

Indice

I Dati multimediali.....	5
Il testo.....	5
Codifica di Huffman.....	6
Codifica run-length.....	8
Codifica LZW.....	8
L'audio.....	9
ADC.....	9
DAC.....	10
Teorema di Nyquist.....	11
Frequenze di campionamento.....	11
Bassa frequenza di campionamento e ricostruzione errata (Aliasing).....	11
Companding.....	12
Predictive Coding DPCM.....	12
MiDi.....	13
Le Immagini.....	13
Sistemi di rappresentazione dei colori.....	14
Correzione della Gamma.....	14
Problemi di indicizzazione e ricerca legati allo spazio di colore.....	15
Immagini digitali.....	15
Modello di colore.....	16
Compressione immagine.....	17
Sottocampionamento - Ridondanza spaziale.....	18
Predictive coding.....	18
Codifica mediante trasformazione.....	19
Immagini JPEG.....	20
Immagini Vettoriali.....	21
Immagini frattali.....	22
Il video.....	23
Codifica intraframe / interframe.....	23
Compressione video.....	24
II Database multimediali (MIRS).....	24
Architettura dei MIRS.....	24
MIRS Inserimento.....	25
MIRS Recupero.....	25
Modello dei dati.....	25
Requisiti di un MIRS.....	26
Layer: Oggetto, tipo e formato.....	27
Interfaccia utente.....	28
Fase di ricerca.....	28
Raffinamento di una Query.....	28
Estrazione delle Feature.....	28
Tipi di Features.....	29
Indicizzazione dei dati.....	29
QoS Quality of Service.....	30
Multimedia data compression.....	30
III Indicizzazione e recupero.....	32
Differenze tra IR e DBMS.....	32
Indicizzazione e Ricerca del testo.....	33

Indicizzazione automatica e modello booleano per il retrieve.....	33
Inverted file.....	33
Retrieval con modello spazio vettoriale.....	35
Relevance Feedback.....	35
Modello Probabilistico e modello basato su Cluster.....	35
Modello Cluster.....	36
Misurazione delle prestazioni.....	36
Crawler.....	37
Indicizzazione e Ricerca dell'audio.....	37
Tecniche di studio di un file audio.....	37
Caratteristiche derivate dal Dominio temporale.....	38
Caratteristiche derivate dal Dominio delle Frequenze.....	39
Metodi di classificazione dell'audio.....	40
Riconoscimento del parlato.....	41
Altre tecniche.....	42
Indicizzazione e Ricerca delle immagini.....	43
Annotazioni libere.....	44
Estrazione Feature di basso livello.....	44
Algoritmi basati sull'analisi del colore.....	44
Indicizzazione basata sull'istogramma di colore (Color-Based).....	45
Problematiche.....	45
Soluzione 1 – Distanza tra i bins.....	45
Soluzione 2 – Iстограмма cumulativo.....	45
Soluzione 3 – PWH.....	46
Limite dell'approccio Color-Based.....	46
Indicizzazione basata sulla forma.....	47
Calcolo della similarità.....	48
Indicizzazione basata sulle Texture.....	48
Indicizzazione e Ricerca del video.....	49
Indicizzazione basata sugli Shot.....	49
Segmentazione automatica in Shots.....	50
Transizioni tra Shots.....	50
Segmentazione a due soglie.....	51
Panning e Zoom.....	52
Cambi di illuminazione.....	52
Utilizzo degli Shot.....	52
Individuazione di r-Frame.....	53
Motion Vector.....	53
Riconoscimento degli oggetti.....	53
Micon - Motion Icon.....	54
Browsing Video.....	54
Altre rappresentazioni compatte.....	54
Tipi di Query e Vettori.....	55
IV Strutture dati per la ricerca delle similarità.....	56
Alberi B.....	56
Alberi B+.....	56
Alberi MB+ - B+ Multidimensionali.....	57
Cluster.....	57
Alberi K-d.....	58
Grid files.....	58

Alberi R.....	58
Operazioni.....	59
Efficienza di ricerca.....	59
V Watermark.....	60
Classificazione dei Watermark.....	61
Proprietà dei Watermark.....	62
Watermarking su testi.....	62
VI GIS - Geographic information system.....	63
Organizzazione.....	63
Visualizzazione.....	63
Interrogazione.....	64
Combinazione.....	64
Analisi.....	64
Predizione.....	65
Caratteristiche.....	65
Modello spaziale del mondo reale.....	66
Oggetto GIS.....	66
Oggetti base di un GIS.....	67
Georeferenziazione.....	68
Modelli della terra.....	68
Datum Geodetici.....	69
Latitudine e Longitudine.....	69
Proiezioni.....	70
Classificazione proiezioni.....	70
Proiezione UTM (Universale Trasversa di Marcatore).....	71
Le mappe nei GIS.....	71
Mappa Tridimensionale.....	71
Esplorazione dello spazio.....	72
Modelli di ripartizione territoriale.....	72
Indici statistici geospaziali.....	73
Statistiche centrografiche.....	74
Analisi dei quadranti.....	74
L'analisi del vicinato - Nearest Neighbour Analysis.....	75
Interpolazione.....	75
VII GPS – Global Positioning System.....	76
Navstar.....	76
Sistema GPS.....	77
Segmento spaziale.....	77
Segmento di controllo.....	77
Segmento utilizzo.....	77
Trilaterazione.....	78
Triangolazione.....	78
Funzionamento del GPS.....	78
Rifrazione e riflessione.....	79
Effemeridi.....	79
Fonti di errore nella determinazione della posizione.....	80
Time Division Multiplexing.....	80
Time Zones.....	81
UTC (Coordinated Universal Time).....	81
Standard Spazio/Tempo.....	82

Sincronizzazione Orologi.....	82
A-GPS – Assisted GPS.....	82
D-GPS – Differential GPS.....	83
DOP – Dilution of precision.....	83

I Dati multimediali

Il **Multimedia Indexing and Retrieval Systems MIRS** ha l'obiettivo di Indicizzare e Ricercare I dati multimediali tra cui testi, immagini, audio e video.

Per la progettazione di un MIRS è necessario conoscere la struttura e le caratteristiche peculiari dei dati multimediali.

In Questo capitolo verranno trattati gli argomenti relativi alla struttura di un file multimediale, per informazioni sui MIRS vedere il [capitolo relativo](#).

Il testo

Il testo piano è un testo liscio privo di ulteriori attributi, esso è codificato mediante il **Codice ASCII**. Il Codice ASCII era inizialmente codificato con 7 bit, solo successivamente è stato esteso ad 8 bit (Extended ASCII), per permettere di rappresentare caratteri accentati e simboli.

Negli anni il codice ASCII ha dovuto cedere il passo all'**UNICODE** codificato inizialmente a 16 bit e poi esteso a 21 bit consentendo la rappresentazione di circa 1 milione di caratteri.

Il Testo strutturato si distingue dal testo piano nella sua forma; un testo strutturato può essere formattato, avere titoli, caratteri in grassetto, etc. Non si visualizza quindi solo il semplice testo ma anche i suoi relativi **attributi**.

Il testo compresso consente di salvare spazio nella memorizzazione dell'informazione, nonostante il testo non richieda molto spazio a differenza degli altri media, conviene intervenire con **algoritmi di compressione** soprattutto quando il numero di testi da archiviare è elevato.

La compressione, come per qualsiasi media, aggiunge complessità in fase di memorizzazione ed in fase di lettura. Il tipo di compressione effettuata è definito **LOSSLESS** ovvero senza perdita, dato che a differenza di una fotografia o un video non è possibile perdere parte dell'informazione per salvare spazio.

La compressione del testo è giustificata da:

- Caratteri che appaiono **più di frequente** sul testo rispetto ad altri
- **Gruppi** di caratteri molto frequenti

Infatti questi possono essere compressi mediante algoritmi quali:

- **Huffman**
- **Run length**
- **Lempel-Ziv-Welch (LZW)**

Codifica di Huffman

Questa codifica **LOSSLESS** è basata sull'analisi statistica del dato da comprimere, in particolare sulla frequenza con la quale si ripetono i suoi elementi.

Consideriamo come esempio il testo “Ciao_mamma”, che convertito in ASCII ha grandezza pari ad 80 bit, ovvero 10 caratteri x 8 bit.

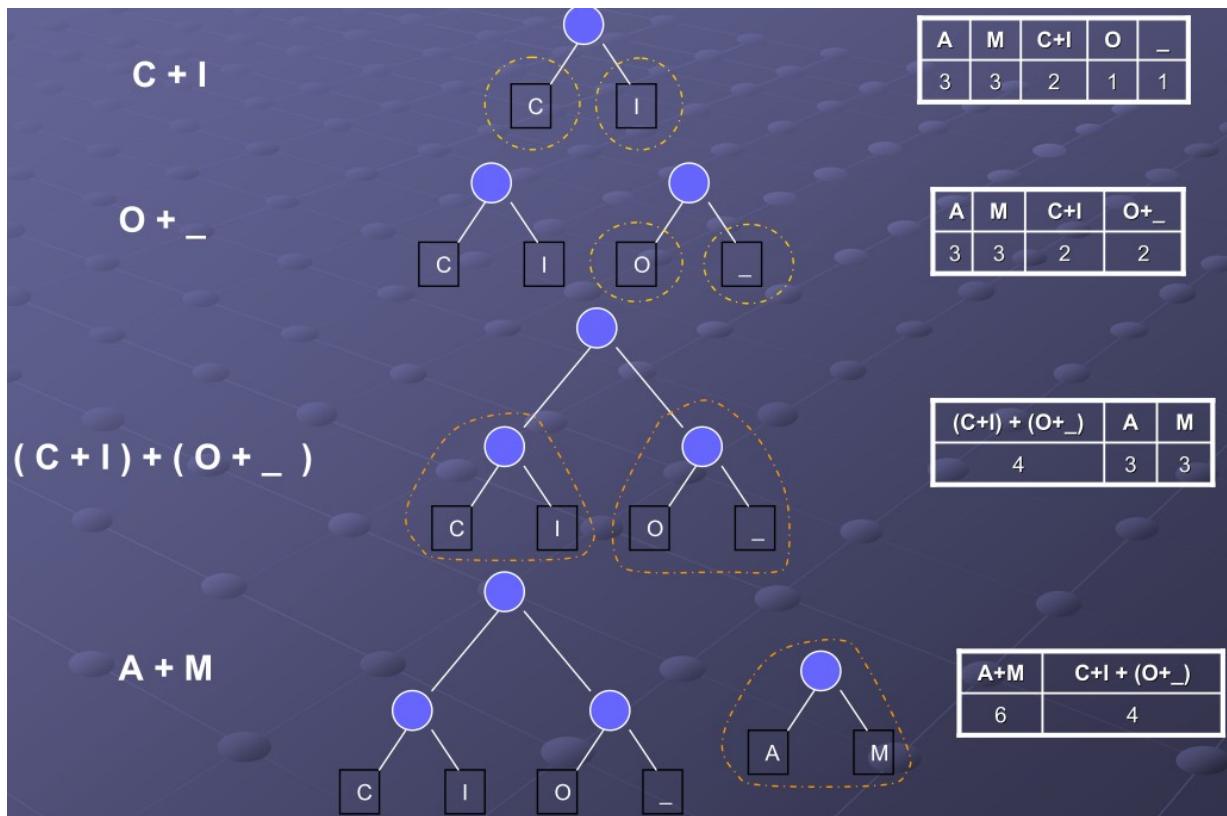
Le fasi della compressione:

1. Analisi della frequenze dei caratteri

- Viene costruita una tabella con il conteggio dei caratteri presenti nel testo
- Successivamente viene ordinata la tabella in ordine decrescente

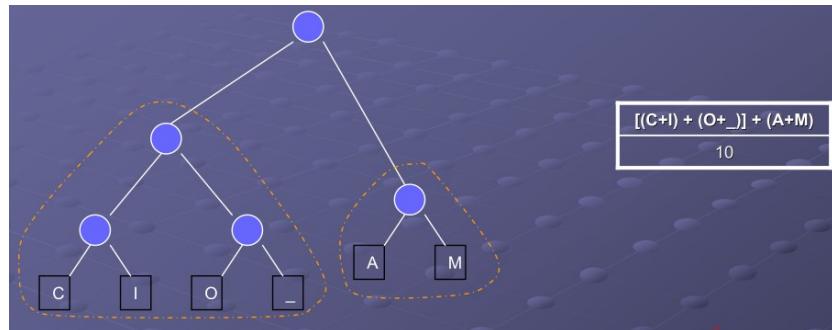
2. Raggruppamento ricorsivo per coppie con frequenza minore

- “C” ed “I” compaiono lo stesso numero di volte

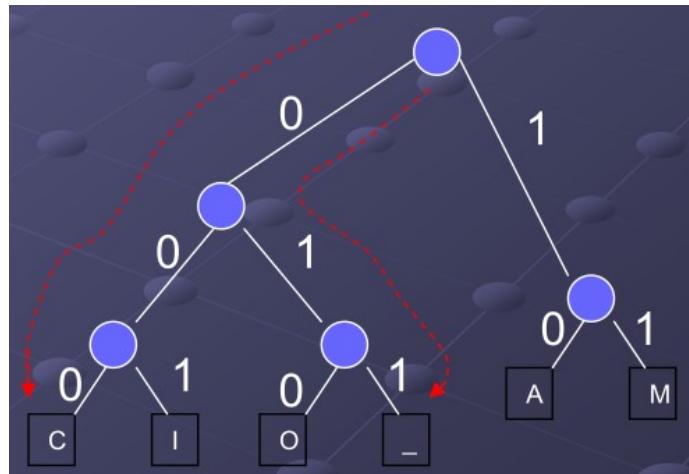


3. Costruzione del codice

- Si costruisce un albero totale fondendo i sotto alberi delle occorrenze, per convenzione a sinistra si avranno i caratteri a frequenza minore e a destra quelli a frequenza maggiore



- I rami di sinistra avranno l'etichetta 0, destra invece 1 (Per convenzione)



- La lettera "I" sarà quindi codificata come 001
- La lettera "A" sarà quindi codificata con 10, quindi 1 bit in meno alle lettere meno frequenti

Codificando abbiamo quindi ottenuto una stringa compressa

C	I	A	O	-	M	A	M	M	A
000	001	10	010	011	11	10	11	11	10

$S' = 000\color{magenta}{001}\color{green}10\color{black}010\color{red}011\color{blue}111\color{green}10\color{black}111\color{blue}111\color{green}10$ → 24 bit

Passando a un 80 bit necessari alla rappresentazione a solo 24 bit, ottenendo quindi un guadagno del 70%.

Il codice che si ottiene non è univoco, dato che si possono adottare diverse convenzioni.

Il codice ottenuto ha la caratteristica di essere **univocamente decifrabile**, ovvero non presenta ambiguità nella sua decifrazione. Nessuna parola sarà prefisso dei un'altra parola (pescatore = pesca + tore), per tanto non si verifica alcuna ambiguità per quanto riguarda la separazione di un gruppo di bit da un altro.

Codifica run-length

Codifica **LOSSLESS** che riduce le ripetizioni di caratteri sostituendo un RUN, ovvero un insieme di caratteri ripetuti, con il carattere ed il suo numero di ripetizioni.

Rappresentazione del RUN :



Sc = carattere speciale per l'identificazione della codifica

X = il carattere che viene ripetuto nel RUN

C = numero di ripetizioni del carattere.

Questa codifica ha una buona efficienza per ripetizioni di carattere superiore a 3.

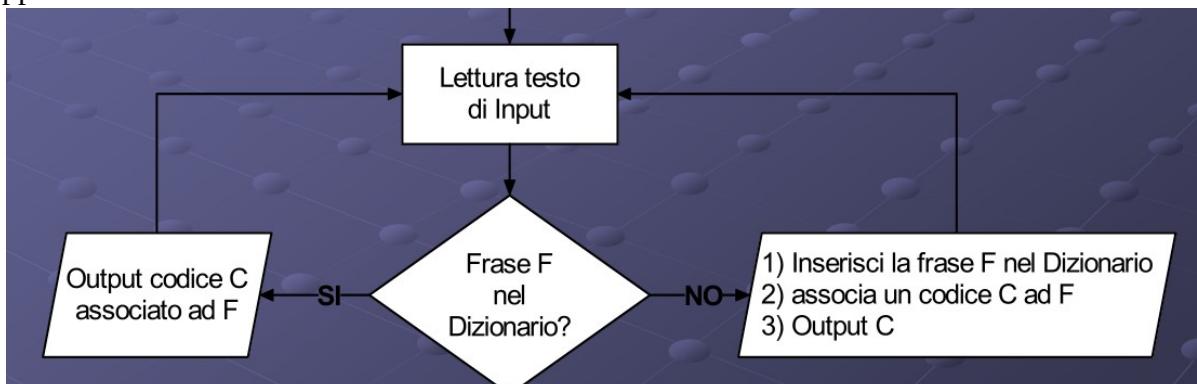
ES:

eeeeeeetnnnnnnnn → @e7t@n8

Questa tecnica era molto utilizzata per la comunicazione via FAX, per indicare la sequenza di punti neri e bianchi trovati in una riga (adatta anche ad immagini con pochi colori).

Codifica LZW

Questa codifica è una estensione della codifica **RUN Length**, conteggiando le ripetizioni di più gruppi di caratteri e frasi.



Le sue prestazioni sono ottime per testi in linguaggio naturale.

L'audio

L'audio è causato da variazioni di pressione dell'aria, le frequenze udibili dall'uomo variano nell'intervallo tra **20 – 20.000Hz**.

I principali descrittori del suono sono: **Aampiezza e Frequenza** che variano nel tempo.

Le **curve isofoniche** descrivono quali livelli di pressione sonora (dB) percepiamo con ugual volume (phon) al variare della frequenza. Queste curve descrivono quindi il rapporto tra **Frequenza** (Hz) del suono e **Intensità** (dB).

La maggiore sensibilità è tra i **1.000 - 5.000Hz**, in questa fascia ci vogliono meno dB per avere più phon, questa fascia è particolarmente sviluppata nell'essere umano, grazie al parlato.

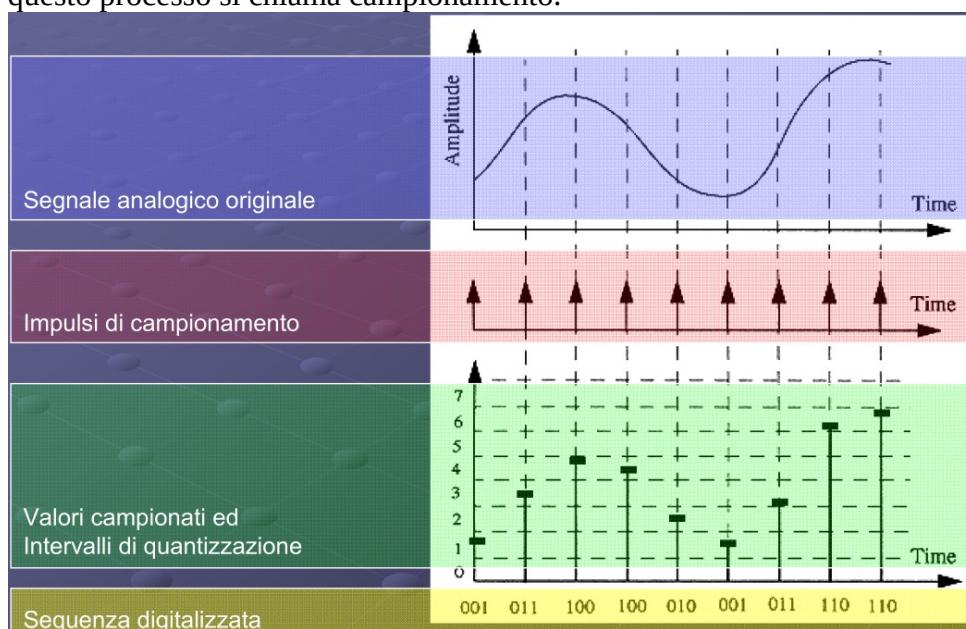
Le **caratteristiche principali** del suono sono:

- Intensità (I = Potenza / Superficie di propagazione)
- Frequenza (Hz)
- Timbro (Ciò che permette di distinguere fonti diverse)

ADC

Conversione del segnale da Analogico a Digitale (ADC - Analog to Digital Converter).

Partendo dal segnale analogico originale, vengono letti ad istanti di tempo prefissati i valori dell'ordinata, questo processo si chiama campionamento.



I valori campionati sono dei punti che, se congiunti, ricostruiscono la curva analogica originale. La sequenza digitale, nasce dalla delimitazione in **intervalli di equalizzazione**, essi vengono approssimati ed associati ad una fascia.

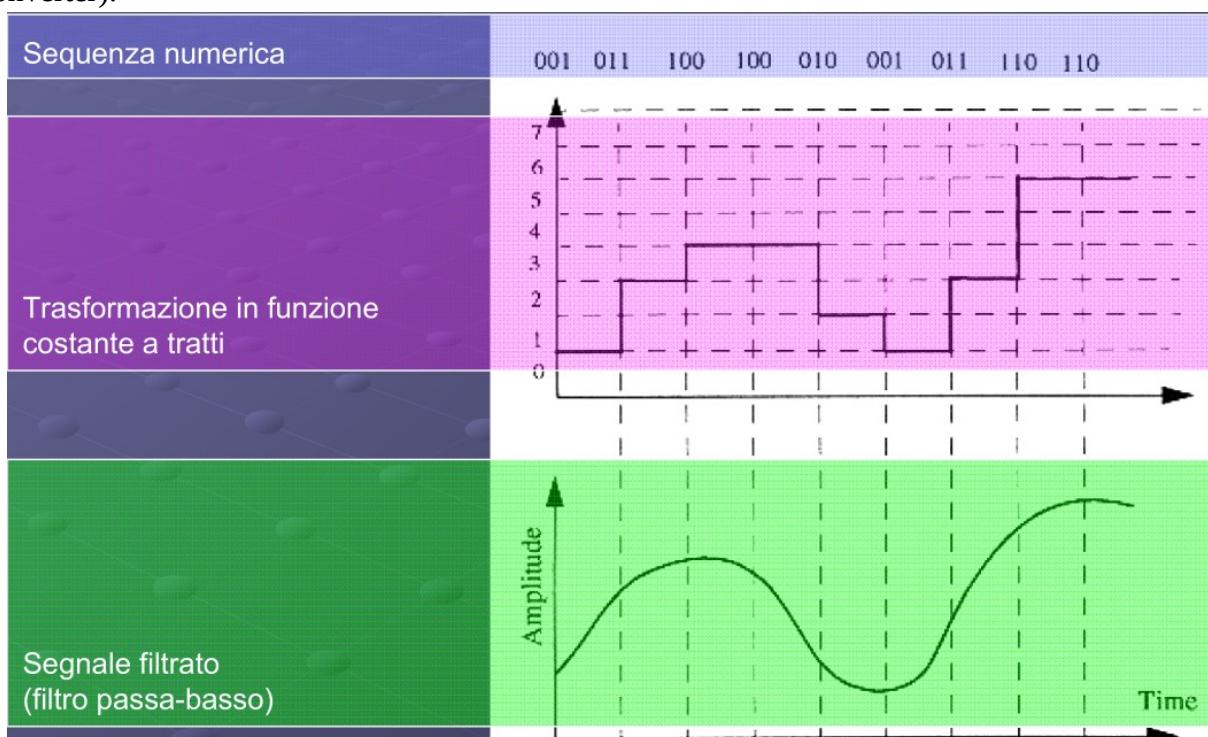
Es. Il secondo valore sulla sinistra nel riquadro verde, è tra la fascia 3 e 4 ed è quindi approssimato a 3.

Riassumendo, le fasi fondamentali per la conversione di un suono da Analogico a digitale, sono 3:

- **Campionamento:** Prelievo dei valori assoluti del segnale ad intervalli discreti di tempo (gestiti da un clock). I campioni prelevati sono ancora di tipo analogico.
- **Quantizzazione:** Processo di conversione dei valori continui in valori discreti. L'intervallo del segnale viene suddiviso in un numero fisso di sotto-intervalli (detti passi di quantizzazione) di uguale dimensione e gli viene associato un valore. A ciascun campione viene assegnato un intervallo per arrotondamento, i possibili valori sono quindi un numero limitato.
- **Codifica:** Processo di rappresentazione numerica dei valori quantizzati. All'aumentare dei livelli di quantizzazione, aumenterà la fedeltà del segnale digitalizzato.

DAC

Trasformazione di un segnale Digitale in un segnale Analogico (DAC – Digital to Analog Converter).



Il filtro **passa-basso** è un filtro che impedisce il passaggio di un insieme di frequenze, a frequenze più elevate. Il filtro passa-basso **modula i cambi di frequenza repentini** (visibili nel riquadro viola) in passaggi di frequenza più morbidi.

