

Module 9: References

Lecture 41: Reading Material

[Prev topic](#)

[Next topic](#)

[Prev page](#)

[Next page](#)

The Lecture Contains:

 [Reading](#)

 [Bibliography](#)

[Prev topic](#)[Prev page](#)

Reading

- Mahalanobis distance [[Mahalanobis 1936](#)]
- Bhattacharyya distance [[Bhattacharyya 1943](#)]
- Match distance [[Werman et al. 1985](#)]
- Earth mover's distance [[Rubner et al. 2000](#), [Peleg et al. 1989](#), [Werman et al. 1985](#), [Ljosa et al. 2006a](#), [Ljosa et al. 2006b](#)]
- Dynamic hashing [[Larson 1978](#)]
- Extendible hashing [[Fagin et al. 1979](#)]
- Linear hashing [[Litwin 1980](#), [Larson 1988](#)]
- Space-filling curves [[Jagadish 1990a](#), [Moon et al. 2001](#)]
- Grid file [[Nievergelt et al. 1984](#)]
- K-d tree [[Bentley 1975](#), [Lee and Wong 1977](#)]
- Quadtree [[Finkel and Bentley 1974](#), [Lee and Wong 1977](#)]
- K-d-B tree [[Robinson 1981](#)]
- R-tree [[Guttman 1984](#)]
- Greene's R-tree [[Greene 1989](#)]
- R*-tree [[Beckmann et al. 1990](#)]
- R+-tree [[Sellis et al. 1987](#)]
- SS-tree [[White and Jain 1996b](#)]
- SR-tree [[Katayama and Satoh 1997](#)]
- P-tree [[Jagadish 1990b](#)]
- X-tree [[Berchtold et al. 1996](#)]
- VAMSplit R-tree [[White and Jain 1996a](#)]

- Curse of dimensionality[Weber et al. 1998, Berchtold et al. 1998, Beyer et al. 1999]
- Pyramid technique [Berchtold et al. 1998]
- VA-file [Weber et al. 1998]
- VA+-file [Ferhatosmanoglu et al. 2000]
- M-tree [Ciaccia et al. 1997]
- FastMap [Faloutsos and Lin 1995]
- Lipschitz embedding [Bourgain 1985, Johnson and Lindenstrauss 1984]
- LLR embedding [Linial et al. 1995]
- Johnson-Lindenstrauss lemma [Johnson and Lindenstrauss 1984]
- SparseMap [Hristescu and Farach-Colton 1999]
- Efficient embedding [Achlioptas 2001, Achlioptas 2003]
- Embedding of QFD [Bhattacharya et al. 2009a]
- Bounds on distortion [Bhattacharya et al. 2009a, Bhattacharya et al. 2009b]
- V-optimal histograms [Jagadish et al. 1998]
- Fagin's algorithm [Fagin 1996, Fagin 1999]
- Threshold algorithm [Fagin et al. 2001, Fagin et al. 2003]
- Incremental nearest neighbor [Hjalason and Samet 1999]
- Skyline queries [Börzsönyi et al. 2001]
- Block-nested-loop algorithm [Börzsönyi et al. 2001]
- Sort-filter-skyline algorithm [Chomicki et al. 2003]
- Skyline bitmap algorithm [Tan et al. 2001]
- Skyline nearest-neighbor algorithm [Kossmann et al. 2002]

Bibliography

- D. Achlioptas.
Database-friendly random projections.
In Proc. Symposium on Principles of Database Systems (PODS), pages 274{281,2001.
- D. Achlioptas.
Database-friendly random projections: Johnson-lindenstrauss with binary coins.
J. Computer and System Sciences, 66(4):671-687, 2003.

