Appendices

A.1 Tests and constraints for centre-biased DIEM

Tests

DIEM is lossless if	DIEM is non-redundant if, when S is odd:	DIEM is non-redundant if, when S is even:
$\frac{(W-1)}{2}\left(\frac{2}{S-1}\right)^{\gamma} \leq 1$	$\frac{(W-1)}{2}\left[1-\left(\frac{S-3}{S-1}\right)^{\gamma}\right] \geq 1$	$(W-1)\left[1-\left(\frac{S-2}{S-1}\right)^{\gamma}\right] \geq 1$

Constraints
(assuming $S > 3$, $W > 3$)

Given S, DIEM is lossless if	Given S, DIEM is non-redundant if, when S is odd:	when S is even:
$\gamma \leq \frac{\log\left(\frac{2}{W-1}\right)}{\log\left(\frac{2}{S-1}\right)}$	$\gamma \geq \frac{\log\left(\frac{W-3}{W-1}\right)}{\log\left(\frac{S-3}{S-1}\right)}$	$\gamma \geq \frac{\log\left(\frac{W-2}{W-1}\right)}{\log\left(\frac{S-2}{S-1}\right)}$
Given γ, DIEM is lossless if	Given γ, DIEM is non-redundant if, when S is odd:	when S is even:
$S \geq 1+2\left(\frac{W-1}{2}\right)^{1/\gamma}$	$S \leq \frac{3-\left(\frac{W-3}{W-1}\right)^{1/\gamma}}{1-\left(\frac{W-3}{W-1}\right)^{1/\gamma}}$	$S \leq \frac{2-\left(\frac{W-2}{W-1}\right)^{1/\gamma}}{1-\left(\frac{W-2}{W-1}\right)^{1/\gamma}}$

A.2 Tests and constraints for periphery-biased DIEM

Tests

DIEM is lossless if, when S is odd:	when S is even:	DIEM is non-redundant if
$\frac{(W-1)}{2}\left[1-\left(\frac{S-3}{S-1}\right)^{\gamma}\right]\leq 1$	$(W-1)\left[1-\left(\frac{S-2}{S-1}\right)^{\gamma}\right]\leq 1$	$\frac{(W-1)}{2}\left(\frac{2}{S-1}\right)^{\gamma}\geq 1$

Constraints
(assuming $S > 3$, $W > 3$)

Given S, DIEM is lossless if, when S is odd:	when S is even:	Given S, DIEM is non-redundant if
$\gamma \leq \frac{\log\left(\frac{W-3}{W-1}\right)}{\log\left(\frac{S-3}{S-1}\right)}$	$\gamma \leq \frac{\log\left(\frac{W-2}{W-1}\right)}{\log\left(\frac{S-2}{S-1}\right)}$	$\gamma \geq \frac{\log\left(\frac{2}{W-1}\right)}{\log\left(\frac{2}{S-1}\right)}$
Given γ, DIEM is lossless if, when S is odd:	**when S is even:**	**Given γ, DIEM is non-redundant if**
$S \geq \frac{3-\left(\frac{W-3}{W-1}\right)^{1/\gamma}}{1-\left(\frac{W-3}{W-1}\right)^{1/\gamma}}$	$S \geq \frac{2-\left(\frac{W-2}{W-1}\right)^{1/\gamma}}{1-\left(\frac{W-2}{W-1}\right)^{1/\gamma}}$	$S \leq 1+2\left(\frac{W-1}{2}\right)^{1/\gamma}$

A.3 A feed-forward neural network universal Turing machine

Turing (1937) showed that it is possible to program a Finite State Machine (FSM) to emulate the behaviour of any other Finite State Machine. The emulating FSM is referred to as a Universal Turing Machine (UTM). If a UTM is given not only input data but also an encoded description of how the emulated machine responds to input, the UTM has all the information it needs to carry out the emulation.

All FSMs operate in the same way. They read input data one symbol at a time. The input data is assumed to reside in a one-dimensional storage medium such as a magnetic tape. On the basis of the symbol read and its internal state, the FSM determines its response. A FSM response consists of three components executed in the following order: 1) write a symbol to the tape or erase the symbol; 2) move left or right, or halt, 3) change the internal state (or leave it unchanged as the case may be).

How the FSM actually determines which of its responses to execute is governed by a set of rules which can be most easily expressed as a table combining all possible input symbols with all possible machine states. For each combination the table shows which symbol to write (if any), which way to move, and which internal state to switch to. Many computer scientists have devised state tables for UTMs, trying to find the most economical way of balancing the number of symbols with the number of internal states required (there is a trade-off). The smallest table so far devised is one of 28 responses (Minsky 1962), which uses four symbols and seven states (see **Table A.1**).

As described, for each combination of symbol and state, the table explicitly shows the write, move and state components of the machine's response. Obviously, after the machine has executed the three components it is in a new position, a new state, and is ready to read the next symbol. It is by combining these simple three-part responses in series that a FSM follows its program. In this way, its behaviour is completely deterministic.

Table A.1: Minsky's 4 x 7 UTM table.

	1	2	3	4	5	6	7
Y	- L 1	- L 1	Y L 3	Y L 4	Y R 5	Y R 6	- R 7
-	- L 1	Y R 2	HALT	Y R 5	Y L 3	A L 3	Y R 6
1	1 L 2	A R 2	A L 3	1 L 7	A R 5	A R 6	1 R 7
A	1 L 1	Y R 6	1 L 4	1 L 4	1 R 5	1 R 6	- R 2

In the special case of the UTM, the machine must read both the input data and the coded response of the emulated FSM. There are different ways of doing this, but the differences are not significant for our purpose: creating a neural network UTM.

