

CC5204

**Búsqueda por Contenido
de Imágenes y Videos**

Juan Manuel Barrios

<http://juan.cl/>

Departamento de Ciencias de la Computación
Universidad de Chile
Semestre Otoño 2016



Datos del Curso

- Áreas involucradas:
 - Computer Vision, Multimedia Information Retrieval, Data Mining, Pattern Recognition.
- Requisitos:
 - Diseño y Análisis de Algoritmos (CC4102)
 - Diferentes algoritmos y estructuras de datos.
 - Programación de Software de Sistemas (CC3301)
 - Conocimientos de programación en C/C++ (obsoleto)
- Líneas de especialización:
(<http://www.dcc.uchile.cl/node/238>)
 - Ingeniería de Datos, Información y Conocimiento.
 - Inteligencia Computacional y Robótica.
 - Ciencia e Ingeniería Computacional.

Nuevas Aplicaciones

Qualcomm presenta una plataforma que permite reconocer personas en tiempo real

La aplicación móvil Shazam permitirá a sus usuarios identificar objetos

Google anuncia servicio de almacenamiento ilimitado de fotos y el futuro de Android

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

Las máquinas con grandes capacidades de aprendizaje son el próximo desafío de la tecnología

Los beneficios de este tipo de avances son importantes, pero también generan un riesgo de que reemplacen al hombre en sus actuales ocupaciones.

