

ESTUDIO DE VIABILIDAD DEL USO DE IMÁGENES COMPRIMIDAS EN PROCESOS DE CLASIFICACIÓN

C. Pérez, D.G. Aguilera y A.L. Muñoz

carpegu@usal.es

Dpto. Ingeniería Cartográfica y del Terreno · Universidad de Salamanca
C/ Santo Tomás s/n. 05003 Ávila

RESUMEN: Se analiza la influencia que las diferentes técnicas de compresión tienen en la calidad y procesamiento de las imágenes de Teledetección. En concreto, la comunicación examina el efecto de la compresión de imágenes multispectrales de cara a su posterior clasificación. Se desarrolla una metodología sistemática tanto de compresión como de clasificación, que permite extraer resultados objetivos sobre la degradación de la clasificación con el uso de imágenes comprimidas.

ABSTRACT: This paper analyses how compressed imagery modifies the quality and processing of remote sensing imagery. In particular, the study examines effects of compression on multispectral images in labelling tasks. A methodology is developed in compression and classification algorithms which allows the extraction of objective results about the degradation of images inside the application context.

Palabras clave: compresión, clasificación, Landsat, teledetección.

INTRODUCCIÓN

En la actualidad, son multitud las líneas de investigación que se abren para intentar determinar la influencia que las técnicas de compresión tienen en la calidad y procesamiento de las imágenes satelitales. En el campo de la Teledetección esto constituye básicamente dos objetivos:

1/ Optimizar el envío de la señal (imágenes) comprimiendo los datos on-board previo a su envío a Tierra. Ello implica que el satélite dispone de un lapso de tiempo para comprimir las imágenes antes de su descarga (lo cual no presenta problemas y es un método perfectamente implementado desde largo tiempo), pero redundante en que en el mismo ancho de banda se puedan enviar más bandas espectrales en el mismo tiempo. De gran utilidad en la teledetección hiperspectral, por citar un ejemplo, donde el satélite puede cubrir un amplio número de bandas (recogiendo centenares de ellas), pero existe un cuello de botella en su transmisión (enviando solamente decenas de ellas) a Tierra.

2/ Mejorar las operaciones que involucren transmisión de datos y almacenamiento dentro de una arquitectura Cliente-Servidor. La conjunción de Internet y Teledetección impone minimizar el tamaño y estructura de las imágenes respetando la calidad radiométrica de las mismas. De utilidad para la descarga de imágenes en tiempo real, o para el procesamiento distribuido de las mismas tal cual se vaticina para el futuro: Las imágenes pueden almacenarse en un ordenador central y ser manipuladas y procesadas en

terminales distribuidos con independencia de la distancia geográfica.

Asumiendo las dos premisas anteriores, el presente trabajo se centra en la viabilidad o no, de la compresión de imágenes en un proceso de clasificación supervisada.

ALGORITMOS DE COMPRESIÓN

Nos apoyamos en diversos algoritmos de compresión, que se subdividen en:

Compresión Sin-Perdida (lossless): Técnica que no implica pérdidas de información respecto al fichero original pero disminuye su tamaño. La información contenida en los ficheros es preservada en todo momento, optimizándose la estructura de almacenamiento para hacerla más eficaz. Algoritmos típicos de compresión lossless son: *Run Length Encode* (RLE), *LZ* y *LZW* usados en formatos tan conocidos como ZIP, GZIP, PKZIP, y en formatos de imágenes TIFF, GIF y PNG. En general, las técnicas sin pérdida consiguen bajos ratios de compresión, lo que repercute en archivos relativamente grandes.

Compresión Con-Pérdida (lossy): Técnica que modifica la estructura y la radiometría de la imagen para que ocupe menor volumen. Produce pérdida de información y la reconstrucción de un fichero comprimido por este método no es exactamente el fichero original. Se consiguen ratios más elevados de compresión que en la compresión sin pérdida.

Algunas técnicas de compresión *lossy* se basan en la Transformada Discreta del Coseno (DCT), teoría Wavelet y Teoría Fractal.

