

# Buran



Número 25 - Año 17 - Marzo 2010 - 1.50 €

<http://ieee.upc.es>

---

El uso de señales de oportunidad GPS para observación de la Tierra

---

Security System Based on Suspicious Behavior Detection

---

GOS: Búsqueda virtual de imágenes

---

High Altitude Platform Stations in Design Solutions for Emergency Services

---

Secure Voting From your Living Room



**IEEE**

Barcelona IEEE Student Branch

# EDITORIAL

Bienvenido querido lector a esta nueva entrega de Buran, que podemos asegurar, está encantada de saludarte de nuevo. Han pasado tres años desde el último número, y mucho ha llovido desde entonces, pero podemos asegurar que las ganas de hacer cosas, y nuestra convicción en llevarlas a cabo, no se han perdido.

Desde el equipo de redacción de Buran, integrado por la Rama de estudiantes en su totalidad, queremos demostrar que las nuevas generaciones vienen con ganas de mejorar, afrontando este nuevo reto con nuestra mayor ilusión.

Hace ya un año que nuestro equipo tomó las riendas. Éramos un pequeño grupo inexperto de colaboradores, estudiantes de la ETSETB de la UPC, que buscábamos una oportunidad de llevar a cabo alguna actividad más allá de lo convencional. Encontramos la Rama del IEEE y comprobamos que nuestro granito de arena se podía convertir en una aportación más a la sección del conocimiento que llamamos progreso.

Con inquietud, trabajo y la impagable colaboración de un sinfín de personas, entre las que se encuentran: Ángel Cardama Aznar, Margarita Cabrera Bean, Elisa Sayrol Clols, Ricardo Blasco Serrano, Guillem Hernández Sola, Daniel Navarro Giménez y todos y cada uno de los articulistas que han escrito en esta publicación, hemos llegado hasta aquí, hasta vuestras manos.

No es sólo Buran lo que queremos volveros a presentar, sino el futuro elenco de actividades que verán la luz el año que viene, tales como la celebración de los 30 años de la Rama de estudiantes de Barcelona, la organización de conferencias para estudiantes y la organización de salidas tecnológicas, eventos a los que esperamos, podáis asistir.

Como nota para la posteridad quedará que esta fue la primera del largo número de publicaciones que verán la luz en el futuro, y que fue la primera por la que empezamos a sentirnos orgullosos de pertenecer a la Rama de Estudiantes del IEEE de Barcelona.

Carlos Ciller Ruiz  
*Presidente de la Rama*

## COORDINACIÓN BARCELONA

Sergio Segura García

## EDICIÓN BARCELONA

David Carmona Torondel  
Carlos Ciller Ruiz  
Llorenç Garcia Casas  
Ana García del Molino  
Manuel Muñoz Fuentes  
Raúl Onrubia Ibáñez  
Sergio Segura García

## REVISIÓN

Llorenç Garcia Casas  
Sergio Segura García  
Ramón Llorca Queralt

## DISEÑO PORTADA

Sergio Segura García

## AGRADECIMIENTOS

Ángel Cardama  
Elisa Sayrol Clols  
Margarita Cabrera Bean

## IMPRESIÓN

Canigó, S.L

## DEPÓSITO LEGAL // ISSN

B-19.950-96 // 1698-7047

La organización se reserva el derecho de publicar los artículos. La opinión expresada en los artículos no tiene por qué coincidir con la de la organización.

Agradecemos las colaboraciones hechas desinteresadamente, y a causa de la falta de espacio, pedimos disculpas a todas aquellas personas a las cuales no se les ha publicado su colaboración. Esperamos que en un próximo número tengan cabida.

# SUMARIO

## 2 Editorial

## 3 Sumario

## 4 Ramas de Estudiantes del IEEE

### 5 El uso de señales de oportunidad GPS para observación de la Tierra

*A. Camps, X. Bosch, I. Ramos, J.F. Marchán, N. Rodríguez, y E. Valencia.*

### 11 Security System Based on Suspicious Behavior Detection

*Enrique Bermejo, Oscar Déniz and Gloria Bueno.*

### 16 Sistema de Reconstrucción de Escenas 3D Paralelizado en GPU para su Aplicación en Tiempo Real

*Enrique Oriol, Jordi Salvador y Josep R. Casas.*

### 25 High Altitude Platform Stations in Design Solutions for Emergency Services.

*Israel R. Palma-Lázgare, José A. Delgado-Penín.*

### 30 Criptografía Basada en Atributos

*Javier Herranz.*

### 36 GOS: búsqueda visual de imágenes

*Silvia Cortés Yuste, Xavier Giró i Nieto y Ferran Marqués Acosta.*

### 45 El software libre y lo sostenible

*Rafael Cubarsi y Miguel Escudero.*

### 49 Hombres escondidos en fórmulas

*Miquel Escudero.*

### 50 Secure Voting From Your Living Room

*David Andreu, Alex Escala, Guillem Caldúch.*

### 56 Searching Representative Phrases in a Musical Score using Fuzzy Logic

*Emerson Castañeda.*

*Buran se fundó en Barcelona en la ETSETB de la UPC, en marzo de 1993.  
Se revisa por la Rama de Estudiantes del IEEE de Barcelona y imprime y edita en Barcelona.*





# RAMAS DE ESTUDIANTES DEL IEEE

## ¿Qué es el IEEE?

IEEE son las siglas de *Institute of Electrical and Electronics Engineers*, organización técnica y profesional de ámbito mundial dedicada a dar soporte a la teoría y aplicaciones de la ingeniería eléctrica, electrónica y de la informática. Desde su fundación, en el año 1884, el IEEE se ha ido expandiendo por más de 150 países en todo el mundo y ha alcanzado un número de asociados de más de 320.000, de los cuales unos 50.000 son estudiantes. El IEEE es actualmente la sociedad técnica con mayor número de socios del mundo, dando cabida a ingenieros, informáticos, físicos y matemáticos.

El IEEE promueve el desarrollo de todos los campos relacionados con la tecnología, la informática y la ingeniería, tanto en sus vertientes teóricas como prácticas. Para facilitar este cometido se divide en sociedades especializadas en los sectores más novedosos en el panorama tecnológico actual. Además de publicar alrededor del 25% de las publicaciones técnicas a nivel mundial en los campos de ingeniería eléctrica, electrónica e informática, organiza conferencias, symposiums y encuentros locales y proporciona programas educacionales para mantener a sus miembros en la vanguardia de los avances tecnológicos y científicos. Con todo ello el IEEE beneficia a la sociedad, incide en la mejora de la profesión y aumenta el nivel de formación de sus asociados.

**La Sección Española del IEEE** pertenece a la Región 8 de las que se divide el organismo, y desde marzo de 2002 su junta

## ¿Qué son las Ramas de Estudiantes?

El IEEE promueve y patrocina más de 700 Ramas de Estudiantes en todo el mundo, dando la posibilidad de futuros profesionales de desarrollar las habilidades que necesitarán para moverse más cómodamente en el complejo mundo laboral actual. Ser miembro de las Ramas de Estudiantes, recibir las publicaciones periódicas del IEEE y colaborar en las actividades de cada Rama durante la época de estudiante proporciona información y experiencia en el estado actual de la tecnología y la industria en todo el mundo, además de poner en contacto a los estudiantes con ingenieros profesionales.

## Buran 25

Buran nació en el seno de la Rama de Estudiantes del IEEE de Barcelona, con el objetivo de divulgar temas tanto científicos como tecnológicos y humanísticos, siempre dentro de la filosofía del IEEE, y de ser un portavoz de las actividades docentes, de investigación o sociales que se producen en la Universidad.

En particular, Buran quería ser la revista para los estudiantes, y es por ello que, en este número nos sentimos orgullosos de que tanto profesores como compañeros nuestros colaboren con nosotros.

El próximo número de Buran saldrá en Abril o Mayo de 2010, así que desde aquí queremos animaros a escribir. Podéis obtener la información necesaria en el Call for Papers que se incluye en este número. También podeis escribirnos a: [buran@ieee.upc.edu](mailto:buran@ieee.upc.edu)

Muchas gracias por vuestro apoyo incondicional. Esperamos seguir mereciendo vuestra confianza.

## Para contactar con la Rama de Estudiantes de Barcelona

*Rama de Estudiantes del IEEE de Barcelona.*

Despacho S-105. Edificio Omega

Campus Nord de la UPC

Tel: 93 401 76 56

C/Jordi Girona, 1-3 s/n

[ieee@ieee.upc.es](mailto:ieee@ieee.upc.es)

08034 BARCELONA

<http://ieee.upc.es>

Información de Archivo de BURAN:



Codified with EAN-13

# El uso de señales de oportunidad GPS para observación de la Tierra

A. Camps, X. Bosch, I. Ramos, J.F. Marchán, N. Rodríguez, y E. Valencia

Remote Sensing Lab., Dept. de Teoria del Senyal i Comunicacions, Universitat Politècnica de Catalunya  
UPC Campus Nord, D4-016, 08034 Barcelona. E-mail: camps@tsc.upc.edu

## INTRODUCCIÓN

La falta de observaciones globales y frecuentes desde el espacio es actualmente un factor limitante en muchas misiones de Observación de la Tierra (EO: *Earth Observation*). Por ejemplo, el tsunami del Océano Índico el 26 de diciembre de 2004 pudo ser detectado casualmente por radares altímetros, pero la probabilidad de detección es de hecho muy baja ya que las señales del tsunami son muy débiles en el océano abierto (sólo cuando se acercan a la costa la altura de las olas crece) y el satélite debe sobrevolarlo casi simultáneamente [[http://earth.esa.int/brat/html/appli/geodesy/tsunami\\_en.html](http://earth.esa.int/brat/html/appli/geodesy/tsunami_en.html)]. Para poder estudiar las variaciones de mesoscala de éste y de otros fenómenos oceanográficos las futuras misiones espaciales deberán proporcionar una mejor cobertura espacial y temporal. Dos posibles técnicas que se han propuesto hoy son la utilización de constelaciones de satélites, y la utilización de señales de los satélites de navegación global (GNSS: *Global Navigation Satellite Signals*), como “señales de oportunidad”, es decir, señales que están ahí y que se utilizan para un propósito diferente de aquél para el que fueron concebidas [[http://earth.esa.int/brat/html/alti/future/welcome\\_en.html](http://earth.esa.int/brat/html/alti/future/welcome_en.html)].

Si bien el uso de constelaciones de satélites es hoy en día un hecho en sistemas de comunicación (p. ej. Inmarsat e Iridium) y navegación (p. ej. GPS, Glonass, y el futuro Galileo), su uso en observación de la tierra es mucho más limitado por evidentes restricciones económicas, y su uso sólo se llegará a generalizar cuando el coste unitario de cada satélite incluyendo su lanzamiento disminuya suficientemente, lo cuál pasa por reducir su tamaño, peso, y el de las cargas útiles que llevan.

Por otra parte, las señales GNSS son omnipresentes, están perfectamente caracterizadas, y seguirán existiendo durante las próximas décadas, lo que las hace candidatas idóneas para sistemas de EO. Aunque el uso de las señales GPS reflejadas sobre la superficie del océano como señales de oportunidad fue originalmente propuesto ya en 1993 para aplicaciones altimétricas [1], no fue hasta el año 2002 que dichas señales se detectaron desde el espacio [2,3]. Hoy en día, su uso no sólo se limita a altimetría, sino que también se ha estudiado su aplicabilidad a la determinación del estado del mar o a la medida del viento sobre el mismo [4-10]. Más recien-

temente su uso se ha estudiado a la corrección directa (sin modelos intermedios) del estado del mar en las variaciones de la temperatura de brillo en banda L producidas por los cambios de estado del mar [11], lo cuál es de vital importancia si se quiere mejorar la calidad de las recuperaciones de la salinidad superficial en las misiones continuación de SMOS [<http://www.esa.int/esaLP/LPsmos.html>], en la que la Universitat Politècnica de Catalunya (UPC) ha participado activamente junto con el Instituto de Ciencias del Mar (ICM-CSIC) de Barcelona. Asimismo, recientemente se ha propuesto el uso de las señales GNSS reflejadas sobre la superficie del suelo para medir la humedad superficial y la altura de la capa de vegetación, caso de que la hubiera [12-14].

Este artículo describe brevemente las actividades del Remote Sensing Lab de la UPC en el uso de las técnicas GNSS-R (*GNSS-Reflectometry*) para la monitorización del océano y de la tierra, incluyendo una posible implementación como carga útil de un pico-/nano-satélite.

## I. USO DE SEÑALES GNSS-R PARA APLICACIONES OCEANOGRÁFICAS

La recuperación de la salinidad superficial del mar se puede llevar a cabo por medio de radiómetros de microondas en banda L (1.400-1.427 MHz) [15]. Los radiómetros de microondas son receptores de radio muy sensibles que miden la potencia de ruido captada por una antena en una banda que, en principio, está libre de interferencias electromagnéticas. Esta potencia es equivalente al “volumen” del nivel de ruido captado por el radiómetro, y está relacionada con la llamada “temperatura de brillo” (TB), la cuál depende de la temperatura y de la constante dieléctrica del agua del mar, la cuál a su vez depende de la temperatura y la salinidad. Sin embargo, la TB también depende de la “rugosidad” de la superficie del mar, que no puede ser parametrizada únicamente en términos de la velocidad del viento, la altura significativa de las olas, o de cualquier otro parámetro disponible actualmente. Por ello, a pesar de los experimentos de campo realizados en los últimos años para mejorar nuestra comprensión de este efecto [16], la corrección de la rugosidad de la superficie del mar sigue siendo una de las correcciones necesarias más importantes para poder recuperar la salinidad con la precisión requerida.



En la actualidad, está previsto lanzar dos misiones espaciales para la medida de la salinidad oceánica:

- la misión SMOS de la ESA [<http://www.esa.int/esaLP/LPsmos.html>], con el radiómetro MIRAS de apertura sintética, en forma de Y, y
- la misión AQUARIUS/SAC-D de NASA/CONAE, usando un radiómetro tipo “push-broom”.

En la misión SMOS, se espera que las capacidades de observación multiangular permitan la recuperación no sólo de la salinidad, sino también de un viento “efectivo” que minimice los errores entre los modelos y las observaciones. Por otra parte, la misión AQUARIUS dispone de un dispersómetro en banda L, para realizar dichas correcciones.



Fig. 1a) Visión artística de la misión SMOS de la ESA. Credits ESA - AOES Medialab [<http://www.esa.int/esa-mmg/mmgpl?b=b&keyword=SMOS&start=3>]

En el año 2003 se propuso a la European Science Foundation el desarrollo de unos instrumentos (PAU: *Passive Advanced Unit for ocean monitoring*) para el uso de las señales de oportunidad GNSS-R para realizar la corrección del estado del mar en la temperatura de brillo, proyecto que fue financiado para el período 2/2005-2/2010. El concepto es simple y puede aplicarse a cualquier señales GNSS: GPS, GLONASS, Galileo... Los satélites GPS se transmiten señales electromagnéticas polarizadas circularmente a derechas (RHCP) a la frecuencia de 1575.42 MHz. Al dispersarse sobre la superficie del mar, se convierten mayoritariamente en polarización elíptica a izquierdas. Para un mar en calma (perfectamente plano), la dispersión de la señal proviene de un único punto de reflexión especular, determinado por la distancia más corta entre la transmisión de satélite GPS y el receptor. Sin embargo, cuando el mar está agitado, la

dispersión de las señales proviene de una región más amplia que crece cuanto más agitado está el mar, de la misma manera que el Sol reflejándose en el mar es una especie de elipse cuyo tamaño crece con el estado del mar (Figs. 2a, 2b).

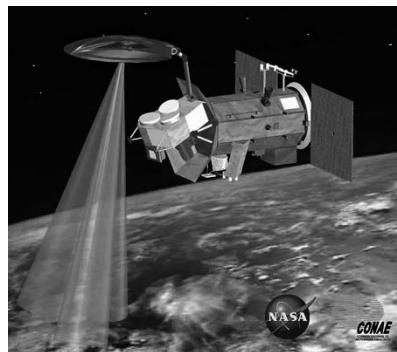


Fig. 1b) Visión artística de la misión AQUARIUS/SAC-D de NASA/CONAE [<http://aquarius.nasa.gov/>]

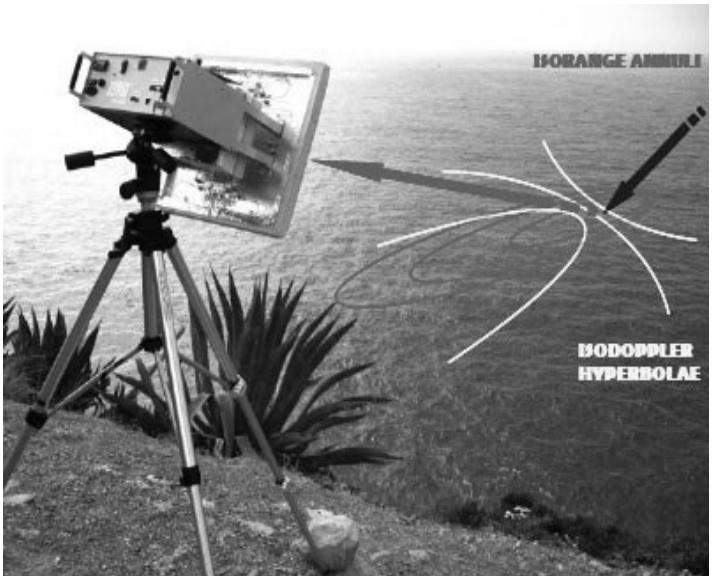


Fig. 2. a) Reflejo del Sol sobre mar en calma, b) Reflejo del Sol sobre una mar agitado. La zona donde se produce dispersión de la luz del Sol está relacionada con el estado del mar. c) Las señales GNSS ofrecen la posibilidad de ser utilizadas como señales de oportunidad para la determinación del estado del mar y aplicaciones altimétricas mediante la medición de la extensión de la zona de dónde llega dispersión.

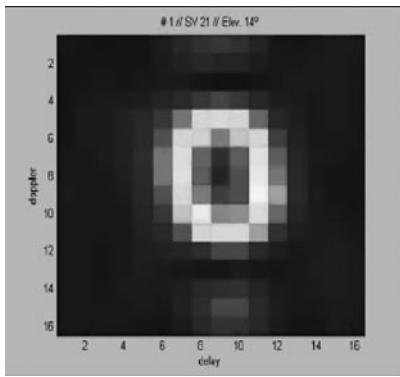


Fig. 2.d) Ejemplo de DDM medido sobre la superficie del mar.

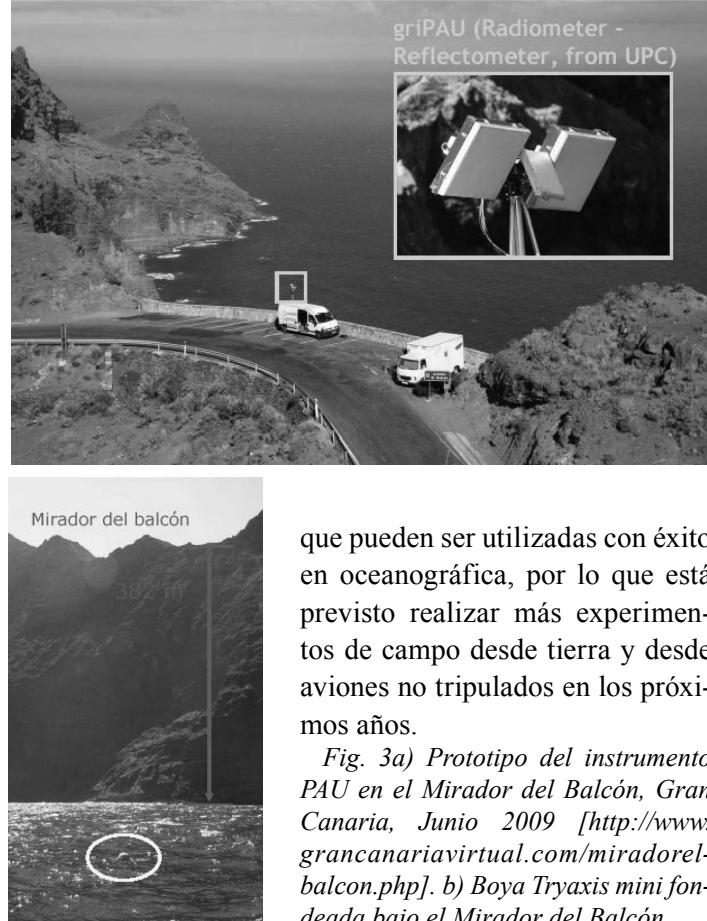
Esta zona puede medirse a partir de la correlación de la señal recibida con una réplica centrada a diferentes frecuencias Doppler (hipérbolas de Doppler constante) y retardos de tiempo diferentes (elipses de retardo constante), como se muestra en la Fig. 2c. Por lo tanto, cada uno de los puntos sobre la superficie tiene un retardo y un Doppler determinados (en realidad, hay dos puntos con el mismo retardo y Doppler). En radar, la función de ambigüedad proporciona una medida de la similitud entre una señal y una versión retardada de la misma que puede incluir un efecto Doppler. En reflectometría GNSS se utiliza el llamado Delay-Doppler Map (DDM, Fig. 2d) que es equivalente a la función de ambigüedad en radar. Una de las principales ventajas de las técnicas GNSS-R radica en la capacidad de obtener información del estado del mar desde muchos puntos simultáneamente (tantos como puntos de reflexión sobre el mar provenientes de diferentes satélites), lo cuál no es posible con altímetros convencionales, por ejemplo. Esto repercute en un menor tiempo de revisita, y por tanto datos globales más frecuentes.

El concepto del instrumento PAU es, pues, sencillo y se compone de: 1) un radiómetro en banda L (PAU-RAD) y 2) un reflectómetro GPS capaz de medir DDMs en tiempo real (PAU-GNSSR), y 3) PAU-IR un radiómetro de infrarrojos para la medida de la temperatura superficial del mar. Tanto PAU-RAD como PAU-GNSSR comparten el mismo front-end de microondas, simplificando el diseño de radio frecuencia.

Durante Junio 2008 y 2009, versiones preliminares del instrumento PAU se desplegaron en la costa Noroeste de Gran Canaria (Fig. 3a) durante las campañas ALBATROSS (Advanced L-BAnd emissiviTy and Reflectivity Observations of the Sea Surface) para medir por vez primera vez la TB en banda L y DDMs, junto con datos oceanográficos: temperatura superficial del mar la superficie del mar, espectro direccional boyas... (Fig. 3b).

Se puede decir que en la actualidad, empezamos a com-

prender la relación entre el estado del mar y los observables GNSS-R (DDMs) y los cambios en el brillo de la temperatura. Sin embargo, la ciencia subyacente todavía necesita mejoras significativas para poder extraer cantidades físicas

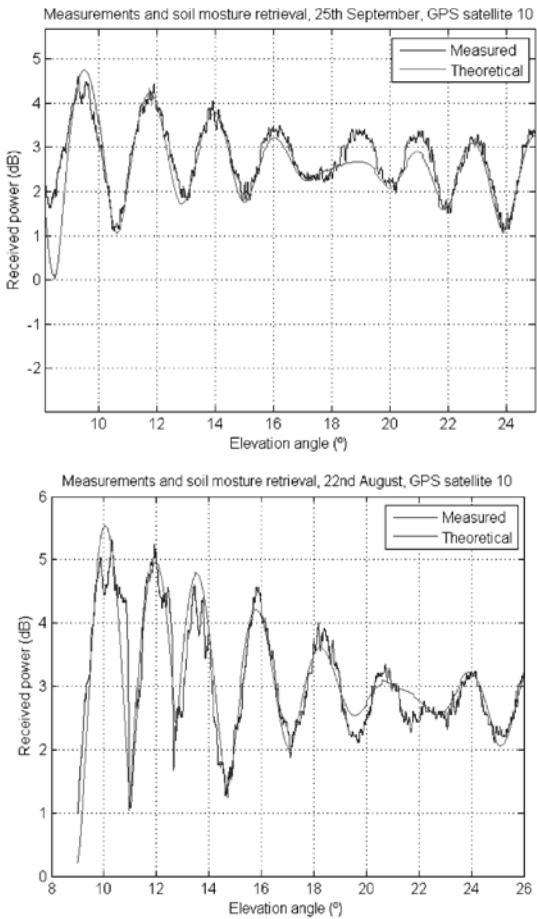


que pueden ser utilizadas con éxito en oceanográfica, por lo que está previsto realizar más experimentos de campo desde tierra y desde aviones no tripulados en los próximos años.

Fig. 3a) Prototipo del instrumento PAU en el Mirador del Balcón, Gran Canaria, Junio 2009 [<http://www.grancanariavirtual.com/miradorel-balcon.php>]. b) Boya Tryaxis mini fonda debada bajo el Mirador del Balcón

## II. USO DE SEÑALES GNSS-R PARA APLICACIONES TERRESTRES

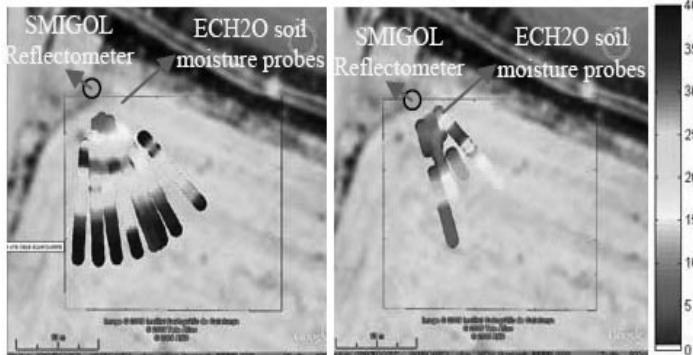
Para aplicaciones terrestres se ha desarrollado una técnica muy sencilla, también basada en el uso de señales GNSS-R, denominada Técnica del Patrón de Interferencia (IPT: *Interferente Pattern Technique*). Ésta consiste en estudiar las fluctuaciones de la potencia recibida por un receptor GPS comercial orientado de tal manera que recibe simultáneamente las señales directa del satélite, y la reflejada en el suelo. Cuando el suelo está húmedo, la reflectividad es mayor y se produce un notch (nulo) en las fluctuaciones de la potencia recibida para ángulos de elevación menores (ángulos de incidencia mayores). La profundidad del notch aumenta para terrenos más lisos ya que hay menos dispersión y la amplitud de la señal reflejada es mayor, y disminuye para terrenos más rugosos ya que hay más dispersión y la amplitud de la señal reflejada es menor. La Fig. 4 presenta dos medidas de la fluctuación de la potencia recibida para un día a) húmedo, y b) seco, obtenidas en Palau d'Anglesola,



Lleida.

*Fig. 4. Modelo teórico y resultados experimentales obtenidos en Palau d'Anglesola, Lleida, España el a) 10 de Septiembre de 2008 (terreno húmedo), y b) 22 de Agosto de 2008 (terreno seco), para el satélite GPS número 10.*

De esta manera, a medida que los diferentes satélites GPS se van desplazando, describen trazas sobre el suelo que permiten determinar después de unas horas un mapa completo de humedad del terreno (Fig. 5). La extensión (y la resolución espacial) de este mapa vendrá determinada por la altura del reflectómetro sobre el suelo, pero ésta en ningún caso podrá superar los ~150 m, a fin de garantizar que tanto la señal directa, como la reflejada se reciben dentro del tiempo



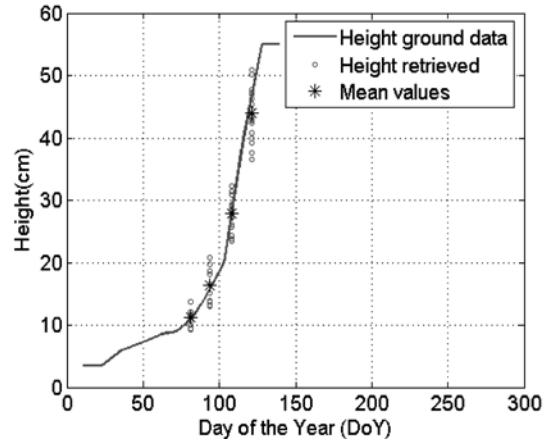
de coherencia de la señal GPS (tiempo de chip = 1 ms).

*Fig. 5. Mapas de humedad del terreno obtenidos con el reflectómetro SMIGOL (Soil Moisture Interference-pattern GNSS Observations at L-band) correspondientes al 25 de Septiembre (izquierda) y 22 de Agosto (derecha).*

Cuando existe una cubierta vegetal, se producen reflexiones múltiples en la interfaz vegetación-suelo, que producen más de un notch. Su número exacto y posición (ángulo de elevación) dependen de la altura de la vegetación en términos de la longitud de onda, que en este caso es fija, lo que permite inferir la altura de la misma. Por supuesto, para que esto sea posible, hay que suponer que la altura es aproximadamente la misma para todos los puntos de la traza del satélite, lo cual es bastante razonable en el caso de los cultivos. La Fig. 6 muestra el radiómetro LAURA (L-band AUTomatic RAdiometer) utilizado anteriormente en numerosas campañas preparatorias de la misión SMOS [<http://www.tsc.upc.edu/prs>], junto con SMIGOL, durante la campaña GRAJO (GPS and Radiometric Joint Observations) en Vadillo de la Guareña, Zamora, donde tendrán lugar actividades de la calibración / validación de SMOS.



*Fig. 6. LAURA y SMIGOL desplegados en la campaña GRAJO 2008-2010 (Zamora).*



*Fig. 7. Crecimiento de la cebada durante el experimento GRAJO en 2009. Recuperaciones de altura de la vegetación los días 22 de Marzo, 4 de Abril, 18 de Abril y 1 de Mayo, valores medios y valores medidos in-situ (ground-truth).*

La Fig. 7 muestra los resultados obtenidos en 4 días: 22 de Marzo, 4 de Abril, 18 de Abril y 1 de Mayo de 2009 para diferentes recuperaciones con diferentes satélites GPS, así como el valor medio de los mismos, y el ground-truth medido in situ por le CIALE/Universidad de Salamanca. Como se puede observar, la correspondencia entre los resultados es excelente.

### III. CONCLUSIONES Y LÍNEAS FUTURAS

En este artículo se han descrito brevemente las dos técnicas desarrolladas en la UPC para utilizar las señales GNSS-R como señales de oportunidad. Para aplicaciones oceanográficas se ha diseñado un reflectómetro GNSS-R que es capaz de generar DDMs completos y complejos cada 1 ms, y de integrarlos coherente y/o incoherente durante el tiempo que se haya predeterminado. Se espera que este observable sea más robusto que la simple correlación en retraso, ya que es insensible a muchos efectos acimutales. En la actualidad se está trabajando en el procesado de los datos de la campaña ALBATROSS 09, para correlacionarlo con los cambios instantáneos en TB. Para aplicaciones terrestres se ha desarrollado el Interferente Pattern Technique, técnica realmente sencilla y potente a la vez, que permite, a partir del análisis de la evolución temporal de las fluctuaciones de la potencia de la señal interferente (directa + reflejada) obtener la información de humedad superficial del suelo y altura de la vegetación.

Las líneas futuras de trabajo que se están siguiendo en el grupo abarcan:

1) en el campo de las aplicaciones terrestres (humedad del terreno y altura de la vegetación), la automatización de todo el procesado, de tal manera que dichas recuperaciones las puedan realizar usuarios no expertos. Así mismo, el desarrollo de sensores miniatura, alimentados con paneles solares y con enlaces radio de medio-largo alcance, para su mejor desarrollo en el campo.

2) en el campo de la oceanografía operacional se pone de manifiesto que son necesarios datos obtenidos desde satélite, ya que hasta la fecha, sólo algunos segundos de datos sobre mar, hielo y tierra están disponibles del satélite UK-DMC, y éstos son netamente mejorables en cuanto a relación señal a ruido.

Los recientes avances tecnológicos en sensores permiten hoy en día llevar a cabo misiones espaciales con satélites de “bajo coste”, que pueden desarrollarse en uno o dos años.

