

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 Primavera 2017



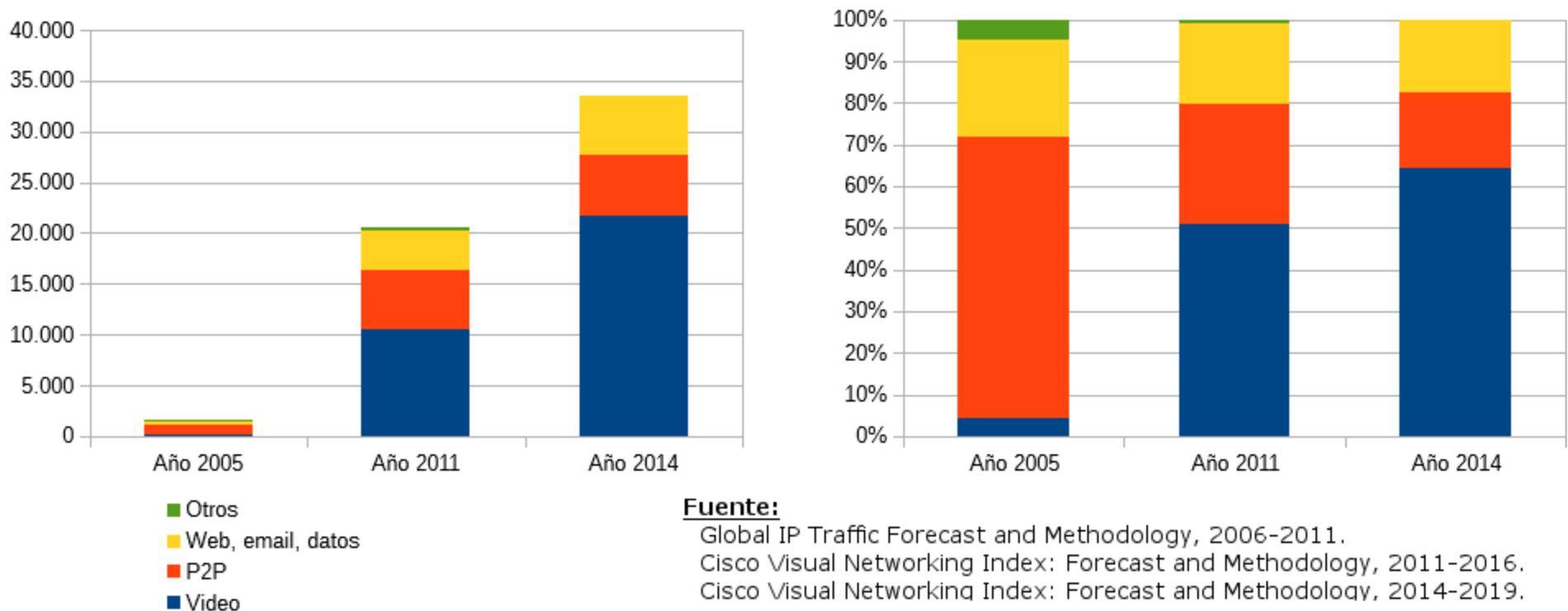
Curso CC5213

- **Recuperación de Información Multimedia**
 - **Multimedia Information Retrieval**
- **Martes y Jueves 16:15 – 17:45**
- **10 Unidades Docentes**
- **Requisito (basta uno):**
 - CC5206 – Introducción a la Minería de Datos
 - EL4106 – Inteligencia Computacional
 - CC4102 – Diseño y Análisis de algoritmos
- **Reemplaza a CC5204 – Búsqueda por Contenido de Imágenes y Videos (código equivalente)**
 - Líneas de especialización (<http://www.dcc.uchile.cl/node/238>)

Evolución del uso de Internet

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

Tráfico de datos de usuarios de Internet [PB mensual]





Nuevas Aplicaciones

Qualcomm presenta una plataforma que permite reconocer personas en tiempo real

Con esta tecnología en el momento de

02 de Marzo de 2015 | 16

Apple recluta a especialistas en inteligencia artificial para reducir brecha con Google

La compañía ha estado donde ha perdido terr

07 de Septiembre de 2015 | 14

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

La empresa enfocará gran parte de sus recursos en la identificación de videos compartidos a través de YouTube, una de sus plataformas con mayor popularidad.

