

Appunti di Elaborazione di Segnali Multimediali

a.a. 2017/2018

La Trasformata Wavelet

L. Verdoliva

La trasformata wavelet è un potente strumento per l'analisi e l'elaborazione dei segnali e risulta estremamente efficiente in diversi campi di applicazione, come la compressione e il denoising, e in generale quando si ha a che fare con segnali non stazionari, come sono le immagini. Queste infatti sono caratterizzate da andamenti di lungo periodo alle basse frequenze (sfondi, background) detti anche *trend*, e andamenti di breve periodo alle alte frequenze (discontinuità, bordi) detti anche *anomalies*. Questi ultimi, sebbene occupino una percentuale relativamente piccola nell'immagine, presentano un elevato contenuto informativo, e vanno quindi rappresentati adeguatamente. La trasformata wavelet presenta diverse interessanti proprietà:

- è in grado di analizzare segnali con caratteristiche variabili nel tempo, grazie alla sua buona capacità di localizzazione tempo-frequenza;
- fornisce una rappresentazione su diverse scale (rappresentazione multirisoluzione);
- è facilmente realizzabile attraverso un banco di filtri.

Di seguito vengono esposti tali concetti sviluppando la trattazione analitica per segnali monodimensionali, per semplicità, per poi estenderla alle immagini. Infine viene descritto un algoritmo molto efficiente per la codifica dei coefficienti wavelet detto Embedded Zerotree Wavelet (EZW).

1 L'analisi tempo-frequenza

Allo scopo di introdurre il concetto di localizzazione tempo-frequenza partiamo dall'analisi nel dominio della frequenza di un segnale analogico $x(t)$ e ricordiamo che la sua trasformata di Fourier si definisce come:

$$FT[x(t)] = X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (1)$$

A questo punto consideriamo due casi in cui i segnali da trasformare sono un tono puro alla frequenza f' : $x(t) = \sin(2\pi f't)$ e un impulso di Dirac centrato all'istante t' : $x(t) = \delta(t - t')$. In figura 1 sono mostrati i moduli delle relative trasformate di Fourier per questi due segnali. Si può notare come la sinusoide risulti perfettamente localizzata in frequenza, dato che è individuata da due impulsi centrati alla frequenza f' , invece per rappresentare l'impulso sono necessarie tutte le componenti frequenziali. Ciò non deve sorprenderci data la natura periodica delle funzioni

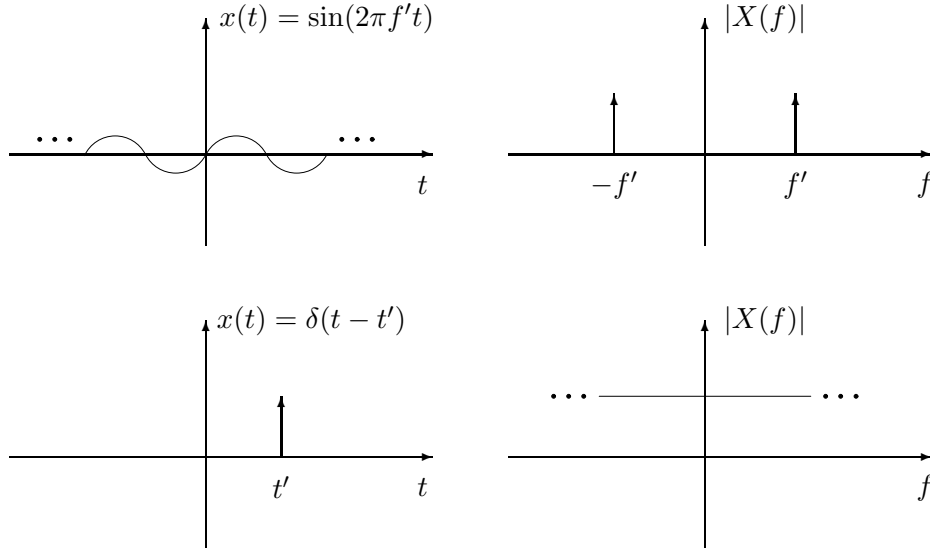


Figura 1: Segnale sinusoidale e impulso di Dirac con le relative trasformate di Fourier.

utilizzate nel kernel della trasformazione, che rendono la trasformata di Fourier un potente strumento per l'analisi dei segnali stazionari. Tuttavia se il segnale presenta caratteristiche variabili nel tempo la trasformata riesce con difficoltà a catturare gli effetti locali (transitori), che tendono a disperdersi su larghe regioni nel dominio trasformato.

Evidentemente se l'obiettivo è quello di rappresentare una brusca transizione è conveniente analizzare il segnale nel dominio del tempo in cui è possibile localizzare perfettamente un impulso ideale, infatti risulta (proprietà di riproducibilità di $\delta(t)$):

$$x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau \quad (2)$$

Se indichiamo con $\mathcal{T}_x(\gamma)$ la generica trasformazione che vogliamo utilizzare per $x(t)$:

$$\mathcal{T}_x(\gamma) = \int_{-\infty}^{\infty} x(t) \phi_{\gamma}^*(t) dt \quad (3)$$

la capacità di localizzazione di una trasformata è legata proprio alla durata nel tempo e in frequenza delle funzioni $\phi_{\gamma}(t)$. Per la trasformata di Fourier risulta $\phi_{\gamma}(t) \equiv \phi_f(t) = e^{j2\pi ft}$, per cui le funzioni hanno durata temporale infinita ($\Delta t = \infty$) e durata frequenziale nulla ($\Delta f = 0$), il contrario accade per la trasformata identità per cui $\phi_{\gamma}(t) \equiv \phi_{\tau}(t) = \delta(t - \tau)$. Di conseguenza, la trasformata di Fourier presenta una *localizzazione in frequenza* perfetta (è sempre possibile discriminare nel dominio trasformato due toni sinusoidali caratterizzati da frequenze comunque vicine tra loro); analogamente, la trasformata identità presenta una *localizzazione nel tempo* perfetta, essendo sempre possibile distinguere due impulsi a distanza piccola a piacere.

Se il segnale che bisogna analizzare presenta sia andamenti lentamente variabili (simili a delle sinusoidi) che brusche variazioni (simili a delle delta), sarebbe desiderabile avere una trasformata

che presenti sia una buona localizzazione temporale che frequenziale. Dato un segnale $x(t) \in L^2(\mathbb{R})$, ci proponiamo di individuare una trasformata tempo-frequenza:

$$\mathcal{T}_x(f, \tau) = \int_{-\infty}^{+\infty} x(t) \phi_{f\tau}^*(t) dt \quad (4)$$

caratterizzata da funzioni $\phi_{f\tau}(t) \in L^2(\mathbb{R})$, ben localizzate sia nel tempo che in frequenza, note anche come *atomi tempo-frequenza*, con energia unitaria:

$$E_\phi = \|\phi_{f\tau}(t)\|^2 = \int_{-\infty}^{+\infty} |\phi_{f\tau}(t)|^2 dt = 1$$

Si noti che l'integrale espresso nella (4) può anche essere interpretato, in $L^2(\mathbb{R})$, come prodotto scalare e misura proprio l'indice di somiglianza tra i due segnali:

$$\mathcal{T}_x(f, \tau) = \langle x(t), \phi_{f\tau}(t) \rangle \quad (5)$$

Il nostro obiettivo è individuare degli atomi tempo-frequenza con durata temporale Δt e durata frequenziale Δf quanto più piccole possibile, in modo che siano caratterizzate da una buona localizzazione tempo-frequenza. Ovviamente è anche importante che queste funzioni siano relazionate tra loro da semplici operazioni elementari, in modo che l'operazione di trasformazione possa essere realizzata in pratica con un accettabile trade-off costo-prestazioni.

Per poter valutare Δt e Δf è necessario innanzitutto definire cosa si intende per durata di un segnale. E' vero, per esempio, che la durata di alcune funzioni è infinita, tuttavia spesso il contributo risulta essere trascurabile all'infinito. Allora si è soliti misurare l'intervallo temporale o frequenziale in cui è concentrata la maggior parte dell'energia del segnale attraverso il concetto di durata efficace:

$$(\Delta t)_{f\tau}^2 \triangleq \frac{\int_{-\infty}^{+\infty} (t - m_{f\tau})^2 |\phi_{f\tau}(t)|^2 dt}{\int_{-\infty}^{+\infty} |\phi_{f\tau}(t)|^2 dt} \quad (6)$$

e banda efficace:

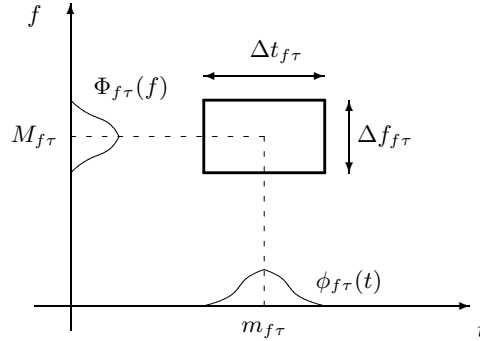
$$(\Delta f)_{f\tau}^2 \triangleq \frac{\int_{-\infty}^{+\infty} (f - M_{f\tau})^2 |\Phi_{f\tau}(f)|^2 df}{\int_{-\infty}^{+\infty} |\Phi_{f\tau}(f)|^2 df} \quad (7)$$

dove $m_{f\tau}$ e $M_{f\tau}$ rappresentano i valori medi definiti come:

$$m_{f\tau} \triangleq \int_{-\infty}^{+\infty} t |\phi_{f\tau}(t)|^2 dt \quad M_{f\tau} \triangleq \int_{-\infty}^{+\infty} f |\Phi_{f\tau}(f)|^2 df$$

Le quantità definite nella (6) e (7) forniscono un indice di dispersione della distribuzione di energia di $\phi_{f\tau}(t)$ e $\Phi_{f\tau}(f) = FT[\phi_{f\tau}(t)]$ intorno al valore medio. Poiché considereremo sempre funzioni tempo-frequenza con energia unitaria, il denominatore delle equazioni (6) e (7) sarà sempre pari a 1.