It is possible to express the task of the UTMNN as a series of data pairs that need to be matched. For example, if presented with the input data Y 1 the network is expected to produce the output - L 1 in accordance with the table. Each input is paired with a single output. Given a training schedule consisting of the 28 combinations of symbol and state that go to make up possible input patterns, and a single response is determined for each of these patterns. The next question is how to encode the information for the network.

To reduce the input and output of the network to very simple binary form, eleven input units are necessary, a group of four which encode the symbol, and a group of seven which encode the state (there are more economical ways of doing this, but they are perhaps less clear). Fourteen units provide the output: four for the symbol, three for the direction of movement (left, halt, right), and seven for the state. The network will expect two of its eleven input cells to fire and will normally be expected to fire three of its output cells (with the one exception of the HALT result, which should arise in response to input - 3).

INPUT UNITS			Weights Between Input and Hidden Units					
Tape Symbols	Y		7.729032	-3.870370	-7.660379	13.399901	25.710225	-18.606300
	-		-16.693832	-18.669296	4.106279	-17.529072	5.172025	1.427290
	1	1	8.574345	11.762844	12.123982	6.188542	-15.616526	-18.597000
	A		-2.126480	3.570529	-20.256548	-4.316801	-6.984530	-18.603700
Machine States	1		-18.116674	-10.990840	-15.500270	25.778959	3.639780	-18.605200
	2		-18.210676	-3.784772	-5.980641	-5.374546	15.942346	1.430979
	3		6.204913	-12.453465	12.896041	5.005033	-3.748457	-18.606900
	4	1	20.861471	-9.636881	-12.826670	6.136215	2.410542	-18.601600
	5		21.274579	6.931598	19.071245	-19.909859	-6.297978	-18.601200
	6		-11.860240	4.346012	3.802850	-23.873782	-15.823005	-18.602300
	7		-7.089428	18.459587	-19.176890	6.433266	14.813735	-18.603500

	Bias Unit						
HIDDEN UNITS	1	1.00000000	.89340114	.33121637	.99999556	.00000184	.00000000

OUTPUT UNITS			Weights Between Hidden and Output Units						
Tape Symbols	Y		-115.596600	180.534727	-90.172925	-81.539323	-153.223230	206.500813	71.999680
	-		-190.188558	-92.310897	48.351965	-53.149334	119.632556	119.569355	-72.000400
	1	1	53.237468	15.070859	21.875877	-138.400808	12.834001	-141.890045	-72.000000
	A		-119.838413	-51.369787	57.662229	168.495201	19.788807	-125.492939	-72.000000
Left	L	1	38.895885	38.521211	-199.658550	12.928770	135.899409	-124.843490	-71.999600
Halt	H		-152.263007	-73.718474	-51.569600	110.601593	-38.176754	111.207769	-144.000000
Right	R		-45.445175	-9.247729	179.621666	-73.773954	-113.195620	100.477936	72.000360
Machine States	1		-42.552041	-99.476411	-126.848434	-53.541602	88.346679	6.858076	-143.999600
	2		-165.563989	-185.030925	122.838242	21.129527	123.251368	-12.527964	360.000600
	3		-137.169274	42.903832	-135.728222	181.889457	78.330039	-113.712299	-72.000000
	4		-110.685406	121.363200	-121.864001	-97.792401	74.763779	-43.137080	-72.000800
	5		-125.502322	162.623872	80.628741	-75.157787	-127.602178	21.935530	-72.000400
	6		-51.952973	-120.974391	133.189972	-20.001459	-192.736410	31.807559	-143.999600
	7	1	-237.192615	95.924049	120.201246	-72.834226	111.975492	-14.548476	-72.000000

Figure A.1: A Universal Turing Machine (UTM) neural network.

A simple three-layer network configured as shown is capable of performing as a UTM, and is therefore capable of carrying out any computation performed by any other Turing machine.

Although there are 28 possible inputs, there are only 20 possible outputs. For a network to have the ability to make binary distinctions between 20 different outputs it must have at least five hidden units ($2^5 = 32 > 20$). A bias unit in the hidden layer allows the network to respond more flexibly. Note it is not necessary to have a bias unit in the input layer as there are always exactly two input units firing. This means the bias unit value (that is, the threshold for each hidden unit) can be spread across the other input unit weights. In the case of hidden units, it is uncertain how many will be firing at any one time, and therefore it is not possible to determine how to spread the output unit threshold values.

The network is shown in **Figure A.1**. The 164 (11 × 6 + 7 × 14) weights have been calculated using standard back-propagation and the network is shown converting the input pattern 1 4 to the output pattern 1 L 7, which can be verified against Minsky's table.

Though five hidden units plus one bias unit are theoretically sufficient, there are six hidden units in addition to the bias unit. During training the network converged on a set of weights which produced 27 correct responses and one incorrect response. This was pragmatically solved by hand-wiring a sixth hidden unit into the system (the far right one). This unit is triggered by only one particular input pattern and in turn triggers only one output unit.

McCulloch and Pitts (1943) showed that, in principle, a neural network could function as a universal Turing machine. This simply demonstrates that it is possible in practice.

A.4 Glossary of terms

WHERE I have used words in a very specific sense, or have used words that are not well known outside a particular discipline (some neuroscience terms, for example) I have included their definition below. Words with **bold** typeface in the glossary definitions can themselves be found as glossary entries. In order to avoid ambiguity I have endeavoured at all times to use words and phrases consistently. It is a feature of all areas of research and learning that individuals adapt language to their own purposes and the consequence is often equivocation, ambiguity and conflation. Attempts to standardise a language are usually short-lived but may occasionally help to inject some order into linguistic chaos. Many of these definitions have been chosen from other researchers, for example, Polana and Nelson (1993).

Active vision, n, a closed-loop form of robot vision in which the camera's direction of gaze is determined by the content of the images it receives.

