

Súper – Resolución Aplicada a Imágenes Satelitales

Paula M. Tristán[†], Ruben S. Wainschenker[‡], Jorge H. Doorn^{*}

Resumen

En la actualidad muchas de las aplicaciones visuales requieren imágenes de alta resolución que permitan realizar una adecuada interpretación de los datos almacenados en las mismas, ejemplo de estas se encuentran en medicina, seguridad, imágenes satelitales, entre otras. En el campo de observación remota, la resolución espacial de las imágenes satelitales se ha ido mejorando con la puesta en órbita de plataformas con instrumentos de medición más precisos. No obstante, este aspecto también puede ser abordado utilizando métodos numéricos y aprovechando otras propiedades de las imágenes. Uno de estos se denomina Súper-Resolución. Este surgió ante el interés de mejorar la definición de videos de baja definición obtenidos a través de cámaras de seguridad de forma tal de permitir por ejemplo reconocer el rostro de una persona o la patente de un vehículo y actualmente se esta extendiendo su uso a otros ámbitos. Este trabajo presenta un algoritmo de *Super-Resolución* como un intento de mejorar la resolución espacial de imágenes satelitales Landsat 5 y 7.

Palabras Clave: *Teledetección, Sensores Remotos, Resolución espacial, Súper-Resolución.*

Abstract

Everyday, more and more visual applications require high resolution images to allow an adequate interpretation of the information embedded in them. Many examples can be found in medicine, security, remote sensing, among others. In the field of remote sensing, spatial resolution of satellite images has notoriously improved with the launch of new platforms with more precise instruments. However, further improvements are sometimes needed and can be obtained by numerical methods using other properties of the images. One of these methods, called hyper-resolution, is being currently explored to be used in this field. It was developed to increase the quality of security video cameras, in order to allow identification of human faces or to be able to read the license plate of a vehicle. This paper presents an algorithm to apply hyper-resolution, to increase the spatial resolution of satellite images. It was successfully applied to Landsat 5 and 7 images.

Keywords: *Remote Sensing, Spatial Resolution, Satellite Image, Hyper-Resolution*

1 Introducción

El desarrollo de la teledetección ha crecido vertiginosamente en los últimos 50 años, y se espera aún un mayor crecimiento en el futuro. Existen numerosas aplicaciones basadas en el análisis de imágenes satelitales que abarcan campos científicos y tecnológicos muy variados como cartografía, agricultura, forestación y logística militar, y prospección petrolera, entre otros [1].

En áreas como la agricultura, se requiere complementar la escala y la cobertura tradicional de los estudios de campo con nuevas herramientas que permitan observar territorios más amplios y más frecuentemente. Una buena alternativa son las imágenes satelitales ya que permiten

[†] INTIA, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires, Paraje Arroyo Seco, Campus Universitario (7000) Tandil Argentina. Becario Consejo Nacional de Investigaciones Científicas y Técnicas – Conicet - ptristan@exa.unicen.edu.ar

[‡]INTIA, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires, Paraje Arroyo Seco, Campus Universitario (7000) Tandil Argentina. rfw@exa.unicen.edu.ar

^{*}INTIA, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires, Paraje Arroyo Seco, Campus Universitario (7000) Tandil Argentina. jdoorn@exa.unicen.edu.ar

visualizar porciones amplias de territorio (del orden de cientos de kilómetros). No obstante, cuando se pretende utilizar este tipo de imágenes para algún estudio que requiera más nivel de detalle, indefectiblemente habrá que recurrir a imágenes de satélites de mayor resolución u otras fuentes provocando un aumento significativo en los costos.

Una de las principales características de los sistemas de teledetección es su resolución espacial la que se define como la menor separación entre dos objetos que puede diferenciar el sensor, este concepto está relacionado con el IFOV, (Instantaneous Field of View). El IFOV es el área de la superficie representada en cada pixel en una imagen satelital y por consiguiente está relacionado con la escala con la que se representan las imágenes [2].

