

Correzione geometrica di immagini satellitari ad alta risoluzione: metodi non parametrici

Piero Boccardo, Enrico Borgogno Mondino e Fabio Giulio Tonolo

Politecnico di Torino, Dipartimento di Georisorse e Territorio, Corso Duca degli Abruzzi, 24 -
- 10129 Torino. E-mail: fabio.giuliotonolo@polito.it

Riassunto

La correzione geometrica di immagini satellitari può essere affrontata utilizzando metodi non parametrici, ovvero generalizzati, che permettono di determinare una relazione tra le coordinate immagine e le coordinate terreno. Nel presente articolo vengono illustrati due differenti metodi non parametrici: i polinomi razionali e le reti neurali. Per i primi ne viene evidenziata l'instabilità numerica e proposta, oltre all'utilizzo dell'algoritmo di stabilizzazione di Tikhonov, l'applicazione dell'analisi di adeguatezza statistica del modello per la determinazione del grado dei polinomi da utilizzare. Per quanto riguarda l'approccio neurale viene presentata una procedura prototipale di correzione geometrica di immagini basata su una rete MLP (*Multi Layer Perceptron*) opportunamente addestrata con algoritmo *Back Propagation* Levenberg-Marquardt. Le prestazioni dei due approcci sono state valutate attraverso l'analisi dell'accuratezza di posizionamento planimetrico calcolata su diversi set di immagini satellitari utilizzate per le prove (Eros A1, QuickBird, Spot5).

Abstract

The geometric correction of high resolution satellite images can be carried out through universal non parametric models that relates the image coordinates to the terrain coordinates. In this paper two different non parametric methods are shown: the rational function model (RFM) and the neural net (MLP). As far as RFM are concerned attention is focused on their numerical instability, strictly related to the polynomials order: Tikhonov regularization algorithm and overparametrization tests are proposed in order to determine the number of coefficients to be used. As far as neural network approach is concerned it is shown an innovative procedure for image geometric correction based on an MLP (Multi Layer Perceptron) network opportunely trained with a Back-Propagation Levenberg-Marquardt algorithm. Performances of these two methods have been evaluated through planimetric positioning accuracy analysis on the residuals calculated using satellite images from different sensors (Eros A1, QuickBird, Spot5).

Introduzione

La correzione geometrica delle scene acquisite da piattaforma satellitare risulta di fondamentale importanza in quegli ambiti applicativi che impongono requisiti di accuratezza metrica. La calibrazione risulta indispensabile qualora l'immagine sia utilizzata come base cartografica o per la derivazione di tematismi. In letteratura questo tipo di operazione può

essere effettuata secondo due approcci opposti: modelli rigorosi e modelli non parametrici. I modelli rigorosi si basano sulle equazioni di collinearità [Toutin, 2004] adattate alla tecnica di acquisizione *pushbroom*, utilizzata da tutti i satelliti ottici ad alta risoluzione geometrica. In questo caso i parametri di orientamento esterno vengono modellizzati come polinomi tempo dipendenti di grado superiore al primo: la stima delle incognite richiede in questo caso valori iniziali approssimati che vengono dedotti dai file di metadati generalmente forniti con le immagini stesse.

Non sempre però le società distributrici di immagini sono disponibili a divulgare all'utente finale informazioni tecniche dettagliate relative alla piattaforma utilizzata o alle caratteristiche del sensore necessarie per l'implementazione di modelli rigorosi. Per questo motivo sono stati introdotti i modelli non parametrici, ovvero generalizzati, indipendenti sia dal tipo di sensore utilizzato sia dalla metodologia di acquisizione (Fig. 1).

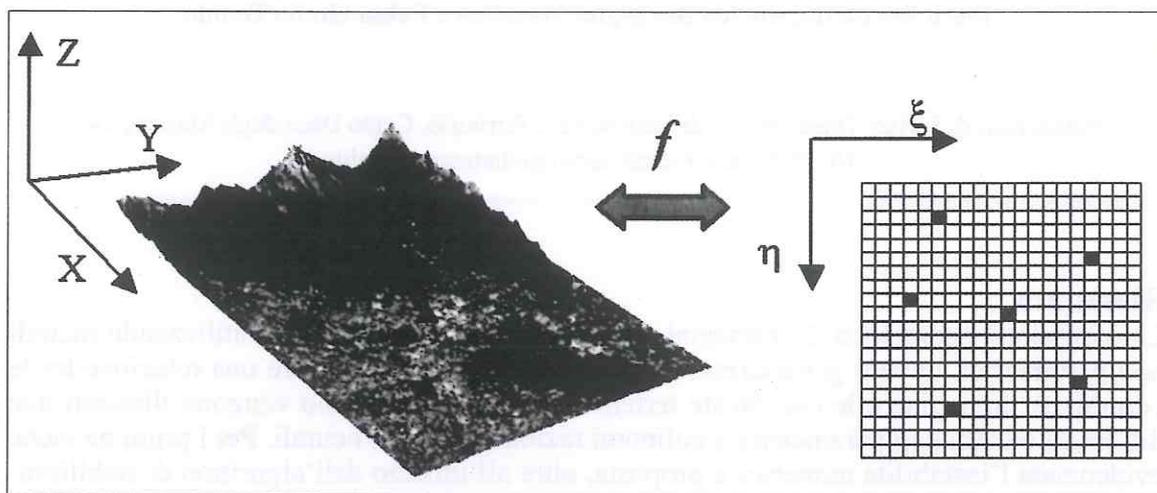


Figura 1 – Schematizzazione di un metodo non parametrico: relazione generalizzata f tra coordinate terreno ed immagine.

Il metodo dei rapporti polinomiali (RFM)

Il modello non parametrico maggiormente diffuso, implementato in quasi tutti i software per il trattamento geometrico di immagini satellitari, è quello dei rapporti polinomiali (in letteratura anglosassone denominati *Rational Function Model*, *RFM* - *Rational Polynomial Coefficient*, *Rational Polynomial Camera*, *RPC* - *Rational Function Coefficients*, *RFC* [Dowman e Tao, 2002]). Questo tipo di approccio viene utilizzato dai rivenditori di immagini per permettere agli utenti finali di ottenere prodotti a valore aggiunto quali ortoproiezioni senza la necessità di dover disporre del modello del sensore ma allegando unicamente i coefficienti dei polinomi che definiscono la relazione tra coordinate immagine e coordinate terreno.

