



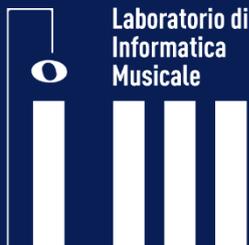
UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI INFORMATICA

Music Fingerprinting

The Shazam magic!

G. Presti

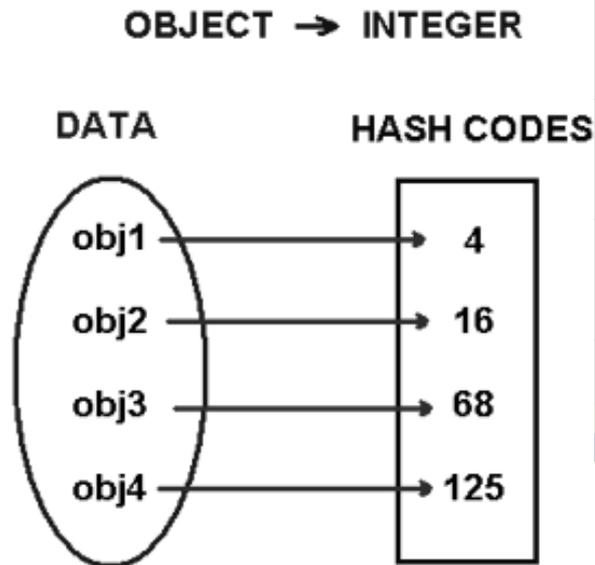
10/06/2016



Hash

Definizione

- L'hash è una funzione non iniettiva che mappa un dato di dimensione arbitraria in uno di lunghezza predefinita
- Questo deve essere calcolato in modo che dati simili generino hash molto diversi



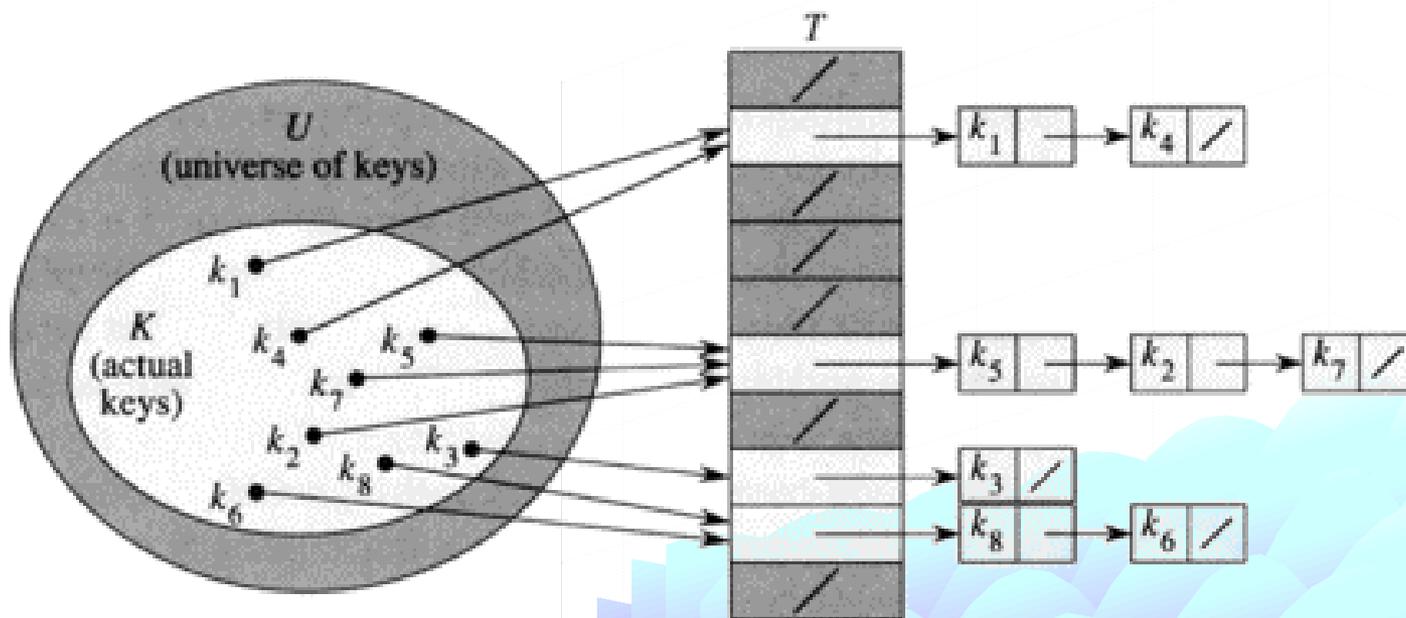
Uso tipico: in crittografia
per il controllo dei dati

