude error). Fig-5.18 shows the encoding performance where patterns with n=1000 and SD=1.0 (k=.6) have been encoded with 5 iterations. It plots three situations for load factor L=.02, .04 and .08. Fig-5.19 shows the same characteristics for 3 different data distributions SD=.8,1.2, and 3, with other parameters remaining the same. Similarly, Fig-5.20 plots the performance for four different training iterations I=5,10,20,40 (all of these overlap, signifying the stability of training just after 5 iterations). These curves show typical zoning that has been observed with many other similar simulations. However the most important thing to notice in these graphs (from ORS objective) is the focus boundary of these regions. As can be observed in these plots, (a) the high-performance zone extends from f=1.0-0.9, (b) linear-zone extends from f=0.9-0.1, and cut-off zone is from f=0.1-0.0.
Training Range: For this experiment, several holographs have been generated each containing various numbers of patterns. Each of these is then tested after each encoding iteration using a part of the originally stored patterns as query inputs.

Fig-5.21 shows a typical training characteristics plot obtained in this manner. It plots the performance as a function of iteration (primary parameter). In this graph, the left y-scale shows the signal-to-noise ratio, and the right y-scale shows the error in percentage of dynamic range. The three sets of curves (secondary parameter) correspond to the three RCA type-A retrieval results with focus strengths \( f = 0.75, 0.50, \) and \( f = 0.25 \). These focus strengths correspond to the two boundaries and the mid-point of the linear-zone of operational focus range (as observed in Fig-5.18). A moderate load (\( L = 0.02 \)) with moderate asymmetric distribution (SD=1.0) has been used for this experiment.
A typical training characteristics curve shows that the performance of holograph improves as the training progresses. The improvement is steep during the first few iterations. After that, it gradually stabilizes. The flatness of the curves represents the stabilization. The performance (SNR) at stabilization varies with the focus strength. The higher is the focus strength, the better is the stabilized accuracy of retrieval (as can be expected from Fig-5.18 results).

**Fig-5.22 ORS(V)**

As demonstrated in Fig-5.21, holographic memory tends to converge (stabilize) within 5-10 iterations inside ORS. Experiments have shown that the number of iterations required to converge apparently is not sensitive to the increase in loading or skew in data-distribution. Fig-5.22 shows training characteristics at another boundary of ORS, with much higher loading (L=.08), and asymmetry (SD=3). As can be seen, in this case also, the holograph converges in 5-10 iterations.
These measurements show two important features of MHAC. First, that the convergence of MHAC is characteristically faster (less than 10) compared to that of conventional scalar valued ANNs (often more than 100). Secondly, MHAC is potentially scalable. The required number of iterations does not grow with training set.

**Loading Range:** To determine the ORS boundaries of the load factor range, first a pool of clipped gaussian patterns has been generated (all with a fixed length). Then, different holographs have been trained each time taking a different number of patterns from this pool. Each time, sufficient iterations have been performed to ensure the stabilization.

**LOADING CHARACTERISTICS**

![](Fig-5.23 ORS(VI))

Fig-5.23 shows usual loading performance. It plots the SNR (y-axis) against various load factors (x-axis) for three RCA type-A cases with focus strengths \(f=.75, f=.50, \) and \(f=.25\). The pattern set for this case has a typical asymmetry corresponding to standard deviation \(SD=1\) (\(k=.3\)).
As shown in this plot, a typical loading characteristics curve shows monotonically decreasing performance with increased load factor. Quantitatively, for $f=.25$, the RCA type-A retrieval accuracy drops to 20 db, while the load factor reaches .07. Typically, a load factor of .03 to .10 can be reached maintaining 20 db performance with 0.3-0.1 focus strength within this range space. This range defines the load factor boundary of ORS.

**Data Distribution Range:** The ORS boundaries of asymmetry parameter can be observed through the projection of range-space where the continuous parameter is asymmetry. To perform this experiment, several sets of patterns have been generated with varying standard deviations. These are then encoded into different holographs. Variation in the standard deviation (of clipped gaussian distribution) generates data sets with various asymmetries. The narrower is the deviation, the higher is the asymmetry. Later, some example will be given showing asymmetry values of some typical images.

![DATA DISTRIBUTION CHARACTERISTICS](image)

Fig.5.24 ORS(VII)
Fig-5.24 shows a typical data distribution characteristics plot. It plots the SNR (y-axis) against the primary parameter asymmetry (x-axis). Three RCA type-A test results with secondary parameter focus strength f=.75, f=50, and f=.25 have been shown by the three curves. For these experiments, each of these holographs has been loaded with a load factor =.02, and has been trained with 5 iterations.

Typically, as the asymmetry increases, the performance of MHAC decreases. As shown in Fig-5.24, MHAC can tolerate up to .6 asymmetry of the data distribution and can still maintain 20 db performance within the operational range-space.

Below, two important relevant questions are answered with experimentations.

**Question (hard training): Can hard training compensate for bad distribution?**

The answer is yes. If, the distribution is significantly asymmetric then hard training can improve the performance. Fig-5.25 plots the performance of an identical MHAC as characterized in Fig-5.24 except it has been trained with 100 iterations, instead of 5 iterations. It shows that the performance with >.6 asymmetry can be improved significantly with harder training. For all three curves, the asymmetric sides (towards the right) have been raised gradually by the added training. However, the improvement is more pronounced for retrieval with larger focus. Fig-5.26 plots the training characteristics of this data set. It shows that a distribution of about asymmetry k=.6 requires about 20 iterations to converge at its maximum performance state\(^{26}\).

\(^{26}\) However this result does not necessarily extend the training range of ORS beyond usual 5-10 iterations. Because, the symmetry of a data can be improved by preprocessing.
DATA DISTRIBUTION CHARACTERISTICS

\[ n=1000, p/n=.02, I=100, f=100-25\% \]

\[ f=.75 \quad f=.50 \quad f=.25 \]

Fig-5.25 ORS(VIII)

TRAINING CHARACTERISTICS

\[ n=1000, L=.02, SD=.5, f=.75, .5, .25 \]

\[ f=.75 \quad f=.50 \quad f=.25 \]

Fig-5.26 ORS(IX)
**Question (scalability):** Are the results obtained in the above plots applicable to larger holographic memory systems?

The above experiments have been performed with holographs of quite moderate size such as $n=1000$, and $p=100$. However there is theoretical as well as empirical evidence that the result does not depend on the scale. As analytically shown in this section, most of the performances are related to the load ratio $p/n$, rather than $n$. This result has been verified by repeated simulation for all the above plots for other scales. As expected, these results conformed that the performance does not depend on $n$. Fig-5.27 is such a graph. It plots the SNR performance against focus strength $f$ for $n=400,800,1600,3200$ and 6400. Although, the scale varies widely, these curves overlap on each other demonstrating the performance invariancy of focus characteristics with respect to the scale of $n$. Similarly, Fig-5.28 shows the scale invariance of training characteristics, and Fig-5.29 shows the scale invariancy of loading characteristics. As evident, all the plots are overlapping, suggesting almost perfect scalability.

![Focus Characteristics Graph](image)

**Fig-5.27 ORS(X)**
**TRAINING CHARACTERISTICS**

\( n=1000, L=0.2, SD=3, f=0.25 \)

![Graph of training characteristics](image)

Fig. 5.28 ORS (XI)

**LOADING CHARACTERISTICS**

\( L=10, SD=3, f=1.0 \)

![Graph of loading characteristics](image)

Fig. 5.29 ORS (XII)
5.4.3 Summary of ORS Boundaries

The quantitative results of ORS experiments are summarized in the following range table (table 5.1) suggesting an operating range-space for the holographic memory system. The range-space guarantees an accuracy in the range of 20 db. It shows the asymmetry, load factor, focus and iteration ranges required to achieve this target performance.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>Operational range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>SNR (db)</td>
<td>30-20</td>
</tr>
<tr>
<td>Asymmetry</td>
<td>k</td>
<td>0.0-.6</td>
</tr>
<tr>
<td>Load factor</td>
<td>L</td>
<td>0.01-.08</td>
</tr>
<tr>
<td>Focus</td>
<td>f</td>
<td>1.0-0.1</td>
</tr>
<tr>
<td>Training</td>
<td>I</td>
<td>5-20</td>
</tr>
</tbody>
</table>

In Table-5.1, accuracy is the target parameter. Asymmetry is a semi-controllable constraint in a given application. It represents the characteristics of data that is enfolded in the memory. It is possible to improve symmetry through various smoothing techniques as a part of data preprocessing (given in next chapter). Table-5.2 shows some asymmetry measure of few example images.
Table-5.2 Typical Asymmetry

<table>
<thead>
<tr>
<th>Image</th>
<th>$k$ (1st order)</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>lake.im</td>
<td>.23</td>
<td>512x512</td>
</tr>
<tr>
<td>tree.im</td>
<td>.19</td>
<td>256x256</td>
</tr>
<tr>
<td>lena.im</td>
<td>.38</td>
<td>256x256</td>
</tr>
<tr>
<td>house.im</td>
<td>.27</td>
<td>256x256</td>
</tr>
<tr>
<td>pepper.im</td>
<td>.17</td>
<td>512x512</td>
</tr>
</tbody>
</table>

Loading and training are two controllable design parameters. Loading is closely tied to the space efficiency of a MHAC. The dimension of a holograph is determined by the length of the stimulus ($n$) and response patterns ($m$). Load factor provides an estimate how many such patterns can be enfolded on a single holographic memory. Table-5.3 shows typical estimates on the number of patterns that can be stored (and queried) for few image sizes. Load factor is not necessarily a hard limitation. For patterns with a given $n$, the number of stored patterns $p$ can be increased by higher order encoding.

Table-5.3 Typical Memory Loading

<table>
<thead>
<tr>
<th>$n$</th>
<th>$L$</th>
<th>$p$ (1st order)</th>
</tr>
</thead>
<tbody>
<tr>
<td>160x120</td>
<td>.04</td>
<td>768</td>
</tr>
<tr>
<td>256x256</td>
<td>.02</td>
<td>1310</td>
</tr>
<tr>
<td>512x512</td>
<td>.02</td>
<td>5120</td>
</tr>
<tr>
<td>1024x1024</td>
<td>.01</td>
<td>10240</td>
</tr>
<tr>
<td>1024x1024</td>
<td>.02</td>
<td>20480</td>
</tr>
</tbody>
</table>

148
The above operational range-space provides a quick means for predicting the performance and estimating design parameters when ever a new application is considered. For example, if an associative memory with CT-scan images is to be constructed, all that is required is to estimate the asymmetry characteristics of the images. If, the asymmetry is within the range space \((k<.6)\), then it is possible to predict the required dimension and other parameters for the corresponding MHAC system. On the other, hand if \(k>.6\), even then it is possible to estimate how much smoothing is needed to obtain the target performance.

Appendix-F shows performance predictions for some typical image database search cases based on this range-space experimentation.

### 5.5 Chapter Summary

The chapter presents the principal characteristics of holographic associative computation. It shows both, analytically and experimentally, the superior capacity and accuracy of the memory. Even when the memory is used as a conventional associative memory of type-U, its capacity, convergence speed and accuracy surpasses the performance of most conventional AAMs. More, importantly, it also demonstrates the ability of MHAC to perform search with changeable attention on the element space. This is the unique capability of holographic memory. It also demonstrates the other unique aspect of this memory, its ability to convey its confidence of match or MNC characteristics.

Finally, this chapter also presents the results of extensive operational range-space experimentation for the proposed MHAC memory system which provides a reliable means for estimation and performance prediction of any application on MHAC, from engineering point of view.
6.1 Introduction

The objective of this chapter is to present the critical design issues involved in the process of translating the technique of holographic computing and its ability to perform associative retrieval with changeable attention, into a search mechanism for content based querying into image database. A prototype system called Content Associative Image Database Search System (CAIDSS) has been implemented for this purpose. This chapter presents the system and explains its critical design issues.

