

Spiral Architecture for Machine Vision

by

Phillip Sheridan

Thesis submitted for the degree of

Doctor of Philosophy

UNIVERSITY OF TECHNOLOGY, SYDNEY

School of Computing Sciences

February 29, 1996

© Copyright 1996 by Phillip Sheridan

All Rights Reserved

Preface

In a time not too long ago and in a land not too far away, there lived a simple people who sought evolvement by application of their sacred myths and symbols. They held that Art and Science were servants of nature and that the medium through which the energies of these two servants flowed and mixed was a third servant known to folks of other lands as games. This manuscript is, in part, a record of selected conversations between two of its inhabitants which took place in the unfolding of one such game.

The story begins one fine morning when Pekanini goes in search of her mentor, Kimo. She is full of excitement with the desire to tell Kimo of the announcement of this year's Challenge of the Third Servant and to begin their assault on it.

As Pekanini stepped onto the beach which had been left smooth by the receded tide, she began her spiral dance. This dance, an art form unique, by all accounts, to this Land, is performed with strong and decisive movements over a carpet of sand. The movements must deliver bird-like flight while simultaneously sketching a pattern of artistic expression by dragging a combination of toes over the surface of the sand.

Half way up the beach, Pekanini broke off the dance to survey her etchings glistening in the moist sand. Only partially pleased with her efforts, she turned and out of the corner of her eye, she saw sitting cross legged in the sand the familiar figure of Kimo, who had been watching her dance from its beginning.

Acknowledgements

The ideas expressed within this thesis are largely a result of a life long exercise in *learning to see*: one in which many people have exerted important influences. It is with considerable pleasure that I take this opportunity to thank those people who have contributed to the exercise. I gratefully acknowledge the valued contributions that resulted from the numerous conversations held with my supervisor, Tom Hintz. This thesis also contains influences that range from those of a subtle nature with a long gestation period to those that are obvious and had an immediate consequence. Although this is not the place to specify the particular contribution of each person listed below, each in their own way has made an input of significance which I value. I have two regrets in attempting to make individual acknowledgements to the following people. The first is that some will never come to know of this fact; the other is that I will undoubtedly think of others too late to be included.

Mildred Asato, Terry Bossomaier, Jo Burns, Ed Burtram, Robert Chen, Errol Chopping, John Coen, Kathleen Collard, Ken Court, John Crossley, Brad Curry, Alex Depodalinsky, Bernard Ellem, Jim Geyer, Orm Gould, Jean Haakenson, Tim Harata, Carel Harris, Burtram Hinke, Jerry Huff, Ann Ito, Ron James, Harry Kenaman, Bruce Keolani, Landy Kia, Mike Low, Hans Lausch, Glenn Lean, Ken Lodge, Graeme MacDougall, Valerie MacDougall, Shauna McKenna, John McKenzie-Low, Xain Ma, Sue Madeley, Wayne Moore, Adele Morey, Anant Nachimuthu, Ed Niske, Vivian Niske, Ed Niven, Ron Niven, Sue Niven, Nancy Pang, Nigel Pearson, Jim Pratley, Gordon Preston, Yan Rek, Ted Rocavert, Emma Sheridan, Mary Jane Sheridan, Walter Sheridan, Edward Silva, Stanly, Ury Szewcow, Linda Tada, Kees Tesselaar, Ed Toma, Marilyn Van Dyke, Lary Walin, Laison Watai, Ed Watanabe.

I would also like to extend a special thanks to my wife, Manon.

Contents

Abstract	1
1 Introduction	3
1.1 The Perplexities of Vision	4
1.2 Machine Vision	5
1.3 Goal of this Thesis	6
1.4 Observation of this Thesis	8
1.5 Significance of this Thesis	11
1.6 Organization of this Thesis	12
2 A Detailed Look at the Theory	14
2.1 Introduction	15
2.2 Related Other Works	17
2.3 The Spiral Honeycomb Mosaic (SHM)	19
2.3.1 Elements of SHM	20
2.3.2 Definitions and Terminology	22
2.3.3 Spiral Counting	25
2.3.4 Operations on SHM	30
2.3.5 Algebraic Properties of SHM	35
2.3.6 Algebraic Significance of SHM	36

