

# Compresión sin pérdida aplicada a imágenes de microscopía electrónica

V. Ruiz, J. J. Fernández, I. García

Departamento de Arquitectura de Computadores y Electrónica

Universidad de Almería

04120 Almería. SPAIN

e-mail: {vruiz@gogh, jose@maiden, inma@iron}.ualm.es

## Resumen

Presentamos una técnica basada en redes neuronales para la compresión sin pérdida de imágenes de microscopía electrónica de alta resolución. La red se utiliza como predictor adaptativo en una sola pasada. A continuación la imagen residuo se comprime utilizando un codificador de longitud variable. El resultado es que para el tipo de imagen en cuestión se obtienen mejores razones de compresión que utilizando los estándares de compresión de imágenes.

**Palabras clave:** compresión sin pérdidas de imágenes, predicción adaptativa, redes neuronales BPN.

## 1 Introducción

La compresión de imágenes está dominada por algoritmos de compresión con pérdidas de información, tales como JPEG (Joint Photographic Experts Group). Estos algoritmos de compresión están basados en la eliminación de la menos relevante para el sistema visual humano, junto con una codificación eficiente de la información que sí lo es. Gracias a esta pérdida controlada, la razón de compresión obtenida aumenta alrededor de un orden de magnitud respecto a los niveles alcanzados por métodos que recuperan la imagen original tras la descompresión y que son llamados sin pérdidas o totalmente reversibles.

El uso de compresores de imágenes sin pérdidas está justificado en aquellos casos en los que toda la información almacenada en la imagen es necesaria. Las imágenes tomadas desde satélite, las de origen médico o las obtenidas a partir de muestras microscópicas son ejemplos claros en los que la degradación de la imagen no es deseable. Generalizando, podemos decir que si el fin de la imagen no es ser visualizada por un humano, entonces los métodos de compresión irreversibles no son interesantes.

En este documento se presenta un método reversible de compresión de imágenes procedentes del campo de la microscopía electrónica empleadas para abordar estudios estructurales de alta resolución en especímenes biológicos, basado en la compresión-texto (por ejemplo, mediante un codificador aritmético y un modelo probabilístico) de la imagen residuo (también llamada imagen error) resultante de la diferencia existente entre la imagen a comprimir y la imagen generada por un predictor no lineal que utiliza una red neuronal de retropropagación (BPN: Back Propagation Network). Los procesos de compresión y expansión se realizan en una sola pasada lo que permite leerlas y almacenarlas en dispositivos de almacenamiento secuencial (en general cintas) sin utilizar otros sistemas de almacenamiento intermedio (en general discos) consiguiéndose mejorar los niveles de compresión alcanzados por otros compresores de imágenes.

## 2 Características de las imágenes de microscopía electrónica del alta resolución

La microscopía electrónica, en conjunción con sofisticadas técnicas de procesamiento de imagen y de reconstrucción tridimensional, permite obtener información estructural cuantitativa sobre especímenes biológicos con el fin último de conocer en detalle su función biológica. La acción destructiva de la radiación electrónica sobre la muestra biológica constituye el mayor inconveniente de la microscopía electrónica en lo que se refiere a estudios estructurales de alta resolución. Con el objetivo de preservar estos detalles,

imagen	tamaño	entropía	nivel de compresión
3950	6000 × 6000	5.63	25335000 1.421:1
0697	6000 × 6000	6.18	27810000 1.294:1
9645	5000 × 5000	5.90	18437500 1.356:1

Tabla 1: Imágenes de microscopía que han servido como “corpus” de prueba. Se presenta las dimensiones, la entropía y el nivel de compresión esperado (longitud esperada de la imagen comprimida (en bytes) y un factor de compresión).

