Desarrollo de Algoritmos para Muestreo Compresivo Aplicado a Señales de Audio

Rodolfo G. Moreno^{(1),(2)}, Mauricio A. Martínez⁽¹⁾, Mariko Nakano⁽²⁾, Héctor M. Pérez⁽²⁾

(1) Universidad la Salle, Facultad de Ingeniería, Benjamín Franklin 47, C.P. 06140, México D.F.-México

(2) Escuela Superior de Ingeniería Mecánica y Eléctrica, Instituto Politécnico Nacional. Av. Santa Ana 1000, C.P. 04430 México D.F. México (e-mail: rg.moreno@lasallistas.org.mx; mauricio.martinez@lasallistas.org.mx; mnakano@ipn.mx; hmperez@ipn.mx)

Recibido Dic. 4, 2013; Aceptado Ene. 24, 2014; Versión final recibida Mar. 12, 2014

Resumen

En este artículo se propone la aplicación de la transformada discreta de coseno (DCT) y el muestreo compresivo (MC) para la compresión de señales de audio. Usando el análisis espectral y las propiedades de la DCT se obtiene una señal de audio dispersa en el dominio de la frecuencia, lo cual permite el empleo de técnicas de MC para este tipo de señales en el dominio de la DCT. Así, se propone el uso de la DCT y el MC para obtener una representación eficiente de las señales de audio donde la DCT opera como un módulo de pre-procesamiento que permite obtener una representación dispersa de la señal en el dominio de la frecuencia, permitiendo la posterior aplicación del MC a las señales de audio. Con esto es posible obtener una representación de las señales de audio con menos muestras que las requeridas por el conocido teorema de muestreo.

Palabras clave: muestreo compresivo, señales de audio, transformada discreta de coseno, dispersión en frecuencia, compresión

Development of compressive sampling algorithms applied to audio signals

Abstract

In this paper, it is proposed to apply the discrete cosine transform (DCT) and the compressive sampling (CS) techniques to audio signal compression. Using spectral analysis and the properties of the DCT, audio signals can be treated as sparse signals in the frequency domain. This is especially true for sounds representing tones. Thus, it is proposed using the DCT as a preprocessor to obtain a sparse representation in the frequency domain, combined with CS to obtain an efficient representation of the signals. It is shown that the subsequent application of CS represents the signals with less information than the well-known sampling theorem.

Keywords: audio signals, compressive sampling, discrete cosine transform, frequency sparse, compression

INTRODUCCIÓN

En los últimos años, se ha incrementado considerablemente el interés en el estudio del muestreo compresivo (Eldar y Kutyniok, 2012), siendo esta, una nueva forma de muestrear y comprimir cierto tipo de señales. En el muestreo compresivo, el modelo tradicional limitado en banda planteado por el teorema de Shannon-Nyquist-Whittaker-Kotelnikov es remplazado por un modelo disperso, el cual, asume que una señal puede ser eficientemente representada usando solamente un número pequeño de proyecciones lineales no adaptativas presentes en el dominio de alguna transformada. Los innovadores trabajos escritos por (Candes y Col 2006; y Donoho 2006) muestran que ciertas señales consideradas dispersas pueden ser precisamente reconstruidas a partir de un pequeño conjunto de mediciones lineales lo cual implica una potencial reducción de las tasas de muestreo, bajo consumo de energía y complejidad computacional en la adquisición de datos digitales, cabe mencionar que debido a la naturaleza dispersa de la señal, esta técnica se aplica directamente a imágenes, radar y corrección de errores, siendo para otras aplicaciones necesario un pre procesamiento de la señal como es el caso de las señales de audio.

Debido al avance en el campo del procesamiento de señales el cual apunta hacia una mayor calidad en la señal adquirida y una alta eficiencia en la compresión de datos, el muestreo compresivo es muy atractivo especialmente para dispositivos con baja resolución o simplemente cuando es demasiado costoso tomar mediciones en aplicaciones que involucran el empleo de frecuencias de muestreo extremadamente altas. Por lo anterior, desde el descubrimiento de la teoría del muestreo compresivo se han desarrollado diversos algoritmos que toman mediciones en el dominio de la transformada de Fourier como sucede en aplicaciones de resonancia magnética MRI (Lusting y Col., 2008).

