

RISPOSTE ALLE DOMANDE DI SIM

1. 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:

Quindi, al centro della Terra, la forza di Gravità è zero.

La risposta è che fluttueremmo nella stanza perché la quantità di massa da cui saremmo attratti è la stessa per ogni lato.

2. Differenza lossy/lossless + esempi di entrambe

Risposta:

Compressione **lossless** è un tipo di compressione **senza perdita** ovvero permette di ricostruire perfettamente l'oggetto originale.

Esempi: Hauffman, RUNlength, LZW

Compressione **lossy** è un tipo di compressione **con perdita** di informazione, permette di ricostruire solo in parte l'oggetto originale. Questa compressione non è applicabile al testo in quanto perderebbe di significato.

Esempi: Mpeg Audio(mascheramento), Compressione con sottocampionamento, Predictive Coding, Codifica mediante trasformazione (jpeg,frattali modello lossy)

Compressione con sottocampionamento (immagini): grazie alla **ridondanza spaziale** (due pixel adiacenti hanno colore simile) vengono considerati solo alcuni pixel, come ad esempio un pixel ogni due.

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 (immagini): 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.

Predictive Coding (audio): 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.**

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

Codifica mediante trasformazione: Una immagine viene suddivisa in sottoimmagini rettangolari su cui si applica una **trasformazione unitaria dal dominio spaziale al dominio delle frequenze.**

I due tipi di trasformazione principalmente utilizzati sono: **DFT** e **DCT**.

DFT: Trasformata di Fourier: Fourier dimostrò che qualsiasi segnale periodico 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.

3. Come si sceglie il frame rappresentativo di un video?

Per ogni Shot si determinano uno o più **Frame rappresentativi (r-frame)** che vengono usati per l'indicizzazione e la ricerca.

Quanti r-frame occorre utilizzare per ogni Shot?

1 r-frame oppure **N r-frame** oppure suddividere in **sotto-shot**

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
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.**

4. - Cos'è l'eccentricità?

Indicizzazione basata sulla forma:

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

Una volta identificato il soggetto dell'immagine, esso viene incasellato 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.

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 e Scala**
- Per ogni forma si conosce **l'eccentricità**

5. Alberi B

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**.

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 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.**

Inserimento 3 casi:

1. SE la foglia **F** in cui deve essere piazzata **K** ha ancora spazio **ALLORA** si piazza **K**.

2. SE la foglia **F** in cui deve essere piazzata **K** è piena **ALLORA**:

Si crea una nuova foglia **G** e metà delle chiavi di **F** vengono spostate in **G**;

Una chiave di **F** (per es. l'ultima oppure la mediana) si sposta nel proprio padre **P**;

Nel padre **P** si introducono i puntatori di **G**

3. SE la radice **R** dell'albero è piena **ALLORA** si crea una nuova radice **S** ed un nuovo fratello **M** della radice **R**.

6. 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.

7. 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.

Range Query: Tutti gli oggetti che hanno una distanza dalla query inferiore o uguale a una certa distanza specificata vengono considerati.

8. SISTEMI UTC

Motivazione: il termine 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.

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.

9. 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**.

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).

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.

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**.

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. 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 perdersi e deve **essere recuperabile in caso di alterazione**.

Si definisce Watermark robusto quando è **in grado di resistere ad attacchi volti alla sua rimozione**.

Altra importante caratteristica è l'**autonomia** di un Watermark, che si distingue in 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.

10.D-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 fornitagli dal satellite. Questa **differenza** può essere usata per **correggere le rilevazioni errate dei terminali nel suo intorno**.

11.A-GPS

L'A-GPS è un sistema **che consente di abbattere i tempi necessari alla prima localizzazione** durante l'uso di un terminale GPS. Mostra la sua utilità soprattutto nei

"canyon" urbani, quali vie strette o viali notevolmente alberati, in cui è difficile stabilire con precisione la lista di satelliti in vista al terminale.

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.

11.SEGMENTAZIONE A DUE SOGLIE

Spesso in un video non sono utilizzate transazioni a taglio netto, ma si utilizzano delle dissolvenze incrociate. Se bene all'occhio sia molto più gradevole, per gli algoritmi di segmentazione automatica in Shots crea un bel problema.

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.