Descrizione del modello

Il metodo dei rapporti polinomiali permette di determinare una relazione tra le coordinate immagine (ξ, η) e le coordinate tridimensionali oggetto (X, Y, Z) mediante rapporti di polinomi, come indicato nella [1]:

$$\begin{aligned} \xi &= \frac{P_a(X, Y, Z)}{P_b(X, Y, Z)} \\ \eta &= \frac{P_c(X, Y, Z)}{P_d(X, Y, Z)} \end{aligned} \quad [1]$$

P_a, P_b, P_c, P_d sono generalmente polinomi di grado massimo pari a 3 (20 coefficienti), esprimibili mediante le Equazioni [2], [3] o [4]:

$$P_a(X, Y, Z) = a_0 + a_1X + a_2Y + a_3Z + a_4X^2 + a_5XY + \dots + a_{17}Y^2Z + a_{18}YZ^2 + a_{19}YZ^3 \quad [2]$$

$$P_a(X, Y, Z) = \sum_{i=0}^{m_1} \sum_{j=0}^{m_2} \sum_{k=0}^{m_3} a_{ijk} X^i Y^j Z^k$$

$$0 \leq m_1 \leq 3 ; 0 \leq m_2 \leq 3 ; 0 \leq m_3 \leq 3 \text{ e } m_1 + m_2 + m_3 \leq 3 \quad [3]$$

$$\begin{aligned} P_a &= P \cdot a^T \\ P &= (1, X, Y, Z, X^2, XY, \dots, Z^2, X^3, X^2Y, \dots, YZ^2, Z^3) \\ a &= (a_0, a_1, \dots, a_{19}) \end{aligned} \quad [4]$$

Sebbene i modelli non parametrici non siano direttamente correlati alla fisica del problema, nella scelta del grado dei polinomi da utilizzare è comunque possibile tenere in conto le seguenti considerazioni [Tao e Hu, 2001]: le distorsioni prospettiche sono modellabili attraverso i termini di primo grado mentre le distorsioni dovute alla curvatura terrestre, alla rifrazione atmosferica ed alle distorsioni dell'ottica sono approssimabili mediante i termini di secondo grado. Ulteriori distorsioni incognite, quali le vibrazioni della piattaforma satellitare, possono essere invece corrette mediante termini di terzo grado. Le Equazioni [1] sono note in letteratura come *RFM Upward* (verso l'alto) in quanto permettono di ricavare le coordinate immagine a partire dalle coordinate tridimensionali di un punto terreno. In alcuni casi, ad esempio per la determinazione degli spigoli della matrice immagine corretta, è necessario utilizzare le Equazioni inverse [5] che prendono il nome di *RFM Downward* (verso il basso):

$$\begin{aligned} X &= \frac{P'_a(\xi, \eta, Z)}{P'_b(\xi, \eta, Z)} \\ Y &= \frac{P'_c(\xi, \eta, Z)}{P'_d(\xi, \eta, Z)} \end{aligned} \quad [5]$$

Stima dei coefficienti

Per poter utilizzare le Equazioni [1] devono essere stimati i coefficienti polinomiali a_i, b_i, c_i e d_i ($i = 0 \div 19$). Considerando polinomi di grado massimo pari a 3 i coefficienti incogniti sono al massimo 78, in quanto i termini noti dei polinomi a denominatore b_0 e d_0 sono posti uguali ad 1. E' quindi necessario collimare un numero minimo di punti di appoggio a terra (*GCP, Ground Control Point*) pari a 39 ($N \geq 39$), in quanto ogni punto genera 2 equazioni. Poiché le Equazioni [1] non sono lineari è innanzitutto necessario linearizzarle rispetto alle incognite mediante uno sviluppo in serie di Taylor arrestato al primo ordine:

$$\begin{aligned} \xi &= \xi^0 + \left(\frac{\partial \xi}{\partial a_0} \right) (a_0 - a_0^0) + \left(\frac{\partial \xi}{\partial a_1} \right) (a_1 - a_1^0) + \dots + \left(\frac{\partial \xi}{\partial a_{19}} \right) (a_{19} - a_{19}^0) + \\ &+ \left(\frac{\partial \xi}{\partial b_1} \right) (b_1 - b_1^0) + \dots + \left(\frac{\partial \xi}{\partial b_{19}} \right) (b_{19} - b_{19}^0) + \left(\frac{\partial \xi}{\partial c_0} \right) (c_0 - c_0^0) + \dots \\ &+ \left(\frac{\partial \xi}{\partial c_{19}} \right) (c_{19} - c_{19}^0) + \left(\frac{\partial \xi}{\partial d_1} \right) (d_1 - d_1^0) + \dots + \left(\frac{\partial \xi}{\partial d_{19}} \right) (d_{19} - d_{19}^0) \\ \eta &= \eta^0 + \left(\frac{\partial \eta}{\partial a_0} \right) (a_0 - a_0^0) + \left(\frac{\partial \eta}{\partial a_1} \right) (a_1 - a_1^0) + \dots + \left(\frac{\partial \eta}{\partial a_{19}} \right) (a_{19} - a_{19}^0) + \\ &+ \left(\frac{\partial \eta}{\partial b_1} \right) (b_1 - b_1^0) + \dots + \left(\frac{\partial \eta}{\partial b_{19}} \right) (b_{19} - b_{19}^0) + \left(\frac{\partial \eta}{\partial c_0} \right) (c_0 - c_0^0) + \dots \\ &+ \left(\frac{\partial \eta}{\partial c_{19}} \right) (c_{19} - c_{19}^0) + \left(\frac{\partial \eta}{\partial d_1} \right) (d_1 - d_1^0) + \dots + \left(\frac{\partial \eta}{\partial d_{19}} \right) (d_{19} - d_{19}^0) \end{aligned} \quad [6]$$

dove: $\bar{0} = (a_0^0, a_1^0, \dots, a_{19}^0, b_1^0, \dots, b_{19}^0, c_0^0, c_1^0, \dots, c_{19}^0, d_1^0, \dots, d_{19}^0)$

I valori approssimati iniziali delle incognite a_i^0, b_i^0, c_i^0 e d_i^0 vengono calcolati dalle Equazioni [1] mediante la soluzione del seguente sistema lineare:

$$\begin{aligned} \xi \cdot P_b &= P_a \\ \eta \cdot P_d &= P_c \end{aligned} \quad [7]$$