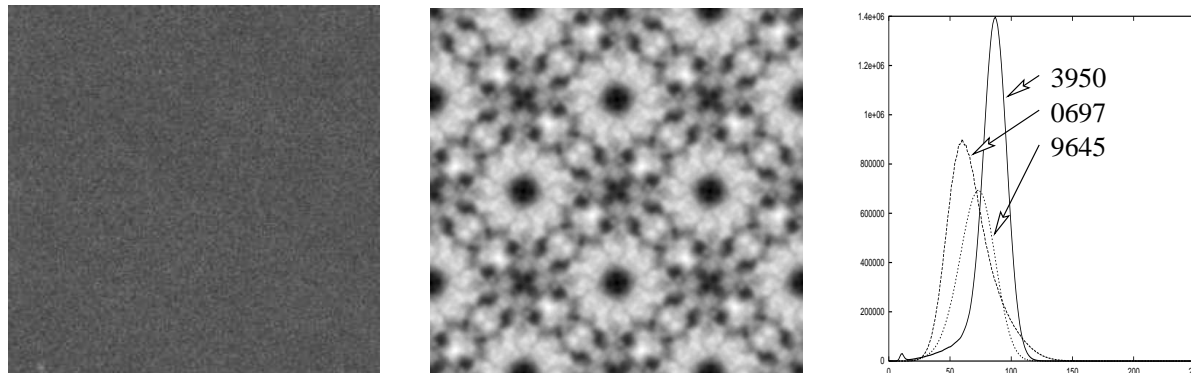


Figura 1: Ampliación de  $256 \times 256$  puntos de la zona central de una de las imágenes (izquierda), el contenido (ampliado) típico de una imagen procesada (centro) y los histogramas de las tres imágenes (derecha).

se emplean bajas dosis de radiación lo que ocasiona que las imágenes obtenidas sean extremadamente ruidosas y la relación señal-ruido sea enormemente pobre. La forma de alcanzar una relación señal-ruido estadísticamente significativa consiste en disponer de una cantidad suficiente de redundancia estructural. Esto implica la toma de imágenes de amplios campos de microscopía que contengan un gran número de muestras del espécimen biológico.

El estudio estructural a alta resolución de especímenes biológicos, por tanto, parte de un conjunto de imágenes de gran tamaño, enormemente ruidosas, a partir de las cuáles se extrae la información estructural[4]. El tratamiento de ese conjunto de imágenes implica un enorme volumen de información, como consecuencia de lo cual surge la necesidad de la compresión. Sin embargo, el carácter extremadamente ruidoso de las imágenes (por tanto, con escasa correlación espacial) las convierte en difíciles de comprimir por la mayoría de los métodos de compresión actualmente existentes, ideados para codificar otro tipo de imagen más correlacionada.

La tabla 1 presenta las imágenes de prueba utilizadas mostrando el tamaño de las imágenes (todas utilizan 8 bits/punto), la entropía<sup>1</sup> (expresada en bits/punto) que alcanzan respecto de un modelo probabilístico de orden 0 y el nivel de compresión esperado<sup>2</sup> si aplicamos un compresor-texto que utiliza un modelo probabilístico de orden 0 y un codificador de longitud variable.

Por otra parte, la figura 1 muestra una imagen típica (sólo un detalle es presentado) usada en un análisis estructural de alta resolución, donde el carácter ruidoso es claramente evidente. Esta figura también muestra el mapa de proyección (y en consecuencia, el motivo medio de todas las muestras del espécimen en la imagen) resultante del procesamiento de una de las imágenes siguiendo la metodología descrita en [4]. Ese mapa de proyección muestra  $2 \times 2$  muestras del espécimen repitiéndose a lo largo de la imagen. Con el objetivo de comparar, una imagen típica contiene entre 5000 y 10000 muestras de el espécimen repetido por la imagen, dependiendo del tamaño de la imagen y de la frecuencia de muestreo. Finalmente, la figura 1 presenta los histogramas<sup>3</sup> de las imágenes de prueba que han sido utilizadas.

