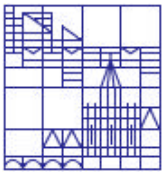
Advanced Technology Seminar
Similarity Search in Multimedia Databases

Daniel A. Keim and Benjamin Bustos

Databases, Data Mining, and Visualization
University of Konstanz, Germany



Universität Konstanz

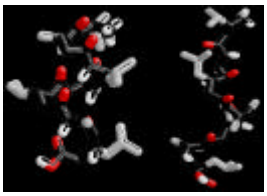
E-mail: {keim|bustos}@informatik.uni-konstanz.de
http://dbvis.inf.uni-konstanz.de/

Overview

1. Introduction
2. Efficiency
3. Effectiveness
4. Applications
5. Future research

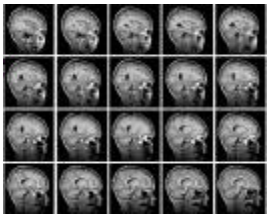
Introduction

- Many application domains

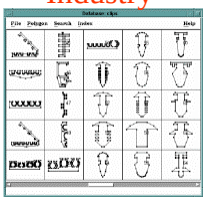


Molecular Biology

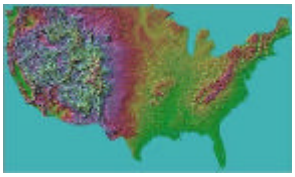
Medicine



Manufacturing Industry



Geography



And many others…

Introduction

- Multimedia data: Heterogeneous!

Image

Audio & video

Text

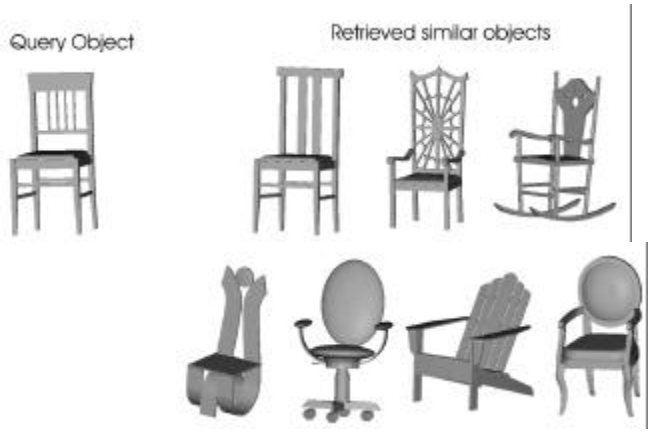
Introduction

- Content-based retrieval in multimedia databases [YI99]
  - Two approaches for retrieval in multimedia databases:
    - *Object Annotation (Meta Information)*: Describes the content of the multimedia object
    - *The object itself*: Representation is the multimedia object itself
  - Exact search is not meaningful

→ Similarity Search!

Introduction

- Example of content-based retrieval

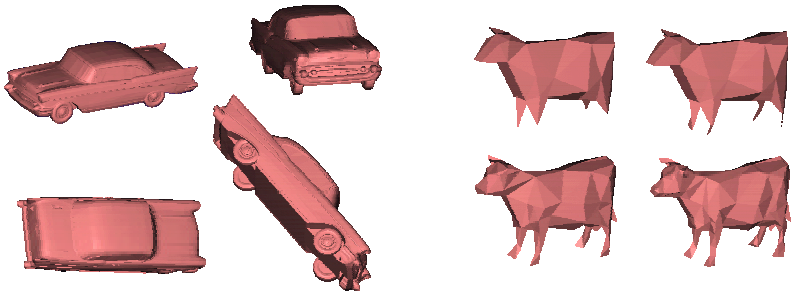


Introduction

- Content-based retrieval in multimedia databases is a difficult problem!

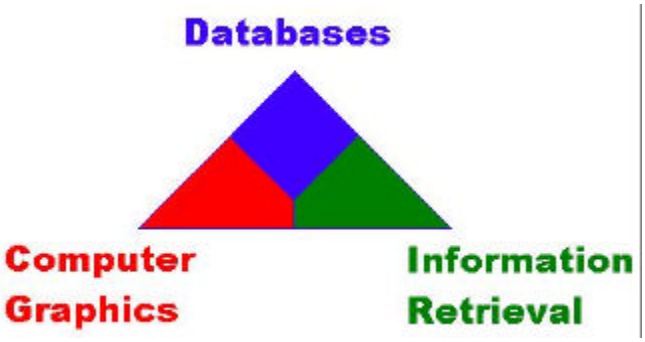
Orientation

Level-of-detail



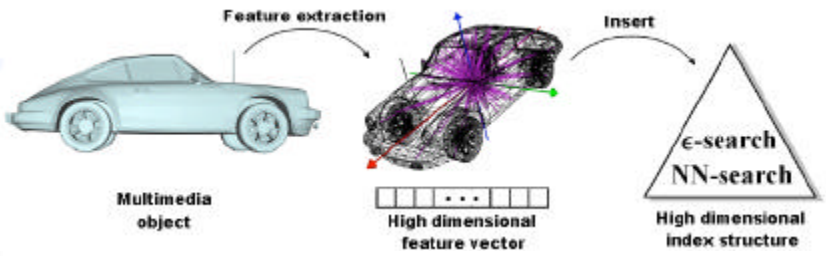
Introduction

- Multimedia databases: Involves different areas in Computer Science



Introduction

- Basic Approach to Similarity Search



Introduction

- Modeling multimedia data
  - Metric space [CNB+01]
  - Vector space [BBK01]
- Nomenclature
  - X: Universe of valid objects
  - $U \subset X$: Database
  - $n \in \mathbb{N}$: Size of U
  - $d(x, y)$: Distance function
  - $q \in X$: Query object

Introduction

- Modeling multimedia data: Metric space
  - Measure of distance between objects

$$d : X \times X \rightarrow \mathbb{R}^{+}$$

  - Properties of a metric:
    1. Positiveness:
       $\forall x, y \in X, d(x,y) \geq 0 \wedge d(x,y) = 0 \Leftrightarrow x = y$
    2. Symmetry: $\forall x, y \in X, d(x,y) = d(y,x)$
    3. Triangle inequality:
       $\forall x, y, z \in X, d(x,z) \leq d(x,y) + d(y,z)$

Introduction

- Distance functions: Minkowski

$$l_s(\vec{x}, \vec{y}) = \left(\sum_{i=1}^{t} |x_i - y_i|^s\right)^{1/s}, s \geq 1$$

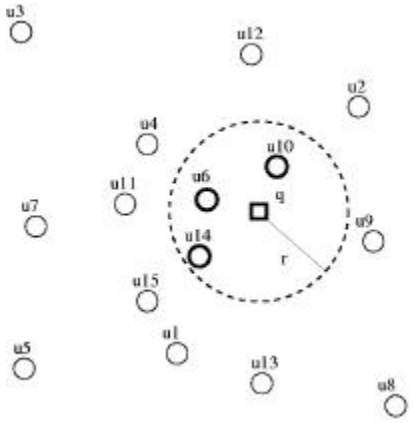
- $l_1(\vec{x}, \vec{y}) = \sum_{i=1}^{t} |x_i - y_i|$ "Manhattan"
- $l_2(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^{t} |x_i - y_i|^2}$ "Euclidean"
- $l_\infty(\vec{x}, \vec{y}) = \max_{i=1}^{t}\{|x_i - y_i|\}$ "Maximum"

- Weighted Minkowski, Mahalanobis, etc.