This chapter is organized as follows. It begins with an overview of the CAIDSS. The steps necessary to convert images into multidimensional complex numeric (MCN) representation is described next in section 6.3. The suggested methods of data preprocessing, index formulation, holographic encoding, and holographic memory maintenance steps are explained here. Section 6.4 finally describes the query formalism of CAIDSS that accepts users subjective interpretation of the query scene and performs direct content based search in the encoding of images.

A separate chapter (chapter 7) has been dedicated to a command based advanced query method, named as compound query formalism (CQF), which combines the fuzzy logical constraint based reasoning with direct associative search of CAIDSS.
The performance of CAIDSS has been characterized with several experimental image archives. The description of CAIDSS is illustrated appropriately with examples from prototype archive managed by CAIDSS. The computational complexities of the procedures involved in CAIDSS have also been derived.

6.2 System Overview

CAIDSS involves three principal components; (i) storage and access subsystem (SAS), (ii) associative encoding subsystem (AES), and (iii) associative query subsystem (AQS). Fig-6.1 presents the schematic.

The storage and access subsystem (SAS) is concerned with space efficient compact storage of the images and access speed. CAIDSS isolates the query mechanism from the image archiving and storage. As far as the search mechanism is concerned, whenever a user or the search mechanism is required to retrieve any image physically, it passes an appropriate registration index (called RLP) to the storage mechanism. Otherwise, it is functionally independent from the rest of the search system, and will not be discussed further.

The associative encoding subsystem (AES) is concerned with the generation and maintenance of the holographic archive abstract of the images. The principal responsibilities of it are (i) image pre-processing, (ii) associative learning of the image frame set, and (iii) maintenance of the holograph. Section 6.3 discusses each of these sub-stages in further details.

The associative query subsystem (AQS), allows a direct content based search into the image archive. The principal responsibilities of this subsystem are; (i) user interface, (ii) search optimization, and (iii) associative recall. The user generally specifies his/her expectation through a pictorial image example and a focus specification on it. It then translates them into a set of independent search templates and performs associative recall.
Fig-6.1 CAIDSS Overview
into the holographic archive abstract. It also communicates with the SAS through registration indices, when physical access to the images is demanded. Section 6.4 discusses each of these sub-stages in further details.

6.3 Encoding Process

The encoding process of CAIDSS is performed by its AES blocks. It involves the following major steps, (i) pre-processing and stimulus pattern (S) generation, (ii) Assignment of response label patterns (RLP), (iii) Memory training, and (iv) Memory characterization. These steps are explained below.

6.3.1 Stimulus Generation

This section presents the process of obtaining a complex stimulus pattern (S) from a given image. The principal design issues in this stage are the normalization and mapping of input data. For image application a frame of image corresponds to a pattern. Pixels correspond to the pattern elements of MHAC. Specifically, pixel intensity corresponds to the measurement component of elements.

Conventional ANNs employ a wide variety of pre-processing techniques on input data. Preprocessing serves two purposes: (i) domain knowledge based data-refinement, and (ii) network constraint based data-mapping.

Data refinement: With sufficient a-priori domain knowledge about the type of contaminant, bias in the input data, and the interest of the inquirers, it is often possible to semantically refine desired information content of input data. The objective generally is to remove the undesirable components and to enrich the desirable components. Bias filtering, gray scale manipulation [GoWi87], moment transformation [Hu1962, Kulk90], removal of deterministic trends [Mast94], etc. are examples of such special purpose semantic transformations.
The objective of CAIDSS search system is to provide a direct content based search mechanism into the pristine image frames. Therefore, within its limited scope, no such semantic refinement of input information has been performed. Nevertheless, the proposed search technique is quite general and is applicable on any semantically transformed data sets.

**Data mapping**: The second reason for preprocessing originates from the representational constraints and efficiency consideration of specific ANN architecture. It involves a mapping of the data. Normalization, scaling, centering, differencing, etc. [Mast94] are examples of such preprocessing. This step is important for multidimensional complex numerical representation (MCN).

The MCN representation specifically requires; (i) dynamic range of the input data distribution to be mapped into the 0-2π range, and (ii) sufficient symmetrical distribution. The first specification is fundamental to the proposed paradigm. The second constraint is intended for improving the storage efficiency. The effect of input data symmetry on MHAC performance has been explained already in chapter 5.

In the raw format, an image is represented by a two dimensional array of pixels with discrete intensity values in the range of 0-255. Generally, a typical image demonstrates normal distribution of intensity. AES uses the following spiral mapping transform to obtain symmetric distribution in the range of -π to +π. Fig-6.2 shows the spiral mapping function.

\[
\theta_i = m(I_i) = I_i \cdot \text{spread} \cdot \frac{2\pi}{I_{max} - I_{min}} - \pi
\]

For example, a typical Backpropagation works well if the discriminating ranges in the dataset corresponds to the non-saturation ranges of activation function. Or Kohonen network requires data to be mapped in 0-1 scale.
$k_{spread}$ is the spreading coefficient. It spirally maps intensity sub-ranges repeatedly over the 0 to $+2\pi$ range.

Besides (6.1), other functions can also be used as the mapping function $m()$. The only constraint is that the functions are required to be single valued and continuous. For discrete-valued input, it is required that each of the input measurements should be mapped uniquely. A normal distribution can be ideally transformed into a uniform distribution by using sigmoidal mapping [Suth90]. Histogram equalization [Kulk84] can also help in achieving uniform distribution. However, ideal mapping functions are generally dependent on the data (depend on the mean and standard deviation of each encoded image frame). But, mapping functions which are dependant of global distribution of data are ill-suited for RCA retrieval.

In the RCA scenario, one of the basic assumption is that the query image frame may differ considerably from the encoded one outside the region of focus. Therefore, global characteristics of an encoded image are expected to vary significantly from that of the query image. Therefore, such data dependent pre-processing leads to incoherent equalization.

![Fig-6.2 Spiral Spreading Function](image)
To maintain the uniqueness and equidistant distribution of mapping, the spiral spreading factors are chosen to be odd numbers (256 has only one prime factor 2). The function is simple and easy to compute. Fig-6.3 shows the shifted spreading mapping for $k_{\text{spread}}=7$. All quantization steps have a distance of .0245 radian, which is ideal (the plot shows only the points between 0 to .5 radian).

**SPIRAL INTENSITY MAPPING**

Fig-6.3 Spiral Mapping ($k_{\text{spread}}=7.0$)

**Input significance measure:** In the encoding stage, CAIDSS sets $\lambda=1.0$ for all elements. However attention based discriminatory encoding can be performed by allowing variation of $\lambda$ over the pixels of a frame. Such encoding is specially suitable for the applications where the relative importance of various sub-regions of encoded images is known a-priori. In such a scheme, the regions with higher matching importance are encoded with higher $\lambda$, and vice versa.
6.3.2 Response Label Associations

This section discusses how the Response Label Patterns (RLP) are generated. RLPs are internal index used by the Associative Query System (AQS) to identify images. The principal design issues are (i) the length of RLPs, (ii) and the assignment of RLPs.

**RLP Length:** The length of the RLP \( m \) is decided by number of image frames \( P \) and the quantization interval in the dynamic range of \( \theta \). Given a quantization interval \( q \), the length is given by:

\[
m = \log_{\left(\frac{\pi}{q}\right)}(P)
\]

...(6.2)

The quantization interval \( q \) depends on the expected loading of the holograph.

**RLP Assignment:** A holograph is generally designed for an expected maximum number of images. However at any one time, the holograph may contain less than the designed limit. A sequential assignment (such as assigning intervals 1,2,3,... to images 1,2,3) tends to asymmetrically overcrowd the dynamic range of \( \theta \). Therefore, a Reverse Modulus Weight code (RMWC) is used to assign RLPs to the images. Fig-6.4 shows a valid assignment sequence. Let us consider a case by assigning only 1 RLP element which has a dynamic \( \theta \)-range from 0-2\( \pi \). Let the circular \( \theta \)-range be divided into 8 intervals, each being \( \pi/4 \). According to RMWC, the first and second images are assigned the intervals 1 and 4, instead of 1 and 2. As a result, when a holograph is loaded with only 2 images, the effective interval becomes \( \pi \) instead of the designed \( \pi/4 \). When more images are to be encoded, accordingly next two images are assigned intervals 3 and 7, to maintain the symmetrical balance.

One way of obtaining such symmetric assignment is to maximize the distance between each of the consecutive code pairs. Thus, RMWC alters the maximally weighted digit. RMWC is a symmetric mapping. If \( u \) is a sequence number then it is mapped to a polar phase \( v \) by the following transform:
RMWC code provides the phases of the RLP elements. The magnitudes of all $\lambda$ are set to 1.0. Table-6.1 shows a code conversion.

$$u = a_2 d^2 + a_1 d^1 + a_0 d^0 \Rightarrow (a_2 d^2 + a_1 d^1 + a_0 d^0)^* \frac{2\pi}{d} = v$$  

...(6.3)

### Table-6.1 A RMWC Coding Example

<table>
<thead>
<tr>
<th>$u$</th>
<th>$a_2$ $a_1$ $a_0$</th>
<th>$a_2$ $a_1$</th>
<th>$v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>000</td>
<td>000</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>001</td>
<td>100</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>010</td>
<td>010</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>011</td>
<td>110</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>001</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>101</td>
<td>101</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>110</td>
<td>011</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>111</td>
<td>111</td>
<td>7</td>
</tr>
</tbody>
</table>
6.3.3 Training Algorithm

The training algorithm involves three stages, and corresponds to equation (4.14) presented in chapter 4. In the first stage, the stimulus \( S \) is decoded. Then a difference between the recalled and expected RLP is estimated. Finally, an association between the difference response and the stimulus is computed and superimposed on the existing holograph. The algorithm is presented below:

1. Set \( t = 0; \ iter = 1 \)

2a. Read the \( t \)th RLP \( RR = \text{ReadPattern}(t) \);

2b. Read the \( t \)th image \( IMG = \text{ReadImage}(t) \);

3. Convert the Image \( IMG \) to stimulus pattern \( S \).

4a. Decode for \( R \) by applying \( S \) to \( H \).

4b. Compute the difference \( R_{\text{diff}} = RR - R \).

4c. Generate the difference association \( A_{\text{diff}} = ST.R_{\text{diff}} \).

4d. Superimpose the association on the holograph \( H_{\text{new}} = \alpha H_{\text{old}} + A_{\text{diff}} \).

5. Set \( t = t+1 \); Repeat from step 2 until all images are encoded.

6. Compute average recall error \( E \) for all images.

7. Set \( t = 0; \ iter = \text{iter}+1 \); Repeat from step 1 until the average \( E < \text{EPSILON} \).

Fig-6.5 and Fig-6.6 present the training performance and recall convergence of individual associations for a prototype holograph. It encodes 64 color images of size 160x140 pixels. It has been trained with 30 iterations (although it converged in 10 iterations). Fig-6.5 shows the convergence of average recall SNR and average MNC as the training progressed. Fig-6.6 plots the type-U retrieval accuracy (left y-axis) and MNC (right y-axis) of SEQ001 test archive for each of the individual images (x-axis is image index).
TRAINING SUMMARY

IDB: SEQ001

Fig-6.5 Training Characteristics for SEQ001

RLP RECALL TRACE

IDB: SEQ001

Fig-6.6 Call Characteristics for SEQ001
6.3.4 Holograph Characterization

The status of holograph is tracked by two tests known as ZOOM test, and SCAN test. ZOOM test probes the distinguishability among the encoded images by applying focus windows of decreasing dimensions at the center of the template. SCAN test on the other hand probes the distinguishability at various spatial locations with a fixed size focus window. Fig-6.7 illustrates the windowing used in these tests.

