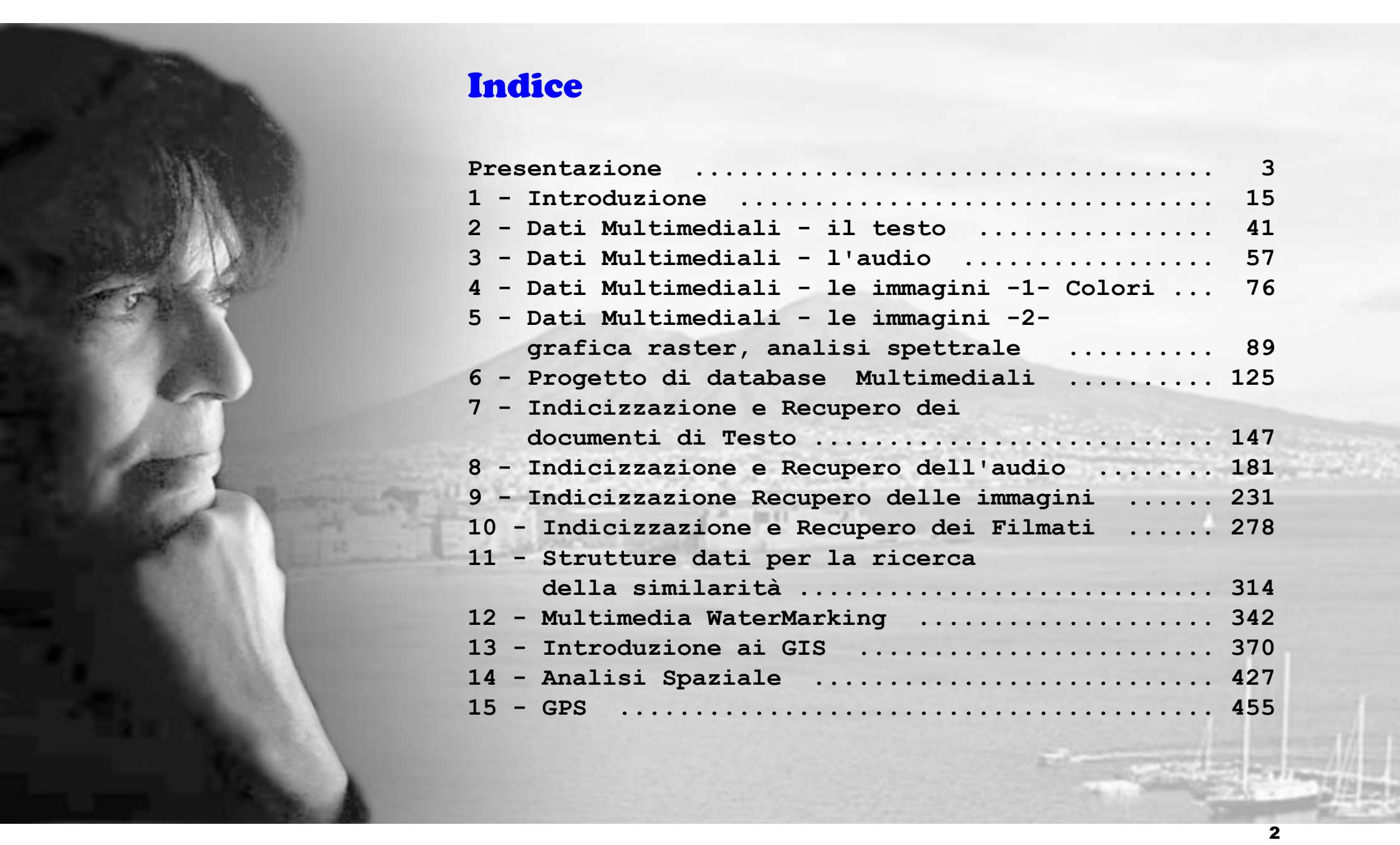




Walter Balzano
Università Federico II
Napoli

Sistemi Informativi Multimediali



Indice

Presentazione	3
1 - Introduzione	15
2 - Dati Multimediali - il testo	41
3 - Dati Multimediali - l'audio	57
4 - Dati Multimediali - le immagini -1- Colori ...	76
5 - Dati Multimediali - le immagini -2- grafica raster, analisi spettrale	89
6 - Progetto di database Multimediali	125
7 - Indicizzazione e Recupero dei documenti di Testo	147
8 - Indicizzazione e Recupero dell'audio	181
9 - Indicizzazione Recupero delle immagini	231
10 - Indicizzazione e Recupero dei Filmati	278
11 - Strutture dati per la ricerca della similarità	314
12 - Multimedia WaterMarking	342
13 - Introduzione ai GIS	370
14 - Analisi Spaziale	427
15 - GPS	455

Welcome...!

Sistemi Informativi Multimediali

Walter Balzano

Dipartimento di Ingegneria Elettrica
e Tecnologie dell'Informazione

Web: <https://sites.google.com/site/walterbalzano/>

E-mail: walter.balzano@gmail.com



walter.balzano@gmail.com

Organizzazione del Corso

Lezioni

- Esame
 - Prova scritta (vedi web site)
 - Prova orale (vedi web site)
- Ricevimento
 - Per un appuntamento inviare una E-mail



walter.balzano@gmail.com

Info

Tutte le principali informazioni saranno costantemente aggiornate sul sito web del corso. Quindi:

- 1) Vai alla URL:
<https://sites.google.com/site/walterbalzano>
- 2) poi Clicca su **Didattica**
- 3) poi Clicca su **Sistemi Informativi Multimediali**



walter.balzano@gmail.com

Registrazione al Corso

La registrazione al corso è
obbligatoria !



walter.balzano@gmail.com

PROGRAMMA - 1

Architettura dei Sistemi Informativi Multimediali

- Tipi di dati, formati e standard
- Analisi della percezione
- Indicizzazione e Ricerca
- Misure di efficacia e sicurezza
- Linguaggi per i metadati
- Multimedialità e semantica
- Ontologie Multimediali
- Multimedia Data Mining
- Definizioni



walter.balzano@gmail.com

PROGRAMMA - 2

Sistemi Informativi Territoriali

- Definizioni
- Struttura di un GIS
- Aggregazione dei dati
- Cartografia digitale
- Modelli di GeoReferenziazione
- DATUM Geodetici e mappe
- Principali metodi di analisi: Voronoi, NNI, Variogrammi, Interpolazioni



walter.balzano@gmail.com

PROGRAMMA - 3

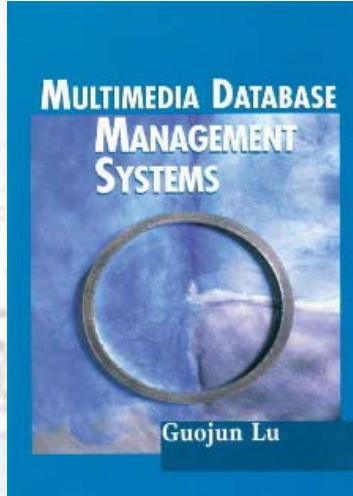
Il Global Positioning System

- Definizioni
- La Trilaterazione
- Segmenti GPS: Spazio, Controllo, Uso
- Sinc. Spazio/Tempo
- Il modello D.O.P.
- Errori relativistici e fonti minori
- Varianti GPS: A-GPS e D-GPS
- Tracking-Logs: file gpx
- Applicazioni ed esempi



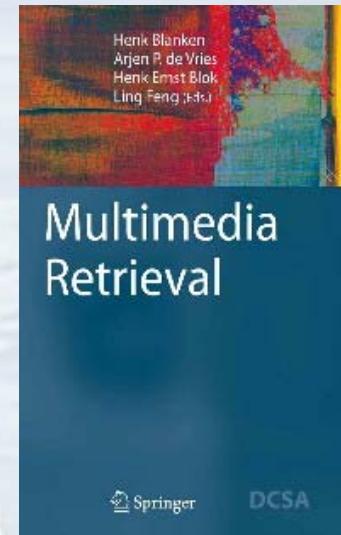
walter.balzano@gmail.com

Libri di Testo consigliati - 1



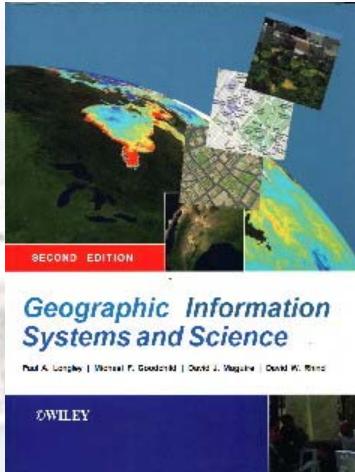
Guojun Lu
Multimedia Database Management Systems
Artech House
ISBN: 1-89006-342-7