Sin embargo, a pesar de los trabajos mencionados aun existe una brecha muy amplia entre la teoría del muestreo compresivo y las aplicaciones a señales de audio (Sreenivas y Kleijn, 2009; Giacobello y Col., 2008). Particularmente se desconoce cómo obtener una buena representación dispersa de una señal de audio, lo cual es pieza fundamental ya que el muestreo compresivo se fundamenta en dos propiedades: Dispersión e Incoherencia, donde la dispersión pertenece a la señal de interés, y la incoherencia a la modalidad en la cual se capta la señal (Candes y Romberg, 2007; 2006; Khayam, 2003). Como una posible solución al problema relativo a la creación de una representación dispersa de la señal de audio, se ha propuesto el uso de la transformación coseno discreta (DCT), la cual es ampliamente usada en algoritmos para la compresión de imágenes, video y es la herramienta más utilizada para obtener una representación dispersa de la señal de voz (Haiyan y Col, 2012; Kassim y Col 2012; Moreno-Alvarado y Martinez-Garcia 2011). Su popularidad se basa en el hecho de que ésta proporciona una compactación de datos relativamente buena debido a que concentra la mayor parte de la información en unos pocos coeficientes (Khayam, 2003; Rao y Yip, 1990). Esto significa que se puede lograr una versión comprimida de la señal de audio obteniendo en primera instancia una representación dispersa de la misma en el dominio de la frecuencia, lo cual permite su procesamiento mediante algoritmos de muestreo compresivo (Candes y Wakin, 2008).

Por tal motivo en este artículo se presenta la aplicación de la transformación coseno discreta en tramas a dos segmentos de señales de audio (voz y música), con el fin de obtener una representación dispersa de las señales de audio y aplicar el algoritmo de muestreo compresivo, en las siguientes secciones: muestreo compresivo y transformada coseno discreta se presentan algunos conceptos básicos necesarios para la realización del método desarrollado, en la sección de algoritmo propuesto se detalla la metodología y las herramientas empleadas, los resultados mas relevantes se detallan en la sección de resultados obtenidos y finalmente se presentan algunas conclusiones del trabajo realizado.

MUESTREO COMPRESIVO

El desarrollo del muestreo compresivo ha sido objeto de una intensa investigación durante los últimos años, los resultados se encuentran publicados en una serie de artículos que desarrollan una teoría de reconstrucción de señales a partir de información incompleta (Candes y Col., 2006; Donoho, 2006; Lustig y Col. 2008; Candes y Romberg, 2006, 2007). Los resultados centrales de esta teoría establecen que un vector disperso de longitud N , $X^0 \in \mathbb{R}^N$ puede ser recuperado desde un conjunto pequeño de K mediciones lineales $y = \Phi X^0 \in \mathbb{R}^K$, con la condición $K \ll N$.

Considerando una señal X de longitud N y suponiendo que su base Ψ provee una representación K dispersa de ésta, esto implica que sólo presenta K -términos distintos a cero, entonces en términos de una notación matricial tenemos que $X = \Psi F$, donde F puede ser aproximada usando solo $K \ll N$ entradas distintas de cero, siendo Ψ la base dispersa de la matriz (Eldar y Kutyniok G 2012.). La teoría del muestreo compresivo establece que una señal puede ser reconstruida tomando solo $M=O(K \log N)$ mediciones lineales de la siguiente forma:

$$y = \phi X = \phi \psi F \quad (1)$$

donde y representa un vector de $M \times 1$ muestreado mientras que ϕ representa una matriz de medición de $M \times N$ la cual es incoherente con la base Ψ (Candes y Col., 2006; Donoho, 2006). Finalmente, con esta información se procede a recuperar la señal por medio de la minimización de la norma l_1 de manera que, si F es suficientemente dispersa, la recuperación por medio de la minimización de la norma l_1 es probablemente exacta (Candes y Col., 2006). Esto se debe a que la norma l_1 intrínsecamente favorece la reconstrucción de señales, por lo que para una matriz ϕ de dimensión $M \times N$ y un vector $y \in \mathbb{R}^M$ la solución de:

$$\min_{X \in \mathbb{R}^N} \|X\|_{l_1} \text{ sujeto a } y = \phi X \quad (2)$$

nunca deberá tener más de N -términos distintos a cero. De una manera concreta el algoritmo selecciona una base, la cual consiste de un conjunto de N vectores linealmente independientes, a partir del diccionario ϕ , para efectuar la reconstrucción de la señal (Candes y Romberg, 2007).

