

Sobre el uso de la descomposición en valores singulares para mejorar la compresión de imágenes mediante transformadas wavelet

Francisco Pedroche Sánchez¹

Resumen

En esta comunicación se presenta un método original para la mejora del *ratio* de compresión de imágenes mediante el uso de la descomposición en valores singulares (DVS). El método propuesto es susceptible de ser usado como una técnica pre y post-compresión que puede combinarse con cualquier método de compresión basado en matrices. Como ejemplo de aplicación mostramos como combinarlo con una compresión basada en transformadas wavelet.

Introducción

Los primeros estudios sobre la utilización de la DVS de una matriz para el tratamiento de imágenes se remontan a la década de los 70 (v. [1]). Últimamente, debido a las mejoras en los métodos numéricos para el cálculo de los valores y vectores singulares, se vuelve a reconsiderar su capacidad para el tratamiento de señales unidimensionales e imágenes (v. [2], [6], [8], [9]) incluyendo aspectos novedosos como el ocultamiento de unas imágenes en otras y las marcas de agua (*watermarking*) [5].

La estructura de la comunicación es la siguiente. Primero, a partir de los resultados conocidos sobre la DVS se muestra como una matriz A de tamaño $m \times n$ y de rango $r \leq \min\{m, n\}$ puede ser descrita completamente mediante la pareja formada por sus r coordenadas respecto a una base formada a partir de $2r$ vectores singulares (r a derecha y r a izquierda) de A , y las $(m+n)r$ coordenadas de estos vectores singulares.

A continuación se detalla la idea original del artículo, consistente en un algoritmo que permite la creación de una matriz C_k de tamaño $(m+n)k$, con $k < r$, prefijado *a priori* por el usuario, según la pérdida de información que admita en la pre-compresión. Este control de la pérdida de información se relaciona con magnitudes usadas en el campo del tratamiento de imágenes (v. [7], [10]). Hay que resaltar que la matriz C_k no es una descomposición parcial de A , aunque tiene los datos necesarios para una tal reconstrucción.

Esta matriz C_k , que es de menor tamaño que A bajo ciertas condiciones, puede entonces comprimirse por cualquier procedimiento, por ejemplo mediante una transformada wavelet. En la compresión de imágenes mediante wavelets, los coeficientes que se transmiten son las coordenadas de la matriz respecto a una base formada por

matrices que dependen del tipo de base (o *función madre* wavelet, desde un punto de vista continuo) que se utilice para la compresión (Haar, Coiflet, Daub, etc) (v. [3]).

Si $\mathcal{W}(C_k)$ es la transformada wavelet de C_k , para la reconstrucción de C_k se han de transmitir ciertos coeficientes de $\mathcal{W}(C_k)$ así como información sobre el tipo de compresión realizada (nombre de la transformada, mapa de significancia, etc). En nuestro caso, junto a esta información, se han de transmitir también los k valores singulares σ_i de A .

Sea \widehat{C}_k la matriz reconstruida de C_k a partir de la transformada wavelet realizada en el paso anterior. Para la post-compresión final, a partir de \widehat{C}_k y de los k coeficientes σ_i se puede reconstruir una matriz \widehat{A}_k , de tamaño $m \times n$ y de rango $\widehat{r} \leq k$. Entonces, la matriz \widehat{A}_k es la versión reconstruida de la matriz inicial.

Como ejemplo de aplicación de la técnica propuesta se usa la imagen test llamada *Lena*, almacenada en una matriz de 512×512 coeficientes.

Descomposición en valores singulares

Recordemos unos resultados básicos [4].

Teorema 1. Dada una matriz real A , de tamaño $m \times n$, existe una factorización real:

$$A = USV^T \quad (1)$$

en donde U es una matriz ortogonal real de tamaño $m \times m$, V es una matriz ortogonal real de tamaño $n \times n$ y la matriz S es una matriz $m \times n$, real, pseudo-diagonal con elementos no negativos en la diagonal, que denotaremos por σ_i , $i = 1, 2, \dots, \min\{m, n\}$, que se llaman valores singulares de la matriz A .