Si el tamaño de la superficie representada por un pixel es grande comparado con la del objeto de interés, seguramente la señal capturada corresponderá a más de un tipo de cubierta, por ejemplo en una zona donde hay un sembrado próximo a un camino seguramente algún pixel estará afectado por el flujo emitido tanto desde del cultivo como del camino. Por el contrario cuanto menor sea la superficie representada, habrá más probabilidad de que cada pixel albergue una única cubierta [1]. Esta característica se supone una cuestión fundamental en la selección del sensor más conveniente para cada objetivo de estudio.

En los años 70 las imágenes de los primeros como Landsat tenían un IFOV de 79m x 79 m, en la actualidad con el avance de la tecnología se han logrado imágenes con un IFOV de 0,62m x 0,62m. No obstante aún para las imágenes con la máxima resolución espacial, siempre es una buena noticia avanzar un paso más. Persiste el problema de mejorar la resolución de aquellos satélites que no tienen buen IFOV y de mejorar la resolución de las imágenes capturadas anteriormente con sensores antiguos de baja resolución.

El resto del artículo se organiza como sigue. En la sección 2 se describe la técnica de súper-resolución, en la que se presenta el problema y la forma clásica de resolverlo y además se describe su aplicación a imágenes satelitales. En la sección 3 se explica el algoritmo propuesto para mejorar la resolución de imágenes satelitales, más adelante se presentan los resultados obtenidos utilizando imágenes de baja resolución y por último se plantean las conclusiones y trabajos futuros.

2 Estado del Arte

La técnica de súper-resolución se basa en aprovechar la información no redundante presente en una secuencia de imágenes. Cada imagen debe diferir respecto de las otras en un simple vector desplazamiento. Si hay correspondencia entre la secuencia de baja resolución en unidades enteras de pixeles, entonces cada imagen contendrá la misma información, solamente desplazada, con lo cual no existe nueva información que permita construir una imagen de mayor resolución. Pero si las imágenes tienen desplazamientos a nivel sub-píxel entre ellas, entonces existe información adicional y se puede mejorar la resolución [3], [4], [5], [6], [7],[9], [10], [11], [12].

2.1 El Modelo Estándar

El primer paso necesario para analizar el problema de súper-resolución es formular el modelo de observación que relaciona la escena original X con las imágenes de baja resolución Y_K para $1 \leq K \leq T$ observadas.

El tratamiento clásico del problema presenta la K -ésima imagen de baja resolución Y_K , obtenida por un sistema de captura estándar [3] como muestra la Fig. 1, como una transformación de la imagen original X afectada por diferentes fenómenos. Se puede definir entonces el modelo de la siguiente manera.

$$Y_K = D_K B_K M_K X + n_K \quad (1)$$



Fig. 1. Modo y problemas en la adquisición de una toma de video

La escena digital original se la puede definir como $X = [X_1, X_2, \dots, X_N]^T$ donde $N = L_1 N_1 \times L_2 N_2$ con N_1, N_2 definiendo la dimensión de cada Y_K y L_1, L_2 los factores de incremento de resolución. La matriz M_K , de dimensiones $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, representa las operaciones geométricas realizadas sobre la imagen original X que conducen a la imagen obtenida Y_K . Esta distorsión es causada por la óptica del sistema sensor (fuera de foco, límite de difracción, aberraciones, etc.). Es sabido que cuanto más se aparte del eje óptico, mayor será la deformación geométrica y cromática que introduce el sistema óptico sobre la imagen.

La matriz B_K representa el blur o borronamiento y es introducido básicamente por el tamaño del diafragma. Esta matriz esta definida de tamaño $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$.

La matriz D_K , de dimensión $(N_1 N_2) \times L_1 N_1 L_2 N_2$, representa el muestreo y la cuantificación establecidos por las características propias del CCD. Y por último n_K es el ruido presente en las imágenes de baja resolución.

2.2 El Modelo Planteado

Este modelo para imágenes satelitales tiene características especiales que lo diferencian del existente para una secuencia de video. En primer lugar en una secuencia de video hay un *frame* disponible cada 0.05 segundos mientras que las imágenes satelitales están disponibles cada varios días o semanas. Esta característica, denominada resolución temporal, conlleva a otros inconvenientes como el cambio de estado los objetos observados, el más común se da con los cultivos, cuestión que se desarrollará en secciones posteriores.