TRANSFORMADA COSENO DISCRETA

La transformación coseno discreta es ampliamente usada en el campo de procesamiento de señales e imágenes debido a sus propiedades de decorrelación y compactación de energía que la hacen una herramienta sumamente útil para el desarrollo de algoritmos de compresión, particularmente de imágenes y video. La transformación coseno discreta en una dimensión, con una longitud de N se define como (González y Wintz, 2008; Rao y Yip, 1990):

$$x(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{\pi(2x+1)u}{2N} \right], \quad u = 0, 1, 2, \dots, N-1, \quad (3)$$

Mientras que la transformada inversa está dada por (González y Wintz, 2008; Rao y Yip, 1990):

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) c(u) \cos \left[\frac{\pi(2n+1)u}{2N} \right], \quad n = 0, 1, 2, \dots, N-1. \quad (4)$$

En ambas ecuaciones (3) y (4) $\alpha(u)$ está definida como $\alpha(u) = \sqrt{1/N}$ para $u=0$ y $\alpha(u) = \sqrt{2/N}$ para $u \neq 0$, respectivamente. De (3) es claro que para $u = 0$ se obtiene:

$$x(u) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} f(x) \quad (5)$$

lo que implica que el primer coeficiente de la transformada es el valor promedio de la secuencia de entrada, también llamado coeficiente de DC siendo los demás coeficientes de esta transformación conocidos como los coeficientes de AC (Khayam, 2003).

Propiedades de la Transformación Coseno discreta

Algunas de las propiedades mas importantes de la transformación coseno discreta, al menos desde el punto de vista del muestreo compresivo en señales de audio, son la de-correlación de las señales de entrada y su capacidad de compactación de la energía (Rao y Yip, 1990). Una de las mayores ventajas de emplear la transformación coseno discreta es la remoción de redundancia entre los valores vecinos, lo que se conoce como propiedad de de-correlación, ya que esto permite que los coeficientes de la transformada, que están aproximadamente no correlacionados entre si puedan ser codificados de forma independiente. Por otro lado la eficacia de un esquema de transformación se mide directamente por la habilidad que tenga este de empaquetar los datos de entrada en tan pocos coeficientes como sea posible, lo que se conoce como compactación de energía. Esto permite descartar coeficientes con amplitudes relativamente pequeñas sin introducir una distorsión apreciable en la imagen o señal reconstruida. La Transformación coseno discreta exhibe una excelente compactación de energía para señales correlacionadas y una buena capacidad de de-correlacionar las señales de entrada, aunque sus coeficientes se concentran en las frecuencias bajas.

ALGORITMO PROPUESTO

En esta sección se presenta el algoritmo de muestreo compresivo para señales de audio, empleando las técnicas descritas anteriormente se describe la manera en que es posible representar las señales de audio en una forma dispersa, para posteriormente analizar la aplicación del algoritmo de muestreo compresivo. Considerando que una señal dispersa puede ser recuperada aproximadamente sin pérdidas a partir de solo unas pocas mediciones, se analiza la manera de representar las señales de audio en una forma dispersa, para posteriormente aplicar el algoritmo de muestreo compresivo a estas señales de audio resultantes, las cuales serán aproximadamente dispersas. El punto fundamental del algoritmo es obtener una reconstrucción aproximada de las señales de audio a partir de pocas mediciones, las cuales tendrán una tasa de muestreo mucho menor a la de Nyquist. Idealmente, sabemos que se miden todos los N coeficientes de la señal $f(x)$ dada por (4), sin embargo en el marco del muestreo compresivo solamente será necesario observar un conjunto menor de estos datos recolectados. Este proceso se ilustra en la figura 1.

Fig. 1. Sistema propuesto para el procesamiento de señales de audio usando muestreo compresivo.

