

Domanda 1-

Cosa sono le Armoniche ???

I suoni armonici sono suoni piu' acuti e meno intensi che accompagnano il suono fondamentale prodotto da un corpo vibrante,
e sono fondamentali nella determinazione del timbro e degli intervalli musicali

Domanda 2-

Feature Vector ???

E' un insieme di vettori di caratteristiche prodotto dalla fase di indicizzazione, ciascun vettore a seconda della complessita' dell'oggetto contiene un numero grande di componenti(testo , audio). La fase di retrieving e' pertanto caratterizzata da un gran numero di confronto di caratteristiche tra le query Q e le caratteristiche del vettore

Domanda 3

Cluster ???

Insieme di oggetti con caratteristiche omogenee. Viene costruito attraverso un'analisi statistica che consente di accostare oggetti con caratteristiche omogenee rispetto a variabili predefinite. Le ricerche della query risulteranno piu' efficienti, la query e' confrontata con i centroidi e seleziona il cluster piu' simile

Domanda 4

Differenze tra dominio Temporale e delle Frequenze ???

Dominio temporale descrive il segnale rispetto al tempo, nel dominio della frequenza e' descritto attraverso le sue frequenze , composti da molti segnali sinusoidali con frequenze differenti. Si puo' convertire il dominio temporale nel dominio delle frequenze applicando la trasformata di Fourier.

Domanda 5

Distanza Euclidea ???

E' la distanza tra due punti ossia la misura del segmento avente come estremi i due punti. Istogramma cumulativo...

Domanda 6

Prodotto Scalare ???

Il prodotto scalare si utilizza nel retrieval con modello Spazio Vettoriale, dato un documento con componenti $D=[d_1, d_2, \dots]$ e un insieme di termini $Q=[q_1, q_2]$. Si usa il prodotto scalare che ha come input 2 vettori e output uno scalare e si calcola $\sum Q_j * D_j$
Il grado di similarita' viene quantificato utilizzando il coseno dell'angolo tra i 2 vettori.

Domanda 7

DTW ???

Ogni pezzo del parlato e' rappresentato da un vettore P di caratteristiche, il quale per essere riconosciuto e' confrontato con altri vettori memorizzati nella fase di training ----> Al sistema vengono fornite le componenti di base del parlato fonemi, parole e frasi. Il sistema le rappresenta tramite vettori delle caratteristiche. Si suddivide il parlato in singole unita' che verranno confrontate con i vettori delle caratteristiche raccolti nella fase di training. Training basato su HMM HIDDEN MARKOV MODELS , viene costruita un hmm che per ogni fonema considera variabilita' degli speaker rumore di fondo e le differenze temporali.

Domanda 8

Reti Neurali Artificiali ANN ??

Simula i processi cognitivi del cervello impiegando un gran numero di elementi di calcolo. ognuno di essi rappresenta un neurone connesso all'esterno tramite connessioni alle quali arrivano gli input dati. Ogni connessione è caratterizzata da un peso, le informazioni che il neurone porta ad un neurone destinatario mediante la connessione. L'insieme dei pesi rappresenta la conoscenza totale della rete

Domanda 9

Istogramma del Colore ???

È un vettore $H(M) = (h_1, \dots, h_n, \dots, h_m)$ ogni h_j rappresenta il numero di pixel dell'immagine M che ricadono nel bin j
(bin=numero totale di combinazione di colore).

Domanda 10

Key-Frame fotogramma iniziale che permette di descrivere i cambiamenti che occorrono tra un fotogramma ed il successivo descritto con codifica Interframe

R-Frame frame rappresentativi di uno shot possono essere -1 il primo frame dello shot

-2 Si calcola il frame medio facendo la media dei colori di tutti i frame

-3 Si calcola la media di tutti gli istogrammi dei frame e si sceglie come rappresentativo il cui istogramma è più vicino alla media

Domanda 11

Tutte le strutture dati

Domanda 12

GPS Quanti ce ne sono ???

Esistono 5 sistemi navstar (usa), glonass(RUSS), galileo (EU), beidou(CINA), irnss (INDIA)

Domanda 13

Quali segmenti è composto ???

