

# ANÁLISIS DE COMPONENTES PRINCIPALES.

## APLICACIONES EN TELEDETECCIÓN

<b>Apellidos, nombre</b>	Ruiz Fernández, Luis Ángel (laruiz@cgf.upv.es)
<b>Departamento</b>	Ingeniería Cartográfica, Geodesia y Fotogrametría
<b>Centro</b>	ETSI. Geodésica, Cartográfica y Topográfica



## 1 Introducción

Cuando se trabaja con una elevada cantidad de datos y variables, una de las principales decisiones que el analista debe hacer es la selección de las variables que optimizan la solución a un determinado problema. En ocasiones, un planteamiento teórico inicial y un estudio exhaustivo de las variables puede conducir a la selección adecuada de los datos. Sin embargo, cuando se trabaja con numerosas variables, como es el caso de los datos suministrados por los sensores multiespectrales e hiperespectrales (estos últimos con información de hasta cientos de longitudes de onda diferentes), la correlación existente entre ellas puede ser muy elevada y el estudio, análisis y comprensión de todas las correlaciones posibles resultaría inviable.

En este tema se describe y analiza el procedimiento estadístico de análisis multivariante conocido como análisis de componentes principales, concebido fundamentalmente como un método de reducción de datos. Se tratarán, inicialmente, los aspectos teóricos del método, particularizándolo a nuestro tipo de datos, las imágenes, y de variables, la respuesta espectral en las distintas longitudes de onda del espectro electromagnético. Finalmente, se repasarán las principales aplicaciones en el campo de la teledetección. El esquema según el cual se va a desarrollar el tema es el siguiente:

Objeto>Datos> Procedimiento> Interpretación geométrica> Ejemplo> Aplicaciones

## 2 Objetivos

Una vez que el estudiante lea con detenimiento este documento, será capaz de:

- Entender en qué consiste el método de análisis multivariante de análisis de componentes principales y cómo se aplica en imágenes con varias bandas.
- Interpretar el resultado de aplicar la transformación de componentes principales sobre imágenes multiespectrales.
- Conocer y entender el uso y aplicaciones fundamentales de esta técnica en teledetección.

## 3 Desarrollo

### 3.1 Objeto

El método de análisis de componentes principales consiste en la transformación de un conjunto original de variables en otro conjunto de variables o componentes, obtenidas mediante **combinación lineal** de las anteriores, de forma que los primeros nuevos componentes generados expliquen el **máximo** de la **variabilidad** total, mientras que el resto vaya explicando menor cantidad de varianza de forma sucesiva. Además, las sucesivas combinaciones lineales se extraen de forma que **no** exista **correlación** entre ellas.

Así pues, el objetivo principal del análisis de componentes principales es el de determinar nuevos factores (componentes principales) que expliquen la mayor



cantidad de la variabilidad existente en los datos utilizando, para ello, el menor número de factores posible.

### 3.2 Datos

Partimos de una matriz de datos  $X$ , de  $n$  individuos por  $p$  variables, donde  $X$  podría representar una imagen multiespectral con  $n$  píxeles y  $p$  bandas espectrales, y un elemento genérico de dicha matriz,  $x_{ij}$ , representaría el nivel de gris del píxel  $i$  en la banda  $j$ . De esta forma, la matriz se puede descomponer en  $p$  vectores columna, donde cada uno de ellos representa una banda espectral de la imagen.

$$X = (x_1, \dots, x_j, \dots, x_p) = \begin{pmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1p} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i1} & \dots & x_{ij} & \dots & x_{ip} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & \dots & x_{nj} & \dots & x_{np} \end{pmatrix} \quad (1)$$

Además, sea  $C$  la matriz de varianzas y covarianzas de  $X$ , matriz cuadrada cuyas dimensiones coinciden con el número de variables o bandas de la imagen:

$$C = \begin{pmatrix} \sigma_{11}^2 & \dots & \sigma_{1p}^2 \\ \dots & \dots & \dots \\ \sigma_{p1}^2 & \dots & \sigma_{pp}^2 \end{pmatrix} \quad (2)$$

el término general de la matriz de covarianzas para dos variables genéricas  $j$  y  $j'$  vendrá dado por la expresión

$$\sigma_{jj'}^2 = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \mu_j)(x_{ij'} - \mu_{j'}) \quad (3)$$

donde  $\mu_j$  es la media de los  $n$  individuos para la variable genérica  $j$ :

$$\mu_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (4)$$

### 3.3 Procedimiento

Dadas las variables iniciales  $x_1, \dots, x_j, \dots, x_p$ , las nuevas variables o componentes principales se obtienen combinando linealmente las variables originales de forma que satisfagan dos condiciones:

- 1) que **maximicen la varianza**, (de forma que cada componente sucesivo explique el máximo de variabilidad posible).