Uno de los principales problemas a resolver para emplear técnicas de muestreo compresivo en señales de audio es el hecho de que estas no son señales dispersas por lo cual en el sistema propuesto, inicialmente la señal a ser procesada se transforma al dominio de la transformada coseno discreta (DCT) la cual proporciona una expansión de la señal de entrada en términos de una serie de funciones ortogonales en el intervalo $[0, N]$, como se muestra en (3) y (4) (Khayam, 2003). El empleo de la DCT, como se mencionó en las secciones anteriores, proporciona una decorrelación aproximada de los datos de entrada, por lo cual, después de esta decorrelación, cada coeficiente de la transformada puede ser codificado independientemente sin perder eficiencia en la compresión o compactación de la energía. Esta compactación de la energía de las señales de audio proporciona una representación dispersa en el dominio de la frecuencia lo que permite cumplir con el principio de dispersión, el cual consiste en tener una representación de la señal donde k - términos sean distintos de cero, siendo este principio una de las bases sobre el cual subyace el muestreo compresivo (Donoho., 2006; Eldar Y. y Kutyniok G., 2012). Considerando las propiedades de la DCT mostradas en el apartado anterior, la señal discreta dispersa, $X = [x(0), x(1), x(3), \dots, x(M-1)]^T$, se obtiene aplicando la transformada coseno discreta a la señal de entrada $f(n)$, la cual está dada por (3).

Una vez realizado el proceso de dispersión, se procede a tomar las muestras de forma aleatoria para lo cual se deben cumplir, tanto la propiedad de Isometría Reservada (RIP) como la de Incoherencia. La primera propiedad mide el grado en el cual cada subconjunto de k vectores dispersos columna, con no más de K elementos diferentes de cero, de una matriz A se comportan como ortonormales (Candes, E y Tao, T., 2008; Foucart 2010). Esta es una condición necesaria y suficiente para garantizar una eficiente recuperación de la señal de entrada, como una solución del problema inverso. Así específicamente una matriz satisface la RIP de orden k si existe un valor $\delta_k \in (0, 1)$ tal que

$$(1 - \delta_k) \|X\|_2^2 \leq \|\Phi X\|_2^2 \leq (1 + \delta_k) \|X\|_2^2 \quad (6)$$

Así, si Φ satisface la RIP de orden $2K$, entonces la condición (6) establece que Φ aproximadamente preserva la distancia entre cualquier par de K vectores dispersos. Esto tiene claramente implicaciones fundamentales relativas a la robustez al ruido (Davenport y Col., 2012).

En lo que respecta a la propiedad de incoherencia esta se refiere a que entre la base usada para la toma de mediciones y la señal no debe haber elementos correlacionados (Candes 2008; Donoho y Huo 2001; Foucart 2010) así, el muestreo compresivo aprovecha el hecho de que algunas señales presentan una estructura dimensional menor, comparada con el ambiente dimensional elevado del espacio en el que residen, como lo establece el teorema de Johnson–Lindenstrauss (1984). En este caso, se está interesado en recolectar información a partir del vector $X \in \mathbb{R}^N$ la cual está dada por $y = \Phi X$, donde Φ es una matriz de $n \times N$ donde $y \in \mathbb{R}^n$, por lo que la matriz Φ mapea $\mathbb{R}^N$ en donde N es generalmente mayor que n . Seguidamente para extraer la información que y tiene acerca de X se debe usar un decodificador Δ el cual debe mapear de $\mathbb{R}^n$ a $\mathbb{R}^N$, siendo el rol de Δ el de proveer una aproximación para $\bar{x} := \Delta(Y) = \Delta(\Phi x)$, por lo

que se utiliza una matriz aleatoria con el fin de cumplir con el criterio de isometría reservada (RIP) el cual está establecido por (6). De esta forma la matriz aleatoria tiene la capacidad de recuperar las muestras tomadas en forma incoherente de nuestra señal dispersa la cual está dada por la transformación coseno discreta de la señal de audio.

Reconstrucción de la señal de entrada

En muchas aplicaciones se requiere recuperar la señal original una vez que ésta ha sido transmitida o procesada, para lo cual se consideran los aspectos descritos anteriormente. Así dadas las muestras dispersas tomadas de forma aleatoria, esto es, muestras que satisfacen la condición de incoherencia, se puede reconstruir la señal empleando algunos de los métodos de reconstrucción propuestos tales como: a) El método BP (Basis Pursuit). b) Método OMP (Orthogonal Matching Pursuit). En el método BP, el proceso se reduce a un problema de optimización dado por

$$\min \|X\|_{l_1} \text{ sujeto a } \phi X = b \quad (7)$$

Este proceso recupera X exactamente bajo las condiciones siguientes: a) Incoherencia mutua de ϕ : Dado ϕ podemos recuperar X si X es lo suficientemente dispersa (Donoho y Huo 2001). b) Dada la dispersión de X , X se puede recuperar tomando como base las condiciones RIP de ϕ .

