



ITALIAN NATIONAL RESEARCH COUNCIL
"NELLO CARRARA" INSTITUTE FOR APPLIED PHYSICS
CNR FLORENCE RESEARCH AREA
Italy

TECHNICAL, SCIENTIFIC AND RESEARCH REPORTS

Vol. 1 - n. 55-1 (2009)

Bruno Aiazzi

**Analisi di Immagini SAR: stima di parametri
spaziali e temporali del rumore di speckle
per il filtraggio multitemporale
multirisoluzione
e di parametri tessiturali utili per la
classificazione**

CNR-IFAC-TR-10/000

ISSN 2035-5831

Analisi di immagini SAR: Stima di parametri spaziali e temporali del rumore di speckle per il filtraggio multitemporale multirisoluzione e di parametri tessiturali utili per la classificazione; studio di un metodo di compressione near-lossless e virtually-lossless di immagini SAR che non alteri i parametri di classificazione.*

Bruno Aiazzi

Ottobre 2000

1 Introduzione

La possibilità di usufruire di sensori remoti posti su piattaforme quali satelliti ed aerei ha aumentato molto la disponibilità di immagini della superficie terrestre che possono essere utilizzate per vari scopi quali, ad esempio, il monitoraggio delle risorse ambientali e lo studio di siti di interesse archeologico.

In questo campo possono essere usati sistemi di tipo incoerente che sfruttano la radiazione solare riflessa dal terreno oppure le emissioni termiche della superficie terrestre e sensori coerenti che provvedono essi stessi alla propria

*Relazione di assegno per la collaborazione ad attività di ricerca (Progetto ASI Modellistica di emissione e backscattering di microonde da superfici naturali: indagini sperimentali e teoriche).

illuminazione, con il vantaggio di poter essere utilizzati anche di notte o in condizioni meteorologiche avverse, e di poter rilevare informazioni relative a scatteratori posti a varie profondità nel terreno (e che quindi possono essere invisibili per un sensore ottico) inviando segnali a frequenze diverse.

In questa categoria rientrano le tecnologie basate sui *radar ad apertura sintetica*, o *SAR*, che riproducono il comportamento di un'antenna di grandi dimensioni grazie alla velocità della piattaforma rendendo possibile un aumento della risoluzione in azimuth che entro certi limiti risulta anche indipendente dalla distanza dalla superficie terrestre [1].

Lo svantaggio presente nei SAR è che la generazione dell'immagine avviene processando in modo coerente i segnali riflessi da molti scatteratori indipendenti elementari presenti in una cella di risoluzione, che possono trovarsi a distanze diverse dal radar a seconda della rugosità della superficie (modello dello speckle *pienamente sviluppato*). Ciò fa sì che le onde ricevute non sono più coerenti in fase e quindi si ottiene un segnale forte o debole, a seconda che le riflessioni si sommino costruttivamente oppure no. Questo effetto causa una variazione di intensità tra pixel adiacenti anche appartenenti a zone omogenee, che si manifesta come un rumore granulare *signal-dependent* chiamato *speckle* [2], complicando in modo notevole l'estrazione di informazioni utili dall'immagine [3]. Se con x e y indichiamo la parte reale ed immaginaria del segnale complesso risultante dalla somma vettoriale delle onde riflesse dai vari scatteratori si possono ottenere due tipi di immagine:

- in *intensità* considerando $I = x^2 + y^2$, in cui lo speckle risulta avere una distribuzione esponenziale con rapporto tra standard deviation e media della distribuzione, detto coefficiente di variazione teorico, $C_v = 1$.
- in *ampiezza* considerando $A = \sqrt{I}$, in cui lo speckle risulta avere una distribuzione di *Rayleigh* con $C_v = 0.5227$.

2 Stima dei parametri di rumore

2.1 Parametri del rumore nelle zone omogenee

Essendo costante nella distribuzione dello speckle il rapporto tra media e standard deviation, si può assumere per lo speckle un modello moltiplicativo

$$g(m, n) = f(m, n) \cdot u(m, n) \tag{1}$$

dove il termine $g(m, n)$ rappresenta l'immagine osservata, $f(m, n)$ l'immagine priva di rumore, mentre $u(m, n)$ è il rumore di speckle con media unitaria, varianza σ_u^2 , e spazialmente correlato con coefficienti di correlazione lungo le direzioni x e y dati da ρ_x and ρ_y . Notiamo fra l'altro che il modello moltiplicativo (1) può essere visto come un caso particolare con $\gamma = 1$ di un modello più generale di rumore *signal-dependent* dato da

$$g(m, n) = f(m, n) + f^\gamma(m, n) \cdot v(m, n) \quad (2)$$

in cui γ è una costante positiva che tiene conto della dipendenza del rumore dal segnale, mentre $v(m, n) \equiv u(m, n) - 1$, e $f^\gamma(m, n) \cdot v(m, n)$ è un rumore a media nulla dipendente dal segnale e additivo.

Calcoliamo la varianza del modello moltiplicativo (1):

$$\sigma_g^2(m, n) = \sigma_f^2(m, n) \cdot (1 + \sigma_u^2) + \mu_f^2(m, n) \cdot \sigma_u^2 \quad (3)$$

in cui $\mu_f(m, n) \equiv E[f(m, n)]$ indica la media spazio-variante di f . Ovviamente $\mu_f(m, n) = \mu_g(m, n) \equiv E[g(m, n)]$ dalla (1). Dei due termini a destra di (3), il primo prevale sulle aree tessiturate e tende a scomparire in quelle omogenee, in cui si ha

$$\sigma_g(m, n) = \sigma_u \cdot \mu_g(m, n) \quad (4)$$

mentre nelle aree tessiturate vale

$$\sigma_g(m, n) \approx \sigma_f(m, n) \cdot \sqrt{1 + \sigma_u^2} \quad (5)$$

L'equazione (4) fornisce una stima di σ_u , ossia $\hat{\sigma}_u$, come il coefficiente di regressione di σ_g versus μ_g , stimato in zone omogenee, mentre l'equazione (5) mostra che in zone altamente tessiturate la standard deviation locale tende a diventare indipendente dalla media. In più, la (5) non risulta completamente vera laddove il modello di rumore non è più moltiplicativo, a causa del fatto che lo speckle non è pienamente sviluppato.

