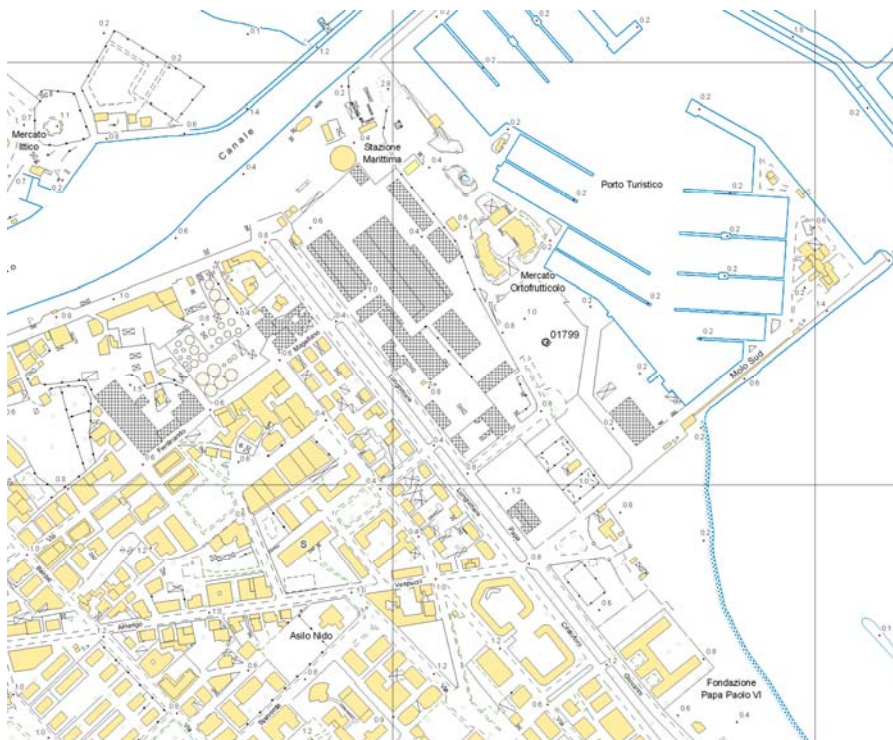


# Cartografia numerica

La cartografia numerica (CN) prevede la creazione di archivi di coordinate che descrivono la geometria degli oggetti rappresentati e delle codifiche che ne individuano la tipologia e/o geometria. Si tratta di dati di tipo vettoriale che possono essere stampati su carta o visualizzati a video e si contrappongono ai dati raster, tipici delle immagini digitali<sup>11</sup>.

La lettura del dato cartografico-numerico è basata sulle proprietà numeriche e geometriche associate allo stesso dato, mentre nel caso della cartografia tradizionale cartacea (ma anche rasterizzata tramite scanner) è l'utente che interpreta gli aspetti grafici grazie alla simbologia associata. In figura l'esempio di una stampa di cartografia nel formato numerico.



La transizione verso la cartografia numerica risolve alcuni aspetti negativi della cartografia tradizionale tra i quali:

- ✓ *univocità del contenuto metrico*: non c'è alcun margine di interpretazione del dato, eliminando i problemi connessi alla soggettività della misura per determinare le coordinate di punti ed alla deformazione del supporto cartaceo sul quale è stampata la mappa;
- ✓ *nessuna esigenza di limitare l'estensione dell'area rappresentata*: in questo caso il limite risiede nella possibilità di memorizzare quantità più o meno voluminose di dati;
- ✓ *rappresentazione tridimensionale del dato*: il dato planimetrico e quello altimetrico possono essere gestiti allo stesso tempo attraverso l'inserimento di coordinate tridimensionali. Ovviamente si può utilizzare anche CN solo planimetrica;

---

<sup>11</sup> Per una definizione di dato raster e vettoriale vedere l'ALLEGATO 1 in fondo alle dispense

- ✓ *possibilità di eseguire calcoli*: il dato numerico consente per via automatica di calcolare superfici, volumetrie, parametri urbanistici, calcoli statistici e selezione di dati con interrogazione dell'archivio (database);
- ✓ *aggiornamento delle informazioni*: i dati di CN possono essere continuamente integrati con aggiornamenti successivi purché il dato aggiunto rispetti gli standard richiesti sulla qualità dei dati e sul formato.

Le *elementi geometrici* fondamentali nella CN sono punti, linee o poligoni (con relativa codifica) che possono essere organizzati per rappresentare oggetti di varia complessità (*entità*). Si definisce poi *oggetto* un qualunque elemento naturale o artificiale che non sia ulteriormente suddivisibile.

La CN, con la corretta strutturazione dei dati in termini geometrici e di codifica, rappresenta l'elemento base per la creazione dei Sistemi Informativi Geografici che verranno discussi nell'ultima parte di queste dispense. In questo caso diventa basilare l'utilizzo di coordinate planimetriche espresse in un ben definito Sistema Cartografico (ad esempio GB o UTM). L'evoluzione di questa disciplina ha poi portato alla creazione di basi di dati che, per sfruttare al massimo le potenzialità della CN nell'ambito dei Sistemi Informativi, sono di tipo "object oriented" ovvero non più basati su entità geometriche con attributi associati ma su entità autoconsistenti che includono (per definizione) gli attributi relativi agli oggetti stessi. Per esempio di un edificio potrà sempre essere calcolata la volumetria o visualizzata la tipologia in quanto tale dato è già contenuto nella nuova "entità oggetto".

## ***Il concetto di scala nella CN***

Nella CN il concetto di scala sembrerebbe superato in quanto i punti sono memorizzati in coordinate assolute (di un qualche sistema di riferimento cartografico o arbitrario). Pertanto il rapporto di scala è sempre del tipo 1:1. In realtà il dato di CN è sempre associato a parametri di accuratezza e precisione che dipendono dalla tecnica che ha fornito il dato originale e dalle modalità operative che lo hanno generato. In particolare le tecniche che producono CN sono:

- ✓ *restituzione fotogrammetrica*;
- ✓ *digitalizzazione di mappe cartacee*;
- ✓ *digitalizzazione a video di prodotti raster* (dispositivi di puntamento e digitizers);
- ✓ *digitalizzazione con inseguimento automatico*;
- ✓ *rilievo topografico diretto sul terreno* (tradizionali e Global Positioning System, GPS<sup>12</sup>);
- ✓ *immagini telerilevate*

alcune delle quali verranno descritte successivamente. Pertanto l'informazione sull'origine del dato numerico non si deve assolutamente perdere in quanto da questa si può derivare l'accuratezza del dato e quindi il suo potenziale utilizzo.

Ad esempio, la possibilità di ingrandire l'immagine (zoom) a piacimento non assegna al dato un maggiore potere di definizione delle informazioni. Non avrebbe senso quindi plottare alle scale 1:2000 o 1:1000 un dato di CN ottenuto da un rilievo concepito per cartografia alla scala 1:5000.

---

<sup>12</sup> Ulteriori informazioni sul sistema di posizionamento GPS si trovano nell'ALLEGATO 2 alle dispense.

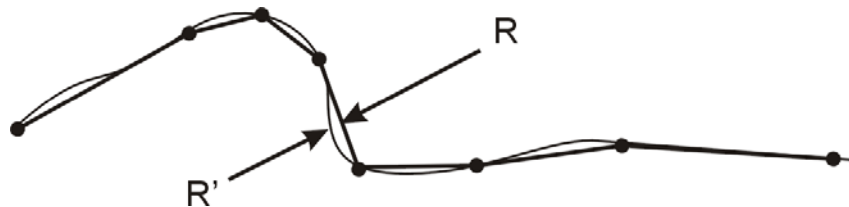
Il concetto di scala, quindi, non si perde totalmente nella CN, ma potrà essere definito come il massimo rapporto di scala col quale plottare una carta in modo tale che essa abbia gli stessi requisiti qualitativi e metrici di una carta tradizionale avente la medesima scala. Tale rapporto prende il nome di **scala nominale**.

La descrizione della qualità del dato è possibile attraverso alcuni parametri.

- ✓ *Precisione geometrica*: differenze tra la posizione di un punto nella c.n. e la sua posizione reale nel sistema prescelto;
- ✓ *Risoluzione*: dimensione del più piccolo particolare rappresentabile;
- ✓ *Precisione semantica*: corrispondenza tra attributo descrittivo dell'oggetto e denominazione reale;
- ✓ *Congruenza logica*: assenza di incongruenze rilevabili (ad esempio nelle curve di livello).
- ✓ *Congruenza geometrica*: assenza di errori nella forma (geometrica) degli oggetti, errori riscontrabili solo attraverso il confronto con dati rilevati.

## **Discretizzazione delle linee nella CN**

Gli elementi grafici che sono contenuti nella cartografia tradizionale (costituiti da un numero molto elevato di punti) devono essere discretizzati, nel modo più accurato ed affidabile, in elementi numerici. Infatti nella CN i concetti di linea e curva non esistono più e si parla invece di "spezzata" o "polyline". La spezzata in CN (immagine di una curva nella cartografia tradizionale) è identificata da un certo numero di vertici scelti in modo tale che un punto R sulla curva disti dal corrispondente R' sulla spezzata meno di un valore prefissato (tolleranza planimetrica).



L'aspetto delle linee spezzate può essere migliorato attraverso tecniche di "smoothing" che interpolano il dato fornendo un aspetto continuo della curva. Tuttavia questa tecniche presentano degli svantaggi tra cui l'alterazione del dato originale e la non univocità del prodotto finale (infatti possono essere applicati smoothing con parametri o forzature differenti in modo del tutto soggettivo).

