

Sintonización de la cuantización uniforme en compresores perceptuales de imagen basados en la transformada wavelet

Martínez-Rach, M. O.; López Granado, O.; Piñol Peral, P.; Pérez Malumbres, M.
Departamento de Física y Arquitectura de Computadores de la Universidad Miguel Hernández
{mmrach, otoniel, pablop, mels}@umh.es

Resumen — Cuando se usan esquemas de cuantización uniforme con dead zone, el rendimiento R/D final puede verse afectado por dos parámetros, el tamaño del dead zone y la ubicación del punto de reconstrucción en cada intervalo de cuantización. En este trabajo analizamos cómo afecta el ancho del dead zone a la calidad de la imagen reconstruida en codificadores wavelet que han sido mejorados perceptualmente usando la función de sensibilidad al contraste del sistema visual humano. Puesto que la aplicación de la CSF altera la distribución de los coeficientes wavelet en los intervalos de cuantización y afecta a los coeficientes que entran en el dead zone, una elección apropiada de estos parámetros de cuantización puede mejorar el rendimiento R/D. Tras un completo estudio de los efectos de estos parámetros del cuantizador, observamos que ajustándolo de manera óptima para cada imagen, se pueden obtener ahorros en tasa de bits de hasta el 7,7% respecto a la utilización de esquemas tradicionales de cuantización uniforme o con dead zone fijo.

Palabras clave—Cuantización perceptual; Codificación perceptual de imágenes; análisis del rendimiento R/D; Contrast Sensitivity Function; Codificadores de Imagen basados en wavelets.

I. INTRODUCCIÓN

EN un codificador de imagen o video con pérdida, el cuantizador es la etapa en la que se produce la pérdida de información. Obviamente la pérdida de información está estrechamente relacionada con la calidad de la imagen reconstruida y con la tasa de bits objetivo. Por tanto el cuantizador debe diseñarse con cuidado para obtener la mejor calidad posible para una determinada tasa de bits o para obtener la mínima tasa de bits posible para una misma calidad de la imagen reconstruida, es decir, debe ser diseñado para obtener la mejor relación rate-distortion (R/D).

Los esquemas de cuantización más utilizados en los estándares de codificación de imagen y video son: (a) Uniform Scalar Quantizer (USQ) usado en el JPEG, SPITH, MPEG-2, MPEG-4 y JPEG2000 Part I entre otros; (b) Uniform Scalar Dead Zone Quantizer (USDZQ) usado en los codificadores H.263, H.264/AVC y en HEVC; (c) Universal Coded Trellis Quantizer (UTCQ) usado en JPEG2000 Part II y finalmente (d) Uniform Variable Dead Zone Quantizer (UVDZQ) usado también en el codificador JPEG2000 Part II.

Todos estos esquemas de cuantización eliminan la información de la imagen que aportan los coeficientes localizados en la zona alrededor del cero, conocida como Dead Zone (DZ). La diferencia entre USQ y UDSZQ es básicamente el Dead Zone Size (DZS) como se muestra en la Fig. 1, donde además se puede observar que el tamaño de los intervalos de cuantización, permanece constante o uniforme, con valor Δ . Estos cuantizadores ubican el punto de reconstrucción en el centro del intervalo. Por ejemplo, en la Fig. 1, el punto de reconstrucción para cada coeficiente dentro del primer intervalo, entre d_1 y d_2 , es r_1 .

Un cuantizador UVDZQ, variable o parametrizado, puede actuar como un USQ o como un USDZQ. Se pueden modelar variando el paso de cuantización Δ y la ubicación del punto de reconstrucción, denotado por δ . El DZS se expresa normalmente como múltiplo de Δ y por tanto un USQ tiene un DZS DE 1Δ y un USDZQ tiene un tamaño fijo de DZ de 2Δ .

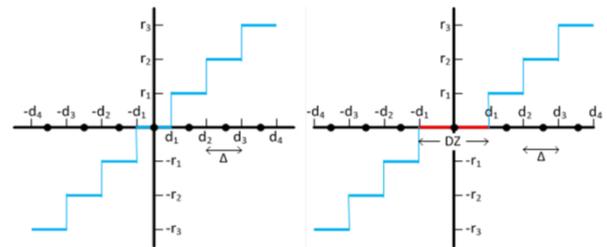


Fig. 1. Esquemas de cuantización uniforme. Izquierda USQ: DZS= 1Δ ; Derecha USDZQ: DZS= 2Δ

Para determinar qué coeficientes deben anularse, es decir, entrar en el dead zone, hay que establecer un compromiso entre calidad y rate. Variaciones en el DZS y en Δ afectan a la calidad de la reconstrucción y al rate obtenido, sin embargo, variaciones en δ sólo afectan a la calidad de reconstrucción y no a la tasa de bits, ya que este parámetro se usa sólo en el decodificador. Elegir una combinación óptima de estos tres parámetros al codificar una determinada imagen es difícil y más aún si queremos obtener un estimador de éstos que obtenga resultados similares al óptimo para cualquier imagen. La principal motivación de este trabajo es estudiar y analizar el papel que tiene el dead zone y el punto de reconstrucción en el rendimiento R/D de codificadores de imagen basados en wavelets y mejorados perceptualmente.

