

# Contents

Preface	xv
<b>PART I BACKGROUND</b>	<b>1</b>
<b>1 About Face</b>	<b>3</b>
1.1 The Visual Face . . . . .	3
1.2 The Changing Face . . . . .	4
1.3 Computing Faces . . . . .	6
1.4 Biological Perspectives . . . . .	11
1.5 The Approach . . . . .	12
<b>2 Perception and Representation</b>	<b>15</b>
2.1 A Distal Object . . . . .	16
2.2 Representation by 3D Reconstruction . . . . .	18
2.3 Two-dimensional View-based Representation . . . . .	19
2.4 Image Template-based Representation . . . . .	21
2.5 The Correspondence Problem and Alignment . . . . .	22
2.6 Biological Perspectives . . . . .	26
2.7 Discussion . . . . .	29
<b>3 Learning under Uncertainty</b>	<b>31</b>
3.1 Statistical Learning . . . . .	32
3.2 Learning as Function Approximation . . . . .	33
3.3 Bayesian Inference and MAP Classification . . . . .	36
3.4 Learning as Density Estimation . . . . .	37
3.4.1 Parametric Models . . . . .	37
3.4.2 Non-parametric Models . . . . .	39
3.4.3 Semi-parametric Models . . . . .	40
3.5 Unsupervised Learning without Density Estimation . . . . .	41

3.5.1	Dimensionality Reduction . . . . .	43
3.5.2	Clustering . . . . .	43
3.6	Linear Classification and Regression . . . . .	44
3.6.1	Least-squares . . . . .	45
3.6.2	Linear Support Vector Machines . . . . .	45
3.7	Non-linear Classification and Regression . . . . .	48
3.7.1	Multi-layer Networks . . . . .	48
3.7.2	Support Vector Machines . . . . .	50
3.8	Adaptation . . . . .	51
3.9	Biological Perspectives . . . . .	53
3.10	Discussion . . . . .	54

## PART II FROM SENSORY TO MEANINGFUL PERCEPTION 57

<b>4</b>	<b>Selective Attention: Where to Look</b>	<b>59</b>
4.1	Pre-attentive Visual Cues from Motion . . . . .	60
4.1.1	Measuring Temporal Change . . . . .	61
4.1.2	Motion Estimation . . . . .	62
4.2	Learning Object-based Colour Cues . . . . .	65
4.2.1	Colour Spaces . . . . .	66
4.2.2	Colour Density Models . . . . .	68
4.3	Perceptual Grouping for Selective Attention . . . . .	71
4.4	Data Fusion for Perceptual Grouping . . . . .	73
4.5	Temporal Matching and Tracking . . . . .	76
4.6	Biological Perspectives . . . . .	77
4.7	Discussion . . . . .	78
<b>5</b>	<b>A Face Model: What to Look For</b>	<b>81</b>
5.1	Person-independent Face Models for Detection . . . . .	82
5.1.1	Feature-based Models . . . . .	82
5.1.2	Holistic Models . . . . .	83
5.1.3	The Face Class . . . . .	85
5.2	Modelling the Face Class . . . . .	86
5.2.1	Principal Components Analysis for a Face Model . . . . .	87
5.2.2	Density Estimation in Local PCA Spaces . . . . .	88
5.3	Modelling a Near-face Class . . . . .	89
5.4	Learning a Decision Boundary . . . . .	90

5.4.1	Face Detection in Dynamic Scenes . . . . .	91
5.4.2	Normalisation . . . . .	93
5.4.3	Face Detection using Multi-layer Perceptrons . . . . .	95
5.4.4	Face Detection using Support Vector Machines . . . . .	97
5.5	Perceptual Search . . . . .	99
5.6	Biological Perspectives . . . . .	101
5.7	Discussion . . . . .	102
<b>6</b>	<b>Understanding Pose</b>	<b>103</b>
6.1	Feature and Template-based Correspondence . . . . .	105
6.2	The Face Space across Views: Pose Manifolds . . . . .	105
6.3	The Effect of Gabor Wavelet Filters on Pose Manifolds . . . . .	111
6.4	Template Matching as Affine Transformation . . . . .	113
6.5	Similarities to Prototypes across Views . . . . .	118
6.6	Learning View-based Support Vector Machines . . . . .	121
6.7	Biological Perspectives . . . . .	123
6.8	Discussion . . . . .	123
<b>7</b>	<b>Prediction and Adaptation</b>	<b>125</b>
7.1	Temporal Observations . . . . .	128
7.2	Propagating First-order Markov Processes . . . . .	129
7.3	Kalman Filters . . . . .	131
7.4	Propagating Non-Gaussian Conditional Densities . . . . .	132
7.4.1	Learning Priors using HMMs and EM . . . . .	132
7.4.2	Observation Augmented Density Propagation . . . . .	133
7.5	Tracking Attended Regions . . . . .	134
7.6	Adaptive Colour Models . . . . .	136
7.7	Selective Adaptation . . . . .	140
7.8	Tracking Faces . . . . .	145
7.9	Pose Tracking . . . . .	147
7.9.1	Person-specific Pose Tracking . . . . .	149
7.9.2	Person-independent Pose Tracking . . . . .	150
7.10	Biological Perspectives . . . . .	156
7.11	Discussion . . . . .	157