Activity, n, a motion which exhibits temporal or spatial repetition, such as walking, running, animals flying, machinery rotating or reciprocating, etc. *c.f.* **motion event**, **temporal texture**.

Adaptation, n, a qualitative metric for evaluating the breadth of performance of intelligent systems. *c.f.* **attainment**, **autonomy**.

Afferent, adj, input, incoming, applied to signals, esp. neural.

Affordability, n, a feature of **DIEM**.

Agent, n, a relatively discrete entity that transforms **afferent** information into **efferent** information, or sensation into action.

Animal vision, **biological vision**.

Artificial vision, n, manufactured processes that appear to emulate important characteristics of animal vision.

Attainment, n, a qualitative metric for evaluating the height of performance of intelligent systems. *c.f.* **adaptation**, **autonomy**.

Autonomy, n, a qualitative metric for evaluating the depth of performance of intelligent systems (your attainment minus the attainment of your props). *c.f.* **adaptation**, **attainment**.

Behavioural, adj, a level of analysis that asks 'why are you doing this?'.

Biological vision, n, the process by which optical information is used to guide the drives of animals towards their *de facto* goals.

Centroid of interest, n, an algorithm for selecting a single direction of gaze.

Computer vision, **artificial vision**.

Dædalus, n, 1. A computer that helped. Part of **WRAITH** at one time. 2. An art journal.

Dennett, n, aka Garfield, a computer that helped. Neither name being the title of an art journal.

Dimensionally independent exponential mapping, n, a method of sampling data with many advantages over standard sampling techniques.
Directed Perception, n, the manufacturers of the pan-tilt hardware and software used in **WRAITH**.
DIEM, **Dimensionally independent exponential mapping**.
Dynamic receptive field, n, the track described in a scene by a pixel contained in a randomly moving platform (e.g., A shaky robot-mounted camera).
Efferent, adj, output, outgoing, applied to signals, esp. neural.
Elementary motion detector, n, an abstract unit of motion detection (biological or artificial).
EMD, Elementary motion detector.
Existential, adj, a level of analysis that asks 'is there *anything* out there?'
Fixation, n, a short period during which the eyes are still and focused on something.
Focus, 1. n, v, the state of holding a clear image, or of making an image clear, 2. n, v, a concentration of attention.
Fovea, n, a small area in the centre of the retina, capable of the highest spatial, temporal and spectral acuity.
Foveate, v, to turn the eye or camera so that a particular object is perceived by the fovea.
Functional, adj, a level of analysis that asks 'what *are* you doing?'
Ganglion, n, (pl. ganglia), an inter-neuron in the central nervous system that combines the signals of other neurons and relays its signal to somewhere else.
Gaze control, n, the mechanics and ability to direct a camera or eye in a desired direction.
Human motion suitability, n, a feature of **DIEM**.
Incremental unsupervised learning, n, learning without a pre-established categorisation of phenomena, and without any differentiation between learn mode and operate mode.
Intentional, adj, a level of analysis that asks 'what's *that* supposed to mean?'
Interpretive unit, n, any computational unit with a specific function (e.g., detecting colour, edges, motion). Such a unit has a spatial reference defined by its receptive field, a set of inputs (either pixel values or the output of other units, or a mixture of both), and creates an output of variable strength.
Invertibility, n, a feature of **DIEM**.
JIGSAW, n, an algorithm for developing fundamental spatial mapping.
Lateral geniculate nucleus, n, the area in the brain where the optic nerve terminates on other neurons. The LGN's (one in each half of the brain) in turn send signals to the visual cortex (area V1).
Leonardo, n, 1. A computer that helped. Part of **WRAITH** at one time. 2. An art journal.
LGN, **lateral geniculate nucleus**.
Locus, n, (pl. loci), a curve made up of a set of points that all obey some rule.
Machine vision, **artificial vision**.
Motion event, n, an isolated simple motion which exhibits no temporal or spatial repetition, such as opening a door, starting a car, throwing a ball, etc. *c.f.* **activity**, **temporal texture**.
Motorium, n, (pl. motoria) all that is moved by an agent, the sum of all **efferent** signals, or the apparatus that delivers those signals. *c.f.* **sensorium**.
Natural vision, **biological vision**.
Ontological, adj, a level of analysis that asks if 'is there *anybody* there?'

Optic nerve, n, a nervous tract joining the eye to the (rest of) the brain. It would be more appropriately called the 'optic tract' since it connects ganglia to ganglia, not sensors to ganglia.

Opusculi, n, meaning 'small works', a general purpose term including both publications and unpublished matter (such as paintings).

Perseveration, n, a psychological term given to a form of behaviour in which the subject realises it is ineffective, but is unable to forsake it. Not to be confused with the more commonly used perseverance.

Photoreceptor, n, cells in the retina, which transduce light, turning it into small electrical impulses that trigger all visual processes.

Physical, adj, a level of analysis that asks 'does this belong to *you*?'

Physical world, n, the 'real world', the world that humans and animals inhabit, without restriction or constraint for the purposes of ensuring experimental success.

Pixel value, n, a local measure of light intensity in monochrome images, a local measure of primary colour intensities in colour images.

Random dot kinematogram, n, a tool used by vision researchers, a sequence of images in which many dots move randomly, but dots in specific areas within the field move in some coherent way.

RDK, **random dot kinematogram**.

Real world, **physical world**.

Receptive field, n, the area within the visual field to which a particular interpretive unit responds at a given time.

Retina, n, the thin membrane at the back of the eye that contains all the photoreceptors.

Robot vision, **artificial vision**.

Saccade, n, a rapid eye movement, enabling the eye to quickly foveate objects lying in different directions.

Sensorium, n, (pl. sensoria) all that is sensed by an agent, the sum of all **afferent** signals, or the apparatus that delivers those signals. *c.f.* **motorium**.