Segmento Spaziale ---> Prevede 24 satelliti, disposti su 6 piani orbitali ugualmente spazati tra di loro, inclinati di

55° rispetto all'equatore. Sono posti in orbita circolare 20200km dalla terra e compiono una rivoluzione in 12 ore.

È dotato di orologio atomico basati su 4 oscillatori ad altissima precisione di cui 2 al cesio e 2 al rubidio.

Segmento di controllo ---> Gruppo di stazioni a terra con il compito di Seguire in maniera continua i satelliti

imporre correzioni d'orbita, sintonizzare gli orologi atomici a bordo dei satelliti, scaricare i dati per la trasmissione

tramite i satelliti

Segmento d'utilizzo sono i dispositivi che usano il gps

Domanda 14

Differenze Efemeridi/Almanacco

Almanacco -> Parametri orbitali approssimati dell'intera costellazione

Effemeridi -> Tabelle che contengono un insieme di parametri sintetici necessari e sufficienti per calcolare

la posizione del satellite, ne esistono due tipi: effemeridi trasmesse e quelle precise, le prime calcolate dal

satellite, le seconde calcolate da organizzazioni governative, e variano tra rapide (ritardo 1 giorno) e finali (ritardo di 14 giorni).

Domanda 15

Trilaterazione vs Triangolazione ???

La triangolazione si effettua scegliendo 3 punti opportuni sul terreno, considerati vertici di un triangolo uno di questi lati

viene misurato direttamente (la base geodetica). Da ognuno degli estremi della base si misurano gli angoli sotto i quali viene visto l'altro punto e con calcoli trigonometrici se ne determina la distanza.

La trilaterazione metodo usato per il calcolo della posizione, Un primo satellite individua la nostra posizione e sulla

superficie di una sfera, l'intersezione di due sfere è una circonferenza. Una seconda misura indica la nostra posizione sull'intersezione delle due sfere

Una terza misura individua 2 punti ottenuti dall'intersezione di due sfere.

Domanda 16

Perché si usa 4 sfere ???

La quarta misura identifica un solo punto (Se non dovesse esserci la quarta sfera, la stessa Terra diventa la quarta misura)

Domanda 17

Differenza tra riflessione e rifrazione ???

Quando delle onde di tipo qualsiasi incidono su un ostacolo piano, rigenerano onde che si allontanano dall'ostacolo. Questo fenomeno è chiamato riflessione.

La rifrazione è la deviazione che un raggio luminoso subisce nel passare da un mezzo trasparente a un altro.

Nel GPS i fenomeni di riflessione e rifrazione si verificano

vicino a specchi d'acqua, o in presenza di ostacoli, quindi il segnale rettilineo è solo apparente, e la posizione inviata dal satellite non è del tutto precisa.

Domanda 18

AGPS E DGPS ???

AGPS --> Uno dei principali problemi è la prima localizzazione in quanto

un terminale deve ricavare la lista dei satelliti a cui agganciarsi.

Il sistema a-gps permette di abbattere tali costi anche in previsione di un uso effettuato su terminali con basse capacità

di elaborazione o risorse energetiche limitate.

Ogni cella di telefonia presente sul territorio ha una posizione fissa e ricava i satelliti ad essa in vista, quando il terminale

a-gps vuole conoscere la posizione si collega ad un assistance server al quale viene inviata

l'informazione sulla cella. Il terminale vede i medesimi satelliti della cella pertanto il server elabora una lista

e le invia tramite la rete cellulare al terminale che può ricavare subito la sua posizione.

DGPS-> Poiché il gps è soggetto ad errori nel calcolo della posizione, il GPS differenziale è un sistema che permette di migliorare il posizionamento basato sul Global Positioning System.

Questa tecnica si basa sulla combinazione dei dati osservati contemporaneamente da due (o più) ricevitori GPS, ottenendo l'eliminazione o la significativa riduzione dei principali errori nel posizionamento GPS.