Un modo per visualizzare la capacità di localizzazione tempo-frequenza è attraverso l'uso del piano tempo-frequenza, che fornisce una rappresentazione astratta bidimensionale della dispersione tempo-frequenza delle funzioni utilizzate nella trasformata. In particolare, ogni funzione è rappresentata generalmente da una regione di forma rettangolare (detta *cella di risoluzione*), la cui larghezza e altezza sono proprio pari alla durata e alla banda della funzione, rispettivamente. In figura 2 si mostra tale rappresentazione per il generico atomo $\phi_{f\tau}(t)$.

Figura 2: Cella di risoluzione relativa a $\phi_{f\tau}(t)$

1.1 La Short-Time Fourier Transform (STFT)

Un modo molto semplice per ottenere una trasformata tempo-frequenza è quello di valutare la trasformata di Fourier su intervalli temporali di durata finita, ottenuti moltiplicando il segnale per una fissata finestra $g(t)$. Quindi prima si individua la porzione di segnale da elaborare e poi si effettua la trasformata di Fourier. Questo equivale a considerare atomi tempo-frequenza del tipo:

$$\phi_{f\tau}(t) \equiv g_{f\tau}(t) = g(t - \tau) e^{j2\pi ft} \quad (8)$$

Possiamo così definire la *Short Time Fourier Transform* (STFT):

$$STFT[x(t)] = STFT[f, \tau] = \int_{-\infty}^{+\infty} x(t) g^*(t - \tau) e^{-j2\pi ft} dt \quad (9)$$

Gli atomi tempo-frequenza definiti nella (9) rappresentano versioni traslate e modulate di $g(t)$, detta finestra elementare. Generalmente si preferisce non utilizzare una finestra rettangolare, perché determina delle discontinuità “artificiali”, che non sono presenti nel segnale originale. Spesso si sceglie una finestra gaussiana, $g(t) = \alpha e^{-\beta t^2}$, (figura 3) e si parla in tal caso di trasformata di Gabor. Si può dimostrare che la trasformazione definita nella (9) è invertibile e che la formula di ricostruzione è:

$$x(t) = \int \int_{-\infty}^{+\infty} STFT[x(t)] g_{f\tau}(t) df d\tau \quad (10)$$

Per determinare la capacità di localizzazione tempo-frequenza di questa trasformata determiniamo la durata e la banda efficace. Se facciamo riferimento a $g(t)$ otteniamo dalle (6) e (7):

$$\begin{aligned} (\Delta t)^2 &= \int_{-\infty}^{+\infty} t^2 |g(t)|^2 dt \\ (\Delta f)^2 &= \int_{-\infty}^{+\infty} f^2 |G(f)|^2 df \end{aligned}$$

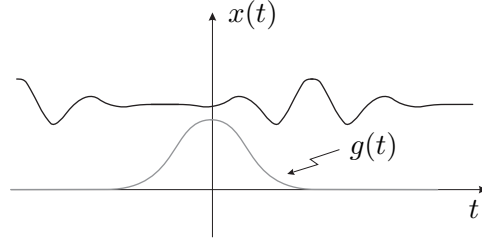


Figura 3: finestra gaussiana

con $G(f) = FT[g(t)]$. Se invece consideriamo un generico atomo centrato in $\tau = t_0$ e con $f = f_0$, otteniamo per la durata:

$$\begin{aligned} (\Delta t)_{f_0 t_0}^2 &= \int_{-\infty}^{+\infty} (t - t_0)^2 |g_{f_0 t_0}(t)|^2 dt \\ &= \int_{-\infty}^{+\infty} (t - t_0)^2 g^2(t - t_0) dt = \int_{-\infty}^{+\infty} \tau^2 g^2(\tau) d\tau \equiv (\Delta t)^2 \end{aligned}$$

Analogamente per la banda risulterà:

$$(\Delta f)_{f_0 t_0}^2 \equiv (\Delta f)^2$$

Quindi tutti gli atomi hanno nel piano la stessa risoluzione tempo-frequenza pari a Δt e Δf , come si mostra in figura 4. In effetti, dal momento che gli atomi sono costruiti attraverso traslazioni nel tempo e modulazioni in frequenza della stessa funzione $g(t)$, la durata e la banda non vengono modificate. Un vantaggio evidente della STFT è che, a differenza della FT, è in grado di localizzare il contenuto frequenziale in un intervallo di durata Δt , d'altra parte non è in grado di discriminare due toni puri la cui distanza in frequenza è inferiore a Δf . In effetti, una buona trasformata tempo-frequenza dovrebbe avere Δt e Δf molto piccoli, tuttavia il principio di incertezza esclude la possibilità di avere una risoluzione arbitrariamente elevata sia nel tempo che in frequenza, dal momento che limita inferiormente il prodotto durata-banda delle funzioni nel modo seguente:

$$\Delta t \Delta f \geq \frac{1}{4\pi} \quad (11)$$

Questa relazione è anche detta *disuguaglianza di Heisenberg*.

Il calcolo della STFT prevede di far variare con continuità i parametri τ e f , producendo così una trasformata estremamente ridondante. Ciò non è conveniente qualora la STFT debba essere usata nell'ambito della compressione dati, pertanto i parametri di tempo e frequenza vengono discretizzati nel seguente modo:

$$\begin{cases} \tau = n\tau_0 \\ f = mf_0 \end{cases}$$

ottenendo le funzioni:

$$g_{mn}(t) = g(t - n\tau_0) e^{j2\pi mf_0 t}$$

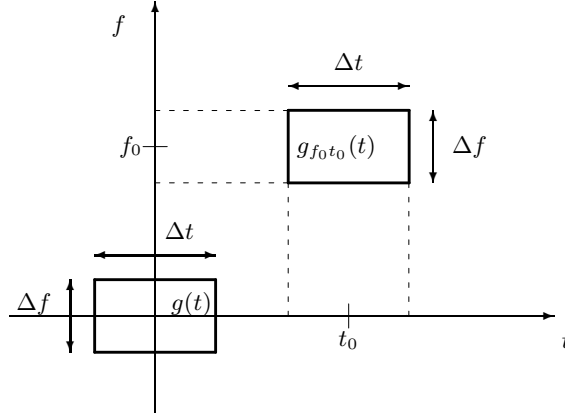


Figura 4: Celle di risoluzione per la STFT.

in questo modo la trasformata diventa una serie, i cui coefficienti possono essere determinati come:

$$\begin{aligned} STFT[m, n] &= \int_{-\infty}^{+\infty} x(t) g^*(t - n\tau_0) e^{-j2\pi m f_0 t} dt \\ &= \langle x(t), g_{mn}(t) \rangle \end{aligned} \quad (12)$$

Si può dimostrare che scegliendo opportunamente τ_0, f_0 e $g(t)$ è ancora possibile a partire dai coefficienti trasformati ricostruire il segnale $x(t)$ come:

$$x(t) = \sum_n \sum_m \langle x(t), g_{mn}(t) \rangle g_{mn}(t) \quad (13)$$

In figura 5 è mostrata la copertura del piano tempo-frequenza relativa alla trasformata in cui sono stati discretizzati i parametri.

Esempio: Consideriamo il segnale

$$x(t) = \sin(2\pi f' t) + \delta(t - t')$$

e calcoliamo i coefficienti della STFT secondo la (12), quindi disegniamo delle zone d'ombra in corrispondenza dei coefficienti che presentano ampiezza maggiore. Evidentemente in questo caso, supponendo che

$$3f_0 - \Delta f/2 \leq f' \leq 3f_0 + \Delta f/2$$

e

$$2t_0 - \Delta t/2 \leq t' \leq 2t_0 + \Delta t/2$$

le zone d'ombra sono quelle mostrate in figura 6, in cui si evidenzia come sia possibile localizzare la frequenza della sinusoide con un'incertezza pari proprio a Δf e la posizione temporale

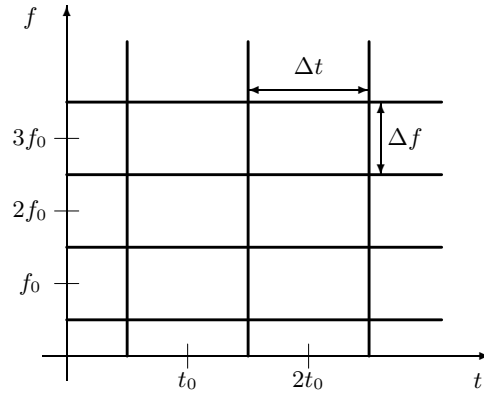


Figura 5: Copertura del piano tempo-frequenza per la STFT.

dell'impulso con incertezza pari proprio a Δt . Ovviamente se volessimo migliorare la risoluzione nel tempo quella in frequenza peggiorerebbe e viceversa.

Confrontiamo questa situazione con ciò che otterremmo con la trasformata di Fourier e con la trasformata identità. Discretizzando i parametri di tempo e frequenza, considerando cioè le serie corrispondenti, avremo le situazioni mostrate in figura 7. Queste possono essere interpretate come due situazioni limite: nel primo caso è possibile individuare la sinusoide, ma si perde la localizzazione dell'impulso, nel secondo invece accade il contrario.

1.2 La Trasformata Wavelet (CWT)

Da quanto detto nel paragrafo precedente si può concludere che la STFT risulta essere inadeguata per l'elaborazione di segnali naturali, come le immagini, dato che non è possibile localizzare

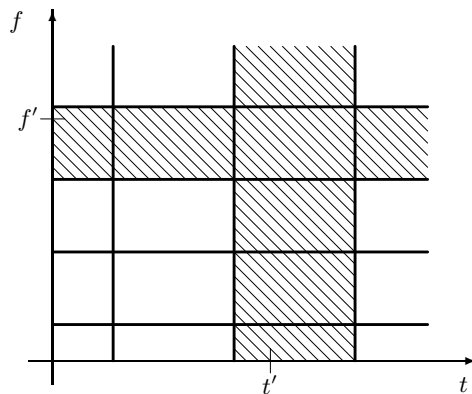


Figura 6: Esempio di localizzazione tempo-frequenza per la STFT.