Per conservare il carattere di fedeltà all'aspetto geometrico originale, la CN deve contenere in forma esplicita tutti i punti necessari per la sua corretta visualizzazione. Questi andrebbero separati dal set di punti ottenuti, per esempio, tramite tecniche di infittimento.

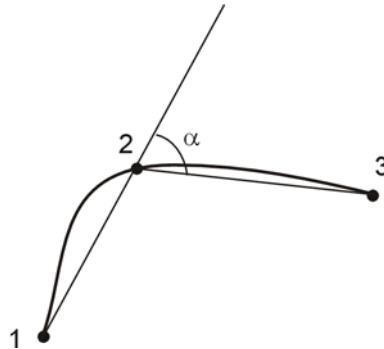
Nei metodi fotogrammetrici esistono tre modalità di discretizzazione delle linee curve:

**a spazio prefissato**: i punti che discretizzano una linea curva sono acquisiti ad intervalli spaziali regolari lungo direttrici preferenziali (anche nelle tre dimensioni).



**a tempo prefissato:** i punti che discretizzano una linea curva vengono campionati ad istanti temporali precisi (impostati dall'utente in funzione dell'accuratezza voluta).

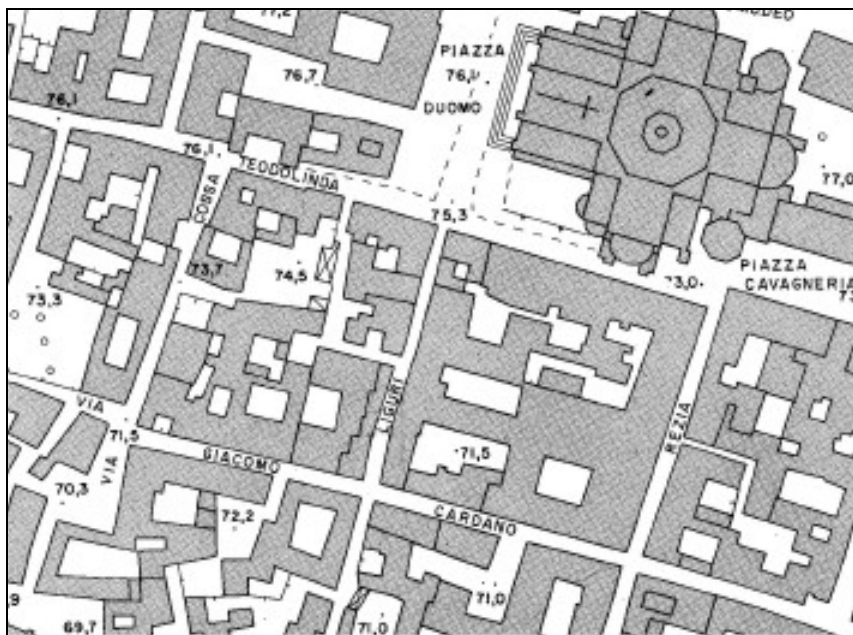
**a spazio e direzione prefissata:** in questa modalità possono essere discretizzate curve dall'andamento disomogeneo. Pochi punti saranno necessari per tratti rettilinei mentre punti a maggiore densità descriveranno le porzioni curvilinee. Con questa tecnica viene registrato un punto ogni volta che l'angolo formato tra il segmento già registrato ed il successivo preso in esame (per esempio a spazio prefissato) supera un certo limite.



A questo criterio ne possono essere associati altri quali ad esempio un minimo valore di spazio percorso necessario per procedere all'inserimento dei punti.

## ***Contenuto planimetrico della cartografia numerica***

Nella CN tutte le entità che venivano riconosciute ed interpretate dall'utente nei prodotti tradizionali, devono essere rappresentate da geometrie ben definite ed associate a codifiche prestabilite. Nella memorizzazione di uno spazio urbanizzato occorrerà che strade e incroci vengano memorizzati secondo geometrie codificate (in base alla tipologia) come "tronco" e "nodo" rispettivamente.



Non devono esistere entità non codificate e le aree possono essere suddivise in superfici omogenee in base all'utilizzo che viene fatto. Tra le varie categorie si possono trovare:

ISOLATI: superficie delimitata da una spezzata chiusa rappresentativa dei lati di edifici, muri di recinzione, marciapiedi ecc.

SERVIZI: zone destinate ad uso particolari e stabili (stazioni ferroviarie, impianti industriali, centrali elettriche, centri commerciali ecc.

AREE VERDI: aree destinate al verde urbano, e non, ed alle iniziative che qui possono svolgersi (sportive, ricreative, ecc.)

SPAZI APERTI: sono le aree destinate all'uso agricolo o incolte, ad attività di estrazione (cave), zone di esondazione ecc.

SPECCHI D'ACQUA: superfici con corpi idrici quali fiumi, laghi, invasi, darsene, porti ecc.

TRONCHI: tratto di strada delimitato da spazi di altro tipo

NODI: superficie relativa ad una piazza, un largo, incroci di più vie ecc.

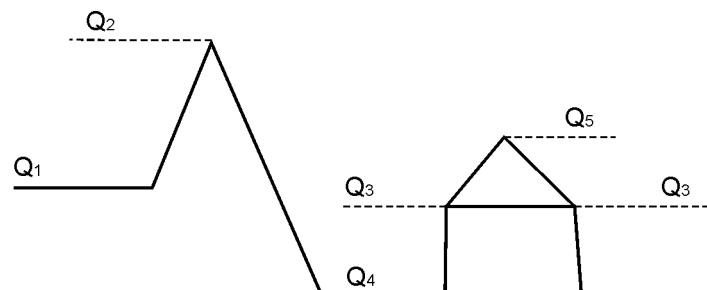
Il nodo dovrebbe essere memorizzato come spezzata chiusa con lati reali (perimetri di superfici che lo delimitano) o virtuali (linee che delimitano i tronchi o nodi adiacenti).

## **Contenuto altimetrico della cartografia numerica**

Uno dei vantaggi principali della CN è la possibilità di superare la divisione tra dato planimetrico e dato altimetrico che invece caratterizza la cartografia tradizionale. Si ricorda che le superfici di riferimento utilizzate nella cartografia classica sono diverse nel caso planimetrico (con coordinate definite sull'ellissoide di rotazione o sulla sua proiezione sul piano) ed in quello altimetrico (con quote ortometriche definite rispetto al geoide ovvero al livello medio marino).

Nella CN l'informazione altimetrica può esprimersi come *punti quotati*, *curve di livello* o come *attributi* associati ad elementi geometrici planimetrici. Per punti quotati e curve di livello l'analogia con la cartografia tradizionale è immediata mentre la presenza di una quota espressa come attributo di altri elementi geometrici rappresenta una novità.

I punti che definiscono un prodotto di CN sono memorizzati con una quadrupletta di numeri ovvero [Codice Oggetto, Est. Nord, Quota]. Anche se l'oggetto rappresentato ha estensione verticale, la quota che lo rappresenta deve essere quella del terreno. Per alcune entità si possono prevedere delle eccezioni prevedendo una quotatura sia al piede che sulla sommità. Nella figura sotto le quote Q1 e Q2 sono riferite al terreno mentre per l'entità cartografica "edificio" si possono prevedere anche quotature intermedie (Q3) e della sommità (Q5).



Inoltre per le unità volumetriche si possono indicare le quote di gronda come entità puntuali all'interno dell'unità volumetrica con coordinate Est e Nord uguali al perimetro esterno e quote relative alla gronda.

## ***Rappresentazione dell'altimetria nei formati digitali***

L'altimetria di una porzione di territorio viene rappresentata nel formato numerico attraverso i *modelli digitali del terreno*. Questi sono formati da insiemi di punti per i quali sono note le coordinate cartografiche planimetriche (ma anche quelle geografiche ellissoidiche) e la quota ortometrica (oppure quella ellissoidica se i dati provengono, ad esempio, da un rilievo GPS).

La disponibilità del modello numerico del terreno per un'area di interesse è spesso fondamentale in quanto consente di rappresentare, in modo più o meno dettagliato, l'orografia della zona e quindi ottenere elaborati grafici quali curve di livello, profili altimetrici del terreno e sezioni, viste prospettiche e di effettuare inoltre determinazioni volumetriche richieste negli interventi di ingegneria sul territorio (infrastrutture, nuovi insediamenti, cave ecc.).

Occorre ricordare che la rappresentazione del territorio attraverso modelli numerici del terreno avviene sempre tramite un numero discreto di punti e quindi questa sarà sempre una approssimazione della realtà. Quanto più questi punti saranno densi ed accurati, tanto migliore sarà la definizione delle reali caratteristiche morfologiche del terreno rappresentato. I punti che vanno a formare il modello numerico del terreno in origine possono essere "sparsi" in modo casuale o "regolari" a seconda del metodo di rilevamento che ha generato i punti stessi.

Da un modello di dati sparsi è sempre possibile ricavarne uno a maglia regolare, attraverso tecniche di interpolazione dei dati, che rendono più efficiente la memorizzazione dei dati.