Los formatos de compresión elegidos para hacer el estudio, fueron tanto sin pérdida (*TIFF-LZW* y *LuraWave-Lossless*) como con pérdida (*MrSID*, *LuraWave-Lossy*, *JPG2000* y *JPG*)

Cada uno de ellos presenta características propias en el desarrollo de su algoritmia, lo que confiere diferentes peculiaridades tanto estéticas como cuantitativas en la imagen comprimida resultante. Como ejemplo, *JPG* basa su fundamento en la Transformada de Fourier a bloques de 8x8 píxeles. Ello repercute en una buena relación de compresión de la imagen; pero en contra, es un formato de compresión destructivo (*lossy*), que presenta un 'efecto bloque' o mosaico, resultante de la división de la imagen en los bloques de píxeles, y de la gran diferencia existente entre los coeficientes de cuantificación (escalón).

El acercamiento basado en Wavelets (propio de *MrSID*, *JPG2000* o *LuraWave*) parametriza la imagen en variables de frecuencia, energía y tiempo, para posteriormente estudiar cada componente con una resolución acorde a la escala. El nivel de detalle con que la imagen se discretiza no es constante; algo lógico, pues no será necesaria la misma cantidad de información para describir un campo de cultivo (área homogénea), que para observar una zona urbana (área heterogénea).

MÉTODO DE TRABAJO

Ubicación y selección de la zona de trabajo.

Para el estudio se seleccionó un área lo suficientemente extensa y geográficamente diversa como para contrastar la presencia de diferentes categorías de usos del suelo.

En función de la disponibilidad de imágenes y documentación complementaria, de la proximidad y el conocimiento del territorio objeto de estudio, se decidió delimitar una zona de trabajo que abarcaba el conjunto de las superficies correspondiente a cuatro hojas del Mapa Topográfico Nacional a escala 1:50.000 centradas sobre la capital de Ávila. Hojas 505, 506, 530 y 531.

El espacio geográfico comprende, en su parte más meridional, la fosa del Valle Amblés y las estribaciones septentrionales de la Paramera. La zona central y superior de la imagen de satélite reflejan el zócalo paleozoico en la Sierra de Ávila y los materiales sedimentarios de la comarca de la Moraña.

La imagen de partida para el estudio fue la

escena 202-32 del satélite Landsat 7 fechada en agosto de 1999, con cobertura nubosa nula.

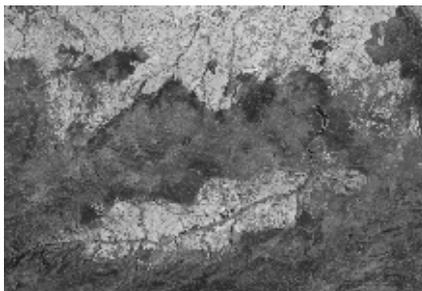


Figura 1._ Imagen Landsat ETM+ de la zona de estudio.

Georreferenciación y recorte

En primer lugar se seleccionaron como bandas de trabajo (*inputs* en la georreferenciación, en la clasificación y en la compresión) todas las proporcionadas por el sensor ETM+ excepto las correspondientes al infrarrojo térmico y al pancromático por ofrecer resoluciones espaciales no homogéneas con el resto de bandas del espectro óptico.

Se georreferenció la imagen utilizando puntos de control procedentes de cartografía oficial y empleando un ajuste polinómico de tercer grado con remuestreo por vecino más próximo.

Se consiguió un error medio cuadrático por debajo del tamaño del píxel que garantizaba la correcta geometría y la posibilidad de superposición con capas ráster y vectoriales (recintos de usos del suelo del proyecto CORINE Land Cover)

Compresión de las imágenes

El recorte georreferenciado se transformó a formato TIFF, asumiéndolo como formato de intercambio estándar entre los softwares específicos de compresión de imágenes, y los de clasificación.

A partir de las imágenes originales se generaron tres grupos de imágenes por cada tipo de compresor, utilizando tres ratios de compresión que permitieran contrastar la cantidad de información perdida con la disminución del tamaño de los ficheros. Los porcentajes de compresión escogidos fueron del 100% (denominados A), 75% (B) y 50% (C) respectivamente.

Se consideró a priori, que emplear ratios más allá del 50%, supondría pérdidas radiométricas demasiado acentuadas para los propósitos de la clasificación.

Los porcentajes de compresión denotan una mayor reducción del tamaño del fichero original cuanto menor sea el porcentaje referido. Es decir, la

menor compresión la presenta el 100% y la mayor, de entre los seleccionados, el 50%.