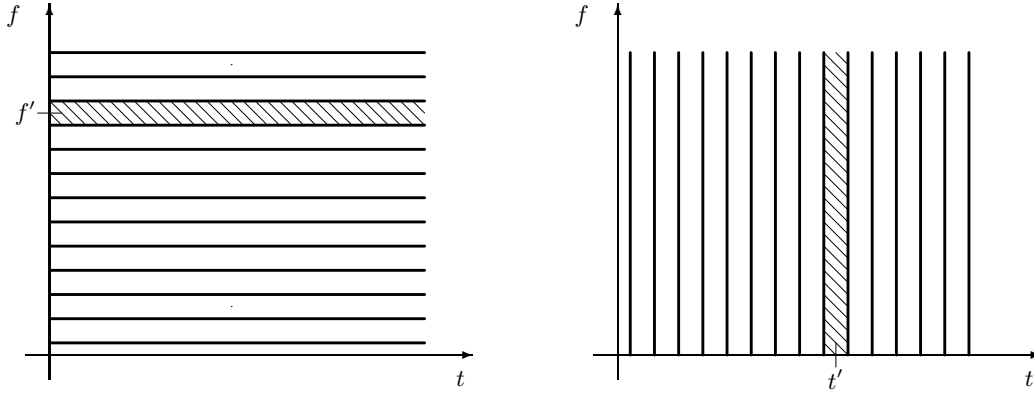


Figura 7: Esempio di localizzazione tempo-frequenza per la FT e la trasformata identità.

con la stessa precisione fenomeni che evolvono rapidamente (anomalies) e fenomeni che invece evolvono lentamente (trend). Per ottenere un'analisi a risoluzione variabile occorre far in modo che all'aumentare della frequenza f la banda Δf aumenti in modo proporzionale. Ciò si può ottenere costruendo un insieme di funzioni base mediante traslazione e cambiamento di scala di un'unica funzione, $\psi(t)$, detta *mother wavelet*:

$$\phi_{f\tau}(t) \equiv \psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (14)$$

A partire dalla mother wavelet è possibile definire la *trasformata wavelet continua* (CWT, Continuous Wavelet Transform) del segnale $x(t)$:

$$\begin{aligned} CWT[x(t)] = CWT[a, b] &= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-b}{a}\right) dt \\ &= \langle x(t), \psi_{ab}(t) \rangle \end{aligned} \quad (15)$$

Si può dimostrare che la relazione che permette di ricostruire il segnale originario a partire da quello trasformato è la seguente (*Resolution of identity*):

$$x(t) = \frac{1}{C_\psi} \iint_{-\infty}^{+\infty} \frac{1}{a^2} CWT[a, b] \psi_{ab}(t) da db \quad (16)$$

Per garantire l'invertibilità della trasformata la costante C_ψ deve essere una quantità finita:

$$C_\psi = \int_0^{+\infty} \frac{|\Psi(f)|^2}{f} df < \infty \quad (17)$$

dove $\Psi(f) = FT(\psi(t))$, e decade rapidamente a zero grazie alla proprietà di localizzazione in frequenza della wavelet. Quindi l'unica condizione da imporre per garantire che la costante C_ψ non diverga è $\Psi(0) = 0$, che equivale a riconoscere che il valor medio della mother wavelet deve essere nullo dato che:

$$\Psi(0) = \int_{-\infty}^{+\infty} \psi(t) dt$$

Da quanto detto è possibile giustificare il nome della trasformata wavelet, in quanto gli atomi tempo-frequenza devono oscillare (come un'onda, in inglese *wave*) e allo stesso tempo devono essere ben localizzati nel tempo (*small wave*). In figura 8 si mostra un esempio di mother wavelet e la corrispondente versione espansa ($a > 1$) e compressa ($a < 1$).

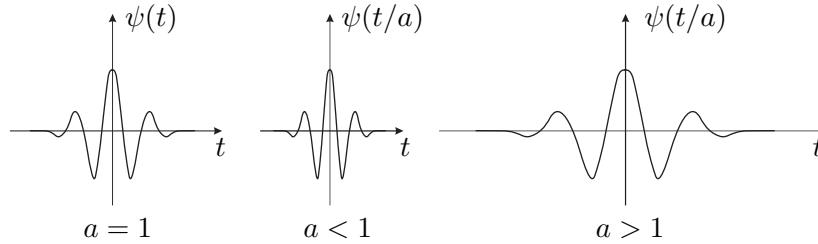


Figura 8: Esempi di atomi tempo-frequenza per la Wavelet.

Valutiamo anche in questo caso la capacità di localizzazione tempo-frequenza della trasformata, determinando Δt e Δf , tenendo presente che la funzione wavelet ha un comportamento passabanda. Per la mother wavelet risulta:

$$(\Delta t)^2 = \int_{-\infty}^{+\infty} t^2 |\psi(t)|^2 dt$$

$$(\Delta f)^2 = \int_{-\infty}^{+\infty} (f - f_0)^2 |\Psi(f)|^2 df$$

Considerando un generico atomo centrato in $b = b_0$ e con $a = a_0$, otteniamo per la durata:

$$\begin{aligned}
 (\Delta t)_{a_0 b_0}^2 &= \int_{-\infty}^{+\infty} (t - b_0)^2 |\psi_{a_0 b_0}(t)|^2 dt \\
 &= \int_{-\infty}^{+\infty} (t - b_0)^2 \frac{1}{a_0} \psi^2 \left(\frac{t - b_0}{a_0} \right) dt \\
 &= a_0^2 \int_{-\infty}^{+\infty} \tau^2 \psi^2(\tau) d\tau = a_0^2 (\Delta t)^2
 \end{aligned} \tag{18}$$

Nella (18) è stato fatto il cambio di variabili: $\tau = \frac{t-b_0}{a_0}$. In conclusione:

$$(\Delta t)_{a_0 b_0} = a_0 \Delta t \tag{19}$$

Analogamente si può far vedere che:

$$(\Delta f)_{a_0 b_0} = \frac{\Delta f}{a_0} \tag{20}$$

Questo significa che le celle di risoluzione per la trasformata Wavelet sono variabili nel piano tempo-frequenza così come è mostrato in figura 9.

Anche per la wavelet siamo interessati alla versione discreta di tale trasformata. In questo caso, affinché il piano tempo-frequenza sia completamente coperto, la discretizzazione del parametro b deve essere legata a quella del parametro a perchè funzioni compresse sull'asse temporale

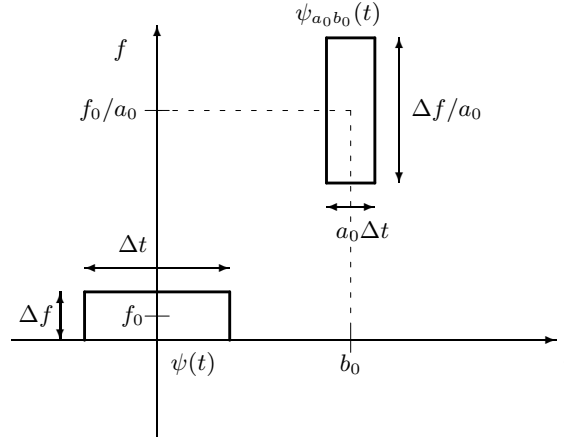


Figura 9: Celle di risoluzione per la CWT.

devono essere traslate con passo piccolo per coprire l'intero range temporale, mentre funzioni espanse sull'asse dei tempi devono essere traslate con passo grande. La discretizzazione quindi è la seguente:

$$\begin{cases} a = a_0^{-m} & a_0 > 1, \quad m \in \mathbb{Z} \\ b = n b_0 a_0^{-m} & b_0 > 0, \quad n \in \mathbb{Z}. \end{cases} \quad (21)$$

Sostituendo nella (14) i parametri così definiti, si ottengono le seguenti funzioni:

$$\psi_{m,n}(t) = a_0^{\frac{m}{2}} \psi(a_0^m t - n b_0) \quad (22)$$

Di particolare interesse nell'elaborazione di immagini è la discretizzazione su una griglia diadica, che si ottiene per $a_0 = 2$ e $b_0 = 1$:

$$\psi_{m,n}(t) = 2^{\frac{m}{2}} \psi(2^m t - n) \quad (23)$$

e che conduce al piano tempo frequenza mostrato in figura 10, dove si osserva che le funzioni wavelet presentano una buona localizzazione frequenziale alle basse frequenze, e allo stesso tempo, una buona localizzazione temporale alle alte frequenze. Anche in questo caso si può dimostrare che, scegliendo opportunamente $\psi(t)$, è valida la seguente formula di ricostruzione:

$$x(t) = \sum_n \sum_m \langle x(t), \psi_{mn}(t) \rangle \psi_{mn}(t) \quad (24)$$

dove le $\psi_{mn}(t)$ costituiscono un insieme di funzioni base ortonormali.

Esempio: Consideriamo nuovamente il segnale

$$x(t) = \sin(2\pi f' t) + \delta(t - t')$$

dove questa volta facciamo l'ipotesi che la frequenza della sinusoide sia piccola, in modo da considerare un segnale costituito da una sinusoide lentamente variabile e un impulso.

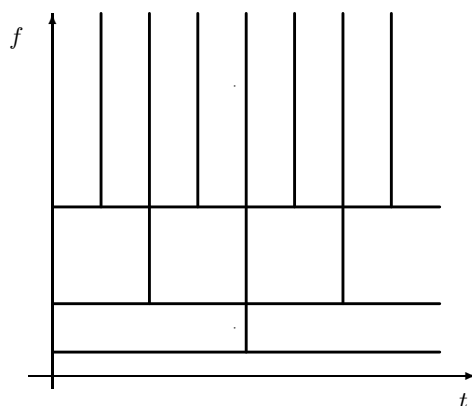


Figura 10: Copertura del piano tempo-frequenza per la wavelet.