Da queste considerazioni deriva che è possibile sfruttare per la stima della standard deviation σ_u dello speckle metodi basati sugli scatterplot *media vs. standard deviation*. Infatti, gli scatterplot relativi all'intera immagine possono essere visti come la PDF congiunta delle variabili aleatorie media e standard deviation stimate attraverso una maschera spaziale centrata sul pixel corrente. In assenza di zone tessiturate e assumendo lo speckle stazionario, ossia se l'immagine è composta solamente da zone rumorose aventi

statistiche *spazialmente omogenee*, la PDF sarà data dalla sovrapposizione di tante distribuzioni unimodali quante sono le zone omogenee. Ogni distribuzione viene ad essere una gaussiana 2D attorno al suo valore atteso, e tutti i valori attesi risultano allineati lungo una retta passante per l'origine avente pendenza pari al parametro desiderato. Se quindi all'interno dello scatterplot si riesce a riconoscere le distribuzioni allineate che corrispondono alle zone omogenee, ne deriva che può essere stimato σ_u .

In modo analogo, se calcoliamo la covarianza spazio-variante di lunghezza 1 lungo m and n , considerando solo i contributi presenti nelle zone omogenee otteniamo

$$\begin{aligned}
 C_g(m, n; 0, 1) &= \rho_x \cdot \sigma_u^2 \cdot \mu_f(m, n) \cdot \mu_f(m, n + 1) \\
 &= \rho_x \cdot \sigma_g(m, n) \cdot \sigma_g(m, n + 1) \\
 C_g(m, n; 1, 0) &= \rho_y \cdot \sigma_u^2 \cdot \mu_f(m, n) \cdot \mu_f(m + 1, n) \\
 &= \rho_y \cdot \sigma_g(m, n) \cdot \sigma_g(m + 1, n)
 \end{aligned} \tag{6}$$

Quindi, ρ_x and ρ_y possono essere stimati a partire da quei punti, giacenti sullo scatterplot *covarianza-varianza*, corrispondenti ad aree con coefficiente di backscattering costante.

2.2 Determinazione automatica delle zone omogenee

Vari metodi basati sugli scatterplot sono stati utilizzati per la determinazione dei parametri dello speckle [3, 4]. Il punto cruciale è però come correttamente identificare le zone omogenee attraverso un procedimento automatico. Un sistema molto potente consiste nel supporre che i punti aventi backscatter costante originano nello scatterplot dei *cluster* ben identificabili [5]. I cluster risultano allineati lungo una retta passante attraverso l'origine e con pendenza pari a σ_u , nel caso della standard deviation sulla media, ed uguale a ρ_x oppure a ρ_y , nel caso della covarianza sulla varianza. Al contrario, la presenza di point targets, contorni e tessiture originano punti sparsi sullo scatterplot.

Per determinare i punti dello scatterplot, vengono calcolate le statistiche locali dell' $M \times N$ immagine rumorosa all'interno di una finestra mobile di dimensione $(2p + 1) \times (2p + 1)$ per stimare le statistiche d'insieme spazio-varianti:

- media $\hat{\mu}_g(m, n)$

$$\hat{\mu}_g(m, n) = \frac{1}{(2p+1)^2} \sum_{k=-p}^p \sum_{l=-p}^p g(m+k, n+l) \quad (7)$$

- deviazione RMS dalla media $\hat{\sigma}_g(m, n) = \sqrt{\hat{\sigma}_g^2(m, n)}$, con

$$\hat{\sigma}_g^2(m, n) = \frac{1}{(2p+1)^2 - 1} \sum_{k=-p}^p \sum_{l=-p}^p [g(m+k, n+l) - \hat{\mu}_g(m, n)]^2 \quad (8)$$

- cross-deviazioni dalla media $\hat{C}_g(m, n; 0, 1)$ e $\hat{C}_g(m, n; 1, 0)$

$$\begin{aligned} \hat{C}_g(m, n; 0, 1) &= \frac{1}{(2p+1)^2 - 1} \sum_{k=-p}^p \sum_{l=-p}^p \\ &\quad [g(m+k, n+l) - \hat{\mu}_g(m, n)] \\ &\quad \cdot [g(m+k, n+l+1) - \hat{\mu}_g(m, n+1)] \\ \hat{C}_g(m, n; 1, 0) &= \frac{1}{(2p+1)^2 - 1} \sum_{k=-p}^p \sum_{l=-p}^p \\ &\quad [g(m+k, n+l) - \hat{\mu}_g(m, n)] \\ &\quad \cdot [g(m+k+1, n+l) - \hat{\mu}_g(m+1, n)] \end{aligned} \quad (9)$$

A questo punto si hanno a disposizione $(M - 2p) \times (N - 2p)$ punti per $\hat{\mu}_g(m, n)$ e $\hat{\sigma}_g(m, n)$, mentre si ottengono $(M - 2p) \times (N - 2p - 1)$ e $(M - 2p - 1) \times (N - 2p)$ punti per $\hat{C}_g(m, n; 0, 1)$ e $\hat{C}_g(m, n; 1, 0)$, rispettivamente. Notiamo che al contrario di [3, 4] dove i blocchi sono disgiunti producendo così un minor numero di punti, in questo lavoro si calcola un valore per ogni punto dell'immagine, esclusi i bordi, mediante la sovrapposizione delle finestre. In tal modo si ha il vantaggio di catturare in modo migliore l'omogeneità delle varie parti dell'immagine, ottenendo cluster più densi che possono essere individuati meglio dalla procedura.

Si possono così disegnare gli scatterplot di:

- $\hat{\sigma}_g(m, n)$ vs. $\hat{\mu}_g(m, n)$;
- $\hat{C}_g(m, n; 0, 1)$ vs. $\hat{\sigma}_g(m, n) \cdot \hat{\sigma}_g(m, n+1)$;

- $\hat{C}_g(m, n; 1, 0)$ vs. $\hat{\sigma}_g(m, n) \cdot \hat{\sigma}_g(m + 1, n)$.

Gli scatterplot vengono suddivisi in $L \times L$ blocchi rettangolari, che sono ordinati in ordine decrescente di popolazione, ossia di numero di scatterpoint. Si calcola poi una successione di rette di regressione partendo dai punti contenuti nel blocco più popolato e aggiungendo via via i punti degli altri blocchi secondo la successione precedente. La successione dei valori stimati $\{\hat{\sigma}_u(k)\}$, $\{\hat{\rho}_x(k)\}$ e $\{\hat{\rho}_y(k)\}$ approssima il valore vero del parametro dopo un certo numero di termini che dipende dalla percentuale di punti omogenei presente nell'immagine. Gli esperimenti hanno permesso di stabilire criteri di stop che variano a seconda del parametro da stimare [5].