Scrivendo le Equazioni [6] per ognuno degli N punti di appoggio collimati, denominando A la matrice disegno, V il vettore degli scarti, L il vettore dei termini noti e Q la matrice dei pesi, si ottiene il seguente sistema in forma matriciale:

$$V = Q \cdot A \cdot X - Q \cdot L \quad [8]$$

dove:

$$L^{(2N \times 1)} = \begin{bmatrix} \xi_i - \xi_i^0 \\ \vdots \\ \eta_i - \eta_i^0 \end{bmatrix} \quad Q^{(2N \times 2N)} = I \quad \text{alla prima iterazione}$$

$$A^{(2N \times 78)} = \begin{bmatrix} A_\xi & 0 \\ 0 & A_\eta \end{bmatrix}$$

$$A_\xi^{(N \times 39)} = \begin{bmatrix} \left(\frac{\partial \xi}{\partial a_0} \right)_{0,1} & \left(\frac{\partial \xi}{\partial a_1} \right)_{0,1} & \dots & \left(\frac{\partial \xi}{\partial b_{19}} \right)_{0,1} \\ \vdots & \vdots & \dots & \vdots \\ \left(\frac{\partial \xi}{\partial a_0} \right)_{0,N} & \left(\frac{\partial \xi}{\partial a_1} \right)_{0,N} & \dots & \left(\frac{\partial \xi}{\partial b_{19}} \right)_{0,N} \end{bmatrix}$$

$$A_\eta^{(N \times 39)} = \begin{bmatrix} \left(\frac{\partial \eta}{\partial c_0} \right)_{0,1} & \left(\frac{\partial \eta}{\partial c_1} \right)_{0,1} & \dots & \left(\frac{\partial \eta}{\partial d_{19}} \right)_{0,1} \\ \vdots & \vdots & \dots & \vdots \\ \left(\frac{\partial \eta}{\partial c_0} \right)_{0,N} & \left(\frac{\partial \eta}{\partial c_1} \right)_{0,N} & \dots & \left(\frac{\partial \eta}{\partial d_{19}} \right)_{0,N} \end{bmatrix}$$

In relazione all'Equazione [4] la matrice disegno A assume la seguente forma (maggiormente indicata per la successiva implementazione dell'algoritmo):

$$A^{(2N \times 78)} = \begin{bmatrix} \left(\frac{1}{P_b} \cdot P \right)_0 & \left(-\frac{P_a}{P_b^2} \cdot P \right)_0 & 0 & 0 \\ 0 & 0 & \left(\frac{1}{P_d} \cdot P \right)_0 & \left(-\frac{P_c}{P_d^2} \cdot P \right)_0 \end{bmatrix}$$

Risolvendo iterativamente il sistema [8] ai minimi quadrati al fine di minimizzare il vettore degli scarti V si ottiene quindi:

$$X = (A^T \cdot Q \cdot A)^{-1} \cdot A^T \cdot Q \cdot L \quad [9]$$

Implementazione dell'algoritmo

L'algoritmo dei polinomi razionali sopra descritto è stato implementato in linguaggio IDL (*Interactive Data Language*) in modo da poter supplire alla poca trasparenza dei software commerciali.

Innanzitutto è necessario, per evitare problemi di calcolo numerico (errori di troncamento, *under/overflow*), normalizzare sia le coordinate immagine sia quelle oggetto nell'intervallo $(-1;+1)$ [OpenGIS-OCG, 1999] in accordo con la seguente formula (k generica coordinata in ingresso):

$$\begin{aligned} k_{NORM} &= \frac{k - k_{Offset}}{k_{Scale}} \\ k_{Offset} &= \frac{\max(k_i) + \min(k_i)}{2} \\ k_{Scale} &= \frac{\max(k_i) - \min(k_i)}{2} \end{aligned} \quad [10]$$

Inoltre poiché i denominatori dei polinomi assumono valori molto differenti in funzione della distribuzione dei GCP e dell'escursione altimetrica è probabile che la matrice disegno A risulti mal condizionata. Conseguentemente la matrice $A^T \cdot Q \cdot A$ può risultare singolare, in particolare quando vengono utilizzati polinomi di grado elevato. Il processo iterativo in questo caso risulta spesso non convergente. Proprio per evitare questo problema i coefficienti RPC distribuiti insieme alle immagini vengono calcolati sfruttando la conoscenza del modello del sensore. Non vengono infatti collimati manualmente dei punti di appoggio ma viene generato un grigliato regolare in coordinate immagine e successivamente vengono ricavate, mediante il modello rigoroso, le relative coordinate oggetto su differenti livelli di quota (metodo *Terrain-Independent*).

Ipotizzando che il *camera model* non sia a disposizione degli utenti risulta necessario scegliere i punti di appoggio in maniera convenzionale, ovvero attraverso collimazione di punti omologhi su cartografia/DEM o mediante apposite campagne di rilievo GPS. Non essendo possibile ottenere una distribuzione regolare dei GCP è quindi necessario implementare un algoritmo di regolarizzazione numerica al fine di rendere il processo iterativo convergente. Uno degli algoritmi di regolarizzazione maggiormente usati per la risoluzione di sistemi mal condizionati è quello di Tikhonov [11], che consiste nell'aggiunta di una costante λ^2 arbitrariamente piccola alla diagonale della matrice $A^T \cdot Q \cdot A$, al fine di migliorarne il numero di condizionamento:

$$X = (A^T \cdot Q \cdot A + \lambda^2 \cdot I)^{-1} \cdot A^T \cdot Q \cdot L \quad [11]$$

La scelta del parametro di Tikhonov non è univoca e viene fatta empiricamente elaborando numerose soluzioni al variare del parametro λ , scegliendo quella che minimizza gli scarti sui punti di controllo.

Le numerose prove effettuate hanno evidenziato come, sebbene il processo iterativo converga, la successiva fase di ortoproiezione presenti, in alcuni casi, problemi di distorsioni delle immagini generate dovute con molta probabilità all'utilizzo di polinomi di grado troppo elevato.

In Figura 2 viene riportata a titolo di esempio una porzione di immagine Eros A1 ortoproiettata mediante il metodo dei polinomi razionali in cui si è notata la presenza di un asintoto. Per verificare la correttezza di tale ipotesi e per evitare di dover decidere a priori il grado dei polinomi da utilizzare (ovvero il numero di parametri per ogni polinomio) è stata imple-

mentata un'analisi di adeguatezza del modello, basata su due differenti test statistici, al fine di determinare automaticamente se e quali coefficienti non sono necessari.



Figura 2 – Porzione di immagine Eros A1 ortoproiettata. Si noti come la strada si disponga erroneamente lungo un asintoto evidenziato dalla linea rossa.

In primo luogo viene utilizzato il test del χ^2 a due code per verificare se il modello è sovrapparametrizzato. Vengono calcolati il parametro K definito nella [12] e, fissato un indice di significatività α (es: $\alpha = 5\%$), i valori $K_1 = \chi_{\frac{\alpha}{2}}^2$ e $K_2 = \chi_{1-\frac{\alpha}{2}}^2$ per $(n - r)$ gradi di libertà :

$$K = \frac{(n-r) \cdot \hat{\sigma}_0^2}{\sigma_0^2} \quad [12]$$

dove: $\hat{\sigma}_0^2$ = varianza stimata, σ_0^2 = varianza a priori (posta pari ad 1), n = numero di equazioni, r = numero di incognite.