In funzione della struttura dei dati contenuti nel modello e degli elementi del territorio che essi rappresentano si possono avere i<sup>15</sup>:

- Modelli Digitali della Superficie (DSM, *Digital Surface Model*)
- Modelli Digitali del Terreno (DTM, *Digital Terrain Model*)

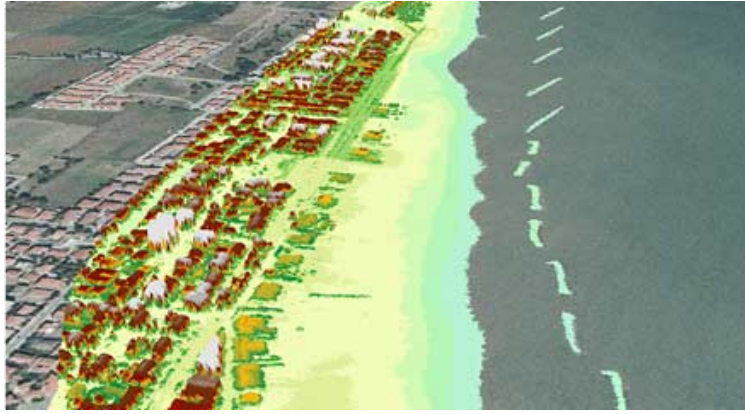
## **DSM e DTM**

Il DSM rappresenta la parte superiore di tutti gli oggetti presenti sul territorio, sia naturali che antropici. Quindi la quota dei punti che fanno parte di un DSM potrà essere riferita a tetti, manufatti, chiome di alberi oppure al suolo se questo non è coperto da vegetazione o altro. Questo tipo di prodotto è generato dai metodi di indagine laser da aereo (LIDAR) oppure da rilievi aerofotogrammetrici.

---

<sup>15</sup> Esperienze significative nella descrizione dei modelli digitali di elevazione in termini di definizione, produzione, accuratezza ed utilizzi sono state condotte da organizzazioni ed Agenzie di diversi paesi principalmente coinvolte nelle discipline del rischio idrogeologico e della modellizzazione del regime idraulico costiero e di ambienti continentali. Tra i principali contributi si segnalano quelli della *Federal Emergency Agency*, FEMA (*Guidelines and Specifications for Flood Hazard Mapping Partner*), *National Digital Elevation Program*, NDEP (*Guidelines for Digital Elevation Data*), *Ecole Polytechnique Federale de Lausanne*, EPFL (*Technical Specification for the Elaboration of Digital Elevation Model*).



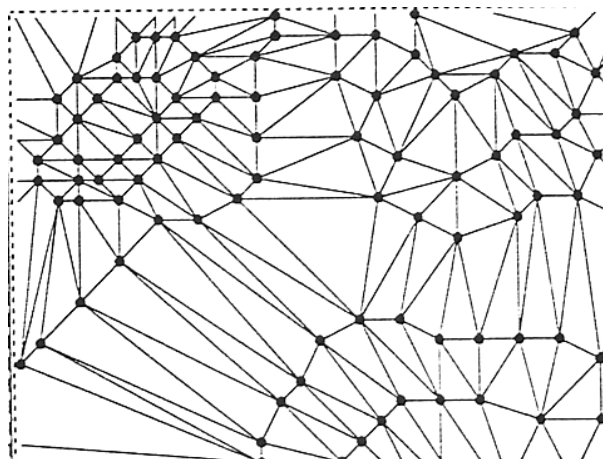


Il DTM invece rappresenta la quota del terreno nudo dopo che elementi quali edifici, vegetazione, manufatti ecc. sono stati rimossi. I corpi idrici sono rappresentati attraverso la quota della superficie dell'acqua, anche se spesso in ambiente costiero è richiesta la batimetria dei fondali. In alcuni casi al DTM si possono aggiungere le superfici dei manufatti, ottenendo un prodotto misto. E' il caso delle indagini sui rischi legati all'esondazione di corpi idrici dove i manufatti sono considerati degli ostacoli alla circolazione dell'acqua.

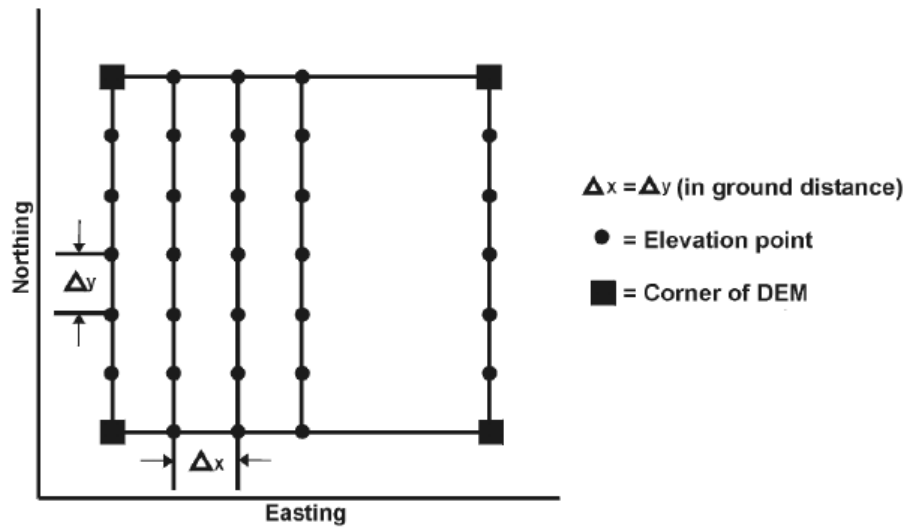
I DTM e DSM possono essere rappresentati in vario modo tra cui: punti sparsi, TIN, grid.

La rappresentazione attraverso "dati sparsi" può rivelarsi utile quando occorre definire con maggior dettaglio alcune zone rispetto ad altre. Se invece sono presenti elementi di particolare interesse nella definizione del territorio (aste fluviali, sede stradali rialzate, discontinuità naturali..) questi saranno inseriti separatamente attraverso le cosiddette *breaklines*, elementi geometrici vettoriali che descrivono accuratamente quegli elementi che potrebbero non essere inseriti nel DTM o DSM con sufficiente accuratezza (i due sono solo una rappresentazione discreta del territorio).

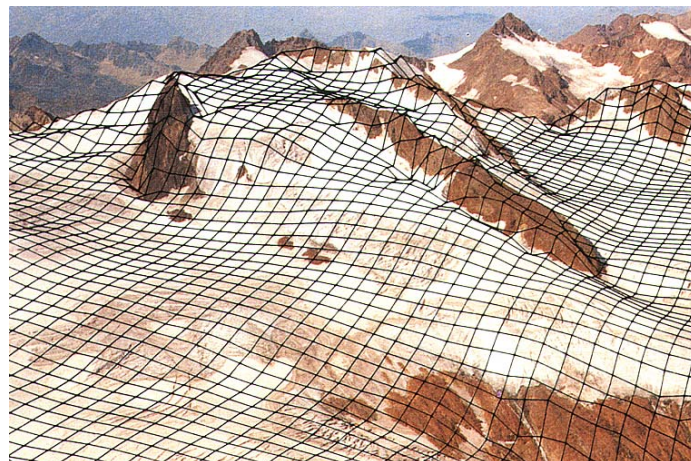
Da un gruppo di dati sparsi è possibile creare una superficie attraverso la tecnica TIN (*Triangulated Irregular Network*) che prevede la creazione di superfici triangolari tra i punti sparsi senza operazioni di interpolazione che "falsificano" l'informazione originale.



Attraverso i *grid* invece i dati vengono inseriti secondo una disposizione a maglia regolare. Ogni nodo contiene l'informazione della quota e sarà separato dal nodo adiacente da una certa distanza (pixel spacing o post-spacing). Questa distanza definisce la risoluzione del modello e si quantifica in metri- terreno.

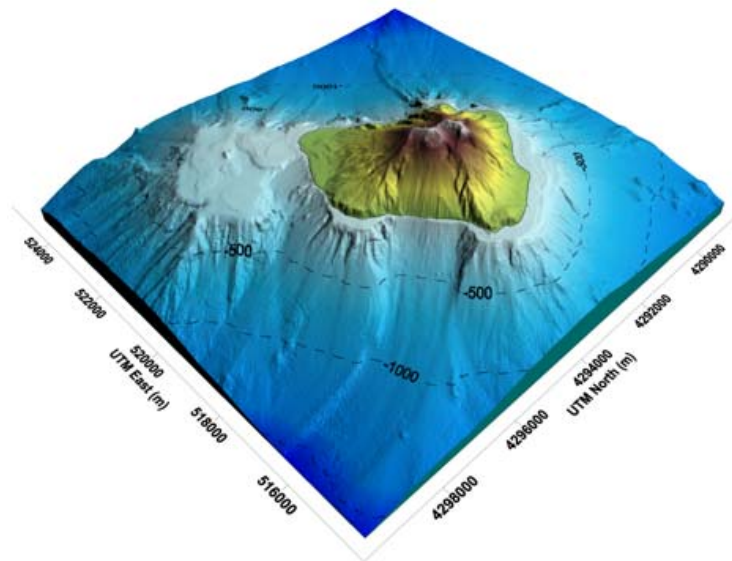


Una maglia regolare di questo tipo può essere visualizzata in 3D e fornire un modello altimetrico del terreno in grado di simulare la realtà e visualizzabile secondo una certa prospettiva.



I software per il trattamento dei dati altimetrici consentono di produrre modelli 3D, inquadrati in sistema di riferimento cartografico, applicando anche *palette* di colori ed effetti di ombreggiature (tecnica *shaded relief*) che rendono più accattivante la visualizzazione del modello accentuando gli aspetti morfologici. In figura il modello 3D (anche batimetrico) dell'isola di Stromboli (arcipelago delle Isole Eolie).