19 de Junio de 2017 | 16:39 | DPA

La aplicación móvil Shazam permitirá a sus usuarios identificar objetos

La herramienta que reconoce la música y identificar productos en un supermercado

05 de Marzo de 2015 | 16:42 | Reuters

El streaming de música vive su mejor momento: Consumo creció un 10% en EE.UU. en lo que va del año

Cada semana 10.800 millones de canciones son reproducidas a través de servicios como Spotify, Apple Music y Tidal, que a su vez representan el 78% del mercado.

05 de Julio de 2017 | 16:40 | EFE



Búsqueda de Información

- **Information Retrieval (IR)** búsqueda de documentos relevantes a una necesidad de información del usuario.
 - Tipos de Documentos: páginas web, libros, e-mails, etc.
 - Tipos de Consulta: palabras, frase, pregunta.
 - Casos de uso típico: Buscador web, Catálogo de biblioteca.
- **Multimedia Information Retrieval (MIR)** búsqueda de documentos multimedia.
 - Tipos de Documentos: audio, imagen, video sin metadatos.
 - Fotos no etiquetadas, videos sin descripción, música sin ID3, etc.
 - Tipos de Consulta: palabras o “by-example”.

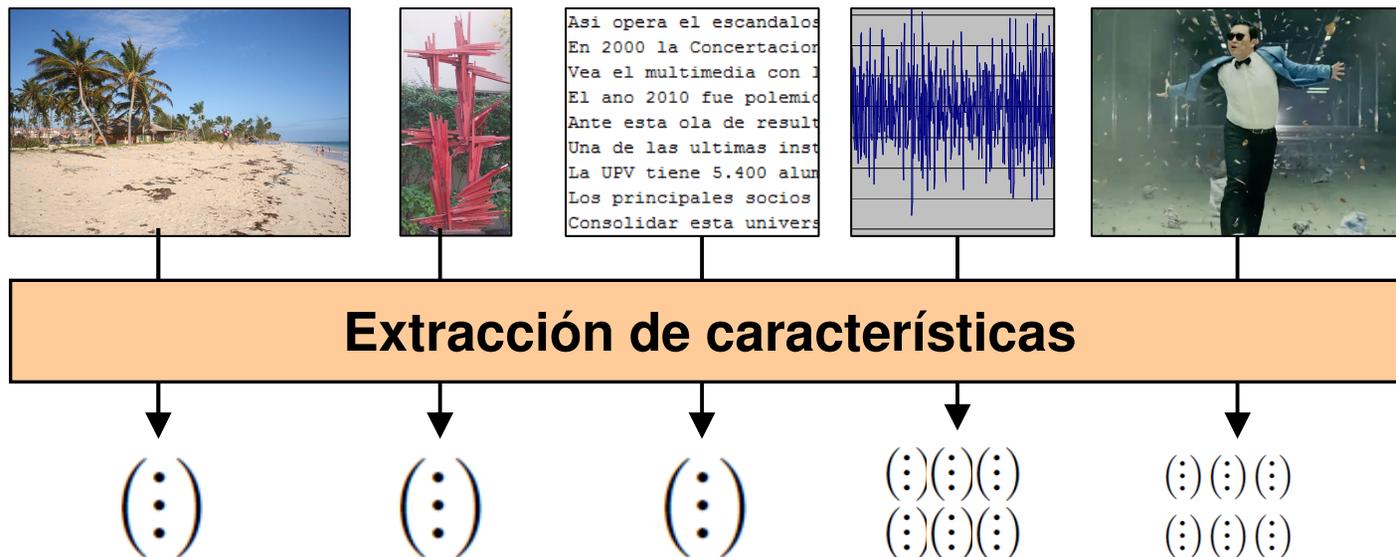


Áreas Involucradas en MIR

- Information Retrieval
- Computer Vision
- Pattern Recognition
- Data Mining
- Machine Learning
- Data Visualization
- Recommender Systems