Otros trabajos en la literatura se han propuesto y analizado el diseño y el rendimiento de diferentes cuantizadores uniformes. En [1] los autores compararon el rendimiento de los esquemas USQ, USDZQ y UTCQ con codificadores wavelet sin codificación perceptual. Sus resultados muestran que aunque los errores de reconstrucción fueron menores con el UTCQ, cuando se combina la cuantización uniforme con codificadores entrópicos, el USDZQ fue la mejor opción habiendo seleccionado con cuidado el DZS. Los resultados muestran que el USDZQ puede reducir efectivamente la entrada al codificador entrópico mejorando por tanto la relación R/D. Por tanto, los autores mostraron que un USDZQ parametrizado, es decir, un UVDZQ es apropiado para sistemas de compresión de imagen basados en transformadas wavelet o DCT.

Está muy extendido en la literatura usar el centro del intervalo de cuantización para ubicar δ [2], o usar el centroide de la distribución de coeficientes en cada intervalo de cuantización. Sin embargo algunos autores recomiendan otras ubicaciones de δ para cuantizadores basados en la transformada DCT [3].

Algunos trabajos analizan específicamente la importancia de estos parámetros (DCZ y δ). En el estándar H.264/AVC se propone un parámetro de redondeo f que controla la ubicación del punto de reconstrucción dentro del intervalo de cuantización, especificando $f=\Delta/3$ cuando se usa codificación intra y $f=\Delta/6$ cuando se usa codificación inter. En [4] los autores aplican un esquema de cuantización con dead zone variable para el H.264/AVC utilizando un parámetro de desplazamiento (offset) adicional para mejorar simultáneamente el tamaño del dead zone y la ubicación del punto de reconstrucción. De esta forma, argumentan que δ se ajusta mejor a la distribución de coeficientes dentro de los intervalos de cuantización.

En [5] se realizaron estudios analíticos para obtener un DZS óptimo para un rango de tasas de bits de hasta 1bpp. Los autores propusieron un algoritmo para obtener el DZS y Δ óptimos. Un cuantizador uniforme con esos óptimos minimiza el error cuadrático medio de la fuente cuantizada. Usaron una GGD (Generalized Gaussian Distribution) para probar el algoritmo con diferentes tipos de distribuciones de coeficientes, como gaussiana, laplaciana y otras con colas más extensas. En todos los casos el autor mantiene sin embargo δ en el centro del intervalo de cuantización.

También Marcellin et al. en [2,6] mostraron la influencia del DZS en el rendimiento R/D del codificador JPEG2000. También usaron una GGD con distribuciones gaussianas, laplaciana y otras para cubrir la variabilidad observada en las distribuciones de coeficientes wavelet para las imágenes comúnmente usadas. Los autores proponen un DZS de 1.5Δ , el cual puede proporcionar un ligero descenso en el MSE generando sin embargo una mejor percepción de la imagen reconstruida en zonas con poca textura. Como no hay un δ óptimo para todas las imágenes, el estándar JPEG2000 permite la elección del valor de δ entre 0 y 1, usando $\delta=1/2$ como valor recomendado para la mayoría de las imágenes.

El rendimiento del codificador puede aumentarse utilizando cuantizadores con dead zone y ajustando su

tamaño. En [7] el autor hizo un experimento con una única imagen y un codificador de imagen basado en wavelets para determinar cuál era el DZS que obtenía mejor rendimiento. Midió la ganancia en calidad al reemplazar un USQ por un USDZQ en el codificador wavelet. El estudio midió el rendimiento R/D utilizando el PSNR como métrica de distorsión. Para dicha imagen obtuvieron un DZS óptimo de 1.9Δ que proporcionaba un incremento de calidad de 0,5 dBs.

En los estudios mencionados se han utilizado diferentes cuantizadores uniformes, remarcando la influencia de ambos parámetros (DZS y δ) en el rendimiento R/D de codificadores wavelet. Como se observa, estos estudios se hicieron con condiciones heterogéneas que proporcionan diferentes propuestas de cuál es el DZS óptimo, el punto de reconstrucción sugerido se sitúa en el centro del intervalo y no consideran ninguna aproximación de codificación perceptual.

En este trabajo realizaremos un extenso estudio para determinar cómo afecta el tamaño del dead zone DZS y la ubicación del punto de reconstrucción δ , en el rendimiento R/D de codificadores wavelet que han sido mejorados perceptualmente con la inclusión de la curva de sensibilidad al contraste, Contrast Sensitivity Function (CSF) del sistema visual humano.

Cuando se aplican técnicas de codificación perceptual como la CSF los coeficientes transformados se ponderan conforme a la importancia perceptual de la información que aportan. Se puede incluir la CSF en los codificadores de diversas formas como se expone en [8] pero la idea fundamental es que aquellos coeficientes localizados en las frecuencias espaciales para las que el sistema visual humano es más sensible sean elevados o ponderados en mayor amplitud que otros menos importantes perceptualmente.