Hash table

Definizione

Occhio alle collisioni!
Si sacrifica spazio per evitarle

Una hash table è un array, in cui ogni elemento è posizionato nell'indice che corrisponde al suo hash.

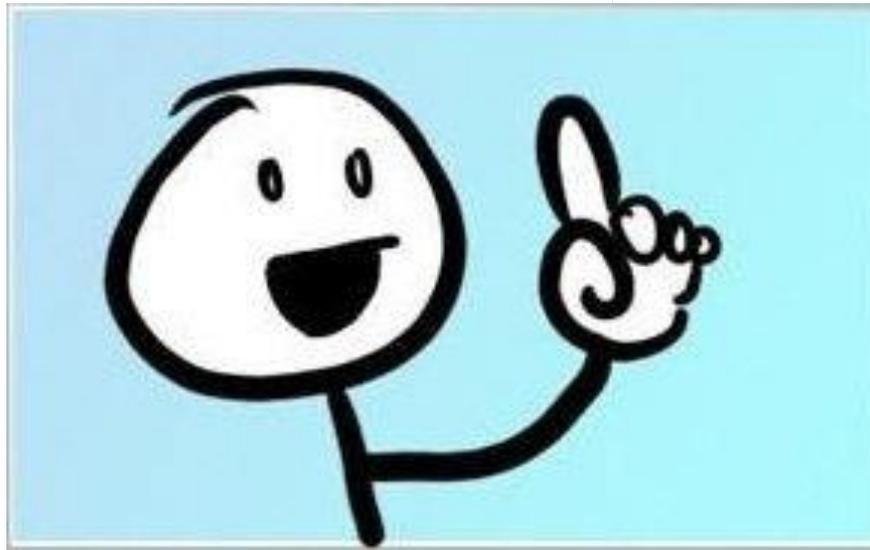


Questo rende possibile ricerche molto rapide!

Fingerprint

Definizione

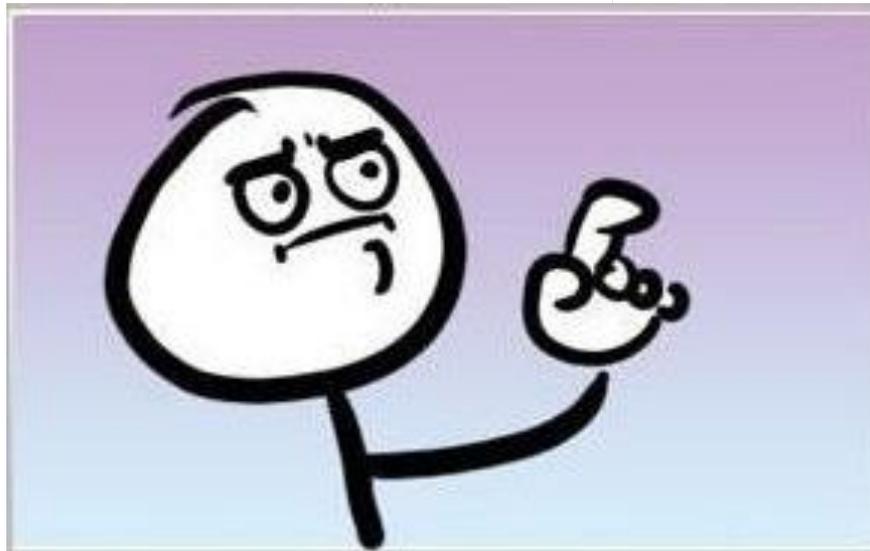
- Un algoritmo di **fingerprinting** mappa un dato molto grande su una stringa di bit più breve, che identifica in modo univoco i dati originali