Domanda 19

Parallasse -> Sulla terra la base da usare per le triangolazioni non può essere più grande del diametro terrestre, ma non è possibile determinare la distanza delle stelle. Si usa una base molto più grande pari al diametro dell'orbita della Terra attorno al Sole. Si definisce parallasse di una stella l'angolo sotto cui un osservatore che si trovasse sulla stella vedrebbe il raggio dell'orbita terrestre

SECONDA PARTE

Domanda Esame 1

:"Se l'istogramma fosse una linea piatta (i colori sono sparpagliati in maniera uniforme) quale sarebbe il corrispondente cumulativo?"
Ogni volta che mi sposto incremento sempre di parti uguali, mi porta ad una linea retta, non posso affermare che corrisponde ad un'immagine che abbia un'uniformità di colori, dunque gli intervalli contengono diverse sfumature. Quanto più approssimo, più sono larghi gli intervalli. In oltre da una parte mi facilita il compito, dall'altra mi può dare più errori

Domanda Esame 2

Orbita Geostazionaria ???

Un'orbita di un satellite viene definita 'geostazionaria' se per un osservatore a terra il satellite appare fermo in cielo, sospeso sempre al di sopra del medesimo punto dell'equatore poiché ruota alla stessa velocità angolare della Terra.

Per pianeti diversi dalla Terra quest'orbita è anche detta isosincrona.

Non per tutti i pianeti esiste un'orbita stazionaria in quanto la loro velocità di rotazione può richiedere che il satellite stia in un'orbita troppo vicina oppure troppo lontana per essere stabile

Domanda Esame 3

Cometa ???

La corona, la parte più esterna dell'atmosfera solare, si espande e investe con una corrente continua lo spazio interplanetario. Questa corrente è il vento solare, composto da protoni ed elettroni con tracce di nuclei di elio. Il vento, che provoca alterazioni nei campi magnetici della Terra e degli altri pianeti, "spinge" in direzione opposta al Sole stesso la chioma delle comete.

Domanda Esame 4

Quando il gps ha difficoltà di ricezione ???
All'alba e al tramonto

Domanda Esame 5

Segmentazione a due soglie CI STA su marco

Domanda Esame 6

Formato Matroska -> scopo di contenere oggetti e sincronizzarli e un contenitore tipo Avi non un formato, al suo interno ci sono più formati sia audio che video

DOMANDE SIM

INDICIZZAZIONE

DOMANDA: Cos'è il Feature Vector

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

L'indicizzazione di dati multimediali implica la considerazione di:

- spazi multidimensionali di caratteristiche
- definizione di una metrica in tale spazio.

Questi spazi multidimensionali di caratteristiche prendono il nome di feature vectors ovvero vettori di caratteristiche, e per estrarre queste caratteristiche ciascun oggetto multimediale gestito dal database deve essere preprocessato.

I requisiti per l'estrazione delle feature, che poi andranno a comporre il vettore, sono:

- le feature estratte devono essere complete e rappresentare il contenuto e l'informazione presente nel dato
- le feature devono essere memorizzate in maniera compatta(altrimenti sarebbe più veloce ricercare e confrontare i dati originali)
- il calcolo della distanza tra le feature deve essere veloce in modo che siano bassi i tempi di risposta del sistema

TESTO

DOMANDA: Precisione e Recall riferite a documenti di testo

RISPOSTA: 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 maggiore 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.

Al numeratore c'è l'intersezione ovvero gli elementi recuperati dal sistema mentre nel denominatore per la precisione ci stanno gli elementi pertinenti alla richiesta e per la recall invece ci stanno i documenti che sono corrispondenti a quelli richiesti.

DOMANDA: Perché è più importante la precisione dei media invece della quantità in un motore di ricerca

RISPOSTA: In un motore di ricerca è più importante la precisione dei media perché nel caso di una query su un argomento generico è meglio prendere meno documenti che qualitativamente sono migliori che prenderne un sacco ma che qualitativamente non sono poi così buoni questo anche perché i documenti che poi uno va ad analizzare non saranno mai tutti quanti siccome un motore di ricerca recupera parecchi documenti su un determinato argomento.