Social-cultural, adj, a level of analysis that asks 'what *will* people think?'

Space-variant, adj, mapping of space inside vision systems, including biological ones, is not linear, it varies across space.

Spatio-temporal, adj, a level of analysis that asks 'what have you been doing with yourself?'

Spatio-temporal cube, n, the three-dimensional data structure formed by the two spatial dimensions (x and y), and the single temporal dimension of a sequence of images.

ST cube, **spatio-temporal cube**.

Striate cortex, n, the part of the brain that does most of the vision processing, so named because it is multi-layered. It appears that there is considerable specialisation of function within the layers.

Superior colliculus, n, the visual reflex centre of the brain.

SURPRISE, n, an algorithm for deciding what is interesting and what is not, without really knowing enough to say either way.

Synecdochism, n, the act of letting the part stand for the whole, such as calling a worker a 'hand'.

Temporal texture, blown trees or grass, turbulent flow in cloud patterns, ripples on water, the motion of a flock of birds, etc. *c.f.* **activity**, **motion event**.

Unsupervised incremental learning, **incremental unsupervised learning**.

V1, **striate cortex**.

Visual cortex, **striate cortex**.

Visual robotics, n, biologically-informed vision for robots.

Visual servoing, n, the act of moving the camera with motors.

Winner-take-all, adj, n, a type of network that modulates incoming values such that the highest value is made non-zero and all others are made zero.

WRAITH, n, 1. The system comprising a video camera, frame grabber, computer, pan-tilt controller, pan-tilt motors, tripod, various cables and the software that inhabited this machinery; 2. The algorithms used to self-organise, sample, evaluate, and respond to information, namely, **JIGSAW**, **DIEM**, and **SURPRISE**. So named because the first implementations of the program displayed eerie and ephemeral ghost-like images of moving people on an RGB monitor. A built-in image subtraction analogue function in the DT frame grabber used at the time produced these images in real-time.

WTA, **Winner-take-all**.

A.5 Relevant opusculi by the author

Peters, M.W. (1992a). *Fantagma V*. Collection of the artist, Sydney.

Peters, M.W. (1992b). *Triangle (seed 1-1-1 base 4 generation 7)*. Collection of the artist, Sydney.

Peters, M.W., & Sowmya, A. (1996). Active vision and adaptive learning. D.P. Casasent (Ed.), *Intelligent Robots and Computer Vision XV*, Vol. 2904 (pp. 413-424), Boston: SPIE.

Peters, M.W., & Sowmya, A. (1997). WRAITH: ringing the changes in a changing world, *Fourth Conference of the Australasian Cognitive Science Society*, Newcastle, Australia: (available on CD, ISBN 0 7259 1059 3).

Peters, M.W. (1998a). Spatial competence via self-organisation: an intersect of perception and development, *Conference of the Cognitive Science Society 1998* (pp. 830-835), Madison, Wisconsin: Lawrence Erlbaum Associates.

Peters, M.W. (1998b). Towards artificial forms of intelligence, creativity, and surprise, *Conference of the Cognitive Science Society 1998* (pp. 836-841), Madison, Wisconsin: Lawrence Erlbaum Associates.

Peters, M.W., & Sowmya, A. (1998a). A real-time variable sampling technique for active vision: DIEM, *International Conference on Pattern Recognition 1998*, Vol. 1 (pp. 316-321), Brisbane, Australia: IEEE.[1]

Peters, M.W., & Sowmya, A. (1998b). Integrated techniques for self-organisation, sampling, habituation, and motion-tracking in visual robotics applications, *Machine Vision Applications 1998* (pp. 199-202), Makuhari, Japan: IAPR.

---

[1] Winner of the ICPR 1998 Best Student Paper prize.

A.6 References

Abbott, A.L., & al. (1993). Promising directions in active vision. *International Journal of Computer Vision, 11*(2), 109-126.

Akita, K. (1984). Image sequence analysis of real world human motion. *Pattern Recognition, 17*(1), 73-83.

Allmen, M., & Dyer, C. (1993). Computing spatiotemporal relations for dynamic perceptual organisation. *Computer Vision, Graphics and Image Processing, 58*(3), 338-351.

Ancona, N., & Poggio, T. (1995). Optical flow from 1D correlation: application to a simple time-to-crash detector. *International Journal of Computer Vision, 14*(2), 131-146.

Assereto, M., Figari, G., & Tesei, A. (1994). Robust approach to tracking human motion in real scenes. *Electronic Letters, 30*(24), 2013-2014.

Badler, N.I. (1976). A concept model for the description of image sequences, *The Milwaukee Symposium on Automatic Computation and Control* (pp. 377-381), Milwaukee: IEEE.

Bartolini, F., Cappellini, V., & Mecocci, A. (1994). Counting people getting in and out of a bus by real-time image-sequence processing. *Image and Vision Computing, 12*(1), 36-41.

Bederson, B.B., Wallace, R.S., & Schwartz, E. (1995). A miniaturized space-variant active vision system: Cortex-1. *Machine Vision and Applications, 1995*(8), 101-109.

Beer. R.D., Quinn, R.D., Chiel, H.J., Ritzmann, R.E. (1997). Biologically inspired approaches to robotics: what can we learn from insects? *Communications of the ACM, 40*(3), 31-38.

Ben-arie, J., & Z.-Q., Wang. (1998). Pictorial recognition of objects employing affine invariance in the frequency domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(6), 604-618.

Berger, J. (1972). *Ways of seeing*. London: British Broadcasting Corporation.

Berthouze, L., Rougeaux, S., Chavand, F., & Kuniyoshi, Y. (1996). A learning stereo-head control system. In M. Jamshidi, F. Pin, & P. Dauchez (Eds.), *Robotics and manufacturing: recent trends in research and applications* (pp. 43-47). New York: ASME Press.