In questo caso la copertura del piano tempo frequenza è mostrata in figura 11, in cui appare evidente la capacità della trasformata di localizzare con un'incertezza piccola sia la sinusoide che l'impulso.

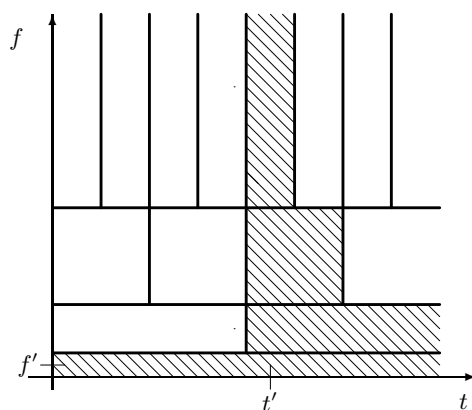


Figura 11: Esempio di localizzazione tempo-frequenza per la wavelet.

La più semplice mother wavelet a supporto compatto è quella di Haar, definita come:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{altrimenti} \end{cases}$$

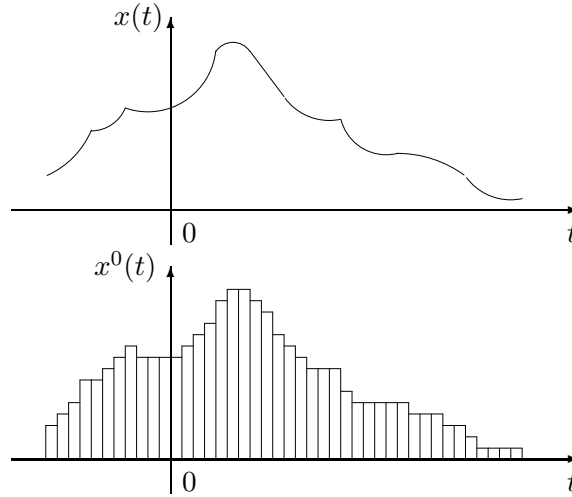


Figura 12: Segnale $x(t)$ e sua rappresentazione mediante funzioni costanti a tratti.

2 Analisi multirisoluzione

In questo paragrafo descriviamo il legame tra la trasformata Wavelet e l'analisi multirisoluzione, nel caso più semplice in cui le funzioni wavelet siano di tipo Haar. Consideriamo allora un generico segnale $x(t)$ (figura 12) e ricaviamone una rappresentazione più grossolana, $x^{(0)}(t)$, utilizzando la seguente famiglia di funzioni costanti a tratti $\phi_{0k}(t) = \phi(t - k)$ (figura 13). L'espansione in serie del segnale $x(t)$ può essere scritta nella seguente forma:

$$x^{(0)}(t) = \sum_k c_{0,k} \phi_{0,k}(t) \quad (25)$$

con i coefficienti c_{0k} dati dalla relazione¹:

$$c_{0,k} = \langle x(t), \phi_{0,k}(t) \rangle = \int_{-\infty}^{\infty} x(t) \phi_{0,k}(t) dt = \int_k^{k+1} x(t) dt \quad (26)$$

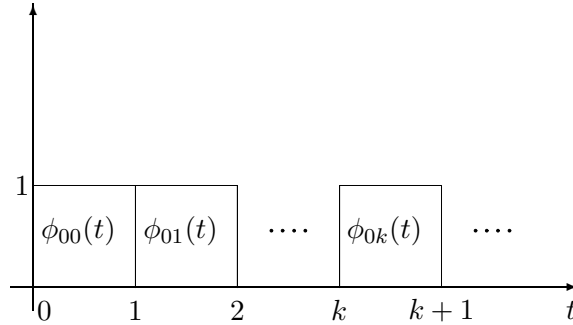
Se volessimo ottenere una rappresentazione del segnale $x(t)$ a risoluzione maggiore dovremmo considerare delle funzioni di scaling più strette: $\phi_{1,k}(t) = \sqrt{2} \phi(2t - k)$ (il fattore $\sqrt{2}$ fa sì che le funzioni di espansione abbiano tutte energia unitaria), in questo modo otteniamo una versione più accurata di $x(t)$:

$$x^{(1)}(t) = \sum_k c_{1,k} \phi_{1,k}(t) \quad (27)$$

¹Quando si considerano espansioni in serie del tipo: $x(t) = \sum_n \alpha_n \phi_n(t)$, dove $\phi_n(t)$ rappresentano un insieme di funzioni ortonormali tra loro, allora risulta:

$$\langle x, \phi_k \rangle = \sum_n \alpha_n \langle \phi_n, \phi_k \rangle = \sum_n \alpha_n \delta_{n-k} = \alpha_k$$

cioè i coefficienti di espansione sono dati proprio dal prodotto scalare del segnale con le funzioni ϕ .

Figura 13: Funzioni di base $\phi_{0,k}(t)$.

dove i coefficienti c_{1k} questa volta sono dati da:

$$c_{1,k} = \int_{-\infty}^{\infty} x(t) \phi_{1,k}(t) dt = \sqrt{2} \int_{\frac{k}{2}}^{\frac{k+1}{2}} x(t) dt \quad (28)$$

La rappresentazione $x^{(1)}(t)$, a risoluzione maggiore, contiene sicuramente la rappresentazione $x^{(0)}(t)$, a risoluzione più bassa, il che ci consente di relazionare i coefficienti $c_{0,k}$ e $c_{1,k}$. A tal fine ricaviamo i coefficienti $c_{1,2k}$ e $c_{1,2k+1}$:

$$c_{1,2k} = \sqrt{2} \int_k^{k+\frac{1}{2}} x(t) dt, \quad (29)$$

$$c_{1,2k+1} = \sqrt{2} \int_{k+\frac{1}{2}}^{k+1} x(t) dt \quad (30)$$

da cui si ottiene

$$c_{0,k} = \int_k^{k+1} x(t) dt = \int_k^{k+\frac{1}{2}} x(t) dt + \int_{k+\frac{1}{2}}^{k+1} x(t) dt = \frac{c_{1,2k} + c_{1,2k+1}}{\sqrt{2}}. \quad (31)$$

Più in generale, possiamo considerare la seguente famiglia di funzioni base per l'espansione del segnale:

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k), \quad (32)$$

con i coefficienti c_{jk} dati dalla relazione seguente:

$$c_{j,k} = 2^{j/2} \int_{\frac{k}{2^j}}^{\frac{k+1}{2^j}} x(t) dt \quad (33)$$

Le funzioni $\phi_{j,k}(t)$ sono dette *funzioni di scaling*, perché, al variare di $j = 0, \dots, n$, forniscono una rappresentazione del segnale $x(t)$ a diverse risoluzioni: $x^{(0)}(t), x^{(1)}(t), \dots, x^{(n)}(t)$, e generano

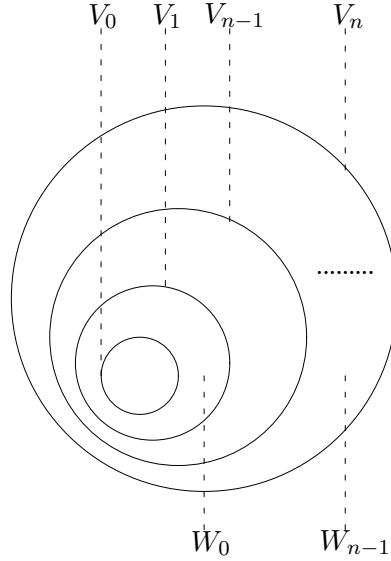


Figura 14: Rappresentazione mediante gli spazi vettoriali.

i seguenti spazi vettoriali:

$$\begin{aligned}
 V_0 &= \overline{\text{span}\{\phi_{0,k}(t)\}}, \\
 V_1 &= \overline{\text{span}\{\phi_{1,k}(t)\}}, \\
 \dots &= \dots\dots\dots \\
 V_n &= \overline{\text{span}\{\phi_{n,k}(t)\}}.
 \end{aligned} \tag{34}$$

Il generico spazio V_j è quindi costituito da tutti questi segnali che possono essere espressi mediante combinazione lineare, secondo opportuni coefficienti, delle funzioni $\phi_{j,k}(t)$. La sequenza di segnali $x^{(0)}(t), x^{(1)}(t), \dots, x^{(n)}(t)$ per come è stata costruita rappresenta una gerarchia di segnali a diversa risoluzione, che però risulta essere fortemente ridondante, essendoci una relazione di inclusione tra gli spazi (figura 14):

$$V_0 \subset V_1 \subset \dots\dots\dots \subset V_n \tag{35}$$

In effetti è possibile ottenere una rappresentazione priva di ridondanza se facciamo il seguente ragionamento. Supponiamo di avere un segnale $x(t)$ appartenente allo spazio V_1 , vale a dire:

$$x(t) \in V_1 \implies x(t) \equiv x^{(1)}(t)$$

Risulta:

$$x(t) = x^{(0)}(t) + (x^{(1)}(t) - x^{(0)}(t))$$

Consideriamo allora la differenza $D^0(t) = x^{(1)}(t) - x^{(0)}(t)$, e vediamo come può essere espressa focalizzando l'attenzione all'intervallo $[k, k+1)$:

$$x^{(1)}(t) - x^{(0)}(t) = \begin{cases} \sqrt{2} c_{1,2k} - c_{0,k} & k \leq t < k + \frac{1}{2} \\ \sqrt{2} c_{1,2k+1} - c_{0,k} & k + \frac{1}{2} \leq t < k + 1 \end{cases}$$

Sostituendo la (31), si ha:

$$x^{(1)}(t) - x^{(0)}(t) = \begin{cases} \frac{1}{\sqrt{2}} (c_{1,2k} - c_{1,2k+1}) & k \leq t < k + \frac{1}{2} \\ \frac{1}{\sqrt{2}} (c_{1,2k+1} - c_{1,2k}) & k + \frac{1}{2} \leq t < k + 1 \end{cases}$$

che possiamo riscrivere nel seguente modo:

$$x^{(1)}(t) - x^{(0)}(t) = \begin{cases} d_{0,k} & k \leq t < k + \frac{1}{2} \\ -d_{0,k} & k + \frac{1}{2} \leq t < k + 1 \end{cases}$$

