Revisión de métodos de representación geométrica de imágenes basados en *bandlets*

César Martínez
Informe del curso: Análisis y Procesamiento Avanzado de Señales
Doctorado en Ingeniería - FICH, UNL.
31 de Octubre de 2007

I. Introducción

El artículo [1] presenta una revisión de técnicas para la representación de las imágenes más complejas de analizar hoy en día: los paisajes, tomas de exterior, fotografías de interior en ambientes no controlados, con combinación de texturas, etc., denominadas en general *imágenes naturales*. Al contrario de las imágenes controladas, o aquéllas generadas mediante síntesis no fotorealística, las imágenes naturales poseen bordes irregulares y/o difusos, altas frecuencias (no necesariamente ruido), y otras características que exigen al máximo a los algoritmos de procesamiento actuales. En estas condiciones, para el avance de las técnicas se hace necesario sacar partido de las estructuras geométricas presentes, ya que son las que definen las zonas de cambio que proveen pistas para la percepción humana.

Una primera aproximación a la geometría de una imagen está dada por el conjunto de curvas sobre las cuales la imagen es singular. En imágenes naturales, esta geometría no es regular (mal condicionada) por lo que el análisis armónico es la herramienta que aporta las técnicas más adecuadas para la representación, y a ellas se dedica la revisión del artículo.

Dada una base $\mathcal{B}$ ortogonal y una función f, se define su representación f_M con los M coeficientes de mayor amplitud y se busca la aproximación *óptima* que maximice el decaimiento del error al incrementar M de forma $||f - f_M||^2 = O(M^{-\beta})$ para el β más alto posible.

En 1D, la transformada wavelet (WT) descompone una función f en una base ortogonal que logra una aproximación $||f - f_M||^2 = O(M^{-2\alpha})$, siendo f una función C^α sin singularidades o con un número finito de éstas (regular por tramos). La extensión a 2D de la WT es, sin embargo, subóptima ya que si la función f tiene discontinuidades a lo largo de un borde, siendo C^α fuera de las curvas, entonces solamente se logra $||f-f_M||^2 = O(M^{-1})$. Así, al contrario del caso 1D, la presencia de singularidades ralentizan el decaimiento del error.

En cuanto a las aproximaciones *no-wavelets*, una serie de técnicas fueron propuestas tratando de capturar la geometría según una base de funciones elongadas y orientadas según porciones de curvas. Las *curvelets* logran aproximaciones óptimas para funciones C^2 por tramos, pero no existe todavía una base ortogonal, lo que les juega en contra respecto a la WT para la compresión de imágenes naturales [2]. Las *wedgelets*, por su parte, adapta la representación a la geometría de cada imagen, pero es sólo aplicable en bodes simples y bien definidos [3], con algunas mejoras propuestas en [4], [5].

Los coeficientes de la WT muestran altos valores justamente a lo largo de todos los bordes, introduciendo una regularidad que disminuye la capacidad de compresión de la representación. La *bandletización*[1] consiste en un algoritmo que remueve la redundancia geométrica mediante una familia de operadores que adaptan los coeficientes a la geometría de la imagen, a fin de capturar las singularidades de los contornos. Específicamente, luego de realizar la descomposición wavelet de la imagen, se calculan los coeficientes bandlet sobre los detalles que contengan un segmento de borde, como los productos internos $\langle f, b_l \rangle$ donde las funciones bandlet b_l son combinaciones lineales de las wavelets: $b_l(x) = \sum_p a_l[p]\psi_p(x)$. Aquí, los $a_l[p]$ son los coeficientes de la transformada Alpert [6], los cuales reorientan los coeficientes wavelet de acuerdo al flujo geométrico.

Para funciones C^α la base bandlet ortogonal obtiene una aproximación $||f - f_M||^2 = O(M^{-\alpha})$, siendo $f_M = \sum_{|\langle f, b_\nu \rangle| > T} \langle f, b_\nu \rangle b_\nu$, con buenos resultados en compresión, denoising y restauración.