<sup>1</sup>Como una medida de la cantidad de información o dificultad de compresión asociada.

<sup>2</sup>Expresado como la longitud calculada multiplicando el tamaño de la imagen original por la entropía, dividiendo entre 8 ya que la longitud se expresa en bytes y como un factor  $X : 1$  donde  $X$  es el cociente entre la longitud de la imagen sin comprimir y la longitud de la imagen comprimida.

<sup>3</sup>El histograma nos da también una idea de la cantidad de compresión que puede sufrir la fuente cuando se codifica mediante un código de longitud variable. Como puede verse, se trata de ruido gaussiano.

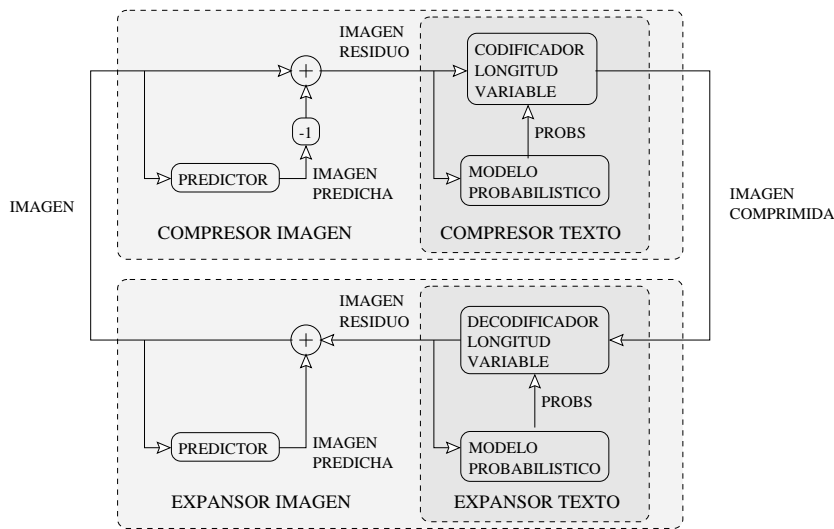


Figura 2: Modelo de compresión y expansión DPCM.

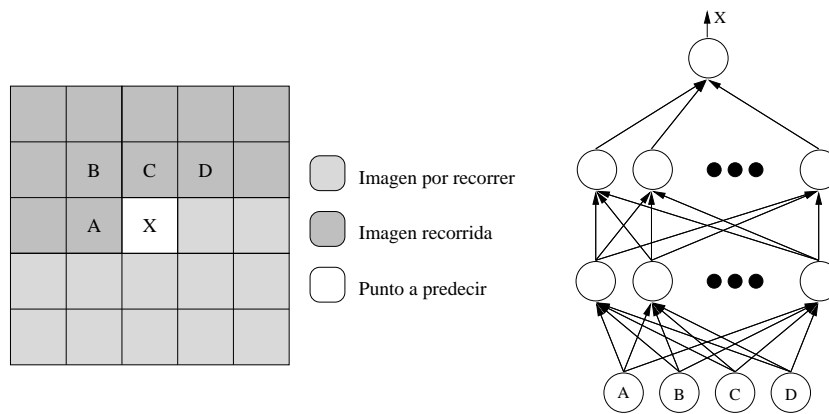


Figura 3: Puntos utilizados para realizar la predicción y estructura de la red.

### 3 La codificación predictiva de imágenes

Es posible comprimir (de forma reversible) las imágenes en cuestión utilizando un compresor-texto basado en un codificador de longitud variable ya que los histogramas no son planos (ver figura 1) aunque con niveles de compresión no superan el 50% (ver tabla 1). Una forma de mejorarlos consiste en utilizar un predictor especializado en los datos a comprimir que permita enviar al compresor-texto la mínima cantidad de información (entropía) posible. Si esto ocurre, el histograma de la imagen enviada es más estrecho y por tanto más comprimible. El proceso que recibe el nombre de modulación por codificación diferencial de pulsos (DPCM: Differential Pulse Code Modulation) puede verse descrito en la figura 2.