Un'alternativa consiste nell'effettuare un'analisi tridimensionale dello scatterplot, in modo da considerare all'interno del procedimento precedente solo i blocchi corrispondenti alle mode (i massimi locali dello scatterplot), che individuano in modo più esatto i cluster presenti.

2.3 Stima della correlazione temporale dello speckle

La disponibilità di immagini multitemporali della stessa area di studio rende possibile il filtraggio dello speckle traendo vantaggio dalla correlazione temporale del rumore stesso.

A questo fine si opera dapprima una trasformazione reversibile che mantiene le proprietà del modello moltiplicativo, consistente nella media geometrica e nel rapporto tra le immagini, poi successivamente viene stimato il coefficiente temporale ρ_t dello speckle [6] partendo dalla varianza di rumore dei dati trasformati, che può essere ottenuta dal metodo degli scatterplot descritto in precedenza. Successivamente è possibile filtrare le immagini trasformate tenendo conto del valore di ρ_t , per poi ricombinarle attraverso la trasformazione inversa.

Partendo dalle immagini multitemporali $f_0(m, n)$ e $f_1(m, n)$ si ottengono le immagini trasformate della media geometrica e del rapporto $F_0(m, n)$ e $F_1(m, n)$ dalle relazioni

$$\begin{aligned} F_0(m, n) &= \sqrt{f_0(m, n) \cdot f_1(m, n)} \\ F_1(m, n) &= \frac{f_0(m, n)}{f_1(m, n)} \end{aligned} \tag{10}$$

mentre la trasformazione inversa è data da

$$\begin{aligned} f_0(m, n) &= F_0(m, n) \cdot \sqrt{F_1(m, n)} \\ f_1(m, n) &= \frac{F_0(m, n)}{\sqrt{F_1(m, n)}}. \end{aligned} \tag{11}$$

La media geometrica rafforza le componenti temporalmente correlate, ciò che si traduce in un SNR più alto che nelle singole osservazioni. Al contrario, il rapporto rimuove tali componenti evidenziando i cambiamenti stagionali e il rumore. Tale immagine mostra una media approssimativamente costante, in modo che il modello moltiplicativo del rumore conserva validità.

Consideriamo un termine di correlazione nella coppia di osservazioni $s(m, n)$ e un termine di cambiamento $c(m, n)$, entrambi processi spazio varianti aventi valori positivi e quindi media non nulla. Inoltre consideriamo due termini di rumore $u_0(m, n)$ e $u_1(m, n)$, entrambi con media unitaria, varianza σ_u^2 , e correlati fra loro attraverso il coefficiente di correlazione ρ_t , legati in modo moltiplicativo ai segnali *noise-free*:

$$\begin{aligned} f_0(m, n) &= s(m, n) \cdot \sqrt{c(m, n)} \cdot u_0(m, n) \\ f_1(m, n) &= \frac{s(m, n)}{\sqrt{c(m, n)}} \cdot u_1(m, n) \end{aligned} \tag{12}$$

L'equazione (12) definisce un modello moltiplicativo dei cambiamenti temporali. Se (10) è applicata alla (12), le immagini di uscita possono essere scritte come

$$\begin{aligned} F_0(m, n) &= s(m, n) \sqrt{u_0(m, n) \cdot u_1(m, n)} \triangleq s(m, n) \cdot u_G(m, n) \\ F_1(m, n) &= c(m, n) \cdot \frac{u_0(m, n)}{u_1(m, n)} \triangleq c(m, n) \cdot u_R(m, n) \end{aligned} \tag{13}$$

in cui le varianze dei termini di rumore $u_G(m, n)$ and $u_R(m, n)$ possono essere approssimate prendendo gli sviluppi al primo ordine della radice quadrata e del rapporto nella (13):

$$\begin{aligned} \sigma_G^2 &= \frac{\sigma_u^2}{2} \cdot (1 + \rho_t) \\ \sigma_R^2 &= 2 \cdot \sigma_u^2 \cdot (1 - \rho_t) \end{aligned} \tag{14}$$

Prendendo il rapporto fra le radici in (14) si ottiene ρ_t come

$$\rho_t = \frac{4\sigma_G^2 - \sigma_R^2}{4\sigma_G^2 + \sigma_R^2} \quad (15)$$

Il valore di ρ_t può quindi essere stimato dalle varianze di rumore σ_G^2 e σ_R^2 , ottenute operando con la procedura degli scatterplot.

3 Filtraggio dello speckle basato sulle statistiche locali dell'immagine

I metodi comunemente utilizzati per ridurre lo speckle possono essere suddivisi in varie categorie:

- si effettua una media incoerente di N stime dell'immagine ottenute dividendo in N parti la lunghezza d'apertura sintetica (o equivalentemente la banda spettrale disponibile); i vari segmenti sono processati indipendentemente l'uno dall'altro per ottenere un'immagine in intensità oppure in ampiezza, poi le diverse immagini sono sommate assieme per ottenere un'immagine SAR a N -look; in questo modo si riduce la standard deviation dello speckle di un fattore \sqrt{N} , a prezzo però di una perdita in risoluzione di un fattore N [1].
- si combinano assieme immagini SAR ottenute con diverse polarizzazioni e/o diverse bande di frequenza.
- nel caso siano disponibili immagini multitemporali della stessa area si può trarre vantaggio dalla *correlazione temporale* fra le immagini.
- si effettua un filtraggio spaziale del rumore dopo che l'immagine è stata formata basandosi su un modello dello speckle come rumore moltiplicativo [7].

Noi considereremo algoritmi per il filtraggio spaziale attraverso una decomposizione multirisoluzione dell'immagine SAR che fa uso del noto filtro *LLMMSE* di Kuan (Local-Linear, Minimum Mean-Squared Error) [8]. Per ottenere risultati ottimali è necessario considerare anche l'eventuale *correlazione spaziale* dello speckle, dovuta ad eventuali sovracampionamenti effettuati per evitare l'aliasing. Il coefficiente di correlazione e la standard

deviation dello speckle vengono stimati dal metodo basato sugli scatterplot descritto precedentemente.

