

# EDITORIAL

## «Decíamos ayer...»

Con esta célebre frase de Fray Luis de León resumimos el ánimo con que reemprendemos el camino de la revista BURAN. No podemos negar que Buran 10 fue para nosotros un gran hito, no sólo por la importancia de este número en nuestro sistema de numeración (os referimos al artículo escrito por Miguel Escudero), sino también por el punto de inflexión que supuso en la redacción: incorporación de nuevas caras y despedida de los viejos «buraneros». Pero nosotros seguimos ahí, al pie del cañón.

Estamos orgullosos de presentaros esta nueva Buran, la 11. Como en el primer numero, seguimos fieles al planteamiento original de nuestra revista: **Comunicación y Formación**.

Queremos que nuestra revista sea un nexo de **comunicación** entre los estudiantes y profesores universitarios, donde puedan expresar sus ideas, presentar sus trabajos, o sencillamente comentar el estado del arte en Ingeniería.

**Comunicación** también entre las distintas Ramas. Buran se convierte en una excusa para mantener la relación entre las Ramas que participamos en su edición: Barcelona, Valencia y la incorporación de Las Palmas, que en este número colabora en la obtención de artículos.

Estas nuevas aportaciones evidencian el incremento de actividad que está experimentando la sección estudiantil del IEEE en España, que esperamos siga en aumento, con la creación de nuevas Ramas (como la de las Palmas o la futura en Zaragoza) o Capítulos específicos de las diversas sociedades del IEEE (por ejemplo, la apertura del Capítulo de Comunicaciones en Barcelona). Desde aquí queremos animarles a continuar con esta tarea recién iniciada, y que vaya aumentando número a número.

Y **Formación**, tanto en el aspecto científico, con artículos que abarcan las distintas disciplinas en el campo de la Tecnologías de la Información y las Comunicaciones, como en el humanístico, con las ya habituales aportaciones de temática no estrictamente técnica.

Sirva la Torre de Collserola, motivo de nuestra portada, de metáfora de lo que queremos seguir representando para la comunidad estudiantil: referencia en el avance tecnológico y nodo de comunicación.

**La Redacción**

## EDICIÓN

### BARCELONA

Estíbaliz Bear  
Daniel González  
Antoni Homs  
Mariano Korman  
Xavier Lago  
Carlos López  
José A. López  
David Rincón  
Miguel A. Sastre  
Imma Valls

### VALENCIA

Erika Moreno  
Eduardo Pascual  
Aitana Sempere  
Rafael Vidal

### LAS PALMAS

Santiago González  
Jorge Isern  
Humberto José Ayala

## AGRADECIMIENTOS

Il. Dir. Antoni Elias Fusté, Elisa Pla, Ángel Cardama, Pere Camps, Jaume Clavera Ortiz y a los puntos de distribución en la UPC (Abacus, CPET y Kiosk Campus Nord).

## IMPRESIÓN

RET, s.a.l.

## FOTOMECÁNICA

Sistemes d'Edició

## DEPÓSITO LEGAL

B-19.950-96

## COORDINACIÓN

### BARCELONA

Estíbaliz Bear

### VALENCIA

Aitana Sempere

### LAS PALMAS

Humberto José Ayala

## REVISIÓN

Estíbaliz Bear  
Mariano Korman  
Josep Pegueroles  
David Rincón  
Imma Valls

La portada muestra la torre de Telecomunicaciones de Collserola, en Barcelona. Agradecemos al "Centre de Telecomunicacions de la Generalitat de Catalunya" la cesión de la fotografía.

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.



## PERDER LA CABEZA EN EL 2000

Miguel Escudero

Profesor del Departamento de Matemática Aplicada y Telemática de la UPC  
escudero@mat.upc.es

Existen miles de lenguas y decenas de alfabetos, pero podría hablarse de las cifras como del único y auténtico lenguaje universal. Georges Ifrah, autor de una reciente y enciclopédica historia de las cifras, ha destacado la ausencia de una torre de Babel de los números, los cuales, una vez asimilados en un sistema, han sido comprendidos en todas partes de la misma manera. El sistema de numeración que tenemos ha sido denominado perfecto por ser el más breve

*El sistema de numeración que tenemos ha sido denominado perfecto por ser el más breve y eficaz en signos de los que ha habido y podamos imaginar, y su clave de bóveda es el cero.*

y eficaz en signos de los que ha habido y podamos imaginar, y su clave de bóveda es el cero. Puede decirse del cero que es un hallazgo tardío y que está dotado de dos sentidos: el propiamente numérico de cantidad nula y el de 'vacío'. Por cierto que *sifra* significa vacío en árabe, y cabe imaginarse al gran Carl Gauss escribiendo sus obras en latín y emplear, todavía en 1801, la palabra *cifra* con el significado de cero.

Autores arábigos-musulmanes fueron los llamados a vincular numerosos conocimientos de la India con los de Occidente, entre ellos el revolucionario cero, y mostraron al mundo que la ciencia es de quien la trabaja y no es propiedad de una casta ni de un pueblo *potentado*. Pero trabajar la ciencia exige saber dudar, lo cual supone la necesidad de hacerse con un método, Ortega dijo de la duda que es «la hermana bizca que tiene la ciencia». Para ser fiel a sí misma, la ciencia tiene que saber cuestionarse su propio quehacer y requiere una dimensión externa para proyectar su mejor luz. La ciencia es necesaria pero no suficiente en nuestra existencia: no puede hacerse cargo, por ejemplo, de la realidad fundamental que es la vida *humana*, en la que radican todas las demás, y la cual sólo es posible cuando se interpreta y entiende como tal vida y como un área, por tanto, en donde cohabita el pasado con el futuro. Cuando los hombres y los pueblos están seguros de lo que no

pueden estarlo pierden la razón por partes. Y este desarme es la mayor amenaza que pende sobre la condición humana.

El propio Ortega escribió en su *Renan* que «la verdad tiene este privilegio eucarístico de vivir a un tiempo e igualmente en cuantos cerebros se lleguen a ella». No ha lugar a razas superiores con las que justificar la explotación de seres «inferiores», desposeídos de su carácter personal. Hoy como ayer, hay - ¿quién se atreverá a negarlo, cómo podría hacerlo? - talentos desigualmente repartidos y concentrados en las distintas sociedades. Pero hay que recordar que no hay superioridad 'natural' ni automática que sea *venerable*, la que merece reconocimiento y *emulación* es la que deriva de un esfuerzo exigente, coherente e imaginativo, y parte de saber que la inteligencia que la sustenta siempre es problemática y desde luego insegura.

Roma fue un pueblo que alcanzó un nivel técnico muy alto, con una aceleración de progreso superior a la de otras sociedades. Sin embargo, nos legó un sistema de numeración inútilmente complicado, huella indudable de una forma de pensar arcaica. (¿Alguien pretende ser el primero de la clase en todas las asignaturas y de manera sempiterna? Si lo hubiera,

*Trabajar la ciencia exige saber dudar, lo cual supone la necesidad de hacerse con un método*

habría que repetirle lo que escribió Quevedo: «en todo mi propia ignorancia me sirve de penitencia».) De modo fatal, aquellos tiempos de esplendor y miseria ya pasaron y no volverán, pues el tiempo ni vuelve ni tropieza. Podemos, no obstante, indagar la clave de aquella forma de pensar y de vivir, y si la captamos acaso encontraremos afinidades que nos sirvan de referencia y de guía en nuestro camino personal, pero nunca una *identidad* que haga que aflore o se libere nuestra intimidad. (El pensador italiano Vattimo ha denominado a Occidente como «lugar de la negación de la identidad», pero en nuestra Península hay políticos con cotas de poder empeñados en contradecirlo

y que declaran la identidad como mito justificador de sus proyectos; por eso pugnan por hacernos vivir en países *virtuales*, impregnados de ficciones.)

Si contemplamos a las distintas sociedades occidentales, se observa, al lado de un desarrollo tecnológico deslumbrante, un innegable auge social de la superstición -refugio de frustraciones y aburrimiento. La tendencia a dar excesivo valor a las nociiones científicas o pretendidamente científicas se conoce como *cientificismo* y es una enfermedad del espíritu científico, la cual puede afectar tanto a legos como a eruditos. El mencionado Ifrah destaca la extraña paradoja de una civilización milenaria como la nipona, actualmente en la cima de la ciencia y de la tecnología, pero que ha conservado supersticiones y temores que nadie pensaría subvertir desde dentro. Así, recuerda el temor de los japoneses hacia ciertos números, que ha llegado al punto de eliminar de hoteles y clínicas el 304 y el 309; a excluir de los aparcamientos los guarismos 4, 9, 14, 19 o 24, o bien suprimir de los aviones el 4 (se debe indicar que en el sistema sino-japonés el cuatro se denota *shi*, «la muerte», y esta agorera etimología ha sido reprimida y reemplazada por otro término *inofensivo*).

*Ifrah recuerda el temor de los japoneses hacia ciertos números, que ha llegado al punto de eliminar de hoteles y clínicas el 304 y el 309; a excluir de los aparcamientos los guarismos 4, 9, 14, 19 o 24, o bien suprimir de los aviones el 4*

---

En el origen de las grafías de las cifras, se hallan vestigios de un pasado inmemorial con un pensamiento lógico muy retrasado. Abreviaturas que además de servir para anotar y retener los números, parece que buscaban confundir a los profanos y ocultarles ciertos significados de fuerzas e ideas mágicas que permitían conectar con una *realidad profunda*. En una curiosa relación entre cifras y letras, la civilización mesopotámica adoptó para el número diez mil la etimología de la cantidad «más grande» y, por otro lado, la sutil cultura india denominaba el billón como *parardha* («mitad del más allá») y al cinco lo expresaba por pavana («purificación»).

De los dígitos de la mano se pasó a los ábacos (que la Revolución francesa prohibiría). De éstos se saltó a las máquinas mecánicas de calcular, y luego a los ordenadores, siempre a partir de unos mecanismos de pensamiento que desbloquearon la imaginación. Por cierto que el año que viene se celebrará la muerte

de Ramón Verea a los sesenta y seis años de edad. Verea fue un gallego que se embarcó hacia Cuba en 1854, después de abandonar sus estudios eclesiásticos comenzados cinco años antes. Este hombre de letras trabajó durante diez años en la Perla del Caribe como maestro de escuela y como periodista, hasta que se trasladó a Nueva York. Familiarizado con las máquinas impresoras, se dispuso a construir una calculadora y en 1878 patentó una pequeña máquina capaz de efectuar directamente todas las multiplicaciones con números de dos cifras, por la que fue galardonado y aún hoy día recordado. Cuando unos periodistas neoyorquinos le preguntaron cómo iba a realizar la comercialización de su invento, él contestó que esa cuestión no le interesaba, ya que su único objetivo había sido probar que el problema podía ser resuelto, y sobre todo demostrar que un hispano era tan capaz como un norteamericano de dar pruebas de astucia e imaginación.

Este final de siglo ha ofrecido un error de extensión universal. Hace cosa de dos decenios, con el objeto de ahorrar memoria, se eliminaron de numerosos programas informáticos los dos primeros dígitos del número del año. En lugar de anotar los ocho dígitos de una fecha, correspondientes a año, mes y día (por este orden), se anotaron los seis últimos; obsérvese la conservación de dos cifras para los meses: febrero, por ejemplo, viene dado por 02, el cual no se trata de un cero a la izquierda al llevar en esa secuencia el año por delante. Este ahorro generalizado ha costado caro, puesto que «la ciudad alegre y confiada» de los ordenadores se ha encontrado con el paso cambiado ante la llegada del nuevo milenio, y se ve obligada a rectificar. En efecto, en caso contrario, al llegar al año 2000, los programas pasarían a marcar la fecha 00, dos cifras que nos devolverían a comienzo del siglo XX. El problema no es que sean «ceros a la izquierda», sino que se manifestaría al medir el tiempo y tomar distancias entre fechas: el 2000 menos el 1998 señala un salto de dos años, pero el 00 menos el 98 resulta -98, con lo que el caos estaría organizado. Todavía no podemos manejar el tiempo al buen tuntún. Curiosamente, este inesperado problema de programación no admite, hoy por hoy, una solución universal y requiere actuar con correcciones particulares y hechas a la medida para solventarlo. Por cierto que no poca gente anda especializada en remediar estos entuertos. Son ironías de los tiempos, que encierran su moraleja: hay que tomar siempre la distancia adecuada para no perder el sentido histórico. La imaginación es una fuerza imperecedera que hay que cultivar incansablemente. A menudo nos apresuramos a menospreciar o tirar cosas que por pereza o poco juicio, no llegamos a comprender suficientemente bien. En este denominado «efecto 2000» mucha e importante gente perdió la cabeza por mimesis colectiva, recuperarla lo tienen que hacer ahora de prisa, por partes y de forma individual. ■





## VISITA A LA TORRE DE COLLSEROLA

Carlos López, David Rincón

Capítulo de la Communications Society. Rama de Estudiantes del IEEE de Barcelona.  
ieee@casal.upc.es

El pasado día 1 de abril el Capítulo de Comunicaciones de la Rama de Estudiantes del IEEE de Barcelona realizó una visita a las instalaciones de la Torre de Comunicaciones de Collserola, gracias a la gentileza del Centre de Telecomunicacions de la Generalitat de Catalunya. En este artículo se van a describir las instalaciones que posee el complejo.

### LA TORRE DE COLLSEROLA

La Torre de Telecomunicaciones de Collserola, diseñada por el arquitecto británico Sir Norman Foster, inició su construcción en febrero de 1.990 y entró en servicio el mes de junio de 1.992, con motivo de la celebración en Barcelona de la XXV edición de los Juegos Olímpicos.

La promotora del proyecto y actual propietaria es la Sociedad Anónima Torre de Collserola cuyos accionistas son Telefónica S.A. con un 42%; Retevisión S.A. con un 36%; Centre de Telecomunicacions de la Generalitat de Catalunya con un 17% y la Entidad Metropolitana del Transporte con un 5%. Puede decirse que es el resultado del esfuerzo conjunto de los grandes operadores de radiocomunicaciones del país, y también se cuenta con la participación de los gobiernos locales.

La Torre tiene una altura total de 288 m. y está situada en la Sierra de Collserola, que limita con la ciudad de forma paralela al Mar Mediterráneo. El lugar del emplazamiento se denomina Turó de la Vilana, a 445 m. sobre el nivel del mar, y la construcción emerge de esta altura un total de 268 m.

El complejo está compuesto por la Torre y un edificio subterráneo destinado a soporte de instalaciones. La Torre propiamente dicha consta de tres partes: un fuste de hormigón de 205 m. de altura y tan sólo 4,5 m de diámetro (con un hueco interior de 3m.), un mástil tubular de acero de 38 m. (2.7 m. de diámetro), y un tramo de celosía de 45 m (0.7 m. de diámetro) que culmina la Torre.

Entre la cota 84 y la 152 se ha construido una estructura metálica, con un peso total de 3000 Tm, compuesta por 13 plataformas, cuya superficie total es de 5.550 m<sup>2</sup>, siendo la separación entre plataformas de 5,62 m. La décima plataforma está destinada a Mirador Público situado a una altura sobre el nivel de mar de 560 m., lo que permite una excelente vista sobre la ciudad, su

área metropolitana y hasta más de 70 Km. de distancia. A este Mirador se accede desde el interior de la montaña mediante un elevador totalmente panorámico, exterior al fuste cuya velocidad de 1 m/s lleva al público al observatorio en menos de 2 minutos.



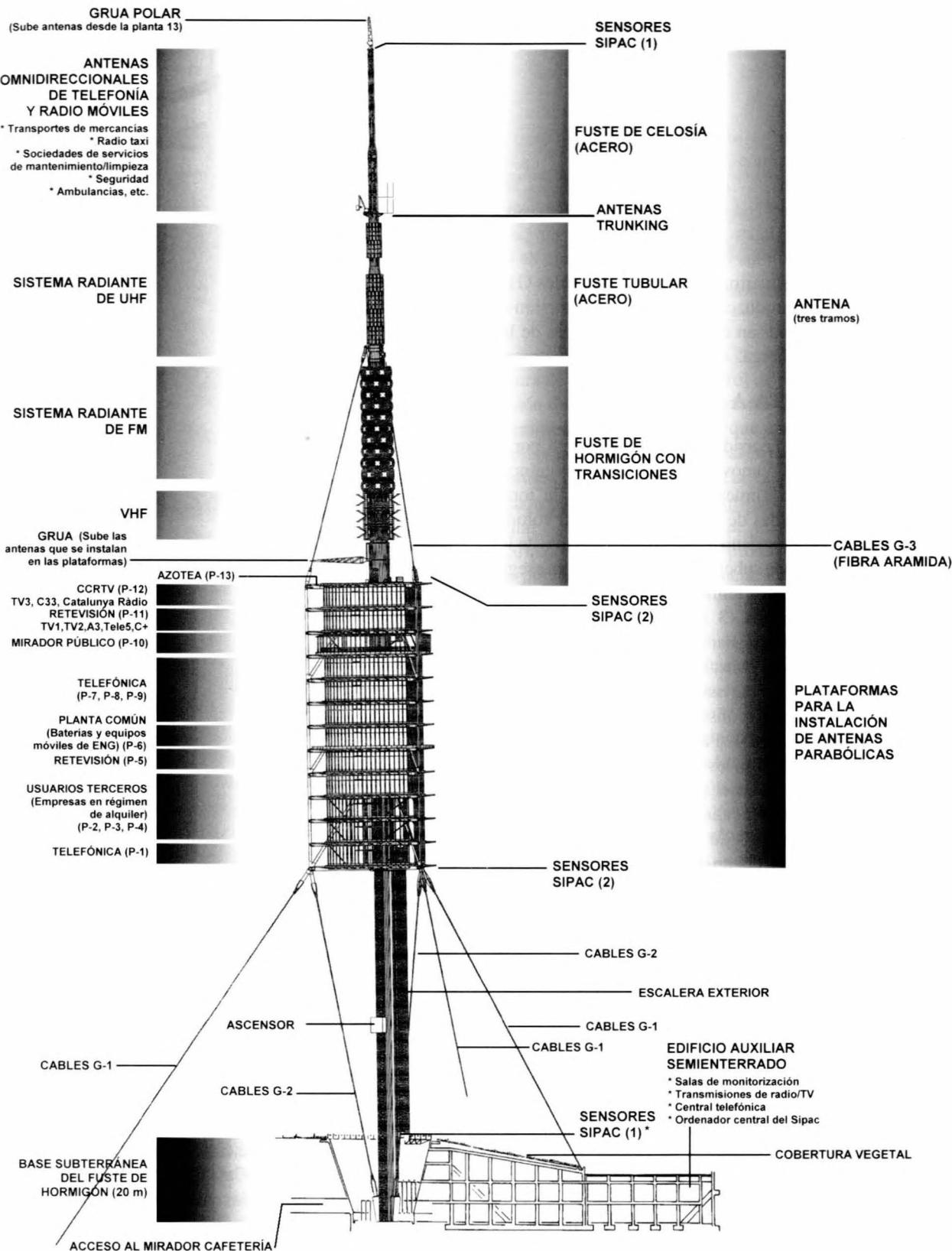
Vistas desde el mirador. A la izquierda se observan las instalaciones de Telefónica en la montaña de Sant Pere Martir.

Se dispone de un segundo ascensor de servicio igual que el primero, pero no panorámico, destinado a los servicios de telecomunicaciones y su mantenimiento.

El edificio auxiliar subterráneo dispone de una superficie útil de 4.360 m<sup>2</sup>, y alberga los equipos emisores de radio y televisión, la sala de control y la administración de la torre. El criterio básico para el diseño del proyecto fue la minimización del impacto ambiental y visual. Por ello se decidió que el edificio no fuera visible desde la ciudad, permitiendo así la conservación del perfil original de la montaña.

### LA CONSTRUCCIÓN DE LA TORRE

Una vez adjudicado el proyecto definitivo al arquitecto Norman Foster y la construcción a la empresa Cubiertas y MZOV, se inician en el mes de marzo de 1990 los primeros trabajos de hormigonado. El primer paso fue construir el fuste de hormigón mediante la técnica de encofrado deslizante hasta la cota 185m. (siempre estabilizado mediante cables de atracción). Al llegar a la cota 27 metros, se introdujeron dentro del fuste de hormigón los tramos del mástil de acero y el de celosía, que una vez acabada la construcción de la estructura de la torre serían izados de forma telescopica.



\* El número entre paréntesis indica la cantidad de sensores SIPAC instalados en cada nivel.

Diagrama cedido por la revista "Barcelona Metrópolis Mediterrània"

La estructura de plataformas se construyó simultáneamente, al pie del fuste. Tras la finalización, se procedió a su elevación mediante 9 gatos hidráulicos situados en el extremo superior del fuste. El recorrido total de la operación fue de 77 metros, a una velocidad de 2 m/hora, lo que supuso una duración total de 3 días, en junio de 1991.

Posteriormente, se montaron los cables definitivos que unen el fuste con la plataforma 13 (cables G3) y los que unen la plataforma 1 con tierra (cables G1 y G2). Los cables G3 están fabricados de kevlar (material transparente a la radiación electromagnética) para evitar cualquier modificación en los diagramas de radiación de las antenas de la parte superior de la torre. Los cables G1 y G2 impiden que se produzcan movimientos de torsión de la estructura, que podrían causar desalineaciones de los radioenlaces que alberga.

### VISITA A LA TORRE

Iniciamos el recorrido entrando al recinto por el edificio auxiliar. Allí vimos, en primer lugar, los equipos relacionados con la alimentación general de la torre. Existen dos acometidas de alta tensión (11.000 voltios) que proporcionan el suministro eléctrico, cada una de las cuales proviene de una subestación diferente, para asegurar la continuidad de la alimentación. En esta etapa se convierte a 380 voltios (3 fases más neutro). Como ejemplo, se puede mencionar que los equipos del Centre de Telecomunicacions consumen unos 600 amperios, y los pertenecientes a Retevisión, unos 1200 amperios, siendo actualmente el consumo total de la torre de aproximadamente 1,5 MW. Para el caso en que se produjese un fallo simultáneo en las dos acometidas, se dispone de cuatro generadores diesel marinos. Dichos generadores son motores de cárter caliente para reducir el periodo de calentamiento propio de los motores diesel y los problemas derivados de un arranque en frío. Dichos motores se usan al 60 % de su capacidad total para evitar situaciones peligrosas, pudiendo generar cada uno de ellos 1 MW. Debido a estas características los motores son capaces de estar a pleno rendimiento en unos 5 segundos. En las condiciones actuales dos motores bastan para proporcionar el consumo total de la torre. Para evitar cualquier discontinuidad en las emisiones, se dispone de un sistema de alimentación ininterrumpida capaz de cubrir las necesidades de la torre durante un periodo aproximado de 25 minutos.

Posteriormente se pasó a ver la sala de refrigeración de agua. Debido a la gran cantidad de potencia que no se radia, sino que se disipa en forma de calor, es necesario el uso de un eficiente sistema de refrigeración de forma permanente. Dicho equipo de refrigeración consta de un sistema de distribución formado por una potente bomba y el sistema de tuberías que llevan el agua hasta los sistemas a refrigerar (tanto en el edificio como en la parte superior de la torre), y del sistema que mantiene el agua a una temperatura adecuada (10 grados centígrados) para llevar a cabo su función.

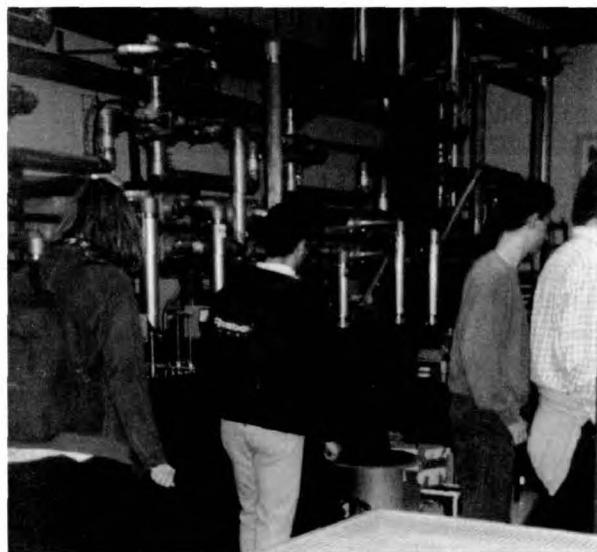


Equipos de los radioenlaces ubicados en la planta 12

Acto seguido nos dirigimos hacia el ascensor panorámico para subir hasta el mirador en la planta 10, a 135,5 m. de la base de la torre. En dicha planta, con una vista de 360 grados, pudimos admirar una hermosa panorámica de la Ciudad Condal y la comarca del Vallès gracias a unas inmejorables condiciones meteorológicas. Desde dicho mirador pudimos observar dos de los cuatro radioenlaces de Retevisión que constituyen la conexión con las islas Baleares. Dichos radioenlaces, situados en las plataformas 5 y 11, utilizan las técnicas de diversidad en espacio y en frecuencia para asegurar una correcta recepción de las señales pese a la enorme distancia a salvar (180 kilómetros). Después subimos por la escalera de servicio hasta la planta 12, ocupada por el Centre de Telecomunicacions. Allí se encuentran equipos relacionados con diversos tipos de servicios como radioenlaces para la Corporació Catalana de Ràdio i Televisió, Trunking (Policía Autonómica, empresas de mensajería, etc...), Digital Audio Broadcast (DAB, también conocida como «FM digital»), Digital Video Broadcast (DVB, televisión digital terrena), enlaces de datos para el CESCA (Centre de Supercomputació de Catalunya) y las conexiones a Internet de las universidades catalanas, así como diversos equipos emisores de televisiones locales (Barcelona Televisió, Televisió de Rubí, Televisió de l'Hospitalet, etc.). Allí se nos explicó que los arrays de dipolos situados en la parte superior del fuste, encargados de difundir las señales de televisión y radio a una población de 4 millones de habitantes, se apuntan de forma electrónica (mediante un desfase de alimentaciones). El ángulo de inclinación del haz es mayor en la parte encargada de iluminar

Barcelona que en la parte orientada hacia el Vallès, obteniendo en el segundo caso un mayor alcance.

Seguidamente bajamos en el ascensor de servicio de vuelta al edificio auxiliar para visitar la zona ocupada por los equipos emisores pertenecientes al Centre de Telecomunicacions. En los pasillos pudimos observar los cables coaxiales de 6 pulgadas (15,24 cm) encargados de transportar la señal, ya modulada y amplificada a las antenas situadas 200 m. más arriba con unas pérdidas no superiores a 1 dB. Acto seguido entramos en la sala en la que estaban ubicados los equipos moduladores y amplificadores. En dicha sala, como en el resto de la torre, todos los equipos están duplicados para asegurar un servicio continuo. Allí pudimos observar los emisores de las cadenas de TV y radio de la Generalitat (TV3, Canal 33, Catalunya Ràdio, Ràdio Associació, Catalunya Música, Catalunya Informació). La potencia de equipo total de la torre es de unos 240 kW en FM y unos 140 kW en televisión. Especialmente espectacular es la instalación de microondas encargada de multiplexar los diferentes canales en una sola señal, así como la carga de 50 ohmios (puramente resistiva), refrigerada por líquido, capaz de disipar 30 kW. Esta sala alberga los commutadores ATM y routers para transmisión de datos.



Sala de amplificadores. Al fondo, los multiplexadores de microondas.

Finalmente accedimos a la sala de control, presidida por un panel de monitores que permiten visualizar las diversas señales provenientes de los centros emisores secundarios de la red de televisión (Rocacorba en Girona, Alpicat en Lleida, etc). En esta sala se reciben y distribuyen las señales provenientes de diversas cámaras utilizadas en los boletines meteorológicos («beauty shots»), así como las desconexiones locales de los centros territoriales y los intercambios de noticias con redes exteriores (Retevisión, FORTA, etc). Se dispone de un sistema de telecontrol capaz de gestionar todos los centros emisores del territorio catalán desde un único punto. Este sistema

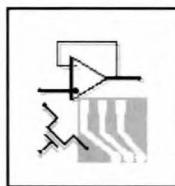


Sala de control. Se pueden observar los equipos de control y monitorización.

permite monitorizar variables como el estado y la temperatura de los equipos, el control de entradas o la comunicación de las señales transportadas por los radioenlaces, permitiendo de esta forma que no sea necesaria una presencia constante de personal en todos y cada uno de los centros. El control de calidad de la emisión queda garantizado por otro sistema automático que monitoriza cada señal a su entrada y salida, de manera que se disparan alarmas en el caso de que se produzca alguna incidencia no prevista (bajada de potencia, desaparición de señal, etc).

En una dependencia anexa a la sala de control, se encuentran las instalaciones del Servicio Automático de Información Hidrológica (SAIH). Este sistema, gestionado por la Junta d'Aigües de Catalunya, tiene como finalidades principales prevenir y minimizar los daños causados por las inundaciones y optimizar la gestión de los recursos hídricos. En la Torre de Collserola se centraliza la recepción de los datos sobre el caudal de los ríos y el nivel de los pantanos catalanes. La Torre también aloja el Sistema de Predicción Atmosférica de Catalunya (SIPAC), que permite prever, con un margen de 48 horas, los niveles de contaminación y los fenómenos de inversión térmica en Barcelona y las comarcas de los alrededores. Dispone de un ordenador y un conjunto de sensores para medir datos meteorológicos (nivel de oscilación, humedad, dirección y velocidad del viento, pluviometría, presión, temperatura...) en cinco niveles distintos de la torre. A partir de esta información y de la enviada por otras estaciones situadas en el área metropolitana, se elabora una predicción de la evolución de los episodios de contaminación atmosférica. Una vez vista la sala anexa de control, se dio por concluida la visita.

Tan sólo nos queda agradecer al señor Jaume Clavera Ortiz, Jefe del Centro de Operaciones y Control de Red del Centre de Telecomunicacions, la colaboración y disposición mostradas durante la realización de esta visita. También agradecemos a Joan Anton Benach, director de la revista Barcelona Metrópolis Mediterránea, la cesión de material gráfico.



# COMO CONVERTIR LA SOUND BLASTER EN UN INSTRUMENTO VIRTUAL

Manuel Vargas

Profesor Titular del Departament d'Enginyeria Electrònica de la UPC  
mvargas@eel.upc.es

Mediante el software adecuado es posible convertir la tarjeta de sonido Sound Blaster en un instrumento de laboratorio (osciloscopio, analizador de espectros, frecuencímetro, ...). Este artículo va especialmente dirigido a aquellos aficionados a la Electrónica que siempre habían soñado con tener esos instrumentos en casa pero no disponían de los medios económicos necesarios. ¡Ahora ese sueño puede convertirse en realidad!

## INTRODUCCIÓN

Recientemente se ha puesto de moda el concepto de instrumento virtual. Podríamos definir un instrumento virtual como un software que presenta en la pantalla del ordenador una imagen parecida al panel frontal de un instrumento clásico. Por ejemplo en un osciloscopio virtual aparecen en la pantalla del ordenador los típicos mandos de control de la amplitud y de la base de tiempos que se pueden rotar mediante el ratón. Naturalmente este software ha de controlar también un hardware que es el que efectivamente realiza la medida o genera la señal. Este hardware puede consistir en instrumentos con capacidad de comunicarse con el ordenador mediante un bus adecuado o bien en tarjetas de adquisición de datos conectadas directamente al bus del ordenador. Hay que aclarar que las medidas obtenidas con un instrumento virtual son tan reales como las obtenidas con un instrumento de medida clásico. En el curso de laboratorio de Instrumentación Virtual que imparto en la E.T.S. de Ingeniería de Telecomunicación de Barcelona se contro-

lan diversos instrumentos (osciloscopio, multímetro, generador de funciones) así como una tarjeta de adquisición de datos mediante un ordenador PC. Algunos estudiantes se muestran realmente interesados en la instrumentación virtual, hasta el punto de desear adquirir instrumentos de ese tipo. Sin embargo su precio resulta prohibitivo. Esto me hizo pensar si no habría una forma barata de que los estudiantes pudieran tener sus propios

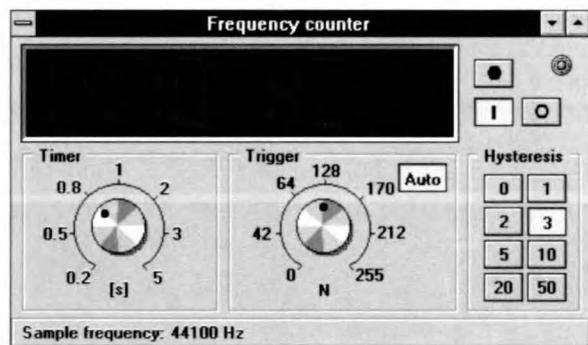


Figura 2. Frecuencímetro virtual. La frecuencia representada corresponde a la señal de la figura 1.

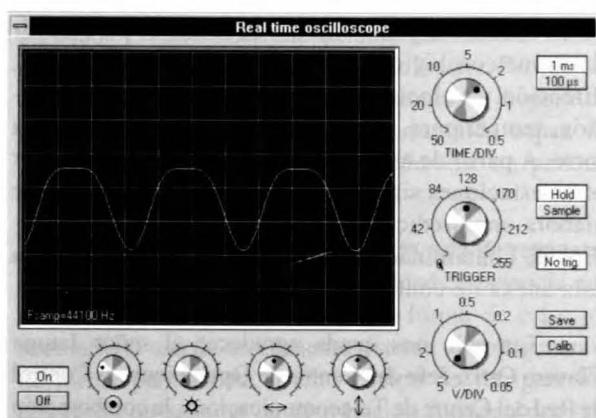


Figura 1. Osciloscopio virtual. La señal representada se ha obtenido acercando el micrófono al altavoz de graves.

instrumentos. Encontré la solución en la tarjeta de sonido Sound Blaster. Actualmente muchos PC ya vienen equipados con una tarjeta de sonido compatible Sound Blaster. Pensé que la adquisición y generación de señal con esta tarjeta no tenía porque limitarse a la voz sino que en principio tendrían que poder medirse y generarse tensiones de todo tipo siempre que estuvieran en la banda de audio. Sólo quedaba encontrar el software adecuado. Así que empecé a buscar en Internet y efectivamente encontré lo que buscaba.