12.ALGORITMO ADC

Conversione del segnale da Analogico a Digitale.

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.

13. ORBITA GEOSTAZIONARIA

Orbita geostazionaria, orbita che fa sembrare che un satellite si trova a una posizione fissa nel cielo, mentre invece si trova in un'orbita più bassa poiché più veloce, per sfuggire alla forza di gravità ho bisogno di maggiore velocità di rotazione attorno alla terra.

14. ERRORE NEI GPS

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**

15. Trilaterazione e Triangolazione

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.

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.

16. PRECISIONE E RECALL

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.

17. DOMINIO TEMPORALE file audio

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

- Dominio temporale (Tempo/Ampiezza)
- Dominio delle frequenze (Frequenza/Magnitudine)

La rappresentazione **Time Domain** è la tecnica più immediata ed intuitiva per la rappresentazione di un segnale la cui ampiezza varia nel tempo.

Il silenzio è rappresentato dallo **Zero**

I valori del segnale possono essere positivi o negativi a seconda se la pressione d'aria provocata dall'onda sonora risulta essere superiore o inferiore alla pressione atmosferica in condizioni di silenzio.

Caratteristiche derivate dal Dominio temporale:

Average Energy: 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.

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

Silence Ratio: Indica la quantità di silenzio nel brano.

18. DOMINIO DELLE FREQUENZE

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**.
- Per il testare se un suono contiene armoniche si controlla che le **frequenze di componenti dominanti** siano multiple di una **frequenza fondamentale**.

19. Come vengono selezionati gli shot per l'indicizzazione dei video

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:

$SD_i =$ sommatoria che parte da j del modulo della differenza tra $H_i(j)$ e $H_{i+1}(j)$

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**.

20. INDICIZZARE LE IMMAGINI

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:

21.DIFFERENZA TRA TACKER E LOGGER

Strumenti gps logger serve per collezionare informazione di spostamento del gps quindi tutti gli eventi successi, esempio durante jogging.. i dati si conoscono a posteriori esempio cheap per sapere il cane dove e stato

Tracker è un logger che ha una sim e i dati vengono trasferiti a un telefono tramite sms, quindi in qualsiasi momento posso sapere la posizione del mio cane. Utile per antifurti satellitari.

22.FILE 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

23.TEOREMA NQUIST

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**.

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

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.

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.

24.ERRORE DI QUANTIZZAZIONE

Errore (o rumore) di quantizzazione:

$\text{Max} \{ \text{Campione_quantizzato} - \text{segnale_analogico} \}$

Se l'errore di quantizzazione supera il valore della soglia uditiva allora viene avvertito.

25.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

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.

Riflessione: riflesso del flusso di luce

26.ARMONICA

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**.
- Per il testare se un suono contiene armoniche si controlla che le **frequenze di componenti dominanti** siano multiple di una **frequenza fondamentale**.

27. FEATURE VECTOR

Per un insieme di oggetti multimediali, la fase di indicizzazione produce un insieme di vettori di caratteristiche (**feature vector**).

Ciascun vettore V , a seconda della complessità dell'oggetto indicizzato, può contenere un **numero grande di componenti** (caratteristiche testo, audio, immagini, video,...)

28. 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.

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**.

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 su 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.

29. PRODOTTO SCALARE

Il **prodotto scalare** di due vettori è uguale al prodotto dei loro moduli, moltiplicato per il coseno dell'angolo compreso tra di essi.

30. RETI NEURALI ARTIFICIALI

Le ANN (Artificial Neural Networks), largamente impiegate per il riconoscimento, simulano i processi cognitivi, in particolare l'apprendimento, del cervello umano.

Una ANN è costituita da due fasi:

Training (addestramento): i vettori di caratteristiche ottenuti durante l'addestramento di parlato servono per tarare i pesi dei link della rete.

Recognition: l'ANN seleziona il fonema più verosimile basandosi sulle caratteristiche dei vettori.

Le ANN sono basate su implementazioni che prevedono l'utilizzo di un grande numero di elementi di calcolo (PE – Processing Elements) molto semplici e interconnessi tra di loro.

Ogni PE implementa una semplice funzione matematica di tipo non lineare e rappresenta un **NEURONE**.