La aplicación de la CSF modifica la distribución de los coeficientes wavelets proporcionando uniformidad perceptual, de forma que cuando posteriormente se aplique un cuantizador uniforme, éste afecte perceptualmente a todos los coeficientes en igual medida. Para maximizar el rendimiento R/D para un amplio rango de tasas de bits es importante elegir la relación óptima entre del DZS y Δ , conjuntamente con el mejor valor de δ .

Para determinar la magnitud de mejora del rendimiento R/D usaremos un UVDZQ en un codificador wavelet que incluye la CSF midiendo el rendimiento en términos de la métrica perceptual PSNR-HVS puesto que como ciertos estudios sugieren [9,10,11] las comparaciones de rendimiento de codificadores que usen técnicas de codificación perceptual deben ser hechas con métricas objetivas de valoración perceptual de la calidad (perceptual quality assessment metrics). Cubriremos un amplio rango de tasas de bit, hasta 3 bpp, es decir, desde muy baja calidad hasta niveles de calidad perceptual sin pérdida (visually lossless), incrementando el valor de Δ .

El resto del artículo está organizado como sigue: en la sección II se repasa brevemente la formulación de los diferentes esquemas de cuantización y cómo se pueden relacionar entre sí. En la sección III se expone la metodología utilizada para realizar el trabajo. En la

sección IV presentamos los resultados de las comparaciones y en la sección V abordamos las conclusiones y líneas futuras.

II. ESQUEMAS DE CUANTIZACIÓN

En esta sección repasamos brevemente la formulación para los cuantizadores USQ, USDZQ y UVDZQ y cómo el UVDZQ puede ser considerado un cuantizador uniforme universal al ser posible funcionar como un USQ o un USDZQ ajustando apropiadamente los parámetros de cuantización.

Cualquier cuantizador puede ser descompuesto en dos etapas distintas, referidas como la etapa de clasificación (o cuantización directa) y la etapa de reconstrucción (o cuantización inversa). Las ecuaciones (1) y (2) corresponden a la cuantización directa e inversa para un USQ. Las ecuaciones (3) y (4) representan esas etapas para un USDZQ y finalmente las ecuaciones (5) y (6) corresponden a un UVDZQ.

$$C' = \text{sign}(C) \left\lfloor \frac{|C|}{\Delta} + \frac{1}{2} \right\rfloor \quad (1)$$

$$\hat{C} = \Delta C' \quad (2)$$

$$C' = \text{sign}(C) \left\lfloor \frac{|C|}{\Delta} \right\rfloor \quad (3)$$

$$\hat{C} = \text{sign}(C')(|C'| + \delta)\Delta \quad (4)$$

$$C' = \begin{cases} \text{sign}(C) \left\lfloor \frac{|C| + \xi\Delta}{\Delta} \right\rfloor & \text{if } |C| \geq -\xi\Delta \\ 0 & \text{if } |C| < -\xi\Delta \end{cases} \quad (5)$$

$$\hat{C} = \begin{cases} \text{sign}(C')(|C'| - \xi + \delta)\Delta & \text{if } C' \neq 0 \\ 0 & \text{if } C' = 0 \end{cases} \quad (6)$$

Donde C es el coeficiente transformado antes de la cuantización. C' es el coeficiente cuantizado justo tras la etapa directa y \hat{C} es el valor reconstruido después de la etapa de cuantización inversa. El USQ recupera el valor del coeficiente en el centro del intervalo. La constante δ , usada en los otros esquemas, establece la ubicación del punto de recuperación. Los valores permitidos para δ están en el rango $[0..1]$. La constante ξ define el tamaño del dead zone, permitiéndose valores en el rango $(-\infty..1]$, y finalmente el paso de cuantización Δ determina la cantidad de cuantización y por tanto el nivel de compresión.

El parámetro ξ tal que $(\xi \leq 1)$, determina el tamaño del dead zone, el DZS, en un cuantizador UVDZQ. Dependiendo de su valor el dead zone se fija de la siguiente manera:

- $\xi < 0$ incrementa el tamaño típico de un USDZQ, es decir, $DZS > 2\Delta$
- $\xi = 0$ produce un dead zone exactamente de $DZS = 2\Delta$, siendo Δ el primer punto de decisión o umbral de cuantización, es decir d_1 en Fig.1.
- $0 < \xi < 1$ reduce el tamaño del dead zone por debajo del típico, es decir, $DZS < 2\Delta$, donde un valor de $\xi = 0.5$ corresponde con un USQ con $DZS = 1\Delta$.
- Conforme ξ se aproxima a 1 el DZS se reduce hasta 0, por tanto cuando $\xi=1$ entonces $DZS=0$.