El modelo de observación presentado para imágenes satelitales que relaciona la escena original X con las distintas imágenes de baja resolución Y_K observadas es el definido por la siguiente ecuación.

$$Y_K = D_K M_K X + n_K \quad (2)$$

Este modelo tiene menos términos que la (1) debido al modo de captura, ya que la medición sobre cada pixel es independiente de los vecinos. A diferencia del modelo anterior, en este no existe la matriz B_K . Estos problemas en teledetección no se encuentran, ya que cada pixel representa la energía recibida desde la superficie apuntada por el sistema sensor. La matriz M_K , de dimensión $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, contempla en este caso, los errores introducidos por el sistema sensor a bordo del satélite durante la captura de cada pixel. La mayoría de estos problemas son corregidos por las estaciones receptoras, ya que antes de comercializar las imágenes satelitales se les realiza un procesamiento tal que elimina o corrige defectos incorporadas por el sensor, como movimientos y cambios de ángulos, por el satélite, como desviaciones y por la Tierra como rotación y curvatura. En la Fig. 2 se muestra el proceso de captura de imágenes del satélite ETM del Landsat.

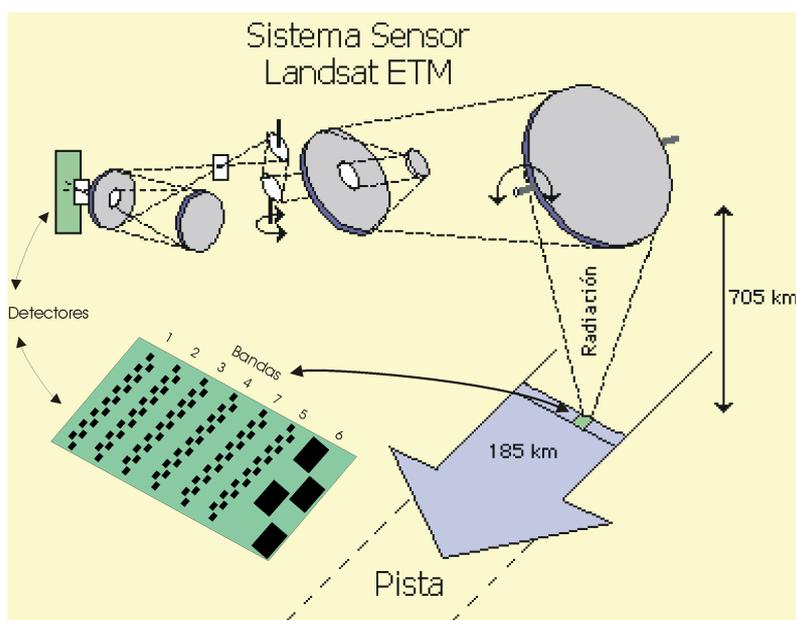


Fig. 2. Sistema de adquisición de Imágenes Satelitales Landsat

2.3 Traslaciones

Aplicar súper-resolución a una secuencia de video implica conocer el desplazamiento de los objetos presentes en cada cuadro de la secuencia. En imágenes satelitales el movimiento de los objetos de escena en escena se debe a características propias de la órbita y giro de cada satélite además de los movimientos propios, como alabeo y cabeceo, lo que hace imposible que un satélite observe dos veces exactamente igual una misma porción de territorio.

El movimiento entre dos imágenes satelitales implica realizar dos tipos diferentes de traslaciones, uno es a nivel pixeles enteros y otro es a nivel interior del pixel, es decir un movimiento inferior al tamaño de un pixel o simplemente sup-pixel.

2.3.1 Traslación a Nivel Pixel

Es la más intuitivo y simple, y se refiere a cuando existe una diferencia entre una imagen a otra de una cierta cantidad entera de pixeles. Por ejemplo, en la Fig. 3, el pixel $Y_1[10,15]$ se corresponde al $Y_2[12,8]$, esto significa que existe un movimiento entre estas imágenes definida por el vector $t = (2, -7)$. Esto se corrige con una traslación.

