




CC5213
Recuperación de
Información Multimedia

Juan Manuel Barrios

jbarrios@dcc.uchile.cl

<http://juan.cl/>

Departamento de Ciencias de la Computación
Universidad de Chile
Semestre Otoño 2018



Curso CC5213 (2018)

- **Recuperación de Información Multimedia**
 - **Multimedia Information Retrieval**
- 10 Unidades Docentes
- Martes y Jueves 12:00 – 13:30
- Requisito (basta uno):
 - CC5206 – Introducción a la Minería de Datos
 - EL4106 – Inteligencia Computacional
 - CC4102 – Diseño y Análisis de algoritmos



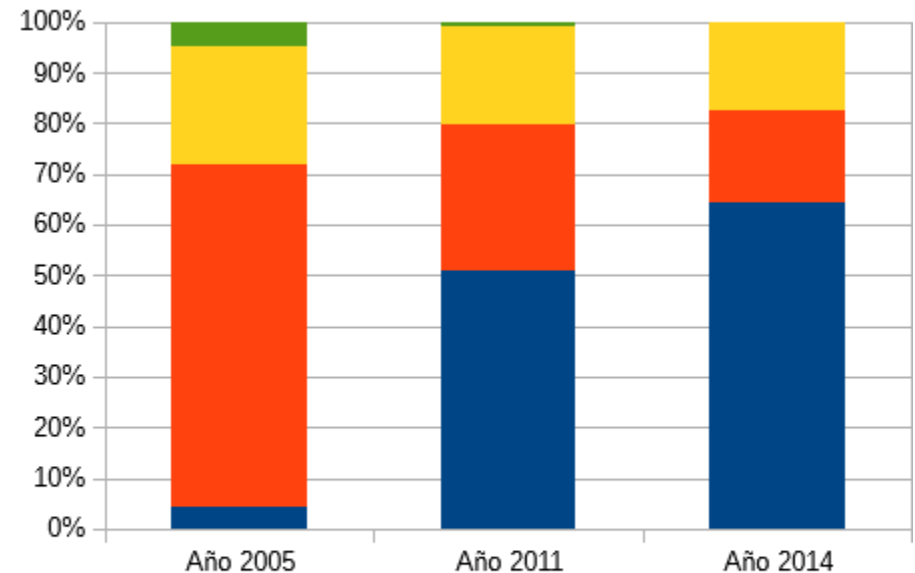
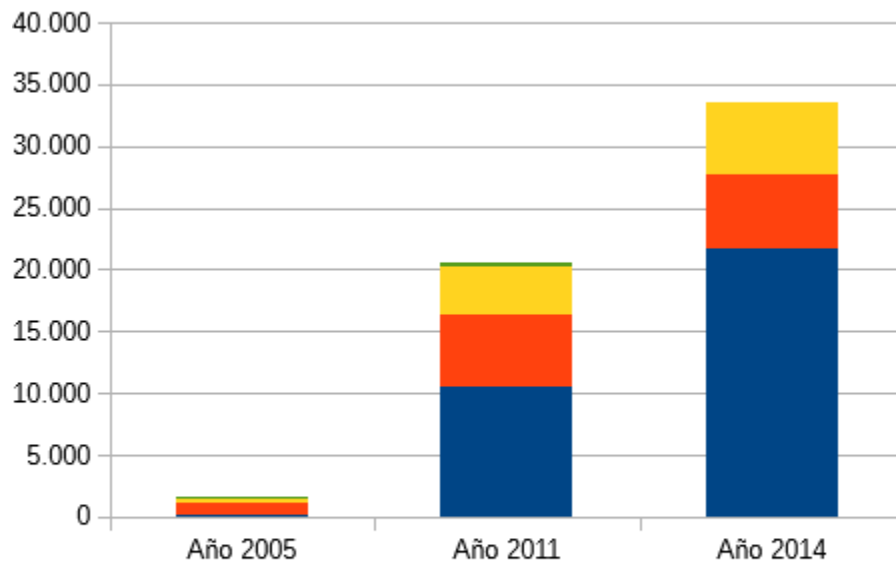
Recuperación de Información