2.4	Spiral Honeycomb Lie Algebra (SHLA)	37
2.4.1	Transformations on SHM	38
2.4.2	Lie Multiplication	39
2.5	Interpretation of SHLA	40
2.5.1	Meaningful Transformations	40
2.5.2	SHLA is Computational	50
2.5.3	Invariant Properties	51
2.6	Summary	52
3	Implementation of the SHLA	53
3.1	Introduction	54
3.2	Transputers	55
3.3	Related Other Works	56
3.4	Data Structure	57
3.4.1	Elements of the Data Structure	57
3.4.2	Operations	60
3.5	Algorithms	63
3.6	Configuration	65
3.7	Simulation	66
3.7.1	Description of the Simulator	67
3.7.2	Simulation Results	68
3.8	Interpretation of Results	72
4	A Segmentation Algorithm for SHLA	74
4.1	Introduction	75
4.2	Related Other Works	78
4.3	Definitions and Terminologies	78

4.4	Informal Statement of 2DSA	81
4.4.1	The Local Actions	82
4.4.2	The Global Actions	84
4.5	Formal Statement of 2DSA	88
4.6	A Transputer Implementation of 2DSA	91
4.7	Simulation Results	93
4.7.1	Synthetic images	93
4.7.2	Natural image	102
4.7.3	Timing Results	110
4.8	Higher Level Segmentation Algorithms	112
4.9	Discussion	112
	Perception of Time and Space	115
5.1	Introduction	116
5.2	Related Other Works	118
5.3	Perception of Depth	118
5.3.1	A 3-D View of SHLA	121
5.3.2	Segmenting the 3-D SHM	127
5.4	Perception of Motion	133
5.4.1	A Representation of Time	133
5.4.2	Transformations of Time and Space	134
5.4.3	Perception of Depth from Parallax	137
5.5	Implementation of 3DSA	138
5.6	Error and Illusions	139
5.6.1	Error from Spatial Mismeasurement	140
5.6.2	Error from Time Mismeasurement	142
5.7	Discussion	146

6	Future Directions	148
6.1	Akamai	149
6.1.1	Overview of Akamai	150
6.1.2	Image Acquisition	151
6.1.3	Image Analysis Unit	157
6.1.4	Locomotion Unit	162
6.1.5	Hardware for Akamai	162
6.1.6	Methodology	163
6.2	Biological Considerations	164
6.2.1	Understanding Biological Vision	164
6.2.2	The Analytic Component of SHM	166
6.2.3	The Retinal Cone Mosaic	169
6.2.4	Binocular Vision	170
6.2.5	Irregularities in the Cone Mosaic	171
6.3	Summary	173
7	The Next Higher Octave	174
7.1	Summary	175
7.2	Conclusion	176
	Appendix A	178
	Glossary	184
	Other Publications by the Author	187
	Bibliography	188
	Bibliography	190

List of Tables

1.1	Burt's list of elements that make up a machine vision system	7
3.1	Spiral Addition and Multiplication Tables.	60
3.2	Time (measured in ticks) required to complete transformations of translation, rotation and scaling on transputer networks configured in a pipeline of various lengths.	69
4.1	Cluster Computations.	82

List of Figures

1.1	Distribution of cones on a primate's retina, taken from [23].	9
1.2	Distribution of hexagonal shaped petals on the flower of a chrysanthemum.	10
1.3	Nerve fibre around the fovea and point of exit at the optic disc showing the spiralling effect, taken from [90] page 63.	10
1.4	The fundamental unit of vision. Seven hexagons arranged such that a centre hexagon is adjacent to six other hexagons.	11
2.1	A fundamental unit of vision. A collection of seven hexagons where the centremost is labelled with address 0 and each neighbour is labelled consecutively.	21
2.2	Step 1. Each of the original addresses from 2.1 are multiplied by 10.	21
2.3	Step 2 . The cluster from Step 1 is dilated so that six additional hexagons can be placed about each of the addressed hexagons.	22
2.4	Step 3. Each new group of seven hexagons from Step 2 is labelled consecutively.	23
2.5	Collection of $7^3 = 343$ hexagons with labelled address from 0 to 666 in base seven.	24
2.6	An octave.	25
2.7	Octave established by key of 1.	27