Si presentano quindi i seguenti tre possibili casi:

$K_1 < K < K_2$	corretto
$K < K_1$	sovrapparametrizzazione
$K > K_2$	errori grossolani

Nel caso in cui venga evidenziata una sovrapparametrizzazione del modello utilizzato si procede ad un ulteriore test di significatività dei coefficienti per determinare quanti e quali coefficienti non siano necessari.

Viene calcolato il parametro standardizzato Z secondo la seguente relazione:

$$Z_i = \frac{\hat{x}_i}{\hat{\sigma}_{x_i}} \quad [13]$$

dove: \hat{x}_i = i-imo coefficiente stimato, $\hat{\sigma}_{x_i}$ = i-imo sqm stimato. Nel caso risulti verificato il seguente test a due code [14] basato sulla distribuzione di Student il relativo coefficiente x_i viene posto uguale a zero:

$$|Z_i| \leq t_{\alpha/2} \quad [14]$$

dove: $t_{\alpha/2}$ = valore della distribuzione di Student per il relativo valore di ridondanza ($n - r$). Il coefficiente viene annullato mediante l'aggiunta di una nuova equazione di condizione nel sistema [6] inserendo un peso elevato nella posizione corrispondente della matrice Q . In Tabella 1 sono evidenziati i risultati dell'analisi di adeguatezza del modello: risulta evidente l'incremento del numero di condizione della matrice normale. Si sottolinea inoltre come il numero di coefficienti scartati risulti minimo per le immagini del sensore Eros A1, in quanto acquisite in modalità asincrona (velocità del satellite maggiore della velocità di scansione a terra) e caratterizzate da un elevato angolo di *Off-Nadir* (circa 23°).

Tabella 1 – Incremento del numero di condizione della matrice normale e numero di coeff. scartati in seguito all'analisi di adeguatezza del modello.

Satellite	N° condizione iniziale	N° coeff. scartati	N° di condizione finale
Eros A1	0.0003	4	18
QuickBird	0.0001	26	79
Spot5	0.3000	24	392

Il metodo a Reti Neurali

L'approccio neurale per l'ortoproiezione di immagini satellitari ed aeree a scansione può essere considerata un innovativo tentativo di soluzione del problema della correzione di immagini mediante metodi non parametrici.

Le reti neurali costituiscono modelli matematici la cui filosofia operativa si ispira alle dinamiche biologiche cerebrali, schematizzando il processo di calcolo come un flusso di informazioni distribuite la cui elaborazione avviene all'interno di unità di calcolo dedicate dette "nodi" o "neuroni" della rete. Alcune di queste ricevono informazioni dall'ambiente esterno, altre restituiscono risposte all'ambiente e altre ancora, se ve ne sono, comunicano solamente con le unità all'interno della rete (Fig. 3): esse sono dette rispettivamente unità di ingresso (*input*), unità di uscita (*output*) e unità nascoste (*hidden*).

Il panorama neurale è assai vasto ed algoritmi neurali sono stati sviluppati per far fronte alle più diverse applicazioni: sono queste che determinano la scelta dell'uno o dell'altro algoritmo. Per la correzione geometrica di immagini satellitari l'attenzione è stata posta sull'algoritmo MLP (*Multi Layer Perceptron*), le cui caratteristiche di approssimazione e stima di funzioni (non lineari) sono già state ampiamente dimostrate.

L'idea base è quella di sostituire il modello proiettivo *upward* che relaziona le coordinate immagine (ξ, η) e quelle oggetto (X, Y, Z) con una rete neurale MLP opportunamente addestrata sulla base dei punti di appoggio (GCP).

Le motivazioni della scelta nascono dall'analisi dei problemi connessi all'approccio RFM precedentemente descritto. Le reti neurali preservano dalla forzata linearizzazione delle

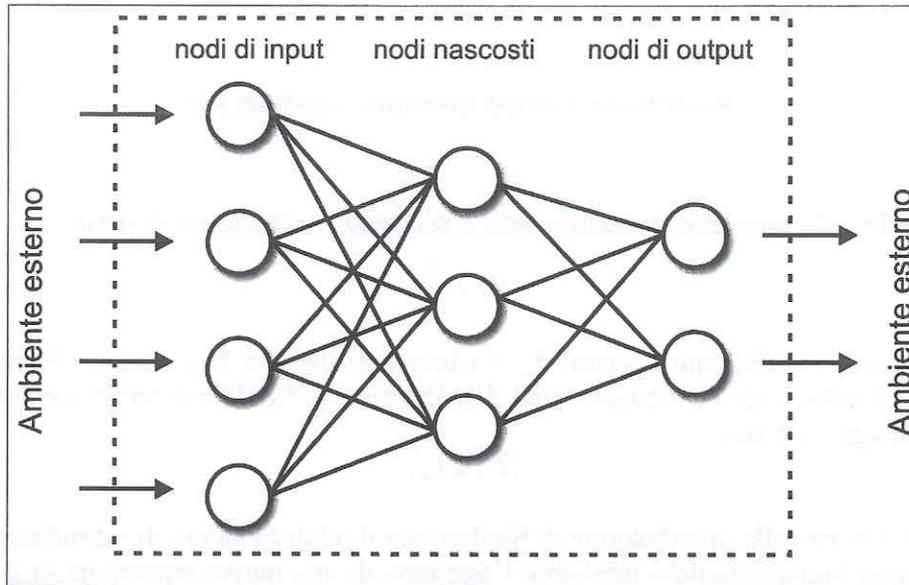


Figura 3 - Schema concettuale di rete neurale MLP a due layer computazionali (*hidden e output*).

equazioni nell'intorno di una soluzione approssimata. Esse costituiscono una risposta non lineare ad un problema non lineare, la cui efficacia aumenta, al pari degli RFM, all'aumentare del numero di GCPs ed al diminuire delle deformazioni originarie dell'immagine.

Reti neurali MLP

Nelle reti MLP ciascun nodo svolge un'operazione molto semplice che consiste nel generare, attraverso una opportuna funzione, detta di trasferimento, una risposta ai segnali che su di esso convergono attraverso i canali di comunicazione. Tali canali simulano le sinapsi biologiche ed il loro compito consiste nel "pesare" l'intensità dei segnali trasmessi: per questo essi vengono definiti con il nome di "pesi sinaptici" o semplicemente "pesi". Formalmente, il segnale di risposta (u_i) restituito dal generico nodo i -esimo è uguale a:

$$u_i = f\left(\sum_{j=1}^N w_{ij} p_{ij} + b_i\right) \quad [15]$$

dove f è la funzione di trasferimento, che normalmente assume la forma di una tangente iperbolica [16] o di una sigmoide logica [17], w_{ij} sono i pesi del nodo i -esimo, p_{ij} sono gli ingressi al nodo i -esimo (in numero di N) e b_i sono scalari additivi detti *bias*, considerabili come i pesi di un ingresso aggiuntivo unitario (Fig. 4).