Blanken, de Vries, Blok, Feng
Multimedia Retrieval
Springer
ISBN: 978-3-540-72894-8



walter.balzano@gmail.com

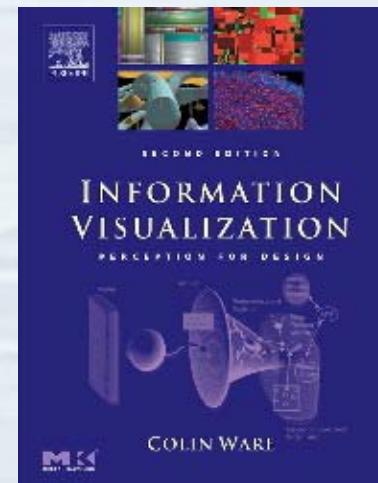
Libri di Testo consigliati – 2



Longley, Goodchild, Maguire, Rhind
*Geographic Information Systems
and Science*
Wiley
ISBN: 0-470-87001-X

Colin Ware
*Information Visualization
Perception for design*

Morgan Kaufmann
ISBN: 1-55860-819-2



walter.balzano@gmail.com

Altre fonti...

- **Piattaforma di E-learning
Federica**

<http://www.federica.unina.it>



- **Link diretto:**

<http://www.federica.unina.it/corsi/sistemi-informativi-multimediali>



walter.balzano@gmail.com

Altre fonti...

The screenshots show:

- Home page:** Profile picture of Walter Balzano, course title 'Sistemi informativi multimediali', contact information (phone 081 879310, email walter.balzano@unisa.it), and a brief bio.
- Anagrafica del corso:** Course details including name (Anagrafica del corso), code (Scienze Matematiche Fisiche e Naturali), year (2009/2010), and professor (Walter Balzano).
- iTools:** RSS feed information for the course.
- Federica sul tuo iPod Touch / iPhone:** Information about the mobile version of Federica, showing a smartphone icon.
- Walter Balzano > 1) INTRODUZIONE:** A diagram titled 'Fasi dell'era digitale' (Phases of the digital era) showing a flow from a stack of books to a computer monitor, labeled 'Digitalizzazione' and 'Management'.
- MIRS: funzionamento:** A flowchart titled 'MIRS (Multimedia Indexing and Retrieval Systems)' showing the process: Queries → Query Features → Processing and Feature extraction → Similarity computation → Retrieval of Similar Items, and parallel paths for Information Items and Preprocessing and Indexing.



walter.balzano@gmail.com

NOTA BENE:

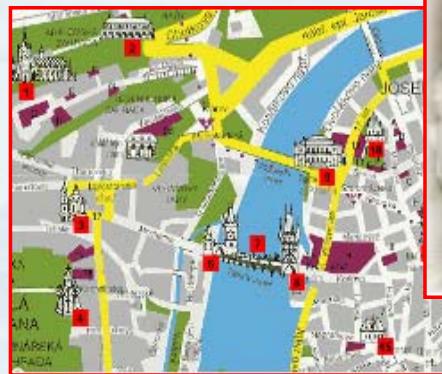
**Le slides ed i video
non sostituiscono i libri di testo**