Teorema di Nyquist

La frequenza di campionamento è strettamente dipendente dalla frequenza massima del segnale analogico da convertire.

Se in un segnale analogico c'è una componente con frequenza **f** Hz allora la frequenza di campionamento dovrebbe essere almeno **2f** Hz.

Prendere un numero **troppo elevato** di campioni, potrebbe essere poco efficiente, in quanto una volta capito l'andamento della curva la parte scartata è possibile che non sia necessaria per diversi fattori che variano dall'udito umano al dispositivo di riproduzione.

Prendere un numero **troppo basso** di campioni, di contro è critico dato che non consente di ricostruire in modo accurato la curva.

Il Teorema dai Nyquist stabilisce il **passo di campionamento**, più è elevata la frequenza più campioni devono essere prelevati su quella fascia.

Frequenze di campionamento

La qualità standard di campionamento per:

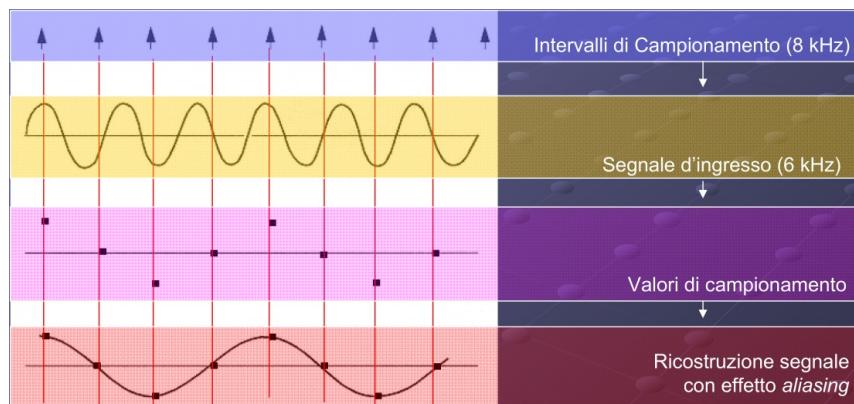
- CD-Audio è di 44.1 kHz
- DAT (Digital Audio Tape) è di 48 kHz
- Telefono è di 8 kHz

E' facile notare come detto prima la voce si aggira tra i 1.000 e 5.000 Hz, il telefono non rispetta a pieno il Teorema di Nyquist, campionando il segnale a 8 kHz e non 10 kHz.

Maggiore è la frequenza di campionamento maggiore sarà la richiesta di spazio/banda di quel segnale (per questo motivo i Provider telefonici cercano di contenere i consumi di banda).

Bassa frequenza di campionamento e ricostruzione errata (Aliasing)

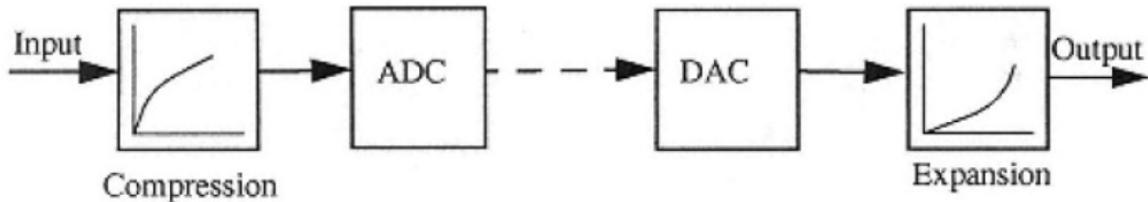
Se si utilizza una frequenza troppo bassa di campionamento non sarà possibile ricostruire la curva in modo corretto, parte dell'informazione andrà persa e si otterrà quindi una **ricostruzione errata** in cui i cambi di frequenza non sono morbidi ma spigolosi (Aliasing).



Il segnale ricostruito soffrirà dell'**effetto Aliasing** e la curva sonora ricostruita risulterà essere completamente diversa dall'originale.

Companding

In italiano, compansione è una combinazione di **compressione** ed **espansione**, è una sorta di compressione analogica del segnale che rende possibile riprodurre un segnale ad 8bit con la stessa qualità di segnale che si avrebbe con 12bit.



L'uso della compansione permette ai segnali con un **ampio intervallo dinamico** di essere trasmessi su attrezzature che hanno una **minore capacità** dell'intervallo dinamico. Ad esempio, si impiega nei radiomicrofoni professionali poiché l'intervallo dinamico dello stesso segnale audio del microfono è più grande dell'intervallo dinamico fornito dalla trasmissione radio. La compansione riduce anche i livelli di rumore e di diafonia nel ricevitore.

Questa tecnica consente di **dettagliare maggiormente determinati cambi di frequenza**, perdendo meno informazione possibile, senza ledere sullo spazio occupato.

- **Quantizzazione variabile:** Gli intervalli di frequenza non hanno tutti la stessa grandezza, alcuni intervalli possono essere più brevi di altri, di conseguenza il campione su quell'intervallo sarà più dettagliato conservando un informazione meno approssimata e più dettagliata. (Distanze diverse sull'asse y, frequenza)
- **Campionamento variabile:** Il segnale non è campionato ad intervalli di tempo sempre uguali, nelle zone con cambi di frequenza più elevati si sceglie di avere più campioni, accorciando l'intervallo tra un campione e l'altro. (Distanze diverse sull'asse x, tempo)

Predictive Coding DPCM

Questa tecnica prevede anziché di codificare il valore del campione, di codificare la **differenza tra la predizione del valore del campione successivo ed il valore del campione attuale (DPCM – Differential pulse-coded modulation)**.

Il valore della predizione si ricava dai valori precedenti assunti dal segnale.

L'efficacia del Predictive Coding si basa sul fatto che:

- Campioni vicini sono significativamente correlati
- Per codificare una differenza occorre un numero minore di bit

Per differenze molto grandi vengono applicati algoritmi correttivi.

MiDi

Acronimo di **Musical Instrument Digital Interface**, è un protocollo standard per l'interazione degli strumenti musicali elettronici.

Il MiDi **non contiene musica pre-registrata**, ma le direttive e le specifiche per riprodurla, è come un libro che contiene la storia e non come il narratore la leggerà. Il MiDi è una sorta di spartito musicale.

Esegui la nota **N** con un durata **T** e con lo strumento **S**

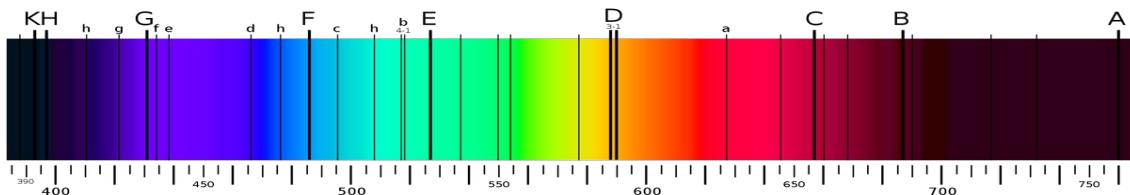
La grandezza di un file MiDi è esigua.

Le Immagini

Le immagini e i colori non sono percepite da tutti gli esseri viventi o macchine capaci di elaborarli, allo stesso modo, per questo motivo non per tutti i fruitori è necessario lo stesso livello di dettaglio.

Gli algoritmi di indicizzazione e di ricerca delle immagini sono basati su di una **Rappresentazione dei colori**; differenti lunghezze d'onda della luce producono differenti sensazioni di colore.

Lo spettro visibile dall'uomo va dai 750nm a 450nm, al di sopra dei 750 c'è l'infrarosso al di sotto dei 450 c'è l'ultravioletto.



Le **caratteristiche principali** dell'immagine sono:

- Luninanza (Illuminazione)
- Tinta (Colore)
- Saturazione (Purezza)

I colori sono riproducibili mediante **Sintesi additiva** o **Sintesi sottrattiva** dei colori primari:

- **Sintesi Additiva:** Partendo dalle frequenze di colore **Rosso Verde e Blu, RGB**, è possibile ottenere altre frequenze di colore.

ES. Il disco di Newton colorato metà di Rosso e metà Verde, fatto girare ad alta velocità l'occhio umano alla vista percepisce il colore giallo.

- **Sintesi Sottrattiva:** Considerando il **fenomeno della radiazione assorbita**, alcuni materiali possono **assorbire** parte dello **spettro luminoso** consentendo comunque il passaggio dello stesso. Tale fenomeno provoca una variazione della frequenza luminosa e di conseguenza una percezione diversa di colore.

Lo **spazio di colori** RGB non è percettivamente uniforme ed è fortemente dipendente dal dispositivo di riproduzione. La Commissione Internazionale per l'illuminazione, **CIE** propone gli spazi di colore CIERGB, CIEXYZ, CIELUV e CIELAB per risolvere tale problematica.

Lo spazio RGB è uno spazio Tridimensionale, gli spazi di colore proposti dal CIE sono spazi Bidimensionali.

Sistemi di rappresentazione dei colori

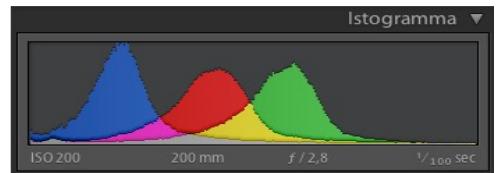
Rappresentazione SPD (Spectral Power Distribution), ogni colore fisico può essere rappresentato mediante la distribuzione dell'energia radiante delle varie lunghezze d'onda.

Il principale **vantaggio** è l'**accuratezza** con la quale si specifica un colore

I principali **svantaggi** sono che:

- Non descrive la relazione tra le proprietà fisiche del colore e la sua percezione visiva
- Ha una complessa rappresentazione numerica

Grazie a questa rappresentazione è possibile per un Computer trovare similitudini nelle immagini, questo avviene confrontando l'**istogramma** delle varie immagini proposte.



Correzione della Gamma

Partendo dall'ipotesi che ogni dispositivo ha la sue caratteristiche e un certo **margine d'errore** sulle frequenze luminose, conoscendo questo margine è possibile **bilanciare le frequenze** per ottenere un segnale quanto più fedele possibile.

Ogni strumento fisico di acquisizione o riproduzione dei colori applica una funzione non lineare all'intensità di luce catturata o emessa in relazione al segnale in **Volts** emesso o applicato.

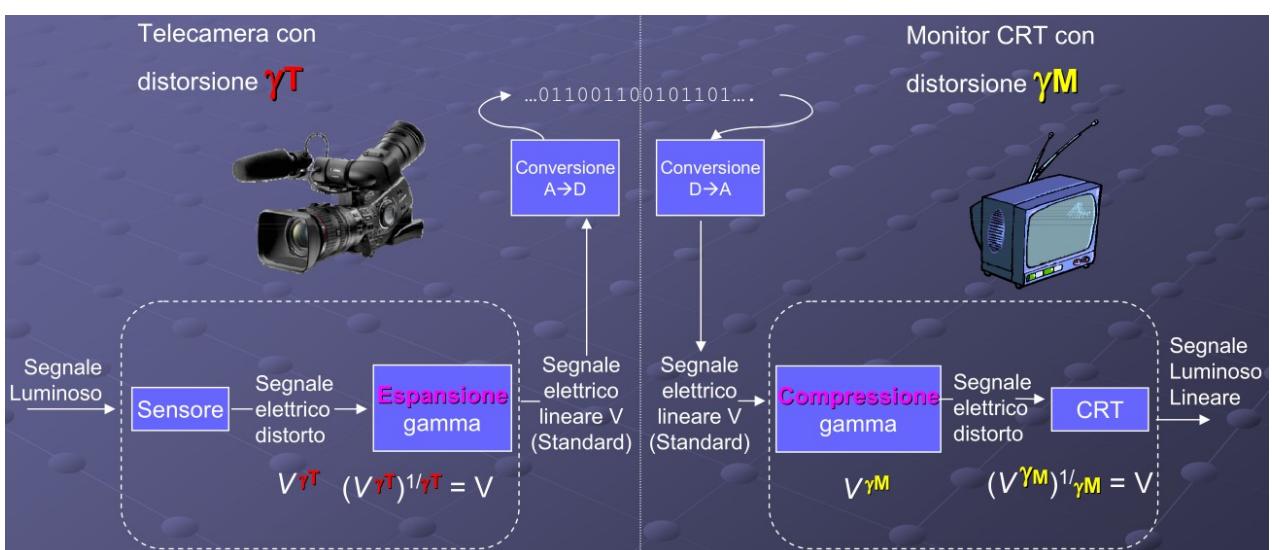
$$\text{Luminanza} = V^y \quad \text{Correzione gamma} = 1/y$$

La **correzione della gamma** è quindi un'operazione non lineare usata per codificare e decodificare la luminanza o i valori tristimolo in un sistema video o fotografico.

Nel caso più semplice la correzione di gamma è definita dalla legge potenziale:

$$V_{\text{out}} = V_{\text{in}}^y$$

Questa funzione di trasferimento serve per neutralizzare la non linearità del dispositivo; V indica il segnale elettrico analogico che il dispositivo preleva.



Problemi di indicizzazione e ricerca legati allo spazio di colore

Nella ricerca e nell'indicizzazione delle immagini basate sul colore è necessario assumere che tutte le immagini siano **rappresentate nello stesso spazio di colore**.

Data un'immagine in formato RGB non è possibile interpretarla correttamente dato che:

- Non vengono indicate le **definizioni delle primitive RGB** e quindi non è possibile interpretare l'immagine in uno spazio comune.
- Non viene indicato il **valore della correzione gamma applicata dal Device** che ha prodotto l'immagine.

Tutti i **formati immagine più usati** (GIF, JPEG, TIFF) **non contengono informazioni sulle primitive** dello **spazio di colore** e sulla **correzione gamma** usata; il nuovo TIFF 6 contiene anche queste informazioni.

Immagini digitali

Un'immagine digitale può essere costruita da Macchine fotografiche, scanner, fotogrammi di filmati, disegni elettronici, etc...

Un'immagine digitale può essere:

- **Raster o Bitmap:** Raster in italiano è **griglia**, questa tecnica definisce una **scacchiera** (griglia ortogonale) in cui ogni **elemento/slot** è chiamato **pixel** a cui è associato un **colore**.
 - La sua memorizzazione può essere **LOSSY** o **LOSSLESS**.
- **Vettoriali:** Immagine definita da un **insieme di primitive geometriche**, esse descrivono punti, linee, curve e poligoni a quali sono associati diversi attributi come il colore. Questo tipo di immagini hanno una **qualità altissima** ed un **peso irrisorio** (Come i MiDi, le immagini vettoriali non contengono l'immagine ma le informazioni per crearla).

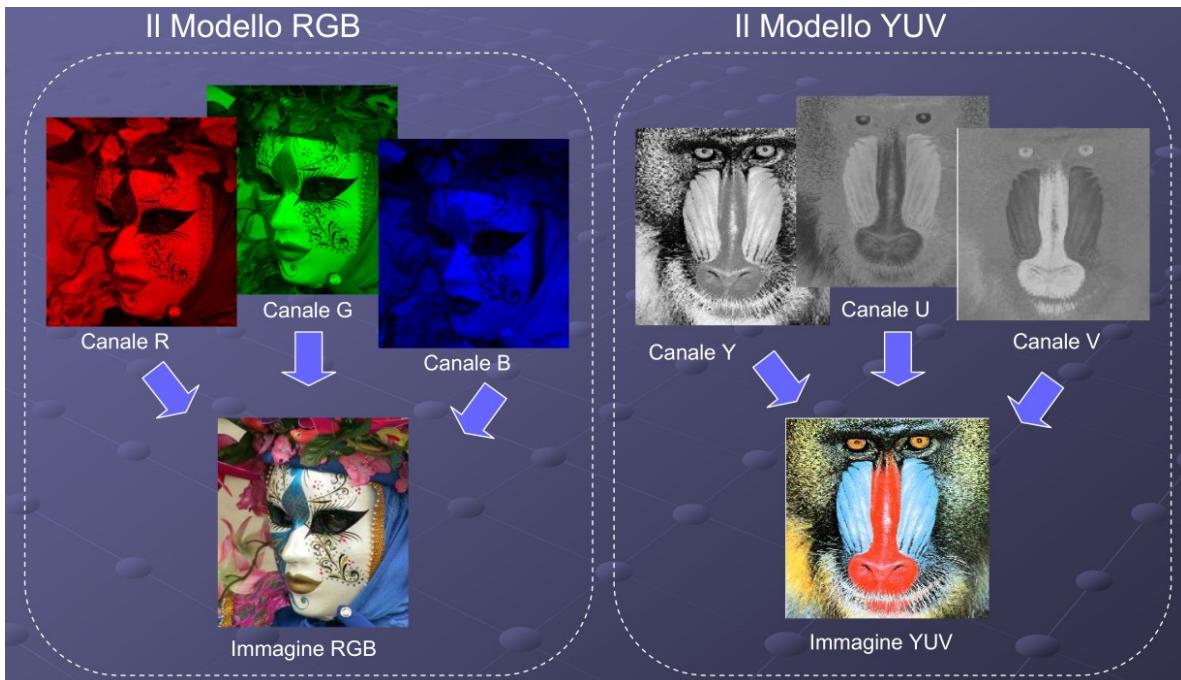
La Super-Resolution è una tecnica che serve a migliorare la qualità delle immagini Raster. L'SR ha due modi di agire:

- **Single-Frame:** Mediante algoritmi di intelligenza artificiale riesce a ricostruire parte dell'informazione ed applicando filtri rende l'immagine più nitida ([waifu2x-caffe](#) o [Waifu2x](#)).
- **Multi-Frame:** Partendo da più foto uguali, vengono applicati meccanismi di correzione basati sulle differenze tra i frame. In astronomia questa tecnica è molto utilizzata. Anche i moderni cellulari con camere multiple sfruttano questa tecnica per correggere le ombre e le sfocature.

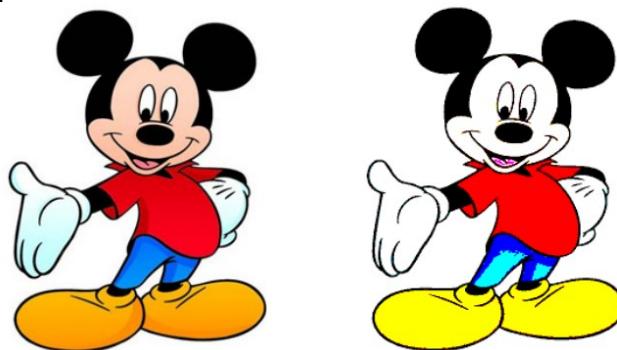
Modello di colore

Un modello matematicamente astratto che permette di rappresentare i colori delle immagini in forma numerica, tra i modelli più in uso troviamo:

- **RGB** (Red Green Blue - Sintesi Additiva)
- **CMYK** (Cyan Magenta Yellow Black - Sintesi sottrattiva)
- **YUV** (Ambito analogico)
- **YcbCr** (Equivalente digitale di YUV, Y=Luminanza, Cb=ChromaBlu, Cr=CromaRed)
- **HSV** (Hue, Saturation, Value)



Altro aspetto molto importante quando si parla di immagini è il numero di Bit assegnati ad ogni pixel per definire il colore. Con 8 bit/pixel si possono avere $2^8 = 256$ colori per pixel rappresentabili, con 4 bit/pixel invece solo $2^4 = 16$.



L'occhio umano non è capace di cogliere qualsiasi sfumatura di colore, per questo motivo non è necessario avere più sfumature di colore di quelle che l'occhio umano può vedere, in ambiti come la fotografia. La storia cambia in settori, come la medicina, in cui le variazioni di colore anche leggere possono avere rilevanza, come in una radiografia.

Compressione immagine

Partendo dal presupposto che **l'occhio umano è limitato** e che quindi molta dell'informazione non viene percepita, è possibile **comprimere le immagini rimuovendo i punti adiacenti** ad un determinato pixel.

Questa tecnica di compressione è possibile in quanto pixel adiacenti hanno colore simile, tale similitudine è detta **Ridondanza Spaziale**.

La perdita dell'informazione non altera il significato dell'immagine e per tanto è **accettabile ai fini della compressione**; tuttavia, anche se l'immagine non risulterà differente all'occhio umano, in caso di Zoom l'informazione scartata potrebbe essere rilevante.

L'occhio umano non percepisce ultravioletto ed infrarosso, per questo motivo le Camere presentano dei filtri per scartare i range di frequenza non necessari. Rimuovendo tali filtri, fisicamente da una Camera, è possibile usare tale camera al buio o in contesti di scarsa luminosità (Camere ad infrarossi).

I tipi di compressione possono essere:

- **LOSSLESS**: Permette di ricostruire perfettamente l'oggetto originale es. codifica di Huffman
- **LOSSY**: Permette di ricostruire solo in parte l'oggetto originale
 - Compressione con [sottocampionamento - ridondanza spaziale](#)
 - [Predictive coding](#)
 - Codifica mediante [trasformazioni](#)

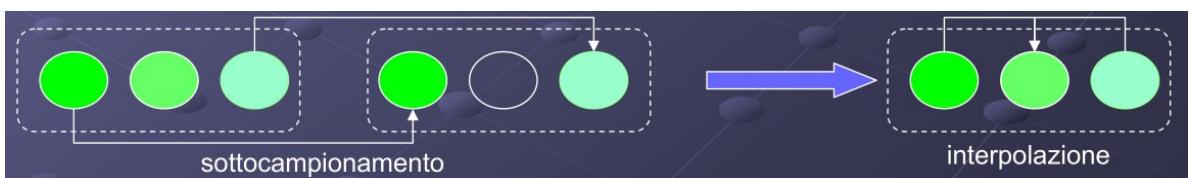
Sottocampionamento - Ridondanza spaziale

Grazie alla **ridondanza spaziale**, ovvero la similitudine tra i Pixel limitrofi, vengono considerati solo alcuni pixel, come ad esempio un pixel ogni due.

In **fase di decodifica** i pixel *mancanti* vengono ricostruiti mediante approssimazione tramite **interpolazione**.

Questo metodo può essere reso più efficiente scegliendo di sottocampionare solo le componenti dell'immagine per le quali il nostro **occhio è meno sensibile**.

- Un'immagine può essere decomposta in due matrici **Luminanza e Crominanza**
 - Il nostro occhio è molto più sensibile alla Luminanza che alla **Crominanza**, quindi è possibile applicare su quest'ultimo:
 - Un livello di **sottocampionamento maggiore**
 - Una **quantizzazione meno raffinata**

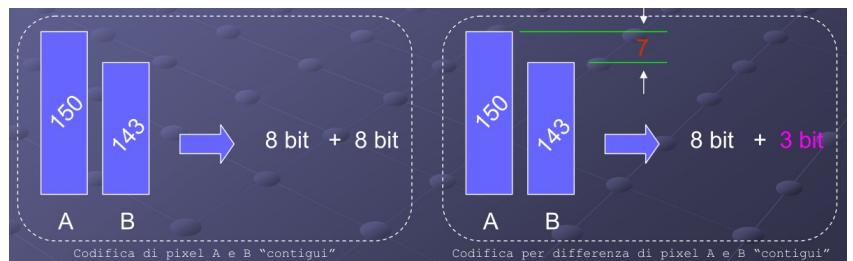


Predictive coding

Analogamente al [Predictive Coding audio](#), valori spazialmente vicini sono fortemente correlati **ridondanza spaziale**.

Supponiamo A e B essere due Pixel vicini e che A sia già stato codificato, allora anziché codificare interamente anche B , è possibile codificare la differenza tra A e B . Tale differenza è inferiore al valore di B e quindi codificabile con meno bit.

Tale predizione può essere fatta sia per **Righe** che per **Colonne** dell'immagine Raster.



Codifica mediante trasformazione

L'idea dietro questo algoritmo di compressione, consiste considerare ogni **pixel** dell'immagine in modo indipendente; gli elementi correlati dell'immagine saranno poi convertiti in un **insieme minimale di coefficienti indipendenti**, così da eliminare la ridondanza.

Un immagine viene quindi **suddivisa in sotto immagini rettangolari** su cui si applica una **trasformazione unitaria dal dominio spaziale**.

SE nel *dominio spaziale* i dati sono fortemente correlati, allora i dati risultanti nel *dominio delle frequenze* sono adatti ad una fase di **compressione** (Huffman, Run-Length, etc).

I due tipi di trasformazione principalmente utilizzati sono:

- **DFT** Discrete Fourier Transform

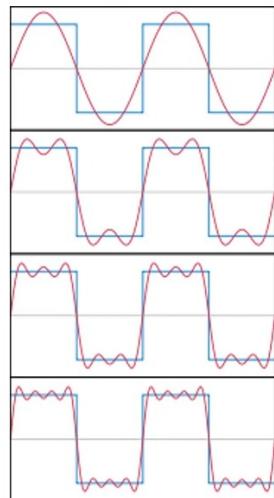
Fourier dimostrò che *qualsiasi segnale periodico* $f(x) = f(x+T)$ può essere scomposto in una somma di infiniti segnali sinusoidali.