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad \text{Tangente iperbolica} \quad [16]$$

$$f(x) = \frac{1}{1 + e^{-\alpha x}} \quad \text{Sigmoide logica} \quad [17]$$

Questo tipo di algoritmo appartiene alla famiglia delle reti neurali *feed-forward*, reti cioè in cui l'informazione viaggia in modo parallelo e monodirezionale.

La rete MLP costituisce dunque un modello matematico i cui parametri sono i pesi e i *bias* dell'*hidden* e dell'*output layer*. Stimare i valori di tali parametri sulla base di opportuni campioni detti *pattern* costituisce la fase di addestramento della rete.

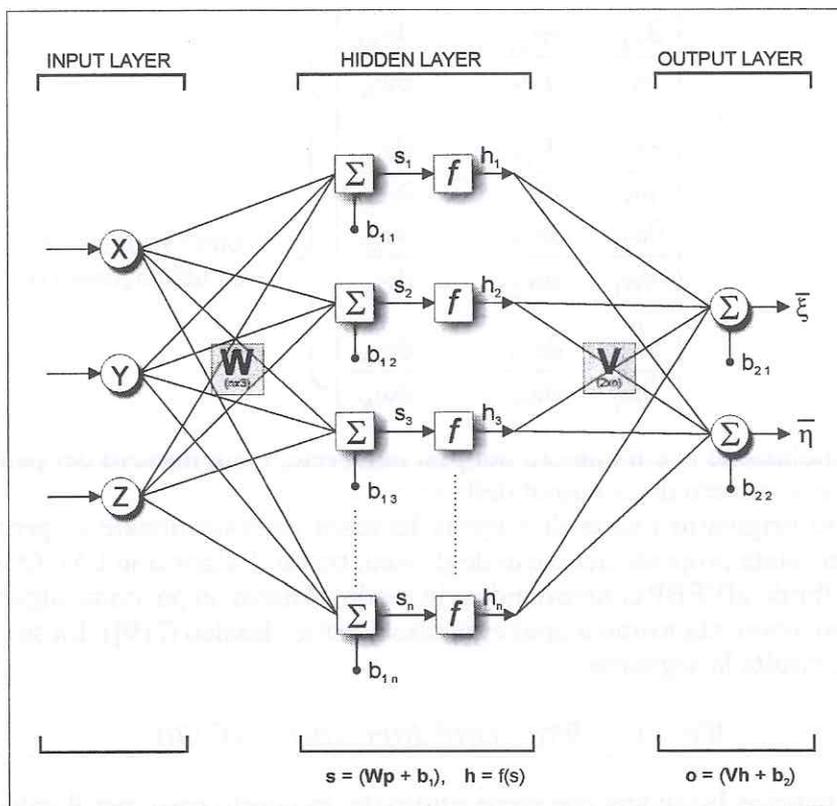


Figura 4 - Modello matematico di rete neurale MLP a due layer computazionali (hidden e output).

Questo avviene, nel caso considerato, attraverso un algoritmo *Error Backpropagation* (EBP) ottimizzato (per una maggiore velocità di convergenza), denominato Levenberg-Marquardt (LM). L'algoritmo EBP [Werbos, 1988] prevede che i pesi della rete assumano valori che minimizzano (minimi locali) la Funzione di Prestazione o *Performance Function* (PF). Questa viene definita, per un addestramento di tipo *batch* (in cui i pesi vengono aggiornati, ad ogni iterazione, sulla base degli errori commessi additivamente da tutti i *pattern* di addestramento) come:

$$PF(W(t)) = \sum_{p=1}^P \sum_{k=1}^K (d_{kp} - f_{kp})^2 = E(t)^T E(t) \quad [18]$$

dove $W(t) = [w_1, w_2, \dots, w_N]^T$ è il vettore dei pesi della rete all'epoca t , t conta le iterazioni (epoche) del processo di addestramento fissate dall'operatore, d_{kp} è il valore atteso (*target*) dell'output k -esimo relativo al pattern di addestramento p -esimo, f_{kp} è il valore dell'output k -esimo calcolato dalla rete, $E(t) = [e_{1p}, e_{2p}, \dots, e_{kp}, e_{12}, \dots, e_{k2}, e_{1p}, \dots, e_{kp}]^T$, in cui $e_{kp} = (d_{kp} - f_{kp})$, $k=1, \dots, K$, $p=1, \dots, P$, è l'errore cumulativo di un addestramento *batch*.

La legge di aggiornamento dei pesi, o regola di apprendimento della rete, nella sua formulazione più nota (gradiente discendente) relativa all'algoritmo EBP è la seguente:

$$W(t+1) = W(t) - \alpha \nabla PF(t) = W(t) - \alpha \cdot J^T E(t) \quad [19]$$

dove $\alpha = \text{learning rate}$ (numero costante o variabile che stabilisce l'entità della variazione), $\nabla PF(t) = J^T E(t)$ è il gradiente della PF esprimibile come prodotto vettoriale tra la matrice Jacobiana J delle derivate prime ([20]) della PF ed il vettore dell'errore cumulativo $E(t)$.

$$J = \begin{bmatrix} \frac{\partial e_{11}}{\partial w_1} & \frac{\partial e_{11}}{\partial w_2} & \dots & \frac{\partial e_{11}}{\partial w_N} \\ \frac{\partial e_{21}}{\partial w_1} & \frac{\partial e_{21}}{\partial w_2} & \dots & \frac{\partial e_{21}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{K1}}{\partial w_1} & \frac{\partial e_{K1}}{\partial w_2} & \dots & \frac{\partial e_{K1}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{1P}}{\partial w_1} & \frac{\partial e_{1P}}{\partial w_2} & \dots & \frac{\partial e_{1P}}{\partial w_N} \\ \frac{\partial e_{2P}}{\partial w_1} & \frac{\partial e_{2P}}{\partial w_2} & \dots & \frac{\partial e_{2P}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{kP}}{\partial w_1} & \frac{\partial e_{kP}}{\partial w_2} & \dots & \frac{\partial e_{kP}}{\partial w_N} \end{bmatrix} \quad \begin{array}{l} \left. \begin{array}{l} \text{Errori sugli output relativi} \\ \text{al primo pattern} \end{array} \right\} \\ \left. \begin{array}{l} \text{Errori sugli output relativi} \\ \text{all'ultimo pattern (P)} \end{array} \right\} \end{array} \quad [20]$$

In questa formulazione N è il numero dei pesi della rete, P è il numero dei pattern di addestramento, K è il numero degli output della rete.

