

# Fusión de imágenes con múltiples puntos de enfoque basado en sensado compresivo

Erick Quezada | Roger Arribasplata

Estudiantes, Universidad Tecnológica de Panamá  
{erick.quezada, roger.arribasplata}@utp.ac.pa

Fernando Merchan

Profesor, Universidad Tecnológica de Panamá  
fernando.merchan@utp.ac.pa

**Tipo de artículo:** original

**Fecha de recepción:** 9 de julio de 2013

**Fecha de aceptación:** 12 de noviembre de 2013

**Resumen:** *en este artículo se presentan resultados en la aplicación del principio del sensado compresivo (SC) al problema de fusión de imágenes con múltiples puntos de enfoque. Se presenta las bases teóricas e implementación de un algoritmo de fusión basado en SC. Adicionalmente, se presenta un estudio comparativo con otros algoritmos de fusión de imágenes que pone en evidencia las propiedades y el potencial del método presentado.*

**Palabras claves:** *procesamiento de imágenes, sensado compresivo, fusión de imágenes.*

**Title:** *Multi-focus image fusion based on compressive sensing.*

**Abstract:** *in this paper we present the application of the principle of compressive sensing in the fusion of multi-focus images. We present the theoretical principles and the implementation of an image fusion algorithm. A comparative study with other image fusion algorithm is presented to point out the properties and advantages of the proposed method.*

**Keywords:** *image processing, compressive sensing, image fusion.*

## 1. Introducción y motivación

El teorema de Shannon y Nyquist nos enseña que la tasa de muestreo necesaria para la reconstrucción exacta de una señal, debe ser superior al doble de su ancho de banda. En muchas aplicaciones de imágenes digitales y video-cámaras, la tasa de Nyquist es tan alta que el número de muestras resultantes causa que la compresión de estas señales sea una necesidad previa al almacenamiento o transmisión; y en otras aplicaciones, tener una alta tasa de muestreo resulta ser muy costosa. Alrededor del año 2006, emergió una teoría alternativa a este teorema propuesta por David L. Donoho y Emmanuel J. Candès, llamada "Sensado Compresivo" [1,2]. El principio del Sensado Compresivo (SC) propone que una señal

o una imagen puede ser reconstruida por un número de medidas/muestras mucho menor de lo que señala el teorema de Shannon y Nyquist. Nosotros basamos nuestro estudio en el procesamiento digital de imágenes, específicamente en técnicas para la fusión de imágenes de múltiples puntos de enfoque utilizando SC [3]. Fusión de imágenes es un proceso que combina información de múltiples imágenes en una imagen fusionada que provee mejor capacidad de interpretación. La imagen fusionada resulta ser mucho más útil para consecuentes aplicaciones de procesamiento de imágenes. Hoy en día muchos algoritmos de fusión han sido propuestos [4]. Un trabajo reciente [3, 5] demuestra la posibilidad de fusionar imágenes con una cantidad menor de muestras de las imágenes originales, si ellas son adquiridas utilizando el principio de SC. Una ventaja clave de este enfoque es que las muestras necesarias para el procesamiento pueden ser recolectadas sin asumir información previa de la imagen que se está observando. Además, SC requiere menos espacio de almacenamiento, a pesar del mayor costo en proceso computacional. Por lo tanto, esta técnica de SC ha sido muy atractiva para aplicaciones de imágenes digitales.

Nuestro objetivo en esta documentación es presentar nuestros primeros estudios de los conceptos de fusión de imágenes con múltiples puntos de enfoque usando sensado compresivo, el cual esperamos que sea útil en todas las técnicas de video-procesamiento y fotografía en general, en áreas de video-vigilancia, procesamiento de imágenes médicas e imágenes satelitales, adquisición de imágenes en cámaras digitales y videocámaras. Esperamos que con esta publicación podamos informar e inspirar a más personas para contribuir con los estudios de procesamiento digital de imágenes y sensado compresivo. A continuación, presentamos una breve reseña de la teoría, conceptos y términos mencionados y utilizados para el desarrollo de nuestros algoritmos de procesamiento. Luego procederemos a presentar aspectos teóricos de SC; presentaremos un método convencional de fusión de imágenes basadas en wavelet, y un algoritmo de fusión basado en SC. Por último comparamos el método de fusión basado en SC con el método convencional de wavelets y con el método de promedio.

