

A FRAMEWORK FOR DISTRIBUTED MACHINE
PERCEPTION IN UNDERSTANDING STATIC
NATURAL SCENES

By

ARUN AGARWAL

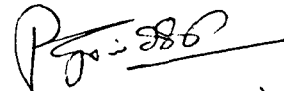
Department of Computer Science
and Engineering

Thesis submitted in fulfilment of the
requirements of the Degree of
DOCTOR OF PHILOSOPHY

TO THE
INDIAN INSTITUTE OF TECHNOLOGY, DELHI
MARCH, 1989.

CERTIFICATE

Certified that the thesis entitled " A
FRAMEWORK FOR DISTRIBUTED MACHINE PERCEPTION IN
UNDERSTANDING STATIC NATURAL SCENES," being
submitted by Mr. Arun Agarwal, is a record of
work carried out by him under my supervision.
The matter embodied in this thesis has not been
submitted for the award of any other degree.



(Dr. P.G. REDDY)

Professor of Computer Science.

DEDICATED TO MY PARENTS

ACKNOWLEDGEMENTS

I am grateful to my supervisor Professor P.G. Reddy for his personal and untiring help during the course of this work. Without his constructive criticisms and constant encouragement this work would never have been polished.

My thanks are due to Prof. M.K. Jain for his help and encouragement during my stay at Indian Institute of Technology, Delhi and to Prof. R.C. Malhotra and Prof. P.C.P. Bhatt for providing the sort of environment which is very vital for research. I would like to acknowledge the contribution of faculty members and also the staff of the Computer Centre, for their cooperation and help.

I would like to express my gratitude to Dr. B.E.Prasad, Dr. S.P. Rana, Dr.Subhash Bhalla, Dr. Niraj Sharma, Nirmal Jain and Rekha Rani who have helped me in several ways.

I would also like to express my gratitude to my friends Ashok Gupta, Arvind K. Jain, Dr. T.K. Dutta, Mahalovya Gauba, S.M.K. Rehman, P.K.Sharma, A. Datta and all others who have provided a kind of company and help which I can never forget.

I would like to thank Anjali, Ashok and everyone else for being so nice throughout the duration of this work.

Finally, I would like to thank Mr.E.V. Suryanarayana Murty for transforming the conglomeration of hand-written text into this document. Thanks are also due to Ms.Manjula and Puspha for neat drawing work.



(ARUN AGARWAL)

ABSTRACT

Humans pursue many goals in an unconstrained, constantly changing world with many objects. For a vision system to perform with so much of generality and flexibility the system should perform a broad range of tasks which involves different object classes, which have different contextual information and which vary greatly at the image-level because of varied viewpoint, illumination, sensors, weather and obscuration and camouflage.

Also, with a variety of information and domain independent knowledge modules available i.e. edge finding, region finding, texture, shape from shading, shape from texture, shape from contours etc and a strong desire for applications which demand speed, it is becoming increasingly important to reconsider our conventional approaches to Model-based Image Understanding Systems (IUS) along the lines of distributed problem solving framework. The traditional and recent systems do not talk about how such a demand can be met in conventional centralized system, although some have talked on schema communicating with knowledge sources and a scheduler controlled knowledge sources.

This thesis attempts to provide a framework for Distributed Machine Perception (DMP) in understanding static natural outdoor scenes. The major motivation for this lies in the potential it offers for making available more problem solving power, by applying a collection of processors to the solution of a single- problem. This would provide a step towards a general image understanding system rather than a method for solution of a specific problem.

We have taken an organizational view to the problem of vision task distribution. This paradigm allows us to look at problems which are extremely complex, resource limited etc in an effective manner, as these theory themselves provide substantial basis for constructing human organization. We try to fit together the progress in computer vision into such an organizational framework and then look beyond for a new distribution of vision system.