Le connessioni rappresentano le **SINAPSI** (strutture nervose di collegamento tra neuroni che si occupano di trasportare segnali elettrici che agiscono da inibitori o eccitatori tra neuroni diversi).

In una rete neurale artificiale ogni connessione è caratterizzata da un PESO (w) in $[-1..1]$:

Il peso rappresenta l'influsso che un neurone porta sul neurone destinatario della connessione

La rete viene inizializzata con pesi random e, durante l'esecuzione, l'insieme dei pesi rappresenta la "conoscenza" detenuta dalla rete.

I neuroni sono in genere organizzati in livelli (**livello di input**, **livelli nascosti e livello di output**) e le connessioni sono stabilite tra neuroni appartenenti a livelli differenti.

Ogni neurone della rete (PE) effettua una somma pesata **INTEGRAZIONE** degli input derivanti dalle connessioni con gli altri neuroni.

L'input pesato viene poi valutato da una funzione detta di **TRASFORMAZIONE** che determina l'output del singolo PE.

31.ISTOGRAMMA DI COLORE

I **bins** sono una **quantizzazione di un gruppo di sfumature** 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 su istogramma di colore:

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

Per la ricerca delle immagini nel DB serve definire una misura di distanza tra l'istogramma dell'immagine query e quelli delle immagini contenute nel database.

Date 2 immagini A e B, la **misura di distanza** più semplice è data da:

Sommatoria di i che va da uno a n del modulo della differenza tra a_i e b_i

in cui: a_i e b_i = il numero pixel delle immagini A e B che ricadono nel bin i-esimo

32.DIFFERENZA TRA J-FRAME E KEY-FRAME

Codifica Interframe

Descrizione dei cambiamenti che occorrono tra un fotogramma ed il successivo partendo da un fotogramma iniziale descritto con codifica intraframe.

Sequenze video con **pochi elementi che cambiano nella scena**.

Metodo preferito per la fase **distribuzione di un filmato**.

Complesso accesso diretto ad una scena generica;

tecnica dei keyframes: per la creazione di fotogrammi successivi

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

34 GPS

Il GPS (Global Positioning System) è **un sistema di posizionamento** su base satellitare, a copertura globale e continua, gestito dal dipartimento della difesa statunitense.

Il GPS permette a piccoli ricevitori elettronici di determinare la loro posizione (longitudine, latitudine e altitudine) con un errore di pochi metri usando segnali radio trasmessi in linea ottica da satelliti.

Navstar (Navigation System with Timing And Ranging Global Positioning System) fu concepito dal Ministero della Difesa statunitense come mezzo universale per: determinare con precisione il punto in cui un ricevitore si trova sulla terra ottenere un'indicazione oraria molto precisa.

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

- **Il segmento spaziale** (24 Satelliti)
- **Il segmento di controllo** (5 centri di controllo)
- **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.

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.

Colorado Springs

- Raccoglie i dati delle altre stazioni
- Compensazione dei dati

Segmento utilizzo

Definito da ogni ricevitore, da aerei ed automobili fino ai cellulari. Essi ricevono il segnale, lo elaborano, ricavano le informazioni a proposito della velocità, posizione e tempo del veicolo. Essi sono basati su clock al quarzo.

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

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

ΔT = orario di arrivo - orario di invio

Distanza = velocità (velocità della luce) * tempo (ΔT)

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

35. EFFEMERIDI E ALMANACCO

Effemeridi

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.

Almanacco

l'almanacco (parametri orbitali approssimati) dell'intera costellazione.

36. PARALLASSE

Metodo parallasse calcolare distanza tra terra e stella, uso la triangolazione costruendo un triangolo osservando la stella da due punti (Napoli, Milano) ma l'angolo è troppo stretto, per avere un angolo maggiore faccio la misurazione a distanza di sei mesi così la Terra sta nella parte opposta dell'orbita intorno al Sole quindi creo un triangolo molto più ampio. Unità astronomica e distanza tra Terra e Sole.

37. 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 su:

- Limite della percezione visiva dell'occhio umano
- Analisi spettrale dell'immagine

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 DCT 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 la altre frequenze.
- **Codifica:** L'immagine viene scansionata a Zigzag, eliminando così le ridondanze mediante codifica Run-Length e Huffman.

ALTRE DEFINIZIONI

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.

Companding: Trasformazione di un segnale da lineare a non lineare.