Se definiamo:

$$\psi_{0,k}(t) = \begin{cases} 1 & k \leq t < k + \frac{1}{2} \\ -1 & k + \frac{1}{2} \leq t < k + 1 \end{cases}$$

possiamo ricavare la differenza come:

$$D^{(0)}(t) = x^{(1)}(t) - x^{(0)}(t) = \sum_k d_{0,k} \psi_{0,k}(t), \quad (36)$$

dove le funzioni $\psi_{0,k}(t)$ sono costituite dalle traslazioni della mother wavelet di Haar. Scopriamo, in questo modo, che l'insieme delle funzioni base che genera $D^{(0)}(t)$ sono rappresentate proprio dalle funzioni wavelet ortogonali di Haar. In conclusione la rappresentazione $x^{(1)}(t)$ risulta essere data da:

$$x^{(1)}(t) = x^{(0)}(t) + D^{(0)}(t) = \sum_k c_{0,k} \phi_{0,k}(t) + \sum_k d_{0,k} \psi_{0,k}(t) \quad (37)$$

cioè il segnale è espresso attraverso la rappresentazione a risoluzione inferiore cui vanno aggiunti i dettagli determinati proprio mediante un'espansione su base wavelet. Iterando questo procedimento si ottiene la generica rappresentazione di $x^{(j+1)}(t)$:

$$\begin{aligned} x^{(j+1)}(t) &= x^{(j)}(t) + D^{(j)}(t) = \\ &= x^{(j-1)}(t) + D^{(j-1)}(t) + D^{(j)}(t) = \\ &= \dots\dots\dots = \\ &= x^{(0)}(t) + D^{(0)}(t) + D^{(1)}(t) + \dots + D^{(j)}(t) = \\ &= x^{(0)}(t) + \sum_j \sum_k d_{j,k} \psi_{j,k}(t). \end{aligned} \quad (38)$$

Questa rappresentazione non è ridondante e mostra che l'insieme di funzioni $\psi_{j,k}(t)$ rappresentano una base dello spazio vettoriale W_j (figura 14), cioè

$$W_j = \overline{\text{span}\{\psi_{j,k}(t)\}} \quad (39)$$

Infine, notiamo che la funzione di scaling, detta anche father function, $\phi(t) \equiv \phi_{0,0}(t) \in V_0 \subset V_1$ quindi può essere espressa come combinazione lineare secondo opportuni coefficienti indicati con h_n delle funzioni base che generano V_1 :

$$\phi(t) = \sum_n h_n \phi_{1,n}(t) = \sqrt{2} \sum_n h_n \phi(2t - n). \quad (40)$$

dove

$$h_n = \langle \phi(t), \phi_{1,n}(t) \rangle = \sqrt{2} \int_{-\infty}^{+\infty} \phi(t) \phi(2t - n) dt \quad (41)$$

In modo analogo, per la mother function si ha:

$$\psi(t) = \sum_n g_n \phi_{1,n}(t) = \sqrt{2} \sum_n g_n \phi(2t - n). \quad (42)$$

dove

$$g_n = \langle \psi(t), \phi_{1,n}(t) \rangle = \sqrt{2} \int_{-\infty}^{+\infty} \psi(t) \phi(2t - n) dt \quad (43)$$

La (40) rappresenta la prima equazione dell'analisi multirisoluzione (MRA), mentre la (42) la seconda equazione MRA.

3 Banco di filtri

In questo paragrafo determineremo la struttura mediante banco di filtri che ci permette di calcolare con un algoritmo veloce i coefficienti wavelet per un segnale discreto, ottenendo in questo modo ciò che è nota in letteratura come DWT (Discrete Wavelet Transform). Consideriamo, ancora una volta, un segnale $x(t) \equiv x^{(1)}(t) \in V_1$, per cui risulta:

$$x(t) = \sum_n c_{1,n} \phi_{1,n}(t) = \sum_n c_{0,n} \phi_{0,n}(t) + \sum_n d_{0,n} \psi_{0,n}(t). \quad (44)$$

Vogliamo ricavare la relazione tra i coefficienti $c_{1,n}$ ed i coefficienti $c_{0,n}$ e $d_{0,n}$. Si ha:

$$\begin{aligned} c_{0,n} = \langle x(t), \phi_{0,n}(t) \rangle &= \langle \sum_k c_{1,k} \phi_{1,k}(t), \phi_{0,n}(t) \rangle \\ &= \sum_k c_{1,k} \langle \phi_{1,k}(t), \phi_{0,n}(t) \rangle \end{aligned}$$

dove:

$$\langle \phi_{1,k}(t), \phi_{0,n}(t) \rangle = \int_{-\infty}^{\infty} \sqrt{2} \phi(2t - k) \phi(t - n) dt \quad (45)$$

Effettuando il cambio di variabili $t \rightarrow t + n$ e confrontando la relazione ottenuta con la (41) si ottiene:

$$\langle \phi_{1,k}(t), \phi_{0,n}(t) \rangle = \int_{-\infty}^{\infty} \sqrt{2} \phi(2t - (k - 2n)) \phi(t) dt = h_{k-2n}. \quad (46)$$

In conclusione:

$$c_{0,n} = \sum_k c_{1,k} h_{k-2n} = \sum_k c_{1,k} \tilde{h}_{2n-k} \quad (47)$$

dove abbiamo definito $\tilde{h}_n = h_{-n}$. Pertanto la sequenza di coefficienti $c_{0,n}$ si ottiene semplicemente mediante filtraggio e decimazione per 2 della sequenza $c_{1,n}$. Utilizzando la II equazione MRA e ripetendo i calcoli in maniera analoga, si ricava:

$$d_{0,n} = \sum_k c_{1,k} g_{k-2n} = \sum_k c_{1,k} \tilde{g}_{2n-k} \quad (48)$$

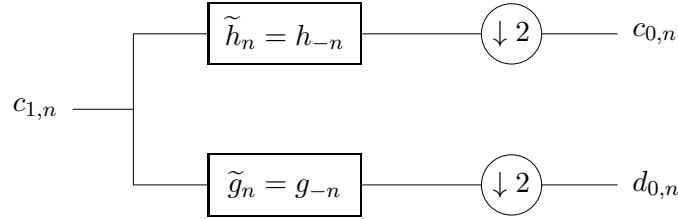


Figura 15: Implementazione con banco di filtri (fase di analisi).

Il legame tra i coefficienti è rappresentato nello schema a blocchi mostrato in figura 15.

A questo punto ripetiamo il ragionamento appena fatto nell'ipotesi in cui il segnale $x(t) \equiv x^{(2)}(t) \in V_2$, per cui risulta:

$$x(t) = \sum_n c_{2,n} \phi_{2,n}(t) = \sum_n c_{1,n} \phi_{1,n}(t) + \sum_n d_{1,n} \psi_{1,n}(t). \quad (49)$$

e vediamo che tipo di relazione c'è tra i coefficienti $c_{2,n}$ ed i coefficienti $c_{1,n}$ e $d_{1,n}$. Si ha:

$$\begin{aligned} c_{1,n} &= \langle x(t), \phi_{1,n}(t) \rangle = \langle \sum_k c_{2,k} \phi_{2,k}(t), \phi_{1,n}(t) \rangle \\ &= \sum_k c_{2,k} \langle \phi_{2,k}(t), \phi_{1,n}(t) \rangle \end{aligned}$$

dove:

$$\langle \phi_{2,k}(t), \phi_{1,n}(t) \rangle = \int_{-\infty}^{\infty} 2 \phi(4t - k) \sqrt{2} \phi(2t - n) dt \quad (50)$$

Effettuando questa volta il cambio di variabili $t \rightarrow (t + n)/2$ e confrontando la relazione sempre con la (41) si ottiene:

$$\langle \phi_{2,k}(t), \phi_{1,n}(t) \rangle = \int_{-\infty}^{\infty} \sqrt{2} \phi(2t - (k - 2n)) \phi(t) dt = h_{k-2n}. \quad (51)$$

Scopriamo così che anche modificando la scala non varia il legame tra i coefficienti:

$$c_{1,n} = \sum_k c_{2,k} h_{k-2n} = \sum_k c_{2,k} \tilde{h}_{2n-k} \quad (52)$$

Si ottiene allora lo schema di decomposizione a due livelli riportato in figura 16. Ovviamente lo schema può essere iterato fino al numero di livelli di decomposizione desiderati.

Vediamo, adesso, quali sono i passi da seguire in fase di ricostruzione. Vogliamo cioè ricavare i coefficienti $c_{1,n}$ a partire da $c_{0,n}$ e $d_{0,n}$ nell'ipotesi in cui $x(t) \in V_1$. Ricordando la (44), si ha:

$$\begin{aligned} c_{1,n} &= \langle x(t), \phi_{1,n}(t) \rangle = \langle \sum_k c_{0,k} \phi_{0,k}(t) + \sum_k d_{0,k} \psi_{0,k}(t), \phi_{1,n}(t) \rangle = \\ &= \sum_k c_{0,k} \langle \phi_{0,k}(t), \phi_{1,n}(t) \rangle + \sum_k d_{0,k} \langle \psi_{0,k}(t), \phi_{1,n}(t) \rangle \end{aligned} \quad (53)$$

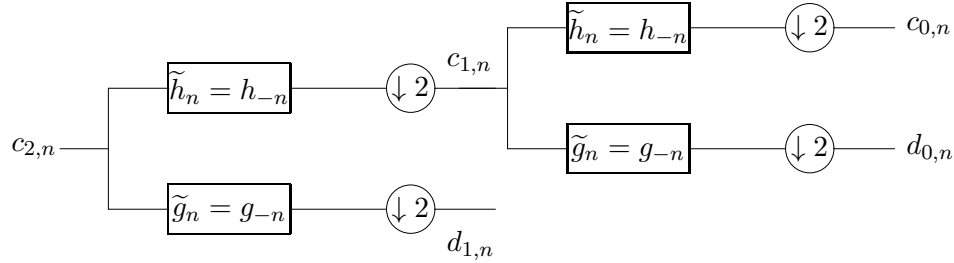


Figura 16: Decomposizione wavelet su due livelli.

dove:

$$\langle \phi_{0,k}(t), \phi_{1,n}(t) \rangle = \int_{-\infty}^{+\infty} \phi(t-k) \sqrt{2} \phi(2t-n) dt = h_{n-2k}$$

e

$$\langle \psi_{0,k}(t), \phi_{1,n}(t) \rangle = \int_{-\infty}^{+\infty} \psi(t-k) \sqrt{2} \phi(2t-n) dt = g_{n-2k}$$

In conclusione:

$$c_{1,n} = \sum_k c_{0,k} h_{n-2k} + \sum_k d_{0,k} g_{n-2k} \quad (54)$$

Questa relazione mostra che è necessario realizzare un'espansione per 2 seguita da un filtraggio, i cui coefficienti sono proprio quelli determinati dalle due equazioni MRA (figura 17). Lo schema può essere iterato così come già visto in fase di analisi all'aumentare del numero di livelli di decomposizione. Chiaramente è necessario imporre un vincolo alla lunghezza del segnale che deve essere una potenza di 2.