Introduction

- Similarity queries: Range query

$$(q, r)_d = \{u \in U | d(q, u) \leq r\}$$



Introduction

- Similarity queries:
  k-Nearest Neighbor Query
  – Returns an answer set *C* such that

$$|C| = k$$

and

$$\forall x \in C, y \in U - C, d(x, q) \leq d(y, q)$$

Introduction

- Multimedia Content Descriptor Interface (MPEG-7)
  - MPEG-7 is a standard that describes multimedia content data

Introduction

- Main elements of MPEG-7 standard
  - Description tools
    - Descriptors
    - Description schemes
  - Description definition language (DDL)
  - System tools
    - Text format (searching and editing)
    - Binary format (efficient storage and transmission)
- URL: http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

Overview

1. Introduction
2. Efficiency
3. Effectiveness
4. Applications
5. Future research

Overview

1. Introduction
2. Efficiency
   i. Efficiency considerations
   ii. Spatial access methods
   iii. Metric indices
   iv. Approximate and probabilistic approaches
3. Effectiveness
4. Applications
5. Future research

Efficiency Considerations

- Effects in high-dimensional spaces [BBK01]
  - Exponential dependency of measures on the dimension
  - Boundary effects
  - No geometric imagination

  ⇒ Intuition fails

"Curse of dimensionality"

Efficiency Considerations

- Notations and assumptions
  - *D* dimensions
  - Size of the database = *N*
  - Data space normalized to $[0,1]^D$
  - Uniformly distributed data

Efficiency Considerations

- Exponential growth of volume
  - Hypercube

$$Vol_{cube}(edge, D) = edge^D$$

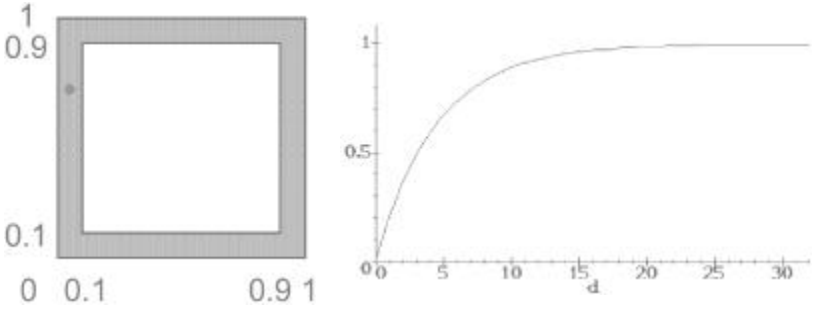
$$Diag_{cube}(edge, D) = edge \cdot \sqrt{D}$$

  - Hypersphere

$$Vol_{sphere}(radius, D) = radius^D \cdot \frac{\sqrt{\pi^D}}{\Gamma\left(\frac{D}{2}+1\right)}$$

Efficiency Considerations

- The surface is everything!
  - Probability that a point is closer to 0.1 to a (*D*-1)-dimensional surface

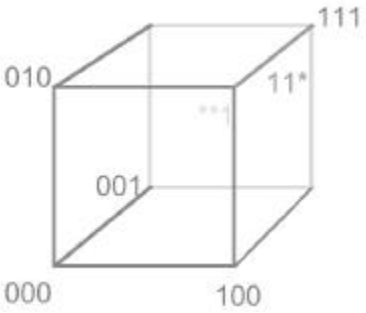


Efficiency Considerations

- Number of surfaces
  - How many *k*-dimensional surfaces has a *D*-dimensional hypercube $[0..1]^D$?

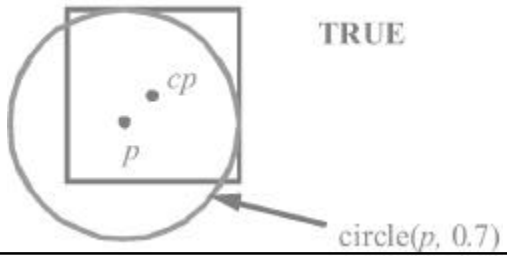
$$\binom{d}{k} \cdot 2^{(d-k)}$$

***



Efficiency Considerations

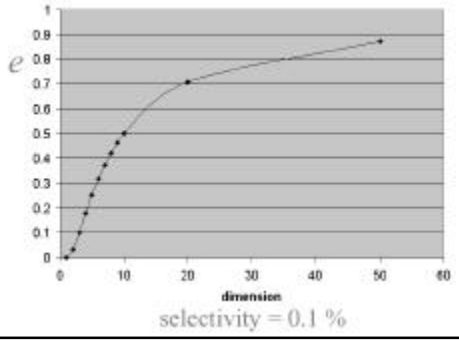
- "Each circle touching all boundaries includes the center point" ➡ False!
  - *D*-dimensional cube $[0,1]^D$
  - cp=(0.5, 0.5, ..., 0.5), p=(0.3, 0.3, ..., 0.3)
  - 16-D: circle (p, 0.7), distance (p, cp)=0.8!!!



Efficiency Considerations

- Database specific effects
  - Selectivity of range queries: Depends on the volume of the query

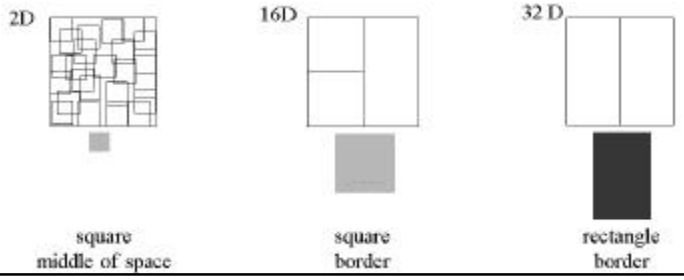
$$e = \sqrt[d]{Vol_{cube}}$$



selectivity = 0.1 %

Efficiency Considerations

- Database specific effects
  - Data pages have large extensions
  - Most of data pages touch the surface of the data space on most sides

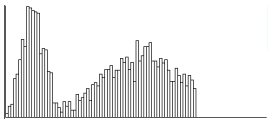


Efficiency Considerations

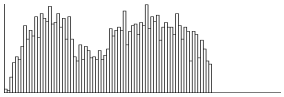
- How to express useful queries in high-dimensional spaces?
  - Histograms describing some statistical properties
    - Medium - very high dimensionality (20-1000)
    - Meaningful queries are ***difficult*** to express
  - Observations
    - Not all dimensions are equally relevant for a given query
    - Multiple meaningful NNs exist for different search metrics

Efficiency Considerations

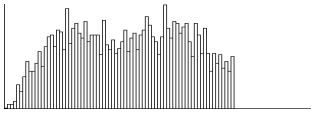
- How do meaningful distance distributions look like?



All 10 dimensions are relevant



9 of 10 dimensions are relevant



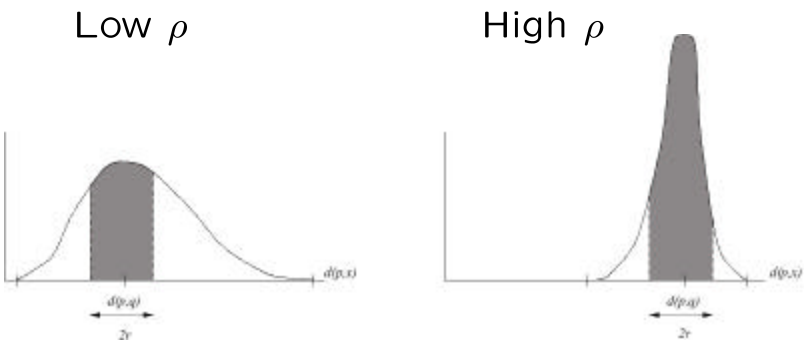
8 of 10 dimensions are relevant

Efficiency Considerations

- Effects in metric spaces [CNB+01]

  Intrinsic dimension $\rho = 2\mu^2/\sigma^2$

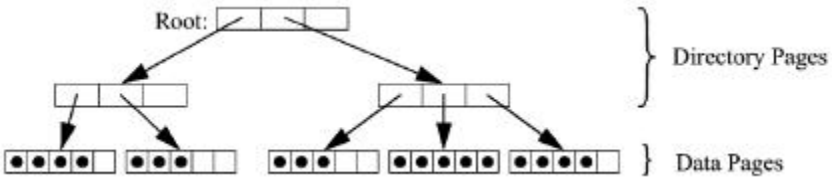
Low ρ High ρ



Efficiency: Spatial access methods

- High-dimensional indexing methods [BBK01]

Hierarchical index structures



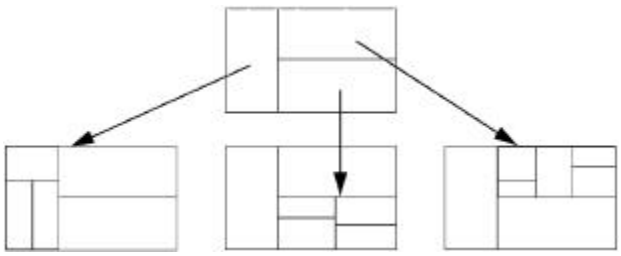
Efficiency: Spatial access methods

- Minimum bounding rectangles
  - kd-tree directory
    - kd-B-tree [Rob81]
    - LSD^h-tree [Hen98]
  - R-tree variations
    - R-tree [Gut84]
    - R$^+$-tree [SRF87]
    - R*-tree [BKS+90]
    - X-tree [BKK96]

Efficiency: Spatial access methods

- kd-B-tree [Rob81]

•Hyperrectangle-shaped page regions
•An adaptive kd-tree is used for space partitioning
•Complete and disjoint partitioning



