

## TECNOLOGÍA

## Bits, Ciencia y Sociedad



## Jun 15 Navegando en el mar de datos multimedia

Posteado a las 16:52 | 0 comentario

Por Benjamín Bustos, profesor del Departamento de Ciencias de la Computación, FCFM, U. de Chile.



La cantidad de información multimedia (por ejemplo, imágenes, audio y video) en la Web ha tenido un crecimiento muy importante en los últimos años. Actualmente, se estima que más del 95% de los datos almacenados en la Web corresponde a contenido multimedia, lo cual se debe entre otras razones al incremento del poder computacional que permite manejar este tipo complejo de información, a una mayor disponibilidad de ancho de banda en Internet (que permite por ejemplo ver videos en línea), y a la masificación de dispositivos de captura como cámaras digitales y smartphones. Hoy en día, ¡casi todos nosotros somos productores de datos multimedia! A esto se le debe agregar el auge de las redes sociales (Facebook, Twitter) y de sitios especializados para compartir imágenes y video (YouTube, Flickr, Picasa).

Un problema que surge a partir de la disponibilidad de estos grandes repositorios de datos multimedia, en particular imágenes, es cómo buscar y navegar en este mar gigante de datos. No sólo son datos complejos (una imagen es una matriz de píxeles), sino que además ocupan mucho espacio (usualmente un par de megabytes por cada imagen). Además, en la Web una imagen no solamente se muestra y describe por su contenido, la matriz de píxeles, sino que puede haber mucha información contextual a su alrededor: descripciones y anotaciones en forma de palabras claves o tags, de comentarios puestos por el creador de la imagen o sus amistades, información geográfica, temporal, etc. Toda la información del contexto de la imagen se puede recuperar e indexar para facilitar búsquedas textuales de las imágenes (como en Google Images). Este método tiene algunas desventajas, por ejemplo pueden haber muchas imágenes sin ninguna anotación (pregúntese a cuántas de sus fotos de las vacaciones le agregó tags), las anotaciones son por lo general subjetivas y varían entre distintos usuarios (salvo datos que puedan extraerse en forma automática, como fecha de la foto y posición geográfica), y por lo general son incompletas (muchos se acordarán del dicho "una imagen vale más que mil palabras").

Otro enfoque que permite realizar búsquedas sin utilizar información contextual se basa en usar el contenido mismo de la imagen para realizarla. Existen variadas técnicas para analizar en forma automática los colores, bordes, texturas y formas en una imagen, y usando esta información se puede realizar la búsqueda. En este caso, en vez de colocar una o varias palabras como términos de consulta (por ejemplo, escribir "torre Eiffel" en el buscador) se coloca una imagen de consulta (por ejemplo, una foto de la torre Eiffel), la cual es analizada en su contenido y luego se realiza la búsqueda con dicha información. Los resultados usando este tipo de búsqueda pueden ser bastante buenos, como lo muestra el siguiente ejemplo tomado desde el Buscador de Imágenes PRISMA, desarrollado en el DCC (<http://prisma.dcc.uchile.cl/ImageSearch/>), la imagen que se busca está a la izquierda, los resultados se muestran a la derecha):



Sin embargo, el problema es bastante difícil y está lejos de estar resuelto. Una de las grandes dificultades es el llamado "Gap Semántico": existe una brecha entre la información que uno puede capturar de los píxeles de la imagen y su verdadero significado semántico. Por ejemplo, a pesar de que el computador es capaz de analizar en detalle los

colores, bordes, texturas y formas en una imagen, no puede hacer distinción entre las siguientes imágenes (fuente: Google):



Actualmente, para disminuir la brecha semántica en la búsqueda de imágenes en la Web se están proponiendo varias alternativas. Una es tratar de combinar información contextual y de contenido en la búsqueda, de manera que ambas fuentes de información se complementen. Aún no está claro cuál es la mejor forma de realizar esta combinación, y si bien para aplicaciones específicas es posible lograr resultados realmente notables (por ejemplo, ya es fácil detectar todas las apariciones de un mismo objeto en una película), aún falta mucho para resolver el problema en forma general.

Tags: benjamin bustos, datos multimedia



PERMALINK

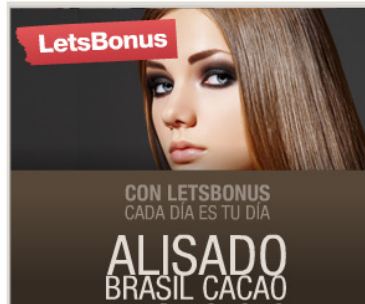


BLOGS QUE CITAN ESTE POSTEO

comentarios 0 comentario

COMENTAR

POSTEO MÁS ANTIGUO ▶



## perfil del autor



El blog Bits, Ciencia y Sociedad de la sección de Tecnología de Terra es un espacio donde académicos del Departamento de Ciencias de la Computación de la Universidad de Chile hablarán de la Tecnología y su impacto

político y social en nuestro país".

Aquí escribirán semanalmente José Miguel Piquer, Claudio Gutiérrez, Juan Álvarez, Tomás Barros, Nancy Hirschfeld, Benjamin Bustos, Alejandro Hevia, Pablo Barceló y Cecilia Bastarica."

## posteos

VER: MÁS RECIENTES MÁS COMENTADOS

## A 40 años de la primera carrera de Ingeniería del área de Computación en Chile

8 de Junio de 2011 - 14:25

## Valiant, o cómo seguir siendo renacentista hoy en día

1 de Junio de 2011 - 12:30

## La química de la comunicación

25 de Mayo de 2011 - 15:19

## ¿Para qué sirven las universidades estatales?

17 de Mayo de 2011 - 15:23

BUSCAR