3.1 Filtro di Kuan LLMMSE

Riprendiamo per lo speckle il modello moltiplicativo dato dalla (1):

$$g(m, n) = f(m, n) \cdot u(m, n) \quad (16)$$

E' noto che la stima MMSE di f nota g è data da $E[f|g]$, che, comunque, richiede di conoscere la PDF dell'immagine *noise-free* almeno al secondo ordine. In alternativa può essere adottata una strategia subottima, che consiste nello sviluppare solo statistiche del primo ordine, più facili da calcolare in un intorno locale del pixel corrente. Questo tipo di approccio è noto come filtro LLMMSE. Procedendo così, la stima ottima di f , ossia \hat{f} , può essere approssimata come:

$$\hat{f}(m, n) = \hat{\mu}_g(m, n) + [g(m, n) - \hat{\mu}_g(m, n)] \times \left(1 - \frac{\sigma_u^2 \cdot \hat{\mu}_g^2(m, n)}{\hat{\sigma}_g^2(m, n)}\right) \cdot \left(\frac{1}{1 + \sigma_u^2}\right) \quad (17)$$

in cui $\hat{\mu}_g(m, n)$ e $\hat{\sigma}_g^2(m, n)$ rappresentano le stime della media locale e della varianza locale di g nella posizione (m, n) , calcolate su intorni di larghezza opportuna, tipicamente 5×5 o 7×7 .

L'approssimazione di considerare le statistiche locali al posto della media e della varianza dell'immagine, entrambe *spazio-varianti*, può essere inadeguata, specialmente in presenza di contorni, dove la capacità di filtraggio di (17) diventa assai scarsa. Questo è il motivo per cui si ottengono risultati migliori con lo schema del filtro di *Lee Refined* [9], che opera considerando le ampiezze e le orientazioni dei gradienti locali [7].

3.2 Lo schema di Lee Refined

Questo filtro fu progettato per superare il problema del rumore presente nei contorni e non risolto in [10], dove, nella formulazione originale, il filtraggio veniva realizzato linearizzando il modello di rumore nell'intorno di $\mu_g(m, n)$. In questo modo, veniva ottenuta una soluzione approssimata ma soddisfacente, identica a (17), tranne il termine $(1 + \sigma_u^2)^{-1}$, il cui contributo può

essere trascurato per immagini multi-look, dove $\sigma_u^2 \ll 1$ [7, 11]. L'algoritmo determina poi l'orientazione dei contorni presenti mediante il calcolo dei gradienti locali in una finestra mobile 7×7 . A questo punto il calcolo delle statistiche locali viene effettuato in finestre non quadrate che tengono conto dell'orientazione del contorno. Se non viene individuato alcun contorno, la stima viene effettuata sull'intera finestra 7×7 . In questo modo il filtraggio viene notevolmente migliorato sulle aree ad alto contrasto, ma può segmentare eccessivamente le aree tessiturate producendo artefatti in tali zone, dovute anche al fatto di lavorare con una finestra di dimensioni fisse 7×7 .

3.3 Il filtro di Meer

Un filtro orientato su uno schema multirisoluzione è stato recentemente proposto da Meer *et al.* [12], considerando intorno locali costituiti da un set di tre finestre mobili concentriche di dimensioni 7×7 , 5×5 , e 3×3 . Il coefficiente di variazione C_v viene calcolato su tutte le finestre e usato come indice di omogeneità, in modo da individuare la finestra più larga che soddisfi un criterio di omogeneità, ottenuto da una sogliatura del C_v . Se tale finestra esiste, l'uscita è costituita dalla media spaziale calcolata su di essa, altrimenti dalla stima LLMSE (17) sulla finestra 3×3 più interna. Questo filtro è ottimo nel preservare i *point target* e i contorni, grazie alla sua adattabilità nel variare le dimensioni della finestra.

4 Filtraggio dello speckle mediante la piramide a rapporto

4.1 Le piramidi gaussiane e laplaciana

La piramide gaussiana (GP) [13] è una rappresentazione a multirisoluzione dell'immagine ottenuta tramite una riduzione ricorsiva, costituita da un filtraggio passa-basso e un sottocampionamento. Sia l'immagine d'ingresso $G_0(m, n) \equiv g(m, n)$, $m = 0, \dots, M - 1$, $n = 0, \dots, N - 1$, $M = p \times 2^K$, $N = q \times 2^K$; con p , q , e K interi positivi. La GP, di cui k identifica il livello corrente e $K > 0$ la radice di dimensioni $p \times q$, risulta definita da

$$G_k(m, n) \triangleq \sum_{i=-L_r}^{L_r} \sum_{j=-L_r}^{L_r} r(i) \times r(j) \cdot G_{k-1}(2m + i, 2n + j) \quad (18)$$

per $k = 1, \dots, K$, $m = 0, \dots, M/2^k - 1$, $n = 0, \dots, N/2^k - 1$. Il filtro di riduzione 2D è di tipo separabile e quindi ottenuto dal prodotto di un kernel lineare simmetrico dispari $\{r(i)\}$, di lunghezza $2L_r + 1$.

A partire dalla GP viene ottenuta la ELP (*enhanced Laplacian pyramid*) [14],

$$L_k(m, n) \triangleq G_k(m, n) - \tilde{G}_{k+1}(m, n) \quad (19)$$

per $k = 0, \dots, K-1$, in cui $\tilde{G}_{k+1}(m, n)$ indica il livello $(k+1)$ della GP espanso alla dimensione del sottostante livello k :

$$\tilde{G}_{k+1}(m, n) \triangleq \sum_{\substack{i=-L_e \\ (j+n) \bmod 2=0 \\ (i+m) \bmod 2=0}}^{L_e} \sum_{j=-L_e}^{L_e} e(i) \times e(j) \cdot G_{k+1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right) \quad (20)$$

per $m = 0, \dots, M/2^k - 1$, $n = 0, \dots, N/2^k - 1$, e $k = 0, \dots, K-1$. Il filtro separabile 2D deriva da un kernel lineare simmetrico dispari $\{e(i)\}$ di lunghezza $2L_e + 1$.

La ELP è una descrizione *passa-banda* dell'immagine, ossia si ha una suddivisione della banda dell'immagine in ottave disgiunte, con la radice della GP, G_K , che corrisponde alla banda base.

L'attributo *enhanced* deriva dal fatto che il filtro di espansione è forzato ad essere a *mezza banda*, ossia un interpolatore per 2, ed è indipendente dal filtro di riduzione, che può essere a mezza banda oppure no [14]. Un esempio di filtro usato sia per la riduzione che per l'espansione è il kernel a *mezza banda* a 15 coefficienti ($L_r = L_e = 7$):

$$\begin{aligned} E(z) = 2 \cdot R(z) = & 1 + 0.61(z + z^{-1}) - 0.145(z^3 + z^{-3}) \\ & + 0.042(z^5 + z^{-5}) - 0.007(z^7 + z^{-7}). \end{aligned} \quad (21)$$

che presenta un buon compromesso tra selettività e costo computazionale.