Para poder ajustar un UVDZQ que funcione como un USQ tenemos que usar $\xi = 0.5$ y $\delta=0.5$, y para que funcione como un USDZQ con el punto de reconstrucción en el centro del intervalo de cuantización tenemos que usar $\xi = 0.0$ y $\delta=0.5$.

III. MÉTODOS

Como mencionábamos, el objetivo de este estudio es analizar cómo afecta el tamaño del dead zone, DZS, y la ubicación del punto de reconstrucción δ , al rendimiento R/D de codificadores wavelet con mejoras perceptuales, concretamente la inclusión de la CSF.

Por tanto en este estudio utilizamos un codificador de imagen descrito en [12] y una métrica perceptual o QAM, en concreto la PSNR-HVS, para medir la calidad perceptual de cada uno de los esquemas de cuantización [13] por medio de la implementación en C++ de dicha métrica hecha en la herramienta Video Quality Measurement Tool (VQMT) [14]. El codificador utiliza una matriz de pesos optimizada que implementa uno de los modelos más utilizados de la CSF [15].

Las variaciones en estos dos parámetros, DZS y δ , producen diferentes rendimientos o curvas R/D. Analizaremos el rendimiento R/D desde dos perspectivas. La primera desde la ganancia de calidad perceptual, es decir, comparamos la calidad para la misma tasa de bits, y en la segunda comparamos las tasa de bits para un mismo nivel de calidad en la imagen reconstruida.

Para proporcionar un valor de ahorro de tasa de bits y ganancia en calidad usamos el método Bjontegaard [16] para representar los resultados como porcentajes de ganancia para distintos rangos de tasa de bits y distintos intervalos de calidad. Adicionalmente proporcionaremos valores absolutos de la tasa de bits expresados en bits por pixel (bpp).

Hemos definido un conjunto representativo de imágenes de entrenamiento (768x512) compuesto por las 23 imágenes del Kodak Set. Para cada imagen del conjunto de entrenamiento obtendremos el par de parámetros ξ y δ que maximizan el área de la curva R/D. Para hacer esto creamos un espacio 2D de valores (ξ, δ) para los rangos que se muestran a continuación. Para cada combinación de estos parámetros codificamos y decodificamos la imagen en valores crecientes de stepsize, Δ . Medimos la calidad en PSNR-HVS dBs y la tasa de bits en bpp.

- $-0.250 \leq \xi \leq 1$ en saltos de 0.010 ξ obteniendo 126 valores diferentes. Este rango produce DZS que varían desde 2.5Δ a 0 en saltos de -0.02Δ
- $0 \leq \delta \leq 1$ usando saltos de 0.1 δ para obtener 11 posiciones diferentes variando de izquierda a derecha en el intervalo de cuantización.

Se han calculado un total de 1.386 permutaciones (ξ, δ) . Para cada una de ellas hemos creado una curva R/D al usar 13 valores de Δ diferentes, distribuidos uniformemente en el rango de tasa de bits. Esto produce un total de 18.018 imágenes comprimidas para cada una de las 23 imágenes del conjunto de entrenamiento.

La Fig. 2 muestra el comportamiento R/D para la imagen '01' del conjunto de entrenamiento. Cada curva corresponde con uno de los valores de ξ dentro del rango

de evaluación, pero sólo la curva con mejor rendimiento de δ se ha dibujado. La curva con mejor rendimiento (ξ, δ) se ha remarcado con una línea más gruesa. Se ha elegido un rango tan grande de valores de ξ para evitar máximos locales en el cálculo del área de las curvas y para tener una perspectiva general de cómo es el comportamiento R/D de las curvas en el rango.

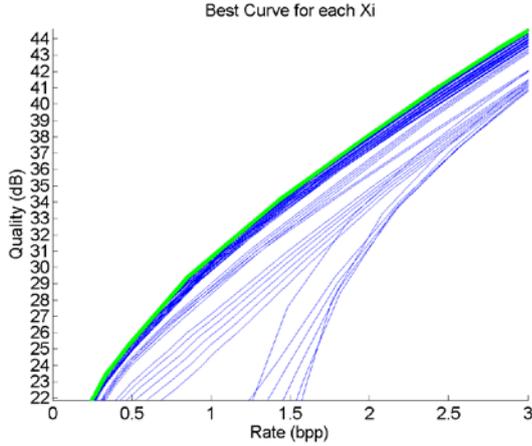


Fig. 2. Mejor curva R/D para cada ξ del conjunto de entrenamiento. En esta figura sólo el mejor δ es mostrado para cada ξ

TABLA I

RANGO DE TASAS DE BITS BAJO ESTUDIO

Rate Ranges	Rate (bpp)		Quality (dB)	
	Low	high	low	high
<i>ALL</i>	0.0	3.0	25.0	50.5
<i>L</i>	0.0	0.5	25.0	32.1
<i>M</i>	0.5	1.0	32.1	37.6
<i>H</i>	1.0	1.5	37.6	41.7
<i>VH</i>	1.5	3.0	41.7	50.5