![ZOOM TEST](image1)

![SCAN TEST](image2)

Fig-6.7 Characterization Tests

Fig-6.8-6.11 presents ZOOM and SCAN test results for the prototype archive SEQ001. The MATCH graph shown in Fig-6.8 demonstrates the accuracy of each individual recalls in ZOOM test. Each jump signifies a wrong detection. As shown, it has been possible to retrieved each of these images when window is 25x25 pixels wide. Even, when the window dropped to 15x15 there has been only one wrong detection (image 47) out of 64 images. The performance graph in Fig-6.8 demonstrates the average response over all the images for various focus window sizes. The bar plot (left y-axis) shows the average accuracy. The line plot in the inset (right y-axis) shows the attention strengths of the windows.
Similarly Fig-6.10 and Fig-11 respectively plot the result of final match, and average performance for SCAN test, which probes the distinguishability at nine spatial locations with fixed 20x20 pixels wide focus window. The legends of Fig-6.10 show the exact window centers. As can be seen, the distinguishability among the images varies at various spatial locations.

**ZOOM-TEST MATCH**

*Fig-6.8 Pattern Distinguishability in ZOOM test*
ZOOM-TEST PERFORMANCE

PAN TEST for HIDB SEQ001 and SEQ002 (64 Images)

Fig-6.9 ZOOM test for SEQ001 Archive

SCAN-TEST MATCH

HIDB SEQ001

Fig-6.10 Pattern Distinguishability in SCAN test

163
6.4 Query Process

This section presents the query mechanism for direct associative search into CAIDSS. AQS block is responsible for query processing in CAIDSS. The formalism is based on the object models defined by the user (interrogator). Users’ expectation is conveyed to the AQS through a combination of pictorial examples, focus shields and logical connectives.

In the first level, the user specifies a set of elementary objects through defining a focus shield on various query image frames. The user can construct complex objects by logically combining these elementary objects.

The following sub-sections one by one discuss (i) the procedure for defining the elementary objects, (ii) specification of complex objects, and finally, (iii) the search mechanism.
6.4.1 Segmentation and Elementary Objects

An elementary object \( OBJ \) in any image is specified by an image and a mask pair \([I,M]\). \( I \) represents the pictorial example. The mask represents the object definitions inside this example image.

A mask is a matrix which has the same dimensions as that of the image. Mask elements are assigned values between \([0,1]\). A value 1.0 in mask pixel \( M(x,y) \) indicates that the image pixel \( I(x,y) \) is a member pixel of the object \( OBJ \). In the other extreme a value of 0.0 indicates that corresponding image pixel does not contribute to the concept specified. In general the mask values are analog, although in elementary mask they are generally assigned a uniform value. At the very beginning the user is required to specify the elementary masks.

<table>
<thead>
<tr>
<th>object</th>
<th>red</th>
<th>green</th>
<th>blue</th>
<th>x</th>
<th>y</th>
<th>density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nbody</td>
<td>180,3</td>
<td>150,30</td>
<td>20,0</td>
<td>141,12</td>
<td>92,48</td>
<td>0.056</td>
</tr>
<tr>
<td>Nhead</td>
<td>80,0</td>
<td>90,0</td>
<td>60,1</td>
<td>137,112</td>
<td>110,97</td>
<td>0.015</td>
</tr>
<tr>
<td>Narm</td>
<td>60,0</td>
<td>100,20</td>
<td>80,0</td>
<td>111,94</td>
<td>93,63</td>
<td>0.013</td>
</tr>
<tr>
<td>Nlam</td>
<td>60,0</td>
<td>100,20</td>
<td>80,0</td>
<td>159,141</td>
<td>97,66</td>
<td>0.019</td>
</tr>
<tr>
<td>Neyes</td>
<td>200,50</td>
<td>90,0</td>
<td>90,0</td>
<td>135,116</td>
<td>117,110</td>
<td>0.005</td>
</tr>
<tr>
<td>Nleg1</td>
<td>60,0</td>
<td>100,20</td>
<td>80,0</td>
<td>122,104</td>
<td>60,38</td>
<td>0.012</td>
</tr>
<tr>
<td>Nleg2</td>
<td>70,0</td>
<td>100,20</td>
<td>80,0</td>
<td>150,134</td>
<td>64,46</td>
<td>0.012</td>
</tr>
<tr>
<td>Nleg2</td>
<td>60,0</td>
<td>100,20</td>
<td>100,0</td>
<td>117,104</td>
<td>37,27</td>
<td>0.004</td>
</tr>
<tr>
<td>Nknee</td>
<td>180,30</td>
<td>50,0</td>
<td>100,10</td>
<td>112,102</td>
<td>41,34</td>
<td>0.002</td>
</tr>
<tr>
<td>Nknee</td>
<td>180,30</td>
<td>50,0</td>
<td>100,10</td>
<td>153,141</td>
<td>47,39</td>
<td>0.003</td>
</tr>
<tr>
<td>Choy</td>
<td>255,80</td>
<td>180,80</td>
<td>120,20</td>
<td>144,122</td>
<td>41,24</td>
<td>0.008</td>
</tr>
<tr>
<td>Ccar</td>
<td>210,50</td>
<td>30,0</td>
<td>100,10</td>
<td>148,98</td>
<td>28,3</td>
<td>0.041</td>
</tr>
<tr>
<td>Shody</td>
<td>255,0</td>
<td>255,0</td>
<td>10,0</td>
<td>105,0</td>
<td>55,0</td>
<td>0.164</td>
</tr>
<tr>
<td>Grass</td>
<td>200,0</td>
<td>220,100</td>
<td>170,40</td>
<td>159,0</td>
<td>22,0</td>
<td>0.039</td>
</tr>
</tbody>
</table>

This is performed by specifying a rectilinear window as filter. The mask specification filter for CAIDSS requires 10 parameters to specify space (\( x_{\text{max}}, x_{\text{min}}, y_{\text{max}}, y_{\text{min}} \)) and color plane (\( r_{\text{max}}, r_{\text{min}}, g_{\text{max}}, g_{\text{min}}, b_{\text{max}}, b_{\text{min}} \)) range boundaries. Table-6.1 shows a set
of filters for objects segmented from the scene XQ1 in Fig-D(a) (Appendix-D) as per interpretation of Fig-D(b). A second user is entitled to subscribe a totally different interpretation of the same scene. However such flexibility must be complemented with the appropriate object filter table. Mask density refers to the attention strength of individual segmented objects.

\[
\rho = \frac{\sum_{i} \lambda_i}{n}
\]

...(6.4)

CAIDSS use the above described relatively simple space-intensity window based segmentation tool. However more sophisticated image segmentation tools can be used to improve the user convenience, as well as accuracy while specifying the filter, with the objective of reducing user's routine efforts. Image segmentation itself constitutes a major area of current image understanding research.

6.4.2 Specification of Complex Objects

Once the elementary objects are specified, more complex objects can be defined using logical connectives. The following are a few examples of constructing complex objects:

\[(Q1.OBJ1) \cup (Q2.OBJ2) \rightarrow (Q3.OBJ3)OBJ3\]

\[(Q1.OBJ1) \cup (\text{NOT} \ (Q2.OBJ2)) \rightarrow (Q4.OBJ4)\]

In this formalism OBJ1, OBJ2 are component objects and OBJ3 and OBJ4 are complex objects. Q1, Q2, ... are fuzzy quantifiers. These are assigned values in the range of [0,1]. In reality they can be assigned linguistic multi-valued quantifiers such as HIGH, LOW, etc. (with membership function set which can translate the quantifiers into analogue values).

The logical union, intersection and complementation operations are defined according to fuzzy set functions given in equation-6.4.
\[ f_{A \lor B}(x) = \max(f_A(x), f_B(x)) \]
\[ f_{A \land B}(x) = \min(f_A(x), f_B(x)) \]
\[ f_{\bar{x}}(x) = 1 - f_A(x) \]

The elementary objects can be considered as composed of (union of) individual pixels. Therefore to compute the mask of the complex object, each of the values in the elementary masks is first multiplied by the corresponding membership value, and then the fuzzy inference rules (6.4) are applied to obtain the mask values of the complex mask. Fig-6.12 shows an example of computing a composite mask (c) from the elementary masks (a) and (b).

![Fig-6.12 Complex Mask Composition](image)

**6.4.3 Search with Complex or Elementary Objects**

The search process involves the following three stages: (a) generation of query template pattern, (b) computation of recalled response label pattern (RLP), and (c) matching the recalled RLP with the response pattern set. Each of these stages is described now.

**Query template generation:** The elementary objects are defined with a pictorial image and a mask [I,M] pair. These pairs are used to translate a query template into a query pattern with complex elements. The transformation for phase elements is identical to the encoding process. The only difference is that now the magnitudes of the complex elements are not 1. Instead, the mask elements are used as the magnitude of the pattern elements. Once, the
template patterns are obtained the search process involves associative recall computation and noise filtration. Fig-D(c) of appendix-D shows an example where three different templates have been generated with three subjective focus windows specified by a user.

**RLP computation:** The recall response is computed by equation (4.5). Computationally the process is equivalent to an inner product between the holograph and query stimulus.

**RLP matching:** The recalled RLPs obtained from previous stage generally are approximations of the actual RLPs and are contaminated by the crosstalk components. Therefore, a linear search is performed among the RLPs. The search produces a list of matching RLPs in order of their numerical proximity. This ordered list is the possible range of answers. The following RCA matching criterion is used to evaluate the distance:

$$D(R^Q, R^n, A^Q) = \left[ \sum_i^N \lambda_i^Q \text{dist}(r_i^Q, r_i^n) \right]$$  \hspace{1cm} \ldots (6.5)

Even at this level, the matching criterion has the property, that it puts more emphasis on the RLP elements which demonstrates higher recalled magnitude.

Fig-D(d), (e), and (f) respectively shows the three matched images retrieved from holograph based on the three focus regions specified by the user in Fig-D(c). Table-6.3 shows the accuracy of retrieved RLP corresponding to these templates. The, rightmost column shows that despite some error in the raw retrieved RLP, after noise suppression, each of the queries has been satisfied correctly in the first guess.
Table 6.3 Retrieval Performance

<table>
<thead>
<tr>
<th>Object</th>
<th>Density</th>
<th>SNR (db)</th>
<th>Correct Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-PATCH-OF-BKGR D</td>
<td>.108</td>
<td>9.73</td>
<td>1st (A1)</td>
</tr>
<tr>
<td>POND</td>
<td>.208</td>
<td>24.37</td>
<td>1st (A1)</td>
</tr>
<tr>
<td>SIMBA</td>
<td>.193</td>
<td>19.10</td>
<td>1st (A4)</td>
</tr>
<tr>
<td>NINJA</td>
<td>.144</td>
<td>16.93</td>
<td>1st (A6)</td>
</tr>
<tr>
<td>FRED-ON-CAR</td>
<td>.039</td>
<td>16.43</td>
<td>1st (A5)</td>
</tr>
<tr>
<td>A-PATCH-OF-JUNGL E</td>
<td>.09</td>
<td>10.65</td>
<td>1st (A7)</td>
</tr>
</tbody>
</table>

6.5 Search Complexities

**Time Complexity:** For CAIDSS, the RCA search complexity with \( p \) patterns (images) each with \( n \) elements (pixels), and RLP length \( d \), is \( O(d.n) = O(n \log p) \).

Derivation: Let us consider, that each image has \( n \) pixels, and there are \( p \) such images. Let us also consider that the length of RLP is \( m \). The search process involves (a) computation of pattern, (b) holographic decoding, and (c) RLP matching. The complexities of the corresponding stages are:

- The cost of pattern computation = \( O(n) \).
- The cost of holographic decoding = \( O(nm) \). This is an inner product matrix operation involving complex matrix multiplication.
- RLP matching = \( O(mp) \). It is a linear search with relatively very small pattern length.

Thus, the cost is negligible.

Generally, \( m = O(\log p) \). Thus the complexity of the overall search process is.

\[
O(n) + O(n \log p) + O(p \log p) = O(n \log p)
\]  \[\text{...(6.6)}\]
**Space Complexity:** For CAIDSS, the RCA search with \( p \) patterns (images) each with \( n \) elements (pixels), and RLP length \( d \), requires \( M = m \cdot d \cdot (n + p) \) space for holographic encoding.

