DBMS multimediali

ALBERTO BELUSSI

BASI DI DATI
ANNO ACCADEMICO 2012/2013

DBMS multimediali

Def: Sono DBMS che consentono di memorizzare e recuperare dati di natura multimediale: testo, immagini, suoni, animazioni, voce, video, ecc..

Introducono rispetto ai DBMS tradizionali problematiche diverse:

- Dimensione dei dati molto maggiore
- Gestione dei media continui (video e audio)
- Complessità della modellazione
- Interrogazioni basate su predicati non “esatti”: interrogazioni approssimate

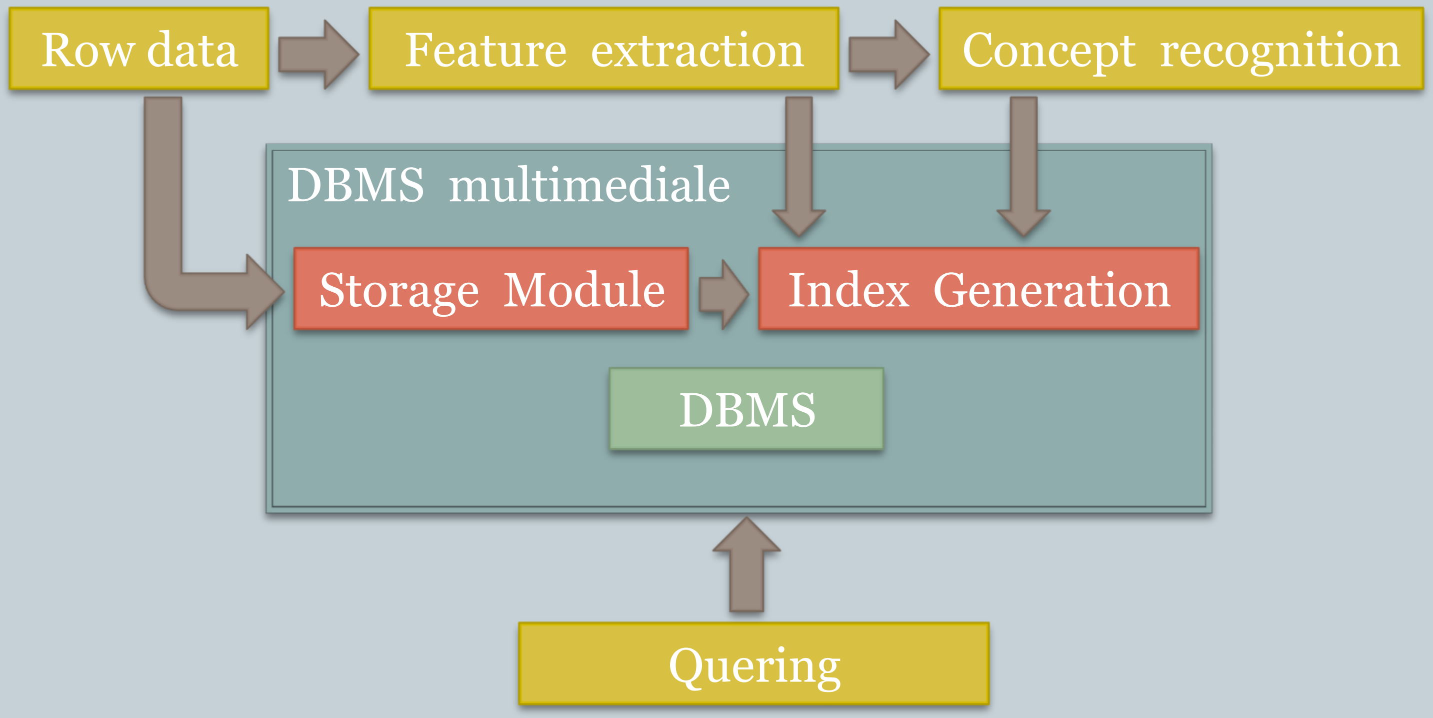
DBMS multimediali

Interessano diversi settori applicativi:

- **Applicazioni mediche** (TAC, RX, ecc…)
- **Applicazioni scientifiche** (dati biologici, meteorologici, spettrografici, ecc…)
- Applicazioni legali (impronte digitali, volti, ecc…)
- Applicazioni educative o di intrattenimento (TV on demand, distance learning)
- **Applicazioni geografiche**
- …

Gestione dei dati multimediali in un DBMS

Schema di riferimento



Dati multimediali

Il dato multimediale si presenta in diverse forme: testo, audio, immagini e video.

La caratteristica condivisa da tutte le forme di dato multimediale è il fatto che la generazione del dato multimediale (nella sua forma finale di natura digitale) si produce dal campionamento di un segnale analogico.

Tale fase di rilevazione del dato ("*data capture*") può essere eseguita in modi diversi e con diversi livelli di accuratezza producendo dati di diversa qualità.

Dati multimediali

Il primo passaggio che introduce approssimazione è quello da *segnale analogico* alla *rappresentazione digitale*.

Il ben noto **teorema di Nyquist** indica che per riprodurre accuratamente un segnale analogico di frequenza massima f con un segnale digitale è necessario campionare il segnale analogico con una frequenza pari almeno a 2f.