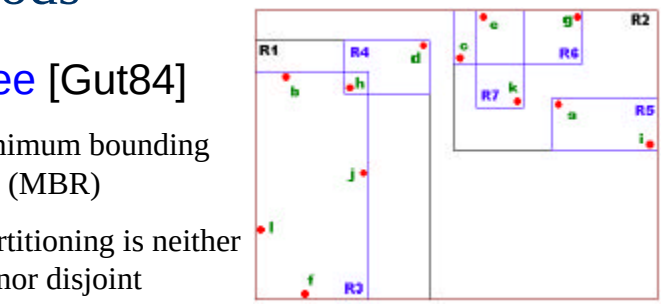
Efficiency: Spatial access methods

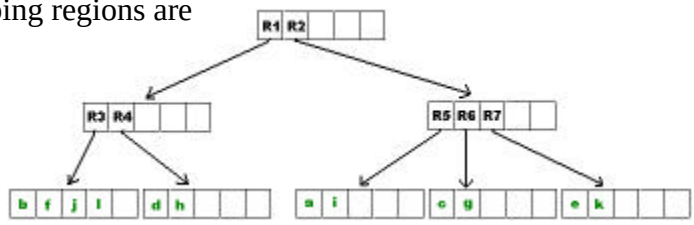
- R-tree [Gut84]

•Solid minimum bounding rectangles (MBR)

•Space partitioning is neither complete nor disjoint

•Overlapping regions are allowed



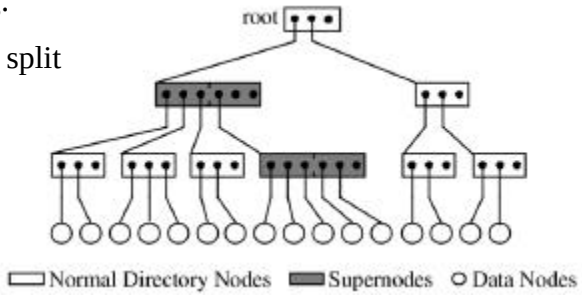


Efficiency: Spatial access methods

- X-tree [BKK96]

•Avoids overlap in the directory by using:

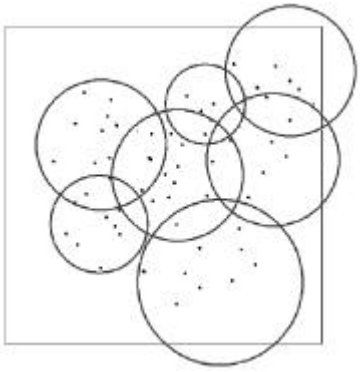
- Overlap-free split
- Supernodes



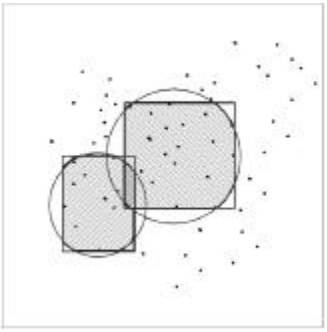
Efficiency: Spatial access methods

- Bounding spheres and combined regions

SS-tree [WJ96b]



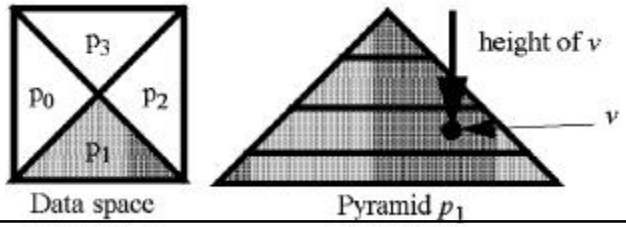
SR-tree [KS97]



Efficiency: Spatial access methods

- Other structures
  - TV-tree [LJF94]
  - Space filling curves [Sag94]
  - Pyramid technique [BBK98]

Example: Pyramid technique

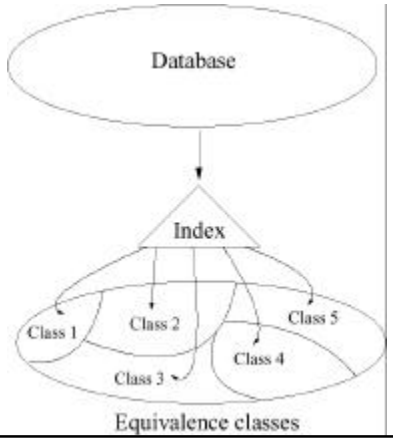


Efficiency: Spatial access methods

- GEMINI: Generic Multimedia object INdexING [Fal96]
  1. Determine distance function *D* between two objects
  2. Find numerical feature-extraction functions
  3. Prove that distance in feature space is a lower-bound of *D*
  4. Use an index to store and retrieve feature vectors

Efficiency: Metric indices

- Indexing metric spaces [CNB+01]



Querying:
•Traverse index and discard classes (internal complexity)
•Search in candidate classes (external complexity)

Efficiency: Metric indices

- Complexity of the search
  - Usually measured as the *number of distance computations*
  - Other costs (I/O, CPU) are neglected
- Two main indexing approaches
  - Pivot-based indexing
  - Indexing based on compact partitions

Efficiency: Metric indices

- Pivot-based indexing

•Set of *k* pivots
$\{p_1, \ldots, p_k\}, \ p_i \in U$

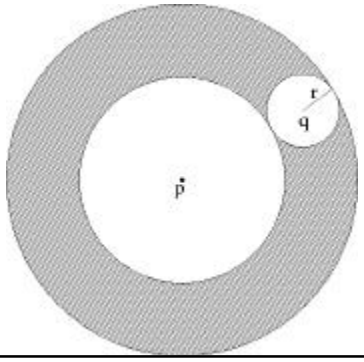
•Distance lower bound

$$d(q,x) \geq |d(p_i,x) - d(p_i,q)|$$

•Exclusion condition for (q,r)

$$|d(p_i,x) - d(p_i,q)| > r$$

Example using 1 pivot



Efficiency: Metric indices

- Metric trees based on pivots
  - Burkhard-Keller Tree [BK73]
  - Vantage Point Tree [Yia93]
  - Fixed Queries Tree [BCM+94]
  - Fixed-Height Queries Tree [BCM+94]
  - Multi Vantage Point Tree [BO97]
- Array representations of trees
  - Spaghettis [CMB99]
  - Fixed Queries Array [CMN01]

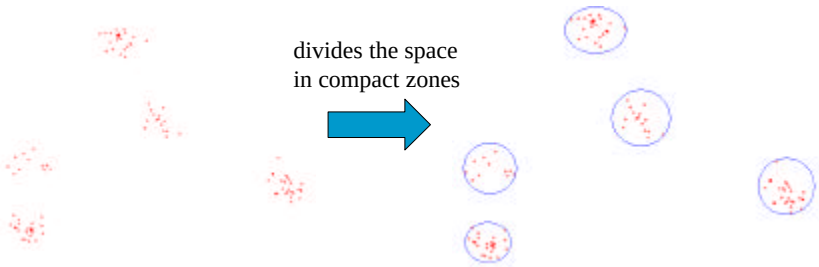
Efficiency: Metric indices

- Other structures
  - Approximating and Eliminating Search Algorithm (AESA) [Vid86]
  - Linear AESA [MOV94]
- Pivot selection techniques [BNC03]
  - Random selection
  - Maximize mean distribution of

$$D_{p_1,\ldots,p_k}(x,y) = \max_{i=1}^{k} |d(p_i,x) - d(p_i,y)|$$

Efficiency: Metric indices

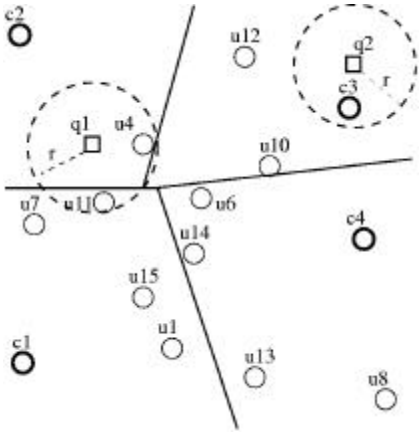
- Indexing based on compact partitions



- Criteria for partitioning the space
  - Hyperplane partition
  - Covering radius

Efficiency: Metric indices

- Hyperplane criterion



Search algorithm for *(q,r)*:

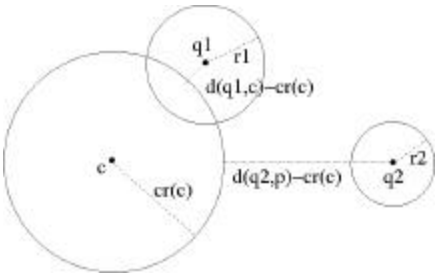
- •Compute distances between centers and *q*
- •Let *c* be the closest center to *q*
- •*Exclusion condition*:

$$d(q, c_i) > d(q, c) + 2r$$

- •For *q1*, the algorithm discards the zone of *c4*
- •For *q2*, the algorithm discards the zones of *c1* and *c2*