- **Information Retrieval (IR)** estudia cómo representar, organizar, almacenar y acceder a **información** existente en documentos
 - Objetivo: Recuperar (*retrieve*) documentos **relevantes** a la necesidad de **información** del usuario
 - Documentos: página web, texto de emails, conjunto de tags, fichas de libros, etc.
 - Consulta: frase, keywords, preguntas

Evolución del uso de Internet

- 2014: más del 60% del tráfico corresponde a videos

Tráfico de datos de usuarios de Internet [PB mensual]



■ Otros
■ Web, email, datos
■ P2P
■ Video

Fuente:

Global IP Traffic Forecast and Methodology, 2006-2011.
Cisco Visual Networking Index: Forecast and Methodology, 2011-2016.
Cisco Visual Networking Index: Forecast and Methodology, 2014-2019.



Recuperación de Información Multimedia

■ Multimedia Information Retrieval (MIR)

- Buscar documentos **multimedia relevantes** a la necesidad de **información** del usuario
- Documentos: Audio, Imagen, Video, Objetos 3D, etc.
Ej: fotos personales, películas, grabación de cámaras de seguridad, música sin ID3, etc.
- La búsqueda **NO requiere metadatos ni etiquetas**, si no que debe analizar **contenido multimedia** (píxeles, samples de audio, frames de video)
- Consulta: keywords, “by-example” (buscar algo parecido a un documento modelo), “by-sketch” (buscar usando bosquejos)

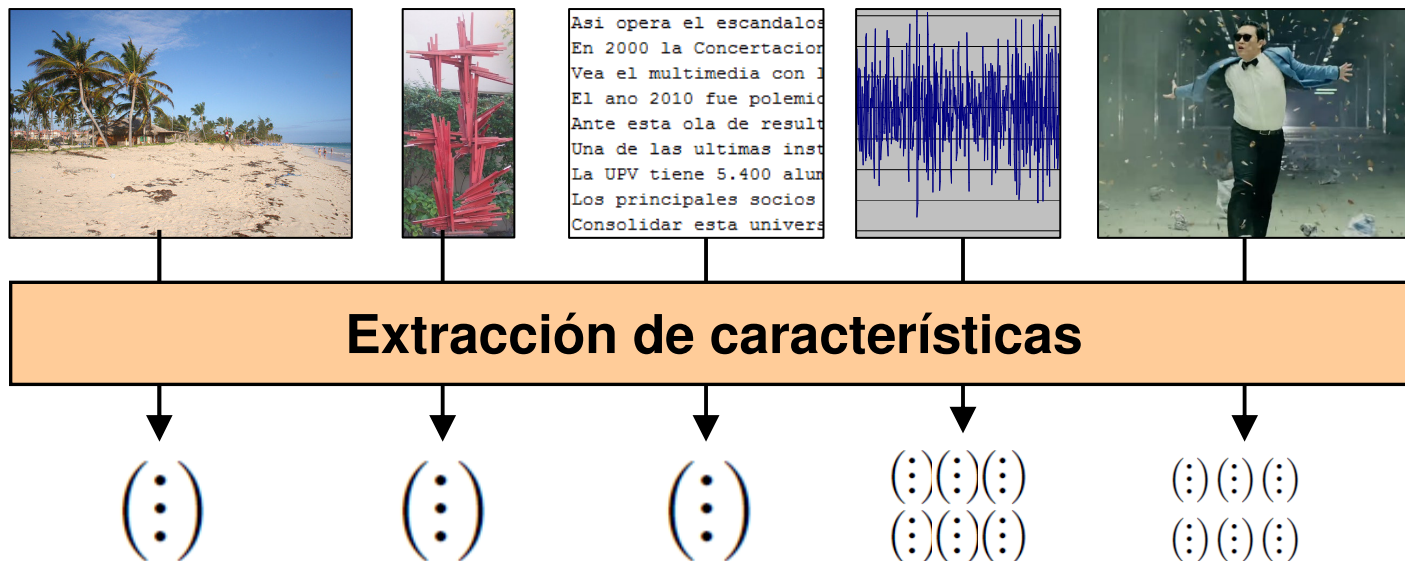


Áreas Involucradas

- Information Retrieval
- Data Mining
- Machine Learning
- Pattern Recognition
- Computer Vision
- Artificial Intelligence
- Data Visualization