We have visualized our task of understanding along product lines (i.e. objects) where respective methods or strategies tailored for the object is locally available with it, while the other knowledge sources for scene analysis can be contracted to build the object description. Object expert selection is

arbitrated by the Gist Analyzer. In its capacity the Gist Analyzer monitors the current state of knowledge about the scene under consideration and is responsible for directing the object experts attention to a specific part of the image for the accomplishment of the task. This we call as the focus of attention. Gist Analyzer directs this operation by means of Schema's which are both specific/general, as well as viewpoint dependent/independent, while scene synthesizer constructs the conceptual model of the 3-D world being viewed. This interaction between Object Experts, Gist Analyzer and Scene Synthesizer constitutes our High-level Interpretation Components.

To affect the interaction between the high-level system components we have elucidated high level protocols that facilitates distributed control of cooperative task execution with efficient internode communication. We have identified possible message types that captures the gist of interaction that arises in our DMP environment. The message slots have been so selected that they may reflect the types of information that must be passed between nodes to make these interactions viable. The protocols are also designed to aid in maintaining the focus of the problem-solving.

We have also reported here a new segmentation technique based on spatial clustering and region growing, which constitutes the low-level component of our DMP System. We have given a new set of transformations for image feature vector to its cartesian form. This leads us to evolve a representation which has a greater cluster resolving power and provides basis for understanding meaningfully the 'meaning of the color.' We also see how this representation gives a new sense of direction for achieving segmentation of the image.

Lastly, the thesis outlines the experimental methodology of increasing complexity, so that the interpretation and understanding of the image with generic object classes and generic viewing conditions may be tuned in a way that is generalizable.

C O N T E N T S

Acknowledgements	iv
List of Figures	xv
Chapter	
I. INTRODUCTION	1
1.1. Difficulties associated with visual process	1
1.2. Applications	3
1.3. Progress in Computer Vision	5
1.3.1. Computational Approach	5
1.3.2. model-based Image Understanding System	7
1.3.3. Robot Vision	8
1.3.4. Architectural Consideration	8
1.3.5. Summary of the Progress in Computer Vision	10
1.4. The elements of IUS	11
1.4.1. Acquisition and Visualization	11
1.4.1.1. Image Sensors	11
1.4.1.2. Display Devices	12
1.4.2. An Image	13
1.4.3. Simple Constituents	13
1.4.3.1. Edges	15
1.4.3.1.1. Local	16
1.4.3.1.2. Regional or Best fit Model	22
1.4.3.1.3. Global methods	23
1.4.3.1.4. Heuristic Methods	24
1.4.3.1.5. Relaxation Methods	24

1.4.3.2.	Regions	26
1.4.3.2.1.	Histogramming	27
1.4.3.2.2.	Region growing	27
1.4.3.2.3.	Spatial Clustering	28
1.4.3.2.4.	Split and Merge	28
1.4.4.	Representation and description of shape	29
1.4.5.	Interpretation	30
1.5.	Our goals and the thesis organization	30
II.	STATIC COMPUTER VISION SYSTEMS	33
2.1.	Literature Review	34
2.1.1.	Block World or Early Computer Vision Systems	34
2.1.2.	Real-World Computer Vision Systems	37
2.1.3.	Query-directed Computer Vision Systems	42
2.1.4.	General-purpose Computer Vision Systems	43
2.2.	Organization and Control of Static Vision Systems	47
2.2.1.	The problem of control	48
2.2.2.	Common Control Structures	50
2.2.2.1.	Hierarchical Control	50
2.2.2.1.1.	Bottom-up Control	51
2.2.2.1.2.	Top-down Control	53
2.2.2.2.	Top-down + Bottom-up Control	55
2.2.2.3.	Heterarchical System	55
2.2.2.4.	Blackboard Model	57