TABLA II

ξ Y δ ÓPTIMOS PARA EL CONJUNTO DE ENTRENAMIENTO

Image	Optimum ξ	Optimum δ	DZS
1	0.460	0.4	1.08 Δ
2	0.400	0.4	1.20 Δ
3	0.300	0.5	1.40 Δ
4	0.450	0.4	1.10 Δ
5	0.450	0.5	1.10 Δ
6	0.390	0.5	1.22 Δ
7	0.350	0.5	1.30 Δ
8	0.390	0.5	1.22 Δ
9	0.380	0.4	1.24 Δ
10	0.400	0.4	1.20 Δ
11	0.380	0.5	1.24 Δ
12	0.450	0.4	1.10 Δ
13	0.450	0.4	1.10 Δ
14	0.490	0.4	1.02 Δ
15	0.380	0.4	1.24 Δ
16	0.370	0.5	1.26 Δ
17	0.400	0.4	1.20 Δ
18	0.450	0.4	1.10 Δ
19	0.350	0.5	1.30 Δ
20	0.350	0.5	1.30 Δ
21	0.380	0.5	1.24 Δ
22	0.460	0.4	1.08 Δ
23	0.380	0.4	1.24 Δ
<i>Min</i> (ξ, δ)	0.300	0.40	1.40 Δ
<i>Mean</i> (ξ, δ)	0.403	0.44	1.19 Δ
<i>Max</i> (ξ, δ)	0.490	0.50	1.02 Δ

Medimos el área de cada una de estas curvas respecto al eje de abscisas pudiendo, por tanto, determinar que par de valores (ξ, δ) obtienen la mayor área. La curva con mayor área es la que tiene el mejor rendimiento R/D dentro del rango estudiado. Dicho rango va desde 0 bpp hasta 3.0 bpp. La calidad, expresada en PNSR-HVS dBs varía para todo el conjunto de entrenamiento desde 17.3 dBs hasta 57.7 dBs, es decir, desde muy mala calidad respecto al original hasta visualmente imperceptible. El rango dinámico de calidades para cada imagen varía dependiendo de su contenido.

Para cada imagen escogemos los parámetros (ξ, δ) con mejor rendimiento, los que nos proporcionan la curva R/D óptima. Llamamos a esa curva C_{OPT} y la comparamos con las curvas correspondientes a los parámetros para USQ y USDZQ, llamadas C_{USQ} y C_{USDZQ} con ($\xi=0.5, \delta=0.5$) y ($\xi=0.0, \delta=0.5$) respectivamente. En esta comparación medimos la ganancia en tasa de bits y calidad para C_{OPT} respecto a C_{USQ} y C_{USDZQ} utilizando el método Bjontegaard.

Analizamos estas diferencias para todo el rango de tasa de bits y también para diferentes subrangos que etiquetamos como bajos, medios, altos y muy altos. La tabla I muestra estos rangos con una aproximación de sus correspondientes intervalos de calidad. Estos rangos son ALL para todo el rango, L para tasas bajas, M para medias, H para altas y VH para tasa de bit muy altas. Dependiendo del contenido de la imagen, la tasa de compresión es diferente para cada imagen, por tanto, las columnas para calidad en la tabla I muestran valores promedios de todas las imágenes para los límites de calidad alto y bajo.

Se calcula la ganancia en calidad o tasa de bits para C_{OPT} respecto a C_{USQ} y C_{USDZQ} en cada uno de estos rangos.

De esta forma tenemos la ganancia máxima para cada imagen y rango de tasa. Las curvas C_{OPT} , C_{USQ} y C_{USDZQ} se han estimado con un proceso automático de ajuste de curvas (curve fitting) que busca el mejor ajuste usando modelos polinómicos y racionales proporcionados por la herramienta Matlab Curve Fitting Toolbox. Una vez tenemos los parámetros que modelan la curva podemos obtener las diferencias absolutas en tasa de bits y calidad para cualquier rango. Se obtienen también los valores absolutos medios de ahorro en tasa de bits.

Tras analizar el conjunto de entrenamiento usaremos un nuevo conjunto de test compuesto por las siguientes imágenes: (512 x 512): Balloon, Barbara, Boat, Goldhill, Horse, Lena, Mandrill y Zelda; (2048 x 2560): Bike, Cafe y Woman. Como el par (ξ, δ) óptimo es diferente para cada imagen, obtenemos un valor representativo, es decir un par (ξ, δ) subóptimo basado en los parámetros del conjunto de entrenamiento.

IV. RESULTADOS

La tabla II muestra los parámetros (ξ, δ) óptimos para cada una de las imágenes del conjunto de entrenamiento. Como podemos ver, el parámetro δ , que fija la ubicación del punto de recuperación en el intervalo, permanece casi constante cerca del centro de éste con un valor medio de $\delta = 0.44$. Por otro lado, el valor óptimo de ξ varía en un rango desde 0.30 a 0.49 unidades que

corresponde a dead zones de 1.02Δ a 1.40Δ respectivamente.