Derivation: Let us consider, that each image has \( n \) pixels, and there are \( p \) such images. Let us also consider that the length of RLP is \( m \) and that each complex element requires \( d \) bytes for representation.

The space, required by the holograph is \( m \cdot n \cdot d \). Some additional space is also required by the RLPs. Which is \( p \cdot m \cdot d \). Thus the total space requirement is:

\[
M = m \cdot d \cdot (n + p) \quad \text{...(6.7)}
\]

In practice \( p = n \), and 4-12 bytes are sufficient for images with 256x256x256 full colors.

### 6.6 Chapter Summary

This chapter has presented the detailed schemes for an attention based search system. This chapter specifically presented the design issues such as the techniques for data preparation, RLP assignment, dynamic characterization techniques for holograph, and RLP matching techniques for noise reduction at the end of holographic search. In short, this chapter fills the gap between the holographic computing mechanism presented in earlier chapters and an effective mechanism to perform content associative search based on user specifiable attention.
CHAPTER 7

COMPOUND QUERY

7.1 Introduction

The holographic method emphasizes the flexibility of accommodating the user's subjective interpretations in the query process, instead of hard encoding users' subjectivity into any intermediate rigid representation during storage of image information. This increased emphasis on user oriented dynamic image interpretation capability suggests the need for an advanced user definable platform for specification, interpretation, and reasoning during query. The objective of the work presented in this chapter is to enhance the CAIDSS model towards a sophisticated query platform, which is named as Compound Query Formalism (CQF). The query method described earlier in section-6.4 is a special case of the CQF formalism.

The reasoning and interpretation framework of human users are generally symbolic in nature. In addition to the symbolism, the human interpretation of image information is also ambiguous, vague and imprecise from the perspective of classical black and white logic\(^{28}\). Such, bi-valued symbolic logic appears brittle while managing information with such undulating characteristics. Therefore, CQF uses fuzzy logic [Zadeh78] based symbolic inference mechanism to model complex queries, although the underlying search is asso-

---

\(^{28}\) As pointed out in the introductory chapters, classical system of bi-valued logic enjoyed a great popularity among the early IDB researchers just because it was readily available. Nevertheless, it is ill-suited to manage image information. One of the key reasons, that the conventional archive technology fails to manage image information, perhaps, lies right in their inference mechanism which is entirely based on the rigid system of bi-valued symbolic logic.
ciative in nature. The proposed fuzzy logic based CQF not only provides an efficient mechanism for inference with imprecise information [KBGV93], but also allows intervention at specific stages of the overall inference process to accommodate subjectivity.

The CQF formalism can be considered as a two tier process. On the outer level, CQF supports complex associative search operations in the image archive through a query language. On the inner level, it also enables user to attain flexible control over the interpretation and inference process that translates his/her query language into the actual associative search. Such a platform can be viewed as a 'soft' inference mechanism, as opposed to the conventional inference mechanism, which is 'hard' in the sense that the user has very little control over the pre-encoded inference process. These two tiers of CQF together act as a query shell over the holographic content associative search kernel. Fig-7.1 shows the conceptual organization the overall query process.

CQF consists of (i) a compound query language (CQL), and (ii) an inference mechanism. This chapter first explains the language and then explains the inference mechanism.

The chapter is organized as follows. First, section-7.2 briefly outlines the CQF formalism. Section-7.3 then describes the inference mechanism that supports CQF. Section-7.4
specifically addresses the intervention points of CQF which allows the user to define his own subjective framework of reasoning with flexibility. Finally, section-7.5 illustrates the query and inference process through a set of examples.

In the subsequent discussion the term 'decoding' will be used only to refer to the memory decoding operation specified by equation (4.5), and the term 'search' will refer to the more general AQS process which involves one or more 'decoding'.

7.2 Compound Query Language

A typical compound query in CQF characteristically handles initiation of multiple decoding into holographic memory. Need for multiple decoding originates from query which is based on complex logical inferences or involves some form of transformation. Thus, the compounding process can be of two types; namely (a) logical compounding, and (b) transform compounding. These are explained below:

7.2.1 Logical Compounding

A query can be logically compounded by integrating logical connective AND, OR and NOT. The specification language allows logically compound queries to be chained together. The query language is composed of (i) objects, (ii) logical connectives, and (iii) MNC linguistic quantifier (MNC-LQ).

Format: The query formalism is of the following format:

\[ \text{COM-QUERY: (MNC-LQ.OBJECT).LOGICAL-CONNECTIVE. (MNC-LQ.OBJECT)} \]

Example: The following is an example of logically compounded query. It defines its complex query objective OBJ-NINJA as:

\[ \text{OBJ-NINJA: (HIGH.NINJA-HEAD).OR.(HIGH.NINJA-CHEST)} \]
This example searches for object NINJA in the archive. The above command specifies that the presence of compound object NINJA can be confirmed by the strong presence of either, NINJA-HEAD or NINJA-CHEST. The quantifier HIGH is a linguistic quantifier of the required detection strengths of NINJA-HEAD and NINJA-CHEST objects to confirm the presence of NINJA.

**Explanation:** The query generally invokes independent search corresponding to each of the component objects. The result of each search is given in a *Search Summary Table* (SST). Each SST summarizes (i) the occurrence index, and (ii) corresponding MNC. The possibility of occurrence of the compound object is computed by using these MNC values, and min, max, complement rules of fuzzy logic.

Corresponding to each MNC quantifier MNC-LQ, there is a fuzzy set function \( y = f_{MNC,y}(s) \), which maps the MNC strength \( s \) of the component object to the MNC strength of the quantified concept object \( y \).

A final SST corresponding to the query is compiled, which sorts the occurrences in order of final MNC values.

### 7.2.2 Transform Compounding

The mechanism of transform compounding allows users to perform searches involving variations based on transformations. Transforms are used when the query image differs from the expected image in some procedurally definable abstract sense (such as the sought object may be approximately half the size than the sample). It is also used when searching in based on invariancy with respect to some procedurally definable abstract feature (such as size, translation, size). Translation in space, rotation, all are examples of transforms. In general, the nature of transforms are highly abstract and domain dependent. For example, a spatial translation on an image is fundamentally different when the reality of the space depicted in
the image is 3-D instead of 2-D. Same is true of lighting, intensity and other invariances. Therefore the CAIDSS approach is to allow users to define the transforms. The mechanism of transform compounding permits the user to simulate individualized realities with his domain knowledge.

Each transform has a set of transform parameters to quantify the abstract concept it defines. The general formalism of transform compounding accepts a specification of a range of these transform parameters on a given object. Each query is composed of (i) objects, (ii) transforms (XFORM(parameter-range)), (iii) transform-relations (XFORM-RELATION), and (iv) transform-relation-quantifiers (XFORM-REL-LQ or LQ).

**Format:** The format is specified below (here p,q and r are indices to the keywords):

```
COMP-OBJ:  (OBJECT1(XFORMp(range-parameters)). XFORMp-RELq-LQr.XFORMp-RELq.
(OBJECT1(XFORMp(range-parameters))
```

**Example:** Below, an example of transform query is provided (the prepositions are just for the ease of reading and can be ignored):

```
HUMAN: (NOSE(in LOCATION(middle of the picture)) JUST ABOVE (MOUTH( in LOCATION(lower half of the picture))
```

In this example, LOCATION is a transform (SIZE, CONTRAST, etc. are also other possible transforms). The range-parameters of LOCATION is generally a region in the picture frame, where the object NOSE is expected to exist. ABOVE is a relation concept.

---

29 The concept of 'translation invariancy' (or as such any other invariancy) is deceptively simple. In fact, the effect on image, and therefore, the subsequent search process for each of these transforms can be widely different depending on the reality of the space which is depicted in the image. Therefore, there may not be any single technique which might ever be able to perform searches which can be labelled generically as 'rotation invariant' or 'translation invariant'.

175
which is associated with LOCATION (LEFT, DISTANCE, are possible relations which can be associated with the LOCATION transform). JUST is a linguistic quantifier associated with the ABOVE concept (FAR is another example of such quantifier).

**Explanation:** A transform (such as XFORMp) represents any pre-defined processing of the query stimulus pattern associated with an object [I,M]. Transforms require parameters. However, in the compound query, instead of any particular value for the parameters, a range is specified. The compound search engine automatically generates a series of search patterns for each parameter value within this range. (The optimization interface allows the specification of resolution).

The results are presented as a **Search Summary Table** (SST). Each SST specifies (i) the occurrence index, (ii) the transform parameter for which a match has been found, and (iii) corresponding MNC of search.

Each transform generally, comes with several transform relations (such as XFORMp-RELq). These concepts connect corresponding transform parameters. Also, each of these relations comes with a set of linguistic quantifiers (XFORM-RELq-LQr).

In effect, each transform-relation is an evaluation of a relation strength measure (s). The evaluation function operates on the transform parameters listed in the search summary of the associated objects.

Corresponding to each of the transform-relation-quantifier values, there is a fuzzy set function\(^{30}\) \(y=f_{p,q,r}(s)\), which maps the relation strength \(s\) to the strength of the quantified concept \(y\).

---

\(^{30}\) As evident by now, the MNC obtained from decoding and directly used during logical compounding is a special transform relation, purporting the concept of existence.
7.2.3 The Language (CQL)

Now, the complete grammar for Compound query language (CQL) is given. The basic types are (i) OBJECT, (ii) XFORM, (iii) XFORM-RELATION, (iv) linguistic quantifier LQ, and can be compounded with logical operators.

\[ \text{object} \leftarrow \text{OBJECT} \]
\[ \leftarrow \text{object(tranform)} \]
\[ \leftarrow \text{LQ.object} \]
\[ \leftarrow \text{object.} \land \text{object} \lor \text{object} \leftarrow \text{object} \]
\[ \leftarrow \text{object.relation.object} \]
\[ \text{relation} \leftarrow \text{LQ.XFORM-RELATION} \]
\[ \leftarrow \text{relation.} \land \text{relation} \lor \text{relation} \leftarrow \text{relation} \]
\[ \text{transform} \leftarrow \text{XFORM(parameter - ranges)} \]
\[ \leftarrow \text{transform.} \land \text{transform} \leftarrow \text{transform} \]

Grammar set (a) provides the compounding rules for objects. Objects can be compounded directly or through compound relations. However the direct logical compounding of objects can be interpreted as a special case of transform-relational compounding with a default transform-relation purporting the sense of existence with a linear evaluator function \( s=MNC \). Grammar set (b), and (c) respectively provide the rules for compounding relations and transforms. In the syntax, the LQ and transform-range can be omitted, implying a pre-declared corresponding default value.

7.2.4 Few Examples of compound queries

**Spatial invariant reasoning:** A host of predefined transforms can be used to obtain templates from the given object template. Below two transform examples are given, corresponding to the spatial concepts distance and size:
OBJECTIVE1: $OBJECT1(\text{ in REGION}(x_{\text{max}},x_{\text{min}},y_{\text{max}},y_{\text{min}}))$

OBJECTIVE2: $OBJECT2(\text{ of SIZE}(s_{\text{max}},s_{\text{min}}))$

When no transform is specified a default parameter is assumed. The default values are defined during the definition of original search. For example:

$OBJECT1 = OBJECT1(\text{ in REGION}(0,0,0,0) \text{ of SIZE}(1,1))$

However the declaration of following default results in a search which is always somewhat flexible from the very definition allowing for little tolerance.