Un segnale quindi può essere decomposto per ottenere **solo l'informazione principale**, sostanzialmente **eliminando la portante** che è *periodica e non informativa*.

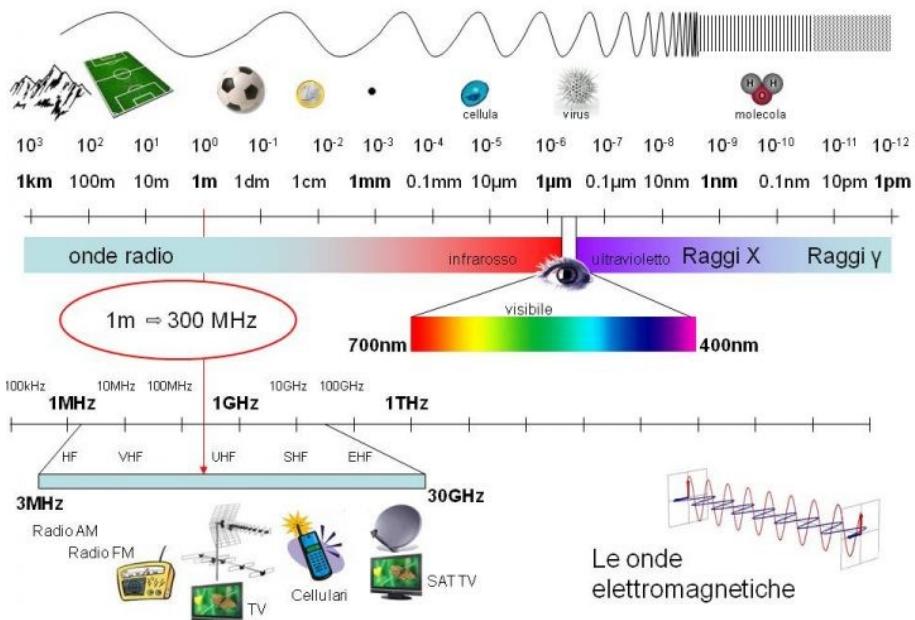
La trasformata di Fourier crea una **decomposizione** che è **unica**.

La **FFT** *Fast Fourier Transform* è una semplificazione della Trasformata di Fourier che è caratterizzata da una complessità di calcolo molto inferiore.

In sintesi: Analizzando lo spettro di un'immagine o di un suono, è possibile decomporlo in frequenze ed eliminare la portante.



- **DCT** Discrete Cosine Transform



Immagini JPEG

Standard fotografico Raster **Joint Photographic Experts Group**, è un modello di rappresentazione dell'immagine **LOSSY** a **colori** a **24 bit**. JPEG è un modello Open Source.

Il modello JPEG ha un **ottima performance** qualità/peso, esso si basa sul:

- Limite della percezione visiva dell'occhio umano
- Analisi spettrale dell'immagine (DCT Discrete Cosine Transform).



Lena Forsén

Fasi principali del modello JPEG:

- **Preparazione in blocchi:** Viene decomposta l'immagine in blocchi 8x8 e viene convertito lo spazio di colore da RGB a YUV per ottenere le matrici di:
 - Luminanza **Y**
 - Crominanza Blu **Cb**
 - Crominanza Rossa **Cr**
- **Passaggio al dominio di frequenze:** Si applica la **DTC** alle 3 matrici **Y Cr e Cb**.
- **Quantizzazione:** Le frequenze vengono **quantizzate in modo non lineare**, ovvero, sono utilizzati opportuni valori per arrotondare con una **precisione più alta le basse frequenze e con precisione inferiore le altre frequenze**.
- **Codifica:** L'immagine viene scansionata a Zigzag, eliminando così le ridondanze mediante codifica Run-Length e Huffman.

Immagini Vettoriali

Tecnica che descrive le immagini mediante primitive geometriche a cui possono essere attribuiti colori e sfumature. Questo tipo di immagini sono generalmente definite mediante insiemi di equazioni matematiche.

Immagine vettoriale



Immagine raster



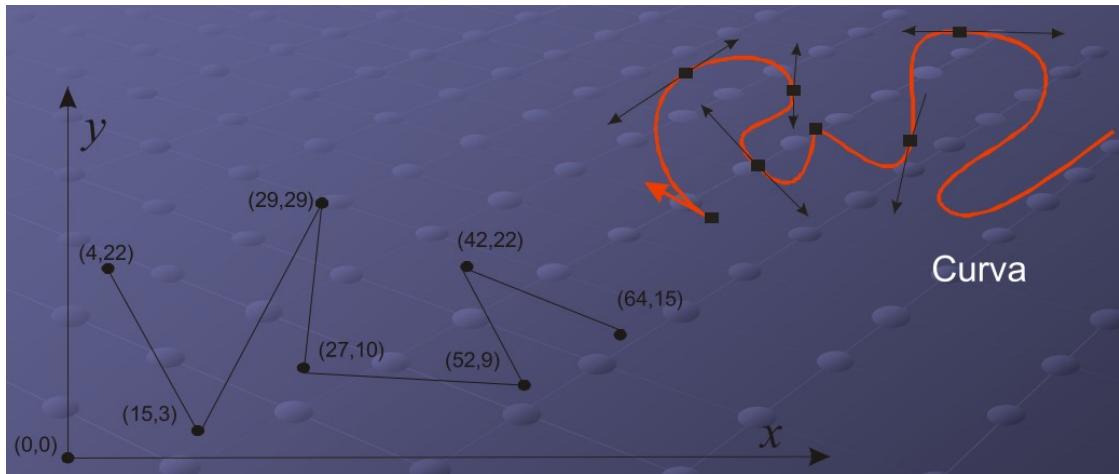
Differenza tra Vettoriale e Raster in seguito ad un ingrandimento

Vantaggi:

- **La rappresentazione è indipendente dalla risoluzione:** dato che ogni primitiva geometrica è memorizzata mediante equazioni matematiche specificando punto di inizio e fine; pertanto queste immagini sono ingrandibili all'infinito.
- **Peso irrisorio:** La sua memorizzazione consiste nel salvare un insieme di formule; ingrandimento e riduzione non pesano in maniera significativa sul peso dell'immagine.

Svantaggi:

- Richiede maggiore capacità di elaborazione rispetto ad un'immagine Raster



Es. Le cartine stradali mostrate dai vari navigatori come Maps e Waze sono realizzate in formato vettoriale.

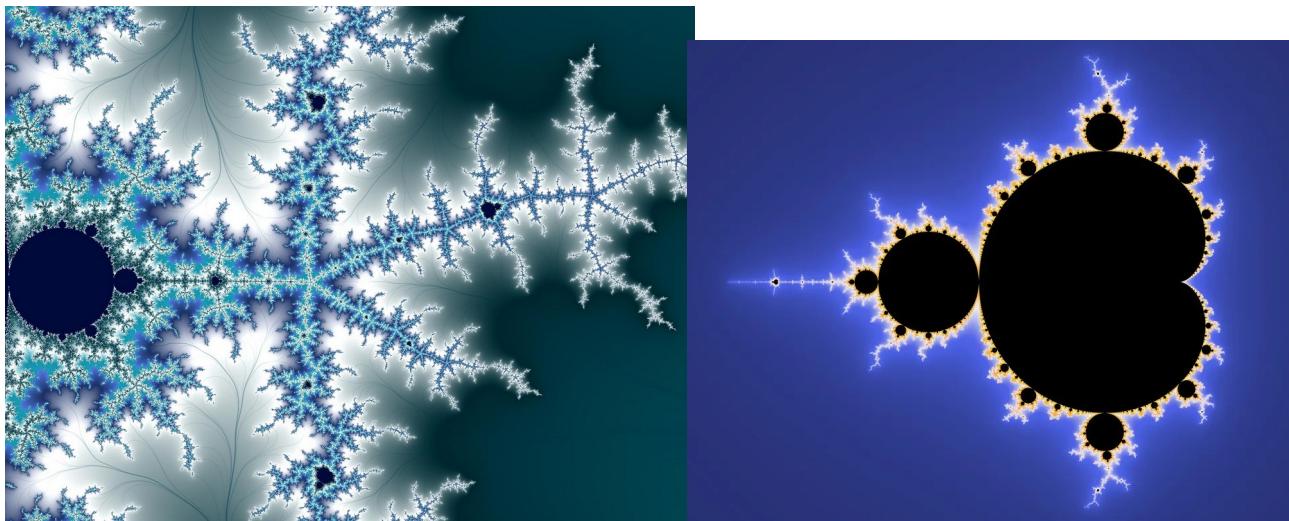
Immagini frattali

Così dette per la loro somiglianza ai Frattali, rappresentano un **intermedio** tra **Raster** e **Vettoriale** per la memorizzazione dell'immagine.

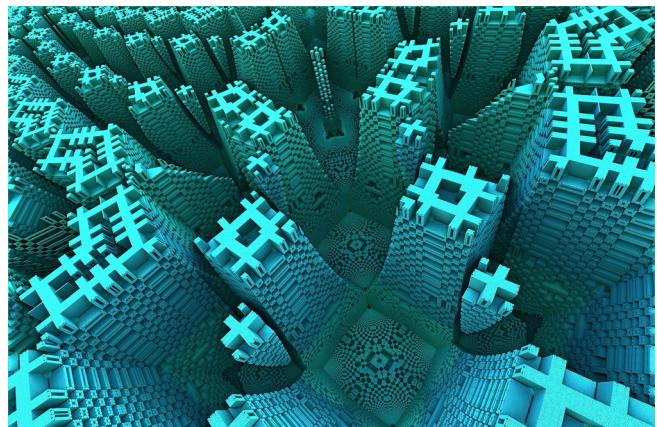
Fondamentalmente risulta essere un'immagine vettoriale con le proprietà di ripetizione di dettagli Raster.

Il metodo **Frattale** è un metodo **LOSSY** che cerca di decomporre l'immagine in **parti elementari** da memorizzare insieme alle **regole** che ne guideranno la ricomposizione.

L'idea di base è la **scomposizione** dell'immagine in **parti più piccole** cercando di trasformarle in **forme geometriche elementari**, che applicate iterativamente alle parti precedenti, permettono di ricostruire l'immagine originale.



Queste immagini si basano fortemente su **pattern ricorrenti**.



Il video

Una **sequenza di fotogrammi** o immagini visionate a frequenza detta Frame Rate.

Il **Frame Rate** è la velocità di scorrimento dei singoli Frame (immagini) in un secondo.

Un elevato numero di Frame in un secondo rende il filmato più **fluido** ledendo la necessità di banda per la trasmissione a causa del peso del filmato stesso, che risulterà essere più elevato.

Esistono due tecniche di composizione dell'immagine:

- **Progressiva:** Prevede la composizione dell'immagine pixel per pixel partendo dall'angolo superiore sinistro e procedendo per righe
- **Interlacciata:** Prevede la composizione dell'immagine componendo prima tutte le righe dispari e successivamente quelle pari.

Il video è un media che per sua natura richiede molto spazio, per tale motivo risulta necessaria la sua ottimizzazione e compressione.

Codifica intraframe / interframe

Codifica **Intraframe**

- Codifica e decodifica del flusso video descrivendo **ogni singolo Frame**, rispetta quindi l'approccio tradizionale di **quantizzazione** come sequenza di immagini statiche
- Utile per sequenze **video particolarmente movimentate**
- Metodo preferito per la fase di **Authoring** di un filmato
- L'accesso diretto ad una generica scena è **semplificato**

Codifica **Interframe**

- Descrizione dei **cambiamenti** che occorrono tra un **Frame ed il successivo** partendo da un **fotogramma iniziale** descritto con codifica **intraframe** (analogo al [Predictive Coding](#))
 - Tecnica dei **Keyframes** usata per ricostruire i fotogrammi successivi
- Utile per sequenze video con **pochi frame che cambiano nella scena**
- Metodo preferito per la fase di **distribuzione** di un filmato
- L'accesso diretto ad una generica scena risulta essere **complesso**

Compressione video

Un buon approccio di compressione video si basa sulla **ridondanza temporale** che viene trattata mediante la **similitudine dei fotogrammi adiacenti**.

Ogni fotogramma si divide in blocchi e si cerca la **maggior corrispondenza** tra i blocchi dei **fotogrammi adiacenti**.

Per ogni **coppia di blocchi** simili si determina

- Lo **spostamento** del blocco (Motion Vector)
- La **differenza** tra i due blocchi

Ciò che sarà trasmesso o conservato è il Motion Vector e la relativa differenza.

II Database multimediali (MIRS)

Il **Multimedia Indexing and Retrieval Systems MIRS** ha il compito di Indicizzare i dati multimediali, comprendendone le caratteristiche e catalogandole.

Architettura dei MIRS

Modularità:

L'insieme delle funzionalità fornite è appunto modulare così da garantire:

- Flessibilità ed adattività (Librerie di funzioni)
- Gestione Aggiornamenti
- Aggiunta di moduli Opzionali come:
 - **Thesaurus manager:** contiene sinonimi e relazioni tra parole per effettuare Query
 - **Integrity rule base:** Testa l'integrità di una data applicazione
 - **Context manager:** Tiene traccia del contesto dell'applicazione



Distribuzione:

- Gestione dei dati multimediali (Client-Server)
- Gestione accessi simultanei (Librerie digitali e video-on demand)

Il MIRS è utilizzato in due determinate circostanze:

- **Indexing:** Vengono fornite al MIRS le chiavi per il recupero di un determinato Media
- **Retival:** Vengono prelevate le chiavi per la fruizione di un Media

MIRS Inserimento

1. L'utente specifica il tipo di input (CD, Videocamera, Microfono, File, ...)
2. Il contenuto del dato multimediale viene estratto
3. Gli oggetti multimediali estratti con le relative caratteristiche vengono inviati al/ai server
4. Gli oggetti multimediali ricevuti vengono organizzati
5. Viene eseguita l'indicizzazione delle informazioni per il recupero
6. Vengono memorizzate le informazioni di indicizzazione degli oggetti multimediali

MIRS Recupero

1. L'utente con una Query richiede un elemento al Database
2. Vengono estratti i contenuti della Query per effettuare una ricerca
3. I dati dall'estrattore della Query vengono inoltrati al motore di indicizzazione e ricerca
4. Il motore di indicizzazione e ricerca reperisce l'elemento richiesto (se esiste) che meglio si adatta alla Query fornita
5. L'oggetto richiesto viene inoltrato all'utente

Modello dei dati

In un MMDBMS o in un MIRS le finalità della modellizzazione comprendono una **specifica dei diversi livelli di astrazione dei dati multimediali**.

Un modello di dati per un MIRS deve descrivere proprietà:

- **Statiche:** Riguardanti gli oggetti che costituiranno i dati multimediali, le loro relazioni e gli attributi.
- **Dinamiche:** Le interazioni tra gli oggetti, le operazioni disponibili su di essi e l'interazioni con gli utenti.

L'usabilità di un MIRS è fortemente condizionato dal modello dei dati che ospita.

Requisiti di un MIRS

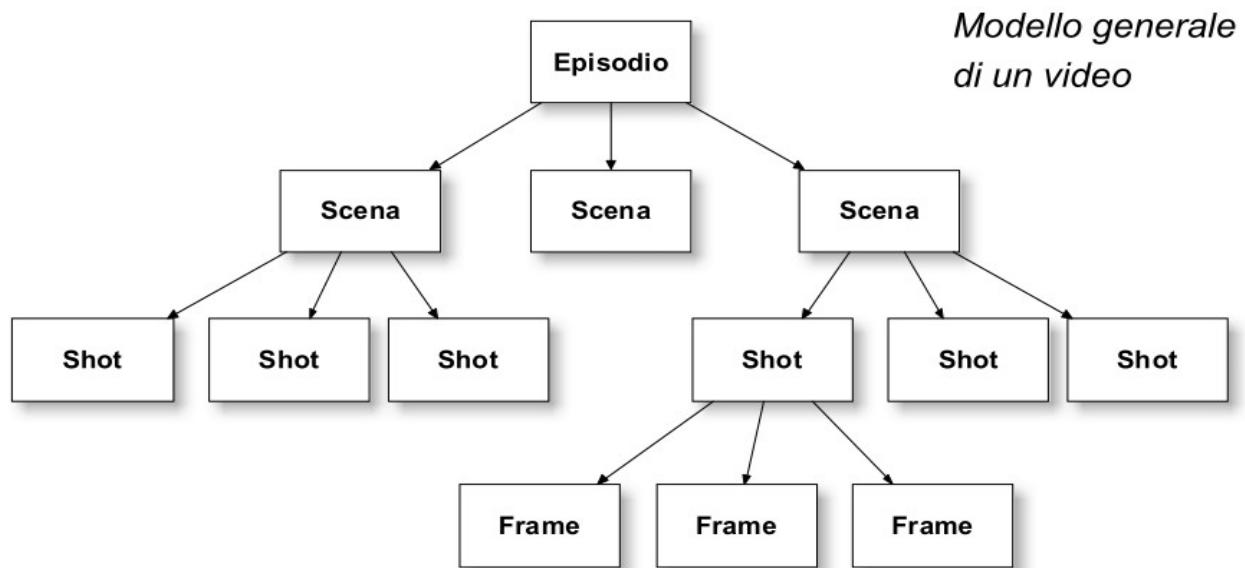
Un MIRS deve rispettare dei requisiti fondamentali per poter operare ed essere funzionale nel tempo, tra i requisiti principali troviamo:

- **Estensibilità:** Capacità di essere adoperato per nuovi tipi e formati di dati anche non previsti dal principio.
- **Flessibilità:** Permettere l'inserimento e la ricerca a vari livelli di astrazione.
 - Strutture di tipo Multilayer.
- **Predisposizione:** Per la rappresentazione dei dati multimediali semplici e composti e per le relazioni spaziali e temporali che intercorrono tra essi.
- **Efficienza:** Per le strategie di memorizzazione e ricerca.

Il paradigma **Object Oriented** (OO) è adattabile al meglio alla modellizzazione dei dati multimediali; l'efficienza trova un valido compagno negli **Alberi**, che consentono un'elevata efficienza per il recupero e l'inserimento.

Un dato multimediale indicizzato in un MIRS è catalogato mediante le sue caratteristiche intrinseche, così da permetterne la ricerca ed il recupero su più livelli.

Esempio di un modello di **file Video**:



Episodio: Autore, data di creazione, tipo di video, produttore, etc...

Scena: Semantica comune di tutti gli Shot che contiene

Shot: Frame chiave, oggetto ripreso, data, luogo, etc...

Frame: Statistiche sull'immagine, distribuzione del colore, etc...

Un MIRS non si limita a conservare e sistemare i dati ma li comprende analizzandoli. Per fare il classico esempio della libreria, un MIRS cataloga i "libri" analizzando il contenuto del testo delle pagine, il tipo di carta usata ed il font con il quale il libro è stampato, oltre ai classici attributi dell'oggetto più esplicativi come il titolo e l'autore (o Metadati).

Layer: Oggetto, tipo e formato

Layer OGGETTO: Costituito da uno o più media con specifiche relazioni spaziali e temporali

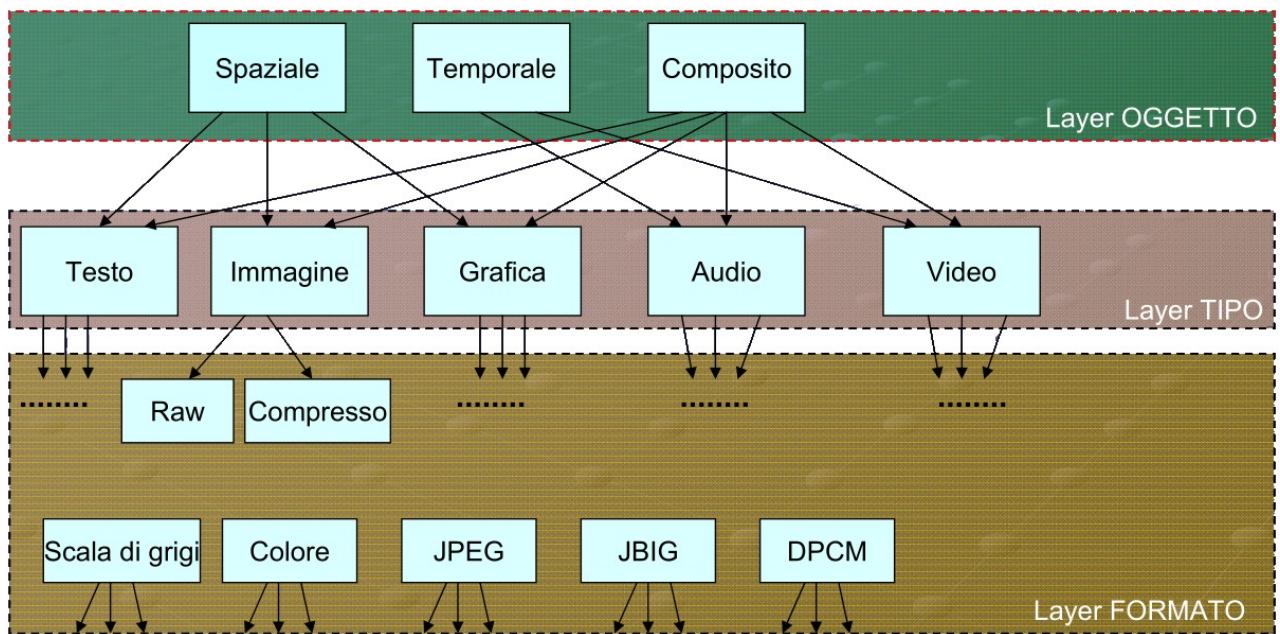
- Relazioni **spaziali**: dimensione di ogni immagine
- Relazioni **temporali**: Tempo di inizio e permanenza, sincronizzazione con il contenuto audio

Layer TIPO: Contiene i tipi comuni di media

- Informazioni usate per il calcolo delle similarità ed in fase di ricerca
- Ogni tipo ha associata una classe media astratta
- A questo livello sono specificati gli attributi del media (Es. immagine):
 - Dimensione in pixel
 - Istogramma del colore
 - Oggetti contenuti nell'immagine

Layer FORMATO: Specifica il formato in cui il media è memorizzato

- Informazioni usate nelle fasi di decodifica, analisi e presentazione
 - Tipo RAW, Compresso, Tipo di compressione



Interfaccia utente

Descriviamo i principali requisiti di un interfaccia utente per gli utenti di un MIRS.

Un interfaccia utente MIRS deve:

- Fornire strumenti per **inserire** oggetti nel database
- Fornire strumenti per definire efficacemente le **Query**
- **Presentare** i risultati in maniere efficiente ed efficace

L'interfaccia deve quindi essere user-friendly.

Fase di ricerca

La ricerca è una fase fondamentale per i sistemi di gestione dei database. Una ricerca può essere effettuata:

- **Per Specifica:** Una Query che va interpretata per passare da testo a definizione di attributi e features.
ES. Auto rossa = Range di pixel a forma di oggetto auto contenuto nell'immagine
- **Per Esempi:** Strumenti di Authoring multimediale che consentono l'inserimento di un esempio come un'immagine, un disegno, un video, una traccia audio, etc...

Raffinamento di una Query

Spesso la Query di un utente è **incerta**, l'utente sa cosa vuole ma non sa come descriverlo; richiesta vaga che presenta un vasto numero di risultati vagamente correlati tra loro.

Data l'**incertezza dell'utente** esso deve poter essere in grado di **raffinare** la sua richiesta in base ai risultati ottenuti dalla ricerca iniziale. Se l'utente si vede restituito un elemento abbastanza simile a quello voluto, il sistema deve permettergli di riutilizzarlo per una nuova ricerca più raffinata.

Anche la **conoscenza dell'utente** può servire per **raffinare la Query**, conoscendo i suoi **interessi** e le sue **abitudini** è possibile effettuare una prima **scrematura** sui risultati di una ricerca, proponendo principalmente **risultati più affini all'utente**.

Estrazione delle Feature

Come già spiegato, gli oggetti indicizzati in un MIRS sono categorizzati anche per **valori intrinseci** come *features* ed *attributi*. Il processo di **ricerca** si basa sulla **comparazione** di tali **features**.

L'estrazione delle **Features** da un media non è effettuato in tempo reale, tali caratteristiche sono state già **estratte al momento dell'inserimento**. Questa pratica rende il sistema più **reattivo**.