Como hemos mencionado anteriormente, valores menores de ξ producen mayores DZS, por tanto en la tabla II podemos ver que para el menor ξ (0.30) obtenemos el DZ más ancho, 1.40Δ . Para el todo el conjunto de entrenamiento obtenemos un DZS medio de 1.19Δ que corresponde con un dead zone un 19% más ancho que el de un cuantizador USQ y un 40.5% más estrecho que el de un USDZQ. Por tanto, una primera conclusión es que puesto que el valor óptimo de DZS es más próximo al de un USQ que al de un USDZQ y por tanto un USQ proporcionará mejor rendimiento R/D que un USDZQ.

TABLA III

RESULTADOS PROMEDIO PARA EL CONJUNTO DE ENTRENAMIENTO

Rate Ranges	Quality Gain (dB)		% of Rate Saving	
	USQ	UDZQ	USQ	UDZQ
ALL	0.15	0.36	1.48	3.31
L	0.12	0.25	2.01	4.12
M	0.12	0.32	1.55	3.96
H	0.17	0.38	1.74	3.76
VH	0.17	0.40	1.38	2.96

TABLA IV

GANANCIAS MÁXIMAS PARA EL CONJUNTO DE ENTRENAMIENTO

Rate Ranges	Quality Gain (dB)		% of Rate Saving	
	USQ	UDZQ	USQ	UDZQ
ALL	0.30	0.53	3.23	4.55
L	0.26	0.41	5.30	6.64
M	0.25	0.60	3.53	7.67
H	0.34	0.60	4.54	5.92
VH	0.35	0.59	3.07	4.27

TABLA V

RESULTADOS PROMEDIO PARA EL CONJUNTO DE TEST.
PSNR-HVS dBs PARA CALIDAD Y % DE AHORRO EN TASA DE BITS

Rate Ranges	Quality Gain (dB)		% of Rate Saving	
	USQ	USDZQ	USQ	USDZQ
512 x 512				
ALL	0.22	0.34	2.17	3.28
L	0.17	0.17	2.90	3.02
M	0.16	0.32	2.17	4.02
H	0.23	0.36	2.52	3.69
VH	0.25	0.38	2.02	3.00
2048 x 2560				
ALL	0.16	0.41	1.43	3.62
L	0.08	0.25	1.29	3.97
M	0.11	0.40	1.30	4.53
H	0.15	0.40	1.48	3.75
VH	0.19	0.45	1.45	3.36

Como los resultados dependen del contenido de la imagen, mostramos los valores medios y máximos para el conjunto de entrenamiento. En la Tabla III se muestran los valores medios de ganancia en tasa de bits y calidad de las imágenes, y en la tabla IV se muestran las correspondientes ganancias máximas. Como se ve, con el UVDZQ optimizado se obtienen ahorros de tasa de bits de hasta 6.64% en el rango L y hasta un 7.67% en el rango M. Respecto a las ganancias en calidad, se

obtienen incrementos de calidad de 0.41dBs y 0.60dBs en los rangos L y M respectivamente.

La Fig. 3 muestra la comparación de las curvas R/D para la imagen '01' del conjunto de entrenamiento, considerando un rango de tasa de bits desde 0 hasta 1.5 bpp (rangos L,M y H). Como se muestra, el ahorro de tasa se mantiene para todo el rango. La Fig. 4 se centra sólo en las curvas R/D en el rango H, en este caso para la imagen '14' del conjunto de entrenamiento. Se obtiene una ganancia de calidad de 0.60dB al usar el óptimo respecto a USDZQ. Esto corresponde con un ahorro en tasa de bits de un 5.9%.

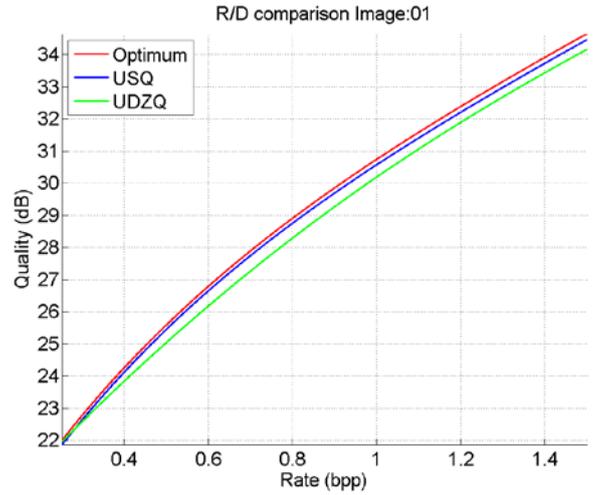


Fig. 3. Comparación R/D entre los parámetros óptimos, USQ y USDZQ para la imagen '01' en los rangos de tasa de bits L,M y H