Un aspecto prometedor son las perspectivas de las constelaciones de nano-satélites por una fracción del coste de una sola misión de microsatélites. En este sentido, la National Science Foundation de los EE.UU. ha iniciado ya un programa (NSF 08-549, 09-523) para financiar este tipo de misiones [<http://www.nsf.gov/pubs/2008/nsf08549/nsf08549.htm>], basándose en el standard CubeSat, un picosatélite de 10 cm x 10 cm x 10 cm, que pesa menos de 1 kg. Y aunque hoy en día todavía no sea factible integrar un reflectómetro GNSS completo en un picosatélite, sí que parece posible (jaunque difícil!) en el hermano mayor de CubeSat que mide 10 cm x 10 cm x 30 cm, y que pesa menos de 3 kg. Estas estructuras están disponibles comercialmente, así como algunos de los subsistemas requeridos, como el ordenador de a bordo, transmisor/receptor, la fuente de alimentación, los paneles solares, el control de actitud etc., y otros nuevos como paneles solares desplegables, star-trackers, sistemas de propulsión ... estarán disponibles en breve (Fig. 8). En el Remote Sensing Lab se está trabajando en la actualidad en el desarrollo de la tecnología de estos pico-satélites, y en paralelo en el desarrollo de payloads GNSS-R que pudieran ser embarcadas en los mismos.

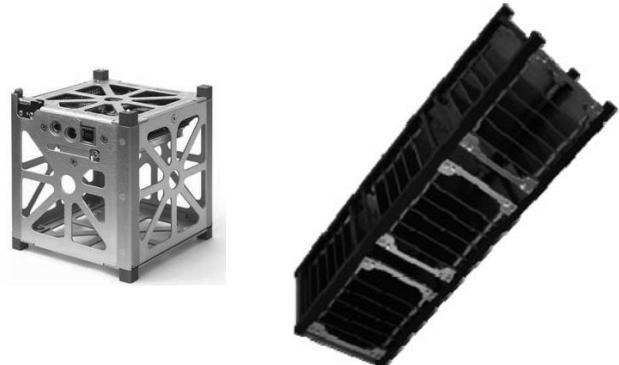


Fig. 8. Modelos comerciales de CubeSat de a) 1 unidad (10 x 10 x 10 cm) y de b) 3 unidades (10 x 10 x 30 cm)

### REFERENCIAS

- [1] Martín-Neira, M., A Passive Reflectometry and Interferometry System (PARIS): Application to Ocean Altimetry. *ESA Journal* 1993, vol. 17. pp 331-355.
- [2] Lowe, Stephen T.; LaBrecque, John L.; Zuffada, Cinzia; Romans, Larry J.; Young, Larry E.; Hajj, George A., First spaceborne observation of an Earth-reflected GPS signal, *Radio Science*, Vol. 37, No. 1, 07 February 2002.
- [3] Gleason, S.; Hodgart, S.; Yiping Sun; Gommen-

ginger, C.; Mackin, S.; Adjrad, M.; Unwin, M.; Detection and Processing of bistatically reflected GPS signals from low Earth orbit for the purpose of ocean remote sensing; IEEE Transactions on Geoscience and Remote Sensing, Vol. 43 (6), pp. 1229 – 1241, June 2005, DOI 10.1109/TGRS.2005.845643

[4] Soulat, F., M. Caparrini, O. Germain, P. Lopez-Dekker, M. Taani, G. Ruffini, Sea state monitoring using coastal GNSS-R Geophys. Res. Lett., Vol. 31 (21)

[5] Garrison, J.L., Katzberg, J.L., Effects of sea roughness on bistatically scattered range coded signals from the Global Positioning System, GRL, Vol. 25 (13), 1998.

[6] Komjathy, A., V. Zavorotny, P. Axelrad, G.H. Born, and J.L. Garrison, GPS signal scattering from sea surface: Wind speed retrieval using experimental data and theoretical model. Rem. Sens. Env., 73:162–174, 2000.

[7] Garrison, J.L.; Komjathy, A.; Zavorotny, V.U.; Katzberg, S.J.; Wind speed measurement using forward scattered GPS signals; IEEE Transactions on Geoscience and Remote Sensing, Vol. 40 (1), pp. 50 – 65, January 2002, DOI 10.1109/36.981349

[8] Cardellach, E., G. Ruffini , D. Pino , A. Rius , A. Komjathy , and J.L. Garrison, Mediterranean Balloon Experiment: ocean wind speed sensing from the stratosphere, using GPS reflections Remote Sensing of Environment Vol. 88 (3) , pp. 351-362 , 15 December 2003,

[9] Rius, A., J.M. Aparicio, E. Cardellach, M. Martín-Neira, and B. Chapron, “Sea surface state measured using GPS reflected signals”, Geophys. Res. Lett., 29(23), 2122, 2002.

[10] Marchan-Hernandez, J. F., N. Rodriguez-Alvarez, A. Camps, I. Ramos-Perez, E. Valencia, X. Bosch-Lluis, “Ground-Based GNSS-R Measurements with the PAU Instrument and Their Application to the Sea Surface Salinity Retrieval: First Results,” International Geoscience and Remote Sensing Symposium 2008, IGARSS 2008, 7-11 July, 2008

[11] Valencia, E., J.F. Marchan-Hernandez, A. Camps, N. Rodriguez-Alvarez, J. Miguel Tarongi, M. Piles, I. Ramos-Perez, X. Bosch-Lluis, M. Vall-llossera, P. Ferré, “Experimental Relationship Between the Sea Brightness Temperature Changes and the GNSS-R Delay-Doppler Maps: Preliminary Results Of The Albatross Field Experiments,” International Geoscience and Remote Sensing Symposium

2009, IGARSS 2009, 13-17 July, 2009

[12] Rodríguez-Álvarez, N., J.F. Marchán, A. Camps, E. Valencia, X. Bosch-Lluis, I. Ramos-Pérez, J.M. Nieto, “Soil Moisture Retrieval Using GNSS-R Techniques: Measurement Campaign In A Wheat Field,” International Geoscience and Remote Sensing Symposium 2008, IGARSS 2008, 7-11 July, 2008

[13] Rodriguez-Alvarez, N., J. F. Marchan-Hernandez, A. Camps, X. Bosch-Lluis, E. Valencia, I. Ramos-Perez, M. Vall-llossera, S. Monerris, J. Martinez-Fernandez, C. Perez-Gutierrez, G. Baroncini-Turricchia, N. Sanchez-Martin, J.M. Nieto, “Topographic Profile Retrieval using the Interference Pattern GNSS-R Technique,” International Geoscience and Remote Sensing Symposium 2009, IGARSS 2009, 13-17 July, 2009

[14] Rodriguez-Alvarez, N., S. Monerris, X. Bosch-Lluis, A. Camps, J. F. Marchan-Hernandez, I. Ramos-Perez, E. Valencia, M. Vall-Llossera, J. Martinez-Fernandez, N. Sanchez-Martin, G. Baroncini-Turricchia, C. Perez-Gutierrez, “Soil Moisture and Vegetation Height Retrieval using GNSS-R Techniques,” International Geoscience and Remote Sensing Symposium 2009, IGARSS 2009, 13-17 July, 2009

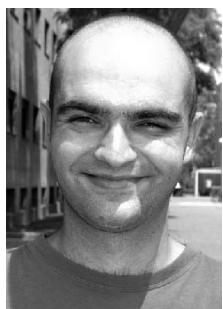
[15] Swift, C. T., and R. E. McIntosh, Considerations for microwave remote sensing of ocean surface salinity, IEEE Transactions on Geoscience and Remote Sensing., Vol. 21 (4), pp. 480–491, April 1983.

[16] Camps, A., J. Font, M. Vall-llossera, C. Gabarró, I. Corbella, N. Duffo, F. Torres, S. Blanch, A. Aguasca, R. Villarino, L. Enrique, J. Miranda, J. Arenas, A. Julià, J. Etcheto, V. Caselles, A. Weill, J. Boutin, S. Contardo, R. Niclós, R. Rivas, S.C.Reising, P. Wursteisen, M. Berger, and M. Martín-Neira, The WISE 2000 and 2001 field experiments in support of the SMOS Mission: Sea Surface L-Band Brightness Temperature Observations And Their Application to Multi-Angular Salinity Retrieval, IEEE Transactions on Geoscience and Remote Sensing, Vol. 42 (4), pp. 804-823, April 2004

## AUTHORS



Adriano Camps was born in Barcelona, Spain, in 1969. He received the degree in telecommunications engineering and Ph.D. degree in telecommunications engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1992 and 1996, respectively. In 1991 to 1992, he was at the ENS des Télécommunications de Bretagne, France, with an Erasmus Fellowship. Since 1993, he has been with the Electromagnetics and Photonics Engineering Group, Department of Signal Theory and Communications, UPC, where he was first an Assistant Professor, an Associate Professor in 1997, and where he has been a Full Professor since 2007. In 1999, he was on sabbatical leave at the Microwave Remote Sensing Laboratory, of the University of Massachusetts, Amherst. Since 1993, he has been deeply involved in the European Space Agency SMOS Earth Explorer Mission, from the instrument and algorithmic points of view, performing field experiments, and more recently studying the use of GNSS-R to perform the sea state correction needed to retrieve salinity from radiometric observations. His research interests are focused in microwave remote sensing, with special emphasis in microwave radiometry by aperture synthesis techniques and remote sensing using signals of opportunity (GNSS-R).



Xavier Bosch was born in Palau d'Anglesola, Spain. He received the degree in Telecommunication Engineering from the Universitat Politècnica de Catalunya (UPC) in May 2005, and after that he joined the Passive Remote Sensing Group of the Signal Theory and Communications (TSC) department of the UPC. His current research activities involve primary hardware (L band radiometer), developing calibration algorithms and characterization analysis for digital radiometers, new digital radiometric concepts such as digital beamforming or polarization synthesis and, he is also involved in algorithms development for a light radiometer aboard a radio-control aircraft.

He is finishing his Electronic Engineering degree. He is IEEE Student Member since 2004.



Isaac Ramos Pérez was born in Barcelona, Spain. He received the degree in Telecommunication Engineering from the Universitat Politècnica de Catalunya

(UPC) in May 2005, and after that he joined the Pasive Remote Sensing Group of the Signal Theory and Communications (TSC) department of the UPC. His primary areas of research include L band real aperture radiometry and L band synthetic aperture radiometry for sea salinity measures.



Juan Fernando Marchán-Hernández (S'04) was born in Barcelona, Spain. He received the M.S. degree in telecommunications engineering from the Polytechnic University of Catalonia (UPC), Barcelona, in 2004, and his PhD in May, 2009. In 2003, he was with the Laboratory of Space Technology, Helsinki University of Technology (TKK), Espoo, Finland. His current research interests are radiometry and signal reflections of GNSS.



Nereida Rodríguez Álvarez was born in Barcelona, Spain. She received the degree in Telecommunication Engineering from the Universitat Politècnica de Catalunya (UPC) in February 2007, and after that she joined the Passive Remote Sensing Group of the Signal Theory and Communications (TSC) department of the UPC. Her current research activities involve the use of different Global Navigation Satellite Signal Reflectometry (GNSS-R) techniques in order to retrieve geophysical parameters from the land and sea surfaces, as soil moisture or surface roughness, using simulators and developed instruments. She is IEEE Student Member since 2007.



Enric Valencia i Domènech was born in Sabadell, Catalunya. He received the degree in Electronic Engineering from the Universitat Politècnica de Catalunya (UPC) in April 2007, and after that he joined the Passive Remote Sensing Group of the Signal Theory and Communications (TSC) department of the UPC. His current research activities involve the use of different Global Navigation Satellite Signal Reflectometry (GNSS-R) techniques in order to retrieve geophysical parameters from the land and sea surfaces, as soil moisture or surface roughness, using simulators and developed instruments. He is IEEE Student Member since 2007.

# Security System Based on Suspicious Behavior Detection

Enrique Bermejo, Oscar Déniz and Gloria Bueno  
enric2186@gmail.com Universidad de Castilla-La Mancha, Spain.

## ABSTRACT

In recent years, the demand for image analysis applications of video surveillance has grown rapidly. The latest advances in video surveillance have aimed at automating the monitoring itself, so that it is a computer (not the security personnel) what observes the images and detects suspicious behavior or events. In this context, we present system for the automatic detection of suspicious behavior in public buildings, that obtains high resolution image of the individual or individuals who have activated the alarm in the system.

## 1. INTRODUCTION

Most current video surveillance systems share one characteristic: they need human operator to constantly watch monitors that show the images captured by the cameras [1]. The effectiveness of these systems is not determined by their technological characteristics but by the person who is monitoring the system [2]. Today, thanks to advances in many fields of computer vision, these systems are evolving to become virtually automatic. It is not human observer who detects suspicious situations, but algorithms that process the captured images and detect suspicious behavior or events [3].

The purpose of this article is to describe semiautomatic video surveillance system that is able to detect suspicious situations using artificial vision, as well as to facilitate the operator's work by generating visual and audible alerts. Most surveillance cameras installed today have static location and capture low quality images (we all have once seen on television news the images of an assault at shop where the action and the individual appears in small area of the image). Despite the high quality of the available tools for image processing, typical captured images are not very useful for crime investigation. The increasing need for security in public spaces makes real-time detection of suspicious behavior essential, rather than simply recording them [4].

This system aims to control the camera movement and zoom in order to obtain higher quality pictures for further investigations. The system relies on two physical compo-

nents: security camera to get pictures of the monitored environment and motorized Pan-Tilt.

The motorized Pan-Tilt allows the use of *patrol mode*. This patrol mode is sequence in which the camera remains static during configured time monitoring defined area/position. When this time is over, the system changes the position of the camera and re-starts the monitorization (of the new area). Figure 1 represents the patrol mode described.

Section 2 describes the image processing, behavior analysis and alerting methods developed. Then, Section 3 shows the results obtained with the system. Finally, Section 4 details the most important conclusions.

## 2. DEVELOPED SYSTEM

### 2.1 Image Processing

During the initial processing of the images obtained by the surveillance camera, the system performs a segmentation of objects in order to obtain a data set that provide the position and movement they make. In order to get the data set, we use back-ground subtraction techniques ([5], [6]) and motion analysis ([7]), using the OpenCV library [8].

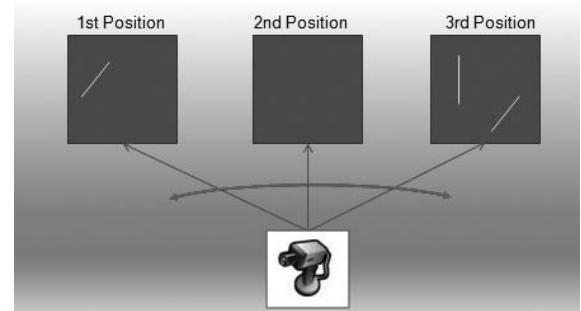


Figure 1. Patrol Mode Representation.

#### 2.1.1 Background Subtraction

For modeling and object detection the system uses the background subtraction method known as Running Average. This method involves the creation of a background model from the average of the  $n$  previous frames, as described

in [9]. Although this method is highly sensitive to changes in light and noise, it requires little memory and processing time, which is ideal for real-time operation.

In order to clean up the raw segmented image obtained during this process, the system realizes a connected-components analysis, in which it takes in a noisy image and then uses the morphological operations of opening and closing to eliminate noise and segment the objects that are large enough. After the analysis it is possible to find the contours of the objects and retrieve all contours of size above a defined threshold.

### 2.1.2 Motion Detection

After object segmentation, the system analyzes the motion generated during the sequence captured by the camera. The technique used is Motion Templates[8], which allows recording the movement generated by the displacement of the object's different shapes. This generates a motion history image (MHI) of the different silhouettes of the object in motion.

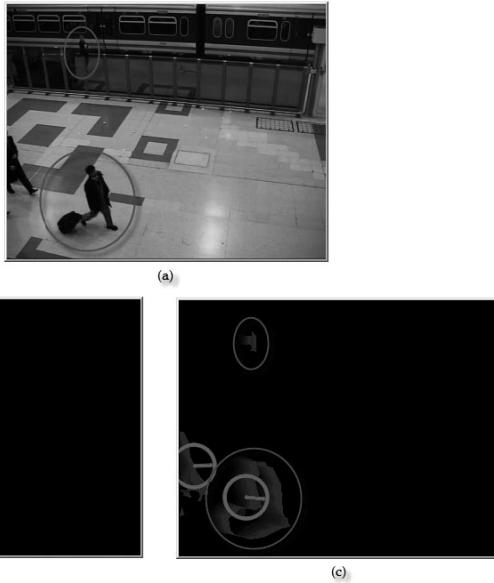


Figure 2. Motion direction (360x288). (a) Original Image.  
(b) Foreground. (c) MHI.

Thanks to the intensity information provided by the timestamps in the MHI, it is possible to calculate the motion gradient. That is, by analyzing the pixel values of the MHI the system is able to extract a vector indicating the flow of movement. This topic is described in detail in [10], with an analysis of the calculation of speed and orientation histograms.

This technique allows the system to obtain data that can be used for further behavioral analysis.

These data include:

1. Object's position.
2. Orientation angle.
3. Movement direction.

Figure 2 shows the results of background subtraction processes (b) and the motion template (c).

### 2.2 Behaviour Analysis

After obtaining the data referring to all objects that are in each frame, the system analyzes the data looking for suspicious actions such as:

- Trespassing of imaginary lines.
- Fights.
- Running people or Riots.

#### 2.2.1 Restricted Area Detection

The trespass detection consists of the constant evaluation of the moving objects that are detected from MHI. This makes use of simple planar geometry, namely the equations of the line, the relative positions and the intersection point between two lines.

The data that allow the system to make these checks are, first, the restricted crossing lines stored in a XML configuration file (created during system configuration), and second, the motion lines of the objects detected during the motion history processing.

The first step is to obtain information from both the configuration file and the history of motion, then estimate the line parameters and analyze, line by line, the possibility that one of the lines detected intersects with one of the restricted crossing lines. If there is any chance of intersection, the system checks whether the trespass direction is allowed or prohibited (this is also specified in the configuration). If it detects a forbidden trespass, it calculates the exact spot of intersection.

If the intersection point is within the detection limits (start point and end point) of the restricted line drawn, the motion detected is considered suspicious and, therefore, there is a trespass risk. In this case, the system generates an alert only when the moving object is at a distance less than a defined threshold. Moreover, with the information of the object that triggered the alarm, a region of interest (ROI) is established around this object and tracking is activated.

## 2.2.2 Races, Fights and Riot Detection

The detection of races, fights and riot consists of the evaluation of positions and motion angles of the moving objects. In order to do this, for each frame captured by the surveillance camera, the system calculates the change, from one frame to the next, the positions of every object and the change of motion angles associated with those objects. To perform these calculations we made use of the Hausdor metric.

The Hausdor metric measures the distance between two compact subsets of a metric space. This metric basically represents the maximum distance of a set of points with the closest point of another set. The formal definition of the Hausdor distance,  $d_H(X, Y)$ , is as follows:

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}$$

Where X and Y are two subsets of points in the metric space M.

This metric has many applications in computer vision and can be used to locate a particular template (pattern) in an image [11], in tracking and classifying objects, in comparison of 2D images with 3D objects. In this work we use it to obtain the degree of movement of an object in an image sequence. More specifically, the Hausdor distance is used for calculating the deviation of the positions and motion angles of the moving objects detected with the Motion Templates algorithm. The system uses a modified version of the Hausdor algorithm, where the values are normalized by calculating the average of the minimum distances obtained.

When the distance measured by this algorithm is small, it means that the positions and directions of movement components have not changed much from one frame to another. When the distance is large, the variation is abrupt and may be indicative of rapid linear movement and/or changing direction, which we considered indicative of possible run, riot or fight.

## 2.3 Alert Generator

In the case that suspicious activity is detected, either an trespass of prohibited areas, fight, riot or run, the system will manage a series of alarms in order to store a record of the action, and alert the operator of the system. Four types of alerts have been defined.

### 2.3.1 Visual Alerts

Visual alerts are divided into two groups. First we have real-time alerts, where a circle is drawn around the individual

who caused the suspicious action, as well as a line indicating its motion direction. These alerts are useful for the operator to pay attention to the area of interest in the images of the video sequence. Secondly, we have storage alerts, which are images in high resolution and with more zoom than the typical video capture. These images may be useful for further investigation.

### 2.3.2 Sound Alerts

The sound alerts consist of a brief sound reproduction that alerts the operator, so that he can focus on the monitor. The sound is played whenever the system detects a suspicious action.

### 2.3.3 Logs

Besides the generation of visual and audible alerts, when the system detects suspicious behavior it stores a record indicating the date and time, as well as the type of behavior that has been detected. This is useful because the date and type of action facilitate the reconstruction of events from the images stored as visual alerts.

### 2.3.4 Tracking

If the action detected is a trespass, the system enters a state in which tracks the object that generated the alert. During this period, the system estimates the new position of the object relative to the center of the image, performs the necessary movement of the Pan-Tilt to center the object in the image, and, then, increases the camera zoom and performs a high resolution capture of the image. Tracking is done via the mechanism of Camshift Tracking (Continuously Adaptive Mean-Shift) [12]. In Figure 3 it is shown a sequence of images obtained during the tracking process.

## 3. EXPERIMENTAL RESULTS

The images below show some of the results obtained during the operation of the monitoring system. For those examples we have used images obtained from a FireWire surveillance camera and a video from the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), that refers to a train station monitoring.

First, Figure 4 shows the result of a trespass (a), in which a man is crossing the imaginary line defined, the template of motion (b), the region of interest defined on the individual before start tracking on this individual (c) and, finally, the high-resolution image stored as a log for further investigation(d).



Figure 3: Tracking a moving object (512x384).

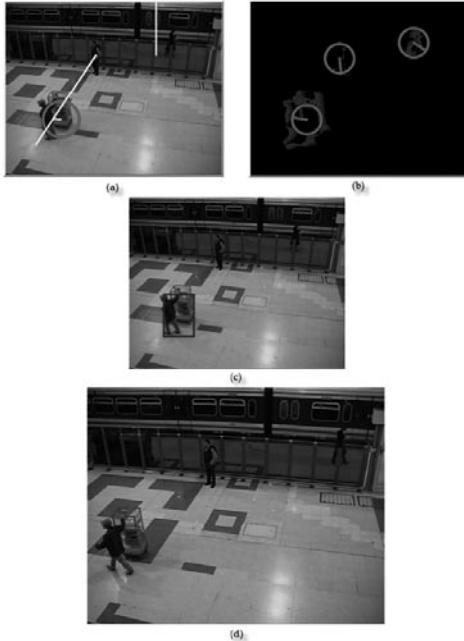


Figure 4. Trespass detected.(a),(b) y (c) (360x288). (d) (720x576).

In Figure 5 it is observed a trespass detected by the surveillance camera. Image (a) shows a snapshot of the moment in which the trespass occurs. The sequence of images (b) through (e) shows the zoom progress during tracking. After the tracking, the system also stores a high resolution image as it is shown in image (f). In this image we can observe that the individual appears at the center of the image.

Finally, Figures 6 and 7 show the results obtained using the Hausdorff distance for the detection of runs and ghting, respectively. As shown in the images, the system stores images as evi-

dence that the action has been identified as suspicious. In Figure 7 we can observe that the individuals who have activated the alarm appears at the center of the image with an increasing zoom.



Figure 5. (a) Overrun (512x384). (b), (c), (d) y (e). Tracking (512x384). (f) Detection Proof (1024x768).

## 4. CONCLUSION

We built a prototype of video surveillance hardware-software system to detect potentially dangerous events in real time and alert the human operator.

In addition, we have successfully fulfilled the goal of obtaining a view with more detail of the area or individual that has generated an alarm thanks to the pan-tilt-zoom mechanism.

The system deployment is very simple, since it enables real-time threshold setting that allow to modify the degree of detections made by the system during its execution, thus adapting the system to the conditions of the different surveillance areas.

## REFERENCES

- [1] M. Shah, O. Javed, and K. Shaque. Automated visual surveillance in realistic scenarios. IEEE Computer Society, 14:30-39, Jan.-March 2007.
- [2] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. Se-

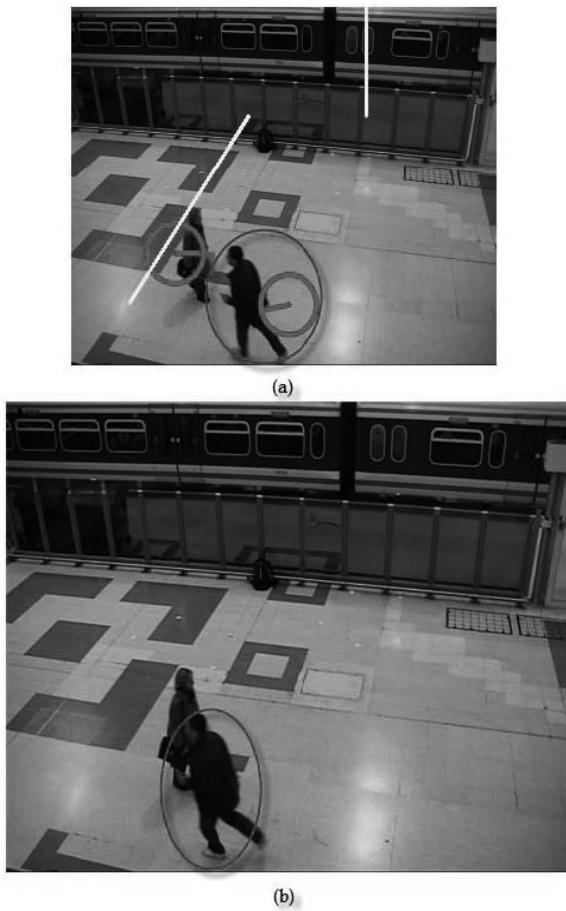


Figure 6. (a) Detected run (360x288). (b) Proof image (720x576).

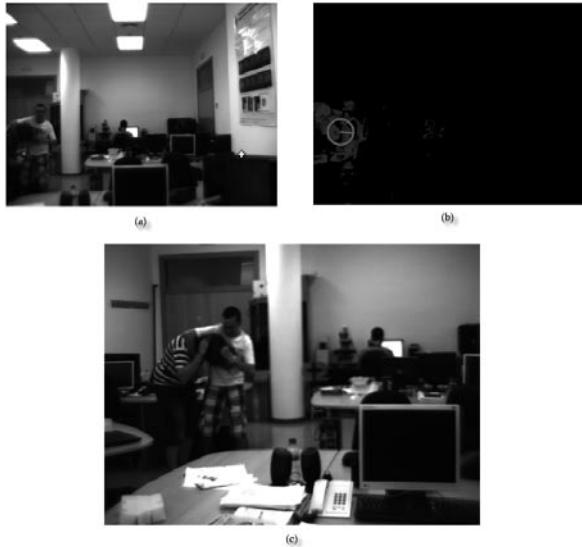


Figure 7. (a) Detected ght (512x384). (b) MHI (512x384).  
(c) Stored proof (1024x768).

nior, and Y. Tian. Smart video surveillance: Applications, technologies and implications. In IEEE Pacific-Rim Conference On Multimedia, Singapore, December, 2:1133-1138, 2003.

[3] V. Gouaillier and A. Fleurant. Intelligent video surveillance: Promises and challenges. Technological and Commercial Intelligence Report, March 2009.

[4] Anthony R.Dickand and Michael J.Brooks. Issues in

automated visual surveillance. School of Computer Science, University of Adelaide, 2003.

[5] Nathan Johnson. Background subtraction: Various methods for different inputs. CiteSeerX Scientific Literature Digital Library and Search Engine [<http://citeseerx.ist.psu.edu/oai2>] (United States), 2008.

[6] Sen-Ching S. Cheung and Chandrika Kamath. Robust techniques for background subtraction in urban traffic video. Video Communications and Signal Based Surveillance, IEEE Conference on, 0:30, 2006.

[12] Gary R. Bradski. Real time face and object tracking as a component of a perceptual user interface. Applications of Computer Vision, IEEE Workshop on, 0:214-219, 1998.

## AUTHORS



Enrique Bermejo: Degree in Computer Engineering. Ph.D. Student and associated researcher at Ingeniera de Sistemas y Automatica (ISA) department at E.T.S. Ingenieros Industriales, Universidad de Castilla-La Mancha, Spain.  
Enrique.Bermejo@uclm.es



Oscar Deniz: He received his MsC and PhD from Universidad de Las Palmas de Gran Canaria, Spain, in 1999 and 2006, respectively. He has been associate professor at Universidad de Las Palmas de Gran Canaria from 2003 to 2007 and currently at Universidad de Castilla-La Mancha, Spain. His main research interests are human-robot interaction and computer vision. He is a research fellow of the Institute of Intelligent Systems and Numerical Applications in Engineering and member of IEEE, AEPIA and AERFAI. Oscar.Deniz@uclm.es



Gloria Bueno: She received her MsC from Universidad Complutense de Madrid in 1993, and her PhD from Coventry University in 1998. From 1998 to 2000 Gloria worked as a postdoctoral researcher at Université Louis Pasteur, Strasbourg. In 2000-2001 she worked at CNRS-Institut de Physique Biologique-Hopital Civil and from 2001 to 2003 she was a senior researcher at CEIT (Centro de Estudios e Investigaciones Técnicas de Gipuzkoa), San Sebastián, Spain. She is currently an Associate Professor at Universidad de Castilla-La Mancha, Spain. Her main research interests include image processing particularly for biomedical engineering applications- computer vision, artificial intelligence, modeling and simulation. Gloria.Bueno@uclm.es

# Sistema de Reconstrucción de Escenas 3D Paralelizado en GPU para su Aplicación en Tiempo Real

Enrique Oriol, Jordi Salvador y Josep R. Casas

Image and Video Processing Group, Universidad Politécnica de Cataluña

E-mail: enriqueoriol@gmail.com,{jordi.salvador, josep.ramon.casas}@upc.edu

## ABSTRACT

Se presenta una adaptación de algoritmos clásicos de reconstrucción 3D en escenarios multi-cámara a un entorno multiprocesador que permite aprovechar la enorme capacidad de cálculo de los nuevos procesadores gráficos (GPUs), con el objetivo de lograr *frame rates* adecuados para aplicaciones en tiempo real.

El caso de estudio recoge el proceso que va desde la captura de imágenes de las diversas cámaras hasta la reconstrucción 3D de los objetos de interés en la escena, mediante algoritmos de *foreground segmentation*, *shape from silhouette* y *space carving*.

La adaptación de dichos algoritmos a la tecnología CUDA permite lograr unos tiempos de ejecución muy inferiores a los que se obtienen en una CPU tradicional, tal y como reflejan los experimentos realizados.

## I. INTRODUCCIÓN

Las exigencias actuales de videojuegos y entornos de diseño gráfico han llevado a los fabricantes de tarjetas gráficas a aumentar drásticamente su capacidad de cálculo. Al verse limitados por la ley de Moore en la velocidad de los procesadores y dado que el tratamiento gráfico en general no está ligado a demasiadas exigencias de procesado secuencial, los fabricantes han evolucionado hacia la parallelización de su hardware, diseñando dispositivos con un número cada vez más elevado de procesadores.

Esta tendencia ha llevado a convertir las tarjetas gráficas de última generación en verdaderos supercomputadores a un precio razonablemente asequible, cosa que no ha pasado desapercibida para los fabricantes. Aprovechando esta ventaja de sus dispositivos, Nvidia ha creado un conjunto de herramientas de desarrollo y una arquitectura basada en lenguaje C, denominada CUDA (Compute Unified Device Architecture), con el propósito de facilitar la programación de carácter paralelo para sus GPUs, sin la necesidad de aprender un lenguaje completamente nuevo desde cero.

Por otro lado, la similitud que presenta con el estándar de computación paralela OpenCL<sup>1</sup> en desarrollo, convierten a CUDA en un paso intermedio muy cómodo para quienes pretenden llegar a esta última tecnología de software libre.

### A. Entendiendo la paralelización

Para hacernos una idea de las mejoras que implica trabajar sobre estos dispositivos, debemos pensar que, mientras una CPU típica Intel Core 2 Duo logra entre 15 y 25 Gflops/s, una GPU actual de gama alta como la GeForce GTX 295 ofrece cerca de 1788 Gflops/s. Es decir, estamos hablando de una capacidad de cálculo 70 veces superior a la de una CPU, y por cierto, sólo unas 10 veces inferior al sistema que ocupa el último puesto en la lista de los 500 supercomputadores más potentes del mundo<sup>2</sup>.