Los σ_i se suelen ordenar de manera decreciente: $\sigma_1 \geq \sigma_2 \geq \dots$. El conjunto de valores singulares $\{\sigma_i\}$ se denomina espectro singular de la matriz A . Las columnas \mathbf{u}_i de U (\mathbf{v}_i de V) se llaman vectores singulares a izquierda (a derecha) de A . Para cada i , la terna $(\mathbf{u}_i, \sigma_i, \mathbf{v}_i)$ se llama i -ésimo triplete singular de A .

Lema 1. El número de valores singulares no nulos coincide con el rango de la matriz A .

Lema 2. Mediante la DVS, cualquier matriz real A puede escribirse como la suma de $r = \text{rango}(A)$ matrices de rango uno, en la forma:

$$A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T. \quad (2)$$

en donde $(\mathbf{u}_i, \sigma_i, \mathbf{v}_i)$ es el i -ésimo triplete singular de A . La representación (2) se suele denominar como *expansión DVS*, *expansión de producto exterior* o *descomposición diádica* de A . Cada una las matrices $\sigma_i \mathbf{u}_i \mathbf{v}_i^T$ se llama *capa significativa*. Denotemos [4]:

$$A_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad k < r. \quad (3)$$

que llamaremos matriz de descomposición A_k o matriz *DVS truncada* de parámetro k . Esta matriz puede usarse como una aproximación a la matriz A en el campo de imágenes

digitales. Para evaluar el error cometido cuando se utiliza A_k como aproximación a A usaremos el concepto siguiente [10]

Definición 1. Dadas las matrices $A = (a_{ij})$ y $B = (b_{ij})$ de tamaño $m \times n$, llamaremos *diferencia relativa* entre ellas al número real:

$$D(A, B) = \sqrt{\frac{\sum_{i,j=1}^{m,n} (a_{ij} - b_{ij})^2}{\sum_{i,j=1}^{m,n} a_{ij}^2}} \quad (4)$$

Como regla práctica, si $D(A, B) \leq 0.05$, entonces la imagen almacenada en B puede considerarse como una aproximación aceptable (desde el punto de vista de nuestra percepción) a la imagen almacenada en A . Es fácil demostrar [7] que si A_k se toma como aproximación a A , entonces:

$$D(A, A_k) = \sqrt{\frac{\sum_{i=k+1}^r \sigma_i^2}{\sum_{i=1}^r \sigma_i^2}} \quad (5)$$

Por tanto, $D(A, A_k)$ queda determinado por el espectro de A .

Construcción de C_k

Definición 2. Dada $A \in \mathbb{R}^{m \times n}$, de rango r , para $k < r$ definimos la matriz de tamaño $k(m+n)$:

$$C_k = \begin{pmatrix} \mathbf{u}_1^T & \mathbf{v}_1^T \\ \mathbf{u}_2^T & \mathbf{v}_2^T \\ \vdots & \vdots \\ \mathbf{u}_k^T & \mathbf{v}_k^T \end{pmatrix} \quad (6)$$

donde $(\mathbf{u}_1, \dots, \mathbf{u}_k)$ son las k primeras columnas de la matriz U dada por el teorema 1 y $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ son las k primeras columnas de la matriz V dadas por el mismo teorema.

Lema 3. Sean m y n números naturales tales que $1 \leq m < n$. Entonces:

$$(m+n) \frac{m}{2} < mn \quad (7)$$

Demostración: Al ser $1 \leq m < n$, existe un único número natural no nulo p tal que $m+p = n$, de donde $m+n = n-p+n = 2n-p < 2n$, con lo cual $\frac{1}{m+n} > \frac{1}{2n}$ y multiplicando ambos miembros por mn tenemos que

$$\frac{mn}{m+n} > \frac{mn}{2n} = \frac{m}{2} \quad (8)$$

de donde se deduce la desigualdad buscada.

