

# Sistema di riconoscimento automatico di aree tematiche per immagini telerilevate

Cosimo DISTANTE (\*), Alessandro LEONE (\*), Lorenzo VASANELLI (\*\*),  
Marco PALAZZO (\*\*), Stefano ROCCA (\*)

(\*) CNR – IMM Sezione di Lecce, Via Provinciale per Arnesano, 73100 Lecce

(\*\*) DII – Università di Lecce, Via Provinciale per Arnesano, 73100 Lecce

Il presente lavoro propone una tecnica di classificazione automatica di aree tematiche di particolare interesse nel Salento (uliveto, seminativo e zone abitate) a partire da immagini telerilevate a diverse scale di risoluzione acquisite dal satellite *QuickBird*. L'approccio proposto utilizza le informazioni tessiturali di piccole regioni nelle diverse bande dell'immagine, al fine di determinare un modesto numero di *features* essenziali nell'attività di classificazione supervisionata basata su rete neurale auto-organizzante SOM. I descrittori che sono stati adottati esprimono l'informazione tessiturale sia in termini statistici (legati ai momenti del primo e secondo ordine), sia sfruttando le capacità di localizzazione spaziale/frequenziale esibite da un limitato insieme di funzioni 2D di *Gabor* opportunamente individuate mediante una strategia di *Pursuit*. La fase di sperimentazione ha evidenziato la validità dell'approccio proposto garantendo la discriminazione delle aree tematiche in maniera affidabile, precisa e con un esiguo dispendio computazionale.

## Abstract

This work presents a new approach for automatic classification of thematic regions of the south part of Salento (olive grove, sown grove and populated regions), by using images at different spatial resolutions acquired by Quickbird. The proposed scheme describes textural information of little patches of one (or more) band of the multispectral image, in order to define few features that we can use during the supervised classification process based on SOM neural network. The extracted features describe textural information in terms of statistical parameters (mean, contrast and energy) and as linear combination of a restricted set of 2D Gabor functions selected by using a Pursuit scheme. Experimental results prove that the used features and the classification process are able to describe correctly 3 clusters in a good way and in a low-power manner.

## Introduzione

L'attività di classificazione/riconoscimento delle aree tematiche in immagini telerilevate è indispensabile in diversi settori applicativi che vanno dal monitoraggio del territorio alla pianificazione urbana, dal censimento e controllo di piantagioni sottoposte a vincoli (ad esempio l'ulivo secondo le norme dell'Unione Europea) all'individuazione dell'abusivismo edilizio, ecc.. I soggetti interessati a disporre di strumenti di classificazione delle immagini da satellite si individuano sia nelle Pubbliche Amministrazioni che in enti ed aziende private. A tutti i livelli, però, si necessitano automatismi per l'analisi e l'interpretazione della grande mole di dati forniti dai diversi *Imagery Provider*.

Il presente lavoro affronta il problema della classificazione di 3 aree tematiche quali l'uliveto (codici *Corine* 223), il seminativo (codici *Corine* 2111 e 2112) e le zone abitate (codici *Corine* 111, 112, 121, 122, 131, 133), tipiche del territorio nelle immediate vicinanze del Comune di Surbo nel sud Salento (figura 1).

In particolare sono state utilizzate immagini acquisite dal satellite *QuickBird* fornite dal *Provider DigitalGlobe*, utilizzando, a diversi livelli, l'informazione delle 5 bande (una pancromatica a risoluzione spaziale di 0.7m, tre multispettrali RGB ed il vicino infrarosso, queste ultime a risoluzione spaziale di 2.8m). Si precisa che i dati *QuickBird* in possesso sono del tipo *Basic* per cui provvisti di correzioni radiometriche e di sensore.



Figura 1 - Caratteristiche strutturali delle 3 aree tematiche di interesse (seminativo, zone abitate ed uliveto in banda pancromatica) acquisite da *QuickBird* e relative all'agro di Surbo (LE)

### La metodologia

L'approccio proposto in questo lavoro effettua la classificazione automatica delle 3 aree tematiche utilizzando una rappresentazione semplice e compatta della tessitura di piccole regioni centrate in ciascun pixel di una banda (o più) dell'immagine multispettrale considerata. I primi tre descrittori della tessitura di cui ci si è avvalsi sono definiti sulla base di informazioni statistiche locali delle diverse regioni in accordo alle seguenti relazioni:

$$\mu_{(x,y)}^n = \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m I_{(x,y)}^n(i, j) \quad [1]$$

$$C_{(x,y)}^n = \sqrt{\frac{\sigma_{(x,y)}^n}{\mu_{(x,y)}^n}} \quad [2]$$

$$E_{(x,y)}^n = \sum_{i=1}^m \sum_{j=1}^m [I_{(x,y)}^n(i, j)]^2 \quad [3]$$

dove  $\mu_{(x,y)}^n$  rappresenta la media dei livelli di intensità nella n-esima banda della regione quadrata  $I_{(x,y)}^n$  di dimensione  $m \times m$  centrata nel pixel a coordinate  $(x, y)$ ,  $C_{(x,y)}^n$ ,  $\sigma_{(x,y)}^n$  e  $E_{(x,y)}^n$  rispettivamente il contrasto, la varianza e l'energia per la medesima regione.