Con la tecnología CUDA se pretende aprovechar la arquitectura multiprocesador de estas tarjetas para realizar de forma masivamente paralela cálculos idénticos que hasta el momento se han realizado de forma serializada. Para ilustrar el tipo de cálculos que son susceptibles de aprovechar esta arquitectura, podemos pensar en un procesado de imagen elemental a nivel de píxel, como pueda ser la inversión de colores o negativo. Pongamos por caso que queremos invertir una imagen de 1200x800 píxeles. El procedimiento consiste en reemplazar el valor de cada píxel por la resta de éste al máximo. Es decir, en RGB con 8 bits por píxel necesitaríamos calcular para cada píxel:

$$\text{pixel}_{\text{new}} = (255, 255, 255) - \text{pixel}_{\text{old}}$$

<sup>1</sup> Open Computing Language: propuesta de estándar abierto para la programación paralela en sistemas heterogéneos.

<sup>2</sup> Listado Junio 2009 de Top500.org (Financial Services, Blade Center HS21 Cluster: 17.080 Gflops/s)

La diferencia entre el caso serializado y el paralelo radica en que mientras el primero realizaría 960.000 veces el mismo cálculo en un único procesador (para simplificar), el segundo podría realizar, por ejemplo, el mismo cálculo sólo 2000 veces por procesador pero en 480 procesadores a la vez (GTX 295). No es difícil imaginar que cuando el volumen de los datos aumenta, la computación paralela se perfila como una solución mucho más rápida que su análogo serializado.

### B. Programar en CUDA

A continuación veremos las diferencias fundamentales entre programar en CUDA y en C:

- En CUDA hay 2 entornos diferenciados de ejecución. Por un lado está el *device*, que se corresponde con el lado de la GPU y por tanto es donde se realizan los cálculos complejos de forma paralela en el seno de una función que denominaremos *kernel*; mientras que por el otro está el *host*, que se corresponde con el lado de la CPU, desde donde se accede a ficheros de datos y se llama a los distintos kernels.
- La ejecución de un kernel se lleva a cabo por un conjunto numerado de *threads*. Al lanzar el kernel se indican los tamaños de un *array* de dimensión 1 o 2 (grid) que contiene a su vez arrays (bloques) de hasta dimensión 3 de *threads*. De esta manera, habitualmente se sustituyen los clásicos bucles en C para cálculos repetitivos (con índices representados por nuestras variables) por un cálculo repetitivo (*kernel*) ejecutado por varios *threads* (con índices representados por las variables *threadIdx.x*, *threadIdx.y*, etc. que nos proporciona CUDA).
- Los *threads* de cada bloque, se ejecutan en grupos de 32, denominados warp, trabajando primero con la mitad inferior y luego con la superior (*half-warp*). En general, si un half-warp accede a memoria global de forma alineada y ordenada, bastará con una lectura conjunta del espacio de memoria, en vez de necesitarse 16 lecturas separadas, una por *thread*.
- Al haber dos entornos de ejecución, se necesita duplicar estructuras de datos, de forma que generalmente tendremos un equivalente al contenido de la memoria del *host* en la memoria del *device*.
- La memoria del *device* es un recurso muypreciado y variado. Cuanto más escaso es el tipo de memoria (empe-

zando por los registros) más rápido es el acceso a ésta. La alta latencia que implica la memoria más abundante de la GPU (denominada global), suele orientar el estilo de programación y es el punto más importante a optimizar ya que es un factor determinante en el tiempo de ejecución del código.

- Cuando un *thread* se queda bloqueado por un acceso a memoria, otro *thread* ocupa su lugar en el procesador, de manera que la latencia de los accesos a memoria se puede ocultar parcialmente en *kernels* con un alto contenido computacional, así como con un número elevado de *threads*.

Debemos entender que parte de estas limitaciones de memoria se deben precisamente a la naturaleza del dispositivo. Mientras que en una CPU el propósito general es poder ejecutar muchas aplicaciones al mismo tiempo y acceder a un gran volumen de datos, en la tarjeta gráfica lo que prima es la velocidad con la que se procesan los datos, siendo preferible un *streaming* ligero de datos y varias ejecuciones sobre estos datos.

### C. Caso de estudio

Una vez determinada la naturaleza y los beneficios de la tecnología que proponemos utilizar, veamos cómo puede ayudar en la situación que se estudia. El escenario en el que pretendemos aplicar esta tecnología es el de una sala equipada con múltiples cámaras y micrófonos, cuyo objetivo es el análisis, por parte de un sistema cognitivo, de las acciones que se desarrollan en la sala. La parte de este sistema en la que trabajaremos es la encargada de procesar los datos visuales capturados por las cámaras previamente calibradas. Existen dos metodologías para trabajar con datos *multi-cámara*: por un lado, existe la posibilidad de analizar el contenido de la escena sobre cada una de las imágenes y validar los resultados a partir de las redundancias presentes entre vistas. La alternativa, con la que trabajaremos, es la fusión de los datos multi-cámara en un soporte común, obtenida mediante la reconstrucción 3D a partir del contenido de las imágenes de cada cámara.

Dado que el objetivo es analizar las interacciones entre personas y objetos utilizados por éstas en el entorno controlado de la sala, nuestro sistema se divide en dos etapas: la primera dedicada a la detección de personas y objetos utilizados mediante la segmentación de primer plano (*foreground segmentation*) y la segunda encargada de obtener la reconstrucción de estos elementos de primer plano.

## Foreground segmentation

En primer lugar, determinaremos las zonas de interés en cada imagen, es decir, aquellas zonas en las que la cámara ve un objeto que se quiere representar. Para ello, haremos una extracción de elementos de primer plano, o *foreground segmentation*. Considerando un modelo de fondo de la escena, obtenido a partir de una secuencia de imágenes de entrenamiento sin objetos en la sala, se detectan objetos de primer plano en aquellas áreas de la imagen que no se ajustan a dicho modelo. Este proceso, de manera muy simplificada, se reduce a restar, píxel a píxel, la imagen actual de cada cámara con una imagen conocida del fondo, de manera que aquellos píxeles que no hayan sufrido variación alguna quedarán con valor cero. De esta manera, es fácil ver que los píxeles de la imagen que se correspondan con algún objeto a representar en la sala, tendrán valor diferente de cero. Como se puede ver, igual que en el ejemplo comentado anteriormente, este algoritmo puede sacar mucho provecho de CUDA, realizando la misma operación con varios píxeles de forma simultánea.

## Shape from Silhouette

A continuación, podemos estimar la ocupación volumétrica del espacio de la sala a partir de los datos disponibles en cada vista, que son las siluetas obtenidas en la etapa anterior. Así, obtenemos una fusión de los datos de partida (siluetas de cada cámara) a un marco común (ocupación volumétrica) que nos proporciona una información más completa que la ofrecida por las siluetas. Partiendo de las siluetas obtenidas con anterioridad y el conocimiento de los datos de calibración de las cámaras, determinaremos el máximo volumen posible contenido en la intersección de las proyecciones de las siluetas, también denominado *Visual Hull* [3].

Para lograr este objetivo aplicamos un algoritmo de *Shape from Silhouette* [4] voxelizado. La idea básica de este algoritmo consiste en dividir el espacio 3D a representar en una matriz tridimensional de pequeños cubos (vóxeles). Consideramos que un vóxel está ocupado cuando visto desde la perspectiva de cada cámara se corresponde con píxeles de primer plano. Llegado este punto, ya somos capaces de representar una primera aproximación al volumen de interés en nuestra escena representado por un conjunto de vóxeles ocupados, aunque, en general, el volumen obtenido será mayor que el que ocupan realmente los objetos. No obstante, podemos asegurar que los objetos reales están contenidos en su interior: esto se debe a que al encerrar un volumen a par-

tir de la intersección de la proyección de siluetas, si existen espacios vacíos ocultos entre éstas también los incluiremos. Además, el desconocimiento previo de la distancia entre cámara y objetos nos lleva muchas veces a incluir un espacio mayor, como se puede ver en la Figura 1. Es por esto que la cantidad de falsos positivos vendrá altamente determinada por el número de cámaras que se hayan empleado. La idea que esconde este proceso se aprecia de forma simplificada en la visión aérea que observamos en la Figura 1, donde el cuadrado verde representa la figura real, la zona verde en los rectángulos blancos es la zona de la imagen donde hay silueta y de la intersección de líneas verdes, obtenemos el Visual Hull. En su interior, la parte anaranjada representa los falsos positivos que encontramos en este caso.

De nuevo, podemos ver como la repetición de un mismo proceso sobre todos y cada uno de los vóxeles en los que dividimos el espacio, nos conduce a pensar que la paralelización de este proceso será realmente eficiente.

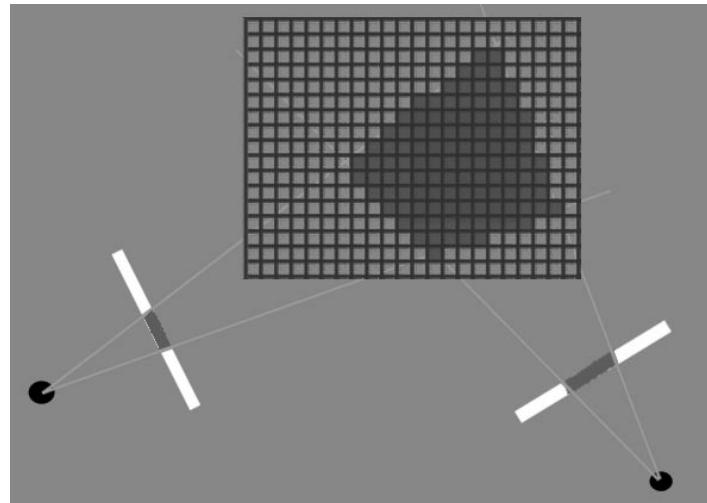


Figura 1. Visual Hull voxelizado de un cubo, a partir de 2 cámaras

## Space Carving

Un método más costoso, pero que nos permite mejorar en precisión la ocupación de los vóxeles es el análisis de fotoconsistencia (Space Carving [5]), que se puede aplicar sobre la estimación del Visual Hull obtenida anteriormente. A grandes rasgos, este algoritmo tiene un funcionamiento muy similar al anterior, con la diferencia de que ahora para cada vóxel de la escena se comprueba si el color con que se ve desde cada una de las cámaras es similar. En caso de que los colores no coincidan, se considera que ese vóxel está en realidad vacío. Así mismo, se aprovecha para determinar el color final de representación de cada vóxel ocupado, obteniendo lo que denominamos *Photo Hull*. Podemos ver como

funcionaría el proceso sobre el ejemplo anterior en la Figura 2, donde hemos añadido un cambio de color al objeto y podemos ver como los véxeles en amarillo se han descartado y para el véxel en negro se está analizando la fotoconsistencia, que también dará como resultado el descarte del véxel.

De nuevo, el análisis sistemático de todos los véxeles del espacio, nos hace pensar en la idoneidad de la paralelización de este algoritmo.

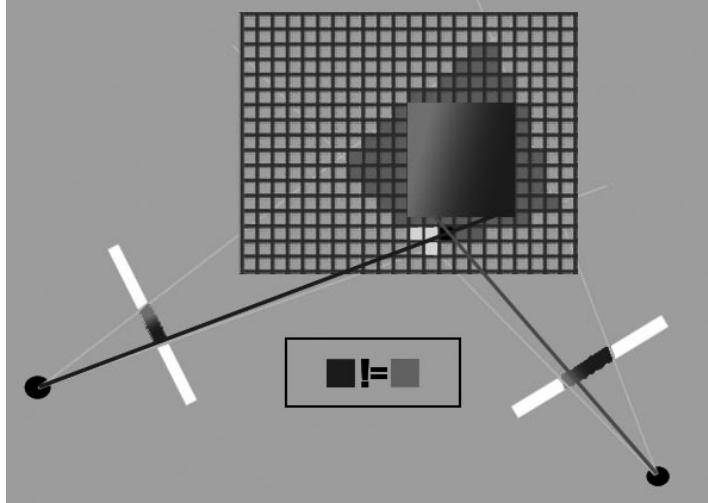


Figura 2. Proceso de space carving con 2 cámaras, vista cenital

## II. IMPLEMENTACIÓN PROPUESTA

### A. Foreground segmentation – Stauffer & Grimson

En la introducción hemos explicado que hacia falta conocer el fondo para poder determinar los píxeles que no le pertenecen y que forman parte, por tanto, de la silueta del objeto a representar. Los cambios de intensidad de la luz a lo largo del tiempo, la proyección de sombras, etc. implican variaciones en el color de las imágenes captadas por las cámaras que hay que tener en cuenta para identificar correctamente el fondo, por lo que, basándonos en la propuesta de Stauffer & Grimson [2], modelamos el fondo que ve cada cámara como un conjunto de Gaussianas tridimensionales en el espacio RGB, una por píxel, con una cierta varianza. Inicialmente, durante unos cuantos segundos de entrenamiento con la sala vacía, se determina dicho modelo calculando los valores RGB de cada píxel como el valor medio de la secuencia así como la varianza de ésta. Al final del proceso, lo que obtenemos es una matriz que contiene el valor medio del color de cada píxel –en adelante la *media*– y su varianza a lo largo de la secuencia de entrenamiento.

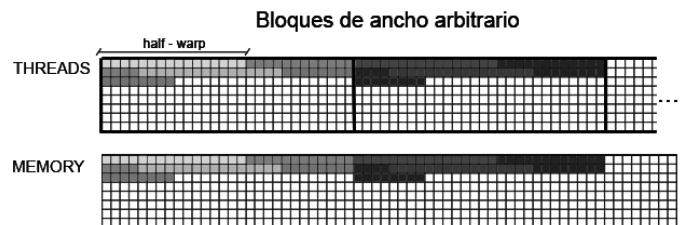
Una vez tenemos la media y la varianza, lo que haremos es comprobar para cada nueva imagen, si los colores de sus píxeles tienen un valor cercano al de la media o no. En caso de que su distancia supere un cierto umbral, se decide que no forman parte del fondo y por tanto se marcarán como primer plano, en color blanco, en la imagen de salida, como vemos en la Figura 3. Para los píxeles que sí se considere que forman parte del fondo, se tomará su valor para actualizar las imágenes de media y varianza de éste.



Figura 3. Imagen original y segmentación de siluetas

El acceso ordenado a memoria en la GPU nos proporciona grandes beneficios, por lo que siempre que sea posible se intentará leer y escribir en memoria de forma ordenada (para *threads* consecutivos). En este caso, la idea es crear un *grid* de ejecución bidimensional, de forma que la estructura en *threads* sea equivalente a las dimensiones de la imagen y podamos acceder a ésta ordenadamente, fila por fila. No obstante, para obtener un alto rendimiento del código en CUDA, debemos hacer dos optimizaciones en este punto.

En primer lugar, y esto será una pauta común para todos los accesos ordenados a memoria, los bloques de *threads* se definen con un ancho múltiple del tamaño de medio *warp* (16 *threads*). Con esto logramos que, a la hora de acceder a memoria de forma ordenada en el bloque, se haga línea por línea, sin tener saltos de línea en un mismo *half-warp*. Debemos recordar que se necesitarán varios bloques para completar la imagen, por lo que dicho salto, significaría introducir un offset de tamaño equivalente al ancho de la imagen entre dos espacios de memoria de dicho *half-warp*. La Figura 4 ilustra la diferencia entre ambos casos.



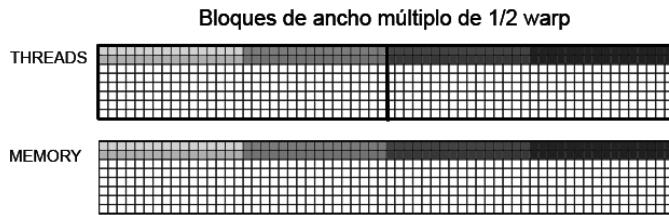


Figura 4. Offset en accesos en 1 mismo half-warp debido a un ancho arbitrario de bloque

En segundo lugar, al reservar memoria para la imagen, no lo haremos con las dimensiones justas de ésta, sino que alargaremos su anchura para que sea también múltiplo de 16, de modo que el ancho de la imagen en memoria sea un número entero de *half-warps*. De no hacerlo, un ancho arbitrario de la imagen podría provocar la falta de alineación que queremos evitar. Nos ayuda a comprender este funcionamiento la Figura 5 .

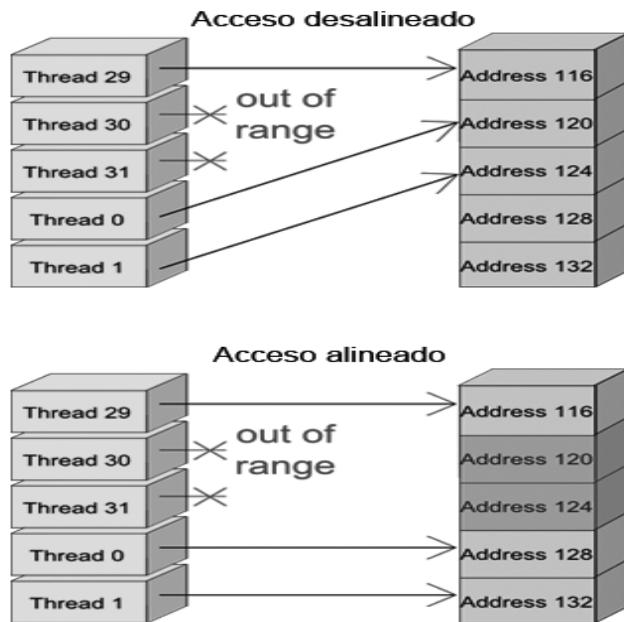


Figura 5. Asignación alineada de memoria

### B. Visual Hull - Shape from silhouette

Una vez se tienen las siluetas de las distintas cámaras, procedemos a la reconstrucción 3D.

Para ello, dividimos el espacio a representar en un conjunto ordenado de cubos (vóxeles), con la idea de repetir el mismo procedimiento para cada voxel y con cada una de las cámaras. A partir de las coordenadas 3D del voxel y los datos de calibración de la cámara, obtenidos mediante [1], se determina a qué píxel de la imagen corresponde mediante su proyección.

De forma similar al primer algoritmo, ejecutaremos estas operaciones con kernels numerados y ordenados ahora en las tres dimensiones del espacio, para poder acceder a la memoria donde guardaremos los datos del volumen de forma óptima. Como en el caso anterior, deberemos modificar el ancho de la memoria a reservar para coincidir con un múltiplo del *half-warp* y evitar así perdidas de alineamiento. Al ejecutarse cada thread, obtenemos de forma casi inmediata las coordenadas 3D del voxel correspondiente y lo proyectamos para obtener las coordenadas en la imagen. Llegado este punto, sencillamente se comprueba si dicho píxel corresponde o no a una silueta. Este paso requiere forzosamente una lectura desalineada de la imagen, ya que la correspondencia entre coordenadas 3D consecutivas en general no se mantiene al proyectar sobre la imagen 2D, aunque una cantidad elevada de threads ocultará parcialmente la demora introducida por estas lecturas.

Tras repetir el proceso para todas las cámaras, en una nueva iteración parecida a la anterior se contabiliza el número de cámaras que ven la proyección de cada voxel como perteneciente a silueta y en caso de que sean todas, se decide, finalmente, que el voxel pertenece a un objeto, pintándolo de blanco. Un ejemplo de los resultados obtenidos con éste método, para voxels de 2 cm de lado, se puede ver en la Figura 6. Es importante fijarse en el detalle de falsos positivos bajo la mano de la figura de la derecha, debido a que la configuración de cámaras utilizada permite que todas ellas vean la proyección de dichos voxels como pertenecientes a una silueta (en todas las imágenes quedan ocultos por un objeto).



Figura 6. Reconstrucción basada en Shape from Silhouette

### C. Photo Hull – Space Carving

Finalmente, proponemos eliminar falsos positivos comprobando la consistencia de color de cada vóxel mediante la distancia euclídea entre los colores de los píxeles en que se proyecta. Este algoritmo, a pesar de ser similar al anterior, es el más costoso de todos, dado que se necesitan varias iteraciones en las que solamente se determina la consistencia de color para aquellos vóxeles que se encuentran en la superficie del volumen y son, por tanto, visibles desde las cámaras. Además, también requiere de más memoria dado que necesitamos 3 nuevos canales de salida para poder extraer los colores RGB correspondientes a cada vóxel ocupado del volumen, así como una imagen en escala de grises con la profundidad vista por cada cámara, útil para determinar la visibilidad de cada vóxel.

En primer lugar, necesitamos obtener el Visual Hull con el algoritmo anterior. A continuación necesitamos conocer el mapa de profundidad observado por cada cámara para poder determinar qué vóxeles son visibles para una determinada cámara, y cuáles se encuentran tapados por otros vóxeles. Con este propósito reservamos espacio alineado para tantas imágenes como cámaras haya, de la misma manera que en el algoritmo de segmentación de primer plano. Se inicializa cada píxel con un valor de distancia máxima arbitrariamente elevado.

A continuación, se ejecuta un kernel que recorrerá todo el volumen y comparará la distancia de cada vóxel a la cámara con la distancia almacenada en la imagen. En caso de ser menor, la actualizará al valor actual. Al acabar el proceso tendremos una imagen representando la distancia mínima del volumen observado a cada píxel. Una vez se dispone de las distancias mínimas a cada cámara, se recorre de nuevo todo el espacio 3D, como en el *Shape from Silhouette*, y para aquellos vóxeles que se encuentren a distancia mínima de todas las cámaras en las que forma parte de su campo de visión, se comprueba que la distancia euclídea entre las componentes RGB de sus respectivas proyecciones no supere un cierto umbral. En caso de superarlo, se considera el vóxel como no fotoconsistente y se descartará como parte del objeto a representar, dando lugar a la visibilidad de vóxeles previamente ocluidos por éste.

Tras esta criba, es necesario recalcular los mapas de distancias mínimas antes de poder repetir la operación hasta encontrar que el volumen reconstruido es fotoconsistente para todos los vóxeles que lo componen.

Después de las iteraciones necesarias, el grueso de falsos positivos se habrá descartado y el volumen que tendremos será más ajustado a la realidad que en el caso de *Shape from Silhouette*. Llegado este punto, realizaremos una última iteración, en la que sencillamente se determinará el color de cada vóxel de superficie que aún pertenezca al objeto a representar, eligiendo como criterio el de mayor saturación. Por otro lado, cabe comentar que todos aquellos vóxeles que son visibles desde una sola cámara se clasificarán directamente como fotoconsistentes y se tomará el color de su píxel correspondiente. Podemos ver un ejemplo del resultado de *Space Carving* con vóxeles de 2 cm de lado en la Figura 7.

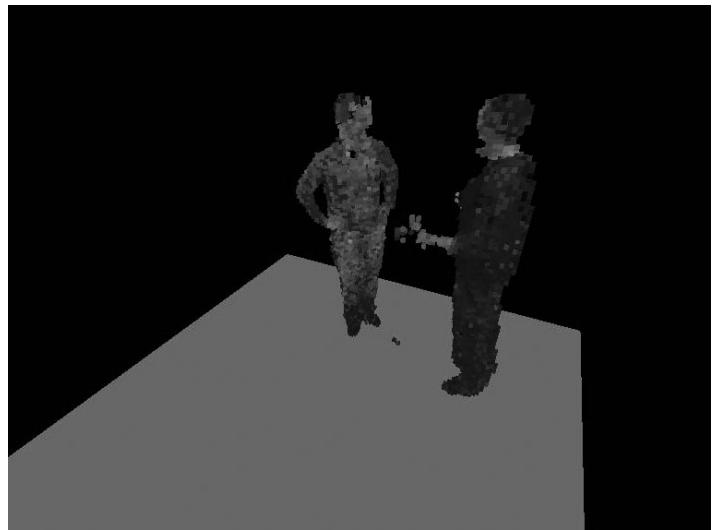


Figura 7. Reconstrucción basada en space carving

## III. RESULTADOS

La tarjeta gráfica con la que se ha trabajado es una Nvidia Quadro FX 3700 (112 procesadores a 1.3GHz, 512 MB). Pese a no ser una de las tarjetas más indicadas para supercomputación, ni por procesadores ni por memoria debido a un tiempo de existencia relativamente largo en el mercado, los resultados son suficientemente reveladores como para entender el beneficio de esta tecnología. Las comparaciones con CPU han sido ejecutadas sobre un procesador Intel Core 2 Duo E7300 (2.66 GHz).

Respecto al escenario multi-cámara, se trata de una sala rectangular equipada con un total de cinco cámaras: una en cada esquina y una quinta cámara en posición cenital.

Antes de empezar a detallar los resultados, hay que considerar que las funciones para controlar el tiempo de ejecución en C y en CUDA son distintas, teniendo la primera una resolución de aproximadamente 10 ms, mientras que la segunda

tiene una resolución inferior a 1 ms. Es por eso que en determinados momentos puede parecer que el algoritmo en CPU no varía entre dos ejecuciones donde se ha incrementado el volumen de datos ligeramente, mientras que en GPU incrementa su tiempo de ejecución. Esto se debe sencillamente a una limitación en la resolución de la medida en CPU.

#### A. Comparativa foreground segmentation

Para comparar la rapidez de ambas opciones con el algoritmo de segmentación de primer plano, se propone trabajar con una secuencia de 300 imágenes, donde las 10 primeras son de entrenamiento para generación del modelo del fondo. Se determinará el tiempo total que se tarda en procesar cada una de las imágenes y se dividirá por el número total de imágenes para obtener un promediado del tiempo de ejecución por imagen. El mismo proceso se repetirá con imágenes de proporciones diferentes para ver como varía el tiempo de ejecución en función del volumen de datos. Los resultados los podemos observar en la Figura 8.

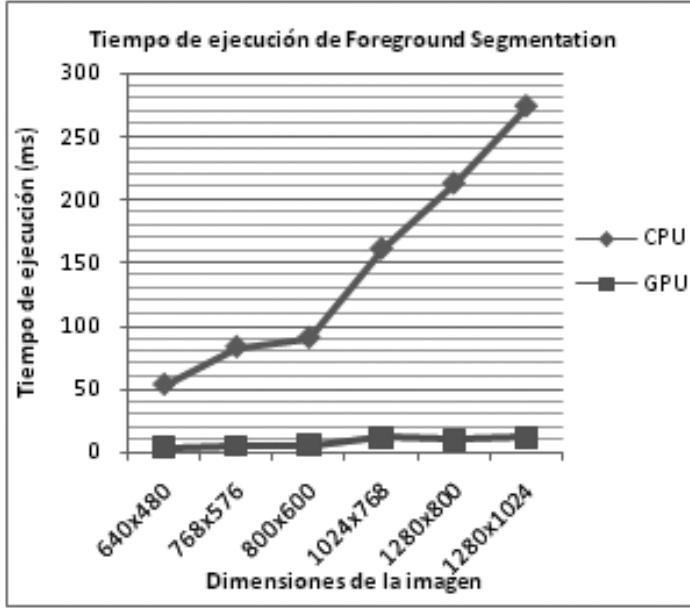


Figura 8. Comparativa CPU-GPU de ejecución de Foreground Segmentation

Podemos apreciar en el gráfico una enorme diferencia entre ambas opciones, especialmente cuando se trabaja con un gran volumen de datos. Como se verá en los siguientes apartados, es con este algoritmo donde se determina más claramente la diferencia que llega a existir entre ambas soluciones y donde se demuestra la importancia de los accesos a memoria en GPU de la que hablábamos anteriormente. En este algoritmo, el 100% de los accesos se realizan de forma alineada y ordenada, consiguiendo en todos los casos lectu-

ras conjuntas para cada half-warp de threads. La ventaja que obtenemos con este diseño, junto con un número elevado de operaciones por thread, esconden prácticamente por completo la latencia que se deriva de los accesos a memoria.

#### B. Comparativa shape from silhouette

Para comparar los tiempos de ejecución al realizar el algoritmo de Shape from Silhouette se ha propuesto un conjunto de ejecuciones donde se reduzca progresivamente el tamaño del voxel desde 4 cm de lado, a 0.5 cm de lado, a intervalos de 0.1 cm. El escenario a analizar se ha reducido a un cubo de 2 metros por lado (8 m<sup>3</sup>).

Hay que recordar que un decremento de X en el lado del cubo, en realidad representa un incremento de X3 en el número total de voxels con el que trabajamos. Del experimento realizado se puede constatar cómo el algoritmo sobre CPU se distancia del de GPU de forma incremental, conforme se aumenta el número de voxels de trabajo.

No obstante, antes de entrar en comparaciones es bueno comentar que en el caso de CPU, previamente al procesado en streaming, se crea una *Look-Up-Table* con la correspondencia de la proyección de cada voxel sobre cada una de las cámaras. Este cálculo, bastante costoso, no se tiene en cuenta a la hora de medir los tiempos, puesto que sólo se realiza una vez y lo que pretendemos demostrar aquí es que el algoritmo sobre GPU es más eficiente de cara al procesado en *streaming*. En el caso de GPU no se realiza este cálculo previo debido a las restricciones de memoria de las que hemos hablado, por lo que se debe en cuenta el tiempo que se tarda en calcular las proyecciones. Aun así, el algoritmo sobre GPU sigue siendo más rápido, como se ve en la Figura 9.

#### C. Comparativa space carving

Para comparar los tiempos de ejecución al realizar el algoritmo de space carving (con Shape From Silhouette previo incluido), se ha propuesto un conjunto de ejecuciones donde se reduzca progresivamente el tamaño del voxel desde 4 cm de lado a 0.6 cm de lado, a intervalos de 0.2 cm, mientras que se aumenta progresivamente el número de iteraciones a realizar. El escenario a analizar se ha reducido, de nuevo, a un cubo de 2 metros por lado. Pese a ver cómo realmente el algoritmo sobre GPU es mucho más rápido, también podemos observar las limitaciones a las que está sometido, siendo imposible ejecutarlo para voxels de 0.6 cm de lado debido a no disponer de suficiente memoria.

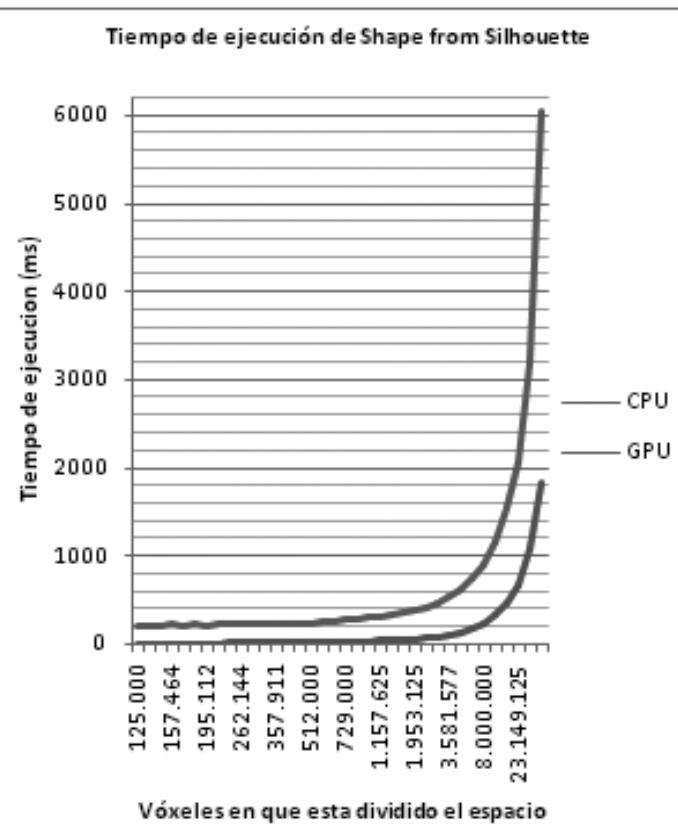


Figura 9. Comparativa ejecución Shape from Silhouette CPU-GPU

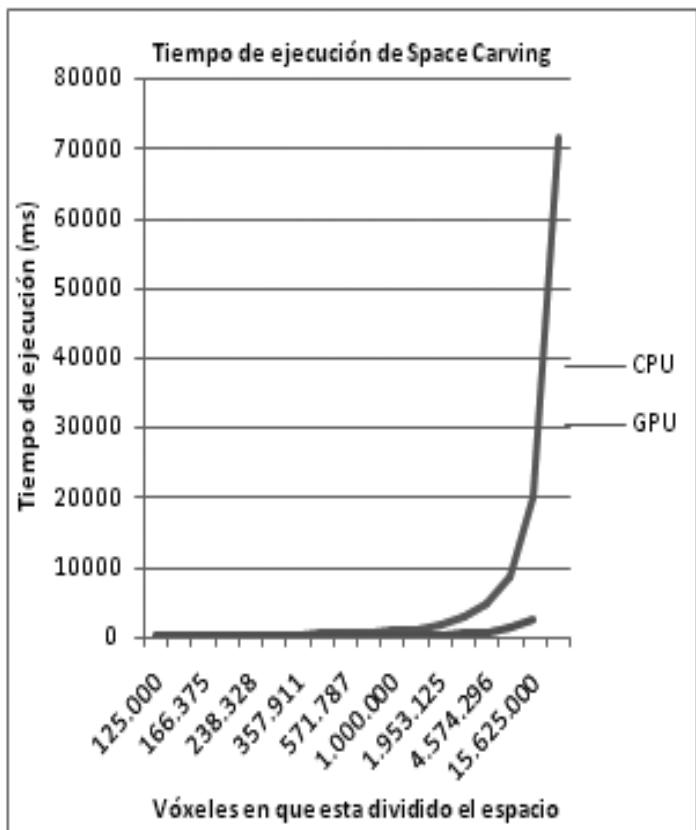


Figura 10. comparativa CPU-GPU con Space Carving

Como podemos ver en la Figura 10, la diferencia entre ambos procedimientos se dispara de forma exponencial conforme aumenta el volumen de datos. Pese a que los procesadores de la CPU son más rápidos que los de la GPU, cuando el volumen de datos es suficientemente grande, el limitado número de procesadores en la CPU se convierte en un cuello de botella, cosa que no pasa en el otro caso.