Bhanu, B., & Poggio, T. (1994). Introduction to the special section on learning in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 16*(9), 865-868.

Bischof, W.F., & Caelli, T. (1994). Learning structural descriptions of patterns: a new technique for conditional clusterings and rule generation. *Pattern Recognition, 27*(5), 689-697.

Blake, A., & Yuille, A. (1992). *Active vision*. Cambridge, Massachusetts: MIT Press.

Blewitt, M. (1980). *Celestial navigation for yachtsmen* (8th. ed.). London: Stanford Maritime.

Boden, M. (1992). *The creative mind*. London: Abacus.

Boyce, J.F., & Toulson, D.L. (1993). Implementation of an automatic object tracking system onto a transputer surface, *Fourth Annual Conference of the Meiko User Society*, Southampton.

Brooks, R.A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation, RA-2*(1), 14-23.

Brooks, R.A. (1989). The whole iguana. In M. Brady (Ed.), *Robotics science* (pp. 432-456). Cambridge, Massachusetts: MIT Press.

Brooks, R.A., & Stein, L.A. (1994). Building brains for bodies. *Autonomous Robots, 1*(1), 7-25.

Bruce, V., & Green, P.R. (1990). *Visual perception: physiology, psychology and ecology* (2nd ed.). Hove: Lawrence Erlbaum Associates.

Caelli, T., & Dreier, A. (1994). Variations on the evidence-based object recognition theme. *Pattern Recognition, 27*(2), 185-204.

Campbell, R., Landis, T., & Regard, M. (1986). Face recognition and lipreading: a neurological dissociation. *Brain, 109*(3), 509-521.

Camus, T., Monsarrat, J., & Dean, T. (1993). Planning and selective perception for target retrieval, *Image Understanding Workshop* (pp. 593-597), Washington, D.C.: Morgan Kauffman.

Carlson, N.R. (1991). *Physiology of behavior* (4th ed.). Boston: Allyn and Bacon.

Cedras, C., & Shah, M. (1995). Motion-based recognition: a survey. *Image and Vision Computing, 13*(2), 129-155.

Chella, A., Frixione, M., & Gaglio, S. (1997). A cognitive architecture for artificial vision. *Artificial Intelligence, 89*(1-2), 73-111.

Chen, Z., & Lee, H.-J. (1992). Knowledge-guided visual perception of 3-D human gait from a single image sequence. *IEEE Transactions on Systems, Man, and Cybernetics, 22*(2), 336-342.

Clancey, W.J. (1993). Situated action: a neuropsychological response to Vera and Simon. *Cognitive Science, 17*(1), 87-116.

Crevier, D. (1993). *AI: the tumultuous history of the search for artificial intelligence*. New York: Basic Books.

Crick, F., & Koch, C. (1992). The problem of consciousness. *Scientific American* (September), 110-117.

Crick, F. (1994). *The astonishing hypothesis: the scientific search for the soul*. New York: Macmillan.

Crisman, J.D., & Thorpe, C.E. (1993). SCARF: a color vision system that tracks roads and intersections. *IEEE Transactions on Robotics and Automation, 9*(1), 49-58.

Dance, S., & Caelli, T. (1993a). A symbolic object-oriented picture interpretation network: SOO-PIN, *Advances in Structural and Syntactic Pattern Recognition: Proceedings of the International Workshop*. Bern: World Scientific Publishing.

Dance, S., & Caelli, T. (1993a). On the symbolic interpretation of traffic scenes, *Asian Conference on Computer Vision* (pp. 798-801), Osaka: ACCV.

Dance, S., Caelli, T., & Liu, Z.-Q. (1996). A concurrent, hierarchical approach to symbolic dynamic scene interpretation. *Pattern Recognition, 29*(11), 1891-1903.

Dawkins, R. (1987). *The blind watchmaker: why the existence of evolution reveals a universe without design*. New York: Norton.

Dawkins, R. (1989). *The selfish gene* (new ed.). Oxford: Oxford University Press.

Day, R.H. (1972). *Human perception*. Sydney: John Wiley.

Dennett, D.C. (1987). *The intentional stance*. Cambridge, Massachusetts: MIT Press.

Dennett, D.C. (1991). *Consciousness explained*. London: Penguin.

Dennett, D.C. (1996). *Kinds of minds*. New York: Basic Books.

Dong, D.W., & Atick, J.J. (1995a). Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus. *Network: Computation in Neural Systems, 6*(2), 159-178.

Dong, D.W., & Atick, J.J. (1995b). Statistics of natural time-varying images. *Network: Computation in Neural Systems, 6*(3), 345-358.

Dreyfus, H.L. (1993). *What computers still can't do: a critique of artificial reason*. Cambridge, Massachusetts: MIT Press.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Ellis, A.W., & Young, A.W. (1988). *Human cognitive neuropsychology*. Hove: Lawrence Erlbaum Associates.

Epstein, R. (1992) Can machines think? *AI Magazine, 13*(2), 80-95.

Ernst, B. (1986). *The eye beguiled* (Karen Williams, Trans.). Germany: Taschen.

Fearnley-Sander, D. (1999). Mathematical structures for rational discourse, *Fourth Asian Technology Conference in Mathematics*, (pp. 311-319), Guangzhou, People's Republic of China.

Feller, M.B., Wellis, D.P., Stellwagen, D., Werblin, F.S., & Shatz, C.J. (1996). Cholinergic synaptic transmission is required for wave propagation in developing retina. *Science, 272*, 1182.

Fermüller, C., & Aloimonos, Y. (1995). Vision and action. *Image and Vision Computing, 13*(10), 725-744.