Efficiency: Metric indices

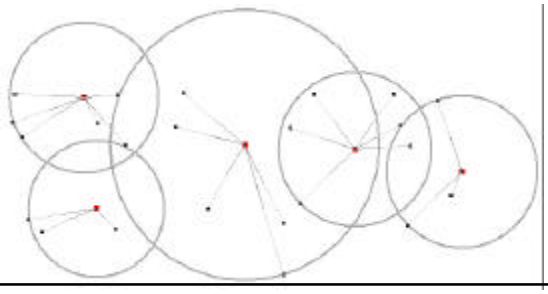
- Covering radius criterion
  - *Covering radius*: Maximum distance from a center to an object from its zone.
  - Exclusion criterion: $d(q, c) - cr(c) > r$



Example: For *q1*, the zone of *c* cannot be discarded, but for *q2* it is discarded

Efficiency: Metric indices

- M-tree [CPZ97]
  - Based on the covering radius criterion
  - Good I/O performance and few distances computations



Efficiency: Metric indices

- Hyperplane criterion
  - Generalized-Hyperplane Tree [Uhl91]
- Covering radius criterion
  - Bisector Tree (BST) [KM83]
  - Voronoi Tree [DN87]
  - Monotonous BST [NVZ92]
  - M-Tree [CPZ97]
  - List of Clusters [CN00]
- Mixed criteria
  - Geometric Near-neighbor Access Tree [Bri95]
  - Spatial Approximation Tree [Nav02]

Efficiency: Approximate and probabilistic approaches

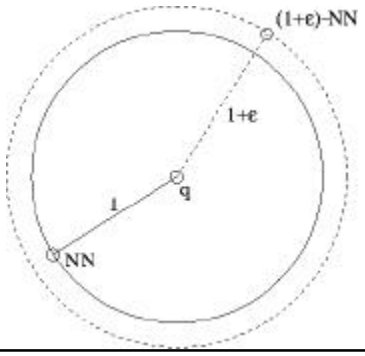
- Approximate and probabilistic approaches

–Trade off between performance efficiency and quality of the approximation

–*(1+e)*-approximate NN: Distance is within a factor *(1+e)* of the distance to the true NN

–Time-bounded search: Retrieve similar objects in a fixed amount of time

Approximately correct NN



Efficiency: Approximate and probabilistic approaches

- Classification schema [CP01]

• Data type:
  -Metric spaces
  -Vector spaces
•Error metrics:
  -Changing space
  -Reducing comparisons

•Quality guarantees:
  -Deterministic
  -Probabilistic (parametric and non-parametric)
•User interaction:
  -Static
  -Interactive

Efficiency: Approximate and probabilistic approaches

- Approaches for vector spaces
  - Approximate range search [AM95]
  - Algorithms and strategies for similarity retrieval [WJ96a]
  - Optimal approximate NN search [AMN+98]
  - Limited radius NN-search [Yia00]
  - Approximate similarity queries [CP01]

Efficiency: Approximate and probabilistic approaches

- Approaches for metric spaces
  - Approximate *k*-NN queries [ZSA+98]
  - Approximate NN-search [Cla99]
  - Probabilistic Approximately Correct (PAC) NN-search [CP00]
  - Probabilistic pivot-based range search [CN03]
  - Probabilistic algorithms based on compact partitions [BN04]

Overview

1. Introduction
2. Efficiency
3. Effectiveness
   i. Effectiveness measures
   ii. User-oriented measures
   iii. Reference collections
4. Applications
5. Future research

Effectiveness: Effectiveness measures

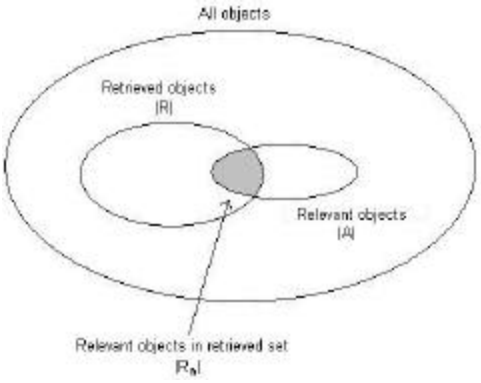
- Retrieval performance evaluation: *Effectiveness measures*
- Ground truth: *Test reference collection*
- Evaluation measure: Quantifies *similarity* between retrieved objects and relevant objects

Effectiveness: Effectiveness measures

- Precision and recall [BR99]

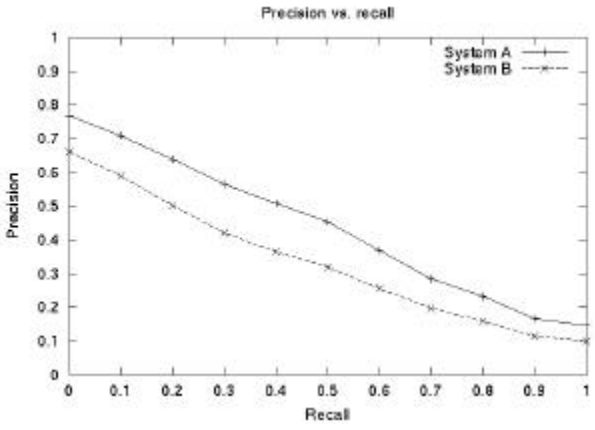
$$Precision = \frac{|R_a|}{|A|}$$

$$Recall = \frac{|R_a|}{|R|}$$



Effectiveness: Effectiveness measures

- Precision vs. recall figure



System A is more effective than system B

Effectiveness: Effectiveness measures

- Single values summaries
  - R-precision (first tier) [BR99]: Precision computed when

$$|R| = |A|$$

  - Bull-Eye Percentage (second tier) [ZP01]: Recall computed when

$$|R| = 2|A|$$

Effectiveness: Effectiveness measures

- Alternative measures
  - Harmonic mean [SBP97]; E measure [Rij79]

$$F(j) = \frac{2}{\frac{1}{r(j)} + \frac{1}{P(j)}}$$

$$E(j) = 1 - \frac{1+b^2}{\frac{b^2}{r(j)} + \frac{1}{P(j)}}$$

$r(j)$: recall j^{th} object in the ranking

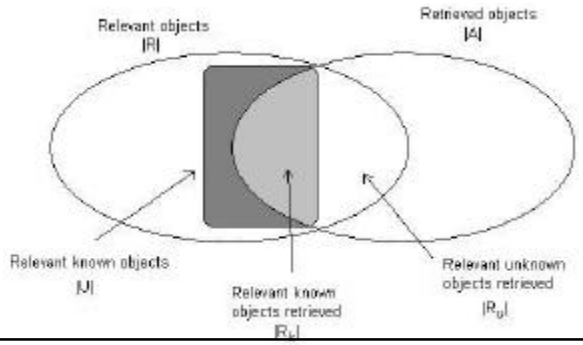
$P(j)$: precision j^{th} object in the ranking

b: user specified parameter

Effectiveness: User-oriented measures

- Coverage and novelty [Kor97,BR99]

$$Coverage = \frac{|R_k|}{|U|} \qquad Novelty = \frac{|R_u|}{|R_u| + |R_k|}$$



Effectiveness: User-oriented measures

- Relative recall, relative effort [Kor97]

$$Relative\ recall = \frac{\#\ \text{relevant objects retrieved}}{\#\ \text{relevant objects desired}}$$

$$Relative\ effort = \frac{\#\ \text{relevant objects desired}}{\#\ \text{objects examined to find them}}$$

Effectiveness: User-oriented measures

- Satisfaction and frustration [Kor97]
  - Objects judged on a 5-point scale
  - {0,1}: non-relevant; {2,3,4}: relevant

$Retrieved = \{3, 0, 4, 2, 1\}$

$Ideal = \{4, 3, 2, 1, 0\}$

(a) $Satisfaction = \{3, 3, 7, 9, 9\}$

(b) $Frustration = \{0, 2, 2, 2, 3\}$

$Total = \alpha Satisfaction - \beta Frustration$

$\alpha = 1 \wedge \beta = 1 \Rightarrow \{3, 1, 5, 7, 6\}$

$Total\ ideal = \{4, 7, 9, 8, 6\}$

Effectiveness: Reference collections

- Reference collection
  - "A collection of documents used for testing IR models and algorithms" [BR99]
  - Usually includes:
    - Set of objects
    - Set of queries
    - Set of objects known to be relevant to each query

Effectiveness: Reference collections

- TREC collection
  - Text REtrieval Conference, started in 1992.
  - TREC document collection
    - Several Gigabytes of data
    - Documents come from diverse sources
    - Set of relevant documents obtained via *pooling method*.
  - URL: http://trec.nist.gov/