En resumen, se dispone finalmente de doce conjuntos de imágenes comprimidas más el original (véase Tabla 1).

Original	13.2 MB		
Sin-Pérdida (Lossless)			
LZW	11.0 MB (0.83)		
Lurawave	8.57 MB (0.64)		
Con-Pérdida (Lossy)			
	A (100%)	B (75%)	C (50%)
JPG	9.6 MB (0.72)	2.14 MB (0.16)	1.37 MB (0.10)
Lurawave	Vease Lura_lossless	0.05MB (0.01)	0.03MB (0.01)
MrSID	6.34MB (0.47)	1.12 MB (0.08)	0.63 MB (0.01)
JPG2000	7.95 MB (0.60)	0.05 MB (0.01)	0.03MB (0.01)

Tabla 1.- Tamaños de las imágenes (en MB) para cada porcentaje de compresión (A, B, C) y ratio de compresión (en tanto por 1).

Clasificación

Para cada uno de los grupos de imágenes multibanda, se realizaron diversas pruebas de clasificación no supervisada utilizando el algoritmo ISODATA, con el fin de obtener una primera aproximación sobre las categorías informacionales susceptibles de ser segmentadas.

A continuación, se optó por utilizar como método de clasificación el supervisado apoyado en el algoritmo de Máxima Probabilidad y con estableciendo la clase nula.

Con este procedimiento, se definieron, inicialmente, campos de entrenamiento para un total de veinte clases apoyándose en la leyenda del proyecto CORINE. Para en el proceso de postclasificación subsiguiente, reagruparlas en las ocho clases: superficies de agua, zonas urbanas, pastizal-matorral, coníferas, tierras de labor en seco, suelos desnudos, y encinares

Con objeto de no perjudicar las posteriores comparaciones entre las imágenes clasificadas no se aplica filtro alguno que suavice los resultados, o que generalice los contornos de los recintos delimitados por el algoritmo de clasificación

Tras la clasificación de la imagen original se procede, de igual modo, a la clasificación de cada uno de los grupos de imágenes comprimidas. En esta fase del trabajo se utilizan los mismos campos de entrenamiento en todos los grupos de imágenes

Comparación de los resultados

Con todas las imágenes reagrupadas en las ocho categorías mencionadas se procede a la comparación de los resultados.

Se confeccionaron matrices de confusión que permitan cruzar cada una de las doce clasificaciones obtenidas a partir de grupos de imágenes comprimidas con la clasificación de la imagen original. Véase Tabla 2.

Del análisis de dichas matrices se dedujeron los resultados y conclusiones que se detallan en el apartado siguiente.

	A	B	C
JPG	97.1%	70.41%	67.54%
Lurawave	100%	52.12%	48.31%
MrSID	90.91%	68.27%	63.07%
JPG2000	93.14%	56.02%	50.85%

Tabla 2.- Fiabilidad de las clasificaciones en valores porcentuales para cada grupo de imágenes.

Cada valor es el resultado del sumatorio de la diagonal ($\sum a_{ii}$) de la matriz de confusión.

CONCLUSIONES

Se desaconseja la utilización de imágenes con compresión media (B) o alta (C). El tamaño de las imágenes se reduce sustancialmente con estos niveles de compresión, pero también lo hace la fiabilidad de la clasificación. Para la compresión B y C, JPG_B es el que menor variación produce con respecto a la clasificación original. Aún así, 3 de cada 10 píxeles resultan mal clasificados. En el peor de los casos (JPG2000_C y LuraWave_C) se llega a fiabilidades globales en el que 1 de cada 2 píxeles son erróneamente asignados.

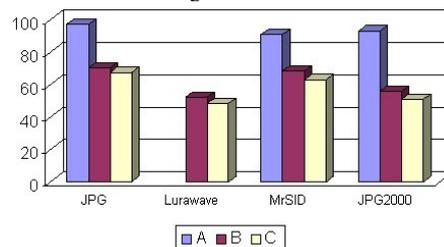


Gráfico 1.- Fiabilidad resultante para cada ratio de compresión de los diferentes formatos.

Como compresor más eficiente para los cometidos aquí reseñados se encuentra el Lura_A que al ejecutar una compresión sin pérdida reduce la imagen a ratios similares a los de JPG y JPG2000 pero no altera, como es lógico, en nada la clasificación.