$OBJECT1 = OBJECT1(\text{ in REGION}(-.01,.01,-.01,.01) \text{ of SIZE}(.99,1.01))$

**Mixed logical and spatial reasoning:** Examples below mix logical compounding with transform search:

$COBJECT3: \text{ HIGH OBJECT1( in REGION}(x_1,x_2,y_1,y_2))$

$COBJECT1: \text{ (HIGH OBJECT1( in REGION}(x_1,x_2,y_1,y_2)) \text{ NEAR of (OBJECT2}(x_1,x_2,y_1,y_2))$

$COBJECT4: \text{ (HIGH OBJECT2) FAR LEFT of (OBJECT2 in REGION}(x_1,x_2,y_1,y_2))$

**7.3 Inference Mechanism for Compound Queries**

The inference mechanism for processing compound queries is shown in Fig-7.2. In this process the compound search command issued by the user is sent to the search controller unit. The search controller unit decomposes the command into four components. These are respectively, (i) canonical object specifications, (ii) transform specifications, (iii) decoder specifications, and (iv) quantifier specifications.

Processing of compound query involves three stages:

1. Generation of templates.
3. Assimilation of match.
Fig-7.2 Compound Query System
**Generation of templates:** In the first stage, search controller reads the object specifications in the command and isolates the component objects which are to be searched independently to satisfy the query. The pattern generator generates these canonical patterns from the object base. The objects can be either elementary or complex and are expected to be defined by the user prior to initiating the search in the object base.

**Matching of templates:** In the second stage, each canonical pattern is replicated into a set of templates which are parametrically translated from the canonical pattern. The parametric translations are generated using pre-declared transforms from transform library. The transform driver performs parametric translations according to the range specification provided by the search controller. In general, one transformed set of pattern templates is obtained from each of the elementary canonical patterns.

The next step is to perform associative decoding. The tolerance, and thresholds during decoding can be controlled by the decoder specifications set by the search controller. Each query pattern template in this set can be decoded independently. The process corresponds to AQS that has been described earlier in section-6.4. The associative decoder logs its result in the Search Summary Table (SST). It prepares one SST for each canonical search pattern. The search summary table contains all suggested matches for the canonical form. It lists (i) the suggested matching RLPs, (ii) corresponding MNC, and (iii) the parameter values, at which the match has been registered.

**Assimilation of match:** The final stage is the logical assimilation of the information obtained in the SSTs. In this stage, first, the transform-relation evaluator functions (TREF) are invoked to compute the strength ($s$) of the transform relations.
Each of the transformed-relation concepts is linguistically quantified into several quantifier levels or fuzzy sets (such as, HIGH, MODERATE, LOW). For, a given s, the membership strength of each quantified level can be determined by the fuzzy sets membership functions (FSMF). These functions return the strength of the quantified-transform-relations. One function is required for each linguistic quantifier level for each relation.

The individual membership strengths are gradually assimilated to satisfy the logical constraints specified in the command (and extracted by search controller) using fuzzy inference rules. Standard fuzzy transformation functions and operations are used to compute the confidence values for each possible answer. The inference results are presented in the assimilated SST which lists the possible occurrences of the query, and corresponding MNC values. The assimilated SST is presented to the user interface, sorted in accordance of the computed MNCs.

7.4 Modeling of Subjective Reasoning

The incoherency problem originating from such subjectivity has been identified and discussed in section-3.4.

Transitory basis of inference: The inference mechanism of CQF provides a flexible environment for users to set up his/her own basis for the reasoning process. Conceptually, a search involves the following steps, (a) a basis concept, (b) specification of the concept in a query format, (c) search for the concept in image archive, (d) measurement of concept strength, and (e) verification of the measurements with expectation stated in the query. In the context of image information, each of these steps is inherently subjective in the sense of concreteness of definitions and precision of measurements. In contrast, crisp symbolic
information can assume much firmer objectivity in both of the above senses and thus, a
search mechanism can assume almost a stationary basis of inference. But, for image
information these are more like assumptions and are characteristically transient.

CQF is an experimental search formalism that attempts to include and manage such
transiences. It assumes much less concreteness of the underlying definitions and measure­
ments than the symbolic query languages used by other IDBs. It is specially alert in offering
the users as much flexibility as possible to customize, alter or redefine these basis concepts
and in providing subsequent inference support from such nonstationary 'tectonic' basis.

Subjectivity and CQF: To reformulate the incoherency problem in terms of CAIDSS
types, the subjectivity, that is inherent in the image information, is principally associated
with the multiplicity and versatility of (i) the possible definitions of basis objects, (ii) their
classification and compounding, (iii) relational concepts, and (iv) the quantifications
associated with each of them.

Even, the terminologies used in the specification language of quantification may
purport different meaning under different context. Interpretations differ not only between
querying entities, but may also differ between image to image.

Control points: CQF offers user definability at each of the above four potential sources
of subjectivity. The very first step of customization involves the definition of the objects.
This process is an overt pre-requisite to the query language. Section-6.4.1 already explains
how user can define objects (both basic and compound). It will not be repeated here. Current
section explains the other three intervention points of the compound query processing system
(Fig-7.2) which allows modeling of users’ expectations with greater sophistication. These
control points are:

(b) Transform (XFORM) generation routines.

(c) Transform-relation (XFORM-REL) strength evaluation routines.
(d) Transform-relation-quantifier (LQ) membership functions.

Current section explains the definition process of these routines. Together, these processes constitute the inner tier of CQF formalism, which makes the ‘assumptions’ underlying the interpretation and inference of outer tier (CQL) programmable.

**Transform customization:** The first step of the customization process is the definition of *transform routine*. It can be viewed as a routine which obtains a specific set of parameters during a search. These parameters appear as a column in the SST. For example, if a user wants to perform search on the notion of distance (or size) between objects he must design the suitable transform like location (or scale), which can measure the relevant parameters for the intended concept during a search. In fact, a single transform can be used to measure parameters for several close transform-relations.

**Evaluator customization:** In the second stage, a *transform-relation evaluator function* finally compares these SST parameters to actually detect and measure the *strength* of the transform-relations on a numerical scale. In fact, user quantifies a transform-relation by defining evaluation routine.

**Evaluation scale customization:** The next step involves interpretation of these measurements. User can specify the expected quantities using linguistic quantifiers along with various search constraints in the query invocation command. The numerical interpretation of these linguistic quantifiers can be customized by modifying the *membership functions*. The process involves adjusting the shapes of fuzzy membership functions corresponding to each linguistic quantifier on the scale of corresponding evaluator *strength* parameter.
7.5 Examples of Inferences

This section illustrates the process of fuzzy-logic based compound query processing described in section-7.3 through few examples.

7.5.1 Example #1

\[ OBJECT3: \ (HIGH.\text{OBJECT2}).\text{AND}.(\text{NOT}.(HIGH.\text{OBJECT2})) \]

In the first stage, the search is decomposed into two independent searches, involving the two canonical objects OBJECT1 and OBJECT2. There is no transform involved. So, these canonical patterns are directly passed on to the associative decoder as templates.

The SSTs return the RLPs and corresponding MNCs of detection. The MNC value returned by the search mechanism provides a measure of confidence for the existence of each object. In the assimilation stage, each of these values is connected with min-max rule specified by fuzzy logic. Fuzzy membership function \( y=f_{MNC}(x) \)

![Fuzzy Sets Functions](image)

Figure-7.3 explains the inference process, which involves the computation used in the process. Fig-7.3(a) and (b) show typical SSTs obtained from the two decoding operations corresponding to OBJECT1 nd OBJECT2. The fuzzy set membership functions for quantifiers HIGH and LOW are shown in (c). In this particular set, the MNC shifts from 0 to 1.0,

184
membership in LOW decreases linearly, and membership in HIGH increases linearly. The MNC of both the objects for their occurrences at image associated with RLP 2401 are given in SST. The strengths of the quantified assertions HIGH OBJECT1 and HIGH OBJECT2 are evaluated through the fuzzy membership function shown in (c). Finally, \( \min \) and complement rules are applied to compute the SST of OBJECT3 to account for logical compounding (AND in this case). SST of OBJECT3 (Fig-7.3(d)) indicates that the IMAGE associated with RLP2401 is the answer with MNC .7, which is the correct answer in this case.

7.5.2 Example #2

\[ \text{OBJECT3: (HIGH.OBJECT1) OR (HIGH.OBJECT2)} \]

This search is identical to the canonical object generation stage. But the elementary SSTs, constructed by the fuzzy assimilation algorithm, interprets the result in a different manner. Fig-7.4 shows the example SSTs. The final SST is evaluated using fuzzy max rule, assuming linear set functions as in previous example.

7.5.3 Example #3

\[ \text{OBJECT2: OBJECT1 (in LOCATION(xmax,xmin, ymax, ymin)) or OBJECT1 (in LOCATION(0-1, 2-2))} \]
This example involves generation of transform patterns from the canonical object OBJECT1. LOCATION(xmax,xmin,ymax,ymin) is a transform with two parameters. Transform LOCATION() can be used as a basis for spatial reasoning in images of 2D scene.

The transform driver applies the transform LOCATION() from the transform library to generate a set of transform patterns for each of the canonical objects. Transform LOCATION(0-1,2-2) results in two searches (assuming a search resolution 1) with canonical pattern translated at locations (0,2) and (1,2). (A spatial search LOCATION(1,0,1,0) results in 4 template search).

The SST generated for this search not only lists all the occurrences and corresponding MNCs, but also the transform parameter values at which matches have been registered.

7.5.4 Example #4

OBJECT2: LOW.OBJECT1( in LOCATION(xr2)(yr2))

This is an example of simple transform compounding. The query generates (xr-x2)(yr-y2) search for OBJECT1. The SST corresponding to OBJECT1 is obtained exactly like example #3. However after that, the MNCs are recomputed through fuzzy membership function for logical quantifier LOW. The results are compiled in a new SST with the detected matching RLPs, match locations, and new MNCs.

7.5.5 Example #5

OBJECT3: (HIGH.OBJECT1( in LOCATION(xr-x2)(yr-y2))). MODERATELY.LEFT-OF. (HIGH.OBJECT2( in LOCATION(xr-x2)(yr-y2))

This example involves transforms as well as fuzzy spatial reasoning. The example is explained in Fig-7.5. As, before, the OBJECT1 and OBJECT2 are searched for all their occurrences in respective regions specified. The search results in two SSTs, which include
the location of match in the image (figures (a) and (b)). Each of the SSTs is mapped with fuzzy set functions to evaluate the MNC for the concepts HIGH.OBJECT1 and HIGH.OBJECT2. A linear function is assumed in this case which leaves the MNC values intact.

In this command, LEFT-OF is a transform-relation defined with the parameters of transform LOCATION(). It invokes a LEFT-OF() strength evaluator function, which operates on the LOCATION parameter (x,y) columns of the SSTs of OBJECT1 and OBJECT2. The user is entitled to have any interpretations of the linguistic concept LEFT-OF(). A typical definition will be (the user is free to define his/her own):

\[ s = (x_1 - x_2) \text{ when, } x_1 > x_2 \]

\[ = 0 \text{ when, } x_1 < x_2 \]

The evaluation suggests that RLPs 1012 and 2017 are the possible matches with respective normalized strengths .568 and .110. The next step is to compute the strength (y) of the quantified-relation MODERATELY. LEFT-OF. This is evaluated by the fuzzy set function associated with the linguistic quantifier MODERATELY as defined in (Fig-7.5(c)).

The evaluation suggests that RLPs 1012 and 2017 are the possible matches with respective normalized strengths .568 and .110. The next step is to compute the strength (y) of the quantified-relation MODERATELY. LEFT-OF. This is evaluated by the fuzzy set function associated with the linguistic quantifier MODERATELY as defined in (Fig-7.5(c)).
Finally, the three concepts HIGH.OBJECT1, NEAR.DISTANT-FROM and HIGH.OBJECT2, are combined together by min-rule to obtain the MNC of OBJECT3 (Fig-7.5(d)).

7.5.6 Example #6

OBJECT3: (HIGH.OBJECT2). FAR. DISTANT-FROM. (OBJECT2 in REGION(x1,x2,y1,y2))

This example is almost identical as of previous example, except, it is defined on the basis of a different transform-relation concept DISTANT-FROM defined on the same transform parameters. A Typical evaluator will be:

\[ s = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \]

This relation also has its own set of linguistic quantifiers like FAR. The evaluation process is also identical.