2) que sean **ortogonales** entre sí. (condición para que no exista correlación entre componentes).

Llamando  $z_f$  a las nuevas variables ( $f = 1, \dots, p$ ),  $\bar{z}_f$  a sus respectivos vectores de datos y  $\bar{u}_f$  a sus respectivos vectores de coeficientes, tenemos que

$$\bar{z}_f = X \bar{u}_f \quad (5)$$

$(n \times 1)$      $(n \times p)$      $(p \times 1)$

donde:

$$\bar{u}_f = \begin{pmatrix} u_{1f} \\ \dots \\ u_{pf} \end{pmatrix} ; \quad \bar{z}_f = \begin{pmatrix} z_{1f} \\ \dots \\ z_{nf} \end{pmatrix}$$

es decir, cada nuevo componente vendrá dado por la expresión:

$$z_f = x_{i1} u_{1f} + x_{i2} u_{2f} + \dots + x_{ip} u_{pf} \quad (6)$$

La varianza del conjunto de datos representado por la matriz X y proyectado sobre la dirección  $\bar{u}$  viene dada por

$$\text{Varianza}(X\bar{u}) = \bar{u}' C \bar{u} \quad (7)$$

Así pues, el problema que se plantea es el de determinar el eje  $\bar{u}$  que pasa por el origen y para el cual es máxima la varianza del conjunto de datos proyectado. Matemáticamente, debemos hallar los p componentes  $u_1, \dots, u_j, \dots, u_p$  de

un vector unitario  $(\sum_{j=1}^p u_j^2 = 1)$  de tal forma que maximice la expresión:

$$\bar{u}' C \bar{u} \quad (8)$$

con la restricción  $\bar{u}' \bar{u} = 1$  (\*) (vector unitario) (9)

**- Cálculo del primer componente principal:**

---

\* (NOTA: esta forma cuadrática aumentaría indefinidamente si  $\bar{u}$  fuera arbitrariamente grande, por lo que se limita su longitud a la unidad)



Llamando  $\bar{z}_1 = X\bar{u}_1$  al primer componente principal y  $\bar{u}_1$  a su vector de coeficientes, las coordenadas de  $\bar{u}_1$  pueden calcularse por el método de los multiplicadores de Lagrange, considerando la función  $h(\bar{u}_1, \lambda_1)$ :

$$h(\bar{u}_1, \lambda_1) = \bar{u}_1' C \bar{u}_1 - \lambda_1 (\bar{u}_1' \bar{u}_1 - 1) \quad (10)$$

derivando e igualando la función a cero,

$$\begin{aligned} \frac{\partial h}{\partial \bar{u}_1} &= 2C\bar{u}_1 - 2\lambda_1 \bar{u}_1 = 0 \\ C\bar{u}_1 &= \lambda_1 \bar{u}_1 \end{aligned} \quad (11)$$

por lo que  $\bar{u}_1$  es el primer vector propio de C, y  $\lambda_1$  su primer valor propio.

Premultiplicando por  $\bar{u}_1'$ , se obtiene

$$\bar{u}_1' C \bar{u}_1 = \lambda_1 \bar{u}_1' \bar{u}_1 = \lambda_1 \quad (12)$$

por lo que dicho valor propio representa la varianza máxima buscada.

#### **- Cálculo del segundo componente principal (y sucesivos):**

El segundo componente principal,  $\bar{z}_2 = X\bar{u}_2$ , se calcula de forma análoga, añadiendo una nueva condición:

$$\bar{u}_1' \bar{u}_2 = 0 \quad (13)$$

esto es, ambos vectores son ortogonales.

Por consiguiente, buscaremos  $\bar{u}_2$  maximizando la expresión  $\bar{u}_2' C \bar{u}_2$ , con las condiciones  $\bar{u}_2' \bar{u}_2 = 1$  y  $\bar{u}_1' \bar{u}_2 = 0$ .