## 2. Teoría del sensado compresivo

SC es un paradigma de muestreo que nos permite capturar y representar señales a una tasa de muestreo significativamente menor a la tasa de muestreo de Nyquist tomando ventaja de la estructura esparsa de la señal. Convencionalmente una señal  $x$  muestreada con un número de elementos  $N$  atraviesa por un proceso de compresión en donde la señal resultante tiene un número de elementos menor  $M$ ; es decir  $N \gg M$ . El SC nos enseña que es posible muestrear y comprimir datos simultáneamente, es decir adquiriendo un número de muestras  $K$  donde  $N \gg K > M$ .

Una señal  $x$ , considerada como un vector en un sub-espacio de dimensiones finitas de  $R^N$ ,  $x = [x[1] \dots x[N]]$ , es estrictamente o exactamente esparsa si la mayoría de sus elementos son iguales a cero; es decir, si se comprueba que  $\Delta(x) = \{1 \leq i \leq N \mid x[i] \neq 0\}$  es de cardinalidad  $k \ll N$ . Una señal  $k$ -esparsa es aquella para la cual exactamente  $k$  número de elementos tienen un valor distinto de cero.

La experiencia indica que las señales e imágenes naturales son esparsas en algún dominio.

El principio del SC establece que una señal discreta con una representación esparsa en ciertas bases, puede ser recuperada a partir de un número pequeño de proyecciones lineales de dicha señal sobre una base arbitraria (Fourier, wavelets, Gabor, etc); en otras palabras, cuando se habla de una señal esparsa se hace referencia a aquellas que sólo presentan pocos valores distintos de cero en algún dominio de transformada.

### Planteamiento del problema de SC en imágenes

Primero, suponga que representamos una imagen en tonos de grises como una matriz de  $N \times N$  de elementos en el rango de 0 a 255, donde 0 es el color negro, 255 es blanco, y los valores intermedios son los tonos de grises. Esa matriz se puede expresar como un vector  $x \in R^n$  de dimensiones  $n \times 1$ , donde  $n = N \times N$ . Consideremos que este vector es  $k$ -esparso en una base representada por la matriz  $\Psi = [\psi_1 | \psi_2 | \dots | \psi_N]$  de dimensiones  $n \times n$  con los vectores  $\{\psi_i\}$  como columnas. Esto significa que tiene  $k$  número de coeficientes diferentes de cero en esta base. Esto es algo que se observa en las transformadas de imágenes naturales, por ejemplo, la transformada de wavelets. Ahora podemos expresar  $x$  como:

$$x = \Psi\theta \quad (1)$$

donde  $\theta$ , de dimensiones  $n \times 1$ , representa un vector de coeficientes de transformada con solamente un número  $k$  de elementos diferentes de cero, donde  $k$  es menor que  $n$ . Por lo tanto,  $x$  se le puede considerar  $k$ -esparso en la base ortonormal  $\Psi$ . Estudios recientes han demostrado que una señal puede ser reconstruida fielmente con este número relativamente pequeño de muestras [1, 2]. Este muestreo puede ser expresado en el vector de medición  $y \in R^m$ , con  $k < m < n$ , que es obtenido por el siguiente sistema lineal:

$$y = \Phi x = \Phi\Psi\theta \quad (2)$$

donde  $\Phi$  es una matriz de medición de tamaño  $m \times n$ . Contrario a métodos tradicionales de muestreo, las medidas contenidas en  $y$  son suficientes para recuperar  $x$ . La condición para esto es que se verifique una condición de incoherencia entre la base de medición  $\Phi$  y la base de representación  $\Psi$ . Esto significa que las filas de  $\Phi$  no puedan representar de manera esparsa las columnas de  $\Psi$  [1,2].