4.2 La piramide laplaciana a rapporto

Essendo lo speckle modellato come rumore moltiplicativo, non è possibile applicare la ELP per il filtraggio. Infatti, il rumore è proporzionale alla media del segnale, che però viene rimossa dalla differenza in (19), in modo che il filtraggio *local-statistics* (17) non risulta più praticabile. Per superare tale inconveniente, viene definita una LP a rapporto (RLP), che possa conformarsi

alla natura moltiplicativa del rumore, a partire dalla GP (18). La nuova piramide viene così definita dal rapporto tra il livello k della GP e la versione espansa del livello $(k + 1)$, ottenendo

$$L_k^*(m, n) \triangleq \frac{G_k(m, n)}{\tilde{G}_{k+1}(m, n)}. \quad (22)$$

La RLP è una rappresentazione approssimativamente *passa banda* dell'immagine, dando vita ancora a un processo a multirisoluzione. In più, il vantaggio principale di (22) è che i suoi livelli sono caratterizzati da una media d'insieme approssimativamente unitaria. In questo modo, i filtri che usano le statistiche locali, come (17) possono essere riportati in una struttura a multirisoluzione processando i vari livelli della RLP, introducendo due tipi di benefici: da un lato, si ottiene una stima migliore delle statistiche locali grazie ad una media praticamente *spazio-invariante*; dall'altro, le statistiche del primo ordine stimate su più risoluzioni danno informazioni sulla covarianza, ossia una statistica del secondo ordine. In questo modo, il filtraggio LLMMSE sub-ottimo (17) viene a basarsi anche su stime robuste della covarianza spazio-variante approssimando il risultato ottimo ottenuto mediante la stima MMSE [15].

4.3 Schema piramidale di filtraggio

Per immagini naturali, ossia correlate, ogni livello della RLP risulta caratterizzato da un SNR che cresce al diminuire della risoluzione spaziale. Per questo motivo, il filtraggio può essere ottimizzato calcolando la varianza del rumore presente nei vari livelli della piramide (ossia L_k^* , $k < K$, oppure G_K), in modo da preservare meglio i vari dettagli nei livelli meno rumorosi, raggiungendo così un'adattività anche lungo le diverse scale.

La figura 1 mostra il diagramma a blocchi della procedura di filtraggio, in cui la versione filtrata del livello G_k della piramide, ossia \hat{F}_k , $k = K - 1, \dots, 0$, può essere ricostruita con una procedura ricorsiva interpolando il precedente livello ricostruito della GP e successivamente moltiplicando per il livello corrente della RLP filtrato con il filtro LLMMSE di Kuan (KF). Il livello filtrato della RLP è denominato \hat{L}_k^* in Fig. 1. Questo procedimento

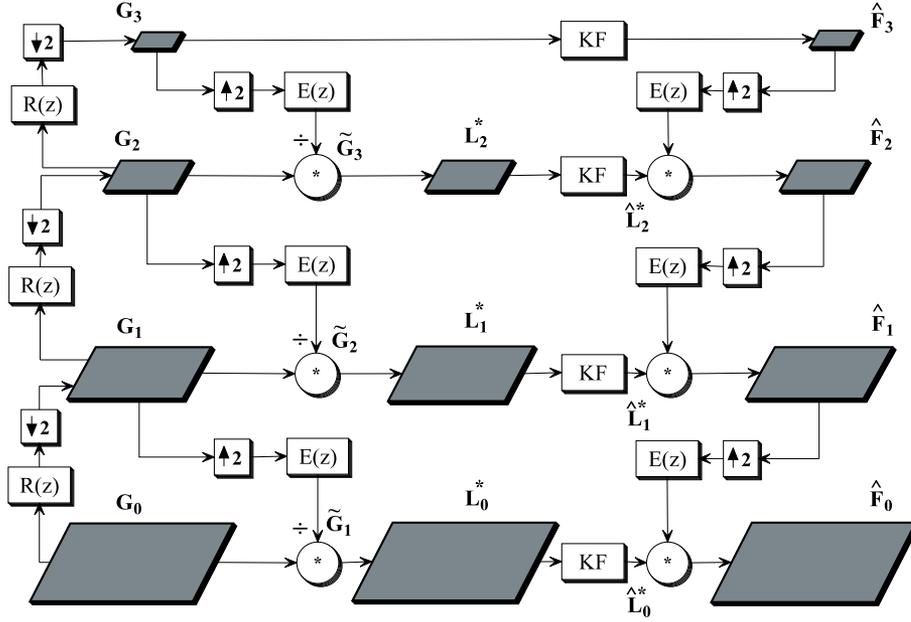


Figura 1: Schema di filtraggio con la piramide a rapporto.

continua fino a giungere ad ottenere l'immagine originale filtrata $\hat{F}_0(m, n)$:

$$\hat{F}_k(m, n) = \hat{L}_k^*(m, n) \times \sum_{\substack{m=-L_e \\ (j+n) \bmod 2=0 \\ (i+m) \bmod 2=0}}^{L_e} \sum_{\substack{n=-L_e \\ (j+n) \bmod 2=0 \\ (i+m) \bmod 2=0}}^{L_e} e(i) \times e(j) \times \hat{F}_{k+1} \left(\frac{i+m}{2}, \frac{j+n}{2} \right). \quad (23)$$

La dimensione della finestra di processazione al livello k dovrebbe essere progressivamente ridotta al crescere di k , ossia per esempio, 9×9 , 7×7 , 5×5 , 3×3 , for $K = 3$. Si noti che per $K = 0$, lo schema di Fig. 1 si riduce al filtro LLMMSE a singola scala (17).

4.4 Calcolo della varianza del rumore sui livelli della piramide per rumore bianco e colorato

La varianza di rumore sul livello k della RLP può essere calcolata a partire dalla varianza sull'immagine d'ingresso σ_u^2 , e considerando le risposte in fre-

quenza $R(\omega)$ del filtro di riduzione $\{r(i)\}$ ed $E(\omega)$ dell'interpolatore $\{e(i)\}$. Il procedimento viene descritto nel lavoro allegato [15], in cui però la trattazione si riferisce ad un rumore bianco. Nel caso di rumore correlato con coefficiente di correlazione sull'immagine di partenza pari a ρ_x e ρ_y nelle due direzioni, il procedimento deve essere modificato. Ciò perchè lo spettro del rumore colorato risulta spostato verso le basse frequenze, con la conseguenza di una diversa distribuzione nei livelli della GP e della RLP (nella radice, che rappresenta la banda base, troviamo una maggior quantità di rumore a spese di un minor livello presente nella RLP).