Fingerprint

Definizione

- Un algoritmo di **fingerprinting** mappa un dato molto grande su una stringa di bit più breve, che identifica in modo univoco i dati originali
- E' auspicabile tolleranza al rumore e alle distorsioni



L'hash si riferisce ai dati,
il fingerprint alle informazioni

Fingerprint audio

Il problema

Un algoritmo di ricerca audio deve:

- Funzionare con spezzoni brevi di un brano
- Essere insensibile al rumore di fondo e al possibile riverbero
- Essere robusto alle distorsioni non lineari
- Generare rapidamente la risposta, ricercando all'interno di un database di riferimento molto grande (milioni di brani).

Fingerprinting + hashing + Hash table

Fingerprint audio

Le caratteristiche

Le fingerprint devono essere:

- **Temporalmente localizzate:** eventi distanti non devono influenzare la tabella di hash
- **Invarianti alle traslazioni:** riproducibili indipendentemente da quale segmento è stato registrato
- **Robuste:** riproducibili anche se in ambienti rumorosi
- **Sufficientemente entropiche:** minimizzare la probabilità di falsi match

Fingerprinting di Shazam

Riferimenti

- **Algoritmo:**

Wang, Avery. "An Industrial Strength Audio Search Algorithm." ISMIR. 2003.

- **Codice Matlab basato su:**

D. Ellis (2009), "*Robust Landmark-Based Audio Fingerprinting*", web resource, available: <http://labrosa.ee.columbia.edu/matlab/fingerprint/>

Fingerprinting di Shazam

Architettura

L'idea è di estrarre una **fingerprint**¹

su cui calcolare più **hash**² da usare per la ricerca

e ritornare il brano con più **match**³

Fingerprinting di Shazam

Architettura

Server side: fase preliminare

1. Calcolo delle fingerprint del database
2. Calcolo degli hash

Client side:

1. Calcolo delle fingerprint del frammento
2. Calcolo degli hash del frammento
3. Invio al server degli hash da cercare

Server side: runtime

1. Lookup nella hash table
2. Conteggio dei match

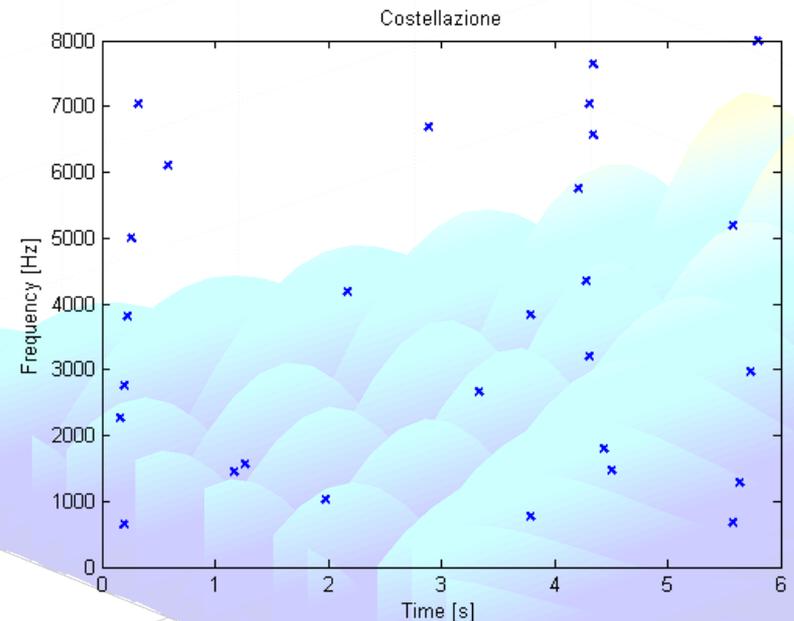
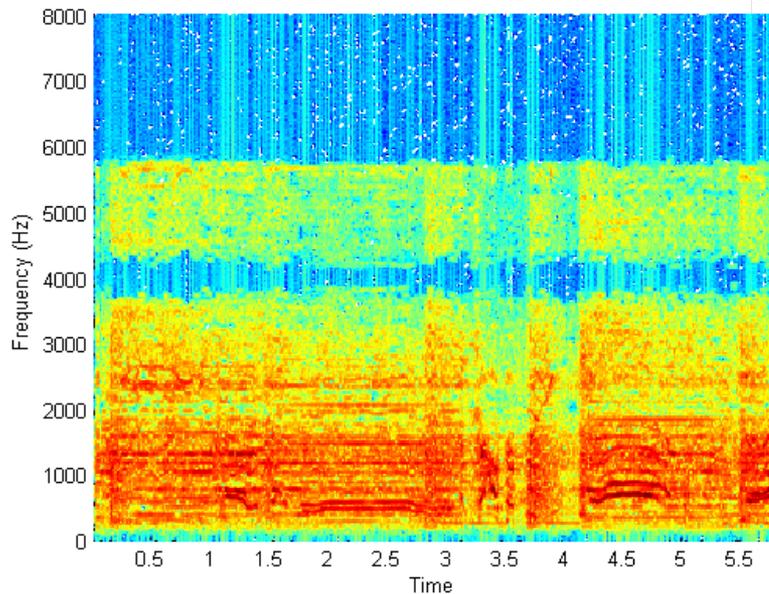
Fingerprinting di Shazam

1 Fingerprinting

*Dei picchi si salva solo f e t
Si potrebbe già cercare con queste informazioni?*

Fingerprint: Costellazione di landmarks

- Un punto tempo-frequenza è un **picco** se presenta una energia elevata rispetto ai punti vicini
- I picchi sono selezionati con un criterio di **densità**