DPA

sábado, 26 de diciembre de 2015 13:19

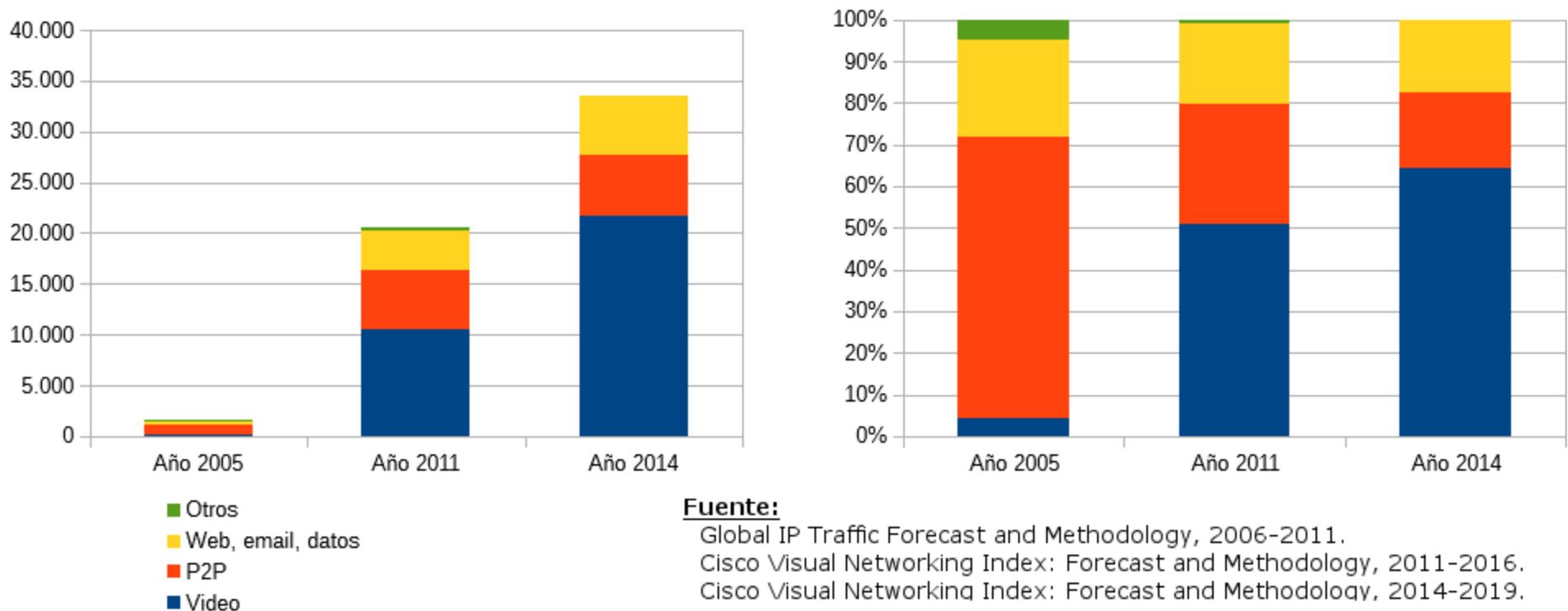


BERLÍN.- Los computadores son cada vez más inteligentes, reconocen idiomas y escrituras, objetos en fotos, proponen respuestas adecuadas para e-mails, traducen, informan del estado

Evolución del uso de Internet

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

Tráfico de datos de usuarios de Internet [PB mensual]



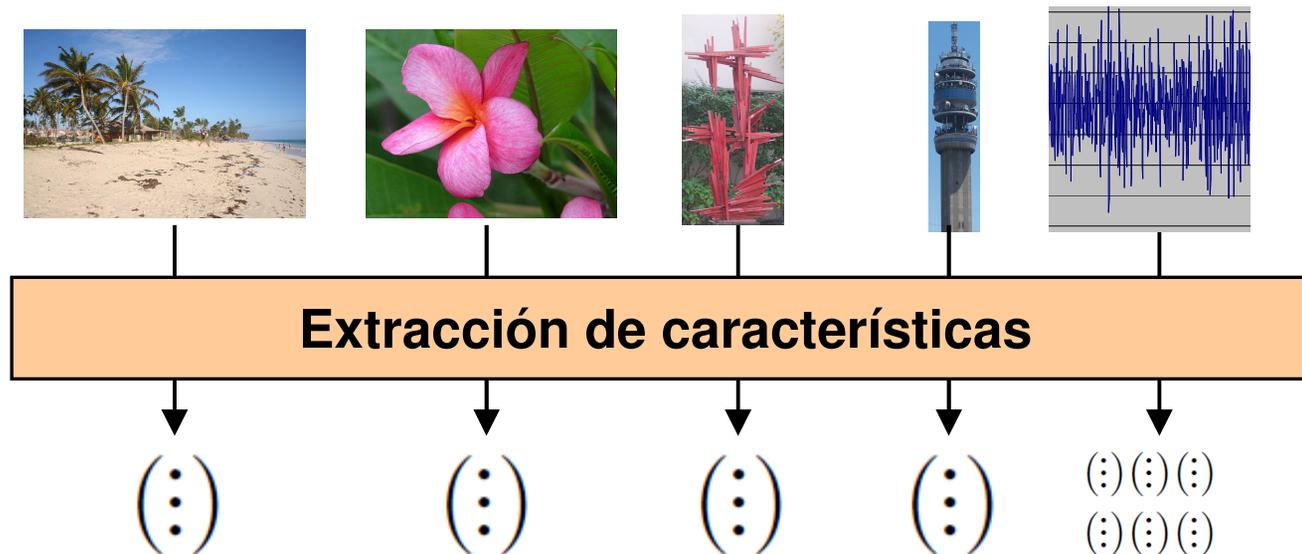


¿Cómo hacer uso de estos datos?

- Algoritmos para procesar datos multimedia.
- Dos etapas principales:
 - Descripción del Contenido
 - Búsqueda por Similitud
- Minería de datos sobre grandes volúmenes de datos multimedia.

