

## TÉCNICAS DE SÚPER-RESOLUCIÓN PARA LA MEJORA DE SECUENCIAS DE VÍDEO COMPRIMIDO

Gustavo Marrero Callicó,  
Dácil Barreto Dos Santos,  
Sebastián López Suárez,  
Carlos Tarajano Beracoechea,  
José López Feliciano,  
Roberto Sarmiento Rodríguez,  
Antonio Núñez Ordóñez

Instituto de Microelectrónica  
Aplicada. ULPGC

El mercado de consumo actual ha incrementado de forma espectacular la demanda de sistemas que mejore la calidad de la imagen para aplicaciones que van desde el vídeo doméstico a las militares, pasando por la teledetección, la seguridad, la astronomía o las aplicaciones médicas. Para una determinada secuencia de imágenes, el proceso de súper-resolución permite mejorar la resolución de dicha secuencia, incrementando de forma notable su calidad visual. Este trabajo muestra una novedosa técnica basada en la aplicación de una etapa de pre-procesado que permite aplicar el algoritmo de súper-resolución para incrementar la calidad de la imagen sólo en aquellas regiones donde la calidad puede efectivamente ser mejorada, disminuyendo de esta forma el coste computacional total. Al mismo tiempo, también se presenta una aplicación que persigue disminuir el tamaño del vídeo comprimido mientras que mantiene su calidad visual.

*Nowadays, the markets that demand improvements in the image quality have spectacularly grown, ranging from domestic video to military, teledetection, security, astronomical or medicine applications. For a given image sequence, the super-resolution process allows the resolution enhancement of such sequence, notably increasing the perceived visual quality. This work exposes a novel technique based on the application of a pre-processed stage which performs the super-resolution algorithm to increase the image quality only in those regions where the quality can be improved, with the subsequent computation saves. At the same time, an application intended to decrease the size of the compressed video while keeping its visual quality is also presented.*

### INTRODUCCIÓN

En estas últimas décadas, la digitalización de la imagen ha supuesto un espectacular aumento de las aplicaciones y servicios en los que se demanda mejor calidad, no sólo a nivel visual en el mercado de consumo, sino también para su análisis en otros mercados como el militar, la teledetección, la astronomía o la medicina. La súper-resolución (SR) permite, dada una determinada secuencia de imágenes, obtener dicha secuencia con una resolución espacial mayor, mejorando de forma importante la calidad visual final. Para ello se utilizan las diferencias que existen entre imágenes adyacentes a consecuencia del movimiento de la cámara que las captura o del movimiento propio de los objetos que aparecen en la secuencia. De esta forma, es posible

recuperar detalles que se han perdido en el proceso de adquisición de la imagen.

A continuación se describen algunas de las aplicaciones más frecuentes donde se utiliza la súper-resolución, analizando en detalle uno de los entornos más atractivos de aplicación de la SR en la actualidad: la compresión de vídeo.

### **Aplicaciones de la súper-resolución**

El proceso de SR (Tsai y Huang, 1984) se basa en el hecho de que muchas veces es posible tener acceso a múltiples vistas de una misma imagen, como ocurre en las secuencias de vídeo, donde se capturan de 25 a 30 imágenes por segundo. En casos de movimiento normal (es decir, sin cambios demasiado brus-

En este artículo se presentan nuevas técnicas para mejorar la calidad de secuencias de imágenes usando algoritmos de súper-resolución y su aplicación en entornos de compresión de vídeo

cos de una imagen a la siguiente, como por ejemplo en el caso de los cohetes balísticos o el lanzamiento de objetos en eventos deportivos), estas frecuencias de captura permiten obtener varias vistas similares de determinadas zonas de interés de la imagen. El algoritmo de SR determina qué información adicional puede obtenerse de las diferentes vistas de la imagen (o de zonas de la imagen) y la combina en una nueva imagen con mayor resolución espacial, es decir, con un mayor número de detalles o, lo que es lo mismo, con un mayor número de píxeles por pulgada (*ppp*), entendiendo que un píxel es la unidad mínima de información visual en el momento de muestrear la imagen. Decimos que un sensor es capaz de capturar imágenes de mayor resolución cuanto mayor sea el número de píxeles muestreados.