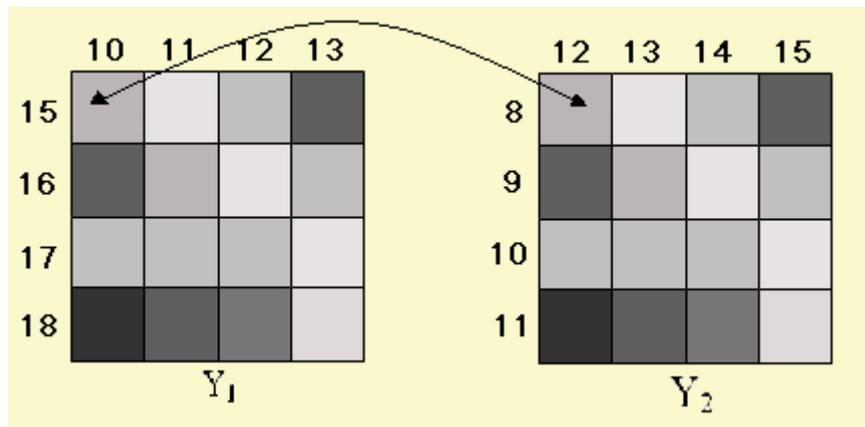


Fig. 3. Sector de un par de imágenes hipotéticas en la que se muestra la correspondencia de pixeles entre ambas. Traslación a nivel pixel entre dos imágenes diferentes

2.3.2 Traslación a Nivel Sub-Pixel

La Fig. 4 muestra dos cuadrículas que representan sectores de dos imágenes capturadas en diferentes momentos de una misma zona, Y_1 e Y_2 respectivamente, en ella se puede observar que el pixel $Y_2[47,40]$, representa una porción de territorio cuya intensidad representa la energía recibida desde esa porción de territorio en esa fecha. La imagen Y_1 muestra una posible distribución de pixeles en donde $Y_1[47,40]$, $Y_1[47,41]$, $Y_1[48,40]$ y $Y_1[48,41]$ recibirán cada uno parte proporcional de la energía emitida por la superficie correspondiente al pixel $Y_2[47,40]$ de la imagen Y_2 más otra parte correspondiente a la porción de terreno adyacente.

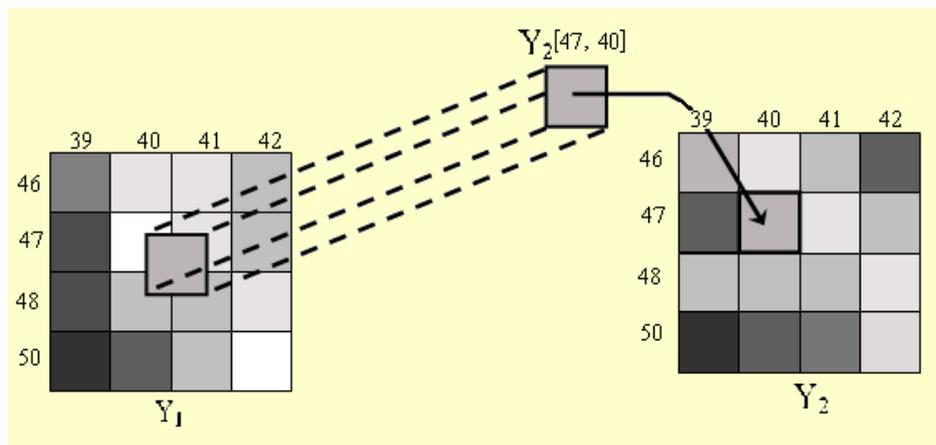


Fig.4. Sector de un par de imágenes hipotéticas en la que se muestra la correspondencia de pixeles entre ambas. Traslación a nivel sub-píxel entre dos imágenes

Esto es lo que se presenta como movimiento sub-píxel de una imagen satelital respecto de otra. Este movimiento no es conocido y necesita ser encontrado para poder aplicar la técnica de súper-resolución.