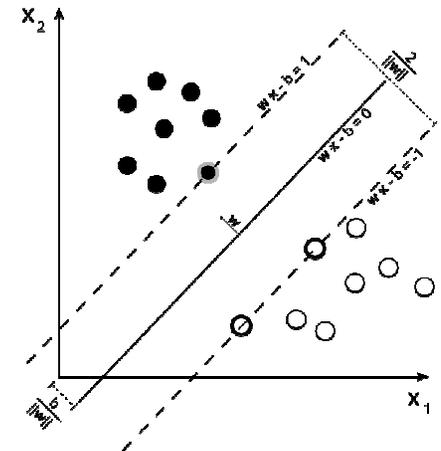
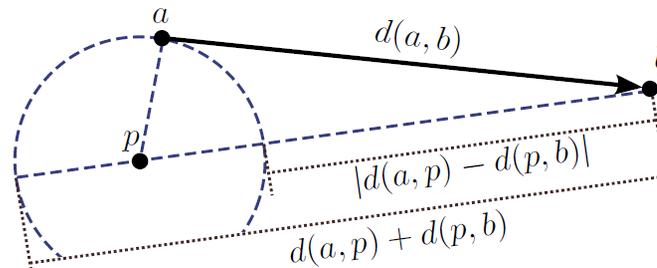
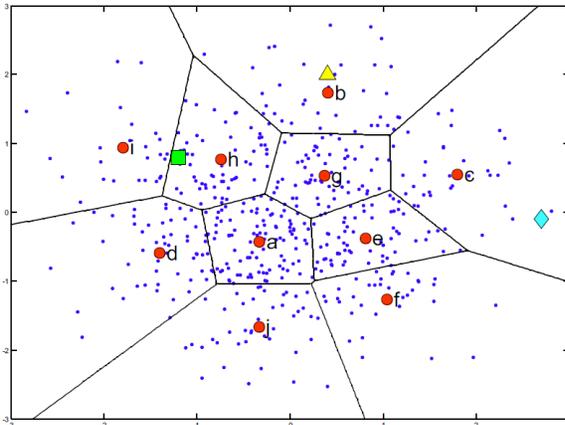
Descripción del contenido

- Representar una foto, una canción, un video, un email, etc. por uno o más vectores.



Búsqueda por similitud

- Resolver eficientemente búsquedas en espacios vectoriales y métricos.
- Algoritmos para análisis automático de millones de elementos.





Problemas a estudiar

- Se estudiarán soluciones para problemas como:
 - Búsqueda por similitud de imágenes.
 - Etiquetado automático de imágenes y videos.
 - Búsqueda de videos duplicados.
 - Clasificación automática de canciones.
 - Detección de eventos en videos.
 - Minería de Bases de Datos de videos.

Caso de Estudio #1

- Dada una imagen de consulta, buscar más imágenes parecidas según algún criterio (colores, formas, etc.)

Imagen de
consulta



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

Caso de Estudio #1

- Dada una imagen de consulta, buscar más imágenes parecidas según algún criterio (colores, formas, etc.)

Resultados



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

Caso de Estudio #2

- Dado un catálogo de objetos, reconocer la ocurrencia de un objeto.



Imagen de consulta



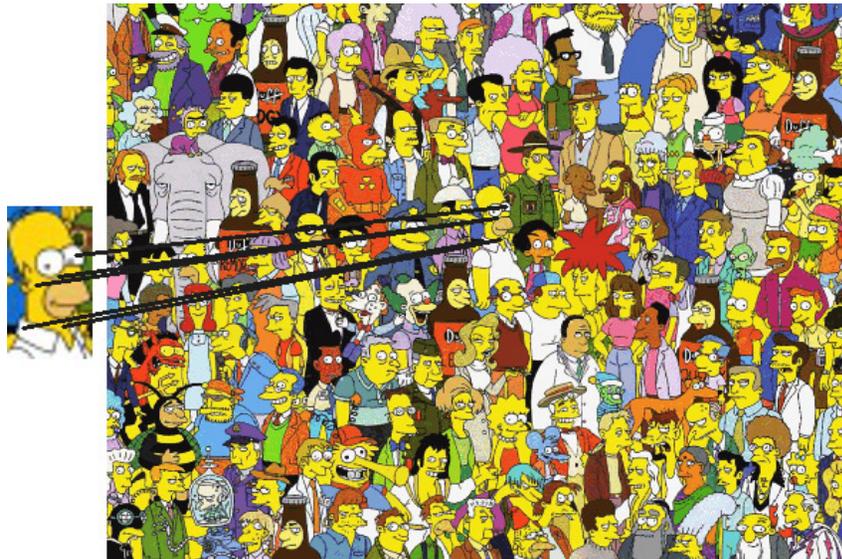
Ver: <http://api.impresee.com/DemoSearch/>

A screenshot of a mobile application interface titled 'Demo Product Search'. It shows a query image of a hand holding an 'afe' bottle. Below the query image, the text 'Query (2165):' is displayed. A table lists search results with columns for Product, ID, and Score.