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.

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.

Le Immagini

Gli algoritmi di indicizzazione e di ricerca delle immagini sono basati su di una **Rappresentazione dei colori**;

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

- Luminanza (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.

Sintesi Sottrattiva: 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.

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.

$$V_{out} = V_{in}$$

Requisiti di un MIRS

Estensibilità a nuovi tipi e formati di dati.

Flessibilità per permettere l'inserimento e la ricerca a vari livelli di astrazione

Predisposizione per la rappresentazione dei dati multimediali semplici e composti comprese le relazioni spaziali e temporali che intercorrono tra di essi.

Efficienza nelle strategie di memorizzazione e ricerca.

Layer: Oggetto, tipo e formato

Fase di ricerca: Per specifica, per esempio(authoring->inserimento)

Tipi di Features: metadata, annotazioni testuali, feature a basso livello(dati e statistiche), feature a alto livello(contenuto)

QoS Quality of Service

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

- Qualità preferibile
- Qualità accettabile

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

Multimedia data compression: 3 metodi: anteprima(copia), ridimensionamento(originale), compressione(metodi decompressione)

Indicizzazione del testo:

Inverted file: termine, puntatore

Retrieval con modello spazio vettoriale: A questo punto la similitudine tra un file ed una Query è semplicemente l'angolo formato dai due vettori. Quanto più l'angolo θ è piccolo, tanto più Query e File si somigliano.

Modello Cluster

Indicizzazione e Ricerca dell'audio: Parlato, Musica, Rumori

Metodi di classificazione dell'audio: step by step, caratteristiche vettoriali

Problematiche

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

Due Bins adiacenti sono considerati completamente diversi.

3 soluzioni: distanza tra i bins, istogramma cumulativo, PHW(da rgb a cielv)

Calcolo della similarità:

- 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 e Ricerca del video

Le principali metodologie impiegate per l'indicizzazione ed il recupero di un video sfruttano: metadati, testo, audio, 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.

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**.

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.

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

Tipi di Query e Vettori

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.

Alberi MB+

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.

Grid files

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

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**.

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, insert, delete.

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.

GIS - Geographic information system

Un GIS può essere definito come un **Sistema informatico per la gestione dei dati geografici**; più propriamente è una superclasse di sistemi informatici più completi per il Rilevamento Geografico.

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:** Un GIS colleziona **grandi quantità di dati da sorgenti eterogenee**. Due tipi ben distinti tra loro: • **Spaziale** (Geografico) • **Alfanumerico** (Attributi)

2. **Visualizzazione:** Per tanto un GIS ha **la capacità di presentare graficamente** ed in **maniera organizzata**, le informazioni raccolte mediante la creazione di **mappe**.

3. **Interrogazione:** Le informazioni presenti in un GIS sono memorizzate attraverso **oggetti grafici**, ovvero elementari detti Features. Deve permettere la richiesta dei dati in 3 modi: **interrogazione degli attributi**, **interrogazione spaziale**, **interrogazione topologica**.

4. **Combinazione:** deve permettere **la combinazione di informazioni provenienti da fonti diverse**, così da permettere di scoprire correlazioni o di generare dati derivati.

5. **Analisi:** Partendo da dati grezzi un GIS può elaborare informazioni utili al sistema, mediante l'uso di strumenti **Statistici e di Modellazione**.

6. **Predizione:** 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?**

Caratteristiche

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

CAD: per immagini vettoriali(mappe)

Image Processing: per immagini raster(foto aeree, immagini satellitari)

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**.

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

Oggetti base di un GIS

Punto: Entità **Adimensionale** può essere 3 tipi: entità , area , nodo

Arco o Linea: Entità **Unidimensionale**

Anello: **Insieme di archi** che costituiscono un anello chiuso.

Poligono: Insieme di uno o più anelli. Entità **Bidimensionale**

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

Modello **Terra Piatta:** Adatto per mappe di dimensioni non superiori ai 2-3Km e pianeggianti 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.

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.

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

- **Paralleli o linee di Latitudine**

- Cerchi paralleli all'equatore.
- Parallelo 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.

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.

Non esiste alcuna **proiezione o trasformazione** capace di creare un modello che rappresenti fedelmente una superficie curva su di una piana senza introdurre distorsioni

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.**