Contents

1	App	roach	and scope	1
	1.1	The app	proach	1
		1.1.1	Data, models, and theory	1
		1.1.2	From physiology to behavior and back via theory and models	3
	1.2	The pro	oblem of vision	4
		1.2.1	Visual tasks and subtasks	5
		1.2.2	Vision seen through visual encoding, selection, and decoding	7
		1.2.3	Visual encoding in retina and V1	10
		1.2.4	Visual selection and V1's role in it	12
		1.2.5	Visual decoding and its associated brain areas	14
2	A ve	ery brie	ef introduction of what is known about vision experimentally	16
	2.1	Neuron	s, neural circuits, and brain regions	16
		2.1.1	Neurons, somas, dendrites, axons, and action potentials	16
		2.1.2	A simple neuron model	17
		2.1.3	Random processes of action potential generation in neurons	18
		2.1.4	Synaptic connections, neural circuits, and brain areas	18
		2.1.5	Visual processing areas along the visual pathway	19
	2.2	Retina		22
		2.2.1	Receptive fields of retinal ganglion cells	22
		2.2.2	Sensitivity to sinusoidal gratings, and contrast sensitivity curves	25
		2.2.3	Responses to spatiotemporal inputs	30
		2.2.4	P and M cells	34
		2.2.5	Color processing in the retina	35
		2.2.6	Spatial sampling in the retina	37
		2.2.7	LGN on the pathway from the retina to V1	38
	2.3	V1		39
		2.3.1	The retinotopic map	39
		2.3.2	The receptive fields in V1—the feature detectors	40
A. 1		2.3.3	Orientation selectivity, bar and edge detectors	41
1 4		2.3.4	Spatial frequency tuning and multiscale coding	42
		2.3.5	Temporal and motion direction selectivity	43
		2.3.6	Ocular dominance and disparity selectivity	46
		2.3.7	Color selectivity of V1 neurons	48
		2.3.8	Complex cells	48
		2.3.9	The influences on a V1 neuron's response from contextual stimuli outside the	
			receptive field	52

riigilei visual aleas	54	3.6.6 Coupling the spatial and color coding in V1	154
2.4.1 Two processing streams	54	3.6.7 Coupling spatial coding with stereo coding in V1—coding disparity	161
2.4.2 V2	55	3.6.8 Coupling space and time coding in the retina and V1	164
2.4.3 MT (V5)	57	3.6.9 V1 neurons tuned simultaneously to multiple feature dimensions	167
2.4.4 V4	59	3.7 The efficient code, and the related sparse code, in low noise limit by numerical simulation	10/
2.4.5 IT and temporal cortical areas for object recognition	60	3.7.1 Sparse coding	
Eye movements, their associated brain regions, and links with attention	60	3.8 How to get efficient codes by developmental rules and unsupervised learning	171
2.5.1 Close link between eye movements and attention	62	3.8.1 Learning for a single encoding neuron	173
Top-down attention and neural responses	63	3.8.2 Learning simultaneously for multiple encoding neurons	174
Behavioral studies on vision	65		175
e efficient coding principle	67	4 V1 and information coding	177
A brief introduction to information theory	68	4.1 Pursuit of efficient coding in V1 by reducing higher order redundancy	177
3.1.1 Measuring information	68	4.1.1 Higher order statistics contain much of the meaningful information about visu.	al
3.1.2 Information transmission, information channels, and mutual information		objects	178
3.1.3 Information redundancy, representation efficiency, and error correction	70	4.1.2 Characterizing higher order statistics	180
Formulation of the efficient coding principle	74	4.1.3 Efforts to understand V1 neural properties from the perspective of reducing highs	er
3.2.1 An optimization problem	77	order redundancy	183
3.2.2 Exposition	77	4.1.4 Higher order redundancy in natural images is only a very small fraction of th	ie
	79	total redundancy	185
Efficient neural sampling in the retina 3.3.1 Contrast sampling in a fly's compound eve	83	4.2 Problems in understanding V1 solely based on efficient coding	186
	83	4.3 Multiscale and overcomplete representation in V1 is useful for invariant object recognition	n
and the state of t	85	from responses of selected neural subpopulations	187
and the second	89	4.3.1 Information selection, amount, and meaning	188
Efficient coding by visual receptive field transforms	90	5 The V1 hypothesis—creating a bottom-up saliency map for preattentive	
3.4.1 The general analytical solution for efficient coding of Gaussian signals	91	selection and segmentation	
Case study: stereo coding in V1 as an efficient transform of inputs in the dimension	on of	5.1 Visual selection and visual saliency	189
ocularity	96		189
3.5.1 Principal component analysis K _o	98	top down and bottom-up selections	189
3.5.2 Gain control	102	5.1.2 A brief overview of visual search and segmentation—behavioral studies o	f
3.5.3 Contrast enhancement, decorrelation, and whitening in the high S/N regime	105	saliency	195
3.5.4 Many equivalent solutions of optimal encoding	106	5.1.3 Saliency regardless of visual input features	197
3.5.5 Smoothing and output correlation in the low S/N region	108	5.1.4 A quick review of what we should expect about saliencies and a saliency map	200
3.5.6 A special, most local, class of optimal coding	110	5.2 The V1 saliency hypothesis	201
3.5.7 Adaptation of the optimal code to the statistics of the input environment	110	5.2.1 Detailed formulation of the V1 saliency hypothesis	202
3.5.8 A psychophysical test of the adaptation of the efficient stereo coding	117	5.2.2 Intracortical interactions in V1 as mechanisms to compute saliency	204
3.5.9 How might one test the predictions physiologically?	120	5.2.3 Reading out the saliency map	206
The efficient receptive field transforms in space, color, time, and scale in the retina a	and	5.2.4 Statistical and operational definitions of saliency	207
V1	120	5.2.5 Overcomplete representation in V1 for the role of saliency	208
3.6.1 Efficient spatial coding in the retina	123	5.3 A hallmark of the saliency map in V1—attention capture by an ocular singleton which is	
3.6.2 Efficient coding in time	134	barely distinctive to perception	209
3.6.3 Efficient coding in color	138	5.3.1 Food for thought: looking (acting) before or without seeing	215
3.6.4 Coupling space and color coding in the retina	142		215
3.6.5 Spatial coding in V1	147		216
		5.4.2 Calibration of the V1 model to biological reality	222