3 Algoritmo Propuesto

Como se mencionó en la sección anterior, no se conoce el movimiento existente de una imagen a otra posterior, la serie de satélites Landsat posee una resolución temporal de 16 días, esto significa que transcurrido este período el satélite pasará por el mismo lugar brindando así imágenes de la misma región, el movimiento existente entre estas es el mayor problema a abordar.

La Fig. 5 muestra un ejemplo del desplazamiento sub-píxel entre dos imágenes consecutivas Y_k e Y_{k+1} . El vector traslación está definido por los valores α y β , donde α define el desplazamiento del píxel en sentido horizontal y β define el desplazamiento del píxel en sentido vertical.

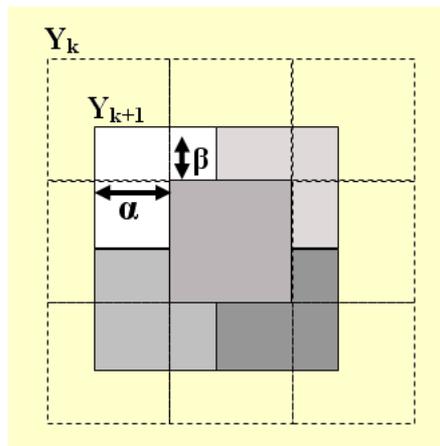


Fig. 5. Desplazamiento a nivel sub - píxel entre Y_k e Y_{k+1}

El algoritmo propuesto busca la mejor aproximación del vector de traslación de una imagen respecto a otra utilizando el concepto de correlación entre imágenes. Se prueba con varias combinaciones de α y de β y se elige aquel par que maximice la correlación.

El incremento de la resolución de una imagen a partir de un conjunto de éstas por medio de la técnica de súper-resolución implica la creación, a través de cálculos matemáticos, de una nueva imagen con más píxeles que la original. En otras palabras poner cuatro, nueve y quizás 16 píxeles de la misma área donde el medio de adquisición provee solo uno. Si existe sólo una imagen disponible, cualquier intento de obtener varios píxeles de uno adquirido producirá varios píxeles idénticos. Lo mismo sucedería si existen varias imágenes disponibles pero con desplazamientos en píxeles enteros. En el contexto de imágenes satelitales esto no sucede, ya que cuando se tienen varias imágenes existen indefectiblemente entre estas, desplazamientos relativos a nivel sub-píxel.

Al aplicar la técnica, cada píxel de la secuencia de imágenes debe ser dividido en cuatro o nueve píxeles conteniendo cada uno de ellos exactamente la misma información que el píxel

original. Es importante destacar que las intensidades de cada una de las imágenes utilizadas en este proceso han sido todas previamente normalizadas para evitar diferencias estacionales ocasionales.

Este proceso se aplica sobre una imagen tomada como original y se “trasladan” respecto de ella todas las imágenes restantes llevadas al nuevo muestreo, así hasta obtener la imagen de salida con una resolución mayor a las imágenes utilizadas durante este proceso.

Podemos expresar entonces el método de la siguiente manera: Sea ${}^F Y_i^L$ la fracción de la imagen Y_i que contiene la región de interés con un factor L de incremento en el muestreo y trasladada a nivel pixel respecto de la imagen inicial ${}^F Y_0$. Podemos expresar la salida ${}^F X_n$ como sigue:

$${}^F X_0 = {}^F Y_0^L \quad (3)$$

$${}^F X_{n-1} = \frac{1}{n} \sum_{i=1}^n T_{\alpha^i \beta^i} ({}^F Y_i^L) \quad (4)$$

Donde $T_{\alpha\beta} ({}^F Y_i^L)$ representa la traslación a nivel sub-pixel de ${}^F Y_i^L$ en (α^i, β^i) respecto de ${}^F Y_i^L$ utilizando interpolación bilineal, los valores de α^i y de β^i son aquellos que maximizan la correlación entre Y_0 y $T_{\alpha\beta} ({}^F Y_i^L)$.

4 Resultados

Debido a que algunas cubiertas cambian su comportamiento espectral con el correr del tiempo, esta técnica se probó sobre imágenes con objetos invariantes en el tiempo dentro de un contexto con baja variación. Ejemplo de estos son caminos, rutas y construcciones edilicias.