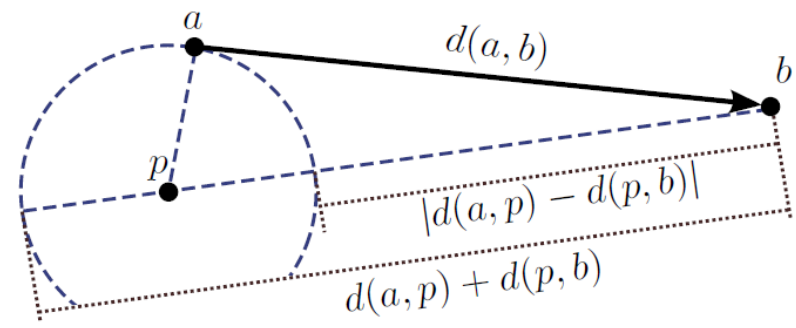
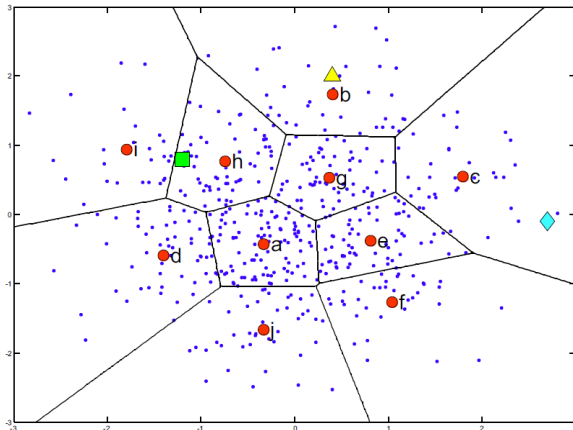
Descripción del Contenido

- Representar el contenido de cada documento multimedia (foto, canción, video, email, etc.) por uno o más vectores.



Búsqueda por Similitud

- Resolver eficientemente búsquedas de vectores entre millones de elementos
- Índices para espacios vectoriales y métricos





Temas a estudiar en el semestre

- Parte 1: Descripción de contenido multimedia
 - Métodos para analizar contenido de imágenes, audio, videos y texto.
- Parte 2: Búsqueda por similitud e Indexamiento
 - Métodos para resolver eficientemente búsquedas en espacios vectoriales y métricos.
- Parte 3: Aplicaciones e Investigación
 - Técnicas actuales de descripción y búsqueda usando Codebooks y Deep Learning.

Casos de Estudio

- Durante el semestre se estudiarán métodos para resolver diferentes problemas tipo
- Problema 1: Dada una imagen de consulta, buscar otras imágenes parecidas

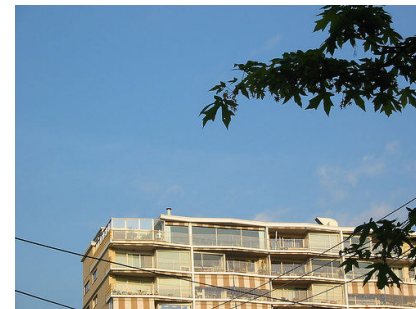
Imagen de consulta



Casos de Estudio

- Buscar imágenes parecidas según algún criterio (colores, formas, etc.)

Resultados



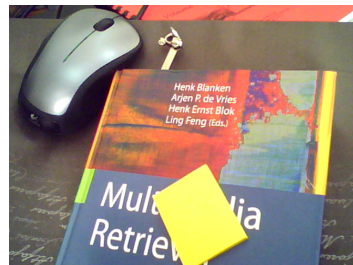
Ej: Dada la imagen de una playa encontrar otras imágenes parecidas.