No obstante, el beneficio no se limita sólo a volúmenes realmente grandes de datos, como podría interpretarse a partir de la gráfica anterior. Como podemos observar en el zoom de dicha gráfica en la Figura 11, los beneficios se obtienen también con menores volúmenes, haciendo posible el uso de Space Carving orientado a tiempo real con tamaños de voxel que permitan apreciar detalles de la imagen.

Podemos ver como con voxelés de 1,8 cm por lado (5,832 cm<sup>3</sup>/voxel) lograríamos el mismo *frame rate* (4 frames/s aprox.) en GPU que con voxelés de 3,2 cm por lado (32,768 cm<sup>3</sup>/voxel) en CPU. O lo que es lo mismo, con un mismo *frame rate*, en GPU estaríamos trabajando con aproximadamente 1,367 millones de voxelés, mientras que en CPU sólo llegaríamos a trabajar con 238 mil.

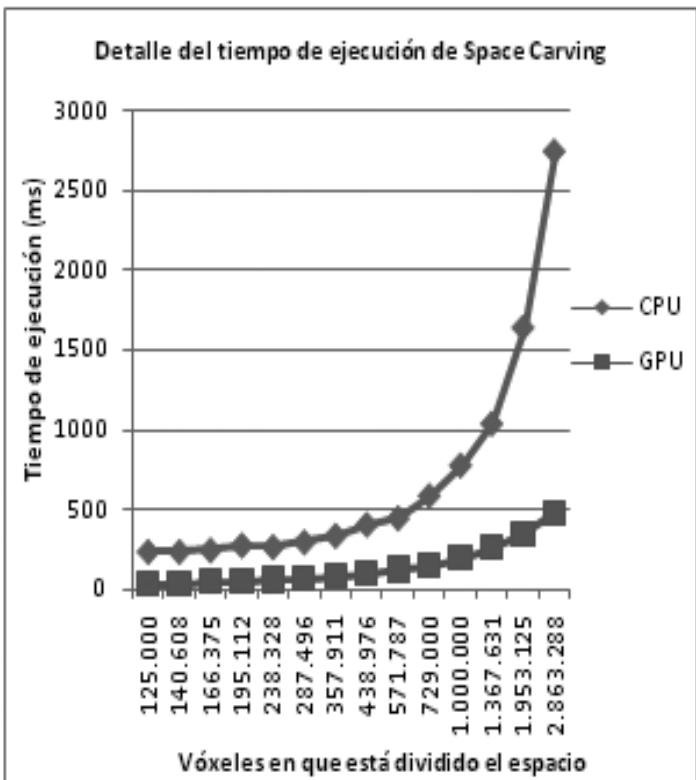


Figura 11. Detalle de comparativa CPU-GPU

## IV. CONCLUSIONES

En el caso que se estudia, podemos observar como con la computación sobre GPU realmente obtenemos tiempos de ejecución menores para un mismo coste computacional, de modo que nos permite extrapolar el algoritmo de reconstrucción a una situación de tiempo real con un *frame rate* muy superior al que nos ofrecería una CPU, o a mismo *frame rate*, con un nivel de detalle muy superior.

A la vista de los resultados obtenidos y especialmente, teniendo en cuenta que el equipo utilizado no es el que saca un mayor rendimiento a las capacidades de supercomputación que ofrece CUDA, podemos afirmar que la paralelización de algoritmos sobre GPU tiene un gran potencial y ofrece una solución económicamente viable para aquellos que requieren mejorar la velocidad de ejecución de aplicaciones no secuenciales con un alto volumen de datos.

Aunque la tecnología aún se encuentra en fase de expansión, gracias a su gran capacidad de cómputo y su facilidad de aprendizaje por su similitud con C, es probable que se acabe consolidando como una alternativa adecuada para necesidades de pequeños supercomputadores en el ámbito universitario y profesional. Por otro lado, podríamos decir que Nvidia ha abierto la brecha del potencial que supone la computación paralela de datos y es probable que en adelante se inicie una tendencia mucho más extrema que la actual hacia el hardware multiprocesador, como método para aumentar la capacidad total de cómputo del equipo, con iniciativas como el ya mencionado OpenCL y soluciones como los procesadores de 8 y 12 núcleos que propone AMD para el 2010<sup>3</sup> o algo más llamativo aún, como el procesador de 100 núcleos para entornos profesionales del fabricante Tilera<sup>4</sup>.

## REFERENCIAS

- [1] Bouguet, J.-Y., Camera Calibration Toolbox for Matlab, [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/)
- [2] C. Stauffer and W. Grimson. Adaptive background mixture models for realtime tracking. In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pages 252–259, 1999.

<sup>3</sup> Prototipos Sao Paolo y Magny-Cours de AMD

<sup>4</sup> Tilera GX100

[3] A. Laurentini, The Visual Hull concept for silhouette-based image understanding, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (2) (1994) 150–162.

[4] Cheung, G., Kanade, T., Bouguet, J.-Y. & Holler, M. (2000), ‘A real time system for robust 3d voxel reconstruction of human motions’, in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR’00), Vol. 2, Hilton Head Island, South Carolina (US), pp. 714–720

[5] Kutulakos, K.N., Seitz, S.M (2000), ‘A Theory of Shape by Space Carving’, *International Journal of Computer Vision* 38 (3), pp. 199–218

## AUTORES



Enrique Oriol es estudiante de Telecomunicaciones en la ETSETB y desarrolla su proyecto final de carrera en el Departamento del TSC en el ámbito de la reconstrucción 3D en entornos multi-cámara sobre GPU.



Jordi Salvador es estudiante de Doctorado en Teoría de la Señal y Comunicaciones del Departamento de TSC (ETSETB) e investiga en tecnologías de reconstrucción 3D en entornos multi-cámara.



Josep R. Casas es Profesor Titular del Departamento de TSC e imparte docencia en Procesado de Imagen y Televisión (ETSETB). Coordina las actividades de investigación que se realizan en el laboratorio “Smart Room” del Campus Nord de la UPC.

# High Altitude Platform Stations in Design Solutions for Emergency Services

*Israel R. Palma-Lázgare, José A. Delgado-Penín,*

CMC Group, TSC Department, Universitat Politècnica de Catalunya, Barcelona, Spain

E-mail: {ipalma, delpen}@tsc.upc.edu

## ABSTRACT

High Altitude Platform Stations (HAPS) are expected to conform a third major infrastructure for communications and broadcasting, after terrestrial and satellite systems. The proposal, which is maintained by many authors, is the use of HAPS as alternative wireless network provider that can partially replace or add capacity to damaged or overloaded wireless networks during a man-made or large- and small-scale natural disaster. During these critical phenomena, the telecommunications infrastructure and the required coverage for the emergency service operations might be unavailable due to the destroyed area or overloading by the excessive communications demand. Along with satellites, high altitude platforms (HAPs) will be completely isolated from the effects of disasters on the ground. A couple of stratospheric-based network scenarios are considered as examples for a HAPS-aided disaster deployment assessing communication viability and outlining issues in interoperability with existing networks.

## 1. INTRODUCTION

The big three options in the use of telecommunications are by means of the terrestrial wire systems (copper wire, coax and fibre optic cable), terrestrial wireless systems and satellite communication systems [5]. In the last few years, a new set of options have been added in the form of high altitude platform stations (HAPS) and unmanned aerial vehicles (UAVs) that operate as mobile telecom providers, mostly operating at stratospheric altitudes of ~20 km, or other low altitudes, above earth; in our days, it has been deeply studied that near-space platforms could be safely deployed at altitudes in the 17-25 km range. The origin of HAPS-based wireless systems comes principally from the military research, in such case are called UAVs, such as the Rover III that have assisted within the disaster recovery from Katrina, but these latter systems are being withdrawn from as commercial operation. Now, the issues of how HAPS-based systems, or solar-powered lighter-than-air (LTA) aircrafts, can be used to supplement communications and air survei-

llance during emergencies is examined under the point of view of specialised organisations and research groups, e.g., Department of Homeland Security [15], and CAPANINA [9].

The purpose of this paper is to describe how HAPS can be used to provide emergency telecommunication coverage in a disaster scenario. A general classification is defined identifying the main points that need to be considered for such application. An empirical GSM-TETRA-UMTS HAPS-based disaster layout is introduced and together with their suitability is discussed. Finally, the conclusion of the paper is given.

## 2. HAPS & EMERGENCY AID LIAISON

### 2.1. HAPS Survey

One major intention of this paper is to provide a contribution in the direction of describing the research activity performed (in particular at the European level) on HAPS-based communications [12], with particular attention to the two projects HeliNet and CAPANINA [9], both financed by the European Commission, in order to explore the use of HAPS for integrated communication services and broadband transmission.

In the broad multiform communications, HAPS possesses all the potentialities to be proposed as a novel stratospheric segment in the wireless communications market. They are able to overcome the main drawbacks of satellite technology due to its reduced distance from the ground with respect to satellites, and its quasi-stationary situation in the sky. Additionally, their costs for construction, deployment, launch, and maintenance can be kept in lower orders of magnitudes than those of satellites can, all these at an environmental sustainable effort because of their solar power supplies. On the other hand, it is evident that HAPS cannot replace satellites, nor terrestrial radio links, for reasons of coverage, reliability, safety, and cost. In fact, satellites, HAPS, and terrestrial broadband systems have different but complementary characteristics. While satellites are more suited for coverage of very large areas and broadcast applications, HAPS are able to cover remote or sparsely populated areas at reduced costs, and to

offer broadband services to mobile users; meanwhile, terrestrial infrastructures are advantageous for interactive services in densely populated areas. For these reasons the final design goal must be a flexible and synergic integration between satellite, stratospheric, and terrestrial segments, which can lead to a truly evolutionary scenario.

As far as services integration is concerned, we can point out that the establishment of the stratospheric network is a challenging task, but whose positive benefits should be exploited in the best possible way. Since the weight constraints for HAPs are generally less critical than those for satellites, many different payloads may be carried by the same platform, thus exploiting the favourable transmitting position of HAPs for improving performance and coverage of a variety of services. Among those possible services that can benefit from a stratospheric segment, we can cite broadband communications, environmental surveillance, video surveillance, reconnaissance, meteorological and atmospheric measurements, localisation and positioning services, emergency cellular telephony services, broadcasting, and emergency communication services.

For now, in our emergency case, such emergencies can happen any time and any place and emergency communications in the damaged field requires a global infrastructure accessible 24 hours a day from any place [1]; HAPS-based constellations, and satellite-based networks, are proposed as the innovative aiding technologies. We can mention that all member countries in the International Telecommunication Union (ITU), who have allocated a wide frequency band of more than 500 MHz in total [1], have accepted such proposals.

Hence, HAPS bring the idea of a global emergency communications and information infrastructure defining the main objective in which integrates service operations in a cost-effective way. In a disaster scenario, HAPS can be deployed quickly over the emergency area carrying telecommunications payloads. What is more, from their operational stratospheric altitude they have a wide coverage area and can provide additional capacity to the existing network or replace the coverage holes left due to damaged masts and repeaters. And, besides to the superior flight endurance relative to airplanes and helicopters, HAPS have the potential for a variety of uses at the time of a natural disaster [3] –such as monitoring of the disaster area, traffic control and guidance for emergency vehicles, provision of search-related information, mobile communications, and emergency broadcasting. –

## 2.2. HAPS-Disaster Relief Assessment

Some brief terms of global emergencies are described in [6] that can be applied into HAPs emergency considerations. The global natural disaster assessment involves three types of data that human is exposed to: risk, hazards and vulnerability. For instance, the risk assessments (of mortality and economic losses) are based on two data sets characterizing elements at risk: population and Gross Domestic Product (GDP) per unit area. Global hazard data were compiled from multiple sources. Drought, flood and volcano hazards are characterized in terms of event frequency, storms by frequency and severity, earthquakes by frequency and probability of exceeding a set threshold of peak ground acceleration, and landslides by an index derived from probability of occurrence. For a given hazard, vulnerability will vary across a set of similar elements and from one element to the next.

After, the next items undertake a comprehensive review of all types of telecommunications systems in the context of upgrading emergency communications systems within appropriate technologies [5]. We consider the resilience, flexibility, upgrade-ability, modularity and interoperability for emergency communications standards inside the all possible relevant Information and Communications Technology (ICT): fibre optic networks, satellites (fixed, mobile, broadcasting and military), broadband wireless, WLANs, 802.11 & 802.16 systems, power line communications (PLC) systems, HAPS systems and UAVs, voice over IP-capable systems. We can have, additionally, every emergency communications systems working at 700 MHz, 800 MHz frequency bands and/or at newly designated 4.9 GHz bands.

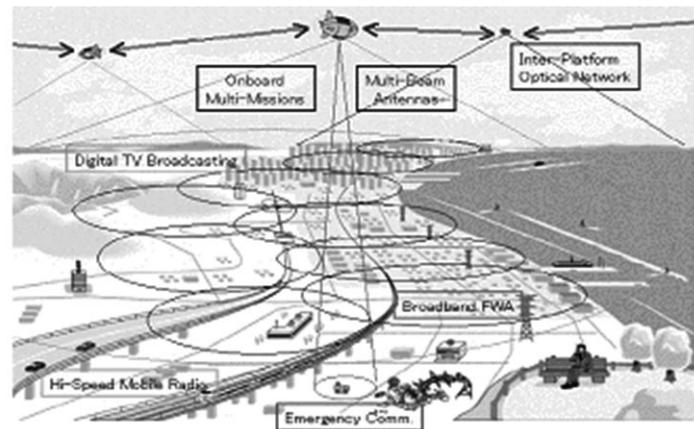


Figure 1. General HAPS-network (HAPN) architecture and the possible communications scenario [11].

Concerning the HAPS-based systems [2], essentially, HAPS-based wireless communications work even under the worst conditions, including natural disasters and emergencies. This potential, however, needs to be fully recognized and exploited. Current emergency communications within the ‘active field’ bases on an old narrowband technology created to transmit mainly voice messages over restricted areas only. Therefore, it is considered the use of HAPS that is found fully suitable for operational communications especially in the initial phase of an emergency, when real-time exchange of information is most essential.

Various potential resolution specifications and disaster-assistance functions are to be aimed from these stratospheric platforms [4, 5, 6]:

- HAPS will be maintained over the critical region for exchanging real-time information on operations within the disaster framework laws to facilitate decision-making.

- Interoperability, interconnection and credentials for all types of first responders (police, fire, utility repair and installation, dispatchers).

- Rapid information exchange with key partners and permanent monitoring of media resources (TV, radio); prompt notification and dissemination of emergency information in the event of an emergency.

- Creation of a basis for timely mobilization of international assistance to countries facing environmental emergencies to minimize the environmental impacts from brokerage of information and assistance.

- Establishment of an inventory of possible emergency information resources and tools, and facilitation of access to these during emergency incidents: better quality emergencies management, mitigation and response through rapid access for vital emergency tools and resources.

- Continuous assessment and implementation of best practices, improved use of HAPS-technology and more extensive reliance on automated system: improved response, more efficient and less time dissemination of information, and overall emergencies coordination and management.

- Efficient mobilisation and coordination for international assistance to requesting countries: delivery of specialised service to affected countries, and supplementing national efforts to respond to environmental emergencies.

- Development of post-incident measurements for the assessment of internal procedures following incidents: identification of lessons learned and areas for improvement, to increase efficiency and better coordinated emergency response management.

- Plan, organize and conduct periodical meetings of groups on environmental emergencies: further development and harmonisation of the international response to environmental emergencies, and better transparency of the joint environments Units.

### **3. GSM-TETRA-UMTS HAPS-BASED SCENARIO FOR EMERGENCY OPERATION, A CASE STUDY**

The deployment of GSM, UMTS and TETRA services was examined in [13] with the help of an emergency scenario (Fig. 2). An island was assumed as an operational area of 16 km by 18 km. Due to a natural disaster all mobile communication nodes have failed. An emergency mobile network was deployed over the area to assist emergency services and served the public until the restoration of the terrestrial network was possible. As the HAPS flew over the area there were different important decisions that needed to be considered: service area needs, cell layout needs, backhaul needs, and operational current systems that could interfere with the HAPS network.

There were two different deployment scenarios: a single 25 km diameter cell covered the entire area during the early hours of an emergency to allow quick deployment, and within the allowed limits and no synchronisation problems were anticipated. Since capacity was limited in such deployment, only the emergency services have served; later on, extra smaller cells (2 km diameter) could be deployed over the residential areas, road links, airport, and other important facilities. When the smaller links were in place, the large coverage cell could be shut down leaving the next cellular deployment.

Small cells could provide the same data rates as the large cells use only a fraction of the power. This is due to the smaller antenna beamwidth required to create the footprint, which in turn resulted in a bigger antenna gain. Handovers, however, were needed as users moved between cells. These handovers increased the payload processing power and complexity. A cell larger than the coverage area could guarantee that there were no users at the edge of coverage and therefore they could be insensitive to small platform movements.

In the cases of GSM and TETRA protocols, they could face problems with long link lengths due to synchronisation problems inherited from their TDMA nature. The GSM stack included guard bands, which compensate for the propagation delays up to a theoretical maximum of 37.8 km. Having considered that HAPS operates from a 17 km altitude, HAPS-GSM cell was limited to 30.5 km radius. Taking into account the ITU-R recommendation and allowing 500 m of azimuth platform movements, the resulting GSM coverage was limited to a 60 km diameter cell assuming that HAPS-antennae were stabilised to ensure a fixed coverage area. Similarly, TETRA was also limited to a 60 km diameter cell. There was the option to use only 4 out of the 8 available time slots in each TDMA frame and hence doubled the guard band but also halved the capacity. In such case the maximum cell size could be doubled and a coverage area of 140 km diameter could be achieved. This scheme has been successfully running in Australian rural areas where traffic is limited. Successful trials increasing the theoretical maximum cell radius from 35 km to 121 km have carried out by Motorola's, Cellular Infrastructure Group (CIG), and Spanish GSM Operator Telefonica, at the expense of limited capacity.

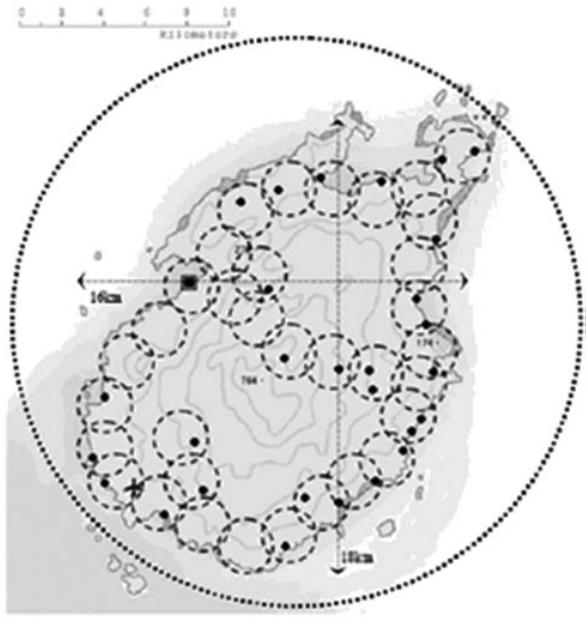


Figure 2. GSM-TETRA-UMTS scenario [13].

Studies and simulations in [13] showed that HAPS-antenna roll-off rate is a trade-off between having sufficient gain at edge of coverage and reducing unwanted interference. To meet the required specification, a roll-off factor  $n=15$  has been chosen for the 25 km diameter cell while  $n=2000$  has been chosen for the 2 km diameter cell; aperture efficiency of 70% has been decided for link budget calculi.

So far, the empirical scenario in [13] assumed that no other mobile network was active at the emergency area. This can rarely be realistic since most of Europe enjoys some kind of wireless telephony service.

To ensure the operability of the provided services, a satisfactory link budget needs to be achieved for both the uplink and downlink for GSM/TETRA (~900 MHz) and UMTS (~2 GHz) correspondingly. It is noted that similar link budgets can be produced for TETRA services, which shares the same principles as GSM due to their common TDMA nature. UMTS can re-use an already used frequency due to the CDMA nature of the multiple access method. The limiting factor is that with the addition of every new user the noise level of the cell rises. UMTS can support different types of services, each with different Eb/N0 requirements. In [13] already presents a scenario with the speech service (up to speeds of 120 km/h) and the 384 kbps real-time data service (up to 5 km/h).

The link budget for the HAPS-downlink has also been considered in [12, 14]. Due to the system and regulatory constraints, some extra parameters must have been stated. The use of a high-gain multi-beam on-board antenna was discussed. The HAPS-altitude and the broadband service provision over a region of about 35 km radius on the ground were thought about; the region covered by the HAPS was divided into cells and frequency reuse was adopted in order to avoid inter-cell interference. A uniform cellular structure was obtained by illuminating the covered area with multiple antenna elliptic beams so a circular footprints on the ground were shown.

Towards the end, regarding to enable co-existence with the existing GSM/TETRA network, in [13] a second layer of GSM cells could be laid by the HAPS overlapping the terrestrial cells. The overlapping cells used different frequencies from the terrestrial ones to avoid interference. Alternatively, HAPS could project cells over the coverage gaps produced by the non-operational base stations. The problem in the latter deployment was how to determine, in a short time, which base-stations were not operational and which frequencies were not used. Getting up-to-date information from mobile providers during a major disaster could be impossible, so an alternative way of determining which base stations were not in service should have been devised. Finding available frequencies for operation is mainly a GSM specific problem. TETRA is not used widely and it was highly unlikely that all the frequencies in [13] the area would be occupied. In this unlikely event, authors in [13] proposed the techniques of area scan using a selective frequency scanner for locating an unused frequency can apply.

Finally, [13] addressed the implications arising from the coexistence between the terrestrial UMTS network (2km diameter cell) and the HAPS-based UMTS network (25km diameter cell). The effects on the HAPN from the interference caused by the terrestrial network showed that the operational threshold (10dB below the noise floor, as specified by the ITU-R) was reached when the HAPS-cell completely overlaps the terrestrial 2km cell. As the terrestrial cell moved towards the centre of the HAPS-cell, the interference level raised and system was saturated.

## 4. CONCLUSIONS

This paper has briefly presented the idea of the use of HAPS as base stations to provide and face emergency services. Advantages of such an application include rapid deployment, large coverage area and certain immunity to most catastrophes. A case study of an emergency scenario for GSM, TETRA, and UMTS, and coexistence with the HAPS network, have been implicated to show the evaluations to be followed for a viable disaster planning and crisis-services.

## REFERENCES

- [1] Peha J. M., Struzak, R., "Public safety and emergency communications," *IEEE Communications Magazine*, March 2005.
- [2] Struzak R.: "Evaluation of the OCHA (DRB) Project on Emergency Telecommunications With and In the Field," *UN-OCHA Geneva- New York*, 2000.
- [3] "NICT (National Institute of Information and Communications Technology) News," *ISSN 1349-3531, No. 349, April 25, 2005*.
- [4] "Programme Update-Disaster Management & Coordination," *Appeal No. 05AA086, Programme Update no. 2, Period covered*, June to December, 2005.
- [5] "White Paper on Emergency Communications," *Prepared by the Space & Advanced Communications Research Institute (SACRI)*, George Washington University, January 5, 2006.
- [6] Dilley, M., et al., "Natural Disaster Hotspots: A Global Risk Analysis, Synthesis Report," *International Bank for Reconstruction and Development/The World Bank and Columbia University*, March 2005.
- [7] [www.hapcos.org](http://www.hapcos.org)
- [8] [www.usehaas.org](http://www.usehaas.org)
- [9] [www.capanina.org](http://www.capanina.org)
- [10] [www.elec.york.ac.uk/comms](http://www.elec.york.ac.uk/comms)
- [11] Tozer, T., "High Altitude Platforms for Broadband - Pie in the sky?" IEEE Wireless Broadband Conference 2006, The University of York and SkyLARC Technologies, 2006.
- [12] Falletti, M., et al., "Integrated Services from High-Altitude

Platforms: A Flexible Communication System," *IEEE Communications Magazine*, February 2006.

[13] Akalestos, K., et al., "Emergency Communications from High Altitude Platforms," International Workshop on High Altitude Platform Systems - Athens 2005.

[14] D. Grace, et al., "Integrating Users Into the Wider Broadband Network Via High Altitude Platforms," IEEE Wireless Communications, Oct. 2005.

[15] M. Stephen, T. Makrinos, "High Altitude Airships for Homeland Security - Commercial and Military Operations," CACI Technologies Incorporated, NJ, USA, Feb. 2005.

## AUTHORS



**IsraelRomualdo Palma Lázgar** was born in Mexico, in 1980. He received his Electronic Engineer degree from the Instituto Tecnológico de Celaya, Mexico, in 2002, and the Diploma in Networks & Telecommunications Services at La Salle-Universitat Ramon Llull, in Barcelona, in 2004. Since 2004, he has been with the Signal Theory & Communications Department

of Universitat Politècnica de Catalunya, and member of the Control, Monitoring, & Communications Group, in Barcelona, Spain. He is working towards his PhD degree with focus on radiowave propagation, wireless systems, satellite communications, and multicarrier transmission. He is mainly interested in developing wireless propagation channel modelling simulation of broadband communication systems based on HAPs – as an integration to radio-access technologies for the design solutions over heterogeneous wireless networks). Member of the IEEE.



**José Antonio Delgado Penín** (IEEE M.1971;S M 1989) Full Professor (Catedrático de Universidad 1984) in the UPC. He held different positions related with the teaching and research in UPM ( Madrid), Polito (Politecnico di Torino), CNET (Lannion and Paris), Philips (Eindhoven), Kent and Manchester Universities (UK), Univ. California (UCLA). He was visiting Prof. in Cinvestav (Mexico) and PUC (Chile), and Polito (Italia) long time. At the present, he is interested in Wireless communications, HAPS, statistical channel modelling, and performance radio communications systems simulation. He holds several awards collectives and personal (IEEE Millennium Medal 2000).

# Criptografía Basada en Atributos

Javier Herranz

Dept. Matemàtica Aplicada IV, Universitat Politècnica de Catalunya

E-mail: jherranz@ma4.upc.edu

## RESUMEN

Con el progreso constante de las tecnologías digitales, se produce un rápido incremento de la información confidencial que debe gestionarse de manera correcta. La criptografía ofrece herramientas seguras y eficientes para asegurar autenticidad, integridad y confidencialidad en el mundo de la información digital. Sin embargo, la criptografía habitual considera un escenario concreto de comunicación entre dos usuarios, un emisor y un receptor.

Este escenario no cubre algunas de las situaciones prácticas que pueden aparecer en aplicaciones reales. Por eso, se están desarrollando nuevas técnicas criptográficas para hacer frente a estos nuevos escenarios. Un ejemplo es la criptografía basada en atributos, que fue introducida en 2005. En este trabajo hablaremos sobre este nuevo concepto: qué es, qué aplicaciones tiene, qué protocolos concretos se han propuesto, qué resultados hemos obtenido en la UPC, qué puntos quedan por resolver, etc.

## 1. INTRODUCCIÓN

La criptografía es una rama de la ciencia que puede situarse entre las matemáticas y la informática. El objetivo de la criptografía es proporcionar y analizar métodos que ofrezcan confidencialidad, integridad y autenticidad a las comunicaciones (digitales) entre usuarios. Hasta 1976, la criptografía que se utilizaba era simétrica: las dos personas que se comunican de manera confidencial / autenticada deben compartir una clave secreta común,  $K$ , que deben haber acordado previamente, de manera segura. Para cifrar y/o autenticar un mensaje, el emisor aplica un algoritmo que depende de  $K$ . Para descifrar y/o verificar la autenticación, el receptor aplica otro algoritmo que también depende de  $K$ .

La criptografía simétrica tiene ventajas, como su simplicidad conceptual o la existencia de protocolos muy eficientes (por ejemplo, AES) para llevar el concepto a la práctica. Sin embargo, tiene varios inconvenientes que limitan su aplicación directa en un escenario digital tan amplio y ubicuo como el que ofrece Internet. El primer inconveniente tiene que ver con la necesidad de acordar de manera secreta una

clave común: ¿cómo puede acordarse una clave secreta, sin un encuentro físico real, en un entorno tan hostil como Internet, donde cada comunicación corre el riesgo de ser escuchada por usuarios deshonestos? El segundo inconveniente, quizás más serio aún, está relacionado con la cantidad de información secreta que debe almacenar cada usuario: si un usuario  $A$  debe compartir una clave secreta  $K_{AB}$  con cada usuario  $B$ , entonces deberá almacenar tantas claves secretas como usuarios con los que quiera comunicarse. En un mundo digital tan global como la Internet actual, esto sería completamente inviable.

Para evitar estos problemas, Diffie y Hellman [3] introdujeron en 1976 el concepto de criptografía *asimétrica* (o *criptografía de clave pública*). Cada usuario  $A$  genera un par de claves asociadas: una clave secreta  $sk_A$  y una clave pública  $pk_A$ . El usuario  $A$  almacena  $sk_A$  de manera segura (de hecho, ésta es la única información secreta que deberá almacenar), y publica  $pk_A$  para que el resto de usuarios tengan acceso a ella. Paralelamente,  $A$  debe registrar  $pk_A$  ante una autoridad de certificación, que asegura que la clave es válida mediante la emisión de un certificado digital. Si un usuario  $B$  quiere enviar un mensaje confidencial a  $A$ , después de verificar que la clave pública  $pk_A$  tiene un certificado válido, aplica un algoritmo de cifrado que depende de  $pk_A$ . El texto cifrado resultante sólo puede ser descifrado mediante un algoritmo que depende de la clave secreta  $sk_A$ ; por tanto,  $A$  es el único que podrá acceder al mensaje original. De manera similar, si  $A$  quiere autenticar un mensaje, puede aplicar un protocolo de firma que depende de  $sk_A$ ; la firma resultante, como sólo puede haber sido calculada por  $A$ , autentifica el mensaje. La corrección de la firma puede ser verificada por cualquier usuario, mediante un protocolo que depende de  $pk_A$ .

La aparición de la criptografía de clave pública supuso una auténtica revolución, puesto que permitió solucionar los problemas más graves de la criptografía clásica. Sólo faltaba encontrar protocolos concretos que llevasen el concepto a la práctica. La primera propuesta, en 1978, fue el sistema RSA [4]. Desde entonces, han aparecido muchos otros protocolos de cifrado de clave pública (confidencialidad) y de firma digital (integridad y autenticidad). Todos ellos deben satisfacer unas ciertas propiedades de seguridad; por ejemplo, que sea imposible (o inviable con la capacidad de cálculo actual)



descifrar un texto cifrado sin el conocimiento de la clave secreta correspondiente. Los protocolos no pueden satisfacer estas propiedades de manera absoluta; lo que se demuestra es que romper estas propiedades de seguridad (y atacar así el protocolo) es equivalente a resolver algún problema matemático computacional que se considera extraordinariamente difícil / costoso. Por ejemplo, el problema de factorizar un número entero que sea el producto de dos números primos muy grandes, o el problema de calcular el logaritmo discreto en un grupo cíclico con orden muy grande.

