

CC52V BASES DE DATOS MULTIMEDIA

10 UD

Prof. Benjamín Bustos
Semestre 2007/1

1. Requisitos: CC30A, CC42A.

2. Objetivos:

Introducir los conceptos y algoritmos principales para la implementación de bases de datos multimedia. Al finalizar el curso se espera que el alumno conozca los distintos tipos de datos multimedia (imágenes, objetos 3D, grafos, etc.) así como sus respectivos modelos de similitud, los algoritmos de búsqueda de objetos similares y los índices existentes para realizar estas búsquedas en forma eficiente.

3. Programa:

- Introducción (3 horas). Principios básicos de las bases de datos multimedia. Tipos de datos multimedia. Aplicaciones prácticas.
- Búsqueda por similitud en bases de datos multimedia (6 horas). Modelamiento de datos multimedia en espacios métricos y multidimensionales. Transformación en vectores característicos. Métricas y funciones de distancia. Búsqueda por rango y vecinos más cercanos. Efectividad y eficiencia de la búsqueda por similitud.
- Modelos de similitud para datos multimedia (15 horas). Modelos de similitud para imágenes. Modelos de similitud para objetos geométricos en 2-D y 3-D. Búsqueda por similitud parcial en aplicaciones CAD. Búsqueda por similitud en bases de documentos. Modelos de similitud para grafos.
- Algoritmos de búsqueda por similitud (6 horas). Principios generales de los índices multidimensionales. Búsqueda basada en filtrar-y-refinar. Algoritmos para búsquedas por rango. Algoritmo de Roussopoulos, Kelley y Vincent para el vecino más cercano. Algoritmo de Hjaltason y Samet para vecinos más cercanos. Modelos de costo de la búsqueda.
- Índices para bases de datos multimedia (9 horas). Índices multidimensionales (R-tree y variantes, X-tree, SS-tree, SR-tree, métodos de filtrado). Índices para espacios métricos (índices basados en pivotes, M-tree, List of Clusters, índices basados en una partición de Voronoi).
- Tópicos avanzados (3 horas). Espacios multimétricos. Ingeniería de vectores característicos.

4. Metodología y evaluación:

Se realizarán dos clases de cátedra semanales, más una clase auxiliar donde se resolverán ejercicios prácticos. La evaluación consistirá en dos controles y un examen (2/3 de la nota final) más una nota correspondiente a tareas (1/3 de la nota final).

5. Bibliografía:

- [1] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. Modern Information Retrieval. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.
- [2] Christian Böhm, Stefan Berchtold, and Daniel Keim. Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases. *ACM Computing Surveys*, 33(3):322–373, 2001.
- [3] Edgar Chávez, Gonzalo Navarro, Ricardo Baeza-Yates, and José Luis Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, 2001.
- [4] Christos Faloutsos. Searching Multimedia Databases by Content. Kluwer Academic Publishers, Norwell, MA, USA, 1996.
- [5] V. S. Subrahmanian. Principles of Multimedia Database Systems. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998.
- [6] Pavel Zezula, Giuseppe Amato, Vlastislav Dohnal, and Michal Batko. Similarity Search: The Metric Space Approach (Advances in Database Systems). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.