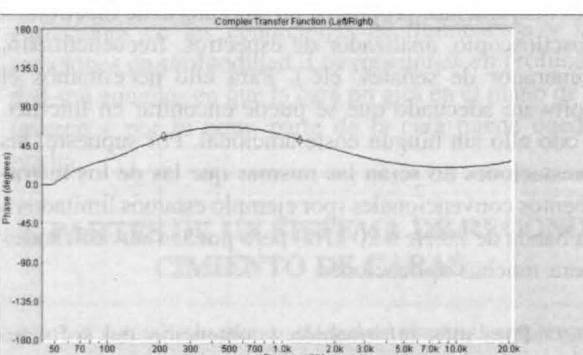
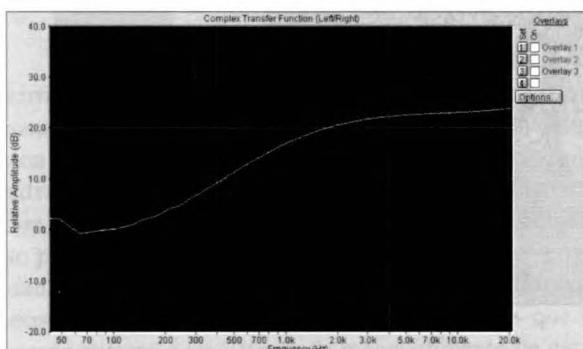
## EJEMPLOS

En la figura 1 se observa un osciloscopio virtual. Aparece en pantalla apretando simplemente sobre el icono correspondiente. Mediante el ratón se pueden ajustar los controles de amplitud (V/div) y de la base de tiempo (time/div) de manera similar a la de un osciloscopio convencional. La señal se obtuvo acercando el micrófono al altavoz de graves hasta que se produjo el típico pitido. Si queremos medir la frecuencia con mayor exactitud abrimos el frecuencímetro virtual (figura 2) apretando sobre el icono correspondiente. También aquí observamos algunos de los controles típicos en un frecuencímetro.



**Figura 3.** Analizador de espectros virtual. La señal de entrada se ha obtenido acercando el micrófono al altavoz de agudos. Se observa el pico a 2358 Hz correspondiente a la frecuencia fundamental así como los armónicos.

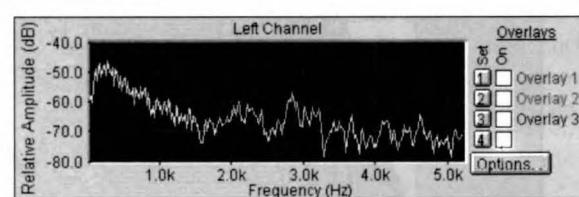
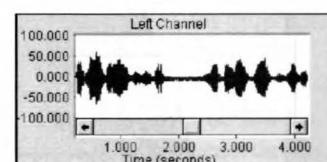
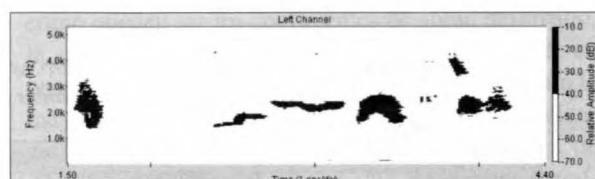
El analizador de espectros virtual de la figura 3 trabaja en tiempo real pues el código es muy rápido (la FFT está escrita en ensamblador, el resto en C++). También tiene algunos de los controles típicos en un analizador de espectros: escala de amplitud logarítmica o lineal, elección del tipo de ventana, detección de pico, etc. También tiene un control para la Sound Blaster que permite seleccionar las entradas (micrófono, línea y/o CD ROM) así como controlar la amplificación de cada una. El espectro representado corresponde a una señal



**Figura 4.** Módulo (arriba) y fase (abajo) de la función de transferencia de un filtro.

obtenida al acercar el micrófono al altavoz de agudos. Puede observarse un pico a la frecuencia fundamental (2358 Hz) así como otros picos menores correspondientes a los armónicos (múltiplos enteros de la frecuencia fundamental). Este software tiene la ventaja de que se proporciona el código fuente (Borland C++) con lo que puede ser muy útil para aquellos que deseen crear sus propios instrumentos virtuales para la Sound Blaster.

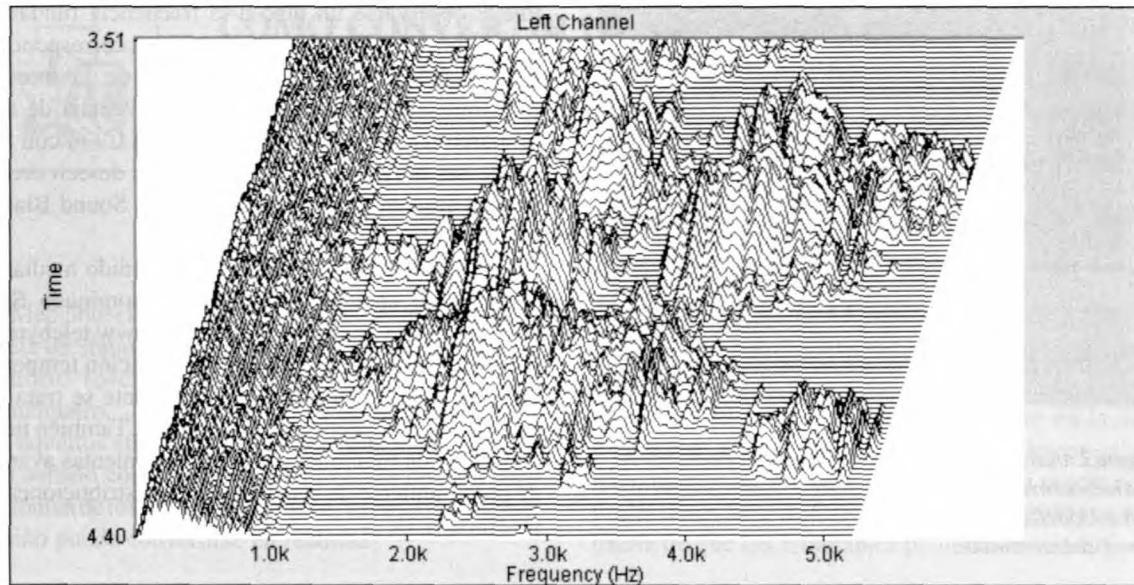
Las figuras 4, 5 y 6 se han obtenido mediante un software de carácter profesional denominado Spectra Plus. Se puede bajar de la dirección [www.telebyte.com/pioneer](http://www.telebyte.com/pioneer) y obtener una llave de autorización temporal sin cargo ni obligación alguna. Básicamente se trata de un analizador de espectros de dos canales. También tiene un generador de funciones. Incluye herramientas avanzadas de procesamiento de señal, como las distribuciones tiempo-frecuencia.



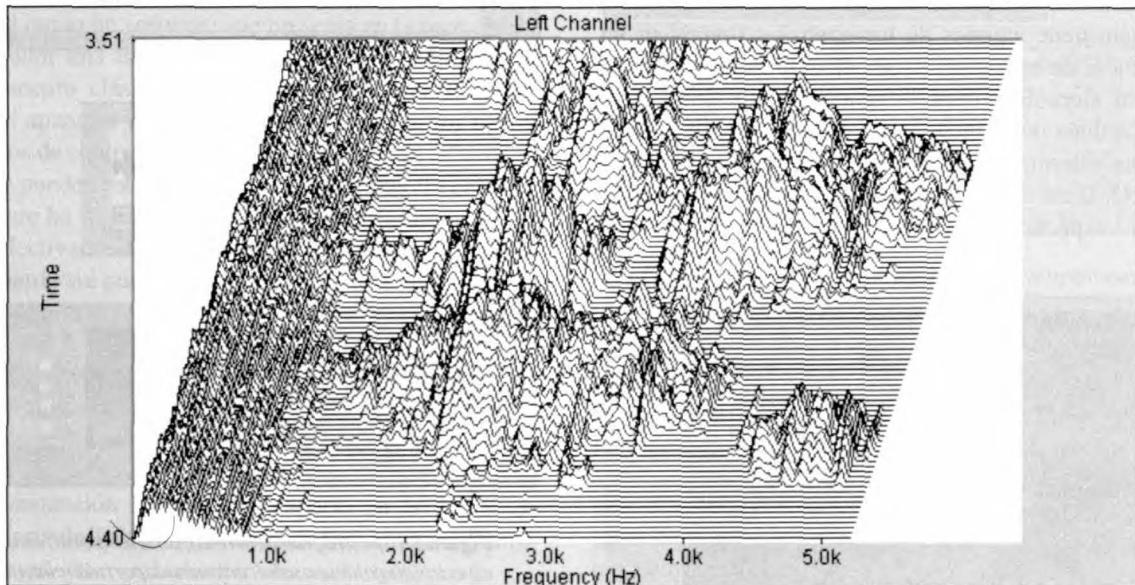
**Figura 5.** Espectrograma (arriba), serie temporal (centro) y espectro (abajo) de una señal contaminada por ruido e interferencia.

En la figura 4 se ha obtenido el módulo y la fase de la función de transferencia de un filtro. En la figura 5 se ha obtenido el espectrograma (el eje horizontal representa el tiempo, el eje vertical frecuencia y las zonas oscuras representan una amplitud mayor), la serie temporal y el espectro de una señal contaminada por ruido e interferencia. En la figura 6 tenemos la representación de la variación del espectro respecto al tiempo en 3D de la señal anterior. Tanto el espectrograma como la superficie 3D son útiles en el caso de señales cuya frecuencia varía en función del tiempo. En este caso la FFT no da resultados correctos. Se observa claramente el ruido en la banda inferior a 1 kHz. La señal (en este caso corresponde al canto de pájaros) está situada por encima de 1 kHz. Se ha filtrado la señal con un filtro paso alto con frecuencia de





**Figura 6.** Representación 3D del espectro en función del tiempo de una señal contaminada por ruido de baja frecuencia (hasta 1 kHz) antes (arriba) y después (abajo) de aplicar un filtro paso alto.



corte a 1 kHz. El efecto se observa visualmente en la figura. También podemos oír el efecto reproduciendo la señal mediante la tarjeta de sonido.

**Una advertencia:** si en lugar de utilizar el micrófono queremos medir tensiones directamente, conviene aislar galvánicamente la entrada de la tarjeta de sonido de la señal a medir, por ejemplo mediante optoacopladores, para evitar dañar la tarjeta en caso de que la tensión aplicada sea excesiva.

## CONCLUSIONES

Si poseemos un PC con tarjeta de sonido compatible Sound Blaster tenemos a nuestra disposición diver-

sos instrumentos propios de un laboratorio de Electrónica (osciloscopio, analizador de espectros, frecuencímetro, generador de señales, etc.). Para ello necesitamos el software adecuado que se puede encontrar en Internet. Todo ello sin ningún coste adicional. Por supuesto, las prestaciones no serán las mismas que las de los instrumentos convencionales (por ejemplo estamos limitados a la banda de 20 Hz a 20 kHz) pero pueden ser suficientes para muchas aplicaciones.

Para más información y obtención del software puedes consultar Buran en internet en la dirección:

**<http://citel.upc.es/users/buran/>**  
**<http://citel.upc.es/users/buran11/index11.htm>**



# REPRESENTACIÓN DE CARAS MEDIANTE EIGENFACES

Luis Lorente Giménez

Proyectista del Dpto. de Teoría del Señal y Comunicaciones, Grupo de Procesado de Imagen,  
Universidad Politécnica de Catalunya  
alorente@gps.tsc.upc.es

## 1. INTRODUCCIÓN

El proyecto de fin de carrera en el que estoy trabajando desde hace casi un año consiste en el desarrollo de un sistema de reconocimiento de caras. Las tecnologías de reconocimiento de caras (*face recognition technologies* o FRT) son un área de investigación muy activa en los últimos años que engloba varias disciplinas como el procesado de imagen, redes neuronales, *pattern recognition* y visión por ordenador. Tiene varias aplicaciones potenciales, tanto comerciales como de seguridad, como por ejemplo la identificación de fotos de carnets, pasaportes o identificación en tiempo real de personas desde imágenes de cámaras de vigilancia. El reconocimiento de caras, aunque fácilmente realizable por las personas, es difícilmente implementable de una manera totalmente automatizada en ordenador.

El objetivo de un sistema de reconocimiento de caras es, generalmente, el siguiente: dada una imagen de una cara “desconocida” (o imagen de test) encontrar una imagen de la misma cara en un conjunto de imágenes “conocidas” (conjunto de entrenamiento).

Esto, que es tan sencillo de decir, plantea unas terribles dificultades para su implementación en un sistema automático. Las personas reconocemos las caras con gran facilidad, rapidez e incluso en circunstancias adversas de iluminación, orientación y otros factores. Sin embargo los sistemas automáticos de reconocimiento por ordenador todavía están muy lejos de la efectividad del cerebro humano, aunque los resultados son esperanzadores. Las mayores dificultades a las que se puede enfrentar un sistema de reconocimiento de caras son: las variaciones en la expresión de las caras, las variaciones en las condiciones de iluminación y las rotaciones en profundidad. Las rotaciones en profundidad son aquellas en que la cara no gira en el plano de la imagen y por lo tanto parte de la cara puede quedar oculta.

## 2. PARTES DE UN SISTEMA DE RECONOCIMIENTO DE CARAS

A pesar de la gran variedad de sistemas existentes, todo sistema de FR puede dividirse en tres partes:

■ preprocessado de las imágenes: consiste en intentar compensar todo lo que puede provocar que dos imá-

genes de la misma cara sean diferentes. Esto incluye normalizar el tamaño y el contraste de la imagen. A veces también se intentan compensar los cambios de iluminación, la rotación y otras características de la imagen que pueden perjudicar los resultados del sistema. Existe una buena cantidad de investigaciones dedicadas únicamente a esta parte, debido a su complejidad.

■ extracción de características: en esta fase se extraen una serie de valores característicos de cada imagen, como pueden ser los coeficientes de algún desarrollo, la salida de un filtro, etc. Independientemente de su origen estos valores deben intentar caracterizar con la mayor exactitud cada cara (lo que se considera eficiencia) y, al mismo tiempo, deben tener capacidad de discriminación. Esto significa que los valores extraídos de las imágenes de una cara y los de las imágenes de otras cara deben formar dos grupos lo más compactos y separados posible

■ comparación de características: se comparan los valores característicos de la imagen de test (la que se quiere reconocer) con los de las imágenes de entrenamiento y se calcula una medida de semejanza. Los métodos van desde la distancia euclidea (considerando que el conjunto de valores característicos forman un vector) a otros mucho más sofisticados. La imagen de entrenamiento que más semejante sea a la de test se considerará que es de la misma persona

En una primera fase que llamaremos fase de entrenamiento, las dos primeras partes se aplican a las imágenes del conjunto de entrenamiento, para extraer las características que caracterizan a las imágenes de este conjunto. Posteriormente, en la fase de test, las imágenes del conjunto de test que queremos reconocer pasan por las dos primeras partes y la fase de comparación determina finalmente la imagen de entrenamiento que más se parece a la de test.

La parte más importante es la de extracción de características. En ella, se intenta caracterizar una cara con independencia de factores externos como la iluminación, expresión, etc. En general el objetivo es hallar una serie de características de la imagen que pueden ser desde las dimensiones de los elementos de la cara como la boca, ojos, etc. hasta los coeficientes que nos da un filtro o una transformada al aplicarlo a la imagen de la cara. Lo importante es que caractericen con gran fidelidad a cada cara y que sean invariantes a los cambios en

los factores externos que se han mencionado anteriormente.

### 3. TIPOS DE TÉCNICAS DE RECONOCIMIENTO DE CARAS

Las técnicas de reconocimiento de caras se pueden clasificar en tres tipos, en función de los métodos utilizados para la extracción de las características de la cara:

- técnicas basadas en la extracción de características geométricas de la cara como las posiciones relativas y dimensiones de ciertos elementos de la cara (cejas, ojos, nariz, boca y contorno de la barbilla principalmente)
- técnicas basadas en procedimientos de *template matching* (emparejamiento de plantillas), en las que zonas de la imagen son comparadas con zonas equivalentes de otra imagen utilizando alguna distancia (pe la distancia Euclídea) para poder calibrar su grado de semejanza o discrepancia. Los sistemas que utilizan este tipo de técnicas carecen de la fase de extracción de características, ya que utilizan directamente la información de niveles de gris (o color) de los pixels de la imagen
- técnicas basadas en la aplicación de transformadas: las características de la imagen son los coeficientes resultantes de aplicar algún procedimiento numérico o alguna transformada, como la transformada Karhunen-Loeve

A pesar de esta sencilla división, existen gran cantidad de técnicas y no resulta sencillo escoger entre ellas. Al comenzar el proyecto se definieron una serie características que debía cumplir el sistema y que limitaron las alternativas entre las que escoger. Se decidió buscar un sistema que utilizara una sola vista frontal de cada cara como conjunto de entrenamiento. En otras palabras, el sistema debía ser capaz de poder identificar a una persona teniendo sólo una imagen (una vista frontal) de ésta. Sin embargo se pueden encontrar técnicas de reconocimiento de caras que utilizan vistas laterales o perfiles, información 3D de la cara u otros tipos de información.

Tras una intensa etapa de documentación y búsqueda se decidió implementar un sistema de reconocimiento basado en la representación mediante *eigenfaces*. Una de las razones es que este tipo de sistemas tiene probablemente el mejor compromiso entre complejidad, rapidez de ejecución y resultados.

### 3. REPRESENTACIÓN DE CARAS MEDIANTE EIGENFACES

#### 3.1. Introducción

El origen de los sistemas de FR que utilizan *eigenfaces* se remonta a unos trabajos de Sirovich y Kirby ([1]). El objetivo de Sirovich y Kirby era caracte-

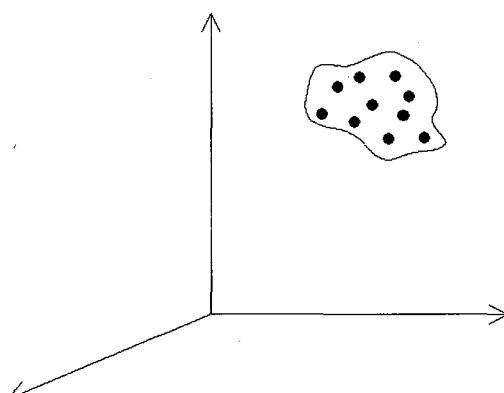
rizar un conjunto de caras con un mínimo número de parámetros. Para ello, se utilizaba el análisis de componentes principales o PCA (*principal components analysis*) para generar unas imágenes semejantes a caras llamadas *eigenpictures*. Las imágenes de caras se caracterizaban como una combinación lineal de *eigenpictures*. Fueron Turk y Pentland ([2]) quienes utilizaron esta representación de la imagen, mucho más compacta que los valores de los pixels de la imagen, para implementar un sistema de FR, además de rebautizar las *eigenpictures* como *eigenfaces*.

El primer paso para comprender los sistemas de reconocimiento basados en *eigenfaces* es conocer la filosofía que se oculta detrás del PCA, que se utiliza para generarlas, y que permite un mejor entendimiento de la representación de las imágenes de caras mediante *eigenfaces*.

#### 3.2. Motivación

Consideremos el conjunto de todas las imágenes posibles cuyas dimensiones son  $w$  pixels de ancho por  $h$  pixels de alto. El número de componentes de estas imágenes es  $w \cdot h$ . Si concatenamos las filas de estas imágenes podemos considerarlas como vectores de dimensión  $w \cdot h$ . Utilizando esta representación vectorial, el conjunto de imágenes forman un espacio vectorial de dimensión  $w \cdot h$ . Este espacio se puede generar a partir de la base canónica  $(1, 0, 0, \dots, 0)$ ,  $(0, 1, 0, 0, \dots, 0)$ , ...,  $(0, 0, \dots, 0, 1)$ .

Consideremos ahora el conjunto de las imágenes de caras. En relación al resto de todas las posibles imágenes son todas relativamente semejantes: todas tienen los mismos elementos, situados de forma semejante, con textura parecida, etc. Además, los pixeles están altamente correlados con sus vecinos. Es razonable, por lo tanto, suponer que el conjunto de las imágenes de caras ocupa sólo una pequeña región de este espacio. Una consecuencia de esto es que la representación de las



**Figura 1.** Ejemplo simplificado del espacio de las imágenes (en este caso de dimensión 3) y región ocupada por las imágenes de caras.

imágenes de caras utilizando la base del espacio de las imágenes es claramente ineficiente.

Además, debido a que todas las imágenes de caras son relativamente semejantes, podemos suponer que las imágenes de caras están en una zona reducida del espacio de las imágenes (llámémoslo el espacio de las caras). En ese caso, podríamos intentar hallar una base de vectores cuyo número sería mucho menor que el de una base del espacio de las imágenes. A pesar de que no existe ninguna evidencia al respecto, la rapidez con la que el cerebro humano es capaz de reconocer las caras sugiere que su representación en el cerebro es de una dimensionalidad baja, aunque nadie sabe cuál es ni como se maneja la información. En la Figura 1 se muestra un ejemplo simplificado en el que el espacio de las imágenes es de dimensión tres (imágenes de tres pixels). Finalmente, llegamos a la conclusión debería ser posible generar las imágenes de caras (o al menos una muy buena aproximación) con muchos menos vectores que los necesarios para representar el espacio de las imágenes completo.

El objetivo pasa a ser encontrar esta base de vectores. El PCA (análisis de componentes principales o *principal components analysis*) permite hallar una nueva base con un máximo de eficiencia. El PCA fue desarrollado por estadísticos, aunque también ha sido formulado en el campo de las redes neuronales. En este proyecto se utiliza la concepción estadística del PCA.

### 3.3. El PCA

Como ya se ha dicho, las componentes (los pixels) de las imágenes de caras están altamente correladas entre sí. El PCA se basa precisamente en las propiedades estadísticas de las imágenes y es un método óptimo para reducir el número de dimensiones necesarias para representar un conjunto de vectores (en nuestro caso, imágenes de caras).

Consideremos un vector aleatorio  $\mathbf{X}$  de dimensión  $n$ , con su correspondiente media  $\mu$

$$\mu = E\{\mathbf{X}\}$$

La matriz de covariancia se define como

$$\Sigma = E\{(X - \mu)(X - \mu)'\} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

donde  $\sigma_{ij}$  representa la covariancia entre el componente número  $i$  y el número  $j$  del vector  $\mathbf{X}$ . El análisis de componentes principales consiste en hallar los vectores propios de esta matriz y expresar  $\mathbf{X}$  en función de estos vectores. De esta manera, el vector aleatorio  $\mathbf{X}$  se puede

representar sin error mediante una combinación lineal de vectores de la forma

$$\mathbf{X} = \sum_{i=1}^n y_i \mathbf{A}_i$$

donde los vectores  $\mathbf{A}_i$  son los vectores propios de la matriz de covariancia. La cualidad más importante de esta representación es que si queremos representar  $\mathbf{X}$  con sólo  $m$  (menor que  $n$ ) componentes la mejor elección posible en términos del error que se comete son los vectores propios de la matriz de covariancia con mayores valores propios asociados.

### 3.4. Aplicación del PCA para la representación de caras: las eigenfaces

Como se acaba de ver, el PCA se reduce básicamente a encontrar los vectores propios de la matriz de covariancia  $\Sigma_x$  del vector al que aplicamos el PCA. Apliquemos toda esta teoría a nuestro problema particular de reconocimiento de caras: nuestro objetivo es caracterizar un conjunto de caras (que serán las del conjunto de entrenamiento) en un espacio de menor dimensionalidad.

Tal y como se ha expuesto anteriormente, consideremos las imágenes de caras de este conjunto (de anchura  $w$  y altura  $h$ ) como realizaciones del vector aleatorio  $\mathbf{X}$  de dimensión  $w \times h$ , con su correspondiente vector media  $\mu$ . A este vector, al que llamaremos vector cara, deseamos aplicarle el PCA.

La media de este vector cara no la podemos obtener al no conocer la función de densidad de probabilidad, pero la podemos estimar a partir las imágenes que tenemos:

$$\mu \approx \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i$$

La matriz de covariancia tampoco la conocemos pero se también se puede estimar:

$$\Sigma_x \approx \frac{1}{N} \sum_{i=1}^N (\mathbf{X}_i - \mu)(\mathbf{X}_i - \mu)'$$

Llegados a este punto, en teoría sólo queda el cálculo de los vectores propios de la matriz de covariancia, para lo que existen varios algoritmos automáticos fácilmente implementables. Pero la implementación de este algoritmo plantea un problema: la cantidad memoria necesaria para almacenar la matriz es demasiado grande. Si los vectores cara de la base de entrenamiento son de dimensión  $w \times h$ , las dimensiones de la matriz de covariancia estimada  $\Sigma_x$  serán de  $w \times h$  de ancho y  $w \times h$  de alto. Por ejemplo, si utilizamos imágenes de entrenamiento con unas dimensiones contenidas de  $w=128$  y  $h=128$ , el número de elementos de la matriz de covariancia



sería de  $128^4$ , o sea más de 268 millones, lo que está fuera del alcance de las memorias de los ordenadores actuales.

La solución es aplicar una SVD (descomposición en valores singulares o *singular value decomposition*) a la matriz  $\mathbf{M}_x$ . Esta operación descompone una matriz en un producto de tres matrices

$$\mathbf{M}_x = \mathbf{U} \mathbf{A}^{1/2} \mathbf{V}^T$$

donde si  $\mathbf{M}_x$  es una matriz de  $w \times h$  filas y  $N$  columnas con  $w \times h$  mayor que  $N$  (lo que siempre será nuestro caso) la matriz  $\mathbf{U}$  es de las mismas dimensiones y las matrices  $\mathbf{V}$  y  $\mathbf{A}^{1/2}$  son cuadradas de dimensiones  $N$  por  $N$ . Además, la matriz  $\mathbf{A}^{1/2}$  es diagonal.

La propiedad que nos es útil de la SVD es que las  $N$  columnas de la matriz  $\mathbf{U}$  son  $N$  vectores propios de la matriz  $\mathbf{X}\mathbf{X}^T$ , y los elementos de la matriz  $\mathbf{A}^{1/2}$  son las raíces cuadradas de los valores propios correspondientes a estos vectores propios. Si recordamos la estimación de la matriz  $\Sigma_x$  vemos que esta no era exactamente  $\mathbf{X}\mathbf{X}^T$ , sino que además se dividía este producto por  $N$ . Pero esto no importa ya que los vectores y valores propios de una matriz y los de la misma matriz dividida por un escalar son iguales, salvo por el hecho de que los valores propios también quedan divididos por el mismo número.

Este método está al alcance de un ordenador, ya que la matriz  $\mathbf{M}_x$  es de dimensiones  $w \times h$  por  $N$ , y  $N$  (que es el número de caras del conjunto de entrenamiento) suele estar alrededor de un centenar, lo que la convierte en una matriz manejable por un ordenador. Sin embargo este método no calcula todos los vectores propios de  $\mathbf{S}_x$  ya que esta matriz, al ser real y simétrica, tiene siempre  $w \times h$  vectores propios (en la práctica muchos más que  $N$  que es el número de imágenes del conjunto de entrenamiento). Pero se puede demostrar que los  $N$  vectores propios conseguidos con la SVD son los únicos vectores propios de  $\Sigma_x$  que pueden tener un valor propio distinto de cero. Desde el punto de vista teórico del PCA, esto quiere decir que el resto de vectores propios (los que siempre tienen un valor propio igual a cero) no capturan ninguna información del vector  $\mathbf{X}$  al que estamos aplicando el PCA, y por lo tanto los eliminaríamos del análisis tal y como se ha explicado en el punto anterior.

Volviendo a nuestro problema de representación, todo esto significa que los  $N$  vectores propios que genera la SVD son capaces de caracterizar a los  $N$  vectores correspondientes al conjunto de imágenes de caras de entrenamiento, y el resto de vectores propios sólo nos podrían servir para abarcar el resto del espacio original de  $\mathbf{X}$  (el espacio de las imágenes de  $w$  por  $h$  pixels), pero no nos servirían para representar las caras. De entre estos  $N$  vectores propios el número de ellos que contienen información útil (valor propio mayor de cero) es igual al número de vectores linealmente independientes del conjunto de entrenamiento menos uno (ya que al restar la media se reduce siempre en uno el número de vectores linealmente independientes).

De esta manera obtenemos los deseados vectores propios que definen nuestro nuevo espacio, el espacio de las caras. Estos vectores tienen realmente apariencia de caras, y por ello han sido llamados primero *eigenpictures* (imágenes propias) y posteriormente *eigenfaces* (caras propias). En la figura 2 se muestran las 5 primeras *eigenfaces* generadas con un conjunto de 57 imágenes de entrenamiento.



**Figura 2.** Cara media y 5 primeras eigenfaces de un total de 56 generadas a partir de 57 imágenes de caras sin barba ni gafas. La posición de los ojos y la boca está normalizada y los componentes vectores propios han sido remapeados en el rango [0,255] para poder visualizarlos

### 3.5. Reducción de la dimensionalidad

Llegados a este punto, la situación es la siguiente: tenemos un conjunto de  $N$  imágenes de caras  $\mathbf{X}_i$  (que utilizamos como conjunto de entrenamiento) que queremos representar en un espacio de menor dimensión. Para ello hemos considerado estas imágenes como realizaciones del proceso aleatorio vectorial "imágenes de caras" y hemos aplicado el PCA a este vector. Como resultado, obtenemos  $N$  vectores propios (*eigenfaces*) ortonormales y con  $N-1$  de ellos se pueden generar sin error las  $N$  caras del conjunto de entrenamiento. En la figura 3 se muestra de una manera gráfica parte de la representación de una de las caras utilizadas para generar las *eigenfaces* de la figura 2.

Para hallar las coordenadas  $\mathbf{Y}_i$  de una cara  $\mathbf{X}_i$  sobre este nuevo espacio no tenemos más que proyectar  $\mathbf{X}_i$

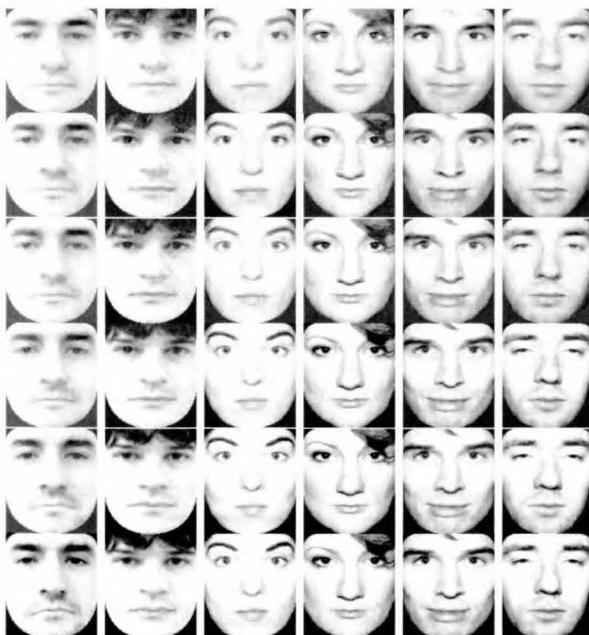
$$\begin{array}{c} \text{[face image]} \\ - \end{array} = -4.9e3 * \begin{array}{c} \text{[eigenface 1]} \end{array} + 3.0e3 * \begin{array}{c} \text{[eigenface 2]} \end{array} + \dots$$

**Figura 3.** Representación de una cara a partir de las eigenfaces: a la cara se la resta la cara media y esta diferencia se puede expresar como una combinación lineal de las eigenfaces. Para que la representación sea exacta la imagen debe estar incluida en el conjunto al que se la aplica el PCA (el conjunto de entrenamiento)

sobre las *eigenfaces*. Para hallar la proyección de  $\mathbf{X}_i$  sobre cada *eigenface* bastará con calcular el producto escalar, ya que las *eigenfaces* son ortonormales. Por lo tanto, si las columnas de la matriz  $\mathbf{A}$  son las *eigenfaces*, hallaremos las nuevas coordenadas en el espacio de las caras de la siguiente manera:

$$Y_i = P^T X_i$$

Al haber  $N-1$  *eigenfaces* la dimensión de los vectores  $\mathbf{Y}_i$  es  $N-1$  (recordemos que  $N$  es el número de imágenes de entrenamiento). Esto supone una fuerte reducción de dimensionalidad frente al espacio de las imágenes ( $w \times h$ ) pero esto no es ninguna sorpresa ya que como ya se ha comentado obviamente para representar a  $N$  vectores a los que se ha restado su media basta con  $N-1$  vectores. La verdadera reducción de dimensionalidad de la representación mediante *eigenfaces* consiste en utilizar sólo las *eigenfaces* con mayor valor propio asociado. Recordemos que el valor propio asociado a una *eigenface* es igual a la variancia de la proyección de las caras sobre esa *eigenface*. De esta manera se puede reducir notablemente el número de dimensiones de la representación sin perder apenas calidad en la representación de las caras. En este aspecto, las *eigenfaces* consiguen la máxima eficiencia que se puede conseguir ya que las primeras *eigenfaces* (las de mayor valor propio asociado) consiguen capturar la información más importante de las caras, la que explica la máxima variancia, mientras que las últimas apenas aportan información (sus valores propios son muy bajos). Dicho de otra manera: la mejor representación posible de un conjunto de  $N$  imágenes de caras utilizando sólo  $M$  dimensiones (con  $M < N$ ) se consigue proyectándolas



**Figura 4.** Representaciones de 6 imágenes de caras utilizando las primeras 8, 16, 24, 32, 40, 48 y 56 *eigenfaces* de un total de 56 *eigenfaces*, generados a partir de un PCA de 57 imágenes de caras y ordenadas de mayor a menor valor propio. La reconstrucción incluye la suma de la cara media para poder apreciar mejor la calidad



**Figura 5.** Imágenes de personas no incluidas en el conjunto de entrenamiento representadas con un conjunto de 56 *eigenfaces*. El error es considerable debido a que el número de *eigenfaces* es demasiado bajo

sobre las  $M$  *eigenfaces* con mayor valor propio. En la figura 5 se puede ver la evolución de la calidad en la representación de algunas caras del conjunto de entrenamiento utilizado para generar las *eigenfaces* de la figura 2. Como se puede apreciar, con un número muy bajo de *eigenfaces* se consigue una representación con una notable calidad, lo que demuestra que las primeras *eigenfaces* consiguen capturar la mayor parte de la información de las caras.

