

Computational Neuroscience of Vision

EDMUND T. ROLLS

University of Oxford, Department of Experimental Psychology, Oxford, UK

and

GUSTAVO DECO

Siemens, Munich; and University of Munich, Germany

OXFORD
UNIVERSITY PRESS

Contents

1	Introduction	1
1.1	Introduction and overview	1
1.2	Neurons	2
1.3	Neurons in a network	2
1.4	Synaptic modification	4
1.5	Long-Term Potentiation and Long-Term Depression	7
1.6	Distributed representations	11
1.6.1	Definitions	11
1.6.2	Advantages of different types of coding	12
1.7	Neuronal network approaches versus connectionism	13
1.8	Introduction to three neuronal network architectures	14
1.9	Systems-level analysis of brain function	16
1.10	The fine structure of the cerebral neocortex	21
1.10.1	The fine structure and connectivity of the neocortex	21
1.10.2	Excitatory cells and connections	21
1.10.3	Inhibitory cells and connections	23
1.10.4	Quantitative aspects of cortical architecture	25
1.10.5	Functional pathways through the cortical layers	27
1.10.6	The scale of lateral excitatory and inhibitory effects, and the concept of modules	29
1.11	Backprojections in the cortex	30
1.11.1	Architecture	30
1.11.2	Learning	31
1.11.3	Recall	33
1.11.4	Semantic priming	34
1.11.5	Attention	34
1.11.6	Autoassociative storage, and constraint satisfaction	34
2	The primary visual cortex	36
2.1	Introduction and overview	36
2.2	Retina and lateral geniculate nuclei	37
2.3	Striate cortex: Area V1	43
2.3.1	Classification of V1 neurons	43
2.3.2	Organization of the striate cortex	45
2.3.3	Visual streams within the striate cortex	48

2.4	Computational processes that give rise to V1 simple cells	49
2.4.1	Linsker's method: Information maximization	50
2.4.2	Olshausen and Field's method: Sparseness maximization	53
2.5	The computational role of V1 for form processing	55
2.6	Backprojections to the lateral geniculate nucleus	55
3	Extrastriate visual areas	57
3.1	Introduction	57
3.2	Visual pathways in extrastriate cortical areas	57
3.3	Colour processing	61
3.3.1	Trichromacy theory	61
3.3.2	Colour opponency, and colour contrast: Opponent cells	61
3.4	Motion and depth processing	65
3.4.1	The motion pathway	65
3.4.2	Depth perception	67
4	The parietal cortex	70
4.1	Introduction	70
4.2	Spatial processing in the parietal cortex	70
4.2.1	Area LIP	71
4.2.2	Area VIP	73
4.2.3	Area MST	74
4.2.4	Area 7a	74
4.3	The neuropsychology of the parietal lobe	75
4.3.1	Unilateral neglect	75
4.3.2	Balint's syndrome	77
4.3.3	Gerstmann's syndrome	79
5	Inferior temporal cortical visual areas	81
5.1	Introduction	81
5.2	Neuronal responses in different areas	81
5.3	The selectivity of one population of neurons for faces	83
5.4	Combinations of face features	84
5.5	Distributed encoding of object and face identity	84
5.5.1	Distributed representations evident in the firing rate distributions	85
5.5.2	The representation of information in the responses of single neurons to a set of stimuli	90
5.5.3	The representation of information in the responses	