Ad esempio, la voce umana occupa un intervallo di frequenze da 0 a 7 kHz. Di conseguenza dovrebbe essere campionata con una frequenza di 14 kHz (circa ogni 70 μs).

Dati multimediali

Il dato digitale che ottengo con il campionamento può comunque essere "*molto grande*" rispetto ai dati che tradizionalmente tratta una base di dati, pertanto si introduce un ulteriore livello di approssimazione a seconda dei requisiti dell'applicazione.

Per esempio per le applicazioni telefoniche il range di frequenze può essere ridotto all'intervallo 300-3400 Hz, quindi il campionamento in questo caso può essere ridotto ad una frequenza di circa 8 kHz (ogni 125 μs).

Dati multimediali

Inoltre la dimensione della rappresentazione digitale finale dipende anche dal numero di cifre binarie che uso per rappresentare il valore campionato.

Per le **immagini** il campionamento dipende dalla griglia di **pixel** che si adotta per la rappresentazione della stessa e dalla dimensione del pixel.

La dimensione finale dipende poi da cosa viene memorizzato nel pixel: un solo valore di **intensità di luce**, oppure la rappresentazione di un **colore** che può richiedere ad esempio 24 bit (8 bit per ogni colore fondamentale – per un totale di circa 16 milioni di colori).

Dati multimediali

Quindi una delle caratteristiche comuni ai dati multimediali è la **grande quantità di bit necessari per rappresentarli**.

Esempio su dati video

Durata	640 x 480	320 x 240	160 x 120
1 secondo	27 Mb	6,75 Mb	1,68 Mb
1 minuto	1,6 Gb	400 Mb	100 Mb
1 h	97 Gb	24 Gb	6 Gb

Dati multimediali: compressione

Per rimediare parzialmente a questo problema si applicano **compressioni** della rappresentazione digitale con vari algoritmi.

Sulla compressione vedremo:

- Alcuni concetti generali
- Due tecniche in dettaglio: Huffman e Lempel-Ziv-Welch

Tecniche di compressione: concetti generali

Proprietà statistiche di una sorgente S di simboli statisticamente indipendenti (emissione casuale)

- $A=\{a_1,a_2,\ldots,a_\alpha\}$ (alfabeto sorgente)
- α (cardinalità dell'alfabeto)
- $p(a_i)$ (probabilità del simbolo a_i)
- $I(a_i) = \log_2(1/p(a_i)) = -\log_2(p(a_i))$ (quantità di informazione di a_i espressa in bit)

$$I(a_i\, a_j) = \log_2(1/(p(a_i)\, p(a_j))) = \log_2(1/p(a_i))+\log_2(1/p(a_j))=I(a_i)+I(a_j)$$

Entropia

<u>Entropia</u> della sorgente S di alfabeto A={a_1,a_2,…,a_α}

$$H = \sum_{i=1}^{\alpha} p(a_i)\log_2\frac{1}{p(a_i)} = \sum_{i=1}^{\alpha} p(a_i)I(a_i)$$

Proprietà:

- $H \geq 0$
- $H = 0$ *solo se c'è un solo simbolo con probabilità 1*
- $H \leq \log_2(\alpha)$
- $H = \log_2(\alpha)$ *solo se tutti i simboli sono equiprobabili* ($p(a_i)=1/\alpha$)

Messaggi

Messaggio su alfabeto $A=\{a_1,a_2,\ldots,a_\alpha\}$

- $m_i=(a_1,a_2,\ldots,a_N)$ (messaggio di lunghezza N)
- N ($N>>\alpha$) (numero simboli del messaggio)
- $p(a_i)$ (probabilità del simbolo a_i)
- $p(m_i)= [p(a_1)]^{N\times p(a1)} \times\ldots\times [p(a_\alpha)]^{N\times p(a\alpha)}$ (probabilità del messaggio m_i)

Entropia media dei messaggi m_i di lunghezza N

$$I(m_i)=\log_2\left(\frac{1}{p(m_i)}\right)=N\sum_{i=1}^{\alpha}p(a_i)\log_2\left(\frac{1}{p(a_i)}\right)=H_N$$

$$H_N=N\times H_{sorgente}$$

Messaggi

Codifica di un messaggio su alfabeto $A=\{a_1,a_2,\ldots,a_\alpha\}$

- $p(a_i)$ (probabilità del simbolo a_i)
- $code(a_i)$ (codifica del simbolo a_i)
- $length(code(a_i))$ (lunghezza della codifica di a_i)
- $m_i=(a_1,a_2,\ldots,a_N)$ (messaggio)

Lunghezza media della codifica di un messaggio m con N simboli:

$$L_N = \sum_{i=1}^{\alpha} N \times p(a_i) \times length(code(a_i))$$

Teorema di Shannon

Primo teorema di Shannon

La minima lunghezza media L_N per la codifica in bit di un messaggio di N simboli su alfabeto A (dove siano assegnate le probabilità per ogni simbolo) è pari all'entropia dei messaggi di N simboli (anche detta quantità di informazione media dei messaggi di lunghezza N):

$$H_N \leq L_N$$