Effectiveness: Reference collections

- Cystic Fibrosis Database [SWW+91]
  - 1,239 documents published from 1974 to 1979 discussing Cystic Fibrosis Aspects
  - A set of 100 queries with the respective relevant documents as answers
  - Set of relevance scores generated by experts (0 to 8 points)
  - URL: http://www.sims.berkeley.edu/~hearst/irbook/cfc.html

Effectiveness: Reference collections

- Princeton Shape Benchmark [SMK+04]
  - Database and tools for 3D objects retrieval
  - 1,814 3D models:
    - Base training classification, 90 classes, 907 models
    - Base test classification, 92 classes, 907 models
  - URL: http://shape.cs.princeton.edu/benchmark/index.cgi

Overview

1. Introduction
2. Efficiency
3. Effectiveness
4. Applications
   i. Text
   ii. Images
   iii. Computer Aided Design (CAD)
   iv. 3D objects
   v. Audio
   vi. Video
5. Future research

Applications: Text

- Text retrieval
  - Document: Paragraph, chapter, web page, book…
  - *Term*: word whose semantics defines the main theme of a document
  - Goal: Search in unstructured documents

Google http://www.google.com

Applications: Text

- Vector model for documents [BR99]
  - Term *i* is associated with a positive weight
  - *t* total number of terms ⟹ *t* features
  - Similarity between two documents
    - *Cosine similarity*:

$$sim(\vec{d_1}, \vec{d_2}) = \frac{\vec{d_1} \cdot \vec{d_2}}{||\vec{d_1}||_2 \cdot ||\vec{d_2}||_2}$$

    - Metric [FL95]

$$dist(\vec{d_1}, \vec{d_2}) = \sqrt{2 \cdot (1 - sim(\vec{d_1}, \vec{d_2}))}$$

Applications: Text

- Vector model for documents (cont.)
  - Term weights: *tf-idf* schema

$$w_{ij} = \frac{freq_{i,j}}{\max_{k=1}^{t}(freq_{k,j})} \cdot \log\left(\frac{N}{n_i}\right)$$

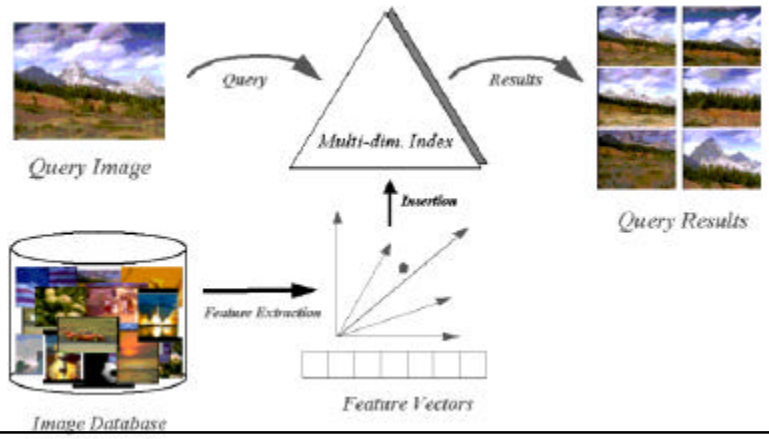
- N: Number of documents
- n_i: # of docs. where term i appears
- $freq_{i,j}$: Frequency of i^{th} term in doc. j

Applications: Text

- Approximate string matching [Nav01, NR02]
  - Given a word, retrieve all words close to it
  - Metric function: *Edit distance*
  - Applications:
    - OCR errors
    - Correcting misspelled words
    - Search of DNA sequences

Applications: Images

- Similarity search in images databases



Applications: Images

- Similarity search in images databases
  - Goal:
    - Content-based similarity search in large image DB
    - Improved recall without explicit object recognition
  - General approach:
    - Feature transformation to extract compact feature vectors
    - Post-processing in feature space (e.g., clustering of feature vectors)
    - Search on feature vectors using index structure
    - Support of different similarity measures

Applications: Images

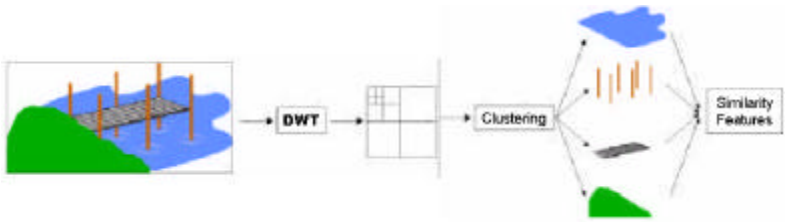
- Feature extraction
  - Color histograms [SK97]
  - Contour descriptors [Jag91,AKS98]
  - Texture similarity [VB98]
  - Color similarity matching [LTO+01]
  - Multiresolution similarity search [HHK+02]
- Retrieval systems
  - QBIC [AFH+95]

Applications: Images

- WALRUS [NRS04]
  - Invariance w.r.t. translation and scaling of regions in image
  - Approach:
    - Haar Wavelet Transform of sliding window of varying size
    - Clustering of signatures in wavelet space (BIRCH) => variable number of signatures per image
    - Storage of centroids of clusters in index structure (R*-tree)
    - Similarity search: Matching pairs of signatures (largest overlap)

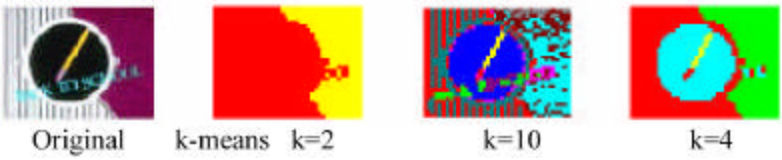
Applications: Images

- Windsurf [ABP99]
  - Wavelet-based INDexing of imageS Using Region Fragmentation
  - Partial similarity based on image regions



Applications: Images

- Windsurf (cont.)
  - Approach:
    - Haar Wavelet transformation of each color channel
    - Partitioning of the image based on clustering the three color coefficients (k-means clustering on 3rd subband wavelet coefficient)



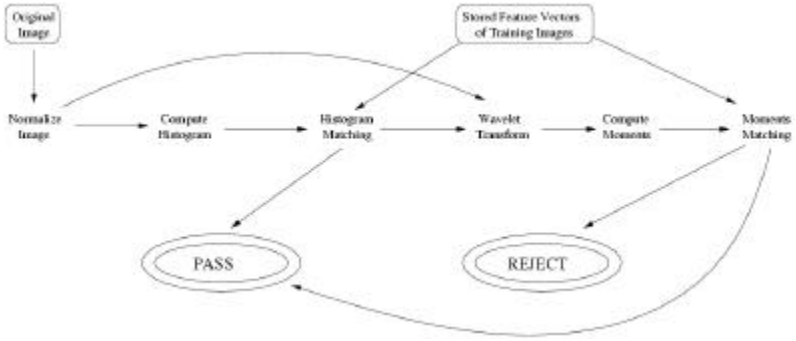
    - Feature vectors correspond to regions found in clustering step: (Size, Centroids, Covariance matrix of pixels in region)
    - Similarity retrieval based on matching regions

Applications: Images

- WIPE [WWF98]
  - Wavelet Image Pornography Elimination
  - Fast special purpose image filtering
  - Approach
    - Normalization of images to standard size
    - Wavelet transformation using Daubechies-3 wavelets
    - Edge detection in different subbands of wavelet transformation
    - Feature vectors used for similarity matching: Central moments, invariant moments, and color histograms
    - Filtering based on training the search for the desired filtering

Applications: Images

- WIPE (cont.)
  - Schematic approach:



  - Results: 95% correct images found with 10% wrong rejects

Applications: Images

- Similarity measure learning [BVG+99]
  - Interactive learning of the similarity measure
  - Approach:
    - Vector median filtering (to reduce noise)
    - Haar wavelet transform
    - Storage of 128 largest coefficients (quantized to +1 / -1)
    - Supervised learning to find similarity measure to find weighting for feature vector comparison

Applications: Images

- Partial image retrieval and sketch retrieval [WWF+97a,WWF+97b]
  - Improved content-based retrieval

low resolution | partial image | block sketch | hand-drawn sketch

Applications: Images

- Partial image retrieval (cont.)
  - Approach:
    - Wavelet transformation for each color component using Daubechies-8 wavelets
    - Low frequency wavelet coefficients and their variance are stored as feature vectors
    - 2-step retrieval:
      - Pre-selection (filtering) based on variance (candidates)
      - Similarity computation based on full feature vectors of candidates
    - Extension: Two-level multi-resolution similarity search