of a population of inferior temporal visual cortex neurons	94
5.5.4 Advantages for brain processing of the distributed representation of objects and faces	98
5.5.5 Should one neuron be as discriminative as the whole organism, in object encoding systems?	103
5.5.6 Temporal encoding in the spike train of a single neuron	105
5.5.7 Temporal synchronization of the responses of different cortical neurons	108
5.5.8 Conclusions on cortical encoding	111
5.6 Invariance in the neuronal representation of stimuli	112
5.6.1 Size and spatial frequency invariance	112
5.6.2 Translation (shift) invariance	113
5.6.3 Reduced translation invariance in natural scenes	113
5.6.4 A view-independent representation of objects and faces	115
5.7 Face identification and face expression systems	118
5.8 Learning in the inferior temporal cortex	120
5.9 Cortical processing speed	122
5.10 Conclusions	125
6 Visual attentional mechanisms	126
6.1 Introduction	126
6.2 The classical view	126
6.2.1 The spotlight metaphor and feature integration theory	126
6.2.2 Computational models of visual attention	129
6.3 Biased competition – single cell studies	132
6.3.1 Neurophysiology of attention	133
6.3.2 The role of competition	135
6.3.3 Evidence of attentional bias	136
6.3.4 Non-spatial attention	136
6.3.5 High-resolution buffer hypothesis	139
6.4 Biased competition – fMRI	140
6.4.1 Neuroimaging of attention	140
6.4.2 Attentional effects in the absence of visual stimulation	141
6.5 The computational role of top-down feedback connections	142
7 Neural network models	145

7.1	Introduction	145
7.2	Pattern association memory	145
7.2.1	Architecture and operation	146
7.2.2	The vector interpretation	149
7.2.3	Properties	150
7.2.4	Prototype extraction, extraction of central tendency, and noise reduction	151
7.2.5	Speed	151
7.2.6	Local learning rule	152
7.2.7	Implications of different types of coding for storage in pattern associators	158
7.3	Autoassociation memory	159
7.3.1	Architecture and operation	160
7.3.2	Introduction to the analysis of the operation of au- toassociation networks	161
7.3.3	Properties	163
7.3.4	Use of autoassociation networks in the brain	170
7.4	Competitive networks, including self-organizing maps	171
7.4.1	Function	171
7.4.2	Architecture and algorithm	171
7.4.3	Properties	173
7.4.4	Utility of competitive networks in information pro- cessing by the brain	178
7.4.5	Guidance of competitive learning	180
7.4.6	Topographic map formation	182
7.4.7	Radial Basis Function networks	187
7.4.8	Further details of the algorithms used in competitive networks	188
7.5	Continuous attractor networks	192
7.5.1	Introduction	192
7.5.2	The generic model of a continuous attractor network	195
7.5.3	Learning the synaptic strengths between the neu- rons that implement a continuous attractor network	196
7.5.4	The capacity of a continuous attractor network	198
7.5.5	Continuous attractor models: moving the activity packet of neuronal activity	198
7.5.6	Stabilization of the activity packet within the continu- ous attractor network when the agent is stationary	202
7.5.7	Continuous attractor networks in two or more dimen- sions	203

7.5.8	Mixed continuous and discrete attractor networks	203
7.6	Network dynamics: the integrate-and-fire approach	204
7.6.1	From discrete to continuous time	204
7.6.2	Continuous dynamics with discontinuities	205
7.6.3	Conductance dynamics for the input current	207
7.6.4	The speed of processing of one-layer attractor networks with integrate-and-fire neurons	209
7.6.5	The speed of processing of a four-layer hierarchical network with integrate-and-fire attractor dynamics in each layer	212
7.6.6	Spike response model	215
7.7	Network dynamics: introduction to the mean field approach	216
7.8	Mean-field based neurodynamics	218
7.8.1	Population activity	218
7.8.2	A basic computational module based on biased competition	220
7.8.3	Multimodular neurodynamical architectures	221
7.9	Interacting attractor networks	224
7.10	Error correction networks	228
7.10.1	Architecture and general description	229
7.10.2	Generic algorithm (for a one-layer network taught by error correction)	229
7.10.3	Capability and limitations of single-layer error-correcting networks	230
7.10.4	Properties	234
7.11	Error backpropagation multilayer networks	236
7.11.1	Introduction	236
7.11.2	Architecture and algorithm	237
7.11.3	Properties of multilayer networks trained by error backpropagation	238
7.12	Biologically plausible networks	239
7.13	Reinforcement learning	240
7.14	Contrastive Hebbian learning: the Boltzmann machine	241
8	Models of invariant object recognition	243
8.1	Introduction	243
8.2	Approaches to invariant object recognition	244
8.2.1	Feature spaces	244
8.2.2	Structural descriptions and syntactic pattern recognition	245