2.2.3.	Summary on Organization and Control structures for Vision System	60
2.3.	Conclusion and Summary	60
III.	DISTRIBUTED PROBLEM SOLVING: A REFLECTION	64
3.1.	Motivation	65
3.1.1.	Social Organization	65
3.1.2.	Cognitive Organization	68
3.1.3.	Further Biological and Neurophysiological Evidences	70
3.1.4.	Technological Advancements	72
3.2.	Advantages of Distribution	72
3.3.	Disadvantages of Distribution	74
3.4.	Main concepts	75
3.5.	Relevance to our work	77
3.6.	Relevance to AI research	79
3.7.	Conclusions and Summary	80
IV.	FRAMEWORK FOR DISTRIBUTED MACHINE PERCEPTION	82
4.1.	Organizational View of the DMP Framework	84
4.1.1.	Choosing an Organization	85
4.1.1.1.	Scenerio 1	86
4.1.1.2.	Scenerio 2	88
4.1.1.3.	Scenerio 3	91
4.1.1.4.	Scenerio 4	94
4.1.1.5.	Scenerio 5	95
4.1.1.6.	Scenerio 6	97
4.2.	An experimental Methodology	106
4.2.1.	Mode 1: Specific scene schema and viewpoint known	107
4.2.2.	Mode 2: General scene schema and viewpoint known	107

4.2.3.	Mode 3: Specific scene schema known and viewpoint unknown	108
4.2.4.	Mode 4: General scene schema known and viewpoint unknown	108
4.2.5.	Mode 5: General scene schema unknown	108
4.2.6.	Mode 6: No schema's given	109
4.3.	Description of the low-level system component	110
4.3.1.	READ	110
4.3.1.1.	Symbolization of Segmented Image	111
4.3.1.1.1.	Region	113
4.3.1.1.2.	Boundary Segments	118
4.3.1.1.3.	Line Segments	119
4.3.1.2.	Color maps	120
4.4.	Conclusion and Summary	120
V.	REPRESENTATION AND SEGMENTATION OF THE IMAGE USING COLOR TRIANGLE	125
5.1.	Theory and Analysis of Color	128
5.2.	Representation	133
5.2.1.	Basic Representation	135
5.2.2.	Advance Representation	138
5.3.	Region Segmentation	144
5.4.	Conclusion	154
VI.	HIGH-LEVEL INTERPRETATION SYSTEM	155
6.1.	Issues in Knowledge representation: A critical Review	157

6.2.	High-level System Components	164
6.2.1.	Gist Analyzer	164
6.2.1.1.	Instantiation of a schema	167
6.2.1.1.1.	Evaluation of plan by property-rules	170
6.2.1.1.2.	Evaluation of plan by Relational rules	173
6.2.1.1.3.	Evaluation of plan using viewpoint knowledge	175
6.2.2.	Object Experts	175
6.2.2.1.	Representation	176
6.2.2.1.1.	Levels of Abstraction	178
6.2.2.1.2.	Relation with other objects and Scene	180
6.2.2.2.	Object Understanding	180
6.2.2.2.1.	Description phase	181
6.2.2.2.2.	Matching phase	185
6.2.2.2.3.	Prediction phase	186
6.2.2.2.4.	Interpretation phase	188
6.2.3.	Scene Synthesizer	194
6.3.	Scene Interpretation Mechanism	196
6.4.	Summary and Conclusions	203
VII.	HIGH-LEVEL PROTOCOLS FOR DMP	205
7.1.	Design Goals for problem solving protocols	208
7.2.	Study of Message Types by node interaction in DMP environment	209
7.3.	Protocol Message Format	211

7.4.	Description and Encoding of the Protocols	214
7.4.1.	The issue of the Language	214
7.4.2.	The message-types	215
7.4.3.	Protoccl Specification	217
7.4.3.1.	For object Experts	217
7.4.3.1.	For Gist Analyzer	221
7.4.3.3.	For Scene Synthesizer	223
7.4.3.4.	For Knowledge Sources	224
7.5.	Conclusions and Discussions	225
VIII	CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS	227
8.1.	Conclusions	227
8.2.	Future work	234
	BIBLIOGROPHY	237
	GLOSSARY	260