Casos de Estudio

- Problema 2: Dado un catálogo de objetos, reconocer el objeto fotografiado



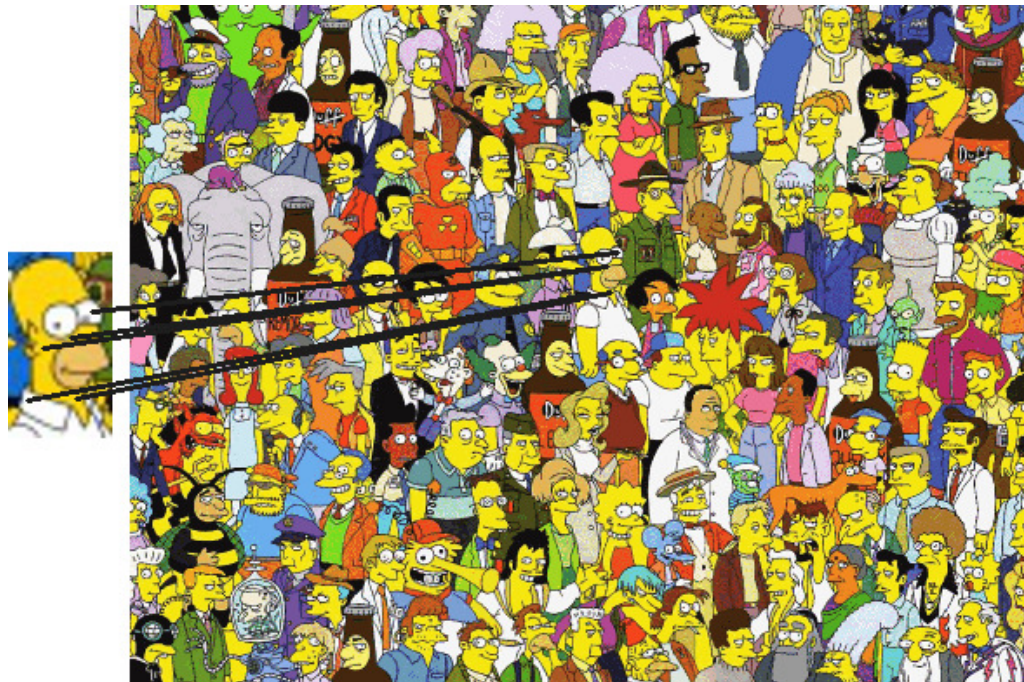
Imagen de consulta



Ver: <http://www.impresee.com/>

Casos de Estudio

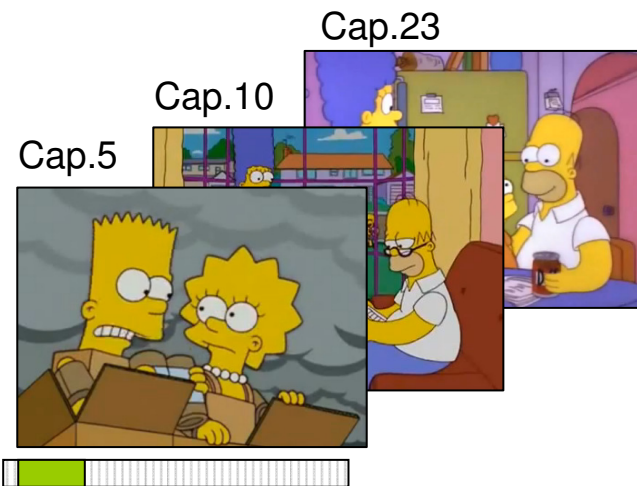
- Además de la ocurrencia, determinar la ubicación y pose del objeto encontrado.



Casos de Estudio

- Problema 3: Dado un video de Internet determinar la escena original de la que proviene.
- Problema 3: Reconocer una canción según un trozo de audio (“Shazam”)