Considerando nuevamente una función  $h(\bar{u}_2, \lambda_2)$ , sujeta a las dos condiciones mencionadas:

$$\begin{aligned} h(\bar{u}_2, \lambda_2) &= \bar{u}_2' C \bar{u}_2 - \lambda_2 (\bar{u}_2' \bar{u}_2 - 1) - \beta \bar{u}_1' \bar{u}_2 \\ \frac{\partial h}{\partial \bar{u}_2} &= 2C\bar{u}_2 - 2\lambda_2 \bar{u}_2 - \beta \bar{u}_1 = 0 \end{aligned} \quad (14)$$

Premultiplicando ambos miembros por  $\bar{u}_1'$  y teniendo en cuenta las dos condiciones impuestas, se obtiene

$$\begin{aligned} 2\bar{u}_1' C \bar{u}_2 - 2\lambda_2 \bar{u}_1' \bar{u}_2 - \beta \bar{u}_1' \bar{u}_1 &= 0 \\ 2\bar{u}_1' C \bar{u}_2 - \beta &= 0 \end{aligned} \quad (15)$$



Como se ha visto antes,  $C\bar{u}_1 = \lambda_1\bar{u}_1$  (ecuación 11), multiplicando esta expresión por  $\bar{u}_2'$  y considerando la condición de ortogonalidad entre ambos vectores, obtenemos que  $\bar{u}_2' C\bar{u}_1 = 0$  (\*\*), de lo cual se deduce, sustituyendo en la ecuación anterior, que  $\beta=0$ .

Volviendo a la ecuación (14) y sustituyendo  $\beta$ :

$$C\bar{u}_2 - \lambda_2\bar{u}_2 = 0 \quad (16)$$

por lo que  $\bar{u}_2$  es el segundo vector propio de la matriz C , asociado al valor propio  $\lambda_2$  ( $\lambda_2 \leq \lambda_1$ ). El mismo planteamiento puede aplicarse para la obtención de los valores y vectores propios restantes.

Generalizando:

$$C\bar{u}_f = \lambda_f\bar{u}_f \quad (17)$$

Dado que todos los vectores propios son ortogonales, las varianzas son aditivas y

$$\frac{\lambda_f}{\lambda_1 + \dots + \lambda_p}$$

representa la proporción de la varianza total explicada por el componente f.

### 3.4 Interpretación geométrica

La transformación de componentes principales puede entenderse fácilmente considerando únicamente dos variables. Si representamos gráficamente los valores de los puntos de la imagen en el espacio bidimensional definido por las dos variables como ejes de abscisas y de ordenadas, obtendremos el diagrama de dispersión de la figura 1. En él, la nube de puntos presenta una forma típicamente alargada, poniendo en evidencia cierta correlación o dependencia entre ambas bandas. Si estas fueran totalmente independientes, su covarianza y su correlación serían nulas y la mencionada nube tendría una forma circular. Cuanto mayor sea la correlación entre dos bandas, más alargada será la nube de puntos y mayor será la covarianza entre las bandas (en valor absoluto).

Al realizar el cálculo de los componentes principales, los vectores propios de la matriz de covarianzas nos indican la dirección de los ejes principales o ejes de mínima inercia del conjunto de datos (PC1 y PC2 en la figura 1) y los valores propios son proporcionales a la variabilidad de los datos según los respectivos ejes. Por ello el primer componente, cuyo valor propio  $\lambda_1$  es máximo, explica la mayor variabilidad de la información contenida en la imagen, mientras que el segundo componente explicará el resto.

---

\*\* (ya que , al ser **C** una matriz simétrica, se cumple que  $\bar{u}_1' C\bar{u}_2 = \bar{u}_2' C\bar{u}_1$ )

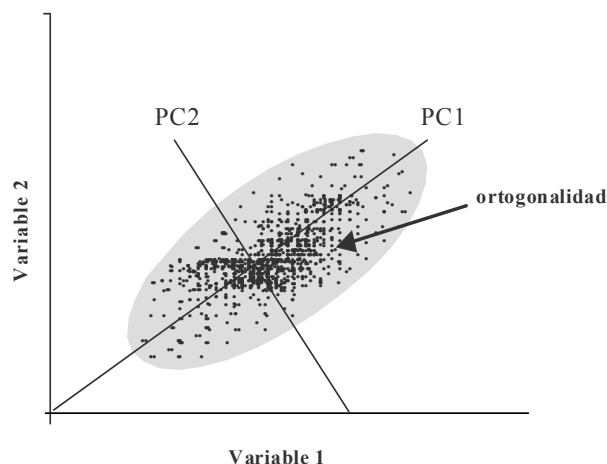


Figura 1.- Diagrama de dispersión de dos variables y direcciones de los componentes principales resultantes (PC1 y PC2).