In aggiunta a tali descrittori, l'approccio caratterizza l'informazione tessiturale in termini di combinazione lineare di un numero ristretto di funzioni 2D di *Gabor* (Weldon et al., 1996) estratte da un grande insieme di funzioni *kernel* opportunamente generato (dizionario overcompleto). Un approccio *greedy* di *Pursuit* (Vandergheynst P., Frossard P., 2001) è stato adottato nell'ottica di limitare le funzioni di base capaci di rappresentare correttamente qualsivoglia tessitura: l'algoritmo *Matching Pursuit* MP (Phillips P.J., 1998 – Mallat S., Zhang Z., 1993 – Bergeaud F., Zhang Z., 1995) consente la selezione di poche funzioni appartenenti al dizionario capaci di descrivere la tessitura delle piccole regioni in input allo schema con una bassa perdita di informazione. Tale procedura viene eseguita *offline* una sola volta, al fine di definire il sottoinsieme (sub-dizionario) delle funzioni che meglio descrivono le caratteristiche tessiturali delle regioni in input. Definito il sub-dizionario, l'estrazione dei descrittori della tessitura di una regione richiede un'esigua richiesta computazionale poiché ottenuti come prodotto interno tra la regione ed il sub-dizionario (o una sua parte). Le successive sezioni di questo lavoro forniranno indicazioni in merito sia alle caratteristiche delle funzioni di *Gabor* adottate per la generazione dei dizionari overcompleti, sia ad una nuova versione generalizzata dell'algoritmo MP usato per la selezione delle funzioni 2D di *Gabor* maggiormente caratterizzanti le tessiture delle regioni considerate.

## I dizionari overcompleti 2D di Gabor

Negli ultimi dieci anni si è manifestato un progressivo interesse nei confronti dell'analisi basata sull'uso di funzioni 2D di *Gabor*, soprattutto nell'ambito dell'*image processing*, nei contesti di analisi della tessitura, di *pattern recognition* e di analisi/sintesi di immagini. Nel dominio spaziale, la funzione di *Gabor*  $g(x, y)$  è definita come una funzione Gaussiana modulata da un esponenziale complesso orientato:

$$g(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left[-\frac{1}{2}\left(\frac{\tilde{x}^2}{\sigma_x^2} + \frac{\tilde{y}^2}{\sigma_y^2}\right)\right] \exp[j(\omega_x^0 x + \omega_y^0 y)] \quad [4]$$

dove:

$$\begin{cases} \tilde{x} = x \cos \theta + y \sin \theta \\ \tilde{y} = -x \sin \theta + y \cos \theta \end{cases} \quad [5]$$

con  $\sigma_x, \sigma_y$  le deviazioni standard della funzione di *Gabor* lungo gli assi coordinati,  $(\omega_x^0, \omega_y^0)$  il vettore frequenza angolare e  $\theta$  la direzione lungo la quale la funzione Gaussiana è orientata. Papoulis A. (1998) ha dimostrato che per una funzione 2D di *Gabor* con banda passante di B ottave, i valori di  $\sigma_x$  e  $\sigma_y$  sono:

$$\sigma_x = \frac{(2^B + 1) \sqrt{2 \ln 2}}{(2^B - 1) 2\pi F^0} \quad [6]$$

$$\sigma_y = \frac{\sqrt{2 \ln 2}}{2\pi F^0 \tan \frac{\theta}{2}} \quad [7]$$

con:

$$F^0 = \sqrt{(f_x^0)^2 + (f_y^0)^2} \quad f_x^0 = \frac{\omega_x^0}{2\pi} \quad f_y^0 = \frac{\omega_y^0}{2\pi} \quad [8]$$

Un dizionario overcompleto è la collezione di versioni traslate (denominate atomi) e campionate (in una griglia diadica) di  $g(x, y)$ , al variare dei suoi parametri caratteristici. La figura sottostante riporta le caratteristiche di 3 differenti dizionari di Gabor ottenuti variando le dimensioni degli atomi, considerando 4 frequenze centrali, 8 frequenze angolari ed 1 ottava.