Como se ha dicho anteriormente, si las *eigenfaces* se generan a partir de un conjunto de entrenamiento suficientemente grande de imágenes entonces son capaces de representar con fidelidad imágenes de caras de personas que no están en el conjunto de entrenamiento. Para ello el tamaño del conjunto de entrenamiento debe ser como mínimo de unas 100 imágenes. En el presente proyecto no se disponía de tal número y por lo tanto no se consigue una buena representación. En la figura 5 se muestran algunos ejemplos de reconstrucción de caras de personas no incluidas en el conjunto de entrenamiento.

Sin embargo, lo que siempre se consigue es una buena representación de imágenes de caras que tienen alguna otra imagen (de la misma cara o persona) en el conjunto de entrenamiento, aunque estas tengan variaciones en la expresión respecto a la imagen del conjunto de entrenamiento. En estos casos, aunque la imagen representada con las *eigenfaces* tenga un error considerable la identidad de la cara siempre se preserva con gran fidelidad, lo que es especialmente importante de cara a su aplicación al reconocimiento de caras. En la figura 6 vemos unos ejemplos que ilustran este hecho.

La figura 4 da una idea visual de la información capturada por las primeras *eigenfaces* pero ¿hasta qué punto consiguen concentrar esta información en un bajo número de *eigenfaces*? En la figura 7 se muestra la gráfica de la evolución de los valores propios (iguales a la variancia de la proyección de las caras sobre la *eigenface* asociada), ordenados de mayor a menor. La gráfica sigue una evolución semejante a una exponencial,



**Figura 6.** Primera fila: imágenes pertenecientes al conjunto de entrenamiento. Segunda fila: imágenes de las mismas personas con cambios de expresión. Tercera fila: imágenes de la segunda fila representadas a partir de las eigenfaces generadas con las imágenes de la primera

gráfica sigue una evolución semejante a una exponencial, lo que demuestra la gran concentración de variancia que se produce en las primeras *eigenfaces*. Esto se refleja claramente en la calidad obtenida en la representación. Para expresar esta calidad numéricamente debemos hallar una medida del error cometido al representar una cara. Si la representación de una cara  $\mathbf{X}$  con  $M$  *eigenfaces* (de un total de  $N$ ) es

$$\tilde{\mathbf{X}} \approx \boldsymbol{\mu} + \sum_{i=1}^M y_i \mathbf{A}_i$$

donde  $\boldsymbol{\mu}$  es el vector media  $\mathbf{A}_i$  son las *eigenfaces* ordenadas de mayor a menor valor propio y  $y_i$  la proyección de  $\mathbf{X}$  sobre  $\mathbf{A}_i$ , el error de representación se define como

$$e = \left( \frac{\|\mathbf{X} - \tilde{\mathbf{X}}\|}{\|\mathbf{X}\|} \right)^2$$

El error está normalizado y su rango de valores va de cero a uno. Sin embargo, debido a que la media es una buena estimación de todas las caras en la práctica los valores siempre son mucho más reducidos. La evolución de sus valores se ve en la figura 8 para los tres casos de representación vistos. Como se podía esperar, esta gráfica confirma que con un bajo número de *eigenfaces* se consiguen representaciones con bajos porcentajes de error.

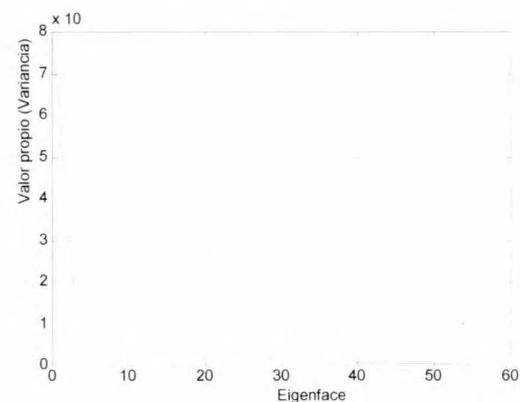
#### 4. EL SISTEMA DE RECONOCIMIENTO DE CARAS

En el capítulo anterior se ha visto la conveniencia de representar las imágenes de caras mediante *eigenfaces*. En este capítulo se verá como se aplica esta representación al reconocimiento de caras. Recordaremos el planteamiento del problema a resolver por un sistema de reconocimiento de caras: dado un conjunto de imágenes de caras que llamaremos conjunto de entrenamiento, el

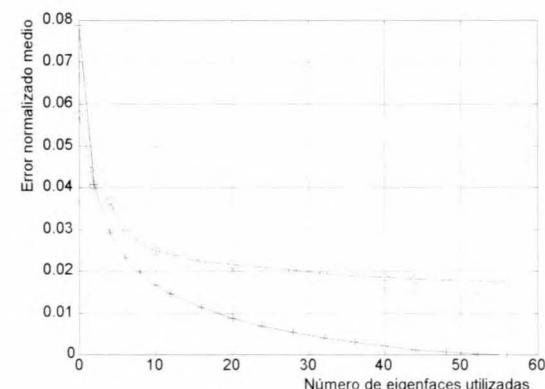
sistema debe ser capaz de emparejar una nueva imagen con la imagen del conjunto de entrenamiento que pertenece a la misma persona.

##### 4.1. Etapas de entrenamiento y de test

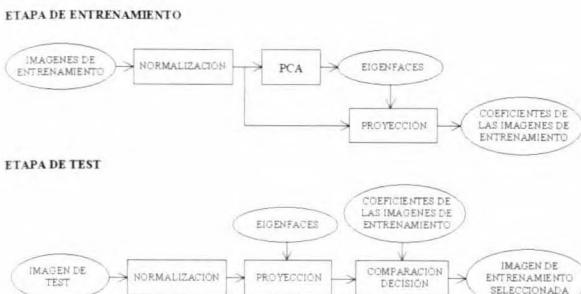
Para que el sistema esté preparado para realizar la parte de comparación es necesario disponer de las características de las imágenes de entrenamiento. Por ello es necesario realizar una vez una etapa de entrenamiento. Esta consiste en aplicar las dos primeras fases a las imágenes del conjunto de entrenamiento. A partir de entonces se puede efectuar el reconocimiento con un conjunto de imágenes de test. El esquema general se muestra en la Figura 9.



**Figura 7.** Gráfica de los valores propios asociados a las eigenfaces ordenados de mayor a menor. La disminución es muy rápida, lo que indica una gran concentración de la variancia de las imágenes en las primeras eigenfaces



**Figura 8.** Error de representación de las caras del conjunto de entrenamiento (media de las 57) en función del número de eigenfaces utilizadas para representar imágenes del conjunto de entrenamiento (+), otras imágenes de personas incluidas en el conjunto de entrenamiento con diferentes expresiones (X) y imágenes de caras “desconocidas” (O). El error decréce rápidamente con las primeras eigenfaces y la reducción de error debida al resto de eigenfaces es mínima



**Figura 9.** Esquema del sistema de reconocimiento de caras.

## 4.2. Implementación de las partes del sistema de reconocimiento

A continuación se describe la implementación concreta de las partes del sistema de reconocimiento.

### 4.2.1. Normalización o preprocessado

La fase de normalización es de una gran importancia en los sistemas que utilizan *eigenfaces*. Como se ha visto anteriormente, la potencia de las *eigenfaces* reside en su gran capacidad para caracterizar con un mínimo error una imagen de una cara. Esta capacidad siempre es muy alta con las imágenes con las que se han generado las *eigenfaces*, pero con otras imágenes de caras (pe. las que se tendrán que reconocer, que llamaremos conjunto de test) depende mucho del procesado que se le aplique a la imagen antes de proyectarla sobre las *eigenfaces*.

Nuestra implementación de esta fase intenta compensar las diferencias en los siguientes aspectos:

- tamaño: se normaliza la distancia entre los ojos y entre los ojos y la boca. En otras palabras: se normaliza la altura y la anchura de la cara. Como se ha comentado antes, esto supone una pequeña distorsión en la forma de la cara, pero los resultados demuestran una mejora. La localización de los ojos y la boca se realiza manualmente al no ser un objetivo del proyecto, pero existen métodos automáticos muy eficientes ([3])
- rotación en el plano de la imagen: utilizando las posiciones de los dos ojos, se efectúa una rotación de la imagen de manera que los dos ojos queden en horizontal
- zona de la cara utilizada: la cara se recorta de manera que no aparezcan zonas del fondo de la imagen ni zonas del pelo, cuyo aspecto es muy variable y que perjudicaría la robustez del sistema
- contraste y nivel de iluminación: se expande el rango de niveles de gris de la imagen al máximo. Al estar representadas en 8 bits, esto hace que el valor mínimo de los pixels pase a cero y el máximo a 255. Esta etapa sólo debería realizarse después de la de recorte de la cara, de lo contrario el pelo y el fondo podrían distorsionar los resultados

La normalización del tamaño y de la rotación se consigue realizando un nuevo muestreo sobre la imagen en el que las posiciones de los ojos y la boca sean fijas. El proceso se puede ver en la figura 10.



**Figura 10.** El proceso de remuestreo.

### 4.2.2. Cálculo de las eigenfaces

En esta fase se realiza el PCA con el que se extraen las eigenfaces, mediante los siguientes pasos:

- se calcula de la cara media y se resta de todas las caras normalizadas
- se forma la matriz  $\mathbf{M}_x$  cuyas columnas son las imágenes del conjunto de entrenamiento normalizadas
- se realiza la SVD de la matriz  $\mathbf{M}_x$ , dando como resultado tres matrices, una de ellas conteniendo los vectores propios (las *eigenfaces*) y otra las raíces cuadradas de los valores propios
- se guardan las *eigenfaces* así como la proyección de las imágenes de entrenamiento sobre ellas (coeficientes de las imágenes de entrenamiento)

### 4.2.3. Proyección sobre las eigenfaces

Las *eigenfaces* forman un conjunto ortonormal. Por lo tanto, para hallar la proyección de una imagen sobre ellas basta con realizar el producto escalar de la imagen sobre cada una de las *eigenfaces*.

### 4.2.4. Fase de comparación/decisión

El objetivo es determinar que imagen del conjunto de entrenamiento es más parecida a la imagen de test, a partir de sus representaciones mediante las eigenfaces (sus proyecciones). Para ello se compara el vector  $\mathbf{y}_{TEST}$  formado por las proyecciones de la imagen de test sobre las eigenfaces con cada uno de los vectores  $\mathbf{y}_{ENTI}$ . El criterio que se utiliza es el de la menor distancia euclídea, es decir, menor  $|\mathbf{y}_{TEST} - \mathbf{y}_{ENTI}|$



**Figura 11.** Ejemplo de la base de datos de Berna



## 5. RESULTADOS

### 5.1. Imágenes utilizadas

Para probar el sistema se utilizaron dos conjuntos de imágenes. La base de datos de Berna consiste en imágenes de 30 personas (todas de sexo masculino). De cada persona se seleccionaron una vista frontal, otra vista frontal con ligeros cambios de expresión y una vista lateral con un ligera rotación en profundidad. Varias personas llevan gafas y algunas barba. En la figura 11 vemos un ejemplo.

La base de datos de Stirling consiste en imágenes de 17 mujeres y 15 hombres, sin gafas ni barba. De cada persona se seleccionaron tres vistas equivalentes a las anteriores pero con un grado de dificultad más elevado que en la anterior base de datos: en la segunda vista frontal los cambios de expresión son más acusados y en la vista lateral la rotación hace que parte de la cara no se vea.



**Figura 12.** Ejemplo de imágenes de la base de datos de Stirling

### 5.2. Pruebas

Se realizaron dos pruebas. En ambas se utilizaron como conjunto de entrenamiento las 62 vistas frontales de personas diferentes.

En la primera prueba se utilizaron las 62 segundas vistas frontales como conjunto de test, mientras que en la segunda se utilizaron las vistas laterales. Los resultados se muestran en la figura 13. También se indican los resultados parciales con las imágenes de cada base de datos (con el conjunto de entrenamiento total de 62 imágenes).

Como se puede ver, los cambios de expresión apenas afectan a la eficacia del sistema. Sin embargo, las rotaciones en profundidad producen tal cambio en la imagen que la representación mediante *eigenfaces* deja

PRUEBA	CONJUNTO DE ENTREN.	CONJUNTO DE TEST	TOTAL	BERNA	STIRLING
Prueba 1	62 frontales ( 1 por pers.)	62 frontales ( 1 por pers.)	60/62	30/30	30/32
Prueba 2	62 frontales ( 1 por pers.)	62 laterales ( 1 por pers.)	29/62	25/30	4/32

**Figura 13.** Resultados del sistema

de ser adecuada y se producen errores. Particularmente con las imágenes de la base de datos de Stirling se producen muy malos resultados debido a que la rotación es muy pronunciada

## 6. CONCLUSIONES

El sistema que se ha presentado obtiene unos buenos resultados, pero con ciertas limitaciones. En concreto las rotaciones en profundidad producen cambios que la representación mediante *eigenfaces* no puede asimilar. Este problema es común a casi todos los sistemas de reconocimiento, y de momento no se ha conseguido resolver utilizando una sola imagen por persona en el conjunto de entrenamiento. Las únicas soluciones son utilizar un múltiples vistas de cada persona en el conjunto de entrenamiento, pero estas vistas normalmente no están disponibles y se necesitan muchas para poder abarcar todas las rotaciones posibles, necesitándose además muchas más memoria y cálculos independientemente del sistema utilizado.

Otro problema que no se ha tratado en este artículo pero que también produce dificultades son los cambios de iluminación, en concreto de la dirección de iluminación. En algunos trabajos de nuevo se recurre a tener múltiples vistas de cada persona (una para cada condición de iluminación diferente) para poder saber compensar esos cambios, pero los problemas siguen siendo los mismos, siendo el principal que normalmente sólo se va a disponer de una vista (frontal) para el conjunto de entrenamiento.

Como conclusión se puede decir que aunque están en continuo avance y los resultados son aceptables, los sistemas de reconocimiento de caras todavía deben superar algunas limitaciones.

## 7. BIBLIOGRAFÍA

- [1] L. SIROVICH AND M. KIRBY, "Low-dimensional procedure for the characterization of faces", J. Opt. Soc. Am. A, vol. 4, num. 3, 1987, pp. 519-524.
- [2] M. TURK AND A. PENTLAND, "Eigenfaces for recognition", J. Cognitive Neuroscience, vol. 3, num. 1, 1991.
- [3] A. PENTLAND, B. MOGHADDAM AND T. STARNER, "View-based and modular eigenspaces for face recognition", IEE Conf. On Computer Vision & Pattern Recognition, Seattle, WA, July 1994.



## APOLO 13, ¡TENEMOS UN PROBLEMA!

*Manuel Montes Palacio*

*Editor del boletín Noticias del Espacio*

<http://www.amazings.com/notesp/index.html>

e-mail: montes@ctv.es

Pocas veces en la historia de la Ciencia y en particular de la Astronáutica un fracaso ha despertado tanta expectación. Pocas veces un viaje espacial nos ha abierto tanto los ojos, haciéndonos ver cuán peligrosa es la aventura tripulada.

En efecto, el vuelo del Apolo-13, la tercera excursión americana sobre la Luna, acabó convirtiéndose en la más tensa y agonizante misión que el joven programa espacial había vivido hasta entonces. Un viaje que, merced a su final feliz, fue pronto olvidado entre la euforia de los emocionantes alunizajes, pero que aún así había contado con todos los elementos que pueden convertir a un arriesgado suceso en una historia periodística de gran éxito.

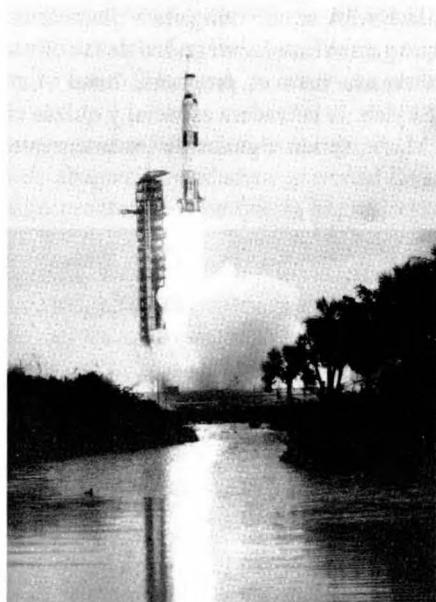
Lo ocurrido con el Apolo-13 pudo haber cambiado radicalmente el futuro de la exploración espacial. La muerte de su tripulación hubiera obligado a la NASA, la agencia americana, a realizar un completo replanteamiento de su programa, tal y como tuvo que hacerse tras el trágico accidente del Apolo-1, en enero de 1967. Los soviéticos, por su parte, luchando para superar a sus rivales en la carrera lunar, estaban desarrollando su propio proyecto de alunizaje tripulado, una gigantesca empresa formada por el mastodóntico cohete N-1 y la familia de cosmonaves L-3. Diversos problemas técnicos y de organización hundieron las expectativas de la URSS en este terreno, pero si América hubiera tenido que frenar sus actividades sobre nuestro satélite, la nación comunista se hubiera de pronto encontrado con la oportunidad de recuperar el terreno perdido, pisar Selene y prepararse para el ya propuesto viaje a Marte.

Nadie duda ya que la carrera lunar nació profundamente enraizada en diversas consideraciones políticas, económicas y militares. Un cambio de este calibre en el resultado final de la contienda hubiera producido también grandes transformaciones en dichas áreas, con imprevisibles resultados.

¿Por qué, pues, ante la trascendencia de esta misión, su historia ha permanecido relativamente desconocida para el gran público? Es difícil contestar a esta pregunta, pero parece claro que uno de los motivos podría estar relacionado con la desenfrenada cadencia de espectaculares resultados en los que incurrió, antes y después, el programa americano. El Apolo fue un proyecto de gran

éxito, y en este contexto, el Apolo-13 (como antes el Apolo-1) desapareció ante el enorme triunfo global del sistema.

Es evidente que muchas personas recuerdan la emoción de las noticias que, poco a poco, desgranaban la odisea del Apolo-13. Pero su recuerdo, volátil, es similar al de otros muchos acontecimientos que, a pesar de su dramatismo, han acabado bien. Sólo los fanáticos del vuelo espacial tienen presente, con todo detalle, lo que ocurrió durante aquella semana.



*Lanzamiento del Apolo 13, el 11 de abril de 1970, a las 13:13 horas (hora de Houston)*

Veinticinco años después, una película americana realizada con un gran presupuesto y la participación de actores de primera fila, puso de nuevo de actualidad aquello que la mayoría ya había olvidado. Con un tratamiento claramente cinematográfico, el film puso de manifiesto lo que los interesados por la astronáutica ya sabían: que el viaje del Apolo-13 tuvo todos los ingredientes que han caracterizado desde siempre a las buenas historias, a las aventuras que saben captar la atención del espectador, y que su plasmación en el celuloide era sólo una cuestión de tiempo. Como dijo el astronauta Jim Lovell, la película está bien hecha, tiene altas dosis de realismo y lo mejor: lo que sucede en ella, ¡es verdad!



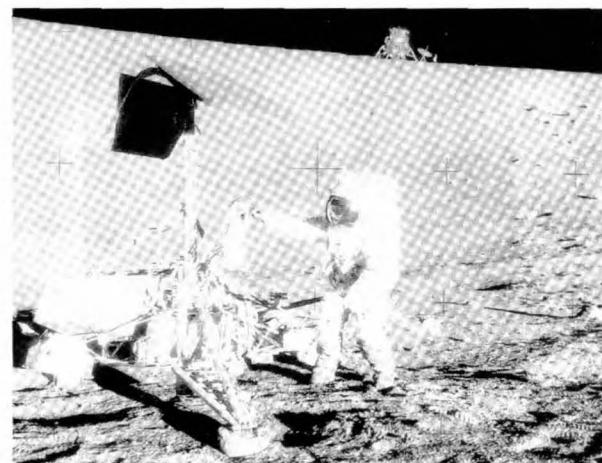
Sin embargo, lo que a nosotros más nos interesa es el relato de lo sucedido desde la objetividad histórica, no desde el dramatismo y el espectáculo que sin duda impregnán el metraje de la película, un producto, al fin y al cabo, comercial.

## UN VIAJE DE IDA Y VUELTA

Después de la apoteosis del Apolo-11, el gran objetivo del programa se había visto cumplido con creces. El mandato del fallecido John F. Kennedy no sólo se había hecho realidad en el plazo previsto sino que además América había vencido a la URSS en la carrera lunar.

Fue entonces cuando, como una olla a presión que ha alcanzado su máximo punto de ebullición y que por fin empieza a dejar escapar el vapor, el proyecto Apolo perdió su consideración de prioridad nacional, convirtiéndose sólo en el más caro juguete de la NASA.

La agencia, conocedora de este hecho, fue advertida del próximo cese de las misiones tripuladas a la Luna incluso antes del vuelo del Apolo-11. Ante estas expectativas, la NASA se vio obligada a plantear un plan de futuro que garantizase la integridad de la colossal infraestructura creada para el programa lunar. La estación orbital Skylab, la lanzadera espacial y quizás un posible viaje a Marte, serían algunos de los integrantes de este plan.



Un tripulante del Apolo 12 examina la sonda Surveyor 3

Por tanto, los vuelos posteriores al Apolo-11 no deberían sólo repetir el éxito de su histórico antecesor, sino también, en calidad de su probable categoría de misiones únicas, aglutinar todos los objetivos que fuera posible discernir en tan escaso margen de tiempo.

Así lo hizo el Apolo-12, en noviembre de 1969, posándose en el Océano de las Tempestades, a sólo 183 metros de distancia de donde descansaba la sonda automática Surveyor-3. El Apolo-13, por su parte, debía ser el primero de tres viajes a la Luna dedicados en su totalidad a la investigación geológica de la superficie.

Para ello debía aterrizar en Fra Mauro, una altiplanicie especialmente interesante donde los astronautas desplegarían el ALSEP-3, un paquete compacto de experimentos que efectuaría gran cantidad de mediciones de todo tipo.

La tripulación debía estar compuesta por James A. Lowell (comandante), Thomas K. Mattingly (piloto del Módulo de Mando) y Frederick W. Haise (piloto del Módulo Lunar). El despegue, previsto para el 12 de marzo de 1970, fue retrasado hasta el 11 de abril (la próxima oportunidad o ventana de lanzamiento lunar) para dar tiempo a planear la misión con mayor detalle.

*Pocas veces en la historia de la Ciencia y en particular de la Astronáutica un fracaso ha despertado tanta expectación*

---

Por desgracia, un problema inesperado complicó las cosas apenas una semana antes de la partida. Los astronautas, de algún modo, quedaron expuestos a una enfermedad infecciosa (una variedad del sarampión), concluyéndose después de varias pruebas que Mattingly tenía posibilidades de contraerla durante el viaje. La NASA, que no quería retrasar la misión otro mes, decidió sustituirle por su astronauta de reserva, John L. Swigert. Debió ser ésta una sorpresa considerable para Swigert, que tuvo que enfrentarse en sólo dos días a una intensa batería de ensayos y conferencias para ponerse al día en los últimos detalles.

Las operaciones técnicas previas al lanzamiento del cohete y la nave espacial se llevaron a cabo sin mayores problemas, aunque una de ellas, como veremos, marcaría de forma decisiva el futuro de la misión.

Superadas todas las dificultades, el despegue se produjo puntualmente el 11 de abril de 1970, desde el Centro Espacial Kennedy, en Florida. No todo fue bien durante la ascensión a través de la atmósfera, sin embargo: el motor J-2 central de la segunda fase del cohete Saturno-V se apagó dos minutos antes de lo previsto. Para compensar la pérdida de empuje, los otros cuatro motores del escalón tuvieron que funcionar durante más tiempo. Lo mismo ocurriría con el único motor J-2 de la tercera etapa. Como resultado, el Apolo-13, aún unido a esta última, fue colocado en órbita alrededor de la Tierra con algunos segundos de retraso.

La enorme flexibilidad del sistema había permitido superar los obstáculos iniciales y prepararse para emprender el viaje a la Luna. Así, la tripulación ordenó el encendido definitivo de la tercera etapa S-IVB, operación

que les colocaría en una trayectoria de transferencia hacia nuestro satélite. La ruta, denominada de retorno-libre, permitiría rodear la Luna y volver a la Tierra sin ninguna intervención por parte de los astronautas en caso de que los sistemas propulsivos fallaran. Poco después, la nave Apolo-13, formada por el cónico Módulo de Mando y el cilíndrico Módulo de Servicio, se separó del cohete y dio un giro de 180 grados para extraer del interior de la zona de carga al Módulo Lunar. Ya unido a él, el conjunto volvió a girar sobre sí mismo, colocándose en configuración de crucero.

Para evitar interferir en la misión, la etapa S-IVB evacuó todo el combustible residual, situándose en una



De arriba a abajo: Jim Lovell, comandante; Fred Haise, piloto del módulo lunar; Jack Swigert, piloto suplente del módulo de mando.

ruta de impacto contra la Luna. Dos días después, el motor principal del Módulo de Servicio funcionó durante apenas 3 segundos, situando al Apolo-13 en una trayectoria de retorno híbrido, es decir, de retorno no libre. En caso de tener que abortar el alunizaje, los astronautas deberían usar sus motores, pero ésta era la única opción disponible si lo que se quería era entrar en órbita alrededor de Selene.

Hasta ese momento, los puntos del plan de vuelo se habían cumplido con tanta escrupulosidad que la opinión pública empezó a perder interés. No había emoción, se sabía lo que iba a ocurrir en cada instante durante las próximas jornadas.

## LA EXPLOSIÓN

Fred Haise y John "Jack" Swigert se encontraban en el interior del Módulo Lunar "Aquarius", realizando diversas comprobaciones rutinarias, cuando un sonido apagado inundó la nave. Eran las 04:08 de la mañana, hora de Florida, del 14 de abril.

Lowell, en el Módulo de Mando, informó de inmediato al centro de control en Houston, Texas, pronunciando su famosa frase: "Houston, ¡tenemos un problema!".

Aunque ni unos ni otros fueron capaces en primera instancia de abarcar la verdadera gravedad de la situación, resultó pronto evidente que la explosión se había producido en el Módulo de Servicio, donde se almacenaban los sistemas de propulsión, producción eléctrica y provisión de oxígeno. Los astronautas, desde su posición, eran incapaces de observar visualmente lo que estaba pasando debido a la arquitectura de la nave, pero con el transcurrir de los minutos, los instrumentos indicaron las consecuencias del misterioso estallido.

Tras él, efectivamente, se produjo un fugaz aumento de la tensión eléctrica, seguido por el cese del funcionamiento de una de las células de combustible que producían la electricidad. Al mismo tiempo, y debido a ello, los índices de energía y presión del oxígeno empezaron a descender en la cabina del Módulo de Mando. Posteriormente, observaciones visuales de la tripulación confirmaron un escape de oxígeno líquido, así como la eyeción de varios fragmentos pertenecientes a la zona afectada, que pronto se alejaron de la nave.

El hecho de que los sistemas de control ambiental quedaran muy afectados por lo sucedido, obligó al centro de control a ordenar a los astronautas la activación de sus homólogos en el Módulo Lunar. Mientras, los cálculos indicaron la magra existencia de apenas 15 minutos de energía eléctrica y oxígeno para alimentar al Módulo de Mando.

Dos horas después de la explosión, el Módulo Lunar se había convertido ya en una especie de balsa de salva-

mento improvisada. La nave abandonó su estado de hibernación para pasar a proporcionar el oxígeno que necesitaría la tripulación, así como una presión y temperatura aceptables.



Vista del control de misión, durante la transmisión televisada del 13 de abril de 1970, momentos antes de la explosión.

Al principio, un recuento de los recursos disponibles contabilizó sólo 38 horas de energía, oxígeno y agua (menos de la mitad del tiempo necesario para volver a la Tierra). Posteriores estimaciones, que tenían en cuenta reducir la temperatura al mínimo soportable y la desactivación de todos los sistemas no esenciales, incrementaron esta cifra hasta hacerla compatible con el objetivo último: la vuelta a casa.

Por supuesto, la dirección del proyecto anunció la cancelación del alunizaje, otorgando máxima prioridad al regreso de los tres astronautas. Existían para ello diversos problemas, no obstante. La pérdida de control en el Módulo de Mando elevaba ciertas dudas sobre la capacidad de la nave de efectuar una reentrada atmosférica segura. Asimismo, la inoperatividad de las células de combustible implicaba que la tripulación tendría muy poca agua, un subproducto de su funcionamiento durante la producción de electricidad a partir de la unión química de oxígeno e hidrógeno líquidos.

La temperatura era también baja. La falta de potencia impedía calentar la atmósfera artificial, lo que a su vez provocaba la aparición de humedad sobre los instrumentos. Al mismo tiempo, la no retirada del dióxido de carbono producido por la respiración de los astronautas amenazaba con asfixiarlos.

Tras numerosas deliberaciones, el centro de control comunicó al Apolo-13 un plan improvisado que debía situarlos de regreso a la Tierra. Sin saber si el motor del Módulo de Servicio estaba dañado, la NASA prefirió usar el de la etapa de descenso del Módulo Lunar, efectuando con él una breve corrección de curso. Ésta devolvería al conjunto a una trayectoria de regreso libre. Sin la potencia necesaria para frenar y volver directamente a su punto de partida, el Apolo debía rodear primero la Luna y emprender después el retorno, una maniobra que duraría al menos cuatro días. Éste era el tiempo que los astronautas debían sobrevivir en esta situación si querían

alcanzar nuestro planeta con posibilidades de ser rescatados.

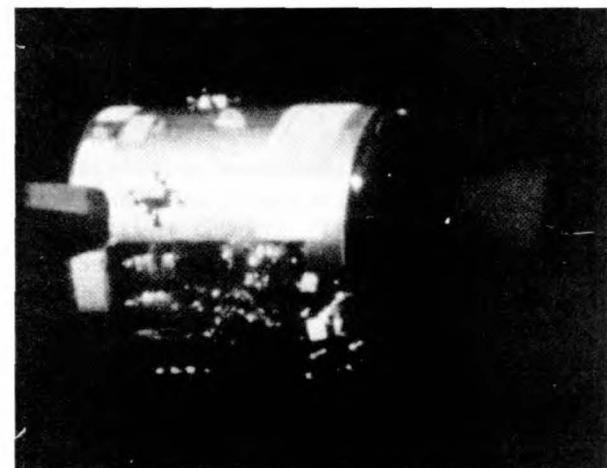
El encendido del motor del Módulo Lunar, diseñado en su día para posar a dos hombres sobre la Luna, se desarrolló según lo convenido. El Apolo-13 se vio entonces en una trayectoria que lo llevó a unos 254 km de distancia de la cara oculta. El sobrevuelo, sin embargo, se produciría sin la asistencia de la Tierra. Durante la media hora en la que permanecieron tras la mole rocosa de Selene, las señales de radio no podían llegar al Apolo-13.

*Un recuento de los recursos disponibles contabilizó sólo 38 horas de energía, oxígeno y agua*

---

A la 01:51 del 15 de abril, las antenas de la NASA volvían a entrar en contacto con el tren espacial, precisamente a la hora y en la posición esperados. Los astronautas informaron de la toma de fotografías de la superficie, el único resultado científico que podría salvarse de la fallida misión.

Un poco más tarde, a las 03:41, el motor del Aquarius entró de nuevo en acción. Funcionando durante 4 minutos y 21 segundos a menos de la mitad de su potencia, situó a la nave en una ruta tranterrestre. El resto del viaje consistiría en mantenerse en las mejores condiciones posibles y en controlar la trayectoria para asegurar la intersección con nuestra atmósfera.



Daños producidos en el módulo de servicio por la explosión del tanque de oxígeno.

En este sentido, fueron necesarias sendas correcciones el 16 y el 17 de abril, sin las cuales la nave hubiera perdido la necesaria alineación. Los gases en expansión del destrozado Módulo de Servicio provocaron inopportunos movimientos de rotación, que en ocasiones dificultaron las comunicaciones con la Tierra.

## FASE FINAL

Asegurado el viaje de regreso, la NASA se entregó a la tarea de mantener con vida a la tripulación. Swigert permanecería en el Módulo de Servicio durante una buena parte del viaje, vigilando los poquísimos sistemas de control que se habían dejado en marcha para monitorizar las fases más críticas. Sus compañeros, mientras, estarían en el Módulo Lunar. Éste sólo había sido diseñado para dos personas y no permitía excesiva libertad de movimientos.

Swigert tuvo que soportar temperaturas bajísimas, que sólo pudo combatir a base de acumular toda la ropa disponible. Aunque en general intentaron dormir para estar alerta en los momentos esenciales, les resultó difícil ante el desagradable ambiente que reinaba en el interior de la nave. Uno de los máximos problemas, la intoxicación por dióxido de carbono, fue resuelta gracias al ingenio del equipo de tierra. Utilizando sólo elementos que se sabían disponibles a bordo, construyeron una serie de conductos que facilitarían la acción de los purificadores, su intercambio y su distribución en zonas en las que no estaba previsto su uso.