En la Fig. 6 se muestra una secuencia de 10 imágenes Landsat de un sector de una pista de aterrizaje y varias construcciones edilicias típicas de un aeropuerto como hangares, oficinas, etc.

Aplicando las ecuaciones (3) y (4) a la secuencia de imágenes en el orden mostrados en la Fig. 6 se obtiene en pasos sucesivos la secuencia mostrada en la Fig. 7 donde cada imagen ilustra el resultado que se obtiene al incorporar una nueva imagen a la resultante del paso anterior.

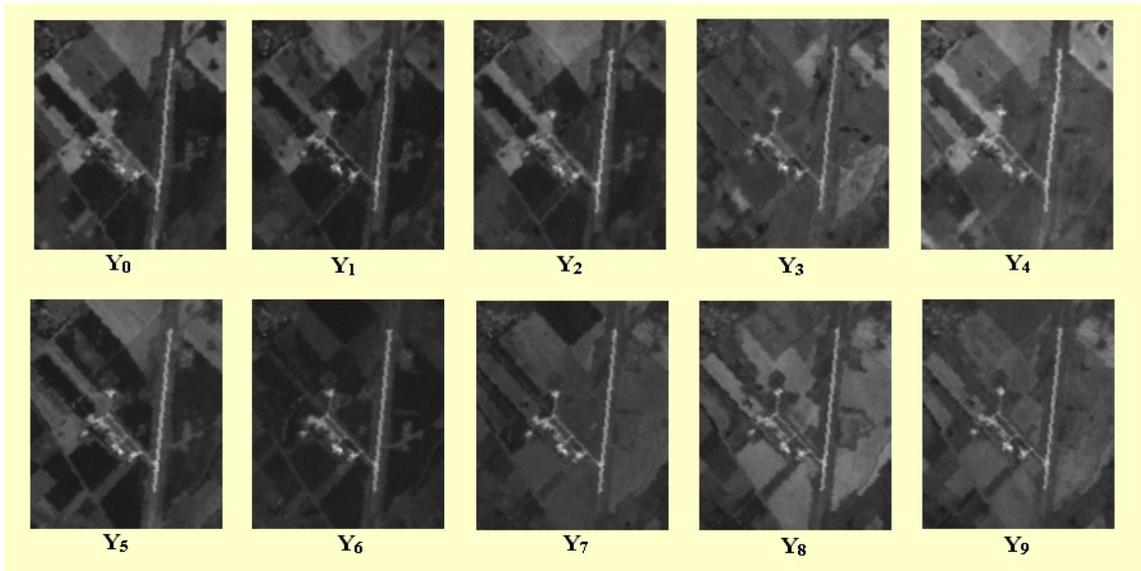


Fig. 6. Imágenes satelitales correspondientes a diferentes fechas obtenidas para el mismo sector con el muestreo original

En esta figura, a simple vista, se puede observar claramente las mejoras obtenidas dentro de la secuencia.

La forma de determinar el valor óptimo de L no ha sido estudiada aún, sin embargo algún conocimiento en este aspecto se ha ganado. Si el valor de L elegido es demasiado grande, no causará ningún problema en la calidad de la imagen creada y el único precio pagado será el costo de memoria y procesamiento. Por otro lado, la utilización de un L pequeño reduce potencialmente el incremento de calidad de la imagen creada utilizando súper-resolución. Para el caso estudiado aquí con imágenes satelitales donde el número de imágenes disponibles es pequeño en la mayoría de los casos el valor $L = 2$ es suficiente.

