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

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
### Contents

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 autoassociation 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 processing 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 neurons 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 continuous attractor network when the agent is stationary 202  
   7.5.7 Continuous attractor networks in two or more dimensions 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
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.2.3</td>
<td>Template matching and the alignment approach</td>
<td>247</td>
</tr>
<tr>
<td>8.2.4</td>
<td>Invertible networks that can reconstruct their inputs</td>
<td>248</td>
</tr>
<tr>
<td>8.2.5</td>
<td>Feature hierarchies</td>
<td>249</td>
</tr>
<tr>
<td>8.3</td>
<td>Hypotheses about object recognition mechanisms</td>
<td>253</td>
</tr>
<tr>
<td>8.4</td>
<td>Computational issues in feature hierarchies</td>
<td>257</td>
</tr>
<tr>
<td>8.4.1</td>
<td>The architecture of VisNet</td>
<td>258</td>
</tr>
<tr>
<td>8.4.2</td>
<td>Initial experiments with VisNet</td>
<td>266</td>
</tr>
<tr>
<td>8.4.3</td>
<td>The optimal parameters for the temporal trace used in the learning rule</td>
<td>274</td>
</tr>
<tr>
<td>8.4.4</td>
<td>Different forms of the trace learning rule, and their relation to error correction and temporal difference learning</td>
<td>275</td>
</tr>
<tr>
<td>8.4.5</td>
<td>The issue of feature binding, and a solution</td>
<td>284</td>
</tr>
<tr>
<td>8.4.6</td>
<td>Operation in a cluttered environment</td>
<td>295</td>
</tr>
<tr>
<td>8.4.7</td>
<td>Learning 3D transforms</td>
<td>301</td>
</tr>
<tr>
<td>8.4.8</td>
<td>Capacity of the architecture, and incorporation of a trace rule into a recurrent architecture with object attractors</td>
<td>307</td>
</tr>
<tr>
<td>8.4.9</td>
<td>Vision in natural scenes — effects of background versus attention</td>
<td>313</td>
</tr>
<tr>
<td>8.5</td>
<td>Synchronization and syntactic binding</td>
<td>319</td>
</tr>
<tr>
<td>8.6</td>
<td>Further approaches to invariant object recognition</td>
<td>320</td>
</tr>
<tr>
<td>8.7</td>
<td>Processes involved in object identification</td>
<td>321</td>
</tr>
<tr>
<td>9</td>
<td>The cortical neurodynamics of visual attention — a model</td>
<td>323</td>
</tr>
<tr>
<td>9.1</td>
<td>Introduction</td>
<td>323</td>
</tr>
<tr>
<td>9.2</td>
<td>Physiological constraints</td>
<td>324</td>
</tr>
<tr>
<td>9.2.1</td>
<td>The dorsal and ventral paths of the visual cortex</td>
<td>324</td>
</tr>
<tr>
<td>9.2.2</td>
<td>The biased competition hypothesis</td>
<td>326</td>
</tr>
<tr>
<td>9.2.3</td>
<td>Neuronal receptive fields</td>
<td>327</td>
</tr>
<tr>
<td>9.3</td>
<td>Architecture of the model</td>
<td>328</td>
</tr>
<tr>
<td>9.3.1</td>
<td>Overall architecture of the model</td>
<td>328</td>
</tr>
<tr>
<td>9.3.2</td>
<td>Formal description of the model</td>
<td>331</td>
</tr>
<tr>
<td>9.3.3</td>
<td>Performance measures</td>
<td>336</td>
</tr>
<tr>
<td>9.4</td>
<td>Simulations of basic experimental findings</td>
<td>336</td>
</tr>
<tr>
<td>9.4.1</td>
<td>Simulations of single-cell experiments</td>
<td>337</td>
</tr>
<tr>
<td>9.4.2</td>
<td>Simulations of fMRI experiments</td>
<td>339</td>
</tr>
<tr>
<td>9.5</td>
<td>Object recognition and spatial search</td>
<td>341</td>
</tr>
<tr>
<td>9.5.1</td>
<td>Dynamics of spatial attention and object recognition</td>
<td>343</td>
</tr>
<tr>
<td>9.5.2</td>
<td>Dynamics of object attention and visual search</td>
<td>345</td>
</tr>
</tbody>
</table>
9.6 Evaluation of the model
  9.6.1 Spatial attention and object attention
  9.6.2 Translation-invariant object recognition
  9.6.3 Contributions and limitations