En el contexto de la compresión de imágenes, el predictor intenta predecir el valor del siguiente punto de la secuencia normal de recorrido en función de sus vecinos conocidos, aprovechándose de la correlación espacial existente[9]. La imagen predicha se resta punto a punto a la imagen original para formar una imagen error o imagen residuo que sólo debe almacenar la redundancia estadística no explotada por el predictor y que debe ser eliminada por el compresor-texto. Nótese que el tamaño de la imagen residuo es el mismo que el de la imagen original y por tanto, el predictor no comprime ni expande. Es el compresor-texto el que genera realmente una secuencia comprimida.

#### 3.1 El predictor

El predictor es una red neuronal como la estructura descrita en la figura 3, que también muestra los 4 puntos A,B,C y D de entrada a la red y el punto a predecir X<sup>4</sup>. El algoritmo de predicción de un punto

<sup>4</sup>Con el objetivo de explotar al máximo la correlación espacial, los puntos de entrada rodean al punto a predecir, teniendo en cuenta la parte conocida de la imagen expandida si se recorre por filas.

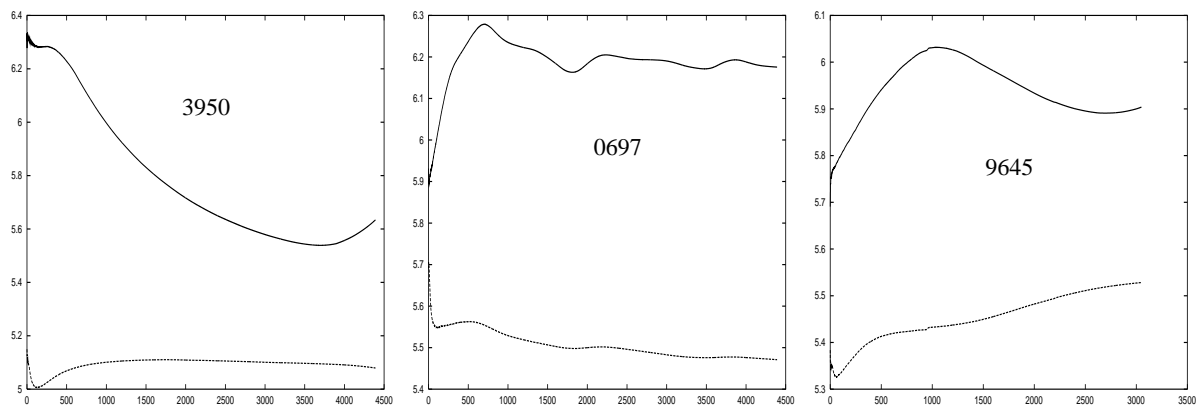


Figura 4: Entropías parciales de las imágenes originales (curvas superiores) y las imágenes residuo (curvas inferiores).

consiste en: (1) propagar la señal a partir de A,B,C y D hasta la neurona de salida, (2) calcular el error de predicción (imagen residuo) usando X y (3) actualizar los pesos con un factor de aprendizaje  $\alpha$  y un momento  $\eta$  utilizando el algoritmo de retropropagación[1]. Este proceso se realiza para todos los puntos de la imagen.

La predicción es adaptativa ya que la red se entrena simultáneamente, por lo que el método de compresión puede ser clasificado como ADPCM (Adaptive DPCM). Gracias a ésto, la compresión/expansión puede realizarse en una sola pasada y no es necesario transmitir los pesos de las conexiones de la red. Evidentemente, la predicción de los puntos de las primeras filas será de baja calidad (alta entropía), pero como muestra la figura 4, ésta situación cambia rápidamente (dependiendo notablemente de  $\alpha$  y  $\eta$ ).