Tipi di Features

- **Metadata:** Informazioni di contesto che **NON descrivono o interpretano il contenuto** del dato stesso, ma contengono: *L'autore, data di creazione, titolo, etc...*
- **Annotazioni testuali:** Descrizioni testuali del contenuto di un dato multimediale, queste descrizioni sono **soggettive ed incomplete**.
- **Feature di basso livello:** Sono estratte automaticamente e **descrivono i dati e le statistiche** di un oggetto come le relazioni spazio temporali tra le parti dell'oggetto.
 - Audio: Volume, frequenze
 - Immagini: Distribuzione colore, forma degli oggetti
 - Video: Struttura temporale, feature dei frame chiave
- **Feature di alto livello:** Generalmente estratte mediante l'intervento umano o di complessi algoritmi di intelligenza artificiale già istruiti, **descrivono il contenuto di un media** analizzandolo.

ES. Se il media contiene parlato e/o musica e che musica è contenuta nel media, chi è il soggetto in primo piano, di cosa il media tratta, quali parole chiave sono contenute nel media.

Indicizzazione dei dati

Le features necessarie per la ricerca sono estratte in fase di indicizzazione, ovvero al momento dell'inserimento di un media nel Database. Le features ovviamente devono essere **memorizzate** ed organizzate **categorizzandole**.

Date le diversa natura delle varie caratteristiche è necessario munirsi di **strategie di indicizzazione** adeguate.

ES. la distribuzione del colore di un immagine può essere memorizzata attraverso un istogramma.

QoS Quality of Service

I sistemi MIRS sono generalmente **sistemi distribuiti**, così da garantire ridondanza in caso di problemi ad un server e prestazioni migliori in diverse regioni e stati distanti. I dati multimediali richiedono requisiti stringenti per una adeguata fruizione come un'ambita larghezza di banda.

Il **QoS** specifica un insieme di parametri e requisiti divisi in due gradi:

- Qualità *preferibile*
- Qualità *accettabile* (Minimo indispensabile)

La **QoS** è negoziata tra client e server e sottoscritta tramite contratto che garantisce i parametri in uno dei seguenti modi:

- **Deterministico:** La qualità richiesta è **garantita pienamente**
- **Statistico:** La qualità richiesta è garantita con una **certa probabilità**
- **Best-Effort:** La qualità **non è garantita**, si fa il meglio che si può

Il server cercherà sempre di destinare al singolo utente il minimo necessario, così da poter gestire più client possibili in contemporanea.

Il client richiederà sempre il massimo possibile che esso si può permettere di gestire.

Multimedia data compression

Le gran parte dei dati Audio, Video, Musicali è saltava in maniera compressa, per l'estrazione delle caratteristiche è quindi necessario **decomprimere** l'oggetto per elaborare le specifiche da conservare.

L'operazione di estrazione delle caratteristiche è un processo molto frequente e richiede una elevata capacità computazionale, per questo motivo occorre uno schema efficiente sia per la trasmissione che per la presentazione:

Metodo 1 - Anteprima:

- Sul server per ogni grande immagine si salva anche una **copia ridotta**
- La **Query** dell'utente **recupera** sempre la **copia ridotta**
- Se **occorre un maggior dettaglio** allora anche l'immagine **originale** viene inviata al richiedente
- **Svantaggio:** Ridondanza di dati sul server

Metodo 2 - Ridimensionamento:

- Ogni Query dell’utente recupera direttamente l’immagine **originale**
- L’immagine originale viene **ridotta** per poter essere gestita dal client
- Se è **necessario maggior dettaglio** allora il server invia un altra copia **meno ridotta** dell’immagine
- **Svantaggio:** Spreco di banda per la trasmissione

Metodo 3 - Compressione:

- Uso di metodi di decompressione scalabili, progressivi e gerarchici (gif e Jpeg progressive)

Nessuno di questi metodi dovrebbe essere realizzato in modo tale da portare ad alterazioni sostanziali del media, tagli di frequenze di un file audio o di un’immagine portano un’alterazione del dato. Tali alterazioni devono sempre rientrare in dei **margini di tollerabilità**.

III Indicizzazione e recupero

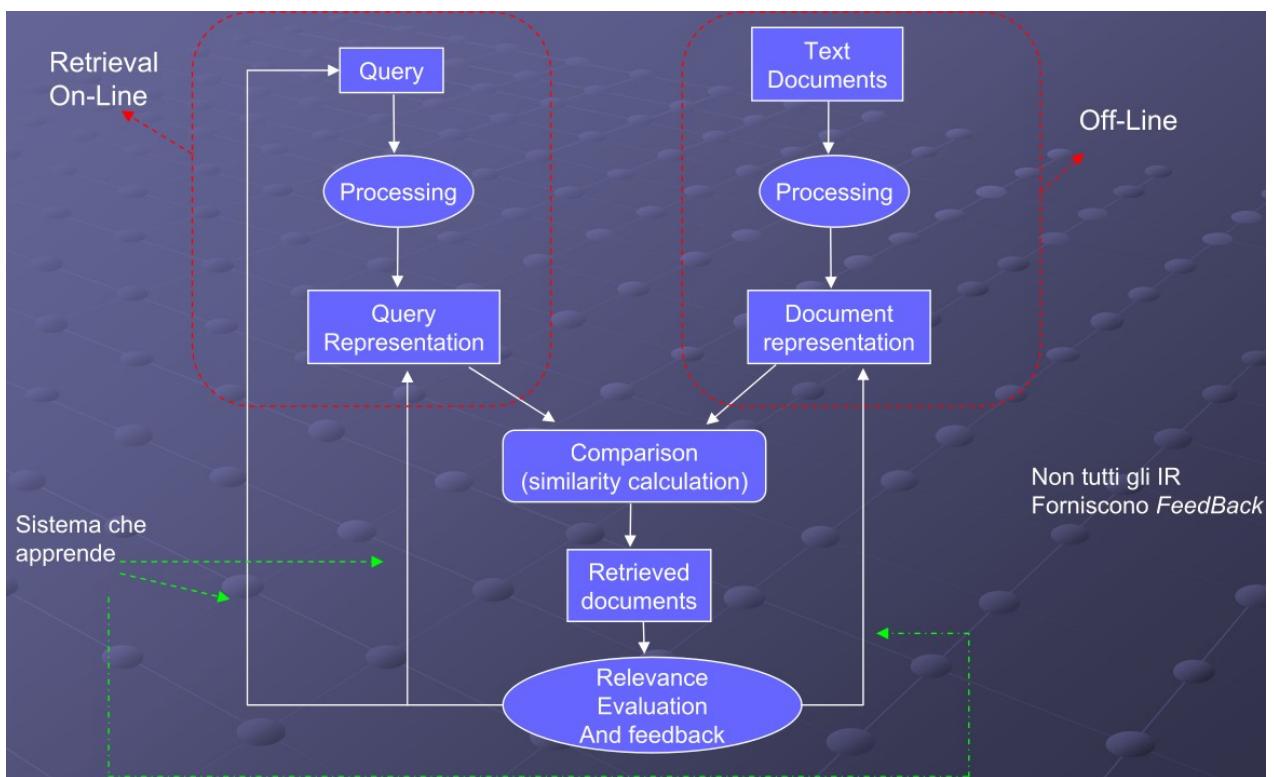
IR: **Indexing and Retrieval**, tecniche di recupero dell'informazione.

Differenze tra IR e DBMS

In un **DBMS** la struttura è **omogenea** e i componenti, ovvero i **record** sono omogenei. Ogni record in un DBMS è definito completamente ed univocamente dai propri attributi.

Il **recupero** di informazioni da un **DBMS** avviene tramite **match esatto**: Query effettuata su tabelle con un determinato tipo omogeneo di record.

In un **IR**, invece, i record non sono strutturati e gli attributi non sono prefissati. Ogni oggetto è indicizzato mediante **keywords**, **descrittori** e **indici soggettivi** ed approssimativi. Il recupero non avviene mediante un match esatto, il **match sarà approssimativo o parziale**.



Online: La Query scritta dall'utente viene **processata** e **rappresentata** in una determinata forma.

Offline: Quando un oggetto viene **caricato** o **indicizzato** da un motore di ricerca, esso viene **processato**, ne vengono comprese le **caratteristiche**, ed una sua **rappresentazione** viene salvata sul Server. Al sopraggiungere di una richiesta queste informazioni verranno **valutate** come idonee per il risultato o meno.

Confronto: La **Query** dell'utente viene **confrontata** con determinate **rappresentazioni** di oggetti noti, alla ricerca di oggetti **idonei** per il risultato. In determinati casi il **Feedback** dell'utente viene conservato per migliorare le ricerche future.

Indicizzazione e Ricerca del testo

Indicizzazione automatica e modello booleano per il retrieve

L'indicizzazione ed il recupero viene effettuato mediante espressioni regolari e stringhe. Ciò può avvenire con strumenti come: **grep**, **egrep**, **awk**, etc...

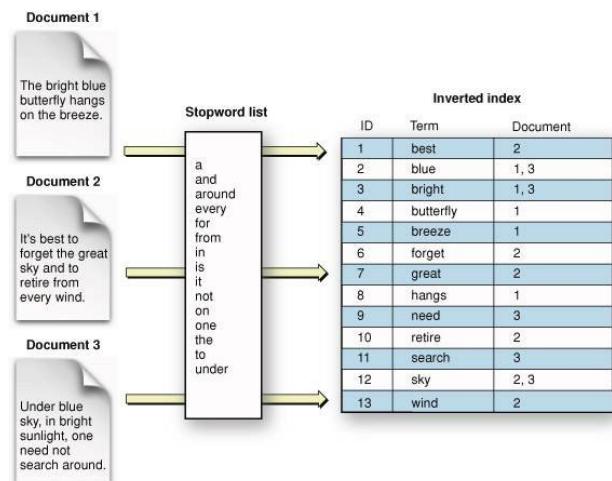
Le espressioni regolari vengono concatenate con delle espressioni booleane per raffinare la ricerca.

Inverted file

L'inverted file contiene un insieme di **righe di testo**, ogni riga contiene:

- Il **termine** che si vuole cercare
- Un **puntatore** a documenti e record che contengono quel termine

Il termine inverted spiega l'inversione del verso di ricerca: prima la chiave e poi il file che la contiene.



ES. Butterfly e breeze sono stringhe contenute solo nel documento 1, una ricerca che contiene queste due chiavi porta al documento 1.

Rispetto alla ricerca su **Flat File**, ovvero la ricerca di un elemento sull'intero file, la ricerca su **Inverted File** è estremamente **più efficiente**.

Una ricerca può essere ulteriormente **raffinata** dando un **peso** a determinate parole e valutando anche la **posizione** e la **frequenza** di determinate stringhe in un testo.

Definiamo quindi delle fasi di filtraggio del testo al fine di indicizzarlo:

- **Stop Words:** Si **escludono** tutti gli **elementi insignificanti** come: articoli, congiunzioni, pronomi, etc...
- **Stemming:** Se nel testo compaiono più volte **parole analoghe**, come “Pescare”, “Pescatore”, “Pesce” allora si considera il **prefisso comune** “Pesc”. Anche se l’operazione rende il file più compatto spesso la brevità del termine fa perdere significato alla ricerca.
- **Thesaurus:** Possibilità di **sostituire termini simili** che compaiono nel testo con un unico termine, un sinonimo comune.

ES. Lavare, pulire, detergere, sgrassare → Lavare

- **Weighting:** I **termini** che compaiono in un testo **non hanno tutti la stessa rilevanza** al fine della ricerca; la loro **rilevanza può essere valutata** in base al significato, frequenza ed ordine nelle frasi. Termini con frequenza inferiori sono considerati più importanti.
 - $W_{ij} = tf_{if} * \log(N/Df_j)$
 - **W_{ij}**: Peso del termine “j” nel documento “i”
 - **tf_{if}**: Frequenza del temine “j” nel documento “i”
 - **N**: Numero totale documenti nel database
 - **df_j**: Numero dei documenti nel database che contengono il termine “j”

ES. Se la parola “metri” compare molteplici volte nel testo e la maggior parte delle volte è adiacente a numeri, probabilmente quella parola non ha un’elevata rilevanza nel testo.

Definiamo anche due nuovi **operatori**:

- **WITHIN SENTENCE**: I termini “i” e “j” devono essere presenti nello stesso paragrafo
- **ADJACENT**: I termini “i” e “j” devono essere confinanti

Le **Query** di un utente saranno quindi comprese per **elaborarne una forma che sfrutti i vari operatori** messi a disposizione dal motore di ricerca.

Retrieval con modello spazio vettoriale

Le **Query** e le **caratteristiche** di un file possono **formare un vettore** in uno spazio vettoriale.

Supponiamo un **vettore** "A" le cui componenti sono: a_1, a_2, \dots, a_n

Questo vettore "A" rappresenta un file, le **componenti** che lo descrivono sono le **caratteristiche** di questo file.

Supponiamo allo stesso modo di **estrarre da una Query dei dettagli** che descrivono un file che si vuole ricercare. Con **questi dettagli estratti si può formare un vettore "B"** le cui componenti sono: b_1, b_2, \dots, b_n

A questo punto la **similitudine tra un file ed una Query** è semplicemente l'angolo formato dai due vettori. Quanto più l'angolo teta è **piccolo**, tanto più Query e File si **somigliano**.

$$\mathbf{a} = [a_1, a_2, \dots, a_n]$$
$$\mathbf{b} = [b_1, b_2, \dots, b_n]$$
$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + \dots}$$

Il coseno di teta sarà:

0: Se l'angolo è di 90° - File e Query non si somigliano

1: Se l'angolo è di 0° - File e Query massima similitudine

Relevance Feedback

L'idea di base di questo modello è quella di **sfruttare le valutazioni degli utenti** per raffinare le ricerche future.

I **termini** che in passato si sono rivelati **meno d'interesse** dagli utenti, vengono **diminuiti di peso** o **scartati**, viceversa **Query** con **termini simili apprezzati in passato**, saranno **corrette** al fine di fornire un **risultato migliore**.

I risultati più cliccati passano in primo piano, quelli meno cliccati scendono in classifica.

La regola applicata è la formula di **Rocchio**.

Modello Probabilistico e modello basato su Cluster

Modello **probabilistico**:

- Considera la **dipendenza** dei termini e le loro **principali relazioni**
- Ha avuto **scarso successo** a causa della **difficoltà di calcolo** delle probabilità

Modello basato su **Cluster**:

- Basato sull'idea di **definire un insieme di gruppi** contenenti elementi simili
- Le **Query cercano all'interno dei Cluster**

Modello Cluster

Generazione del Cluster:

- **Similarità per coppie:** Ogni documento è analizzato ed inserito in una matrice, coppie di elementi molto vicini vengono sostituiti da un unico elemento. Questo processo itera fino a formare un unico Cluster.
- **Clustering euristico:** Il primo file forma un cluster, successivamente ogni nuovo documento verrà confrontato con i cluster presenti e nel caso di similitudine verrà aggiunto al cluster.

Nel caso in cui il nuovo file non abbia similitudini rilevanti con cluster già formati, esso stesso forma un nuovo cluster.

Retrieval dei dati:

Ogni cluster è caratterizzato da un **Centroide**, calcolato mediando tutti i vettori del Cluster. Una Query quando sopraggiunge viene confrontata con i centroidi al fine di selezionare il Cluster più simile alla Query passata.

L'**output** sarà costituito dai documenti del **Cluster più simile individuato**.

Misurazione delle prestazioni

Il web è ormai talmente vasto che analizzare qualsiasi risultato di una determinata Query sarebbe impossibile. Per questo motivo si preferisce restituire **meno elementi**, per una data Query, con il **maggior dettaglio** possibile.

Le prestazioni di un motore di ricerca si basano principalmente su:

- **Velocità di Ricerca**
- **Recall:** Capacità di recuperare solo le informazioni rilevanti per una data Query. Si definisce come un **rapporto** tra il numero degli **elementi rilevanti recuperati** ed il numero totale di **elementi rilevanti presenti nel database**.
- **Precisione:** Accuratezza degli elementi recuperati. Si definisce come il **rapporto** tra il **numero di elementi rilevanti recuperati** ed il **numero totale di documenti recuperati**, per una data Query.

$$precisione = \frac{|doc_ret \cap doc_rel|}{|doc_ret|} \quad recall = \frac{|doc_ret \cap doc_rel|}{|doc_rel|}$$

Quanto maggiore è il recall tanto minore risulta essere la precisione e viceversa.

Si tende a prediligere sempre un'alta precisione.

Crawler

I Crawler anche doti come spider, sono dei bot che analizzano il contenuto di una rete in modo automatizzato. Solitamente gli elementi vengono analizzati per essere inseriti all'interno di un indice.



Indicizzazione e Ricerca dell'audio

Per l'orecchio umano risulta semplice distinguere differenti **tipi di audio, velocità e toni**, inoltre per l'essere umano risulta semplice trovare similitudini tra file audio.

Il modo più utilizzato per identificare file audio è si basa sui suoi **metadati**, ma questo metodo è facilmente **aggirabile** nel caso in cui si cerchi di aggirare normative sul **Copyright**.

Un'ulteriore complicazione risiede nella **mancanza di standardizzazione** per la memorizzazione di file audio; possono essere adottati diversi metodi di campionamento, frequenze di campionamento, ampiezze, etc... Ciò comporta **problematiche per il confronto**.

Tecniche di studio di un file audio

Lo studio di un file audio consiste nel cercare di identificare 3 principali fonti audio:

- **Parlato**
- **Musica**
- **Rumori**

Distinte le tre fonti, esse possono essere trattate in modi differenti per comprenderle al meglio. Il parlato, ad esempio, può essere analizzato e con software di **Speech recognition** per essere convertito in testo da analizzare in seguito con i tradizionali [metodi descritti prima](#).

I file **audio** sono **caratterizzati** da:

- **Dominio temporale** (Tempo/Aampiezza)
- **Dominio delle frequenze** (Frequenza/Magnitudine)

Ogni rappresentazione è particolarmente idonea per l'estrazione di determinate caratteristiche.

Caratteristiche derivate dal Dominio temporale

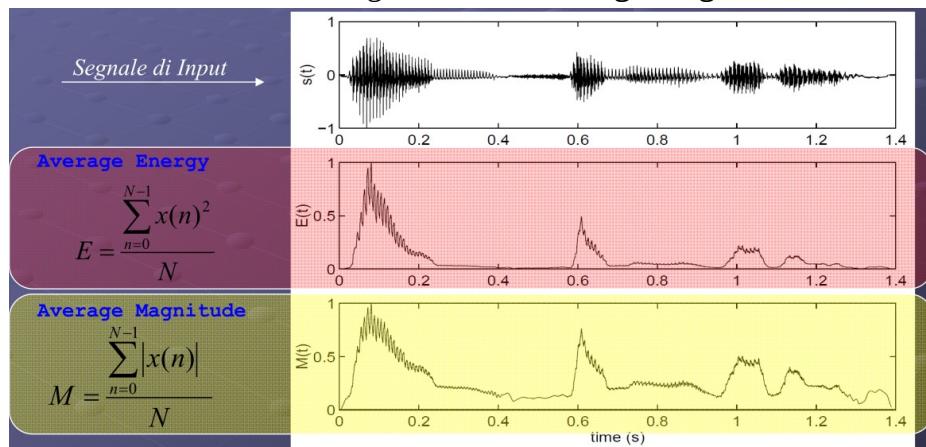
Average Energy: Nel dominio temporale è possibile analizzare l'energia media, ovvero la media della "rumorosità" del segnale audio per la sua durata.

La somma di tutti i campioni al quadrato, così da eliminare i valori negativi, fratto il numero di campioni, restituiscono l'Average Energy.

$$E = \frac{\sum_{n=0}^{N-1} x(n)^2}{N}$$

In cui:
 E = energia media del brano audio
 N = numero totale dei campioni valutati
 $x(n)$ = valore del campione n -esimo

Dato che l'Average Energy aumenta fortemente per grandi ampiezze di segnale (campioni elevati al quadrato), conviene introdurre una nuova grandezza: l'**Average Magnitude**.



Zero Crossing Rate: Frequenza di passaggi per lo zero, indica con quale frequenza cambia segno l'ampiezza del segnale.

$$ZCR = \frac{\sum_{n=1}^N |\operatorname{sgn} x(n) - \operatorname{sgn} x(n-1)|}{2N}$$

$\operatorname{sgn} x(n) = \begin{cases} 1 & \text{se } x(n) > 0 \\ -1 & \text{se } x(n) < 0 \end{cases}$

Silence Ratio: Indica la quantità di silenzio nel brano.

$$\text{Silence Ratio} = \frac{\text{Somma dei periodi di silenzio}}{\text{Lunghezza totale del brano}}$$

Caratteristiche derivate dal Dominio delle Frequenze

Un segnale può essere **scomposto nelle frequenze che lo compongono**, tale scomposizione deriva dalla rappresentazione del Domini temporale a cui si applica la [trasformata di Fourier](#).

Il Dominio delle Frequenze mostra il modo in cui è distribuita l'energia alle varie Frequenze. Questa rappresentazione è comunemente detta **Spettro del Segnale**.

Dalla trasformata di Fourier si derivano:

- **Bandwidth:** Range delle frequenze del suono
 - Tipicamente la musica ha un range più ampio rispetto al parlato, frequenze inferiori ai **7kHz** indicano file audio contenenti parlato.
 - Il **Centroide Spettrale** indica il punto medio della distribuzione dell'energia sonora, da questo possiamo dire che il **Centroide Parlato è inferiore al Centroide musicale**.



- **Armoniche:** Un suono prodotto da un **corpo vibrante non è mai puro**. Gli armonici sono suoni alti e bassi più o meno intensi che si accavallano.
 - Hanno una grande rilevanza per la **determinazione del timbro di uno strumento** e nella determinazione di **intervalli musicali**.
 - In linea di massima la **musica contiene più armoniche** di altri suoni quali rumori e parlato.
 - Per il **testare** se un suono contiene armoniche si controlla che le frequenze di **componenti dominanti** siano **multiple** di una **frequenza fondamentale**.

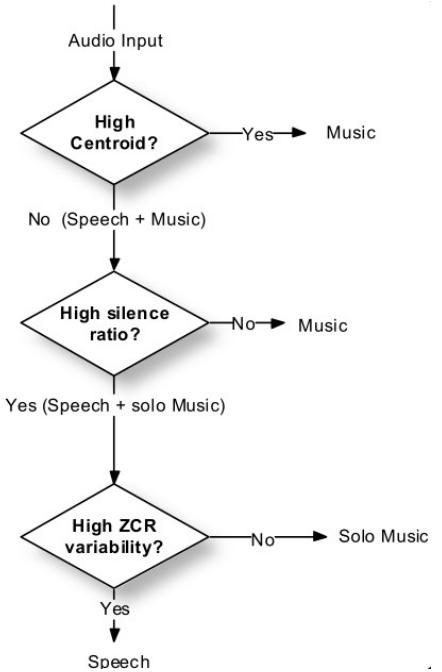
Caratteristiche	Parlato	Musica
Larghezza di Banda	0-7 kHz	0-20 kHz
Centroide Spettrale	Basso	Alto
Quantità di silenzio	Alto	Basso
Frequenza di passaggio per Zero	Molto variabile	Meno variabile
Ritmo regolare	No	Si

Metodi di classificazione dell'audio

Ciascuna **caratteristica** viene **valutata singolarmente**, per poter determinare se si tratti di un **brano musicale** o di **parlato**. Ogni caratteristica è un possibile criterio di selezione che può essere in grado di stabilire la tipologia di un pezzo audio.

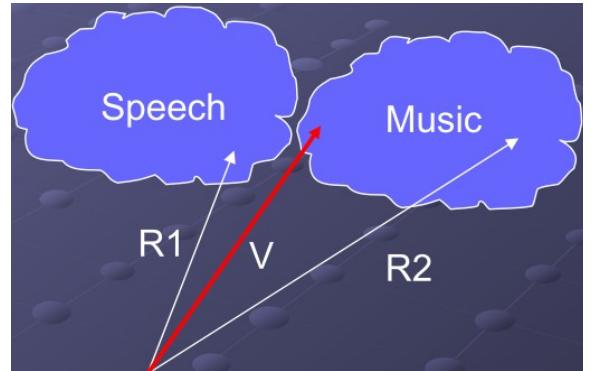
Classificazione Step By Step:

- Il primo filtro valuta la posizione del **Centroide**: quando il valore **superà una certa soglia** prefissata allora si presume che si tratti di un **brano musicale**.
- Si valuta la **quantità di silenzio** (**Silence ratio**), tipicamente un **basso silenzio** implica un **brano musicale**.
- Si valuta la frequenza di passaggi per lo Zero ovvero il **Zero Crossing Rate**.



Classificazione basata su **caratteristiche vettoriali**:

- I valori di un insieme di **caratteristiche del suono**, costituiscono le componenti di un **Vettore** che verrà **confrontato** con un altro **Vettore di Riferimento**.