10 Visual search: Attentional neurodynamics at work
  10.1 Introduction
  10.2 Simple visual search
  10.3 Visual search of hierarchical patterns
    10.3.1 The spatial resolution hypothesis
    10.3.2 Neurodynamics of the resolution hypothesis
    10.3.3 Visual search in the framework of the resolution hypothesis
  10.4 Visual conjunction search
    10.4.1 The binding problem
    10.4.2 The time course of conjunction search: experimental evidence
    10.4.3 Extension of the computational cortical architecture
    10.4.4 Computational results
  10.5 Conclusion

11 A computational approach to the neuropsychology of visual attention
  11.1 Introduction
  11.2 The neglect syndrome
    11.2.1 A model of visual spatial neglect
    11.2.2 Spatial cueing effect on neglect
    11.2.3 Extinction and visual search
    11.2.4 Effect on neglect of top-down knowledge about objects
  11.3 Hierarchical patterns – neuropsychology
  11.4 Conjunction search – neuropsychology
    11.4.1 Simulations and predictions
    11.4.2 Experimental test of the predictions in human subjects
  11.5 Conclusion

12 Outputs of visual processing
  12.1 Visual outputs to Short Term Memory systems
    12.1.1 Prefrontal cortex short term memory networks, and their relation to temporal and parietal perceptual net-
works
12.1.2 Computational details of the model of short term memory
12.1.3 Computational necessity for a separate, prefrontal cortex, short term memory system
12.1.4 Role of prefrontal cortex short term memory systems in visual search and attention
12.1.5 Synaptic modification is needed to set up but not to reuse short term memory systems
12.2 Visual outputs to Long Term Memory systems in the brain
12.2.1 Effects of damage to the hippocampus and connected structures on object-place and episodic memory
12.2.2 Neurophysiology of the hippocampus and connected areas
12.2.3 Hippocampal models
12.2.4 The perirhinal cortex, recognition memory, and familiarity
12.3 Visual stimulus–reward association, emotion, and motivation
12.3.1 Emotion
12.3.2 Reward is not processed in the temporal cortical visual areas
12.3.3 Why the reward and punishment associations of stimuli are not represented early in information processing in the primate brain
12.3.4 Amygdala
12.3.5 Orbitofrontal cortex
12.3.6 Effects of mood on memory and visual processing
12.4 Output to object selection and action systems
12.5 Visual search
12.6 Visual outputs to behavioral response systems
12.7 Multimodal representations in different brain areas
12.8 Visuo–spatial scratchpad, and change blindness
12.9 Conscious visual perception

13 Principles and Conclusions
13.1 Transform invariance in the inferior temporal visual cortex
13.2 Representation of information in IT
13.3 IT information processing is fast
13.4 Continuous neuronal dynamics allows fast network processing
13.5 Hierarchical feature analysis
13.6 Trace learning rule for invariant representations
13.7 Spatial feature binding by feature combination neurons
13.8 IT provides a representation for later memory networks
13.9 Face expression and object motion
13.10 Attentional mechanisms
13.11 Visual search
13.12 Egocentric vs allocentric representations
13.13 Short term memory as the controller of attention
13.14 Output to object selection and action systems
13.15 'What' versus 'where' processing streams
13.16 Short term memory must be separated from perception
13.17 Backprojections must be weak
13.18 Long-term potentiation and short-term memory
13.19 "Executive control" by the prefrontal cortex
13.20 Reward processing occurs after object identification
13.21 Effects of mood on memory and visual processing
13.22 Visual outputs to Long Term Memory systems
13.23 Episodic memory and the operation of mixed discrete and continuous attractor networks
13.24 Visual outputs to behavioural response systems
13.25 Multimodal representations in different brain areas
13.26 Visuo–spatial scratchpad and change blindness
13.27 Invariant object recognition and attention
13.28 Conscious visual perception
13.29 Attention – future directions
13.30 Integrated approaches to understanding vision
13.31 Apostasis

A Introduction to linear algebra for neural networks
A.1 Vectors
   A.1.1 The inner or dot product of two vectors
   A.1.2 The length of a vector
   A.1.3 Normalizing the length of a vector
   A.1.4 The angle between two vectors: the normalized dot product
   A.1.5 The outer product of two vectors
   A.1.6 Linear and non-linear systems