El método OMP, por su parte, es una alternativa viable para recuperar un vector K -disperso X sujeto a la condición $\phi X = b$, donde $\phi \in \mathbb{R}^{n \times N}$. El método OMP, el cual se aproxima iterativamente a la solución tiene la ventaja de ser muy rápido y fácil de implementar; y con frecuencia sus resultados son muy cercanos a aquellos obtenidos por medio de la minimización de la norma l_1 . Una realización del método OMP se muestra a en la Tabla 1 (Edar y Kutyniok, 2012).

Tabla 1. Método OMP para obtener la solución aproximada de X^k

Datos de entrada
Matriz $\phi = ((\phi_i)_{i=1}^n) \in \mathbb{R}^{n \times N}$ y el vector $X \in \mathbb{R}^N$
Umbral de error ε
1. $K=0$
2. $X^0=0$ Condicion inicial
3. $r^0=Y-\phi X^0$ Residual inicial
4. $S^0=\text{supp}X^0=\emptyset$
5. Repetir
6. $k=k+1$
7. Seleccione
i_0 tal que $\min_c \ c\phi_{i_0}-r^{k-1}\ _2 \leq \min_c \ c\phi_{i_0}-r^{k-1}\ _2$ para todo i
8. Defina $S^k=S^{k-1} \cup \{i_0\}$
9. Calcule $X^k=\text{argmin}_x \ X-Y\ _2$ sujeto a $\text{supp } X=S^k$
10. Calcule $r^k=Y-\phi X^k$
11. Hasta que $\ r^k\ _2 < \varepsilon$
Salida
Solución aproximada de X^k

Dado que la recuperación de la señal de entrada difiere en complejidad dependiendo de n , se analizan los dos casos para cada algoritmo de reconstrucción. Para el primer caso si $N = n$ no existe propiamente dicho una compresión debido a que la señal original N y el vector de medición n tienen la misma longitud, sin embargo ya que los datos de los átomos ϕ , o señales discretas de longitud n que representan formas de onda (Chen, y Donoho, 1999), constituyen un conjunto ortonormal, la señal puede ser recuperada como: $X_i = \langle Y, \phi_i \rangle$; $i=1,2,\dots,n$ donde la notación $\langle \cdot, \cdot \rangle$ indica el producto interno en el espacio $\mathbb{R}^n$. En este caso la representación del vector y sobre los átomos ϕ permite recuperar de forma unívoca la señal X . Para el segundo caso si $n < N$ existe compresión adicional, sin embargo los átomos ϕ no constituyen un conjunto ortonormal. Por tal motivo la recuperación de la señal X consiste en encontrar la señal $\bar{X}$ que minimice la energía de error entre y y la combinación lineal de átomos ϕ utilizando $\bar{X}$, es decir se debe resolver el problema de optimización descrito anteriormente. De lo anterior es evidente que el éxito de la compresión y la recuperación de la señal X depende del equilibrio existente entre los siguientes factores: a) La relación entre N/n la cual determina el grado de compresión de la información. b) la matriz aleatoria ϕ , c) el nivel de dispersión de la señal X .

RESULTADOS

En esta sección se presentan los resultados obtenidos de la evaluación del sistema propuesto el cual emplea tramas de DCT para pre-procesar las señales de audio con el fin de obtener una representación dispersa de las mismas en el dominio de la frecuencia, para la subsecuente aplicación del algoritmo de muestreo compresivo. Para llevar a cabo la evaluación del proceso de compresión y recuperación de las señales de audio, se emplearon señales que consisten en vocales naturales, así como una pieza musical, la suite #1 para Cello de Bach (Preludio), esta pieza musical cuenta con una duración de 15 segundos.

Para cada señal se realizan dos experimentos, el primer experimento consiste en aplicar el muestreo compresivo a distintas tramas de DCT tomando las mismas longitudes de esta para su muestreo, esto significa que $N=n$ o bien $DCT=CS$, lo que significa que la señal original N y el vector de medición n tienen la misma longitud, el segundo experimento consiste en aplicar el muestreo compresivo a tramas de DCT de 1000 muestras con diferentes longitudes de medición como son: 128, 256, 512, 700. Debido a que una forma de medir la convergencia de las señales es por medio de la energía del error de aproximación para todos los experimentos se utiliza la siguiente medida de similitud:

$$\text{Similitud} = e^2 / E^2 \quad (8)$$

Donde e^2 es la energía del error de aproximación dado por la norma $\|X - X_p\|$ dividida por la energía de la señal original, E^2 , dada por la norma de $\|X\|$. El rango para el cual está definida la similitud es $[0, 1]$ donde valores cercanos a 0 representa que las señales son casi idénticas y los valores cercanos a 1, representa que las señales son distintas.