No siempre es posible obtener mejoras de calidad en el proceso de aplicar la SR a una determinada secuencia de imágenes (Alam et al., 2000). Las principales limitaciones se deben al proceso de captura y a la naturaleza de la secuencia tomada. Dentro de las limitaciones debidas al proceso de captura es obvio que, si el sensor o cámara captura una imagen cada 10 segundos, es muy poco probable que varias imágenes consecutivas tengan información que compartir. Decimos entonces que su correlación temporal es baja y el proceso de SR no es útil. Si la cámara está situada sobre un trípode, este sistema de sujeción evitará la introducción de movimiento debido a la cámara, eliminando una de las posibles fuentes de información al proceso de SR. Sin embargo, todavía es posible que exista nueva información proveniente del movimiento de la propia escena. Dentro de las limitaciones debidas al tipo de escena, la más importante es que la imagen en sí misma no tenga detalles que recuperar. Una zona de la imagen con un cielo azul puede ser de gran belleza, pero difícilmente se podrán recuperar detalles de donde no los hay. Al

mismo tiempo, es posible que la escena no tenga movimiento, como suele ser el caso de las cámaras de vídeo-vigilancia durante la mayor parte de su tiempo de funcionamiento. Si además la cámara está sujeta a una base fija, por muchos detalles que tenga la escena y mucha correlación que exista entre las imágenes, no se podrá obtener nueva información debido a la carencia absoluta de todo tipo de movimiento. En ocasiones, aunque exista movimiento, el nivel de detalle es tan escaso que una cámara de alta resolución (varios mega-píxeles) puede capturar cómodamente todos los detalles de la imagen, no dejando nada que mejorar al proceso de SR. Esta última limitación es una combinación de los dos factores limitantes anteriormente comentados: excesiva calidad de la cámara (limitación del sistema de captura) y poco nivel relativo de detalle de la escena (limitación por la naturaleza de la secuencia grabada).

Las aplicaciones de la SR se extienden prácticamente a todos los campos donde se capturen imágenes y se quiera o se necesite mejorar su calidad. En algunos casos, este proceso sólo persigue el aumento de la calidad perceptual, como en las aplicaciones de fotografía o vídeo domésticos. El proceso de SR podría ayudar de forma importante a otros sistemas de procesamiento de imágenes, como podría ser el caso de los reconocedores ópticos de caracteres (OCR, *Optical Character Recognizer*) donde una mejor calidad de la imagen del texto a reconocer mejora de forma importante la cantidad de caracteres correctamente reconocidos. En aplicaciones de astronomía, la SR puede emplearse para distinguir detalles tales como un pequeño planeta orbitando alrededor de una estrella lejana o los extremos del disco de crecimiento que suele formarse en las cercanías del horizonte de sucesos de un agujero negro. En aplicaciones de radar, la SR permite localizar mejor objetos pequeños, al igual que en aplicacio-

**La súper-resolución permite incrementar la resolución espacial, mejorando la calidad visual de una imagen o una secuencia de vídeo**

**Gracias a aplicaciones de súper-resolución se puede reconocer una cara o recuperar la matrícula de un vehículo en películas que han sido grabadas por una cámara de seguridad de escasa calidad o a mucha distancia**





cuestiones, pero antes es necesario comentar un último aspecto relacionado con las comunicaciones de bajo ancho de banda. Normalmente, un paso previo a la compresión es el diezmo espacial de las imágenes de una secuencia. Esto significa que, si por ejemplo el factor de submuestreo es 2, se tomará una de cada dos muestras en las direcciones vertical y horizontal. De esta manera, una imagen de 288 filas y 352 columnas pasaría a tener un tamaño de 144 filas y 176 columnas. El procedimiento de submuestreo descrito se representa en la Figura 3.

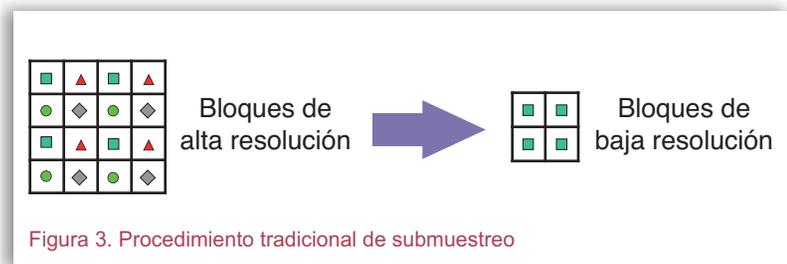


Figura 3. Procedimiento tradicional de submuestreo

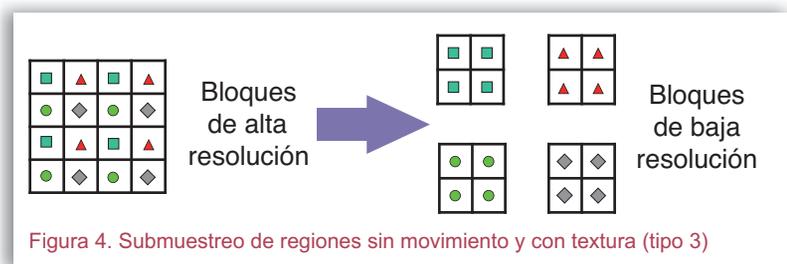


Figura 4. Submuestreo de regiones sin movimiento y con textura (tipo 3)