Tale approccio originario risulta di estrema lentezza computazionale e, pertanto, alcune modifiche sono state proposte nel corso degli anni, tra cui l'algoritmo LM. Questo realizza un approccio ibrido all'EBP comportandosi in modo adattivo un po' come algoritmo Gauss-Newton, un po' come algoritmo a gradiente discendente classico ([19]). La sua formulazione matriciale, risulta la seguente:

$$W(t+1) = W(t) - (J(t)^T J(t) + (t)I)^{-1} J(t)^T E(t) \quad [21]$$

J è ancora la matrice Jacobiana che viene utilizzata, in questo caso, per il calcolo delle derivate seconde (matrice Hessiana) della PF grazie alla formulazione, che questa possiede, come somma di quadrati. La matrice Hessiana ($H = J^T J$) viene utilizzata dagli algoritmi Gauss-Newton come parametro di apprendimento dinamico della rete. L'algoritmo LM è ampiamente riconosciuto come il più efficiente dal punto di vista del raggiungimento dell'accuratezza delle stime. E' un buon compromesso tra la velocità di convergenza dei metodi Gauss-Newton ($\mu = 0$) e la stabilità del gradiente discendente (μ molto grande). $\mu(t)$ risulta in questo caso un parametro di apprendimento dinamico che viene aggiornato (secondo modalità più o meno complesse qui taciute), nel corso delle iterazioni. L'operatore ha facoltà di impostarne il valore iniziale. Il metodo Gauss-Newton ($\mu=0$) risulta più veloce e più accurato nell'intorno del minimo della PF. Lo scopo è dunque quello di muovere la soluzione verso un approccio Gauss-Newton il più in fretta possibile [Demuth e Beale, 2000]. μ viene decrementato dopo ogni iterazione in cui la PF risulti diminuita ed incrementato solo quando la PF cresca. L'algoritmo LM presenta problemi di memoria computazionale qualora si lavori con grandi matrici Jacobiane, la cui inversione risulta onerosa. Si ricorda che il rango della matrice da invertire ad ogni iterazione è pari al numero dei parametri da stimare. Se le dimensioni della rete (numero di nodi) crescono si assiste ad un'esplosione dei costi computazionali.

Una rete per la correzione geometrica di immagini

L'approccio neurale che viene qui presentato utilizza una rete MLP addestrata con algoritmo EBP LM. Le routine che la realizzano sono quelle presenti all'interno del Neural Network

Toolbox di MATLAB 5.3. L'approccio di "ortoproiezione" adottato è quello *upward* tale per cui le coordinate oggetto (X,Y,Z) costituiscono gli *input* della rete e le coordinate immagine (ξ, η) gli *output*. E' stata prevista una rete a due soli *layer* computazionali ritenendo eventuali livelli *hidden* aggiuntivi non significativi per le prestazioni (da letteratura). Sono stati previsti dunque un solo *hidden layer* ed un *output layer*. Inoltre sono state verificate ed implementate due configurazioni di reti rispetto alle possibili funzioni di trasferimento utilizzabili. Nel primo caso, risultato il migliore per il trattamento di immagini *pushbroom* (cui si riferiscono i risultati qui presentati) la funzione di trasferimento adottata per l'*hidden layer* è una tangente iperbolica [16] mentre per l'*output layer* è una funzione lineare semplice (pura somma pesata).

Nel secondo caso, considerato il più idoneo al trattamento di immagini *whiskbroom*, è stata prevista, per l'*hidden layer*, una funzione di trasferimento a sigmoide logica [17] mentre per l'*output layer*, anche in questo caso, una semplice funzione lineare.

Il numero di nodi (dell'*hidden layer*) che realizza le migliori prestazioni è da determinare di volta in volta sulla base di prove ripetute, che, nella routine sviluppata in ambiente MATLAB, è il calcolatore a condurre in modo autonomo.

Si ricorda che una stima approssimata (giacchè si opera in ambito non lineare queste considerazioni risultano puramente indicative) del numero massimo di nodi ammissibile può essere ottenuta confrontando il numero dei pattern di addestramento (i GCP) con quello dei parametri da stimare (pesi e bias). Questi risultano in numero pari a

$$N_{param} = (3_{(X,Y,Z)} \cdot M_{PESHidden}) + M_{BIASHidden} + 2_{(\xi,\eta)} \cdot M + 2 \quad [22]$$

dove M è il numero dei nodi dell'*hidden layer*.

Indicazioni più precise possono derivare comunque da un'attenta analisi dei risultati (residui) verificando l'eventuale comparsa di fenomeni di sovrapparametrizzazione (*overfitting*), in prima approssimazione identificabili nel progressivo ampliarsi della differenza tra i residui sui GCP e quelli sui *Check Point* (CHK).

L'accuratezza della soluzione (valutabile appunto in termini di residui sui GCP e sui CHK) varia sensibilmente al variare dell'inizializzazione pseudo-casuale dei pesi della rete e del numero dei nodi, mentre il valore iniziale del parametro di apprendimento μ risulta abbastanza trascurabile anche se valori prossimi a 10^{-3} risultano consigliabili. Le configurazioni di rete (architetture) che la procedura sviluppata è in grado di verificare dipendono da alcuni parametri che l'operatore deve fornire:

- a) intervallo di variabilità del numero dei nodi: deve essere definito il numero massimo e minimo di nodi da testare. Tale valore risulta influenzato dal numero di GCP forniti come *pattern* di addestramento. Un numero troppo elevato di nodi pregiudica la capacità di generalizzazione della rete, un numero troppo basso non approssima la funzione in modo adeguato;
- b) numero di inizializzazioni successive per ogni configurazione di nodi: deve essere indicato quante volte l'addestramento debba essere ripetuto per ogni architettura. I risultati ottenibili, con pari numero di nodi e di μ , possono differire anche molto tra di loro a seconda di come i pesi vengano inizializzati alla prima iterazione. Il rischio, per una sfortunata inizializzazione, è quella che l'algoritmo di ottimizzazione si blocchi all'interno di minimi locali, della funzione d'errore, troppo pronunciati. La ripetizione per un numero di volte sufficientemente alto (10 per questo lavoro) dell'addestramento preserva da questo rischio.
- c) livello di accuratezza richiesto ai residui sui GCP e sui CHK: la configurazione ottimale viene identificata sulla base del soddisfacimento simultaneo di alcune condizioni:

$$\begin{aligned}
 RMSE_{GCP} &< soglia \\
 RMSE_{CHK} &< N \cdot soglia
 \end{aligned}
 \tag{23}$$

dove

$$RMSE = \sqrt{\frac{\sum_{i=1}^P RMS_i}{N-1}}$$

con

$$RMS_i = \sqrt{\Delta\xi_i^2 + \Delta\eta_i^2}$$

rispettivamente errore sul singolo GCP o CHK. Il valore di *soglia* dipende dalla scala attesa per l'ortofoto. La simultaneità del soddisfacimento delle 2 condizioni serve ad evitare, in modo semplice, problemi di sovrapparametrizzazione della rete (troppi nodi) rilevabili, in prima approssimazione, come differenza tra $RMSE_{GCP}$ e $RMSE_{CHK}$.