Sin embargo, ya que la bandletización evalúa el flujo geométrico localmente sobre los bordes que aparecen en los coeficientes wavelet, tiene una geometría restringida que la hace ineficiente para manejar regularidades de largo alcance. Esta cuestión es afrontada mediante un campo de asociación multiescala que aplica una transformada *grouplet* sobre los coeficientes wavelet [7]. Esta transformación ortogonal calcula promedios de coeficientes pares y variaciones de coeficientes impares en diversas escalas, logrando resultados del estado del arte en aplicaciones de denoising, superresolución, inpainting y otras.

II. Análisis general

A. *Estructura general*

El resumen es conciso (100 palabras), pero suficiente para introducir la idea de la bandletización como aproximación óptima para representación geométrica de imágenes. Asimismo, introduce también el concepto de agrupamiento

[1] Libertad de traducción del término inglés *bandletization*.

de coeficientes bandlet en los denominados gruplets, en la búsqueda de una representación más flexible y global sobre la imagen.

Respecto a los contenidos, se realiza una adecuada presentación del problema actual en algoritmos de procesamiento de imágenes: el tratamiento de la geometría –los bordes–, importante para la mejora en el desempeño de estos procesos. Luego plantea correctamente una medida de eficiencia de las representaciones de imágenes en alguna base de funciones, con la cual en el resto del artículo comparan las diferentes alternativas propuestas en la literatura.

En la presentación de todos los métodos de representación se resumen: el decaimiento en el error de aproximación, las restricciones en la continuidad paramétrica de las funciones y si se disponen o no de bases ortogonales y algoritmos eficientes de cálculo.

Luego de exponer la bandletización, el artículo destaca las limitaciones de este método en el análisis global de la imagen y dedica la última sección a la presentación de los grouplets, técnica de post-procesamiento de los coeficientes wavelets que intenta encontrar regularidades geométricas en toda la imagen.

Atendiendo al tipo de investigación presentada, es de esperar que el artículo sea descriptivo de las técnicas revisadas. Desde el punto de vista del contenido expuesto de cada técnicas, el artículo es correcto en líneas generales. No obstante, resulta a veces muy somero en las descripciones presentadas, llegando a haber resumido en sólo un párrafo los aportes de 3 trabajos diferentes de otros autores.

Un último comentario crítico se dedica a la presentación de los grouplets en el artículo. Si bien la técnica mencionada –de acuerdo a lo expuesto– es una mejora respecto a los bandlet, el objetivo general del artículo es la revisión del estado del arte en la representación geométrica de imágenes. En este contexto se considera inadecuada la inclusión de esta técnica ya que constituye una propuesta aún no sometida a referato por pares (estado de *preprint*).

Finalmente, es de notar que el artículo no contiene una sección de conclusiones, dado el tipo de investigación presentada.

B. *Presentación del estado del arte*

El artículo no contiene una sección introductoria, y dentro de ésta el estado del arte en el tema de manera resumida, como es usual en otros trabajos. El aporte mismo del artículo es la revisión del estado del arte, por lo que a lo largo del trabajo se van presentando las características de las principales aproximaciones.

En este contexto, es correcta la estructuración y secuencia de la presentación del problema de representación geométrica de imágenes, la aproximación wavelet 2D y sus inconvenientes, las técnicas alternativas a la wavelet, la técnica bandlet y su extensión grouplet.

Las críticas a los trabajos revisados son correctas desde el punto de vista metodológico, ya que utiliza el mismo conjunto de parámetros para evaluar los métodos: la bondad de la aproximación (error y continuidad paramétrica), disponibilidad de bases ortogonales, eficiencia de algoritmos y resultados visuales.

En la revisión presentada, sin embargo, se extrañan referencias a otras aproximaciones en bases de funciones que comparten el mismo objetivo que las bandlet, las cuales se mencionan en el apartado siguiente.

C. *Bibliografía*

Las principales referencias de cada aproximación son actuales (en general, 2002 en adelante), coincidente con el interés de la comunidad científica en el problema tratado, publicadas en revistas internacionales indexadas. Existen tres referencias a artículos en estado de *preprint* y si bien dos de ellas son complementarias a técnicas con referencias en revistas, la única referencia de la técnica grouplet se da a un artículo en *preprint*.