8.2.3	Template matching and the alignment approach	247
8.2.4	Invertible networks that can reconstruct their inputs	248
8.2.5	Feature hierarchies	249
8.3	Hypotheses about object recognition mechanisms	253
8.4	Computational issues in feature hierarchies	257
8.4.1	The architecture of VisNet	258
8.4.2	Initial experiments with VisNet	266
8.4.3	The optimal parameters for the temporal trace used in the learning rule	274
8.4.4	Different forms of the trace learning rule, and their relation to error correction and temporal difference learning	275
8.4.5	The issue of feature binding, and a solution	284
8.4.6	Operation in a cluttered environment	295
8.4.7	Learning 3D transforms	301
8.4.8	Capacity of the architecture, and incorporation of a trace rule into a recurrent architecture with object attractors	307
8.4.9	Vision in natural scenes — effects of background versus attention	313
8.5	Synchronization and syntactic binding	319
8.6	Further approaches to invariant object recognition	320
8.7	Processes involved in object identification	321
9	The cortical neurodynamics of visual attention – a model	323
9.1	Introduction	323
9.2	Physiological constraints	324
9.2.1	The dorsal and ventral paths of the visual cortex	324
9.2.2	The biased competition hypothesis	326
9.2.3	Neuronal receptive fields	327
9.3	Architecture of the model	328
9.3.1	Overall architecture of the model	328
9.3.2	Formal description of the model	331
9.3.3	Performance measures	336
9.4	Simulations of basic experimental findings	336
9.4.1	Simulations of single-cell experiments	337
9.4.2	Simulations of fMRI experiments	339
9.5	Object recognition and spatial search	341
9.5.1	Dynamics of spatial attention and object recognition	343
9.5.2	Dynamics of object attention and visual search	345

9.6	Evaluation of the model	348
9.6.1	Spatial attention and object attention	348
9.6.2	Translation-invariant object recognition	350
9.6.3	Contributions and limitations	351
10	Visual search: Attentional neurodynamics at work	353
10.1	Introduction	353
10.2	Simple visual search	354
10.3	Visual search of hierarchical patterns	358
10.3.1	The spatial resolution hypothesis	358
10.3.2	Neurodynamics of the resolution hypothesis	361
10.3.3	Visual search in the framework of the resolution hypothesis	363
10.4	Visual conjunction search	369
10.4.1	The binding problem	369
10.4.2	The time course of conjunction search: experimental evidence	371
10.4.3	Extension of the computational cortical architecture	373
10.4.4	Computational results	376
10.5	Conclusion	381
11	A computational approach to the neuropsychology of visual attention	383
11.1	Introduction	383
11.2	The neglect syndrome	383
11.2.1	A model of visual spatial neglect	384
11.2.2	Spatial cueing effect on neglect	388
11.2.3	Extinction and visual search	390
11.2.4	Effect on neglect of top-down knowledge about objects	392
11.3	Hierarchical patterns – neuropsychology	398
11.4	Conjunction search – neuropsychology	400
11.4.1	Simulations and predictions	400
11.4.2	Experimental test of the predictions in human subjects	401
11.5	Conclusion	403
12	Outputs of visual processing	404
12.1	Visual outputs to Short Term Memory systems	406
12.1.1	Prefrontal cortex short term memory networks, and their relation to temporal and parietal perceptual net-	