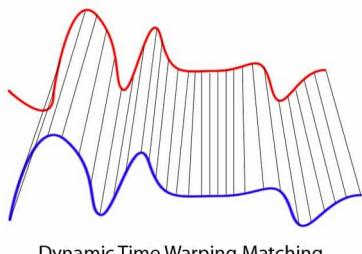
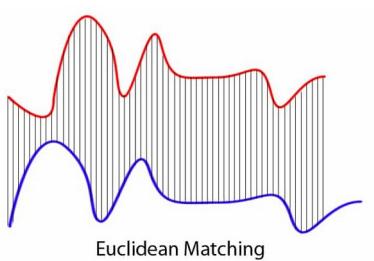
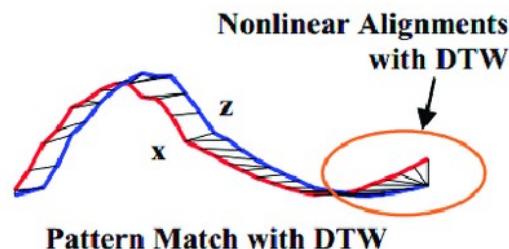
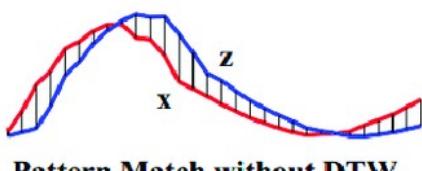
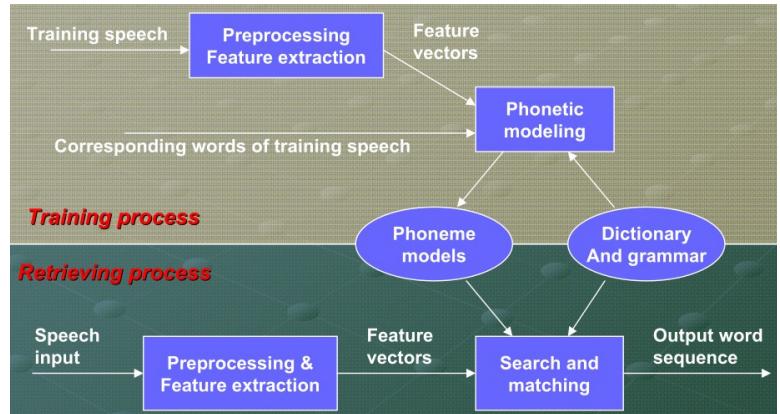
Riconoscimento del parlato

Come anticipato prima, il parlato una volta che è stato identificato, viene convertito in testo.

L'Automatic Speech Recognition (ASR) prende un file audio e genera un testo, l'IR si occuperà quindi di estrarre ed indicizzare i contenuti.

Un sistema ASR viene **istruito** mediante dei **campioni audio** a riconoscere determinate parole ed il tono con le quali vengono pronunciate.

In generale un sistema ASR ben performante, suddivide il parlato in **singole unità** che vengono **confrontate con i Feature Vectors** raccolti durante la fase di **Training**.



Dynamic Time Warping: Questa tecnica cerca di **normalizzare** la durata dei Frame audio del parlato, così da poterli confrontare con i vettori memorizzati in fase di Training.

Il **DTW** è un noto algoritmo applicato in **diverse campi** che esulano anche dall'ambito audio come: Riconoscimento della scrittura, gesti, animazioni, movimento, etc...

ES. Per le sequenze di proteine vengono ricercati pattern comuni da allineare.

Altre tecniche

HMM: Tecnica molto usata e con ottimi risultati, sia per il riconoscimento dello scritto che del parlato. Hiden Markov Models è costituito da: Stati, Probabilità di transizione e probabilità di generazione dei simboli.

ANN: Artificial Neural Network largamente impiegate per il riconoscimento simulando i processi cognitivi del cervello umano.

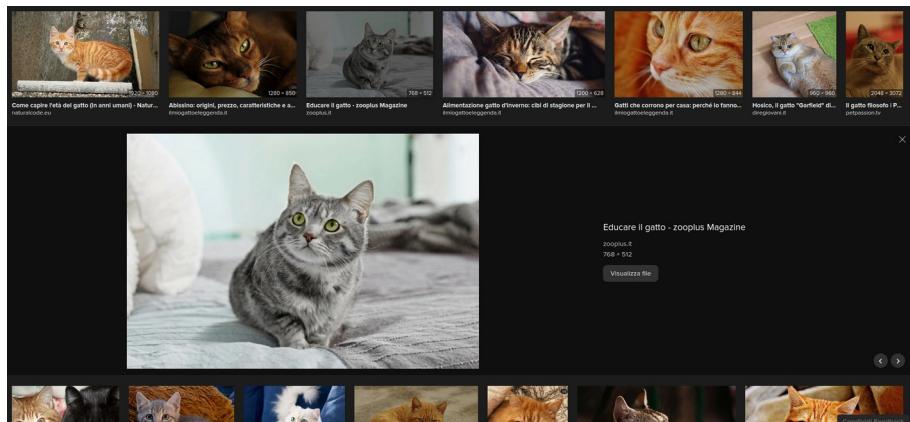
Le ANN sono basate su implementazioni che prevedono l'uso di un grande numero di elementi di calcolo (Processing Elements **PE**), semplici ed interconnessi tra loro, come dei **neuroni**. Le connessioni tra le PE rappresentano le **sinapsi**.

Le sinapsi processano i dati passati, sfruttando input ed output dei neuroni:

- Ogni neurone PE della rete effettua una somma pesata, **Integrazione** degli input derivati dalle connessioni con gli altri neuroni.
- L'input passato viene poi valutato da una funzione detta **Trasformazione** che determina l'output del singolo PE.

Indicizzazione e Ricerca delle immagini

La stragrande maggioranza dei media presenti su internet non sono taggati manualmente, quando si ricerca la foto di un gatto su internet, nella maggior parte delle volte la foto non si chiama gatto.png ma ha magari il nome che la macchina fotografica allo scatto gli ha assegnato.



Il link dell'immagine è: <https://www.zooplus.it/magazine/wp-content/uploads/2020/10/1-32-768x512.jpg> il nome dell'immagine è “1-32-768x512.jpg”, nonostante questo l'immagine è stata fornita dalla ricerca.

Sono 4 i principali approcci per l'indicizzazione e ricerca delle immagini.

- **Tradizionali:**

- Metadati: Nome file, Categoria, Data creazione, Autore, etc...
- Testo: Annotazioni

Non possono descrivere in maniera completa le immagini, le Query restano quindi limitate ai soli attributi del file

- **Basati sui contenuti:**

- Caratteristiche a basso livello: Colori, Texture
- Riconoscimento degli oggetti: Computer Vision, [OCR \(Optical character recognition\)](#)

Computazionalmente molto pesanti dato l'uso di algoritmi estremamente sofisticati.

Annotazioni libere

Le immagini sono descritte con un **testo libero non controllato**. Le **Query** sono sotto forma di **parole chiave** concatenate con operatori booleani.

La ricerca sfrutta quindi i **tradizionali algoritmi di IR** basati sulla ricerca di similarità tra Query e testo descrittivo.

Questa è una **tecnica manuale** ed ha quindi tutti gli **aspetti negativi che ne conseguono** come la **soggettività** della descrizione, **l'incompletezza** ed **inconsistenza** e la necessità di integrazione di un dizionario per il riconoscimento dei sinonimi usati nella descrizione.

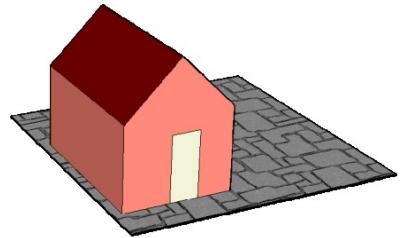
Ha però il grosso **vantaggio** di poter descrivere ed indicizzare **concetti di alto livello**, come: Persona che sorride o che è triste.

Estrazione Feature di basso livello

L'estrazione di Feature di basso livello è definita su tecniche di indicizzazione e ricerca basate sul contenuto (**Content-Based**).

Possono essere prese in considerazione una o più delle seguenti caratteristiche:

- Colore
- Forma
- Texture



Algoritmi basati sull'analisi del colore

Ogni immagine è memorizzata assegnando ai pixel tre valori numerici corrispondenti tipicamente ai tre dell'**RGB** che possono assumere valori da 0 a 255.

Ogni canale è discretizzato in m intervalli **quantizzazione dei colori**.

I **bins** sono una **quantizzazione di un gruppo di sfumature** (Classi di colore); anziché considerare ogni singola sfumatura di un colore, esse vengono aggregate in un bin.

L'**istogramma di colore** è definito dal vettore:

- $H(M) = (h_1, h_2, \dots, h_j, \dots, h_n)$
 - h_j : Rappresenta il numero di pixel dell'immagine M che ricadono nel bin j

Indicizzazione basata sull'istogramma di colore (Color-Based)

Per ogni immagine si calcola l'**istogramma di colore** $H(M)$ che verrà utilizzato come **indice** dell'immagine M .

Sostanzialmente per effettuare la **ricerca di immagini simili ad un'immagine fornita come Query**, basta definire una misura di **distanza** tra i due istogrammi.

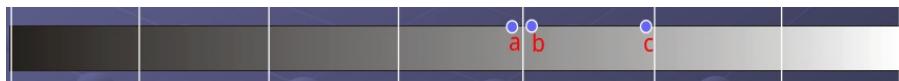
Date due immagini **A** e **B**:
$$d(A, B) = \sum_{i=1}^n |a_i - b_i|$$

a_i e **b_i**: Il numero di pixel dell'immagine A e B che ricadono nell'i-esimo bin.

Problematiche

La discretizzazione dello spazio di colori in Bins (Classi) **non tiene conto della similarità di colori**.

Due **Bins adiacenti** sono considerati completamente **diversi**.



Anche se simili, i punti "a" e "b" sono considerati completamente diversi, invece "b" e "c" sono considerati simili, anche se "b" somiglia molto più ad "a" che a "c".



Il punto 11 sarà considerato uguale al punto venti ma diverso dal punto 10, questo esempio è ovviamente un'esasperazione del concetto, ma rende l'idea.

Soluzione 1 – Distanza tra i bins

La differenza tra Bins si definisce come misura di similarità calcolata bin per bin.

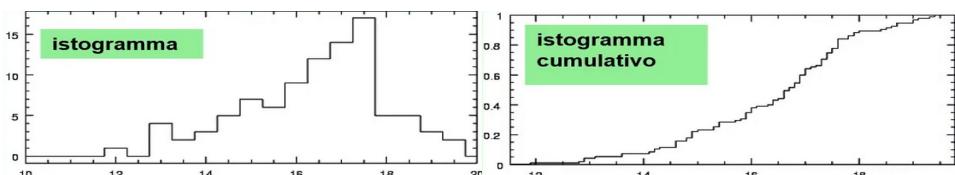
$$\|Z\| = Z^T A Z$$

Soluzione 2 – Iстограмма cumulativo

Non considera la distanza tra i bin ma crea delle classi cumulative per aggirare il problema.

L'istogramma cumulativo è l'istogramma i cui valori sono ottenuti associando a ciascuna classe la **somma dei valori, della stessa classe e delle classi che la precedono**:

$$CH(M) = (ch_1, ch_2, \dots, ch_n) \text{ in cui } ch_i = \sum_{j \leq i} h_j$$



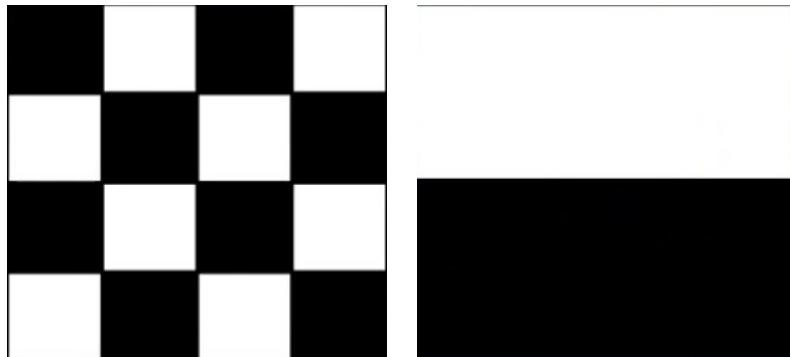
Soluzione 3 – PWH

Perceptually Weighted Histogram si basa nel **cambiare lo spazio** di rappresentazione dei colori da **RGB** a **CIELuv**. Questo comporta che ogni **singolo colore RGB** può avere **10 diverse corrispondenze in CIELuv**.

Si costruisce quindi l'istogramma dell'immagine nel dominio CIELuv.

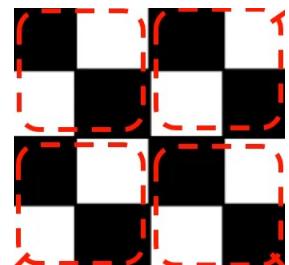
Limite dell'approccio Color-Based

I sistemi di indicizzazione Color-Based hanno il grosso **limite di ignorare le relazioni spaziali** tra pixel.



Queste due immagini hanno la stessa quantità di pixel bianchi e neri, quindi sono **uguali** per l'approccio **Color-Based**.

Per **risolvere** questa problematica si possono utilizzare **tecniche** che prevedono la **suddivisione dell'immagine in regioni più piccole**, il **confronto** quindi sarà effettuato basandosi sugli istogrammi delle **regioni**.



Altra grossa **problematica** è legata al **background** dell'immagine, il background potrebbe influenzare una ricerca Color-Based anche se si è interessati al Soggetto e non allo sfondo. Tale problematica è risolvibile attraverso tecniche di segmentazione automatica, [dividendo il soggetto dell'immagine dallo sfondo](#).

Indicizzazione basata sulla forma

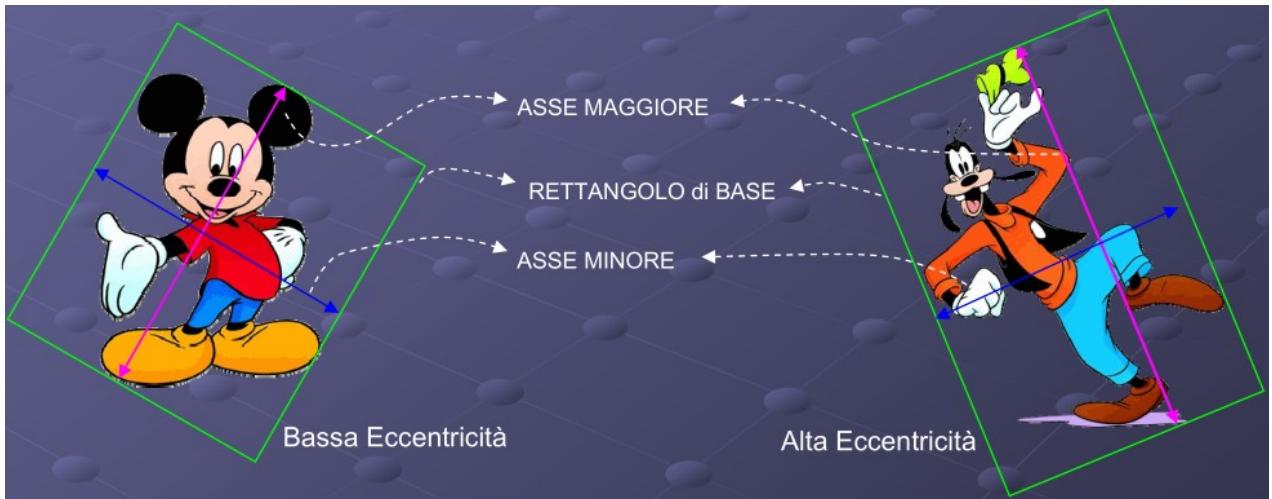
Si basa su **algoritmi di segmentazione** dell'immagine in grado di **suddividere una immagine in singoli oggetti** che la compongono.

I **sistemi efficienti** devono essere capaci di riconoscere ed identificare oggetti anche a seguito di **Traslazioni, Rotazioni e modifiche di scala**.



Una volta identificato il soggetto dell'immagine, esso viene inscatolato in un **rettangolo di base** che cerca di contenere solo il soggetto in analisi eliminando tutto il resto.

All'interno del **rettangolo di base**, vengono tracciati **asse maggiore** ed **asse minore**, così da calcolare l'**eccentricità** dell'immagine, ovvero il rapporto tra la lunghezza dell'asse maggiore e la lunghezza dell'asse minore.



Durante i confronti, immagini con eccentricità diversa da quella ricercata, vengono scartate.

Quando un immagine viene ricercata, è preferibile effettuare una **normalizzazione** di scala dell'immagine e delle immagini da confrontare, in questo modo avranno tutte lo stesso **asse maggiore**.

Il criterio di **univocità di rappresentazione di una forma** è pienamente determinato se:

- Viene fissata la **CellSize**
- Si presume che l'**asse maggiore** sia univoco
- Si effettuano operazioni di normalizzazione di **Rotazione**
- Si effettuano operazioni di normalizzazione di **Scala**
- Per ogni forma si conosce l'**eccentricità**

Calcolo della similarità

Un sistema di indicizzazione e ricerca di forme basato su Regioni può essere realizzato in modo che sia **invariante a rotazioni, traslazioni, modifiche di scala e trasformazioni speculari**, nel seguente modo:

- Si determinano gli **assi maggiore e minore** e quindi si calcola l'**eccentricità**, si ruota la forma in modo che l'asse maggiore coincida con l'asse delle X.
- Si **scala** la forma in modo che l'asse maggiore abbia una **lunghezza prefissata**.
- Si sovrappone una **griglia regolare** alla forma.
- Si **assegna 1** (dove la figura è presente) o **0** (dove è presente il background) ad ogni **cella della griglia** così da ottenere una **stringa binaria** leggendo la griglia da SX a DX partendo dall'alto.
- L'**indice** della forma è quindi costituito dalla **stringa binaria** composta e dal numero di celle che occupa.

Indicizzazione basata sulle Texture

La texture descrive **aree** di un'immagine **caratterizzate** da una **trama comune**. Una texture è **identificabile** nella sua struttura per:

- Granularità
- Contrasto
- Direzionabilità (direzione dominante)
- Regolarità
- etc...

Sono possibili misurazioni matematiche atte ad trovare ed indicizzare tali immagini.

Indicizzazione e Ricerca del video

Il video è un flusso dati è ricco di informazioni, come:

- **Sottotitoli** (Testo)
- **Audio** come musica e parlato
- **Frame** Immagini riprodotte a velocità costante
- **Metadati** Titolo, Autore, Produttore, etc...

Le principali metodologie impiegate per l'indicizzazione ed il recupero di un video sfruttano:

- **Metadati**
- **Testo:** Sottotitoli o trascrizione dei dialoghi.
- **Audio:** Segmentazione in parlato e non parlato, così da applicare tecniche di IR sul testo ottenuto mediante Speech Recognition.
- **Video:**
 - **Singolo fotogramma:** Il video viene considerato come una serie di **immagini indipendenti** su cui applicare le tecniche di indicizzazione delle immagini.
 - **Gruppi di fotogrammi:** I frame vengono suddivisi in **Shots** e l'indicizzazione si basa sui frame rappresentativi di ogni Shot.

Il risultato migliore, come al solito, lo si ottiene combinando più tecniche.

Indicizzazione basata sugli Shot

Gli **Shot** sono gruppi di frame che rappresentano lo **stesso contesto**:

- Frame della **stessa scena**.
- Frame non interrotti da uno **stacco di camera**.
- Frame che contengono lo **stesso evento** o azione.

In molti casi non è solo necessario individuare quale video contiene una determinata **informazione**, ma anche quale **parte del video** la contiene. Per far sì che ciò sia possibile è necessario **segmentare** i video in **Shots**.

Segmentazione automatica in Shots

Generalmente Frame consecutivi **non hanno cambiamenti radicali** di colore, sfondo, soggetti; quando **avviene un cambio radicale** allora è giusto assumere che i frame appartengano a **Shot** distinti.

Per individuare gli Shot occorre quindi definire una **misura quantitativa** che catturi le differenze tra **coppie di Frame**:

- Data una certa soglia se la **differenza** tra il frame **attuale** ed il **successivo** supera tale valore, il punto viene considerato come **interruzione di Shot**.

Tipicamente la suddivisione di un video in Shots può essere effettuata con 2 metodi:

1. Si calcola la somma **pixel per pixel** delle differenze tra 2 frame consecutivi
 - Questo metodo ha scarsi risultati, dato che la presenza di oggetti in movimento causa grandi differenze in una scena.
2. Si calcola la **differenza tra gli istogrammi** di colore di 2 frame consecutivi
 - Il movimento di un oggetto all'interno di una scena **non altera eccessivamente l'istogramma** di colore tra i frame.
 - In particolare si definisce la distanza tra i frame come:
$$SD_i = \sum_j |H_i(j) - H_{i+1}(j)|$$
 - Se **SD_i** supera una certa soglia allora si verifica un cambio di Shot
 - La **scelta della soglia** per l'individuazione degli Shot è critica per l'ottenimento di risultati corretti. Una buona tecnica sta nello scegliere un **valore più alto della media** di tutte le differenze tra un frame ed il successivo.

Transizioni tra Shots

Spesso in un video non sono utilizzate **transazioni a taglio netto**, ma si utilizzano delle **dissolvenze incrociate o Fade-in e Fade-out**. Se bene all'occhio sia molto più **gradevole** una transizione di questo tipo, per gli **algoritmi** di segmentazione automatica in Shots crea un bel **problema**.

Questo tipo di problematiche possono essere risolte con tecniche di **segmentazione a due soglie**.

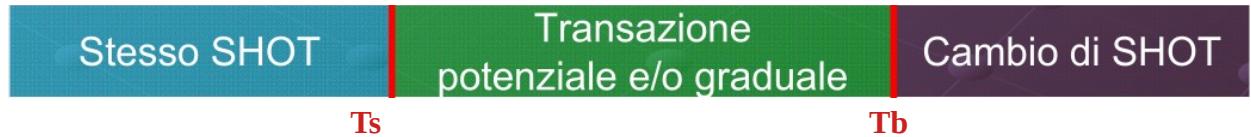


Segmentazione a due soglie

Un meccanismo utilizzato per risolvere la problematica precedentemente descritta consiste nel uso di **due soglie distinte** per determinare i cambi di scena:

- La soglia **Tb** è utilizzata per determinare i cambi di camera.
- La soglia **Ts** per determinare i frame nei quali avviene una transazione graduale.

Durante il confronto tra frame se la differenza è maggiore di Tb si introduce un **cambio di Shot**, se invece è minore di Tb ma superiore a Ts, il frame è dichiarato come **potenziale frame di transizione**.



Se si ottengono **molti potenziali frame di transizione consecutivi**, la loro differenza viene sommata e quando raggiunge la soglia Tb, si genera un cambio di Shot.

Le percentuali di riconoscimento corretto degli Shots si abbassano radicalmente con questa tecnica.



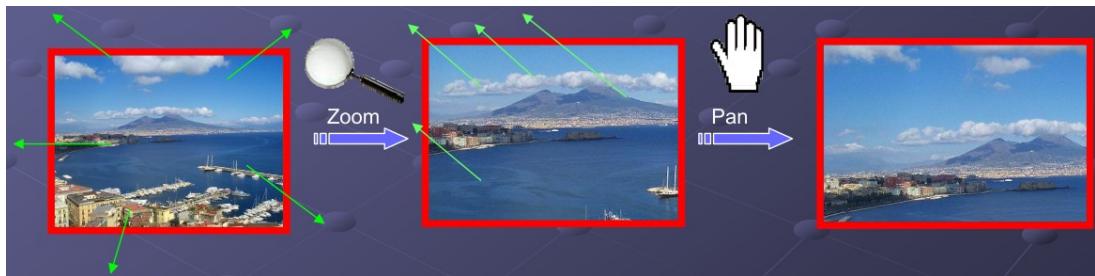
*ES. Immaginiamo di star inquadrando i **giardini** di Buckingham Palace a Londra e di spostare l'inquadratura progressivamente sempre più verso la **residenza**; l'istogramma di colore tra i Frame varierà dal **Verde al grigio** gradualmente sempre di più, i Frame verranno segnalati come **potenziali frame di transizione**, una volta inquadrata la residenza la soglia Tb sarà stata superata e si otterrà un **cambio di Shot non voluto**.*

Panning e Zoom

Le operazioni di **Panning** e **Zoom** effettuate dalle telecamere possono generare cambiamenti graduali come le transizioni discusse prima. Per limitare questo fenomeno si introducono tecniche d'analisi per **identificare il movimento di blocchi di pixel** nei frame.