La criptografía de clave pública ofrece soluciones a muchas situaciones prácticas de hoy en día. Por ejemplo, el sistema RSA se utiliza en programas para cifrar y firmar correos electrónicos, para pago seguro por Internet, para conexión segura a servidores web, etc. Sin embargo, el uso masivo de las tecnologías de la información digitales, en ámbitos muy diversos, da pie a nuevas situaciones y nuevos problemas en los que hay que proteger información confidencial, pero en un escenario diferente al clásico escenario de un emisor y un receptor. Consideremos el siguiente ejemplo. Un consorcio de hospitales desea mantener un sistema informático para almacenar y gestionar los datos médicos de los pacientes: sus diagnósticos, tratamientos, resultados de análisis, etc. Evidentemente, se trata de información confidencial que debe estar al alcance de algunos actores (médicos, enfermeros) solamente, y no siempre: un enfermero de un hospital no debería tener acceso a la información sobre pacientes que no estén a su cargo, o incluso a alguna parte de la información de sus propios pacientes. La solución debe utilizar algún tipo de mecanismo criptográfico para cifrar la información. Sin embargo, si se consideran esquemas de cifrado de clave pública convencionales, únicamente, el sistema resultante es muy ineficiente. Veamos un ejemplo: supongamos que la totalidad de la información relativa a un paciente  $P_i$ , con problemas coronarios, ingresado en un cierto hospital  $H_j$  debe estar al alcance de:

- el médico de  $H_j$  que es responsable de  $P_i$
- médicos de cualquier otro hospital del consorcio que sean especialistas en enfermedades coronarias,
- los directores de todos los hospitales del consorcio.

A su vez, una parte de la información relativa a  $P_i$  (por ejemplo, el tratamiento a seguir) debe estar al alcance de los enfermeros que se ocupan de  $P_i$ , y también de los encargados de la sección farmacéutica del hospital  $H_j$ .

Si pensamos en una solución basada en criptografía de clave pública convencional, donde cada actor tiene su par de claves secreta y pública, cada vez que se introdujese información en el sistema, debería cifrarse con las claves públicas de todos los actores autorizados. Es decir, los resultados de los análisis de  $P_i$ , su diagnóstico, su tratamiento, etc. deberían cifrarse con las claves públicas del médico responsable de  $P_i$ , de todos los especialistas en enfermedades coronarias, de todos los directores de hospitales. La información relativa al tratamiento también se debería cifrar con las claves públicas de los enfermeros que se ocupan de  $P_i$  y de los farmacéuticos de  $H_j$ . Como resultado, el sistema debería almacenar una cantidad de información cifrada demasiado elevada, lo cual haría bastante inviable su implementación.

Para resolver este tipo de problemas, se ha introducido recientemente el concepto de criptografía basada en atributos. La idea es la siguiente: cada usuario del sistema tiene un conjunto de atributos  $AT$ , y recibe de un servidor central una cierta clave secreta  $sk_{AT}$  que depende de dichos atributos. Por ejemplo, un médico  $M_i$ , especialista en enfermedades coronarias que trabaje en el hospital  $H_s$  recibirá una clave secreta para los atributos  $at_1 = \text{'identidad Mt'}$ ,  $at_2 = \text{'hospital } H_s\text{'}$ ,  $at_3 = \text{'especialista enfermedades coronarias'}$ . Después, cuando haya que cifrar la información confidencial de los pacientes, se escogerá para cada tipo de información una política de descifrado: qué atributos (como mínimo) debe poseer un usuario para poder descifrar y obtener la información confidencial. En nuestro ejemplo, para los resultados de los análisis médicos del paciente  $P_i$ , si el médico encargado de  $P_i$  es  $M_p$ , la política sería: ‘identidad  $M_p$ ’ Ó ‘especialista enfermedades coronarias’ Ó ‘director de hospital’. Para el tratamiento a seguir por el paciente  $P_i$ , si  $E_a$  y  $E_b$  son los enfermeros a cargo de  $P_i$ , a la política de descifrado previa habría que añadir: Ó ‘enfermero  $E_a$ ’ Ó ‘enfermero  $E_b$ ’ Ó ‘farmacéutico’ Y ‘hospital  $H_j$ ’.

En este trabajo discutiremos el concepto de criptografía basada en atributos, dando un enfoque que pretende ser mínimamente formal y divulgativo a la vez. Repasaremos los protocolos que forman parte en un sistema de cifrado basado en atributos; explicaremos qué requisitos de seguridad deben exigirse a estos sistemas; haremos un resumen de las propuestas que se han hecho hasta ahora, enfatizando los aspectos a mejorar; mencionaremos los resultados sobre este tema que hemos obtenido desde el grupo MAK-UPC, así como las líneas de investigación que tenemos abiertas actualmente; por último, mostraremos cómo el concepto de criptografía basada en atributos podría utilizarse para implementar sistemas de

control de acceso que preserven el anonimato de los clientes.

## 2. CIFRADO BASADO EN ATRIBUTOS

Un sistema de cifrado basado en atributos (para simplificar, escribiremos ABE, del inglés ‘attribute-based encryption’) consiste en los siguientes protocolos.

- Inicialización: se generan los parámetros públicos  $pms$  que van a ser comunes a los usuarios del sistema (por ejemplo, el conjunto total de atributos  $U$ ), así como la clave secreta,  $msk$ , del servidor autorizado que repartirá las claves secretas a los usuarios.

- Obtención de claves: un usuario demuestra al servidor autorizado que posee un subconjunto  $AT$  de atributos. Como respuesta, obtiene una clave secreta  $sk_{AT}$ .

- Cifrado: un usuario que quiere esconder una información o mensaje confidencial  $m$  escoge una política de descifrado, que siempre puede describirse como una familia  $\Gamma$  de subconjuntos de  $U$ : aquellos subconjuntos de atributos que, en caso de ser poseídos por un usuario, permitirán el descifrado correcto. El resultado del protocolo de cifrado es un texto cifrado  $C$ .

- Descifrado: un usuario cuyo subconjunto de atributos  $AT$  pertenece a la familia  $\Gamma$  puede utilizar su clave secreta  $sk_{AT}$  para descifrar  $C$  y recuperar la información original  $m$ .

### 2.1. Propiedades de Seguridad Requeridas

Como todos los sistemas de cifrado, un esquema ABE debe satisfacer una condición básica de seguridad: los usuarios que en teoría no están autorizados a descifrar un texto cifrado no deben obtener ninguna información sobre el mensaje que ha sido cifrado. Para formalizar este requisito, se considera la situación más ventajosa para un adversario que intentase atacar el sistema de cifrado: se supone que el adversario escoge dos mensajes  $m_0$  y  $m_1$ , y que el texto cifrado corresponde a uno de esos dos mensajes; el adversario, que no conoce las claves secretas necesarias para descifrar correctamente el texto cifrado, debe intentar adivinar cuál de los dos mensajes ha sido cifrado. Si existe algún adversario que acierta con probabilidad significativamente superior a  $1/2$  (esta probabilidad se obtiene trivialmente con una elección al azar), eso quiere decir que el sistema de cifrado está filtrando alguna información, y por tanto se considera que no es seguro.

En el caso del cifrado basado en atributos, un buen sistema debe ser capaz de resistir *ataques por coalición*. Es decir, si

varios usuarios no cumplen, por separado, la política de descifrado, pero juntando sus atributos sí que se cumple dicha política, tampoco deben ser capaces de descifrar un texto cifrado que corresponda a esa política, aunque comparten sus claves secretas. En nuestro ejemplo de la Sección 1, supongamos que un enfermero que trabaja en el hospital  $H_j$ , pero que no es ni  $E_a$  ni  $E_b$  (los enfermeros a cargo del paciente  $P_j$ ) confabula con un farmacéutico del hospital  $H_s$ , donde  $s \neq j$ . Por separado, ninguno de estos dos usuarios tienen derecho a acceder al tratamiento del paciente  $P_j$ , pero si juntasen sus atributos, se obtendría el subconjunto ‘farmacéutico’ Y ‘hospital  $H_j$ ’, que sí que pertenece a la política de descifrado. El diseño del sistema debe asegurar que estos dos usuarios no serán capaces de obtener ninguna información sobre el tratamiento del paciente  $P_j$ , aunque se intercambien sus claves secretas.

Estos requisitos que hemos explicado informalmente en estos dos párrafos se pueden formalizar con un experimento entre un retador y un adversario  $Adv$ . El adversario puede obtener claves secretas para todos los subconjuntos de atributos que quiera, siempre que ninguno de estos subconjuntos pertenezca a la política de descifrado, naturalmente. Al otorgar esta información al adversario, se modela la situación de un ataque por coalición. El objetivo es que el adversario sea incapaz de adivinar qué mensaje se ha cifrado, entre los dos que escoge él mismo. El juego es el siguiente:

- (1) El adversario  $Adv$  escoge una política de descifrado  $\Gamma$  que quiere atacar.
- (2) El retador ejecuta el protocolo de Inicialización, mantiene  $msk$  en secreto y envía a  $Adv$  la información pública  $pms$ .
- (3)  $Adv$  puede pedir al retador las claves secretas correspondientes a subconjuntos de atributos  $AT_i$  de su elección, siempre que  $AT_i$  no pertenezca a  $\Gamma$ . El retador ejecuta el protocolo de obtención de claves, y envía a  $Adv$  las claves  $sk_{AT_i}$  correspondientes.
- (4)  $Adv$  escoge dos mensajes  $m_0$  y  $m_1$ .
- (5) El retador escoge al azar un número  $b$  (ó el 0 ó el 1), y ejecuta el protocolo de cifrado para el mensaje  $mb$  y la política de descifrado  $\Gamma$ . El texto cifrado resultante,  $C^*$ , se envía a  $Adv$ .
- (6) El paso (3) se repite.
- (7) Finalmente,  $Adv$  devuelve un valor  $b'$  (ó 0 ó 1).

Un adversario tiene éxito en dicho experimento si su probabilidad de acertar, es decir, de obtener  $b' = b$ , es significativamente mayor a  $1/2$ . Si un sistema ABE cumple que ningún adversario con capacidad de cálculo razonable puede tener éxito en este

experimento, entonces el sistema ABE se considera seguro.

### 3. ESTADO DEL ARTE EN SISTEMAS ABE

El concepto de cifrado basado en atributos fue introducido por Sahai y Waters en [5]. En ese artículo, propusieron un sistema ABE pero que funciona sólo para políticas de descifrado de tipo umbral (es decir, se necesita poseer como mínimo  $t$  atributos para poder descifrar) y además el umbral  $t$  se fija al principio y no se puede cambiar según la información que se desee cifrar. Por tanto, el sistema no tiene suficiente flexibilidad como para ser usado en situaciones prácticas (como nuestro ejemplo hospitalario).

El primer sistema ABE que admitía políticas de descifrado más flexibles fue propuesto por Bethencourt, Sahai y Waters en [1]. Las políticas que admite este esquema son las que pueden representarse mediante un árbol en el que las hojas representan los atributos, y cada nodo interno representa una puerta de tipo umbral. De hecho, las políticas de descifrado de nuestro ejemplo hospitalario pertenecen a este tipo, puesto que un ‘Ó’ puede representarse mediante una puerta de umbral en la que el umbral es igual a 1, y un ‘Y’ puede representarse como una puerta de umbral en la que el umbral es igual al número de inputs que tiene la puerta. Sin embargo, este sistema tiene la limitación que, en los árboles que representan la política de descifrado, cada atributo (hoja del árbol) puede colgar sólo de un nodo-puerta de umbral. Esto hace que sea imposible, por ejemplo, realizar una política de acceso como  $\Gamma = (at_1 \text{ Y } at_2 \text{ Y } at_3) \text{ Ó } (at_2 \text{ Y } at_4 \text{ Y } at_5) \text{ Ó } (at_3 \text{ Y } at_5)$ .

Recientemente, otros artículos han propuesto sistemas ABE que solucionan este problema, puesto que permiten realizar políticas de descifrado más generales, de hecho cualquier política de descifrado  $\Gamma$  que sea monótona creciente: si un subconjunto  $AT_1$  pertenece a  $\Gamma$ , y  $AT_2$  contiene a  $AT_1$ , entonces  $AT_2$  debe pertenecer a  $\Gamma$ , también. Estos sistemas ABE, propuestos por Waters [7] y por Daza, Herranz, Morillo y Ràfols [2], utilizan como herramienta el concepto de los esquemas para compartir secretos [6].

#### 3.1. Problemas Abiertos

Todos los sistemas ABE que hemos mencionado en la sección anterior tienen diversos inconvenientes, entre los que se pueden destacar tres.

(1) La longitud de los textos cifrados  $C$  siempre depende (como mínimo, de manera lineal) del número total

de atributos que aparecen en la política de descifrado  $\Gamma$ . Por ejemplo, si una política umbral de descifrado se define como ‘poseer al menos  $t$  atributos de entre una lista (amplia) de  $n$  atributos’, entonces los textos cifrados de los esquemas existentes hasta ahora contendrán al menos  $n$  elementos ( $2(n-t)$  elementos, en el caso del esquema en [2]). Esto es un problema en el caso de querer cifrar información para políticas de descifrado complejas en las que estén involucrados muchos atributos.

(2) En el diseño de todos los sistemas ABE propuestos hasta ahora, se usa un tipo de objeto matemático, los emparejamientos bilineales (*bilinear pairings*, en inglés). Son un tipo de aplicación entre grupos matemáticos, e:  $G \times G \rightarrow G_p$  donde  $G$  y  $G_p$  son grupos con el mismo orden  $q$ , que cumplen que  $e(g^a, g^b) = e(g, g)^{ab}$ , para cualquier elemento  $g \in G$  y cualquier par de números  $a, b \in \{0, 1, \dots, q-1\}$ . Los emparejamientos bilineales se han utilizado mucho en los últimos años para diseñar nuevos protocolos criptográficos, ya que permiten obtener funcionalidades que no se saben implementar sin ellos. El cifrado basado en atributos es un ejemplo. Sin embargo, los emparejamientos bilineales tienen aspectos negativos de cara a la implementación de estos sistemas: sólo se conocen ejemplos de emparejamientos bilineales en grupos asociados a algunas curvas elípticas, y son conceptualmente muy complicados. Esta complicación se traduce en el hecho de que calcular un emparejamiento para dos elementos de  $G$  es bastante ineficiente, y por tanto la eficiencia global de los sistemas ABE actuales no es del todo satisfactoria.

(3) La seguridad de los sistemas ABE propuestos hasta ahora se demuestra en relación a la dificultad de resolver algunos problemas matemáticos computacionales (como siempre en criptografía moderna) que no son muy conocidos, sino que han ido apareciendo ‘on-line’ a la vez que se diseñaban los sistemas ABE. Por tanto, aún no ha habido tiempo para estudiar estos problemas matemáticos detenidamente, para convencerse de que son realmente difíciles de resolver.

Desde el grupo de Matemática Aplicada a la Criptografía (MAK), <http://www-ma4.upc.edu/mak>, del Departamento de Matemática Aplicada IV de la Universitat Politècnica de Catalunya (UPC) estamos trabajando para intentar solucionar alguno de estos inconvenientes. En un trabajo que está actualmente bastante avanzado, junto con Carla Ràfols (MAK-UPC) y Fabien Laguillaumie (Université de Caen, Francia), estamos diseñando un sistema ABE que produce

textos cifrados C de longitud constante, independientemente del número de atributos involucrados en la política de descifrado. Disponemos ya de un sistema seguro para el caso de políticas de umbral (con un umbral t flexible que puede elegir la persona que cifra cada vez), y ahora estamos intentando adaptarlo para que funcione con políticas de descifrado más expresivas, por ejemplo las que se pueden representar mediante un árbol con puertas de umbral.

Además, junto con Paz Morillo, Carla Ràfols, Àlex Escala y Carlos Luna (MAK-UPC) estamos intentando diseñar un sistema ABE que no utilice emparejamientos bilineales. Da la impresión que se trata de un problema realmente difícil, todo un reto para nosotros, pero el tema es apasionante y los conceptos matemáticos que estamos investigando en nuestros intentos son realmente interesantes.

#### **4. OTRA APLICACIÓN DE LOS SISTEMAS ABE: CONTROL DE ACCESO ANÓNIMO**

En esta última sección del artículo vamos a explicar otro ejemplo de una situación real en la que se pueden utilizar los sistemas de cifrado basados en atributos para obtener una solución satisfactoria. Se trata de diseñar un sistema de control de acceso en el que se preserve el anonimato del cliente (autorizado) que accede a determinados recursos. Por ejemplo, en muchos sitios web (clubs, tiendas, foros), los usuarios tienen acceso a diferentes partes / recursos del sitio, en función de su status, de la cuota que hayan pagado, de su edad, etc. La solución habitual es que cada usuario debe completar un proceso de registro, en el que se verifica que dicho usuario cumple los requisitos que dice cumplir. Se guarda un perfil del usuario, de manera que en el futuro, cuando el usuario accede al sitio, debe identificarse (normalmente mediante un nombre de usuario y una contraseña). El servidor detecta el perfil del usuario y automáticamente el sitio web queda personalizado para que el usuario tenga acceso sólo a los recursos para los que está autorizado según su perfil. Desafortunadamente, esta solución no proporciona ningún nivel de anonimato / privacidad al usuario, puesto que el servidor del sitio web controla completamente qué usuario está accediendo en ese momento a qué recursos.

En un sistema ideal, el servidor no debería ser capaz de identificar al usuario; sólo debería estar seguro de que los usuarios que están accediendo a determinados recursos tienen realmente derecho a hacerlo. Es decir, un usuario que accede a un sitio web no tendría que dar más información que la estrictamente necesaria, y que en este caso no es otra

que el hecho de cumplir los requisitos necesarios para acceder a esos recursos. Por ejemplo, para acceder a recursos reservados a personas mayores de edad, un usuario sólo tendría que demostrar que tiene más de 18 años, sin tener que difundir su nombre ni su DNI ni su fecha de nacimiento.

La criptografía basada en atributos ofrece una solución a este problema. El servidor encargado del sitio web hace el trabajo del servidor autorizado que reparte las claves secretas. Cada usuario completa inicialmente un proceso de registro con el servidor, en el que se le asigna una clave secreta en función de sus atributos; algunos ejemplos de atributos en esta situación podrían ser ‘mayor de edad’, ‘miembro del club con categoría A / B / C’, ‘persona con tarjeta de crédito registrada’, etc. Después, para acceder a cada recurso protegido del sitio web, un usuario deberá superar un reto propuesto por el servidor (de manera automatizada, claro): el servidor habrá definido anteriormente una política de acceso para ese determinado recurso, incluyendo los atributos mínimos que debe poseer un usuario para poder acceder. Entonces, se escoge un mensaje  $m$  al azar y se cifra mediante el sistema ABE, tomando como política de descifrado la política de acceso correspondiente. Si el usuario es capaz de descifrar y dar como respuesta el mensaje  $m$ , entonces el servidor estará convencido que el usuario posee los atributos necesarios, y por tanto está autorizado a acceder a ese recurso. De esta manera, el servidor no obtiene ninguna información sobre la identidad de los usuarios que están intentando acceder al sitio, ni sobre qué atributos en concreto posee cada uno de ellos.

A pesar de que hemos usado como ejemplo ilustrativo el caso de un sitio web, existen muchas otras situaciones en las que un sistema de control de acceso anónimo es deseable, por ejemplo para gestionar el control de acceso a infraestructuras que pueden ser críticas: empresas, aeropuertos, edificios gubernamentales, instalaciones militares, etc. En estos casos, el sistema podría implementarse mediante lectores de tarjetas inteligentes o dispositivos RFID que se repartirían entre los usuarios, con las claves secretas ya integradas, dependiendo de sus atributos. Puesto que estos dispositivos tienen unas capacidades computacionales y de memoria bastante más limitadas que un ordenador convencional (usado por los clientes en el ejemplo de los sitios web), cualquier avance en el diseño de sistemas ABE más eficientes será bienvenido. La eficiencia se mide por la longitud de los parámetros públicos  $pms$ , la longitud de las claves secretas  $sk_{Ap}$ , la longitud de los textos cifrados



C, y el número de operaciones para cifrar y descifrar.

## 5. CONCLUSIONES

Este artículo pretende ser una aproximación divulgativa pero precisa al concepto de la criptografía basada en atributos. Como hemos podido observar, este tipo de esquemas puede ser una herramienta muy potente para implementar soluciones satisfactorias a problemas de la vida real que tratan con información y recursos confidenciales.

La criptografía basada en atributos es un concepto que ha sido introducido recientemente, y por tanto hay muchas posibilidades abiertas para trabajar y obtener nuevos resultados que mejoren el estado del arte. Es lo que estamos haciendo (o intentando) desde el grupo de investigación MAK de la Universitat Politècnica de Catalunya.

## REFERENCIAS

[1] J. Bethencourt, A. Sahai y B. Waters. ‘Ciphertext-policy attribute-based encryption’. En Proceedings of IEEE Symposium on Security and Privacy, IEEE Society Press, pág. 321-334 (2007).

[2] V. Daza, J. Herranz, P. Morillo y C. R\`afols. ‘Extended access structures and their cryptographic applications’. Manuscrito disponible en <http://eprint.iacr.org/2008/502> (2008).

[3] W. Diffie y M.E. Hellman. ‘New directions in cryptography’. En IEEE Transactions on Information Theory, vol. 22, núm. 6, pág. 644-654 (1976).

[4] R.L. Rivest, A. Shamir y L. Adleman. ‘A method for obtaining digital signatures and public key cryptosystems’. En Communications of the ACM, vol. 21, pág. 120-126 (1978).

[5] A. Sahai y B. Waters. ‘Fuzzy identity-based encryption’. En Proceedings of Eurocrypt’05, Lecture Notes in Computer Science 3494, Springer-Verlag, pág. 457-473 (2005).

[6] J.L. Villar, C. Padró y G. Sáez. ‘Compartición de secretos en criptografía’. En la revista BURAN, Student Section of IEEE (1997).

[7] B. Waters. ‘Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization’. Manuscrito disponible en <http://eprint.iacr.org/2008/290> (2008).

org/2008/290 (2008).

## AUTOR



Javier Herranz es licenciado en Matemáticas por la Universitat Politècnica de Catalunya (FME, 2000) y Doctor en Matemática Aplicada por la misma universidad (FME, 2005). Ha trabajado como investigador post-doctoral en la École Polytechnique (LIX; Palaiseau, Francia, 2005), en el Centrum voor Wiskunde en Informatica (CWI;

Amsterdam, Holanda, 2006) y en el Institut d’Investigació en Intel·ligència Artificial (IIIA-CSIC; Bellaterra, España, 2007-08). Desde enero de 2009 trabaja como profesor e investigador en el grupo de Matemàtica Aplicada a la Criptografia (MAK) del Departament de Matemàtica Aplicada IV de la UPC, con un contrato Ramón y Cajal. Su investigación se centra en aspectos relacionados con la criptografía, en particular con el diseño y análisis de protocolos criptográficos de firma y cifrado que tengan algunas propiedades especiales: firmas distribuidas, firmas de anillo, cifrado basado en atributos, cifrado con propiedades homomórficas, etc. Más información sobre sus trabajos puede encontrarse en su web personal: <http://www-ma4.upc.edu/~jherranz/>



# GOS: búsqueda visual de imágenes

Silvia Cortés Yuste

Licenciada en Comunicación Audiovisual (UAB)

Ingeniera Técnica de Telecomunicaciones, especialidad Sonido e Imagen (EUETIT –UPC)

Xavier Giró i Nieto

Grupo de Procesado de la Imagen, Teoría del Señal y Comunicaciones, UPC

Ferran Marqués Acosta

Grupo de Procesado de la Imagen, Teoría del Señal y Comunicaciones, UPC

*Las imágenes televisivas utilizadas en este artículo son propiedad de TVC, Televisió de Catalunya, SA, y contienen copyright.*

## RESUMEN

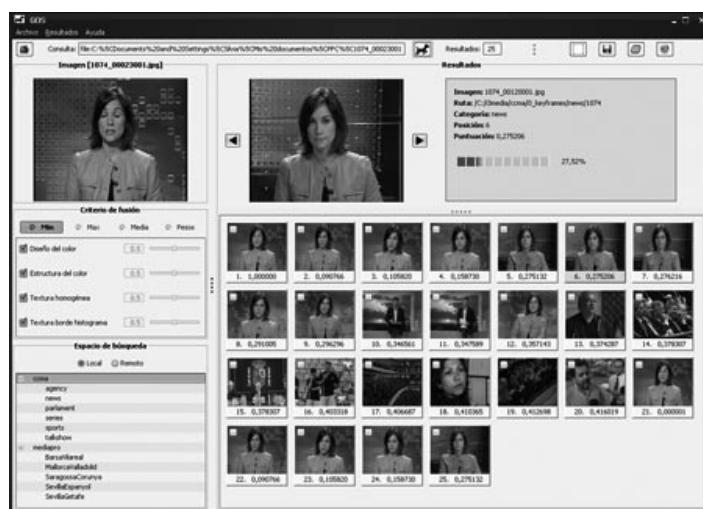
No es fácil encontrar aquello que buscamos. Los sistemas de búsqueda automatizados son máquinas que deben interpretar la información introducida por el usuario para ejecutar una petición y recuperar la información deseada. La tarea del “traductor” es esencial en el proceso, por tanto, una buena interfaz de usuario (GUI) es determinante en el éxito de la búsqueda. Os presentamos el GOS (Graphic Object Searcher), una aplicación que utiliza un sistema de búsqueda visual, para recuperar imágenes similares a otra imagen usada como ejemplo en la consulta. La industria del sector audiovisual está especialmente interesada en el desarrollo de este tipo de herramientas de gestión de contenidos que han de facilitar su trabajo diario.

## INTRODUCCIÓN

Dicen que “una imagen vale más que mil palabras”. Cuántas veces hemos tratado de describir algo muy concreto que buscamos y no hemos encontrado las palabras acertadas para hacer entender a alguien qué demonios queremos. Qué fácil sería tener una fotografía, esa idea hecha imagen y mostrarla: “Esto, esto es exactamente lo que busco”. Si ese “alguien” es una máquina, todavía es más complicado.

En el siglo XXI la imagen le ha ganado el pulso a la palabra. Tanto en el ámbito profesional como en el doméstico, la generación de contenido audiovisual ha aumentado de forma vertiginosa en los últimos años gracias a la digitalización, dificultando el acceso a una información que no da tiempo a ordenar ni catalogar, y por lo tanto, no es fácil de encontrar. A los consumidores de información nos resulta imposible absorber todos estos contenidos, y se evidencia la necesidad de desarrollar nuevas técnicas y herramientas de gestión de información audiovisual, que nos permitan acceder de forma rápida y eficiente a cualquier contenido que nos interese.

Con este objetivo nace el **GOS (Graphic Object Searcher)**, una interfaz gráfica (GUI - Graphic User Interface) destinada a realizar búsquedas de imágenes alojadas en grandes bases de datos a partir de una imagen ejemplo y de unos criterios de búsqueda establecidos por el usuario. El GOS es una aplicación enmarcada dentro de la iniciativa **i3media** [1], proyecto dedicado a la investigación y desarrollo de tecnologías para la creación y la gestión automatizada de contenidos audiovisuales inteligentes. Como herramienta de soporte para estas tecnologías, el GOS proporciona un entorno gráfico amigable para la utilización de nuevos sistemas de indexación y recuperación de contenido audiovisual. Empresas líderes del sector media participan en el consorcio i3media con el objetivo de impulsar diversas áreas de investigación asociadas al contenido audiovisual y mejorar el papel de la industria audiovisual española en el mercado global. Para ello cuentan con la colaboración de grupos de investigadores expertos de universidades y centros tecnológicos que llevan años trabajando con éxito en este campo.



Interfaz del GOS

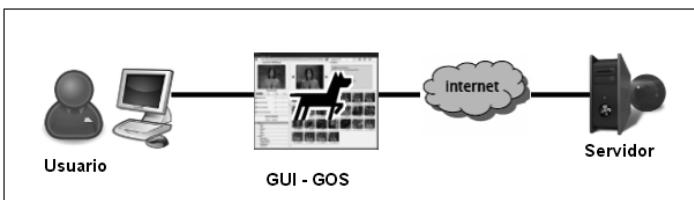


## DIÁLOGO HOMBRE-MÁQUINA: PAPEL DE LA INTERFAZ GRÁFICA DE USUARIO

Hoy en día, personas y ordenadores están condenados a “entenderse”. Las interfaces de usuario son las herramientas clave para establecer esa comunicación entre el hombre y la máquina. El usuario debe transmitir a la máquina lo que desea conseguir, la máquina debe entender la orden y ejecutarla, y finalmente responder al usuario con el resultado del proceso o acción ejecutada.

Todo este diálogo se realiza a través de la interacción que se establece entre ambos actores y se gestiona a través de un elemento intermedio, **la interfaz de usuario (GUI)**. En cualquier sistema de recuperación de información o contenido audiovisual, este diálogo es imprescindible, y de la precisión del “traductor” depende el éxito de los resultados.

El GOS es una GUI desarrollada por el Grupo de Procesado de Imagen (GPI) de la Universidad Politécnica de Cataluña (UPC), pensada para facilitar la utilización de métodos de búsqueda inteligente de imágenes desarrollados por el grupo. La GUI se encarga de realizar la tarea de intermediario entre el usuario y el sistema de búsqueda de imágenes. Por un lado, la interfaz recoge los datos de entrada facilitados por el usuario y necesarios para realizar la consulta, y se pone en contacto con el motor de búsqueda, alojado en un servidor remoto, para que la ejecute. Por otro, una vez realizada la búsqueda,



la interfaz presenta al usuario los resultados obtenidos.

*Funcionamiento del GOS*

La historia de las GUI va unida indiscutiblemente a la evolución de la tecnología. Antes de la explosión tecnológica que se dio en los años 70, los ordenadores eran manejados por personas muy especializadas. El hombre se adaptaba en su totalidad a la máquina, y no al revés (la antítesis de la usabilidad). A partir de los años 80, la proliferación del ordenador personal dirigido a un usuario final no experto en informática (para uso comercial, administrativo y empresarial), supone una gran revolución. Nace la necesidad de crear herramientas que faciliten el trabajo con el ordenador: “*el usuario no quiere utilizar una aplicación, quiere*

*hacer su trabajo de la forma más sencilla y rápida posible*” [2].

De ahí, que la disciplina **Interacción entre Humanos y Máquinas (HCI-Human Computer Interaction)** esté adquiriendo cada vez más importancia en el desarrollo de GUI. Enmarcada dentro de las ciencias documentales, la HCI nace con la voluntad de ayudar a mejorar la comunicación entre los usuarios y los sistemas de documentación (informátizados), y se ocupa del análisis y diseño de interfaces entre hombre-máquina, estudiando la creación de productos informáticos que ayuden en la realización de tareas a sus usuarios atendiendo a la facilidad de uso, al tiempo de ejecución, a la evitación de los posibles errores y, en consecuencia, a su satisfacción [3]. Así, la HCI tiene un carácter interdisciplinar, que abarca aspectos humanos, tecnológicos y la comunicación entre ambos.

## ¿CÓMO BUSCAMOS?

La interfaz del GOS utiliza un sistema de búsqueda inteligente para recuperar imágenes similares a la utilizada en la consulta. Este tipo de técnicas de búsqueda pertenecen al conjunto de sistemas conocidos como sistemas CBIR (*Content-Based Image Retrieval*) y permiten recuperar imágenes digitales a partir de atributos visuales, como los colores, las formas o texturas. Dichos atributos se extraen y se representan automáticamente a través de estructuras de datos numéricas, y las imágenes son indexadas sin necesidad de tener asociada ningún tipo de anotación manual. Gracias a estos métodos de anotación e indexación automática, es posible realizar búsquedas sin necesidad de utilizar palabras, usando en su lugar paletas de colores, dibujando o seleccionando una imagen a partir de la cual el sistema pueda recuperar otras similares. La mayoría de estos sistemas utilizan esta última técnica, la **consulta mediante ejemplo (QbE - Query by Example)**, que consiste en realizar la búsqueda a partir de una imagen inicial de consulta y una serie de criterios establecidos por el usuario para encontrar imágenes parecidas en grandes bases de datos.