*...ed ora si inizia,
Buon Lavoro !*



walter.balzano@gmail.com

Premessa L'era digitale



walter.balzano@gmail.com

Premessa L'era digitale



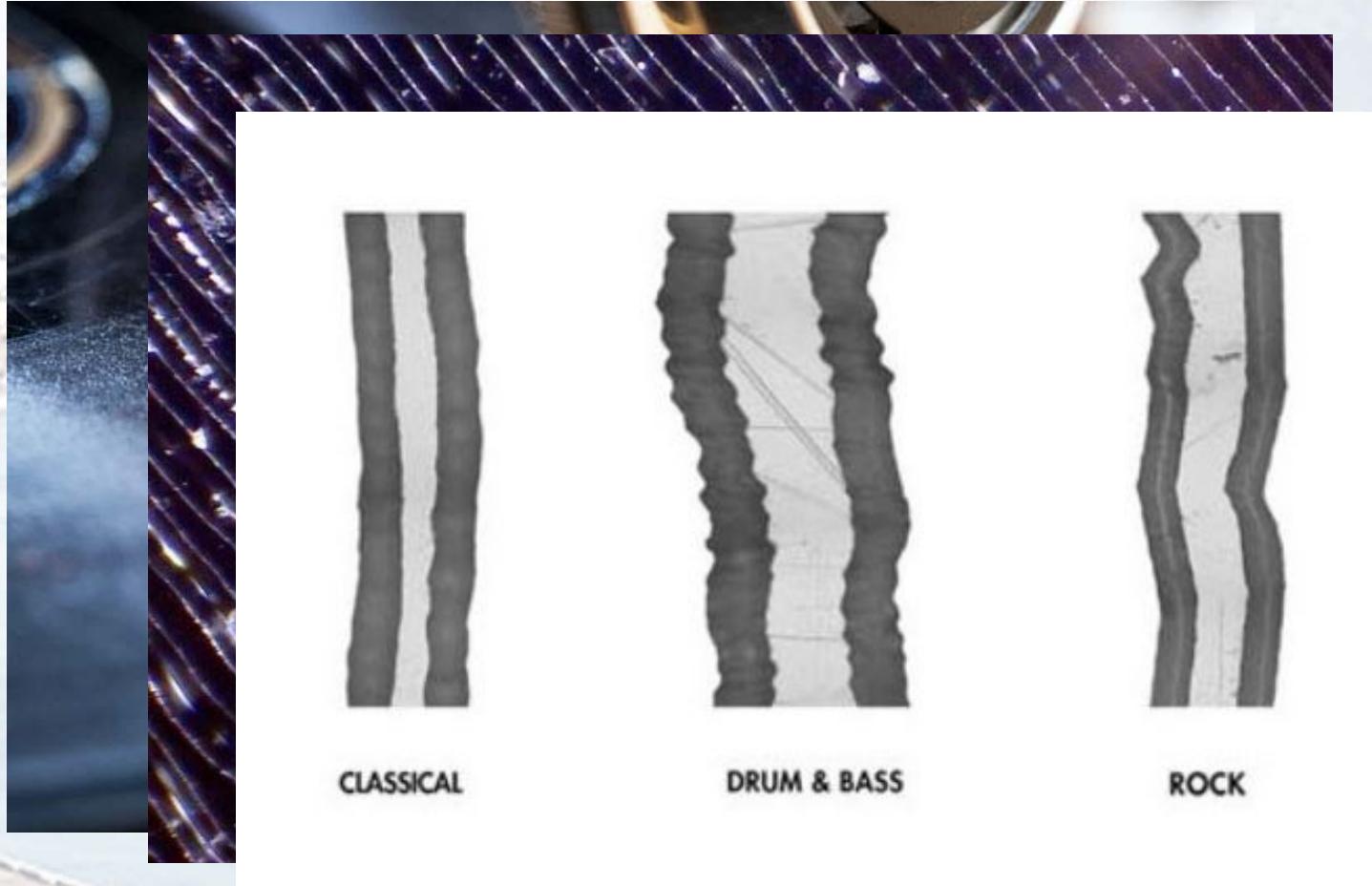
walter.balzano@gmail.com

Premessa L'era digitale



walter.balzano@gmail.com

Premessa L'era digitale



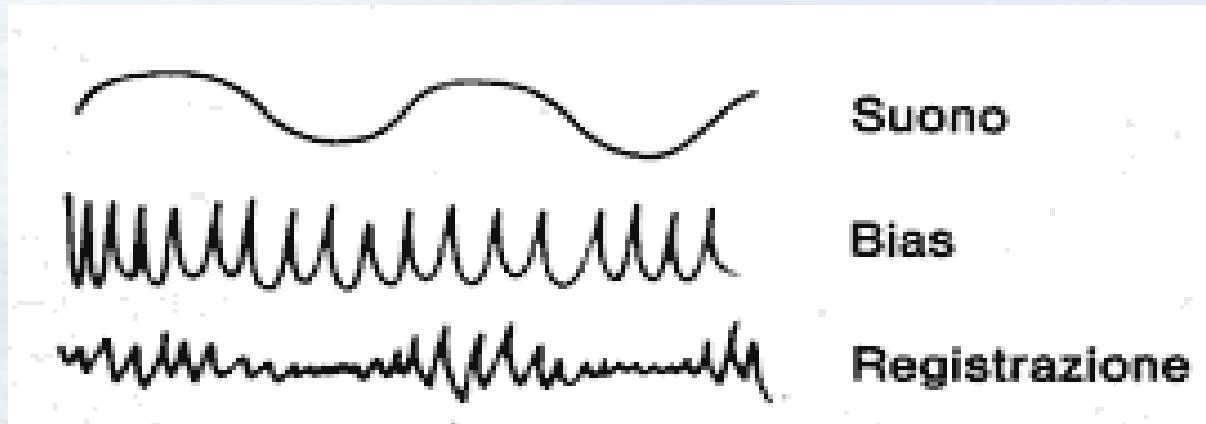
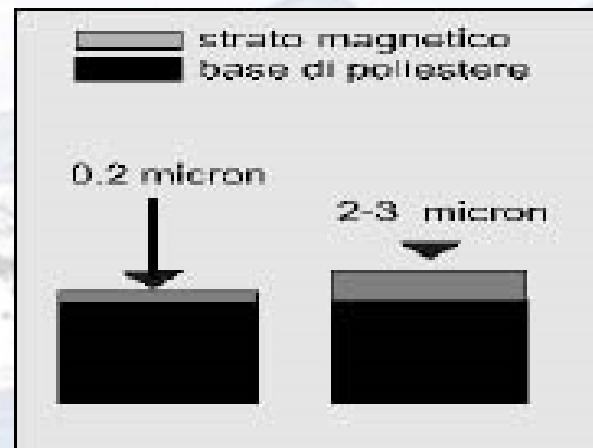
walter.balzano@gmail.com

Premessa L'era digitale



walter.balzano@gmail.com

Premessa L'era digitale



walter.balzano@gmail.com



walter.balzano@gmail.com

Premessa L'era digitale

Analogico



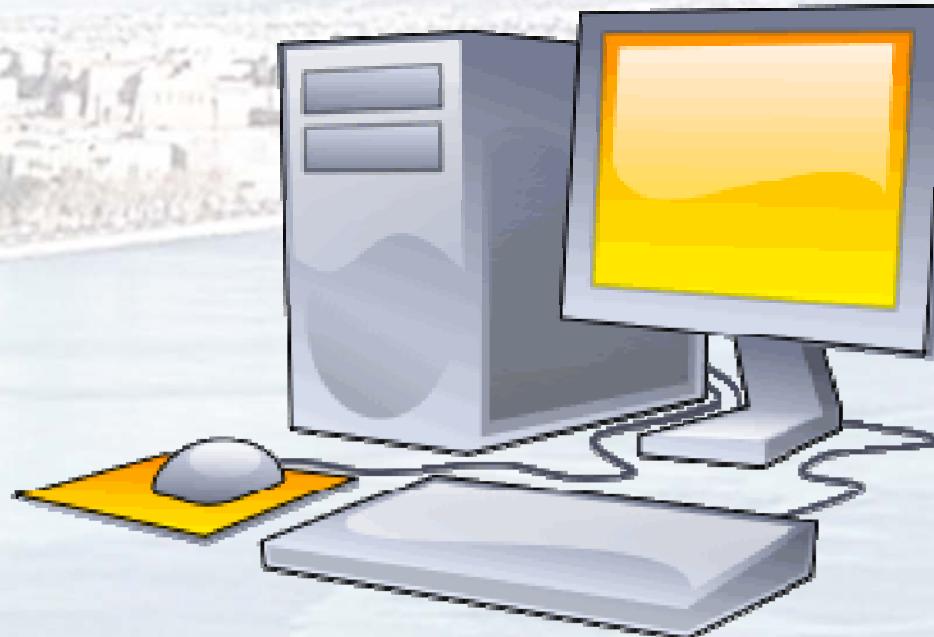
Di gi tale



walter.balzano@gmail.com

Premessa L'era digitale

Fase 1 → digitalizzazione dei dati



...10101100100100....



walter.balzano@gmail.com

Premessa L'era digitale

Fase 1 → digitalizzazione dei dati



walter.balzano@gmail.com

Premessa L'era digitale

Fase 1 → digitalizzazione dei dati



walter.balzano@gmail.com

Premessa L'era digitale

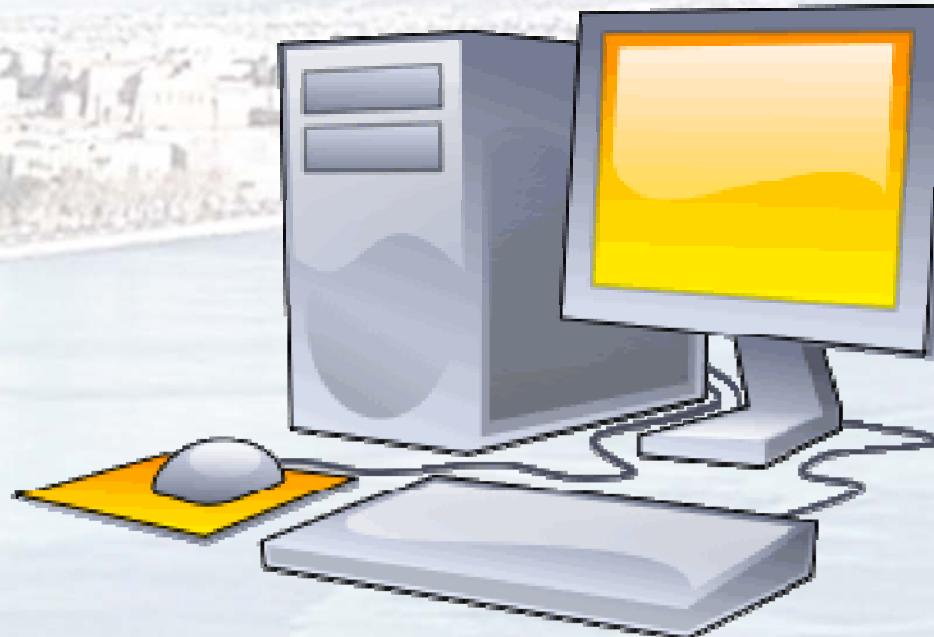
Fase 1 → digitalizzazione dei dati



walter.balzano@gmail.com

Premessa L'era digitale

Fase 1 → digitalizzazione dei dati



...10101100100100....



walter.balzano@gmail.com

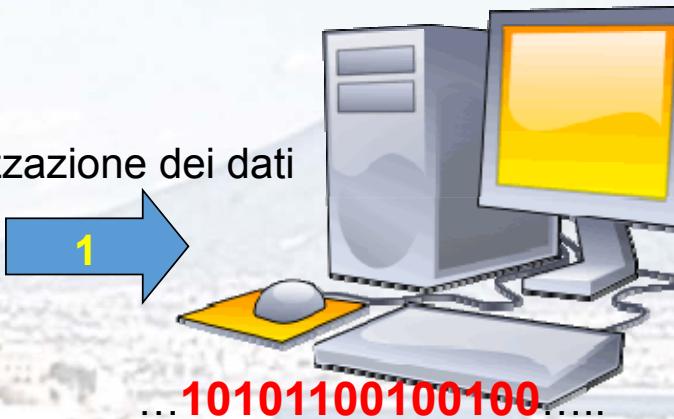
Premessa



walter.balzano@gmail.com

Premessa

Fase 1 → digitalizzazione dei dati



Fase 2 → Management dei dati



walter.balzano@gmail.com

Definizioni ed Evoluzione dei D.B.

	DBMS (DataBase Management Systems)	MMDBMS (MultiMedia DBMS)
DOMINIO	AlfaNumerico	MultiMediale
MATCH	Testuale esatto	Similitudine Caratteristiche
PRESTAZIONI	■ Velocità di risposta	■ Velocità di risposta ■ Capacità di confronto caratteristiche

M.I.R.S.
Multimedia
Indexing and
Retrieval
Systems



Query

- Qual è la data di nascita di Giuseppe Garibaldi?
- Quali sono i titoli dei Film in cui Robert De Niro ha una camicia verde?
- Qual è il titolo del brano musicale di cui possiedo solo un file mp3 senza altre info?
- Nel mio D.B., quali sono le fotografie in cui piove?
- Nel mio D.B., quali sono le fotografie in cui c'è un cane?
- In quale video ci sono persone che si muovono in modo circospetto?