Nel caso di Panning e Zoom i **vettori di movimento** hanno un preciso comportamento:

- Nel **Panning**, vanno tutti nella **stessa direzione** (Traslano).
- Nello **Zoom convergono o divergono** dal punto centrale dell'immagine.



Cambi di illuminazione

Cambi di illuminazione repentina e radicali come un **Flesh**, alterano **l'istogramma di colore** dell'immagine e possono quindi portare l'algoritmo a credere che si tratti di uno **Shot diverso**.

Tale problematica oltre a poter essere risolta mediante l'analisi del soggetto del Frame ([analisi della Forma](#)) è anche più facilmente risolvibile **Normalizzando i colori** del Frame.

ES. Aumentando la luminosità dell'immagine di k volte, il pixel $P1(R1,G1,B1)$ si trasforma in $P2(kP1, kG1, kB1)$. $P1$ e $P2$ normalizzati coincidono.

Utilizzo degli Shot

Dopo aver identificato gli Shots, occorre **rappresentarli** ed **indicizzarli** così da poterli gestire in modo efficiente in fase di Ricerca.

Per ogni Shot si determinano uno o più **Frame rappresentativi** (r-frame) che vengono usati per l'indicizzazione e la ricerca con [tecniche simili a quelle usate per le Immagini Statiche](#).

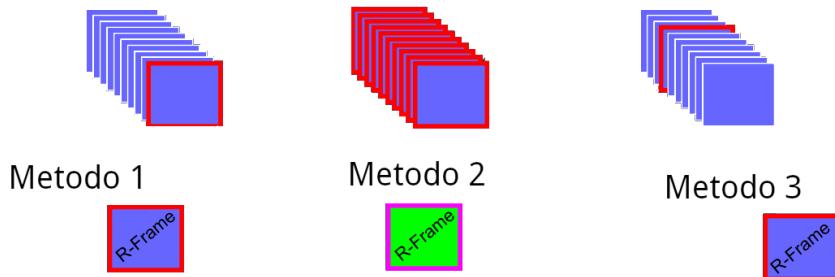
Quanti r-frame occorre utilizzare per ogni Shot?

- **1 r-frame:**
 - Non vengono considerate correttamente le differenze di lunghezza tra i vari Shot
 - Non è catturato completamente il contenuto di uno Shot
- **N r-Frame:**
 - Gestisco la lunghezza ma non il contenuto
 - Sovrabbondanti per Shot uniformi

- Suddivisione in **sotto-shot**
 - Mediante le stesse tecniche utilizzate per la determinazione degli Shot, si introduce un **nuovo livello** così da generare un r-frame per ogni sotto-shot.

Individuazione di r-Frame

L'individuazione di un r-Frame può variare in base alla tipologia del video; se il video è abbastanza **statico**, ogni Frame può essere rappresentativo, d'altro canto se ci sono **oggetti in movimento** la scelta dell'r-Frame ha un'importanza rilevante per ottenere buone performance in fase di ricerca.



Possibili metodi di individuazione:

1. Si considera come r-Frame il **primo Frame** dello Shot.
2. Si considera un **Frame medio** facendo la media dei colori di tutti i frame dello Shot (Per l'elaborazione ha il suo significato, all'occhio umano risulta essere una **macchia di colori**).
3. Si calcola la **media di tutti gli istogrammi di colore dei Frame** dello Shot e si sceglie come r-Frame il **Frame più vicino all'istogramma medio**.

Motion Vector

Ogni Shot può essere indicizzato attraverso i Motion Vectors che lo caratterizzano. Questo rende possibile la ricerca di un video per elementi che lo compongono ed il loro comportamento.

Riconoscimento degli oggetti

Così come per le immagini, gli oggetti focus di uno Shot possono essere identificati. Nel caso dei video è possibile identificare il movimento degli oggetti identificati nell'arco del tempo.

Micon - Motion Icon

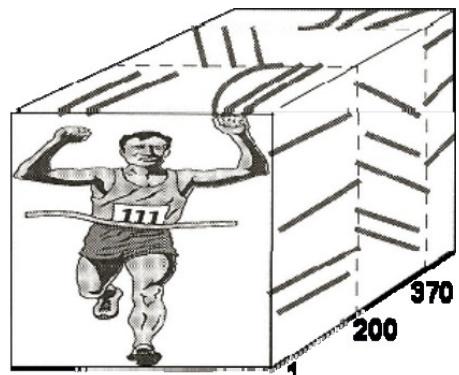
Rappresentazione compatta di uno Shot caratterizzata da 3 dimensioni:

- Un r-frame
- La profondità (Durata)
- I pixel presenti sui bordi orizzontali e verticali (Danno l'idea del movimento)

Il Micron può essere utilizzato per implementare operazioni come:

Browsing: Portando in primo piano qualsiasi frame dello Shot.

Taglio: Tagliando orizzontalmente il cubo si possono ottenere informazioni sul suo contenuto.



Browsing Video

Strumento che consente all'utente di **ispezionare le varie parti della struttura video**, saltando da un punto all'altro senza rispettare una precisa sequenza.

La struttura video è così composta ed inscatolata:

- Livello Video (Titolo video)
 - Livello Scene
 - Livello Shots
 - Frames

Altre rappresentazioni compatte

Storyboard: Collezione di r-Frame che rappresentano le porzioni più importanti del video.

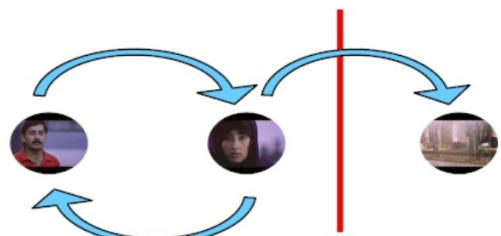
Mosaicatura: Unione di Frame che descrivono nel loro insieme qualcosa di più complesso.

Scene Transition Graph (STG): Struttura che cattura in modo compatto sia il contenuto che il flusso temporale.

Questa tecnica utilizza una **struttura a grafo orientato** per unire tra loro diversi **r-Frame** in successione temporale.

Un grafo $G = (V, E)$ in cui:

- **V:** Vertici che rappresentano cluster di Shots simili.
- **E:** Archi che uniscono coppie di vertici se esistono Shot che si precedono.



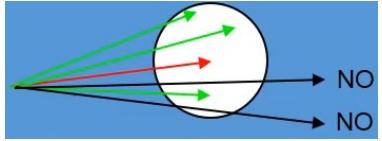
Tipi di Query e Vettori

Una Query per la ricerca è [rappresentata da un vettore](#), la ricerca di un vettore corrispondente può quindi avere diverse forme:

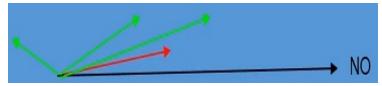
Point Query: Quella di trovare un vettore delle Feature che corrisponda esattamente con la Query (Exact Match).



Range Query: Tutti gli oggetti la cui distanza dalla Query è inferiore o uguale ad una certa distanza specificata sono presi in considerazione.



K-Nearest Neighbour: I K oggetti con le distanze minori del vettore Query sono presi in considerazione.



Il filtraggio basato sulla **disuguaglianza triangolare** consiste nel calcolare la distanza tra i vari file indicizzati salvando i risultati in una matrice. Una volta eseguita tale operazione, quando si esegue un **Recupero** sarà inutile calcolare la distanza di un oggetto, se si è già calcolato che un suo simile è troppo distante dalla ricerca. Con questa tecnica si risparmiano confronti velocizzando le operazioni di Recupero, comunque sia questa operazione pesa computazionalmente nella fase di Indicizzazione.

IV Strutture dati per la ricerca delle similarità

La fase di indicizzazione genera dei Features Vector, ovvero dei vettori formati dalle caratteristiche di un determinato file.

La fase di Recupero è quindi caratterizzata da un gran numero di **confronti** di caratteristiche, tra la **Query** e le caratteristiche degli **oggetti** memorizzati.

NON è quindi neanche pensabile eseguire dei confronti in modo lineare su tutti gli oggetti.

Pertanto è necessario l'uso di **strutture dati** adatte ad una ricerca rapida per oggetti **non omogenei**.

Alberi B

Un albero B di ordine m (m rappresenta il **massimo numero di figli** che un nodo può avere), è un generico albero di ricerca che gode delle seguenti proprietà:

- La **radice** ha almeno 2 sotto alberi, a meno che non sia foglia
- Ogni **nodo interno** (non radice non foglia) contiene k – 1 chiavi e k riferimenti a sotto alberi, in cui $\lceil M/2 \rceil \leq k \leq m$
- Ogni nodo **foglia** contiene k – 1 chiavi, in cui $\lceil m/2 \rceil \leq k \leq m$
- Tutte le **foglie** si trovano sullo **stesso livello**

Per la definizione data, un albero B è sempre pieno almeno per metà, ha pochi livelli ed è perfettamente bilanciato. Sintetizziamo infatti dire che un albero B:

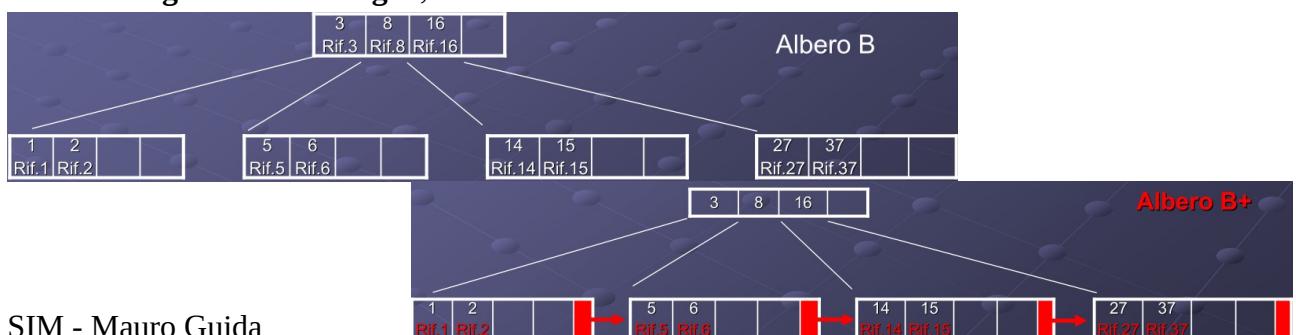
- E' un albero **Bilanciato**
- La **complessità** delle operazioni di **ricerca e/o attraversamento** dell'albero sono **prevedibili**

Le operazioni fondamentali per gli alberi B sono: **Creazione, Inserimento, Cancellazione, Ricerca**.

Alberi B+

L'albero B+ è un'evoluzione dell'albero B, in particolare in un albero B+:

- I **riferimenti** ai dati sono contenuti **solo nelle foglie** anziché essere contenuti in qualsiasi nodo.
- Le foglie di un albero B+ contengono anche un campo **Puntatore** aggiuntivo **per la navigazione delle foglie**, come una Linked List.



Alberi MB+ - B+ Multidimensionali

Gli alberi Multidimensionali MB+ sono una estensione degli alberi B+ Standard monodimensionale.

Questo tipo di alberi supporta le **Similarity Query**, ovvero Query per **intervalli e prossimità**.

- Ogni **Feature Vector** è un punto nello spazio 2D
- L'intero spazio delle feature è una **Bounding Box**, una scatola con dei limiti ben definiti
- La **Bounding Box** è divisa in **regioni** contenenti **Feature Vector** simili per caratteristiche

Ogni regione contiene una lista di **vettori di caratteristiche** e non dati.

Cluster

Come già discusso il modello basato su Cluster ottimizza i tempi di ricerca delle Features aggregando file simili in blocchi detti Cluster.

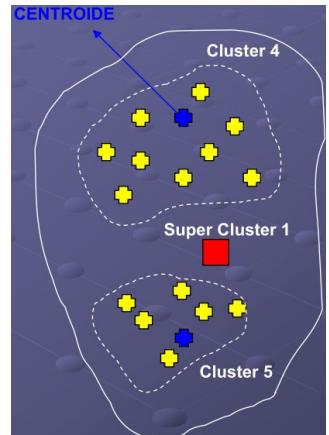
Vettori di Features simili vengono raggruppati in cluster in base ad un **calcolo di similarità**, se un vettore è troppo distante da qualsiasi altro Cluster presente, allora formerà da se un **nuovo Cluster**.

Il calcolo delle similarità di una Query con un Cluster avviene **confrontando la Query al Centroide** di ogni Cluster (Il **Centroide** è calcolato mediando tutti i vettori del Cluster). Vengono quindi presi in **considerazione solo i Cluster più simile al Vettore Query**, così da analizzare nel dettaglio meno informazioni possibili.

Va da sé che se un Cluster ha una **dimensione troppo elevata**, contenendo troppi elementi, allora questo processo non semplificherà affatto le operazioni di ricerca.

Per ovviare alla problematica descritta si creano **Super Cluster**, ovvero Cluster che contengono altri Cluster più piccoli.

Ogni **Super Cluster** avrà il proprio **Centroide**, ottenuto **mediando** i Centroidi dei Cluster che contiene.



Alberi K-d

I K-d Trees sono una estensione degli Alberi Binari.

Ogni **chiave** è costituita da un **Vettore K-Dimensionale** anziché da un singolo valore.

L'inserimento e la Ricerca si basano quindi sulle **componenti del vettore**.

Questa struttura rende molto semplice l'implementazione delle [Range Query](#).

Grid files

Costituiscono una modalità di **indicizzazione e ricerca semplice**, questo tipo di strutture sono molto utilizzate nel mondo reale.

I **Grid Files** consistono nella suddivisione dello spazio n-dimensionale in **iper cubi aventi tutti la stessa dimensione**, ogni ipercubo contiene zero o più **Feature Vector**.

Questa struttura rende molto semplice l'implementazione di [Point Query](#) e [Range Query](#).

Questa struttura ha **ottimi risultati** se i vettori delle feature sono **distribuiti abbastanza uniformemente** all'interno dello spazio dei valori, altrimenti alcune griglie saranno vuote ed altre sovraffollate.

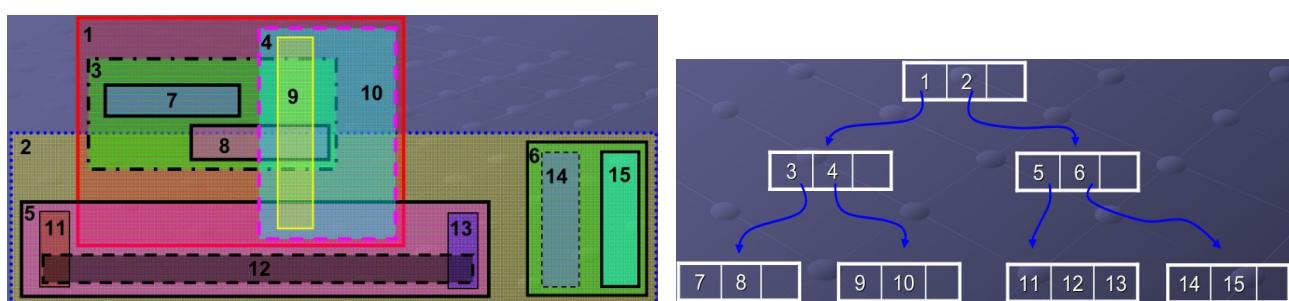
Una **Griglia sovraffollata** avrà un puntatore ad una **lista di vettori molto lunga** e quindi **difficile da scorrere**.

Per risolvere tale problematica, invece di utilizzare griglie di dimensione fissata, si crea una **suddivisione adattiva** cercando di **bilanciare il contenuto delle diverse griglie**.

Nelle **zone densamente popolate** si utilizzano **griglie più piccole** mentre nelle zone scarsamente popolate si fa il contrario.

Alberi R

Gli alberi Rettangoli sono multidimensionali ed una generalizzazione degli [alberi MB+](#). Questa struttura dati divide lo spazio in **MBR** Minimum Bounding Rectangles.



Operazioni

Query:

- La regione da cercare viene caratterizzata dal suo MBR (minimum bounding-box)
- A partire dalla radice si attraversa l'albero cercando i rettangoli che intersecano l'MBR (possono essere più di uno ad ognuno dei livelli)
- Raggiunti i nodi foglia, si calcola l'intersezione tra l'MBR e il rettangolo collegato

Insert:

- Si attraversa l'albero selezionando il rettangolo più piccolo che include l'oggetto da inserire o quello che richiederebbe l'allargamento minore per "coprire" il nuovo oggetto.
- L'inserimento comporta l'allargamento del nodo padre per fare in modo che il suo rettangolo includa completamente il nuovo oggetto
- Se il nodo nell'albero è già pieno per più di metà, occorre procedere alla operazione di splitting in maniera analoga a quanto avviene sugli Alberi MB+
- Lo splitting si può ripercuotere ricorsivamente verso l'alto fino a quando l'aggiunta di un nuovo rettangolo non comporta un riempimento eccessivo

Delete:

- Si utilizza un procedimento di attraversamento dell'albero simile a quello della ricerca
- Se l'eliminazione di un oggetto comporta che un nodo dell'albero contiene troppo pochi elementi, il nodo viene eliminato e gli oggetti che conteneva vengono reinseriti nell'albero

Data la natura dell'albero risulta evidente che la ricerca **K-nearest Neighbour** sia facilmente implementabile.

Efficienza di ricerca

Dipende da 2 concetti (definiti per ognuno dei livelli dell'albero):

- **Coverage:** L'area totale di tutti i rettangoli associati ai nodi del livello
- **Overlap:** L'area totale coperta da due o più nodi

Un Albero R è **efficiente** se sia la **Coverage** che l'**Overlap** sono minimizzati; in particolare, l'**overlap comporta problemi in fase di ricerca**.

Inoltre è cruciale l'ordine di inserimento per ottenere un albero maggiormente bilanciato.

V Watermark

Un insieme di **strumenti** e **metodi** per marcare i **qualsiasi file digitale**, così da indicare a chi ne fruisce, chi ne sia il proprietario o anche per **garantirne l'originalità** e la **non alterazione**.

Un esempio di Watermark fisco può essere visto sulle banconote, questa complessa figura difficilmente riproducibile è posta per impedirne la riproduzione non autorizzata.



Il watermark può essere:

- **Evidente:** Spesso accade per immagini scaricate da internet che riportano grosse scritte non cancellabili, per evitarne l'uso non autorizzato.
- **Latenti:** Nascoste all'interno del file come una sorta di **steganografia** (Filigrana digitale).



Esempio di Watermark evidente



Esempio di Watermark Latente

Ciao, questo è un
messaggio segreto

Gli obiettivi del **Watermarking** sono molteplici:

1. Manifestare a tutti gli utenti chi sia il **proprietario**.
2. Dimostrare **l'originalità** ed impedirne la contraffazione.
3. Evitare la **distribuzione non autorizzata**.
4. Marcare determinate **caratteristiche**.
5. **Segnare il processo di vendita** del documento utilizzando marchi differenti per ciascun utente.

Classificazione dei Watermark

I Watermark come già accennato possono essere caratterizzati da delle proprietà come la visibilità, che dipendono dallo scopo della marcatura.

Un Watermark può essere

- **Visibile:** Utilizzato solitamente per rendere note alcune informazioni all'utente finale.
- **Invisibile:** Utilizzano nei contesti in cui il proprietario legittimo di un dato documento vuole garantirsi i diritti d'autore. La copia marcata è uguale all'originale alla percezione umana.

Un Watermark può avere diversi **tipi di resistenza** a tentativi di contraffazione, esso può infatti essere:

- **Fragile:** Può essere **facilmente attaccato, distrutto** e reso irriconoscibile. Watermark di questo tipo sono facilmente rimovibili. L'uso di un Watermark fragile risulta utile quando si vuole disconoscere un'opera in caso di alterazioni effettuate da terzi.
- **Semi-fragile:** Se l'utente arriva ad un **certo livello di alterazione** si distrugge come quello fragile.
- **Robusto:** Resiste alle più comuni manipolazioni del documento in quanto deve provare la proprietà di un dato documento. L'informazione che trasporta non deve perdere e deve essere **recuperabile in caso di alterazione**. Si definisce Watermark robusto quando è in grado di resistere ad attacchi volti alla sua rimozione.

ES. Nel caso di un'immagine, un esempio di Watermark robusto può essere un'alterazione molto latente ed invisibile all'occhio umano ma presente su tutta l'immagine. Un tentativo di alterazione, come la rimozione della firma dell'autore, risulterà quindi evidente se si effettua un'analisi.

Altra importante caratteristica è **l'autonomia** di un Watermark, che si distingue i Ciechi e Non:

- **Ciechi:** Per verificare la loro presenza NON è necessario il documento originale.
- **Non Ciechi:** Sono **più robusti** ma, per essere verificati, è necessario anche il documento originale.

Inoltre un Watermark può anche essere di Dominio Pubblico o Privato:

- **Watermark privato:** Può essere estratto solo quando si conosce il contenuto e si possiede il documento non marchiato.
- **Watermark pubblico:** Rilevabile anche se non se ne conosce il contenuto e senza l'ausilio del documento originale. Il Watermark pubblico è semplice da identificare e alterare, ma in alcuni contesti sono utili per **indicare il legittimo proprietario di un documento**.

Proprietà dei Watermark

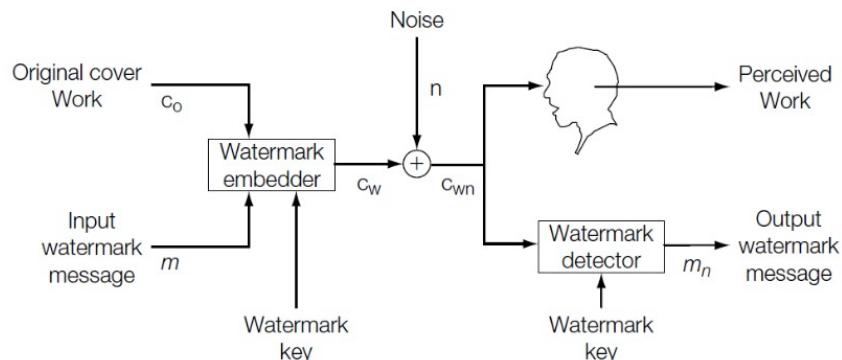
Esistono caratteristiche comuni che tutti i Watermark **devono** rispettare, questo per far si che soddisfino determinate esigenze.

- Il legittimo proprietario o un'autorità di controllo indipendente deve poter **facilmente estrarre le informazioni da un Watermark**.
- Il recupero di un Watermark **NON deve presentare ambiguità** su chi sia il proprietario effettivo di un documento.
- Deve essere possibile **sovrapporre più Watermark**, senza che i precedenti vengano distrutti.
- Il Watermark **deve essere inserito all'interno del segnale da proteggere** per maggiore sicurezza e portabilità.

L'implementazione di queste caratteristiche porta sicuramente ad un **leggero degrado del segnale** che si vuole proteggere. Questo degrado deve essere **scelto dal proprietario**; un elevato degrado rende si più sicuro il documento da eventuali alterazioni, ma comporta anche una minore utilità dello stesso.

Ogni segnale di Watermark è associato ad una **Chiave privata** (Sequenza di bit) posseduta del proprietario del Watermark. Questa chiave serve sia a **firmare il documento** producendo il Watermark, ma anche a **riconoscere eventuali Watermark invisibili** all'interno del documento. Nel caso di **Watermark multipli**, ogni chiave può essere usata per riconoscere la corrispondente firma.

Un Watermark può anche essere **invertibile**, offrendo quindi al legittimo proprietario la possibilità di rimuovere il Watermark dal documento firmato. Tale proprietà va in **confitto con la robustezza**.



Watermarking su testi

Esistono diversi metodi per marchiare un file di testo: **Spaziali, Sintattici, Semantici**.

Una possibile tecnica consiste nell'inserire **spazi al termine delle righe**. Questi spazi sono invisibili ma riconoscibile attraverso la selezione. Anche l'uso di **spazi tra le parole** possono essere usati per costruire un codice.

VI GIS - Geographic information system

Un GIS può essere definito come un **Sistema informatico per la gestione dei dati geografici**; più propriamente è una supercalss di sistemi informatici più completi per il Rilevamento Geografico (Maps, Waze, etc...).

Geographical: Rappresentazione del terreno realizzate attraverso la memorizzazione della localizzazione geografica degli elementi terrestri.

Information: I sistemi GIS contengono informazioni in varie forme (Mappe, Immagini, Relazioni, etc...).

System: Un GIS è un sistema a tutti gli effetti, formato da più componenti Hardware e Software, dati geografici e risorse umane.

Un GIS può essere **scomposto in modo gerarchico**, distinguendo 6 componenti distinte:

1. [Organizzazione](#)
2. [Visualizzazione](#)
3. [Interrogazione](#)
4. [Combinazione](#)
5. [Analisi](#)
6. [Predizione](#)

Organizzazione

Un GIS colleziona **grandi quantità di dati da sorgenti eterogenee**. Affinché il Collage Informativo abbia effettivamente un uso pratico per le funzionalità che un GIS offre, c'è bisogno di una **consistente Organizzazione dei dati**.