Data set

La seguente tabella riporta le principali caratteristiche delle immagini oggetto della sperimentazione:

Tabella 2 – Caratteristiche delle immagini satellitari utilizzate nella fase di test dei metodi non parametrici.

Sensore	Risoluzione (m)	Area (km ²)	Zona
Eros A1	2.3 (off-nadir)	12x19	Cuneo
Spot 5	2.5 (supermode)	60x60	Torino
QuickBird	0.62	12x12	Caselle (TO)

Per la collimazione dei punti di appoggio (GCP) e dei punti di controllo (CHK) relativi alle immagini Eros A1 e Spot 5 sono stati utilizzati come supporto cartografico la Carta Tecnica Regionale 1:10.000 (CTR) e la Carta Tecnica Provinciale 1:5.000 (CTP).

In funzione dell'elevata risoluzione geometrica dell'immagine QuickBird è stata invece progettata un'apposita campagna di rilievo GPS al fine di poter disporre di 100 vertici



Figura 5 – Verifica dell'accessibilità e visibilità dei satelliti GPS di un GCP individuato sull'immagine QuickBird.

omogeneamente distribuiti all'interno di un grigliato con passo di circa 1 km². Inizialmente sono stati individuati 200 potenziali punti sull'immagine che sono stati successivamente oggetto di verifica relativamente all'accessibilità dell'area e alla visibilità dei satelliti GPS (Fig. 5). Tale fase ha permesso l'individuazione di 100 punti che sono stati rilevati in modalità rapido statica (tempo di occupazione di circa 20 minuti).

Il posizionamento dei vertici è stato di tipo relativo, generando le diverse basi tra il vertice, la stazione GPS permanente di Torino ed una stazione Master diposta nel centro della scena (Mappano), con scarti sempre inferiori ai 30 cm (0,5 celle per le immagini QuickBird).

Il modello digitale del terreno della Regione Piemonte, (grigliato regolare con passo di 50m ed accuratezza di $\pm 2.5m$) è stato utilizzato durante la fase di ortoproiezione, in quanto le equazioni [1] richiedono la conoscenza dell'andamento altimetrico del terreno.

Risultati sperimentali

I metodi non parametrici RFM e MLP sono stati testati sulle immagini satellitari del *data set* descritto nel precedente paragrafo. L'accuratezza di posizionamento planimetrico è stata valutata attraverso il calcolo dei residui sia sui GCP utilizzati per la stima dei parametri di trasformazione sia sui CHK (differenti dai precedenti). Sono state inoltre calcolate le medie dei residui in modo da evidenziare eventuali sistematismi. Durante tutte le prove effettuate è stata mantenuta una distribuzione di GCP geometricamente omogenea sull'intera immagine, essendo la validità dei metodi non parametrici decrescente con l'aumentare della distanza dai punti di appoggio.

Le Tabelle 3 e 4 riportano i risultati ottenuti rispettivamente mediante i rapporti polinomiali e le reti neurali.

Tabella 3 – Risultati ottenuti mediante l'applicazione del metodo RFM.

<i>Satellite</i>	<i>N° GCPs</i>	<i>N° CHKs</i>	<i>Media E CHK</i>	<i>Media N CHK</i>	<i>RMSE CHK (pixel)</i>	<i>RMSE GCP (pixel)</i>
Eros A1	51	6	0.00	0.00	3.19	0.83
QuickBird	60	30	-0.09	0.02	2.76	0.86
Spot5	50	5	-0.02	-0.09	2.09	1.01

Tabella 4 – Risultati ottenuti mediante l'applicazione del metodo a reti neurali MLP.

<i>Satellite</i>	<i>N° GCPs</i>	<i>N° CHKs</i>	<i>Media E CHK</i>	<i>Media N CHK</i>	<i>RMSE CHK (pixel)</i>	<i>RMSE GCP (pixel)</i>
Eros A1	51	6	-0.23	-1.10	2.46	1.08
QuickBird	60	30	-0.23	-0.14	2.37	1.40
Spot5	50	5	-2.04	0.02	2.96	1.38

I risultati ottenuti mediante il metodo RFM confermano la validità locale del metodo, evidenziata da residui inferiori al pixel sui GCP e maggiori sui CHK. Si sottolinea inoltre come il valore di RMSE sui punti di controllo sia superiore a quelli ottenuti in precedenti test [Boccardo et al., 2003; Giulio Tonolo e Poli, 2003] in quanto sono state elaborate scene intere e non porzioni limitate. I risultati peggiori sono relativi alle prove effettuate sull'immagine Eros A1, a causa dell'elevato valore dell'angolo di *off-nadir* e dell'acquisizione di tipo *pushbroom* asincrona.