2.8	Octave established by key of 10.	27
2.9	The octave established by the key of 15.	29
2.10	Three examples of Spiral Addition and two examples of Spiral Multiplication.	31
2.11	(left) Pal can of dog food represented by 384x288 pixels, (right) shows the relative position of the 7^6 hexagon SHM to the rectangular figure.	41
2.12	Transformation by A_{1000} . The effect is one of translating the image a small distance along the line segment joining addresses 0 and 1000.	41
2.13	Transformation by A_{10^5} . The effect is one of translating the image along a line segment joining the addresses 0 and 100000. The amount of translation makes visible a <i>wrap around</i> effect.	42
2.14	Transformation by M_2 . The effect is one of rotation by 60 degrees.	43
2.15	Transformation by M_3 . The effect is one of rotation by 120 degrees.	43
2.16	Transformation by M_4 , The effect is one of rotation by 180 degrees.	44
2.17	Transformation by M_5 . The effect is one of rotation by 240 degrees.	44
2.18	Transformation by M_6 . The effect is one of rotation by 300 degrees.	45
2.19	(upper left) original image, (upper right) application of M_{100000} to the original image, (middle left) application of M_{10000} to the original image, (middle right) application of M_{1000} to the original image, (lower left) application of M_{100} to the original image, (lower right) application of M_{10} to the original image. The observable effects in these transformations are 1. rotation, 2. scaling, 3. production of multiple near copies of the original image.	46
2.20	The inverse transformation of M_{15} . The observable effects are 1. rotation, 2. seven contractions of the original image of which six are partial representations of the original image.	48

2.21	The inverse transformation of M_{12} . The observable effects are similar to those of M_{15}^{-1} differing in the amount of rotation, degree of contraction and the number of multiple copies.	48
2.22	A compound transformation, inverse transformation of M_{15} followed by A_{100000} . The partial copies observable in Figure 2.20 have been wrapped around by the translation.	49
3.1	The unwinding of the Spiral Honeycomb Mosaic indicates; i) the one-to-one correspondence between SHM and the integers; ii) the spiral pattern contained within it.	58
3.2	Processors in a pipe, ring and grid.	65
3.3	A flow diagram indicating the communication between the Application Handler and the $n \times H_i$ processes.	71
4.1	A collection of 7 pixels each of which is initialized with four pieces of information as indicated within the detached hexagon.	79
4.2	A flow diagram of the segmentation program.	84
4.3	The diagram indicates the relationship between the active and passive processes of 2DSA for a 49 hexagon SHM. The active processes are marked with a \star	85
4.4	A synthetically generated circle.	94
4.5	The 49 patterns generated by 2DSA for the representation of the circle (Figure 4.4) represented at the lowest resolution.	95
4.6	The 7 patterns generated by 2DSA for the representation of the circle (Figure 4.4) represented at the middle resolution.	96
4.7	The single pattern generated by 2DSA for the representation of the circle (Figure 4.4) represented at the highest resolution.	97
4.8	A synthetically generated annulus.	98

4.9	A synthetically generated trapezoid.	99
4.10	The 7 patterns generated by 2DSA for the representation of the trapezoid (Figure 4.9) represented at the middle resolution.	100
4.11	The single pattern generated by 2DSA for the representation of the trapezoid (Figure 4.9) represented at the highest resolution.	101
4.12	(upper left) original image with 7^6 pixels, (upper right) application of M_{100000} to the original image producing 7 unique copies each of which contains 7^5 pixels, (middle left) application of M_{10000} to the original image producing 49 unique copies each of which contains 7^4 pixels, (middle right) application of M_{1000} to the original image producing 7^3 unique copies each of which contains 7^3 pixels, (lower left) application of M_{100} to the original image producing 7^4 unique copies each of which contains 49 pixels, (lower right) application of M_{10} to the original image producing 7^5 unique copies each of which contains 7 pixels.	103
4.13	The image represents a scaling up of one of the copies from 4.12 (middle right).	105
4.14	The collection of images represent six iterations of the segmentation algorithm as applied to the image from Figure 4.13; (upper left) first iteration at highest δ grey-value producing 2 conjugacy classes, (upper right) second iteration, (middle left) third iteration, (middle right) fourth iteration, (lower left) 8th iteration, (lower right) 12th iteration.	106
4.15	The image represents a scaling up of one of the 49 copies in Figure 4.12 (middle left).	108