29 DOMANDA: Prodotto Scalare

RISPOSTA: Per effettuare il retrieval ovvero un recupero si possono utilizzare varie tecniche ed una di queste è la tecnica dello spazio vettoriale che utilizza il prodotto scalare tra i vettori.

In questo caso i vettori rappresentano sia i feature vectors dei documenti del database sia le query ed è utile fare il prodotto scalare tra questi due perché il risultato di questo prodotto, che è possibile chiamare indice e si basa sull'angolo presente tra i due vettori, rappresenta appunto la similitudine che c'è tra la mia query e un documento del database (tanto più l'angolo presente tra i vettori converge a 0 gradi allora più i documenti saranno simili e viceversa quando l'angolo si avvicina a 90 gradi allora i documenti saranno diversi).

DOMANDA: Cluster

RISPOSTA: I cluster sono gruppi di elementi simili tra loro, per generare un cluster ci sono due principali metodi.

- **Similarità per coppie:** consiste nel prendere una coppia di documenti e costruire una matrice di distanze, la coppia più vicina(coppia con minor distanza possibile) la si elimina dalla matrice sostituendola con un oggetto 'congiunto' che prende il nome di centroide del cluster. Questo processo si ripete fino a che la matrice non diventa di un unico elemento (cluster) che racchiuderà di conseguenza tutti i documenti.
- **Clustering Euristico:** consiste nel considerare ogni documento in un certo cluster e quando arriva un documento dall'esterno, questo documento lo si aggiunge al cluster più vicino che di conseguenza aumenta. Se però il documento che arriva è troppo distante dal cluster allora si fonda un nuovo cluster e si ripete il processo fino ad esaurire i documenti.

Dopo aver generato i cluster sia le ricerche che il recupero dei documenti risulteranno efficienti ed efficaci dal momento che una query può racchiudere un gran numero di documenti del DB da confrontare e la logica dei cluster serve proprio a minimizzare il numero di confronti da effettuare per far sì che la risposta del motore sia più rapida possibile.

AUDIO

DOMANDA: Cos'è un armonica

RISPOSTA: Le armoniche sono una caratteristica tipica dei corpi che vibrano per la quale oltre una base frequenziale ho un insieme di frequenze che sono multiple di questa base(armonica di base).

Se ad esempio abbiamo una frequenza fondamentale del segnale di 440Hz allora le armoniche, essendo multipli del segnale, saranno tutti picchi di segnale con valori come 880Hz, 1320Hz, 1720Hz e 2200Hz.

DOMANDA: Algoritmo ADC (Analogic Digital Conversion)

RISPOSTA: 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.

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.

DOMANDA: File Audio (caratteristiche dominio temporale)

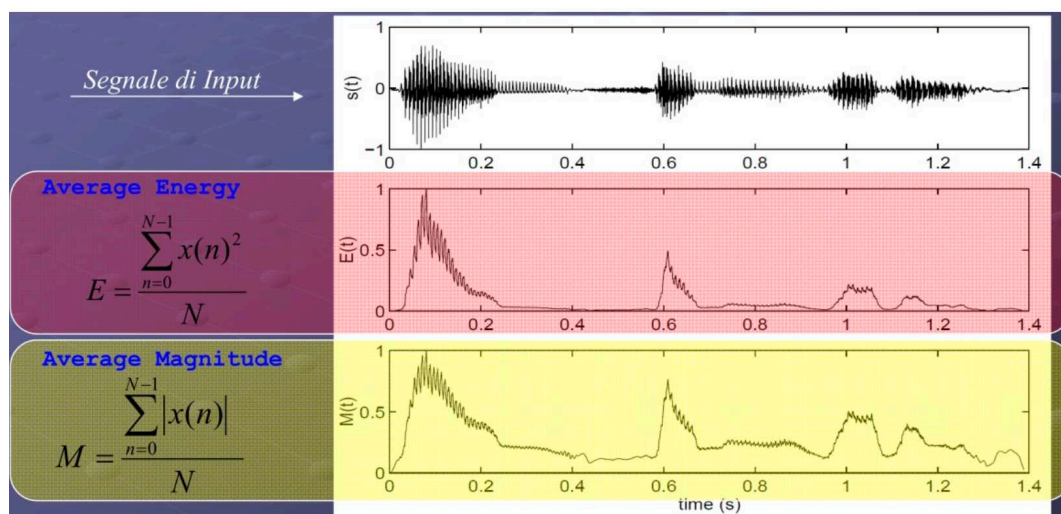
RISPOSTA: Nel dominio temporale è possibile analizzare l'energia media (**Average Energy**), ovvero la media della "rumorosità" del segnale audio per la sua durata.