Existen propuestas recientes en la bibliografía de técnicas que persiguen el mismo objetivo del artículo revisado: la representación geométrica de imágenes por un método más eficiente que la extensión directa a 2D de la transformada wavelet, y que no se encuentran analizadas en el mismo. Entre estas omisiones se pueden mencionar: la transformada *ridgelet* (Do and Vetterli, The finite ridgelet transform for image representation. IEEE Trans. on Image Processing, 2003); la transformada *contourlet* (Do and Vetterli, The contourlet transform: an efficient directional multiresolution image representation. IEEE Trans. on Image Processing, 2005) y la transformada discreta *shearlet* (Easley *et. al.*, Optimally Sparse Image Representations using Shearlets. IEEE ACSSC,2006).

III. DESARROLLO

En la Sección I se introdujo el problema de la representación de la geometría de una imagen en una base wavelet eficiente, en el sentido de que logre una densidad baja de coeficientes en las zonas de borde. Recordemos la expresión de la aproximación de una función f en una base ortonormal $\mathcal{B}$:

$$f_M = \sum_{|\langle f, g_\mu \rangle| > T} \langle f, g_\mu \rangle g_\mu, \tag{1}$$

donde $M \triangleq \#\{\mu \backslash |\langle f, g_\mu \rangle| > T\}$. El error de aproximación está dado por:

$$\|f - f_M\|^2 = \sum_{|\langle f, g_\mu \rangle| \leq T} |\langle f, g_\mu \rangle|^2 \tag{2}$$

En estas aproximaciones se busca encontrar una base que logre hacer decaer el error $\|f - f_M\|^2$ asintóticamente a cero lo más rápido posible a medida que se incrementa M. Así, la base óptima será aquella que logre $\|f - f_M\|^2 = O(M^{-\beta})$ para el β más alto posible.

En 2D, una base wavelet B de $\mathbf{L}^2([0,1]^2)$ se obtiene mediante traslaciones y dilataciones de las wavelets elementales ψ^H, ψ^V, ψ^D en direcciones horizontal, vertical y diagonal respectivamente como:

$$B = \left\{ \psi^k_{jn}(x) = 2^{-j} \psi^k (2^{-j} x_1 - n_1, 2^{-j} x_2 - n_2) \right\}, \tag{3}$$

con: $k = H, V, D$; $j < 0$ y $2^j n \in [0,1]^2$.

Si $f \in \mathbf{L}^2([0,1]^2)$ es una imagen C^α, su aproximación f_M cumple $\|f - f_M\|^2 = O(M^{-\alpha})$. Sin embargo, si la imagen presenta una discontinuidad –borde–, y considerando al borde una función C^α, se tiene entonces que f es regular por tramos fuera del conjunto de bordes. En este caso, el error sólo satisface $\|f - f_M\|^2 = O(M^{-1})$ ya que los bordes del objeto generan una gran cantidad de coeficientes de alta amplitud en todas las escalas y es precisamente esta densidad de coeficientes lo que ralentiza el decaimiento del error.

Diversas aproximaciones, como curvelets, contourlets y otras transformaciones genéricamente denominadas *x-lets* han sido propuestas para capturar las singularidades sobre los bordes de los objetos. El funcionamiento general se basa en dividir el dominio de la imagen en cuadrados con tamaño variable según la continuidad del área cubierta, y aplicar sus potencialidades en aquéllos subdominios que contienen un segmento de borde.

A continuación se presenta la idea conceptual de la transformación bandlet. Este operador comparte con las *x-lets* el esquema de partición diádica de la imagen, pero en cada cuadrado no define la geometría por la localización del borde sino por su orientación. Un campo vectorial denominado *flujo geométrico* define la orientación, la que resulta paralela a la discontinuidad en todo punto, generando una "banda". El subdominio es luego "enderezado" en dirección horizontal o vertical para dejar el flujo geométrico paralelo a un eje. De este modo, la descomposición de la imagen enderezada en una base wavelet anisotrópica con soporte rectangular es equivalente a descomponer la imagen original en bandlets ortogonales paralelas al flujo.