### Pre-procesamiento y súper-resolución

**Para disminuir el coste computacional, la súper-resolución aplicada a la compresión puede utilizar los vectores de movimiento transmitidos por el compresor**

La estimación de movimiento es el proceso más importante de la SR (Callicó et al., 2002). La calidad final que se obtenga va a depender en gran medida de la exactitud de los vectores de movimiento utilizados. Para disminuir el coste computacional, la SR aplicada a la compresión puede utilizar los vectores de movimiento transmitidos por el compresor. Tras realizar un estudio de la estimación de movimiento de ajuste de bloques utilizada en compresión, cuando se aplica a la SR (Barreto et al., 2005), se obtuvieron las siguientes conclusiones:

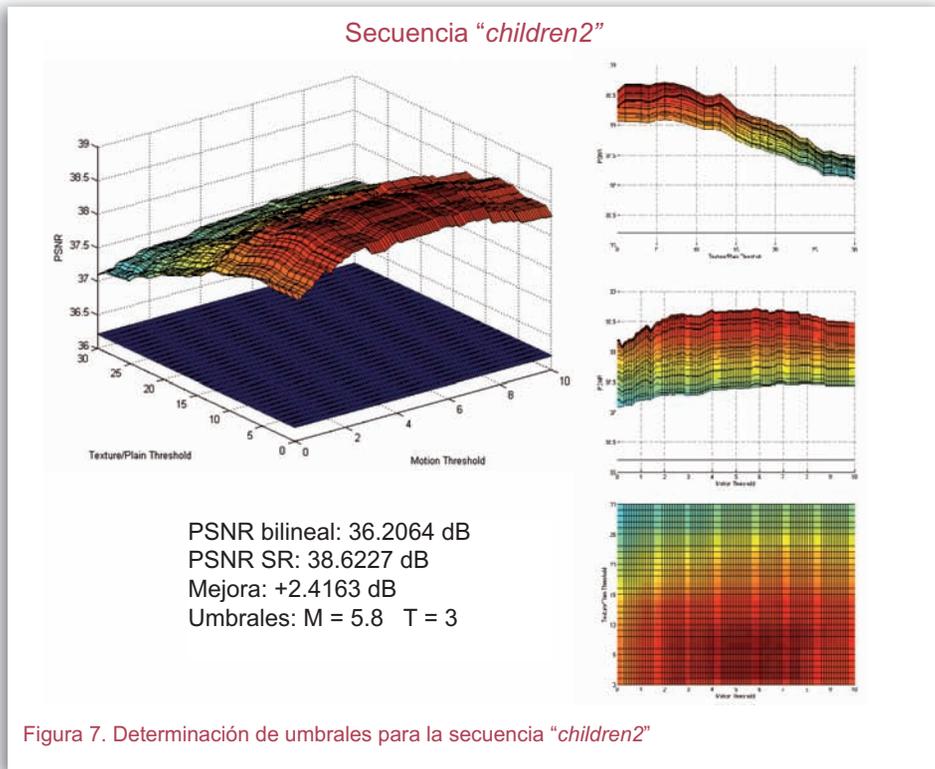
- Las regiones planas de la imagen (bajas frecuencias) suelen producir vectores de movimiento erróneos durante el proceso de estimación de movimiento. Si estas regiones son detectadas y los vectores de movimiento descartados, una simple interpolación bilineal proporciona resultados de buena calidad.
- Las regiones con textura, en las que se encuentran los detalles y bordes de la imagen, producen estimaciones de movimiento muy aproximadas a las reales. Si existe movimiento, estas regiones pueden ser recuperadas a partir de información de imágenes adyacentes.
- Las regiones sin movimiento sólo pueden ser interpoladas. Mientras las regiones planas sí son recuperadas con precisión, las regiones con textura no son mejoradas.