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.

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

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

Silence Ratio: Indica la quantità di silenzio nel brano.

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

DOMANDA: File Audio (caratteristiche dominio delle frequenze)

RISPOSTA: Un segnale può essere scomposto nelle frequenze che lo compongono, tale scomposizione deriva dalla rappresentazione del Dominio 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**.

DOMANDA: Teorema di Nyquist

RISPOSTA: 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 **fino a 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 di Nyquist stabilisce il **passo di campionamento**, più è elevata la frequenza più campioni devono essere prelevati su quella fascia.

DOMANDA: Predictive Coding

RISPOSTA: 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.

DOMANDA: File Midi

RISPOSTA: 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.

DOMANDA: Reti Neurali Artificiali

RISPOSTA: Le **Reti Neurali Artificiali** (Artificial Neural Network) sono 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.

IMMAGINI

DOMANDA: Differenza lossy / lossless + esempi di entrambe.

RISPOSTA: Per risolvere i problemi con le dimensioni elevate di un qualsiasi dato sono state inventate dei processi di compressione che possono sia ridurre le dimensioni mediante la riduzione di bit necessari per codificare una singola informazione (compressione entropica) oppure possono ridurre il numero di informazioni da memorizzare (compressione semantica). La compressione senza perdita di informazione o lossless sfrutta le ridondanze nella codifica del dato attraverso diverse tecniche:

- Codifica differenziale in cui ogni dato è rappresentato come differenza rispetto al dato precedente.
- Codifica di Huffman (vedi PDF) che assegna un numero inferiore di bit alle sequenze più probabili attraverso un vettore di codifica.

Rappresentazione del RUN :

Sc	X	C
----	---	---

Sc = carattere speciale per l'identificazione della codifica
X = il carattere che viene ripetuto nel RUN
C = numero di ripetizioni del carattere.

- Codifica Run-Lenght Encoding (RLE) che riduce le ripetizioni di caratteri sostituendo un RUN, ovvero un insieme di caratteri ripetuti, con il carattere ed il suo numero di ripetizioni.
- Codifica Lempel-Ziv-Welch (LZW) che viene considerata come un'estensione della codifica RLE dal momento che è in grado di conteggiare ripetizioni di più gruppi di caratteri e frasi.

La compressione con perdita di informazione o lossy comporta riduzioni notevoli delle dimensioni ma la ricostruzione dell'informazione da quella complessa non è identica a quella iniziale, i metodi lossy infatti rimuovono parti che non possono essere percepite come avviene nel caso di video, immagini e suoni. Tra le tecniche di compressione lossy si ricordano:

- la compressione JPEG per le immagini che applica una trasformata nel dominio delle frequenze(Discrete Cosine Transform) che permette di sopprimere dettagli irrilevanti riducendo il numero di bit per la codifica.
- la compressione MPEG per i video che codifica parte dei frame come differenza rispetto ai valori previsti in base a una interpolazione.
- la compressione MP3 per l'audio che si basa sulle proprietà psico-acustiche dell'udito umano per sopprimere le informazioni inutili.

DOMANDA: Indicizzazione di immagini

RISPOSTA: 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.

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.

DOMANDA: Indicizzazione basata sull'istogramma di colore

RISPOSTA: 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 :

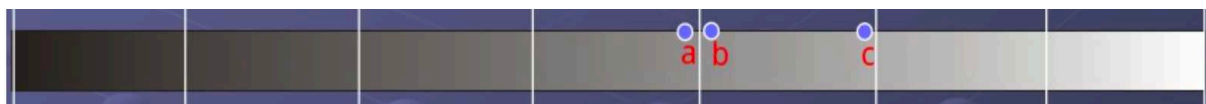
$$d(A, B) = \sum_{i=1}^n |a_i - b_i|$$

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”.