Los sistemas CBIR surgieron a comienzos de la década de los '90, pero no es hasta finales de esa misma década que se produce un auge en los trabajos de investigación dedicados a este ámbito. Sin embargo, la mayoría trata el tema desde el punto de vista técnico y son pocos los trabajos dedicados a aspectos como la evaluación de su eficacia, el diseño de interfaces o estudio de los usuarios [4]: “Mientras las tecnologías base de los sistemas CBIR evolucionan, desarrolladores e investigadores ignoran la importancia de la interacción entre humanos y máquinas, así como el rol crucial de la interfaz de usuario” [5]

El proyecto del GOS se centra en el diseño de una GUI para un sistema CBIR, para ampliar ese pequeño conjunto de trabajos dedicados a este tema. El método de diseño se centra en el usuario (UCD - *User-centered design*), puesto que el objetivo final de todo sistema de recuperación de imágenes es conseguir que el usuario acceda al contenido que busca. La HCI afirma que la forma de presentar la información al usuario en un sistema de recuperación de contenido hace variar su manera de interactuar [3], y por lo tanto influirá en la utilidad de la herramienta y el grado de satisfacción final que se genere. En este sentido, una aplicación de búsqueda de imágenes como el GOS, debe trabajar con minuciosidad la presentación de la consulta a realizar y la presentación de los resultados obtenidos.

## Tipología de GUI para sistemas CBIR

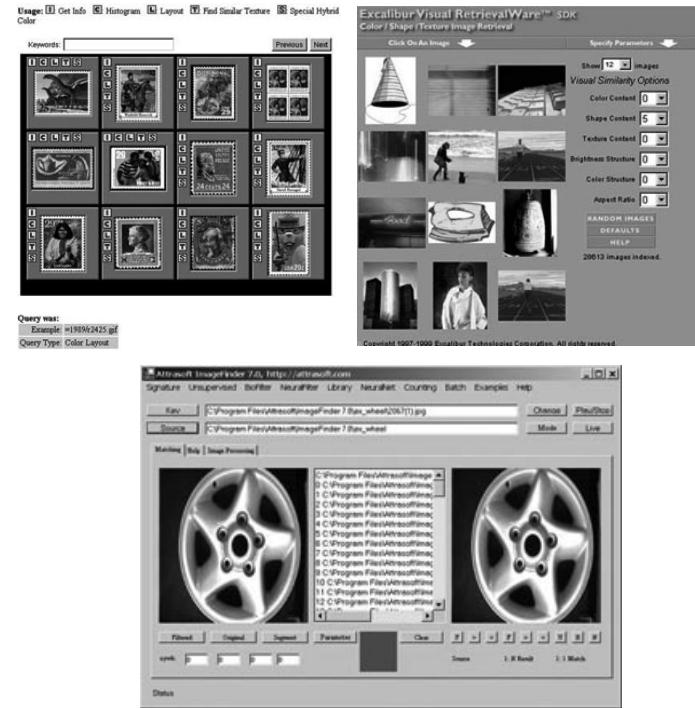
Es precisamente el tipo de usuario al que se dirige la aplicación, lo que marca las dos tendencias existentes en tipos de GUI de sistemas CBIR desarrollados hasta el momento. Por un lado, encontramos las interfaces dirigidas a los académicos e investigadores de este campo, que utilizan interfaces pobres en diseño y usabilidad, pero complejas en su manejo y en sus funcionalidades. Por otro, encontramos aplicaciones destinadas a un público más amplio, productos más comerciales enfocados a un usuario no experto, con diseños más trabajados y atractivos, en general más usables y a su vez más sencillos. Esta diferencia de enfoque en el diseño según el usuario tipo de la aplicación nos sirve para establecer una clasificación de las interfaces de sistemas CBIR, que debido a la coetaneidad de la mayoría de ellas (todas se desarrollaron en los últimos 15 años), es un criterio más interesante que el clásico análisis cronológico.

### · Interfaces comerciales de sistemas CBIR

De manera genérica, vamos a aplicar el calificativo de **interfaces comerciales** a aquellas aplicaciones que están desarrolladas por empresas y destinadas a su comercialización. La mayoría de estas empresas inicialmente se preocupa de implementar los sistemas de búsqueda que posteriormente los clientes adaptarán a sus productos.

El **sistema QBIC** (1995) desarrollado por IBM fue uno de los sistemas CBIR pioneros [6]. El objetivo de IBM era vender simplemente el sistema de búsqueda, por lo que no se preocupó demasiado en diseñar una buena GUI para complementarlo. El punto fuerte de QBIC es su potente motor de búsqueda, que se sigue utilizando en múltiples aplicaciones,

como el buscador de la web del State Hermitage Museum de St. Petersburg de Russia (integrado a través de applets de Java permite realizar búsquedas sobre la colección digital del museo a través del espectro de color (*QBIC Colour Search*) o de formas geométricas (*QBIC Layout Search*) de las obras) [7]. Otras empresas siguieron el ejemplo de IBM, como Excalibur Technologies Corporation con su **Excalibur Visual RetrievalWare** (1997-1999), que desarrolló software para crear aplicaciones de manipulación de imágenes digitales y su contenido visual, extracción de características, indexación y recuperación basada en contenido. Al igual que IBM, Excalibur priorizó las técnicas de búsqueda en detrimento del diseño de una GUI de soporte para su tecnología.

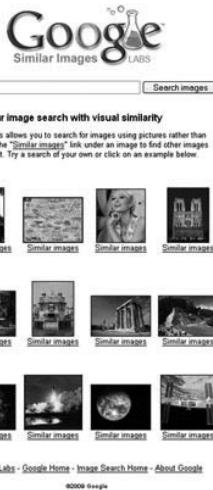


Interfaces QBIC, Excalibur e ImageFinder

La compañía Attrasoft desarrolló un sistema CBIR para Windows que comercializó en tres productos diferentes (*ImageFinder*, *Internet ImageFinder* e *ImageHunt*). La tecnología utilizada era la misma para los tres, pero el diseño de sus interfaces era diferente. Actualmente solo sobrevive **ImageFinder** [8], cuya primera versión data de 1998, y hoy se comercializa la versión 7.0. ImageFinder es una herramienta compleja, no utilizada solamente para realizar búsquedas, sino que también permite el tratamiento y el procesado de imágenes, por lo que el destinatario final es un usuario experto. Así, en su GUI prima la funcionalidad por encima del diseño y la usabilidad, descuidando la estética, que se ha quedado bastante anticuada.

Actualmente una gran parte de estas aplicaciones proliferan en la web, presentadas como valor añadido a los tradicionales buscadores de imágenes que utilizan como método de búsqueda la recuperación basada en texto. Este método anota las imágenes a través del texto HTML del documento que las contiene, basándose en la suposición de que una imagen en una página web está semánticamente relacionada con el texto que la rodea. Las GUI en la web utilizan un formulario para los datos de entrada y una lista o tabla de resultados con las imágenes obtenidas como datos de salida. Las más complejas aumentan las funcionalidades de la aplicación con lenguajes de programación que permiten asociar programas o procesos ejecutables con los componentes de una página.

Google ha lanzado **Similar Images** (2009) en su Google Labs



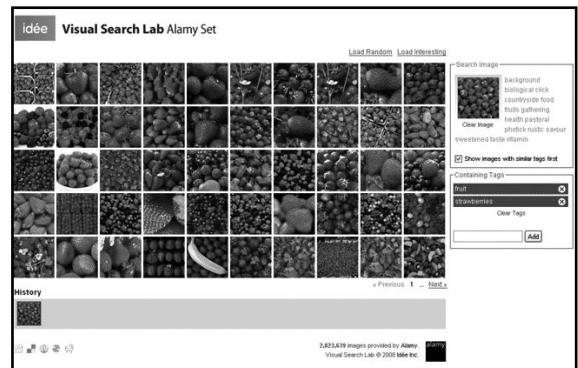
[9], como funcionalidad extra a su clásico buscador de imágenes en la web. Su aparición supone un gran paso en la incorporación de estos sistemas de búsqueda inteligentes en buscadores de imágenes convencionales, por la enorme repercusión de cualquier producto de la marca Google. La gran ventaja de esta herramienta es su fácil manejo, ya que el

usuario está familiarizado con la estética y el tipo de buscador creados por Google (el diseño de la aplicación no ha sufrido ninguna variación respecto al buscador tradicional).

Sin embargo, paralelamente a *Similar Images* se están desarrollando proyectos mucho más interesantes e innovadores como **Picollator** (2008), **Idée Labs** (2008) o **GazoPa** (2008), aunque de momento, con menor repercusión.

**Picollator** es un buscador de imágenes especializado en el reconocimiento facial, está desarrollado por la empresa de origen ruso Recogmission LLC y se encuentra en versión beta en la web [10]. **Piximilares**, **PixID** y **TinEye** son productos creados por la compañía Idée Inc., empresa especializada en software de reconocimiento avanzado de imágenes y búsqueda visual [11]. **PixID** es un servicio de monitorización para identificar las imágenes que se utilizan en las publicaciones impresas y en Internet, especialmente dedicado a la detección de copias (cumplimiento de licencias, malos usos, reclamar pagos por usos no autorizados, etc.). **TinEye** es una herramienta de búsqueda de imágenes en la web

que es capaz de detectar dónde y cómo una imagen es utilizada (incluso si ha sido modificada) [11]. **Piximilares** es una herramienta especializada en búsqueda de imágenes pertenecientes a una colección. Trabaja con el método de búsqueda por imagen similar o por selección de varios colores. **Idée** es otra muestra del creciente interés que los sistemas CBIR están adquiriendo en el mercado. Utilizan un sistema de búsqueda propietario llamado Visual Search.



Interfaces de Picollator e Idée Labs.

**GazoPa**, desarrollado por la empresa Hitachi, es otra de las aplicaciones que utiliza QBIC de IBM como sistema de búsqueda inteligente. Se encuentra en fase beta, pero explota el potencial de este tipo de buscadores en el mercado y la clara tendencia actual a proporcionar herramientas cada vez más elaboradas y orientadas al consumidor de contenidos de la web (un público masivo) [12]. La GUI de GazoPa recuerda ligeramente a Google, pero con un diseño muy superior, elaborado y atractivo. Incorpora una herramienta de dibujo y recorte/retoque de imágenes en Flash, además de la posibilidad de buscar dentro del portal Amazon con un filtro de resultados que permite visualizar productos a la venta en esta tienda on-line con imágenes similares a la utilizada en la consulta.



Interfaz de GazoPa

El negocio en la web viene de la mano de compañías especializadas en proveer de contenido visual a determinados consumidores (generalmente profesionales del sector audiovisual). Estas empresas disponen de grandes colecciones de imágenes, generalmente catalogadas manualmente, y utilizan sistemas propietarios para indexar y recuperar estas imágenes a través de palabras clave (las colecciones suelen ser actualizadas periódicamente). *Getty Images*, *Corbis* o *Masterfile* [13] son algunos ejemplos, y algunas de ellas como *Masterfile*, ya incorporan en sus buscadores el método de consulta por imagen ejemplo. Pero este tipo de empresas es solo una pequeña muestra de las posibilidades de negocio de estos sistemas de búsqueda. Tiendas on-line como **Like.com** [14] son propuestas innovadoras, que apuestan por el uso de este tipo de tecnología orientado a un público masivo, el consumidor particular. El portal Like.com ha acuñado el término de *Visual Shopping* (compra visual) y se describe como “el primer motor de búsqueda realmente visual, donde el contenido de las fotos se utiliza para buscar y recuperar objetos similares”. Su GUI está completamente orientada a facilitar la compra, incluyendo fotografías de personajes famosos, donde aparecen estrellas (ícono del portal) sobre la ropa que visten y que permiten con un solo clic, iniciar la búsqueda de la prenda (puesta a la venta).

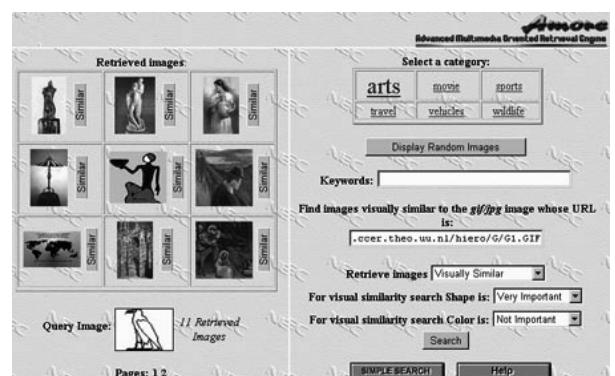
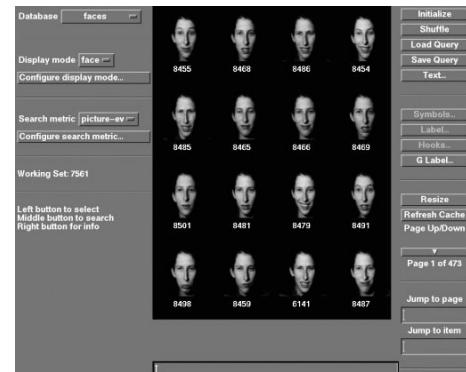


Portal Like.com

## · Interfaces académicas de sistemas CBIR

En contraposición de las interfaces comerciales, otorgamos el calificativo de **interfaces académicas** a aquellas aplicaciones desarrolladas por universidades y centros de investigación, que en principio, no tienen ningún tipo de interés en su comercialización. Generalmente, las interfaces de estos sistemas no son especialmente atractivas en diseño, preocupa más su funcionalidad. Para este tipo de sistemas, la web es el banco de pruebas preferido, porque permite una mayor difusión del trabajo realizado. Además, la mayoría de sistemas CBIR académicos elaboran proyectos completos, que abarcan desde herramientas para la anotación e indexación de las imágenes, hasta las técnicas de búsqueda.

Uno de los pioneros fue **Photobook** (1994), desarrollado en el MIT Media Laboratory de Cambridge [15] como un conjunto de herramientas interactivas para la búsqueda y navegación de imágenes y de secuencias de imágenes. Photobook es el primer sistema que subraya la importancia de la interactividad entre el motor de búsqueda y el usuario para conseguir resultados satisfactorios. Su interfaz utiliza la tecnología Motif [16] para conseguir una interactividad realmente eficaz con el usuario, aunque no presta ninguna atención al diseño de la aplicación. El sistema de búsqueda de Photobook se complementa con la herramienta de anotación **FourEyes**, que utiliza el mismo tipo de interfaz.

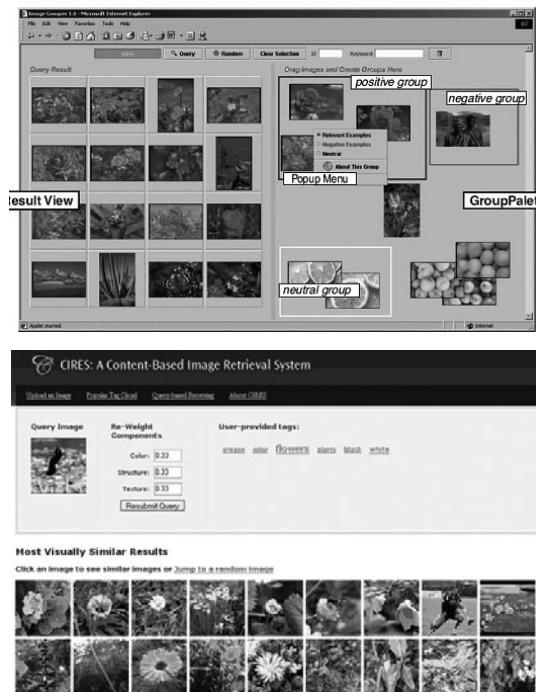


Interfaces de Photobook y Amore



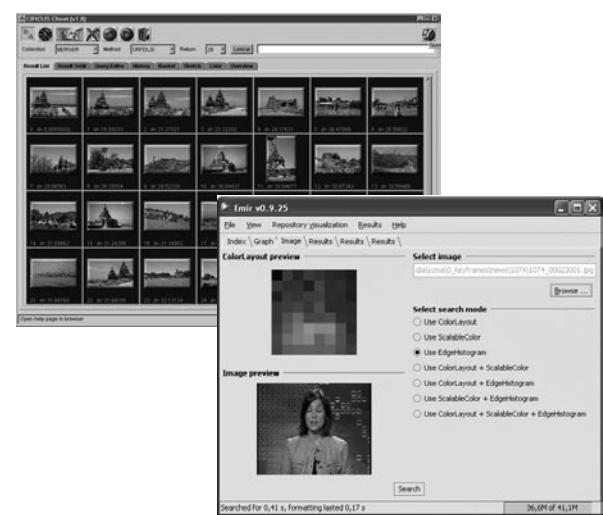
El sistema **AMORE** (*Advanced Multimedia Oriented Retrieval Engine*) (1999), desarrollado por los C & C Research Laboratories NEC USA, es uno de los primeros buscadores de imágenes para web que se preocupó por implementar un método de visualización de resultados más cuidado, tendencia que se extendería a partir del año 2000, cuando los sistemas de búsqueda de imágenes empiezan a mostrar más interés por el uso más eficiente del espacio de la pantalla [17].

Otros proyectos destacados del ámbito académico son **MARS** (Multimedia Analysis and Retrieval Systems) de las Universidades de Illinois y California (1997-2003) [18] y **Cires** (Content Based Image REtrieval System) de la Universidad de Texas (2002-2007) [19]. MARS destaca por ser el primero en introducir técnicas de *relevance feedback* en este tipo de aplicaciones (sistemas de feedback que utilizan el resultado de una primera consulta para refinar las búsquedas con consultas posteriores) y su interfaz supone una evolución en el tratamiento visual de la información en pantalla [17]. Los desarrolladores de MARS también son los responsables de **ImageGrouper**, un sistema CBIR basado en consultas de grupos de imágenes. Cires fue junto con MARS, uno de los sistemas CBIR de referencia por su implementación de métodos *relevance feedback* en sus interfaces, pero no ha evolucionado su propuesta, a diferencia de MARS con su *ImageGrouper*. Cires ha sorprendido con un cambio de su GUI hacia un diseño más elaborado estéticamente pero más simple en funcionalidades (ha eliminado la funcionalidad de *relevance feedback* en su aplicación).



Interfaces de ImageGrouper y Cires

También podemos encontrar proyectos interesantes de código abierto, que promueven la colaboración entre diversos centros de investigación e universidades. Un ejemplo es **CIRCUS** (*Content-based Image Retrieval and Consultation User System*), proyecto dedicado al desarrollo de un sistema para la recuperación de imágenes de colecciones distribuidas, heterogéneas y anotadas, que destaca por su elaborada arquitectura cliente/servidor a través del protocolo abierto de comunicación MRML (*Multimedia Retrieval Markup Language*), disponible bajo licencia pública GNU [20]. El trabajo realizado en CIRCUS (colaboración entre Laboratoire de communications audiovisuelles (LCAV), Ecole Polytechnique Fédérale de Lausanne (EPFL), y el Computer Vision Group University of Geneva) tiene su continuidad en VIPER, un proyecto para la recuperación de información multimedia que ha desarrollado el GIFT, un paquete de código abierto para implementar sistemas CBIR utilizando el método QBE (*Query-by-Example*) y muy fácil de integrar en interfaces web. Y bajo el auspicio de SourceForge.net, se ha desarrollado **Caliph & Emir** [21], un kit de herramientas en Java para la anotación y recuperación de imágenes y fotografías digitales basándose en los descriptores visuales definidos en el estándar MPEG-7. Aunque SourceForge.net es propiedad de la empresa SourceForge Inc, los proyectos que se gestionan en este sitio web no pueden ser considerados comerciales, ya que contribuyen a la difusión de nueva tecnología. Caliph & Emir utilizan la librería **LIRE** (Lucene Image REtrieval), desarrollada también por SourceForge.net dentro del proyecto Apache Lucene, dedicado al desarrollo de software de búsqueda. LIRE proporciona un sistema para crear índices de imágenes, realizar búsquedas, navegar por estos índices y crear mosaicos de imágenes para su visualización. Todo es código abierto.



Interfaces de CIRCUS y Emir

## GOS

El GOS quiere iniciar una nueva tendencia en la tipología de las GUI de sistemas CBIR, un nuevo tipo de GUI donde convergen características de interfaz comercial e interfaz académica, con aplicaciones destinadas a dar servicio tanto a expertos de la materia como a clientes (empresas del sector audiovisual o consumidor particular).

Empresas como **Mediapro S.L.**, líder del proyecto i3media y dedicada a la producción de contenidos y provisión de servicios media, y la **Corporació Catalana de Mitjans Audiovisuals (CCMA)**, empresa de desarrollo de tecnologías de gestión de contenidos, son los principales destinatarios del GOS. La herramienta permite realizar búsquedas con imágenes de ejemplo en sus archivos de vídeo para crear nuevos contenidos, agilizando el proceso de documentación y recuperación de contenido, al visualizar rápidamente los resultados de la búsqueda. A su vez, el GOS también sirve de soporte a la UPC para experimentar y avanzar en el desarrollo de técnicas de gestión y recuperación de contenido audiovisual.

¿Qué puede hacer el GOS? El GOS es la GUI de un buscador de imágenes basada en ejemplos, y como tal, permite formular una consulta, ejecuta la búsqueda y visualiza los resultados. Éstas son las funciones básicas que todo buscador implementa. Sin embargo, el GOS pretende aportar un valor añadido que diferencie la aplicación del resto de productos disponibles hasta el momento. Con el objetivo de facilitar el trabajo de búsqueda y recuperación de imágenes al usuario final, el diseño de la GUI ha cuidado enormemente los tres aspectos fundamentales que dan forma a una aplicación: el diseño visual o gráfico, la usabilidad y la tecnología. La mayor dificultad es aunar funcionalidad y estética, dotar a los elementos gráficos de una “personalidad” propia, un *look&feel* característico de la aplicación, pero sin perder de vista las necesidades reales del usuario, y paralelamente ir sorteando las limitaciones de la tecnología.

Una buena distribución de los elementos en la pantalla es la base para una utilización óptima de la herramienta. ¿De qué nos sirve tener potentes funcionalidades disponibles en la GUI, pero no fáciles de utilizar o de localizar? Como requisito imprescindible, un buscador de imágenes mediante consulta por ejemplos debe diferenciar siempre dos zonas en pantalla, la zona destinada a la formulación de la consulta y la zona destinada a la visualización de los resultados, creando un flujo de acciones coherentes y lógicos en su ejecución.

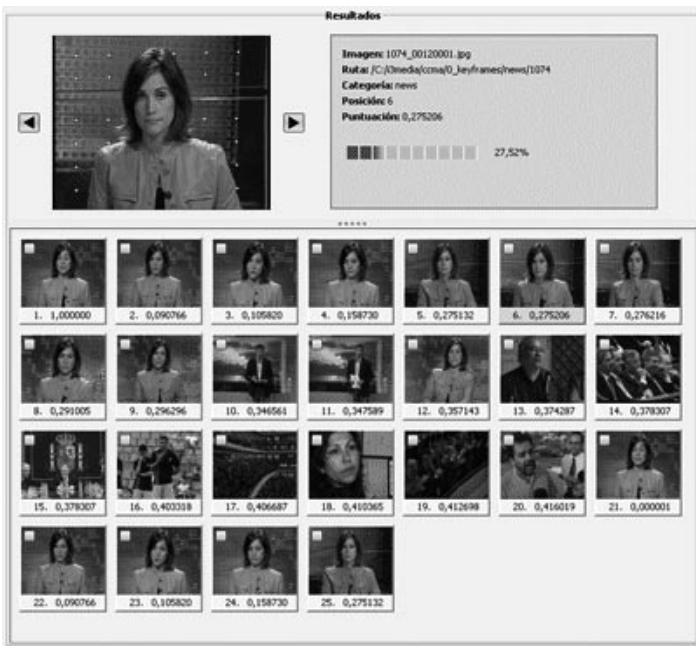


Organización de elementos en pantalla del GOS

En el GOS, se ha destinado la zona superior de la aplicación al menú principal y la barra de herramientas (zona roja de la figura). De esta manera, todas las funcionalidades son accesibles al usuario de forma rápida. En la zona izquierda ubicamos el área de consulta (zona amarilla de la figura), donde el usuario podrá configurar la petición de búsqueda. Siguiendo los pasos verticalmente, la GUI permite cargar la imagen de consulta, seleccionar el criterio de fusión y los descriptores visuales a tener en cuenta en la búsqueda, y el ámbito de la búsqueda, escogiendo la categoría dónde buscar. El resto de pantalla (zona verde de la figura) se utiliza para la visualización de los resultados, que vuelve a reproducir el esquema anterior: en la parte superior del área de resultados se coloca una zona destinada a la presentación de información destacada (una imagen resultado junto a un área de texto con la información asociada a esta imagen), y en la parte inferior se visualizan todos los resultados de la búsqueda en forma de parrilla de imágenes, ordenados según el grado de similitud con la imagen consulta (en orden decreciente).

Además, la aplicación dispone de un método para refinar las búsquedas, ya que cualquier imagen resultado puede ser utilizada para lanzar una nueva consulta haciendo simplemente doble clic sobre ella, y permite navegar cómodamente a través de la lista de resultados obtenidos. Para obtener más información sobre una imagen resultado, el usuario puede darle foco haciendo clic sobre la diapositiva. Cuando una diapositiva obtiene el foco, su fondo cambia de color, de blanco a azul, y su borde de gris a un azul más oscuro, y automáticamente pasa a ocupar el lugar de la imagen destacada del panel superior. La interfaz dispone de dos métodos alternativos

para navegar por los resultados y visualizarlos como imágenes destacadas. Por un lado, encontramos dos típicos botones de navegación a la izquierda y derecha de la imagen destacada, para recorrer los resultados (anterior y posterior respectivamente) con un solo clic, sin necesidad de desplazarse a la zona de la parrilla para clicar sobre la imagen deseada. Por otro, a través de la rueda central del ratón, el usuario puede desplazarse por los resultados de forma rápida y sin necesidad de ir realizando clics.



*Visualización de los resultados en el GOS*

El GOS también permite al usuario seleccionar aquellos resultados que sean de su interés y guardarlos en un archivo XML. Esta selección puede realizarse de modo individual, haciendo clic sobre el checkbox de la diapositiva o en grupo, a través de la herramienta de selección rectangular.

La elección de Java para desarrollar la aplicación, nos permite crear una herramienta muy versátil, con más posibilidades de evolución e integración con el resto de aplicaciones y herramientas, de la UPC y del proyecto i3media. Además, la utilización del estándar MPEG-7 en ficheros XML favorece la interoperabilidad, entre las diferentes piezas que componen el sistema, y la manipulación de la información visual que contienen las imágenes.

El GOS está disponible en línea a través del software de Java Web Start, que permite la descarga y ejecución de la aplicación desde su web [22], para jugar un poco con él y descubrir sus posibilidades.

## CONCLUSIONES

A juzgar por el resultado final, podemos avanzar que el GOS supone un paso más en la evolución de las GUI para sistemas CBIR. La interfaz es fácil de utilizar, bien estructurada, pensada para simplificar el trabajo de recuperación de imágenes al usuario, y a la vez incorpora técnicas de búsqueda y recuperación de imágenes pioneras. Como valor añadido, la aplicación permite su ejecución en entorno remoto y local, aunque el objetivo final es acabar ofreciendo el servicio de búsqueda de manera remota, siguiendo la tendencia de la gran mayoría de proveedores de contenido media. A partir de servicios web, cualquier usuario en red puede tener acceso a múltiples aplicaciones, de forma cómoda y sencilla. El enorme volumen de información disponible, especialmente en la industria audiovisual, precisa de herramientas que se adapten a este tipo de arquitecturas, modulares e integrables.

Pero esto es solo el principio. El GOS se encuentra en una primera versión beta, y su desarrollo continua dentro del proyecto i3media, para ir incorporando nuevas técnicas de recuperación de contenido audiovisual que trabaja la UPC. El siguiente paso que se plantea el GOS es la ampliación de los métodos de búsqueda que soporta la interfaz. Esta ampliación deberá realizarse en dos fases: la primera, y más inmediata, es la implementación de la búsqueda de imágenes basada en regiones, y la segunda, dar la posibilidad de utilizar y combinar más de un tipo de consulta (basada en texto, imágenes y regiones). Será interesante seguirle el rastro. Podéis seguir la evolución del proyecto en el blog: <http://bitsearch.blogspot.com/search/label/GOS>.

## REFERENCIAS

PFC: Cortés Yuste, Silvia. **Interfaz Gráfica de Usuario para la Búsqueda de Imágenes basada en Imágenes**. EUETIT-UPC, Junio-2009.

([http://gps-tsc.upc.es/imatge/\\_Xgiro/teaching/thesis/2008-2009/SilviaCortes/memoria.pdf](http://gps-tsc.upc.es/imatge/_Xgiro/teaching/thesis/2008-2009/SilviaCortes/memoria.pdf))

[1] Web i3media: <http://www.i3media.org>

[2] Roe, Benjamin. **Diseño de interfaces de usuario usables**. Publicado en Mundo Geek. Diciembre 2004

(<http://mundogeek.net/traducciones/interfaces-usuario-usables/gui.html>)

- [3] Marcos, Mari-Carmen. “**HCI (Human computer interaction): concepto y desarrollo**”. En: El profesional de la información, 2001, junio, v. 10, n. 6, pp. 4-16.  
*(http://www.elprofesionaldelainformacion.com/contenidos/2001/junio/1.pdf)*
- [4] Pérez Álvarez, Sara. **Análisis de usabilidad de sistemas CBIR | User friendliness of CBIR systems analysis** -. Abril 2008  
*(http://biblioteca.universia.net/ficha.do?id=34103694)*
- [5] Colin C. Venters. **User Interface Design & Evaluation for a Content-Based Image Retrieval System**. Department of Information & Library Management, University of Northumbria at Newcastle, England.  
*(http://www.bcs-hci.org.uk/hci1998/C51/)*
- [6] Sistema QBIC: *http://www.qbic.almaden.ibm.com/*.
- [7] State Hermitage Museum: *http://www.hermitagemuseum.org/cgi-bin/db2www/qbicSearch.mac/qbic?selLang=English*.
- [8] ImageFinder de Attrasoft: *http://www.attrasoft.com/imagefinder70/*
- [9] Similar Images de Google: *http://similar-images.googlelabs.com/*
- [10] Picollator: *http://www.picollator.com/*.
- [11] Labs Idée: *http://labs.ideeinc.com/*; Tineye: *http://tineye.com/*
- [12] Gazopa de Hitachi: *http://www.gazopa.com*
- [13] Getty Images (*http://www.gettyimages.com*); Corbis (*http://pro.corbis.com/*); Masterfile (*http://www.masterfile.com/info/products/simsearch.html*)
- [14] Like.com: *http://www.like.com/aboutus.py*
- [15] Photobook: *http://vismod.media.mit.edu/vismod/demos/photobook/index.html*
- [16] Pentland, R.W.; Picard, S.; Sclaroff, S. **Photobook: Content-Based Manipulation of Image Databases**. The Media Laboratory, Massachusetts Institute of Technology. June 1995
- [17] Nakazato, Munehiro and Manola, Ljubomir and Huang, Thomas S.(2003) **ImageGrouper: a group-oriented user interface for content-based image retrieval and digital image arrangement**. Journal of Visual Languages & Computing, 14 (4). pp. 363-386. ISSN 1045926X
- [18] MARS: *http://www.ifp.illinois.edu/~qitian/MARS.html*
- [19] CIRES: *http://cires.matthewriley.com/*
- [20] CIRCUS: *http://viper.unige.ch/doku.php/demos#content-based\_image\_retrieval*
- [21] Caliph&Emir de Sourceforge.net: *http://sourceforge.net/projects/caliph-emir/*, *http://caliph-emir.sourceforge.net*
- [22] Web del GOS: *http://gps-tsc.upc.es/imatge/i3media/gos/*

## AUTORES



*Silvia Cortés Yuste*, es licenciada en Comunicación Audiovisual e Ingeniera Técnica de Telecomunicaciones, especialidad Sonido e Imagen. Trabaja en proyectos relacionados con el desarrollo de herramientas de gestión de información y de nuevas aplicaciones y servicios para contenido audiovisual en diferentes medios (televisión, IPTV, web), en materia de diseño, usabilidad y desarrollo de software.



*Xavier Giró i Nieto*, investigador del Grupo de Procesado de la Imagen (GPI) y profesor del grado en Ingeniería de Sistemas Audiovisuales en la Escuela de Ingeniería de Terrassa de la UPC. Es titulado en Ingeniería de Telecomunicaciones por la UPC en 2000 y sigue desde el 2002 el programa de Doctorado del Departamento de Teoría de la Señal y Comunicaciones. Ha trabajado sobre codificación de imágenes volumétricas en la Vrije Universiteit de Bruselas (VUB), televisión digital en el Sony Development Center Brussels, y indexación automática de imágenes en el GPI y en la Columbia University de Nueva York. Su investigación puede seguirse en el blog <http://bitsearch.blogspot.com>.