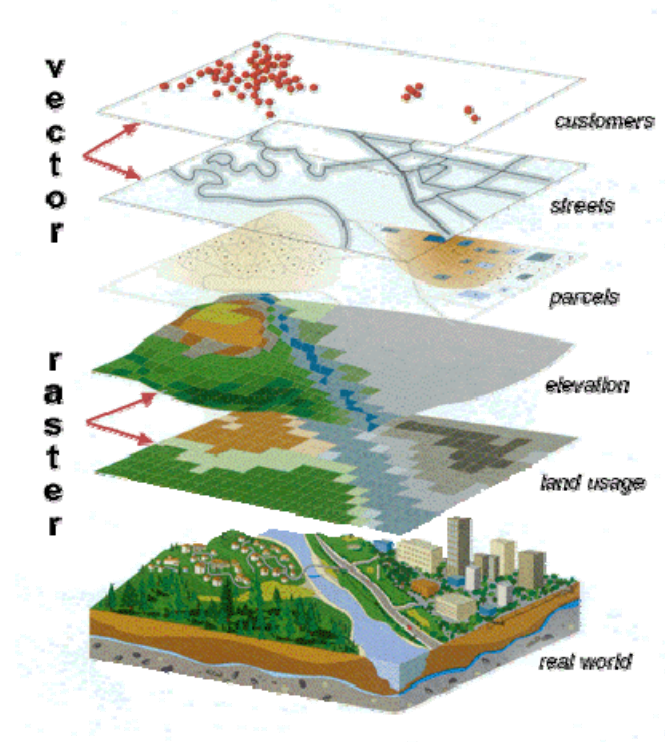
Da questi modelli possono essere ovviamente estratte informazioni come sezioni lungo direttrici particolari, volumi. Analogamente alla cartografia tradizionale, l'altimetria può essere rappresentata tramite isoipse calcolate tramite interpolazione di dati sparsi o regolari. Insieme alla cartografia (raster o vettoriale), il modello altimetrico rappresenta la base di dati geografici per la realizzazione dei Sistemi Informativi Territoriali.

La qualità di un DTM o DSM, oltre ad essere funzione della misura del *post-spacing*, dipende dall'accuratezza del valore della quota che a sua volta dipende dalla tecnica che ha generato il dato. Questo valore può essere valutato sulla base di una serie di punti a quota nota (Ground Control Points), ottenuti per altra via, che vengono confrontati con il modello del terreno. In assenza di errori sistematici i residui ottenuti in corrispondenza dei punti di controllo saranno distribuiti secondo una curva gaussiana con valore medio pari a zero. Si può utilizzare lo scarto quadratico medio come indicatore dell'accuratezza verticale. Evidentemente un modello del terreno dovrà avere valori di *post-spacing* ed accuratezza verticale idonei alle applicazioni che ne prevedono l'uso.

# I Sistemi Informativi Territoriali (SIT)

Come si è visto, l'evoluzione delle tecniche informatiche, nell'ambito delle discipline cartografiche, ha portato all'introduzione della CN ad integrazione di quella tradizionale. Parallelamente a questa rivoluzione sono emersi nuovi approcci e nuove idee sull'analisi spaziale di dati geografici (ovvero riferiti al territorio attraverso la definizione della posizione geografica). Il tutto è reso possibile dal formato digitale (vettoriale o raster) dei dati che si presta, soprattutto se collegato a banche dati, al calcolo.

Questo nuovo approccio alla conoscenza del territorio ha diverse origini; nel 1969 Ian McHarg, nel suo libro "Design with nature", ipotizzava l'uso della sovrapposizione (*overlay*) di dati geografici che venivano strutturati in livelli informativi a singolo tematismo.



Si ricorda che nella cartografia tradizionale invece tutte le informazioni vengono raccolte in un unico livello informativo rappresentato dal supporto cartaceo.

Nel 1967 compare per la prima volta il termine Geographical Information System (Sistemi Informativi Geografici) nel libro "An introduction to the geographic information system of the Canada land inventory" per indicare un sistema informativo sviluppato per gestire l'enorme quantità di dati raccolti sul territorio canadese. Le prime applicazioni furono avviate negli US negli anni '70, quando l'*Harward Laboratory* produsse ODYSSEY, un precursore dei software ad uso commerciale che introduceva il concetto di struttura topologica dei dati e di *overlay mapping*.

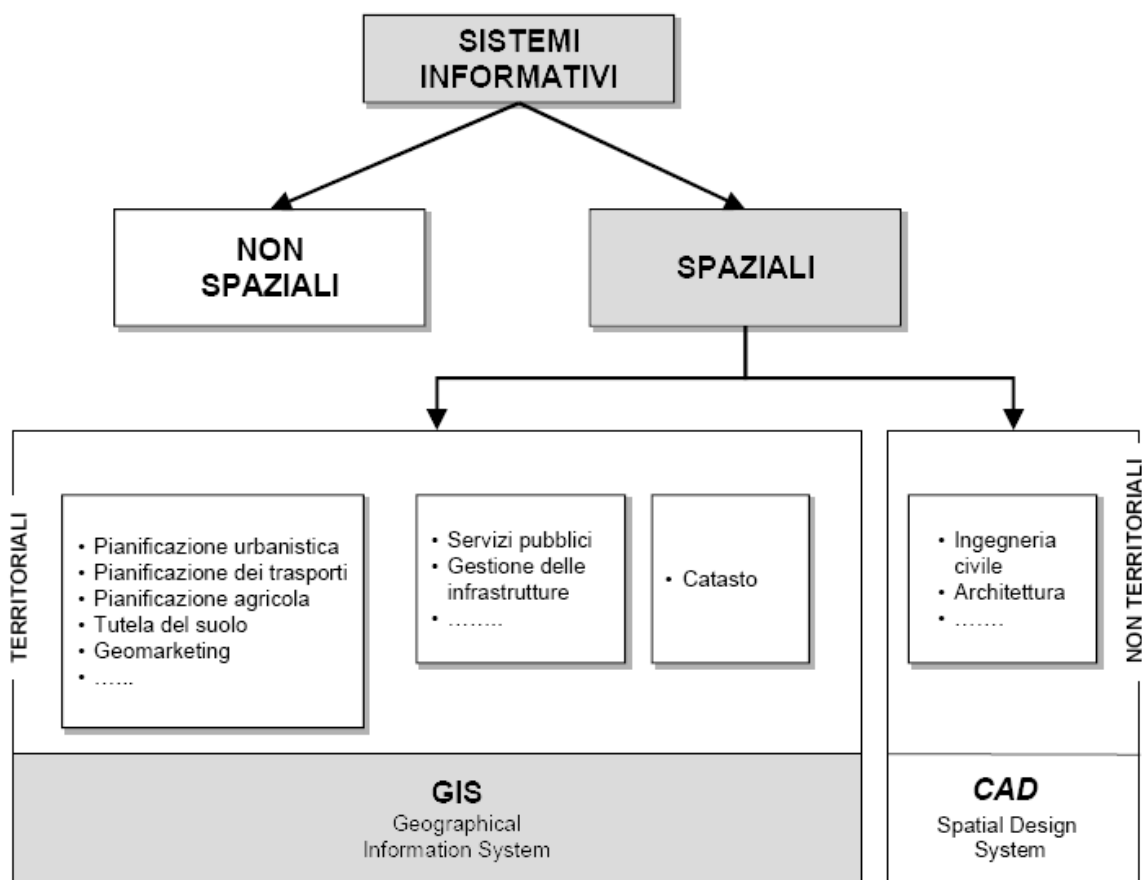
Il *Canadian Geographical Information System* (CGIS), pensato per la creazione di una cartografia tematica idonea alle analisi territoriali, è stato uno dei primi sistemi informativi sviluppati ed è ancora utilizzato (ovviamente con gli aggiornamenti necessari).

A livello di produzione invece gli strumenti pensati per applicazioni GIS venivano di solito utilizzati ai fini della gestione di dati sottoforma di CN (ad esempio la con produzione di

mappe tematiche la cui creazione non richiede necessariamente l'ambiente GIS) senza le funzioni di consultazione ed analisi dei dati, caratteristiche dei GIS.

Lo sviluppo dei Sistemi Informativi è legato anche alla disponibilità di nuove tecniche di acquisizione dei dati territoriali tra cui citiamo il telerilevamento satellitare ed il Global Positioning System (GPS). La cultura dei GIS è diffusa in Italia da quasi un ventennio (notevole l'impulso dato da Internet) con la diffusione capillare presso Enti locali, Università e privati.

Il termine **Sistema Informativo Territoriale** (SIT) non indica la sola traduzione di Geographical Information System (GIS). Il SIT può essere inteso come l'insieme dei dati geografici organizzati per un relativo territorio, mentre il GIS rappresenta lo strumento di lavoro (ambiente) attraverso il quale gestire tali dati. Il GIS in sostanza supporta la creazione del SIT. Il SIT invece viene spesso indicato come una "verticalizzazione" dei GIS.



Entrambi nascono dall'esigenza di disporre di strumenti per la **raccolta, aggiornamento, analisi, visualizzazione ed editing** dei dati. Lo strumento è pensato principalmente come supporto delle fasi decisionali nel governo del territorio (si pensi ad esempio all'utilità nelle fasi di pianificazione urbanistica, dei trasporti e di intervento sul territorio) ma le applicazioni sono infinite.