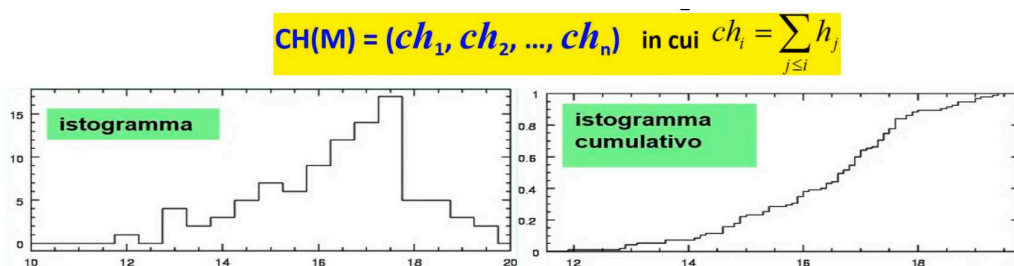
Soluzione 1 – Distanza tra i bins

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

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

Soluzione 2 – Istogramma 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**:



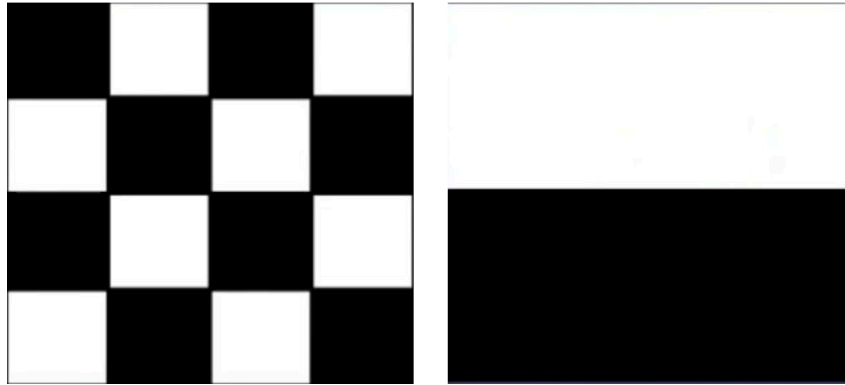
Soluzione 3 – PWH

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.

Limiti sistemi 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**.

DOMANDA: Cos'è l'eccentricità ?

RISPOSTA: I sistemi di ricerca devono essere capaci di riconoscere oggetti anche a seguito di modifiche sulla loro posizione o sulla loro scala, per fare ciò si costruisce un rettangolo intorno al soggetto identificato nell'immagine e in seguito si utilizza un parametro chiamato

eccentricità, questo consiste nel rapporto tra la lunghezza dell'asse maggiore e la lunghezza dell'asse minore. Dunque si facilita la ricerca di un'immagine dal momento che 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.

VIDEO

DOMANDA: Differenza tra j-frame e keyframe

RISPOSTA:

Codifica **Intra**frame (J-Frame)

- 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 **Inter**frame (Keyframe)

- 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**

22 DOMANDA: Segmentazione automatica in shots(come scegliere shot per indicizzazione video)

RISPOSTA: 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:

- 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.

- 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.

DOMANDA: Segmentazione a due soglie

RISPOSTA: 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.

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

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

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

DOMANDA: Come si sceglie il frame rappresentativo(r-frame) di un video?

RISPOSTA: 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:

- Si considera come r-Frame il **primo Frame** dello Shot.
- 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**).
- 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.

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.

DOMANDA: ALBERI B

RISPOSTA: 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. Sintetizzarlo possiamo 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.**

DOMANDA: ALBERI B+

RISPOSTA: 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.

DOMANDA: ALBERI K-D

RISPOSTA: Gli alberi K-d 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.**

GPS

DOMANDA: Quanti GPS ci sono e da quanti segmenti è composto un GPS

RISPOSTA: 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: **GALILEO** (unico non nato in campo militare, è il più preciso)
- Cina: **BEIDOU**
- India: **IRNSS**

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.