Product	ID	Score
	648707	55
	648677	47
		

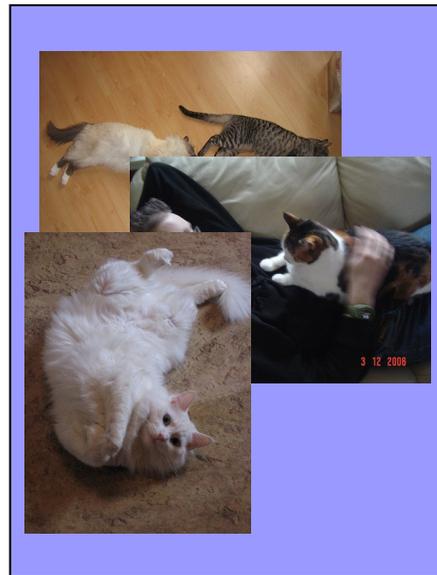
Caso de Estudio #2

- Dado un catálogo de objetos, reconocer la ocurrencia de un objeto.
 - Determinar la ubicación y pose del objeto detectado.



Caso de Estudio #3

- Usar un conjunto de entrenamiento para etiquetar automáticamente fotos y videos.



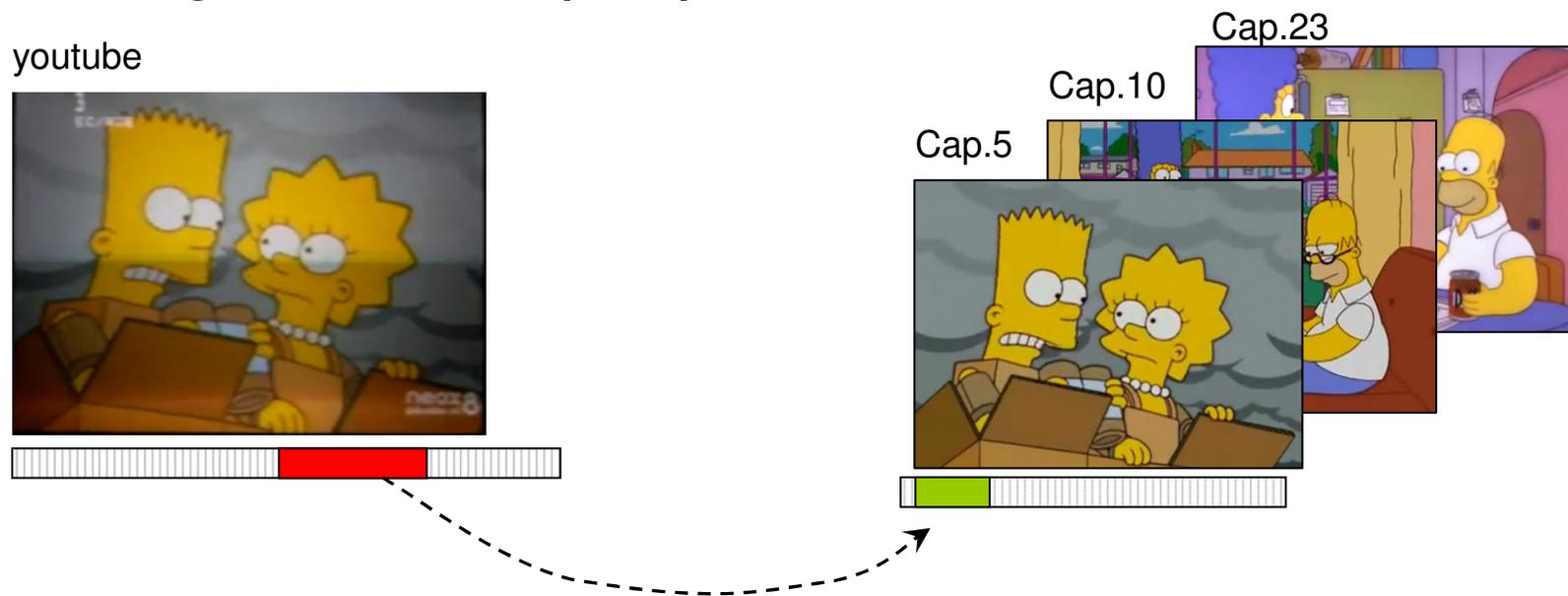
¿Etiqueta?