*Lo ocurrido con el Apolo-13 pudo haber cambiado radicalmente el futuro de la exploración espacial*

Después de la última corrección de ruta, que serviría para ajustar el punto del amerizaje, el Módulo Lunar podía ser eyectado. Su utilización les había permitido sobrevivir, así como ahorrar la energía eléctrica que necesitaría el Módulo de Mando durante los breves minutos de operación en solitario, en la reentrada.

Pero antes, el Módulo de Servicio dañado fue separado del conjunto. Eran las 14:25 del 17 de abril. La NASA quería estar segura de que la maniobra no perjudicaría la orientación de la pequeña cápsula. Los astronautas tuvieron entonces la oportunidad de fotografiar el estado en el que había permanecido la nave tras la explosión. Una explosión que apareció en las imágenes mucho más dramática de lo que todos habían pensado.

Los tres hombres se preocuparon ahora de devolver el pulso a los sistemas de su Módulo de Mando. Usando la limitada energía eléctrica de las baterías, observaron, maravillados, que a pesar de las extremas condiciones vividas, todo funcionaba perfectamente.

Con todos ya en su interior y con la escotilla cerrada, el Aquarius fue eyectado a las 17:43. Un último saludo de agradecimiento hacia la nave que les había salvado la vida precedió a su lenta separación.



Rescate de los tripulantes del Apolo 13, tras el amerizaje del módulo de mando en el Pacífico.

Entonces, la cápsula, bautizada apropiadamente como "Odyssey" mucho antes del lanzamiento, inició la secuencia de reentrada sobre la atmósfera terrestre. Las sospechas de que la explosión podría haber dañado el escudo térmico de protección de la nave probaron ser infundadas y el pequeño vehículo tomó contacto con las capas más externas de la envoltura gaseosa de la Tierra a las 18:43.

Debido a la típica formación de una nube de plasma alrededor de la cápsula a consecuencia de las altas temperaturas alcanzadas, las comunicaciones se interrumpieron durante cuatro tensos minutos. Hasta después de este período, nadie sabría si la tripulación había podido finalmente sobrepasar la fase más crítica del regreso.



Los tripulantes saliendo del helicóptero de rescate, en la cubierta del portaaviones Iwo Jima.

Y por fin, la visión de los tres paracaídas extendidos llenó las pantallas de televisión de los hogares de millones de americanos, que celebraron con euforia el feliz final de esta fallida aventura lunar.

La Odyssey amerizó a sólo 6 kilómetros de distancia del portaaviones USS Iwo Jima, en el océano Pacífico. Varios helicópteros acudieron al rescate, llevando a los hombres y a su nave hasta la cubierta del enorme buque.

Sus rostros mostraban el cansancio de todo lo que habían vivido. Parcialmente deshidratados, sin afeitar, fueron sin embargo recibidos con entusiasmo y trasladados con honores de héroe hasta donde les esperaban sus familias. Sucesivas exploraciones médicas demostraron que ninguno de ellos se encontraba enfermo.



Desfile de bienvenida a los héroes del Apolo 13, en las calles de Chicago.

## QUE OCURRIÓ REALMENTE

Cualquier análisis de lo sucedido, aún en la desgracia, ha mostrado a los historiadores que la tripulación del Apolo-13 tuvo suerte. Si el accidente se hubiese producido tras el alunizaje, con el Módulo Lunar abandonado en la superficie de nuestro satélite, probablemente no habrían podido sobrevivir. ¿Y qué habría pasado si el estallido hubiera afectado al Módulo de Servicio cuando el Odyssey hubiera estado sólo ocupado por una persona, girando alrededor de la Luna?

De un modo u otro, las investigaciones posteriores lo único que buscaron fue la razón del accidente. Era necesario discernir si la nave Apolo tenía algún defecto inherente que pudiese afectar a los demás vuelos de la

serie, o si por el contrario lo ocurrido había sido sólo fruto de la negligencia humana, componentes defectuosos, o cualquier otra cosa.

Las explicaciones que dio la NASA no fueron demasiado claras. La que podemos calificar como versión oficial no se daría a conocer hasta cinco años más tarde, durante la publicación de la historia del programa Apolo financiada por la propia agencia.

*Asegurado el viaje de regreso, la NASA se entregó a la tarea de mantener con vida a la tripulación*

---

Según este texto, el tanque de oxígeno número dos del Módulo de Servicio se recalentó y estalló, destruyendo en el proceso el tanque número uno. El primero había estado instalado en el Apolo-10, pero tuvo que ser desmontado y reemplazado ante la aparición de diversos problemas. La operación de desensamblaje dañó el tanque, que tuvo que ser devuelto a la fábrica. Reconstruido, fue sometido a diversas pruebas ya montado en la propia rampa de lanzamiento. En su interior contenía un sistema calentador eléctrico, que serviría para mantener el oxígeno en el estado más apropiado. Dos interruptores gobernaban este sistema, desconectándolo cuando la temperatura era superior a la permitida, pero durante estas pruebas fueron sometidos a corrientes eléctricas excesivas, circunstancia que acabó con ellos. Como resultado de esto, con el Apolo-13 ya en el espacio, los interruptores nunca funcionaron y el tanque de oxígeno se sobrecalentó. Los cables eléctricos que recorrían el interior del tanque se calentaron también, se destruyó su protección térmica y se inició un incendio que ocasionó la explosión.

El tanque, una vez lleno de oxígeno líquido, se convirtió pues en una bomba potencial desde que fue instalado a bordo del Apolo-13. Una bomba que podía estallar en cualquier momento, incluso en el espacio. Lo hizo finalmente a 55 horas y 54 minutos del despegue, a más de 300.000 km de la Tierra. Afortunadamente, la misma tecnología, la misma inteligencia que había provocado el desastre, se encargaría de devolver a casa a los tres astronautas, en una demostración más de cómo el ingenio humano sabe reaccionar ante la adversidad.

## BIBLIOGRAFÍA

- [1] REGINALD TURNILL. Jane's Spaceflight Directory 1988-1989. Jane's Information Group. 1988. UK.
- [2] P.J. PARKER. We Have a Problem! . Spaceflight Vol. 12-9, pág. 356. BIS. UK.
- [3] SHEPARD/SLAYTON/BARBREE/BENEDICT. Moon Shot. Turner Publishing. 1994. USA.
- [4] MURRAY & COX. Apollo. The Race to the Moon. Seeker & Warburg. 1989. USA.



## DESPUES DE LA MARS

*Marcos Sánchez Sotés*

*Estudiante de la ETSETB y Miembro de AESS Estudiants  
marcos27@aess.upc.es*

Poco a poco los ecos de la Mars Pathfinder se han ido apagando. La sonda continua mandando datos pero su repercusión en la prensa prácticamente es nula. Sin embargo la industria aeroespacial ha proseguido impenetrable su camino.

Por un lado, la incombustible MIR sigue en funcionamiento y resolviendo las continuas averías que parecen estar a la orden del día.

Por otro, mientras la sonda Cassini-Huygens se acerca a Venus en busca del empuje gravitacional que la llevará a Saturno, la sonda Galileo nos ha deleitado con excelentes imágenes de auroras jovianas, volcanes de Io y un minucioso estudio de la luna Europa.

Finalmente, la Voyager-1 se despidió del sistema solar el pasado 16 de Febrero, comenzando así la primera exploración más allá del sistema solar.

Todo parece indicar que en 1998 el espacio se consolidará como una actividad económica y científica importante. Mientras los presupuestos públicos siguen descendiendo, las aportaciones privadas crecen día a día. Esto conllevará inevitablemente una visión más «comercial» de la actividad astronáutica.

La empresa más formidable del presente año es, sin duda alguna, la construcción de la estación espacial internacional ISS, para la que están previstas 2 misiones de la NASA (Nodo-1 más un vuelo logístico) y el lanzamiento de dos módulos construidos en Rusia. Mientras se terminan de discutir los presupuestos (ya que los sobrecostes de la estación ya superan los 1000 millones de dólares), es previsible que para Septiembre estén algunos módulos en órbita.

Para los aficionados a la astronomía destacar que el 9 de Diciembre la ESA publicó el Millenium Star Atlas. Un atlas estelar revolucionario, confeccionado con la información suministrada por la misión Hipparcos. Además del espectacular censo de estrellas (más de un millón) y sus datos básicos, se han preparado una serie de mapas en tres volúmenes de gran utilidad para la comunidad astronómica.

Más información en:

<http://astro.estec.esa.nl/Hipparcos>

El 27 de Febrero la ESA lanzó el satélite Hot Bird-4, a bordo del cual se encuentra el procesador SKYPLEX, capaz de demodular las señales recibidas y combinarlas en una sola señal digital. Este sistema, pionero en el tratamiento digital de la señal, fue diseñado por el Departamento de Teoría de la Señal y Comunicaciones de la UPC.

Tras más de un año de proclamar las ventajas de enviar al espacio a un hombre de más de 70 años, John Glenn, astronauta y senador de 76 ha recibido autorización de la NASA para volver al espacio a finales de año. Como es habitual, las reacciones no se han hecho esperar, tanto a favor como en contra, y se ha creado una verdadera GLENNMANIA, creando más expectación que la que tuvo cuando fue por primera vez al espacio. Además cuenta con un compañero de vuelo muy especial: Pedro Duque, el primer astronauta español, con el que tuvimos el placer de contar en las charlas organizadas por AESS Estudiantes el 4 de Diciembre.

Sin embargo, la noticia que ha causado más revuelo en la prensa estos meses ha sido el descubrimiento de hielo en la Luna. La sonda Lunar Prospector, lanzada el 6 de Enero desde Cabo Cañaveral ha confirmado los rumores de existencia de hielo en los polos lunares. Aunque muy dispersa, el agua en la Luna aumenta la viabilidad de futuras misiones de explotación o colonización, ya que podría proporcionar combustible para las naves y aire para los astronautas, provocando un enorme descenso del coste para futuras misiones.

De nuevo, la Lunar Prospector nos demuestra cómo con un presupuesto de 65 Millones de dólares (menos de la mitad que el de la Mars Pathfinder) se pueden obtener resultados excepcionales.

Ahora más que nunca se impone el principio «Faster, Cheaper, Better».



## MIR: LA GRAN DESCONOCIDA



### MIR/DOS-7 (19 DE FEBRERO DE 1986 HASTA EL PRESENTE)

*Texto de David S. Portree, Traducción de Daniel González Alonso<sup>1</sup>*

<sup>1</sup>Estudiante de Ingeniería de Telecomunicación en la ETSETB  
Presidente de AESE Estudiants  
danig27 casal.upc.es

### INTRODUCCIÓN

La estación espacial MIR ha sido objeto de airadas críticas en los últimos meses debido a una serie de fallos de funcionamiento, fallos achacables a su ya avanzada edad de 12 años, más del doble para el cual fue diseñada.

Pero detrás de esta ya vetusta estructura orbitando a 400 Km por encima de nosotros hay toda una historia de avances tecnológicos aplicados a la investigación espacial y sobre todo, la estación MIR es el mejor laboratorio en órbita para estudiar la adaptación del cuerpo humano al espacio.

Por ello, dedicamos este artículo a conocer algo mejor a la MIR, que será sustituida como estación espacial por la International Space Station en breve.

El texto de este artículo es una traducción de un informe técnico de la NASA realizado por David S. Portree. Asimismo, la fotos provienen de la web de la NASA. Agradecemos la colaboración de ésta en la realización de este artículo.

### MIR/DOS-7

La estación espacial Mir es la pieza central del programa espacial tripulado ruso. Su bloque central ha estado en órbita durante 12 años.

Continuas modificaciones han más que triplicado sus capacidades iniciales, más allá que las de ninguna otra estación espacial.

#### Especificaciones de la Mir

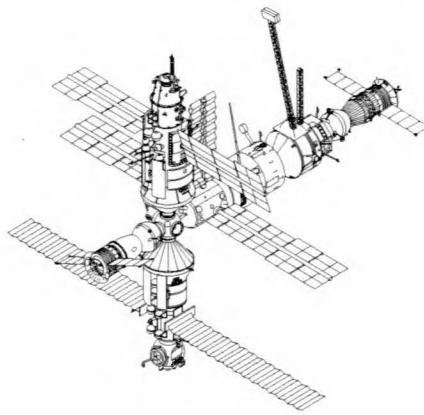
Bloque Central de la Mir

Longitud	13.13 m
Diámetro Máximo	4.15 m
Volumen Habitable	90 m <sup>3</sup>
Peso de Lanzamiento	20,400 kg
Vehículo de Lanzamiento	Proton (3 etapas)
Inclinación Orbital	51.6°
Número de Paneles Solares	2 (3º añadido por EVA)
Longitud de los Paneles	29.73 m
Área de los paneles	76 m <sup>2</sup> (98 con el tercer panel)
Electricidad disponible	9-10 kW a 28.6 V
Naves de abastecimiento	Progress, Progress M
Número de puertos de anclaje	2 longitudinales, 4 radiales
Número de motores principales	2
Empuje de cada motor principal	300 kg

### Complejo Mir

Bloque Central - Kvant, Kvant 2, y Kristall, con las naves Soyuz TM y Progress-M acopladas (noviembre 1994).  
Más adelante se añadió el módulo Spektr.

Longitud	33 m
Máxima longitud a través de los módulos habitables	~28 m
Máximo diámetro de los módulos habitables	4.35 m
Volumen habitable	372 m <sup>3</sup>
Peso	93,649 kg
Inclinación Orbital	51.6°
Número de paneles solares	11
Área de los paneles solares	224 m <sup>2</sup>
Longitud de los paneles solares	29.73 m
Electricidad disponible (máximo teórico)	~27.8 kW
Número de puertos de anclaje	4 longitudinales, 4 radiales
Total de misiones tripuladas	20
Total de misiones de larga duración	17



El complejo Mir, Junio de 1994, con las naves Progress-M23 (derecha) y Soyuz-TM18 (izquierda) acopladas.

de remarcar ya que ha sido ocupado por el módulo Kvant desde Abril de 1987.

Los conectores eléctricos y puertos para transferencias de gas y fluidos que permitieron a la Progress darle servicio a la Mir antes de la llegada del Kvant ahora unen al Kvant al bloque central.

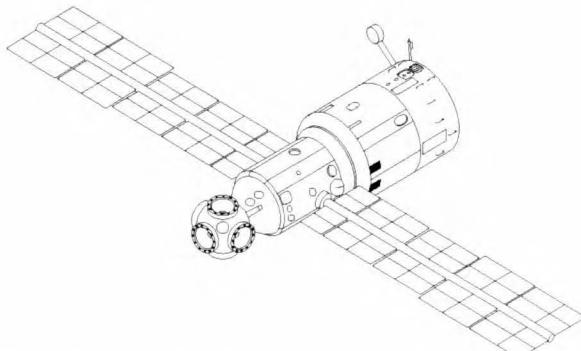
Los vehículos Progress ahora se acoplan con el puerto trasero del Kvant y transfieren gases y fluidos a través del módulo hasta el bloque central a través de estos puertos.

El puerto longitudinal frontal del bloque central está equipado con conductos para transferir combustible y agua desde las naves de abastecimiento Progress-M.

Los paneles solares de Arsenio de Galio producen un aumento del 30% en densidad de potencia sobre los paneles de silicio del Salyut 7 (hasta 120 W/m<sup>2</sup>). Los Soviéticos llevaron a cabo bastantes experimentos con Arsenio de Galio desde 1978 en los Salyut 6 y Salyut 7. Los paneles de la Mir tienen casi el doble de longitud de los del Salyut 7. La Mir fue lanzada con un adaptador encima de su compartimento presurizado de pequeño diámetro para añadirle un panel solar auxiliar.

La Mir fue diseñada para ser utilizada con los satélites Soviéticos geosíncronos Altair/SR de transmisión de voz y datos. Estos son satélites operados bajo la obicua designación Cosmos. El sistema de satélites es a veces designado SDRN (Satellite Data Relay Network) o Luch. Una gran antena para comunicaciones radio con el sistema Altair/SR se extiende desde el extremo posterior de la Mir.

Aunque la mayoría de los desperdicios de la Mir son depositados en los compartimentos de carga de las naves Progress, la Mir, como sus predecesores Salyut de tipo DOS, tiene un pequeño cierre presurizado que puede ser usado para deshacerse de desperdicios. También es usado para experimentos científicos que requieran vacío.



Bloque Central de la Mir: El nodo multipuerto en la parte final anterior (izquierda) tiene un puerto de anclaje longitudinal y cuatro puertos de anclaje laterales.

No hay escotilla para excursiones EVA en el bloque central de la Mir, aunque antes de la llegada del Kvant 2 y Kristall los cosmonautas podían salir por cualquiera de los 5 puertos. El nodo es más corto que el compartimento de transferencia del Salyut 7, explicándose esta circunstancia por la menor longitud total de la Mir.

El sexto puerto de anclaje se encuentra situado en el extremo posterior del bloque central. Se parece bastante a los puertos traseros de los Salyut 6 y Salyut 7. Es digno





Sergei Krikalev, que voló en el transbordador espacial Atlantis en Febrero de 1994 y estuvo en períodos de 2 horas en la Mir, hizo bastantes comentarios comparando las condiciones en el transbordador y en la Mir.

En general, Krikalev dijo que las condiciones de habitabilidad en la Mir eran mejores que las del transbordador.

Él atribuía esto a que la Mir estaba diseñada para vuelos de larga duración, mientras que el transbordador estaba diseñado para mantener a una tripulación durante cortos períodos de tiempo.

Largas estancias en el espacio también afectan al entrenamiento y distribución del tiempo en las tripulaciones de la Mir. Las tripulaciones experimentan más entrenamiento sobre el terreno que las tripulaciones del transbordador, porque es imposible simular una estancia de 6 o 12 meses en una estación y entrenarse para cualquier eventualidad. Sus horarios están mucho menos planificados que los de los astronautas del transbordador, que sólo pasan unos pocos días preciosos en el espacio y deben poner virtualmente cada minuto para uso productivo. Además, los cosmonautas de la Mir tienen tardes y fines de semana de fiesta. (En la práctica, los cosmonautas a menudo trabajan en su tiempo libre, ya sea porque están muy motivados o porque los experimentos lo requieren.)

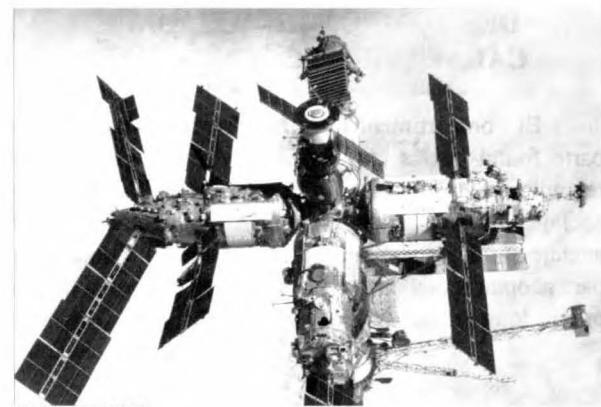
Krikalev dijo también que las condiciones de habitabilidad en la estación dependían enormemente de las preferencias de la tripulación residente. Krikalev dijo que los niveles de limpieza y olores variaban de acuerdo con los estándares que la tripulación residente estaba dispuesta a aceptar.

De manera similar, el nivel de ruido variaba considerablemente. También dijo que, en sus estancias en la Mir, ésta estaba bastante ordenada. Las tripulaciones de las que formó parte intentaban mantener el equipamiento que no usaban detrás de los paneles de la pared, e intentaban evitar poner redes llenas de equipo colgadas del techo de la estación.

La Mir depende mucho más notablemente en la automatización que las estaciones de tipo DOS anteriores,

como parte de la tendencia Soviética a aumentar la automatización en las naves tripuladas. Esto también es evidenciado por las modificaciones en las Soyuz-TM y Progress-M. Una publicación francesa la llamó «La primera estación computerizada en órbita». Además de los ordenadores de control de la estación, cada cosmonauta tenía un ordenador personal. La estación fue lanzada con el ordenador Argon 16B. En 1990, su reemplazo, mucho más capaz, el Salyut 5B, que había sido entregado por el Kvant 2 en 1989, fue conectado.

La Mir fue lanzada con su puerto longitudinal frontal equipado con el sistema de guiado y acoplaje Kurs (Curso) usado por las Soyuz-TM (y ahora también por las Progress-M). El puerto posterior estaba equipado con el antiguo sistema Igla para que los cargueros Progress pudieran seguir anclando allí, y también para permitir el anclaje del Kvant, guiado por el sistema Igla, en 1987. El puerto posterior del Kvant fue equipado con los 2 sistemas, Igla y Kurs. El sistema Igla ya no se utiliza.

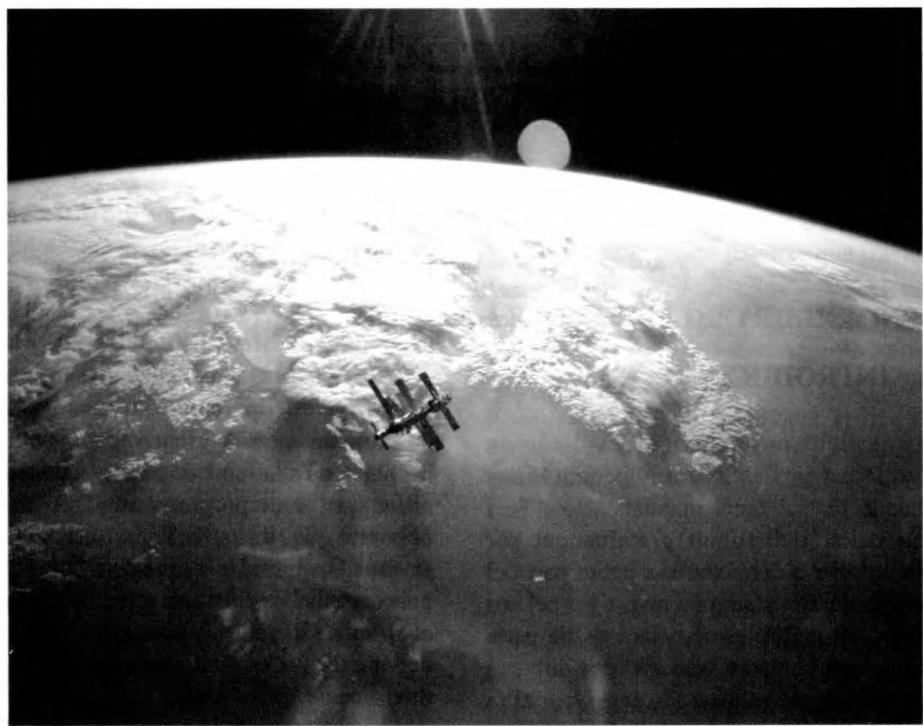


El anclaje de del módulo Kvant bloqueó -aparentemente permanentemente- los motores de mantenimiento de órbita del bloque central de la Mir. Todas las maniobras de mantenimiento de órbita desde 1987 han sido llevadas a cabo por las naves acopladas (Progress, Progress-M, y Soyuz-TM).

El casco de presión de la Mir está compuesto por láminas de aluminio tratadas químicamente con un grosor medio de 2 mm, trenzada en redes de 4 mm de ancho. El casco tiene un grosor de 5mm en el área de la unidad de anclaje multipuerto y de 1.2 mm en el compartimento de trabajo de pequeño diámetro.

El compartimento de gran diámetro está cubierto por un radiador de 2 mm de ancho con una separación de 20 mm del casco.

Otras porciones del casco están cubiertas por un recubrimiento termal multicapa comprendido por térmico medio por 25 capas de Mylar aluminizado y scrim. Cada capa tiene un grosor de 5 micrómetros. Las capas de material de tipo Kevlar cubren el recubrimiento termal.



Los cosmonautas tienen dos cabinas separadas (bolsillos en las paredes del compartimento de gran diámetro) para dormir y privacidad. Un lavabo sellado está situado en la pared trasera de uno de los compartimentos.

Los cajones de almacenamiento ocupan la mayor parte del espacio de pared en tanto los compartimentos de pequeño o gran diámetro.

La Mir tiene varias ventanas de observación, con cierres para protegerlas de impactos de deshechos en órbita y depósitos formados por el uso de los motores de control de actitud. Cada cabina de astronauta tiene una pequeña ventana, y hay una ventana en el «suelo» de la estación para observación de la Tierra.

El lavabo tiene una unidad esférica para lavarse el pelo con entrantes de goma por los cuales se insertan la cabeza y las manos.

El área habitable (secciones de pequeño y gran diámetro) mide 7.6 m de longitud. La sección de pequeño diámetro tiene suelo de color verde oscuro y paredes de color verde claro; la sección de gran diámetro tiene suelo marrón y paredes amarillas.

Ambas secciones tienen techos blancos con luces fluorescentes.

El área de ejercicio de la Mir es también una sala de vídeo para ver cintas y oír música mientras se hace ejercicio. El velo-ergómetro de la Mir (bicicleta estática) puede plegarse y guardarse bajo el suelo. También hay una cinta para correr.

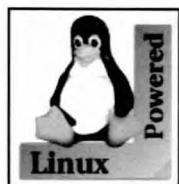
La «enfermería» de la Mir es una cabina situada en el enlace de las secciones de pequeño y gran diámetro del compartimento habitable, cerca del área de ejercicio.

La consola de control de la Mir está orientada hacia la unidad de anclaje frontal, como en anteriores Salyuts de tipo DOS. Dos pantallas de televisión permiten comunicaciones cara a cara con el control de tierra. Cuatro más, agrupadas en pares en cada lado de la escotilla separando el compartimento habitable de la unidad de anclaje multipuerto, permiten monitorizar los módulos anclados al nodo multipuerto (una pantalla por módulo).

Diferentes destinos han sido propuestos para la Mir en los últimos años. Durante un tiempo, una lanzadera de tipo Buran debería haber entregado un nuevo bloque central en 1992. La Buran habría usado un brazo robótico para liberar los enlaces del antiguo bloque central y sustituirlo por el nuevo. El antiguo bloque central habría sido devuelto a la tierra en la Buran. De acuerdo con Yuri Antoshechkin, Director de Vuelo de los Sistemas de la Mir, el final de la vida útil de la estación coincidirá con la puesta en marcha de la Estación Espacial Internacional (ISS), en 1999, con lo cual habrá estado en órbita más de 12 años (más del doble para el que fue diseñado).

Continuará en vuelo no-tripulado durante un año, sirviendo de plataforma experimental para un sistema solar dinámico de potencia conjuntamente desarrollado por los EE.UU. y Rusia.

La estación será probablemente movida de órbita y forzada a reentrar sobre un área preseleccionada del océano pacífico, como fue hecho con varias estaciones Salyut.



# LINUX: EL SISTEMA OPERATIU UNIX PER A PC

Bartomeu Alorda Ladaria  
Jordi Meya Pérez

Estudiants d'Enginyeria Superior de Telecomunicacions  
ala Universitat Politècnica de Catalunya  
[barnif@mx3.redestb.es](mailto:barnif@mx3.redestb.es), [lmeya@pie.xtec.es](mailto:lmeya@pie.xtec.es)

## INTRODUCCIÓ

En aquest article no pretenem explicar en profunditat el funcionament del Sistema Operatiu Linux que per altra banda tants i tants manuals, How-To i llibres se n'han editat i distribuït gratuïtament per internet. Pretenem donar a conèixer una petita part del tema i que sigui el lector qui s'anmi a investigar pel seu compte, ja que trobarà múltiples associacions de parla anglesa i castellana que comparteixen els seus dubtes i ajuden a difondre encara més aquest Sistema Operatiu i en el fons, una altra manera de veure la informàtica.

## INTRODUCCIÓ HISTÒRICA

Per començar a conèixer un Sistema Operatiu és necessari saber els seus orígens i com ha arribat a l'estat actual. Linux té els seus inicis plenament arrelats en l'àmbit universitari i amb el desenvolupament del Sistema Operatiu Unix.

L'any 1968, un grup d'investigadors de General Electric, AT&T Bell i de l'institut Tecnològic de Massachusetts, va realitzar un projecte especial d'investigació de sistemes operatius anomenat Multics. Aquest incorporava molts conceptes nous de multitarea, gestió d'arxius i interacció amb l'usuari. El 1969, Ken Thompson, de AT&T Bell Laboratories (Computing Science Research Laboratories de Murray Hill, New Jersey), va desenvolupar el sistema operatiu UNIX incloent moltes característiques del projecte Multics. Des de la seva concepció, es va veure que Unix era un eficient sistema operatiu multiusuari i multitarea amb un preu accessible.

El 1970, Dennis Ritchie (que va desenvolupar el llenguatge de programació C com una eina flexible pel desenvolupament de programes) va col-laborar amb Ken Thompson en la feina de passar el codi del sistema Unix al llenguatge de programació C. Aquest llenguatge va permetre a D. R. i a K. T. escriure una única versió del sistema Operatiu Unix, que podia després ser compilada pels compiladors de C de diferents computadores. Així, el sistema Operatiu va passar a ser transportable i capaç de ser executat en diverses computadores amb poc o cap modificació de les fonts.

En 1972, els laboratoris Bell començaren a distribuir versions oficials de Unix i a facilitar llicències a diferents usuaris. Un d'aquests va ser la Universitat de Califòrnia, en Berkeley. Aquesta va afegir moltes característiques noves al sistema que, després es convertiren en estàndard. En 1975, Berkeley va llançar la seva pròpia versió de Unix, coneguda per: Berkeley Software Distribution (BSD). En 1980, Microsoft va desenvolupar una versió del sistema pels PC anomenada Xenix. AT&T va desenvolupar diverses versions d'investigació i el 1982 va llançar la seva primera versió comercial, anomenada Sistema 3. A aquesta la va seguir Sistema V, que va comptar amb seguiment empresarial.

El 1983 i com a resultat d'un projecte del DARPA (Agència de Projectes d'Investigació Avançada del Departament de Defensa del EEUU), es va llançar una potent versió anomenada BSD 4.2 que incloïa la possibilitat de treballar amb xarxes basades en la família de protocols TCP/IP.

En 1991, els Unix System Laboratories (organització afí a AT&T) va desenvolupar la versió Sistema V 4 i com a resposta altres companyies (IBM i HP) varen establir la Open Software Foundation per crear la seva pròpria versió estàndard de Unix.

El 1993, AT&T va vendre els drets de Unix a Novell i actualment les versions desenvolupades pels Unix System Laboratories s'anomenen UnixWare.

SunOS va ser desenvolupat per a les estacions de treball Sun i Aix per a les estacions de treball IBM. Així i tot, a mesura que els PC augmentaven el seu potencial es varen desenvolupar versions per aquests. Xenix i System V/386 són versions per aquestes arquitectures. AUX és una versió que s'executa en Macintosh.

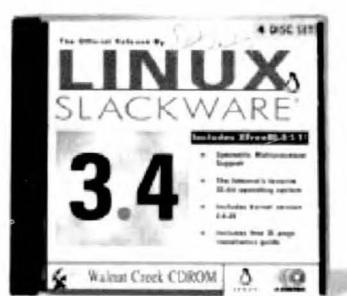


Figura 1: CD-ROM amb la versió Slackware 3.4.

Aquesta seria l'evolució a grans trets del sistema Unix que molts de nosaltres haurem usat alguna vegada en pràctiques, ja que és un dels sistemes més estesos a les Universitats (des dels seus orígens, com hem pogut comprovar). Ara bé, també és un sistema molt usat en empreses que necessiten d'un sistema fiable, potent, amb gran flexibilitat i que no els fa por una interfície poc «amigable».

En aquesta història hi falta l'aparició del Linux, el qual des dels seus orígens va ser dissenyat per PC basats en Intel. Va tenir el seu origen com a projecte personal d'un estudiant de Ciències de la Informàtica de la Universitat de Helsinki anomenat Linus Torvald. En aquesta època els estudiants usaven un programa anomenat Minix, que incorporava diverses característiques Unix. Va ser creat pel professor Andrew Tannenbaum i distribuït per Internet a la comunitat universitària.



**Figura 2:** Símbol de la distribució de BSD

La intenció de Linus era la de crear una versió completa de Unix per PC aprofitant les característiques de commutació de feines en mode protegit del 386. El va escriure tot en ensamblador (Linus Torvalds: «dos mesos de treball fins que vaig tenir un driver de discs -amb nombrosos bugs, però que pareixia funcionar al meu PC -i un petit sistema de fitxers» [1]).

El 1991 va donar a conèixer la versió 0.11 que va ser distribuïda per Internet i durant els anys següents es varen anar fent afegits i desenvolupant característiques que es trobaven al sistema estàndard de Unix. Tots els principals administradors de finestres han sigut traslladats a Linux, la família de protocols TCP/IP i també un complet conjunt d'utilitats de desenvolupament de software, com a compiladors de C i depuradors.

Encara que Linux s'hagi desenvolupat en un entorn lliure i obert com és internet, compleix amb les normes oficials del Unix. L'IEEE ha desenvolupat un estàndard de Unix independent per l'institut ANSI (American National Standards Institute). Aquesta nova versió s'anomena POSIX (Portable Operating System Interface for Computer Environment) i defineix com ha d'operar un sistema tipus Unix especificant detalls com les crides a sistema i les interfícies. Linux ha estat desenvolupat des de l'inici seguint la norma POSIX.

Debగ a que ha tingut el germà gran Unix, ha tingut un creixement més ràpid i en pocs anys ha arribat a l'estat de ser un dels millors sistemes operatius amb

totes les avantatges Unix, però que s'executa sobre màquines menys costoses i amb la característica de ser gratuït.

Si es vol aprofundir més en el tema històric es pot consultar llibres sobre Linux o bé navegar per Internet i visitar les nombroses pàgines que tracten sobre Linux.

