

Dietrich Paulus
Joachim Hornegger

**Pattern Recognition
of Images and Speech
in C++**

To Dorothea, Belinda, and Dominik

In the text we use the following names which are protected, trademarks owned by a company or should be emphasized otherwise: GNU, HP, AT&T, Unix, PostScript \LaTeX , HTK, Entropic, Corel Draw.

Part I	Introductions	5
1	Pattern Recognition	7
1.1	Images and Sound	7
1.2	Applications of Pattern Recognition	9
1.3	Environment, Problem Domain, and Patterns	9
1.4	Characterization of Pattern Recognition	10
1.5	Speech Recording	12
1.6	Video Cameras and Projections	13
1.7	From Continuous to Digital Signals	16
1.8	Sampling Theorem in Practice	18
1.9	Visualization and Sound Generation	19
2	From C to C++	21
2.1	Syntax Notation	21
2.2	Principle of C++ Compilation	23
2.3	Function Calls and Arguments	24
2.4	Declaration and Definition of Variables	25
2.5	Unix-File Access via Standard Functions	27
2.6	Numeric Expressions	29
2.7	Main Program	30
2.8	Function Definition	31
2.9	Scope and Lifetime	33
3	Software Development	37
3.1	Software for Pattern Recognition	37
3.2	Software Development and Testing	38
3.3	Modular and Structured Programming	41
3.4	Comments and Program Layout	41
3.5	Documentation	42
3.6	Teamwork	44
3.7	Efficiency	44
3.8	Tools for Software Development	45
3.9	PUMA	46
4	Control and Data Structures	49
4.1	Structures	49
4.2	Enumerations	50
4.3	Scope Resolution	51
4.4	Unions	51
4.5	Bit- and Shift-Operations and Bit-Fields	52
4.6	Logical Values and Conditionals	54
4.7	Loops	55

4.8	Switches	56
4.9	Exception Handling	57
5	Arrays and Pointers	61
5.1	Vectors and Matrices	61
5.2	Pointers	63
5.3	Vectors vs. Pointers	64
5.4	Pointer Operations and Allocation	65
5.5	Pointer to Structures	66
5.6	Strings	68
5.7	Pointer and Array Arguments	68
5.8	Pointer to Pointer	69
5.9	Command Line Arguments	70
6	Classification and Pattern Analysis	73
6.1	Classification	73
6.2	Preprocessing	75
6.3	Feature Extraction	75
6.4	Analysis	76
6.5	Image Segmentation	76
6.6	Speech Segmentation	79
6.7	Pattern Understanding	80
6.8	Active Vision and Real Time Processing	81
6.9	Software Systems	82
7	C++ as a better C	85
7.1	Type Declaration	85
7.2	Type Conversion for Pointers	86
7.3	Type Specifiers and Variable Declaration	86
7.4	Type-Safe Linkage	88
7.5	Overloaded Function Names	90
7.6	Return Value and Arguments	90
7.7	Macros and Inline Functions	92
7.8	Function Pointers	93
7.9	Comma Operator and Conditional Expressions	94
8	Statistics for Pattern Recognition	97
8.1	Axioms	97
8.2	Discrete Random Variables	98
8.3	Continuous Random Variables	99
8.4	Mean and Variance	101
8.5	Moments of a Distribution	102
8.6	Random Vectors	103

CONTENTS

iii

8.7	Independence and Marginal Densities	105
8.8	Statistical Features and Entropy	106
8.9	Signal-to-Noise Ratio	107

Part II Object–Oriented Pattern Recognition	113
9 Object–Oriented Programming	115
9.1 Object–Oriented Software Techniques	115
9.2 Basic Concepts	117
9.3 Data Abstraction and Modules	118
9.4 Unified Modeling Language	119
9.5 Inheritance	120
9.6 Abstract Classes	122
9.7 Object–Oriented Classification	123
9.8 Polymorphism	124
9.9 C++ as an Object–Oriented Language	124
9.10 Class Libraries	125
10 Classes in C++	127
10.1 Methods and ADT’s	127
10.2 Class Declarations	129
10.3 Object Construction	130
10.4 Destruction of Objects	132
10.5 Operators	133
10.6 User–Defined Conversion	134
10.7 Advanced Methods and Constructors	135
10.8 Vector Class	136
10.9 Class Design	139
11 Representation of Signals	141
11.1 Array Class	141
11.2 Templates	143
11.3 Images	145
11.4 External Data Formats	147
11.5 Binary Images	148
11.6 Color Images	149
11.7 Subimages	151
11.8 Matrix Operations	153
11.9 Speech Signal Class	154
12 Fourier Transform	157
12.1 Introductory Considerations	158
12.2 Fourier Series	159
12.3 Fourier Transform	161
12.4 Discrete Fourier Transform	166
12.5 Complex Number Class	167
12.6 Inverse Discrete Fourier Transform	169

12.7	Fourier Transforms of Speech Signals	170
12.8	Fast Fourier Transform	171
12.9	2-D Fourier Transform	172
13	Inheritance for Classes	177
13.1	Motivation and Syntax	177
13.2	Access to Members of Base Class	179
13.3	Construction and Destruction	180
13.4	Pointer to Objects	181
13.5	Virtual Functions	183
13.6	Abstract Classes	184
13.7	Image Class Hierarchy	185
13.8	Multiple Inheritance	186
13.9	Implementation Issues	188
14	Edge Images	191
14.1	Strategies	191
14.2	Discrete Derivatives of Intensity Functions	193
14.3	Mask Operators	194
14.4	Discrete Directions	196
14.5	Edge Class	197
14.6	Edge Images	199
14.7	Robert's Cross	201
14.8	Second Derivative	201
14.9	Color Edge Operators	203
15	Class Libraries	207
15.1	Stream Input and Output	207
15.2	National Institutes of Health Class Library	209
15.3	Static Class Members	211
15.4	Input and Output for Objects	214
15.5	NIHCL Application Classes	216
15.6	NIHCL Collection Classes	217
15.7	Memory Allocation	219
15.8	Standard Template Library	222
15.9	Templates vs. Inheritance	223

