

Contents

1	Introduction	9
1.1	Contributions of the thesis	11
1.1.1	An architecture for visual recognition	11
1.1.2	Learning from user interaction	15
1.2	Organization of the thesis	16
2	Retrieval as statistical inference	18
2.1	Terms and notation	19
2.2	A Bayesian criteria for image similarity	22
2.3	A unified view of image similarity	25
2.3.1	Bhattacharyya distance	26
2.3.2	Maximum likelihood	28
2.3.3	Minimum discrimination information	28
2.3.4	χ^2 test	31
2.3.5	The Gaussian case	31
2.3.6	L^p norms	34
2.4	Experimental evaluation	37
3	Image representations for retrieval	40
3.1	Bayesian guidelines for image representation	41
3.1.1	Feature transformation	41
3.1.2	Feature representation	42
3.2	Strategies for image representation	44
3.2.1	The color strategy	44
3.2.2	The texture strategy	45

3.2.3	A critical analysis	46
3.3	An alternative strategy	50
4	Feature transformation	53
4.1	Terms and notation	54
4.2	Previous approaches	55
4.3	Minimizing the reconstruction error	56
4.4	Discrete Cosine Transform	57
4.5	Perceptual relevance	58
4.6	Experimental evaluation	62
5	Embedded multi-resolution mixture models	65
5.1	Spatially supported representations	66
5.1.1	Extensions to the Gaussian model	66
5.1.2	Extensions to the histogram model	67
5.1.3	Vector quantization	68
5.2	Mixture models	70
5.2.1	Some obvious relationships	71
5.2.2	Relationship to non-parametric models	71
5.2.3	Relationship to vector quantization	72
5.2.4	Relationship to histograms	74
5.2.5	A unified view of feature representations	74
5.3	Embedded multi-resolution mixture models	77
5.3.1	Invariance	79
5.4	Experimental evaluation	80
5.4.1	Embedded mixtures	81
5.4.2	Perceptual relevance	82
5.4.3	Invariance vs spatial support	84
5.5	Discussion	87
6	From local to global similarity	89
6.1	Local similarity	90
6.1.1	Previous solutions	91

6.1.2	Probabilistic retrieval	93
6.1.3	Experimental evaluation	94
6.2	The complexity of global similarity	97
6.2.1	Computational cost	97
6.2.2	Practical ways to reduce complexity	99
6.2.3	Asymptotic approximations for global similarity	100
6.2.4	Experimental evaluation	113
7	Short-term learning	118
7.1	Prior work	119
7.2	Bayesian relevance feedback	121
7.3	Target search	123
7.4	Negative feedback	125
7.5	Experimental evaluation	129
7.5.1	Experimental setup	129
7.5.2	Positive feedback	131
7.5.3	Negative feedback	135
8	Long-term learning	146
8.1	Probabilistic model	148
8.2	Incomplete queries	149
8.3	Combining different content modalities	150
8.4	Text representation	151
8.4.1	Parameter estimation	152
8.4.2	Long term learning	153
8.5	Experimental evaluation	156
9	The RaBI retrieval system	166
9.1	System overview	166
9.1.1	Client	167
9.1.2	SQL server	169
9.1.3	Image server	170
9.2	Implementation details	172

9.2.1	Feature transformation	172
9.2.2	Feature representation	173
9.2.3	Evaluation of image similarity	174
9.2.4	Vision modules	175
10	Conclusions	176
10.1	Contributions of the thesis	176
10.2	Directions for future work	177
A	Experimental setup	180
A.1	Databases	181
A.1.1	Standard databases	182
A.1.2	Artificial databases	183
A.2	Evaluation criteria	184
	Bibliography	186
	Acknowledgments	203