De estas observaciones se deduce que una clasificación de las imágenes según textura y movimiento puede ser muy beneficiosa no sólo a la hora de mejorar los resultados de la SR sino también para reducir su coste computacional. Esta clasificación se realizará a nivel de macrobloque por compatibilidad con el compresor. Si una región tiene movimiento podrá ser súper-muestreada utilizando SR. El problema se presenta cuando la región carece de movimiento. Si ésta a su vez es plana, una interpolación bilineal será suficiente para aumentar la resolución con una calidad suficiente, al mismo tiempo que se reduce la carga de cómputo. En cambio, si se trata de una región con textura será necesario buscar una solución que permita mejorar los resultados frente a una interpolación bilineal.

El apartado anterior mostraba los principios de funcionamiento de los compresores estándares de vídeo. Una región con textura y sin movimiento sufrirá de manera drástica los efectos del cuantificador, que elimi-







Se observa una dependencia importante del proceso global de súper-resolución con el tipo de región a procesar según su nivel de textura y su cantidad de movimiento

Figura 7. Determinación de umbrales para la secuencia "children2"

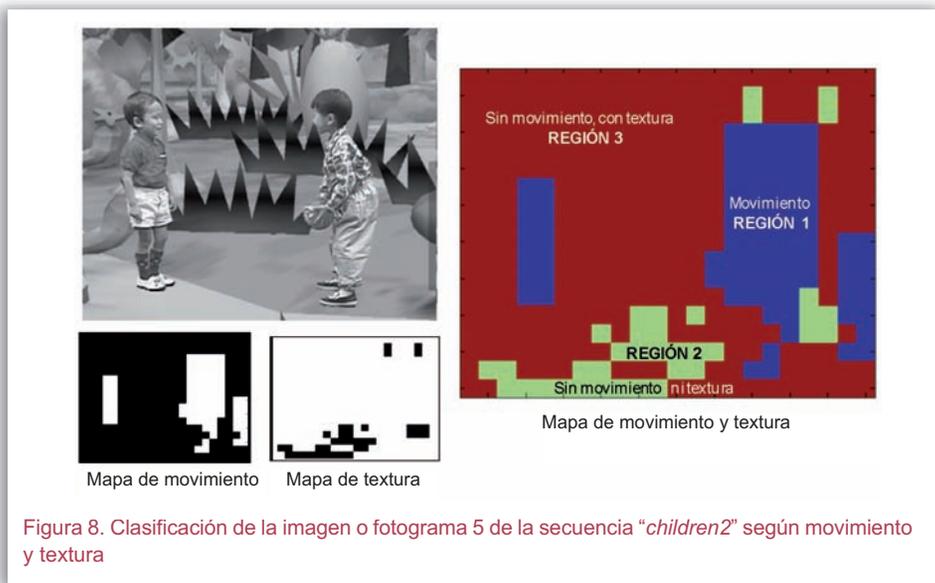


Figura 8. Clasificación de la imagen o fotograma 5 de la secuencia "children2" según movimiento y textura

entre 3.3 y 5.8, mientras que el umbral de textura se debía mantener fijo a 3. Aunque para la textura está claro que conviene dejar el umbral constante e igual a 3, para el movimiento sería mejor emplear un procedimiento dinámico que permitiese variar dicho valor centrado en 4.55 más o menos un factor de corrección de 1.25 en función de la cantidad de mo-

vimiento observado. La selección del umbral sería un procedimiento adaptativo ajustado a la naturaleza de la secuencia a procesar. La Figura 8 muestra la clasificación resultante al aplicar el método propuesto al fotograma 5 de la secuencia "children2", donde se ha obtenido un incremento de calidad de 2.4163 decibelios frente a la imagen interpolada.