Generalizando para un sistema  $n$ -dimensional en el que las distribuciones de los valores para las  $n$  variables sean normales, la nube de puntos vendrá definida por un hiperelipsoide con  $n$  ejes principales.

### 3.5 Ejemplo

Se ha aplicado el método de análisis de componentes principales a una imagen del sensor *Thematic Mapper* (TM) del satélite Landsat 5, con siete bandas espectrales, tres en el visible, tres en el infrarrojo próximo y una en el infrarrojo térmico. La imagen fue adquirida en junio de 1987 y representa una zona de cultivos de regadío de la rivera del río Júcar, en la provincia de Albacete.

La figura 2 muestra la imagen original, visualizada según una combinación en falso color infrarrojo, en la que destacan en tonos rojos las zonas de cultivos de regadío en pleno vigor vegetativo (elevada reflectancia en el infrarrojo próximo).

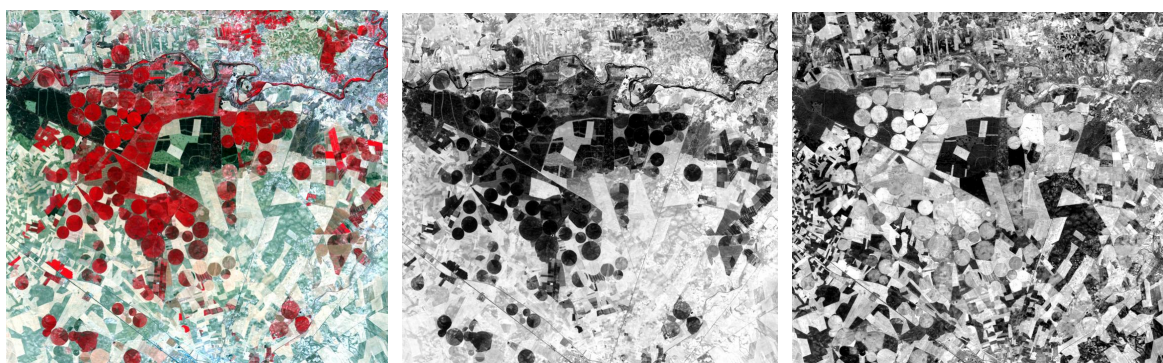


Figura 2.- Combinación en falso color infrarrojo (bandas 432-RGB) de la imagen Landsat-TM de los regadíos del río Júcar (izquierda). Imágenes de los componentes principales 1º y 2º (centro y derecha) de la imagen de las 7 bandas espectrales original.

En la misma figura se muestran los dos primeros componentes resultantes. Paralelamente, en la tabla 2 aparecen los vectores propios obtenidos, que representan los coeficientes de ponderación que han de aplicarse sobre los valores digitales de las bandas originales (filas) para obtener cada uno de los componentes resultantes (columnas) como combinación lineal ponderada de aquéllas.



La interpretación de los nuevos componentes resulta complicada. En este caso, a la vista de los coeficientes mencionados (tabla 2), observamos cómo para la obtención del primer componente se ponderan de forma más o menos similar las bandas originales, por lo que la imagen resultante contendrá principalmente información sobre la intensidad o brillo global de la imagen. Análogamente, en el segundo componente se aprecia claramente una ponderación positiva muy superior para la banda 4 (infrarrojo próximo) que para el resto por lo que, teniendo en cuenta que esta banda está directamente relacionada con un incremento acusado de la reflectancia de la hoja vegetal debido a su estructura morfológica interna, parece lógico asociar el segundo componente al vigor vegetativo y a la presencia de vegetación, lo cual se corrobora examinando la imagen de este componente en la figura 2, en el que las zonas de regadío aparecen con más brillo que el resto.