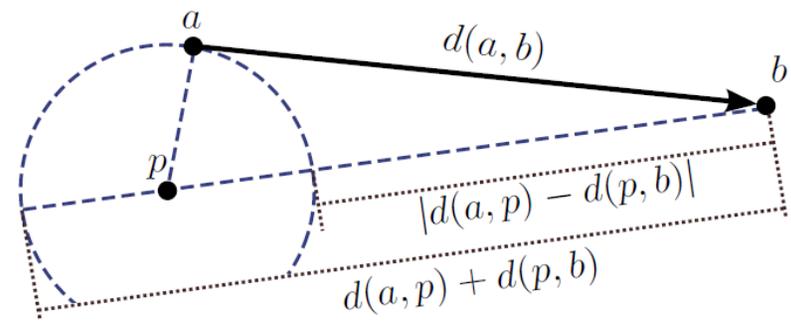
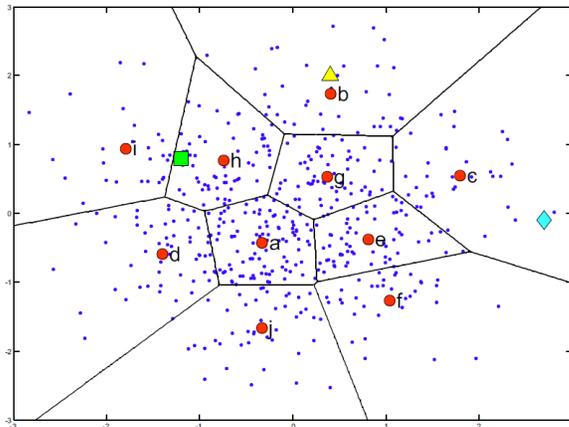
Descripción del contenido

- 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.



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

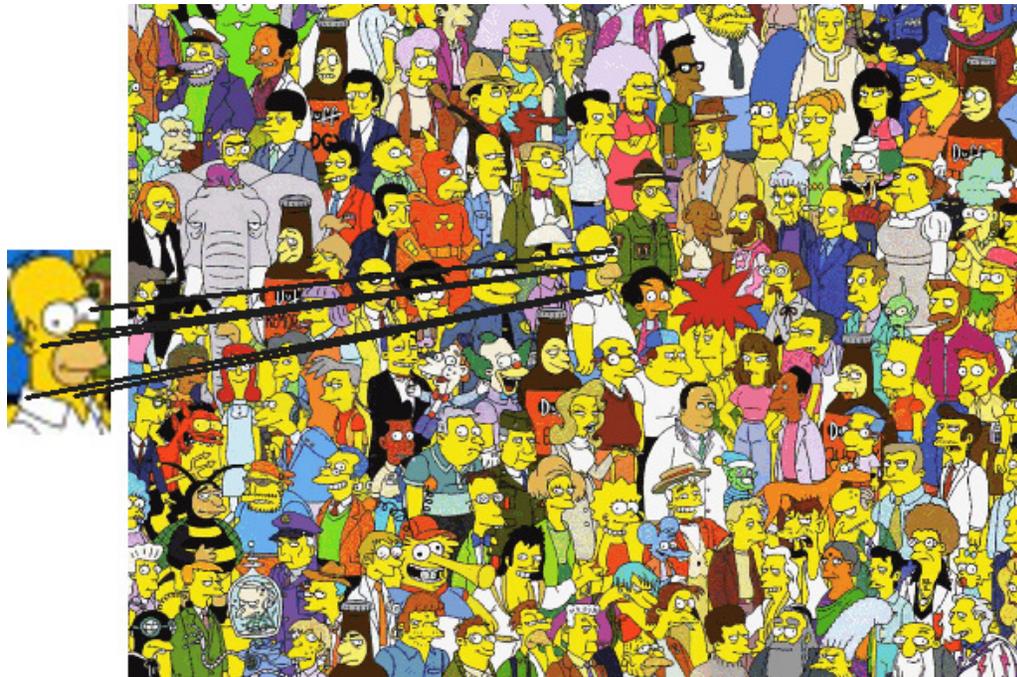


Ver: <http://www.impresee.com/>



Casos de Estudio

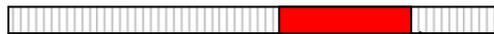
- Además de la ocurrencia, determinar la ubicación y pose del objeto encontrado.