Il valore dell'informazione contenuta in un SIT è legato:

1. al livello di aggiornamento dei dati;
2. alla presenza di un archivio dati (database) correttamente definito per le applicazioni di interesse;
3. alla possibilità di definire entità geografiche ed attributi associati;
4. alla possibilità di mettere in relazioni i dati tra loro;
5. alla coerenza dell'informazione presente nei differenti archivi ed alla corretta descrizione del dato stesso;
6. alla disponibilità di informazioni relative alla qualità ed alla storia del dato (ad esempio la descrizione delle procedure e degli strumenti utilizzati nell'acquisizione dei dati).

Anche se è possibile stabilire varie relazioni tra i dati inseriti in un SIT, non si deve dimenticare che la prima caratteristica del dato è la sua posizione sul territorio che deve essere definita in un certo sistema cartografico come tutte le altre informazioni. Ai dati numerici possono essere assegnate anche informazioni alfa-numeriche, ovvero testo o immagini che descrivono il dato stesso.

## ***Funzionalità e struttura dei Sistemi Informativi Territoriali***

La caratteristica principale dei SIT (e in generale degli ambienti di lavoro GIS) è quella di relazionare dati attraverso il riferimento spaziale, la G di *Geographical*, indica la posizione geografica come principale elemento che contraddistingue un'informazione (entità, geometria, dati alfanumerici...). Le informazioni inserite in un GIS hanno essenzialmente carattere:

1. amministrativo, censuario e socio-economico: con una componente spiccatamente numerica;
2. cartografico/geometrico: per la rappresentazione del territorio e delle opere esistenti con una componente grafica rilevante.

Le informazioni inserite in un SIT sono consultabili in vario modo anche se la forma più efficace è quella delle interrogazioni (*query*). Attraverso le *query* un utente chiede al sistema di estrarre le informazioni che rispondono ad una regola predefinita (si vedranno meglio più avanti).

**ESEMPIO** – Volendo verificare il numero di abitanti che vivono nei Comuni pugliesi della fascia costiera occorrerà, come minimo, la seguente base di dati (geometrici o relativi a database)

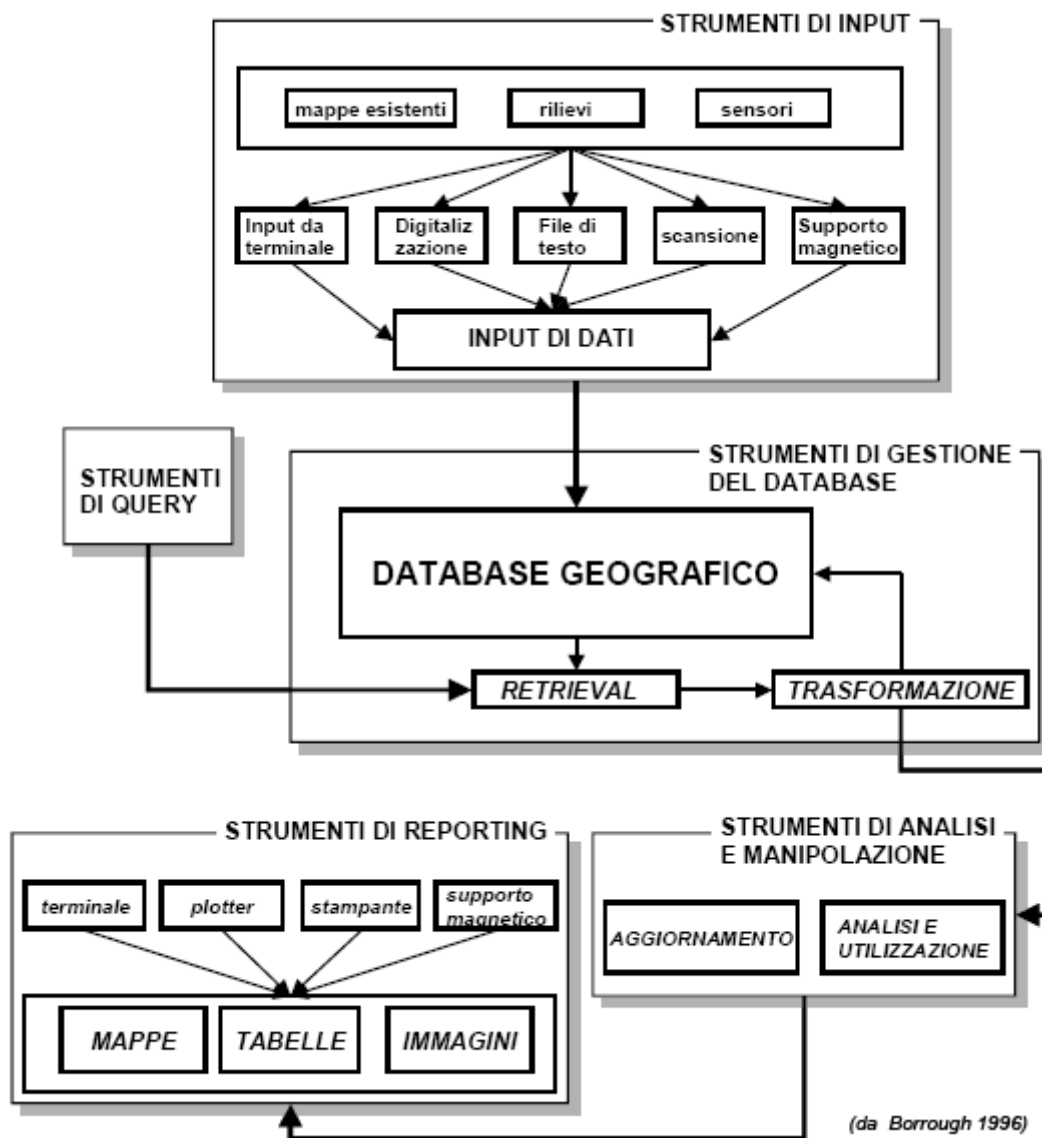
- mappa vettoriale relativa all'estensione dei Comuni regionali che hanno territori costieri;
- mappa vettoriale della linea di costa;
- dati demografici dei singoli comuni (fonte ISTAT)

la *query* da impostare sarà del tipo "*estrai dal database i comuni che si affacciano sul mare e riporta i suoi abitanti*". Tradotta in un linguaggio "macchina" questa diventa "ricerca dei Comuni che hanno confini tangenti alla fascia costiera (geometrie di contatto secondo regole da definire) ed estrazione dell'attributo "popolazione residente" associato al confine comunale. Le stesse informazioni consentiranno altre operazioni come il calcolo della superficie coinvolta, la densità di popolazione ecc. Il risultato potrà essere mostrato come mappa o in forma tabellare.

In generale quindi si può affermare che in ambiente GIS si possono gestire tutte le operazioni di governo del territorio e supportare le attività decisionali (ma anche applicare modelli di analisi più complessi), a patto che i dati inseriti siano idonei all'uso (in termini di ricchezza, accuratezza ed affidabilità) e l'ambiente di lavoro abbia le capacità di analisi richieste. Inoltre deve essere garantita la corretta georeferenziazione delle informazioni (raster, vettoriali ed alfanumerici) che quindi devono essere "riferite" ad un unico sistema di riferimento nel modo più accurato possibile.

Nella produzione di un SIT devono essere disponibili diverse componenti tecnologiche (prerogative dell'ambiente GIS) tra cui:

1. Strumenti per l'input dei dati
2. Strumenti per la gestione delle banche dati
3. Strumenti per le query
4. Strumenti di analisi dati
5. Strumenti per l'*editing* e la redazione di *reports*



# ALL.1: Le strutture dati Raster e Vettoriale

Si è detto che la scansione di immagini fornisce un dato di tipo RASTER mentre il disegno di geometrie particolari, come avviene nella redazione di CN, fornisce dati di tipo VECTOR (vettoriale). In generale il dato di tipo **raster** nasce da dispositivi quali fotocamere digitali, sensori per il telerilevamento, scanner, ecc; mentre il **vettoriale** può derivare sia da un processo di vettorializzazione secondo varie modalità, sia prodotti software di tipo CAD o GIS oppure da restitutori digitali utilizzati nella fotogrammetrie (argomento non trattato).

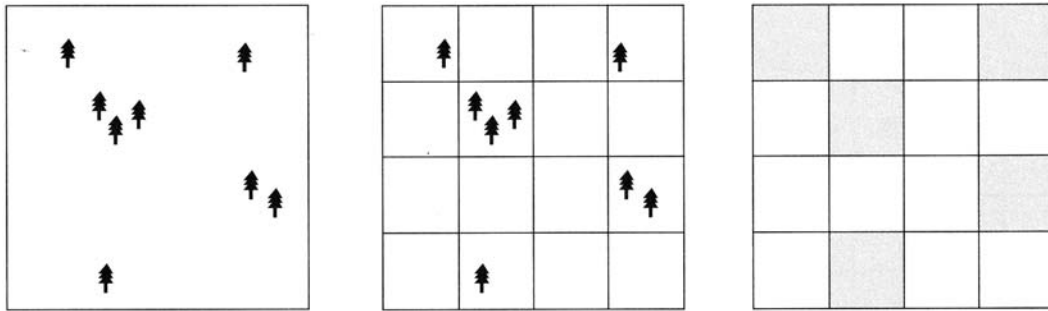
Entrambe le due tipologie di dato offrono informazioni utili nel campo dell'analisi spaziale e vanno considerate complementari e sinergiche: nessuna delle due può sostituire completamente le funzionalità applicative dell'altra. Le due strutture dati sono gestite negli ambienti GIS anche se, come si vedrà, esistono diversi formati disponibili che non sempre sono previsti nei software.