Le tipologie di informazioni contenute in un GIS possono essere prevalentemente di 2 tipi ben distinti tra loro:

- **Spaziale** (Geografico)
- **Alfanumerico** (Attributi)

Visualizzazione

Un GIS sfrutta la capacità umana di comprendere situazioni complesse in maniera migliore attraverso un approccio visuale. Per tanto un GIS ha la **capacità di presentare graficamente ed in maniera organizzata**, le informazioni raccolte mediante la creazione di **mappe**.

Interrogazione

Le informazioni presenti in un GIS sono memorizzate attraverso **oggetti grafici**, ovvero elementari detti **Features**. Ogni feature ha delle informazioni alfanumeriche associate come: Indirizzo, città, etc...

Il GIS deve quindi permettere la richiesta di dati in esso contenuti in 3 modi diversi:

- **Interrogazione degli attributi:** Richiedere una determinata Feature attraverso gli attributi ad essa collegati.
- **Interrogazione spaziale:** Richiedere informazioni su punti geografici dove si verificano una serie di condizioni specificate.
- **Interrogazione topologica:** Richiedere informazioni di carattere topologico come: la distanza, l'adiacenza, l'orientamento, etc...

Combinazione

Un GIS, per essere funzionale, deve permettere la **combinazione di informazioni provenienti da fonti diverse**, così da permettere di scoprire **correlazioni** o di generare **dati derivati**.

La combinazione delle informazioni può avvenire prevalentemente in due modi:

- Visualizzazione **congiunta** di informazioni differenti
- Costruzione fisica di informazioni ottenute come **combinazione** di altre

ES. Combinando diversi Layer informativi come il Layer Satellitare ed il Layer Edilizio/Catastale è possibile trovare degli abusi edilizi commessi (analisi).

Analisi

Partendo da dati grezzi un GIS può elaborare informazioni utili al sistema, mediante l'uso di strumenti **Statistici** e di **Modellazione**. L'analisi permette di scovare aree molto trafficate, rotte aeree efficienti, ottenere statistiche di carattere litologico, etc...

Predizione

L'aspetto più avanzato e complesso del sistema GIS in cui la componente progettistica umana è determinante. La Predizione è **fortemente legata all'analisi**, ma si spinge oltre, facendo **previsioni di scenari evoluti e futuri**.

La predizione si basa sulla domanda **What-if?**

La predizione è una componente fondamentale in campi come:

- Predizione territoriale ed urbanistica
- Valutazioni dell'impatto ambientale
- etc...

Questo aspetto è fortemente legato ad altre discipline come la **Simulazione** e la **Multimedialità**.

Caratteristiche

Gli ambiti **CAD** ed **Image Processing** sono molto vicini ai sistemi GIS.

- **CAD:** Strumenti di supporto al disegno tecnico e di tipo **Vettoriale**. Usati per le mappe stradali ovvero la **cartografia digitale**.
Le mappe geografiche sono realizzate con strumenti CAD, questo comporta risparmi di spazio ed una maggiore efficienza per la visualizzazione.
- **Image Processing:** Strumenti per la visualizzazione, gestione ed elaborazione di immagini **Raster**. Usati per le immagini satellitari e fotografie aeree.

Banalmente anche durante la navigazione con Software come Maps e Waze, l'utilizzo di mappe vettoriali comporta anche un minor consumo di banda rispetto ad immagini satellitari e quindi Raster.

Modello spaziale del mondo reale

Una caratteristica del GIS consiste nell'introdurre la dimensione dello Spazio in un database tradizionale.

Un Database tradizionale è capace di contenere **informazioni territoriali** come: Localizzazione di persone, eventi ed attività, Dati sulle distanze, etc... Ma non è però in grado né di gestire né di derivare **informazioni di tipo spaziale**.

Per fare ciò è necessario introdurre nel DB un **modello spaziale del mondo reale**.

La modellazione del mondo reale e l'organizzazione delle informazioni derivanti, presuppongono un processo diviso in:

- Identificazione delle **entità** che definiscono i fenomeni
- Identificazione degli **attributi** che ne descrivono le caratteristiche rilevanti
- Identificazione della **topologia** e delle **relazioni spaziali** dei fenomeni

I fenomeni del mondo reale tradotti in entità GIS, prendono il nome di **Oggetti**.

Oggetto GIS

Un **oggetto rappresenta in forma semplificata fenomeni del mondo reale**, conservandone le proprietà spaziali. Un oggetto GIS è caratterizzato da:

- Tipo
- Attributi Relazioni spaziali
- Geometria
- Qualità del dato

Ogni **elemento** all'interno di un GIS è costituito da **collezioni di oggetti** rappresentati da **figure geometriche**.

Oggetti base di un GIS

Una rappresentazione spaziale univoca non esiste, la scelta degli oggetti dipende dall'analisi che dovrà essere fatta del fenomeno. Dalle proprietà dei singoli oggetti risulta possibile effettuare un determinato tipo di operazioni anziché un altro.

Punto: Punto isolato. Se fa parte di una linea è detto **Vertice** o **Nodo**.

- Entità **Adimensionale** che specifica la localizzazione di un fenomeno nello spazio, codificandone la posizione.
- Un punto può essere di diversi tipi:
 - **Entità:** Identifica la posizione di un fenomeno nello spazio, senza tener conto della forma e della dimensione.
 - **Area:** Identifica la posizione di un Area come un Centro Urbano.
 - **Nodo:** Collegamento tra due o più punti tramite segmento.

Arco o Linea: Sequenza ordinata di vertici. Il primo punto è detto **Nodo iniziale** l'ultimo **Nodo finale**.

- Entità **Unidimensionale** che ha la capacità di rappresentare un fenomeno reale specificando:
 - Posizione
 - Direzione
 - Lunghezza
- Può essere rappresentata con **diversi spessori** che ne identificano attributi metrici **diversi dalla lunghezza**.
- Può avere un attributo che ne specifica l'orientamento.
- Usato per Strade, Fiumi, etc...

Anello: Insieme di archi che costituiscono un anello chiuso.

Poligono:

- Insieme di uno o più anelli che delimitano un'area chiusa.
- Il primo anello è detto **Esterno** gli altri **Interni**.
- Se ha un solo anello è detto **Semplice** altrimenti **Complesso**.
- Entità **Bidimensionale** che ha la capacità di rappresentare un fenomeno reale evidenziando:
 - Posizione
 - Morfologia
 - Superficie (Dimensione quadrata)
- Usato per **rappresentare strutture territoriali** come Comuni, Quartieri e Regioni.

Georeferenziazione

Tecnica di **localizzazione territoriale** che permette di associare un oggetto ad un particolare punto nello spazio reale. La rappresentazione degli oggetti nello spazio richiede un **sistema di riferimento spaziale**.

Gli oggetti devono conservare: **Posizione, Dimensioni, Relazioni spaziali**.

Si utilizzano diverse **tecniche** di georeferenziazione:

- Georeferenziazione **Continua**: Misura della posizione di un fenomeno rispetto ad un sistema di **riferimento Assoluto**, generalmente **coordinate terrestri**, rilevate mediante sistemi di posizionamento come il GPS.
- Georeferenziazione **Discreta**: La misura della posizione di un fenomeno è ottenuta **indirettamente** rispetto ad **unità territoriali di riferimento** già georeferenziate.

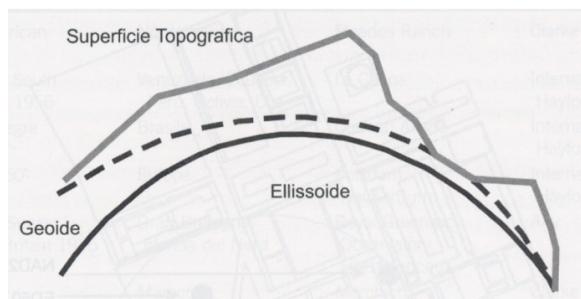
Modelli della terra

La mappa della terra può essere rappresentata in diversi modi:

Modello **Terra Piatto**: Adatto per mappe di dimensioni non superiori ai 2-3Km e pianeggianti (dove la curvatura terrestre non crea grossi problemi), gli oggetti vengono posizionati tutti sullo stesso piano.

Modello **Terra Curva**: Tiene conto della curvatura terrestre, è quindi un modello molto complesso e non univoco, nonostante questo limita notevolmente gli errori nelle relazioni spaziali degli oggetti rappresentati.

- Esistono diversi tipi di implementazione del modello Terra curva:
 - Ellissoide o Sferoide
 - Riproduce lo schiacciamento ai poli ed il rigonfiamento all'equatore.
 - Non tiene conto dell'irregolarità della superficie.
 - **Utilizzato** per la valutazione della posizione geografica di un punto sulla superficie.
 - Geoide
 - Tiene conto delle irregolarità locali della superficie.
 - **Utilizzato** per la valutazione dell'altitudine di un punto.



Datum Geodetici

I sistemi di coordinate attualmente utilizzati si basano sul **Datum Geodetico** che adatta localmente la superficie dell'**ellissoide** a quella del **geoide**.

Originariamente ogni paese aveva definito un proprio Datum al fine di approssimare al meglio la superficie terrestre all'interno della propria area. La diffusione dei sistemi **GPS** ha portato alla definizione di **Datum Globali (WGS84)** che cercano di minimizzare gli errori non al livello locale ma globale.

Datum diversi possono produrre localizzazioni diverse, per questo motivo si è adottato uno standard che dia lo stesso errore a tutti.

Latitudine e Longitudine

Sul modello della terra viene generata una griglia immaginaria al fine di effettuare misurazioni con sistemi di coordinate terrestri.

La griglia è costituita da **linee immaginarie** dette

- **Paralleli** o linee di **Latitudine**
 - Cerchi paralleli all'equatore.
 - Parallello Zero all'equatore.
- **Meridiani** o linee di **Longitudine**
 - Tracciati verticali attraverso il Nord ed il Sud geografico.
 - Viene definito come meridiano Zero quello che passa per l'osservatorio di Greenwich.

All'interno del modello della terra i punti sulla superficie vengono calcolati con strumenti di **geometria sferica**.

Ad ogni punto sulla terra è assegnato una posizione espressa in **Gradi, Minuti e Secondi**.

Proiezioni

La cartografia si occupa di rappresentare su carta la realtà tridimensionale. Il processo di **Proiezione** porta alla costruzione di una mappa 2D della terra.

Ovviamente la costruzione di una mappa comporta il ridimensionamento rispetto alle dimensioni reali; il rapporto tra le dimensioni degli **oggetti riportati** e quelli del **mondo reale** è detto **Scala** ed è espresso mediante frazione numerica.

ES. 1:100 ogni cm su mappa equivale a 100cm reali.

Non esiste alcuna proiezione o trasformazione capace di creare un modello che rappresenti fedelmente una superficie curva su di una piatta **senza introdurre distorsioni**; in particolare le distorsioni possono ripercuotersi su: Distanze, Direzioni, Forme, Superfici, Scala.

Classificazione proiezioni

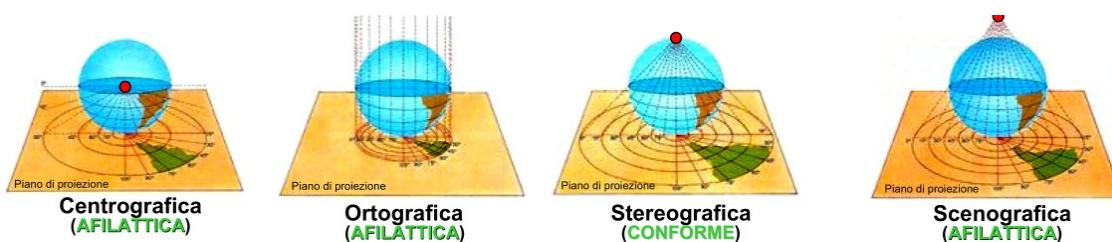
Una proiezione può essere classificata in base a:

Proprietà **fisiche**:

- **Conformi o isogone**: Conservando la forma degli oggetti su scala locale (piccole aree).
- **Ad aree equivalenti**: Riproducono correttamente le superfici in modo proporzionale alla realtà questo a scapito della forma degli angoli e della scala.
- **Equidistanti**: Conservano le distanze tra determinati punti, anche se la scala non è riprodotta correttamente in tutti i punti della mappa. Usata per scenari in cui la distanza da valutare è tra due punti non troppo distanti.

Metodi **geometrici**: (usati per generare la proiezione)

- Proiezione superficie sferica su **Cilindri** o **Coni**.
- **Azimutali** o **piane** la superficie sferica viene riprodotta direttamente sul piano.

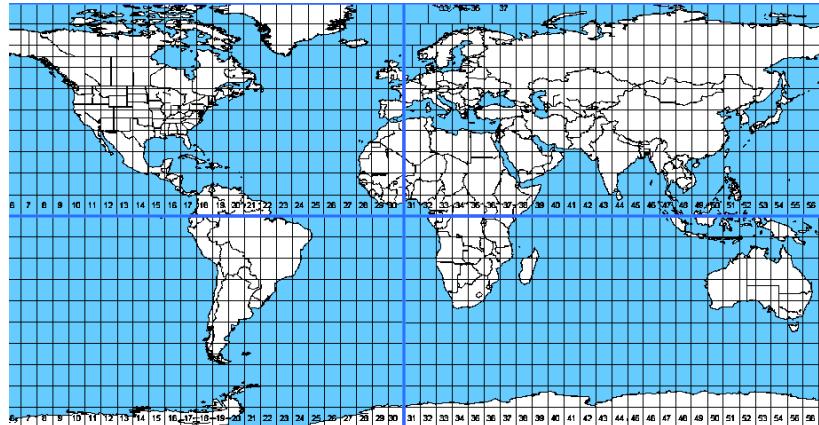


Proiezione UTM (Universale Trasversa di Marcatore)

La proiezione UTM Cassini-Gauss è una modifica della proiezione di Marcatore. L'UTM rappresenta **correttamente la gran parte delle regioni presenti sulla terra.**

Questa proiezione è un **mosaico di proiezioni**, così da minimizzare per ogni elemento gli errori lineari, angolari e superficiali. L'approssimazione ai bordi non è molto precisa ma per quanto riguarda **il centro è molto affidabile.**

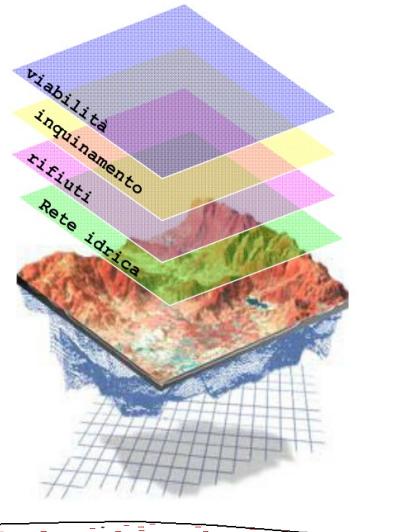
All'interno di ogni zona le coordinate sono descritte in forma cartesiana.



Le mappe nei GIS

In un GIS una mappa è il risultato di sovrapposizioni di più mappe elementari dette **Strati informativi**. Ogni strato informativo ha la caratteristica di essere georeferenziato. I modelli tra di loro sovrapposti sono: Modelli Vettoriali, Raster e Tridimensionali.

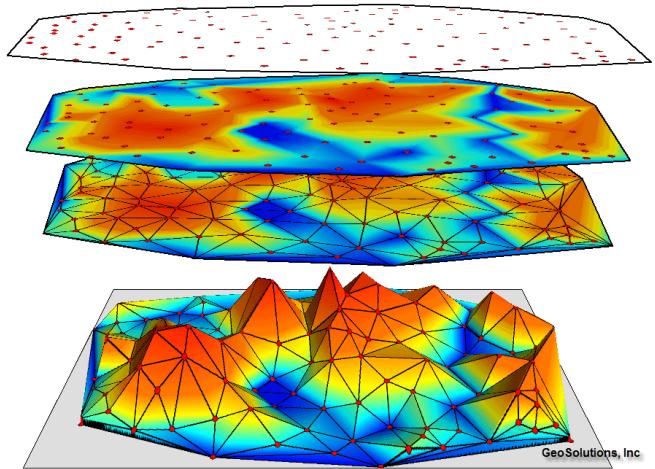
I modelli sovrapposti sono molto eterogenei tra loro e questo comporta una grande difficoltà per un sistema GIS.



Mappa Tridimensionale

TIN Triangulated Irregular Network, è un modello che serve per rappresentare superfici 3D mediante l'uso di **triangoli** che rappresentano in maniera compatta l'andamento della **superficie 3D**.

Il TIN viene costruito rispettando la regola di **Triangolazione di Delaunay**.



Esplorazione dello spazio

Obiettivo della misurazione degli attributi spaziali e non spaziali è quello di **individuare, contare e localizzare** gli oggetti presenti in una regione.

La **tassonomia** dello spazio è la descrizione delle **proprietà Geometriche degli oggetti** e la misurazione delle loro **relazioni**.

Nell'esplorazione dello spazio occorre considerare:

- **Gli attributi spaziali**
 - Geometria degli oggetti
 - Rapporti nello spazio
 - Proprietà topologiche
- **L'analisi multistrato**
 - Analisi territoriale con l'obiettivo di confrontare informazioni contenute in più strati informativi, Overlay di mappe **Raster** e **Vettoriali**.
 - Mappe con lo stesso livello di precisione, possono essere confrontate con **operatori logici**, confrontando cella per cella e pixel per pixel, per generare oggetti più dettagliati ed analizzarli.

Modelli di ripartizione territoriale

Modelli di analisi spaziale che consentono la creazione di **unità territoriali** mediante:

- **Trasformazioni** di entità esistenti.
- **Ripartizione** dello spazio secondo diversi modelli.

I **criteri** adottati per **riorganizzare lo spazio** possono essere di tipo: Sociologico, urbanistico, fisico, etc...

I modelli di ripartizione territoriale sono:

- **Riclassificazione**: Riorganizzazione di oggetti nello spazio, mediante **cancellazione di confini** che delimitano le aree originali, per realizzare nuove figure geometriche prive di divisioni interne.
- **Buffer**: Aree che si estendono **intorno** ad un oggetto, usato per definire oggetti adiacenti a strade e fiumi.
- **Vicinato di Voronoi**: Particolare tipo di **decomposizione di uno spazio metrico** determinata dalle distanze rispetto ad un determinato insieme discreto di elementi dello spazio. Si può sfruttare per scoprire il punto di S più vicino ad un punto dato x senza calcolare ad ogni richiesta la distanza di x da ogni elemento di S. Utilizzato per trovare la stazione di benzina più vicina a noi o l'ospedale meno distante.

- **Gravitazione:** Ripartizione territoriale basata sull'analogia con il modello di **attrazione gravitazionale** proposto da Newton, secondo il quale, due corpi si attraggono in base alla loro massa e dalla distanza che li separa.
 - Oggetti **vicini** esercitano un'attrazione maggiore (Il supermercato più vicino e più appetibile di uno molto distante).
 - Oggetti più **grandi** esercitano un'attrazione maggiore (Il ristorante più importante della città può interessare di più di quello più vicino).

Questo modello di gravitazione genera una **ripartizione completa del territorio in aree di attrattività.**

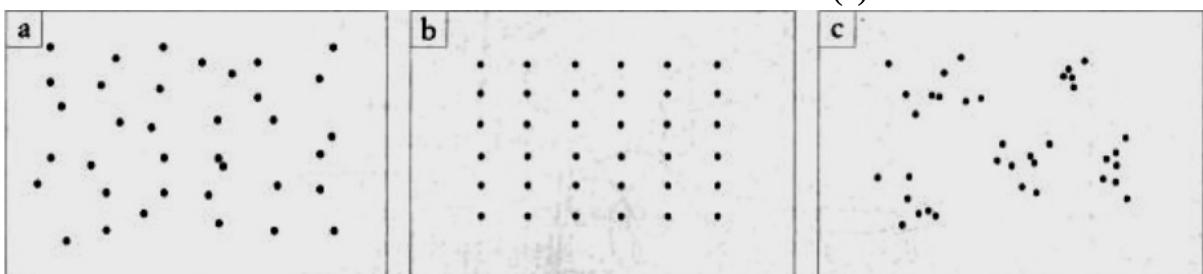
Indici statistici geospatiali

Strumenti statistici per l'analisi spaziale del territorio, che definiscono la disposizione spaziale di un fenomeno. Tali indici danno informazioni riguardo:

- Il **Baricentro** del fenomeno
- La **Dispersione** territoriale
- La **Direzione** prevalente
- La **Regolarità** della distribuzione

Con il termine **disposizione spaziale** o **pattern** si intende una determinata configurazione di oggetti nello spazio, che può essere:

- **Casuale:** Privo di regolarità (a).
- **Uniforme:** Distanza uniforme tra ciascun fenomeno (b).
- **Cluster:** Fenomeni addensati in zone e staccati da zone vuote (c).



Gli indici statistici geospatiali sono usati anche per prevenire o gestire certi fenomeni come:
Concerti, assembramenti, **pandemie**...

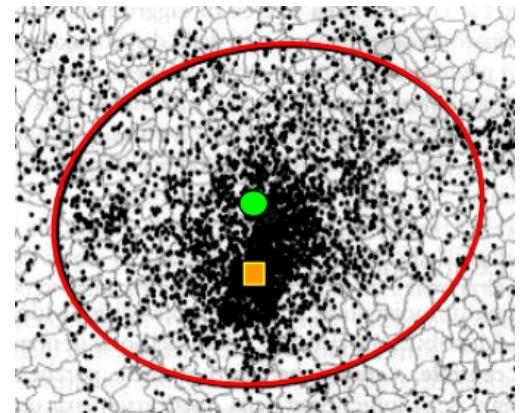
Statistiche centrografiche

L'obiettivo primario è l'individuazione del centro geografico della distribuzione dei dati:

- Centro **medio**: **Media delle coordinate** di latitudine e longitudine delle singole osservazioni (centro di gravità).
- Centro della **distanza massima**: Il punto che **minimizza le distanze** tra tutti i punti della distribuzione (punto di viaggio minimo).
- **Deviazione standard della distanza**: Misura la **dispersione dei fenomeni** sul territorio (distanza rispetto al centro medio).
 - **Cerchio** della deviazione: L'area è evidenziata da un cerchio centrato sul Centro medio il cui raggio è multiplo della deviazione standard
 - **Ellisse** delle deviazioni: Si calcolano le deviazioni lungo l'asse della latitudine e longitudine, definendo gli assi di un'ellisse.

Queste statistiche sono di **rilevanza cruciale** per trovare la collocazione più adatta per le aree di interesse.

- Punto **Verde**: Università
- Quadrato **Arancione**: centro medio
- Ellisse **Rossa**: Area di maggiore concentrazione
- Punti neri: Residenza studenti



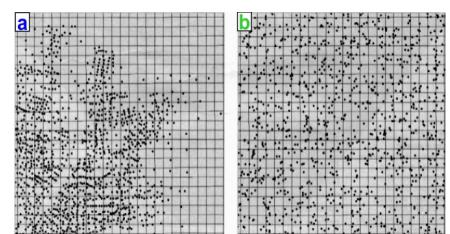
Analisi dei quadranti

Strumento per l'analisi del **pattern**, ovvero la **distribuzione di una variabile sul territorio**. Questa tecnica si basa sul conteggio delle osservazioni su di una **griglia regolare a maglie quadrate** che suddivide l'immagine.

- Per ciascun quadrato viene valutata la frequenza delle osservazioni, ovvero **l'occorrenza**.
- Si calcola la varianza delle frequenze, ovvero la dispersione del fenomeno.
- Dal confronto tra la **distribuzione osservata** (a) e una **distribuzione casuale** (b) si deriva una misura della presenza di **raggruppamenti** della variabile sul territorio.

Punti deboli:

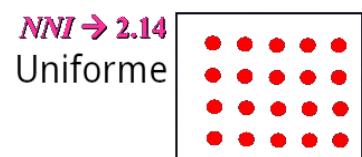
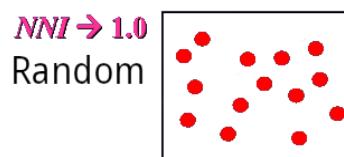
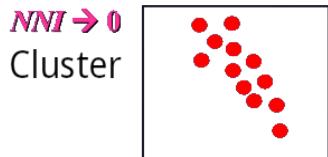
- La **dimensione** della griglia influenza notevolmente il risultato e non esiste una regola per individuare una dimensione ottimale.
- Non tiene conto delle variazioni locali tra celle confinanti.