## VERSIONS I DISTRIBUCIONS ACTUALS

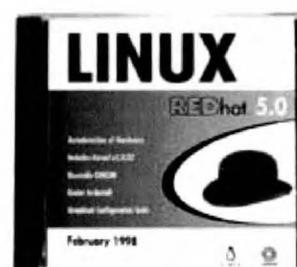
Ara ja sabem que Linux és un sistema en continu desenvolupament, bé realment el que està en continu avanç és el kernel (o nucli) del sistema, actualment ens trobem a la versió 2.0. El número de versió segueix un cert format:

2.x.y

x: Si és parell és una versió estable (versió acabada)  
y: Si és senar és una versió inestable (versió beta)

Indica la correcció de qualche error menor dins la mateixa versió x

Aquesta codificació de les versions permet identificar cada kernel. La versió actual de Linux permet definir-lo com a multitasca i multiusuari (com qualsevol versió de UNIX) capaç d'executar el sistema de finestres X-Window, la família de protocols TCP/IP, l'editor EMACS i software de correu i news. A més, el software freeware que pot executar-se és molt divers, potent i un dels seus millors avantatges, ja que el kernel per si sol no permet grans meravelles (imagineu-vos un Windows 95 sense cap programa addicional). El kernel es desenvolupat quasi íntegrament en C i gestiona totes les tasques relacionades amb el hardware del PC.



**Figura 3:** CD-ROM amb la versió REDhat 5.0.

Actualment han aparegut diferents distribuïdors (empreses que distribueixen el S.O. i que desenvolupen nou software) que posen a la venda per un preu molt reduït un paquet de programes que incloent la darrera versió del kernel de Linux introduceixen més programes de lliure distribució o propietaris dels mateixos distribuïdors que s'han anat implementant per millorar l'accés a l'usuari. Aquestes millores fan que cada vegada a'incrementi més el nombre d'usuaris que descobreixen les grans avantatges de Linux. A més, degut al gran augment d'usuaris d'aquest sistema hi ha empreses comercials que han traduït els seus programes per aquest

sistema. Exemples d'aquest fet són Netscape (<http://www.netscape.com>), WordPerfect 7 (<http://www.sdcorp.com/wplinux.htm>), CorelDraw (<http://www.caldera.com/solutionsCD/Products/Corel/Info/draw35.htm>) i tampoc la poderosa Microsoft ha passat per alt el gran augment d'usuaris de Linux i està a punt de llançar una primera versió alfa del seu Internet explorer 4.0.

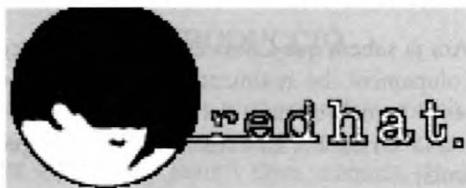


Figura 4: Logotip de la distribució REDhat.

Hi ha nombroses distribucions i cada una d'elles introduceix qualche element o programa diferenciador encara que el kernel sigui el mateix. Això possibilita adquirir una d'aquestes distribucions i fer les renovacions del kernel a través d'Internet de manera gratuïta. També s'han publicat distribucions plenament funcionals a diverses revistes d'informàtica. Però potser la millor forma de buscar informació sobre les diverses distribucions sigui a través d'internet i de les seves pàgines centrals.

Fa uns anys la distribució més coneguda i una de les més extenses internacionalment era la Slackware (el seu distribuïdor oficial es troba a <http://www.cdrom.com>). La seva forma d'instal·lació és relativament senzilla i inclou tot el software que un usuari mig i avançat pugui necessitar. La distribució es divideix en sèries temàtiques, i cada una d'elles en directoris de 1,44 Mbytes de capacitat.

Actualment ha sortit un fort competidor a la Slackware, la Red Hat (el seu distribuïdor oficial és Red Hat Software i es troben a <http://www.redhat.com>). Aquesta distribució es distingeix pel seu attractiu sistema d'instal·lació en mode gràfic i per incloure multitud de programes desenvolupats per facilitar les tasques d'actualització. Aquesta distribució també està organitzada en paquets instal·lables de forma independent i agrupats per temes, encara que no hi ha cap adequació a les capacitats dels disquets, per tant es distribueix amb CDROM o per xarxa. Aquesta distribució no permet la instal·lació a través de disquets i per tant és obligatori tenir una partició «Linux native» (ext2fs) només per ser ocupada pel Linux. La distribució de Caldera (<http://caldera.readysoft.es>) està basada en la Red Hat.

Una de les distribucions més completa i més usada actualment és la Debian GNU/Linux (<http://www.debian.org>) que inclou tant software que arriba a dificultar la seva distribució. Un altre punt al seu favor és la possibilitat de poder usar software de la Red Hat augmentant molt més les possibilitats.

Aquestes podrien ser les distribucions més conegudes i més usades actualment, però degut a l'increment d'usuaris de Linux estan apareixent noves distribucions. A més actualment no és necessari navegar fins la pàgina oficial per obtenir una de les distribucions, podem connectar-nos a FTP anònim de <ftp://ftp.rediris.es/software/linux/distributions/> o simplement usant el servidor FTP anònim de la UPC <ftp://ftp.upc.es> podem trobar un mirall de la part de Linux del servidor Sunsite on ens pot arribar a sorprendre la quantitat d'informació sobre linux i la possibilitat de baixar-nos multitud de software que hi trobarem.

Així i tot una de les adreces més important en el món Linux és <http://sunsite.unc.edu> on es poden trobar distribucions, manuals, HowTo i altres informacions que ens facilitaran el coneixement d'aquest sistema.

## LA INSTAL·LACIÓ DEL SISTEMA

Cada distribució ajuda menys o més en la instal·lació dependent del software de que disposi per aquest efecte. A més, el procés d'instal·lació sol ser la més difícil de l'hora d'introduir-se en el sistema. No podrem ajudar-nos d'una aplicació com la de Windows, ja que serà necessari tenir clares una sèrie de conceptes interns de funcionament.

Encara que Linux pot conviure amb qualsevol sistema operatiu, han de deixar-li lloc per la seva partició. Per aquest motiu totes les distribucions duen un programa DOS anomenat «FIPS» que permet establir noves particions en el disc dur sense necessitat de formatajarlo.



Figura 5: Logotip de l'empresa que distribueix la versió Caldera de REDhat.

També hi ha un programa «UMSDOS» que permet instal·lar el Linux dins un directori de la partició DOS, encara que amb certes limitacions. (Nota: Aquesta no creiem que sigui la millor opció tenint en compte que no hauríem de permetre que un sistema tan bo com Linux hagi de conviure amb Windows 95, pur màrqueting).

La primera passa serà la creació dels disquets d'inici de Linux i iniciar la màquina amb aquells.

S'haurà de definir un espai del disc com a espai d'intercanvi (Swap). Aquesta partició serà usada per Linux com a memòria de disc i és recomanable que sigui d'una grandària aproximadament de 1,5 vegades supe-

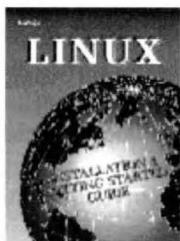


**Figura 6:** Logotip de Debian GNU.

rior la quantitat de RAM. S'ha de dir que Linux no té la famosa limitació dels 640 Kb i fa servir el Swap de manera força eficient.

Seleccionar el software a instal·lar que depenen de la distribució estarà en sèries o directoris. Seria recomanable llegit el HowTo de la distribució que estem instal·lant on es descriuen cada sèrie. Existeix un Manuals traduït al castellà que es pot baixar a l'adreça del projecte LuCAS (Linux en CAStellà) <http://www.infor.es/LuCAS>

La instal·lació del carregador LILO ens permetrà seleccionar el sistema amb el qual volem iniciar el nostre computador. Podrem iniciar amb el sistema operatiu anterior o amb el Linux, d'aquesta manera els dos sistemes podran conviure en el mateix computador i podrem triar-los elegantment (configurable a posteriori al fitxer /etc/lilo.conf).



**Figura 7:** Tapes del llibre "Installation & Getting Started Guide" distribuït per Internet.

Durant la instal·lació serà possible muntar la partició del sistema operatiu anterior com a directori dins la partició de Linux, d'aquesta forma es tindrà accés a tota la informació de les altres particions del disc dur. A senyalar que de moment no accepta noms llargs de Win95 (s'està treballant en això) tot i que el sistema de fitxers del propi Linux és molt més flexible en quant a llargària (fins a 255 caracters) i permet múltiples punts en un fitxer.

Si es vol tenir accés a la informació de Linux des del sistema Operatiu DOS ens haurem de baixar un programa anomenat «alien» de l'adreça <http://kite.ml.org/programs/alien/>.

Amb aquestes idees clares la instal·lació no hauria de resultar molt problemàtica. Així i tot, a Internet podrem trobar tota la informació necessària en forma de FAQ, HowTo, Mini-HowTo o altres documents, molts

traduïts al castellà. Per altra banda, el sistema plug'n'play de Windows 95 no funciona (encara que en Windows 95 tampoc no ho fa massa bé), i s'hauran de muntar (sempre com a root) la resta dels components del sistema (disquetera, CD-ROM, placa de so, impressora, ZIP,



**Figura 8:** Logotip del projecte de documentació LuCAS (LinUx en CAStellà)

scanner i la resta de perifèrics), però això no serà cap problema, ja que la gran Xarxa de la Informació servirà de biblioteca i no s'haurà de fer reserves ni esperar tenir sort a l'hora de trobar el nostre problema documentat, ja que tot el sistema es trobarà totalment documentat, i segur que en tot el món hi ha algú que ha tingut que realitzar les mateixes tasques i les haurà documentat. Així i tot, sempre es tindrà accés a les fonts del sistema i es podrà introduir les modificacions per configurar un nou hardware. D'aquesta forma s'elimina la possibilitat d'errors no documentats, arquitectures desconegudes o bugs impossibles de solucionar per molt aficionat a la informàtica que sigui hom.

Una vegada instal·lat, l'interfície amb l'usuari serà molt semblant a la de Unix on s'haurà de donar un nom d'usuari i un password per entrar al sistema. A més, amb la comanda «adduser» es podrà crear nous usuaris del sistema. Sempre hi haurà el superusuari que s'encarregarà del manteniment de tot el sistema. Aquest superusuari, anomenat root, tindrà tots els privilegis de configuració i manipulació, mentre que la resta d'usuaris només tindran privilegis sobre la seva zona de treball. D'aquesta forma el sistema queda totalment protegit de manipulacions diferents del root.

## ALGUNS TRUCS BÀSICS PER A DUMMIES

Hi ha certes instruccions bàsiques de qualsevol S.O. que en aquest cas són diferents per les de MS-DOS i les de Linux, aquí voldríem comentar-ne algunes, encara que potser resultaran triviales, poden portar maldecaps si no se coneixen. Abans de res recordar que es tracta d'un S.O. multiusuari i el primer que fa el sistema quan s'inicia és demanar un nom d'usuari i una paraula clau o «password». Aquesta paraula clau l'haurà demanada durant la instal·lació, però en cap lloc sol·licita un nom d'usuari. Això passa perquè en el món UNIX sempre hi ha un superusuari anomenat «root». Així per tant, el nom d'usuari ha introduir serà root i la clau secreta la introduïda durant la instal·lació.



Altres comandes bàsiques estan resumides:

## CONCLUSIONS

Comanda	Descripció
mkdir	Crea un directori.
rmdir	Esborra un directori buit.
cd ..	Puja al directori superior (respectar espai)
ls -F	Fer un llistat de fitxers
cp	Copiar fitxers
mv	Moure fitxers
rm	Esborrar fitxers
Maj+PgUp	Puja la visió de la pantalla i permet veure llistats llargs
Maj+PgDown	Baixa la visió de la pantalla i permet veure llistats llargs
halt	Apagar el sistema (sempre com a root)
reboot	Fer un reset (sempre com a root)
Alt+Tecles Funció	Obrir diferents sessions en mode no gràfic
man comanda	Dóna informació detallada de cada comanda

Resum de les comandes més importants del Linux.

Linux ha tingut un inici a l'entorn universitari, el seu desenvolupament es va dur a terme a través de la xarxa d'Internet i encara s'hi continua treballant. Aquesta versió de Unix per PC, que per altra banda és gratuïta, va fer canviar la forma de treballar de molta gent igual que el Unix, però aquesta vegada ja no es tracta tan d'universitats, sinó que comença a introduir-se a la casa d'usuaris no satisfets amb les modes.

Linux és freeware, o sigui, de lliure distribució i per tant el software que és la vida d'un sistema operatiu i que li permet fer tasques avançades, no serà, en la gran majoria dels casos, comercial. Així, la via per aconseguir aquest software és per Internet (els programes de qualsevol altre sistema necessiten emuladors a no ser que estiguin fets amb llenguatge Java), o per revistes que parlin de Linux i incloguin un CD-ROM amb aplicacions -que no seran demostracions-. Ara bé, actualment no tindrem gaires problemes a l'hora de trobar aplicacions per a Linux que facin la gran majoria de tasques que estem acostumats, amb l'única característica de no ser software desenvolupat amb el marketing darrera l'orella.

Ara bé, no tot és de color de rosa a la vida i fins i tot el Linux pot tenir alguna pega que s'ha de comentar. Primerament, el canvi radical que suposa per l'usuari en la concepció de l'ordinador, de la comprensió del funcionament d'aquest i dels molts hàbits que s'agafen

usant Windows. Linux obliga a que s'entengui, per exemple, com funciona la comunicació amb el mòdem si es vol tenir accés a Internet.

També suposa un sistema de fitxers i de usuaris nou, on un usuari té els SEUS fitxers, amb privilegis establerts, amb zones ocultes, amb mesures de seguretat,... i tot això s'ha d'administrar. Aquest fet no suposa un gran trasbals, però cal aconsellar fer dos usuaris, un per només administrar el sistema (root) i l'altre per usar sense perill i evitar, sobretot a l'inici, fer modificacions del sistema irreversibles.

A més i degut a ser gratuït no hi ha serveis oficials per reclamar, amb la qual cosa es pot sentir una mica desemparats al principi. Això es veu compensat amb molts fitxers d'ajuda que es poden trobar per Internet, fins i tot en castellà. Els darrers anys s'han fundat nombrosos grups d'usuaris de Linux en castellà que prenenen ajudar a qualsevol usuari amb llengua castellana. Entre aquests destaquen:

Projecte LuCAS (Linux en CAStellà):

<http://www.infor.es/LuCAS/>

Spanish Linux User Group (SLUG):

<http://slug.ctv.es/>

Grup INSFLUG:

<http://www.insflug.nova.es/>

També hi ha múltiples adreces personals, entre elles:

<http://slug.ctv.es/~acastro/slugfaq>

<http://vivaldi.Princeton.EDU/~angel/LinuxFocus/Castellano>

<http://www.cs.us.es/archive/linux.html>

<http://www.arrakis.es/~barreiro>

<http://slug.ctv.es/~alfredo>

<http://www.arrakis.es/~ulusa/linux/linux.html>

<http://www.ctv.es/USERS/xose/>

<http://www.redestb.es/personal/jordivi/>

<http://www.infor.es/LuCAS/htmls/vendors.html>

## BIBLIOGRAFIA

- [1] REVISTA PC ACTUAL. Any IX, 90. Octubre 1997. «Los primeros pasos»
- [2] LINUX: MANUAL DE REFERENCIA. Ed. MacGraw-Hill. 1994.
- [3] REVISTA PC ACTUAL. Any IX, 91. Novembre 1997. «Configuración y Comandos»
- [4] P.J. BROWN, Starting with UNIX. Ed. Addison Wesley. 1984.
- [5] REVISTA PC ACTUAL. Any IX, 92. Desembre 1997. «El entorno Gráfico X»
- [6] L'accés a totes les adreces esmentades a l'article i més, ha estat la millor i més completa referència bibliogràfica.



## ¿QUÉ ES Y QUE REPRESENTA EL IEEE?

Ángel Cardama

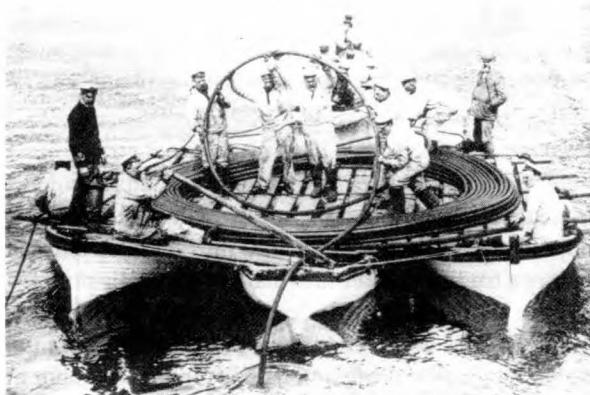
Catedrático del Departamento de Teoría del Señal y Comunicaciones de la UPC

Consejero de la Rama de Estudiantes del IEEE de Barcelona

cardama@tsc.upc.es

El IEEE (Institute of Electrical and Electronics Engineers) es la mayor asociación profesional del mundo. Fue fundada en Estados Unidos en el año 1884 y en la actualidad sus actividades se extienden a 147 países y cuenta con unos 320.000 asociados, unos 1.500 de ellos en España. Las actividades técnicas (y profesionales en EE.UU.) que abarcan la sitúan en el núcleo de lo que hoy en día se denominan genéricamente las tecnologías de la información y de las comunicaciones, es decir la electrónica, la informática y las telecomunicaciones. La afiliación se hace a diferentes niveles, que abarcan desde el de estudiantes hasta el de los profesionales de reconocido prestigio.

Su peso en el mundo científico-técnico es considerable: en sus revistas, simposios y conferencias se publican aproximadamente la cuarta parte de todos los trabajos que se originan en el mundo sobre esas temáticas y es uno de los organismos que está teniendo más impacto en la elaboración de nuevas normas o estándares. La organización de un volumen tan grande de actividades es ciertamente compleja y se hace a través de un conjunto de sociedades técnicas, treinta y siete en la actualidad, centradas en temáticas especializadas (electricidad, comunicaciones, electrónica, informática, educación, etc.).



Los primeros tiempos en la instalación de cable submarino. En la fotografía puede verse el desenrollado de un cable desde una balsa de tres botes.

La estructura actual de la sociedad es el fruto de la disgregación y de la fusión de diversas sociedades a lo largo de los últimos cien años. Sus orígenes están en la aparición de dos nuevas tecnologías comerciales en el siglo pasado: el teléfono y la electricidad, que han conformado en gran medida nuestra vida social actual. La instalación de alumbrado eléctrico en las ciudades fue el



Las comunicaciones telefónicas crecieron rápidamente durante los primeros años de la telefonía, de lo que resultó una auténtica maraña de cables desordenadamente tendidos. Durante los desastres naturales, como las fuertes nevadas de 1888 en la ciudad de Nueva York, muchos cables se caían, interrumpiendo el servicio.

motivo de la aparición de un número importante de técnicos eléctricos y de la creación en 1884 del AIEE (American Institute of Electrical Engineers); entre sus miembros fundadores se encontraban Thomas Alva Edison y Alexander Graham Bell. Una de las primeras actividades del AIEE fue el establecimiento de normas profesionales para los ingenieros eléctricos y de estándares para



Guglielmo Marconi junto al primer receptor de ondas de radio de alcance transoceánico



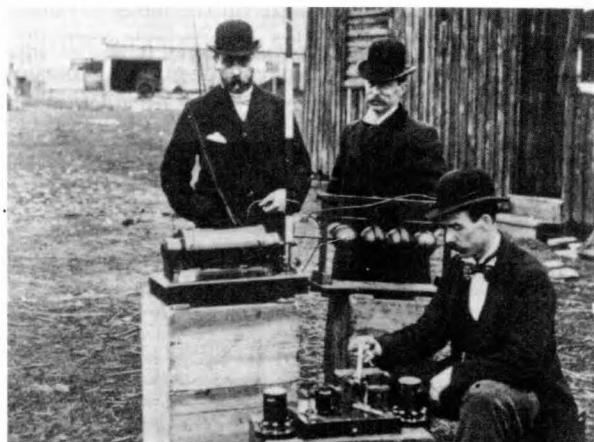
RAMAS DE ESTUDIANTES DEL IEEE



Alexander Graham Bell, inventor del teléfono

la industria eléctrica. A principios de siglo esta sociedad se había consolidado y ejercía una influencia comparable a la de otras sociedades más antiguas (ingeniería mecánica, civil o minas, por ejemplo).

En esa misma época por todos los países se entrecruzaban líneas telegráficas, se habían tendido cables telegráficos submarinos y el teléfono empezaba a tomar un auge considerable, por lo que dentro del AIEE comenzaron a surgir grupos especializados que no encajaban cómodamente en la orientación predominantemente eléctrica de la sociedad. Uno de los grupos más activos era precisamente el de los pioneros de la radio, que finalmente en 1912 crearon su propia sociedad, el IRE (Institute of Radio Engineers). En sus inicios tuvo una afiliación muy escasa, pero de amplia base internacional; precisamente ese fue uno de sus rasgos más característico, comparado con otras sociedades profesionales americanas o europeas. El IRE tuvo un papel preponderante en el establecimiento de la radiodifusión y de la televisión.

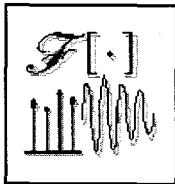


Personal de Correos examinando el aparato de Marconi para experimentos a través del Canal de Bristol, 1897.

A mediados de los años 50 se produjo un vuelco en los tamaños de ambas sociedades y pasó a ser el IRE más numeroso que la AIEE, tendencia que se consolidó en los años siguientes y que concluyó en 1963 con la fusión de ambas para crear el IEEE; fue en gran medida una vuelta a los orígenes, pero con el peso de las actividades transferidas de la ingeniería eléctrica a la electrónica, a diferencia de lo que ocurría a comienzos de siglo.

En contraste con esta situación en norteamérica, el panorama de las sociedades correspondientes en la Europa continental era totalmente diferente; sólo existían colegios profesionales con una perspectiva muy limitada de defensa de las atribuciones profesionales y escasas actividades científicas. El único país europeo en el que existían organizaciones semejantes a las norteamericanas era Gran Bretaña.

Es precisamente ese vacío de actividades tecnológicas en los colegios profesionales europeos, la pujanza tecnológica de EE. UU. y, en cierta manera, la larga tradición internacionalista del IRE, primero, y de su sucesor el IEEE, después, lo que ha llevado a que en países como España se dé con frecuencia una doble asociación profesional de los ingenieros, por un lado a los colegios nacionales correspondientes a las diversas titulaciones, para obtener cobertura profesional y los servicios que ofrecen, y por otro al IEEE, que sirve, principalmente, como vehículo de información tecnológica. Es una combinación que aporta una buena complementariedad y que puede ser muy útil a lo largo de una carrera profesional. Los estudiantes que ya sois miembros del IEEE habéis hecho una buena elección y los que todavía no lo sois animaros a explorar el web del IEEE (<http://www.ieee.org>) y conocer en más detalle lo que os puede ofrecer.



# INTRODUCCIÓN A LOS MODELOS DE TRÁFICO PARA REDES DE BANDA ANCHA

David Rincón Rivera

Estudiante de doctorado

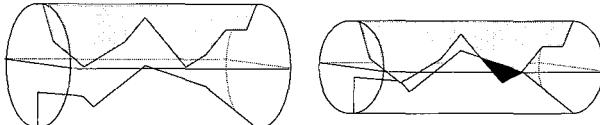
Departament de Matemàtica Aplicada i Telemàtica (UPC)

drincon@mat.upc.es, http://www-mat.upc.es/~drincon

## 1. INTRODUCCIÓN

La necesidad de disponer de redes que integren diversos servicios de telecomunicaciones motivó la aparición del concepto ISDN (RDSI, Red Digital de Servicios Integrados), que consiste en el uso de una única infraestructura para el transporte de datos, voz e imágenes. A mediados de la década de los 80, el CCITT (actual ITU-T, Unión Internacional de Telecomunicaciones) comenzó la estandarización de su modelo de RDSI de banda ancha (*Broadband ISDN*), que permitiría la transmisión de información a altas velocidades. Un punto importante en este proceso fue la elección del mecanismo de transferencia, que se podría definir como el conjunto de mecanismos de multiplexación y conmutación que usa la red. La elección recayó en ATM (*Asynchronous Transfer Mode*).

Las redes ATM permiten multiplexar eficientemente diversas fuentes. La unidad fundamental de transferencia es la celda ATM, de 53 bytes (5 de cabecera y 48 de información producida por las capas superiores). Mediante el concepto de *multiplexación estadística* de fuentes, se consigue una gran eficiencia en el uso de los recursos de la red [Led94]. Supongamos que disponemos de diversas fuentes que emiten a una tasa variable. Para la transmisión de dichas fuentes, las redes "tradicionales" reservan un ancho de banda igual a la tasa de pico, malgastando el ancho sobrante cuando la tasa era inferior. Esta ineficiencia es la que se trata de evitar en ATM, multiplexando las fuentes sobre enlaces con capacidad menor que la suma de tasas de pico. El precio a pagar es que, al ser un método estadístico, la posibilidad de que los picos coincidan no es nula. En este caso, se perderán celdas por *overflow*, con el consiguiente perjuicio en la calidad de servicio (QOS). En la figura 1 se



**Figura 1.** Izquierda: cada fuente dispone de un ancho fijo igual a su tasa de pico, malgastando recursos de transmisión. Derecha: al multiplexar estadísticamente las fuentes sobre un enlace de capacidad menor que la suma de los picos, se produce un conflicto (zona negra), pero se ahorran recursos (se ocupa menos ancho de banda).

pueden observar las diferencias entre la multiplexación en base a la tasa de pico y la multiplexación estadística.

Para minimizar estos efectos de pérdida y mantener la QOS tanto de las nuevas llamadas como de las que ya se están cursando, se necesita disponer de controles de admisión de conexiones (CAC) que decidan si se puede admitir una nueva llamada, así como de mecanismos de control de la congestión. Para ello son necesarias matemáticas que describan el comportamiento de las fuentes; es decir, necesitamos **modelos matemáticos de tráfico**. En este artículo se intentará ofrecer una visión de algunos de los modelos de tráfico propuestos para redes de banda ancha, sin ánimo de ser exhaustiva. Se hará un breve resumen sobre las escalas temporales que se distinguen en el tráfico ATM, se presentarán los modelos de tráfico clásicos y los nuevos modelos basados en la dependencia a largo plazo. El artículo finaliza con algunas reflexiones sobre el impacto de los nuevos modelos y las implicaciones de su aparición en el campo del modelado de fuentes.

## 2. ¿CÓMO SE COMPORTAN LAS FUENTES?

Para llegar a generar buenos modelos, es necesario estudiar y entender el comportamiento estadístico de las fuentes de información multimedia (voz, audio y vídeo) que serán transportadas por las redes ATM.

**Fuentes de datos:** los generadores de datos más habituales serán los servidores de Web, las transferencias de ficheros (FTP), las conexiones remotas (Telnet) y los *chats* (IRC). Es decir, casi todos los servicios presentes hoy por hoy en las redes TCP/IP (Internet). Estos servicios tienen comportamientos muy diferentes. El web, por ejemplo, se distingue por presentar una gran cantidad de transferencias, habitualmente de reducido tamaño (a menos que se esté navegando por una página con ficheros de sonido e imágenes grandes). FTP, por el contrario, presenta largas transferencias de información.

**Fuentes de audio:** lo más habitual es que la señal transmitida sea de voz, por lo que se aprovechan las características del habla humana, que se presenta a ráfagas (*talk spurts*) con silencios intercalados entre palabras y entre frases. Por ello lo más habitual es que los codificadores de voz (GSM, por ejemplo) incorporen detectores de silencios, durante los cuales no transmiten información.



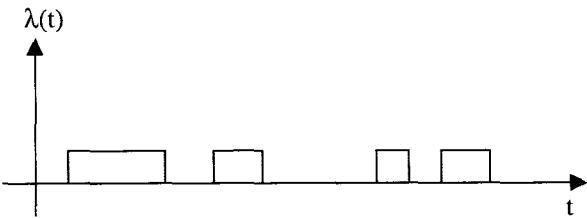


Figura 2. Possible traza de una fuente de voz.

**Fuentes de vídeo de tasa variable (VBR):** son muy dependientes del tipo de codificador usado. En general, presentan una cantidad no despreciable de correlaciones, a diferentes niveles (línea, bloque, cuadro, escena). Los codificadores más modernos, como los MPEG 1 y 2, aprovechan la alta redundancia espacial y temporal de las señales de vídeo para conseguir factores de compresión elevados. Por ello las señales de vídeo de estos codificadores son de tasa variable, y repiten un esquema en el que aparecen imágenes codificadas en el modo *intra* (que contienen toda la información del cuadro, denominadas  $I$ ), e imágenes codificadas a partir de las *intra* por predicción (hacia delante o hacia atrás, denominadas  $P$  y  $B$ ). Los cuadros tipo *intra* son los que provocan picos en la tasa de transmisión, ya que contienen mucha más información que los cuadros  $P$  y  $B$ .

### 3. MODELOS DE TRÁFICO PARA REDES ATM

Como hemos comentado, la principal característica de las redes ATM (y el principal quebradero de cabeza para los implementadores) es su capacidad de multiplexar estadísticamente las fuentes. Otra peculiaridad importante es la jerarquía temporal de tres niveles que presentan sus transmisiones [COST242].

Las redes ATM son **orientadas a conexión**. Eso quiere decir que cada vez que se quiere transferir información se debe establecer un circuito virtual desde el terminal origen al destino. Dentro de esta llamada o sesión, se producen ráfagas (*bursts*) en los que se transmiten celdas, seguidos de intervalos de silencio. A un nivel todavía menor, dentro de cada ráfaga las celdas no han de ser necesariamente transmitidas de modo uniforme. Por tanto, ATM presenta un comportamiento dife-

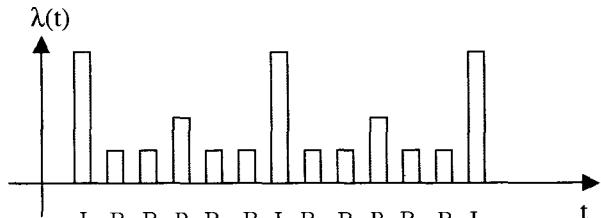


Figura 3. Possible traza de un codificador

renciado en cada una de las tres escalas de tiempo consideradas: llamada, ráfaga y celda (figura 4). Esta distinción es importante si queremos generar modelos de tráfico precisos.

Las escalas de celda y ráfaga son las que nos permiten dimensionar los buffers de los comutadores ATM, así como los mecanismos de admisión (CAC) y los de uso de parámetros (UPC). Por ello son los que han recibido más atención. Vamos a ver las características de cada uno de ellos [CasBlo]:

#### Nivel de celda

- El tráfico está compuesto por entidades discretas: las celdas.
- La mayoría de las fuentes ATM son localmente periódicas a esta escala.
- La aleatoricidad del tráfico agregado proviene de la independencia de las fases de los flujos (localmente periódicos), y no de las fluctuaciones de dichos flujos.
- Las fluctuaciones a escala de celda deben ser suavizadas, por lo que se presenta un problema de dimensionado de buffers.
- La distribución de la longitud de las colas cuando el sistema trabaja próximo a su capacidad máxima es crucial para el dimensionado.

#### Nivel de ráfaga (burst)

- Sólo es relevante para fuentes de tasa variable (VBR, variable bit rate), ya que si la tasa es constante (CBR, constant bit rate) no se producen ráfagas.
- Los posibles problemas provienen de la posibilidad de que el tráfico agregado exceda la capacidad de salida del sistema.
- No es importante conocer el momento exacto de llegada de las celdas.
- Esencialmente, el sistema se comporta como un sistema con pérdidas. Los *overflows* se pueden evitar mediante el uso de mecanismos de control de admisión.
- La prevención de pérdidas provocadas por ráfagas implica forzar una tasa media de entrada muy reducida para el tráfico de alta variabilidad (*bursty traffic*).

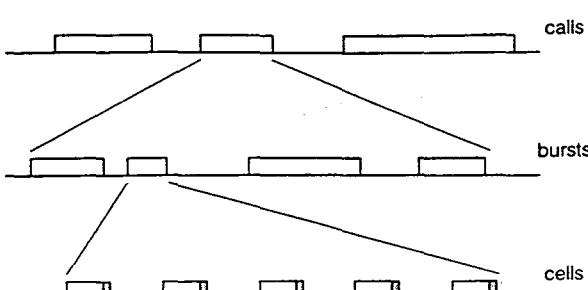


Figura 4. Escalas de tiempo en ATM  
[COST242].

Ya que vamos a presentar diferentes modelos de tráfico, sería interesante disponer de algún criterio que permitiera valorar la bondad de cada modelo. Esta valoración vendría a medir lo “próximo” que es el modelo a una fuente real de tráfico. Los parámetros de bondad de un modelo son la capacidad de capturar la autocorrelación de las fuentes, las distribuciones marginales (las *colas* de

la distribución), y la precisión en el cálculo de retardos y probabilidad de pérdidas [Ada97].

Los modelos de tráfico pueden ser **estacionarios** o **no estacionarios**. Los estacionarios se dividen en dependientes a corto y largo plazo (**short range dependent, long range dependent**). Entre los primeros encontramos los procesos de Markov y los modelos regresivos. Estos modelos presentan correlación sólo en intervalos cortos de tiempo. En cambio, los modelos *long range dependent*, como el F-ARIMA o el Movimiento Browniano Fractal presentan correlaciones incluso en intervalos de tiempo grandes.

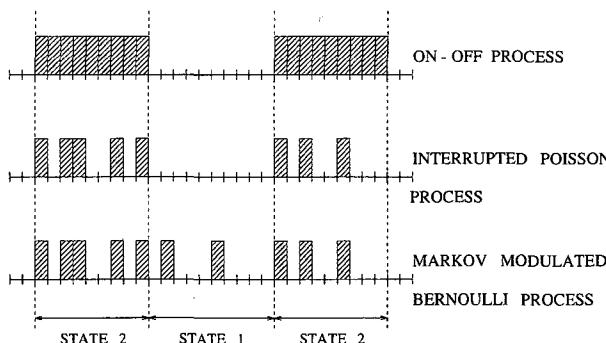
## 4. LOS MODELOS CLÁSICOS

Vamos a denominar como *clásicos* a toda una familia de modelos cuya principal característica es que no contemplan la posibilidad de que las trazas de tráfico procedentes de las fuentes presenten dependencia a largo plazo.

