

Algoritmos Piramidales para Procesamiento Digital de Señales

Ricardo Barrón Fernández
[rbarron@vmredipn.ipn.mx]
Profesor e Investigador del CIC-IPN.
Valente López Muñoz
Profesor e Investigador del CIC-IPN.
Oswaldo Espinosa Sosa
Profesor e Investigador del CIC-IPN.

Los algoritmos piramidales parten de la hipótesis de que, en un objeto en particular, una imagen está constituida por elementos en diferentes escalas; es decir, una imagen no es solo un patrón de diferentes intensidades intercaladas, sino que estas intensidades se pueden agrupar en capas significativas en diferentes escalas llamadas perfiles dimensionales. El tamaño del pixel en cada perfil depende de la función de onda que se use para medir en esa escala; en general, se usa una función prototipo llamada ondeleta (por ser una onda de corta duración), para generar la métrica en cada perfil.

Introducción

La medición se realiza a partir del patrón de interferencia que se origina entre la forma de onda de medición y la franja de la imagen a estimar; este patrón de interferencia se promedia para obtener el valor del pixel del perfil en el centro de la onda de medición. La onda se va recorriendo hasta abarcar toda una matriz cuyos componentes constituyen los elementos de la imagen del perfil correspondiente. A esta generación

de patrones de interferencia en diferentes escalas se le llama análisis, y a la reconstrucción del objeto a partir de los perfiles se le conoce como síntesis.

Análisis Multiresolución

La base de los algoritmos piramidales es lo que se conoce como análisis multiresolución. Este consiste en la descomposición paramétrica de las señales en diferentes escalas de resolución. La escala unitaria está compuesta por una función $\phi(t)$ de corta duración (es decir con un ancho o soporte finito) y sus desplazadas $\phi_n(t) = \phi(t - n)$. Dicha función se conoce como función de dilatación o de muestreo y es tal que:

$$\langle \phi, f \rangle = \int \phi(t) f(t) \approx f(t_0)$$

donde t_0 es el centro de la gráfica de ϕ , es decir la acción de ϕ sobre f consiste en ponderar a f en una vecindad de t_0 de tal manera que $\langle \phi, f \rangle$ es una medición de la amplitud de f en t_0 .

Si se considera la ϕ desplazada $\phi_n(t) = \phi(t - n)$ las mediciones se pueden efectuar en cualquier posición t_0 , por lo tanto con ϕ y sus desplazados ϕ_n se cubren todas las posibilidades de medición en la escala unitaria. Además se cumple que $\langle \phi_n, \phi_k \rangle = \delta_n^k$ es decir la escala unitaria VV_0 forma un espacio vectorial con base ortogonal $\{\phi_n\}$. Nótese

que la resolución de la escala depende del ancho de ϕ , entre más angosta sea la ϕ es mayor la resolución, es decir aumenta el grado de aproximación de $\langle \phi, f \rangle$ con el valor de f en el centro de ϕ . Entonces, se puede aumentar o disminuir la escala incrementando o disminuyendo el ancho de la función de muestreo. A la operación de aumentar el ancho de la ϕ se le conoce como dilatación y esta dado por:

$$(D\phi)(t) = 2^{-1/2} \phi(2^{-1}t)$$

el factor $2^{-1/2}$ es para normalizar, y el operador de compresión D^{-1} es tal que:

$$(D^{-1}D\phi)(t) = (DD^{-1}\phi)(t) = \phi(t).$$

Usando las operaciones de dilatación y translación se pueden considerar a las diferentes escalas VV_m como espacios vectoriales derivados de VV_0 , $V_m = D^m V_0$ y con bases ortogonales $\{\phi_{m,n}\}$, $\phi_{m,n} = D^m T^n \phi$, con el operador de translación dado por $(T\phi)(t) = \phi(t - 1)$.

Se puede probar que las diferentes escalas VV_m forman una partición del espacio de señales

$$L^2(R) = \bigcup_{m \in Z} V_m$$

y las VV_m satisfacen entre si $VV_{m+1} \subset V_m$, esto se puede interpretar intuitivamente como que un objeto grande siempre se puede descomponer en otros más pequeños donde la resolución de los VV_m aumenta a medida que disminuye m .