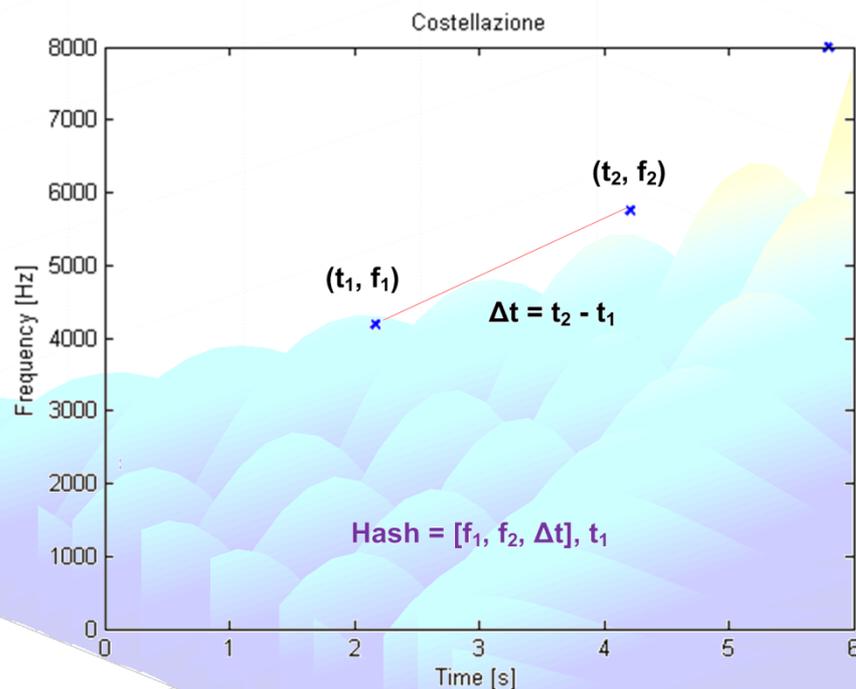
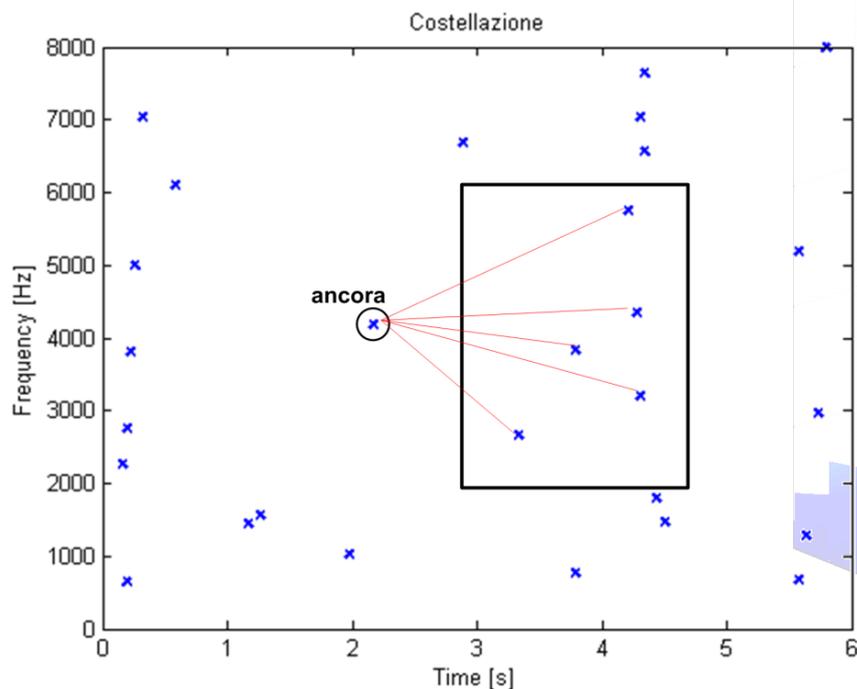


Fingerprinting di Shazam

2 Hashing

- Più hash per canzone
- t_1 come dato satellite (e anche l'ID del brano)

- Ogni picco diventa un'ancora, usata come riferimento per una zona target di picchi successivi
- Gli hash sono generati concatenando le informazioni relative ad una coppia “ancora - picco nel target”



Fingerprinting di Shazam

3 Matching

Attraverso l'ID del brano
poi si risale al titolo

- Dalla **hash table** vengono recuperati tutti i tempi di offset che corrispondono agli hash del campione da cercare
- Per ogni brano si controlla la distribuzione delle differenze tra gli offset: vince chi ha corrispondenze consecutive