5 Stima di parametri di omogeneità per la classificazione

Il coefficiente di variazione C_v è un parametro comunemente usato nella classificazione e nella stima della varianza dello speckle in immagini SAR [7, 12]. Infatti nelle zone omogenee assume un valore molto basso, pari alla standard deviation σ_u dello speckle, mentre tende a crescere nelle regioni tessiturate. Suddividendo il rango di variazione del C_v in tre classi per mezzo di soglie appropriate, è quindi possibile classificare l'immagine in aree omogenee, mediamente tessiturate e altamente tessiturate.

Un metodo alternativo ma molto promettente che è stato sperimentato si basa sugli scatterplot *media vs. standard deviation* usati per la stima dei parametri di rumore. Ogni pixel dell'immagine produce un punto sullo scatterplot, che può venire indicizzato con il valore della 2D-PDF in quella zona del diagramma. Tale valore viene usato per costruire una mappa di *omogeneità* dell'immagine utile per la classificazione, sfruttando il fatto che le zone omogenee tendono a *clusterizzare*. In realtà, diversamente dal C_v che cresce quando aumenta la tessitura dell'immagine, il valore della *densità* è relazionato in modo diretto all'omogeneità locale e la sua dinamica è estremamente alta. Perciò, per poter ottenere valori che possano essere confrontati con quelli del C_v , è necessario effettuare una complementazione comprimendo allo stesso tempo la dinamica in modo non lineare.

Sia infatti $f(x, y)$ la 2D-PDF normalizzata e discretizzata, ossia tale che $\sum_x \sum_y f(x, y) = 1$ con $1 \leq x, y \leq L^2$, dove L^2 è il numero di suddivisioni dello scatterplot. Allora, $f(x, y) \leq 1$, ma per la maggior parte dei punti $f(x, y) \ll 1$. Per questo motivo, il parametro di inomogeneità nel punto

(m, n) dell'immagine viene definito come

$$\mathcal{I}(m, n) = -\alpha \cdot \log f[x(m, n), y(m, n)] \quad (24)$$

in cui $x(m, n)$ e $y(m, n)$ sono i valori di $\hat{\mu}_g(m, n)$ and $\hat{\sigma}_g(m, n)$ dati dalle (7) e (8), opportunamente quantizzati per ottenere valori discreti per x e y . La costante α è ottenuta imponendo che il valore minimo assunto dalla (24) sia uguale al valore teorico del C_v su zone omogenee, ossia $1/\sqrt{N_L}$ e $0.5227/\sqrt{N_L}$, dove N_L indica il numero di look, per immagini in intensità e ampiezza, rispettivamente. Se f_{MAX} indica il massimo di $f(x, y)$, il valore di α sarà dato, per immagini in ampiezza, da

$$\alpha = \frac{0.5227}{\log(f_{MAX}) \cdot \sqrt{N_L}} \quad (25)$$

Le mappe ottenute mostrano una notevole capacità di individuare le aree altamente tessiturate, laddove il C_v si trova invece in difficoltà [16]; in particolare ciò accade nelle zone su cui si trovano costruzioni come ponti, fabbricati, strade ecc., dove lo speckle risulta non completamente sviluppato dato che non vi sono scatteratori, ma si ha completa riflessione del segnale. Tali aree dovrebbero essere classificate come altamente disomogenee, ma poiché il segnale ritornato è molto forte, il C_v risulta alto solo sui bordi della costruzione, e non nella parte centrale, che viene mappata quasi come omogenea; viceversa nello scatterplot questo tipo di riflettori producono *outliers*, ossia punti sparsi che vengono subito correttamente classificati con il massimo grado di disomogeneità.

Questa maggior robustezza può essere attribuita al carattere bidimensionale dello scatterplot che si basa su statistiche congiunte al contrario del C_v . In più il parametro proposto non tiene conto dei valori assoluti delle statistiche locali, ma della probabilità congiunta che tali valori si verifichino, ossia non è un parametro meramente locale come il C_v , ma considera la globalità dell'immagine, e in tal modo opera una classificazione più efficiente.

6 Compressione near-lossless e virtually-lossless per immagini SAR

Il problema della compressione dati per le immagini ottenute da sensori remoti deve da un lato tener conto della grande mole di dati da trasmettere

e/o archiviare, il che richiederebbe un alto livello di compressione che non può essere ottenuto mediante metodi reversibili [14, 17, 18], e dall'altro della necessità di preservare la qualità originale dei dati trasmessi per poter evitare la perdita delle informazioni utili che devono essere estratte dall'immagine stessa [19].

In quest'ambito, notiamo che esistono molti metodi di compressione con perdita, come *JPEG* e più in generale i metodi che fanno uso di trasformate, che riescono a controllare l'errore quadratico medio di ricostruzione (*MSE*) dato da

$$MSE = E[(g(m, n) - \hat{g}(m, n))^2] \quad (26)$$

dove $g(m, n)$ rappresenta l'immagine originale e $\hat{g}(m, n)$ l'immagine ricostruita, ma non quello massimo (*PE*) dato da

$$PE = \max_{(m,n)} \{|g(m, n) - \hat{g}(m, n)|\} \quad (27)$$

che non può essere determinato nel dominio trasformato. Ciò significa che caratteristiche rilevanti dell'immagine possono essere localmente alterate di una quantità imprevedibile. Così solamente i metodi capaci di controllare l'errore di picco (*DPCM causale e non causale*) riescono a garantire una qualità uniforme sull'intera immagine. Tali algoritmi sono detti metodi *near-lossless* [20, 21, 22]. Se poi il PE viene fissato ad un livello inferiore all'errore medio introdotto dal rumore, supposto additivo e indipendente dal segnale, per esempio la metà della standard deviation del rumore stesso, l'immagine ricostruita viene ad essere *virtually-lossless*, ossia praticamente indistinguibile dall'originale, sia visivamente che come risultati ottenuti da postprocessazioni [23].