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)

DOMANDA: Orbita Geostazionaria

RISPOSTA: Per orbita geostazionaria si intende un'orbita precisa che riguarda una sfera concentrica alla terra che permette una stazionarietà apparente di un oggetto nel cielo, questo significa che se un satellite fosse in orbita geostazionaria lo vedremmo in tutti gli istanti sempre nella stessa posizione ben precisa. Per evitare questo e quindi cercare di coprire tutta la terra i satelliti GPS si trovano in orbite più basse e di conseguenza più veloci per bilanciare appunto la forza di gravità e la forza centrifuga (quanto più il satellite è vicino alla terra maggiore è l'attrazione gravitazionale che questa esercita sul satellite e maggiore è la velocità che deve avere il satellite per bilanciare le forze).

I satelliti sono posti in un'orbita circolare a circa 20.200 km dalla terra e compiono una rivoluzione in 12 ore ripassando sullo stesso punto visto da un osservatore terrestre ogni 24 ore circa.

DOMANDA: Trilaterazione e triangolazione

RISPOSTA: 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.

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

DOMANDA: Utilizzo sfere nella trilaterazione

RISPOSTA: La trilaterazione è la misurazione diretta tra un punto di cui si vuole conoscere la posizione e altri punti di cui invece conosciamo la posizione precisa (esempio se ci troviamo tra Napoli Firenze e Roma possiamo calcolare la posizione conoscendo le distanze tra la nostra posizione e le varie città). Nella trilaterazione si utilizzano le sfere poiché non viene considerato uno spazio pianeggiante e di conseguenza bidimensionale ma ci troviamo in uno spazio tridimensionale dal momento che è possibile costruire delle sfere per ogni satellite che viene considerato quando si deve calcolare la posizione.

DOMANDA: Rifrazione e Riflessione

RISPOSTA: Quando un segnale trasmesso tocca un corpo, se il segnale rimbalza e non penetra nel corpo allora si parla di riflessione se invece il segnale attraversa il corpo allora si dice che il segnale viene rifratto, cioè viene curvato rispetto alla sua direzione originaria. La somma energetica tra rifrazione e riflessione restituisce il segnale originario.

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.

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

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

DOMANDA: Differenze tra effemeridi e almanacco

RISPOSTA: L'almanacco è un database che contiene i parametri orbitali approssimati dell'intera costellazione mentre le effemeridi sono sempre un database ma che riguarda solo un satellite ed un insieme di parametri utili a calcolare la posizione di quel preciso satellite.

La principale differenza tra i due database è che l'almanacco riguarda tutta la costellazione dei satelliti (in un certo istante se uno vuole sapere quali satelliti sono visibili e di conseguenza che sono al di sopra di me lo si guarda dall'almanacco) mentre gli effemeridi riguardano solo il satellite che invia il segnale.

DOMANDA: Parallasse

RISPOSTA: Il metodo delle parallasse(**Parallax Metodo**) si può utilizzare quando bisogna misurare la distanza che intercorre ad esempio tra la terra ed una stellano essendo possibile utilizzare un metodo di misurazione diretta come inviare un fascio di luce, immaginiamo di costruire un triangolo tra la stella e due punti della terra(esempio Napoli e Milano) ma siccome la stella è molto lontana gli

angoli alla base tendono a 90 gradi, i lati del triangolo sono quasi uguali e l'angolo sulla stella è tendente quasi a 0 gradi e non disponiamo di uno strumento in grado di calcolare l'angolo allora per rendere più apprezzabile questo angolo bisogna allargare necessariamente la base misurando a distanza di 6 mesi la distanza dalla terra alla stella.

Misurando a distanza di 6 mesi il triangolo che si va a formare ha una base decisamente più grande(2 unità astronomiche "2AU") di quella del triangolo considerato precedentemente e quindi l'angolo al vertice(angolo della stella) essendo a sua volta più grande può essere calcolato.