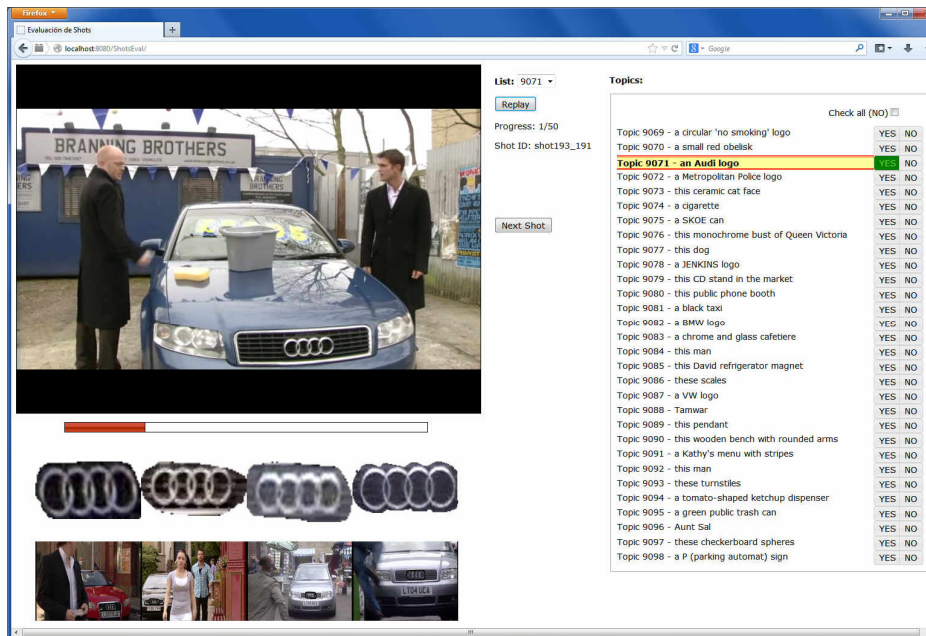
youtube



Ver: <http://sourceforge.net/projects/p-vcd/>

Casos de Estudio

- Problema 4: Buscar la aparición de algún producto, logo, etc. en televisión



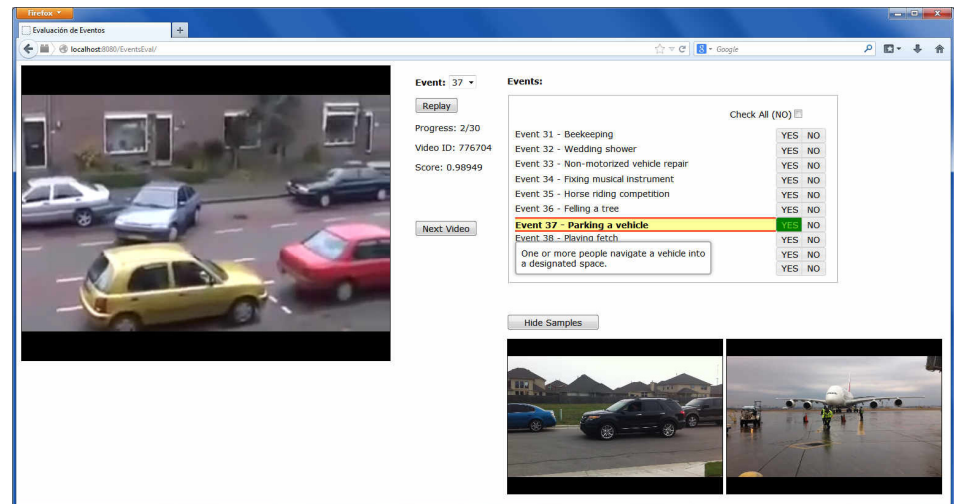
Ej.: Buscar las apariciones del logo Audi
© BBC EastEnders

Ver: <http://www.robots.ox.ac.uk/~vgg/research/vgoogle/>

Casos de Estudio

- Problema 5: Buscar eventos en videos, ingresando texto o imágenes de ejemplo, sin requerir de metadatos ni etiquetado previo de los datos

Ej: Buscar videos donde aparezca un vehículo estacionando



The screenshot shows a web-based interface for video event analysis. On the left, a video player displays a scene with several cars in a parking lot. The interface includes a 'Replay' button, progress information (2/30), video ID (776704), and a score (0.98949). A 'Next Video' button is also present. On the right, an 'Events' section lists various event types with corresponding 'YES' or 'NO' checkboxes. A table below the list shows the classification results for each event.

Event	Classification
Event 31 - Beekeeping	YES NO
Event 32 - Wedding shower	YES NO
Event 33 - Non-motorized vehicle repair	YES NO
Event 34 - Fixing musical instrument	YES NO
Event 35 - Horse riding competition	YES NO
Event 36 - Felling a tree	YES NO
Event 37 - Parking a vehicle	YES NO
Event 38 - Fixing a flat	YES NO
Event 39 - Fixing a flat	YES NO
Event 40 - Fixing a flat	YES NO

Below the table, there are two small thumbnail images showing different scenes: a car in a parking lot and an airplane on a tarmac.