Per quanto riguarda le immagini SAR, l'alto livello di rumore presente rende quasi del tutto inefficienti i metodi *lossless*, anche se consideriamo una singola immagine. In più, la natura moltiplicativa dello *speckle* rende il controllo dell'errore di picco inadeguato, dato che l'errore da controllare non dovrebbe essere quello *assoluto*, ma *relativo*, ossia misurato sul rapporto tra pixel e non sulle differenze. In questo caso infatti la limitazione superiore del PE non garantisce una perdita relativa di qualità uniforme sull'immagine, dal momento che l'immagine è molto più corrotta sulle zone chiare che non su quelle scure. Perciò, la definizione di compressione *near-lossless* dovrebbe essere modificata per le immagini SAR, per indicare che il rapporto tra immagine originale e ricostruita deve trovarsi all'interno di un intervallo prefissato attorno al valore *uno*, che corrisponde all'assenza di errore. Se tale

intervallo si trova compreso ampiamente all'interno della distribuzione dello speckle, allora l'immagine decodificata potrà essere definita *virtually-lossless* [23].

Per poter effettuare una compressione con errore di picco razionale controllato di immagini SAR, viene qui proposto un approccio totalmente originale, basato sulla quantizzazione e successiva codifica della Piramide Laplaciana a Rapporto (RLP) definita dalla (22) nella versione utilizzata per il filtraggio dello speckle.

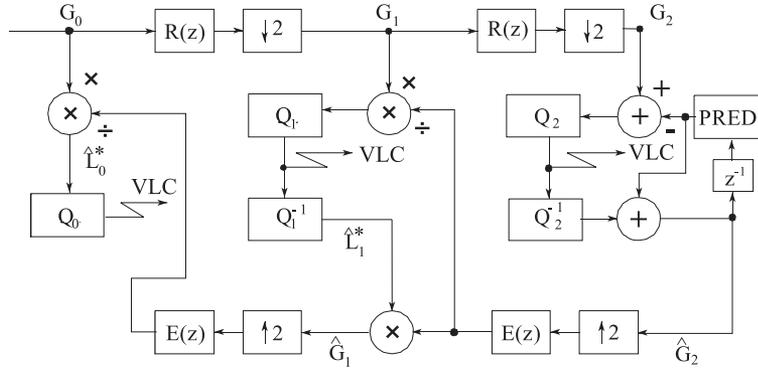


Figura 2: Codificatore ibrido basato sulla piramide a rapporto con anello di retroazione degli errori di quantizzazione.

Come abbiamo visto, la RLP possiede l'importante proprietà che il rumore presente sui suoi livelli, eccetto che per la banda base, tende a divenire indipendente dalla media del segnale. Infatti, la RLP è ricorsivamente definita come il rapporto tra un livello della Piramide Gaussiana (GP) e della sua versione passa-basso, in tal modo adattandosi per i modelli di rumore moltiplicativo. La RLP possiede un comportamento *passabanda*, poiché effettua un'analisi a multirisoluzione e in tal modo presenta diversi vantaggi nel campo della compressione dati. Il metodo proposto consiste nel codificare con una tecnica DPCM la banda base di dimensioni ridotte, quantizzando e codificando poi i livelli della RLP. I passi di quantizzazione possono essere scelti arbitrariamente, tranne quello sulla base della piramide, sfruttando l'anello di retroazione degli errori di quantizzazione al codificatore. Questo infatti provvede a interpolare gli stessi dati affetti da distorsione che saranno poi disponibili al ricevitore.

Per quanto riguarda la banda base ed i livelli intermedi, è stato implementato un quantizzatore lineare, e i passi di quantizzazione sono calcolati

per $k = K, K - 1, \dots, 1$, in modo da minimizzare il bit rate per una data distorsione, come risulta dalla relazione

$$\hat{\Delta}_k = \frac{2 \cdot \sigma_{k-1}^*}{P_E} \quad (28)$$

dove σ_{k-1}^* è la radice quadrata della varianza (media) L_{k-1}^* e P_E è il guadagno in potenza dell'interpolatore 1D [14]. L'ultimo passo di quantizzazione, Δ_0 , così come il tipo di quantizzatore, è cruciale per ottenere uno schema ad errore controllato, dal momento che regola il valore dell'errore di picco *relativo*.

La figura 2 mostra che \hat{G}_k , la GP che verrà ricostruita al decodificatore per $k < K$, è ricorsivamente data dal prodotto della \hat{G}_{k+1} espansa e di una versione approssimata a causa degli errori di quantizzazione della L_k^* , ossia \hat{L}_k^* , in cui $L_k^* = G_k / \hat{G}_{k+1}$ ($L_K^* = G_K$ per la radice della piramide o banda base). G_0 non può essere ricostruita esattamente dalla \hat{G}_K e dalla \hat{L}_k^* , $k = K - 1, \dots, 0$, a meno che la quantizzazione sull'ultimo livello non sia estremamente fine, ciò che implicherebbe un numero di bit di codice molto grande. Così, a differenza della ELP, la RLP è inutilizzabile per ottenere una compressione strettamente reversibile. Per di più, essendo la RLP una piramide *razionale*, l'uso di un quantizzatore lineare non riesce a garantire una compressione ad errore controllato definito dall'utente.

Per tale motivo, dato $\Delta \in \mathbb{R}$, con $\Delta > 0$ e $\Delta \neq 1$, definiamo la quantizzazione *logaritmica* di $t \in \mathbb{R}$, $t > 0$, come

$$\begin{aligned} \mathcal{Q}_\Delta(t) &= \text{round}[\log_\Delta(t)] \\ \mathcal{Q}_\Delta^{-1}(l) &= \Delta^l \end{aligned} \quad (29)$$

Ovviamente, $\mathcal{Q}_\Delta^{-1}(\mathcal{Q}_\Delta(t)) = \mathcal{Q}_{1/\Delta}(\mathcal{Q}_{1/\Delta}(t))$.

Dal momento che il termine \hat{G}_1 tiene conto ricorsivamente degli errori di quantizzazione precedenti e che $L_0^* = G_0 / \hat{G}_1$, la quantizzazione logaritmica (29) di L_0^* con un passo $\Delta_0 \neq 1$ implica che, se definiamo l' *errore di picco razionale* come

$$PRE \triangleq 20 \log_{10} \frac{\max\{G/\hat{G}_0\}}{\min\{G/\hat{G}_0\}} \quad (30)$$

si può facilmente provare che $\Delta_0^{-1/2} \leq G/\hat{G}_0 \leq \Delta_0^{1/2}$; da cui, $PRE = 20 \log_{10} \Delta_0$.