Casos de Estudio

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

youtube



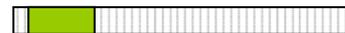
Cap.23



Cap.10



Cap.5



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

Casos de Estudio

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

Firefox - Evaluación de Shots
localhost:8080/ShotsEval/

List: 9071 -
Replay
Progress: 1/50
Shot ID: shot193_191
Next Shot

Topics: Check all (NO)

Topic 9069 - a circular "no smoking" logo	YES	NO
Topic 9070 - a small red obelisk	YES	NO
Topic 9071 - an Audi logo	YES	NO
Topic 9072 - a Metropolitan Police logo	YES	NO
Topic 9073 - this ceramic cat face	YES	NO
Topic 9074 - a cigarette	YES	NO
Topic 9075 - a SKOE can	YES	NO
Topic 9076 - this monochrome bust of Queen Victoria	YES	NO
Topic 9077 - this dog	YES	NO
Topic 9078 - a JENKINS logo	YES	NO
Topic 9079 - this CD stand in the market	YES	NO
Topic 9080 - this public phone booth	YES	NO
Topic 9081 - a black taxi	YES	NO
Topic 9082 - a BMW logo	YES	NO
Topic 9083 - a chrome and glass cafetiere	YES	NO
Topic 9084 - this man	YES	NO
Topic 9085 - this David refrigerator magnet	YES	NO
Topic 9086 - these scales	YES	NO
Topic 9087 - a VW logo	YES	NO
Topic 9088 - Tamwar	YES	NO
Topic 9089 - this pendant	YES	NO
Topic 9090 - this wooden bench with rounded arms	YES	NO
Topic 9091 - a Kathy's menu with stripes	YES	NO
Topic 9092 - this man	YES	NO
Topic 9093 - these turnstiles	YES	NO
Topic 9094 - a tomato-shaped ketchup dispenser	YES	NO
Topic 9095 - a green public trash can	YES	NO
Topic 9096 - Aunt Sal	YES	NO
Topic 9097 - these checkerboard spheres	YES	NO
Topic 9098 - a P (parking automat) sign	YES	NO

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 complejos en videos sin metadatos ni etiquetas.

Ej: Buscar videos donde aparezca un vehículo estacionando

The screenshot shows a web browser window with the URL `localhost:9080/EventsEval/`. The interface is titled "Evaluación de Eventos" and displays a video player on the left showing a street scene with several cars. On the right, there is a control panel with the following information:

- Event: 37
- Replay button
- Progress: 2/30
- Video ID: 776704
- Score: 0.98949
- Next Video button

Below the video player, there is a table of events for evaluation:

Event	YES	NO
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 - Playing fetch	YES	NO
One or more people navigate a vehicle into a designated space.	YES	NO

At the bottom of the interface, there are two thumbnail images: one showing a car parked in a lot, and another showing an airplane on a tarmac.



Temas a estudiar en el semestre

- Parte 1: Descripción de contenido multimedia
 - Métodos para analizar el 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 novedosas de descripción y búsqueda usando Codebooks y Deep Learning.



Evaluación (2017)

- Mini-Controles al inicio de clases
 - Una pregunta escrita, 15-20 minutos, sin apuntes.
- Tareas de Programación
 - Lenguaje C, C++, Python o Java.
- Presentación de Lecturas
 - A elección entre opciones dadas.
- Proyecto Final (grupos de dos personas)
 - Desarrollo de algún problema a elección.
 - Presentación de resultados en fecha del examen.

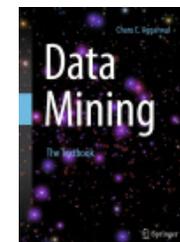
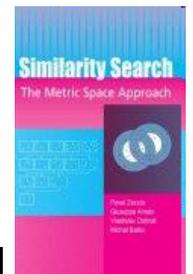
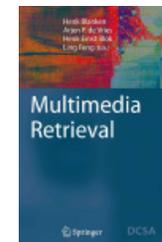
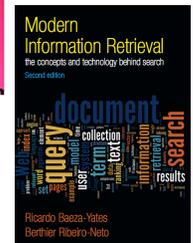
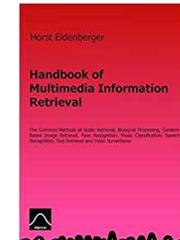
Evaluación (2017)

- 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

Sitio web del curso: <http://juan.cl/cc5213-2017/>

Bibliografía

- **Handbook of Multimedia Information Retrieval.** Eidenberger. 2012.
- **Modern Information Retrieval.** Baeza-Yates, Ribeiro-Neto, 2011.
- **Multimedia Retrieval.** Blanken, de Vries, Blok, Feng. 2007.
- **Similarity Search The Metric Space Approach.** Zezula, Amato, Dohnal, Batko. 2006.
- **Data Mining: The Textbook.** Aggarwal. 2015.



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.
- **Mining of Massive Datasets.** Leskovec, Rajaraman, Ullman. 2014.
- **Computer Vision. Algorithms and Applications.** Szeliski. 2011.
- Muchos artículos de investigación....