## STRUTTURA RASTER

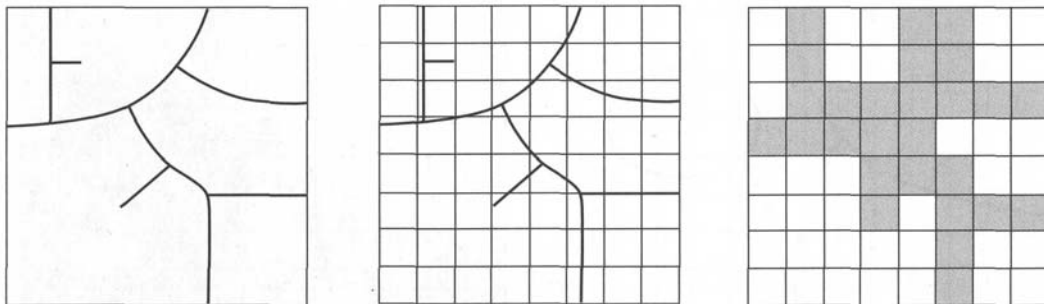
Il *raster* è considerato la più semplice e intuitiva struttura di organizzazione di dati geometrici ed è largamente usata nel campo dei Sistemi Informativi Territoriali. Il dato *raster* è costituito da una matrice di punti (o celle) di grandezza uniforme (*pixel*). A ciascun pixel viene associato un attributo che definisce le caratteristiche dell'elemento rappresentato. Per esempio al pixel potrà essere associato un colore oppure un attributo di altro tipo (come ad esempio la quota media dell'area compresa nel pixel). La matrice potrà essere rappresentata come uno spazio cartesiano piano (x,y), all'interno del quale il singolo pixel viene localizzato dal numero di riga o di colonna dello stesso. Come tutte le matrici, la posizione di ogni cella è riferita ad un'origine che generalmente è la cella in basso a sinistra, anche se talvolta, per esigenze di compatibilità con i sistemi di computer grafica, è considerata quella in alto a sinistra.

Il vantaggio di questa struttura è rappresentato essenzialmente dalla sua semplicità che ne permette una facile gestione ed elaborazione; la struttura matriciale consente infatti di modificare il contenuto o di effettuare operazioni aritmetiche di qualsiasi tipo, anche fra matrici diverse, per ricavarne una nuova che sintetizza il risultato. Alle caratteristiche positive si contrappone l'inevitabile perdita di accuratezza causata dalla rappresentazione in forma discreta di un'informazione che, nella realtà e nel disegno cartaceo, sarebbe continua. Nel dato raster, inoltre, non è possibile discriminare elementi che cadono nello stesso pixel (perché questo ha una sola proprietà associata). Nella figura si vede come elementi definiti nella cartografia di partenza (gli alberi) vengono persi dopo la "rasterizzazione" della carta. La scansione della carta quindi deve essere fatta sempre ad una *risoluzione* (dimensione del pixel) idonea a conservare gli elementi grafici che ci interessano.





Lo stesso esempio si può fare per elementi lineari tipo la viabilità.



La dimensione della cella determina anche quella della matrice e quindi la quantità di memoria necessaria. E' facilmente intuibile come una base di dati raster possa occupare molta memoria: se si rappresenta un territorio di 100×100 Km in una matrice con pixel di 10 m di dimensione, avremo 100 milioni di celle, se in ogni cella scriviamo un valore da 0 a 255 (sono necessari 8 bit per cella), l'occupazione di memoria equivale a 800 milioni di bit (Mb).

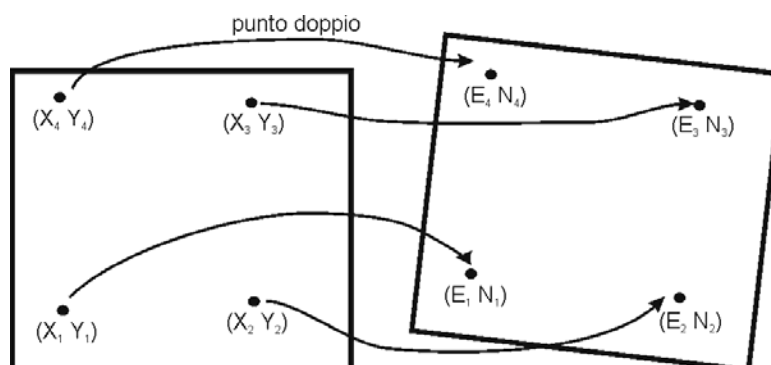
Al fine di ottimizzare la memoria occupata sono stati sviluppati diversi strumenti capaci di comprimere l'informazione. I *formati raster compressi* sono generalmente realizzati con software che comprimono i dati ottimizzando l'organizzazione delle informazioni quando devono essere archiviate. All'apertura del file i dati vengono rigenerati applicando all'inverso la procedura di compressione. Il metodo di compressione più comune archivia la matrice associando i pixel che hanno lo stesso valore di attributo. Procedure di questo tipo diventano inefficienti quando i valori nei pixel cambiano continuamente, come nelle immagini telerilevate da satellite.

## Georeferenziazione di dati raster

La scansione di cartografia, a varia risoluzione, produce un file raster dove la matrice di pixel è caratterizzata da un sistema di coordinate "locale" o "immagine". I pixel vengono semplicemente posizionati a partire da uno spigolo al quale saranno assegnate coordinate immagine ( $x=0$ ,  $y=0$ ). Tutti gli altri pixel della matrice avranno coordinate relative a tale origine.

La georeferenziazione consiste nel riassegnare ad ogni cella un valore di coordinate Nord ed Est in uno dei sistemi di riferimento cartografici piani (UTM, GB nel caso italiano).

$$(x, y) \rightarrow (N, E)$$



Il sistema di coordinate immagine (x, y) è anche paragonale ad un qualunque sistema di riferimento locale. La georeferenziazione si può effettuare anche per passare da un sistema piano ad un altro, ad esempio da GB ad UTM.

Si possono utilizzare algoritmi di georeferenziazione diversi ma che comportano delle trasformazioni, tra sistemi di riferimento, le quali richiedono la determinazione dei parametri relativi alla trasformazione stessa. I parametri pertanto rappresentano l'incognita del problema e vanno determinati a partire dai cosiddetti "punti doppi" ovvero quei punti dei quali si conoscono le coordinate nei due sistemi coinvolti, quello immagine e quello cartografico. Mentre le coordinate immagine sono note dal file raster, quelle cartografiche andranno lette dalla cartografia o rilevate attraverso metodi topografici. Ovviamente i punti doppi dovranno essere ben riconoscibili sia sul raster iniziale sia sulla mappa utilizzata come fonte di coordinate.

Tra i possibili metodi di georeferenziazione citiamo la trasformazione di Helmert piana (rototraslazione piana con variazione di scala). Questa è una trasformazione che sul piano prevede 4 incognite:

- due traslazioni,  $(E_0, N_0)$
- una rotazione,  $\alpha$
- un fattore di scala uniforme nei due assi,  $k$

La trasformazione di Helmert si può scrivere in forma matriciale

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} = \begin{bmatrix} \bar{X}_0 \\ \bar{Y}_0 \end{bmatrix} + \lambda R \begin{bmatrix} X \\ Y \end{bmatrix} \rightarrow \begin{bmatrix} E \\ N \end{bmatrix} = \begin{bmatrix} E_0 \\ N_0 \end{bmatrix} + \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

la quale richiede 4 equazioni per avere una soluzione ovvero la conoscenza di due punti doppi. E' bene avere una ridondanza di punti doppi per arrivare ad una stima dei parametri attraverso il metodo dei minimi quadrati dopo linearizzazione delle equazioni.

Sviluppando la forma matriciale per uno dei punti doppi si ottiene

$$\begin{aligned} E_1 &= E_0 + (k \cos \alpha) x_1 - (k \sin \alpha) y_1 \\ N_1 &= N_0 + (k \sin \alpha) x_1 + (k \cos \alpha) y_1 \end{aligned}$$

si può linearizzare sostituendo  $a = k \cos \alpha$ ,  $b = -k \sin \alpha$  ottenendoli sistema

$$\begin{aligned} E_1 &= E_0 + ax_1 + by_1 \\ N_1 &= N_0 - bx_1 + ay_1 \end{aligned} \rightarrow \begin{bmatrix} E_1 \\ N_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_1 & y_1 \\ 0 & 1 & y_1 & -x_1 \end{bmatrix} \begin{bmatrix} E_0 \\ N_0 \\ a \\ b \end{bmatrix}$$

che corrisponde a due equazioni con 4 incognite  $(a, b, E_o, N_o)$ . Con almeno 3 punti si raggiunge una ridondanza di equazioni e i parametri si possono stimare attraverso la minimizzazione dei residui che insorgono.

Altra possibile trasformazione che si può applicare è la Trasformazione affine a 6 parametri. Si tratta di una rototraslazione con variazione di scala anisotropa (ovvero diversa nelle due direzioni) e scorrimento angolare

$$\begin{pmatrix} E \\ N \end{pmatrix} = \begin{pmatrix} E_o \\ N_o \end{pmatrix} + \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

dove le incognite sono  $E_o, N_o, k_1, k_2, \gamma, \alpha$

Esistono anche altri modelli per la georeferenziazione che non vengono discussi. Tra questi alcuni utilizzano algoritmi polinomiali che negli ordini inferiori si riconducono a casi lineari più semplici.