Como consecuencia de utilizar una red neuronal con 2 capas ocultas, la función de predicción no es lineal[8]. Esto junto con la autoconfiguración dinámica de la red hace que el trabajo de predicción sea mejor que otros métodos no adaptativos y lineales, en aquellas zonas más difíciles de predecir: las fronteras, que son muy frecuentes debido al ruido de alta frecuencia.

El número de neuronas ocultas es un factor a determinar empíricamente. Como veremos, muchas neuronas disminuyen la entropía de la imagen residuo, pero aumentan el tiempo de predicción. Además, este factor depende de la imagen en procesada.

### 3.2 El compresor-texto

Eliminada (supuestamente) toda la correlación espacial gracias al predictor, el siguiente paso es eliminar la redundancia estadística, para así encontrar la representación más reducida posible de la imagen residuo que es ahora tratada como una secuencia unidimensional. Si el predictor fuera perfecto, bastaría con utilizar un modelo probabilístico de orden 0, pues no existiría correlación alguna entre los puntos de la imagen residuo, pero si no lo es, deberemos utilizar modelos de orden superior. Para contrastar estas hipótesis, se han probado diferentes codificadores (Huffman[6] y aritmético[3, 6]) con modelos de diferentes órdenes. Tanto el codificador de Huffman como el aritmético utilizan modelos adaptativos. Además, se han utilizado otros compresores basados en diccionarios (`compress`<sup>5</sup>, `gzip`<sup>6</sup>) y uno basado en la ordenación de bloques (`bzip2`<sup>7</sup>), con el objetivo de hacernos una idea del tipo de redundancia existente en la imagen residuo.

## 4 Otros compresores sin pérdidas de imágenes probados

Para comparar el rendimiento definitivo de compresor hemos probado a comprimir mediante el estándar de compresión de imágenes binarias JBIG[2] (Joint Binary Image Group, también llamada esta técnica ABIC: Adaptive Bilevel Image Compression). Utiliza un predictor bidimensional y codificación aritmética binaria. JBIG Puede ser utilizado para comprimir imágenes de tono continuo aplicándose por separado

<sup>5</sup> `compress` es la implementación más extendida del algoritmo LZW (Lempel-Ziv-Welch) en los entornos UNIX.

<sup>6</sup> Utiliza una variación del algoritmo LZ77 llamado `deflate`. A continuación las referencias al diccionario se comprimen mediante un código de Huffman.

<sup>7</sup> Implementa el algoritmo de ordenación de bloques de Michael Burrows y D. J. Wheeler para la eliminación de la correlación estadística y luego utiliza un código de Huffman.

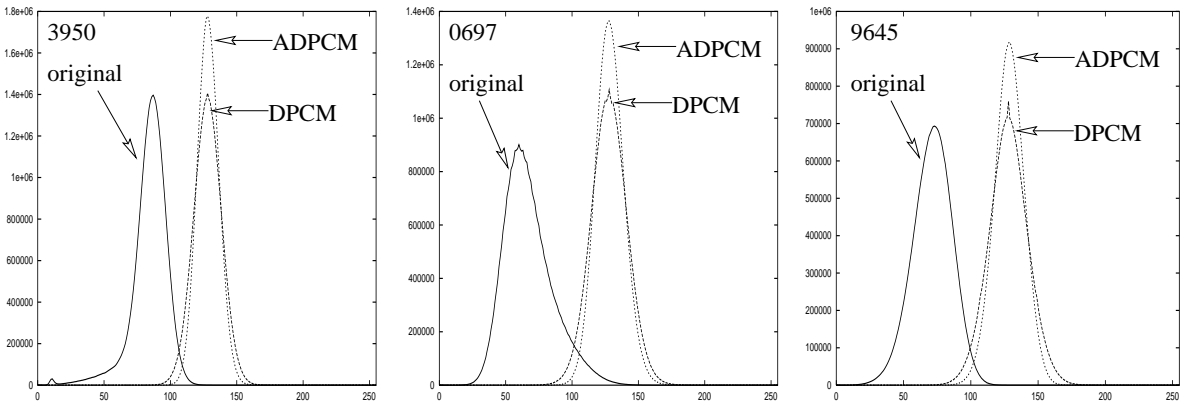


Figura 5: Histogramas de las imágenes originales y residuo, cuando utilizamos un codificador DPCM y una red 4-2-2-1 (ADPCM) con  $\alpha = 1$  y  $\eta = 0.1$ .