En contrapartida las compresiones Lura_B

y Lura_C, que sí tienen pérdida, no resultan nada útiles siendo el peor compresor de todos al mostrar una gran discrepancia (entorno al 50%) con respecto al original.

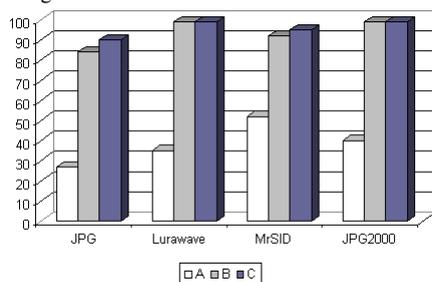


Gráfico 2.- Porcentaje de reducción de tamaño de fichero para cada grupo de imágenes.

Una conclusión destacable es la evolución de la compresión *JPG* al nuevo formato *JPG2000*. Este último conserva mejor la calidad visual (estéticamente) de la imagen original, disminuyendo de forma simultánea el ratio de compresión de la misma. Sin embargo de cara al objeto de nuestro estudio cuantitativo (consecuencias de la compresión en la clasificación), la evolución del nuevo formato *JPG* implica resultados negativos empeorando de modo sustancial la fiabilidad de la clasificación obtenida con el formato más moderno.

A raíz de los resultados, y desechado B y C por su mala fiabilidad, resulta que las imágenes apenas se reducen en tamaño (0.47 en el mejor de los casos), pero modifican significativamente su fiabilidad global.

Por ello se cree conveniente recomendar el uso generalizado de algoritmos sin pérdida, en este caso *LuraWave-Lossless* (Lura_A), pues reducen el tamaño de las imágenes en similares ratios a los formatos de compresión con pérdida pero con la ventaja de no alterar en absoluto los resultados finales de la clasificación.

En cualquier caso, debe valorarse que la compresión de las imágenes multiespectrales se ha llevado a cabo sobre una zona básicamente rural caracterizada por la predominancia de tonos continuos. Ello conduce a que el algoritmo tradicional *JPG*, basado en la Transformada de Fourier, se adapte mejor que los algoritmos de compresión Wavelets, los cuales son más eficientes en zonas heterogéneas, caracterizadas por unas marcadas discontinuidades espectrales (zonas urbanas). Esto es, si la zona de elección hubiera sido eminentemente urbana, los resultados del estudio hubieran variado considerablemente.

Cabe matizar también, que los resultados de este estudio pueden variar significativamente de acuerdo al tamaño físico de la zona de estudio. Así, la ampliación superficial de la imagen conllevaría una mejora en los ratios de compresión mediante una progresión geométrica, mientras que la pérdida de fiabilidad en la clasificación aumentaría en proporción aritmética.

LÍNEAS FUTURAS

Es importante tener presente que la compresión de dichas imágenes se ha realizado banda a banda por no ser adecuados estos algoritmos para matrices *n*-espectrales (excepto $n=1$ ó $n=3$).

Se plantea la necesidad de proseguir los estudios de este tipo con algoritmos de compresión tridimensional (de carácter fractal principalmente). En ellos se asume como característica básica de las imágenes multiespectrales la alta correlación existente entre bandas próximas, como factor propiciatorio de un mayor ratio de compresión de imágenes.

La utilización de otro tipo de imágenes con resoluciones distintas y sobre zonas en las que puedan definirse clases informacionales diferentes a las de la zona estudiada ampliarían el alcance de este trabajo.

BIBLIOGRAFÍA

- Kiema, J.B.K. 2000. *Wavelet Compression and the Automatic Classification of Landsat Imagery*. Photogrammetric Record, 16(96): 997-1006.
- Kiema, J.B.K., BÄHR, H.P. 2001. *Wavelet Compression and the Automatic Classification of Urban Environments Using High Resolution Multispectral Imagery and Laser Scanning Data*. GeoInformática 5:2, 165-179.
- Martín Marcos, A. 1999. *Compresión de Imágenes JPEG*. Editorial Ciencia 3 S.L.
- Unser, M., Aldroubi, A., Line, A. 1999. *Wavelet Applications in Signal and Image Processing*. SPIE The International Society for Optical Engineering. 19-23 July 1999, Denver, Colorado.
- Watt, A., Policarpo, F. 1998 *The Computer Image*. Editorial Addison-Wesley.