## Alcuni tipi di formati RASTER

I dati raster possono essere memorizzati in strutture dati differenti che danno origine a formati diversi del file (il formato del file si desume dall'estensione associata al file stesso). Esistono molti diversi formati, in genere associati ai software per il trattamento dei dati (spesso di elaborazione immagini). Alcuni di questi sono di largo utilizzo e possono essere considerati degli "standard". Molti poi prevedono la compressione dei dati basata su algoritmi che riducono lo spazio di memoria occupato dall'immagine senza compromettere la qualità dei dati. I metodi di compressione sono molti, in genere si basano su procedure che raggruppano i dati simili ottimizzando la fase di memorizzazione dei dati. Ad esempio, considerata una figura in bianco (B) e nero (N), la prima riga di una matrice con BBBBNNNNNNNNNNNNBBBBB, può essere memorizzata come 7B3N1B10N5B senza perdita di informazioni. I principali formati non compressi, di interesse per i sistemi Informativi, sono:

- ✓ **tiff, tiff (+tfw), geotiff**
- ✓ **bmp**

Alcune compressioni invece comportano la perdita di informazioni, riducendo la qualità dell'immagine, che però avrà dimensioni notevolmente minori. Tra i formati compressi più diffusi si ricordano il **jpeg (+jgw)** e l'**ECW**.

## TIFF

Il **Tagged Image File Format** è un formato immagine di tipo raster flessibile, in grado di gestire immagini multiple e dati in un singolo file, attraverso l'inclusione di "tags" (modifiche) nel file header (intestazione del file contenuto nel file stesso). La modifica può indicare la geometria di base dell'immagine, il relativo formato, l'organizzazione dei dati, le eventuali compressioni usate. La capacità di comprimere i dati senza perdita di informazioni rende questo formato utile per l'archiviazione dei dati. Consente di rappresentare immagini in scala di grigio, RGB (Red, Green, Blue) e CMYK (Cyan, Magenta, Yellow, Black). Questo formato è riconosciuto da tutti i più diffusi programmi di elaborazione delle immagini. Il difetto di questo formato è l'elevata memoria richiesta nel salvataggio dei dati. Per questo motivo il formato TIFF non si presta all'utilizzo sul WEB ed all'invio via e-mail come attachment. Rappresentano invece il miglior formato per la pubblicazione di immagini su libri e riviste che richiedono un'alta qualità del dato.

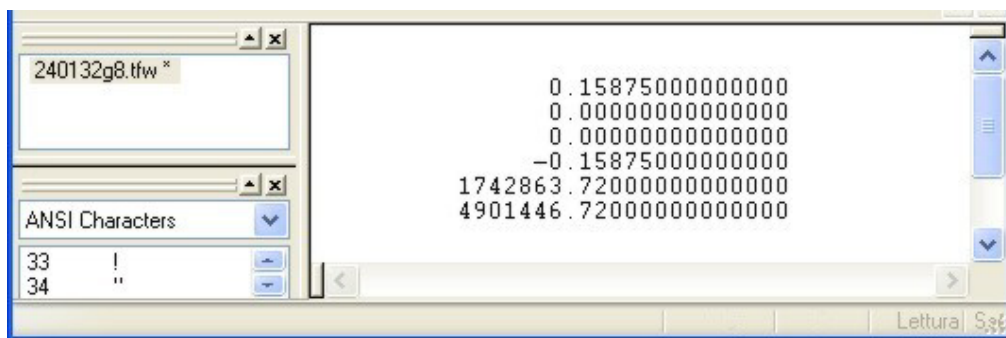
Questo formato non possiede un proprio formato di compressione, ma le immagini vengono in genere compresse con tecniche standard (LZW, RLE), le quali permettono di comprimere e decomprimere senza alcuna perdita di informazione.

### GeoTIFF e TFW

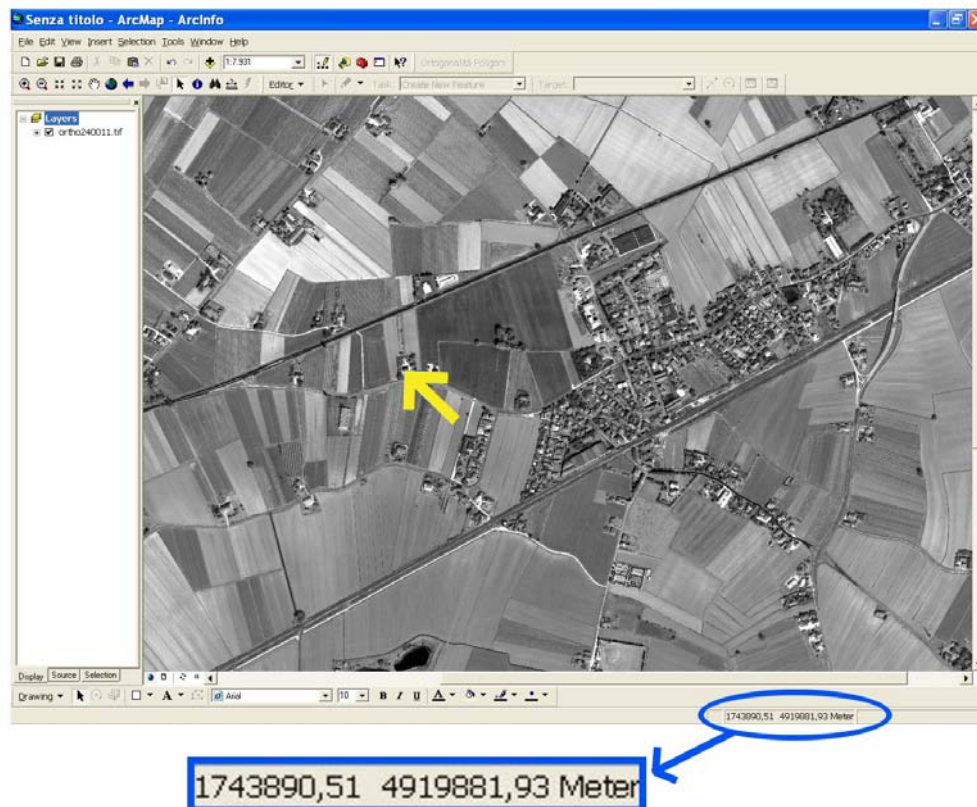
Il **GeoTIFF** rappresenta lo sforzo, compiuto da oltre 160 compagnie e organizzazioni, rivolto a stabilire un formato di scambio per le immagini raster georeferenziate basate sul formato TIFF. La georeferenziazione è data solitamente dalla posizione di un pixel (di solito uno ad un vertice dell'immagine) nel sistema geografico di riferimento; in tal modo è possibile sovrapporre dati geografici raster con quelli geografici vettoriali.

Il formato TIFF georeferenziato si può presentare in due soluzioni diverse:

- ✓ in un unico file con estensione **.tif**: è il formato definito **GeoTIFF** nel quale le informazioni relative alla georeferenziazione si trovano nell'*header* dello stesso file .tif;
- ✓ tramite una coppia di file di cui uno contenente l'immagine (file con estensione **.tif**) e l'altro le informazioni relative alla georeferenziazione (file con estensione **.tfw** chiamato "world file"). Il tfw è direttamente leggibile con un qualunque editor di testo e contiene in sequenza le seguenti informazioni: dimensione orizzontale del pixel (m), effetto di taglio, rotazione, dimensione verticale del pixel, coordinate Est e Nord del pixel in alto a sinistra.



Il vantaggio nell'utilizzo di formati raster georeferenzati è che all'apertura (con programmi in grado di interpretare questo tipo di dato) compariranno le coordinate dei pixel nella rappresentazione cartografica prescelta.



## BITMAP

Le **Bitmap** hanno generalmente l'estensione **.bmp**. Sono state sviluppate tre versioni di questo formato, tra le quali la prima (cioè la 3) è la più comunemente utilizzata. Il bitmap permette operazioni di lettura e scrittura molto veloci e senza perdita di qualità, ma richiede generalmente una maggior quantità di memoria rispetto ad altri formati analoghi.

Le immagini bitmap possono avere una profondità di 1, 4, 8, 16, 24 o 32 bit per pixel. Quelle con 1, 4 e 8 bit contengono una "tavolozza" per la conversione (rispettivamente 2, 16 e 256) dei possibili indici numerici nei rispettivi colori. Nelle immagini con profondità superiore il colore corrispondente ad ogni cella non è indicizzato bensì codificato direttamente nelle sue componenti cromatiche RGB attraverso una "mappa". Tale mappa dei pixel costituisce il corpo vero e proprio della bitmap.

Non è molto usata la possibilità di comprimere le bitmap da 16 e 256 colori con l'algoritmo RLE che non sembra molto efficiente. Il formato .bmp (almeno nelle versioni più diffuse) non supporta alcun tipo di trasparenza. Per ovviare a questo problema sono stati messi a punto diversi stratagemmi, come utilizzare un colore speciale per i pixel trasparenti, oppure servirsi di altre bitmap per mascherare le parti invisibili di un'immagine. Comunque il formato è semplice, ben documentato e non è tutelato da brevetti che ne restringano il libero uso. Soprattutto quest'ultimo punto ha fatto sì che il bitmap sia oggi supportato da quasi tutte le applicazioni grafiche, inclusi molti programmi *open source*.