imagen	DPCM	ADPCM			
		4-2-2-1	4-4-4-1	4-6-8-1	4-8-16-1
3950	5.440995	5.079334	5.079067	5.078683	5.078555
0697	5.796362	5.470796	5.471671	5.466910	5.467077
9645	5.906264	5.528365	5.528328	5.528365	5.527312
3950	24484478	22857003	22855802	22854074	22853498
0697	26083629	24618582	24622520	24601095	24601847
9645	18457075	17276141	17276025	17276141	17272850
3950	23	548	1005	1847	3533
0697	23	547	993	1843	3530
9645	15	375	687	1277	2454

Tabla 2: Entropía (de orden 0), longitud (expresada en bytes) esperada de la imágenes comprimidas y tiempo de cálculo de las imágenes residuo para el codificador DPCM y ADPCM (en 4 configuraciones distintas siendo en todas,  $\alpha = 1$  y  $\eta = 0.1$ ).

a cada plano de bits. Normalmente, los planos de bits son generamos mediante el código Gray[7] ya que genera menos fronteras.

El segundo compresor de imágenes reversible probado es el estándar de compresión de imágenes LS-JPEG (Lossless-JPEG)[5] basado en el algoritmo LOCO-I. Este utiliza un predictor que busca las fronteras de la imagen en la región próxima al punto a predecir mediante la el cálculo del máximo gradiente. La imagen residuo se codifica mediante un código de Golomb-Rice y un modelo probabilístico de orden 5. Se caracteriza por ser un compresor/expansor muy rápido y eficiente.

## 5 Resultados

En primer lugar vamos a contrastar el rendimiento de los dos predictores probados: (1) un predictor DPCM clásico basado en predecir el valor del siguiente punto en función del punto vecino en el recorrido por filas (punto A en la figura 3) y (2) el predictor ADPCM basado en la red neuronal entrenada con el algoritmo de retropropagación del error. Como puede apreciarse en la figura 5 los histogramas de las imágenes originales son en general más anchos que los obtenidos mediante el método DPCM y estos a su vez, más que los obtenidos mediante ADPCM<sup>8</sup>. Este hecho se refleja en las entropías mostradas en la tabla 2. Como en ella se aprecia, la diferencia entre DPCM y ADPCM es notable, pero el aumento del tamaño de la red no mejora significativamente el nivel de compresión esperado y sin embargo, los tiempos<sup>9</sup> crecen proporcionalmente al número de neuronas que forman la red. También son mostradas las longitudes de las imágenes comprimidas mediante DPCM y ADPCM. Comparándolas con las expuestas

<sup>8</sup>La red se describe como 4-X-Y-1 donde X y Y son el número de neuronas de la primera y segunda capa oculta.

<sup>9</sup>Tomados en un Intel Pentium MMX a 200 MHz con 64 MBytes de RAM, corriendo Linux RedHat 5.2 (kernel 2.0.34). Los fuentes (excepto para LS-JPEG que no están disponibles) se compilaron con gcc versión egcs-2.90.29, optimizando en velocidad. Para medir tiempos se utilizó el comando `time` del sistema operativo y están expresados en segundos.