Overview

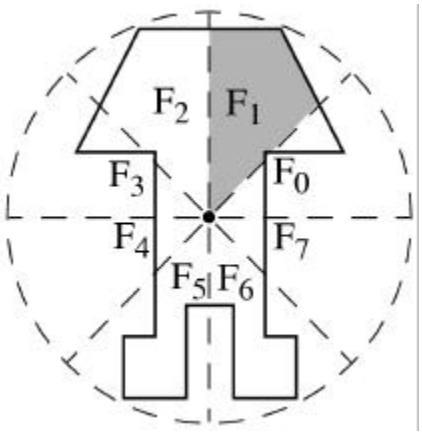
1. Introduction
2. Efficiency
3. Effectiveness
4. Applications
   i. Text
   ii. Images
   iii. Computer Aided Design (CAD)
   iv. 3D objects
   v. Audio
   vi. Video
5. Future research

Applications: CAD

- Geometric similarity
  - Shape similarity [Jag91]
  - Geometric molecular shape [BMH92]
  - Surface segments [KS98]
  - Shape histograms [AKK+99]

Applications: CAD

- Section coding [BK98]

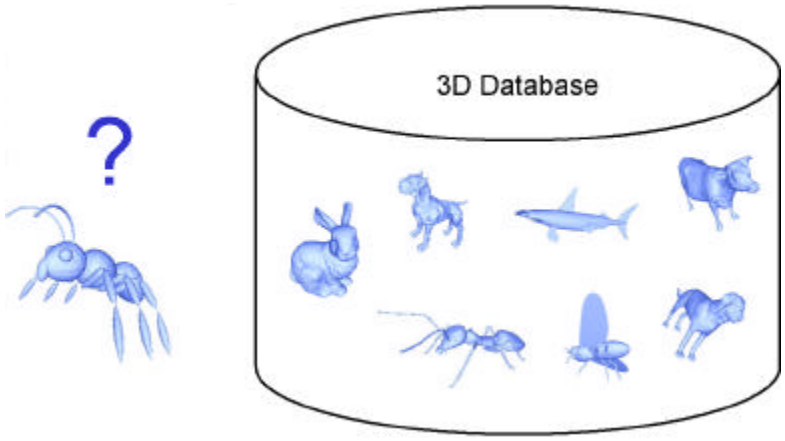


Applications: CAD

- Sets of feature vectors for searching voxelized CAD objects [KBK+03]
  - – Cover sequence model
  - – Set of feature vectors: Minimum Euclidean distance
    - •Minimum weight perfect matching
  - – Filter step using high-dimensional index structures

Applications: 3D objects

- 3D similarity search



Applications: 3D objects

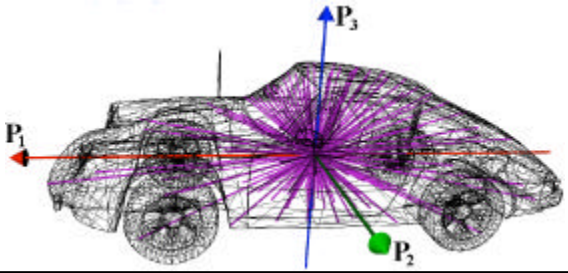
- 3D similarity search
  - Goal:
    - Effective content-based search of 3D objects
  - Requirements:
    - Invariance with respect to translation, rotation, scaling, and reflection
      - Pre-aligning objects (time consuming)
      - Principal component analysis (PCA)
      - Implicit invariance
    - Robustness with respect to level-of-detail and noise
    - Multi-resolution feature representation

Applications: 3D objects

- Classification of 3D feature vectors
  - *Statistics*
  - *Extension-based*
  - *Surface geometry*
  - *Image-based*
  - *Volume-based*

Applications: 3D objects

- Ray-based [VS00]
  - Extension-based descriptor
  - Approach:
    - Sample a 3D model in regularly spaced directions
    - Treat these samples as components for the descriptor



Applications: 3D objects

- Shape distribution with *D2* [OFC+02]
  - Statistical descriptor
  - Approach:
    - Describe the shape of a 3D object as a probability distribution sampled from a shape function
    - Shape function: Euclidean distance between two random points on the surface
    - Construct histograms from random sampling points

Applications: 3D objects

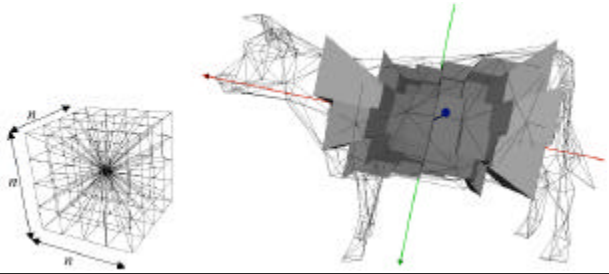
- Depth buffer [HKS+02]
  - Image based descriptor
  - Approach
    - Extend the projections with depth information
    - Code distance surface: View plane in grey values
    - Spatial domain: $6n^2$ values
    - Spectral domain 2D DFT: $6(n^2+n+1)$ values

Applications: 3D objects

- Silhouette [HKS+02]
  - Image-based descriptor
  - Approach
    - Form a silhouette (project model on specified plane)
    - Select contour points (equidistant or equiangular)
    - Form Fourier power spectrum using contour points
    - Take the first *n* coefficients as the descriptor components

Applications: 3D objects

- Volume-based [HKS+02]
  - Approach
    - $6n^2$ pyramid-like segments in the bounding cube
    - Net proportion of volume occupied by the solid object in each segment of the bounding box



Applications: 3D objects

- Other 3D descriptors:
  - "Cords" [PMN+00]
  - "Rotation invariant" [KSO00]
  - Shape spectrum [ZP01]
  - Topology matching [HSK+01]
  - Rotation invariant spherical harmonics [FMK+03]
  - "Lightfield" descriptor [CTS+03]

Applications: Audio

- Content-based audio retrieval

–Audio retrieval
- •Raw audio data [Foo99]
- •Query by humming [KG03]

–Audio content analysis
- •Speech
- •Music
- •Environment sound
- •Silence

Applications: Audio

- Applications
  - Audio classification and segmentation methods [LJZ01]
  - Music retrieval by humming or singing [JL01]
  - Content-based organization of music archives [PRM02]
  - Singer identification and classification of MP3 files [LH02]

Applications: Video

- Content-based video retrieval [PJ04]
  - Video media:
    - Large amounts of information: 1.2 Gb per minute (PAL video)
    - Images, audio, and text
  - Video structure
    - Segmentation (detection of cuts/shots)
    - Time-line models
    - Hierarchical models
  - Video content
    - Feature-based models
    - Annotation-based model

Applications: Video

- COBRA video modeling framework [PJ00]
  - –Raw video data layer
  - –Feature layer
  - –Concept layers
    - •Object layer
    - •Event layer

Applications: Video

- TREC video retrieval evaluation
  - – Objective: “to promote progress in content-based retrieval from digital video via open, metrics-based evaluation”
  - – 2001 and 2002: video “track” in TREC devoted to research in automatic segmentation, indexing, and content-based retrieval of digital video
  - – 2003: Independent evaluation (TRECVID)

Applications: Video

- TRECVID
  - Four main tasks:
    - Shot boundary determination
    - Story segmentation
    - High-level feature extraction
    - Search
  - Video data:
    - 120 hours of ABC World News Tonight and CNN Headline News (late January - June 1998)
    - 13 hours of C-SPAN programming (between 1998 – 2001).
  - URL: http://www-nlpir.nist.gov/projects/trecvid/

Overview

1. Introduction
2. Efficiency
3. Effectiveness
4. Applications
5. Future research

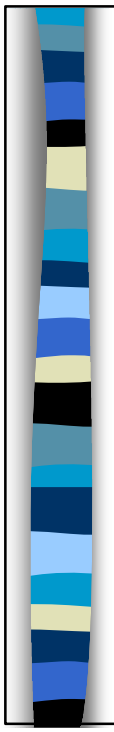
Future research

- New applications:
  - Partial similarity
  - Protein docking
  - Digital mock-up
  - Streams of Multimedia Data

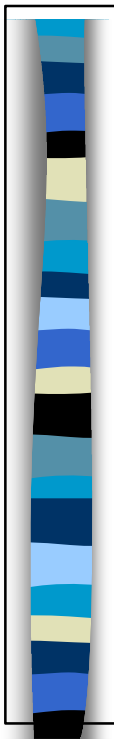
➔ New features transformations necessary

Future research

- Efficiency:
  - Index Structures
  - Query Processing (k-NN)
  - Query Optimization
- Effectiveness:
  - Similarity Measures
  - Evaluation (Ground Truth?)



The End



References

[AKK+99] M. Ankerst, G. Kastenmüller, H.-P. Kriegel, and T. Seidl. 3D shape histograms for similarity search and classification in spatial databases. *In Proc. 6th Symposium on Advances in Spatial Databases*, LNCS 1651, pages 207—226. Springer, 1999.