- N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger.
The r^* -tree: An efficient and robust access method for points and rectangles.
In Special Interest Group on Management of Data (SIGMOD), pages 322-331,1990.
- J. L. Bentley.
Multidimensional binary search trees used for associative searching.
Communications of the ACM, 18(9):509-517, 1975.
- S. Berchtold, D. Keim, and H. P. Kriegel.
The x-tree: An index structure for high-dimensional data.
In Very Large Data Bases Conf. (VLDB), pages 28-39, 1996.
- S. Berchtold, C. Böhm, and H.-P. Kriegel.
The pyramid-technique: Towards breaking the curse of dimensionality.
In SIGMOD, pages 142-153, 1998.
- K. S. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft.
When is "nearest neighbor" meaningful?
In ICDT '99: Proceeding of the 7th Int. Conf. on Database Theory, pages 217-235,1999.
- A. Bhattacharya, P. Kar, and M. Pal.
On low distortion embeddings of statistical distance measures into low dimensional spaces.
In Int. Conf. Database and Expert Systems Applications (DEXA), pages 164-172,2009.
- A. Bhattacharya, P. Kar, and M. Pal.
On low distortion embeddings of statistical distance measures into low dimensional spaces.
arXiv:0909.3169v1 [cs.CG], 2009.
- A. Bhattacharyya.
On a measure of divergence between two statistical populations defined by their probability distributions.
Bulletin of Calcutta Mathematical Society, 35:99-110, 1943.
- S. Börzsönyi, D. Kossmann, and K. Stocker.
The skyline operator.
In ICDE, pages 421-430, 2001.
- J. Bourgain.
On Lipschitz embedding of finite metric spaces in Hilbert space.
Israel Journal of Mathematics, 52(1-2):46-52, 1985.
- J. Chomicki, P. Godfrey, J. Gryz, and D. Liang.
Skyline with presorting.
In ICDE, pages 717-719, 2003.
- P. Ciaccia, M. Patella, and P. Zezula.
M-tree: An efficient access method for similarity search in metric spaces.
In Very Large Data Bases Conf. (VLDB), pages 426-435, 1997.
- R. Fagin, J. Nievergelt, N. Pippenger, and H. R. Strong.
Extendible hashing-a fast access method for dynamic files.

ACM Transactions on Database Systems, 4(3):315-344, 1979.

- R. Fagin, A. Lotem, and M. Naor.
Optimal aggregation algorithms for middleware.
In Symposium on Principles of Database Systems (PODS), pages 102-113, 2001.
- R. Fagin, A. Lotem, and M. Naor.
Optimal aggregation algorithms for middleware.
J. Computer and System Sciences, 66(4):614-656, 2003.
- R. Fagin.
Combining fuzzy information from multiple systems.
In Proc. of the Fifteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS), pages 216-226, 1996.
- R. Fagin.
Combining fuzzy information from multiple systems.
J. Computer and System Sciences, 58(1):83-99, 1999.
- C. Faloutsos and K.-I. Lin.
Fastmap: A fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets.
In Proc. of ACM Special Interest Group on Management of Data (SIGMOD), pages 163-174, 1995.
- H. Ferhatosmanoglu, E. Tuncel, D. Agrawal, and A. E. Abbadi.
Vector approximation based indexing for non-uniform high dimensional data sets.
In Int. Conf. Information and Knowledge Management (CIKM), pages 202-209, 2000.
- R. Finkel and J. L. Bentley.
Quad trees: A data structure for retrieval on composite keys.
Acta Informatica, 4(1):1-9, 1974.
- D. Greene.
An implementation and performance analysis of spatial data access methods.
In Proc. Fifth Int. Conf. on Data Engineering (ICDE), pages 606-615, 1989.
- A. Guttman.
R-trees: A dynamic index structure for spatial searching.
In Special Interest Group on Management of Data (SIGMOD), pages 47-57, 1984.
- G. R. Hjaltason and H. Samet.
Distance browsing in spatial databases.
ACM Trans. Database Syst., 24(2):265-318, 1999.
- G. Hristescu and M. Farach-Colton.
Cluster-preserving embedding of proteins.
Technical Report 99-50, Dept. of Computer Science, Rutgers University, 1999.
- H. V. Jagadish, N. Koudas, S. Muthukrishnan, V. Poosala, K. Sevcik, and T. Suel.
Optimal histograms with quality guarantees.