¿Qué significa esto y qué implicaciones tiene de cara al modelado? La autocorrelación de una serie temporal (en nuestro caso, la serie temporal es la tasa o *throughput* de la fuente medida en bits/segundo o equivalentemente, en celdas/segundo) indica cuál es la dependencia que existe entre los datos de la serie separados en una cierta cantidad de muestras. Es decir, qué grado de similitud existe entre las muestras analizadas en un instante  $t_0$  y las muestras emitidas en  $t_1 = t_0 + \tau$ . Si la autocorrelación de la serie generada por una fuente decae rápidamente con la separación  $\tau$ , las tasas emitidas por la fuente tendrán poco que ver con los emitidos  $\tau$  unidades antes. Si, en cambio, la autocorrelación decae lentamente, existirá una dependencia a largo plazo.

### 4.1 Modelos para fuentes de datos

Tradicionalmente, se ha asumido que en las redes de conmutación de paquetes se daban las condiciones necesarias para suponer que la generación de celdas sigue un proceso de Poisson o de Bernouilli. Sin embargo, como ya se ha mencionado antes al mencionar diversos servicios de Internet, cada aplicación presenta una tasa y



**Figura 5.** Modelos de tráfico a tiempo discreto para fuentes de voz y datos [Led94].

una distribución de generación de celdas diferente, que puede ir desde transmisiones esporádicas y cortas hasta largas transferencias de información (*bulk transfers*).

También es un factor a tener en cuenta el hecho de que los paquetes de las redes de área local (LAN) son mucho más grandes que las celdas ATM, lo que provoca la aparición de pequeñas ráfagas de celdas ATM cada vez que una trama LAN se introduzca en la red. Por ello, los modelos más utilizados son los procesos on-off markovianos, y los interrumpidos de Poisson o Bernouilli (figura 5).

### 4.2 Modelos para fuentes de voz

#### Modelos on-off

Los procesos que más se aproximan al fenómeno *talk spurt* son los que se basan en una cadena de Markov de dos estados (*on* y *off*). Estos modelos, ampliamente utilizados, intentan describir una fuente que emite información a ráfagas, de manera que en el estado *on* se generan paquetes de voz y en el estado *off* hay silencio. Dependiendo de la distribución del proceso de llegadas cuando la fuente se encuentra en estado activo, se pueden considerar dos tipos de procesos: Bernouilli Interrumpido o Poisson Interrumpido.

Si suponemos que las llegadas siguen se producen según Poisson, el tiempo de estancia en los estados está exponencialmente distribuido con medias  $\alpha^{-1}$  y  $\beta^{-1}$ , respectivamente. Los procesos de Poisson interrumpidos son una particularización del caso anterior, donde el estado activo se corresponde con llegadas distribuidas exponencialmente con media  $\lambda$ . Como ejemplo, en [Led94] se presenta el caso de un codificador ADPCM a 32 Kbit/s, con tiempos de paquetización de 12 ms.,  $\alpha^{-1} \approx 650$  ms. y  $\beta^{-1} \approx 350$  ms. La agregación de fuentes de voz se puede modelar como un proceso de nacimiento y muerte con  $N+1$  estados, que resulta de encadenar  $N$  procesos on-off.

#### Procesos de Poisson modulados por cadenas de Markov (MMPP, Markov Modulated Poisson Processes)

Empecemos presentando los procesos de Markov modulados, también llamados procesos doblemente estocásticos, que usan una cadena de Markov que define (según el estado en el que nos encontremos) la distribución de probabilidad del tráfico [Sch96]. Es decir, la cadena “modula” el proceso de generación. Un MMPP usa un proceso de Poisson como proceso modulado, es decir, el que cambia el ritmo de generación de llegadas según el estado en el que nos encontremos (a un estado  $s_k$  le corresponde un proceso de Poisson con media  $\lambda_k$ ).

Mediante los MMPP, se puede conseguir que el modelo de tráfico sea analíticamente abordable. Para ello, se cuantiza la tasa de llegada en diferentes niveles, que se corresponden con los estados de la cadena de Markov. Las probabilidades de transición del estado  $i$  al  $j$ ,  $q_{ij}$ , se pueden estimar a partir de los datos empíricos



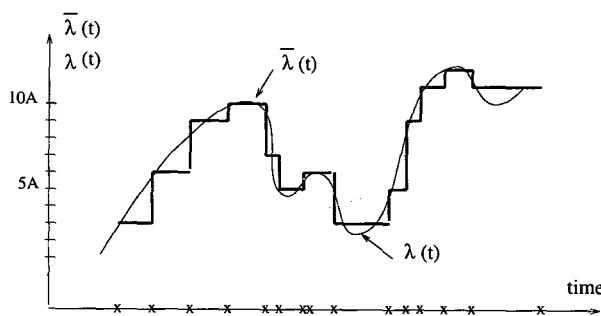
mediante el cálculo de la fracción de la cantidad de veces que se produce dicha transición, respecto a la cantidad total de transiciones. Es fácil observar que un MMPP de  $M+1$  estados se puede formar a partir de la superposición de  $M$  procesos IPP independientes e idénticamente distribuidos. Los MMPP se han aplicado en algunas ocasiones para modelar mezclas de tráfico de voz y datos. Para ello, consideramos que el MMPP modela el tráfico de voz, de manera que el estado  $k$  se corresponde con un tráfico de voz de tasa  $\lambda_k$ , mientras que el tráfico de datos se considera poissoniano con media  $\lambda_d$ . Así, el tráfico total generado es  $\lambda_k + \lambda_d$ .

### 4.3 Modelos para fuentes de vídeo

#### Modelos de fluidos modulados por cadenas de Markov (Markov Modulated Fluid Models)

Los modelos de fluidos se caracterizan por modelar el tráfico como un flujo continuo, siendo especialmente indicados cuando las unidades de tráfico (paquetes o celdas) son muy pequeñas comparadas con el tráfico total. De cara a la simulación, estos modelos son mucho más tratables computacionalmente que los modelos que distinguen cada celda. En los modelos de fluido modulados por Markov, el estado de la cadena determina la tasa del fluido (a un estado  $s_k$  le corresponde una tasa constante  $\lambda_k$ ). Este modelo es usado en fuentes de vídeo del tipo VBR (tasa variable).

En [Mag88], la tasa continua es cuantizada a un número finito de niveles discretos (equivalente al número de estados de la cadena de Markov) y es muestreada temporalmente en puntos aleatorios escogidos según una distribución de Poisson. En este modelo, la tasa en el estado  $i$  es  $iA$ , donde  $A$  se define como el paso de cuantización (figura 6). El modelo no captura exactamente la autocorrelación de los datos empíricos, ya que sólo se permiten transiciones entre estados adyacentes, pero si tuviéramos que tomar en consideración todas las transiciones (necesarias cuando se produce un cambio brusco en la tasa, de más de un nivel), la complejidad de cálculo se haría prohibitiva.



**Figura 6.** Muestreo de Poisson y discretización de la tasa de la fuente. Se permiten saltos de más de un nivel [Led94].

Para mejorar la respuesta del modelo ante los saltos bruscos, en [Mag88] se propone una extensión

bidimensional de la cadena, de manera que existen dos pasos de cuantificación (uno grande,  $A_h$ , y otro pequeño,  $A_l$ ). Para identificar los estados del modelo, se definen unos índices  $i$  y  $j$ ,  $0 \leq i \leq M$ ,  $0 \leq j \leq N$ , de manera que la tasa total generada es  $jA_h + iA_l$ . En [Sen89] se analiza el comportamiento del modelo. Con  $M=1$ , todavía es analíticamente tratable, y presenta una autocorrelación que decae exponencialmente. A medida que se introducen nuevos niveles, el coste computacional se dispara.

#### Modelos regresivos

Los modelos regresivos se caracterizan por predecir la siguiente variable aleatoria de la secuencia, a partir de los valores anteriores, que se encuentran dentro de una "ventana móvil", y de ruido blanco. A continuación se presentan algunos modelos regresivos.

##### - Modelos autoregresivos (AR)

El modelo autoregresivo de orden  $p$ , denominado AR( $p$ ), tiene la forma:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

donde  $\varepsilon_t$  es ruido blanco,  $\phi_i$  son números reales (coeficientes autoregresivos) y  $X_t$  son los valores de la secuencia de variables aleatorias. Si  $\varepsilon_t$  es ruido blanco gaussiano aditivo de varianza  $\sigma^2_\varepsilon$ , las  $X_t$  son variables aleatorias con distribución normal. Si definimos un operador de retardo  $B$  como  $X_{t-1} = BX_t$ , y el polinomio  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ , entonces podemos expresar el proceso AR( $p$ ) como

$$\phi(B) X_t = \varepsilon_t$$

Se puede demostrar que la autocorrelación del proceso AR( $p$ ) será, en general, una exponencial, lo que indica que los modelos autoregresivos presentan dependencia a corto plazo, que como veremos más adelante es una característica que puede marcar su campo de aplicación [Ada97]. Los modelos AR( $p$ ) son especialmente adecuados en el caso de videoconferencia, ya que dicho servicio genera imágenes con poca variabilidad (no suele haber cambios de escena). Aunque es fácil estimar los parámetros del modelo AR, el hecho de que la autocorrelación decaiga exponencialmente no permite capturar funciones de autocorrelación que lo hagan a un ritmo menor, como es el caso de las señales de televisión digital a tasa variable (donde los cambios de escena y los picos de tasa asociados son abundantes).

##### - Modelos autoregresivos de media móvil (ARMA)

Los modelos autoregresivos de media móvil de orden  $p$  y  $q$ , denominados ARMA( $p, q$ ), tienen la forma:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

o alternativamente,

$$\phi(B) X_t = \theta(B) \varepsilon_t \quad \text{donde } \theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$$

Esto es equivalente a filtrar ruido blanco (el proceso  $\varepsilon_t$ ) a través de un filtro lineal y causal (desplazado temporalmente), con  $p$  polos y  $q$  ceros, de la forma:

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{1 - \sum_{k=0}^q \theta_k z^{-k}}{1 - \sum_{k=0}^p \phi_k z^{-k}}$$

Se demuestra que la autocorrelación de los procesos ARMA(p,q) decae exponencialmente, lo que implica que este modelo presenta dependencia a corto plazo.

Los modelos ARMA son capaces de capturar mejor que los AR los picos debidos a los cambios de escena, pero desde el punto de vista práctico, la estimación de los coeficientes  $\phi_k$  implica la resolución de un sistema de ecuaciones no lineales, hecho que complica enormemente su obtención.

#### - Modelos autoregresivos de media móvil integrada (ARIMA)

Los modelos autoregresivos de media móvil integrada de orden p, d y q, denominados ARIMA(p, d, q), son una extensión de los ARMA (p,q). Se obtienen cuando se permite que el polinomio  $\phi(B)$  tenga d raíces unitarias, mientras que el resto permanecen fuera del círculo unidad. La forma general es

$$\varphi(B) \nabla^d X_t = \theta(B) \varepsilon_t$$

donde  $\nabla$  (operador diferencia) se define como  $\nabla X_t = X_t - X_{t-1}$ , y  $\varphi(B)$ ,  $\theta(B)$  son polinomios en B. Los ARIMA(p, d, q) se usan para modelar series temporales homogéneas y no estacionarias.

## 5. NUEVOS MODELOS CON DEPENDENCIA A LARGO PLAZO

Hemos visto, hasta ahora, modelos de tráfico que tenían en común que la autocorrelación decaía exponencialmente; es decir, que la dependencia de la señal consigo misma sólo se da en un margen corto de tiempo. Sin embargo, medidas realizadas sobre tráfico real han llevado a cuestionar la validez de estos modelos [Lel94], y se ha detectado que la autocorrelación decaea a un ritmo menor. Por ello se han adoptado nuevos enfoques y modelos. Vamos a introducir las herramientas matemáticas necesarias para entenderlos.

Definamos  $\{X_t\}$ ,  $t = 0, 1, 2, \dots$  como un proceso estocástico estacionario en sentido amplio, es decir, un proceso con:

- una media estacionaria  $\mu = E[X_t]$
- una varianza estacionaria y finita,  $v = E[(X_t - \mu)^2]$
- una función de autocovarianza estacionaria,  $\gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$ , que sólo depende de k y no de t.

Nótese que  $v = \gamma_0$ . Si definimos la autocorrelación de  $\{X_t\}$  en el instante k como  $\rho_k$ , tenemos que por definición,  $\rho_k = \gamma_k / \gamma_0$ .

Para toda m, definimos  $\{X_j^{(m)}\}$  como una nueva serie obtenida calculando la media de la serie original  $\{X_t\}$  sobre bloques de longitud m, sin superposición:

$$X_j^{(m)} = (1/m)(X_{jm+m+1} + \dots + X_{jm})$$

La varianza de  $\{X_j^{(m)}\}$ ,  $v_m$ , es

$$v_m = E[(1/m)(X_{jm+m+1} + \dots + X_{jm})]^2$$

En el caso de que el proceso sea ruido blanco, las variables aleatorias estarán incorreladas y  $\rho_k = 0$  para  $k > 0$ , y  $v_m = v m^{-1}$ . Si m (longitud de los bloques a promediar) es grande, se puede aproximar la expresión de la autocorrelación por:

$$v_m = V \left[ 2 \sum_{k=1}^m \rho_k \right] m^{-1}$$

Si  $\rho_k \rightarrow 0$  y  $\sum_k \rho_k < \infty$ , la varianza decae asintóticamente a cero de manera proporcional a  $m^{-1}$ , es decir,  $v_m = cm^{-1}$  (donde c es una constante). Procesos como los ARMA o Markov verifican esta propiedad.

Las medidas de tráfico real [Lel94, Ber95] han demostrado que la varianza  $v_m$  decae a un ritmo menor que  $m^{-1}$ . Por ello se introduce un factor de corrección en el modelo, haciendo la varianza proporcional a  $m^{-\alpha}$  con  $\alpha \in (0, 1)$ . Esto implica que

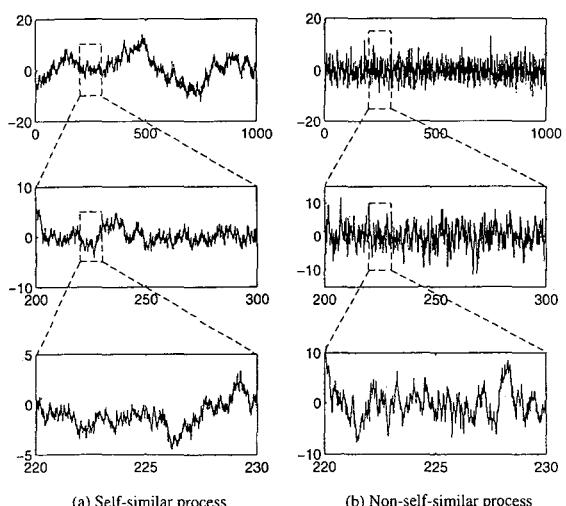
$$\sum_{k=-\infty}^{\infty} \rho_k \rightarrow \infty$$

Por tanto, la autocorrelación decae lentamente y no es sumable. Un ejemplo de este tipo de funciones es:

$$\rho_k = C \rho |k|^{-\alpha} \quad \text{para } k \text{ elevada.}$$

### 5.1 Autosimilitud

Es un concepto muy relacionado con la idea de los fractales, introducida por Mandelbrot. Se dice que un proceso  $\{X_t\}$  es exactamente autosimilar (*self-similar*) si la estadística de la autocorrelación  $\rho_k^{(m)}$  se mantiene en diferentes escalas de tiempo. El proceso se denomina



**Figura 7.** Comparación de un proceso autosimilar y otro no autosimilar. Se puede apreciar que el primero mantiene su apariencia mientras que el segundo tiende a comportarse como ruido blanco cuando está agregado (arriba) [Sta98].



asintóticamente similar si la condición sólo se cumple para  $m \rightarrow \infty$  y  $k \rightarrow \infty$ . Esto se expresa matemáticamente de la siguiente manera:

$$\text{Dist}\{X_{at}\} = \text{Dist}\{a^H X_t\}$$

H es el parámetro de Hurst, que da una medida de “lo autosimilar” que es el proceso. Puede parecer que la definición de autosimilitud es poco rigurosa, al basarse en la apariencia de las gráficas de la función de autocorrelación. Es lo que se llama la “prueba visual” de la autosimilitud (figura 7). Hurst hizo unos estudios hidrológicos que le llevaron a descubrir que muchas series naturales presentan autosimilitud, es decir, que el hecho de cambiar la escala de observación del proceso agregado no cambia la estadística.

## 5.2 Dependencia a corto (*short range*) y largo (*long range*) plazo

Se dice que un proceso  $\{X_t\}$  presenta dependencia a corto plazo si su autocorrelación es sumable:

$$\sum_{k=-\infty}^{\infty} \rho_k < \infty$$

Equivalentemente, la varianza  $v_m$  decrece proporcionalmente a  $m^{-1}$ , la densidad espectral de potencia tiene un valor finito en el cero, y el proceso promediado  $\{X_j^{(m)}\}$  tiende a comportarse como ruido blanco a medida que  $m \rightarrow \infty$ . Los procesos con autocorrelaciones que decaen exponencialmente se denominan dependientes a corto plazo (*short range dependent*).

Se dice que un proceso  $\{X_t\}$  muestra dependencia a largo plazo (*long range dependence*) si su autocorrelación no es sumable:

Equivalentemente,  $v_m$  decrece a un ritmo menor que

$$\sum_{k=-\infty}^{\infty} \rho_k \rightarrow \infty$$

$m^{-1}$ , la densidad espectral de potencia tiene una singularidad en cero y el proceso promediado tiene la misma estadística que el proceso inicial (autosimilitud). Los procesos con  $\rho_k \sim Ck^{-\alpha}$  (para  $k$  grande), presentan dependencia a largo plazo.

A continuación veremos dos casos de procesos autosimilares: uno exactamente autosimilar (Fractional Gaussian Noise) y otro asintóticamente autosimilar (Fractional ARIMA).

### Fractional ARIMA

Los procesos F-ARIMA (*Fractional Autoregressive Integrated Moving Average*) dependen de tres parámetros:  $p$ ,  $d$  y  $q$ , como ya vimos en los ARIMA y se definen como una extensión de éstos:

$$\phi(B) \nabla^d X_t = \theta(B) \varepsilon_t \quad \text{con } 0 < d < 0.5.$$

Expresando el operador  $\nabla^d$  en forma binomial, para el caso F-ARIMA (0,  $d$ , 0) con  $0 < d < 0.5$  ( $d$  puede

tomar valores no enteros, de aquí el adjetivo *fraccional*) y  $k \rightarrow \infty$ , se obtiene que  $\rho_k \sim k^{2d-1}$ . De aquí obtenemos  $d = (1-\alpha)/2 = H-0.5$ . Los procesos F-ARIMA pueden modelar dependencia a largo y corto plazo, según el valor de  $d$ . Por eso se les considera flexibles y potentes, especialmente en el caso del modelado de fuentes de video VBR. Desgraciadamente, la estimación de los parámetros  $p$ ,  $d$  y  $q$ , y de los coeficientes es costosa.

El tema de la estimación de  $d$  a partir de datos reales es de gran importancia. Uno de los métodos más eficientes es el de la gráfica varianza-tiempo. En este método,  $v_m = \text{Var}\{X^{(m)}\}$  es representada respecto a  $m$ , en escalas logarítmicas. La pendiente asintótica es  $-\alpha$ , y de ahí se puede estimar  $d$ .

### Fractional Brownian Motion

El movimiento Browniano es un proceso estocástico  $\{B_t\}$ , para  $t \geq 0$ , con las siguientes características:

- los incrementos  $B_{t+10} - B_{t_0}$  siguen una distribución normal de media 0 y varianza  $\sigma^2 t$ .
- los incrementos en los intervalos temporales no solapados  $[t_1, t_2]$  y  $[t_3, t_4]$ ,  $B_{t_4} - B_{t_3}$  y  $B_{t_2} - B_{t_1}$ , son variables aleatorias independientes.
- $B_0 = 0$  y  $B_t$  es continuo en  $t = 0$ .

El Movimiento Browniano fraccional,  $\{fB_t\}$  es un proceso autosimilar gaussiano con  $H \in [0.5, 1]$ . La diferencia respecto al Movimiento Browniano es que los incrementos tienen una varianza igual a  $\sigma^2 t^{2H}$ . Se demuestra que, para el caso discreto, la autocorrelación normalizada de la secuencia de incrementos (llamada también ruido gaussiano fraccional) es proporcional a  $k^{2H-2}$  cuando  $k \rightarrow \infty$  y, por tanto, presenta dependencia a largo plazo.

El parámetro  $H$  se puede estimar, como en el caso de los F-ARIMA, mediante la gráfica varianza-tiempo.

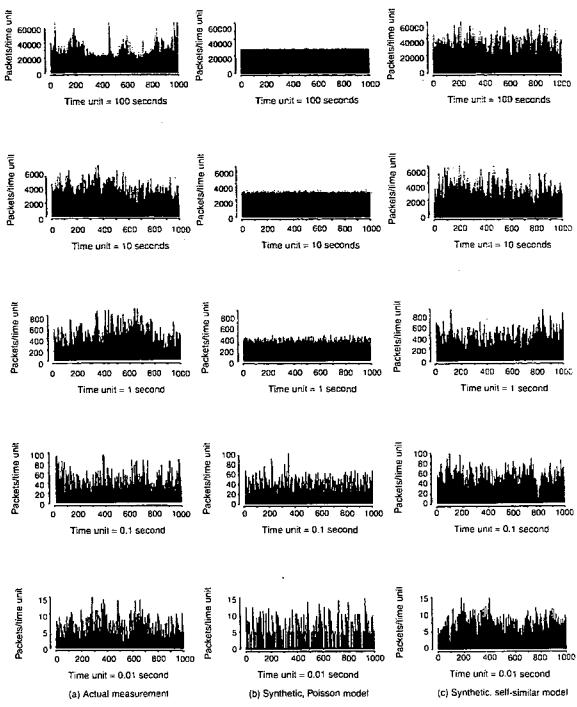
## 6. VALORACIÓN DE LOS MODELOS AUTOSIMILARES

### 6.1 Los estudios pioneros

La voz de alarma llegó en 1994, cuando un equipo de investigadores encabezados por Taqqu, Leland, Willinger y Wilson analizó trazas provenientes de una red Ethernet y constataron que el tráfico seguía un comportamiento autosimilar, en contradicción con los modelos clásicos markovianos (figura 8).

Los mismos autores efectuaron un análisis a nivel de fuentes aisladas, y concluyeron que el tráfico Ethernet se comporta como la agregación de muchas fuentes on-off con distribuciones de alta variabilidad (Pareto, *heavy tails*) [Will97].

Estas pruebas de que el tráfico real no coincide con el predicho por los modelos con dependencia a corto plazo provocaron la sensación de que todo el trabajo



**Figura 8.** Constatación de la autosimilitud del tráfico Ethernet. A la izquierda, las trazas capturadas. En el centro, un modelo clásico. A la derecha, una traza sintetizada a partir de un modelo autosimilar [Will97].

realizado hasta la fecha en el campo del modelado de tráfico se venía abajo. Los estudios sobre dimensionado de buffers, capacidades equivalentes, o control de acceso, parecían quedar invalidados por los nuevos modelos autosimilares. Si bien no se llegó a este extremo, y todavía se está discutiendo si es necesario abandonar los modelos clásicos en todos los casos, sí es cierto que se ha revitalizado la investigación sobre los modelos de fuentes. A continuación expondremos algunos de los resultados obtenidos a partir de los nuevos modelos.

## 6.2 Análisis de diferentes clases de tráfico

Desde los primeros artículos aparecidos en 1994, se ha demostrado que en muchas situaciones el tráfico se puede modelar con procesos autosimilares.

• **HTTP** (World Wide Web). Se puede modelar como la superposición de fuentes on/off de tipo Pareto, con varianza infinita. La justificación se basa en la distribución de los tamaños de los ficheros HTML y multimedia [Sta98].

• **SS7** (señalización). Se demuestra que si se utilizan los modelos poissonianos, el tráfico de señalización se subestima, ya que dicho tráfico presenta distribuciones con colas del tipo *heavy tail*, que indican un fuerte componente de dependencia a largo plazo [Sta98].

• **TCP, FTP, Telnet**. Paxson [Pax95] presenta un estudio exhaustivo sobre los servicios de Internet. Las principales conclusiones son:

– Los modelos poissonianos subestiman la variabilidad (*burstiness*) de las transferencias de paquetes TCP.

– Telnet es bien representado por Poisson a nivel de conexión, pero no a nivel de paquete, donde se detecta autosimilitud.

– Las transferencias en bloque de FTP se desvían mucho del modelo poissoniano, excepto en el nivel de sesión. La distribución del número de bytes en cada paquete sigue un modelo *heavy tailed* (un ejemplo es la distribución de Pareto).

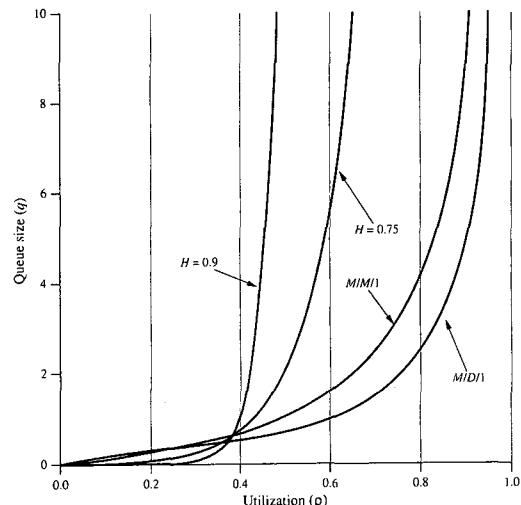
• **Vídeo VBR (Variable Bit Rate)**. Es uno de los temas más estudiados, dada la importancia que este tipo de tráfico tendrá en las redes de banda ancha (es el que más recursos de transmisión necesita). Garrett [Garr94] y Beran et al. [Ber95] presentan trabajos en los que se analizan diferentes películas y programas de televisión, y sus conclusiones son las siguientes:

– Se demuestra que la traza es autosimilar (presenta variaciones lentas de la componente continua) y que la distribución es *heavy-tailed*.

– Se destaca la presencia de componentes con dependencia a corto plazo, no despreciables (por eso son tan interesantes los modelos F-ARIMA, ya que capturan los dos comportamientos).

– La dependencia a largo plazo constituye una característica inherente a la estructura de las películas (planteamiento-nudo-desenlace, con diferentes requerimientos de ancho de banda para cada fase) y a la jerarquía plano-escena-secuencia (las diferentes escalas en las que se da la autosimilitud).

• **Buffers de conmutadores ATM**. Dado que las redes de banda ancha se basarán en la conmutación de celdas ATM, es de vital importancia prever el comportamiento que tendrán los *buffers* de los conmutadores cuando se les inyecte tráfico autosimilar. Es uno de los campos en los que se ha generado más polémica, ya que mientras unos autores destacan que los modelos poissonianos subestiman la varianza del tráfico y pueden llevar a desbordamientos y pérdidas de celdas, otros autores argumentan que en la escala de tiempo a la que trabajan los conmutadores la autosimilitud del tráfico no tiene consecuencias graves.



**Figura 9.** Tamaño de cola. Se presentan curvas de tráfico autosimilar ( $H=0.9, H=0.75$ ) y los modelos clásicos ( $M/M/1$  y  $M/D/1$ ) [Nor94].



– Norros [Nor94, COST242] analiza el comportamiento de un buffer bajo tráfico autosimilar, inyectando un proceso FBM a un buffer infinito con tiempos de servicio constantes. Concluye que se encuentran grandes discrepancias con los modelos M/M/1 i M/D/1, que se necesitan buffers mucho más grandes que los previstos cuando se usan dichos modelos (figura 9), y que también hay diferencias en el retardo de las celdas.

– Heyman et al. [Hey96] argumenta que la LRD puede tener una influencia mínima en el dimensionado de los *buffers* si el factor de Hurst es reducido y la parte SRD es importante. Asimismo, se cuestiona la influencia de la LRD en el cálculo de las pérdidas y se destaca el “*truncating effect of finite buffers*”, que disminuye los efectos debidos a la dependencia a largo plazo (ya que se pierden celdas y la dependencia se reduce).

### 6.3 Reflexiones sobre los modelos autosimilares

Una vez revisado el trabajo realizado en este campo, y después de constatar que hay aspectos todavía abiertos, aparecen lagunas y surgen muchas preguntas:

– ¿Es realmente autosimilar el tráfico de las redes de banda ancha? Parece una buena aproximación, pero podría ser que el tráfico fuera simplemente no estacionario, y no necesariamente autosimilar (que es uno de los casos particulares de no estacionariedad).

– ¿La autosimilitud es debida a las fuentes o a su agregación? Depende del servicio. Por ejemplo, el tráfico LAN es autosimilar sólo cuando hay agregación, mientras que en el caso del vídeo VBR el efecto aparece en el nivel de fuentes individuales.

– ¿Hasta qué punto es la agregación y la propia dinámica de la red la que provoca la autosimilitud? Todavía está por estudiar a fondo la influencia de los protocolos en el comportamiento del tráfico. Por ejemplo, el protocolo TCP introduce una serie de mecanismos de control (ventana adaptativa, lucha por el ancho de banda, control de congestión). ¿Qué influencia podría tener ATM y su dinámica?

– ¿Qué papel juega la presencia *residual* de dependencia a corto plazo (SRD) en las fuentes de vídeo VBR?. ¿Cómo incluir sus efectos? ¿Es suficiente con los modelos F-ARIMA, o es necesaria una nueva familia de procesos que permitan una estimación de parámetros más fácil?

– Si el tráfico es realmente autosimilar, ¿son los modelos propuestos una buena solución para modelarlo? Ciertamente, son mejores que los modelos poissonianos, y proporcionan una explicación física del fenómeno (al menos en el caso del tráfico Ethernet, a través de la agregación de terminales que se comportan según modelos on-off con distribuciones de Pareto). Pero incluso los modelos autosimilares no son capaces de capturar al 100% el comportamiento de las trazas reales, por la presencia de SRD y periodicidades (como el caso de las fuentes MPEG).

– ¿Son realmente importantes los efectos autosimilares en la práctica? Es un tema abierto, y se

ofrecen conclusiones contradictorias. No está claro que a las escalas temporales en que trabajan los buffers aparezca LRD, aunque se habla de importantes errores de subestimación. Stallings [Sta98] ha propuesto el estudio a dos niveles: **aplicación** (autosimilitud inherente a la fuente) y **red** (influencia de la dinámica de los protocolos como el TCP). A nivel de aplicación, la autosimilitud tendría influencia en el control de acceso (CAC) y la asignación de recursos de transmisión, mientras que a nivel de red, influiría en la congestión, el dimensionado de los buffers, etc. Tampoco queda claro qué influencia tiene cada nivel.

Vemos, por tanto, que pese a que parece bastante claro que los modelos autosimilares son prometedores, hace falta mucho más trabajo en este campo para llegar a conclusiones sólidas y útiles.

## 7. REFERENCIAS

- [Ada97] ADAS, A. “Traffic Models in Broadband Networks”, IEEE Communications Magazine, Vol. 35, No. 7, July 1997.
- [Ber95] BERAN, J., SHERMAN, R., TAQQU, M., WILLINGER, W., “Long-range Dependence in Variable-bit-rate Video Traffic.”, IEEE Transactions on Communications, February 1995
- [CasBlo] CASALS, O. AND BLONDIA, C., “Traffic Management in ATM Networks: an overview”, pendiente de publicación en la revista Computer Networks & ISDN Systems.
- [COST242] “Methods for the performance evaluation and design of broadband multiservice networks. The COST 242 Final Report. Part III. Traffic models and queueing analysis”, COST 242 Management Committee, 1996.
- [Led94] LEDUC, J.-P. “Digital Moving Pictures – Coding and Transmission on ATM Networks”, Elsevier, 1994.
- [LeL94] LELAND, W., TAQQU, M., WILLINGER, W., AND WILSON, D., “On the Self-similar Nature of Ethernet Traffic (Extended Version).”, IEEE/ACM Transactions on Networking, February 1994.
- [Mag88] B. MAGLARIS ET AL., “Performance Models of Statistical Multiplexing in Packet Video Communications”, IEEE Trans. Communications, vol. 36, July 1988.
- [Nor94] NORROS, I., “A Storage Model with Self-similar Input”, Queueing Systems, Vol. 16, 1994.
- [Hey96] HEYMAN, D. AND LAKSHMAN, T., “What Are the Implications of Long-Range Dependence for VBR-Video Traffic Engineering?”, IEEE/ACM Transactions on Networking, vol. 4, no.3, June 1996
- [Pax95] PAXSON, V., AND FLOYD, S., “Wide Area Traffic: The failure of Poisson Modelling”, IEEE/ACM Transactions on Networking, June 1995.
- [Sen89] P. SEN ET AL., “Models for Packet Switching of Variable-Bit-Rate Video Sources”, IEEE JSAC, vol. 7, no. 5, 1989.
- [Sch96] M. SCHWARTZ, Broadband Integrated Networks, Prentice Hall, 1996.
- [Sta98] STALLINGS, W., High-speed networks. TCP/IP and ATM design principles, Prentice Hall, 1998.
- [Will97] WILLINGER, W., TAQQU, M., SHERMAN, R., WILSON, D., “Self-similarity through High Variability:Statistical Analysis of Ethernet LAN Traffic at the Source Level.”, IEEE/ACM Transactions on Networking, February 1997.