Fundamentos del Algoritmo

Debido a la propiedad de inclusión entre las escalas se cumple que $VV_1 \subset V_0$ por lo que $D\phi \in V_0$ y por ser $\{\phi_n\}$ base para VV_0 , existe una secuencia de coeficientes $\{h_n\}$ tales que

$$D\phi = \sum_n h_n \phi_n,$$

esta relación se conoce como ecuación de dilatación y es básica en el análisis. Sea WW_1 complemento ortogonal de VV_1 respecto a VV_0 y $\{\psi_{1,n}\}$ base ortogonal de WW_1 , generada a partir de los deslizamientos de $\psi_{1,0} = D\psi$, en particular

$$D\psi \in W_1 \subset V_0$$

se cumple que

$$D\psi = \sum_n g_n \phi_n$$

para una secuencia única $\{g_n\}$.

En general se tiene que

$$VV_{m+1} \subset V_m$$

lo que implica que

$$VV_m = V_{m+1} \oplus W_{m+1}$$

donde WW_{m+1} es el complemento ortogonal de VV_{m+1} .

Esto implica que tenemos dos bases ortogonales para VV_m , a saber:

$$\{\phi_{m,k}\} \text{ y } \{\phi_{m+1,n}\} \cup \{\psi_{m+1,n}\},$$

por tanto, dado que $f \in V_m$, podemos parametrizar de dos maneras diferentes usando coordenadas respecto a cualquiera de estas bases, y además podemos pasar de unas coordenadas a otras. Este paso de coordenadas es lo que se conoce como transformada «wavelet», y el esquema de cálculo adopta la forma de una pirámide o árbol, según se vea, y de ahí el nombre de piramidal. Veamos

como se efectúa el cambio de coordenadas.

Esquema de Cálculo

Sea $f \in V_m$ entonces se cumple que

$$f = \sum_k a_k \phi_{m,k}$$

pero también

$$f = \sum_n b_n \phi_{m+1,n} + \sum_n c_n \psi_{m+1,n}$$

donde a_k, b_k, c_k son las coordenadas de f en VV_m, V_{m+1}, W_{m+1} respectivamente.

Ahora bien, por la propiedad de ortogonalidad de las bases se tiene que

$$\begin{aligned} a_k &= \langle \phi_{m,k}, f \rangle, \\ b_n &= \langle \phi_{m+1,n}, f \rangle, \\ c_n &= \langle \psi_{m+1,n}, f \rangle. \end{aligned}$$

Para obtener b_n en función de a_k , se considera que:

$$\begin{aligned} b_n &= \left\langle \phi_{m+1,n}, \sum_k a_k \phi_{m,k} \right\rangle = \sum_k a_k \langle \phi_{m+1,n}, \phi_{m,k} \rangle \\ b_n &= \sum_k a_k \langle D^{m+1} T^n \phi, D^m T^k \phi \rangle = \sum_k a_k \langle DT^n \phi, T^k \phi \rangle \\ b_n &= \sum_k a_k \langle T^{2n} D\phi, T^k \phi \rangle = \sum_k a_k \left\langle T^{2n} \sum_l h_l \phi_l, T^k \phi \right\rangle \\ b_n &= \sum_k a_k \left\langle \sum_l h_l \phi_{l+2n}, \phi_k \right\rangle = \sum_k a_k \sum_l \bar{h}_l \langle \phi_{l+2n}, \phi_k \rangle \\ b_n &= \sum_k a_k \sum_l \bar{h}_l \delta_{l+2n}^k = \sum_k a_k \bar{h}_{k-2n} \end{aligned}$$

Por tanto, el cambio de coordenadas de $\{a_k\}$ a $\{b_n\}$ esta dado por:

$$b_n = \sum_k a_k \bar{h}_{k-2n}$$

siguiendo un camino similar se tiene :

$$c_n = \sum_k a_k \bar{g}_{k-2n}$$

nótese que h y g son los coeficientes de las ecuaciones de dilatación para ϕ y ψ .

Para obtener a_k en función de b_n y c_n , se procede como sigue:

$$\begin{aligned} a_k &= \langle \phi_{m,k}, f \rangle = \left\langle \phi_{m,k}, \sum_n b_n \phi_{m+1,n} + \sum_n c_n \psi_{m+1,n} \right\rangle \\ a_k &= \sum_n b_n \langle \phi_{m,k}, \phi_{m+1,n} \rangle + \sum_n c_n \langle \phi_{m,k}, \psi_{m+1,n} \rangle \\ a_k &= \sum_n b_n \langle D^m \phi_k, D^m DT^n \phi \rangle + \sum_n c_n \langle D^m \phi_k, D^m DT^n \psi \rangle \\ a_k &= \sum_n b_n \langle \phi_k, DT^n \phi \rangle + \sum_n c_n \langle \phi_k, DT^n \psi \rangle \\ a_k &= \sum_n b_n \langle \phi_k, T^{2n} D\phi \rangle + \sum_n c_n \langle \phi_k, T^{2n} D\psi \rangle \\ a_k &= \sum_n b_n \left\langle \phi_k, T^{2n} \sum_l h_l \phi_l \right\rangle + \sum_n c_n \left\langle \phi_k, T^{2n} \sum_l g_l \phi_l \right\rangle \\ a_k &= \sum_n b_n \left\langle \phi_k, \sum_l \phi_{l+2n} \right\rangle + \sum_n c_n \left\langle \phi_k, \sum_l g_l \phi_{l+2n} \right\rangle \\ a_k &= \sum_n b_n \sum_l h_l \langle \phi_k, \phi_{l+2n} \rangle + \sum_n c_n \sum_l g_l \langle \phi_k, \phi_{l+2n} \rangle \\ a_k &= \sum_n b_n \sum_l h_l \delta_k^{l+2n} + \sum_n c_n \sum_l g_l \delta_k^{l+2n} \\ a_n &= \sum_n b_n h_{k-2n} + \sum_n c_n g_{k-2n} \end{aligned}$$

observamos que las coordenadas no dependen de la escala, dependen únicamente de los coeficientes $\{h_k\}$ y $\{g_k\}$ lo que les da una fuerza tremenda. El esquema de calculo se simplifica si notamos que el paso de las coordenadas en VV_m a las coordenadas en sus componentes ortogonales VV_{m+1} y WW_{m+1} se puede representar por un banco de filtros de análisis (**figura 1**).

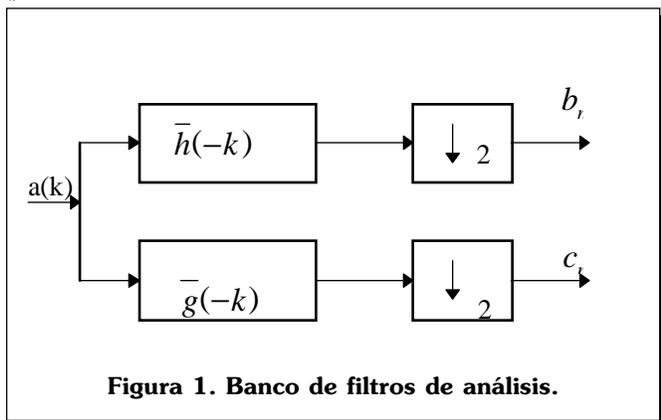


Figura 1. Banco de filtros de análisis.

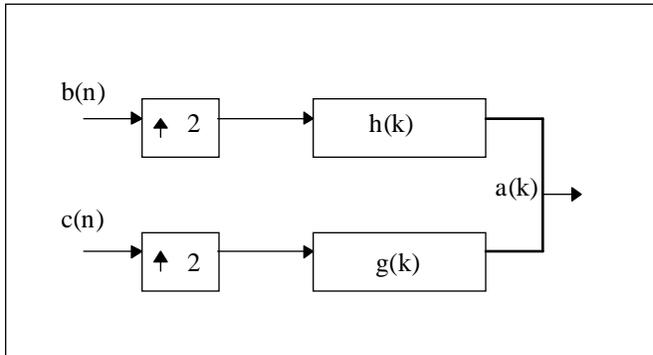


Figura 2. Reconstrucción de coordenadas originales.

La respuesta impulso de los filtros del banco depende directamente de los coeficientes de las ecuaciones de dilatación. Obsérvese que a la salida de cada filtro se tiene un bloque decimador de orden 2, el cual reduce los parámetros a la mitad. Para el proceso inverso, cuando se trata de reconstruir las coordenadas originales, se tiene el esquema de la figura 2.