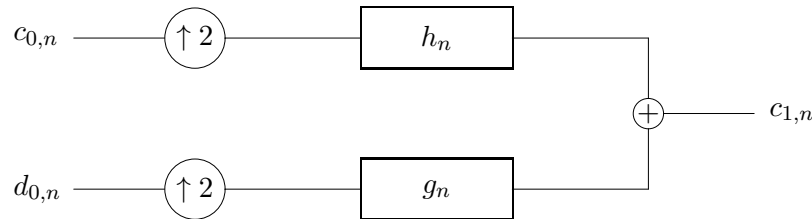


Figura 17: Implementazione con banco di filtri (fase di sintesi).

3.1 Wavelet di Haar

Vediamo come fare per determinare i coefficienti dei filtri nel caso della trasformata Wavelet di Haar. Ricordiamo che la mother function e la father function sono definite come:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{altrimenti} \end{cases} \quad \phi(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{altrimenti} \end{cases}$$

Inoltre, per la (41) risulta:

$$h_n = \langle \phi(t), \phi_{1,n}(t) \rangle = \int_{-\infty}^{+\infty} \sqrt{2} \phi(t) \phi(2t - n) dt$$

Quindi i coefficienti del filtro passa-basso sono (a meno di un ribaltamento):

$$\begin{aligned} h_0 &= \int_{-\infty}^{+\infty} \sqrt{2} \phi(t) \phi(2t) dt = \sqrt{2} \int_0^{1/2} dt = \frac{1}{\sqrt{2}} \\ h_1 &= \int_{-\infty}^{+\infty} \sqrt{2} \phi(t) \phi(2t - 1) dt = \sqrt{2} \int_{1/2}^1 dt = \frac{1}{\sqrt{2}} \end{aligned}$$

Il filtro passa-basso ha quindi lunghezza 2 ed è un filtro simmetrico. Utilizzando invece la (43), si trova che i coefficienti del filtro passa-alto sono:

$$\begin{aligned} g_0 &= \int_{-\infty}^{+\infty} \sqrt{2} \psi(t) \phi(2t) dt = \sqrt{2} \int_0^{1/2} dt = \frac{1}{\sqrt{2}} \\ g_1 &= \int_{-\infty}^{+\infty} \sqrt{2} \psi(t) \phi(2t - 1) dt = -\sqrt{2} \int_{1/2}^1 dt = -\frac{1}{\sqrt{2}} \end{aligned}$$

In realtà i coefficienti del filtro passa-basso e passa-alto sono legati tra loro dalla seguente relazione:

$$g_n = (-1)^n h_{L-1-n} \quad (55)$$

dove L è la lunghezza del filtro. Quindi una volta determinati i coefficienti h_n , basta usare la (55) per trovare i coefficienti g_n . Ricordiamo poi che per la (47) risulta:

$$c_{0,n} = \sum_k c_{1,k} h_{k-2n} = \sum_m c_{1,m+2n} h_m \quad (56)$$

$$= c_{1,2n} h_0 + c_{1,2n+1} h_1 = \frac{c_{1,2n} + c_{1,2n+1}}{\sqrt{2}} \quad (57)$$

Nella (56) è stato fatto il cambio di variabili $m = k - 2n$. Ritroviamo così nuovamente la relazione (31). Analogamente per la (48) si ha:

$$d_{0,n} = \sum_k d_{1,k} g_{k-2n} = \sum_m d_{1,m+2n} g_m \quad (58)$$

$$= d_{1,2n} g_0 + d_{1,2n+1} g_1 = \frac{c_{1,2n} - c_{1,2n+1}}{\sqrt{2}} \quad (59)$$

Esempio: Consideriamo la seguente sequenza monodimensionale:

$$[4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 12, 12, 12, 12, 12]$$

ed applichiamo la wavelet di Haar usando le relazioni (57) e (59). Essendo la lunghezza del segnale pari a $16 = 2^4$ possiamo applicare fino a 4 livelli di decomposizione, come mostrato di seguito:

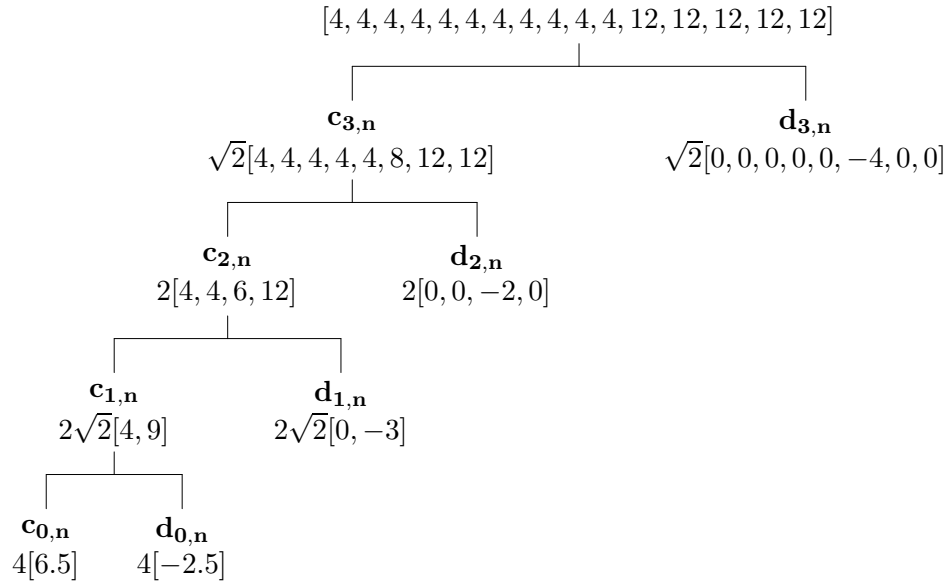


Figura 18: Esempio di decomposizione diadica.

La sequenza trasformata è quindi:

$$[26 | -10 | 0, -8.48 | 0, 0, -4, 0, 0 | 0, 0, 0, 0, -5.65, 0, 0]$$

Questo esempio mostra chiaramente che la trasformata wavelet fornisce una rappresentazione “sparsa” del segnale, in cui sono cioè presenti un gran numero di zeri. Il vantaggio di tale rappresentazione è dato dal fatto che, normalmente, molti dei coefficienti ottenuti dall’operazione di differenza sono nulli, e quelli che assumono valori significativi si trovano soltanto in corrispondenza dei “salti” del segnale originario. In questo senso la rappresentazione del segnale è più compatta, ma non per questo meno significativa, dal momento che non si è persa traccia delle discontinuità. Questa caratteristica rende la trasformata wavelet uno strumento ideale per la compressione dei segnali non stazionari.

3.2 Scelta dei coefficienti dei filtri

Invece di utilizzare le (41) e (43) per determinare i coefficienti dei filtri, in cui si necessita della conoscenza della mother function o della father function, si può dimostrare, a partire dalle

equazioni dell'analisi MRA e sfruttando l'ortogonalità delle funzioni di scaling, che sono valide le seguenti relazioni per decomposizioni wavelet ortogonali:

$$\sum_k h_k = \sqrt{2} \quad (60)$$

$$\sum_k h_k h_{k-2m} = \delta_m \quad (61)$$

Imponiamo allora queste condizioni e determiniamo i coefficienti h_n per diversi valori della lunghezza del filtro.

- Per ottenere filtri di lunghezza $L = 2$, bisogna imporre le seguenti equazioni:

$$\begin{cases} h_0 + h_1 = \sqrt{2} \\ h_0^2 + h_1^2 = 1 \end{cases}$$

Da cui si ottengono proprio i coefficienti del filtro di Haar.

- Per $L = 3$ bisogna che risulti:

$$\begin{cases} h_0 + h_1 + h_2 = \sqrt{2} \\ h_0^2 + h_1^2 + h_2^2 = 1 \\ h_0 h_2 = 0 \end{cases}$$

D'altra parte l'ultima equazione è verificata se e solo se $h_0 = 0$ oppure $h_2 = 0$, questo implica che solo due coefficienti sono diversi da zero, per cui riotteniamo la wavelet di Haar. In effetti, questa situazione si presenta ogni qualvolta si vogliono determinare filtri wavelet ortogonali di lunghezza dispari.