El enderezado se realiza mediante una operación de *warping*[2]. Si consideramos un cuadrado $S \subset [0,1]^2$, la curva del borde puede ser parametrizada por $x_2 = \gamma(x_1)$ y las coordenadas del flujo pueden ser escritas como $(1, \tilde{\gamma}'(x_1))$, donde $\tilde{\gamma}'$ especifica la dirección del flujo. El operador warping w queda definido, entonces por

$$(\tilde{x_1}, \tilde{x_2}) = w(x_1, x_2) \triangleq (x_1, x_2 - \tilde{\gamma}(x_1)). \tag{4}$$

El proceso de bandletización consiste en la transformación de los coeficientes wavelet de acuerdo al flujo geométrico, en un proceso que parte de una segmentación de cada escala 2^j en cuadrados S no solapados de longitud 2^{-k}. En cada cuadrado se evalúa el valor de $\tilde{\gamma}'_S$ y se continúa la subdivisión en aquellos que presenten magnitud considerable hasta cierta profundidad. Este esquema conduce a una estructura jerárquica tipo árbol binario (*quadtree*) $\mathcal{S}_j$. Una vez definido el árbol, el operador w mapea S en $\tilde{S}$ mediante el operador w. A continuación, cada cuadrado modificado $\tilde{S}$ de longitud λ es recursivamente dividido en $2^{-\ell}$ bandas horizontales, por lo que $\tilde{S} = \bigcup_{i=0}^{2^{-\ell}-1} \tilde{\beta}_{\ell,i}$. Esta división es realizada hasta la escala $\ell = L$ tal que $2^L(\lambda 2^{-j})^2 \leq p(p+1)/2$, para un valor dado p.

[2]Proceso de restauración no lineal que cambia la localización de los píxeles, sin alterar su valor.

Teniendo en cuenta la última condición, se construyen las bases ortogonales de Alpert como el conjunto

$$\mathcal{B}(S, \tilde{\gamma}') \triangleq \left\{ a_{l,n} \setminus L \leq \ell \leq 0 \text{ y } 0 \leq m < p(p+1)2^{\ell-1} \right\}, \tag{5}$$

donde m es un entero que indexa las funciones de Alpert y $p-1$ es el grado de los polinomios discretos de Legendre $P_i(\tilde{x}_n)$ que se ajusta sobre cada banda de $\tilde{S}$.

El proceso de bandletización finaliza con el cálculo de los coeficientes bandlet como el producto interno $\langle f, b^k_{j,\ell,n} \rangle$ de f con las funciones bandlet, las que se construyen como las combinaciones lineales dadas por

$$b_\nu(x) = b^k_{j,\ell,n}(x) = \sum_p a_{\ell,n}[p] \psi^k_{j,p}(x), \tag{6}$$

donde $a_{\ell,n}[p]$ son los coeficientes de la transformada Alpert.

En cada cuadrado S se elige, entonces, una base de Alpert $\mathcal{B}(S, \tilde{\gamma}')$ que depende de la geometría $\tilde{\gamma}'$. Es de notar que si la imagen es regular sobre S, una base adaptada a la geometría no debería modificar los coeficientes wavelet. En estos casos, el valor del flujo geométrico no está definido y la proyección sobre $\mathcal{B}(S, \tilde{\gamma}')$ deja los coeficientes wavelet sin cambios.

La segmentación diádica de cada escala 2^j obtiene una geometría diferente Γ_j, por lo que para estos parámetros se define una base bandlet del conjunto de coeficientes wavelet como

$$\mathcal{B}(\Gamma_j) \triangleq \bigcup_{S \in \mathcal{S}_j} \mathcal{B}(S, \tilde{\gamma}'). \tag{7}$$

Definiciones similares pueden obtenerse para cada orientación k de la descomposición. El conjunto de todas las geometrías resultantes puede expresarse como

$$\Gamma = \{\Gamma^k_j\} = (\mathcal{S}^k_j, \{\tilde{\gamma}'\}_{S \in \mathcal{S}^k_j}), \tag{8}$$

que se corresponde con una base bandlet para toda la imagen dada por

$$\mathcal{B}(\Gamma) \triangleq \bigcup_{j \leq 0} \{b_\nu \setminus a_\nu \in \mathcal{B}(\Gamma_j)\}. \tag{9}$$

El conjunto de las potenciales geometrías y bases bandlet define un diccionario bandlet de bases ortogonales $\mathcal{D} = \{\mathcal{B}(\Gamma)\}$.