7.6 Chapter Summary

This chapter has presented the CQF formalism that can support a more complex form of query involving compound objects and various transformations. CQF is fundamentally different from conventional search mechanisms used by other IDBs. It assumes much less concreteness of the underlying definitions and measurements on which a symbolic query is formed. It offers the users as much flexibility as possible to define basis concepts and provides subsequent inference support.

The formalism acts as a fuzzy logic based symbolic query shell over the holographic content associative search kernel. The response MNC, which is unique to holographic associative memory, provides the crucial means for interactive communication between the shell and the kernel. This chapter outlines the methodology, describes an inference mechanism and illustrates the query process with examples.
CHAPTER 8

CONCLUSIONS

This dissertation describes a new associative computation paradigm which can perform associative search with changeable attention over its element field. This dissertation particularly addresses its theory, characterization and a prototype application in image archive query. The contributions and limitations of this method will be reviewed in this concluding chapter.

The characteristics of the proposed complex representation demonstrated in this dissertation, in fact calls forth a massive reinvestigation into the enhancement of almost all existing artificial neural network models. In this final chapter, therefore, few leads will be provided towards the future research directions based on the current understanding of this new computing paradigm.

8.1 Summary: Contributions and Limitations

(1) In this dissertation it has been demonstrated how the capabilities of current parallel and distributed computation can be enhanced to perform search and retrieval with attention. It has been demonstrated that the key to this new capability is hyperspherical representation.
(2) It has been shown that as low as 5-10% cue can be effectively used. What happens beyond this range? Experimental as well as analytical results show that even smaller focus can be used but the loading should be considerably low. Such small load factor makes its space utilization inefficient. However at such low focus another more fundamental problem gradually takes over. The very distinguishability among the dataset tends to diminish\(^3\). 

(3) Traditionally, parallel and distributed processing (PDP) based on neural networks have been used for three major types of applications, (i) associative memory, (ii) adaptive classification (adaptive filter) and (iii) optimization. However from application point of view, most of the success of neuro computing remains confined in the area of adaptive classification, with relatively little progress in the other two areas\(^2\). In this dissertation, one of these less developed areas, associative memory based applications, has been picked, and it has been demonstrated how the new capability in search and retrieval can be immediately applied for associative retrieval of information from massive image archives.

Appendix-F shows some projected retrieval performances for three image archive application scenarios; (i) planetary navigation, (ii) trademark verification and (iii) a typical distributed image archive, each of which requires search and match involving substantial number of images.

However the potential application of attention based retrieval is far broader in scope. In fact, all of the application areas of current artificial neuro-computing can potentially benefit from this work.

(4) For image archives, existing model-based methods for content-based search requires the meaning of the images to be pre-extracted and modeled. The extraction process requires

---

\(^3\) The effect is more pronounced for smaller n.

\(^2\) The limitation of conventional ANN based computing to work with missing information identified in this dissertation also implicitly explains the lack of progress in the area of associative computing.
partial or full human involvement. The process is inexact and highly subjective. On the other hand, the existing procedural search based methods are expensive because multidimensional patterns can not be efficiently pre-ordered to satisfy dynamic attention modulated search.

This dissertation has demonstrated how the holographic associative memory can be used to perform direct search which suffers neither of the problems of the previous approaches. The general problem of content-based image archive query is an extremely complex one involving human cognizance. Any single solution is far from complete. Holographic method provides only a means for efficient large scale search, and may be considered a method that supplements existing ones. From the finer perspective, it does not address the deeper problem of subjectivity rather, it transmigrates the associated responsibility from encoder directly to the end user.

Besides being a complementary approach, holographic search can be used even in conjunction to the current model based approaches. In the model based approaches, the encoding process itself requires heavy search into raw image archive during model generation. CAIDSS search technique can be used to expedite model generation of raw images by conventional encoder, which can be finally used to support sophisticated symbolic query by end users.

(5) As noted before, neither of the principal problems of existing methods addressed in this dissertation are well identified in existing literature. For example, in parallel and distributed computing, very little effort has been made to clarify the concept of robustness. If done it would have clearly shown the limitations of type-U search and that of the existing neuro-computing. Same is true about the limitations of model based search approaches for image archives, with few recent exceptions.
Perhaps, the very identification and characterization of these two problems themselves can be considered a part of the contribution made by this dissertation. In future, there may be other and better solutions available to these problems than the ones proposed in this dissertation.

8.2 Future Research Directions

This section elaborates few of the significant open research directions that naturally follows from the current dissertation.

8.2.1 Scalability and Parallelizability

Research Issue 1: Can parallel processing be used to expedite search with holographic memories?

An algorithm can face potential bottle-neck during parallelization at two stages; in computational scalability, and in hardware scalability. One of the interesting feature of holographic associative memory is its computational scalability. Both analytically and empirical evidences (section-5.4.3) suggest that the performance of the network does not degrade with scale of the problem size (characterized by \( n \) and \( p \)). Such computational scalability is significant specially when compared to other neural models. It is a well-known phenomenon that the performance of even the successful ANN models such as Backpropagation, Counterpropagation networks tend to be severely limited when the problem size increases. Not only the amount of computation, but also the convergence speed, accuracy of any conventional ANN suffers when problem size increases, which makes them difficult to use for large problems.

Computational scalability is a critical aspect of massive data processing algorithms. The lack of such scalability implies that even if large computation hardwares are available, the algorithm itself may degenerate with problem scale.
The characteristics of holographic associative memory are encouraging in this respect. The computational scalability of the method makes it a good candidate for further investigation regarding its architectural scalability. The highly structured heavy grain complex valued matrix operations of holographic associative memory makes it a suitable candidate for parallelization on the current generation parallel computers\textsuperscript{33}.

### 8.2.2 Optical Realizability

**Research Issue 2:** How the current advancement of optical information processing technology will impact holographic model of associative computing?

In recent years the optical mode of information processing has advanced dramatically. Optical communication and storage technologies have already brought in revolutionary breakthroughs (which are commercially viable) in capacity performance in comparison to its electronic counterpart. The optical processor technology is also advancing fast. In this context, it would be timely to investigate the potential of proposed holographic search on emerging optical hardwares. After all the very idea of 'holography' has its origin in optics.

A potential for high gain exists in this investigation. Hyperspherical computations are heavier than scalar computations of current ANN paradigm. This is more so probably due to the physically scalar mode of currently available means of computations. The situation may reverse if the very physical mode of processor's underlying means of computation changes. Recently evolving optical processors use optical inference as a basic means of computation which is inherently complex valued. In such a modality of computation a scalar operation involves more transformations than a complex one. On such a processor, a complex operation would be at least equivalent to a scalar operation if not cheaper. To exploit such

\textsuperscript{33} Current generation parallel computers are characterized by their regular and structured architecture and relatively less efficient communication cost. As a result they always favor computations which are regular and heavy grain.
advantages it might be necessary to look into different architectures of optical processors beyond those which are investigated currently. More research on the optical realizability of holographic computing will be worth investigating.

8.2.3 Applicability to Other Network Models

**Research Issue 3:** How hyperspherical representation will impact other learning models?

The fundamental part of the proposed computing paradigm is its hyperspherical representation, which incorporated attention as an integral aspect of measurement component of information. The particular holographic network, and subsequent learning algorithm, that has been used in this dissertation are more specific towards the intended application.

If one looks into the vast number of existing neural network models which have evolved over the last 50 years, it can be observed that all of them use the same basic cell but largely vary in their inter-connection patterns and learning algorithms. The learning algorithms associated with these networks mostly depend on the constraints of application and that of specific inter-connections. There is no clear reason why the proposed attention modulated representation cannot be used for the vast array of existing models to enhance their representation and retrieval capability.

In this dissertation, the power of complex representation has been demonstrated particularly for crossbar correlation learning. Additional experimentation with two other models of learning, namely self-organization and principal component analysis [KhYu94b] has demonstrated that even if one considers the classical type-U response characteristics, their performance (convergence speed and accuracy) also dramatically improves due to hyperspherical representation. Very recently, Masters [Mast94] has also independently reported similar significant type-U performance improvement for a complex number based on LMS learning (a special case of hyperspherical representation).
As demonstrated in this dissertation, such representation not only improves qualitative performance of type-U search, but provides a fundamentally new capability of type-A retrieval capability which allows attention based retrieval.

This more powerful type-A retrieval capability can potentially open up a large class of new applications of parallel and distributed computing. The results obtained in the course of this dissertation, specially, the causality relationship between hyperspherical representation and attention (explained in section 4.8) strongly suggests that all current models of artificial neuro-computing be reinvestigated in the light of this new representation formalism.

As a concluding remark it would be befitting to say that this dissertation stimulates more curiosity than it answers leading to a new paradigm of parallel and distributed computing based on hyperspherical representation.

34 Artificial neural computing is a prominent paradigm of parallel and distributed computing. There are other modes of it such as genetic algorithms.
# APPENDIX-A
## SURVEY OF ARTIFICIAL NEURAL NETWORKS

### Table-A.1 Neural Network Models

<table>
<thead>
<tr>
<th>PROP.</th>
<th>Perceptron</th>
<th>Madaline</th>
<th>Back Prop</th>
<th>Kohonen SOFM</th>
<th>Counter Prop</th>
<th>Neo Cognitron</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology</td>
<td>multilayer fully connected</td>
<td>multilayer fully connected</td>
<td>multilayer fully connected</td>
<td>sngl. layer close nodes internally connected</td>
<td>2 layers fully connected internally connected</td>
<td>multilayer internally connected</td>
</tr>
<tr>
<td>Learning</td>
<td>Supervised</td>
<td>supervised</td>
<td>supervised generalized delta</td>
<td>unsupervised competitive hebbian</td>
<td>supervised+ unsupervised</td>
<td>unsupervised hebbian</td>
</tr>
<tr>
<td>Convergence</td>
<td>Not proved</td>
<td>Not proved</td>
<td>Conditionally proved</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Efficiency</td>
<td>no</td>
<td>no</td>
<td>Inefficient but practical</td>
<td>efficient practical</td>
<td>efficient practical</td>
<td>some practical</td>
</tr>
<tr>
<td>Rep. mode</td>
<td>binary</td>
<td>analog</td>
<td>analog</td>
<td>analog</td>
<td>analog</td>
<td>binary</td>
</tr>
<tr>
<td>Foundation</td>
<td>none</td>
<td>stat. (LMS)</td>
<td>math (gradient descent)</td>
<td>stat.</td>
<td>stat.</td>
<td>none.</td>
</tr>
<tr>
<td>Motivation</td>
<td>biology</td>
<td>biology</td>
<td>natural</td>
<td>biology</td>
<td>biology</td>
<td>natural</td>
</tr>
<tr>
<td>Tolerance</td>
<td>no</td>
<td>no</td>
<td>noise</td>
<td>noise</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>System design</td>
<td>art</td>
<td>art</td>
<td>science</td>
<td>art+science</td>
<td>science</td>
<td>art</td>
</tr>
<tr>
<td>Application mode</td>
<td>classifier</td>
<td>classifier</td>
<td>classifier</td>
<td>classifier</td>
<td>classifier</td>
<td>classifier</td>
</tr>
</tbody>
</table>