imagen	Huffman		Aritmética		LZ		Ordenación
	orden 0	orden 1	orden 0	orden 1	compress	gzip -9	bzip2 -9
3950	23071605	23045038	22865202	22862495	25918825	24325349	24250350
0697	24742751	24719929	24634652	24620810	28345993	25528744	26056682
9645	17355575	17333791	17272816	17265915	19991201	17863450	18263465
3950	76/89	122/123	266/378	288/436	27/11	59/13	228/84
0697	79/93	136/137	270/382	299/453	28/11	59/13	236/88
9645	56/66	97/97	187/264	209/318	19/8	40/9	164/61

Tabla 3: Tamaño de la imagen comprimida y tiempo de cómputo para los diferentes compresores/expansores texto. La opción -9 en gzip y bzip2 busca la máxima compresión. La imagen residuo comprimida es la generada por el predictor 4-2-2-1.

imagen	Compresores-Imagen			Compresores-texto		
	BPN-ARIT	JBIG	LS-JPEG	compress	gzip -9	bzip2 -9
3950	22862495	24648667	23791705	27897415	25663647	24646073
0697	24620810	26573588	25539314	31143733	28039236	26813798
9645	17265915	23927929	17953425	20639559	18248295	18428099
3950	836/984	238/230	43/48	32/10	58/13	218/82
0697	847/1001	256/249	44/50	33/12	59/13	235/88
9645	757/866	166/160	31/35	25/8	40/9	159/61

Tabla 4: Tamaño de la imagen comprimida a partir de la original y tiempo de compresión/expansión para los tres compresores de texto y de imágenes probados.

en la tabla 1 vemos que la introducción del predictor ADPCM mejora aproximadamente en un 10% las compresiones esperadas. Por otra parte, ADPCM es espera que sea casi un 7% mejor que DPCM<sup>10</sup>.

La tabla 3 compara el rendimiento de los diferentes compresores de texto probados para comprimir la imagen residuo generada con el predictor 4-2-2-1 y que se basan: (1) en la codificación de Huffman, (2) en la codificación aritmética, (3) en métodos LZ (Lempel-Ziv) y (4) un método de ordenación por bloques. De ella se desprende que: (1) que el método de compresión más eficiente es el codificador aritmético, seguido de el codificador de Huffman. Esta diferencia entre el código de Huffman y el aritmético es consecuencia de la redundancia generada por el primero cuando las probabilidades de los símbolos no son una potencia inversa de dos. Los otros métodos no explotan bien la redundancia de la imagen residuo debido ya que éstos buscan especialmente la correlación eliminada por el predictor. Además, podemos ver que el aumento del orden del modelo probabilístico utilizado con el codificador de Huffman y el aritmético no disminuye significativamente el tamaño de la imagen comprimida (la ganancia no supera nunca el 0.02%) lo que nos hace pensar que la cantidad de correlación existente en la imagen residuo es baja y por lo tanto el trabajo del predictor ha sido bueno.

Por último, en la tabla 4 podemos ver el rendimiento del compresor ADPCM propuesto (primera columna) frente a: (1) los compresores de texto y (2) los compresores de imágenes. La utilización de los compresores de texto directamente sobre la imagen original está justificada si tenemos en cuenta que gracias a la baja correlación espacial encontrada, su rendimiento es comparable a los compresores específicos de imágenes. Los datos presentados (1) son la longitud (en bytes) de la imagen comprimida y (2) los tiempos (en segundos) de compresión/expansión. El compresor-imagen escogido es el basado en un predictor 4-2-2-1 y en un compresor aritmético de orden 0 ya que expresa bien la situación de compromiso entre compresión y tiempo de cálculo. Para JBIG se da el tiempo y la compresión alcanzada cuando se comprimen sólo las 4 capas más significativas de la imagen original ya que las 4 restantes son expandidas y por tanto, transmitidas sin comprimir. Tanto para JBIG como para LS-JPEG los parámetros de compresión son los establecidos por defecto. Como conclusión podemos decir que aunque el método propuesto es sin duda el más lento de los probados (aproximadamente el triple de lento que JBIG y unas 19 veces más lento que LS-JPEG), también es el que más comprime (aproximadamente un 4% más que LS-JPEG y un 7.2% más que JBIG).