[AKS98] M. Ankerst, H.-P. Kriegel, and T. Seidl. A multistep approach for shape similarity search in image databases. *IEEE Transactions on Knowledge and Data Engineering*, 10(6):996—1004, 1998.

[ABP99] S. Ardizzoni, I. Bartolini, and M. Patella. Windsurf: Region based image retrieval using wavelets. In *Proc. 10th International Workshop on Database and Expert systems Applications (DEXA'99)*, pages 167—173. IEEE CS Press, 1999.

[AM95] S. Arya and D. Mount. Approximate range searching. In *Proc. 11th Annual ACM Symposium on Computational Geometry*, pages 172—181, 1995.

[AMN+98] S. Arya, D. Mount, N. Netanyahu, R. Silverman, and A. Wu. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *Journal of the ACM*, 45(6):891—923, 1998.

[AFH+95] J. Ashley, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic. The query by image content (QBIC) system. In *Proc. ACM International Conference on Management of Data (SIGMOD'95)*, page 475. ACM Press, 1995.

[BMH92] A. Badel, J. Mornon, and S. Hazout. Searching for geometric molecular shape complementary using bidimensional surface profiles. *Journal of Molecular Biology*, 10:205—211, 1992.

[BCM+94] R. Baeza-Yates, W. Cunto,U. Manber, and S. Wu. Proximity matching using fixed-queries trees. In *Proc. 5th Annual Symposium on Combinatorial Pattern Matching*, LNCS 807, pages 198—212, 1994.

[BR99] R. Baeza-Yates, B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley, 1999.

[BKS+90] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger. The R*-tree: An efficient and robust access method for points and rectangles. In *Proc. ACM International Conference on Management of Data (SIGMOD'90)*, pages 322—331. ACM Press, 1990.

[BBK98] S. Berchtold, C. Böhm, and H.-P. Kriegel. The pyramid technique: Towards breaking the curse of the dimensionality. In *Proc. ACM International Conference on Management of Data (SIGMOD'98)*, pages 142—153. ACM Press, 1998.

References

[BK98] S. Berchtold and D. Keim. Section coding: A similarity search technique for the car manufacturing industry. In *Proc. International Workshop on Issues and Applications of Database Technology (IADT'98)*, pages 256—263, 1998.

[BKK96] S. Berchtold, D. Keim, and H.-P. Kriegel. The X-tree: An index structure for high-dimensional data. In *Proc. 22nd International Conference on Very Large Databases (VLDB'96)*, pages 28—39. Morgan Kaufmann, 1996.

[BBK01] C. Böhm, S. Berchtold, and D. Keim. Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases. *ACM Computing Surveys*, 33(3):322—373, 2001.

[BVG+99], C. Brambilla, A. Ventura, I. Gagliardi, and R. Schettini. Multiresolution wavelet transform and supervised learning for content-based image retrieval. In *Proc. International Conference on Multimedia Computing and systems (ICMCS'99)*, volume 1, pages 183—188. IEEE CS Press, 1999.

[Bri95] S. Brin. Near neighbor search in large metric spaces. In *Proc. 21st International Conference on Very Large Databases (VLDB'95)*, pages 574—584. Morgan Kaufmann, 1995.

[BO97] T. Bozkaya and M. Ozsoyoglu. Distance-based indexing for high-dimensional metric spaces. In *Proc. ACM International Conference on Management of Data (SIGMOD'97)*, pages 357—368. ACM Press, 1997.

[BK73] W. Burkhard and R. Keller. Some approaches to best-match file searching. *Communications of the ACM*, 16(4):230—236, 1973.

[BN04] B. Bustos and G. Navarro. Probabilistic proximity searching algorithms based on compact partitions. *Journal of Discrete Algorithms*, 2(1) :115—134, 2004.

[BNC03] B. Bustos, G. Navarro, and E. Chávez. Pivot selection techniques for proximity searching in metric spaces. *Pattern Recognition Letters*, 24(14):2357—2366, 2003.

[CMB99] E. Chávez, J. Marroquín, and R. Baeza-Yates. Spaghettis: An array based algorithm for similarity queries in metric spaces. In *Proc. 6th International Symposium on String Processing and Information Retrieval (SPIRE'99)*, pages 38—46. IEEE CS Press, 1999.

[CMN01] E. Chávez, J. Marroquín, and G. Navarro. Fixed queries array: A fast and economical data structure for proximity searching. *Multimedia Tools and Applications*, 14(2):113—135, 2001.

References

[CN00] E. Chávez and G. Navarro. An effective clustering algorithm to index high dimensional metric spaces. In *Proc. 7th International Symposium on String Processing and Information Retrieval (SPIRE'00)*, pages 75—86. IEEE CS Press, 2000.

[CN03] E. Chávez and G. Navarro. Probabilistic proximity search: Fighting the curse of the dimensionality in metric spaces. *Information Processing Letters*, 85:39—46, 2003.

[CNB+01] E. Chávez, G. Navarro, J. Marroquín, and R. Baeza-Yates. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273—321, 2001.

[CTS+03] D. Chen, X. Tian, Y. Shen, and M. Ouhyoung. On Visual Similarity Based 3D Model Retrieval. In *Proc. Annual Conference of the European Association for Computer Graphics (Eurographics'03)*, pages 223—233, 2003.

[CP00] P. Ciaccia and M. Patella. PAC nearest neighbor queries: Approximate and controlled search in high-dimensional and metric spaces. In *Proc. 16th International Conference on Data Engineering (ICDE'00)*, pages 244—255, 2000.

[CP01] P. Ciaccia and M. Patella. Approximate similarity queries: A survey. Technical Report CSITE-08-01, Department of Electronics, Computer Science, and Systems, University of Bologna, May 2001.

[CPZ97] P. Ciaccia, M. Patella, and P. Zezula. M-tree: An efiicient access method for similarity search in metric spaces. In *Proc. 23rd International Conference on Very Large Databases (VLDB'97)*, pages 426—435. Morgan Kaufmann, 1997.

[Cla99] K. Clarkson. Nearest neighbor queries in metric spaces. *Discrete Computational Geometry*, 22(1):63—93, 2001.

[DN87] F. Dehne and H Noltemeier. Voronoi trees and clustering problems. *Information Systems*, 12(2):171—175, 1987.

[Fal96] C. Faloutsos. *Searching Multimedia Databases by Content*. Kluwer Academic Publishers, 1996.

[FL95] C. Faloutsos and K. Lin. FastMap: A fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. Technical report CS-TR-3383, UMIACS-TR-94-132; Inst. for Systems Research: TR 94-80. (Also in *Proc. ACM SIGMOD*, pages 163—174, 1995.)

References

[Foo99] J. Foote. An overview of audio information retrieval. *ACM Multimedia Systems*, 7(1):2—10, 1999.

[FMK+03] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1):83—105, 2003.

[Gut84] A. Guttman. R-trees: A dynamic index structure for spatial searching. In *Proc. ACM International Conference on Management of Data (SIGMOD'84)*, pages 47—57. ACM Press, 1984.

[HHK+02] M. Heczko, A. Hinneburg, D. Keim, and M. Wawryniuk. Multi-resolution similarity search in image databases. In *Proc. 8th International Workshop on Multimedia Information Systems*, pages 76—85, 2002.

[HKS+02] M. Heczko, D. Keim, D. Saupe, and D. Vranic. Methods for similarity search on 3D databases. Datenbank-Spektrum, 2(2):54—63, 2002. In German.

[Hen98] A. Henrich. The LSD^h-tree: An access structure for feature vectors. In *Proc. 14th International Conference on Data Engineering (ICDE'98)*, pages 362-369. IEEE CS Press, 1998.

[HSK+01] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In Proc. ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'01), pages 203—212. ACM Press, 2001.

[Jag91] H. Jagadish. A retrieval technique for similar shapes. In *Proc. ACM International Conference on Management of Data (SIGMOD'91)*, pages 208—217. ACM Press, 1991.

[JL01] J. Jang and H. Lee. Hierarchical filtering method for content-based music retrieval via acoustic input. In Proc. *9th ACM International Conference on Multimedia (MM'01)*, pages 401—410. ACM Press, 2001.

[KM83] I. Kalantari and G. McDonald. A data structure and an algorithm for the nearest point problem. *IEEE Transactions on Software Engineering*, 9(5):631—634, 1983.

[KS97] N. Katayama and S. Satoh. The SR-tree: An index structure for high-dimensional nearest neighbor queries. In *Proc. ACM International Conference on Management of Data (SIGMOD'97)*, pages 369—380. ACM Press, 1997.