Dimensione degli atomi	Numero di atomi	$\lambda_{\min}$	$\lambda_{\max}$
8 × 8	2048	3.02	307.61
16 × 16	8192	1.92	563.54
32 × 32	32768	1.15	627.86

Figura 2 – Parametri per la generazione di 3 dizionari di Gabor al variare della dimensione degli atomi

## L'algoritmo Matching Pursuit

In questa sezione sarà descritto un nuovo approccio per l'estrazione di *features* descrittive della tessitura. Il problema è quello di descrivere le strutture fondamentali della tessitura, individuandone una rappresentazione efficiente di basso livello. Il metodo proposto usa i coefficienti della decomposizione di *Matching Pursuit* ricorrendo a dizionari 2D di *Gabor* ed alla teoria dei sistemi overcompleti di funzioni.

Lo schema caratterizza la tessitura in un modo computazionalmente poco oneroso, riducendosi ad una operazione di prodotto interno tra matrici. MP è un algoritmo non-lineare che decompone i segnali in input in espansioni lineari di funzioni elementari appartenenti ad un dizionario overcompleto precedentemente generato, raffinando progressivamente le approssimazioni secondo una procedura iterativa; MP seleziona ad ogni iterazione l'atomo del dizionario che maggiormente

riduce il “residuo” tra l’approssimazione corrente del segnale e il segnale stesso. L’approccio necessita di un processo di *training offline* al fine di definire un solo sub-dizionario di *Gabor* in grado di descrivere correttamente (e compattamente) la tessitura di qualsivoglia piccola regione quadrata di pixels. La semplicità nel processo di estrazione delle features di *Gabor* permette di evitare l’uso di altre misure tipiche caratterizzanti la tessitura quali matrice di co-occorrenza, energia, entropia, ecc..

La variante di MP qui proposta seleziona elementi da un grande insieme  $F = \{\varphi_j\}_{j=1}^{\ell}$  di funzioni di *Gabor* con lo scopo di minimizzare il residuo  $R^i f_p$  di ciascuna training patch (regione quadrata) in input allo schema ( $p \in [1, \dots, m]$  con  $m$  la quantità di training patch  $f_1, f_2, \dots, f_m$ ). La scelta dell’elemento  $\varphi_j$  dipende sia dal residuo  $R^i f_p$ , sia dai coefficienti delle precedenti iterazioni. In particolare, all’ $i$ -esima iterazione è selezionata la funzione  $\varphi_j$  che soddisfa la seguente relazione:

$$\varphi_j = \underset{\varphi_j \in F}{\operatorname{argmin}} C_i(R^i f_1, \dots, R^i f_m, A_{i-1}) \quad [9]$$

dove  $A_i$  è l’insieme dei coefficienti generati all’ $i$ -esima iterazione. Le figure 3 e 4 mostrano lo pseudo-codice per la selezione delle funzioni 2D di *Gabor* e per la valutazione del parametro  $C_i$  rispettivamente. Si osserva che lo schema di MP approssima i segnali in input descrivendo inizialmente il contenuto in bassa frequenza e via via il contenuto in alta frequenza (l’informazione di dettaglio). All’ $i$ -esima iterazione il  $p$ -esimo segnale è approssimato dal prodotto interno tra i coefficienti generati  $c_i^p$  e la matrice di sintesi ottenuta come pseudo-inversa della matrice contenente i primi  $i$ -esimi atomi del sub-dizionario, in accordo a quanto proposto nella teoria dei *frame* (Mallat S., Zhang Z., 1993).

$R^0 f_p = f_p$ con $1 \leq p \leq m$ ed $m$ la quantità di patch di training
For $i = 0$ to numero di funzioni volute
Calcola $C_i$ per ciascun atomo del dizionario $F$
Seleziona $\varphi_j$ che minimizza $C_i$
Aggiorna i coefficienti per ciascuna patch: $c_i^p = \langle R^i f_p, \varphi_j \rangle$
Aggiorna i residui $R^{i+1} f_p = R^i f_p - c_i^p \varphi_j$
End for

Figura 3 - Pseudo-codice per la selezione degli atomi di Gabor mediante lo schema di MP