Considerando nuestras señales de voz con longitudes entre 6055 y 12269 muestras, figura 2 (a), se aplica una ventana de Hamming para normalizar obteniendo una señal con poca distorsión en ambos lados como se muestra en la figura 2 (b). Seguidamente estas señales se subdividen cada una en tramas de 1000 muestras, con el fin de incrementar la precisión del algoritmo de muestreo compresivo. A cada trama se le aplica la DCT con lo cual obtenemos una representación dispersa de cada trama de la señal, figura 3 (a). Considerando las tramas de esta señal de voz como dispersas en frecuencia, podemos aplicar el algoritmo de muestreo compresivo a cada trama. El desempeño del muestreo compresivo es evaluado usando el algoritmo y software de reconstrucción $\text{Min-}l_1$, con restricciones iguales a las que se reportan en $l_1\text{-magic}$, (2010). Como resultado de aplicar el algoritmo de muestreo compresivo obtenemos una señal cercanamente igual a la señal de entrada dispersa, figura 3 (b), aplicando la transformada inversa de coseno (IDCT) se transforma la señal del dominio de la frecuencia al dominio del tiempo, recuperando así la señal de original figura 2 (a).

Fig. 2 Señales de Voz: a)Original b)Ventaneada

Fig. 3 Tramas de Señales de Voz en Frecuencia
a) Original, b) Recuperada

Las tablas 2 y 3 muestran la tabulación de los experimentos 1 y 2 antes mencionados, así como sus respectivas similitudes tras ser reconstruidas las cuales están establecidas por (8). Por su parte las figuras 4 y 5 muestran las gráficas obtenidas de las similitudes entre la señal original y la reconstruida. Como se muestra en la figura 4 podemos observar que a mayor número de observaciones realizadas la similitud entre la señal original y la recuperada disminuye, esta propiedad claramente depende de dos factores: la dispersión de la señal, y la naturaleza de la misma, encontrándose que para tonos agudos (vocal i y vocal u), es necesario tomar un número mayor de observaciones ó alcanzar una mayor dispersión para este tipo de señal, ya que como bien se ha mencionado anteriormente una señal con una mayor dispersión reducirá el número de observaciones que se deben alcanzar para efectuar una reconstrucción óptima. En la figura 5 podemos observar que a mayor número de observaciones realizadas, la similitud entre la señal original y la recuperada aumenta, como se mencionó antes esta propiedad depende de la dispersión de la señal.

Tabla 2. Similitudes en voz experimento 1

Nº de Muestras	DCT	A	E	I	O	U
128	128	4.37×10^{-12}	5.66×10^{-12}	1.75×10^{-11}	1.24×10^{-11}	1.03×10^{-11}
256	256	1.32×10^{-11}	9.84×10^{-12}	3.24×10^{-11}	4.07×10^{-11}	2.98×10^{-11}
512	512	3.31×10^{-11}	2.59×10^{-11}	5.13×10^{-11}	4.73×10^{-11}	8.37×10^{-11}
700	700	2.93×10^{-11}	3.41×10^{-11}	2.39×10^{-10}	7.53×10^{-11}	7.41×10^{-11}

Tabla 3. Similitudes en voz experimento 2

Nº de Muestras	DCT	A	E	I	O	U
128	1,000	0.6506	0.5296	0.1097	0.3611	0.2096
256	1,000	0.2735	0.2810	0.0513	0.0913	0.0337
512	1,000	0.03069	0.0520	0.0254	0.0220	0.0157
700	1,000	0.0152	0.0274	0.0152	0.0115	0.0101