	5.4.3	Computational requirements on the dynamic behavior of the model	225
	5.4.4	Applying the V1 model to visual search and visual segmentation	227
	5.4.5	Other effects of the saliency mechanisms—figure-ground segmentation and the	ne
		medial axis effect	247
	5.4.6	Input contrast dependence of the contextual influences	250
	5.4.7	Reflections from the V1 model	250
5.5		onal psychophysical tests of the V1 saliency hypothesis	252
	5.5.1	The feature-blind "auction"—maximum rather than summation over features	252
	5.5.2	The fingerprints of colinear facilitation in V1	257
	5.5.3	The fingerprint of V1's conjunctive cells	260
	5.5.4	A zero-parameter quantitative prediction and its experimental test	266
	5.5.5	Reflections—from behavior back to physiology via the V1 saliency hypothesis	269
5.6	The ro	les of V1 and other cortical areas in visual selection	269
	5.6.1	Using visual depth feature to probe contributions of extrastriate cortex to atter	1-
		tional control	271
	5.6.2	Salient but indistinguishable inputs activate early visual cortical areas but not th	е
		parietal and frontal areas	275
5.7	V1's ro	le beyond saliency—selection versus decoding, periphery versus central vision	279
	5.7.1	Implications for the functional roles of visual cortical areas based on their repre)-
		sentations of the visual field	281
	5.7.2	Saliency, visual segmentation, and visual recognition	282
8.	Nonline	ear V1 neural dynamics for saliency and preattentive segmentation	285
	5.8.1	A minimal model of the primary visual cortex for saliency computation	286
	5.8.2	Dynamic analysis of the V1 model and constraints on the neural connections	299
	5.8.3	Extensions and generalizations	312
.9	Append	dix: parameters in the V1 model	313
/isi	al rec	ognition as decoding	045
.1		on of visual decoding	315
.2		notable observations about visual recognition	315
	6.2.1		317
	6.2.2	Recognition is after an initial selection or segmentation Object invariance	317
	6.2.3		318
	6.2.4	Is decoding the shape of an object in the attentional spotlight a default routine?	319
	6.2.5	Recognition by imagination or input synthesis	321
	6.2.6	Visual perception can be ambiguous or unambiguous	323
.3		Neural substrates for visual decoding	325
	6.3.1	lecoding from neural responses	326
	6.3.2	Example: decoding motion direction from MT neural responses	327
	6.3.3	Example: discriminating two inputs based on photoreceptor responses	330
	6.3.4	Example: discrimination by decoding the V1 neural responses	332
	6.3.5	Example: light wavelength discrimination by decoding from cone responses	334
	0.0.0	Perception, including illusion, of a visual feature value by neural population de- coding	
		County	338

		6.3.6	Poisson-like neural noise and increasing perceptual performance for strong	or
			visual inputs	34
		6.3.7	Low efficiency of sensory information utilization by the central visual system	345
		6.3.8	Transduction and central inefficiencies in the framework of encoding, attention	al
			selection, and decoding	346
	6.4	Bayes	ian inference and the influence of prior belief in visual decoding	347
		6.4.1	The Bayesian framework	348
		6.4.2	Bayesian visual inference is highly complex unless the number and the dimer	1-
			sions of possible percepts are restricted	349
		6.4.3	Behavioral evidence for Bayesian visual inference	350
	6.5	The ini	tial visual recognition, feedforward mechanisms, and recurrent neural connections	361
		6.5.1	The fast speed of coarse initial recognition by the primate visual system	361
		6.5.2	Object detection and recognition by models of hierarchical feedforward network	
		6.5.3	Combining feedforward and feedback intercortical mechanisms, and recurrer	nt
			intracortical mechanisms, for object inference	363
7	Enil	ogue		
'	7.1	_		364
	7.1		norance of vision viewed from the perspective of vision as encoding, selection, and	b
	7.2	decodir		364
			tational vision	365
Re	feren	ces		367
Ind	lev			
1110	CA			380