Teorema 2. Sea A una matriz real de tamaño $m \times n$ con $3 \leq m \leq n$. Si $r \geq m/2$ entonces existe una matriz C_k dada por (6) que es de menor tamaño que A .

Demostración: Tomando

$$k = \begin{cases} \frac{m}{2} - 1 & \text{si } m \text{ es par} \\ \frac{m-1}{2} & \text{si } m \text{ es impar} \end{cases} \quad (9)$$

se tiene que:

$$k < \frac{m}{2} \quad (10)$$

y si $r \geq m/2$ entonces $k < r$ y según la definición 2, el tamaño de C_k es $k(m+n)$. Ya podemos demostrar el caso $m < n$, puesto que de (10) y el lema 3, se tiene:

$$(m+n)k < (m+n)\frac{m}{2} < mn \quad (11)$$

con lo cual queda demostrado el teorema para este caso. Veamos ahora el caso $m = n$. El tamaño de C_k es ahora $k(m+n) = k2n < \frac{n}{2}(2n) = n^2$ es decir, es menor que el tamaño de A , con lo cual el teorema queda demostrado.

El teorema 2 muestra la existencia de, al menos, una matriz C_k cuando $r \geq m/2$. Ahora bien, esta condición no es muy restrictiva y podremos construir más de una matriz C_k de menor tamaño que A . En concreto, se cumple:

Corolario. Si se cumplen las hipótesis del teorema 2 y si $k \geq 1$, dada por la ecuación (9), entonces existen $i = 1, \dots, k$ matrices C_i de menor tamaño que A .

Nótese que cuando la matriz de partida A sea de tamaño tal que $m > n$, entonces podemos aplicar el teorema 2 a la matriz A^T . Por tanto, dada cualquier imagen (almacenada en una matriz con al menos 3 filas o 3 columnas) siempre podremos construir una matriz A para ver si cumple las condiciones del teorema 2. La condición $r \geq m/2$ se suele cumplir para imágenes naturales. En la figura 1 mostramos algunas imágenes (*mandril, frog, bird, gats*) que la verifican.

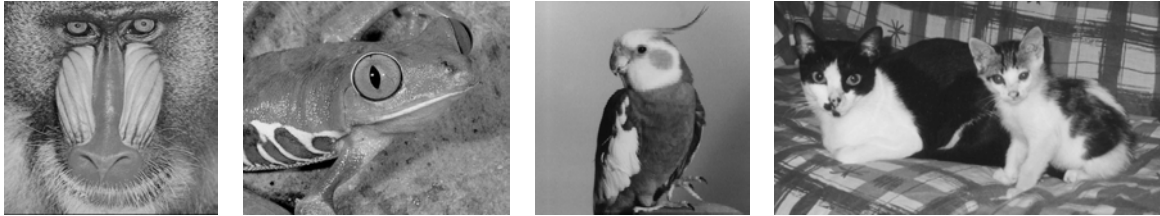


Fig 1. Ejemplo de imágenes que verifican el teorema 2.

Imagen	tamaño	rango
<i>mandril</i>	512×512	512
<i>frog</i>	498×621	498
<i>bird</i>	256×256	239
<i>gats</i>	359×657	359

Tabla 1. Tamaño y rango de las matrices de la figura 1.

Algoritmo

Sea A una matriz que verifica las hipótesis del teorema 2. Sea $\mathcal{W}(\mathcal{A})$ una matriz transformada de A , del mismo tamaño. Sea T_A una matriz poco densa, del mismo tamaño que A , cuyos coeficientes no nulos son un subconjunto de coeficientes de $\mathcal{W}(\mathcal{A})$ de manera que $\hat{A} = \mathcal{W}^{-1}(T_A)$ es una aproximación a A . Como medida de la compresión de la imagen A tomaremos el cociente entre los elementos de A y los elementos no nulos de T_A (que son los que efectivamente transmitimos o almacenamos). En estas condiciones, el algoritmo propuesto para mejorar la compresión ofrecida por \mathcal{W} es el siguiente:

1. Se calcula la DVS de la matriz A .
2. De acuerdo con la definición 2 y el teorema 2, existe una matriz C_k de tamaño menor que A . Para llegar a un compromiso entre calidad de la imagen y ratio de compresión usamos la ecuación (5) y el corolario del teorema 2 para elegir un valor definitivo de $k < r$.
3. Se aplica la transformada \mathcal{W} a la matriz C_k , obteniendo la matriz $\mathcal{W}(C_k)$. Sea T_{C_k} la matriz poco densa formada por ciertos coeficientes de $\mathcal{W}(C_k)$. Los coeficientes no nulos de T_{C_k} son los que se transmiten junto con los k valores singulares σ_i .
4. Para la recuperación de la matriz inicial, se calcula la transformada inversa $\mathcal{W}^{-1}(T_{C_k}) = \widehat{C}_k$ y, asumiendo que \widehat{C}_k conserva la estructura de la definición 2, se utiliza la descomposición (3) para construir una matriz \widehat{A}_K que es una aproximación a A .

En los ejemplos realizados se ha utilizado la misma \mathcal{W} para A y para C_k . Asimismo también se ha usado el mismo criterio de construcción de las matrices T_A y T_{C_K} . De esta manera, como la matriz C_K es de menor tamaño que A , la ganancia en compresión está asegurada. Teóricamente, cuando la transformada $\mathcal{W}(C_k)$ conserve las propiedades de C_k , y el valor de k sea tal que A_k conserve suficiente energía de A , el algoritmo propuesto ofrecerá una buena aproximación a la imagen inicial.

Transformada wavelet 2D

La transformada wavelet 2D se puede definir a partir de las funciones (bases) wavelet y scaling unidimensionales [10]. Para fijar ideas comencemos con la transformada wavelet 2D de nivel 1. Una transformada de este tipo aplicada sobre una matriz A de tamaño $m \times n$ puede simbolizarse de la manera siguiente:

$$A \longrightarrow \mathcal{W}(A) = \left(\begin{array}{c|c} h^1 & d^1 \\ \hline a^1 & v^1 \end{array} \right) \quad (12)$$

en donde las matrices h^1 , d^1 , a^1 , y v^1 son de tamaño $m/2 \times n/2$ y se denominan subimágenes de fluctuaciones horizontales, fluctuaciones diagonales, tendencias y fluctuaciones verticales, respectivamente. La matriz a^1 resulta una imagen promediada, a menor resolución y de menor tamaño, de la imagen inicial. Llamando α_i , $i = 1, \dots, m \times n$ a los elementos de la matriz $\mathcal{W}(A)$, resulta que podemos escribir:

$$A = \sum_{i=1}^{m \times n} \alpha_i W_i = H^1 + D^1 + A^1 + V^1 \quad (13)$$

en donde las matrices W_i , que se obtienen a partir de las funciones wavelet y scaling 1D, forman una base del espacio vectorial $\mathbb{R}^{m \times n}$. La matriz A^1 es una aproximación a la matriz A . Con la notación expuesta en el algoritmo, se tiene:

$$A^1 = \mathcal{W}^{-1} \left(\begin{array}{c|c} O & O \\ \hline a^1 & O \end{array} \right) \equiv \mathcal{W}^{-1}(T_A) \equiv \widehat{A} \quad (14)$$

Si llevamos a cabo la anterior transformada wavelet, el número de coeficientes a transmitir coincide con el tamaño de a^1 , es decir $\frac{mn}{4}$. Análogamente, si la matriz T_A la

construimos a partir de los coeficientes de a^1, v^1 y h^1 , entonces el número de coeficientes para reconstruir A usando A^1, V^1 y H^1 es $\frac{3mn}{4}$.