La fusión de imágenes se logra cuando comparamos y fusionamos los vectores de medidas  $y$  de múltiples imágenes con diferentes puntos de enfoque, para obtener un vector de medición  $y$  resultante, con el cual recuperamos la imagen final fusionada.

Es importante recordar que como  $m \ll n$ , la recuperación de la señal  $x$  a partir de  $y$  resulta imposible por una simple y directa transformada inversa de la Ec. (2). En los trabajos ya mencionados [1, 2] se desarrollan algoritmos de recuperación no lineal que resuelve este problema. Lastimosamente, estos algoritmos pueden exigir muchos recursos computacionales en el proceso.

Con la información de  $m$  (donde  $m \ll n$ ) número de medidas,

nosotros queremos recuperar la señal  $x$ . Esto se puede lograr resolviendo:

$$\hat{\theta} = \operatorname{argmin} \|\theta\|_0 \text{ sujeto a } \Phi\Psi\theta = y \quad (3)$$

La señal  $\hat{x}$  es reconstruida como  $\Psi\hat{\theta}$  donde  $\hat{\theta}$  es el mínimo global del problema de optimización  $l_0$ . Sin embargo, éste es un problema no-convexo altamente combinatorial [7]. Se ha demostrado que es posible obtener resultados similares sustituyendo norma  $l_0$  por la norma  $l_1$ :

$$\hat{\theta} = \operatorname{argmin} \|\theta\|_1 \text{ sujeto a } \Phi\Psi\theta = y \quad (4)$$

Se ha demostrado que cuando  $x$  es lo suficientemente esparsa en  $\Phi$ , la recuperación a través de minimización  $l_1$  es probablemente exacta.

### 3. Fusión de imágenes en el dominio de multiresolución

Cuando estudiamos el tema de fusión de imágenes, observamos que la descomposición en multiresolución ha presentado ventajas significativas en la representación de las señales.

Un esquema de multiresolución descompone la señal o imagen que está siendo analizada en varios componentes, cada uno de los cuales captura la información presente en una escala dada [8]. La noción de resolución o escala está relacionada con la dimensión de los detalles que pueden representarse. Estos conceptos son muy útiles en procesamiento de imágenes por las siguientes razones: 1) los objetos en el mundo real están formados por estructuras a diferentes escalas; 2) el sistema visual humano, procesa la información en multiresolución; 3) reduce la complejidad de los algoritmos. La idea fundamental de estos sistemas consiste en obtener una representación más conveniente (*análisis*) de la señal original sin pérdida de información, de forma que pueda posteriormente reconstruirse (*síntesis*).

La Transformada Discreta de Wavelets (TDW), permite la descomposición de la imagen en diferentes escalas (multiresolución) de coeficientes a la vez que se preserva la información. Podemos observar esquemas de análisis y síntesis en wavelets en la figura 1.

En el contexto de la fusión de imágenes, los coeficientes de wavelets, tanto de aproximación como de detalle, para una escala dada procedente de distintas imágenes pueden combinarse para la obtención de nuevos coeficientes.

Por su diferente significado físico los coeficientes de detalle y aproximación suelen combinarse de manera diferente [8]. Resulta que en las bandas de detalle, la magnitud de los valores de los coeficientes (positivos o negativos) proporciona suficiente información para identificar características relevantes en las imágenes tales como: bordes, líneas o límites de regiones. Es por eso que elegimos los valores máximos de los coeficientes de detalles. Esto se expresa con la regla siguiente:

$$D_F = D_M \text{ con } D_M = \arg \max_{i=1, \dots, I} (|D_i|) \quad (5)$$

donde  $D_F$  son los coeficientes de la imagen fusionada,  $D_M$  es el coeficiente de máximo valor absoluto de las imágenes iniciales o

de entrada, y  $l$  es el total de imágenes iniciales. Los valores de los coeficientes de aproximación son promediados.