7 Risultati sperimentali e conclusioni

L'analisi degli argomenti oggetto degli studi effettuati durante l'assegno per la collaborazione ad attività di ricerca ha permesso di ottenere una gran quantità di risultati, sia per quanto riguarda immagini con rumore simulato, sia riguardo a immagini SAR vere e proprie. Per la discussione dei grafici ottenuti si rimanda ai lavori allegati [15, 24, 25].

Da quanto ottenuto, possiamo concludere che l'uso della piramide laplaciana a multirisoluzione permette sia di ottimizzare il filtraggio delle immagini SAR, ottenendo indubbi vantaggi rispetto agli schemi a singola scala, sia di ottenere una compressione near-lossless e virtually-lossless di tali immagini, ciò che consente di minimizzare il numero di bit trasmessi e quindi l'occupazione del canale trasmissivo, controllando al tempo stesso la qualità dell'immagine ricostruita per le successive elaborazioni di classificazione e interpretazione. L'uso di scatterplot bidimensionali basati sulle statistiche locali dell'immagine rumorosa consente altresì di stimare in modo robusto i parametri del rumore, la cui corretta determinazione risulta essenziale per il procedimento di filtraggio, e allo stesso tempo di ottenere un parametro classificativo che alle prime analisi si dimostra più efficiente rispetto ai classici metodi monodimensionali come il C_v .

Riferimenti bibliografici

- [1] C. Elachi, "Spaceborne radar remote sensing: applications and techniques," *IEEE Press*, 1988.
- [2] J. W. Goodman, "Some fundamental properties of speckle," *J. Opt.Soc.Amer.*, vol.66, pp. 1145–1150, 1976.
- [3] J.-S. Lee, I. Jurkevich, "Segmentation of SAR images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 27, no. 6, pp. 674–680, 1989.
- [4] J.-S. Lee, K. Hoppel, and S. H. Mango, "Unsupervised estimation of speckle noise in radar images," *Int'l J. Imaging Systems & Technology*, vol. 4, pp. 298–305, 1993.
- [5] B. Aiazzi, L. Alparone and S. Baronti, "Reliably estimating the speckle noise from SAR data," *Proc.IEEE Int. Geoscience and Remote Sensing Symposium*, pp.1546–1548, 1999.

- [6] L. Alparone, S. Baronti, and A. Garzelli, "Evaluating time correlation of speckle in ERS-1 SAR images," in *Proc. IEEE IGARSS'98*, pp. 30–32, 1998.
- [7] J.-S. Lee, I. Jurkevich, P. Dewaele, P. Wambacq, and A. Oosterlinck, "Speckle filtering of synthetic aperture radar images: A Review," *Remote Sensing Reviews*, vol. 8, pp. 313–340, 1994.
- [8] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 2, pp. 165–177, 1985.
- [9] J.-S. Lee, "Refined Filtering of Image noise using local statistics," *Comput. Vision, Graphics, Image Process.*, vol. 15, no. 2, pp. 380–389, 1981.
- [10] J.-S. Lee, "Digital Image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, no. 2, pp. 165–168, 1980.
- [11] F. T. Ulaby, F. Kouyate, B. Brisco, and T. H. Lee Williams, "Textural information in SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-24, no. 2, pp. 235–245, 1986.
- [12] P. Meer, R.-H. Park, and K. Cho, "Multiresolution adaptive image smoothing," *CVGIP: Graphic Models Image Process.*, vol. 56, no. 2, pp. 140–148, 1994.
- [13] P. J. Burt, "The pyramid as a structure for efficient computation," in *Multiresolution Image Processing and Analysis*, A. Rosenfeld (Ed.), Springer-Verlag, 1984.
- [14] B. Aiazzi, L. Alparone, S. Baronti, and F. Lotti, "Lossless image compression by quantization feedback in a content-driven enhanced laplacian pyramid," *IEEE Trans. Image Process.* vol. IP-6, no. 6, pp. 831–843, 1997.
- [15] B. Aiazzi, L. Alparone, and S. Baronti, "Multiresolution local-statistics speckle filtering based on a ratio laplacian pyramid," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, no. 5, pp. 1466–1476, 1998.

- [16] A. Lopes, R. Touzi and E. Nezry, “Adaptive speckle filters and scene heterogeneity” *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 992–1000, 1990.
- [17] X. Wu, and N. Memon, “Context-based, adaptive, lossless image coding” *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 437–444, 1997.
- [18] R. E. Roger and M. C. Cavenor, “Lossless compression of AVIRIS images” *IEEE Trans. Image Processing*, vol. 5, no. 5, pp. 713–719, 1996.
- [19] J. A. Benediktsson, J. R. Sveinsson and K. Arnason, “Classification and feature extraction of AVIRIS data” *IEEE Trans. Geosci. Remote Sensing*, vol. 33, no. 5, pp. 1194–1205, 1995.
- [20] K. Chen and T. V. Ramabadran, “Near-lossless compression of medical images through entropy-coded DPCM” *IEEE Trans. Medical Imaging*, vol. 13, no. 3, pp. 19–24, 1994.
- [21] B. Aiazzi, L. Alparone, S. Baronti, G. Chirò, F. Lotti, and M. Moroni, “A pyramid-based error-bounded encoder: an evaluation on X-ray chest images,” *Signal Processing*, vol. 59, no. 2, pp. 173–187, 1997.
- [22] B. Aiazzi, L. Alparone, and S. Baronti, “Near-lossless image compression: a key to high quality data distribution,” *Proc. SPIE , Mathematics of Data/Image Coding, Compression, and Encryption III, with Applications*, vol. 4122, 2000, in press.
- [23] B. Aiazzi, L. Alparone, and S. Baronti, “Information preserving storage of remote sensing data: virtually lossless compression of optical and SAR images,” *Proc. IEEE IGARSS 2000*, pp. 2657–2659, 2000.
- [24] B. Aiazzi, L. Alparone, S. Baronti, and L. Santurri, “Quality issues in remote sensing image compression: near-lossless coding of optical and microwave data,” *Proc. SPIE , Image and Signal Processing for Remote Sensing VI*, S. B. Serpico, ed., vol. 4170, 2000, in press.
- [25] B. Aiazzi, L. Alparone, and S. Baronti, “A hybrid de-speckling filter driven by a novel homogeneity feature,” *Proc. SPIE , SAR Image Analysis, Modelind, and Techniques V*, F. Posa, ed., vol. 4173, 2000, in press.