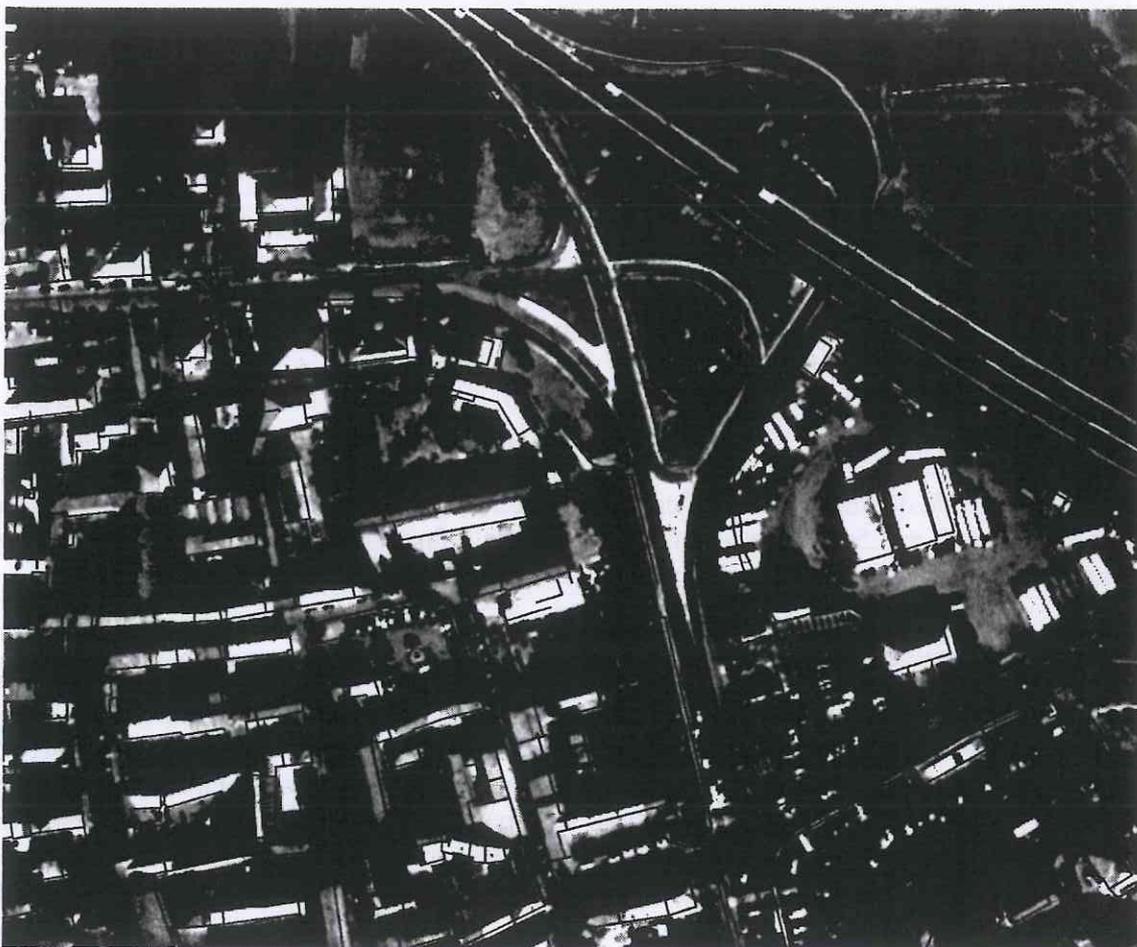


Figura 6 – Ortoproiezione della banda pancromatica QuickBird (0,61m). Metodo RFM. DEM Regione Piemonte.

Le reti neurali confermano ancora una volta la loro grande capacità di generalizzazione per l'approssimazione di funzioni consentendo di raggiungere risultati comparabili, ed a volte superiori, sui *Check Points*, a quelli ottenibili con il metodo RFM. Le reti inoltre consentono di ovviare ai problemi delle deformazioni locali asintotiche proprie degli RFM, producendo immagini corrette senza anomalie particolari.

Al fine di valutare anche visivamente la qualità della trasformazione successivamente alla stima dei parametri sono state ortoproiettate le immagini originali. In Figura 6 è riportata una porzione dell'immagine Quickbird ortoproiettata sovrapposta ad una cartografia numerica 1:1000.

Conclusioni e sviluppi futuri

Entrambi i metodi di calibrazione geometrica illustrati in questo articolo presentano il grande vantaggio di essere indipendenti dal tipo di sensore utilizzato e dalla geometria di presa: nonostante ciò consentono di ottenere accuratezze adeguate alla scala attesa per l'ortofoto nella maggior parte delle immagini elaborate. Non è però possibile ottenere una completa automazione del processo in quanto è necessario procedere ad una scelta empirica di alcuni parametri della trasformazione (in entrambi gli approcci) al fine di ottimizzare la stabilità numerica e minimizzare lo scarto tra soluzione reale e soluzione stimata.

Il modello a reti neurali non ha evidenziato problemi di stabilità o convergenza, ma localmente consente di ottenere accuratezze minori.

Per quanto riguarda il metodo dei rapporti polinomiali lo svantaggio maggiore riscontrato durante i test effettuati è sicuramente la presenza di asintoti e distorsioni che vengono evidenziati nella fase di ortoproiezione delle immagini: tali inconvenienti, sebbene vengano ridotti o eliminati grazie all'analisi di adeguatezza del modello, potrebbero essere controllati solo mediante l'utilizzo di un grigliato regolare di punti di appoggio su diversi livelli di quota.

Per cercare di sfruttare la stabilità numerica delle reti neurali e l'accuratezza locale del metodo RFM è attualmente in fase di sperimentazione un approccio integrato basato sulla successiva applicazione di entrambe le trasformazioni. Attraverso il metodo MLP viene generata una griglia regolare e densa di punti che, opportunamente pesati rispetto ai GCP, vengono utilizzati nella stima dei parametri RFM. I primi risultati hanno evidenziato l'assenza di deformazioni ed un'accuratezza qualitativamente confrontabile con le precedenti prove di ortoproiezione: sono in corso le validazioni statistiche dei risultati.

Le immagini utilizzate sono state acquistate nell'ambito del progetto di ricerca COFIN 2001 "L'uso delle immagini satellitari ad alta risoluzione per le analisi territoriali" (Coordinatore nazionale Prof. Sergio Dequal, Politecnico di Torino).

Bibliografia

- Boccardo P., Borgogno Mondino E., Giulio Tonolo F. (2003) - *High resolution satellite images position accuracy tests*. Atti IGARSS 2003. Toulouse. Atti su CD.
- Dowman I., Tao V. (2002) - *An update on the use of rational functions for photogrammetric restitution*. ISPRS, Vol. 7 (3): 26-29.
- Giulio Tonolo F., Poli D. (2003) - *Georeferencing of EROS-A1 high resolution images with rigorous and rational function model*. Atti "High Resolution Mapping from Space", ISPRS. Hannover. Atti su CD.
- Howard Demuth, Mark Beale - *MATLAB Neural Network Toolbox User's Guide*. The Mathworks, Version 4, pp.5.8-5.30.
- Open GIS (1999) - *Abstract specification. Topic 7: The earth imagery case*. Open GIS Consortium.
- Tao C., Hu Y. (2001) - *A comprehensive study of the Rational Function Model for Photogrammetric processing*. Photogram. Eng. Remote Sensing, December 2001: 1347-1357.
- Toutin T. (2004) - *Geometric processing of remote sensing images: models, algorithms and methods*. Inter. J. Remote Sensing, 24 (to be published).
- Werbos P. J. (1988) - *Back-propagation: Past and future*. Atti International Conference on Neural Networks, San Diego, CA, 1: pp. 343-354.

Manoscritto ricevuto il 20/04/2004, accettato il 03/11/2004.