DOMANDA: Errori nel GPS

RISPOSTA:Per ottenere una sincronizzazione perfetta, ogni orologio a bordo dei satelliti deve essere sincronizzato con tutti gli altri e con quelli sulla Terra.Tuttavia occorre considerare alcuni fattori che costituiscono fonte di errore e sono di diversa natura:

- **Fattori Atmosferici**(ritardo da 1 ns a 100 ns) : causano un rallentamento del segnale al suo passaggio nella ionosfera e nella troposfera.
- **Fattori Elettronici**(ritardo da 1 ns a 100 ns):tempo impiegato dal segnale per il passaggio nella strumentazione e anche il tempo di elaborazione nei computer del GPS.
- **Fattori Relativistici:** Anticipo/ Ritardo del tempo degli stessi orologi atomici. Questi errori sono di quasi tre ordini di grandezza più grandi rispetto alle altre sorgenti di ritardo e pertanto la loro correzione è indispensabile per il corretto funzionamento dei sistemi di navigazione satellitare.

DOMANDA: Qual è la differenza tra logger e tracker

RISPOSTA: Un **GPS tracker** è un dispositivo tecnologico di ultima generazione in grado di monitorare e tracciare gli spostamenti di oggetti o persone sfruttando i segnali della **rete satellitare GPS** (Global Positioning System, è stato sviluppato dal Ministero della Difesa USA) e sofisticati chip elettronici integrati che gestiscono l'invio in remoto (tramite rete GSM/GPRS della telefonia mobile) dei dati satellitari ricevuti e relativi all'ubicazione anche quando i segnali sono deboli. I **logger satellitari GPS** invece registrano nella loro memoria interna le coordinate GPS degli spostamenti di un veicolo. Dopo averli recuperati svelano dove è stato il veicolo, luoghi esatti e durate delle soste, ogni percorso effettuato, ecc. I logger GPS offrono importanti vantaggi: a parità di batteria hanno il 40% di autonomia in più rispetto ai localizzatori GPS, non necessitano di linee telefoniche e di copertura di telefonia mobile, non sono rilevabili con i rilevatori di microspie perchè non trasmettono segnali radio, non si registrano in nessuna rete e non lasciano tracce telematiche di nessun tipo.

DOMANDA: SISTEMA UTC

RISPOSTA: 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 Gennaio 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.

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.

DOMANDA: D-GPS

RISPOSTA: 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 fornitagli dal satellite.

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

DOMANDA: A-GPS

RISPOSTA: 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** 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.

DOMANDE VARIE

DOMANDA: Come calcolare il diametro di una moneta se non si ha uno strumento di misura abbastanza accurato

RISPOSTA: Per calcolare il diametro di una moneta nel caso non si dovesse avere uno strumento di misura abbastanza accurato basta prendere più copie di quella moneta e disporle in maniera tale che sono una attaccata all'altra. Quando avremmo raggiunto una lunghezza sufficiente per misurarle con il nostro strumento allora possiamo effettuare la misurazione e dopo basta semplicemente dividere il risultato per il numero di monete misurate per conoscere il diametro di una singola moneta.

DOMANDA: Se abbiamo due punti e un palazzo in mezzo come facciamo a calcolare la distanza tra i due punti

RISPOSTA: Nel caso dovessimo calcolare la distanza tra questi due punti basta considerare un terzo punto da cui possiamo tracciare dei segmenti che si uniscono agli altri due e di conseguenza costruire una sorta di triangolo. Dalle regole matematiche poi è possibile ricavare il lato che non è possibile conoscere dal momento che si trova il palazzo in mezzo semplicemente conoscendo l'angolo compreso tra i due segmenti che abbiamo tracciato e la lunghezza di uno dei due segmenti.

DOMANDA: Che gravità c'è al centro della Terra? (intende: cosa ci succederebbe se ci trovassimo all'interno di una stanza sicura al centro della Terra).

RISPOSTA: Ci troveremmo a fluttuare nella stanza perché la quantità di massa da cui saremmo attratti è la stessa per ogni lato.

DOMANDA: Cosa sono i watermark

RISPOSTA: I watermark sono 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 fisico 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).