Una vez realizada la mezcla de coeficientes se procede a aplicar la Transformada Discreta Inversa de Wavelets (TDIW) para obtener finalmente la imagen fusionada.

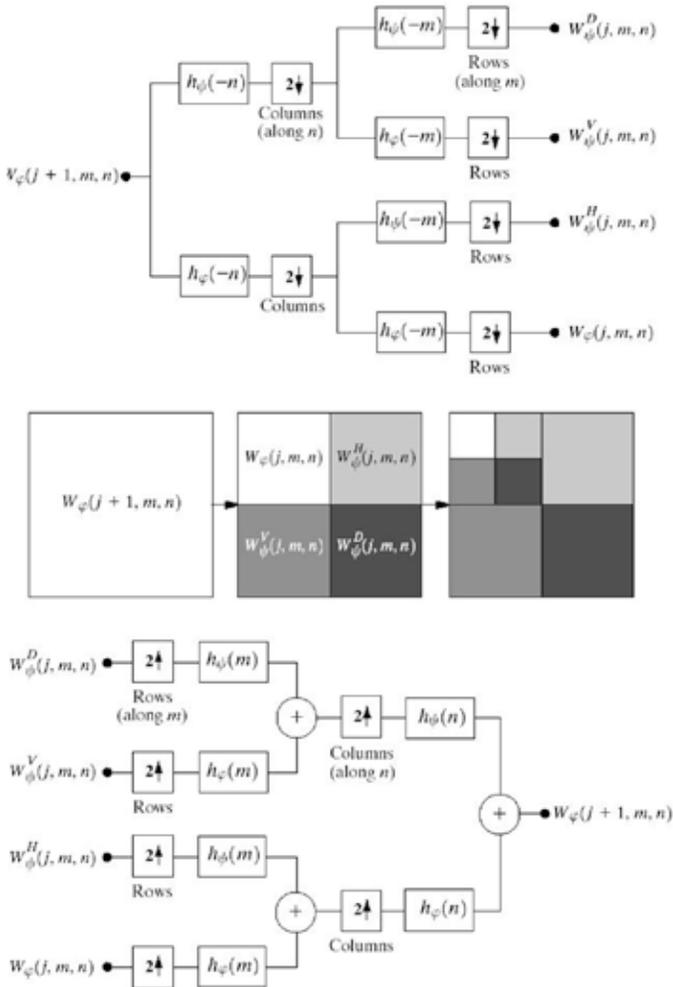


Figura 1. Esquema de descomposición en wavelets: análisis (esquema superior) y síntesis (esquema inferior).

#### 4. Fusión de imágenes en el dominio compresivo

En esta sección, formulamos un algoritmo de fusión de imágenes que usa mediciones compresivas. En éste se aplica un esquema de fusión similar a la usada en el dominio de wavelets en el dominio compresivo. La diferencia es que la fusión es realizada con medidas compresivas, en vez de coeficientes de wavelets. Los pasos básicos del algoritmo se describen a continuación:

#### Algoritmo Fusión de imágenes basadas en SC

- 1) Tomar las medidas compresivas  $Y_i, i = 1, \dots, l$  para la imagen de entrada  $i$  usando el patrón de muestreo con la forma de estrella.
- 2) Calcular las muestras fusionadas utilizando la fórmula:  

$$Y_F = Y_M \text{ con } M = \arg, \text{Max}_{i=1, \dots, l} (|Y_i|).$$
- 3) Reconstruir la imagen fusionada de las medidas  $Y_F$  a través del método de variación total.

En la sección 2 vimos que el vector de medición  $y$  es obtenido a partir de proyecciones lineales no adaptativas de la señal  $x$  pasada por un matriz de medición  $\Phi$ . Wan, Achim y colaboradores (2008) propusieron usar como mediciones algunos coeficientes de la transformada de Fourier en dos dimensiones [5]. Estos coeficientes corresponden por un patrón de muestreo de estrella como se muestra en la figura 2.

