

Bibliography

- [1] C. Aggarwal, A. Hinneburg and Daniel A. Keim. On the surprising behavior of distance metrics in high dimensional spaces. In *Proceedings of the 8th Int. Conf. on Database Theory*, pages 420–434, 2001.
- [2] C. Aggarwal, C. Procopiuc, J. Wolf, P. Yu and J. Park. Fast algorithms for projected clustering. In *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, pages 61–72, 1999.
- [3] Rakesh Agrawal, Christos Faloutsos and Arun N. Swami. Efficient similarity search in sequence databases. In *Proceedings of 4th conference on FODO*, pages 69–84, 1993.
- [4] Jiyuan An, Hanxiong Chen, Kazukata Furuse, Masahiro Ishikawa and Nobuo Ohbo. The convex polyhedra technique: An index structure for high-dimensional space. In *Proc. of the 13th Australasian Database Conference*, pages 33–40, 2002.
- [5] Jiyuan An, Hanxiong Chen, Kazukata Furuse, Masahiro Ishikawa and Nobuo Ohbo. A vector-wise dimensionality reduction for indexing high dimensional data. In *Proc.*

of Pan-Yellow-Sea International Workshop on Information Technologies for Network Era, pages 135–142, 2002.

- [6] M. Ankerst, M. M. Breunig, H-P. Kriegel, R. T. Ng and J.g Sander. Optics: Ordering points to identify the clustering structure. In *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, pages 49–60, 1999.
- [7] N. Beckmann, H. P. Kriegel, R. Schneider and B. Seeger. The R*-tree:an efficient and robust access method for points and rectangles. In *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*, pages 322–331, 1990.
- [8] S. Berchtold, C. Bohm, D. Keim and H.-P. Kriegel. The x-tree: An index structure for high-dimensional data. In *Proceedings of 26th International Conference on Very Large Data Bases*, pages 28–39, 1996.
- [9] S. Berchtold, C. Bohm, D. Keim and H.-P. Kriegel. A cost model for nearest neighbor search in high-dimensional data space. In *ACM PODS Symposium on Principles of Database Systems*, pages 78–86, 1997.
- [10] S. Berchtold, D. Keim and H. P. Kriegel. The pyramid-technique: Towards breaking the curse of dimensional data spaces. In *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*, pages 142–153, 1998.
- [11] K. S. Beyer, J. Goldstein, R. Ramakrishnan and U. Shaft. When is "nearest neighbor" meaningful. In *Proceedings of the 7th Int. Conf. on Database Theory*, pages 217–235, 1999.

- [12] T. Bially. Space-filling curves: Their generation and their application to bandwidth reduction. *IEEE Trans. on Information Theory*, Volume 15, Number 6, pages 658–664, 1969.
- [13] K. Chakrabarti and S. Mehrotra. The hybrid tree: an index structure for high dimensional feature spaces. In *ICDE'99*, pages 440–447, 1999.
- [14] K. Chakrabarti and S. Mehrotra. Locally dimensionality reduction: A new approach to indexing high dimensional spaces. In *Proceedings of 26th International Conference on Very Large Data Bases*, pages 151–162, 2000.
- [15] K. Chan and W. Fu. Efficient time series matching by wavelets. In *ICDE'99*, pages 126–133, 1999.
- [16] Hanxiong Chen, Jiyuan An, Kazukata Furuse and Nobuo Ohbo. C²VA: trim high dimensional indexes. In *Proc. WAIM2002*, pages 303–315, 2002.
- [17] C. Faloutsos, R. Barber and M. Flickner et. al. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, Volume 3, Number 1, pages 231–262, 1994.
- [18] C. Faloutsos and K. I. Lin. Fastmap: A fast algorithm for indexing, data mining and visualization of traditional and multimedia datasets. In *Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data*, pages 163–174, 1995.

- [19] C. Faloutsos, M. Ranfanathan and Y. Manolopoulos. Fast subsequence matching in time series databases. In *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, pages 419–429, 1994.
- [20] C. Faloutsos and S. Roseman. Fractals for secondary key retrieval. In *Proc. 8th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 247–252, 1989.
- [21] T. Feder and D. H. Greene. Optimal algorithms for approximate clustering. In *Proc. 20th Annual ACM Symposium on the Theory of Computing*, pages 434–444, 1988.
- [22] K. Fukunaga. *Statistical Pattern Recognition*. Academic Press, 1990.
- [23] A. Guttman. R-tree:a dynamic index structure for spatial searching. In *Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data*, pages 47–57, 1984.
- [24] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, 1986.
- [25] I. Kamel and C. Faloutsos. On packing r-trees. In *Proc. 2nd. International Conference on Information and Knowledge Management*, pages 490–499, 1993.
- [26] N. Katayama and S. Satoh. The sr-tree: An index structure for high-dimensional nearest neighbor queries. In *Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data*, pages 369–380, 1997.

- [27] Eamonn J. Keogh, K. Chakrabarti, S. Mehrotra and Michael J. Pazzani. Locally adaptive dimensionality reduction for indexing large time series databases. In *Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data*, pages 151–162, 2001.
- [28] Eamonn J. Keogh, K. Chakrabarti, Michael J. Pazzani and S. Mehrotra. Dimensionality reduction for fast similarity search in large time series databases. In *Knowledge and Information Systems Journal*, pages 263–286, 2001.
- [29] Eamonn J. Keogh and Michael J. Pazzani. A simple dimensionality reduction technique for fast similarity search in large time series databases. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'00)*, pages 122–133, 2000.
- [30] G. Moody. Mit-bih database distribution [<http://ecg.mit.edu/index.html>]. In *Cambridge, MA.*, 2000.
- [31] J. Nievergelt, H. Hinterberger and K. C. Sevcik. The grid file: An adaptable, symmetric multikey file structure. *journal of ACM Transactions on Database Systems(TODS)*, Volume 9, Number 1, pages 38–71, 1984.
- [32] J. T. Robinson. The k-d-b-tree:a search structure for large multidimensional dynamic indexes. In *Proceedings of the 1981 ACM SIGMOD International Conference on Management of Data*, pages 10–18, 1981.

- [33] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, Volume 290, Number 1, pages 2323–2326, 2000.
- [34] T. Seidl and H. Kriegel. Optimal multi-step k-nearest neighbor queries. In *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*, pages 154–165, 1998.
- [35] T. Shinohara, J. An and H. Ishizaka. Approximate retrieval of high-dimensional data with l_1 metric by spatial indexing. *New Generation Computing*, Volume 18, Number 1, pages 39–47, 2000.
- [36] J. B. Tenenbaum, V. de. Silva and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, Volume 290, Number 1, pages 2319–2323, 2000.
- [37] W. S. Torgerson. *Theory and Methods of Scaling*. New York:Wiley, 1958.
- [38] R. Weber, H. J. Schek and S. Blott. A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces. In *Proceedings of 24th International Conference on Very Large Data Bases*, pages 194–205, 1998.
- [39] Byoung-Kee Yi and Christos Faloutsos. Fast time sequence indexing for arbitrary l_p norms. In *Proceedings of 26th International Conference on Very Large Data Bases*, pages 385–394, 2000.

- [40] T. Zhang, R. Ramakrishnan and M. Livny. Birch:an effective data clustering method for very large databases. In *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data*, pages 103–114, 1996.
- [41] G. Zipf. *'Human Behavior and Principle of Least Effort: An introduction to Human Ecology*. Addison Wesley, Cambridge, Massachusetts, 1949.