## JPEG

**JPEG** è l'acronimo di **Joint Photographic Experts Group**, un comitato ISO/CCITT che ha definito il primo standard internazionale di compressione per immagini, sia a livelli di grigio che a colori. Attualmente il JPEG è il formato di compressione delle immagini fotografiche più utilizzato. La compressione del jpeg si basa sul principio che alcune gradazioni di colore non sono percettibili dall'occhio umano e di conseguenza possono essere ridotte di numero. Di seguito si riportano i diversi passaggi richiesti per la creazione di un'immagine in formato jpeg:

Alcune estensioni possibili sono le seguenti: .jpeg, .jpg, .jfif, anche se il più diffuso in tutte le piattaforme è **.jpg**.

1. Divisione dell'immagine in blocchi di 8×8 pixel, all'interno dei quali si effettua una media dei valori di radiometria, che viene salvata poi insieme ai numeri indicanti la variazione dalla media per ogni pixel;
2. Ognuno dei 64 valori ottenuti, all'interno di ogni blocco, viene diviso per un coefficiente e trasformato in un numero intero, senza contare eventuali decimali: è il momento in cui si elimina la maggiore quantità di dati ed è, di conseguenza, il processo che causa il maggior degrado dell'immagine;
3. Compressione di tutti i valori ottenuti ricorrendo ad un algoritmo non distruttivo e nella memorizzazione in un file, che contiene anche la descrizione dei coefficienti utilizzati nei processi di divisione e riduzione dei dati.

La compressione in genere riduce le informazioni relative al colore. JPG ha una sua versione georeferenziata rappresentata dall'associazione del .jpg con il relativo world file **.jgw** contenente le informazioni geografiche. Infine è stato recentemente introdotto il formato **JPEG 2000**, che consente di comprimere immagini in modo molto efficiente con una minima perdita di qualità. Nel JPEG2000 esiste anche la possibilità di applicare alla stessa immagine compressioni diverse da una zona all'altra, ad esempio più accurata nei punti d'interesse, più grossolana per le parti circostanti. È possibile poi associare metadati ad un'immagine. Nonostante queste caratteristiche conferiscano al formato JPEG2000 un chiaro vantaggio rispetto agli altri formati, la sua diffusione è ad oggi relativamente limitata perché non viene supportato dalla maggior parte dei browser.

## **ECW**

Il formato di compressione ECW permette di ridurre notevolmente la dimensione di dati raster (fino a 100 volte) e di visualizzare immediatamente le immagini compresse, senza tempi di attesa dovuti alla decompressione, come accade per altri formati. Il formato era previsto inizialmente nel pacchetto software ER Mapper, ma oggi i compressori sono anche esterni e gestiscono file con dimensioni anche dell'ordine di alcuni TeraByte. L'ECW Compressor consente di comprimere immagini anche nel formato **JPEG2000**. Il visualizzatore di immagini ECW è ERViewer, ma è anche possibile aprire immagini con questo formato all'interno dei principali applicativi grazie ai plug-in<sup>16</sup> gratuiti e disponibili on-line. Tra gli altri sono disponibili i plug-in per ArcView, AutoCAD MAP, MapInfo, Photoshop, Internet Explorer, Microsoft Office. Le immagini compresse ECW mantengono la georeferenziazione completa del dato e le informazioni di localizzazione del file. Questo significa che le immagini compresse vengono immediatamente registrate con le altre immagini o con i dati vettoriali, proprio come un dato non compresso.

---

<sup>16</sup> Il Plugin è una applicazione software che viene installata in pacchetti di elaborazioni immagini per ampliarne le capacità di utilizzo. Ad esempio è possibile inserire il plug-in per la visualizzazione di immagini ECW in pacchetti software che on origine non prevedevano tale possibilità.

## STRUTTURA VETTORIALE

La **grafica vettoriale** rappresenta le informazioni attraverso un insieme di primitive di tipo geometrico. Mentre nel formato raster una linea è rappresentata da un insieme di pixel, in quello vettoriale essa viene memorizzata come un'equazione che parte da un punto identificato con delle coordinate iniziali e termina in un altro punto definito da coordinate finali.



Confronto tra le immagini rasterizzate (sinistra) e vettorializzate (destra) di una stessa porzione di cartografia 1:2000.

Tale sistema di descrizione delle informazioni grafiche presenta l'indubbio vantaggio di una descrizione più efficace dei dati e quindi una minore occupazione di memoria rispetto ad una corrispondente raster. La lettura del dato comporta anche un minore utilizzo di RAM (*Random Access Memory*, è la memoria utilizzata per caricare dati in modo temporaneo). E' un formato più facile da gestire e da modificare, essendo minore la quantità di dati coinvolti in ogni singola operazione. Questo rende il vettoriale particolarmente adatto a gestire grandi quantità di dati: i dati cartografici sono tipicamente gestiti in modalità vettoriale e nella Cartografia Numerica il dato vettoriale rappresenta la base.

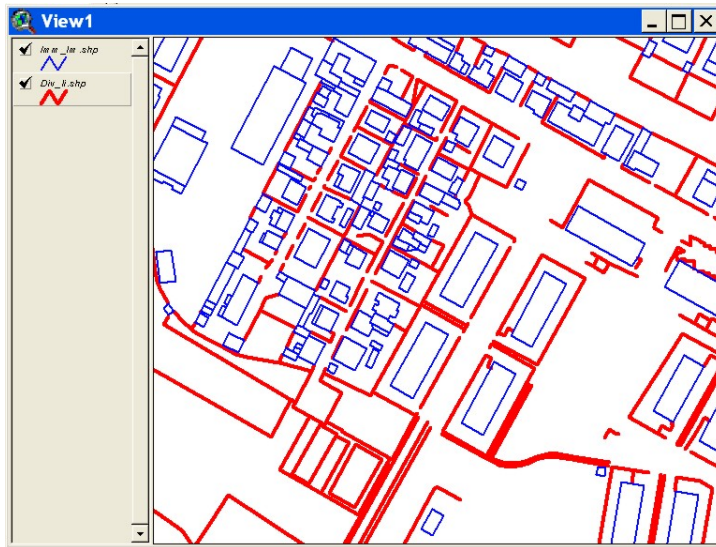
Ricordando che la geometria Euclidea definisce punti e segmenti mediante coordinate, singole nel primo caso e iniziali e finali nel secondo, ne deriva che anche gli oggetti spaziali possono essere rappresentati da un modello matematico simile, il modello vettoriale appunto che vale anche nella terza dimensione. Una descrizione maggiore di questo dato viene fatta nel capitolo relativo alla CN.

Il formato vettoriale è largamente utilizzato per codificare corsi d'acqua, reti di trasporto (strade, ferrovie, ecc.), reti di elettrificazione, ecc. le quali hanno proprietà topologiche di collegamento, orientazione, sequenza e devono essere preservate all'interno del database spaziale.



## Alcuni tipi di formati VETTORIALI

### SHP



Lo Shapefile, sviluppato e gestito da ESRI, è un formato standard per dati geospaziali vettoriali utilizzato dai Sistemi Informativi Geografici (GIS).

Ha la caratteristica di immagazzinare elementi geometrici con associati eventuali attributi. Creato per il software ArcView, è utilizzato anche da ARC/INFO, ArcGIS e altri software GIS.

Con "shapefile" si indica di norma un insieme di file con estensione ".shp", ".shx", ".dbf", ed altre, che hanno in comune il prefisso dei nomi (per es. "laghi.\*").

Spesso con shapefile si indica però solo il file ".shp". Tuttavia l'shp da solo è incompleto poiché interpretazione ed utilizzo dipendono dagli altri file associati:

- ✓ .shp: file che immagazzina le entità geometriche. E' costituito da una serie di record di lunghezza variabile, ciascuno dei quali descrive un'entità attraverso una lista delle coordinate dei suoi vertici. File richiesto.
- ✓ .shx: file che contiene l'indice delle entità geometriche. Tiene conto delle relazioni tra gli oggetti. File richiesto.
- ✓ .dbf: database che immagazzina gli attributi delle entità geometriche. File richiesto.
- ✓ .sbn e .sbx: file che immagazzinano l'indice spaziale delle entità. Possono essere creati in Arcview. Files facoltativi.
- ✓ prj : file che comanda un progetto, non contenente i dati ma le informazioni con le quali istruisce il software su dove cercarli.

Se non è presente anche una struttura topologica, che permette di associare alle entità attributi descrittivi e di relazionare banche dati informative di diversa natura, non è possibile effettuare un'analisi spaziale complessa.

### DXF

DXF era usato originariamente come un formato di scambio e database per AutoCAD. Tuttavia, nel corso del tempo, la facilità con cui può essere interpretato lo ha reso un formato standard, particolarmente nel mondo del PC. Anche se DXF è in grado di descrivere sia dati 2D che 3D, il numero degli elementi definiti è limitato a quei tipi utilizzati in AutoCAD. E' generalmente possibile il trasferimento senza ambiguità. Le difficoltà aumentano per il fatto che lo spessore delle linee e i riempimenti possono essere definiti solamente per specifici elementi. Ad ogni oggetto può essere assegnato un colore, ma il formato DXF non contiene informazioni di *palette*. Conseguentemente, le capacità grafiche sono molto limitate.