196
### Table-A.1 Neural Network Models (continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology</td>
<td>sing. layer</td>
<td>sing. layer</td>
<td>2 layers</td>
<td>2 layers.</td>
<td>2 layers.</td>
<td>multilayers</td>
</tr>
<tr>
<td></td>
<td>internally</td>
<td>fully connected</td>
<td>fully</td>
<td>internally</td>
<td>Fully</td>
<td>Fully</td>
</tr>
<tr>
<td></td>
<td>connected</td>
<td></td>
<td>connected</td>
<td>connected</td>
<td>connected</td>
<td>connected</td>
</tr>
<tr>
<td>Learning</td>
<td>supervised</td>
<td>supervised</td>
<td>supervised</td>
<td>supervised</td>
<td>unsupervised</td>
<td>supervised</td>
</tr>
<tr>
<td></td>
<td>hebbian.</td>
<td>crossbar</td>
<td>crossbar</td>
<td>hebbian</td>
<td>competitive</td>
<td>prob.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hebbian</td>
<td>hebbian</td>
<td></td>
<td>hebbian</td>
<td>reinfrc.</td>
</tr>
<tr>
<td>Convergence</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>(async.)</td>
<td></td>
<td></td>
<td></td>
<td>[unstable]</td>
<td></td>
</tr>
<tr>
<td>Efficiency</td>
<td>questionable</td>
<td>no</td>
<td>time eff.</td>
<td>eff.</td>
<td>eff.</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>(C=.15n)</td>
<td></td>
<td>space ineff.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td>binary</td>
<td>binary</td>
<td>binary</td>
<td>binary</td>
<td>binary</td>
<td>analog</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[analog]</td>
<td></td>
</tr>
<tr>
<td>Foundation</td>
<td>math</td>
<td>physics</td>
<td>matrix</td>
<td>coding theory</td>
<td>diff. eq.</td>
<td>statistics</td>
</tr>
<tr>
<td></td>
<td>[outer prod.]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>physics</td>
</tr>
<tr>
<td>Motivation</td>
<td>biology</td>
<td>energy function</td>
<td>biology</td>
<td>natural</td>
<td>biology</td>
<td>natural</td>
</tr>
<tr>
<td>Tolerance</td>
<td>some</td>
<td>some</td>
<td>some</td>
<td>some</td>
<td>some</td>
<td>some</td>
</tr>
<tr>
<td>System Design</td>
<td>science</td>
<td>science</td>
<td>science</td>
<td>science</td>
<td>art</td>
<td>science</td>
</tr>
<tr>
<td>Application</td>
<td>asso. mem.</td>
<td>asso. mem.</td>
<td>asso. mem.</td>
<td>classifier</td>
<td>classifier</td>
<td>classifier</td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td></td>
<td></td>
<td>mem.</td>
<td>asso. mem.</td>
<td>optimizer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>asso. mem.</td>
</tr>
</tbody>
</table>

197
# APPENDIX-B

## SURVEY OF IDB SEARCH TECHNIQUES

### Table-B.1 Image Database Systems

<table>
<thead>
<tr>
<th>DBASE</th>
<th>IMAID</th>
<th>PICDMS</th>
<th>IIDMS</th>
<th>GRIM DBMS</th>
<th>VIMSYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPL</td>
<td>2D-spatial query.</td>
<td>general</td>
<td>2D-spatial query.</td>
<td>2D-spatial query (fuzzy). Architecture plan.</td>
<td>general PACs, EOS, GIS.</td>
</tr>
<tr>
<td></td>
<td>LANDSAT images.</td>
<td></td>
<td>GIS, maps.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QUERY</td>
<td>condensed+image.</td>
<td>condensed+tagg</td>
<td>condensed.</td>
<td>condensed.</td>
<td>condensed.</td>
</tr>
<tr>
<td></td>
<td>image, computable features.</td>
<td>general.</td>
<td>spatial reasoning.</td>
<td>objects, absolute/relative positions.</td>
<td>computable features.</td>
</tr>
<tr>
<td></td>
<td>table+iconic-ex</td>
<td>command lan-</td>
<td>graphics+ key-</td>
<td>command+fuzzy measure.</td>
<td>command+fuzzy measure.</td>
</tr>
<tr>
<td></td>
<td>ample+preproces-</td>
<td>guage.</td>
<td>word.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>sing.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REPRES.</td>
<td>line-sketch+</td>
<td>stacked images.</td>
<td>2D-string.</td>
<td>indexed access-structure.</td>
<td>four planes (IR),(IO), (DO),(DE)+</td>
</tr>
<tr>
<td></td>
<td>table+text.</td>
<td>relational DBMS.</td>
<td>sigma-tree.</td>
<td></td>
<td>data dictionary.</td>
</tr>
<tr>
<td>META MODEL</td>
<td>designer+ user</td>
<td>encoder+user</td>
<td>designer</td>
<td>designer</td>
<td>designer+user+data model.</td>
</tr>
<tr>
<td></td>
<td>objects.</td>
<td>auto feature.</td>
<td>auto spatial fea-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Purdue University, USA, [ChFu81].</td>
<td>Cardenas, UCLA, USA, [JoCa88, GuWJ92].</td>
<td>DA88].</td>
<td>sa92abc].</td>
<td>nath, Jain, U. of Michigan, USA, [GuWJ92].</td>
</tr>
<tr>
<td>DBASE</td>
<td>PQL/PDL</td>
<td>SKETCH</td>
<td>TRADE MARK</td>
<td>QBIC</td>
<td>RIVAGE</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>------------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>APPL</td>
<td>event query, educational database.</td>
<td>sketch query, chest-xray.</td>
<td>sketch query, trademarks.</td>
<td>iconic query, shape, color, texture.</td>
<td>subjective query, mediaClip.</td>
</tr>
<tr>
<td>QUERY</td>
<td>condensed+tag event, formatted text.</td>
<td>condensed, domain features, pictorial sketch.</td>
<td>condensed, visual similarity, sketches.</td>
<td>condensed+tag color, shape texture, iconic example.</td>
<td>condensed+visual image, subjective content, keyword+relevance feedback.</td>
</tr>
<tr>
<td>REPRES.</td>
<td>semantic net, structured text.</td>
<td>3 bit pixel</td>
<td>clusters based on features.</td>
<td>Starburst relational database.</td>
<td>named fields, field-descriptor+thesaurus.</td>
</tr>
<tr>
<td>META MODEL</td>
<td>narrator.</td>
<td>designer.</td>
<td>designer.</td>
<td>designer+encoder+user</td>
<td>encoder+user</td>
</tr>
</tbody>
</table>
The derivation of a dimension dispersion component in the step-1 of equation (5.16) in section 5.2 is given here. By expanding the \((n-1)\)th and \(n\)th dimensional components of the vector sum according to equation (5.14):

\[
\sum_{k}^{n-1}\left(\sum_{i}^{p}x_{i}^{k}\right)^{2} + \left(\sum_{i}^{p}\prod_{k}^{n}\cos\theta_{k}^{i}\right)^{2}
\]

\[
= \sum_{k}^{n-2}\left(\sum_{i}^{p}x_{i}^{k}\right)^{2} + \left(\prod_{k}^{n-2}\cos\theta_{k}^{i}\sin\theta_{n-1}^{i} + \prod_{k}^{n-2}\cos\theta_{k}^{i}\sin\theta_{n-1}^{i} + \ldots + \prod_{k}^{n-2}\cos\theta_{k}^{i}\sin\theta_{n-1}^{i}\right)^{2}
\]

\[
+ \left(\prod_{k}^{n-2}\cos\theta_{k}^{i}\cos\theta_{n-1}^{i} + \prod_{k}^{n-2}\cos\theta_{k}^{i}\cos\theta_{n-1}^{i} + \ldots + \prod_{k}^{n-2}\cos\theta_{k}^{i}\cos\theta_{n-1}^{i}\right)^{2}
\]

by expanding the last two square terms and rearranging the result,

\[
= \sum_{k}^{n-2}\left(\sum_{i}^{p}x_{i}^{k}\right)^{2} + \left(\prod_{k}^{n-2}\cos\theta_{k}^{i}\right)^{2} + \left(\prod_{k}^{n-2}\cos\theta_{k}^{i}\right)^{2} + \ldots + \left(\prod_{k}^{n-2}\cos\theta_{k}^{i}\right)^{2}
\]

\[
+ 2\cos(\theta_{n-1}^{i}-\theta_{n-1}^{i})\prod_{k}^{n-2}\cos\theta_{k}^{i}\prod_{k}^{n-2}\cos\theta_{k}^{i} + 2\cos(\theta_{n-1}^{i}-\theta_{n-1}^{i})\prod_{k}^{n-2}\cos\theta_{k}^{i}\prod_{k}^{n-2}\cos\theta_{k}^{i} + \ldots
\]

Which can be expressed as:
\[
\begin{align*}
= \sum_{k}^{n-2} \left( \sum_{i}^{p} x_{i} \right)^{2} + \sum_{i}^{p} \left( \prod_{k}^{n-2} \cos \theta_{k} \right)^{2} + \sum_{i}^{p} \sum_{j \neq i} \cos(\theta_{i}^{j} - \theta_{j}^{i}) \prod_{k}^{n-2} \cos \theta_{k}^{i} \cos \theta_{k}^{j}
\end{align*}
\]

\[
= \sum_{k}^{n-2} \left( \sum_{i}^{p} x_{i} \right)^{2} + \sum_{i}^{p} \left( \prod_{k}^{n-2} \cos \theta_{k} \right)^{2} + \sum_{i}^{p} \sum_{j \neq i} \cos \theta_{k}^{i} \cos \theta_{k}^{j} - B_{n-1}
\]

\[
= \sum_{k}^{n-2} \left( \sum_{i}^{p} x_{i} \right)^{2} + \left[ \sum_{i}^{p} \prod_{k}^{n-2} \cos \theta_{k} \right]^{2} - B_{(n-1)}
\]

This corresponds to step-2. The DDC is given by:

\[
B_{n-1} = \sum_{i}^{p} \sum_{j \neq i} \prod_{k}^{n-2} \cos \theta_{k}^{i} \cos \theta_{k}^{j} - \sum_{i}^{p} \sum_{j \neq i} \cos(\theta_{i}^{j} - \theta_{j}^{i}) \prod_{k}^{n-2} \cos \theta_{k}^{i} \cos \theta_{k}^{j}
\]

\[
= \sum_{i}^{p} \sum_{j \neq i} \left[ 1 - \cos(\theta_{i}^{j} - \theta_{j}^{i}) \right] \prod_{k}^{n-2} \cos \theta_{k}^{i} \cos \theta_{k}^{j}
\]
APPENDIX-D
SCENE INTERPRETATION

The figure presented in this appendix shows a typical interpretation of an image. The image XQ1 presented in Fig-D.1 is hierarchically decomposed into components by a human subject according to his/her conceptualization of its meaning. Fig-D.2 shows this interpretation. There may exist numerous equally valid other interpretations.

Conventional image database approaches, which append meaning before encoding, have to adopt one interpretation. Such an adhoc interpretation attached by the encoder may not always match the interpretation of the inquirer. Such occurrence results in incoherence.

CAIDSS approach avoids attaching any meaning at the encoding time. Rather, it allows human subjects to attach meaning during query and to perform meaningful retrieval on the basis of his/her interpretation. Fig-D.3 shows three different pictorial query frames (Pan-A, B and C) based on three meaningful objects chosen from the decomposition of Fig-D.2. Each of them are subjective and depends on how the inquirer views the meaning of XQ1. Each of them should lead to completely different matches.

Fig-D.4 (a), (b) and (c) respectively show the query results corresponding to these three objects (specified by attention windows) those were obtained by CAIDSS.
Fig-D.1 Pictorial Example Scene XQ1

Fig-D.2 A Subjective Interpretation of XQ1
Fig-D.3 Object focus fields

Fig-D.4(a) Retrieved Image (Pan-A)

Fig-D.4(b) Retrieved Image (Pan-B)

Fig-D.4(c) Retrieved Image (Pan-C)
APPENDIX-E

QUERY IN SEQ001 ARCHIVE

This appendix provides some sample images and query examples in the archive SEQ001. SEQ001 archive contains 64 color images with dimensions 160x120. Some samples are shown in Fig-E.1. MAP1, HIM1, HIM2, HIM3, FIN2, MOON1, MOON3, CAR2, MOON2, HNL1, FIN1, HIM4, MOON4 are some example images that have been stored into the SEQ001 archive.