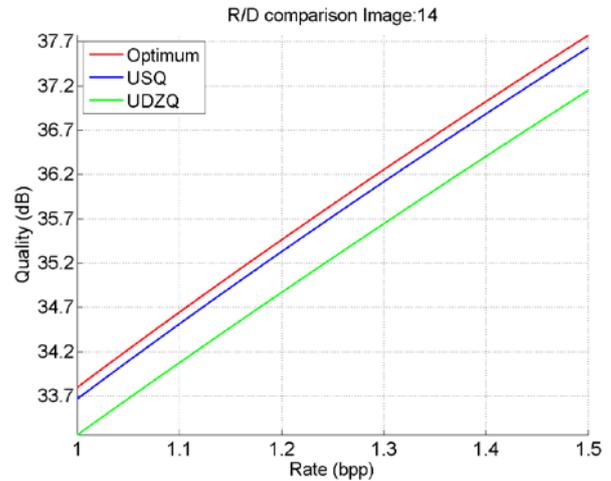


Fig. 4. Comparación R/D entre los parámetros óptimos, USQ y USDZQ para la imagen '14' en el rango de tasa de bits H

Ahora procedemos a evaluar el rendimiento R/D del UVDZQ optimizado con las imágenes del conjunto de test. Como describimos anteriormente, para ello deberíamos calcular el par de parámetros óptimo (ξ, δ) para el nuevo conjunto. Sin embargo, para poder aplicar en la práctica sin necesidad de repetir el proceso, proponemos una aproximación simple eligiendo el centroide del óptimo DSZ y δ obtenidos del conjunto de entrenamiento (ver Tabla II). El par de parámetros resultante estimado (\overline{DSZ}, δ) será el usado en las imágenes del conjunto de test.

Utilizando esta aproximación simple, calculamos para el conjunto de test las curvas para USQ, USDZQ y el

UVDZQ optimizado con los parámetros (\widehat{DSZ}, δ). En la tabla V mostramos los valores medios de ganancia en calidad y ahorro de tasa para las imágenes del conjunto de test. Como puede verse se obtienen resultados similares a los obtenidos con el conjunto de entrenamiento. En particular podemos obtener hasta un 4,02% de ahorro de tasa en el rango M para las imágenes a una resolución de 512x512 y hasta un 4,53% de ahorro para el rango M y las imágenes de resolución 2048x2560.

TABLA VI

RESULTADOS ABSOLUTOS PARA ALGUNOS RANGOS DE TASA DEL CONJUNTO DE TEST.
PSNR-HVS dBs PARA CALIDAD Y % DE AHORRO EN TASA DE BITS

Img. Size	Image	Quality Gain (dB)		% of Rate Saving	
		USQ	USDZQ	USQ	USDZQ
Rate Range: H (1.0 to 1.5 bpp)					
512x512	balloon	0.25	0.31	2.25	2.70
	barbara	0.15	0.36	1.41	3.43
	boat	0.15	0.54	1.36	5.10
	goldhill	0.13	0.47	1.43	5.14
	horse	0.26	0.41	2.59	4.08
	lena	0.32	0.23	4.04	2.97
	mandrill	0.31	0.39	3.30	4.11
zelda	0.27	0.14	3.79	1.98	
Rate Range: M (0.5 to 1.0 bpp)					
2048x2560	bike	0.16	0.28	1.83	3.18
	cafe	0.05	0.54	0.46	5.68
	woman	0.13	0.39	1.61	4.73
Rate Range: L (0.0 to 0.5 bpp)					
2048x2560	bike	0.08	0.18	1.15	2.75
	cafe	0.04	0.29	0.71	4.55
	woman	0.12	0.29	2.01	4.60

En la tabla VI mostramos resultados específicos para diversos rangos de compresión. El UVDZQ optimizado obtiene hasta un 5,14% de ahorro en rate en la imagen 'goldhill' en el rango de rate H y 5,68% para la imagen 'cafe' en el rango M. Además mejora el rendimiento perceptual hasta 0,54 PSNR-HVS dBs para algunas imágenes del conjunto.

V. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo hemos usado un UVDZQ para analizar cómo afectan el ancho del dead zone y la ubicación del punto de reconstrucción al rendimiento R/D de un codificador wavelet mejorado perceptualmente. De este estudio se observa que cada imagen tiene un par de parámetros (ξ, δ) óptimos para los cuales el rendimiento R/D de la imagen se maximiza en un amplio rango de tasa de bits, desde 0 bpp a 3.0 bpp.

Para cuantificar los beneficios por usar el par óptimo de parámetros de cuantización, el rendimiento R/D del UVDZQ optimizado se compara con el obtenido con cuantizadores USQ y USDZQ puesto que son los más comunes en compresión de imagen. Los resultados muestran que se pueden obtener ahorros de tasa de hasta un 4,55% y un 3,23% de media para todo el rango de tasa de bits cuando se compara con el USDZQ y el USQ respectivamente. Sin embargo, como la curva R/D tiene un comportamiento no lineal, el UVDZQ óptimo puede obtener mayores ahorros de rate: 6,64% respecto USDZQ y 5,30% respecto a USQ en el rango de 0 a 0,5

bpp y 7,67% y 3,53% respectivamente en el rango de 0,5 a 1,0 bpp.