Fig. 4 Similitud para experimento 1

Fig. 5 Similitud para experimento 2

Seguidamente considere la siguiente pieza musical suite #1 para Cello de Bach (Preludio) con una duración de 15 segundos y una longitud de 661000 muestras, figura 6 (a), las cuales se subdividen en tramas de 1000 muestras cada una con el fin de incrementar la precisión del algoritmo de muestreo compresivo. A cada trama se le aplica la DCT con lo cual obtenemos una representación dispersa de cada trama de la señal de música como se muestra en la figura 7(a), la cual, es considerada óptima para nuestros experimentos. Una vez que se cuenta con estas tramas, que se pueden considerar como dispersas, se aplica el algoritmo de muestreo compresivo a cada trama, El desempeño del muestreo compresivo es evaluado usando el algoritmo y software de reconstrucción $\text{Min-}l_1$, con restricciones iguales a las que se

reportan en I1-magic, (2010). Como resultado de aplicar el algoritmo de muestreo compresivo obtenemos una señal cercanamente igual a la señal de entrada dispersa, figura 7 (b), aplicando la transformada inversa de coseno (IDCT) se transforma la señal del dominio de la frecuencia al dominio del tiempo, recuperando así la señal de original figura 6 (a).

Fig. 6 Señales en el dominio del tiempo:
a) Original b) Recuperada

Fig. 7 Señales en el dominio de la frecuencia:
a) Original b) Recuperada

Las tablas 4 y 5 muestran la tabulación de los experimentos 1 y 2 antes mencionados, así como sus respectivas similitudes tras ser reconstruidas las cuales están establecidas por (8). Por su parte la figura 8 muestra las gráficas obtenidas de las similitudes entre la señal original y la reconstruida para ambos experimentos: $\diamond$ Experimento 1 $\square$ Experimento 2. Como se muestra en la figura 8 para el experimento 1 podemos observar que a mayor número de observaciones realizadas la similitud entre la señal original y la recuperada disminuye, esta propiedad claramente depende de dos factores: la dispersión de la señal, y la naturaleza de la misma, encontrándose que para componentes de la señal con tonos agudos es necesario tomar un número mayor de observaciones ó alcanzar una mayor dispersión, ya que como bien se ha mencionado anteriormente una señal con una mayor dispersión reducirá el número de observaciones que se deben alcanzar para efectuar una reconstrucción óptima. Dentro de la misma figura 8, para el experimento 2 podemos observar que a medida que aumenta el número de observaciones realizadas, la similitud entre la señal original y la recuperada aumenta, como se mencionó antes, esta propiedad depende de la dispersión de la señal.

Tabla 4. Similitud en señales de música experimento 1

No. de Muestras	DCT	Bach
128	128	1.21×10^{-10}
256	256	4.65×10^{-11}
512	512	1.21×10^{-10}
700	700	1.39×10^{-10}

Tabla 5. Similitud en señales de música experimento 2

No. de Muestras	DCT	Bach
128	1,000	0.1208
256	1,000	0.0234
512	1,000	0.0014
700	1,000	3.12×10^{-4}

Considerando la similitud por trama figura 9, para cada una de ellas, al presentarse una mayor cantidad de tonos agudos como se puede observar en el rango de 300 a 500 la similitud disminuye, aumentando la diferencia entre la señal original y la recuperada, El lector podrá observar que cuando se utilizan las mismas longitudes de la DCT para muestrear, en ambos casos los valores son menores a los que previamente se reportaron en (Moreno-Alvarado, Martínez-García, 2011). lo que da pauta para establecer una mayor compresión en las señales de audio dependiendo de la aplicación.

Fig. 8 Similitud para experimentos
 ◇ Experimento 1 □ Experimento 2

Fig. 9 Similitudes por trama

CONCLUSIONES

En este artículo se propone una eficiente implementación conjunta de la DCT, como método para obtener una representación dispersa de una señal de audio, y del algoritmo de muestreo compresivo, el cual se evaluó con voz natural así como con una señal de música, la suite #1 para Cello de Bach (Preludio). Resultados obtenidos por simulación muestran que la representación dispersa aplicada a nuestras señales de audio a través de la DCT, tiene la habilidad de empaquetar la energía de la señal de entrada en pocos coeficientes (Rao y Yip, 1990) debido a que la mayor parte de la información se concentra en las frecuencias bajas. Esto permite al cuantizador descartar aquellos coeficientes con amplitudes relativamente pequeñas sin introducir distorsión apreciable en la señal reconstruida. Con esto se obtienen resultados razonablemente buenos para la implementación del muestreo compresivo en señales de audio pre-procesando esta señal por medio de la DCT. Lo anterior significa que la técnica propuesta puede reducir significativamente el número de muestras requeridas para representar ciertas señales de audio y así reducir el número de bits necesarios para la codificación.