Si ahora aplicamos \mathcal{W} a la matriz a^1 estaremos construyendo una transformada wavelet de nivel 2, donde el número de coeficientes a transmitir se calcula de manera análoga.

Resultados

Hemos usado una matriz $A = (a_{ij})$ que almacena la imagen test llamada *Lena*. Es una matriz 512×512 , cuyos elementos son números enteros que representan niveles de grises, es decir, $a_{ij} \in [0, 255]$. A tiene rango $r = 507$ y, por tanto, cumple las condiciones del teorema 2.

Hemos usado $k = 80$ ya que la descomposición parcial A_{80} cumple que $D(A, A_{80}) = 0.04$, con lo cual tenemos que A_{80} es una buena aproximación, en terminos de percepción de la imagen, a A (v. [7] para detalles sobre esta elección).

En la tabla 2 se resumen los casos estudiados. Se ha utilizado una transformada wavelet Daub2, según la nomenclatura de MATLAB. Los tres primeros casos son de nivel 1 y los dos restantes son de nivel 2.

En la segunda columna de la tabla 2 indicamos las submatrices usadas para la construcción de T_A y de T_{C_k} . La mejora del ratio de compresión viene dada por el cociente entre el tamaño de estas dos matrices, que coincide con el cociente entre los tamaños de A y de C_k , que es en este caso: $mn/(k(m+n)) = 512 * 512 / (80 * 2 * 512) = 512/160 \approx 3$.

Por tanto, en los cinco casos expuestos conseguimos una mejora de la compresión del orden de 3 veces de la que se consigue usando sólo $\mathcal{W}(\mathcal{A})$. También habría que incluir que hay que enviar los k valores singulares, pero esto no altera significativamente este resultado.

Con objeto de evaluar si las imágenes reconstruidas son una buena aproximación a la imagen inicial, en la tabla 2 se muestran los errores relativos, calculados según la ecuación (4), siguientes: $D(A, \hat{A})$ que da una idea de la diferencia entre la imagen inicial y la transmitida por la transformada wavelet sobre A ; $D(A, \hat{C}_k)$ que compara la imagen inicial y la transmitida por el algoritmo propuesto, y $D(\hat{A}, \hat{C}_k)$ que compara la imagen transmitida por la transformada wavelet sobre A con la transmitida por el algoritmo propuesto.

Caso	Matrices a transmitir	Número de coeficientes no nulos de T_A	$D(A, \hat{A})$	$D(A, \hat{C}_k)$	$D(\hat{A}, \hat{C}_k)$
1	a^1, v^1 y h^1	$3mn/4$	0.013	0.048	0.047
2	a^1 y h^1	$mn/2$	0.035	0.052	0.042
3	a^1	$mn/4$	0.043	0.82	0.82
4	a^2, v^1 y h^1	$9mn/16$	0.070	0.25	0.25
5	a^2, h^2, v^1 y h^1	$5mn/8$	0.058	0.069	0.059

Tabla 2. Resumen de resultados para $k = 80$.



Fig 2. Imágenes reconstruidas a partir de A (izquierda) y de C_{80} para el caso 5.



Fig 3. Imágenes reconstruidas a partir de A (izquierda) y de C_{80} para el caso 4.

En los casos 1, 2 y 5 (v. fig. 2) se observa que el algoritmo propuesto mejora la compresión mientras que no perjudica sustancialmente la calidad de la imagen transmitida. Los casos 3 y 4 (v. fig. 3) ponen de relieve que las propiedades de los coeficientes de $\mathcal{W}(A)$ respecto a las características de la imagen, no son las mismas que las de los coeficientes de $\mathcal{W}(C_k)$. Es decir, una buena estrategia de compresión para A puede no serlo para C_k . La comparación de los casos 4 y 5 abunda en este hecho; La inclusión o no de los detalles horizontales de segundo nivel influye notablemente sobre $\mathcal{W}(C_k)$, mientras que no lo hace sobre $\mathcal{W}(A)$. Esto deja abierto el problema de analizar que estrategias de compresión son las adecuadas para la transformada $\mathcal{W}(C_k)$. Respecto a la medida del error introducido en la compresión se observa que los valores de $D(A, \hat{C}_k)$ son del mismo orden que los de $D(\hat{A}, \hat{C}_k)$, pudiendo, por tanto, usarse indistintamente una u otra magnitud para evaluar el error.