- Per $L = 4$ le condizioni (60) e (61) non sono sufficienti a determinare univocamente i coefficienti h_n , infatti si ottengono 3 equazioni in 4 incognite:

$$\begin{cases} h_0 + h_1 + h_2 + h_3 = \sqrt{2} \\ h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1 \\ h_0 h_2 + h_1 h_3 = 0 \end{cases}$$

E' presente un grado di libertà che si può sfruttare per imporre il vincolo più appropriato. Per esempio, se l'obiettivo è quello di ottenere una rappresentazione sparsa, si può imporre che la wavelet presenti il maggior numero di *momenti evanescenti*, vale a dire che:

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0 \quad 0 < k < p \quad (62)$$

in questo modo se il segnale si potesse modellare come un polinomio di ordine k i coefficienti trasformati sarebbero tutti nulli (si noti che per $k = 0$ la relazione è sempre verificata dalla mother wavelet). Se imponiamo che la (62) sia valida per $k = 1$, si dice che la wavelet presenta un solo momento evanescente. Nel discreto bisogna imporre:

$$\sum_n n g_n = g_1 + 2g_2 + 3g_3 = 0$$

ricordando che vale la (55), questo equivale anche ad imporre:

$$-h_2 + 2h_1 - 3h_0 = 0$$

che rappresenta la quarta condizione che ci permette di risolvere il sistema, ottenendo i coefficienti di lunghezza 4, che portano il nome della matematica francese che li ha scoperti (*Daubechies*):

$$h_0 = \frac{1 + \sqrt{3}}{4\sqrt{2}}, \quad h_1 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, \quad h_2 = \frac{3 - \sqrt{3}}{4\sqrt{2}}, \quad h_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}}$$

Aumentando la lunghezza del filtro è possibile imporre che la wavelet presenti un numero maggiore di momenti evanescenti per cui la rappresentazione risulta essere più sparsa, tuttavia si perde la localizzazione temporale.

3.3 Filtraggio di segnali di durata finita

Per implementare la trasformata wavelet è necessario effettuare filtri passa-basso e passa-alto, vale a dire bisogna realizzare la convoluzione tra sequenze. Questa operazione diventa delicata quando si ha a che fare con segnali di durata finita, perché richiede la conoscenza dei campioni del segnale in punti in cui non è definito. Quindi affinché la trasformata wavelet continui a godere della proprietà di perfetta ricostruzione, è necessario assegnare dei valori al segnale al di fuori del segmento in cui è definito. Data una sequenza di dati di lunghezza finita, si può procedere in tre modi:

1. estensione mediante zero padding;
2. estensione periodica;
3. estensione simmetrica.

Nel primo caso si effettua l'estensione assegnando valori nulli all'esterno del segmento di lunghezza finita. In questo modo si introducono delle discontinuità nella sequenza estesa, che generano componenti alle alte frequenze non realmente presenti nella sequenza originaria. Anche nel secondo caso si possono determinare delle discontinuità nella sequenza estesa, tuttavia si può dimostrare che se si utilizzano espansioni wavelet ortogonali (come visto sinora) è garantita la perfetta ricostruzione.

Per evitare discontinuità spurie si può effettuare l'estensione simmetrica della sequenza originaria, solo che in tal caso è necessario utilizzare filtri wavelet *simmetrici* o *antisimmetrici* (filtri a fase lineare). Poiché l'unica base wavelet ortogonale simmetrica è la wavelet di Haar, bisogna ricorrere a basi biortogonali, che tra l'altro garantiscono risultati migliori di quelle ortogonali nell'ambito della compressione dati, come accade per la wavelet di Daubechies 9/7.

4 Trasformata Wavelet bidimensionale

La trasformata wavelet può essere facilmente generalizzata al caso bidimensionale, quindi alle applicazioni che riguardano l'elaborazione di immagini. La trattazione analitica, fatta per la wavelet nel caso monodimensionale, si estende in modo analogo al caso bidimensionale e si dimostra che la father function e la mother function sono separabili, per cui è possibile operare prima lungo le righe e poi lungo le colonne del segnale.

In figura 19 è riportato lo schema a blocchi per la decomposizione wavelet bidimensionale. La trasformata viene applicata prima lungo le righe, determinando nel ramo superiore i coefficienti ottenuti dal filtraggio passa-basso (L) e in quello inferiore i coefficienti ottenuti dal filtraggio passa-alto (H). In seguito viene applicata la stessa procedura a tutte le colonne. Si determina in questo modo una matrice composta da quattro bande (LL, HL, LH, HH) (figura 20): la banda LL contiene una versione grossolana dell'immagine, la banda HL evidenzia le discontinuità lungo le righe, la banda LH quelle lungo le colonne, infine la HH conterrà le discontinuità presenti lungo la direzione obliqua, come si può vedere in figura 21, dove è mostrato un esempio di decomposizione wavelet di un'immagine naturale su uno e due livelli di decomposizione.

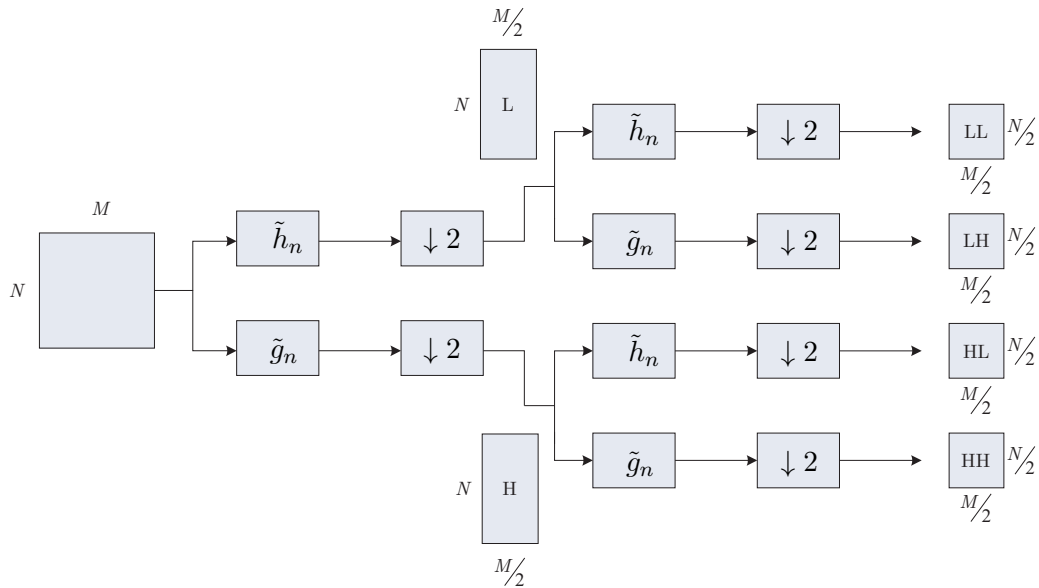


Figura 19: Implementazione wavelet 2D mediante banco di filtri.

5 Strategia di codifica: EZW

EZW (Embedded Zero Tree Wavelet) è una strategia di codifica dei coefficienti ottenuti mediante decomposizione wavelet, introdotta nel 1993 da Jerry Shapiro. L'idea di base di EZW sta nella capacità di riconoscere efficientemente strutture molto grandi che non vanno codificate e che sono etichettate come alberi di zeri (*zerotree*). Ciò è possibile grazie al fatto che i coefficienti wavelet,

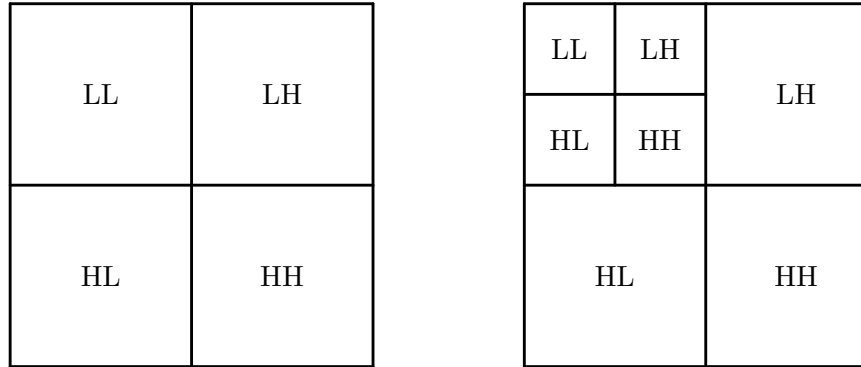


Figura 20: Esempio di decomposizione diadica.

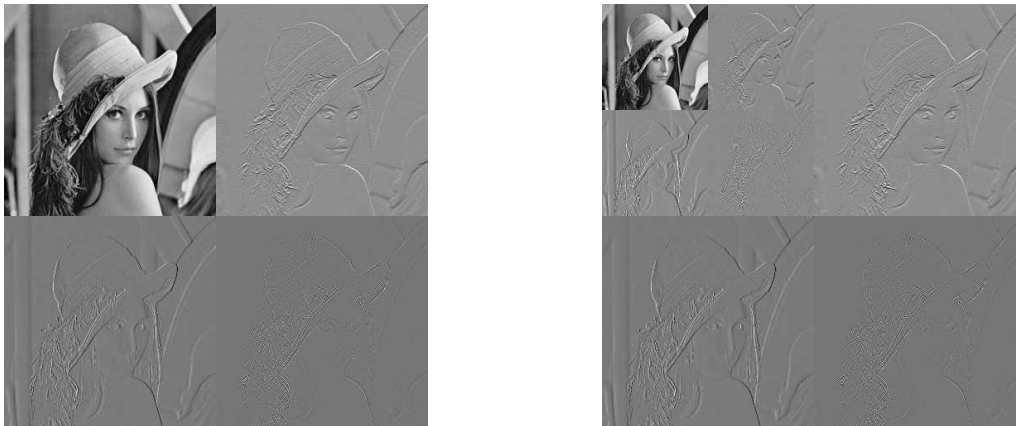


Figura 21: Trasformata Wavelet con uno e due livelli di decomposizione.

nelle diverse sottobande, rappresentano la stessa posizione spaziale dell'immagine e quindi sono fortemente correlati. Considerando la decomposizione mostrata in figura 22, si nota infatti che il coefficiente a , nella banda LL, rappresenta la stessa locazione spaziale dei coefficienti a_1 , nella banda HL, a_2 , nella banda LH, ed a_3 , nella banda HH. A sua volta, a_1 rappresenta la stessa locazione spaziale dei coefficienti a_{11} , a_{12} , a_{13} e a_{14} , ciascuno di questi coefficienti ha quattro discendenti, e così via, fino a formare una struttura ad albero.