*Ferran Marqués Acosta*, realiza su investigación en el Grupo de Procesado de Imagen (GPI) del Departamento de Teoría de la Señal y Comunicaciones (TSC) y su docencia en la Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona (Telecom BCN).

# El software libre y lo sostenible

Rafael Cubarsi\*, Miquel Escudero\*\*

Departament de Matemàtica Aplicada IV, gAGE

Campus Nord, UPC

C/ Jordi Girona, 1-3

08034 Barcelona

E-mail: rcubarsi@ma4.upc.edu\*, escudero@ma4.upc.edu\*\*

Palabras clave: webactores, softdependencia, softcialización, softenibilidad.

## RESUMEN

Del mismo modo que la energía total de un sistema mecánico es la suma de la energía cinética y la potencial, podemos decir que el desarrollo en sostenibilidad es la suma ‘uso de recursos’ más ‘potencialidad de uso’, en la justa medida. El principio de conservación deja de tener sentido si uno de los dos elementos del binomio recursos-desarrollo se lleva a un extremo irreversible. Por consiguiente, sostenible tiene que ver con reversibilidad y equilibrio; pertenece, pues, a la urdimbre de la ecología.

La industrialización marcó una automatización de los procedimientos (a un nivel muy elemental, las máquinas sustituían la mano de obra). Actualmente y en casi cualquier actividad, se ejerce un control informatizado y en red sobre todos los procesos, algo mucho más complejo y opaco que la simple automatización. De aquí procede el gran peligro de emplear programas de los que una sola empresa conozca sus secretos.

De hecho, asistimos a una ruptura acelerada de usos sociales, estimulada por las tecnologías flexibles y abiertas que vamos teniendo en un sistema que está en permanente construcción. Los internautas pasan a convertirse en ‘webactores’. Los datos no dejan de acumularse, compilarse y sintetizarse, todo se pone en relación y deliberación en una clara apuesta por la diversidad. Los usuarios llegan a desplegar una inverosímil capacidad para rebelarse contra las prácticas y las ofertas que no les gustan. En una época en la que muchos, observa Antoine Sire, se preocupan por las consecuencias ecológicas de la actividad industrial, probablemente ha llegado el momento de asociar un verdadero planteamiento de desarrollo sostenible intelectual al desarrollo de la web, en particular, y de Internet, en general. El desarrollo sostenible sólo es posible de la mano del progreso del conocimiento, en una sociedad libre y justa.

Se ha de tener en cuenta que la actividad intelectual –y el software lo es– se ejerce libremente y la hacemos para una comunidad. Por esta razón introducimos los conceptos de *softdependencia*, *softcialización* y *softenibilidad*, los cuales manifiestan la enorme importancia del software en todos los ámbitos actuales.

## 1. PRELIMINARES

### 1.1. Infinitos recursos naturales

La experiencia de la humanidad nos permite decir, casi con toda certeza, que disponemos de un número incontable de recursos naturales. En realidad, estamos abiertos a la sorpresa, la de alcanzar a conocer nuevas e insospechadas posibilidades y aplicaciones de todo género.

Ciertamente las materias primas son, en cantidad y variedad, finitas. Sin embargo, el desconocimiento de sus límites potenciales las hace inagotables; la creatividad humana no tiene límites. Así, por ejemplo, ¿quién habría dicho hace un siglo que de un átomo de hidrógeno se podría obtener tanta energía como se puede conseguir?

Del medio natural, si no se destruye, se puede extraer una energía inagotable. Sólo hace falta experimentar que cada día el sol calienta, que el mar se mueve continuamente, o que el viento no para de soplar.

El uso de los recursos puede implicar desgaste o no. Desgaste si se trata de extracción o transformación de materia, pero no hay desgaste si la actividad no modifica el estado natural original, como en el caso de aprovechar una cascada de agua para hacer girar una rueda de molino.

**Conclusión:** se puede aprovechar los recursos naturales sin limitar las posibilidades actuales y futuras.

## 1.2. Desarrollo

El concepto de desarrollo se puede considerar como la búsqueda de nuevos recursos, o de nuevos usos de los ya existentes. Por tanto, es una combinación de creatividad y de conocimiento. Cuando disminuyen los recursos, se hace inevitable una búsqueda, con lo cual el desarrollo aumenta necesariamente. Por otro lado, cuando los recursos sobran, no se piensa tanto en el propio desarrollo como en la manera de sacarle provecho.

**Conclusión:** así como la energía total de un sistema mecánico es la suma de la energía cinética y la potencial, de manera semejante, el desarrollo en sostenibilidad es la suma ‘uso de recursos’ más su ‘potencialidad de uso’, en la justa medida.

## 1.3. Conservación delicada

El desarrollo sostenible se ve afectado, hasta poder romperse, si las materias primas naturales se usan sin medida, como es bien notorio en el caso de la extracción de petróleo [1].

El desarrollo sostenible puede quebrarse cuando se modifica el estado original de la naturaleza, como sucede con la contaminación de los océanos [2].

Asimismo, el desarrollo sostenible puede desaparecer cuando se bloquea la búsqueda de nuevos usos y recursos, como en el caso de los medicamentos [3].

**Conclusión:** El principio de conservación deja de tener sentido si uno de los dos elementos del binomio recursos-desarrollo se lleva a un extremo irreversible. Por tanto, sostenible tiene que ver con reversibilidad y equilibrio; pertenece, pues, a la urdimbre de la ecología.

# 2. CAUSAS Y EFECTOS

## 2.1. Ruptura del principio de conservación

El principio de conservación puede romperse por causas que afectan al medio, como en el caso del desgaste, contaminación, modificación de las condiciones ecológicas del medio natural [4], tanto por uso como por extracción de los recursos.

Hay que considerar también el abuso de quien explota los recursos por apropiación de los recursos naturales, como sucede con las organizaciones que se quedan todo un territorio para la explotación minera [5], con expropiaciones facilitadas por Estados colaboradores.

Una causa a tener en cuenta es también la limitación de la distribución de los bienes que producen los recursos: Monopolio de la distribución o posesión de materias primas, como los monopolios en el sector de la energía [6].

Y otro modo de bloquear el principio de conservación recursos-desarrollo es mediante patentes o controles excesivos de los procedimientos y conocimientos que permiten el desarrollo, como sucede con los procesos ‘propietarios’ de manufactura de materias primas o de especies transgénicas [7].

## 2.2. Consecuencias catastróficas

Desplazamientos masivos de población y empobrecimiento de grandes extensiones de población y territorio [8]: desertización y desertificación. El primero de estos términos consiste en la transformación de tierras de cultivo en tierras casi desérticas, y se considera a partir de una disminución de la productividad en al menos un 10%. Es el resultado de la destrucción de la cubierta vegetal, de la erosión del suelo y de la falta de agua. Se suele denominar desertificación cuando está provocado por la actividad humana, como por ejemplo, entre otros, el sobrepastoreo (y en consecuencia una vegetación arrancada y pisoteada por los herbívoros y que no se puede recuperar), los incendios provocados o los riegos con agua y sales.

Les patentes en materias de interés general, como las existentes sobre ciertos medicamentos o semillas, encarecen el producto haciéndolo inaccesible a colectividades enteras [9].

Algunas leyes [9] o normativas de ‘copyright’ [10] sobre obras, textos, software, métodos, de interés general, limitan la libertad de conocimiento, investigación y desarrollo, ahora y para futuras generaciones.

Vemos pues que, para la sostenibilidad son condiciones necesarias la ética y la libertad.

## 2.3. Sostenibilidad como punto de equilibrio

Hay que destacar la dimensión vertebradora de los conceptos que estamos considerando. El sociólogo y urbanista François Ascher señala con acierto en su último libro [11] que “la idea de desarrollo sostenible, que se emplea hoy día a diestro y siniestro, contiene el mismo tipo de proyecto de integración económica que los desafíos mediambientales, pues se trata de hacer compatibles –es decir, que converjan– el desarrollo económico, la conservación de los patrimonios natural y cultural, y la igualdad social”.

### **3. EL SOFTWARE LIBRE**

#### **3.1. Softdependencia**

Acaso nos hayamos olvidado de la fallida expectativa del llamado ‘efecto 2000’, previendo serias perturbaciones a causa de unas adaptaciones informáticas. Afortunadamente apenas se produjeron molestias. No obstante, quien actualmente posea el software tiene el poder para controlarlo todo, tanto recursos como desarrollo, lo cual no deja de resultar inquietante. Cabe notar que en la era industrial el control no era tan absoluto ni tan potente como ahora, cuando todo pasa por los ordenadores. De aquí el gran peligro de emplear programas cuyos secretos sólo una empresa los conozca.

Igual que la industrialización marcó una automatización de los procedimientos (a un nivel muy elemental, las máquinas sustituían la mano de obra), actualmente y en casi todas las actividades, se ejerce un control informatizado y en red sobre todos los procesos; algo mucho más complejo y opaco que la simple automatización. La tecnología y la ciencia están relacionadas con el uso de los recursos, y todavía más con la búsqueda de nuevos recursos y nuevos usos, y lo hacen mediante el software. Es decir, el lenguaje codificado, que constituye programas, aplicaciones informáticas, control informatizado de automatismos y sistemas de comunicaciones, bases de datos o gestiones económicas es omnipresente.

#### **3.2. Softcialización**

Se puede decir que asistimos a una acelerada ruptura de usos sociales, estimulada por las tecnologías flexibles y abiertas que vamos teniendo en un sistema que está en permanente construcción. Los internautas (una cuarta parte del mundo utiliza hoy Internet, si bien con una distribución muy irregular: un 2% en África, un 12% en la China; mientras que el 93% de los jóvenes norteamericanos entre 12 y 17 años está conectado a Internet; es curioso que de estos mismos jóvenes sólo el 45% tiene móvil) pasan a convertirse en ‘webactores’ (se ha dicho que nos hayamos en un mundo donde todos somos, casi todo el tiempo, el público de uno u otro medio [12], pero no sólo espectadores, sino actores). Los datos no dejan de acumularse, compilarse y sintetizarse, todo se pone en relación y deliberación en una clara apuesta por la diversidad. Ya se emplea el concepto de alquimia de las multitudes [13], el cual señala el hecho que reunir a un gran número de personas y consultarlas hace posible, en algunos casos, producir oro, pero no siempre. Los usuarios llegan a desplegar una inverosímil capacidad para

rebelarse contra las prácticas y las ofertas que no les gusten. En una época en la que muchos, observa Antoine Sire [13], se preocupan por las consecuencias ecológicas de la actividad industrial, probablemente ha llegado el momento de asociar un verdadero planteamiento de desarrollo sostenible intelectual al desarrollo de la web, en particular, y de Internet, en general.

#### **3.3. Softenibilidad**

Así pues, siempre y cuando el hardware no falle, el software es ‘la palabra’ que se identifica con ‘la acción’. Junto con la técnica y la ciencia, el software es una herramienta de interés general, del cual ellas mismas también dependen [10,14]. Si se limita el uso, la distribución, y el desarrollo del software [15] –o está en manos de quien no procura el interés general- entonces se atenta contra el principio de conservación del desarrollo sostenible. Stallman ha llegado a decir incluso que “una persona que hace valer un ‘copyright’ está dañando a la sociedad en su conjunto, tanto material como espiritualmente; nadie debería hacerlo, aunque la ley lo permita”.

Hay que hacer notar que el software libre es un movimiento social con una base ética, y el ‘open source’ (código abierto) es un método de desarrollo que básicamente busca la eficacia, aceptando a veces compartir espacio con programario de licencia no pública. Son dos movimientos próximos, que se confunden a menudo: Software libre [15] –lo cual no significa que sea gratuito- es el que puede circular y ser modificado con libertad, y hace de ello un principio.

Podemos, por consiguiente, concluir que el desarrollo sostenible sólo es posible de la mano del progreso del conocimiento, en una sociedad libre y justa. Se debe tener en cuenta que el pensamiento, alma del conocimiento, se ejerce libremente y en comunidad -en diálogo-. El software también forma parte de ello.

## **4. CONCLUSIONES**

Los recursos naturales se pueden aprovechar sin limitar las posibilidades actuales y futuras. Bajo esta premisa se puede enunciar un principio de conservación del desarrollo sostenible, que es la suma del ‘uso de recursos’ más su ‘potencialidad de uso’. Y que deja de tener sentido si uno de los dos elementos del binomio recursos-desarrollo se lleva a un extremo irreversible. Por lo tanto, sostenible tiene que ver con reversibilidad y equilibrio.

Los conceptos de *softdependencia*, *softcialización* y *softabilidad*, aquí introducidos, sirven para recalcar la enorme y manifiesta importancia del software en todos los ámbitos sociales. Indican, pues, el carácter sistémico e interdisciplinario del software. Probablemente ha llegado el momento de asociar un verdadero planteamiento de desarrollo sostenible intelectual al desarrollo de la web, en particular, y de Internet, en general. De hecho, se puede decir que asistimos a una ruptura acelerada de usos sociales, estimulada por las tecnologías flexibles y abiertas que vamos teniendo en un sistema que está en permanente construcción.

El desarrollo sostenible sólo es posible de la mano del progreso del conocimiento -y el software también es conocimiento-, en una sociedad entrenada en el diálogo, libre y justa.

Se abre un inmenso territorio interdisciplinario y de participación que debemos cartografiar de manera transparente, inteligente y solidaria.

## REFERENCIAS

1. [http://www.sindominio.net/singuerra/reserves\\_petroli.html](http://www.sindominio.net/singuerra/reserves_petroli.html)
2. <http://www.greenpeace.org/espana/reports/contaminacion-por-plasticos-en>
3. [http://www.redtercermundo.org.uy/texto\\_completo.php?id=2700](http://www.redtercermundo.org.uy/texto_completo.php?id=2700)
4. [http://www.infoforhealth.org/pr/prs/sm13/sm13chap3\\_7.shtml](http://www.infoforhealth.org/pr/prs/sm13/sm13chap3_7.shtml)
5. [http://www.ecoportal.net/Contenido/Contenidos/Eco-Noticias/Manifiesto\\_de\\_CEIMON\\_frente\\_a\\_la\\_explotacion\\_minera\\_metalica\\_en\\_El\\_Salvador](http://www.ecoportal.net/Contenido/Contenidos/Eco-Noticias/Manifiesto_de_CEIMON_frente_a_la_explotacion_minera_metalica_en_El_Salvador)
6. <http://www.rebelion.org/noticia.php?id=75802>
7. [http://www.ecologistasenaccion.org/article.php3?id\\_article=3174](http://www.ecologistasenaccion.org/article.php3?id_article=3174)
8. <http://www.ircamericas.org/esp/5116>
9. [http://www.elpais.com/articulo/sociedad/EE/UU/prohibe/publicar/articulos/cientificos/Cuba/Iran/Libia/Sudan/elpporsoc/20040224elpepisoc\\_2/Tes](http://www.elpais.com/articulo/sociedad/EE/UU/prohibe/publicar/articulos/cientificos/Cuba/Iran/Libia/Sudan/elpporsoc/20040224elpepisoc_2/Tes)
10. <http://www.nature.com/nature/debates/e-Access/Articless/stallman.html>

(<http://www.yukei.net/2005/07/la-ciencia-debe-dejar-de-lado-el-copyright>)

11. Ascher, François. Diario de un hipermoderno. Alianza Ed. Madrid, 2009.

12. Morley, David. Medios, modernidad y tecnología. Gedisa, Barcelona, 2009.

13. Pisani, Francis y Piotet, Dominique. La alquimia de las multitudes (Cómo la web está cambiando el mundo). Ed. Paidós. Barcelona, 2009.

14. <http://www.mmc.igeofcu.unam.mx/LuCAS/Presentaciones/200002hispalinux/conf-18/18-html/ponencia.html>

15. Stallman, Richard M. Software libre para una sociedad libre. Mapas. Madrid, 2007. (<http://biblioweb.sindominio.net/pensamiento/softlibre/index.html>)

## AUTORES



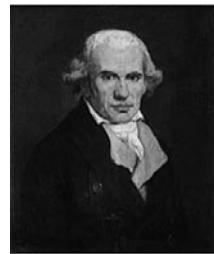
Rafael Cubarsí i Morera es profesor asociado de la Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona y desarrolla su actividad docente en el departamento de Matemática Aplicada IV en UPC desde 1990.

Asimismo, es graduado en Física y Matemática por la Universidad de Barcelona (UB) y Doctor en Astronomía desde 1988.



Miguel Escudero Royo es profesor titular de la Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona, adscrito al Departamento de Matemática Aplicada IV de la Universidad Politécnica de Cataluña. También es Doctor en Filosofía y Letras y ha escrito artículos de opinión en publicaciones como "Cuenta y Razón", el diario "La Vanguardia" o "Buran".

# HOMBRES ESCONDIDOS EN FÓRMULAS



## EL PROFESOR MONGE

*Miquel Escudero*

Departament de Matemàtica Aplicada IV  
E-mail: [escudero@ma4.upc.edu](mailto:escudero@ma4.upc.edu)

El profesor Boyer (¡ojo! No se trata del ex ministro de la abeja de RUMASA. ¿Alguien sabe de lo que hablo?) lo calificó como “el más prestigioso profesor de matemáticas desde los tiempos de Euclides”. Esto nos da ganas de saber de este hombre. Gaspard Monge nació en 1746 en Beaune, un pueblo cercano a Dijon y a unos 300 kilómetros de París. (Cinco años después saldría el primer volumen de la Enciclopedia de la mano de Diderot y D'Alembert, impulsados por el librero Le Breton.) De pequeño destacó por sus notas, y sus maestros lo calificaron de ‘puer aureus’ (niño de oro). A los 18 años hizo un plano topográfico de su pueblo que mereció la admiración y el interés de la Escuela de Ingenieros Militares, la cual lo quiso contratar. Dado que provenía de una familia más bien modesta le correspondió entrar en la sección ‘pobre’, y sólo podía llegar a ser subteniente, algo que le irritó y le dolió profundamente. Al acabar 1771 redactó una memoria matemática sobre un juego de cartas por la cual fue nombrado, al poco, correspondiente de la Academia de Ciencias. Llegó a tener amistad con Vandermonde (el teórico de los determinantes, un método para resolver ecuaciones lineales) y con Lavosier (que sería guillotinado y de quien Lagrange afirmaba que había logrado hacer la química tan fácil como el álgebra). Ya en 1777 –el año en que se casó con una viuda (tuvieron tres hijas, la pequeña murió a los tres años)- hizo una incursión por el mundo de la metalurgia. Digamos que con el tiempo llegó a ser el máximo responsable de la fabricación de armamento en Francia y que una explosión en la fábrica de pólvora de Grenelle produjo más de mil víctimas; él la acababa de supervisar. En 1783 sustituyó a Bézout (recién muerto) como examinador de los alumnos de la Marina. Casi diez años después, y a propuesta de Condorcet, fue nombrado ministro de Marina; lo fue durante sólo ocho meses, llegó a ser jefe del gobierno en funciones cuando Luis XVI fue ejecutado; así, firmó la sentencia de muerte del rey, detalle que he sabido por el libro “Monge. Libertad, igualdad, fraternidad y geometría”, de Antonio Hernández.

Monge promovió en 1794 l’École Centrale des Travaux Publics que al poco cambió de nombre por el de École Polytechnique. Por esas fechas resolvió una ecuación diferencial de segundo orden no lineal que lleva su nombre. No obstante, su máximo prestigio proviene de la geometría descriptiva que diseñó, una técnica por la cual los objetos de tres dimensiones son representados sobre un papel. Monge, de ojos grandes, frente amplia, alto y musculoso, ha sido analizado por Dupin, alumno suyo, de este modo: “Cuando hablaba, nos parecía ver a otro hombre..., y en sus ojos brillaba fuego nuevo; sus rasgos se animaban; su cara se transformaba y parecía ver delante suyo los mismos objetos creados por la imaginación del geómetra”.

En 1798 analizó el fenómeno óptico del espejismo. Lo hizo en la revista ‘La Décade Égyptienne’, que salía cada diez días publicada por el Instituto de Egipto. Esta institución acababa de ser fundada por Napoleón en su célebre expedición a Egipto. Monge, igual que Fourier y Champollion, estuvo entre los asesores científicos de aquel viaje. Allí reorganizaron los servicios públicos, los caminos y las rutas marítimas, con el objeto de modernizar el país. La operación militar de expulsar a los turcos iba dirigida especialmente contra Gran Bretaña, con el fin de asfixiar su comercio en la ruta hacia la India. En aquella época se estableció el museo del Louvre, pero también se desarrolló la egiptología como nueva especialidad histórica, y científica si se atiende al avance de la criptografía.

Napoleón había restablecido la concesión de títulos y honores abolidos por la Revolución. Siempre agradeció el buen trato que, siendo un desconocido, le dispensó una vez Monge en Italia. Lo nombró conde de Péluse en 1806. Monge estaba muy identificado con la política de Bonaparte y le fue leal. Murió en el exilio en Bélgica, el año 1818. Sus restos fueron trasladados a París con la consigna oficial de mantener un silencio absoluto sobre su figura. Pero los estudiantes politécnicos le rindieron homenaje agradecido en su tumba y por suscripción popular se le dedicó un monumento funerario.

# Secure Voting From Your Living Room

David Andreu, Alex Escala, Guillem Caldúch

E-mail:

## ABSTRACT

Nowadays, e-voting is used in more and more countries and in all kind of elections. There's not a single solution to e-voting and each has different characteristics and areas of application. In this article we give a brief description of what we understand by e-voting systems and we introduce the different models used in remote electronic voting. We explain how these systems work and give some ideas of their advantages and drawbacks. Finally we give some examples of elections using e-voting and some ideas of the challenges that researchers try to overcome.

## 1. WHAT IS E-VOTING?

There are several definitions about what we call “e-voting”. In a broad sense, “e-voting” is considered the introduction of electronic systems in an electoral process, without discerning if they take part in the electoral roll, in making the district maps, in the electoral logistics, in the mechanisms of voting or in the count and transmission of the results. However, in this article we will only focus on two areas of application: the emission of the votes and the subsequent count of them.

There is not a unique way to implement an e-voting system. In fact, there are three types of implementations, which basically differ in their benefits and risks. These types are listed below:

A) Optical Mark Recognition (OMR): systems of automatic count of votes by using techniques of optical recognition, which can read marks in the ballots made by the voters. These systems are focused on the count.

B) Direct-Recording Electronic (DRE) Voting System: systems that use digital mechanisms for selecting the vote, such as buttons or a tactile screen. Normally it prevents errors made by the voters because it guides the user step by step. Commonly the votes are registered by the machine but there are some versions that also print a ballot to be placed in a ballot box, usually to check correctness of the election.



Figure 1. A DRE machine.

C) Remote Electronic Voting System: there are some different channels to transmit a vote in a remote way such as Internet (web or e-mail), SMS and others. This kind of systems is more complex than the in-person ones because the electoral authority cannot control all the steps as before.

## 2. PROS AND CONS OF E-VOTING

It's clear that e-voting can be an improvement on traditional elections for its different features: speed, accessibility, error prevention (which implies a reduction of invalid votes), cost reduction and the possibility that voters verify the correct treatment of their votes.

Even more, e-voting can reduce the cost of large scale elections: there will be one paper ballot per voter (or even no paper ballots at all), in contrast to traditional elections where lots of paper ballots are printed, more than the population who can vote.

In spite of the improvements that e-voting presents, there are some vulnerabilities of e-voting systems that have stopped the extensive use of them. For instance, the virtual nature of the ballots makes that they can be added, manipulated or deleted. Another factor is that the systems used may have problems and errors, compromising voter privacy.

To mitigate these risks cryptographic protocols need to be used along with software auditions. With the second ones we can ensure that the system works as intended, with the first ones we can ensure the following requirements: vote integrity, authentication and privacy of voters, accuracy of

election results and prevention of coercion and vote-selling. Even more, we can also make that individual voters can verify their own vote and the correctness of the result.

### 3. ELECTRONIC VOTING PROTOCOLS

Now we will present the electronic voting protocols that use advanced cryptographic techniques to fulfill the requirements listed above. These protocols can be classified by the stage in which the voter anonymity is protected:

- Before the voting process:

- *Pre-encrypted ballots* (also known as pollsterless): the voting options (e.g. candidates) are pre-encrypted generating unique individual ballots for each voter.

- During the voting process:

- *Two-agencies model*: it uses a technique called blind signature in order to split the authentication of the vote from the casting of the (validated) vote.

- After the voting process:

- *Mixing model*: votes are shuffled secretly and, at the end of the election, decrypted, to break any correlation between the encrypted votes and the decrypted ones.

- *Homomorphic model*: its objective is to obtain the tally of the elections without decrypting any single vote.

### 4. PRE-ENCRYPTED BALLOTS

A problem that can arise when using remote electronic voting is that the software used for voting turns malicious. In this way, an intruder could see someone's vote and even send a different vote to the election server, thus breaking privacy and vote integrity.

In 2002, California Internet Voting Task Force and Oppliger presented a solution to that problem: start making use of code sheets (currently known as pre-encrypted ballots). These code sheets contain a voter identifier and codes related to the voting options (each one of these options has a code associated) and, basically, the idea is to deliver a unique code sheet to every voter. The voter will make his selection by introducing the code associated to the alternative he wants to vote for. This way, as nobody except the voter knows the relation between codes and candidates, malicious software won't be able to manipulate the vote.

Name:	Voter Name No Number, No Street No Town.
<b>VoterID: 4545 2321 6742 1209</b>	
Candidates	PCIN RID
Candidate 1	7890 1092
Candidate 2	3417 3417
Candidate 3	8417 8417
Return code (optional)	

Figure 2. One example of a code sheet.

One consequence of not having to compute any encryption is that votes may be cast from devices with low computing power, such as sending an SMS from a mobile phone.

As we can see in the picture above, there are some return codes: these are to verify that the voting server has received his vote properly. This helps to counter software malfunction or attacks on the voter. One example of an attack is the following: when the voter casts the first vote and every time he casts the same vote the attacker denies the communication between the voter and the voting server. The voter may then send another code (to check if the system works) and the attacker would permit the communication, so finally the voter isn't voting what he wanted to vote. Using return codes this attack is countered because the voter will know if he is receiving a denial of service attack.

All the codes are generated using cryptographic and coding tools and a different ballot will then be secretly sent to each voter. So this is the main drawback of the system: we need to deliver the vote in a secure way and there can be logistics problems when distributing the ballots.

### 5. TWO AGENCIES MODEL

One of the requirements of e-voting is voters' privacy, but we need to authenticate them. One way to make both things compatible is to split voters' authentication from the ballot casting process to prevent the association between the voter and the vote. To achieve this, the scheme needs to use two different systems.

On one hand, the validation server authenticates the voters and validates the votes in an anonymous way. On the other, the voting server receives the validated votes without identifying the voter and stores the votes. Normally, this kind of

schemes are based on protocols that use a mechanism called blind signature in order to validate votes without compromising the voter privacy.

The first blind signature protocol was designed by David Chaum in 1982. It allows someone to obtain a message signed by another entity without revealing the message contents to the entity. We require that this signature is able to be verified by others (so it's a signature), like the regular digital signatures, and we also require that whoever has signed the message must not be able to guess the message contents (so we call it blind).

The voting system works as follows: the voter encrypts his vote with the public key of the voting server, blinds the message and sends it to the validation server. The validation server will sign the message and the voter will be able to unblind the message and send it to the voting server, who will check the signature, store the vote and finally decrypt all the votes. To clarify it, we present a diagram of the scheme.

In the figure,  $E_s(m)$ ,  $D_p(m)$  and  $S_s(m)$  stand for encryption using the secret key  $s$ , decryption using the public key  $p$  and signature using the secret key  $s$ , respectively. (Recall that, in asymmetric encryption, a pair of keys {public key, secret key} are used; the first one used to encrypt and the second one to decrypt or make a signature on a message. Decryption and signing are two processes which are usually the same mathematic function).

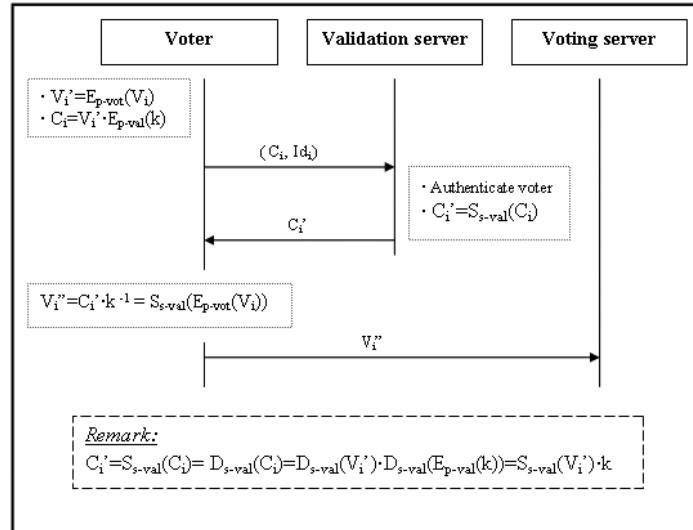


Figure 3. A two-agencies protocol scheme.

Some issues have to be addressed, though. This system breaks the correlation between signature and voting process, so we need to ensure that nobody can monitor both

channels and associate the voter with its vote. Even more, the two trusted authorities may be corrupt, so we must use techniques of secret sharing (for the decryption key) and multiparty computation (for the signature). Finally, nobody should be able to vote twice with the same signed message, we can use some cryptographic techniques to avoid this too.

## 6. MIXING MODEL

Taking another approach to ensure voter's privacy, we could think what happens in traditional elections: when a vote is casted it goes into an urn and when votes are counted there's no way to know the relation between a vote and the voter who cast it because shuffling the urn breaks any correlation between the vote and the voter. This model tries to emulate this behavior.

Actually, this model doesn't come from e-voting. It was conceived to guarantee user anonymity when sending off emails but it has been proved that it can also be applied to electronic voting for the same purpose, guarantee voters' anonymity. In this system we create an anonymous communication channel that will break the correlation between the incoming and outgoing messages.

In mix-nets (a particular case of mixing model) the channel consists of some servers, which are known as mix servers. The first of these servers receives several messages from different senders and applies some sort of transformation on them. Afterwards transformed messages are shuffled randomly and sent to the next server. This one and every remaining server will repeat the same process, delivering messages once more transformed and shuffled to the server that follows them. The reason for not using just one server is that if the server cheats then anonymity is broken.

The messages' transformation can be of two types, which define two types of mixing: decryption mixing and re-encryption mixing. In the first one each mix server decrypts the message that it receives with its private key, so the user has to encrypt the message as many times as the number of servers in the mix-net and a pair of keys must be created for each server. In re-encryption mixing the server just encrypts the message once and each server re-encrypts the message, all encryptions using the key of the recipient. The receiver will then decrypt the message once if a homomorphic cipher is used (which we will explain later).

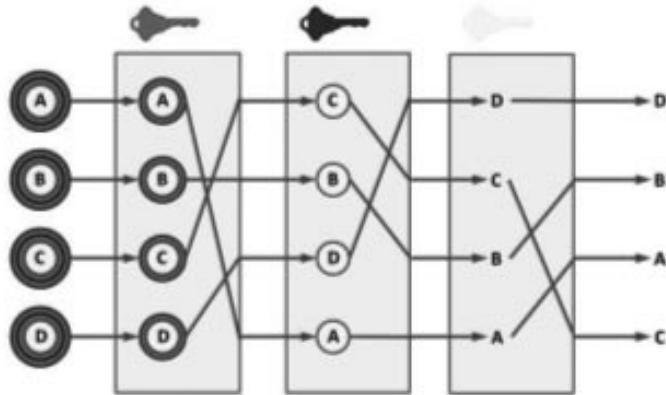


Figure 4. A decryption mix-net.

The main problem of this model is that verifiability by individual voters has to use techniques such as proofs of knowledge (this is, an interactive protocol where the prover convinces a verifier that he knows some values without revealing them), which are computationally expensive.

## 7. HOMOMORPHIC MODEL

The last model we will talk about involves more cryptographic tools than the other does, which makes it an expensive model in terms of computational power. Using this scheme we can obtain a high level of security and a good verifiability by individual voters. As said in the introduction, in this model only the tally is decrypted, while individual votes are kept encrypted.

Before explaining this model we should introduce two concepts: homomorphic ciphers and secret sharing schemes.