<sup>10</sup>Estos porcentajes se han calculado dividiendo la diferencia de las longitudes de las compresiones esperadas entre la longitud más larga, multiplicando a continuación por 100.

## 6 Conclusiones

El estudio estructural a alta resolución de especímenes biológicos requiere el empleo de un enorme volumen de información, procedente de un amplio conjunto de imágenes de gran tamaño, que justifica la necesidad de la compresión. El carácter extremadamente ruidoso de las imágenes empleadas exige el desarrollo de sofisticados algoritmos de compresión. En este trabajo se ha desarrollado una técnica ADPCM basada en una red neuronal entrenada con el algoritmo de retropropagación, que consigue superar los resultados de compresión proporcionados por estándares como JBIG y LS-JPEG.

El mayor inconveniente de la técnica presentada es el elevado tiempo de computación que se requiere frente al requerido por los estándares con los que se ha comparado. Este inconveniente puede hacer que el algoritmo presentado sea de utilidad cuando las restricciones de espacio en disco (en caso de almacenamiento de las imágenes) o de ancho de banda de la red (en caso de transmisión) sean realmente severas.

Alguna de las posibilidades que se están barajando para trabajo futuro es el empleo de algoritmos de compresión con pérdidas de información. El desarrollo de estos algoritmos está orientado a la eliminación de la información menos relevante desde un punto de vista estructural. Por tanto, estos algoritmos estarán basados en heurísticas distintas a las empleadas en los algoritmos estándares de compresión con pérdidas, basados en la eliminación de la información menos relevante para el sistema visual humano.

Por otra parte los tiempos pueden ser reducidos minimizando el tiempo de entrenamiento de la red, bien seleccionando los puntos donde mayor error de predicción se comete, bien escogiendo sólo algunos puntos equidistantes.

## 7 Reconocimientos

Los autores desean agradecer al Dr. J. M. Valpuesta la aportación desinteresada de las imágenes de microscopía electrónica. Este trabajo ha sido financiado por la CICYT, proyecto TIC96-1125-C03-03.

## Referencias

- [1] D. E. Rumelhart and G. E. Hinton and R. J. Williams. Learning Representations by Back-propagating Errors. *Nature*, 323:533–536, October 1986.
- [2] G. Held and T. R. Marshall. *Data and Image Compression*. Wiley, fourth edition, 1996.
- [3] I. H. Witten, R. M. Neal and J. G. Cleary. Arithmetic Coding for Data Compression. *Communications of the ACM*, 30(6):520–540, June 1987.
- [4] J.J. Fernández and J.M. Valpuesta and J.R. Sanjurjo and J.L. Carrascosa and J.M. Carazo. Image Processing of Two-Dimensional Biological Crystals in Transmission Electron Microscopy. In *Proc. VII National Symposium on Pattern Recognition and Image Analysis*, volume 1, pages 221–226, 1997.
- [5] M. J. Weingerger, G. Seroussi and G. Sapiro. LOCO-I: A Low Complexity, Context-Based, Lossless Image Compression Algorithm. In *Proceeding of the IEEE Data Compression Conference*. IEEE, March-April 1996.
- [6] M. Nelson and J. Gailly. *The Data Compression Book*. M&T Books, second edition, 1996.
- [7] R. C. Gonzalez y R. E. Woods. *Digital Image Processing*. Addison-Wesley, 1992.
- [8] R. Lippman. An Introduction to Computing with Neural Nets. *IEEE ASSP Magazine*, 3(4):4–22, April 1987.
- [9] S. Wong, L. Zaremba, D. Gooden and H. K. Huang. Radiologic Image Compression - A Review. *Proceedings of the IEEE*, 83(2):194–219, February 1995.