Alcune tipologie di Query:

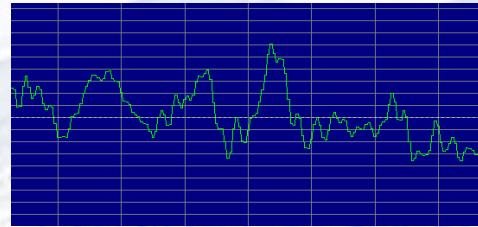
- Spazio/Temporali
- Similitudine



Query – 2

(Query per similitudine)

- (A) •Ricerca il brano musicale in cui
•ci sono 3 colpi di tamburo



(Query per concetto)

- (B) Cercare tutte le sequenze dove compare Maradona che scatta.



(Query spazio/tempo)

- (C) Cercare tutte le sequenze in cui nell'area di rigore Careca effettua un passaggio a Maradona.



- (D) **Combinazione più o meno complesse di (A), (B) e (C)**



Media & Multimedia

- **Media:** tipi di informazione o rappresentazione come dati alfanumerici, immagini, audio, video
- Una comune classificazione è basata su:
 - Formati fisici
 - Relazioni con dimensione temporale
 - Statici:
 - Essi non possiedono una dimensione temporale e quindi il loro contenuto ed il loro significato non ha alcuna dipendenza dal tempo. Esempi di Media statici sono i grafici e le immagini fotografiche
 - Dinamici:
 - Possiedono una dimensione temporale ed il loro significato e la loro correttezza dipende dalla velocità con cui vengono rappresentati. Esempi di Media dinamici sono le animazioni, l'audio ed i video.
 - **MultiMedia:** Raccolta di tipi di media usati contemporaneamente



RDBMS

- Gli **RDBMS** (**Relational DBMS**)
- costituiscono la tipologia maggiormente diffusa di DBMS
- Struttura tabulare (Righe x Colonne):
 - Righe** → Elementi di informazione o RECORD
 - Colonne** → Attributi o CAMPI

Stu#	Name	Address
10	Lew, Tom	2 Main St. Churchill, australia



SQL

- L' **SQL** (Structured Query Language)
- è utilizzato per la creazione delle tabelle, per l'inserimento ed il reperimento degli elementi dal DataBase.

```
select name  
from STUDENT  
where stu#=32
```

Il **BLOB** (Binary Large Objects)

è una stringa binaria di lunghezza variabile che si utilizza per la memorizzazione di dati in formato binario come immagini.

Nota: sul BLOB non è possibile formulare un pattern di ricerca.

```
create table STUDENT (  
    stu# integer,  
    name char(20),  
    address char(100),  
    picture BLOB);
```



OODBMS

- **OODBMS (Object-Oriented DBMS)**
- È un tipo di DBMS orientato ad Oggetti;
- Esso combina le capacità dei DataBase con le note proprietà degli Object Oriented (incapsulamento, eredità,...)

```
create type IMAGE (
private
    size integer,
    resolution integer,
    content float[]],
public
    ...
);
```

```
create table STUDENT (
    stu# integer,
    name char(20),
    address char(100),
    picture IMAGE);
```

NOTA: La differenza sostanziale tra un BLOB ed un “oggetto” è che l’ “oggetto” è propriamente definito e consente pertanto operazioni sulle sue proprietà (il BLOB invece no).



IR

- **IR (Information Retrieval)**

- o Costituiscono una parte rilevante tra i sistemi di recupero delle informazioni.

- o Essi operano prevalentemente in **modalità testuale**, ma possono essere ugualmente impiegati in ambito multimediale poiché:

- Occorre prendere atto della gran mole di documenti testuali già esistenti
 - Mediante il testo è possibile fornire una descrizione (**Annotazione**) ai vari media (Audio, Immagini, Video) per permetterne, così, le successive consultazioni.

- **Limiti degli IR:**

- o L'Annotazione è per la gran parte manuale (e quindi dispendiosa in tempo);

- o L'Annotazione è incompleta e soggettiva;

- o Le query sono esclusivamente testuali;

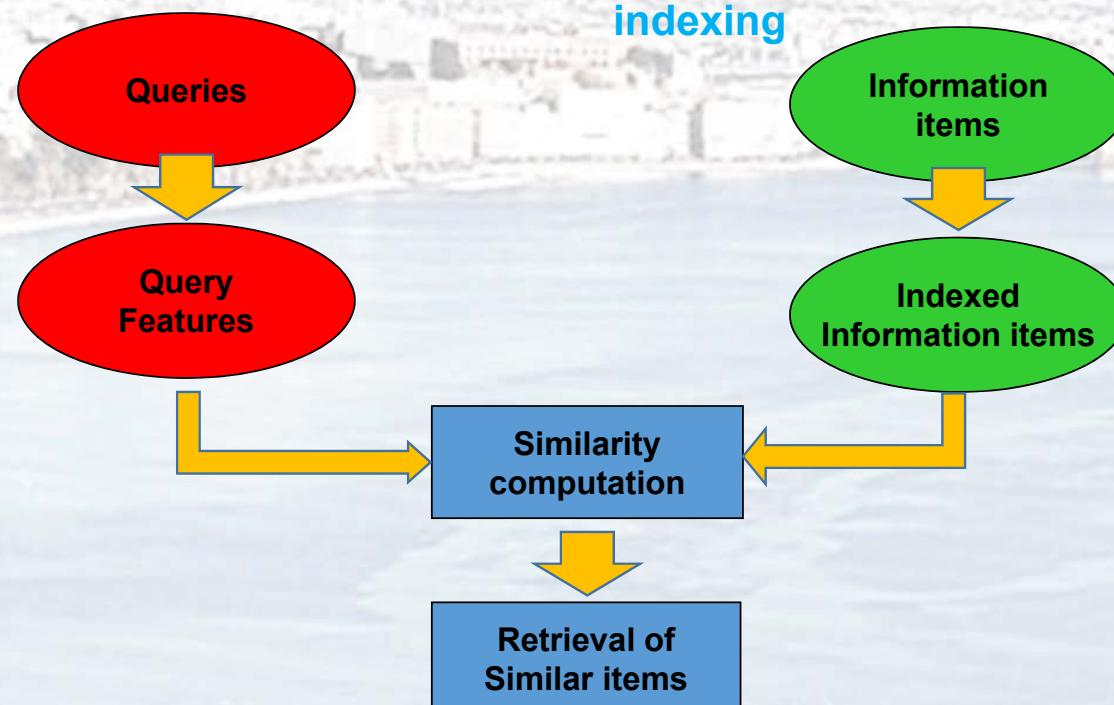
- o Molti oggetti sono difficili da descrivere (Textures, Oggetti di forme varie,...).



MIRS: schema di funzionamento

MIRS (Multimedia Indexing and Retrieval Systems)

Processing and
Feature extraction



walter.balzano@gmail.com

MIRS: Proprietà

- Alcune delle proprietà attese in un MIRS sono:

- o Query su **METADATI**

- Es.: nome_autore, data_creazione,...

- o Query su **ANNOTAZIONI**

- Es.: Con riferimento al testo descrittivo (inserito manualmente) relativo all'oggetto considerato

- o Query su **MODELLI di DATI** o **CARATTERISTICHE**

- Es.: Con riferimento ad una caratteristica come la distribuzione del colore, ...

- o Query su **ESEMPI**

- Es.: ricercare un oggetto multimediale, specificando un oggetto similare (come un'immagine, un suono, un disegno).