L'analisi del vicinato - Nearest Neighbour Analysis

L'indice NNI misura il grado di dispersione spaziale di una distribuzione di punti, questo calcolo si basa sulla **misurazione delle distanze di punti adiacenti**.

La **distanza** di punti raggruppati in un **Cluster** è **inferiore** alla distanza in cui i punti sono distribuiti in modo casuale o uniforme.



Interpolazione

L'**interpolazione spaziale** permette di **stimare** il valore di una determinata variabile presente in una **zona non coperta**, basandosi sulla conoscenza delle variabili in punti noti.

Molti GIS offrono alcuni sistemi di interpolazione, come:

- Poligoni di Thiessen o Voronoi
- [TINs Triangulated Irregular Networks](#)

L'interpolazione spaziale si basa su **due assunzioni fondamentali**:

- Il fenomeno d'interesse si assume abbia natura continua e sia quindi **misurabile** con una variabile metrica.
- Il fenomeno osservato sia **spazialmente dipendente**.

VII GPS – Global Positioning System

Il GPS inizialmente era ad uso esclusivo dell'esercito statunitense, realizzato dal **Dipartimento della Difesa statunitense (NAVSTAR)**, è diventato pienamente operativo nel 1994.

Il segnale civile era intenzionalmente degradato attraverso la **Selective Availability (SA)** che introduceva **errori intenzionali nei segnali satellitari** allo scopo di ridurre l'accuratezza della rilevazione, consentendo precisioni dell'ordine di 900–950 m. Questa degradazione del segnale è stata **disabilitata nel mese di maggio 2000** grazie a un decreto del presidente degli Stati Uniti **Bill Clinton**, mettendo così a disposizione degli usi civili la precisione attuale di circa 10–20 m, anche se **tra i due sistemi**, civile e militare, permangono delle differenze, per impedirne il montaggio su missili. Nei modelli per uso civile sono presenti alcune limitazioni: massimo 18 km per l'altitudine e 515 m/s per la velocità. Questi limiti possono essere superati, ma non contemporaneamente.

Il sistema GPS non è più formato esclusivamente da una rete satellitare gestita dagli americani, negli anni diverse nazioni hanno messo a disposizione propri satelliti ad uso civile e militare per il rilevamento GPS:

- USA: Navstar
- Russia: Glonass
- Europa: Gallileo (L'unico non nato in abito militare, è il più preciso)
- Cina: Beidou
- India: Imss

I dispositivi moderni non adottano solo uno di questi sistemi per il rilevamento, bensì attraverso tecniche di approssimazione è possibile migliorare l'accuratezza sfruttando più sistemi contemporaneamente.

Navstar

Navigation System with Timing and Ranging Global Positioning System fu concepito dal Dipartimento della Difesa USA per:

- Determinare con precisione il punto in cui un ricevitore si trova sulla terra.
- Ottenere un'indicazione oraria molto precisa.

La costellazione di **24 satelliti** che costituisce il sistema GPS Navstar fu completata nel 1993. Di 24 satelliti, **3 sono di scorta** in caso di guasti.

I satelliti sono posti in un orbita circolare a circa 20.200Km dalla terra e compiono una rivoluzione in 12 ore. Il centro di controllo del sistema GPS si trova nei pressi di *Colorado Springs* ed ha il compito di eseguire tutte le misure necessarie per **correggere** le informazioni inviate dai satelliti GPS.

Sistema GPS

Il sistema GPS in generale, è composto da 3 parti:

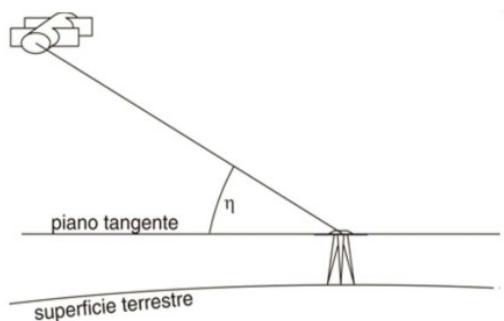
- Il segmento spaziale (Satelliti)
- Il segmento di controllo (Stazioni di terra che effettuano la correzione)
- Il segmento d'utilizzo (I ricevitori GPS)

Segmento spaziale

Ogni satellite è dotato di **Orologi Atomici** con un'elevata precisione con una possibilità di errore di un secondo ogni 30.000 anni. Gli orologi terrestri delle stazioni di riferimento, sono basati sull'idrogeno ed hanno una precisione assoluta ancora più elevata, sono usati per correggere gli orologi in orbita.

Quando si usa il sistema GPS non vengono usati tutti i satelliti in orbita, alcuni satelliti possono essere scartati a causa dell'**errore atmosferico**. Se il satellite è troppo in basso rispetto al punto da calcolare, le onde radio non riusciranno ad attraversare lo strato atmosferico.

L'angolo tra satellite ed il punto da rilevare deve sempre essere maggiore di 10-15 gradi, altrimenti il satellite sarà scartato.

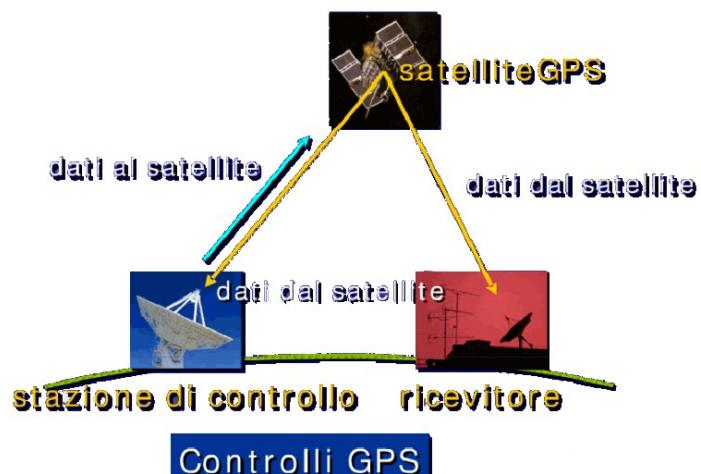


Segmento di controllo

Le stazioni di terra seguono in maniera continua i satelliti ed elaborano i dati per calcolarne la posizione spazio-temporale (**effemeridi**). Gestiscono il segmento spaziale, correggendo eventuali errori degli **Orologi Atomici**, effettuando **correzioni orbitali** e rilevando eventuali problemi ai satelliti.

Segmento utilizzo

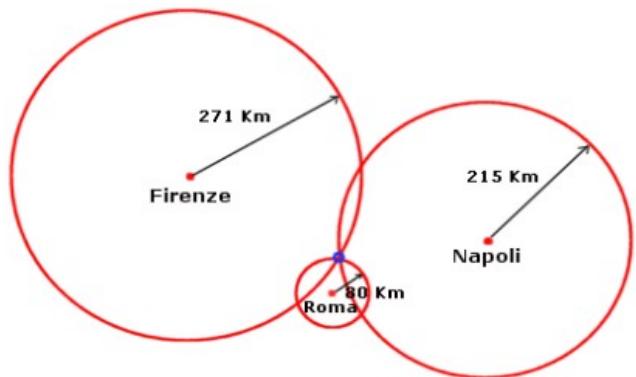
Definito da ogni ricevitore, da aeri ed automobili fino ai cellulari.



Trilaterazione

La trilaterazione è il metodo usato per il calcolo della posizione, derivando la distanza di un punto conoscendo la **distanza di quel punto da tre punti** di coordinate note.

Il modello reale lavora sullo spazio tridimensionale, il concetto resta lo stesso, semplicemente si utilizza un anello in più sul piano tridimensionale.



Triangolazione

La triangolazione in topografia è quel procedimento che permette di **determinare indirettamente distanze tra punti del terreno** e quindi le loro coordinate geografiche.

Tale procedura richiede tre punti, considerati vertici del triangolo, uno di questi lati viene misurato direttamente ed è detto base geodetica misurata. Con la misura degli angoli è possibile calcolare i lati del triangolo.

Funzionamento del GPS

Per misurare distanza tra Receiver ed il satellite viene misurato il **tempo che un segnale impiega ad arrivare a terra**. Il calcolo matematico necessario è semplice ma il livello di **precisione** dipende fortemente dall'accuratezza **dell'orologio** del Receiver (Pochi millisecondi possono portare a Km di errore nella rilevazione).

Nei Receiver si usano **orologi al quarzo** capaci di mantenere una elevata precisione per brevi periodi e quindi sono spesso corretti direttamente dai segnali satellitari.

Quando il satellite invia un segnale invia anche l'orario a cui l'ha inviato, il Receiver calcola:

$$\Delta T = \text{orario di arrivo} - \text{orario di invio}$$

$$\text{Distanza} = \text{velocità (velocità della luce)} * \text{tempo} (\Delta T)$$

Ma il ΔT calcolato è affetto da errore dato che **l'orologio del Receiver è impreciso**.

Il messaggio inviato da un satellite è così strutturato:

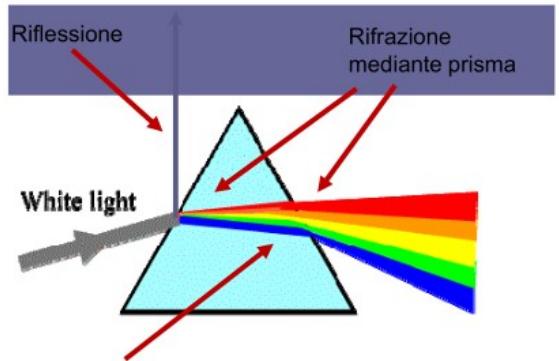
- **Formato:** Il messaggio viene **trasmesso con continuità e su 2 frequenze** contemporaneamente (1,2 e 1,5 GHz) così da eliminare l'errore dovuto alla rifrazione atmosferica.
- **Contenuto:** il messaggio contiene
 - **Almanacco** ovvero i parametri orbitali approssimati dell'intera costellazione
 - **Le effemeridi** del satellite
 - Dati relativi al satellite come **orario e stato** del sistema

Rifrazione e riflessione

La **rifrazione** avviene ogni qualvolta la luce attraversa uno spazio con differente densità, questo avvenimento fa **cambiare la sua velocità di propagazione**.

L'angolo di rifrazione dipende da:

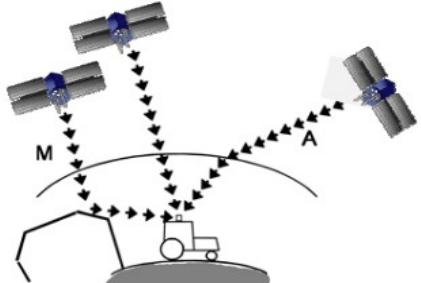
- **Frequenza della luce:** maggiore è la frequenza maggiore è la rifrazione.
- **Densità dello spazio attraversato:** maggiore è la densità maggiore è la rifrazione.



Nota: il blu subisce una "piegatura" maggiore del rosso

Questo fenomeno può essere osservato anche con delle cannucce colorate immerse nell'acqua, colori diversi avranno piegature diverse.

Anche le comunicazioni radio, essendo frequenze, hanno lo stesso problema.



L'atmosfera terrestre non ha tutta la stessa densità, questo porta ad avere **variazioni del segnale radio dovute dalla rifrazione**. Per questo motivo il messaggio viene **trasmesso con continuità su 2 frequenze** contemporaneamente (1,2 e 1,5 GHz). Le due frequenze di trasmissione avranno rifrazioni differenti, fornendo quindi, informazioni in più per la **correzione della rifrazione**.

Il problema della rifrazione era anche presente nelle trasmissioni TV analogiche in presenza di superfici riflettenti.

Effemeridi

Dalla parola greca "ephemeros" = Giornaliero, sono **tabelle** che contengono un insieme di parametri sintetici, **utili per calcolare la posizione del satellite**.

Occorre distinguere le:

- Effemeridi **trasmesse**: Trasmesse dal satellite.
- Effemeridi **precise**: Calcolate a posteriori da diverse organizzazioni governative e di ricerca, e messe a disposizione via web.
 - **Rapide**: Disponibili con 1 giorno di ritardo.
 - **Finali**: Disponibili con 14 giorni di ritardo.

Fonti di errore nella determinazione della posizione

Per ottenere la migliore rilevazione possibile, tutti i satelliti in orbita devono avere **l'orario sincronizzato** tra loro e con tutti gli altri sulla terra.

Tuttavia esistono possibili fonti di errore di diversa natura che possono causare ritardo:

- Fattori **Atmosferici**: Possono causare rallentamenti del segnale al passaggio nella ionosfera e nella troposfera.
- Fattori **Elettronici**: Tempo impiegato nel passaggio nella strumentazione e tempo di elaborazione del segnale.
- Fattori **Relativistici**: Anticipo e ritardo degli orologi atomici.
 - Se paragonato ai due fattori precedenti, i fattori Relativistici sono di quasi **tre ordini di grandezza più elevati**.
 - Velocità e Gravità possono influire sul tempo, secondo la [teoria della Relatività \(Interstellar\)](#):
 - La differenza di gravità tra orbita e suolo terrestre può influenzare sullo scorrere del tempo (I satelliti distano ~20.000Km dalla terra).
 - Più è elevata l'attrazione gravitazionale più lentamente scorrerà il tempo.
 - La velocità quanto più è elevata più il tempo rallenta, quindi la velocità orbitale del satellite (3,87Km/s) può anch'essa influire sul tempo.
 - Quanto più si è prossimi alla velocità della luce tanto più il tempo tende a Zero.

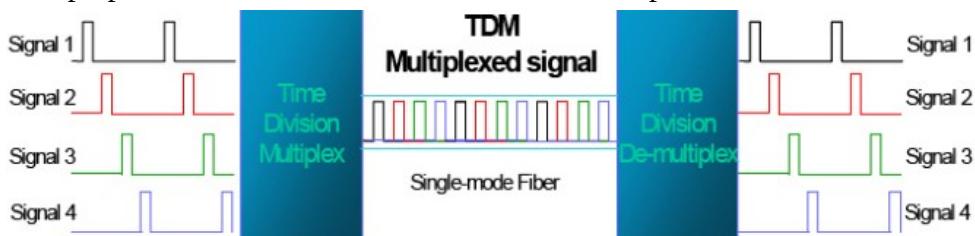
Se non venissero valutati gli errori relativistici, si avrebbe un **errore di precisione di ~11,4Km**

Anche l'attrazione del Sole e della luna influenzano il satellite, così come la pressione della radiazione solare sui pannelli del satellite. Questi errori sono considerati ma sono nettamente inferiori a quelli descritti in precedenza.

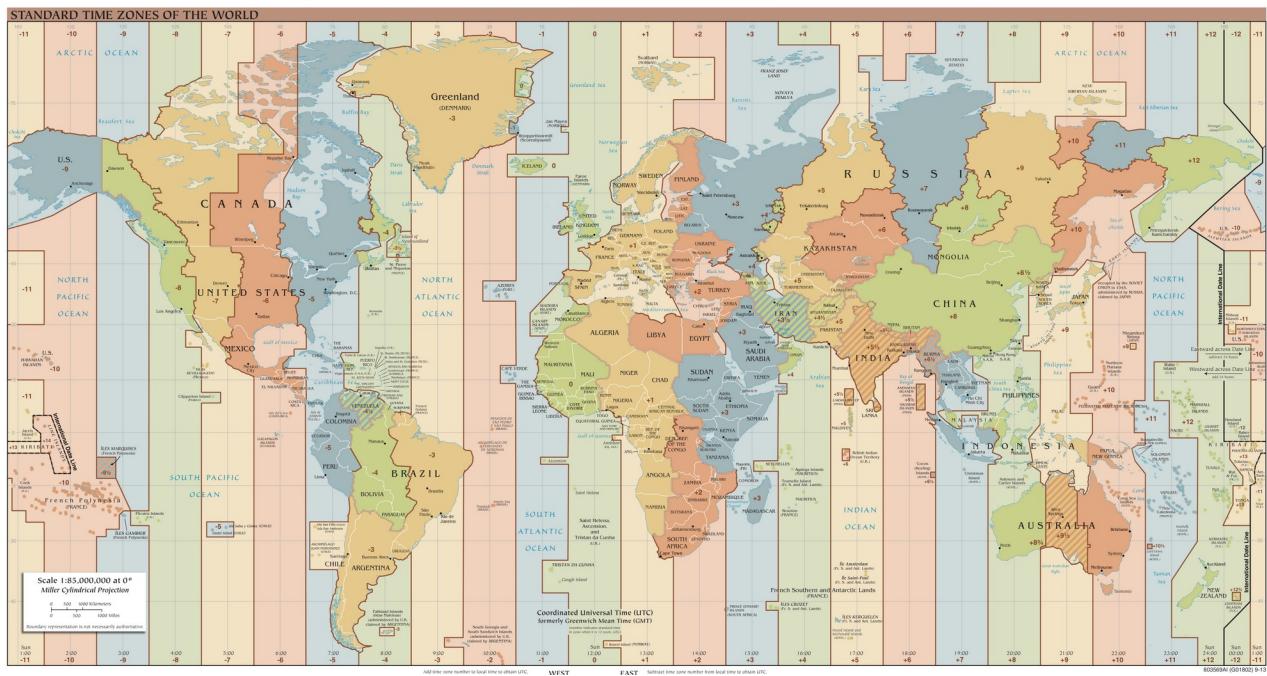
La rilevazione e la correzione di questi errori è indispensabile per il corretto funzionamento della navigazione satellitare.

Time Division Multiplexing

Dato che più satelliti **comunicano contemporaneamente** con più dispositivi, c'è bisogno di dividere il canale di comunicazione trasmettendo a turno. Ad ogni dispositivo è assegnato un intervallo di tempo per la comunicazione. Così come avviene per le infrastrutture di rete.



Time Zones



Partendo dal meridiano di Greenwich, a sinistra si sottrae 1 ora per fascia, a destra invece si aggiunge.

All'estremo destro di questa cartina, la linea nera marcata è la linea di cambiamento di data, dove il fuso orario cambia da +12 ore a -12 ore.

Le isole di **Tonga** e **Samoa** sono vicine ma si trovano ad 1 giorno di distanza.

UTC (Coordinated Universal Time)

Il nome UTC è stato coniato per non dover menzionare una specifica località in uno standard internazionale. L'UTC **si basa su misurazioni condotte da Orologi Atomici** invece che su fenomeni celesti come nel caso del GMT.

L'UTC nei sistemi informatici memorizza la data come **numero di secondi trascorsi dal 1 Gen 1970**, come numero intero con segno a 32 bit. La massima data esprimibile è fino al 19 Gennaio 2038, oltre questa data il bit del segno diventerà negativo segnando come data 13 dicembre 1901.

Un Tool online per calcolare le epoche: [Epoch Converted](#).

Per risolvere il problema del massimo rappresentabile si possono adottare due soluzioni:

- **Cambiare l'intero da con Segno a senza segno**, così da rimandare il problema al 7 Feb 2106.
- **Usare un valore con segno a 64 bit**, rimandando il problema a tra 290 miliardi di anni, oltre alla fine stimata del sistema solare.

Standard Spazio/Tempo

Ricapitolando come standard dello spazio usiamo il [WGS84](#) e per il tempo [l'UTC](#).

Sincronizzazione Orologi

L'efficacia e la precisione dell'intero sistema [GPS dipende fortemente dalla precisione degli orologi.](#)

Gli orologi atomici sono basati su **Cesio** e **Rubidio** e sono eccessivamente costosi, i tradizionali orologi dei ricevitori sono basati sul **Quarzo**, sono economici ma richiedono una continua ricalibrazione.

Un ricevitore, esaminando l'orario di 4 o più satelliti, può misurare la propria imprecisione.

Solo e soltanto un orario T può soddisfare le equazioni geometriche per poter individuare un unico punto nello spazio: le **4 sfere** di distanza individuate dai **4 satelliti** e dal ricevitore **si intersecano in un solo punto** **soltanto se le misurazioni ottenute sono precise**. Il ricevitore R quindi calcola ed imposta il proprio orario a T semplicemente **imponendo che l'intersezione delle 4 sfere coincida in un unico punto.**

A-GPS – Assisted GPS

L'A-GPS è un sistema che **consente di abbattere i tempi** necessari alla **prima localizzazione** durante l'uso di un terminale GPS. Molto utile ed utilizzato soprattutto nei punti in cui è difficile stabilire con precisione la lista dei satelliti in vista di un ricevitore.

Il Sistema A-GPS è stato sviluppato per **abbattere i tempi ed il costo del Fixing** ([download di Almanacco ed Effemeridi](#)) necessario nei normali sistemi GPS per la prima localizzazione. Ciò è possibile chiedendo le informazioni necessarie per il Fixing all'operatore di telefonia mobile.

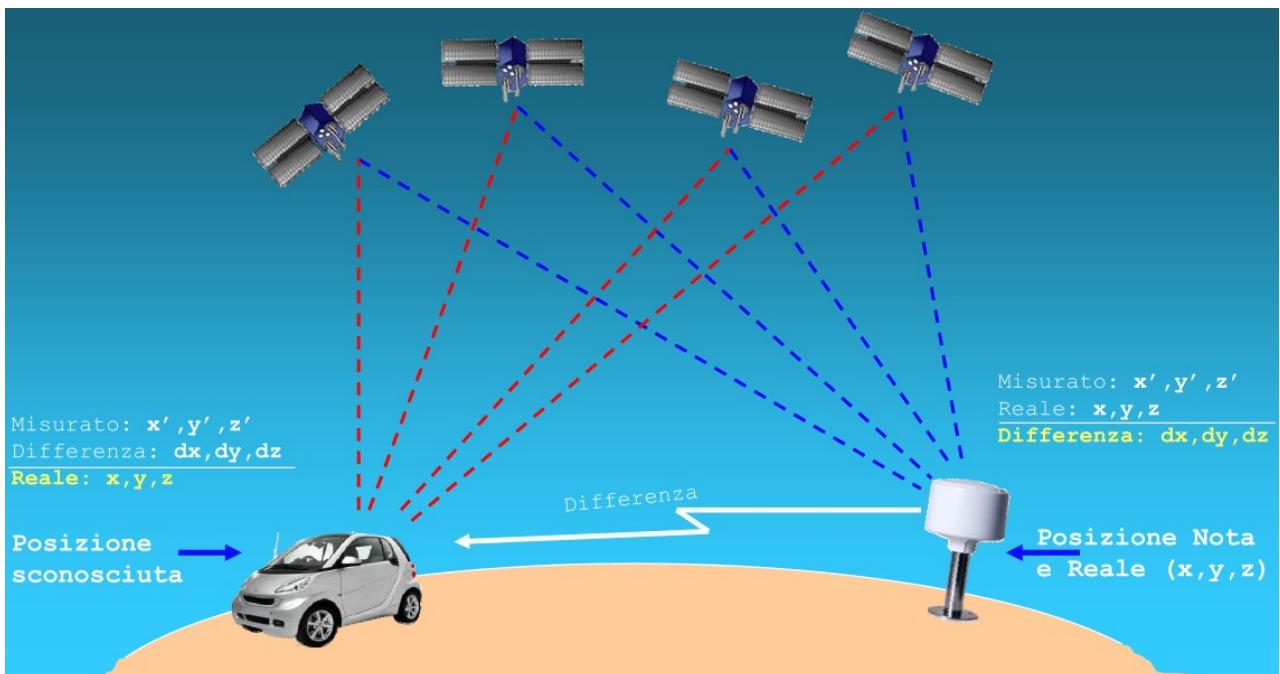
Quando un ricevitore A-GPS vuole conoscere la sua posizione:

- Si collega tramite la rete cellulare ad un **Assistance Server**, al quale viene inviata anche l'informazione sulla cella cellulare a cui l'utente è agganciato.
- Il Server invia al dispositivo le informazioni riguardanti i **satelliti in vista dell'antenna cellulare** a cui l'utente è connesso.

D-GPS – Differential GPS

Il GPS Differenziale cerca di ridurre l'errore atmosferico. Una **stazione meteorologica** o stazione di riferimento, di cui si conosce il punto esatto sulla terra mediante **telemetria terrestre**, può richiedere le proprie coordinate tramite GPS per conoscere la **differenza** tra la sua posizione accurata e quella fornita dal satellite.

Questa **differenza** può essere usata per **correggere le rilevazioni errate dei terminali nel suo intorno**.



DOP – Dilution of precision

Per una maggiore precisione nella rilevazione tramite **trilaterazione**, si scelgono i satelliti che formano l'angolo maggiore sul punto da rilevare ([Parallax Method](#)). Il **Point DOP** misura proprio il grado di **diluizione della precisione**.

- 1: **Ideale**
- 1-2: **Eccellente**
- 2-5: **Buono**
- 5-10: **Medio**
- 10-20: **Basso**
- >20: **Scarso**