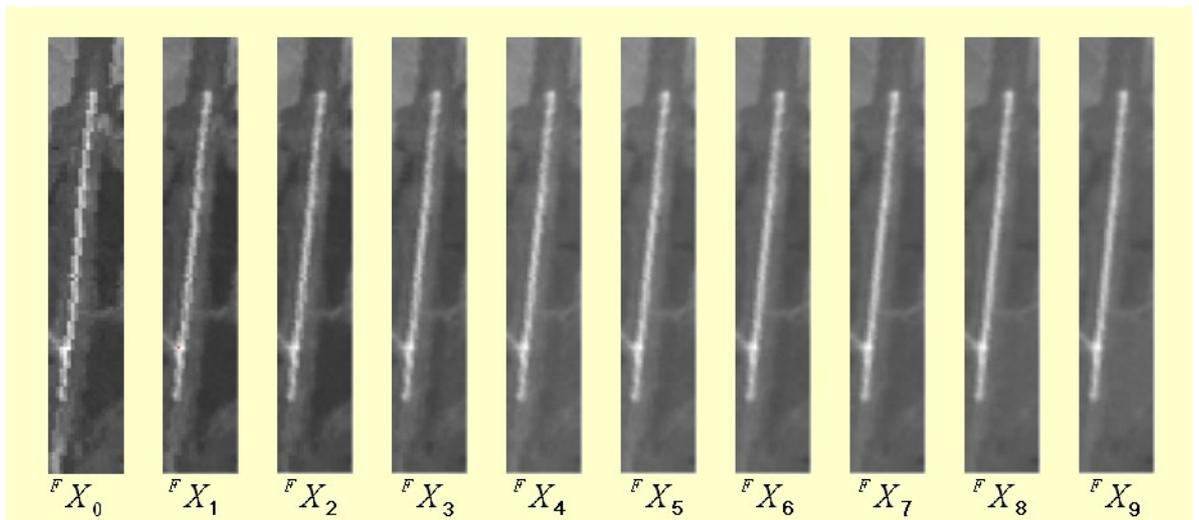


Fig. 7. Resultados sucesivos al aplicar súper-resolución a la pista de aterrizaje y sus alrededores. Este ejemplo tiene un factor de incremento del doble de resolución que las imágenes satelitales originales

Una vez obtenidos los resultados de la técnica de súper-resolución a imágenes satelitales utilizando el fundamento de correlación de imágenes surge la necesidad de medir estos resultados de forma de corroborar lo que visualmente queda más que claro. Para tratar de cuantificar esta mejora se evaluaron varias alternativas. En este caso se decidió analizar el borde de la pista, y a partir de éste, la evolución del perímetro y la longitud del mismo.

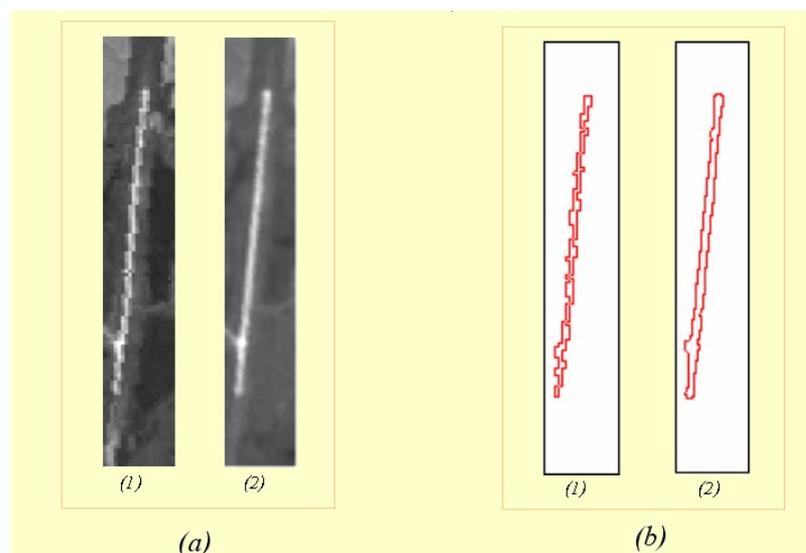


Fig. 8. (a) Muestra la imagen original (1) y la imagen resultante (2). (b) Muestra los bordes de las imágenes mostradas en (a).

La comparación de las Fig. 8a (1) y (2) representan respectivamente, la imagen original con el nuevo muestreo y la resultante al aplicar la técnica. Por otro lado en las Fig. 8b (1) y (2) se muestran los respectivos bordes de la pista obtenidos aplicando el mismo umbral, en esta se ve claramente más recto el borde de la pista Fig. 8a (2), es decir no se visualiza el escalonamiento que se percibe en la original Fig. 8a (1).