Freeman, W.J. (1991). The physiology of perception. *Scientific American* (February), 78-85.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gleick, J. (1988). *Chaos: making a new science*. London: Penguin.

Goddard, N.H. (1992). *The perception of articulated motion: recognizing moving light displays*. PhD, University of Rochester.

Goldstein, E.B. (1989). *Sensation and perception* (3rd. ed.). Pacific Grove: Brooks/Cole Publishing.

Grimson, W.E.L., Ratan, A.L., O'Donnell, P.A., & Klanderman, G. (1994). An active visual attention system to play "Where's Waldo", *Image Understanding Workshop*, Vol. 2 (pp. 1059-1065), Monterey, California: Morgan Kauffman.

Harnad, S. (1991). The symbol grounding problem. In S. Forrest (Ed.), *Emergent computation*. Cambridge Massachusetts: MIT Press.

Healey, G. (1993). Hierarchical segmentation-based approach to motion analysis. *Image and Vision Computing, 11*(9), 570-576.

Hebb, D.O. (1949). *The organization of behavior: a neurophysiological theory*. New York: Wiley.

Held, R., & Hein, A. (1963). Movement produced stimulation in the development of visually guided behaviour. *Journal of Comparative and Physiological Psychology, 56*, 872-876.

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American* (November), 84-94.

Hess, R.H., Baker, C.J., & Zihl, J. (1989). The "motion-blind" patient: low-level spatial and temporal filters. *Journal of Neuroscience, 9*(5), 1628-1640.

Hofstadter, D.R. (1997). *Le ton beau de Marot: in praise of the music of language*. London: Basic Books.

Hogan, J.P. (1997). *Mind matters: exploring the world of artificial intelligence*. New York: Ballantine.

Hogg, D. (1983). Model based vision: a program to see a walking person. *Image and Vision Computing, 1*(1), 5-20.

Hubel, D.H., & Weisel, T.N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London, Series B, 198*, 1-59.

Hubel, D.H., Wiesel, T.N., & Stryker, M.P. (1978). Anatomical demonstration of orientation columns in macaque monkey. *Journal of Comparative Neurology, 177*, 361-380.

Ikeuchi, K., & Suehiro, T. (1992). Towards an assembly plan from observation, *IEEE International Conference on Robotics and Automation* (pp. 2171-2177), Nice, France: IEEE.

Ioerger, T.R. (1994). The manipulation of images to handle indeterminacy in spatial reasoning. *Cognitive Science, 18*(4), 551-593.

Jain, A., & Hoffman, D. (1988). Evidence based recognition of objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 10*(6), 783-802.

Johnson-Laird, P. (1993). *The computer and the mind: an introduction to cognitive science* (2nd ed.). London: Fontana Press.

Kang, S.B., & Ikeuchi, K. (1993). Toward automatic robot instruction from perception - recognizing a grasp from observation. *IEEE Transactions on Robotics and Automation, 9*(4), 432-443.

Kauffman, S.A. (1991). Antichaos and adaptation. *Scientific American* (August), 78-84.

Koestler, A. (1975). *The act of creation*. London: Picador.

Köhler, W. (1927). *The mentality of apes* (E. Winter, Trans.). (2nd ed.). London: Routledge & Kegan Paul.

Kolb, B., & Whishaw, I.Q. (1990). *Fundamentals of human neuropsychology* (3rd ed.). New York: W H Freeman.

Konishi, M. (1993). Listening with two ears. *Scientific American* (April), 34-41.

Kuniyoshi, Y., Inaba, M., & Inoue, H. (1992). Seeing, understanding and doing human task, *IEEE International Conference on Robotics and Automation* (pp. 2-9), Nice, France: IEEE.

Kuniyoshi, Y., & Inoue, H. (1993). Qualitative recognition of ongoing human action sequences, *International Joint Conference on Artificial Intelligence* (pp. 1600-1609), Chambery, France: Morgan Kauffman.

Kuniyoshi, Y., Inaba, M., & Inoue, H. (1994). Learning by watching: extracting reusable task knowledge from visual observation of human performance. *IEEE Transactions on Robotics and Automation, 10*(6), 799-822.

Kuniyoshi, Y., Kita, N., Sugimoto, K., Nakamura, S., & Suehiro, T. (1995). A foveated wide angle lens for active vision, *IEEE International Conference on Robotics and Automation* (pp. 2982-2988), Nagoya, Japan: IEEE.

Land, E.H., & McCann, J.J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America, 61*(1), 1-11.

Lauinger, N. (1996). The inverted retina of the human eye: a trichromatic 4D space-time optical correlator, *Intelligent Robots and Computer Vision XV*, Vol. 2904 (pp. 344-360), Boston: SPIE.

Letang, J.-M., Rebuffel, V., & Bouthemy, P. (1992). Motion detection based on a temporal multiscale approach, *International Conference on Pattern Recognition*, Vol. A (pp. 65-68), The Hague, Netherlands: IEEE.

Leung, M.K., & Yang, Y.-H. (1987). A region-based approach for human body motion analysis. *Pattern Recognition, 20*(3), 321-339.

Levy, S. (1992). *Artificial life: the quest for a new creation*. London: Jonathan Cape.

Linsker, R. (1988). Self organisation in a perceptual network. *Computer* (March), 105-117.

Maes, P. (1992). Invited talk, *Third International Symposium on Electronic Art*, Sydney: Australian Network for Art and Technology.

Mardia, K.V., Ghali, N.M., Hainsworth, T.J., Howes, M., & Sheehy, N. (1993). Techniques for online gesture recognition on workstations. *Image and Vision Computing, 11*(5), 283-294.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society of London, Series B, 200*, 269-294.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. New York: W H Freeman.

McCulloch, W.S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 5*, 115-133.