- o Query su **APPLICAZIONI SPECIFICHE**

- Es.: ricercare qualcosa mediante informazione molto specifiche e dettagliate, come la grandezza di un oggetto oppure l'età di una persona



MIRS: Applicazioni

o Medicina

- Es.: Confronto dell'ecografia di un paziente con tutto il D.B. delle ecografie.

o Sicurezza

- Es.: Riscontro della foto di una persona con un D.B. di ricercati

o Didattica

- Es.: ricercare il nome di un animale di cui si possiede solo una foto e/o registrazione del suo verso

o Editoria

- Es.: ricercare una persona di cui non si conosce il nome ma l'immagine, tra i giornali, foto e filmati degli ultimi 20 anni

o Intrattenimento

- Es.: cercare in un vasto D.B. un video divertente come quello appena visto

o Copyright

- Es: cercare in un D.B. di brevetti eventuali oggetti simili a quello che dovrebbe essere brevettato.



walter.balzano@gmail.com

Tipologia e Formati dei Dati MultiMediali

1. Introduzione
2. Il Testo (testo Piano, Strutturato, Compresso,...)
3. Grafica Vettoriale ed Animazione
4. Audio
5. Immagini Digitali
6. Il Video Digitale
7. Standards per Documenti MultiMediali Complessi
8. Caratteristiche Principali e Requisiti delle Applicazioni e dei dati MultiMediali



walter.balzano@gmail.com

- Obiettivi

- Elaborare
- Indicizzare
- Memorizzare
- Trasmettere
- Presentare
-

- Dominio

- Testi
- Immagini
- Grafici
- Audio
- Video
-



walter.balzano@gmail.com

Introduzione

- L'obiettivo principale di un MIRS è quello di **Indicizzare** e **Ricercare** i dati Multimediali tra cui testi, grafica, immagini, audio, video
- Un **MIRS** è quindi molto diverso da un **DBMS**; per la progettazione di un MIRS è pertanto necessario conoscere la struttura, le caratteristiche e le peculiarità dei dati multimediali. In particolare:
 - In che modo vengono memorizzati i dati;
 - Le tecniche standard di compressione;
 - Modalità e complessità del processo di estrazione ed indicizzazione delle caratteristiche
 - Requisiti di memorizzazione;
 - Requisiti di comunicazione;
 - Requisiti per la presentazione.



walter.balzano@gmail.com

Il Testo - 1

IL TESTO PIANO

- La gran parte delle informazioni sono codificate con testo di caratteri alfanumerici e sono tradizionalmente rappresentati mediante il **codice ASCII** (American Standard Code for Information Interchange – Bob Berner, IBM, 1961).
- Inizialmente era codificato a 7 bit ma poi è stato esteso ad 8 bit (**Extended ASCII**) per rappresentare caratteri quali accentate, simboli semigrafici,... Spesso l'8 bit è impiegato per il controllo di parità.
- I **caratteri stampabili** vanno dal 32 al 126 e gli altri sono caratteri per il controllo.
- Successivamente sono state definite ulteriori estensioni come **l'UNICODE** (1991) per la rappresentazione di caratteri per esempio dell'alfabeto greco, cirillico,....
- Anche l'**UNICODE è stato esteso**: si è passati da una codifica a 16 bit (65.536 caratteri) ad una codifica con 21 bit (circa 1 milione di caratteri). Nella pratica, comunque, è realmente utilizzato solo un sottoinsieme di tali caratteri.



walter.balzano@gmail.com

Il Testo - 2

- Tabelle Ascii

Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	
0 00	00	Null	32 20	20	Space	64 40	40	Ø	96 60	60	`	128 80	80	ç	160 A0	A0	á	192 C0	C0	ł	
1 01	01	Start of heading	33 21	21	!	65 41	41	À	97 61	61	à	129 81	81	ù	161 A1	A1	í	193 C1	C1	ł	
2 02	02	Start of text	34 22	22	"	66 42	42	฿	98 62	62	b	130 82	82	é	162 A2	A2	ó	194 C2	C2	Γ	
3 03	03	End of text	35 23	23	#	67 43	43	₵	99 63	63	c	131 83	83	â	163 A3	A3	ú	195 C3	C3	†	
4 04	04	End of transmit	36 24	24	\$	68 44	44	฿	100 64	64	d	132 84	84	ã	164 A4	A4	ñ	196 C4	C4	-	
5 05	05	Enquiry	37 25	25	%	69 45	45	£	101 65	65	e	133 85	85	à	165 A5	A5	Ñ	197 C5	C5	+	
6 06	06	Acknowledge	38 26	26	&	70 46	46	₣	102 66	66	f	134 86	86	â	166 A6	A6	¤	198 C6	C6	¶	
7 07	07	Audible bell	39 27	27	'	71 47	47	₲	103 67	67	g	135 87	87	ç	167 A7	A7	°	199 C7	C7	॥	
8 08	08	Backspace	40 28	28	(72 48	48	₭	104 68	68	h	136 88	88	è	168 A8	A8	ç	200 C8	C8	₪	
9 09	09	Horizontal tab	41 29	29)	73 49	49	₭	105 69	69	i	137 89	89	ë	169 A9	A9	߱	201 C9	C9	߱	
10 0A	0A	Line feed	42 2A	2A	*	74 4A	4A	₭	106 6A	6A	j	138 8A	8A	è	170 AA	AA	߱	202 CA	CA	߱	
11 0B	0B	Vertical tab	43 2B	2B	+	75 4B	4B	₭	107 6B	6B	k	139 8B	8B	ି	171 AB	AB	ି	203 CB	CB	ି	
12 0C	0C	Form feed	44 2C	2C	,	76 4C	4C	₭	108 6C	6C	l	140 8C	8C	ି	172 AC	AC	ି	204 CC	CC	ି	
13 0D	0D	Carriage return	45 2D	2D	-	77 4D	4D	ℳ	109 6D	6D	m	141 8D	8D	ି	173 AD	AD	ି	205 CD	CD	=	
14 0E	0E	Shift out	46 2E	2E	.	78 4E	4E	₦	110 6E	6E	n	142 8E	8E	ܲ	174 AE	AE	«	206 CE	CE	ܲ	
15 0F	0F	Shift in	47 2F	2F	/	79 4F	4F	ଓ	111 6F	6F	o	143 8F	8F	ܲ	175 AF	AF	»	207 CF	CF	ܲ	
16 10	10	Data link escape	48 30	30	Ø	80 50	50	₱	112 70	70	p	144 90	90	ܲ	176 BO	BO	ܲ	208 DO	DO	ܲ	
17 11	11	Device control 1	49 31	31	1	81 51	51	ܲ	113 71	71	q	145 91	91	ܲ	177 B1	B1	ܲ	209 D1	D1	ܲ	
18 12	12	Device control 2	50 32	32	2	82 52	52	ܲ	114 72	72	r	146 92	92	ܲ	178 B2	B2	ܲ	210 D2	D2	ܲ	
19 13	13	Device control 3	51 33	33	3	83 53	53	ܲ	115 73	73	s	147 93	93	ܲ	179 B3	B3	ܲ	211 D3	D3	ܲ	
20 14	14	Device control 4	52 34	34	4	84 54	54	ܲ	116 74	74	t	148 94	94	ܲ	180 B4	B4	ܲ	212 D4	D4	ܲ	
21 15	15	Neg. acknowledge	53 35	35	5	85 55	55	ܲ	117 75	75	u	149 95	95	ܲ	181 B5	B5	ܲ	213 D5	D5	ܲ	
22 16	16	Synchronous idle	54 36	36	6	86 56	56	ܲ	118 76	76	v	150 96	96	ܲ	182 B6	B6	ܲ	214 D6	D6	ܲ	
23 17	17	End trans. block	55 37	37	7	87 57	57	ܲ	119 77	77	w	151 97	97	ܲ	183 B7	B7	ܲ	215 D7	D7	ܲ	
24 18	18	Cancel	56 38	38	8	88 58	58	ܲ	120 78	78	x	152 98	98	ܲ	184 B8	B8	ܲ	216 D8	D8	ܲ	
25 19	19	End of medium	57 39	39	9	89 59	59	ܲ	121 79	79	y	153 99	99	ܲ	185 B9	B9	ܲ	217 D9	D9	ܲ	
26 1A	1A	Substitution	58 3A	3A	:	90 5A	5A	ܲ	122 7A	7A	z	154 9A	9A	ܲ	186 BA	BA	ܲ	218 DA	DA	ܲ	
27 1B	1B	Escape	59 3B	3B	;	91 5B	5B	[123 7B	7B	{	155 9B	9B	ܲ	187 BB	BB	ܲ	219 DB	DB	ܲ	
28 1C	1C	File separator	60 3C	3C	<	92 5C	5C	\	124 7C	7C		156 9C	9C	ܲ	188 BC	BC	ܲ	220 DC	DC	ܲ	
29 1D	1D	Group separator	61 3D	3D	=	93 5D	5D]	125 7D	7D	}	157 9D	9D	ܲ	189 BD	BD	ܲ	221 DD	DD	ܲ	
30 1E	1E	Record separator	62 3E	3E	>	94 5E	5E	^	126 7E	7E	~	158 9E	9E	ܲ	190 BE	BE	ܲ	222 DE	DE	ܲ	
31 1F	1F	Unit separator	63 3F	3F	?	95 5F	5F	_	127 7F	7F	ܲ	159 9F	9F	ܲ	191 BF	BF	ܲ	223 DF	DF	ܲ	
																			255 FF	FF	ܲ



walter.balzano@gmail.com

Il Testo - 3

Il TESTO STRUTTURATO

- La gran parte dei testi sono “**testi formattati**”: Il titolo del documento si presenta diversamente dai sottotitoli che, a loro volta, si presentano diversamente dagli stessi paragrafi,....
- Esistono **diversi Standard** che rappresentano i diversi formati, tra cui: Formato Microsoft Word (est. **.doc**), Formato Portable Document Format (est. **.pdf**), Formato Latex (est. **.tex**),.....
- Normalmente l'intestazione del file dati contiene le specifiche di formato. Quando il formato è noto, è possibile usare le strutture del documento per estrarre informazioni per la ricerca: le informazioni che compaiono nel titolo sono per esempio più importanti delle informazioni contenute in un paragrafo.