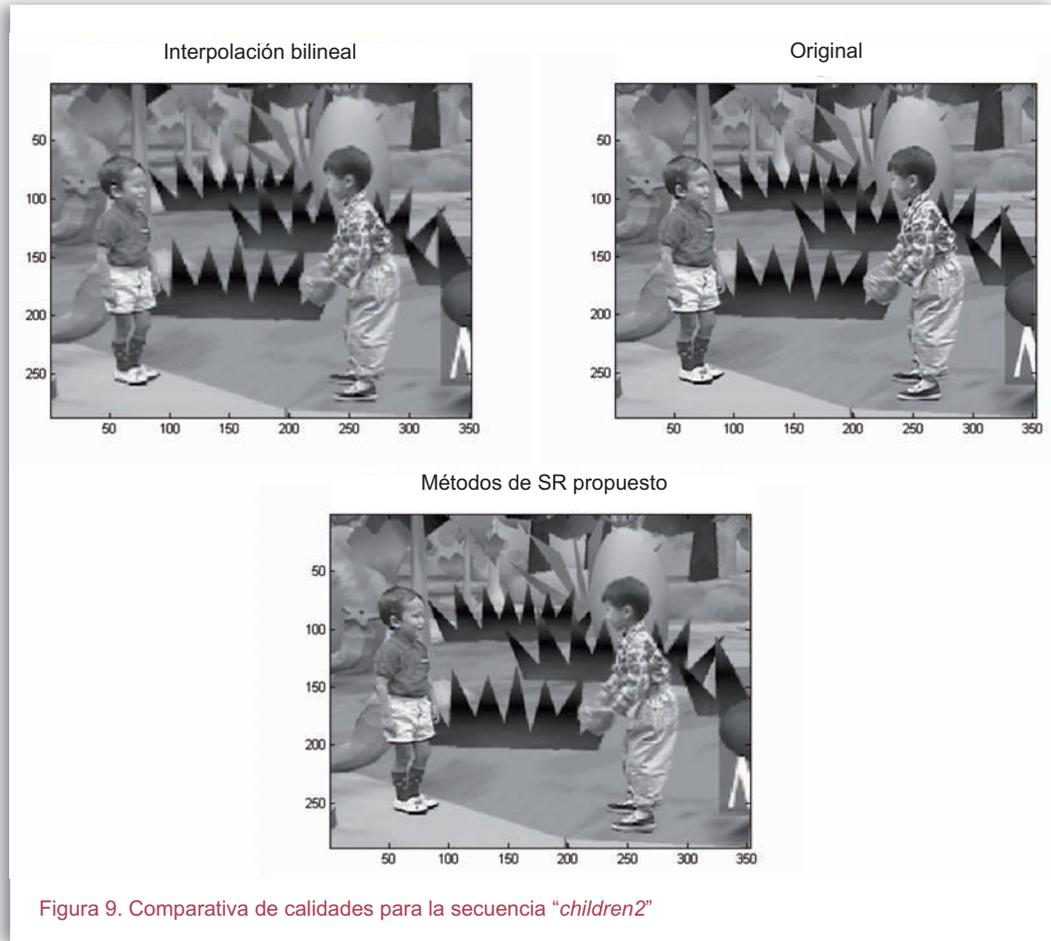


Figura 9. Comparativa de calidades para la secuencia "children2"