Merzenich, M.M., & Kaas, J.H. (1982). Reorganisation of mammalian somatosensory cortex following peripheral nerve injury. *Trends in Neurosciences, 5*(12), 434-436.

Metcalfe, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(2), 288-294.

Metcalfe, J. (1986b). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(4), 623-634.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition, 15*(3), 238-246.

Minsky, M. (1962). Size and structure of universal Turing machines, *Recursive function theory; Symposium in Pure Mathematics* (pp. 229-238). Providence: American Mathematical Society.

Minsky, M. (1985). *The society of mind*. New York: Simon & Schuster.

Mitchison, G. (1991). Neuronal branching patterns and the economy of cortical wiring. *Proceedings of the Royal Society of London, Series B, 245*, 151-158.

Murray, D.W., Bradshaw, K.J., McLaughlan, P.F., Reid, I.D., & Sharkey, P.M. (1995). Driving saccade to pursuit using image motion. *International Journal of Computer Vision, 16*(3), 205-228.

Nilsson, N.J. (1995). Eye on the prize. *AI Magazine, 16*(2), 9-17.

Norman, D.A. (1993). Cognition in the head and in the world: an introduction to the special issue on situated action. *Cognitive Science, 17*(1), 1-6.

Palhang, M. (1998). *Learning object models: an inductive logic programming approach*. PhD, University of New South Wales.

Palhang, M., & Sowmya, A. (1999). Learning discriminatory and descriptive rules by an inductive logic programming system, *International Conference on Machine Learning* (pp. 288-297), Bled, Slovenia: Morgan Kauffman.

Papanikolopoulos, N.P., Khosla, P.K., & Kanade, T. (1993). Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision. *IEEE Transactions on Robotics and Automation, 9*(1), 14-35.

Papanikolopoulos, N.P., & Khosla, P.K. (1993). Adaptive robotic visual tracking: theory and experiments. *IEEE Transactions on Automatic Control, 38*(3), 429-445.

Park, E.H., Ntuen, C.A., Vootla, S., & Park, Y.H. (1994). Adaptive learning of human motion by a telerobot using a neural network model as a teacher. *Computers and Industrial Engineering, 27*(1-4), 453-456.

Park, S.S., & Sowmya, A. (1998). Autonomous robot navigation by active visual motion analysis and understanding, *Machine Vision Applications 1998* (pp. 38-41), Makuhari, Japan: IAPR.

Paulos, J.A. (1980). *Mathematics and humour*. Chicago: University of Chicago Press.

Pearce, A.R., & Caelli, T. (1994). Learning relational structures: applications in computer vision. *Journal of Applied Intelligence, 4*, 257-268.

Pearce, A.R., Caelli, T., & Bischof, W.F. (1994). Rulegraphs for graph matching in pattern recognition. *Pattern Recognition, 27*(9), 1231-1247.

Peitgen, H.-O., & Richter, P.H. (1986). *The beauty of fractals: images of complex dynamical systems*. Berlin: Springer-Verlag.

Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man*. New York: Macmillan.

Penfield, W., & Jasper, H.H. (1954). *Epilepsy and the functional anatomy of the human brain*. Boston: Little & Brown.

Penrose, R. (1989). *The emperor's new mind: concerning computers, minds, and the laws of physics*. London: Vintage.

Poggio, T., & Stringa, L. (1992). A project for an intelligent system: vision and learning. *International Journal of Quantum Chemistry, 42*(4), 727-739.

Polana, R., & Nelson, R. (1993). Detecting activities, *Image Understanding Workshop* (pp. 569-574), Washington, D.C.: Morgan Kauffman.

Pollefeys, M., & Van Gool, L. (1997). A stratified approach to metric self-calibration, *Computer Vision and Pattern Recognition* (pp. 407-412), Puerto Rico: IEEE.

Pollefeys, M., Koch, R., & Van Gool, L. (1998). Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, *International Conference on Computer Vision* (pp. 90-95), Bombay: IEEE.

Pope, A.R., & Lowe, D.G. (1993). Learning object recognition models from images, *International Conference on Computer Vision* (pp. 296-301), Berlin: IEEE.

Pope, A.R. (1994). Model-based object recognition: a survey of recent research. Vancouver: University of British Columbia.

Pribram, K.H. (1991). *Brain and perception*. New Jersey: Lawrence Erlbaum Associates.

Prokopowicz, P.N. (1995). *The development of perceptual integration across eye movements in visual robots*. PhD, Northwestern University.

Prokopowicz, P.N., & Cooper, P.R. (1995a). The Dynamic Retina: contrast and motion detection for active vision. *International Journal of Computer Vision, 16*(3), 191-204.

Prokopowicz, P.N., & Cooper, P.R. (1995b). IRV: learning to integrate visual information across camera movements, *International Joint Conference on Artificial Intelligence*, Vol. 2 (pp. 2051-2052), Montreal: Morgan Kauffman.

Prokopowicz, P.N., & Cooper, P.R. (1996). Integrating visual information across camera movements with a visual-motor calibration map, *Conference of the American Association for Artificial Intelligence* (pp. 1124-1131), Portland, Oregon: AAAI.

Rao, R.P.N. (1994). Top-down gaze targeting for space-variant active vision, *Image Understanding Workshop*, Vol. 2 (pp. 1049-1058), Monterey, California: Morgan Kauffman.

Rao, R.P.N., & Ballard, D.H. (1995). An active vision architecture based on iconic representations. *Artificial Intelligence, 78*(1-2), 461-505.

Raymond, J.E. (1994). Directional anisotropy of motion sensitivity across the visual field. *Vision Research, 34*(8), 1029-1037.

Reiter, R. & Mackworth, A.K. (1989). A logical framework for depiction and image interpretation. *Artificial Intelligence 41*(2), 125-155.