In Very Large Data Bases Conf. (VLDB), pages 275-286, 1998

- H. V. Jagadish.
Linear clustering of objects with multiple attributes.
In Proc. 1990 ACM SIGMOD Int. Conf. Management of Data (SIGMOD), pages 332-342, 1990.
- H. V. Jagadish.
Spatial search with polyhedra.
In Proc. 6th IEEE Int. Conf. Data Engineering (ICDE), pages 311-319, 1990.
- W. Johnson and J. Lindenstrauss.
Extensions of Lipschitz mappings into a Hilbert space.
Contemporary Mathematics, 26:189-206, 1984.
- N. Katayama and S. Satoh.
The sr-tree: An index structure for high-dimensional nearest neighbor queries.
In Proc. of the ACM SIGMOD Int. Conf. on Management of Data, pages 369-380, 1997.
- D. Kossmann, F. Ramsk, and S. Rost.
Shooting stars in the sky: an online algorithm for skyline queries.
In VLDB, pages 275-286, 2002.
- P.-Å. Larson.
Dynamic hashing.
BIT, 18(2):184-201, 1978.
- P.-Å. Larson.
Dynamic hash tables.
Communications of the ACM (CACM), 31(4):446-457, 1988.
- D. T. Lee and C. K. Wong.
Worst-case analysis for region and partial region searches in multidimensional binary search trees and balanced quad trees.
Acta Informatica, 9:23-29, 1977.
- N. Linial, E. London, and Y. Rabinovich.
The geometry of graphs and some of its algorithmic applications.
Combinatorica, 15:215-245, 1995.
- W. Litwin.
Linear hashing: A new tool for file and table addressing.
In Proc. 6th Conference on Very Large Databases (VLDB), pages 212-223, 1980.
- V. Ljosa, A. Bhattacharya, and A. K. Singh.
Indexing spatially sensitive distance measures using multi-resolution lower bounds.
In Int. Conf. on Extending Database Technology (EDBT), pages 865-883, 2006.
- V. Ljosa, A. Bhattacharya, and A. K. Singh.
LB-index: A multi-resolution index structure for images.
In Int. Conf. on Data Engineering (ICDE), pages 144-145, 2006.

- P. C. Mahalanobis.
On the generalised distance in statistics.
Proc. of the National Institute of Science of India, 2:49-55, 1936.
- B. Moon, H. V. Jagadish, C. Faloutsos, and J. H. Saltz.
Analysis of the clustering properties of the Hilbert space-filling curve.
IEEE Transactions on Knowledge and Data Engineering, 13(1):124-141, 2001.
- J. Nievergelt, H. Hinterberger, and K. C. Sevcik.
The grid file: An adaptable, symmetric multikey file structure.
ACM Trans. on Database Systems (TODS), 9(1):38-71, 1984.
- S. Peleg, M. Werman, and H. Rom.
A unified approach to the change of resolution: Space and gray-level.
IEEE Trans. on Pattern Analysis and Machine Intelligence, 11:739-742, 1989.
- J. T. Robinson.
The k-d-b-tree: A search structure for large multidimensional dynamic indexes.
In Proc. of the ACM SIGMOD Int. Conf. on Management of Data, pages 10-18, 1981.
- Y. Rubner, C. Tomasi, and L. J. Guibas.
The earth mover's distance as a metric for image retrieval.
Int. J. Computer Vision, 40(2):99-121, 2000.
- T. K. Sellis, N. Roussopoulos, and C. Faloutsos.
The r+-tree: A dynamic index for multi-dimensional objects.
In Proc. 13th Int. Conf. on Very Large Data Bases (VLDB), pages 507-518, 1987.
- K.-L. Tan, P.-K. Eng, and B. C. Ooi.
Efficient progressive skyline computation.
In VLDB, pages 301-310, 2001.
- R. Weber, H.-J. Schek, and S. Blott.
A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces.
In Very Large Data Bases Conf. (VLDB), pages 194-205, 1998.
- M. Werman, S. Peleg, and A. Rosenfeld.
A distance metric for multi-dimensional histograms.
Computer, Vision, Graphics, and Image Processing, 32(3):328-336, 1985.
- D. A. White and R. Jain.
Similarity indexing: Algorithms and performance.
In SPIE Storage and Retrieval for Image and Video Databases, pages 62-73, 1996.
- D. A. White and R. Jain.
Similarity indexing with the ss-tree.
In Proc. of the Twelfth Int. Conf. on Data Engineering, pages 516-523, 1996.

Thank You