El esquema USQ obtiene mejores resultados que el USDZQ porque su ancho de dead zone es más próximo al del óptimo que se ha obtenido en ambos conjuntos de imágenes, entrenamiento y test y para todos los rangos de tasa. En otras palabras, un dead zone de 1Δ proporciona mejores resultados que uno de 2Δ cuando los coeficientes de la transformada se han ponderado con una matriz de elevación optimizada para la CSF. La CSF proporciona a cada coeficiente su peso o importancia perceptual correcta, de forma que un simple cuantizador uniforme aplicado con posterioridad afecte perceptualmente en la misma medida a todos ellos.

Con los resultados obtenidos podemos asegurar que el uso de un cuantizador UVDZQ optimizado puede en general mejorar el rendimiento R/D de cuantizadores wavelet optimizados perceptualmente. Proponemos una aproximación al par de parámetros óptimo eligiendo el centroide de los DSZ y los δ obtenidos del conjunto de imágenes de entrenamiento, como par sub-óptimo para ser usado en otras imágenes.

Los resultados muestran que incluso con este par sub-óptimo se obtienen buenos resultados para una variedad de imágenes. El UVDZQ obtiene con él hasta 5,68% y 4,73% de ahorro de rate respecto el USDZQ para las imágenes 'cafe' y 'woman' respectivamente en el rango entre 0,5 y 1,0 bpp o hasta 4,5% y 4,6% respectivamente en el rango de 0 a 0,5 bpp.

Aunque se debe realizar más investigación para obtener un estimador del par de parámetros de cuantización óptimo para cualquier imagen basándose en propiedades estadísticas de la imagen o de los coeficientes u otros parámetros como por ejemplo la entropía, los resultados con esta simple aproximación son suficientemente buenos y cercanos al óptimo.

I. AGRADECIMIENTOS

El presente trabajo ha sido financiado por el Ministerio de Economía y Competitividad mediante el proyecto TIN2015-66972-C5-4-R cofinanciado con fondos MINECO/FEDER UE.

II. REFERENCIAS

- [1] Jinhua Yu, "Advantages of uniform scalar dead-zone quantization in image coding system," Communications, Circuits and Systems, 2004. ICCAS 2004. 2004 Int. Conference on, 2004, pp. 805-808 Vol.2.
- [2] Michael W. Marcellin, Margaret A. Lepley, Ali Bilgin, Thomas J. Flohr, Troy T. Chinen, and James H. Kasner. "An overview of quantization in JPEG2000" . Signal Processing: Image Communication, 17:73–84, 2002
- [3] S. Notebaert, J. De Cock, K. Vermeersch, P. Lambert and R. Van de Walle, "Leveraging the quantization offset for improved requantization transcoding of H.264/AVC video," Picture Coding Symposium, 2009. PCS 2009, Chicago, IL, 2009, pp. 1–4
- [4] T. Wedi and S. Wittmann, "Quantization offsets for video coding," 2005 IEEE Int. Symposium on Circuits and Systems, 2005, 324-327 Vol. 1.
- [5] Bo Tao, "On optimal entropy-constrained deadzone quantization," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 4, pp. 560-563, Apr 2001.
- [6] Michael W Marcellin. "JPEG2000: image compression fundamentals, standards, and practice", volume 1. springer, 2002
- [7] Jacob Ström. "Dead zone quantization in wavelet image compression" - mini project in ece 253a, 1996.

- [8] Marcus J. Nadenau, Julien Reichel, and Murat Kunt. "Wavelet-based color image compression: Exploiting the contrast sensitivity function." *IEEE Transactions on image processing*, 12(1), 2003.
- [9] M. Martinez-Rach, O. Lopez, P. Piñol, J. Oliver, and M. Malumbres, "A study of objective quality assessment metrics for video codec design and evaluation," in *Eight IEEE International Symposium on Multimedia*, vol. 1, ISBN 0-7695-2746-9. San Diego, California: IEEE Computer Society, Dec 2006, pp. 517–524
- [10] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. on Image Processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
- [11] Z. Wang, A. Bovik, H. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004.
- [12]
- [13] Martinez-Rach, Miguel Onofre. "Perceptual image coding for wavelet based encoders". PhD thesis, Universidad Miguel Hernández. <http://hdl.handle.net/11000/1764> December 2014.
- [14] Egiazarian, K., Astola, J., Ponomarenko, N., Lukin, V., Battisti, F., & Carli, M. (2006, January). New full-reference quality metrics based on HVS. In *CD-ROM proceedings of the second international workshop on video processing and quality metrics*, Scottsdale, USA (Vol. 4).
- [15] Multimedia Video Processing Group, "VQMT: Video Quality Measurement Tool", <http://mmspg.epfl.ch/vqmt>
- [16] J. Mannos and D. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 525 – 536, Jul. 1974.
- [17] G. Bjontegaard. "Calculation of average psnr differences between rdcurves (vceg-m33)". Technical report, VCEG Meeting (ITU-T SG16 Q.6), Austin, Texas, USA, April 2001.