Percepiva la pensione del padre morto da tre anni

Arrestato uomo di 47 anni a Camerota (Salerno)(ANSA)

CAMEROTA (SALERNO), 19 FEB 2009

Percepiva la pensione del padre morto da tre anni. Protagonista della truffa un uomo di 47 anni di Camerota. Nonostante il padre fosse deceduto tre anni fa, il figlio continuava ad incassare la sua pensione: ha intascato indebitamente quasi 40mila euro. Alla fine, e' stato scoperto dalla Guardia di Finanza ed e' stato denunciato insieme col responsabile dell'ufficio anagrafe per il reato di truffa aggravata per omissione di comunicazione del decesso all'Inps.



walter.balzano@gmail.com

Il Testo - 4

IL TESTO COMPRESSO

- Sebbene tra i vari Media (audio, video,...), il testo richiede meno spazio per la sua memorizzazione, conviene intervenire con algoritmi di compressione soprattutto quando aumenta il numero dei testi da archiviare.
- La compressione effettuata sul testo è **LOSSLESS** (senza perdita): la decompressione ottiene esattamente l'oggetto originario.
- La compressione sul testo è giustificata dal fatto che:
 - Alcuni caratteri appaiono più frequentemente rispetto ad altri.
 - Alcuni gruppi di caratteri appaiono molto frequentemente.
- Metodi noti impiegati per la compressione:
 - **Huffman**
 - **Run length**
 - **Lempel-Ziv-Welch (LZW)**



walter.balzano@gmail.com

- Fondamenti
- Descritto dallo studente David A. Huffman in un articolo pubblicato nel 1952.
- **E' basato sull'analisi statistica del dato da comprimere, in particolare sulla frequenza con la quale si ripetono i suoi elementi.**
- Ha una **prestazione proporzionata alla varianza delle frequenze** con cui compaiono i caratteri del testo da comprimere (maggiore varianza → maggiore prestazione)
- Utilizzabile in combinazione con altre tecniche.



walter.balzano@gmail.com

HUFFMAN – Esempio 1

Stringa S da comprimere

S_{txt} = CIAO_MAMMA

Codice ASCII

A	01000001
C	01000011
I	01001001
O	01001111
M	01001101
_	00100000

Formato non Compresso per la stringa S: 10 caratteri x 8 bit/carattere = 80 bit

S_{0,1} = 01000011 01001001 01000001 01001111 00100000 01001101 01000001 01001101 01001101 01000001

FASE 1: Analisi delle frequenze

Frequenze caratteri

A	C	I	O	M	_
3	1	1	1	3	1

Ordinamento decrescente

A	M	C	I	O	_
3	3	1	1	1	1

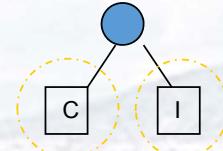


walter.balzano@gmail.com

HUFFMAN – Esempio 2

FASE 2: Raggruppamenti coppie con frequenza minore

C + I



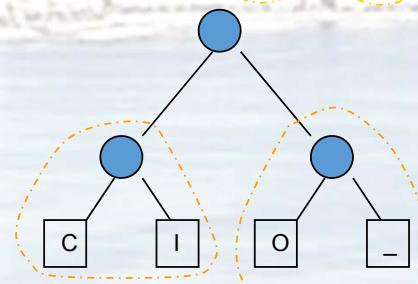
A	M	C+I	O	-
3	3	2	1	1

O + _



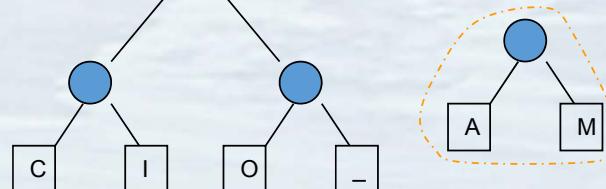
A	M	C+I	O+_
3	3	2	2

(C + I) + (O + _)



(C+I) + (O+_)	A	M
4	3	3

A + M

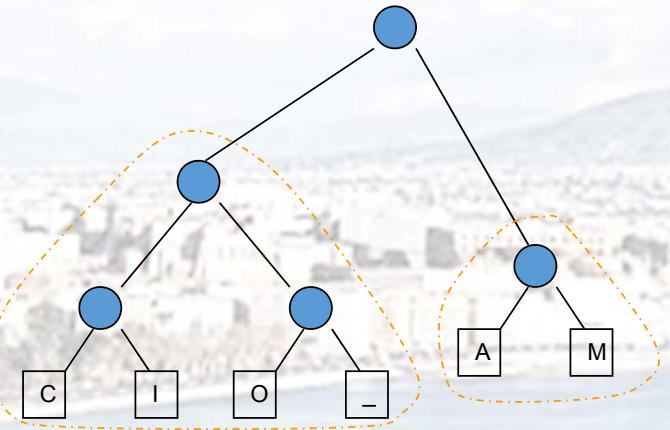


A+M	C+I + (O+_)
6	4



walter.balzano@gmail.com

HUFFMAN – Esempio 3



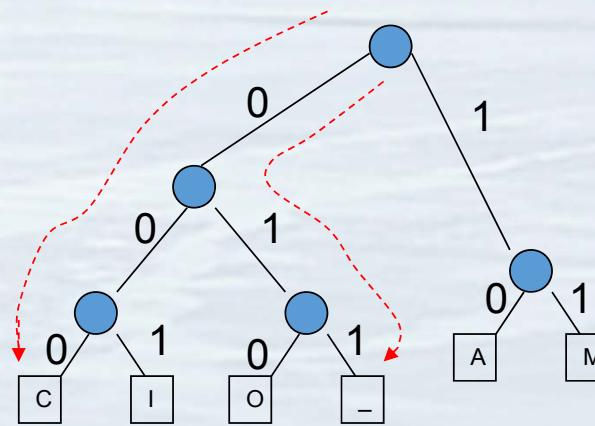
$[(C+I) + (O+_) + (A+M)]$
10

FASE 3: Costruzione del codice

Associazione dei codici agli archi:

0 → sinistra

1 → destra



walter.balzano@gmail.com

HUFFMAN – Esempio 4

Associazione codici \leftrightarrow caratteri

Costruisco il codice leggendo le etichette degli archi dalla radice ad ogni foglia (**CodeBook**):

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

$S' = 000\textcolor{magenta}{001}\textcolor{green}{10}010\textcolor{red}{011}\textcolor{blue}{111}\textcolor{green}{10}\textcolor{blue}{111}\textcolor{blue}{110}$ \longrightarrow 24 bit

Analisi:

- Compressione(S) = S'
- $|S| = 80$ bit
- $|S'| = 24$ bit
- Riduzione = $24/80 = 0,3 \rightarrow 30\%$; **Guadagno = 70%**
- Il codice ottenuto **non è univoco**.



Codifica RUN-LENGTH - 1

RUN-LENGTH è un metodo di compressione lossless che riduce le ripetizioni di caratteri, sostituendo un **RUN** (insieme di caratteri ripetuti) con il carattere che viene ripetuto e con la lunghezza del RUN

RUN: insieme di caratteri ripetuti;

Lunghezza RUN: lunghezza della ripetizioni

Rappresentazione del RUN :



Sc = carattere speciale per l'identificazione della codifica

X = il carattere che viene ripetuto nel RUN

C = numero di ripetizioni del carattere.