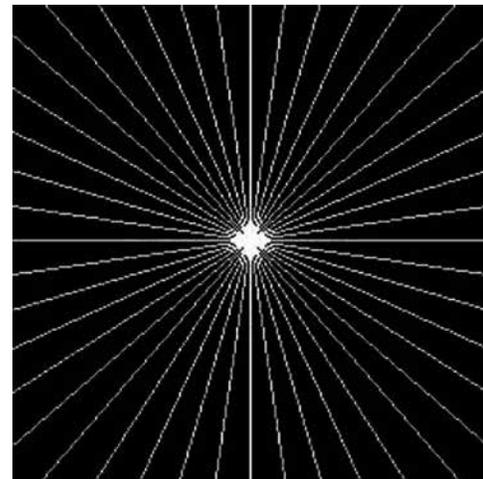


Figura 2. Patrón de Muestreo Estrella.

Este patrón consiste en líneas blancas que indican la localización de las frecuencias usadas para calcular las medidas compresivas  $y$ . Las frecuencias bajas están concentradas en el origen de sistemas coordenadas (centro de la imagen), y las altas frecuencias se encuentran esparcidas lejos del origen. Una vez obtenida  $y$ , se utiliza un algoritmo de reconstrucción para recuperar la señal original  $x$  [6]. Para lograr esto se utilizó un modelo de recuperación alternativo adaptado a las señales en 2D y en el que se busca obtener la imagen cuyo gradiente sea esparso.

Consideremos que  $x_{ij}$  es el pixel en la  $i$ -ésima fila y la  $j$ -ésima columna de una imagen  $x$ , y definamos los operadores  $D_{(h,ij)} x$  y  $D_{(v,ij)} x$  como sigue:

$$D_{(h;ij)} x = \begin{cases} (x_{(i+1,j)} - x_{(i,j)}) & i < n \\ 0 & i = n \end{cases}$$

$$D_{(v;ij)} x = \begin{cases} (x_{(i+1,j)} - x_{(i,j)}) & j < n \\ 0 & j = n \end{cases} \quad (5)$$

Estos operadores pueden interpretarse como los gradientes discretos verticales y horizontales de una imagen. A partir de estos operadores podemos definir la variación total como la suma de las magnitudes de este gradiente discreto en cada punto:

$$TV(x) := \sum_{ij} \sqrt{(D_{(h;ij)} x)^2 + (D_{(v;ij)} x)^2} \quad (6)$$

Así el algoritmo de recuperación busca la solución que verifique:

$$\min TV(x) \text{ sujeto a } y = \Phi x \quad (7)$$

donde  $y$  son las medidas compresivas obtenidas de la imagen  $x$  definidas por la matriz de medición  $\Phi$ . En nuestra implementación, la matriz  $\Phi$  representa una base parcial de Fourier en 2-D definida por el patrón mostrado en la figura 2 y el vector  $y$  los coeficientes de Fourier correspondientes.

### 5. Resultados y perspectivas

Para ilustrar nuestro estudio proponemos dos colecciones de imágenes de prueba para comparar: las imágenes de relojes estándar, y las imágenes de libros tomadas en nuestro laboratorio. En cada imagen de los relojes, uno de los relojes está en foco, mientras que el otro no. El mismo efecto se observa en las imágenes de libros.

Igualmente hacemos una comparación entre los dos algoritmos mencionados previamente, junto con el método de promedio, para comparar resultados entre un método y otro.

Las imágenes originales tienen una resolución de 512 x 512 píxeles; es decir, 262144 píxeles o elementos en los vectores de imágenes iniciales. Usando el algoritmo con *wavelets*, necesitamos analizar este número de coeficientes para el proceso de fusión. En cambio, utilizando el método de SC el número total de muestras en nuestro vector y de medición es de 91785 mediciones. En otras palabras, este proceso de fusión, trabaja con vectores de medición con tamaño 65% menor que aquellos vectores de coeficientes de *wavelets*.