Conclusiones y desarrollos futuros

Se ha presentado un algoritmo de compresión *DVS* combinado con una compresión wavelet, a partir de la definición de una matriz C_k que no es una matriz correspondiente a una imagen ni tampoco es una descomposición parcial. Se han dado condiciones suficientes para que la matriz C_k sea de menor tamaño que la matriz inicial. Los primeros resultados, usando una imagen test, muestran que el algoritmo propuesto consigue un aumento de la compresión de 3 a 1, manteniendo, en ciertos casos, la calidad de la imagen en niveles aceptables. Se ha mostrado también que puede ocurrir que una transformada wavelet apta para A no lo sea para C_k . Como desarrollos futuros se

pretende completar el presente estudio aplicando el algoritmo a otros valores de k , y utilizando transformadas wavelet de más alto nivel. También se hará un tratamiento de más imágenes y nos centraremos entonces en estudiar la posibilidad de aplicar reiteradamente el teorema 2. También se ha de estudiar el tipo de información visual que es susceptible de transmitir C_K en función del tipo de coeficientes de la transformada wavelet que se consideran.

Agradecimientos

Esta publicación está financiada por el *Programa d'Incentiu a la Investigació de la Universitat Politècnica de València*. El trabajo de investigación está financiado parcialmente por el proyecto de investigación DGI número BFM2001-2641.

Referencias

- [1] Andrews, H. C., y Patterson, C. L. (1975) "Outer product expansions and their uses in digital image processing". *American mathematical monthly* **82**, pp. 1-13.
- [2] Callaerts, D., De Moor, B., Vandewalle, J., Sansen, W., Vantrappen, G. and Janssens, J. (1990). "Comparison of the SVD Methods to extract the foetal electrocardiogram from cutaneous electrode signals". *Med. & Biol. Eng. & Comp.*, 28, pp. 217-224.
- [3] Daubechies, I. (1992) *Ten lectures on Wavelets*. SIAM. USA.
- [4] Golub, G.H. and Van Loan, C. F., (1996, 3rd ed.) *Matrix Computations*. The Johns Hopkins University Press. Baltimore, Maryland. USA.
- [5] Gorodetski, V.I. Popyack, L.J., Samoilov, V. and Skormin, V.A. (2001). "SVD-Based approach to Transparent Embedding Data into digital Images". *LNCS* 2052, pp. 263-274. Springer-Verlag.
- [6] Kanjilal, P.P., y Palit, S.(1995) "On multiple pattern extraction using singular value decomposition". *IEEE transactions on signal processing* vol.43, No. 6.
- [7] Pedroche, F. (2002) "On some capabilities of the SVD expansion to handle images". *WSEAS transactions on mathematics. vol 1. issue 2. pp. 67-70*.
- [8] Song, L. y Zhang, S. (1999) "Singular value decomposition-based reconstruction algorithm for seismic travelttime tomography". *IEEE transactions on image processing* vol. 8, No. 8.
- [9] Vidal, V., Ginestar, D., y Verdú, G. (2001) "Análisis de imágenes mediante descomposición parcial en valores singulares", en *XVII Congreso de Ecuaciones Diferenciales y Aplicaciones- VII Congreso de Matemática Aplicada*. L. Ferragut, y A. Santos (Ed). Departamento de Matemática Aplicada, Universidad de Salamanca.
- [10] Walker, J.S. (1999) *A primer on wavelets and their scientific applications*. Chapman & Hall/CRC. USA.

1 Departament de Matemàtica Aplicada. Universitat Politècnica de València. Camí de Vera s/n. 46022. València. pedroche@mat.upv.es