Datos de entrenamiento de “aviones”, “gatos” y “bicicletas”.

Ver: <http://demo.caffe.berkeleyvision.org/>

Caso de Estudio #4

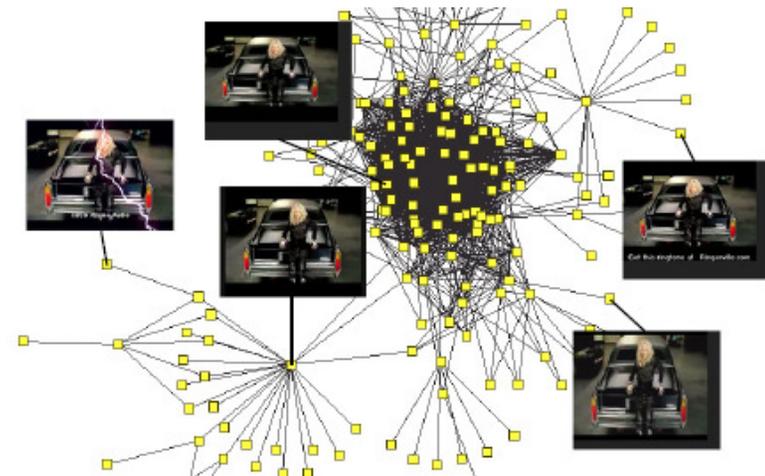
- Minería de Bases de Datos de Videos.
 - Dado un video de Internet determinar la escena original de la que proviene.



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

Caso de Estudio #4

- Minería de Bases de Datos de Videos
 - Detección de comerciales, recomendación de programas, búsqueda de información relacionada, etc.