Muchas Aplicaciones Novedosas y Problemas Abiertos

Google: Flujo de videos en YouTube hace imposible eliminar todo el contenido terrorista

La compañía ju
no se puede de
29 de Enero de 2015 |

Google sigue los pasos de Facebook e implementará inteligencia artificial contra contenidos terroristas

La empresa enfocará g
través de YouTube, una
19 de Junio de 2017 | 16:39 | DPA

Liga inglesa demanda a YouTube por violar derechos de autor

La asociación que maneja el campeonato inglés inició la acción legal en Nueva York, ante los numerosos videos con imágenes de la Premier League que están disponibles en el sitio web.

La aplicación móvil Shazam permitirá a sus usuarios identificar objetos

La herramienta que
identificar producto
05 de Marzo de 2015 | 16:4

Apple confirma la adquisición de Shazam y aventura "interesantes planes" para el futuro

Las compañías debe
comenzar a integrar
11 de Diciembre de 2017 | 15

Facebook reconocerá automáticamente canciones que suenan de fondo

La red social también podrá identificar los programas de televisión que ven sus usuarios mientras escriben mensajes.

22 de Mayo de 2014 | 09:17 | DPA

Aplicación chilena encuentra productos con solo dibujarlos o sacarles una foto

• Creada por investigadores universitarios, la herramienta analiza imágenes y busca los objetos en



De la idea al prototipo

Spotify amplía su alcance más allá de la música, sumando videos y podcasts a su servicio

La compañía de Daniel Ek estrenó hoy un nuevo modelo que busca ayudar al descubrimiento de contenido en la app. También sumó nuevas funciones para escuchar música al hacer deporte.

20 de Mayo de 2015 | 13:33 | Emol



Actividades (2018)

1. Mini-Controles al inicio de clases
 - Una pregunta escrita, 15-20 minutos, sin apuntes
2. Tareas de Programación
 - Python, C++ o Java
3. Presentación de Lecturas
 - A elección entre opciones dadas
4. Proyecto Final
 - Desarrollo de algún problema a elección
 - Presentación de resultados al finalizar el semestre

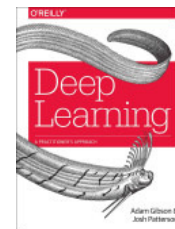
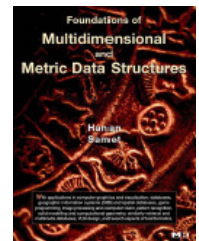
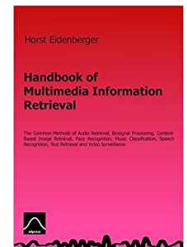
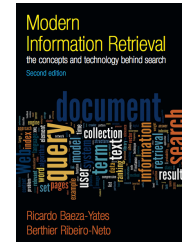


Evaluación (2018)

- Siete mini-controles, se borra uno
 - **NC** = $(C1 + \dots + C7 - \text{Min}\{C1, \dots, C7\}) / 6$
- Tres tareas de programación, una tarea con Informe
 - **NT** = $(T1 + T2 + I2 + T3) / 4$
- Dos presentaciones de lecturas
 - **NL** = $(L1 + L2) / 2$
- Proyecto final (T=trabajo realizado, P=presentación final)
 - **NP** = $(T + P) / 2$
- Para aprobar:
NC \geq 4.0, **NL** \geq 4.0, **NT** \geq 4.0, **NP** \geq 4.0
- Nota Final:
(NC + NL + NT + NP) / 4

Bibliografía

- **Modern Information Retrieval.** Baeza-Yates, Ribeiro-Neto, 2011.
- **Handbook of Multimedia Information Retrieval.** Eidenberger. 2012.
- **Multimedia Retrieval.** Blanken, de Vries, Blok, Feng. 2007.
- **Foundations of Multidimensional and Metric Data Structures.** Samet. 2006.
- **Deep Learning: A Practitioner's Approach.** Patterson, Gibson. 2017.



Bibliografía

- **The Essential Guide to Image/Video Processing.** Bovik. 2009.
- **H.264 and MPEG-4 Video Compression.** Richardson. 2003.
- **Digital Image Processing.** Gonzalez, Woods. 2008.
- **Data Mining: The Textbook.** Aggarwal. 2015.
- **Similarity Search The Metric Space Approach.** Zezula, Amato, Dohnal, Batko. 2006.
- **Computer Vision. Algorithms and Applications.** Szeliski. 2011.