Esta es la parte de síntesis que corresponde fielmente con las ecuaciones obtenidas. Para la parte de descomposición que corresponde con lo que se conoce como la transformada «wavelet», el algoritmo comienza como sigue:

Se toma $f(k\Delta t) = a_k^0$ como las coordenadas iniciales, donde $f(t)$ es la señal a transformar y Δt es el periodo de muestreo, dado que

$$a_k^0 = \langle \phi_k, f \rangle \approx f(k\Delta t)$$

(en este caso $\Delta t = 1$)

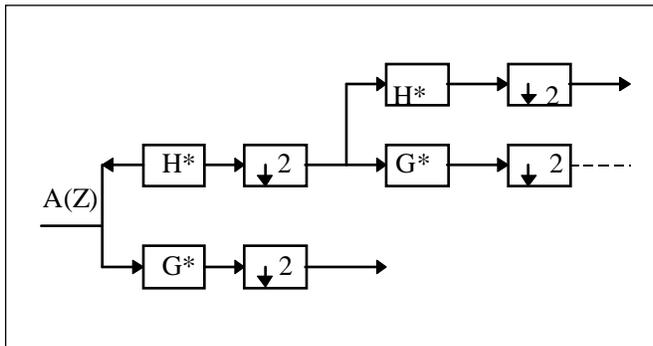


Figura 3. Esquema de una "Transformada wavelet"

es correcto tomar a_k^0 como $f(k\Delta t)$.

El proceso en general se desarrolla en base al diagrama de la figura 3, donde $A(z)$ es la transformada z de las coordenadas en VV_0 y H^* y G^* son las funciones de transferencia.

transferencia diferente; estas funciones de transferencia generan una partición del ancho de banda original, lo que da origen a lo que se conoce como codificación sub-banda.

Conclusiones

En el presente trabajo se analiza la posibilidad y se muestran ejemplos de como abordar los diferentes problemas del procesamiento digital de

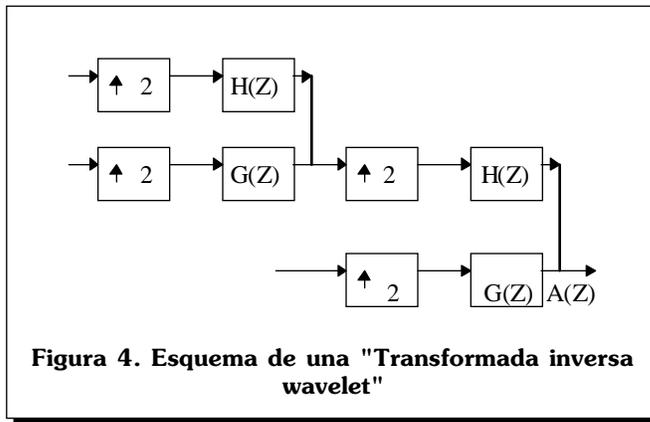


Figura 4. Esquema de una "Transformada inversa wavelet"

Para la transformada inversa, es decir, la síntesis, se tiene el diagrama de la figura 4.

Todo el proceso depende de dos filtros, $H(z)$ y $G(z)$, que tienen como respuesta impulso a los coeficientes de las dos ecuaciones de dilatación h_n y g_n .

La parte interesante consiste en el diseño de estos filtros con las propiedades requeridas. En la parte del análisis en cada etapa se aplica un filtrado, lo que da origen a que en cualquier punto de salida o canal se tenga una función de

imágenes usando técnicas algebraicas multiresolución basadas en ondeletas. También se presenta un marco teórico basado en álgebra de operadores, que permite visualizar conceptualmente los principales resultados teóricos, que son

a su vez la base de los algoritmos desarrollados. El algoritmo presentado no es la única manera de hacer el cambio de coordenadas, y si cada complemento ortogonal WW_{m+1} se descompone a su vez ortogonalmente, las posibilidades de descomposición aumentan y los algoritmos que se obtienen, aunque siguen presentando características piramidales en el dominio de la frecuencia, dividen el ancho de banda cada uno en forma propia. La ventaja de separar una señal en bandas de frecuencia consiste en que cada una de estas bandas, al ser obtenidas por separación ortogonal, no están correlacionadas, por lo que se pueden diseñar técnicas de codificación óptimas por banda, y en los casos de compresión desechar algunas bandas innecesarias.

Bibliografía y Referencias

- [1] Yves Meyer, "*Wavelets Algorithms & Applications*". SIAM, 1993.
- [2] Charles K. Chui, "*An Introduction to Wavelet*". Academic Press, 1992.
- [3] Pamela C. Cosman, "*Vector Quantization of Image Subbands*". IEEE Transaction on image processing, Vol.5, No.2, Febrero 1996.
- [4] Gerald Kaiser, "*A Friendly Guide to Wavelets*". Birkhäuser, 1994.
- [5] Stjephan G. Mallat, "*A Theory for Multiresolution Signal Decomposition: The Wavelet Representation*". IEEE Transaction on pattern analysis and machine intelligence, Vol.II, No.7, Julio 1989.