BANDAS ORIGINALES	COMPONENTES PRINCIPALES						
	CP1	CP2	CP3	CP4	CP5	CP6	CP7
B1	0.32376982	0.16218615	-0.50991653	-0.61961536	-0.01300927	0.35035711	0.31939215
B2	0.23819462	0.13187722	-0.2496522	-0.13367231	-0.02295034	-0.04469158	-0.91823372
B3	0.49060812	0.18133423	-0.43239192	0.49565145	-0.01212206	-0.49385048	0.22305427
B4	0.14463368	0.77924861	0.45482458	-0.06486198	-0.39838987	-0.00325463	0.04533381
B5	0.61794024	-0.07932526	0.38431493	0.21752758	0.46653323	0.44582854	-0.02061078
B6	0.14181706	-0.31521627	-0.10343572	0.33158942	-0.74449061	0.45230894	-0.03203895
B7	0.41816143	-0.45884613	0.35416638	-0.43710559	-0.26176915	-0.47750101	0.03969677

Tabla 1.- Vectores propios de la matriz de covarianzas de la imagen original.

La tabla 2 muestra los valores propios asociados a cada nuevo componente. Dividiendo el valor propio de cada componente entre la suma de todos ellos se obtiene la proporción de varianza total contenida en los datos originales expresada por el componente en cuestión, lo cual proporciona un criterio de selección del número de componentes para utilizar en cualquier procedimiento de análisis posterior, como es el caso de una clasificación espectral.

COMPONENTE	VALORES PROPIOS	% VARIANZA	% VARIANZA ACUMULADA
1	6917.2	90.44	90.44
2	490.1	6.41	96.85
3	127.7	1.67	98.52
4	53.2	0.69	99.21
5	34.3	0.45	99.66
6	24.4	0.32	99.98
7	1.6	0.02	100

Tabla 2.- Valores propios asociados y porcentajes de varianza y varianza acumulada explicada por cada uno de ellos.





## 3.6 Aplicaciones

A continuación se describen las principales aplicaciones del método en teledetección.

### 3.6.1 Reducción del volumen de datos

La principal utilidad del método consiste, como ya se ha visto, en la reducción de la dimensionalidad de los datos, eliminando así la información redundante debida a la correlación entre bandas. Una vez seleccionados los componentes que conlleven la mayoría de la información (suma de varianzas relativas próxima a la unidad) pueden realizarse las mismas operaciones de análisis de los datos (realce, clasificación,...) que utilizando la totalidad de las bandas, consiguiendo resultados similares. Esta aplicación tiene un enorme potencial cuando se trabaja con un número elevado de bandas espectrales tomadas a intervalos de longitud de onda muy cortos, lo que ocasiona una gran correlación entre bandas, haciéndose necesario un proceso estadístico de selección.

### 3.6.2 Reducción del ruido



*Figura 3.- Imagen correspondiente al 7º y último componente principal del ejemplo anterior. Puede apreciarse un gran componente de ruido en la imagen.*

Los valores digitales de las imágenes obtenidas mediante sensores a bordo de satélites se ven especialmente alterados debido tanto a la influencia de los fenómenos de dispersión atmosférica como a deficiencias, a veces muy sutiles, en el funcionamiento de los propios sensores. Al conjunto de alteraciones de la señal que se manifiesta en las imágenes se le denomina ruido.

Los últimos componentes obtenidos en el análisis de componentes principales, aquéllos que contienen la mínima información de los datos, suelen ir asociados a ruido del sistema o a las alteraciones de la imagen debidas a la dispersión atmosférica, por lo que su eliminación contribuye directamente a la reducción del ruido, esto es, a la restauración y realce de las imágenes.

### 3.6.3 Detección de cambios

La utilización de este método para la detección de cambios en el tiempo requiere la integración de dos o más imágenes multispectrales, de una misma zona y adquiridas en fechas distintas, en una misma matriz o archivo de datos sobre el cual se aplica el análisis de componentes principales. Suponiendo que las zonas modificadas sean cuantitativamente menos importantes, esto es, ocupen menos superficie que las zonas que no han evolucionado, los primeros componentes suelen retener la información relativa a las áreas comunes en todas las áreas (las invariantes), mientras que los componentes secundarios registrarán los cambios o evoluciones en el tiempo (características u objetos que sólo aparecen en una de las fechas).