En la figura 3 mostramos los resultados finales. Primero mostramos los dos pares de imágenes de iniciales (figura 4.a-b). Agregamos un resultado extra para comparación: el método promedio (figura 4.c) que consiste en promediar las dos imágenes. Este método es uno de los menos eficientes y sólo lo usamos como ejemplo de control. En los dos últimos resultados, observamos aquellos que fueron obtenido a través de la descomposición en *wavelets* y usando SC, respectivamente. El método de *wavelets* proporciona un resultado bastante satisfactorio. Sin embargo, los resultados obtenidos por SC presentan mayor contraste. Igualmente presentan mayores detalles en los bordes evitando efectos de difusión que podemos observar en el resultado por *wavelets*. Sin embargo, el resultado por SC presenta algunos artefactos y ruido granular.

Como trabajo futuro se considerará un estudio más extenso con diferentes tipos de imágenes para determinar el potencial del método compresivo en la aplicación de fusión de imágenes digitales. Se estudiarán nuevas reglas y dominios para realizar la fusión, y esperamos desarrollar un algoritmo que presente óptimos resultados que aprovechen de las prometedoras propiedades que brinda sentido compresivo.

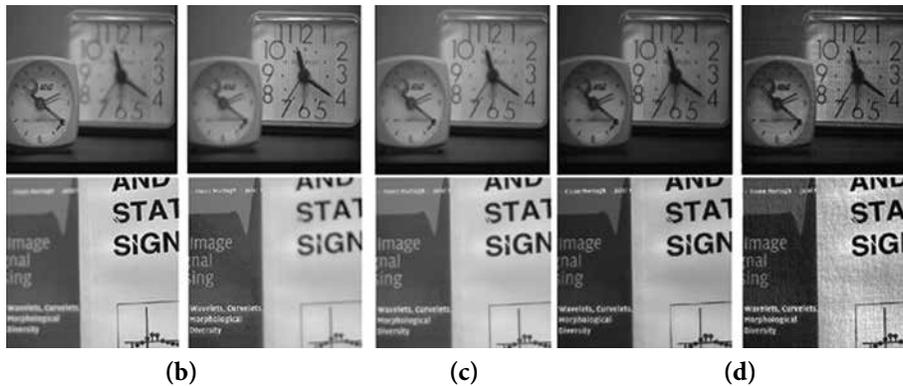


Figura 3. Resultados de fusión de imágenes. (a-b) Imágenes originales. (c) Método por promedio, (d) *wavelets* y (e) SC.

### Referencias

- [1] E. Candés, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 56, No. 2, pp. 489-509, 2006.
- [2] David L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, No. 4, pp. 1289-1306, Apr. 2006.
- [3] T. Wan, N. Canagarajah, A. Achim, "Compressive Image Fusion," *Proc. IEEE Int. Conf. Image Process*, pp. 1308-1311, 2008.
- [4] M. I. Smith, J.P. Heather, "Review of image fusion technology in 2005," *Proceedings on Defense and Security Symposium*, Orlando, FL, March 28–April 1, 2005.
- [5] T. Wan, N. Canagarajah, A. Achim, "Compressive Image Fusion," in *Proc. IEEE Int. Conf. Image Process*, pp. 1308-1311, 2008.
- [6] E. Candés and J. Romberg, **l1-magic: Recovery of sparse signal via convex programming**, code package available at <http://www.l1-magic.org>
- [7] J.-L. Starck, F. Murtagh & J. M. Fadili, *Sparse Image and Signal Processing. Wavelets, Curvelets, Morphological Diversity*, Cambridge Univ. Press, 2010.
- [8] G. Pajares & J. Cruz, *Visión por computador*, 2ed., Alfaomega Ed., México 2008.