<b>PART III</b>	<b>MODELS OF IDENTITY</b>	<b>161</b>
<b>8</b>	<b>Single-View Identification</b>	<b>163</b>
8.1	Identification Tasks . . . . .	163
8.2	Nearest-neighbour Template Matching . . . . .	165
8.3	Representing Knowledge of Facial Appearance . . . . .	166
8.4	Statistical Knowledge of Facial Appearance . . . . .	168
8.4.1	Low Dimensionality: Principal Components Analysis . . . . .	168
8.4.2	Separability: Linear Discriminant Analysis . . . . .	173
8.4.3	Sparseness and Topography: Local Feature Analysis . . . . .	176
8.5	Statistical Knowledge of Identity . . . . .	176
8.5.1	Identification Tasks Revisited . . . . .	177
8.5.2	Class-conditional Densities for Modelling Identity . . . . .	179
8.6	Structural Knowledge: The Role of Correspondence . . . . .	180
8.6.1	Beyond Alignment: Correspondence at a Single View . . . . .	181
8.6.2	Combining Statistical and Structural Models . . . . .	182
8.7	Biological Perspectives . . . . .	183
8.8	Discussion . . . . .	185
<b>9</b>	<b>Multi-View Identification</b>	<b>187</b>
9.1	View-based Models . . . . .	189
9.2	The Role of Prior Knowledge . . . . .	191
9.3	View Correspondence in Identification . . . . .	191
9.3.1	Learning Linear Shape Models . . . . .	192
9.3.2	Nonlinear Shape Models . . . . .	195
9.4	Generalisation from a Single View . . . . .	199
9.4.1	Identity by Linear Combination . . . . .	199
9.4.2	Identity by Similarity to Prototype Views . . . . .	201
9.5	Generalisation from Multiple Views . . . . .	203
9.6	Biological Perspectives . . . . .	205
9.7	Discussion . . . . .	207
<b>10</b>	<b>Identifying Moving Faces</b>	<b>209</b>
10.1	Biological Perspectives . . . . .	210
10.1.1	Learning and Identity Constancy . . . . .	210
10.1.2	The Effect of Temporal Order of Pose on Learning . . . . .	210
10.1.3	The Effect of Motion on Familiar Face Identification . . . . .	211
10.2	Computational Theories of Temporal Identification . . . . .	213

10.2.1	Atemporal Representation . . . . .	213
10.2.2	Spatio-temporal Signatures . . . . .	214
10.3	Identification using Holistic Temporal Trajectories . . . . .	215
10.4	Identification by Continuous View Transformation . . . . .	218
10.5	An Experimental System . . . . .	219
10.6	Discussion . . . . .	225

## PART IV PERCEPTION IN CONTEXT 227

<b>11</b>	<b>Perceptual Integration</b>	<b>229</b>
11.1	Sensory and Model-based Vision . . . . .	230
11.2	Perceptual Fusion . . . . .	231
11.3	Perceptual Inference . . . . .	236
11.3.1	Inference using Hidden Markov Models . . . . .	240
11.3.2	Closed-loop Perceptual Control . . . . .	241
11.4	Vision as Cooperating Processes . . . . .	242
11.4.1	Visual Attention and Grouping . . . . .	242
11.4.2	Face Detection, Tracking and Identification . . . . .	245
11.5	Biological Perspectives . . . . .	247
11.6	Discussion . . . . .	250
<b>12</b>	<b>Beyond Faces</b>	<b>253</b>
12.1	Multi-modal Identification . . . . .	254
12.2	Visually Mediated Interaction . . . . .	255
12.3	Visual Surveillance and Monitoring . . . . .	258
12.4	Immersive Virtual Reality . . . . .	260
12.5	Visual Database Screening . . . . .	261

## PART V APPENDICES 265

<b>A</b>	<b>Databases</b>	<b>267</b>
A.1	Database Acquisition and Design . . . . .	267
A.1.1	Intrinsic Experimental Variables . . . . .	268
A.1.2	Extrinsic Experimental Variables . . . . .	268
A.2	Acquisition of a Pose-labelled Database . . . . .	269
A.2.1	Using Artificial Markers . . . . .	269
A.2.2	Using Sensors and a Calibrated Camera . . . . .	271

A.3	Benchmarking . . . . .	273
A.4	Commercial Databases . . . . .	275
A.5	Public Domain Face Databases . . . . .	277
A.6	Discussion . . . . .	279
<b>B</b>	<b>Commercial Systems</b>	<b>281</b>
B.1	System Characterisation . . . . .	282
B.2	A View on the Industry . . . . .	283
B.2.1	Visionics Corporation . . . . .	284
B.2.2	Miros Inc. . . . .	286
B.2.3	VisionSpheres Technologies . . . . .	287
B.2.4	Eigenface-based Systems: Viisage Technologies, Intelli- gent Verification Systems and Facia Reco . . . . .	289
B.2.5	Systems based on Facial Feature Matching: Plettac Elec- tronic Security GmbH, ZN Bochum GmbH and Eyematic Interfaces Inc. . . . .	291
B.3	Discussion . . . . .	295
<b>C</b>	<b>Mathematical Details</b>	<b>297</b>
C.1	Principal Components Analysis . . . . .	297
C.2	Linear Discriminant Analysis . . . . .	298
C.3	Gaussian Mixture Estimation . . . . .	300
C.3.1	Expectation-maximisation . . . . .	300
C.3.2	Automatic Model Order Selection . . . . .	301
C.3.3	Adaptive EM for Non-stationary Distributions . . . . .	302
C.4	Kalman Filters . . . . .	304
C.4.1	Zero-order Prediction . . . . .	304
C.4.2	First-order Prediction . . . . .	305
C.5	Bayesian Belief Networks . . . . .	306
C.6	Hidden Markov Models . . . . .	309
C.7	Gabor Wavelets . . . . .	312
	Bibliography	315
	Index	339