Este tipo de aplicaciones requiere una **metodología** que consiste en las siguientes fases:

1. Corrección o **ajuste radiométrico** de las imágenes para reducir, en la medida de lo posible, las diferencias radiométricas debidas a las condiciones atmosféricas de cada fecha y a las diferencias de calibración o sensibilidad de los sensores.
2. **Registro geométrico** de las imágenes y **remuestreo** a la misma resolución espacial. Las diferencias geométricas provocarán errores en la localización de los cambios, e incluso harán que se detecte la presencia de cambios donde no existan.
3. Recorte o redimensionamiento de la zona de estudio en cada imagen e **integración** de todas ellas en un solo conjunto de datos o archivo.
4. Aplicación del método de **análisis de componentes principales** sobre la totalidad de las bandas multispectrales y multitemporales.
5. **Interpretación** de los componentes secundarios e **identificación** de las zonas de cambio que hayan sido realizadas.

El resultado puede limitarse al realce de las áreas que hayan sufrido algún cambio, como método de apoyo a la interpretación, o bien pueden aplicarse técnicas posteriores de análisis que permitan la detección automática de las mismas.

### 3.6.4 Ampliación del contraste por decorrelación

Otra aplicación de este método al realce de imágenes se basa en la ampliación del contraste por *decorrelación*, consistente en la conversión de los datos al espacio definido por los componentes principales de las bandas originales (este espacio se caracteriza por la inexistencia de correlación entre los componentes), seguido de la ecualización de los datos según los nuevos ejes, y por último la conversión de los datos al espacio inicial y la combinación de las bandas con los colores primarios RGB. De esta forma, los puntos se distribuyen más uniformemente en el espacio RGB, por lo que la imagen mostrará un contraste muy superior.

Cuando las bandas espectrales de una imagen están altamente correlacionadas entre sí, al visualizarla en el espacio de color RGB (utilizado por los monitores estándar) con frecuencia presenta un elevado componente de gris, lo que supone una carencia de color global en la imagen. Geométricamente, esto se debe a que los valores de los píxeles correlacionados se sitúan, lógicamente, en la zona próxima a la diagonal  $R=G=B$  del cubo de color RGB formando, en este espacio de representación, una concentración de puntos similar a un elipsoide. Las técnicas estándar de realce, tales como la ecualización del histograma o las basadas en transformaciones lineales de las bandas individuales, tienden a alargar dicho elipsoide o bien a orientarlo en otra dirección, pero no consiguen expandirlo de forma que ocupe el espacio de color disponible.

Entre los diversos métodos propuestos para superar esta limitación, uno de los más eficientes se basa en la eliminación de la correlación de los datos en el espacio de coordenadas original mediante el cálculo de los componentes principales. El método consta de tres fases:

1. Cálculo de los componentes principales de la distribución de valores en el espacio 3D original y transformación de los datos a este nuevo espacio.
2. Aplicación de técnicas de ampliación del contraste sobre cada uno de los 3 nuevos ejes, no correlacionados entre sí. Esto se realiza mediante la ecualización o escalado de los nuevos histogramas, consiguiéndose una expansión de los datos en el espacio tridimensional.
3. Transformación de los datos al espacio de coordenadas original (RGB).

Estas operaciones pueden realizarse por medio de una sola transformación lineal para cada banda. Expresándola en forma matricial:

$$X_D = X \cdot U \cdot S \cdot U^{-1}$$

donde  $X$  es la matriz de datos inicial  $n \times 3$  (tres componentes, RGB, para cada uno de los  $n$  píxeles en la imagen);  $U$  es la matriz de la transformación de los componentes principales, cuyas columnas son los vectores propios de la matriz de covarianzas de  $X$ ;  $S$  es una matriz diagonal de escalado, que produce un cambio de escala de los valores (ecualización) en el espacio de los componentes principales; y  $X_D$  es la matriz de datos resultante, también de dimensiones  $n \times 3$ . Cada columna de  $X$  y  $X_D$  representa una banda completa de la imagen.