Prestazioni buone per run-length > 3



Codifica RUN-LENGTH - 2

Esempio:

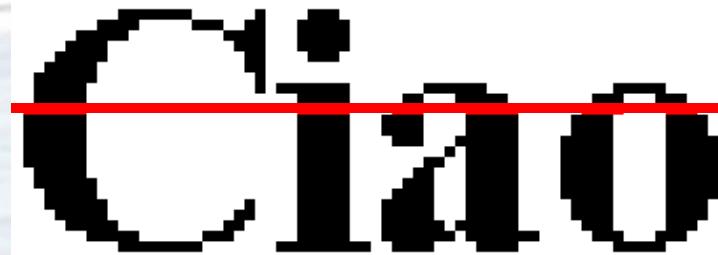
"eeeeeeetnnnnnnnn" → @e7t@n8

16

7

Esempio:

Ciao mamma



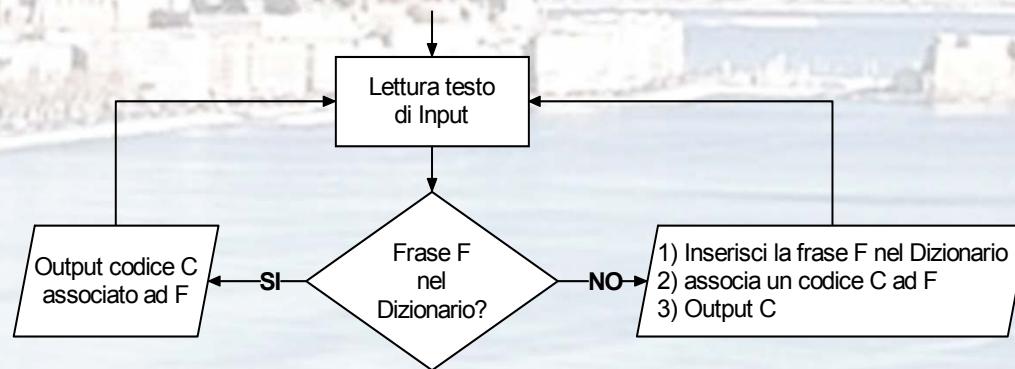
Applicazioni: FAX, immagini con pochi colori,...



walter.balzano@gmail.com

Codifica LZW

- LZW (Abraham **Lempel**, Jacob **Ziv**, Terry **Welch** - 1984) è un metodo di compressione lossless che sfrutta la ripetizione di gruppi di caratteri o frasi:
- Il compressore esamina la presenza delle frasi incontrante con le frasi presenti in un dizionario inizialmente vuoto:



- Prestazioni buone per input di testo con molte ripetizioni: linguaggio naturale (inglese, italiano,...).

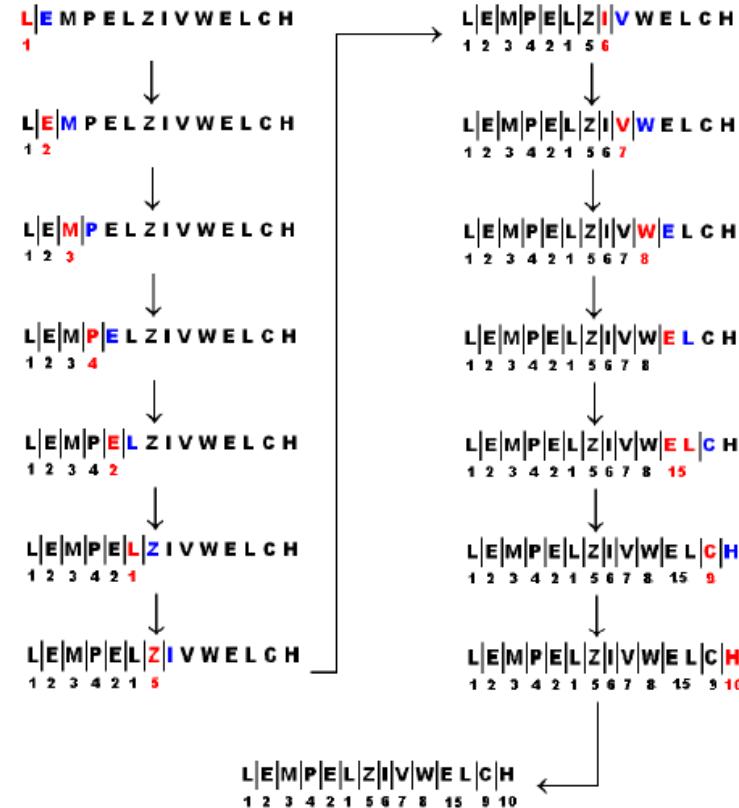


walter.balzano@gmail.com

Codifica LZW: Esempio

Stringa input = "LEMPZIVWELCH"

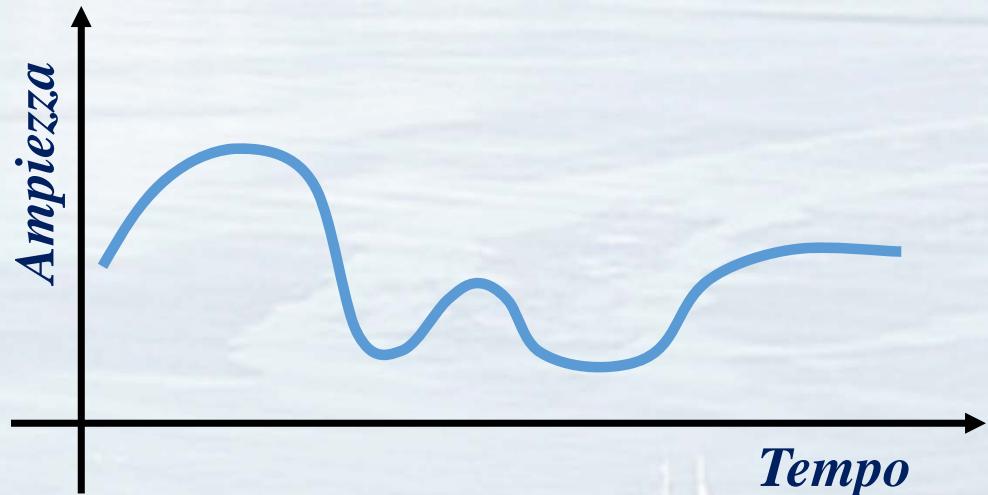
DIZIONARIO			
1	L	12	EM
2	E	13	MP
3	M	14	PE
4	P	15	EL
5	Z	16	LZ
6	I	17	ZI
7	V	18	IV
8	W	19	VW
9	C	20	WE
10	H	21	ELC
11	LE	22	CH



walter.balzano@gmail.com

L' AUDIO

- L'audio è causato da variazioni di pressione dell'aria
- Le frequenze udibili dall'uomo variano nell'intervallo **20 - 20.000 Hz**
- I descrittori fondamentali del suono sono:
 - **Aampiezza**
 - **Frequenza**
- **Aampiezza e Frequenza** del suono variano nel **tempo**