Scatterplot of matching hash locations: No diagonal

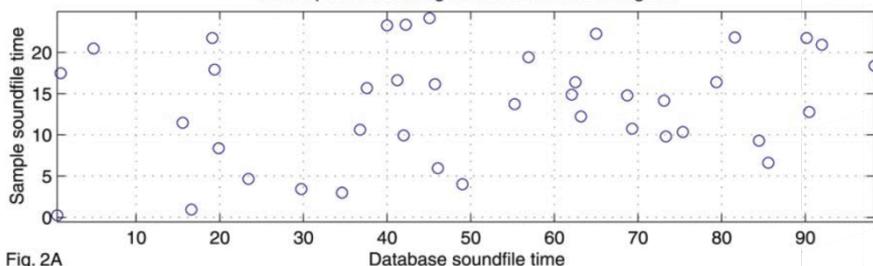


Fig. 2A

Histogram of differences of time offsets: signals do not match

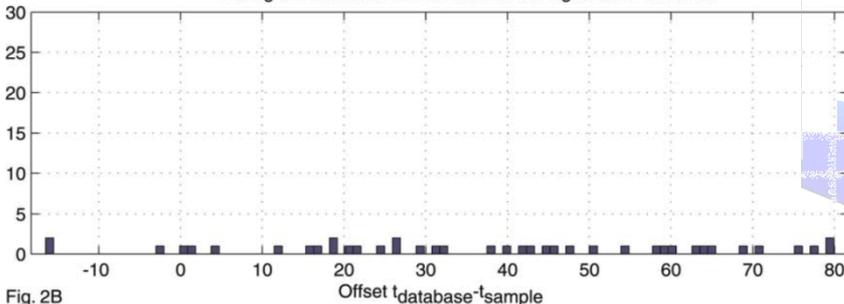


Fig. 2B

Scatterplot of matching hash locations: Diagonal Present

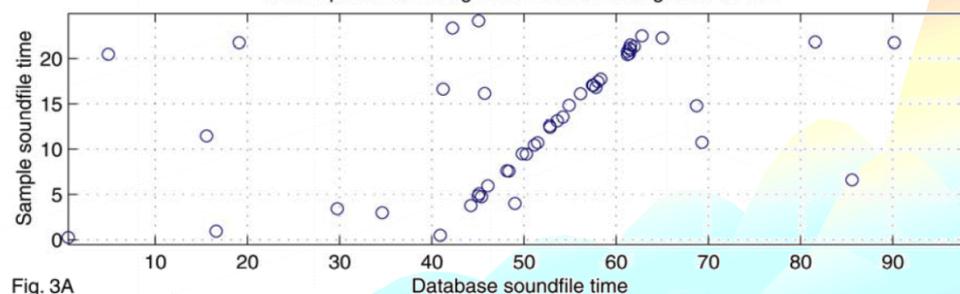


Fig. 3A

Histogram of differences of time offsets: signals match

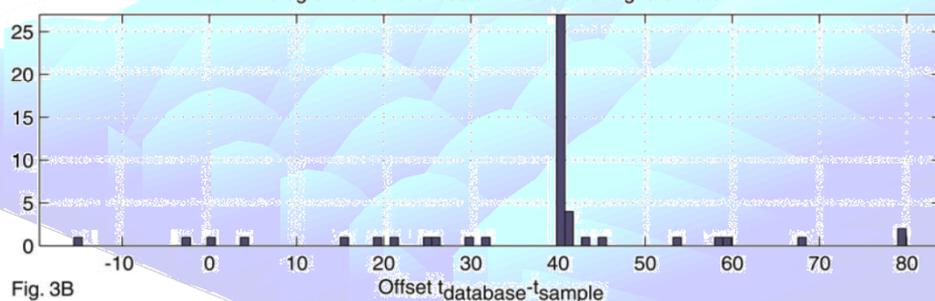


Fig. 3B

Fingerprinting di Shazam

Matlab!

Ecco il ruolo delle funzioni di Matlab che eseguiremo:

demo.m: il main
init_hashtable.m: inizializza la hash table
add_tracks.m: aggiunge brani alla hash table
match_query.m: dato un frammento cerca il brano
illustrate_match.m: plot graziosi

Che si appoggiano a:

listfiles.m: è tipo il dir del dos o ls di unix
find_landmarks.m: calcola i picchi (la fingerprint)
landmark2hash.m: da landmark a hash
hash2landmark.m: da hash a landmark Qui si bara...
record_hashes.m: inserisce gli hash nella hash table
get_hash_hits.m: ritorna le corrispondenze
show_landmarks.m: plot graziosi