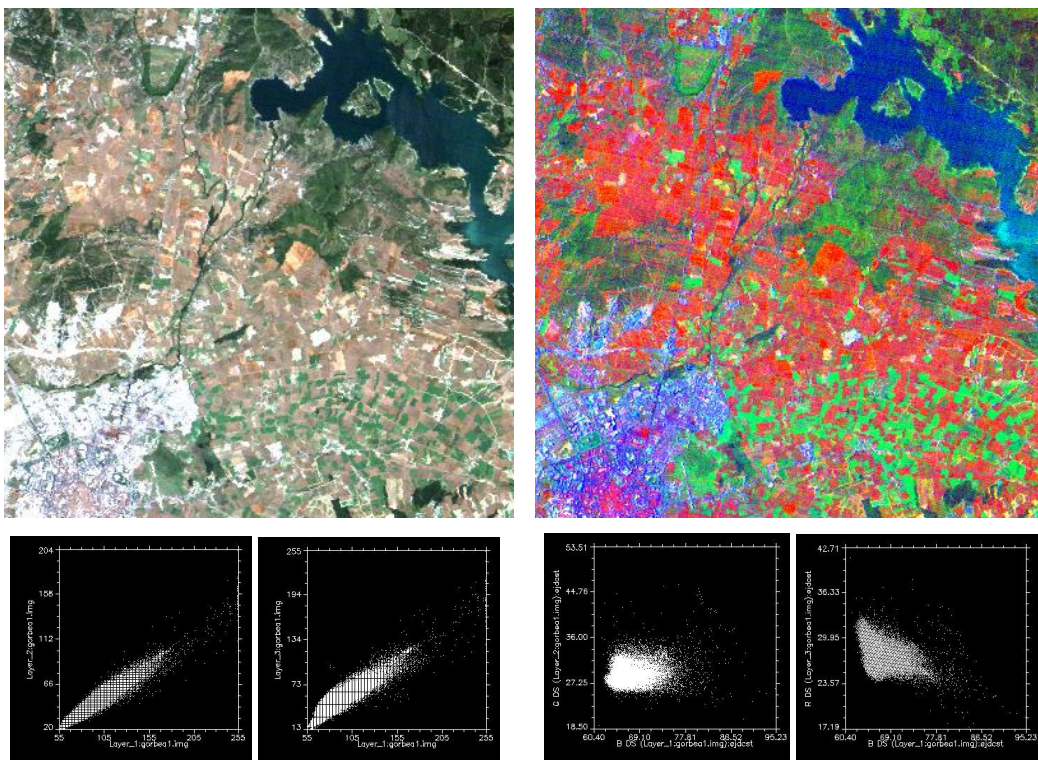


Figura 4.- Ejemplo de aplicación del realce espectral por decorrelación. La imagen de la izquierda representa una combinación en color natural con elevado contraste pero con una gran correlación entre bandas, como puede apreciarse en sus dos diagramas de dispersión (relacionan las bandas rojo/verde y rojo/azul). El resultado de la decorrelación se muestra en la imagen de la derecha,



Como se aprecia en el ejemplo de la figura 4, la imagen resultante adquiere unos colores un tanto exagerados, poco reales, sin embargo, éstos no se reasignan al azar, sino que mantienen su tonalidad inicial (la vegetación sigue siendo verde, el agua azul, etc.), lo cual permite la asociación de cada uno de ellos a las mismas características físicas de la escena inicial. Este método constituye una interesante herramienta de apoyo a la interpretación, especialmente en ciertos tipos de aplicaciones en las que se pretende localizar e identificar objetos poco evidentes (materiales geológicos, etc.).

### 3.6.5 Fusión de imágenes

Consultar el artículo "Métodos básicos de fusión de imágenes"  
(<http://hdl.handle.net/10251/30096>)

## 4 Conclusiones

Se ha descrito el método de análisis de componentes principales aplicado a imágenes multiespectrales, su interpretación y las aplicaciones más comunes en teledetección. Para una descripción más detallada se recomienda consultar la bibliografía que se indica en la sección siguiente.

## 5 Bibliografía

- Multivariate Analysis: Methods and Applications. William R. Dillon - Matthew Goldstein. John Wiley & Sons, 1984.
- Remote Sensing. Models and Methods for Image Processing. Robert A. Schowengerdt. Academic Press, 1997.
- Remote Sensing Digital Image Analysis. An Introduction. John A. Richards. Springer-Verlag, 1994.
- Teledetección Ambiental. Emilio Chuvieco. Ed. Ariel Ciencia, 2002.