$(c_0^p, \dots, c_{i-1}^p) =$ coefficienti calcolati per $(i-1)$ -esima iterazione per la $p$ -esima patch
Calcola il centroide $\mu$ di $(c_0^p, \dots, c_{i-1}^p, \langle R^i f_p, \varphi_i \rangle)$
Calcola la distanza media $C_i$ tra $\mu$ e $(c_0^p, \dots, c_{i-1}^p, \langle R^i f_p, \varphi_i \rangle)$
$C_i = \frac{1}{m} \sum \left\  \mu - (c_0^p, \dots, c_{i-1}^p, \langle R^i f_p, \varphi_i \rangle) \right\ $

Figura 4 - Pseudo-codice per la valutazione del parametro  $C_i$

## Il sistema di classificazione

Per ogni pixel di ciascuna banda dell’immagine multispettrale, il processo di classificazione prevede l’attività di normalizzazione delle *features* precedentemente estratte (media  $\mu_{(x,y)}^n$ ,

contrasto  $C_{(x,y)}^n$ , energia  $E_{(x,y)}^n$  e gli  $h$  coefficienti ottenuti come prodotto interno tra la regione quadrata ed i primi  $h$  atomi del sub-dizionario di *Gabor*) e la successiva definizione del *cluster* di appartenenza, in accordo alla risposta del classificatore usato. In particolare, il sistema di classificazione impiega una rete neurale auto-organizzante SOM (Kohonen T. 1997), operante sul *feature vector* ottenuto come concatenazione delle *features* estratte da 4 delle 5 bande *QuickBird* (un'analisi preliminare ha evidenziato la scarsa capacità discriminativa della banda del vicino infrarosso).

Nello spazio dei vettori di ingresso, la rete SOM definisce una mappa elastica di neuroni che in seguito all'addestramento andranno ad approssimare la distribuzione che caratterizza i dati di *input*: dopo l'addestramento i neuroni saranno disposti in maniera tale che *feature vector* simili siano associati allo stesso neurone (o comunque a neuroni topologicamente vicini), mentre *feature vector* diversi siano associati a neuroni tra loro molto distanti. La fase di *training* è stata condotta fornendo in *input* alla rete le *features* estratte da 3000 regioni aventi caratteristiche tessiturali note.

Durante la fase di classificazione la rete calcola la distanza euclidea tra il *feature vector* e ciascun neurone, individuando come vincente quello a distanza minima (BMU) e considerando come etichetta del *cluster* quella associata al BMU.

### Risultati conseguiti e conclusioni

Durante la fase di sperimentazione si è provveduto a verificare la metodologia proposta inizialmente su immagini sintetiche (immagini per cui per ogni pixel è nota la relativa area tematica), stabilendo la capacità discriminativa delle diverse bande, variando sia la dimensione delle regioni di pixel di cui caratterizzare la tessitura (8×8, 16×16 e 32×32), sia il numero di coefficienti di *Gabor* impiegati. L'approccio ha fornito i migliori risultati caratterizzando la tessitura di regioni di dimensioni 16×16, ricorrendo all'uso di poche funzioni di *Gabor*. Si osserva che un elevato numero di coefficienti di *Gabor* permette una sintesi/ricostruzione precisa della *patch* che rappresenta ma, di contro, è poco adatto nell'attività di classificazione dove si richiedono *feature vectors* di limitate dimensioni.

La tabella seguente mostra i migliori risultati ottenuti in termini di percentuale di corretta classificazione ed in termini di tempi di elaborazione (PC basato su Intel Pentium IV, 512MB) di immagini sintetiche, estraendo le *features* sia su ciascuna banda singolarmente, sia concatenando le *features* estratte da ogni banda.

Banda usata	Coefficienti di Gabor estratti oltre a $\mu_{(x,y)}^n, C_{(x,y)}^n, E_{(x,y)}^n$	Dimensione immagine	Dimensione mappa SOM	% corretta classificazione	Tempo impiegato (sec.)
Pancromatico	256	1416×819	20×8	91	706
R	6	1029×504	29×9	97	272
G	6	1029×504	20×9	96	255
B	6	1029×504	24×8	93	271
R + G + B	6	1029×504	20×9	98	590
R	256	1029×504	29×9	97	4070
G	256	1029×504	20×9	97	3896
B	256	1029×504	24×8	94	4043