Una aproximación de f que provea el mínimo error para un número M de parámetros estaría dada por la aplicación de un umbral en una base bandlet de geometría óptima $\mathcal{B}(\Gamma^\star) = \{b_\nu\}$, de acuerdo a

$$f_M \triangleq \sum_{|\langle f, b_\nu \rangle| > T} \langle f, b_\nu \rangle b_\nu. \tag{10}$$

La cantidad de parámetros está dado por $M = M_B + M_{\mathcal{D}}$, con

$$M_B \triangleq \#\{\nu \setminus |\langle b_\nu, f \rangle| \geq T\}, \tag{11}$$

y $M_{\mathcal{D}}$ es el número de coeficientes necesarios para especificar $\Gamma^\star$, dado por

$$M_{\mathcal{D}} = M_G + M_S. \tag{12}$$

En (12), M_G es el número de parámetros requeridos para describir el *quadtree* y M_S es la cantidad de coeficientes de los polinomios de las bases de Alpert.

Planteando el Lagrangiano

$$\mathcal{L}(f,\mathcal{B}(\Gamma),T) \triangleq \sum_{|\langle b_\nu,f\rangle|<T} |\langle b_\nu,f\rangle|^2 + T^2 M, \qquad (13)$$

la búsqueda de la mejor base se encuentra mediante

$$\mathcal{B}(\Gamma^\star) \triangleq \underset{\mathcal{B}(\Gamma)\in\mathcal{D}_{T^2}}{\arg\min}\ \mathcal{L}(f,\mathcal{B}(\Gamma),T), \qquad (14)$$

que se resuelve mediante un algoritmo rápido [8].

La aproximación f_m obtenida satifsace

$$\|f - f_M\|^2 = O(M^{-\alpha}), \qquad (15)$$

con $\alpha < p$, resultado óptimo aún cuando la imagen está sometida a un suavizado que modele efectos de difracción durante la adquisición.

IV. Análisis de resultados

La presentación de resultados es abordada desde una perspectiva gráfica, mostrando diferentes ejemplos de restauración en presencia de ruido, reconstrucción de imágenes comprimidas y otras aplicaciones de algunas técnicas.

Muchas veces se apela al criterio subjetivo del lector en el análisis visual de los resultados de reconstrucción y denoising, y casi en ningún caso se explicitan tablas, gráficas u otra información más detallada y objetiva que la mencionada. Esta forma de presentación de resultados puede explicarse, en parte, dado el tipo de artículo en cuestión.

Referencias

[1] Mallat S. and Peyré G., A review of bandlet methods for geometrical image representation. Numerical Algorithms, 44:205-234, Springer, 2007.

[2] Candès E. and Donoho D.: Curvelets: A Surprisingly Effective Nonadaptive Representation of Objects with Edges. Vanderbilt University Press, Nashville, TN (1999).

[3] Donoho, D.: Wedgelets: nearly-minimax estimation of edges. Ann. Stat. 27, 353–382, 1999.

[4] Dragotti, P.L., Vetterli, M.: Wavelet footprints: theory, algorithms and applications. IEEE Trans. Signal Process. 51(5), 1306–1323, 2003.

[5] Wakin, M., Romberg, J., Choi, H., Baraniuk, R.: Wavelet-domain approximation and compression of piecewise smooth images. IEEE Trans. Image Proc. 15(5), 1071–1087, 2005.

[6] Alpert, B.K.: Wavelets and Other Bases for Fast Numerical Linear Algebra, pp. 181–216. In: Chui, C.K. (ed.). Academic Press, San Diego, CA, 1992.

[7] Mallat, S.: Geometrical grouplets. (Oct, 2006) (Preprint).

[8] Mallat S. and Peyré G., Orthogonal bandlet bases for geometrical image approximation. (Dec, 2006) (Preprint).