4.16	The collection of images represents six iterations of the segmentation algorithm applied to the image in Figure 4.15; (upper left) first iteration, (upper right) second iteration, (middle left) third iteration, (middle right) fourth iteration, (lower left) fifth iteration, (lower right) sixth iteration.	109
4.17	The image represents a scaling up of one of the 7 copies in Figure 4.12 (upper right).	110
4.18	The collection of images represents four iterations of the segmentation algorithm as applied to the image of Figure 4.17; (upper left) first iteration, (upper right) second iteration, (lower left) third iteration, (lower right) fourth iteration.	111
5.1	The projection of two images at different distances from the observer onto the retina.	116
5.2	A point P at a distance Z from the base line joining points E_l and E_r	119
5.3	The geometry of stereo vision.	120
5.4	The figure indicates the relationship between the positions of seven clusters of the SHM to the seven units of 3-D space that they represent. The optical axis intersects each of the seven clusters.	123
5.5	The optical axis intersects the centre of each hexagonal end of a hexahedron which is a component of the visual world.	124
5.6	The figure indicates the relative movements of the seven clusters from Figure 5.4 when a transformation of rotation is applied to the SHM.	128
5.7	Horizontal cross section of a simplified left eye viewing a patch P in the visual world. The three photo receptors are labelled 1, 2 and 3.	129
5.8	Cross section of the right eye's view of the same patch P displayed in Figure 5.7.	130

5.9	The combined effects of Figure 5.7 and 5.8 shows the patch locked in the intersection of two corridors.	131
5.10	(upper) Trapezoid in a plane which is perpendicular to the optical axis of the observer. (lower) Same trapezoidal shape as in (upper) but with shading, giving the illusion of a rectangle in space tilted away from the perpendicular to the optical axis.	132
5.11	The figure indicates (right) the motion of an object along a circular path and (left) the representation of the position of the object in SHM at seven units of time.	135
5.12	The figure indicates (right) the motion of an object along the path of a vertical line and (left) the representation in SHM of the object's position at seven units of time.	135
5.13	The figure indicates (right) a stationary object and (left) the representation in SHM of the stationary object over seven units of time. . . .	136
5.14	The figure shows an object (darkened hexagon) in the centre of the visual field.	137
5.15	The curvature of the can produces an occluded strip from E_r which is visible to E_l	141
5.16	Binocular view of swinging pendulum with both eyes receiving unfiltered light.	143
5.17	One eye, E_r , receives filtered light. The perceived motion of the pendulum is along the path of an ellipsoid.	144
5.18	Showing the actual position of a swinging pendulum at two points in time and the perceived position of a single point in time.	145
6.1	Architectural overview of Akamai.	151

6.2	Cone distribution along the horizontal meridian in humans. Adapted from [23].	153
6.3	Distribution of a non-uniform sampling scheme. The shaded hexagons of the combined three fundamental units of vision in regions, 1, 2 and 3, indicates the distribution of one such scheme.	154
6.4	A collection of hexagons is partitioned into two equivalence classes. The unshaded hexagons constitute an open class and the collection of shaded hexagons is an example of a closed equivalence class.	158
6.5	Viewing this figure induces the perception of a rectangle due to the presence of a corner etched in each of the four blobs.	159
6.6	Two sinusoidal curves with the same amplitude and different frequencies appear as aliases when sampled by the array of photo receptors of the figure.	169

Abstract

Spiral Architecture for Machine Vision

Phillip Sheridan

This thesis presents a new and powerful approach to the development of a general purpose machine vision system. The approach is inspired from anatomical considerations of the primate's vision system. The geometrical arrangement of cones on a primate's retina can be described in terms of a hexagonal grid. The importance of the hexagonal grid is that it possesses special computational features that are pertinent to the vision process. The fundamental thrust of this thesis emanates from the observation that this hexagonal grid can be described in terms of the mathematical object known as a Euclidean ring. The Euclidean ring is employed to generate an algebra of linear transformations which are appropriate for the processing of multidimensional vision data. A parallel autonomous segmentation algorithm for multidimensional vision data is described. The algebra and segmentation algorithm are implemented on a network of transputers. The implementation is discussed in the context of the outline of a general purpose machine vision system's design.