Figura 5 - Risultati del processo di classificazione ricorrendo a patch 16 × 16

Le percentuali di corretta classificazione sono molto soddisfacenti in tutte le condizioni sperimentali, esibendo valori sempre superiori al 90%; tuttavia si è notato che l'informazione contenuta nella banda pancromatica è meno discriminativa rispetto a quella contenuta nelle tre bande colore RGB. Le migliori performance si hanno in corrispondenza di un esiguo numero di coefficienti di *Gabor* (solo 6), considerando contemporaneamente l'informazione tessiturale di tutte

e tre le bande RGB. Tali risultati sono giustificati dal fatto che i primi coefficienti della decomposizione MP trasportano la più alta quantità di informazione tessiturale (con soli 6 coefficienti è possibile ricostruire una regione  $16 \times 16$  con un errore di sintesi inferiore al 15%), mentre i successivi caratterizzano l'informazione in alta frequenza delle regioni quadrate (tali coefficienti sono scarsamente discriminanti poiché attribuibili a fenomeni di rumore presente nell'immagine). A tal proposito, si osservi che l'estrazione e la classificazione di un numero elevato di descrittori di Gabor (ad esempio 256) richiede tempi di calcolo decisamente elevati, pur tuttavia non garantendo miglioramenti in termini di classificazione.

La figura successiva mostra i buoni risultati conseguiti su una immagine reale (in rosso sono visualizzate le aree abitate, in verde le aree di oliveto ed in blu le colture seminative).

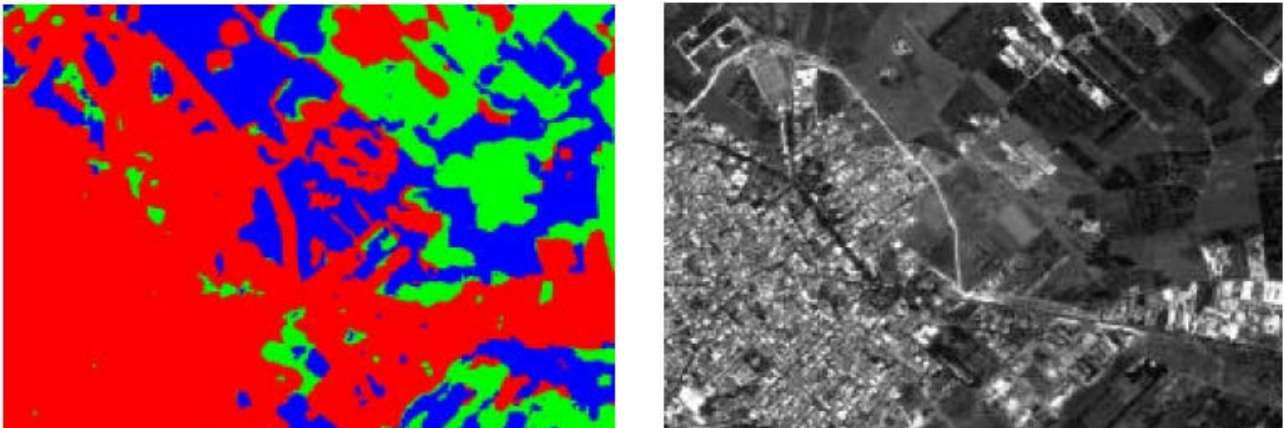


Figura 6 - Risultato della classificazione usando solo 6 coefficienti di Gabor sulla banda B

L'approccio proposto tende tuttavia a generare errori in prossimità di repentini cambi di tessitura (ad esempio in corrispondenza di piccole strade in zone extraurbane) e ciò è dovuto al fatto che tali regioni contengono informazioni tessiturali di più cluster.

In conclusione, la discriminazione delle diverse aree tematiche (a diverse scale di risoluzione) è avvenuta in maniera affidabile, precisa e con un esiguo dispendio computazionale. L'approccio proposto ha validità del tutto generale, per cui è facilmente adattabile alla classificazione di ulteriori aree tematiche previo addestramento del sistema di classificazione supervisionato.

### Riferimenti bibliografici

- Weldon T., Higgins W.E., Dunn D.F. (1996), "Efficient Gabor-filter design for texture segmentation," *Pattern Recognition*, vol. 29, no. 12, 2005-2016
- Vandergheynst P., Frossard P. (2001), "Efficient image representation by anisotropic refinement in matching pursuit," *Proc. IEEE Conference ICASSP Salt Lake City (UT)*, vol. 3, 1757-1760
- Phillips P.J. (1998), "The design of matching pursuit filters," *Computation in Neural Networks*, vol. 9, no. 1, 1-17
- Mallat S., Zhang Z. (1993), "Matching pursuit with time-frequency dictionaries," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, 3397-3415
- Bergeaud F., Mallat S. (1995), "Matching pursuit of images," *Proc. of IEEE International Conference on Image Processing Washington DC*, vol. 1, 53-56
- Papoulis A. (1984), "Signal analysis," *McGraw-Hill International Editions Auckland*
- Kohonen T. (1997), "Self-Organizing Map," *Second Edition, Springer*, 203-217