works	406
12.1.2 Computational details of the model of short term memory	409
12.1.3 Computational necessity for a separate, prefrontal cortex, short term memory system	412
12.1.4 Role of prefrontal cortex short term memory systems in visual search and attention	412
12.1.5 Synaptic modification is needed to set up but not to reuse short term memory systems	413
12.2 Visual outputs to Long Term Memory systems in the brain	413
12.2.1 Effects of damage to the hippocampus and connected structures on object-place and episodic memory	414
12.2.2 Neurophysiology of the hippocampus and connected areas	415
12.2.3 Hippocampal models	418
12.2.4 The perirhinal cortex, recognition memory, and familiarity	421
12.3 Visual stimulus–reward association, emotion, and motivation	424
12.3.1 Emotion	425
12.3.2 Reward is not processed in the temporal cortical visual areas	429
12.3.3 Why the reward and punishment associations of stimuli are not represented early in information processing in the primate brain	430
12.3.4 Amygdala	434
12.3.5 Orbitofrontal cortex	439
12.3.6 Effects of mood on memory and visual processing	447
12.4 Output to object selection and action systems	448
12.5 Visual search	452
12.6 Visual outputs to behavioral response systems	453
12.7 Multimodal representations in different brain areas	453
12.8 Visuo–spatial scratchpad, and change blindness	454
12.9 Conscious visual perception -	454
13 Principles and Conclusions	456
13.1 Transform invariance in the inferior temporal visual cortex	456
13.2 Representation of information in IT	456
13.3 IT information processing is fast	457

13.4	Continuous neuronal dynamics allows fast network processing	457
13.5	Hierarchical feature analysis	457
13.6	Trace learning rule for invariant representations	459
13.7	Spatial feature binding by feature combination neurons	460
13.8	IT provides a representation for later memory networks	461
13.9	Face expression and object motion	462
13.10	Attentional mechanisms	462
13.11	Visual search	464
13.12	Egocentric vs allocentric representations	464
13.13	Short term memory as the controller of attention	465
13.14	Output to object selection and action systems	466
13.15	'What' versus 'where' processing streams	466
13.16	Short term memory must be separated from perception	467
13.17	Backprojections must be weak	468
13.18	Long-term potentiation and short-term memory	469
13.19	"Executive control" by the prefrontal cortex	469
13.20	Reward processing occurs after object identification	470
13.21	Effects of mood on memory and visual processing	471
13.22	Visual outputs to Long Term Memory systems	471
13.23	Episodic memory and the operation of mixed discrete and continuous attractor networks	472
13.24	Visual outputs to behavioural response systems	472
13.25	Multimodal representations in different brain areas	472
13.26	Visuo-spatial scratchpad and change blindness	472
13.27	Invariant object recognition and attention	473
13.28	Conscious visual perception	473
13.29	Attention – future directions	473
13.30	Integrated approaches to understanding vision	475
13.31	Apostasis	475
A	Introduction to linear algebra for neural networks	477
A.1	Vectors	477
A.1.1	The inner or dot product of two vectors	477
A.1.2	The length of a vector	478
A.1.3	Normalizing the length of a vector	479
A.1.4	The angle between two vectors: the normalized dot product	479
A.1.5	The outer product of two vectors	480
A.1.6	Linear and non-linear systems	481

A.1.7	Linear combinations of vectors, linear independence, and linear separability	482
A.2	Application to understanding simple neural networks	484
A.2.1	Capability and limitations of single-layer networks: linear separability and capacity	484
A.2.2	Non-linear networks: neurons with non-linear activation functions	487
A.2.3	Non-linear networks: neurons with non-linear activations	488
B	Information theory	490
B.1	Basic notions	490
B.1.1	The information conveyed by definite statements	491
B.1.2	Information conveyed by probabilistic statements	491
B.1.3	Information sources, information channels, and information measures	492
B.1.4	The information carried by a neuronal response and its averages	494
B.1.5	The information conveyed by continuous variables	496
B.2	The information carried by neuronal responses	498
B.2.1	The limited sampling problem	498
B.2.2	Correction procedures for limited sampling	500
B.2.3	The information from multiple cells: decoding procedures	501
B.2.4	Information in the correlations between the spikes of different cells	504
B.3	Information theory results	507
B.3.1	Temporal codes versus rate codes within the spike train of a single neuron	507
B.3.2	The speed of information transfer from single neurons	509
B.3.3	The information from multiple cells: independent information versus redundancy across cells	512
B.3.4	The information from multiple cells: the effects of cross-correlations between cells	514
B.3.5	Conclusions	517
B.4	Information theory terms — a short glossary	518
	References	520
	Index	565