Las imágenes originales utilizadas tienen una resolución espacial de 30m x 30m mientras que la imagen obtenida como resultado fue construida a partir de píxeles de 15m x 15m, lo que en la práctica significa una cuasi duplicación de la resolución.

No obstante se midió también la longitud del perímetro en cada salida y se comparó con el valor real de 5196m. En la Tabla 1 se puede observar la evolución del perímetro calculado con la imagen mejorada en cada etapa. Se observa que el porcentaje del error relativo disminuye con cada paso.

Tabla 1. Muestra los valores de la longitud del perímetro de la pista para todos los resultados intermedios

| | Perímetro SR (m) | Error (%) |
|----------|------------------|-----------|
| salida 1 | 5713,4 | 10 |
| salida 2 | 5724,3 | 10 |
| salida 3 | 5623,4 | 8,2 |
| salida 4 | 5619,7 | 8 |
| salida 5 | 5511,5 | 6 |
| salida 6 | 5518,8 | 6,2 |
| salida 7 | 5461,8 | 5 |

| | | |
|-----------------|--------|-----|
| salida 8 | 5469,1 | 5 |
| salida 9 | 5423,0 | 4,3 |

5 Conclusiones

Se ha confirmado lo sugerido por diferentes autores, esto es, la súper-resolución se puede aplicar con éxito a imágenes satelitales [3], [6], [7].

Se ha comprobado también que existen particularidades en el algoritmo a la hora de tratar con imágenes satelitales en vez de una secuencia de video.

Se concluye que las mejoras en la calidad de la imagen son evidentes con la sola observación de los resultados. La cuantificación de la mejora corrobora los resultados.

Referencias

- [1] Emilio Chivieco. Fundamentos de la Teledetección Espacial 3era Edicion. Ediciones Rialp S.A Madrid 1996.
- [2] Jensen J. R. Introductory Digital Image Processing. 2th Edition Prentice Hall, New Jersey, 1996.
- [3] S.C. Park, M.K. Park and M.G. Kang. "Super-Resolution Image Reconstruction: A Technical Overview". IEEE Signal Processing Magazine. Vol 20(3), 2003.
- [4] C.A. Segall, R. Molina, and A. K. Katsaggelos y J. Mateos. Reconstruction of high-resolution images frames from a sequence of low-resolution and compressed observations. IEEE International Conference on Acoustics, Speech, and Signal Processing. Vol.2 1701-1704, 2002.
- [5] C. A. Segall, A. K. Katsaggelos, R. Molina and J. Mateos. "Bayesian Resolution Enhancement of Compressed Video". IEEE Transactions on Image Processing. Vol 13. (7) 898-911, 2004.
- [6] M. Elad and A. Feuer. Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images. IEEE Transactions on Image Processing, Vol. 6 (12), 1997.
- [7] B.R. Hunt and Olaf Kübler. Karhunen-Loeve Multispectral Image Restoration, Part I: Theory. IEEE Transactions on Acoustic, speech, and signal processing, Vol ASSP-32, 1984.
- [8] S. Baker and T. Kanade. "Limits on Super-Resolution and How to Break Them". IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol 24 (9) 1167-1183, 2002
- [9] T. Akgun, Y. Altunbasak and R. Mersereau. Super-resolution Reconstruction of Hyperspectral Images. IEEE Transactions on. Vol 14 (11) 1860-1875, 2005.

- [10] Sean Borman, Robert Stevenson. Spatial Resolution Enhancement of Low Resolution Images Sequences. A Comprehensive Review with Directions for future Research.
- [11] P. Vandewalle, S. Susstrunk and M. Vetterli. Superresolution images reconstructed from aliased images. LCAV – School of Computer and Communications Sciences Ecole Polytechnique Fédérale de Lausanne. Ch. 1015 Switzerland.
- [12] R. Schultz y R. Stevenson. Extraction of High-Resolution Frames from Video Aequence. IEEE Transactions on Images Processing. Vol. 5 (6), 1996.