Se encontró que el modelo de compresión de audio propuesto es factible y puede alcanzar una compresión para una señal de música en algunos casos cercana al 50% con una buena calidad dependiendo de la aplicación a la que se destine el sistema en particular. Los valores de compresión obtenidos varían dependiendo de qué tan dispersa sea la señal y de los tonos agudos contenidos en la señal, sin embargo para obtener un nivel de calidad razonable se deben de tomar alrededor de $[DCT/2+100]$ muestras.

AGRADECIMIENTOS

Las autores agradecen al Consejo Nacional de Ciencia y Tecnología (CONACYT), al Instituto Politécnico Nacional y a la Universidad La Salle, instituciones mexicanas que apoyaron la realización de esta investigación.

REFERENCIAS

- Candes E., Romberg J. y Tao T., *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52(2) 489–509, (2006).
- Candes E. and J. Romberg J., *Robust signal recovery from incomplete observations*, Proc. of ICIP, 1281–1284, (2006).
- Candes E. and Romberg J., *Sparsity and incoherence in compressive sampling*, Inverse Problems, 23(6), 969–985, (2007).
- Candes E y Wakin M., *An introduction to compressive sampling*, IEEE Signal Processing Magazine, 25(2) 21–47, (2008).
- Candes, E.J., *The restricted isometry property and its implications for compressed sensing*. Comptes Rendus de l'Academie des Sciences, Serie I, 346, 589–592 (2008).

- Chen, S.S., Donoho, D.L., Saunders, M.A.: Atomic decomposition by Basis Pursuit. SIAM J. Sci. Comput. 20(1), 33-61 (1999).
- Davenport M., Duarte M., Eldar Y., Kutyniok, *Introduction to compressive sensing*, Compressive sensing : Theory and *Applications*, Editors Eldar Y. y Kutyniok G, Cambridge University Press, New York, USA, (2012).
- Donoho D., *Compressed sensing*, IEEE Trans. Inform. Theory, 52(7) 1289–1306, (2006).
- Donoho D.L., y Huo X., Uncertainty principles and ideal atomic decomposition, IEEE Trans. Inform. Theory, vol. 47, no. 7, pp. 2845–2862, Nov. (2001).
- Eldar Y. y Kutyniok G., *Compressive sensing : Theory and Applications*, Cambridge University Press, New York, USA, (2012).
- Foucart, S., Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants. In: Proceedings of the 13th International Conference on Approximation Theory (2010).
- Giacobello D., Christensen M., Dahl J., Jensen S., and Moonen M., *Sparse linear predictors for speech processing*, Proc. Interspeech (2008).
- Gonzalez R. y Wintz P., *Digital Image Processing*, Prentice Hall Englewood Cliffs, NJ, (2008).
- Haiyan Guo., Tianjing Wang., Zhen Yang., Adaptive speech compressed sensing in the DCT domain, Chinese Journal of Scientific Instrument, vol.31, no.6, pp.1262-1268, June (2010).
- l1-magic. [Online]. Available: <http://www.l1-magic.org>, October (2010).
- Khayam A., *The Discrete Cosine Transform (DCT):Theory and Application*, Department of Electrical and Computer Engineering Michigan State University, (2003).
- Kassim Liban A., Khalifa Othman O., Gunawan T. S. Evaluation of sparsifying algorithms for speech signals, International Conference on Computer and Communication Engineering (ICCCE 2012), Kuala Lumpur, pp.308-313, July (2012).
- Lustig M, Donoho D., Santos J. y J. Pauly J., *Compressed sensing MRI*, IEEE Signal Processing Magazine, 25(2), 72-82, (2008).
- Moreno-Alvarado R. G., Martínez-García Mauricio., CONIELECOMP, page 55-59. IEEE, March(2011)
- Moreno-Alvarado R. G., Martínez-García Mauricio., Proceedings of the World Congress on Engineering 2011 Vol II WCE 2011, London, U.K., July 6 - 8, (2011)
- Rao K.R., Yip R.. Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic Press, Inc., San Diego, CA, (1990).
- Sreenivas T. y Kleijn W., *Compressive sensing for sparsely excited speech signals*, Proc. IEEE Int. Conf. Acoustic, Speech and Signal Processing, 4125–4128, (2009).