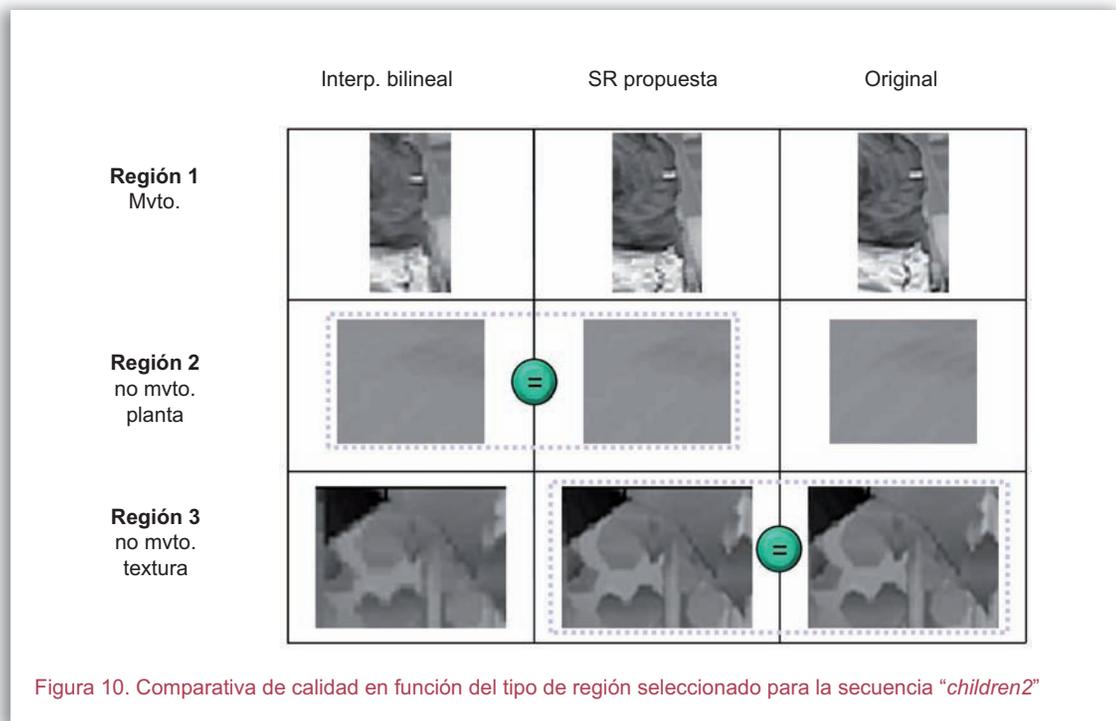


Figura 10. Comparativa de calidad en función del tipo de región seleccionado para la secuencia "children2"



porte recibido a través de los proyectos de investigación TEC2005-08138-C02-01/MIC y TEC2006-13599-C02-02/MIC financiados por el Ministerio de Educación y Ciencia.

## BIBLIOGRAFÍA

Alam M. S., Bogner J. G., Hardie R. C. y Yasuda B. J. (2000): Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames, *IEEE Transactions on Instrumentation and Measurement*, vol. 49, issue 5, págs. 915-923.

Barreto D., Alvarez L.D. y Abad J. (2005): Motion Estimation Techniques in Super-Resolution Image Reconstruction. A Performance Evaluation, *Proceedings of iASTRO MC Meeting & Workshop, Virtual Observatories: Plate Content Digitization, Archive Mining and Image Sequence Processing*, Sofia, Bulgaria.

Bhaskaran V., Konstantinides K. (1997): *Image and Video Compression Standards: Algorithms and Architectures*. Kluwer Academic Publishers (2ª ed.).

Callicó G.M., Peset Llopis R., Núñez A., Sethuraman R. y Marc Op de Beeck (2002): A Low-Cost Implementation of Super-Resolution based on a Video Encoder, *28th Annual Conference of the IEEE Industrial Electronics Society IECON'2002*, vol. 2, Seville, Spain, págs. 1439-1444.

Callicó G.M., Peset Llopis R., Núñez A. y Sethuraman R. (2003): Real-Time and Low-Cost Incremental Super-Resolution over a Video Encoder, *4th International Symposium on Electronic Design ISQED'2003* San Jose, California, USA, págs. 79-84.

López S., Callicó G.M., López J. y Sarmiento R. (2005): Compresión de Vídeo Mediante Técnicas de Post-Procesamiento, *Vector Plus Fundación Universitaria de Las Palmas*, núm. 25, págs. 27-36.

Schulz R. R. y Stevenson R. L. (1996): Extraction of high-resolution frames from video sequences, *IEEE Transactions on Image Processing*, vol. 5, págs. 996-1011.

Tsai R. Y. y T. S. Huang (1984): Multiframe image restoration and registration, *Advances in Computer Vision and Image Processing*, vol. 1, Greenwich, CT: JAI, 1984, págs. 317-339.

Ur H. y Gross D. (1992): Improved resolution from sub-pixel shifted pictures, *CVGIP: Graphical Models and Image Processing*, vol. 54, págs. 181-186.

Rajan D. y Chaudhuri S. (2001): Generalized interpolation and its application in super-resolution imaging, *Image and Vision Computing*, vol. 19, págs. 957-969.

## BIOGRAFÍA

### GUSTAVO IVÁN MARRERO CALLICÓ

Es Ingeniero de Telecomunicación por la ULPGC desde el año 1995 y se doctoró en 2003 en colaboración con Philips Research en Eindhoven (Países Bajos) obteniendo además el título de Doctor Europeo. Es profesor Contratado Doctor en la ULPGC desde 2006, desarrollando su labor investigadora en el seno de la División de Diseño de Sistemas Integrados (DSI) del IUMA. Investiga sobre algoritmos y plataformas *hardware* para incrementar la calidad en secuencias de vídeo, tanto en formato crudo como en entornos comprimidos.  
gustavo@iuma.ulpgc.es  
928 451 271

### SEBASTIÁN LÓPEZ SUÁREZ

Es Ingeniero Electrónico por la Universidad de La Laguna desde el año 2001, y en el año 2006 se doctoró en la Universidad de Las Palmas de Gran Canaria. Actualmente es Profesor Ayudante en la ULPGC, desarrollando su labor investigadora en el seno de la Di-