Abbiamo già visto che la decomposizione wavelet tende a compattare la maggior parte dell'energia nelle bande a risoluzione più bassa, cioè, nella maggior parte dei casi, i coefficienti vicini alla radice dell'albero assumono valori maggiori rispetto ai discendenti. Questo vuol dire che se un coefficiente ha ampiezza minore di una data soglia T , con elevata probabilità tutti i suoi discendenti avranno ampiezze minori di quella stessa soglia. Per comprendere come opera l'algoritmo introduciamo la terminologia introdotta da Shapiro:

- un coefficiente è detto *significativo* se ha ampiezza maggiore di T ;

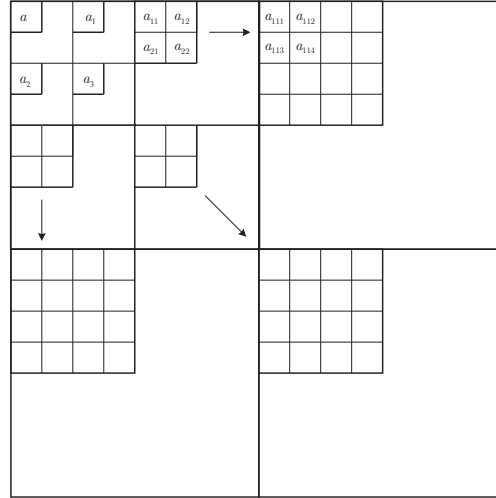


Figura 22: Struttura ad albero di EZW.

- un coefficiente è detto *radice zerotree* se ha ampiezza minore di T , e tutti i suoi discendenti hanno ampiezza minore di T ;
- un coefficiente è detto *isolated zero* se ha ampiezza minore di T , ma esiste almeno un discendente con ampiezza maggiore di T , .

La soglia assume un valore iniziale T_0 pari a:

$$T_0 = 2^{\lfloor \log_2 c_{max} \rfloor}, \quad (63)$$

con c_{max} pari al valore del coefficiente più grande. L'algoritmo prevede essenzialmente due passi:

1. il *passo dominante*, che definisce la mappa di coefficienti significativi rispetto alla soglia corrente T_i . I coefficienti riconosciuti come significativi nel passo dominante si trovano nell'intervallo $[T_i, 2T_i)$ e sono rappresentati da un valore di ricostruzione pari al valore medio di tale intervallo $3T_i/2$. In effetti il passo dominante determina i bit più significativi nella rappresentazione binaria dei coefficienti, rispetto alla soglia corrente. Dopo il primo passo dominante i coefficienti vengono divisi in due liste: la lista dominante, contenente tutti i coefficienti non significativi nel passo precedente, e la lista subordinata, contenente tutti i coefficienti significativi nel passo precedente;
2. il *passo subordinato*, è un passo di raffinamento, nel senso che determina il valore dei nuovi bit più significativi, nella rappresentazione binaria dei coefficienti della lista subordinata.

Ad ogni passo la soglia viene dimezzata, e, dato il valore di T_i , ad ognuno dei coefficienti viene assegnata una tra le seguenti etichette:

- *positivo significativo* (ps, positive significant);
- *negativo significativo* (ns, negative significant);

- *radice zerotree* (zr, zerotree root);
- *zero isolato* (iz, isolated zero).

A questo punto descriviamo l'algoritmo in maniera dettagliata:

1. Inizializzazione

Si inseriscono tutti i coefficienti nella lista dominante, mediante l'ordine di scansione riportato in figura 23, e si definisce la soglia iniziale, come dalla relazione (63).

2. Passo Dominante

Si effettua la scansione della lista dominante e si associa, ad ogni coefficiente, la relativa etichetta ed il relativo valore di ricostruzione. I coefficienti che discendono da uno zerotree non vengono codificati. I coefficienti significativi vengono spostati nella lista subordinata ed i relativi valori nella lista dominante vengono posti a zero.

3. Passo Subordinato

Si effettua la scansione della lista subordinata, e al generico coefficiente c_i di questa lista si associa uno dei seguenti valori:

$$\begin{cases} 1 & \text{se } c_i \geq \text{valore di ricostruzione} \\ 0 & \text{altrimenti} \end{cases}$$

4. Calcolo della soglia

Si effettua il calcolo della nuova soglia mediante la seguente relazione

$$T_i = \frac{T_{i-1}}{2},$$

e, se non è stato raggiunto il bit-rate voluto, si torna al passo 2.

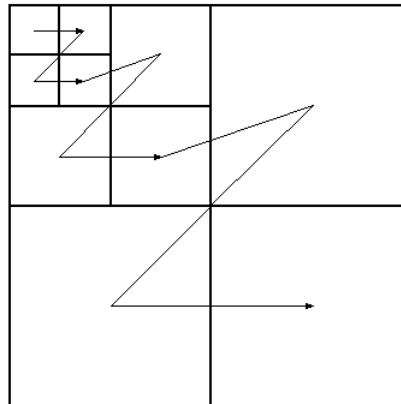


Figura 23: Scansione effettuata nel passo dominante di EZW.

Notate come la codifica della mappa dei coefficienti significativi si possa interpretare come una quantizzazione a tre livelli (con restituzione dello zero). Infatti, per i coefficienti significativi si assegna il valore ricostruito $1.5T_i$ oppure $-1.5T_i$, in base al segno, mentre per quelli non significativi il valore zero. Si tenga poi presente che, avendo realizzato una codifica per bit-plane, è possibile interrompere il processo di codifica in qualsiasi istante, garantendo comunque in fase di decodifica la migliore ricostruzione possibile (scalabilità in SNR).

Esempio: Consideriamo una matrice di coefficienti wavelet 16×16 con 3 livelli di decomposizione (figura 24) e applichiamo l'algoritmo EZW. La soglia iniziale si sceglie pari a

$$T_0 = 2^{\lfloor \log_2 63 \rfloor} = 32$$

In questo modo i coefficienti più grandi si trovano nell'intervallo $(32, 64)$. Si comincia quindi a scandire il coefficiente della banda LL3: essendo la sua ampiezza maggiore della soglia e positiva, esso viene etichettato come *significant positive*, SP, e ricostruito con il valore medio compreso nell'intervallo $(32, 64)$, pari a 48. Si prosegue quindi nella banda HL3 dove si trova il coefficiente -34 , che viene quindi etichettato come *significant negative*, SN, e ricostruito con il valore -48 . Nella banda LH3, invece, è presente il valore -31 , che risulta essere inferiore alla soglia, per questo motivo è necessario esaminare i suoi discendenti: se sono tutti minori della soglia il coefficiente sarà etichettato come *zerotree root*, ZR, altrimenti come *isolated zero*, IZ. In questo caso, data la presenza del coefficiente 47 nella banda LH1 il valore -31 sarà etichettato come IZ. Invece, è facile riconoscere che il coefficiente 23 nella banda HH3 è uno zero-tree, per questo motivo tutti i suoi discendenti in questa prima fase dell'algoritmo non saranno analizzati. Proseguendo in questo modo si costruisce la tabella mostrata a pag.28.

Quello esaminato è il primo passo dell'algoritmo (passo dominante), in cui si codifica la cosiddetta *mappa di significatività* dei coefficienti; successivamente c'è il passo subordinato (detto anche di *refinement*) nel quale viene aggiunto un altro bit di codifica ai coefficienti significativi allo scopo di migliorarne la precisione. In questo modo la matrice dei coefficienti wavelet ricostruiti diventa quella di figura 25, in cui è possibile notare come l'informazione sul dettaglio

63	-34	49	10	7	13	-12	7
-31	23	14	-13	3	4	6	-1
15	14	3	-12	5	-7	3	9
-9	-7	-14	8	4	-2	3	2
-5	9	-1	47	4	6	-2	2
3	0	-3	2	3	-2	0	4
2	-3	6	-4	3	6	3	6
5	11	5	6	0	3	-4	4

Figura 24: Matrice dei coefficienti wavelet con 3 livelli di decomposizione.

contenuta nella banda LH1 non è stata persa. Se si hanno ancora risorse disponibili l'algoritmo ripete i due passi dimezzando via via la soglia. E' chiaro che i coefficienti etichettati come significativi al primo passo non dovranno più essere esaminati nei passi successivi e saranno soggetti solo alla fase di refinement.

sottobanda	valore	simbolo	valore ric.
LL3	63	SP	48
HL3	-34	SN	-48
LH3	-31	IZ	0
HH3	23	ZR	0
HL2	49	SP	48
HL2	10	ZR	0
HL2	14	ZR	0
HL2	-13	ZR	0
LH2	15	ZR	0
LH2	14	IZ	0
LH2	-9	ZR	0
LH2	-7	ZR	0
HL1	7	IZ	0
HL1	13	IZ	0
HL1	3	IZ	0
HL1	4	IZ	0
LH1	-1	IZ	0
LH1	47	SP	48
LH1	-3	IZ	0
LH1	2	IZ	0

56	-40	56	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	40	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 25: Matrice ricostruita dopo il passo dominante e subordinato.

Riferimenti bibliografici

- [1] A.Bovik, *The essential guide to Image Processing*, Academic Press, 2009.
- [2] I.Daubechies, *Ten lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics.
- [3] S.Mallat, *A Wavelet tour of signal processing*, Academic Press, Second Edition.
- [4] K.Sayood: *Introduction to Data Compression*, Second Edition, Morgan Kaufmann 2000.
- [5] G.Strang and T.Nguyen, *Wavelets and filter banks*, Wellesley-Cambridge Press.
- [6] M.Vetterli and J.Kovacevic, *Wavelet and filter banks*, Prentice Hall.
- [7] J.M.Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. on Signal Processing*, pp.3445-3462, Dicembre 1993.
- [8] A.Said, W.A.Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circuits Systems Video Technology*, pp.243-250, Giugno 1996.
- [9] B.E.Usevitch, "A tutorial on modern lossy wavelet image compression: foundations of JPEG2000", *IEEE Signal Processing Magazine*, pp.22-35, Settembre 2001.