# PARALELIZACIÓN DE UN ALGORITMO DE CÁLCULO DE LA RESPUESTA IMPULSIONAL DE UN CANAL IR DIFUSO EN INTERIORES

Elsa Macías<sup>1</sup>, Alvaro Suárez<sup>1</sup>, Rafael Pérez<sup>2</sup>, F.J. López Hernández<sup>3</sup>

<sup>1</sup>GAC (Grupo de Arquitectura de Computadores), Dpto. de Electrónica, Telemática y Automática

<sup>2</sup>GPDS (Grupo de Procesado Digital de la Señal), Dpto. de Señales y Comunicaciones

Universidad de Las Palmas de G.C., e-mail: {alvaro, rperez}@cic.teleco.ulpgc.es

<sup>3</sup>Dpto. de Tecnología Fotónica, Universidad Politécnica de Madrid, email: dxm@tfo.upm.es

En los últimos años ha habido una explosión de la utilización de los sistemas de comunicación no cableados. En este tipo de sistemas, los dispositivos de comunicación pueden ser desde simples emisores y receptores hasta sistemas complejos basados en computadores. Las redes de área local basadas en computadores personales portátiles son un hecho tangible. Una forma de conectar estos sistemas es utilizar comunicación infrarroja. La simulación de la respuesta impulsional de un canal de infrarrojos en un laboratorio es costosa, por ello se recurre al uso de computadores. Sin embargo, debido a la gran cantidad de tiempo necesario para obtener resultados, en algunos casos el empleo de los computadores se puede hacer inviable. Una forma de resolver este problema es ejecutar el programa de simulación en varios procesadores, en paralelo. En este artículo presentamos la simulación en un sistema multicomputador para poder obtener respuestas en un tiempo razonable.

## 1. INTRODUCCIÓN

En la actualidad existen sistemas que integran equipos terminales conectados entre ellos mediante un sistema de comunicación. En algunos casos estos equipos terminales pueden ser simples equipos emisores o receptores de señal y en otros casos pueden estar empotrados en computadores. En cualquier caso, un elemento importante es el tipo de interconexión entre equipos a usar. El método tradicional de interconexión es el cableado. Sin embargo, el uso masivo de computadores portátiles y la comunicación de grandes cantidades de datos, han producido la aparición de las redes de área local no cableadas (denominadas **WLAN**, de las siglas en inglés Wireless Local Area Network). La portadora empleada en el diseño de estas redes puede ser de radiofrecuencia (**RF-WLAN**, del inglés Radio Frequency-Wireless Local Area Network, o redes de área local no guiadas por radiofrecuencia), o por infrarrojo (denominadas **IR-WLAN**, de las siglas en inglés Infrared-Wireless Local Area Network, o redes de área local no guiadas por infrarrojo).

Los sistemas de comunicación no cableados presentan varias ventajas sobre los interconectados por cable. Entre éstas están: la flexibilidad para recolocar equipos terminales y su facilidad de movilidad dentro del entorno en el que se puede emitir y recibir señales [1].

Los *enlaces por infrarrojo* proporcionan comunicaciones no cableadas de alta velocidad sobre distancias relativamente cortas. En la actualidad estas comunicaciones están recibiendo un interés creciente debido a que presentan algunas ventajas sobre las *comunicaciones vía radio*. Ejemplos de estas ventajas son: no necesitan licencia gubernativa, son inmunes a las interferencias producidas por el ruido electromagnético y los componentes empleados (emisores y receptores de infrarrojos que operan a alta velocidad), son baratos y se pueden implantar en hardware de forma sencilla. Por otro lado, las comunicaciones vía infrarrojo están confinadas dentro del entorno cerrado en el que se originan, lo que asegura la privacidad de las comunicaciones y evita interferencias con enlaces que operan en otras habitaciones.

Las comunicaciones mediante canales infrarrojos también presentan algunas desventajas. Entre ellas están: distancia de utilización limitada y no son inmunes a la luz ambiente ni al alumbrado fluorescente e incandescente que inducen ruido sobre el receptor de luz infrarroja. Estas fuentes de ruido están presentes en muchos entornos cerrados como oficinas, empresas, hogares, etc. Además, la presencia de obstáculos en la línea de emisión impide la comunicación de enlaces punto a punto y, por otra parte, la comunicación desde una habitación a otra necesita puntos de acceso de infrarrojos.

Los enlaces infrarrojos se clasifican atendiendo al grado de direccionalidad del transmisor y del receptor y, a la existencia o ausencia de visión directa (**LOS**, del inglés Line Of Sight o **DBC**, de las siglas Directive Beam Configuration) entre el emisor y el receptor [2]. En el primer caso se trata de enlaces directos o no directos según se empleen transmisores y receptores direccionales o de ángulo ancho. Los enlaces no directos se emplean en la comunicación de terminales móviles. En los enlaces directos, la potencia recibida es máxima debido a que se reduce el camino de atenuación de la señal y la recepción de ruido generado por la luz ambiente. También es posible establecer enlaces híbridos que emplean transmisores y receptores con diferentes grados de direccionalidad. El diseño de un enlace LOS maximiza la potencia recibida y disminuye la dispersión por multirayectoria (multipath dispersion). La dispersión por multirayectoria se produce debido a las múltiples reflexiones que sufre la señal emitida hasta alcanzar al receptor.



Un canal de propagación infrarrojo en interiores es, esencialmente, un canal con propagación multirayectoria. Esto significa que si se emite un pulso de potencia óptica de duración muy pequeña, como una función delta de Dirac, la señal recibida en el receptor estaría constituida por un tren de impulsos con diferentes retardos y amplitudes. Este resultado se debe a las múltiples reflexiones que sufre el pulso emitido en las paredes, muebles, personas y cualquier otro objeto situado en el lugar en el que están los emisores y receptores.

Para simular por computador la respuesta impulsional de un canal IR difuso, se parte de un modelo para los emisores, receptores y reflectores que se caracteriza mediante magnitudes escalares y vectoriales. En [3] se presenta un modelo y un algoritmo de simulación en el que se tienen en cuenta el número de reflexiones sobre las paredes (rebotes). En [10] se presenta una mejora del algoritmo anterior introduciendo una matriz en la que se almacenan las interacciones entre las diferentes superficies reflectoras, haciendo una simulación por tiempo. Además se mejora el tiempo de ejecución aprovechando propiedades de simetrías entre las superficies reflectoras para ahorrar cálculos. En [9] se presenta un estudio de los modelos deterministas anteriores y se implanta un modelo estadístico. Con este modelo estadístico se obtiene mayor flexibilidad y tiempos de ejecución muy aceptables. En los algoritmos deterministas comentados existe un problema importante: el número de cálculos que se deben realizar para simular situaciones realistas es extremadamente elevado. Esto significa que el tiempo de ejecución podría llegar a ser de hasta años. Por tanto, estos algoritmos no son viables debido a que para el momento en el que se obtuviese la respuesta de la simulación, probablemente los dispositivos simulados ya estarían obsoletos.

Por tanto es importante reducir este tiempo de ejecución tan elevado. En este artículo se presenta un algoritmo de simulación secuencial. Este algoritmo está inspirado en el algoritmo de [10] si bien se ha programado partiendo de cero, sin aprovechar las propiedades de simetría y sin introducir la matriz que almacena las interacciones entre las superficies reflectoras. Posteriormente se paraleliza aplicando transformaciones sobre el código secuencial para obtener tiempos de ejecución relativamente razonables.

La estructura del artículo es la siguiente: En la sección siguiente presentamos el modelo del sistema, en la sección 3 presentamos los programas de simulación por computador, en la sección 4 se presentan los resultados de la simulación por computador expresados en tiempos de ejecución y eficiencia de uso de los procesadores. Por último, presentamos algunas conclusiones y líneas de trabajo futuras.

## 2. EL MODELO DE SIMULACIÓN

Dado un sistema que consiste en una habitación, en cuyo interior se encuentran uno o varios emisores de

luz (LEDs) y uno o varios receptores, se trata de hallar la distribución temporal de potencia en cada uno de los receptores situados en el interior de la habitación debido a la emisión de señales procedentes de los emisores situados en el interior de dicha habitación.

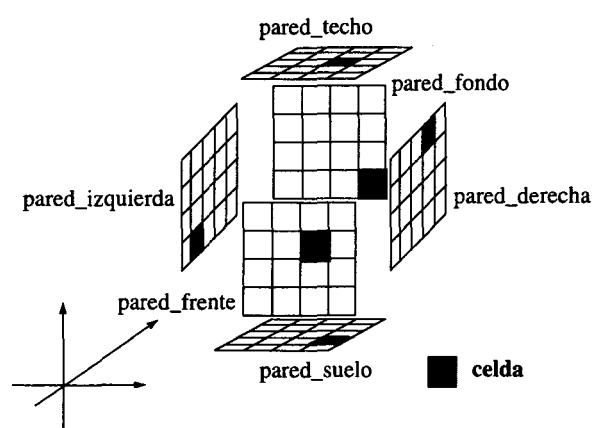
El emisor transmite potencia óptica en el instante de tiempo  $t_0$  con un cierto diagrama de radiación, esto es, la potencia no se transmite por igual en todas las direcciones. Parte de esa potencia llega al receptor y otra parte alcanza a las superficies reflectoras de la habitación. La potencia que llega a las superficies reflectoras rebota hacia otros lugares de la habitación, incluso hacia el receptor. Por tanto, al receptor le llega potencia por dos caminos diferentes:

- A través de una línea recta directa hipotética trazada entre emisor y receptor, y
- A través de un camino hipotético (“emisor pared\_1 >> pared\_2 >> pared\_3 >>...>> receptor”), generado por rebotes en las superficies reflectoras (paredes de la habitación).

La distribución temporal de la potencia que llega al receptor está formada por un pico de potencia correspondiente al enlace directo entre emisor y receptor, seguido de una cola de potencia debida a los rebotes en las paredes de la habitación. La potencia en el receptor disminuye a medida que avanza el tiempo debido a los múltiples rebotes que van atenuando a la señal.

Para considerar las reflexiones en las paredes de la habitación éstas se dividen en elementos discretos a los que denominamos celdas. En la figura 1 se muestran las paredes de una habitación dividida en celdas mediante el trazado de líneas rectas perpendiculares y paralelas a alguna de las aristas que las conforman.

Una celda, con potencia distinta de cero ilumina al resto de celdas que están en paredes diferentes a la suya. El tiempo que tarda en iluminar una celda emisora a una celda receptora, se calcula como donde:  $c$  es la velocidad de la luz y  $d$  es la distancia que existe entre ambas celdas.



**Figura 1.** División en celdas de las paredes de una habitación.

La potencia existente en una celda en el instante de tiempo  $t_i$ , depende de las potencias emitidas por otras celdas y emisores en instantes de tiempo  $t_j$  anteriores ( $j < i$ ). De igual manera, la potencia que llega al receptor en  $t_k$  depende de la potencia emitida por las celdas y emisores en tiempos  $t_l$  inferiores a  $t_k$  ( $l < k$ ). Por tanto, debido a la dependencia temporal que existe en el cálculo de la potencia que llega a una celda o a un receptor, se emplea una simulación dirigida por tiempo.

En la figura 2 se muestra una habitación con un emisor y un receptor. Se dibuja la línea de visión directa entre los dos elementos. También se dibujan los rebotes producidos en varias celdas. Nótese que para cada instante de simulación se debe averiguar qué celdas están iluminadas en ese instante; sin embargo este cálculo es fuertemente dependiente de las celdas que estaban iluminadas en los instantes de tiempo anteriores.

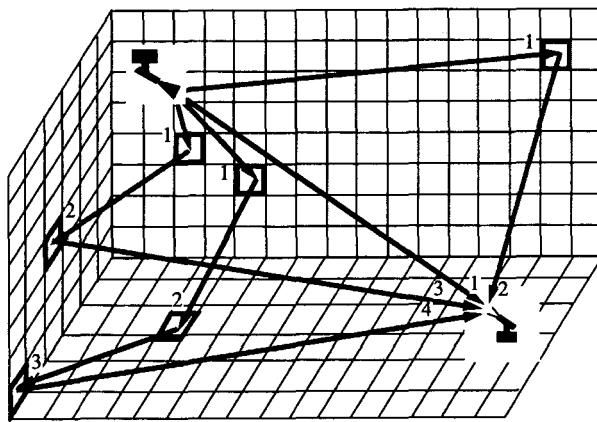


Figura 2. Visión directa y hasta tres rebotes.

Los emisores y receptores se representan mediante las siguientes magnitudes vectoriales y escalares [3]: vector de posición, vectores unitarios que determinan sus orientaciones (perpendiculares a la superficie de cada elemento), potencia óptica de salida (emisor), modo de emisión (emisor), diagrama de emisión lambertiano (emisor), área activa (receptor) y campo de visión (receptor). Las características reflectoras de cualquier superficie dependen del material y de la textura de la superficie. El modelo del reflector correcto, para la mayoría de las superficies reflectoras, debe considerar dos componentes [4]: una no direccional o difusa según un perfil de radiación determinado, y otra direccional o especular. Sin embargo, una buena aproximación para el modelo del reflector consiste en considerar solamente las componentes difusas, con un diagrama de emisión lambertiano ideal, esto es, el modo de emisión de la fuente es igual a 1 [3]. Este modelo simplificado se conoce con el nombre de Modelo de reflexión lambertiano.

La potencia recibida y el tiempo en el que ésta llega al receptor, dependen de la distancia  $d$  entre el emisor y el receptor (visión directa) o entre una celda y el receptor. La distancia  $d$  se calcula considerando la posición del emisor y el receptor (o la celda y el receptor).

### 3. SIMULACIÓN POR COMPUTADOR

Para simular la respuesta al impulso de luz o señal generada por un emisor de luz en una habitación vacía y siguiendo el modelo anterior, se utilizan estructuras de datos para representar los elementos que constituyen el modelo. Además se utiliza una matriz bidimensional, cuyo número de filas está dado por la siguiente expresión:. Donde  $t_{\max}$  y  $t_{\min}$  son el tiempo máximo y mínimo que transcurre para que una celda emisora ilumine a una celda receptora respectivamente y,  $d_{\max}$  y  $d_{\min}$  son las distancias máxima y mínima respectivamente entre cualesquier dos celdas de la habitación. El número de columnas de la matriz de potencias es igual al número de celdas total de la habitación. Esta matriz la denominamos **m\_pot** y las filas representan instantes discretos de simulación, mientras que las columnas representan a cada una de las celdas. Cada elemento de la matriz almacena un valor de potencia producido por la iluminación debida a las celdas o a los emisores.

En el apartado siguiente se describen las acciones secuenciales a realizar en el programa que se diseñó.

#### 3.1. Simulación secuencial

En el programa secuencial básicamente se especifican un conjunto de bucles anidados. Se recorren todos los instantes discretos de simulación y para cada instante se recorren todas las paredes y sus celdas. Para cada celda se comprueba si está iluminada (valor de potencia distinto de cero). En caso afirmativo, se recorren todas las celdas pertenecientes al resto de las paredes (sin considerar la pared a la que pertenece la celda que se está analizando:  $\text{pared\_d} \neq \text{pared\_o}$ ), calculando el efecto y el retardo que existe en la propagación de la señal procedente de la celda iluminada. Para ello hay que tener en cuenta las posiciones de las celdas dentro de la habitación y el diagrama de radiación lambertiano de la celda iluminada. A continuación se actualiza la matriz de potencias (**actualizar m\_pot**). La actualización de esta matriz consiste en acceder a una fila superior al instante de tiempo discreto de simulación que se está considerando ( $d$ ) , y acceder a la columna marcada por la celda a la cual se quiere comprobar si le llega potencia. En ese elemento se almacena el valor recién calculado. El acceso a las filas de **m\_pot** se hace en módulo  $n$  (número de filas de dicha matriz).

Por cada celda iluminada también se comprueba si existe algún receptor iluminado por esa celda. En caso afirmativo, se actualiza la potencia que llega al receptor debido a la iluminación de dicha celda.

El proceso anterior se repite para todas las celdas de la habitación. Acabado este proceso, entonces se pone a cero todos los elementos de la fila  $d$  de **m\_pot**. Esto es necesario porque en la siguiente iteración de  $d$  podría ser posible que se utilizaran algunos de los elementos de esa fila para acumular el valor de la potencia en las celdas.

Por tanto si no se pone previamente a cero entonces se podrían calcular valores incorrectos de la potencia.

A continuación, si se han recorrido todas las filas de `m_pot`, se actualiza la potencia que envía cada uno de los emisores de señal a las celdas en las que se han dividido las paredes de la habitación.

Cuando finaliza esta actualización, se procede con el siguiente valor del tiempo de simulación discreto.

```

for d (1, fin_tiempo_simulación, 1)
  for pared_o (1, max_num_paredes, 1)
    for celda_o (1, fin_pared_o, 1)
      if celda_o iluminada then
        for pared_d (1, max_num_paredes (pared_d<>pared_o), 1)
          for celda_d (1, fin_pared_d, 1)
            Calcular efecto y retardo entre celda origen y destino
            Calcular potencia y actualizar m_pot
        end_for
      end_for
    end_for
    for receptor (0, max_num_receptores-1, 1)
      if receptor iluminado por celda_o then
        Calcular potencia entre celda origen y receptor
      end_if
    end_for
    end_if
  end_for
  end_for
  Actualizar a cero la fila leída de la matriz m_pot
  if recorrida m_pot then
    for emisor(0, max_num_emisores-1, 1)
      if emisor señal then
        Repeat para cada fila de m_pot
          for pared (1, max_num_paredes, 1)
            for celda (1, fin_pared, 1)
              if celda iluminada por emisor then
                Calcular potencia emisor-celda.
              end_if
            end_for
          end_for
        end_repeat
      end_if
    end_for
  end_if
end_for

```

**Figura 3.** Algoritmo secuencial.

En la figura 3 se muestra un pseudocódigo. Se puede comprobar que el código básicamente está constituido por un conjunto de bucles anidados. En estos bucles se actualizan variables vectoriales cuya estructura es relativamente sencilla.

El número de operaciones que se realizan en este algoritmo es muy elevado. Por cada celda iluminada se ha de recorrer el resto de las celdas de la habitación. Y esto se repite para cada instante de simulación. Además, para cada operación se tienen que realizar cálculos costosos en tiempo de ejecución (cálculos de ángulos, distancias, proyecciones, etc.) con números almacenados utilizando la representación en punto flotante (lo que significa que las operaciones en el procesador aún serán más lentas que si se utilizan números en punto fijo). Por

tanto estamos frente a un problema realmente complejo, en el que a medida que el número de celdas aumenta estaríamos frente a posibles tiempos de ejecución de incluso años [3].

Una forma de reducir este elevado coste computacional es repartir el número de cálculos entre diferentes procesadores. Si los procesadores pudieran trabajar al mismo tiempo entonces reduciríamos el tiempo de ejecución considerablemente a medida que empleamos más procesadores.

### 3.2. Paralelización del programa secuencial

El objetivo principal de la paralelización de códigos secuenciales consiste en obtener una nueva ejecución de los cálculos secuenciales de tal forma que algunos de estos se puedan realizar de forma paralela.

Para obtener un programa paralelo, paralelizando un programa secuencial, es necesario particionar y distribuir el código secuencial y los datos entre los procesadores y memorias de la máquina paralela [6]. En los últimos años se han construido potentes máquinas que disponen de más de un procesador (arquitecturas paralelas) y su complejidad se ha incrementado enormemente. La ejecución eficiente de los programas en estas máquinas requiere que se tomen en cuenta las características particulares de las mismas. En la actualidad no existe una herramienta automática que permita realizar el particionado de código y datos entre procesadores. En ausencia de estas herramientas de particionado y distribución, el código y los datos deben ser particionados por el programador especializado.

Los bucles anidados son el núcleo de aplicaciones científicas y de ingeniería que transforman estructuras constituidas por un elevado número de datos. En estos bucles se efectúan una gran cantidad de cálculos y accesos a memoria concentrándose en este tipo de código un gran porcentaje del tiempo de ejecución de estos programas. Por tanto, es en este tipo de códigos en los que se centra el programador para paralelizar los códigos secuenciales transformando el código secuencial para aprovechar la posible ejecución paralela de los cálculos a realizar. A estas transformaciones se les suele denominar reestructuración de código [5].

En el código secuencial de la figura 3 existen varios bucles en los que se efectúan la mayor parte de los cálculos. Para paralelizar este código es preciso centrar nuestra atención en estos bucles. Para paralelizar el programa mediante reestructuración de código, teniendo en cuenta los bucles, hemos de realizar las siguientes fases:

1. *Distribución de datos.* Dado que los bucles que recorren las celdas son los que influyen definitivamente en el tiempo de ejecución, entonces nos centramos en cómo distribuir las celdas entre los procesadores. El objetivo es conseguir que los procesadores realicen los

cálculos del efecto y retardo entre conjuntos de celdas disjuntos. Esto conlleva el que deben trabajar con columnas diferentes de la matriz *m\_pot*.

*2. Distribución de cálculos.* Para conseguir el objetivo del paso anterior se debe modificar el código. Esto es, debemos obtener el código que ejecutarían los procesadores. La forma más sencilla de hacer esto es reestructurar el código secuencial haciendo que los procesadores ejecuten el mismo código, pero parametrizándolo mediante variables que identifiquen las operaciones que ejecuta cada procesador.

Para llevar a cabo este paso se deben aplicar transformaciones de código. Una técnica de paralelización de códigos secuenciales es la denominada *Block Scheduling Equilibrado* [6]. Si cada procesador dispone de una variable local denominada *my\_id* que le identifica únicamente, entonces con dos vectores denominados *imin* e *imax* se pueden obtener los límites inferiores (*imin[my\_id]*) y superiores (*imax[my\_id]*) de los bucles tal que el procesador *my\_id* trabaje con un conjunto de celdas disjunto al del resto de los procesadores.

*3. Paso de mensajes.* Con los pasos anteriores, cada procesador necesita datos que no conoce puesto que han sido calculados por los otros procesadores. Concretamente, en nuestro caso, es necesario que todos los procesadores comuniquen la parte de la matriz de potencia que actualizan para que el resto de los procesadores puedan utilizar estos datos para calcular la potencia que envía su conjunto de celdas hacia el resto de celdas y receptores.

En la figura 4 se muestra el pseudocódigo del programa paralelo resultante. Se puede apreciar que un procesador, después de comprobar todas las celdas que le han sido asignadas dentro de la habitación, realiza una comunicación colectiva en la que se radia la información anteriormente mencionada.

```

for d (1, fin_tiempo_simulación, 1)
    for pared_o (1, max_num_paredes, 1)
        for celda_o (imin[pared_o], imax[pared_o], 1)
            ...
            Actualizar a cero la fila leída de la matriz m_pot
            Comunicar la siguiente fila de la matriz m_pot
            if recorrida m_pot then
                ...
                for pared (1, max_num_paredes, 1)
                    for celda (imin[pared_o], imax[pared_o], 1)
                        if celda iluminada por emisor then
                            Calcular potencia emisor-celda.
                            end_if
                        end_for
                    end_for
                end_if
            end_for
        ...
    end_for
end_for

```

**Figura 4.** Algoritmo Paralelo.

En el código resultante los procesadores efectúan una gran cantidad de cálculo y una vez acabados esos cálculos se efectúa la comunicación. Por cada iteración

del bucle d existe la necesidad de efectuar comunicaciones. Además, si el número de celdas es elevado, los mensajes a comunicar contienen una gran cantidad de datos.

## 4. RESULTADOS DE LA SIMULACIÓN POR COMPUTADOR

En este apartado presentamos algunos resultados de la medida de los tiempos de ejecución y de la eficiencia considerando una habitación de las siguientes características: a) dimensiones 5X5X3 metros cúbicos, b) posición del emisor en (2.5, 2.5, 3), el receptor en (4.5, 1, 0), y c) tiempo de simulación de 60 nsg., utilizando diferentes tamaños de celdas.

Estos algoritmos han sido implantados en el multicomputador SP-2 del CESCA (CEntro de Supercomputación de CAtaluña). En el multicomputador SP-2, los procesadores no comparten memoria física (memoria distribuida), y se comunican usando paso de mensajes a través de un dispositivo de comunicación de alta velocidad, el High Performance Switch (HPS). El HPS es una red de interconexión omega [7], con conmutadores de 8x8 que utilizan encaminamiento buffered-wormhole mediante conmutación de paquetes [7]. Para especificar el paso de mensajes en el programa paralelo se ha utilizado la biblioteca MPI (Message Passing Interface) [8]. La comunicación básicamente se especifica utilizando las llamadas a **MPI\_Bcast** y **MPI\_Allreduce** pertenecientes a la biblioteca MPI.

En la figura 5 se muestra el tiempo de ejecución del programa secuencial y del paralelo, expresado en minutos. El programa paralelo se ha ejecutado en seis procesadores.

La reducción de tiempo de ejecución conseguida por el programa paralelo es considerable. Por ejemplo, para un tamaño de celda igual a 10 centímetros se consigue una reducción aproximadamente de un sexto del tiempo de ejecución del programa secuencial. Además se puede comprobar fácilmente que el tiempo de ejecución disminuye a medida que se aumenta el tamaño de la celdas. Por lo tanto, podemos asegurar que la paralelización del programa secuencial es efectiva.

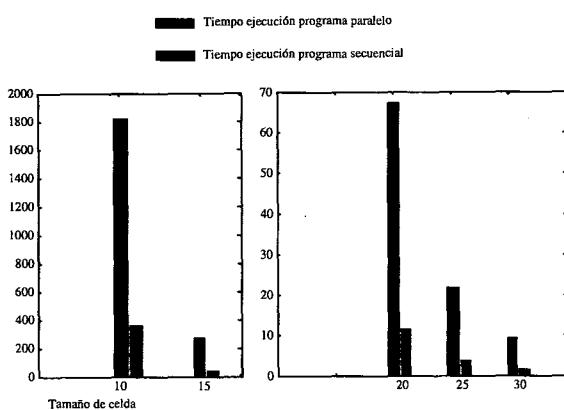
Cuando se genera un programa paralelo es deseable que los procesadores se utilicen durante todo el tiempo de ejecución del programa paralelo. Para obtener información sobre la utilización de los procesadores se utiliza una métrica denominada eficiencia [7]. La eficiencia se calcula como:

$$Eficiencia = \frac{T_{seq}}{n \times T_{par}}$$

donde:  $T_{seq}$  es el tiempo de ejecución del programa secuencial,  $T_{par}$  es el tiempo de ejecución del programa paralelo y  $n$  es el número de procesadores. La eficiencia es una figura de mérito cuyos valores extremos son:

$$0 \leq Eficiencia \leq 1$$





**Figura 5.** Tiempos de ejecución del programa secuencial y del paralelo.

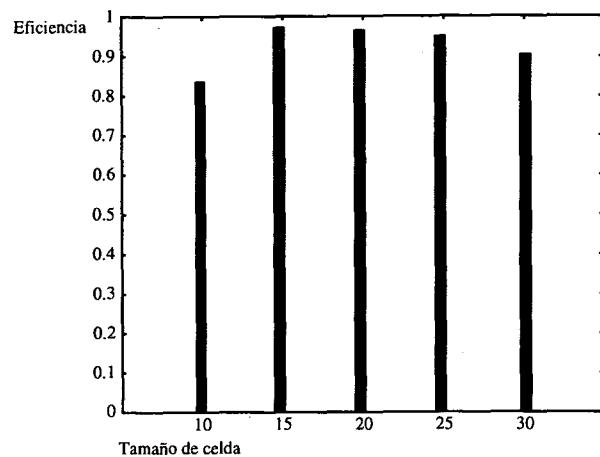
En la figura 6 se muestra la eficiencia del programa paralelo para diferentes tamaños de celdas. La eficiencia tiene un valor elevado lo cual significa que los procesadores han sido utilizados durante todo el tiempo de ejecución del programa paralelo. Esto se debe a que las comunicaciones se han minimizado, haciendo que los procesadores realicen una cantidad de cálculos elevada. En las arquitecturas como el SP-2 esta técnica da buenos resultados de eficiencia.

## 5. CONCLUSIONES Y LÍNEAS DE TRABAJO FUTURO

En las **WLAN** la portadora empleada puede ser de radiofrecuencia (**RF-WLAN**), o de infrarrojo (**IR-WLAN**). Los *enlaces por infrarrojo* proporcionan comunicaciones no cableadas de alta velocidad en distancias relativamente cortas, siendo muy apropiados para aplicaciones de interiores (habitaciones). Para simular, por computador, la respuesta de un canal de infrarrojos en interiores se parte de un modelo para los componentes emisores y receptores así como para el medio físico (canal). La simulación por computador tiene la ventaja de ser menos costosa y más flexible que la simulación real en un laboratorio. Sin embargo, el tiempo de ejecución para la simulación realista por computador puede ser muy elevado.

En este artículo hemos presentado un algoritmo secuencial que tiene un tiempo de ejecución muy elevado (de hasta 30 horas). La parallelización de este programa en un multicomputador conduce a una reducción muy considerable del tiempo de ejecución (hasta 6 horas, para el caso anterior). Se han probado varias configuraciones de emisores y receptores comprobándose la correctitud de la simulación y obteniendo una utilización eficiente de la arquitectura del multicomputador utilizado.

El trabajo presentado en este artículo, se podría ampliar considerando modelos matemáticos más complejos para los emisores. Por otro lado, se podrían considerar habitaciones en L.



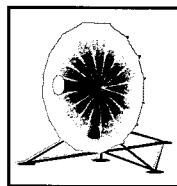
**Figura 6.** Eficiencia del programa paralelo.

## AGRADECIMIENTOS

Agradecemos al personal del *CESCA* (*CEntro de Supercomputación de CAtaluña*) el soporte prestado en la utilización del multicomputador SP-2: sin sus máquinas ni su personal este trabajo no hubiera sido posible.

## REFERENCIAS

- [1] FRITZ R. GFELLER, URS BAPST, Wireless In-House Data Communication via Diffuse Infrared Radiation, Proceedings of the IEEE, vol. 67, no. 11, pag. 1474, Noviembre 1979.
- [2] JOSEPH M. KAHN, JOHN R. BARRY, Wireless Infrared Communications, Proceedings of the IEEE, VOL. 85, NO. 2, Febrero 1997.
- [3] JOHN R. BARRY, JOSEPH M. KAHN, WILLIAM J. KRAUSE, EDWARD A. LEE, DAVID G. MESSERSCHMITT, Simulation of Multipath Impulse Response for Indoor Wireless Optical Channels, IEEE Journal on Selected areas in communicationes, Vol. 11, no. 3, Abril 1993.
- [4] CIPRIANO R.A.T. LOMBA, RUI T. VALADAS, A.M. DE OLIVEIRA DUARTE, Update of Propagation Losses and Impulse Response of the Indoor Optical Channel, IEEE P802.11, Septiembre 1993.
- [5] WOLFE J. M., Optimizing Compilers for Supercomputers, Pitman Publishing, London, 1989.
- [6] STEVEN BRAWER, Introduction to Parallel Programming, Academic Press, Inc., ISBN 0-12-128470-0, 1989.
- [7] HWANG KAI, Advanced Computer Architecture. Parallelism, Scalability, Programability, McGraw-Hill, ISBN-0-07-031622-8, 1993.
- [8] PETER S. PACHECO, Parallel Programming with MPI, Morgan Kaufmann Publishers, Inc., 1997.
- [9] JOSE BERGES COLMENAR, Modelo estadístico para la simulación de redes de datos por enlace óptico no guiado, Proyecto Fin de Carrera, Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad de Las Palmas de Gran Canaria, Noviembre 1996.
- [10] F. J. LOPEZ HERNÁNDEZ, *Dustin: A new algorithm for the calculation of the impulse response in indoor IR*



# SONAR PARA APLICACIONES TERRESTRES

Jordi Gili, Jordi Pera, Lluís Closas

Departament d'Enginyeria Electrònica de la ETSETB (UPC)  
jordi\_gili@es.ibm.com, jgil1210@alu-etsetb.upc.es, jper1201@alu-etsetb.upc.es

## INTRODUCCIÓN

Este trabajo pretende ser una introducción al SONAR con aplicaciones no convencionales, ya que el SONAR se utiliza básicamente en entornos marinos. El motivo que impulsa a la redacción de este artículo es la divulgación de la experiencia realizada por los autores [1].

Convencionalmente se asocia la utilización del sonido como agente de detección en medios acuáticos, debido a su eficiente propagación en estos. Su utilización en medios aéreos (atmósfera) comporta una pérdida importante de esta capacidad de propagación, lo cual ha relegado a este agente a un segundo plano, frente a otros que presentan una propagación más eficiente como las ondas electromagnéticas.

Con estas perspectivas, uno de los pocos huecos que existen para el sonido en la atmósfera es su aplicación al campo de la detección y posicionamiento a corta distancia. Este campo queda vagamente cubierto por las ondas electromagnéticas debido a su gran velocidad de propagación ( $3 \times 10^8$  m/s), la cual impone una imposible capacidad de cálculo con métodos de conteo de tiempo. El sonido, sin embargo, con una velocidad de propagación mucho menor (344 m/s), permite una clara aplicación de estos métodos.

Existen comercialmente dispositivos, TL851[2], que permiten la detección y el cálculo de distancia de objetos situados en el margen de 15.24cm a 10.6 m.

El objetivo del proyecto ha sido desarrollar un sistema capaz de posicionar un objeto, no solamente en distancia, sino también en dirección con respecto a un origen fijado. Para el desarrollo de este sistema se han aplicado conceptos y técnicas ampliamente utilizadas en el campo de las ondas electromagnéticas (agrupaciones de antenas, diagramas de radiación, conformación de haz, etc.).

## TEORÍA BÁSICA

Fundamentalmente existen dos puntos a tratar para realizar la detección y el posicionamiento de un objeto: el cálculo de la distancia y el cálculo del ángulo donde se encuentra.

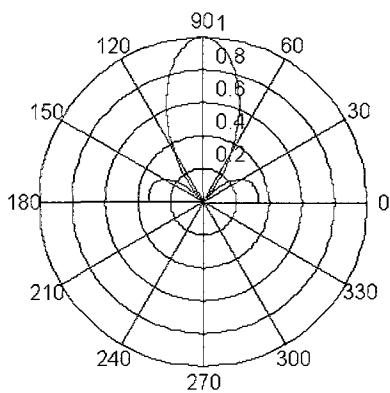
El procedimiento que se utiliza para el cálculo de la distancia se basa en la emisión de un pulso, activando simultáneamente un contador de tiempo. El pulso es reflejado en el objeto y detectado en el receptor. En ese instante se detiene el contador y mediante la simple fórmula  $d = v * t / 2$  ( $d$ : distancia,  $v$ : velocidad de propagación,  $t$ : tiempo medido por el contador) se determina la distancia al objeto.

