

Contents

1	Multidimensional Data and the Concept of Visualization	1
2	Strategies for Multidimensional Data Visualization	5
2.1	Direct Visualization	5
2.1.1	Geometric Methods	5
2.1.2	Iconographic Displays	13
2.1.3	Hierarchical Displays	14
2.2	Dimensionality Reduction	15
2.2.1	Proximity Measures	20
2.2.2	Principal Component Analysis	21
2.2.3	Linear Discriminant Analysis	26
2.2.4	Multidimensional Scaling	27
2.2.5	Manifold-Based Visualization	31
2.2.6	Isometric Feature Mapping	32
2.2.7	Locally Linear Embedding	35
2.3	Quantitative Criteria of Mapping	38
2.3.1	Spearman's Coefficient	39
2.3.2	König's Topology Preservation Measure	40
3	Optimization-Based Visualization	41
3.1	Formulation of Optimization Problems in Multidimensional Scaling	41
3.2	Differentiability Analysis of the Least Squares Stress Function	44
3.3	Optimization Algorithms for Scaling	47
3.4	Hybrid Evolutionary Algorithm for Multidimensional Scaling	51
3.5	Two-Level Optimization of Stress with City-Block Distances	54
3.5.1	Explicit Enumeration in Two-Level Optimization	61
3.5.2	Branch-and-Bound Algorithm for MDS	74
3.5.3	Combinatorial Evolutionary Algorithm	88
3.6	Impact of Used Distance Measure on Visualization	101
3.7	Impact of the Dimensionality of the Projection Space	106

4 Combining Multidimensional Scaling with Artificial Neural Networks	113
4.1 Feed-Forward Neural Networks in Visualization	113
4.1.1 Biological Neuron and Its Artificial Model	114
4.1.2 Artificial Neural Network Learning	116
4.1.3 Visualization by Means of Feed-Forward Neural Networks	122
4.2 Self-Organizing Map and Neural Gas	126
4.2.1 Principles of Self-Organizing Map	127
4.2.2 SOM Training	128
4.2.3 Neural Gas	133
4.2.4 Quality Measures of SOM and Neural Gas	133
4.2.5 Numbers of Winning Neurons in SOM and NG	135
4.2.6 SOM for Multidimensional Data Visualization	135
4.2.7 Comparative Analysis of SOM Software	139
4.3 Combinations of SOM or Neural Gas with MDS	145
4.3.1 Consecutive Combination	145
4.3.2 Integrated Combination	152
4.3.3 Comparative Analysis of Combinations	158
4.4 Curvilinear Component Analysis	168
4.5 The Feed-Forward Neural Network SAMANN	170
4.5.1 Control of the Learning Rate	174
4.5.2 Retraining of the SAMANN Network	175
5 Applications of Visualization	179
5.1 Visual Analysis of Social Data	179
5.1.1 Economic and Social Conditions of Countries	179
5.1.2 Qualitative Comparison of Schools	182
5.2 Multidimensional Data in Medicine and Pharmacology	187
5.2.1 Ophthalmological Data Analysis	187
5.2.2 Analysis of Heart Rate Oscillations with Respect to Characterization of Sleep Stages	190
5.2.3 Pharmacological Binding Affinity	192
5.3 Correlation-Based Visualization	198
5.3.1 Theoretical and Methodological Background	198
5.3.2 Dimensionality Problem in the Visualization of Correlation-Based Data	209
5.3.3 Environmental Data Analysis	210
5.3.4 Visual Analysis of Curricula	218
5.3.5 Analysis of Ophthalmological Features	221
Appendix A: Test Data Sets	227
References	235
Index	247