Remagnino, P., Illingworth, J., Kittler, J., & Matas, J. (1995). Intentional control of camera look direction and viewpoint in an active vision system. *Image and Vision Computing, 13*(2), 79-88.

Restak, R. (1991). *The brain has a mind of its own: insights from a practicing neurologist*. New York: Crown.

Rohr, K. (1994). Towards model-based recognition of human movements in image sequences. *Computer Vision, Graphics, and Image Processing: Image Understanding, 59*(1), 94-115.

Saul, A.B., & Humphrey, A.L. (1990). Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus. *Journal of Neurophysiology, 64*(1), 206-224.

Saul, A.B., & Humphrey, A.L. (1992). Evidence of input from lagged cells in the lateral geniculate nucleus to simple cells in cortical area 17 of the cat. *Journal of Neurophysiology, 68*(4), 1190-1208.

Seyfarth, R.M., & Cheney, D.L. (1992). Meaning and mind in monkeys. *Scientific American* (December), 78-84.

Sharma, R., & Srinivasa, N. (1996). Saccade control for an active stereo head using a learned spatial representation, *IEEE International Symposium on Intelligent Control* (pp. 91-96), Dearborn, Michigan: IEEE.

Sharp, N.G., & Hancock, E.R. (1995). Feature tracking by multi-frame relaxation. *Image and Vision Computing, 13*(8), 637-644.

Simon, H.A. (1969). *Sciences of the artificial*. Cambridge, Massachusetts: MIT Press.

Sims, K. (1992). Interactive evolution of equations for procedural models, *Third International Symposium on Electronic Art* (pp. 128), Sydney: Australian Network for Art and Technology.

Smith, S.M., & Brady, J.M. (1995). ASSET-2: real-time motion segmentation and shape tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 17*(8), 814-820.

Solso, R.L. (1991). *Cognitive psychology*. Needham Heights, Massachusetts: Allyn & Bacon.

Solso, R.L. (1996). *Cognition and the visual arts*. Cambridge, Massachusetts: MIT Press.

Sparks, D.L., Lee, D., & Rohrer, W.H. (1990). Population coding of the direction, amplitude, and velocity of saccadic eye movements in the superior colliculus, *Cold Spring Harbor Symposia on Quantitative Biology XV* (pp. 805-811), Cold Spring Harbor: Cold Spring Harbor Laboratory.

Stadler, M., Richter, P.H., Pfaff, S., & Kruse, P. (1991). Attractors and perceptual field dynamics of homogeneous stimulus area. *Psychological Research, 53*(2), 102-112.

Swain, M.J., & Stricker, M.A. (1993). Promising directions in active vision. *International Journal of Computer Vision, 11*(2), 109-126.

Temple, C. (1993). *The brain: an introduction to the psychology of the human brain and behaviour*. London: Penguin.

Thibadeau, R. (1986). Artificial perception of actions. *Cognitive Science, 10*(2), 117-149.

Todd, J., & Reichel, T. (1989). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review, 96*(4), 643-657.

Tsotsos, J.K. (1984a). The scope of motion research: from image intensity changes to semantic abstractions. *Computer Graphics, 18*(1), 7-11.

Tsotsos, J.K. (1984b). Knowledge and the visual process: content, form and use. *Pattern Recognition, 17*(1), 13-26.

Tsotsos, J.K., Culhane, S.M., Wai, W.Y.K., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence, 78*(1-2), 507-545.

Tsuji, S., Morizono, A., & Kuroda, S. (1977). Understanding a simple cartoon film by a computer vision system, *International Joint Conference on Artificial Intelligence* (pp. 609-610), Cambridge, Massachusetts: Morgan Kauffman.

Turing, A.M. (1937). On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, 2*(42), 230-265.

Turing, A.M. (1950). Computing machinery and intelligence. *Mind, LIX*(236), 433-460.

van der Smagt, P. (1995). *Visual robot arm guidance using neural networks*. PhD, University of Amsterdam.

Vincze, M., & Weiman, C.F.R. (1996). A general relationship for optimal tracking performance, *Intelligent Robots and Computer Vision XV*, Vol. 2904 (pp. 402-412), Boston Massachusetts: SPIE.

Waltz, D. (1972). Understanding line drawings of scenes with shadows. In P.H. Winston (Ed.), *Psychology of computer vision*. Cambridge, Massachusetts, MIT Press.

Waltz, D. (1998) AAAI Presidential Address, The importance of importance. *Fifteenth National Conference on Artificial Intelligence*, Madison, Wisconsin.

Weisberg, R.W. (1992). Metacognition and insight during problem solving: comment on Metcalfe. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 426-431.

Welch, R.B. (1978). *Perceptual modification: adapting to altered sensory environments*. New York: Academic Press.

Wertheimer, M. (1959). *Productive thinking* (enlarged ed.). New York: Harper & Brothers.

Williams, M.-A. (1996). Aesthetics and the explication of surprise. *Languages of design, 3*, 145-157.

Wills, C. (1993). *The runaway brain: the evolution of human uniqueness*. London: HarperCollins.

Wolfram, S. (1983). Cellular automata. *Los Alamos Science, 9*, 2-21.

Wurtz, R.H., & Albano, J.E. (1980). Visual-motor function of the primate superior colliculus. *Annual Review of Neuroscience, 3*, 189-226.

Zanker, J.M., & Hupgens, I.S. (1994). Interaction between primary and secondary mechanisms in human motion perception. *Vision Research, 34*(10), 1255-1266.

Zeki, S. (1992). The visual image in mind and brain. *Scientific American* (September), 42-50.

Zucker, S.W. (1981). Computer vision and human perception: an essay on the discovery of constraints, *International Joint Conference on Artificial Intelligence* (pp. 1102-1116), Vancouver: William Kauffman.