El procedimiento utilizado para el posicionamiento del objeto (ángulo respecto a la vertical) es más complejo y se basa en la utilización de una agrupación de cápsulas ultrasónicas, realizando sobre ellas una conformación electrónica de haz, de forma que se consigue ‘apuntar’ el haz principal de la agrupación en todas las direcciones del plano formado por los elementos de ésta.

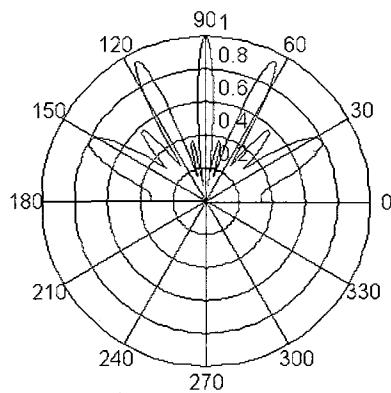
Una agrupación consiste en diferentes elementos situados a distancias conocidas entre ellos. El hecho que radien sonido provoca fenómenos de interferencia constructiva o destructiva en todo el espacio donde existe radiación de forma que algunos puntos del espacio reciben más potencia que otros. Si se mide la potencia recibida en cada uno de los puntos de una circunferencia trazada en el plano que forman los elementos de la agrupación, centrada ésta y con un radio lo suficientemente alejado de ella (campo lejano) se puede medir el diagrama de radiación, obteniendo máximos en las zonas de interferencia constructiva y ceros en las zonas de interferencia destructiva. Para posicionar precisamente un objeto es necesario que el diagrama de radiación de la agrupación tenga un lóbulo principal con una anchura de haz relativamente pequeña.

Para poder detectar objetos en las distintas direcciones la técnica utilizada consiste en modificar a voluntad el diagrama de radiación de la agrupación de forma que el lóbulo utilizado para la detección ‘apunte’ hacia las diversas direcciones, técnica que se conoce como Conformación de Haz o ‘Beamforming’[3]. La conformación del haz se realiza desfasando convenientemente la alimentación de los elementos radiantes consiguiendo así modificar las zonas de interferencia constructiva y destructiva en el espacio con lo que se obtiene un efecto de ‘apuntamiento’ del haz. Este punto constituye una clave del sistema, ya que la implementación del desfase de las alimentaciones no es sencilla.

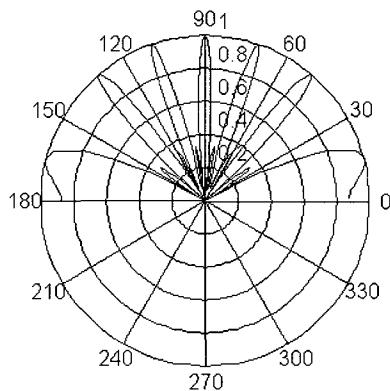




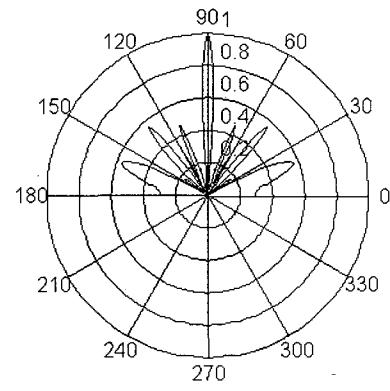
**Figura 1.** Diagrama de radiación de la agrupación en



**Figura 2.** Diagrama de radiación de la agrupación emiso-



**Figura 3.** Diagrama de radiación de la agrupación recepto-  
ra.



**Figura 4.** Diagrama de radiación resultante.

## AGRUPACIÓN DE CÁPSULAS ULTRASÓNICAS

A diferencia de los sistemas que utilizan ondas electromagnéticas, las antenas utilizadas para la emisión de ultrasonidos son distintas a las utilizadas para su recepción, lo cual obliga a la construcción de dos agrupaciones de cápsulas.

En la obtención de un diagrama como el descrito en el apartado anterior se ha utilizado una agrupación lineal de tres elementos en ambos casos. La razón de la elección de este número de elementos es la siguiente: existe un compromiso entre la calidad de conformación del haz y la complejidad de control del sistema; con dos elementos la conformación de haz que se puede realizar es muy pobre, mientras que un número superior de elementos introduce un factor de complejidad al sistema que no compensa el escaso beneficio que aporta.

La agrupación que mejor se ajusta a nuestras necesidades es una agrupación lineal con los elementos equiespaciados una distancia de 0.5, obteniendo así el diagrama de radiación de la figura 1 (en el caso de que los elementos de la agrupación sean isotrópicos). Dada la frecuencia de trabajo de las cápsulas electrónicas

comerciales utilizadas ( $f=40\text{ kHz}$ ), una distancia de 0.5 supone una separación entre cápsulas de 43 mm. El diámetro de estas cápsulas es aproximadamente de 1 cm, con lo que la realización física de la agrupación propuesta resulta inviable. Este hecho obliga a distanciar más los elementos lo que provoca la aparición de lóbulos de difracción en el diagrama de radiación, efecto indeseado ya que produce problemas en la decisión de la posición del objeto.

Para conseguir un diagrama de radiación parecido al propuesto, se ha optado por aprovechar el hecho de tener dos agrupaciones distintas para emisión y recepción y combinar ambos diagramas de radiación. La meta perseguida es la cancelación de los lóbulos de difracción de cada diagrama con los mínimos del otro diagrama, exceptuando el caso del lóbulo principal, que no debe resultar afectado.

Las agrupaciones elegidas presentan los siguientes espaciamientos: 2 y 2.5 para emisión y 3.1 equiespaciada para recepción. Las figuras 2 y 3 muestran los diagramas obtenidos para las agrupaciones emisora y receptora respectivamente (suponiendo elementos isotrópicos). En ellas se puede observar claramente la presencia de los lóbulos de difracción. La figura 4 muestra como con la

combinación de ambos diagramas se consigue disminuir considerablemente la presencia de estos lóbulos. Queremos destacar que no nos hemos planteado en nuestro estudio la búsqueda de la combinación óptima.

## SISTEMA DE CONTROL

El sistema encargado de gestionar las distintas agrupaciones tiene que ser capaz de realizar una serie de funciones que permitan la consecución de las prestaciones esperadas.

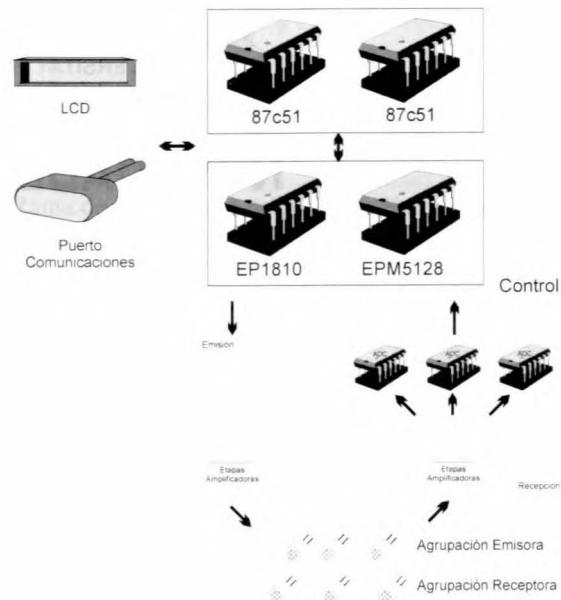
Para el sistema de alimentación hay que resolver dos cuestiones fundamentales: la generación de las señales de alimentación de las cápsulas y el desfase de éstas para conseguir la conformación del haz. Las cápsulas tienen un ancho de banda muy estrecho (aproximadamente 1 kHz)[4], lo cual obliga a que la señal de alimentación tenga una frecuencia prácticamente coincidente con su frecuencia nominal ( $f_{nom}=40$  kHz), cosa que es relativamente sencilla de conseguir con una generación digital de la señal. Las tres señales de alimentación, además, deben de ser desfasadas de forma muy precisa para obtener la conformación de haz deseada.

Para el sistema de recepción, el problema a resolver es básicamente la alineación de las tres señales recibidas por la agrupación para conseguir la conformación del haz deseada.

La complejidad de introducir un desfase variable de forma analógica hace que esta opción sea de difícil realización. Por tanto, la posibilidad de desfasar las señales de forma digital aparece como la más factible. Para el caso de las señales de emisión, el desfase se consigue retardando el instante de inicio de generación de cada una de ellas.

Para el caso de recepción, puesto que las señales son analógicas y se realiza una conversión A/D, se aprovecha el muestreo para introducir el desfase retardando el instante de inicio del mismo, de forma similar a lo realizado en emisión. Como resultado se obtienen las señales digitales desfasadas a voluntad, resultando muy sencillo combinarlas para conseguir el diagrama de radiación deseado.

Un esquema general del prototipo que se ha diseñado se puede observar en la figura 5. En él se pueden distinguir dos partes bien diferenciadas, la generación de las señales de alimentación, y la recepción de las muestras de las señales detectadas y su posterior proceso para la determinación del blanco. En efecto para la generación de las señales se ha empleado uno de los microprocesadores 87C51 [5] y la PLD EPM5128 [6]. Para la parte de recepción de muestras se ha empleado una PLD EP1810 [6] y otro microprocesador 87C51 [5].



**Figura 5.** Esquema del prototipo SONAR desarrollado.

El prototipo dispone además de un puerto de comunicaciones que permite su interconexión con otros sistemas.

## CONCLUSIONES

El prototipo diseñado muestra una clara aplicación comercial en el campo de Sistemas de Seguridad y Vigilancia, dado que el sistema es capaz de discriminar la ubicación de un objeto dentro de una sala puede decidir si éste se encuentra en una zona permitida o prohibida, en función de la distancia y posición calculados. Mediante el puerto de comunicaciones el prototipo puede desencadenar la respuesta de otros sistemas (alarma, circuito cerrado de televisión, etc.).

Existe una gran variedad de campos en los que puede tener aplicación este sistema: desde la domótica hasta la juguetería o ayuda a invidentes.

## REFERENCIAS

- [1] SONAR para detección y seguimiento de blancos móviles. JORDI GILI, JORDI PERA.PFC, ETSETB, 1998.
- [2] Linear Applications. Texas Instruments. 1996.
- [3] LAGUNAS M.A., «Procesamiento de Arrays», Apuntes de la asignatura ARRAYS ADAPTATIVOS, impartida en la ETSETB.1997
- [4] Cápsulas ultrasónicas:  
Frecuencia: 401kHz  
Sensibilidad: -68dB/V/bar  
Tensión máxima a la entrada: 20 VRms.
- [5] Kit de desarrollo para microprocesador 87C51. LLUÍS CLOSAS. EUPBL.
- [6] Sistema de desarrollo con PLD de Altera, MaxPlu.

Para más información, contactar con los autores.

# LOS BORGIA. LA TRINIDAD MALDITA



*Luís A. Alcalde*

*Proyectista del Departamento de Matemática Aplicada y Telemática  
Universitat Politècnica de Catalunya  
lalcalde@mat.upc.es, http://personal.redestb.es/jaimeg/*

Los Borgia fueron una noble familia valenciana (Borgia es la italianización del apellido Borja) que a finales del siglo XV estuvo a punto de someter media Italia al poder de la Santa Sede y convertirla en una pseudo-monarquía hereditaria. Hoy en día Italia es una república unida, pero en el renacimiento era un puzzle de ducados, repúblicas y reinos. Tres fueron los miembros de la familia que han pasado a la historia: el Papa Alejandro VI (Rodrigo Borgia), y sus hijos César y Lucrecia.

Rodrigo Borgia era un mujeriego cardenal originario de Játiva con seis o siete hijos cuando fue elegido Papa (luego tuvo alguno más). Durante su papado tuvo muchos enemigos y fue acusado de simonía y de asesinar mediante «cantarella» (veneno utilizado en Renacimiento) a varios cardenales. Pero la acusación más grave que aún planea sobre él es el posible incesto con su hija Lucrecia. A pesar de todo, fue un gran mecenas y supo mantener la Santa Sede a salvo de las codicias franco-españolas.

César actuó sin escrúpulos para conseguir sus objetivos: utilizó el poder del Papa, su padre; casó a su hermana Lucrecia en favor de sus intereses políticos; asesinó a su hermano Juan de Gandía para obtener la capitánía de los ejércitos pontificios, contrató a Leonardo de Vinci para la fabricación de maquinaria de guerra... pero la ambiciosa idea de César Borgia de unificar Italia topó con la oposición de los propios italianos y de las dos potencias que se la disputaban: España y Francia. La trayectoria política y militar de César fue admirada por el propio Maquiavelo, quien utilizó su figura para inspirarse en su obra más importante, «El príncipe».

La bella Lucrecia fue moneda de cambio en manos de su hermano César y de su padre. Fue utilizada para sellar alianzas. Además fue objetivo de los ataques de los enemigos los Borgia, que la acusaban de ser la prostituta del Papa.

## LOS INICIOS

La historia comienza en Játiva, en el Reino de Valencia, en 1431. En ese año nacía Rodrigo Borja en el seno de una familia noble que había recibido los favores de Jaime I el Conquistador por su lucha contra los moros en la reconquista de Valencia. Rodrigo escogió la carrera eclesiástica, y con sólo veinticinco años fue nombrado vicencanciller de la Iglesia por su tío, el Papa Calixto III

allá por el año 1456. También llegaba al Vaticano su hermano Pedro Luis Borja como gonfalonero (Capitán General) de los ejércitos pontificios. Así comenzaba la historia de los Borgia en Roma, a los que los italianos llamaban 'catalanes'. El pontificado de Calixto III Borgia duró tres años. A su muerte en 1458 la plebe romana asaltó los palacios de los 'catalanes' protegidos del difunto Papa (habitual en aquella época), pero Rodrigo supo mantenerse en su puesto de vicencanciller de los siguientes papas gracias a su astucia. Así hasta el cónclave de sucesión de Sixto IV en 1484, donde empezó a sonar su nombre como nuevo papa. Pero los recelos que levantaba el cardenal Borgia impidieron su elección. Inocencio VIII fue el sucesor.



*LosBorgia*

## LA ELECCIÓN DE ALEJANDRO VI

En agosto de 1492 llegó la segunda oportunidad de Rodrigo Borgia; esta vez se enfrentaba por el papado con Julián de la Rovere (futuro Julio II), quien fue su eterno enemigo. Su elección la noche del 10 de agosto de 1492 le costó decenas de miles de ducados, así como favores y títulos. Los cardenales estaban divididos entre italianos, franceses y españoles. Julián de la Rovere era el protegido de Carlos VIII de Francia, así que contaba con los votos franceses y además era italiano. Rodrigo recibía el respaldo español y tuvo que comprar varios votos entre los cardenales italianos: ofreció la vicecancillería a Ascanio Sforza, con lo que se aseguraba el apoyo de Milán; al cardenal Orsini varios castillos para recibir la ayuda de esta importante familia romana,... y la mañana del 11 de agosto fue proclamado papa con el nombre de Alejandro VI. La elección del nombre de debía a Alejandro Magno, lo que ya hacía presagiar las futuras ansias conquistadoras de los Borgia.

## LOS HIJOS DEL PAPA

Pero el nuevo Papa no llegó solo al Vaticano, al igual que hizo su tío Calixto III, se rodeó de los suyos,

que en su caso eran sus hijos: Juan, César, Lucrecia y Jofré.

Este hecho no era nada extraordinario en una ciudad, Roma, en la que más del diez por ciento de la población eran prostitutas. Nada más instalarse empezó a asegurar el futuro de sus hijos. En 1493 casó a Lucrecia (con trece años) con Juan Sforza, conde de Pesaro y que doblaba en edad a su futura esposa. Con este matrimonio, Alejandro VI se aseguraba la alianza de los Sforza de Milán.

Esta boda enemistada a los Borgia con Nápoles debido a la rivalidad de este reino sureño con Milán, así que se pactó el matrimonio entre Jofré Borgia y Sancha de Aragón, hija de Alfonso II, rey de Nápoles. Boda que se celebró en 1494. Juan fue enviado a España, donde se casó con María Enríquez y adquirió el título de duque de Gandía. Por último, César fue obligado a seguir el oficio eclesiástico y fue nombrado cardenal de Valencia. Pero el joven Borgia no estaba contento con los planes que su padre tenía para él e hizo todo lo posible por cambiarlos...

## LA PRIMERA INVASIÓN FRANCESA

Al morir Ferrante de Aragón en enero de 1494, se desataron las codicias franco-españolas sobre Nápoles. La elección de Alfonso II no frenó al rey Carlos VIII de Francia, que invadió el norte de Italia ese mismo año con la intención de atacar Nápoles atravesando tota Italia si era necesario.

*La trayectoria política y militar de César fue admirada por el propio Maquiavelo, quien utilizó su figura para inspirarse en su obra más importante, «El príncipe».*

Carlos VIII fue un don quijote del renacimiento, con la diferencia que era rey y tenía uno de los ejércitos más poderosos para llevar a cabo sus fantasías aventureras. Su inicial alianza con Ludovico el Moro le permitió atravesar rápidamente Milán y presentarse a las puertas de Roma el 31 de diciembre de 1494. De nada les sirve a los Borgia atrincherarse en Sant Angelo, los franceses tienen el apoyo de los Colonna, familia romana rival de los Borgia.

El rey francés estaba disgustado con el Papa debido al matrimonio entre Jofré y Sancha y, especialmente, por la coronación de Alfonso II como rey de Nápoles; así que exigió a Alejandro VI la entrega de su hijo César como

rehén. Con el hijo del Papa en sus manos, Carlos VIII se aseguraba la ‘colaboración’ de la Santa Sede y dejaba protegida su retaguardia en su viaje hacia el sur; pero el rey francés no contaba con la astucia de los Borgia.

Alejandro VI había pactado con Alfonso II una intervención militar contra los intereses franceses en Italia (ataque marítimo contra Génova). Además César consigue escapar del campamento francés y volver a Roma. Los franceses quedan atrapados en Italia, su retirada es complicada y son finalmente derrotados en verano de 1495 en Fornuove por los ejércitos de la Santa Liga (España, Venecia, Milán, Alemania y la Santa Sede).

Carlos VIII muere en la primavera de 1498 y es nombrado rey Luis XII, quien volverá a intentar la aventura napolitana de su antecesor.

## LOS AÑOS OSCUROS

El periodo de tiempo entre las dos invasiones francesas (1496-1498) son los años más oscuros de los Borgia: divorcios, asesinatos, incesto?, ...

Tras la invasión francesa de Carlos VIII, los barones de la Romaña, en especial la familia Orsini, están enfrentados con Roma. Alejandro VI trae a su hijo Juan de Gandía desde España y le nombre gonfalonero (capitán general) del ejército papal.

La campaña comienza en verano de 1496 con varias victorias papales frente a los Orsini, pero la inutilidad militar de Juan da un respiro a los Orsini, que recobran terreno y fuerzan una paz con el Papa a principios de 1497.

Mientras, César sigue a disgusto con su cargo de cardenal y empieza a maquinar para conseguir sus objetivos. Primero pone en fuga al marido de Lucrecia, Juan Sforza, ya que no interesa la alianza con la familia milanesa.

El divorcio se consigue a final de ese año al reconocer el duque de Pésaro que ha sido incapaz de consumar el matrimonio. Lucrecia se refugia en el convento de San Sixto y en Marzo de 1498... ¡da a luz un niño!. La paternidad de este niño, ‘el infante romano’, no está claro y son dos las hipótesis. La primera y más probable es que el padre fuese Pedro Calderón, camarero del Papa que hacía de intermediario mientras Lucrecia estaba en el convento y que apareció muerto en el Tíber. La segunda, fomentada por los enemigos de los Borgia, apunta a Alejandro VI como padre y abuelo del infante Romano.

César pacta el matrimonio de Lucrecia con Alfonso de Besceglia, hermano de Sancha e hijo del rey de Nápoles, con lo que afianza las relaciones entre los



Borgia y los Aragón de Nápoles. Pero lo que de verdad necesita César para alcanzar sus objetivos es el mando militar del ejército papal y la cobardía de su hermano en la guerra de la Romaña es la gota que colma su paciencia. El 15 de junio de 1497 aparece muerto Juan de Gandía en el Tíber y todas las miradas apuntan a su hermano César.

Según algunas versiones antiborgianas, el cargo de gonfalonero no fue el único motivo que llevó a César a asesinar a su hermano. El amor más allá de lo fraternal que sentía César por Lucrecia podía haberle llevado a asesinar a Juan, a quien Lucrecia admiraba por su campaña militar en la Romaña.

En agosto de 1498 consigue que su padre le libere de su condición de cardenal y recibe el nombramiento de gonfalonero de la Iglesia. En octubre se entrevista con el nuevo rey de Francia, Luis XII, para explicarle sus proyectos. El rey francés le acoge en su corte, le regala el ducado de Valentinois y arregla su boda con Carlota de Albret, hermana del rey de Navarra.

## EL APOGEO DE CÉSAR

La alianza se sella y en octubre de 1499 las tropas de César atacan la Romaña mientras Francia conquista Milán.

El objetivo francés es el de siempre: Nápoles. Pero la alianza entre César y Francia cuenta con un problema: los matrimonios de Lucrecia y Jofré con los hijos del rey napolitano. Victorioso en la guerra de la Romaña, decide eliminar el último escollo que se interpone en su alianza con los franceses: en verano de 1500 ordena el asesinato de Alfonso de Bisceglia, marido de Lucrecia, y el encarcelamiento de Sancha de Aragón en Sant Angelo.

César entra en el siglo XVI en el momento culminante de su vida, para él trabaja Leonardo de Vinci como ingeniero militar y su fama llega a oídos de Nicolás Maquivel, político de la República de Florencia, que es enviado por su país como emisario ante César.

En 1501 César Borgia es nombrado duque de la Romaña, su objetivo parecía cumplido, pero las ambiciones de César amenazan incluso a Florencia. En octubre de 1502 sus capitanes conspiran contra él en Magione, pero les descubre y con una astuta jugada los capturen en Sinigaglia el último día de 1502. Su último gran día...

## CAÍDA Y MUERTE DE CÉSAR

El año 1503 fue catastrófico para César. En primavera los franceses son derrotados en Nápoles por el ejército español capitaneado por el héroe de Granada, 'El Gran Capitán' Gonzalo de Córdoba. Con esta victoria España se anexiona Nápoles y pone fin a las pretensiones francesas. Por otro lado, su padre muere en

Agosto al mismo tiempo que él cae enfermo (lo que hace suponer un envenenamiento). César se queda sin sus soportes: el Papa y el rey de Francia.

Pero la mala fortuna de César no acaba aquí, el día de todos los santos es elegido Papa Julián de la Rovere, el gran enemigo de los Borgia, con el nombre de Julio II. El nuevo Papa hace prisionero a César y lo entrega a Gonzalo de Córdoba para ser juzgado en España.

Después de unos meses en Chinchilla, César es trasladado a la fortaleza de La Mota, en Medina del Campo. En invierno de 1506 consigue escapar y huye

*La bella Lucrecia fue  
objetivo de los ataques de  
los enemigos los Borgia,  
que la acusaban de ser la  
prostituta del Papa.*

hacia el norte, hacia Navarra, donde reina su cuñado Juan de Albret. Navarra estaba en guerra contra el noble Luis de Beaumont y Juan de Albret nombra a César capitán del ejército navarro.

En marzo de 1507 puso sitio a la ciudad de Viana. El día doce de aquel mes amaneció claro después de una noche lluviosa. César recibió la noticia de la llegada de suministros a la ciudad sitiada aprovechando que durante la noche se había retirado la guardia que controlaba el sitio de la ciudad. Enfurecido, montó en su caballo y emprendió el camino de la ciudad al galope, sin preocuparse de si sus hombres le seguían. Cuando se aproximó a la ciudad se topó con veinte jinetes rebeldes que rodearon al solitario jinete y acabaron con su vida.

## LUCRECIA EN FERRARA

Tras el asesinato de Alfonso, el segundo marido de Lucrecia, Alejandro VI y César pactaron el matrimonio de la joven (todavía tenía 21 años) con Alfonso de Este, heredero de Ferrara. A pesar de la mala fama de la borgia, la familia Este aceptó el matrimonio a cambio de ciertas concesiones en su relación con la Santa Sede (y por el poder que César estaba acumulando en la Romaña). Lucrecia se casó en 1501 y se trasladó a Ferrara. Nunca más volvió a ver a los suyos. Fue duquesa de Ferrara y murió en 1519.

## LOS PAISES

### España

Bajo el reinado de los Reyes Católicos se formaba España como la confederación de dos reinos: el reino de Castilla y el reino de Aragón. A principio de 1492,



### *Mapa de la zona italiana*

Castilla conquistaba el reino de Granada, con lo que finalizaba la reconquista; y el 12 de octubre de ese mismo año Cristóbal Colón descubría el Nuevo Continente. Aragón llevaba siglos con los ojos puestos en el Mediterráneo, y al actual objetivo era el reino de Nápoles, donde reinaban descendientes de la misma dinastía aragonesa. Pero el deseo de Francia de anexionarse este reino provocó los roces entre los dos estados, con la victoria final española gracias a la conquista de Nápoles por 'El Gran Capitán' Gonzalo de Córdoba en 1503.

## **Los Estados Pontificios**

Dentro del puzzle italiano se encontraba la Santa Sede. Los dominios del estado papal se expendían por toda la Romaña. Limitaba al sur con Nápoles y al norte con Florencia, los miniestados de Mantua y Módena y el ducado vasallo de Ferrara. La Romaña estaba gobernada por barones vasallos del Papa, pero que habitualmente 'olvidaban' este vasallaje. César Borgia emprendió una guerra contra estos barones que le llevó a la conquista de muchas de sus posesiones: Cesena, Rímini, Pesaro,...

Ferrara

Ferrara era un culto y respetado ducado vasallo del Vaticano y gobernado por la familia Este. Su situación geográfica colocaba a este ducado como ‘marca’ entre la poderosa Venecia y los Estados Pontificios. El matrimonio

nio entre Alfonso de Este y Lucrecia Borgia permitió obtener a los Este su separación de la Santa Sede.

Florencia

Gobernada por los Médicis, era el centro bancario de la Europa renacentista. En 1492 moría Lorenzo el Magnífico, momento en que comienza la decadencia de Florencia. Los Médicis son expulsados y se proclama la nueva República. En esta época destaca el monje Savonarola, serviente predicador que ataca la actuación

*En el Renacimiento Italia era un puzzle de ducados, repúlicas y reinos.*

‘satánica’ de Alejandro VI y que provoca un estado de exaltación religioso en la república. Finalmente es condenado por el Papa y ejecutado. Florencia era enemiga de Milán y se mantuvo neutral en los conflictos debidos a las invasiones francesas debido a sus problemas internos y a su guerra con Pisa. Maquiavelo fue emisario de la república florentina ante César Borgia.

Francia

En los años del papado de Alejandro VI, Francia ‘sufrió’ el reinado de Carlos VIII y Luis XII. El objetivo de ambos fue el reino de Nápoles, lo que les llevó a invadir (más bien atravesar, viendo el resultado final) Italia en dos ocasiones (1494 y 1501) para conseguir su objetivo. De nada sirvieron sus alianzas con César Borgia, pues la Santa Liga (España, Venecia, Milán, Alemania y Los Estados Pontificios) primero, y ‘El Gran Capitán’ después, truncaron sus objetivos.

Milán

Situada al norte de la península italiana, estaba gobernada por los Sforza. En 1494 Ludovico el Moro es nombrado duque de Milán y en un principio el milanesado es un aliado de Francia debido a su enemistad con Nápoles; pero acaba siendo invadido por Luis XII de Francia.

## Nápoles

Reino situado en el sur de Italia y que fue el principal motivo de las invasiones francesas a finales del siglo XV. La casa de Aragón reinaba desde que Pedro III expulsó a los franceses en la primera mitad del siglo. Estos hechos fueron los que dieron pie a que España y Francia exigieran sus derechos sobre Nápoles. Finalmente fue conquistada por Gonzalo de Córdoba en 1504 e incorporada a España.

## Venecia

La república gobernada por los 'dux' basada su poder y su riqueza en el comercio con Oriente. Se mantuvo relativamente al margen de las luchas italianas.

## BIBLIOGRAFÍA BÁSICA

ROBICHON, JAQUES. Los Borgia. La trinidad maldita  
BELLONCI, MARÍA. Lucrecia Borja  
BARBERÁ, CARMEN. Yo, Lucrecia Borgia

## CRONOLOGÍA

- 1453** Toma de Constantinopla por los turcos.
- 1455** Elección de Calixto III Borgia.
- 1456** El cardenal Rodrigo Borgia se convierte en vicecanciller de la Iglesia.
- 1458** Muere Alfonso V de Aragón. Ferrante de Aragón nuevo rey de Nápoles.
- 1459** Rodrigo Borgia conoce a Vannozza Cattanei en el Congreso de Mantua.
- 1468** Nacimiento de Pedro Luis Borgia, primer hijo de Rodrigo Borgia y Vannozza Cattanei.
- 1469** Boda de Fernando de Aragón e Isabel de Castilla.
- 1475** **Nacimiento de César Borgia.**
- 1476** **Nacimiento de Juan Borgia.**
- 1480** **Nacimiento de Lucrecia.**
- 1490** Muerte de Pedro Luis Borgia.
- 1492** 2 de enero. Toma de Granada por Gonzalo de Córdoba. 8 de abril. Muere Lorenzo el Magnífico.
- 11 de agosto. **Elección de Rodrigo Borgia como Alejandro IV.**
- 12 de octubre. Colón descubre América.
- 1493** 3 de marzo. Tratado de Tordesillas.
- 12 de junio. Boda de Lucrecia Borgia y Juan Sforza.
- Agosto. Boda de Juan de Borgia, duque de Gandía con María Enríquez.
- Septiembre. César Borgia nombrado cardenal de Valencia.
- 1494** Muere Ferrante de Nápoles. Franceses y españoles optan al reino de Nápoles. Carlos VIII de Francia invade el norte de Italia.
- Mayo. Boda de Jofré Borgia y Sancha de Aragón. Alfonso V rey de Nápoles.
- Octubre. Ludovico el Moro nuevo duque de Milán.
- Noviembre. Los Medici expulsados de Florencia.
- Fin de año. Las tropas francesas entran en Roma.
- 1495** Febrero. Los franceses invaden Nápoles.
- Abril. Se crea la Santa Liga (España, Alemania, Venecia, Milán y Los Estados Pontificios) contra Francia.
- 6 de julio. Derrota total de los franceses en la batalla de Fornuove.
- 1496** Cuaresma. Predicaciones de Savonarola en Florencia. Juan (Borgia) de Gandía es nombrado gonfalonero de la Iglesia.
- Guerra de la Santa Sede contra los barones de la Romaña.
- 1497** 27 de marzo. Juan Sforza huye de Roma.
- 6 de junio. Lucrecia de refugia en el convento de San Sixto.
- 15 de junio. **Asesinato del duque de Gandía.**
- Noviembre. Divorcio de Lucrecia y Juan Sforza.
- 1498** Febrero. Pedro Calderón aparece muerto en el Tíber.
- Marzo. Nace el Infante Romano.
- 7 de abril. Muere Carlos VIII. Luis XII nuevo rey de Francia.
- Julio. Lucrecia se casa con Alfonso de Aragón, duque de Bisceglia.
- Agosto. César abandona el cardenalato y recibe el ducado de Valentinois del rey de Francia.
- 1499** Mayo. César se casa con Carlota de Albret, hermana del rey de Navarra.
- Octubre. Luis XII de Francia ataca Milán.
- 1 de noviembre. Nace Rodrigo de Bisceglia, hijo de Lucrecia y Alfonso de Aragón.
- César ataca varias ciudades de la Romaña (Imola, Forlì,...).
- 1500** César nombrado gonfalonero de la Iglesia.
- Julio. **Asesinato de Alfonso de Bisceglia**, esposo de Lucrecia.
- Octubre. César reanuda los ataques en la Romaña (Cesena, Pesaro y Rímini).
- Negociaciones secretas entre Francia y España para repartirse el reino de Nápoles.
- 1501** Abril. Rendición a Faenza ante César. César recibe el título de duque de la Romaña. El ejército papal amenaza Florencia.
- Leonardo de Vinci entra al servicio de César.
- Verano. El ejército francés desciende sobre Nápoles.
- 30 de noviembre. Lucrecia se casa con Alfonso de Este, heredero de Ferrara.
- 1502** 2 de febrero. Lucrecia entra en Ferrara.
- Junio. César conquista Urbino. Primer encuentro entre César y Maquiavelo.
- Octubre. Conspiración de Magione contra César.
- 21 de diciembre. Represalia de César en Sinigaglia.
- 1503** Ejecución de los conjurados en Magione.
- 28 de abril. Gonzalo de Córdoba derrota a los franceses en Nápoles. España se apodera del reino napolitano.
- 18 de agosto. **Muere Alejandro VI.**
- Septiembre-Octubre. Papa Pío III.
- 1 de noviembre. Elección del Papa Julio II de la Rovere.
- Maquiavelo presencia la caída de César.
- 1504** César prisionero de Julio II y entregado a los españoles.
- Cautiverio de César en Chinchilla.
- 26 de noviembre. Muere Isabel la Católica en Medina del Campo.
- 1505** Mayo. César es trasladado al castillo de La Mota (Medina del Campo).
- 1506** 25 de octubre. César se escapa de La Mota y se refugia en Navarra, junto con su cuñado.
- 1507** César es nombrado capitán del ejército navarro.
- 12 de marzo. **César muere en Viana.**
- 1519** 24 de junio. **Muere Lucrecia Borgia, duquesa de Ferrara.**