[KSO00] T. Kato, M. Suzuki, and N. Otsu. A similarity retrieval of 3D polygonal models using rotation invariant shape descriptors. *In Proc. IEEE International Conference on Sytems, Man, and Cybernetics*, pages 2946—2952, 2000.

References

[KG03] R. Kline and E. Glinert. Approximate matching algorithms for music information retrieval using vocal input. In *Proc. 11th ACM International Conference on Multimedia (MM'03)*, pages 130—139. ACM Press, 2003.

[Kor97] R. Korfhage. Information Storage and Retrieval. John Wiley & Sons, Inc., New York, 1997.

[KBK+03] H.-P. Kriegel, S. Brecheisen, P. Kröger, M. Pfeifle, and M. Schubert. Using sets of feature vectors for similarity search on voxelized CAD objects. In *Proc. ACM International Conference on Management of Data (SIGMOD'03)*, pages 587—598. ACM Press, 2003.

[KS98] H.-P. Kriegel and T. Seidl. Approximate-based similarity search for 3D surface segments. *Geoinformatica Journal*, 2(2):113—147, 1998.

[LJF94] K.-I. Lin, H. Jagadish, and C. Faloutsos. The TV-tree: An index structure for high-dimensional data. *The VLDB Journal*, 3(4):517—542, 1994.

[LTO+01] S. Lin, M. Tamer, V. Oria, and R. Ng. An extendible hashing for multi-precision similarity querying of image databases. In *Proc. 27th International Conference on Very Large Databases (VLDB'01)*, pages 221—230. Morgan Kaufmann, 2001.

[LH02] C. Liu and C. Huang. Music information retrieval: A singer identification technique for content-based classification of MP3 music objects. In *Proc. 11th International Conference on Information and Knowledge Management (CIKM'02)*, pages 438—445. ACM Press, 2002.

[LJZ01] L. Lu, H. Jiang, and H. Zhang. Audio processing: A robust audio classification and segmentation method. In *Proc. 9th ACM International Conference on Multimedia (MM'01)*, pages 203—211. ACM Press, 2001.

[MOV94] L. Micó, J. Oncina, and E. Vidal. A new version of the nearest-neighbor approximating and eliminating search (AESA) with linear preprocessing-time and memory requirements. *Pattern Recognnition Letters*, 15:9—17, 1994.

[NRS04] A. Natsev, R. Rastogi, and K. Shim. WALRUS: A similarity retrieval algorithm for image databases. IEEE Transactions on Knowledge and Data Engineering, 16(3):301—316, 2004.

References

[Nav01] G. Navarro. A guided tour to approximate string matching. *ACM Computing Surveys*, 33(1):31—88, 2001.

[Nav02] G. Navarro. Searching in metric spaces by spatial approximation. *The VLDB Journal*, 11(1):28—46, 2002.

[NR02] G. Navarro and M. Raffinot. *Flexible Pattern Matching in Strings* – Practical on-line search algorithms for texts and biological sequences. Cambridge University Press, 2002.

[NVZ92] H. Noltemeier, K. Verbarg, and C. Zirkelbach. Monotonous bisector* trees - a tool for efficient partitioning of complex scenes of geometric objects. In *Data Structures and Efficient Algorithms*, LNCS 594, pages 186—203. Springer, 1992.

[OFC+02] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. *ACM Transactions on Graphics*, 21(4):807—832. 2002.

[PRM02] E. Pampalk, A. Rauber, and D. Merkl. Content-based organization and visualization of music archives. In *Proc. 10th ACM International Conference on Multimedia (MM'02)*, pages 570—579. ACM Press, 2002.

[PMN+00] E. Paquet, M. Murching, T. Naveen, A. Tabatabai, and M. Rioux. Description of shape information for 2-D and 3-D objects. Signal Processing: Image Communications, 16:103—122, 2000.

[PJ00] W. Petkovic and W. Jonker. A framework for video modeling. In *Proc. 18th IASTED International Conference on Applied Informatics*, pages 317—322, 2004.

[PJ04] W. Petkovic and W. Jonker. *Content-based Video Retrieval: A Database Perspective*. Kluwer Academic Publishers, 2004.

[Rij79] C. van Rijsbergen. *Information Retrieval*. Butterworths, London, 1979.

[Rob81] J. Robinson. The kd-B-tree: A search structure for large multidimensional dynamic indexes. In *Proc. ACM International Conference on Management of Data (SIGMOD'81)*, pages 10—18. ACM Press, 1981.

[Sag94] H. Sagan. *Space Filling Curves*. Springer-Verlag, New York, 1994.

References

[STT+01] R. Santos-Filho, A. Traina, C. Traina Jr., and C. Faloutsos. Similarity search without tears: The OMNI family of all-purpose access methods. In *Proc. 17th International Conference on Data Engineering (ICDE'01)*, pages 623—630. IEEE CS Press, 2001.

[SK97] T. Seidl and H.-P. Kriegel. Efficient user-adaptable similarity search in large multimedia databases. In *Proc. 23rd International Conference on Very Large Databases (VLDB'97)*, pages 506—515. Morgan Kaufmann, 1997.

[SRF87] T. Sellis, N. Roussopoulos, and C. Faloutsos. The R+-tree: A dynamic index for multi-dimensional objects. In *Proc. 13th International Conference on Very Large Databases (VLDB'87)*, pages 507—518. Morgan Kaufmann, 1987.

[SWW+91] W. Shaw, J. Wood, R. Wood, and H. Tibbo. The Cystic Fibrosis Database: Content and research opportunities. LISR 13, pages 347—366, 1991.

[SBP97] W. Shaw, R. Burgin, and P. Howell. Performance standards and evaluations in IR test collections: Cluster-based retrieval models. *Information Processing & Management*, 33(1): 1—14, 1997.

[SWW+91] W. Shaw, J. Wood, R. Wood, and H. Tibbo. The cystic fibrosis database: Content and research opportunities. *Library and Information Science Research*, 13:347—366, 1991.

[SMK+04] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The Princeton Shape Benchmark. Shape Modeling International, 2004.

[Uhl91] J. Uhlmann. Satisfying general proximity/similarity queries with metric trees. *Information Processing Letters*, 40:175—179, 1991.

[Vid86] E. Vidal. An algorithm for finding nearest neighbors in (approximately) constant average time. *Pattern Recognittion Letters*, 4:145—157, 1986.

[VB98] N. Vujovic and D. Brazakovic. Evaluation of an algorithm for finding a match of a distorted texture pattern in a large image database. *ACM Transactions on Information Systems*, 16(1):31—60, 1998.

References

[VS00] D. Vranic and D. Saupe. 3D model retrieval. In *Proc. Spring Conference on Computer Graphics and its Applications (SCCG'00)*, pages 89—93. Comenius University, 2000.

[WWF98] J. Wang, G. Wiederhold, and O. Firschein. System for screening objectionable images. *Computer Communications*, 21(15):1355-1360, Elsevier, 1998.

[WWF+97a] J. Wang, G. Wiederhold, O. Firschein, and S. Wei. Content-based image indexing and searching using daubechies' wavelets. *International Journal on Digital Libraries*, 1(4):311—328, 1997.

[WWF+97b] J. Wang, G. Wiederhold, O. Firschein, and S. Wei. Wavelet-based image indexing techniques with partial sketch retrieval capability. In *Proc. 4th Forum on Research and Technology Advances in digital Libraries (ADL'97)*, pages 13—24, 1997.

[WJ96a] D. White and R. Jain. Algorithms and strategies for similarity retrieval. Technical Report VCL-96-101, Visual Computing Laboratory, University of California, La Jolla, California, July 1996.

[WJ96b] D. White and R. Jain. Similarity indexing with the SS-tree. In *Proc. 12th International Conference on Data Engineering (ICDE'96)*, pages 516—523. IEEE CS Press, 1996.

[Yia93] P. Yianlos. Data structures and algorithms for nearest neighbor search in general metric spaces. In *Proc.4th ACM/SIAM Symposium on Discrete Algorithms (SODA'93)*, pages 311—321, 1993.

[Yia00] P. Yianlos. Locally lifting the curse of dimensionality for nearest neighbor search. In *Proc.11th ACM/SIAM Symposium on Discrete Algorithms (SODA'00)*, pages 361—370, 2000.

[YI99] A. Yoshitaka and T. Ichikawa. A survey on content-based retrieval for multimedia databases. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):81—93, 1999.

[ZP01] T. Zaharia and F. Prêteux. 3D-shape-based retrieval within the MPEG-7 framework. In *Proc. SPIE Conference on Nonlinear Image Processing and Pattern Analysis XII*, volume 4304, pages 133—145, 2001.

[ZSA+98] P. Zezula, P. Savino, G. Amato, and F. Rabitti. Approximate similarity retrieval with M-trees. *The VLDB Journal*, 7(4):275—293, 1998.