We speak of *homomorphic ciphers* when one can perform some specific algebraic operation on the plaintext by performing a (possibly different) algebraic operation on the ciphertext. For example: multiplying two ciphertexts will give us the encryption of the sum of the plaintexts. While this characteristic isn't desirable in data transmission because homomorphic encryptions are malleable, it's very interesting in e-voting schemes. An example of homomorphic cipher is ElGamal.

Homomorphic ciphers are used in the following way: each voter encrypts its vote and publishes it. When all votes are published everybody can compute the encryption of the tally by multiplying the ciphertexts (which will bring to the sum of the plaintexts, this is, the votes). Later we will show how to decrypt the tally.

Another concept we should introduce is secret sharing schemes. The main idea is divide a secret  $s$  into  $n$  pieces, called shares. We distribute the shares among  $n$  users and

the scheme is made so that a subset can recover the secret if its members collaborate but a non-authorized subset won't learn anything about the secret. One of the first secret sharing schemes was introduced by Shamir in 1979 and it's a threshold secret sharing scheme. That is, a set is able to reconstruct the key if and only if it consists of more than  $t$  people, where  $t$  is the threshold specified.

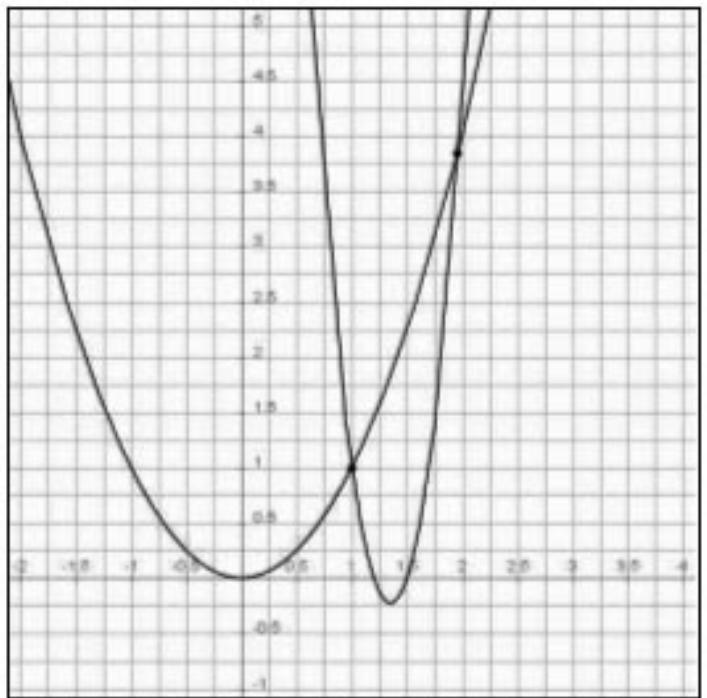


Figure 5. 2 points don't determine a parabola.

The main idea behind Shamir's secret sharing scheme is polynomial interpolation. We know that a polynomial of degree  $t$  is determined by  $t+1$  points, so we can give an evaluation of the polynomial to each participant in the secret sharing. When more than  $t$  people want to get the secret they can interpolate the polynomial which hides the secret but  $t$  people or less won't have any idea of what polynomial is used. This is useful when you need to trust some people and you consider that it's improbable that a large group of people will collude and cheat.

Now we can sketch how the system works. First of all the keys and the shared secret are generated, which can be done by a trusted third party (TTP) or by all the authorities jointly (using multiparty computation). The voters will then encrypt their vote and publish it. When all votes are cast, each authority will compute the encryption of the tally (multiplying the encrypted votes) and will publish their partial decryption with their share of the secret. After that everybody can compute the decrypted tally with the partial decryptions.

Due to the homomorphic properties in this system, we have to check that nobody cheats. For example, a voter could give 100 votes to a candidate (which would add 100 votes to the tally) so we have to avoid this using proofs of knowledge. The same goes for the partial decryptions: we have to make sure that they are done correctly (again, with proofs of knowledge). This means that the computational power needed in this system is higher than in other cases, even for the voters.

## 8. REAL EXPERIENCES AROUND THE WORLD

There have been electronic elections in lots of countries. Most times these were just non-binding tests, because there was no legislation going over it or just to check if the system worked properly.

In the election to the parliament of Catalonia there was a non-binding pilot test, the first in Spain when speaking of public elections. This was made to test advantages, usability and reliability of this technology. It has also been used in non-governmental elections in Catalonia. One example was the election of the labor union of the autonomic police, where e-voting was used successfully.

We can't speak of more experiences in our country because e-voting is still in a very initial state. However in other places there are lots of experiences and problems related to it.

In Brazil, for example, all votes are cast by electronic voting machines. Use of electronic voting technology was authorized in 1996 in the municipal elections. In 1998 this use was extended when 57% of the electorate used electronic voting. By 2000 Brazilian government had converted to fully electronic voting and deployed over 400.000 kiosk-style machines in elections that year.

An interesting example of how things, if they aren't done properly, won't work is Netherlands. Voting machines were introduced there in 1965, with a good acceptance. In 2006 voters from abroad used Internet to vote, and there were plans to use internet for voters within country. But legislation wasn't as hard as it should have been and machines weren't updated with regard to security provisions. An organization against e-voting appealed that the system was easy to hack and six weeks before parliamentary elections in November 2006 they proved that they could manipulate the machines and they could read them from distance (thanks to Tempest). Needless to say, Internet voting was discarded

and classic voting systems were recovered.

But with no doubt, the place where there have been more experiences (and also more problems) is America. On November 4, 2003 in Fairfax County, Virginia machines malfunctioned and collapsed the modems. When 953 voting machines tried to call simultaneously to inform of the results they caused a denial of service accident in the election. Another example: on August 1, 2001 in the Brennan Center at New York University Law School. NY University Law School released a report with more than 60 examples of e-voting machine failures in 26 states in 2004 and 2006. Examples included Spanish language ballots that were cast by voters but not counted in Sacramento in 2004.

In fact, this kind of problems discourages lots of voters, who won't have confidence on this system for many years. For that reason it is important that, when implementing e-voting systems, this is done slowly and carefully to ensure the security of e-voting. There's a need to legislate e-voting with all the details to prevent fraud and malfunctioning of e-voting systems.

## 9. CURRENT RESEARCH ON E-VOTING

E-voting isn't a closed topic: there're still several factors that can be improved.

One of the research objectives is to reduce computational cost on e-voting solutions. This is very important mostly on the homomorphic model as the cost of its protocols is high. For example, there's research on validity check, the main efficiency bottleneck that limits the application of this kind of e-voting. There are proposed solutions making validity highly efficient.

Another example is research done in mixing models or shuffling models. There're some papers proposing new shuffling systems trying to speed up the whole process while still having universal verifiability. Another goal is to make proofs of knowledge to the mixing servers in such way that the computational cost is reduced. As an example, B. Shoenmakers and others proposed a protocol using the DFT to have efficient proofs of knowledge.

Apart from research done by universities there also exist some companies that research and provide solutions to be applied in e-voting. One example in Catalonia is Scytl, created in 2001 as a spin-off from UAB students. It offers solutions for e-voting with several applications that go from

government elections to college elections. Scytl works for the Generalitat de Catalunya, the Ministry of Justice of UK and many others. It also collaborates with universities from Catalonia in their research such as UAB or UPC.

## AUTHORS



Alex Escala was born in Barcelona, on 1986. He is studying the last course of mathematics degree and telecommunication engineering in the Universitat Politecnica de Catalunya. He is also working in a Department of Applied Mathematics, researching on cryptography.



David Andreu was born in Barcelona, on 1986. He is studying the last course of telecommunication engineering at Universitat Politecnica de Catalunya. He is also working in the Theory of Signal and Communications Department as a teaching assistant in mobile communications.



Guillem Caldúch was born in Barcelona, on 1986. He has studied telecommunication engineering at Universitat Politecnica de Catalunya. Currently he is working on his Master Thesis about Android operating system and development of applications for this OS.

# Searching Representative Phrases in a Musical Score using Fuzzy Logic

Emerson Castañeda

Phd Student

Departamento de Ingeniería del Software e Inteligencia Artificial,  
Facultad de Informática, Universidad Complutense de Madrid, 28040-Madrid, Spain  
E-mail: emecas@ieee.org

## ABSTRACT

In this paper a new method to find representative phrases from a musical score is given. A fuzzy proximity relation on a set of phrases is computed as a conjunction of a indistinguishability on the variation of notes of the phrases, where the indistinguishability is the Lukasiewicz t-norm. Different fuzzy logics are used and compared. A method to find the most representative phrase of a musical score is found.

## 1. INTRODUCTION

The concept musical “motif” is related to a short musical phrase on which a composer develops a whole musical score. The “motif” is a melodic element that is important throughout the work and that can be varied to generate more musical phrases. The motif of a score is found using a “fuzzy pattern machine model” that uses indistinguishability operators and proximity fuzzy relations to compare phrases.

The negation of a distance is a indistinguishability operator that can be applied to the variations of consecutive notes for each couple of phrases to compute proximity on the set of phrases in order to obtain how similar the phrases sound. Two phrases are considered ‘similar’ when the variation between the first and the second notes are ‘equivalent’, AND the variation of the second and the third notes are ‘equivalent’, AND ..., so on and so for. That is, two phrases are similar if the distance of their notes is similar and therefore, the two phrases sound the same even if the starting tone is different. Such is modeled using different t-norms.

Once the proximity fuzzy relation on a set of phrases has been computed, a method to automatically select the phrase by computing the fuzzy set ‘similar to the other phrases’ on the set of phrases is applied. The representative phrase is the one with highest membership degree for such fuzzy set.

This method is applied in order to find the representative phrase of the musical score of figure 1.



Figure 1. A few phrases of Inven. # 1 of J. Bach

## 2. SEARCHING ALGORITHM FOR MUSICAL MOTIF

A musical score is separated into phrases that are compared to each other in order to evaluate a proximity degree of every couple of phrases, allowing the identification of the phrases that are candidates to be motifs.

Once we defined a proximity relation on the set of phrases, it will be applied to find musical “motifs” by an algorithm. A pre-searching method is used as a starting point for the musical motifs searching algorithm. The algorithm in pseudo code can be written as,

- a) A score is separated into phrases.
  - b) A proximity degree of every couple of phrases is evaluated.
  - c) A fuzzy set ‘candidate to be a motif’ is computed on the set of phrases by aggregating the proximity degree of each phrase with other phrases.
  - d) The most representative phrase is the one with highest membership degree on the fuzzy set ‘candidate to be a motif’.
- It is possible to improve the results by selecting different sets of phrases, as the process depends on the separation of phrases.
- The pre-searching takes into account the following criteria:
- 1) The comparison of the notes duration.
  - 2) The variation of tones into a phrase.
  - 3) The distance of the intervals between notes into a phrase.

### 3. EXPERIMENTS AND RESULTS

An example is given to show how the algorithm performs. The eleven first phrases of Figure 1 are included to evaluate if there is a possible motif. The score is already divided into phrases of 8 notes.

The first phrase  $P^n_1$  is formed with the 8 first notes in the score, the initial silence is omitted. The rest of the phrases  $P^n_i$  are formed by the set of 8 consecutive notes starting from note i, it is displacing a note forward for every phrase. The eleven initial phrases of the two voices in the score are analyzed in order to simplify the example, but this method should be applied to the whole musical score (more than 400 phrases), using different lengths to find the best motif. The variation points of the first eleven phrases of Inven. # 1 of J. Bach (Figure 1) are:

$$\begin{aligned}
 P^n_1 &= [(C_5, 2), (D_5, 2), (E_5, 1), (F_5, -3), (D_5, 2), (E_5, -4), (C_5, 7), (G_5, 0)] \\
 P^n_2 &= [(C_6, -1), (B_5, 1), (C_6, 2), (D_6, -7), (G_5, 2), (A_5, 2), (B_5, 1), (C_6, -3)] \\
 P^n_3 &= [(B_5, 1), (C_6, 2), (D_6, -7), (G_5, 2), (A_5, 2), (B_5, 1), (C_6, -3), (A_5, 0)] \\
 P^n_4 &= [(C_6, 2), (D_6, -7), (G_5, 2), (A_5, 2), (B_5, 1), (C_6, -3), (A_5, 2), (B_5, 0)] \\
 P^n_5 &= [(D_6, -7), (G_5, 2), (A_5, 2), (B_5, 1), (C_6, -3), (A_5, 2), (B_5, -4), (G_5, 0)] \\
 P^n_6 &= [(G_5, 2), (A_5, 2), (B_5, 1), (C_6, -3), (A_5, 2), (B_5, -4), (G_5, 7), (D_6, 0)] \\
 P^n_7 &= [(C_4, 2), (D_4, 2), (E_4, 1), (F_4, -3), (D_4, 2), (E_4, -4), (C_4, 7), (G_4, 0)] \\
 P^n_8 &= [(D_4, 2), (E_4, 1), (F_4, -3), (D_4, 2), (E_4, -4), (C_4, 7), (G_4, -12), (G_3, 0)] \\
 P^n_9 &= [(E_4, 1), (F_4, -3), (D_4, 2), (E_4, -4), (C_4, 7), (G_4, -12), (G_3, -), (-, 0)] \\
 P^n_{10} &= [(F_4, -3), (D_4, 2), (E_4, -4), (C_4, 7), (G_4, -12), (G_3, -), (-, -), (-, 0)] \\
 P^n_{11} &= [(D_4, 2), (E_4, -4), (C_4, 7), (G_4, -12), (G_3, -), (-, -), (-, -), (G_4, 0)]
 \end{aligned}$$

The n-1 distances  $D(P^n)$  between the variation points of every phrase are computed in the following table:

$D(P^n_1)$	[ 2.00 2.24 4.12 5.83 6.32 11.70 15.65 ]
$D(P^n_2)$	[ 2.24 1.41 9.22 11.40 2.00 2.24 1.41 ]
$D(P^n_3)$	[ 1.41 9.22 11.40 2.00 2.24 4.12 5.83 ]
$D(P^n_4)$	[ 9.22 11.40 2.00 2.24 4.12 5.83 2.24 ]
$D(P^n_5)$	[ 11.40 2.00 2.24 4.12 5.83 6.32 11.70 ]
$D(P^n_6)$	[ 2.00 2.24 4.12 5.83 6.32 11.70 15.65 ]
$D(P^n_7)$	[ 2.00 2.24 4.12 5.83 6.32 11.70 15.65 ]
$D(P^n_8)$	[ 2.24 4.12 5.83 6.32 11.70 20.25 22.47 ]
$D(P^n_9)$	[ 4.12 5.83 6.32 11.70 20.25 16.97 9.00 ]
$D(P^n_{10})$	[ 5.83 6.32 11.70 20.25 16.97 0.00 10.00 ]

Table 1. Distance  $D(P^n)$  between notes of every phrase

The next step is to normalize Table 1 using the maximum value of the distance, which is 22.47. The normalization results are shown in Table 2.

Phrase	Normalized Values							
	1	0.09	0.1	0.18	0.26	0.28	0.52	0.7
2	0.1	0.06	0.41	0.51	0.09	0.1	0.18	0.06
3	0.06	0.41	0.51	0.09	0.1	0.18	0.26	0.26
4	0.41	0.51	0.09	0.1	0.18	0.26	0.1	
5	0.51	0.09	0.1	0.18	0.26	0.28	0.52	
6	0.09	0.1	0.18	0.26	0.28	0.52	0.7	
7	0.09	0.1	0.18	0.26	0.28	0.52	0.7	
8	0.1	0.18	0.26	0.28	0.52	0.9	1	
9	0.18	0.26	0.28	0.52	0.9	0.76	0.4	
10	0.26	0.28	0.52	0.9	0.76	0	0.45	
11	0.28	0.52	0.9	0.76	0	0	0.31	

Table 2. Normalized distances  $D(P^n)$  between notes of every phrase

	Phrase 6							product	min	w
	1	1	1	1	1	1	1			
2	0.99	0.96	0.77	0.75	0.81	0.58	0.37	0.095	0.366	0.000
3	0.97	0.69	0.68	0.83	0.82	0.66	0.56	0.115	0.563	0.000
4	0.68	0.59	0.91	0.84	0.9	0.74	0.4	0.082	0.403	0.000
5	0.58	0.99	0.92	0.92	0.98	0.76	0.82	0.299	0.582	0.000
7	1	1	1	1	1	1	1	1	1	1
8	0.99	0.92	0.92	0.98	0.76	0.62	0.7	0.269	0.620	0.000
9	0.91	0.84	0.9	0.74	0.38	0.77	0.7	0.104	0.380	0.000
10	0.83	0.82	0.66	0.36	0.53	0.48	0.75	0.030	0.358	0.000
11	0.81	0.58	0.28	0.5	0.72	0.48	0.61	0.014	0.282	0.000

Table 3. Indistinguishabilities of the variation points of phrase 6 with the variation points of the other phrases

Table 2 is used to compute the table of indistinguishabilities of every phrase with respect to the other 10 phrases. It is obtained by computing the distance of every variation point of the phrases (values of table 2) with the other 10 rows and applying the usual negation  $N(x) = 1 - x$  to operate indistinguishabilities between variation points of different phrases. For example, Table 3 shows the proximity of phrase 6 with the others using three conjunction operators (t-norms of Table 1).

### 4. PROXIMITY ON THE SET OF PHRASES

Two phrases are considered ‘similar’ when the variation between the first and the second notes are ‘similar’, AND the variation between the second and the third notes are ‘similar’, AND ..., so on and so for. Such concept of ‘similar’ is replaced by ‘indistinguishable’ in our method.

The process of calculating the conjunction of indistinguishabilities of the variation points of each phrase with each phrase define a proximity on the set of phrases. The final proximities on the set of phrases are shown in Tables 4, 5 and 6 where the conjunctions (AND) are the three main t-norms.

	1	2	3	4	5	6	7	8	9	10	11
1	1	0.366	0.563	0.403	0.582	1	1	0.620	0.380	0.358	0.282
2	0.366	1	0.582	0.556	0.542	0.366	0.366	0.063	0.188	0.334	0.509
3	0.563	0.582	1	0.582	0.556	0.563	0.563	0.259	0.198	0.188	0.334
4	0.403	0.556	0.582	1	0.579	0.403	0.403	0.100	0.282	0.198	0.188
5	0.582	0.542	0.556	0.579	1	0.582	0.582	0.380	0.358	0.282	0.198
6	1	0.366	0.563	0.403	0.582	1	1	0.620	0.380	0.358	0.282
7	1	0.366	0.563	0.403	0.582	1	1	0.620	0.380	0.358	0.282
8	0.620	0.063	0.259	0.099	0.380	0.620	0.620	1	0.400	0.099	0.099
9	0.380	0.188	0.198	0.282	0.358	0.380	0.380	0.400	1	0.245	0.099
10	0.358	0.334	0.188	0.198	0.282	0.358	0.358	0.099	0.245	1	0.245
11	0.282	0.509	0.334	0.188	0.198	0.282	0.282	0.099	0.099	0.245	1

Table 4. Proximity of phrases using the t-norm minimum

	1	2	3	4	5	6	7	8	9	10	11
1	1	0.000	0.000	0.000	0.000	1	1	0.000	0.000	0.000	0.000
2	0.000	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	1	0.000	0.000	0.000	0.000	0.000	0.000
6	1	0.000	0.000	0.000	0.000	1	1	0.000	0.000	0.000	0.000
7	1	0.000	0.000	0.000	0.000	1	1	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1

Table 5. Proximity of phrases using the t-norm product

	1	2	3	4	5	6	7	8	9	10	11
1	1	0	0	0	0	1	1	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0
6	1	0	0	0	0	1	1	0	0	0	0
7	1	0	0	0	0	1	1	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	0	0	1	0
11	0	0	0	0	0	0	0	0	0	0	1

Table 6. Proximity of phrases using the t-norm Lukasiewicz

## 5. METHOD TO FIND A REPRESENTATIVE PHRASE

The method proposed tries to find representative phrases from the information of tables 4, 5 and 6. By the aggregation of every row in the proximity matrix, using the arithmetic average a fuzzy set ‘proximity with the rest of phrases’ is defined on the set of phrases. Then the phrase or phrases with maximum membership degree are chosen as the most representative phrases. The normalized average values are presented in Table 7 and Figure 2.

Phrase	Avg Product	Avg Min	Avg W			
1	0.301	1.000	0.555	1.000	0.200	1.000
2	0.094	0.312	0.387	0.697	0.000	0.000
3	0.111	0.369	0.439	0.790	0.000	0.000
4	0.093	0.308	0.369	0.665	0.000	0.000
5	0.152	0.504	0.464	0.835	0.000	0.000
6	0.301	1.000	0.555	1.000	0.200	1.000
7	0.301	1.000	0.555	1.000	0.200	1.000
8	0.104	0.345	0.326	0.587	0.000	0.000
9	0.066	0.220	0.291	0.524	0.000	0.000
10	0.041	0.137	0.267	0.480	0.000	0.000
11	0.038	0.125	0.252	0.453	0.000	0.000

Table 7. Membership degree and normalized “proximity with other phrases” on the set of phrases

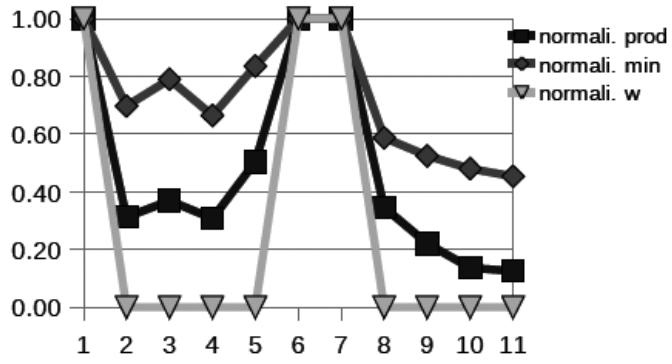


Figure 2. Fuzzy set of values of proximity of every phrase  $i$  with the rest of phrases.

By observing the results it is possible to conclude that the representative phrases are 1, 6, and 7, with average values over 30%, 55% and 20% respective to every t-norm. These phrases are shown in Figure 3. It is also possible to identify a second set of representative phrases looking at Table 7. Phrases 3 and 5 are also representative with values over 40% and 10% for the t-norms Min and Product respectively. A musical representation of phrases 3 and 5 is showed in Figure 4.

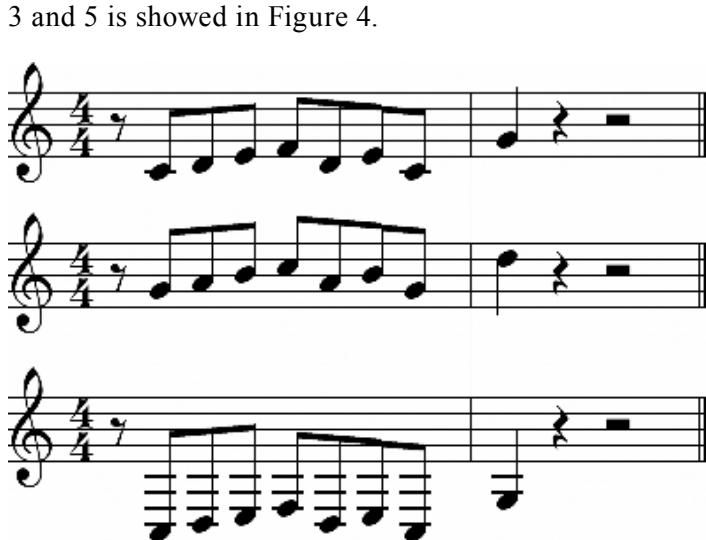


Figure 3. A musical representation of phrases 1, 6 and 7.



Figure 4. A musical representation of phrases 3 and 5.

Phrase 7 is descending an octave and phrase 6 is ascending in 7 semitones (see how they confirm proximity relations in tables 5, 6, 7). On the other hand phrases 3 and 5 have a high level of proximity and in practice it is easy to see that these phrases contain an important part of representative phrases 1, 6 and 7.

## 6. CONCLUSIONS AND FURTHER WORK

A method to search musical representative phrases using a indistinguishability operator and fuzzy proximity relations on the set of phrases is given. Future work will be to consider Yager measures of specificity to decide if the chosen representative phrase is reliable. Some other similarity and indistinguishability operators will be tested, as well as different fuzzy logic operators and distances to compute the proximity.

## REFERENCES

- [1] G. Beliakov, A. Pradera, T. Calvo. Aggregation functions: a guide for practitioners. Series: Studies in fuzziness and soft computing. Springer. 2007.
- [2] E. Klement, R. Mesiar, E. Pap. Triangular norms. Kluwer. Dordrecht. 2000.
- [3] R. E. Overill. On the combinatorial complexity of fuzzy pattern matching in music analysis. Computers and the Humanities, 27, 2, 105-110, 1993.
- [4] B. Schweizer, A. Sklar. Probabilistic metric spaces. North-Holland, Amsterdam, NL, 1983.

[5] E. Trillas, L. Valverde, On mode and implication in approximate reasoning, in: M.M. Gupta, et al., (Eds.), Approximate Reasoning in Expert Systems, Elsevier, North-Holland, Amsterdam, 1985.

[6] L. Valverde. On the structure of F-indistinguishability operators, Fuzzy Sets and Systems 17, 313–328, 1985.

[7] L. A. Zadeh. Fuzzy sets. Inform. and Control 8,338–353, 1965.

## AUTHOR:



Emerson Castañeda se desempeña como consultor en tecnologías de la información. Actualmente es estudiante de Doctorado en la Universidad Complutense de Madrid. DEA en Ingeniería Informática por la Universidad Complutense de Madrid en el año 2006, DEA en Ingeniería Informática en la especialidad Ingeniería del Software por la Universidad Pontificia de Salamanca en el año 2005 e Ingeniero de Sistemas por la Universidad Católica de Colombia en el año 2001.

# NOTA INFORMATIVA

## LA IMPORTANCIA DEL ISSN

La aportación de los autores a la revista y, por ende, al público en general, tiene una recompensa final para todos, el ISSN.

### *¿Qué es el ISSN?*

El ISSN (International Standard Serial Number / Número Internacional Normalizado de Publicaciones Seriadas) es un código numérico reconocido internacionalmente para la identificación de las publicaciones seriadas. El ISSN puede utilizarse siempre que haya que recoger o comunicar información referente a las publicaciones seriadas, evitando el trabajo y posibles errores de transcribir el título o la información bibliográfica pertinente. El ISSN identifica sin ambigüedades ni errores la publicación seriada a la que va asociado. Es el equivalente para las publicaciones seriadas de lo que es el ISBN para los libros.

### *¿Qué ventajas ofrece la posesión de un ISSN?*

· EL ISSN permite identificar en todo el mundo, de una forma única y sin ambigüedades, una publicación seriada, cualquiera que sea el país de edición o la lengua de la misma y sin importar que otras publicaciones seriadas lleven un título igual o parecido. Por ello el ISSN es un elemento básico en todos los procesos de información, comunicación, control y gestión referentes a las publicaciones seriadas.

· Asegura un medio de identificación preciso e inmediato cuando se hace un pedido.

· Permite un método de comunicación rápido y eficaz entre editores, distribuidores, libreros y agencias de suscripción, mejorando los circuitos de venta.

· El ISSN puede servir para la construcción de los códigos de barras de las publicaciones seriadas. (Que ya se incluye en la contraportada de esta misma revista EAN-13).

· La asignación de un ISSN comporta también la inclusión de los datos de la publicación en la Base de datos internacional del ISSN. Esta base de datos es, por su volumen (alrededor de 750.000 registros en 1996), cobertura (mundial) y fiabilidad de sus datos, un recurso informativo esencial sobre las publicaciones seriadas.

· En las bibliotecas y centros de documentación facilita las operaciones de identificación, adquisición y préstamos. Asimismo, la base de datos del ISSN es la fuente más exhaustiva y autorizada para la catalogación de las publicaciones seriadas.

· Su asignación es completamente gratuita.

· Para los docentes, el hecho de publicar un artículo en una revista que posee un ISSN, les reporta más puntos en su investigación. En definitiva, todos aquellos autores que publican sus artículos en Buran, recibirán más puntos de los que recibían antes.

· Se puede encontrar más información sobre la temática relacionada al número de ISSN y las publicaciones seriadas en el página web de la Biblioteca Nacional: <http://www.bne.es>.

### Direcciones de interés:

<http://www.bne.es>

<http://portal.issn.org/cgi-bin/gw/chameleon>

<http://www.barcodeisland.com/ean13.phtml>

<http://www.issn.org/>

### Y por último recordar:

· Si una revista se relanza y/o retoma con un mismo título y con una misma temática, se puede utilizar, en caso de poseerlo, el ISSN anterior, ya que este sigue siendo válido para las futuras ediciones de la publicación.



# BURAN 26

## CALL FOR PAPERS

**DEADLINE: 30 DE JUNIO DE 2010**

### Información general:

Buran es la revista de divulgación científica y cultural, editada por la Branca d'Estudiants de l'IEEE de Barcelona, el primer número de la cual salió en marzo del 1993.

Dirigida tanto a estudiantes como a proyectos, profesores universitarios, personal de investigación o profesionales en general, Buran intenta establecer un canal de comunicación técnico-científico, que permita a cada uno de ellos acceder a un conocimiento, tan aproximado como se pueda, del estado del arte de las tecnologías de la información y comunicación.

El principal objetivo de la revista es la divulgación de opiniones y de los trabajos que se realizan en universidades y empresas, relacionados con cualquiera de las muchas actividades del IEEE o de carácter humanístico.

### Información para los autores:

- Los artículos que se presenten deberán estar escritos en castellano o inglés.
- El formato de entrega de los artículos será primordialmente en archivo de Microsoft Word o de texto, aconsejando a los autores la inclusión de escritos gráficos o fotografías que faciliten la lectura y comprensión del escrito. En el caso de que se incluyan fotografías, se incorporarán en formato TIFF, en un fichero a parte.

- El artículo deberá contener un abstract en el que se haga un pequeño resumen introductorio sobre el tema a tratar i una breve biografía del autor/es.

- Aunque no hay restricción en el número de páginas en que conste el artículo, es recomendable cualquier extensión que comprenda un mínimo de 4 páginas y un máximo de 12.

- Dado que el número de páginas de la revista es limitado, se efectuará un proceso de selección de los artículos recibidos, quedando en archivo para próximos números aquellos que no pudieran ser publicados.

- El autor deberá hacer constar su nombre completo, el departamento al que pertenece (en caso de ser profesor o proyecto), escuela/facultad donde cursa sus estudios (en caso de ser estudiante), o empresa para la cual trabaja, dirección de correo electrónico y una fotografía a tamaño carnet en formato electrónico (preferiblemente en formato TIFF)..

- Se recuerda el carácter divulgativo de los artículos.

### Envío de artículos:

Los artículos se podrán enviar mediante correo electrónico o por correo ordinario a nuestro despacho:

### Branca d'Estudiants de l'IEEE

Edifici OMEGA - S105  
Campus Nord UPC  
c/ Jordi Girona 1- 3  
08034 Barcelona (Spain)

ó:

**buran@ieee.upc.edu**

### Fecha límite de entrega:

**30 DE JUNIO DE 2010**

Para cualquier duda o consulta sobre la recogida de artículos, o para comunicar cualquier sugerencia sobre la revista:

email: [ieee@ieee.upc.es](mailto:ieee@ieee.upc.es)  
web: <http://ieee.upc.es>

e-burancursosfoto  
robótica  
desarrolloWeb...

iecc.upc.es



The image displays two side-by-side screenshots of a computer monitor showing the website for the IEEE Buran program. The left screenshot shows the 'Buran' section, featuring a thumbnail of a magazine cover titled 'Buran'. The right screenshot shows the 'Buran20' section, also featuring a thumbnail of a magazine cover titled 'Buran20'. Both screenshots show a navigation bar at the top with links for 'Archivo', 'Edición', 'Ver', 'Favoritos', 'Herramientas', and 'Ayuda'. Below the navigation bar, there are sections for 'Cursos', 'Actividades', 'Noticias', 'Artículos', 'Eventos', 'Publicaciones', and 'Contacto'. The overall theme of the website is blue and white.



**“Creating  
and Innovating  
the Future”**

Rama de Estudiantes del  
**IEEE de Barcelona**





N90,

una de las regiones de formación estelar en  
la *Pequeña Nube de Magallanes*.

Las poblaciones ricas en estrellas recién  
nacidas que encontramos aquí permitirán a  
los astrónomos estudiar los procesos de  
formación de estrellas en un ambiente muy  
diferente al de nuestra galaxia.



*There are still too many things to know,*



**IEEE**  
BARCELONA STUDENT BRANCH