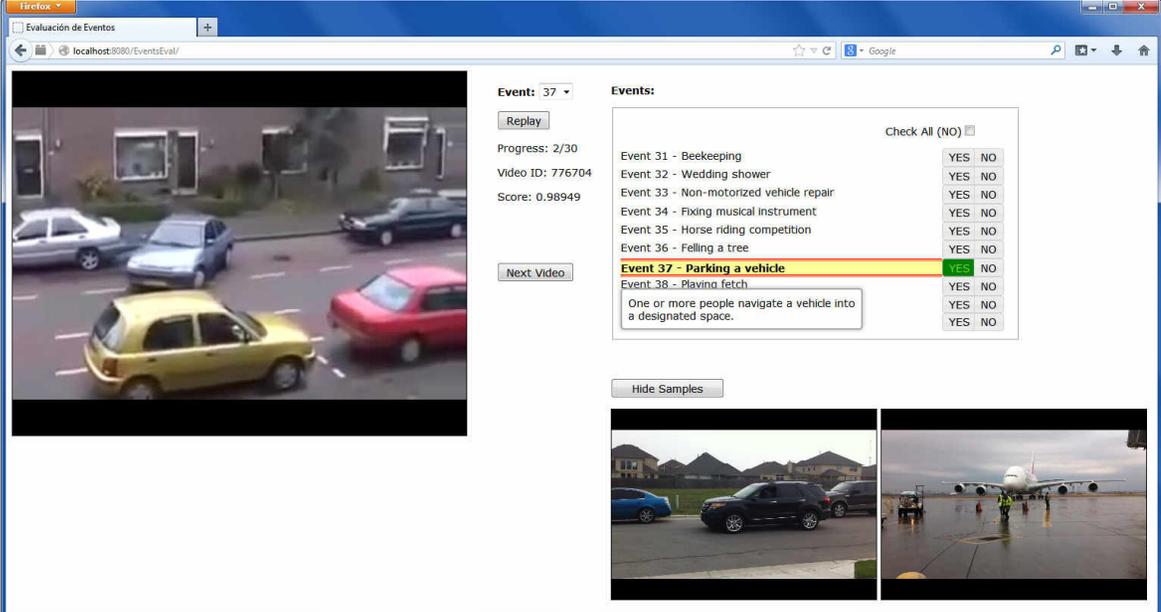


Ver: <http://www.robots.ox.ac.uk/~vgg/research/vgoogle/>

Caso de Estudio #4

- Minería de Bases de Datos de Videos
 - Búsqueda de eventos audiovisuales, sin usar tags ni metadatos.

Ej: Buscar videos donde aparezca un vehículo estacionando.



The screenshot shows a web application interface for video evaluation. On the left, a video player displays 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, there are two small video thumbnails: one showing a car in a parking lot and another showing an airplane on a tarmac.



Temas a estudiar en el semestre

- Parte 1: Descripción del contenido
 - Descripción automática de imágenes, audio, videos y texto.
- Parte 2: Búsqueda por similitud
 - Índices para espacios vectoriales y métricos.
- Parte 3: Aplicaciones
 - Clasificación de imágenes y etiquetado automático.
 - Búsqueda de objetos conocidos en videos.
 - Otras técnicas como Clustering, Deep Learning, Reconocimiento de personas, etc.



Actividades

- Mini-Controles al inicio de clases (7)
 - Una pregunta, 15-20 minutos, sin apuntes.
- Tareas de Programación (3)
 - Lenguajes C/C++/Python/Java.
 - Entregar Informe en PDF y Código fuente por U-cursos.
- Presentación de Lecturas (2)
 - A elección entre opciones dadas.
- Proyecto de Investigación
 - Desarrollo de algún problema a elección.
 - Presentación de resultados en fecha del examen.
 - No es necesario entregar código fuente.
- Las Tareas, Presentación de Lecturas y Proyecto se pueden realizar en grupos de hasta dos personas.

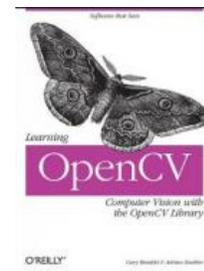
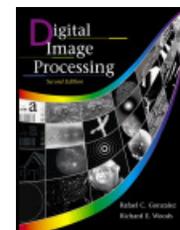
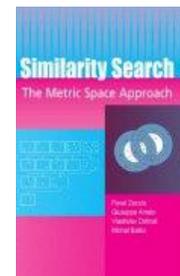


Reglas del Juego (2016)

- Carga académica: 10 UD
- Horario: Martes y Jueves 12:00 – 13:30
- Evaluaciones:
 - **NC**=Siete mini-controles, se borra uno.
 - $NC = (C1 + \dots + C7 - \text{Min}\{C1, \dots, C7\}) / 6$
 - **NL**=Dos presentaciones de lecturas.
 - $NL = (L1 + L2) / 2$
 - **NT**=Tres tareas de programación con código fuente (F) e informe (I).
 - $NT = (F1 + I2 + F2 + I3 + F3) / 5$
 - **NP**=Proyecto personal, presentación inicial (PI) y final (PF).
 - $NP = 0.1 * PI + 0.9 * PF$
- 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

- **Multimedia Retrieval.** Blanken, de Vries, Blok, Feng. 2007.
- **Similarity Search The Metric Space Approach.** Zezula, Amato, Dohnal, Batko. 2006.
- **Digital Image Processing.** Gonzalez, Woods. 2002.
- **Learning OpenCV.** Bradski, Kaehler. 2008.



Bibliografía

- **The Essential Guide to Image/Video Processing.** Bovik. 2009.
- **H.264 and MPEG-4 Video Compression.** Richardson, 2003.
- **Computer Vision: Algorithms and Applications.** Szeliski, 2010.
- **Mining of Massive Datasets.** Leskovec, Rajaraman, Ullman, 2014.
- **Revista Bits DCC, N°8-“Multimedia”, pág. 68.**
<http://www.dcc.uchile.cl/bits-de-ciencia>

