



Intelligent Big Multimedia Databases

Andreas Wichert

INESC-ID/ Instituto Superior Técnico -
Universidade de Lisboa, Portugal

 **World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Contents

<i>Preface</i>	vii
1. Introduction	1
1.1 Intelligent Multimedia Database	1
1.2 Motivation and Goals	5
1.3 Guide to the Reader	6
1.4 Content	7
2. Multimedia Databases	13
2.1 Relational Databases	13
2.1.1 Structured Query Language SQL	15
2.1.2 Symbolical artificial intelligence and relational databases	16
2.2 Media Data	19
2.2.1 Text	19
2.2.2 Graphics and digital images	21
2.2.3 Digital audio and video	23
2.2.4 SQL and multimedia	27
2.2.5 Multimedia extender	27
2.3 Content-Based Multimedia Retrieval	28
2.3.1 Semantic gap and metadata	31
3. Transform Functions	35
3.1 Fourier Transform	35
3.1.1 Continuous Fourier transform	35
3.1.2 Discrete Fourier transform	37

3.1.3	Fast Fourier transform	40
3.1.4	Discrete cosine transform	43
3.1.5	Two dimensional transform	45
3.2	Wavelet Transform	53
3.2.1	Short-term Fourier transform	53
3.2.2	Continuous wavelet transform	57
3.2.3	Discrete wavelet transform	61
3.2.4	Fast wavelet transform	70
3.2.5	Discrete wavelet transform and images	76
3.3	The Karhunen-Loève Transform	78
3.3.1	The covariance matrix	79
3.3.2	The Karhunen-Loève transform	83
3.3.3	Principal component analysis	84
3.4	Clustering	87
3.4.1	k -means	88
4.	Compression	91
4.1	Lossless Compression	91
4.1.1	Transform encoding	91
4.1.2	Lempel-Ziv	93
4.1.3	Statistical encoding	93
4.2	Lossy Compression	96
4.2.1	Digital images	96
4.2.2	Digital audio signal	99
4.2.3	Digital video	101
5.	Feature Extraction	105
5.1	Basic Image Features	105
5.1.1	Color histogram	105
5.1.2	Texture	108
5.1.3	Edge detection	109
5.1.4	Measurement of angle	111
5.1.5	Information and contour	112
5.2	Image Pyramid	113
5.2.1	Scale space	116
5.3	SIFT	116
5.4	GIST	123
5.5	Recognition by Components	123

5.6	Speech	124
5.6.1	Formant frequencies	125
5.6.2	Phonemes	125
5.7	Feature Vector	127
5.7.1	Contours	127
5.7.2	Norm	128
5.7.3	Distance function	128
5.7.4	Data scaling	129
5.7.5	Similarity	130
5.8	Time Series	131
5.8.1	Dynamic time warping	131
5.8.2	Dynamic programming	132
6.	Low Dimensional Indexing	133
6.1	Hierarchical Structures	133
6.1.1	Example of a taxonomy	133
6.1.2	Origins of hierarchical structures	134
6.2	Tree	138
6.2.1	Search tree	138
6.2.2	Decoupled search tree	139
6.2.3	B-tree	140
6.2.4	kd-tree	141
6.3	Metric Tree	147
6.3.1	R-tree	147
6.3.2	Construction	152
6.3.3	Variations	152
6.3.4	High-dimensional space	153
6.4	Space Filling Curves	156
6.4.1	Z-ordering	156
6.4.2	Hilbert curve	161
6.4.3	Fractals and the Hausdorff dimension	167
6.5	Conclusion	169
7.	Approximative Indexing	171
7.1	Curse of Dimensionality	171
7.2	Approximate Nearest Neighbor	173
7.3	Locality-Sensitive Hashing	173
7.3.1	Binary Locality-sensitive hashing	174

7.3.2	Projection-based LSH	176
7.3.3	Query complexity LSH	176
7.4	Johnson-Lindenstrauss Lemma	177
7.5	Product Quantization	178
7.6	Conclusion	180
8.	High Dimensional Indexing	181
8.1	Exact Search	181
8.2	GEMINI	182
8.2.1	1-Lipschitz property	183
8.2.2	Lower bounding approach	185
8.2.3	Projection operators	188
8.2.4	Projection onto one-dimensional subspace	189
8.2.5	l_p norm dependency	194
8.2.6	Limitations	197
8.3	Subspace Tree	198
8.3.1	Subspaces	198
8.3.2	Content-based image retrieval by image pyramid	200
8.3.3	The first principal component	203
8.3.4	Examples	205
8.3.5	Hierarchies	207
8.3.6	Tree isomorphy	208
8.3.7	Requirements	210
8.4	Conclusion	211
9.	Dealing with Text Databases	215
9.1	Boolean Queries	215
9.2	Tokenization	217
9.2.1	Low-level tokenization	217
9.2.2	High-level tokenization	218
9.3	Vector Model	218
9.3.1	Term frequency	218
9.3.2	Information	219
9.3.3	Vector representation	220
9.3.4	Random projection	220
9.4	Probabilistic Model	222
9.4.1	Probability theory	222
9.4.2	Bayes's rule	223

9.4.3	Joint distribution	224
9.4.4	Probability ranking principle	226
9.4.5	Binary independence model	226
9.4.6	Stochastic language models	230
9.5	Associative Memory	231
9.5.1	Learning and forgetting	232
9.5.2	Retrieval	233
9.5.3	Analysis	234
9.5.4	Implementation	235
9.6	Applications	236
9.6.1	Inverted index	236
9.6.2	Spell checker	237
10.	Statistical Supervised Machine Learning	239
10.1	Statistical Machine Learning	239
10.1.1	Supervised learning	239
10.1.2	Overfitting	240
10.2	Artificial Neuron	241
10.3	Perceptron	243
10.3.1	Gradient descent	245
10.3.2	Stochastic gradient descent	248
10.3.3	Continuous activation functions	248
10.4	Networks with Hidden Nonlinear Layers	249
10.4.1	Backpropagation	250
10.4.2	Radial basis function network	252
10.4.3	Why does a feed-forward networks with hidden nonlinear units work?	254
10.5	Cross-Validation	255
10.6	Support Vector Machine	256
10.6.1	Linear support vector machine	256
10.6.2	Soft margin	257
10.6.3	Kernel machine	257
10.7	Deep Learning	258
10.7.1	Map transformation cascade	258
10.7.2	Relation between deep learning and subspace tree	263
11.	Multimodal Fusion	269
11.1	Constrained Hierarchies	269

11.2	Early Fusion	270
11.3	Late Fusion	270
11.3.1	Multimodal fusion and images	271
11.3.2	Stochastic language model approach	271
11.3.3	Dempster-Shafer theory	272
12.	Software Architecture	275
12.1	Database Architecture	275
12.1.1	Client-server system	275
12.1.2	A peer-to-peer	276
12.2	Big Data	276
12.2.1	Divide and conquer	276
12.2.2	MapReduce	276
12.3	Evaluation	278
12.3.1	Precision and recall	279
13.	Multimedia Databases in Medicine	281
13.1	Medical Standards	281
13.1.1	Health Level Seven	281
13.1.2	DICOM	282
13.1.3	PACS	282
13.2	Electronic Health Record	282
13.2.1	Panoramix	283
13.3	Conclusion	289
	<i>Bibliography</i>	291
	<i>Index</i>	301