Fig-E.2, E.3, and E.4 present some examples of queries performed in this archive. HIMQ2, HIMQ3, MOONQ1, MOONQ2 and FINQ1 are several query images. The windows HIMQ2-PEAK, HIMQ2-RIDGE, HIMQ3-CANYON, MOONQ1-BIGHOLE, MOONQ1-PAIR, MOONQ2-BUBBLES, MOONQ2-SHALLOWS and FIN-Q1-M1X in these images show a set of attention frames identified by the inquirer as object of interest in the query images. CAIDSS performs holographic search into the archive on the basis of these objects of interest and retrieves a set of images as closest match. These retrieved images are also shown below each of the query images. The arrows in the retrieved images show the match regions in the target images. For example, archive image HIM4 shows match for both the PEAK and RIDGE of HIMQ3. Similarly, the PAIR of the MOONQ1 can be found in MOON3 image.
Fig -E.1 Sample Images in SEQ001 Archive
Fig. E.1 Sample Images in SEQ001 Archive (continued)
Fig - E.2 Himalayan Landscape Queries in SEQ001 Archive
Fig -E.3 Lunar Landscape Queries in SEQ001 Archive
Fig. E.4 Finger Print Pattern Query in SEQ001 Archive
APPENDIX-F

QUERY IN MEDIA ARCHIVE

This appendix provides first some sample images stored in archive Medical Image Archive (MEDIA). MEDIA stores about 64 CT scan and MRI images. These images are grey scale images with 256 intensities of dimensions 256x256. Fig-F.1 shows some samples. SKL-DN, ABD-RT, ABD-LT, SKL-UL, SKL-UR, and LNG-LT are some example images that have been stored into the MEDIA.

Fig-F.2 and F.3 present some examples of queries performed in this archive. SKL-Q1 and ABD-Q1 are two query images. The windows Basi-occipital (VM-1), Foramen Magnum (VM-2), Foramen Ovale and Spinosum (VM-3), Jugular Foramen and Carotid Canal (VM-5), Kidney (VM-4), and Vertebrae (VM-6) are objects in these images and show a set of attention frames identified by the inquirer as object of interest in the query images. CAIDSS performs holographic search into the archive on the basis of these objects of interest and retrieves a set of images as closest match. These retrieved images are also shown below. The arrows in the retrieved images show the match regions in the target images. Note, some objects have multiple match. For example Jugular Formen and Carotid Canal (VM-5) has been found both in SKL-UR and SKL-UL.
Fig. F.1 Sample Images in MEDIA Archive
Fig. F.2 Sample Queries in MEDIA Archive (Skull)
Fig-F.1 Sample Queries in MEDIA Archive (Abdomen)
APPENDIX-G

PERFORMANCE PROJECTIONS

This appendix provides projected performances of holographic associative search for several image archives based on the results obtained from operation range space (ORS) experiments. Three applications have been considered. These are (a) Trademark Images (b) Planetary Navigation, and (c) a typical distributed image database.

I. TRADEMARK IMAGE DATABASE

Trademark approval requires regular matching of large number of stored images. Currently, the similarity is checked by visual inspection. Here an estimate is presented for a typical trademark image database. The number of images will be assumed as 300,000 (512x512 pixels) patterns (as in Japan Patent Office, 1994 [Kato92]). Other specifications are given below.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern length (n)</td>
<td>512x512</td>
</tr>
<tr>
<td>Pattern/holograph</td>
<td>4096 (L=0.015)</td>
</tr>
<tr>
<td>RLP mod</td>
<td>3</td>
</tr>
<tr>
<td>RLP length (m)</td>
<td>8 bytes</td>
</tr>
<tr>
<td>Complex number b</td>
<td>4 bytes/complex</td>
</tr>
<tr>
<td>Space s</td>
<td>512x512x8x4 = 8 MBYTE/holograph</td>
</tr>
<tr>
<td>Number of holograph h</td>
<td>80 holographs/320,000 images</td>
</tr>
<tr>
<td>Total space S</td>
<td>640 MBYTE/320,000 images</td>
</tr>
<tr>
<td>Raw Space</td>
<td>78 GBYTE</td>
</tr>
<tr>
<td>Space Factor (SF)</td>
<td>0.008</td>
</tr>
<tr>
<td>Retrieval cost Cr</td>
<td>12.5 MFLOP/holograph</td>
</tr>
<tr>
<td>Encoding cost Ce</td>
<td>40 MFLOP/iteration</td>
</tr>
</tbody>
</table>
Table-I.1 Trademark Image Archive

<table>
<thead>
<tr>
<th>TRADEMARK IMAGE DATABASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of frames (P)</td>
</tr>
<tr>
<td>size of each frame (n)</td>
</tr>
<tr>
<td>space</td>
</tr>
<tr>
<td>SF</td>
</tr>
<tr>
<td>speedup</td>
</tr>
</tbody>
</table>

The encoding and retrieval performance for various server systems are given next.

Table-I.2 Server Performance

<table>
<thead>
<tr>
<th>system</th>
<th>no. of processors</th>
<th>power MFLOP/s</th>
<th>Tretrieve sec/loc</th>
<th>Tencode sec/image</th>
<th>Tregular min/loc</th>
</tr>
</thead>
<tbody>
<tr>
<td>HITACHI S-3800/180 (2ns)</td>
<td>1</td>
<td>408</td>
<td>2.41</td>
<td>.97</td>
<td>3.2</td>
</tr>
<tr>
<td>NEC SX-3/14 (2.9ns)</td>
<td>1</td>
<td>314</td>
<td>3.14</td>
<td>1.27</td>
<td>4.16</td>
</tr>
<tr>
<td>FUJITSU VP2600</td>
<td>10</td>
<td>249</td>
<td>3.95</td>
<td>1.6</td>
<td>5.25</td>
</tr>
</tbody>
</table>

II. PLANETARY NAVIGATION

Pattern matching is one of the techniques for interstellar navigation. The direction of the spacecraft can be determined from observed matching stellar map images with previously recorded or simulated maps. If each map covers an area of 10 min x 10 min, then (360x6)x(360x6)=4,665,600 images are required to be matched. Other specifications are given below.

---

35 10 iterations.

36 As System capabilities are based on their performance against LINPACK Benchmark [Dongarra, 1995]
Pattern length (n) = 8192x8192=67,108,864
Pattern/holograph = 531,441 (L=.008)
RLP mod = 3
RLP length (m) = 12 bytes
Complex number size b = 4 bytes/complex
Space s/h = 8192x8192x12x4 = 3.2 GBYTE/holograph
Number of holograph h = 10 holographs/5,314,410 images
Total space S = 32 GBYTE/5,300,000 images
Raw Space = 268 TERABYTE
Space Factor (SF) = .00012
Retrieval cost Cr (6mn) = 4.8 GFLOP/holograph
Encoding cost Ce (15mn) = 12.1 GFLOP/iteration

Table-I.3 Planetary Navigation Archive

<table>
<thead>
<tr>
<th>PLANETARY NAVIGATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of frames (p)</td>
</tr>
<tr>
<td>size of each frame (n)</td>
</tr>
<tr>
<td>space</td>
</tr>
<tr>
<td>SF</td>
</tr>
<tr>
<td>speedup</td>
</tr>
</tbody>
</table>

The encoding and retrieval performance for various server systems are given next.

Table-I.4 PN Server Performance

<table>
<thead>
<tr>
<th>system</th>
<th>no. of processors</th>
<th>power GFLOP³⁸/s</th>
<th>Tretrieve min/loc</th>
<th>Tencode³⁷ sec/image</th>
<th>Tregular min/loc</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRAY C90</td>
<td>16</td>
<td>.479</td>
<td>1.7</td>
<td>25.2</td>
<td>85</td>
</tr>
<tr>
<td>MEIKO COMP. SUR-</td>
<td>32</td>
<td>.210³⁹</td>
<td>3.8</td>
<td>57.5</td>
<td>194</td>
</tr>
</tbody>
</table>

³⁷ 10 iterations.
³⁸ As System capabilities are based on their performance against LINPACK Benchmark [Dongarra, 1995]
³⁹ TPP Best Effort performance
III. TYPICAL IDBMS

This estimate is for a typical image database. Estimates have been given for a medium sized site as well as an Internet like distributed image database. Let each IDBMS site has 2000 color images of size (512x512x3). Other specifications are given below.

- Pattern length (n) = 512x512x3 = 786,432
- Pattern/holograph = 2000 (L = 0.03)
- RLP mod = 2
- RLP length (m) = 11 bytes
- Complex number b = 4 bytes/complex
- Space s/h = 512x512x3x11x4 = 34.6 MBYTE/holograph
- Number of holograph h = 1/2048 images
- Total space S = 34.6 MBYTE/2048 images
- Raw Space = 1.6 GBYTE
- Space Factor (SF) = 0.021

Retrieval cost Cr (6mn) = 52 MFLOP/holograph
Encoding cost Ce (15mn) = 130 MFLOP/iteration

Table I.5 A Typical Image Archive

<table>
<thead>
<tr>
<th>TYPICAL IDBMS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>number of frames (p)</td>
<td>2000</td>
</tr>
<tr>
<td>size of each frame (n)</td>
<td>512x512 pixels</td>
</tr>
<tr>
<td>space</td>
<td>34.6 MBYTE</td>
</tr>
<tr>
<td>SF</td>
<td>0.021</td>
</tr>
<tr>
<td>speedup</td>
<td>31</td>
</tr>
</tbody>
</table>

The encoding and retrieval performance for various server systems are given next.
Table I.6 Server Performance of Single Site

<table>
<thead>
<tr>
<th>System</th>
<th>No. of processors</th>
<th>Power MFLOp/s</th>
<th>Tretieve/sec/site</th>
<th>Tencode/sec/image</th>
<th>Tregular/sec/site</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGI CHALLANGE/Onyx</td>
<td>2</td>
<td>38</td>
<td>1.36</td>
<td>54</td>
<td>42</td>
</tr>
<tr>
<td>SUN SPARC 10-52</td>
<td>1</td>
<td>24</td>
<td>2.2</td>
<td>54</td>
<td>66</td>
</tr>
<tr>
<td>APPLE MAC 6100/60</td>
<td>1</td>
<td>9.8</td>
<td>5.3</td>
<td>132</td>
<td>162</td>
</tr>
<tr>
<td>GATEWAY (P5-60, 60MH z)</td>
<td>1</td>
<td>5.4</td>
<td>9.5</td>
<td>241</td>
<td>296</td>
</tr>
</tbody>
</table>

The above computation will be extended for a distributed image archive system where several sites. Let us assume that each site has a typical IDBMS analyzed above. Let us also assume the following additional specifications:

- No of sites = 10-2
- Type of site system = SUN SPARC 10-52 (1)
- Average poke access = 100
- Transmission time = 10 sec/site
- Our processing time = 220 sec.
- Others processing time = 110 min.
- Receiving time T(r) = 30 sec/site
- Total response time = 620-300 sec = 10-5 min
- Others response time = 116 min.

Table I.7 Server performance for Distributed Archive

<table>
<thead>
<tr>
<th>TYPICAL DISTRIBUTED IDBMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of sites</td>
</tr>
<tr>
<td>Site load</td>
</tr>
<tr>
<td>Type of site system</td>
</tr>
<tr>
<td>Network</td>
</tr>
<tr>
<td>Total response time</td>
</tr>
<tr>
<td>Conventional method</td>
</tr>
</tbody>
</table>

40 10 iterations.

41 As System capabilities are based on their performance against LINPACK Benchmark [Dongarra, 1995]
BIBLIOGRAPHY


[CrSm92] Crehange, M., Malika Smail, "INTERACTIVITY and HYPERTEXT APPROACH in IMAGE RETRIEVAL: the EXPRIM process, the RIVAGE system", SPIE v. 1662, Image Storage and Retrieval Systems, 1992, pp124-133.