Part III Object–Oriented Image Processing	225
16 Hierarchy of Picture Processing Objects	227
16.1 General Structure	227
16.2 HIPPOS Object	228
16.3 Images and Matrices	230
16.4 Chain Code Class	232
16.5 Edges	233
16.6 Polygon Representation	235
16.7 Atomic Objects	236
16.8 Segmentation Objects	238
16.9 External Representation	239
17 An Image Analysis System	245
17.1 Data Flow	245
17.2 Design of ANIMALS	246
17.3 Display and Capture	247
17.4 Geometric Distortions	249
17.5 Polymorphic Image Processing	250
17.6 Efficiency	252
17.7 Command Line Options	252
17.8 Graphical User Interfaces	254
17.9 Image Segmentation Program	254
18 Synthetic Signals and Images	259
18.1 Synthetic Sound	259
18.2 Geometric Patterns	260
18.3 Examples in C++	261
18.4 Pixel Noise	262
18.5 Gaussian Noise	262
18.6 Salt–and–Pepper Noise	263
18.7 2–D Views of 3–D Polyhedral Objects	264
18.8 Single Stereo Images	266
18.9 Textures	269
19 Filtering and Smoothing Signals	271
19.1 Linear Filters	272
19.2 Rank Order Operations	273
19.3 Edge Preserving Smoothing	275
19.4 K Nearest Neighbor Averaging	276
19.5 Conditional Average Filter	276
19.6 Linear Reconstruction	278
19.7 Elimination of Noisy Image Rows	278

19.8	Resolution Hierarchies	280
19.9	Image Operator Hierarchy	281
20	Histogram Algorithms	285
20.1	Discriminant Analysis Threshold	285
20.2	Histogram Entropy Thresholding	288
20.3	Multi-thresholding	289
20.4	Global Histogram Equalization	290
20.5	Local Histogram Equalization	293
20.6	Look up Table Transformation	293
20.7	Histogram Class	295
20.8	Color Quantization	296
20.9	Histogram Back-Projection	297
21	Edges and Lines	301
21.1	More Edge Detectors	301
21.2	Edge Thinning	304
21.3	Line Detection	309
21.4	Hysteresis Thresholds	312
21.5	Closing of Gaps	315
21.6	Zero-Crossings in Laplace-Images	316
21.7	Hough Transform	317
21.8	Circle Detection	319
21.9	Optimal Line Detection	323
22	Chain Codes	327
22.1	Smoothing	327
22.2	Digital Linear Lines	328
22.3	Neighborhood	330
22.4	Contours in Binary Images	331
22.5	Length, Area, and Similarity	332
22.6	Intersections	334
22.7	Rotation	334
22.8	Conversion	335
22.9	Corners of Chain Codes	336

Part IV	Speech and Pattern Analysis	339
23	Spatial and Spectral Features	341
23.1	Different Types of Features	341
23.2	Frames and Blocks	343
23.3	Spatial Features	345
23.4	Short Time Fourier Analysis and Spectral Features	348
23.5	Cepstral Features	353
23.6	Mel Spectral and Cepstral Features	354
23.7	Linear Predictive Coding	357
23.8	Model Spectrum and Cepstrum	360
23.9	Implementation Issues	361
24	Numerical Pattern Classification	365
24.1	General Notes on Classifiers	366
24.2	Design of Classifiers	367
24.3	Linear Discriminants	369
24.4	Polynomial Classifiers	371
24.5	Bayesian Classifiers	372
24.6	Properties of Bayesian Classifiers	373
24.7	From Bayesian to Geometric Classifiers	376
24.8	Nearest Neighbor Classifier	377
24.9	Implementation of Classifiers	379
25	Speech Recognition	383
25.1	Classification of Speech Signals	384
25.2	Dynamic Time Warping	385
25.3	Mixture Densities	391
25.4	Hidden Markov Models	397
25.5	Topological and Statistical Variations	400
25.6	Generalized Hidden Markov Models	402
25.7	Incomplete Data Estimation	402
25.8	Learning from Multiple Observations	405
25.9	An Object–Oriented Implementation of Hidden Markov Models	409
26	Advanced Topics	413
26.1	Segmentation into Lines and Arcs	413
26.2	Stereo Images	415
26.3	Region Segmentation	416
26.4	Active Vision	417
26.5	Semantic Networks	418
26.6	Statistical Object Recognition and Localization	418
26.7	Artificial Neural Networks	420

CONTENTS

ix

26.8 Advanced C++ Features	420
26.9 The Future of C++	421

Part V Appendix	423
A Software Development Tools	425
A.1 Groups and ID's with Unix	425
A.2 Program Building with make	426
A.3 The Use of Libraries	428
A.4 Version and Access Control with rcs	428
A.5 Teamwork	430
B Source Code and Tools	431
B.1 List of Tools	431
B.2 How to get the sources	432
B.3 X11	432
B.4 Slides	432
B.5 Addresses	433
B.6 Headers and Source Files	433
C Formulas	435
C.1 Lookup Table Transformation	435
C.2 Marginal Density	436
C.3 Identity	436
C.4 Property of the H -Function	436
C.5 Lagrange Multiplier	437
D Notation	441
Bibliography	443
Figures	457
Tables	463
List of Programs	465
Index	467