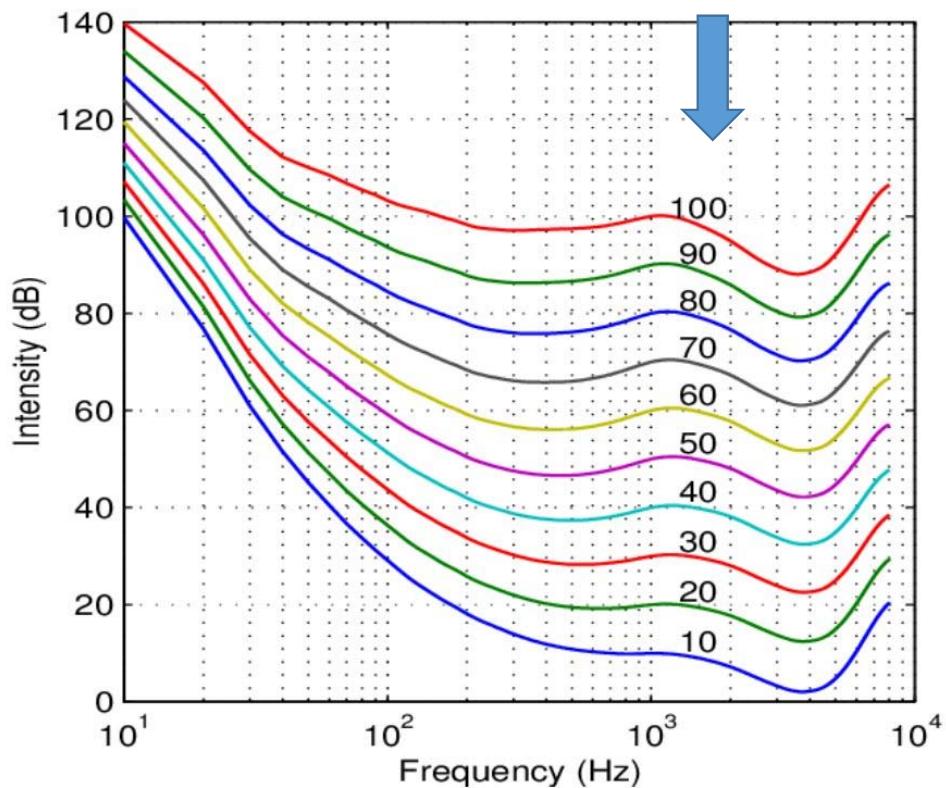
walter.balzano@gmail.com

- Le **capacità uditive dell'uomo non sono lineari** né rispetto all'ampiezza né rispetto alle frequenze
- Le **curve isofoniche** descrivono quali livelli sonori (misurati in dB) percepiamo essere uguali al variare della frequenza, cioè con ugual 'volume' (misurato in phon)
- La **maggior sensibilità è tra i 1.000 e i 5.000 Hz** (in questa regione ci vogliono meno dB per avere più phon): in questa banda di frequenza è concentrato buona parte del contenuto informativo del parlato e il nostro udito si è adeguato e sviluppato di conseguenza.



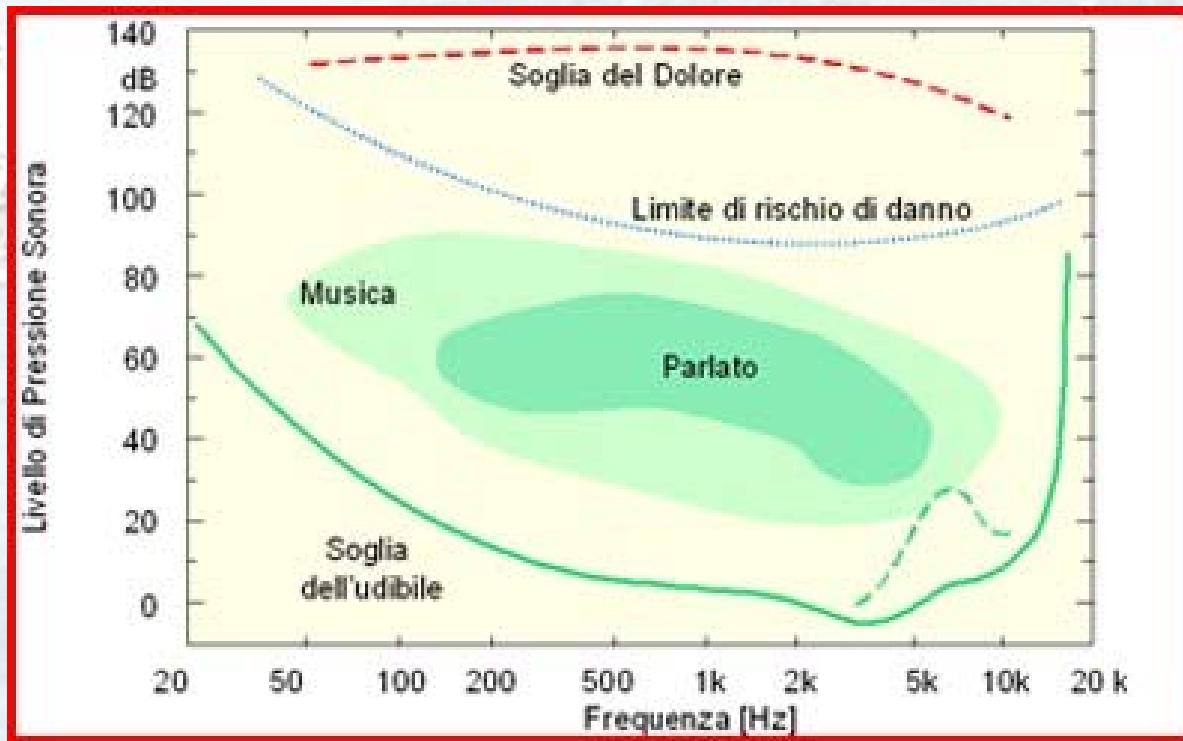
walter.balzano@gmail.com

Curve isofoniche *Fletcher e Munson*



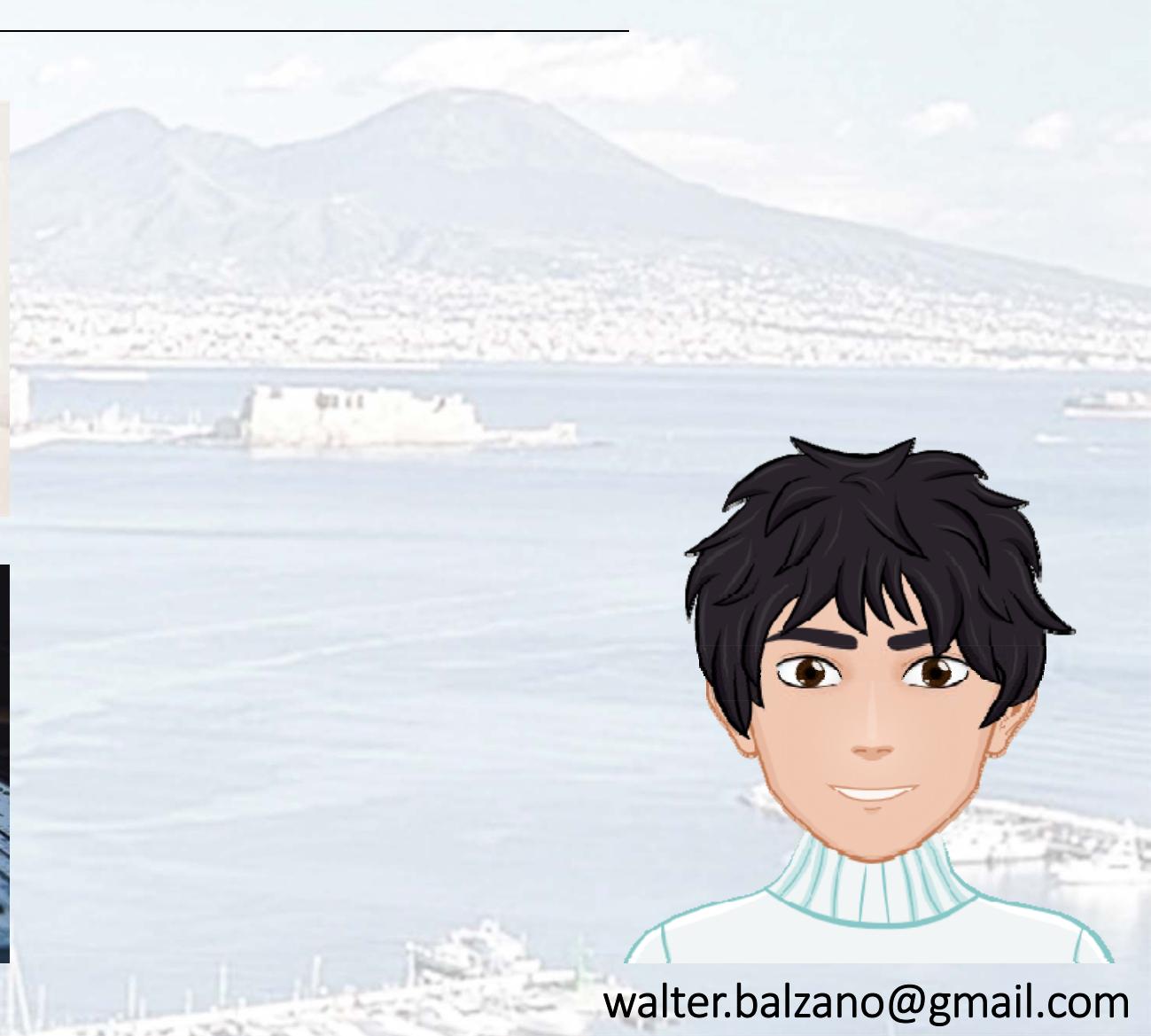
walter.balzano@gmail.com

Zone di percezione



walter.balzano@gmail.com

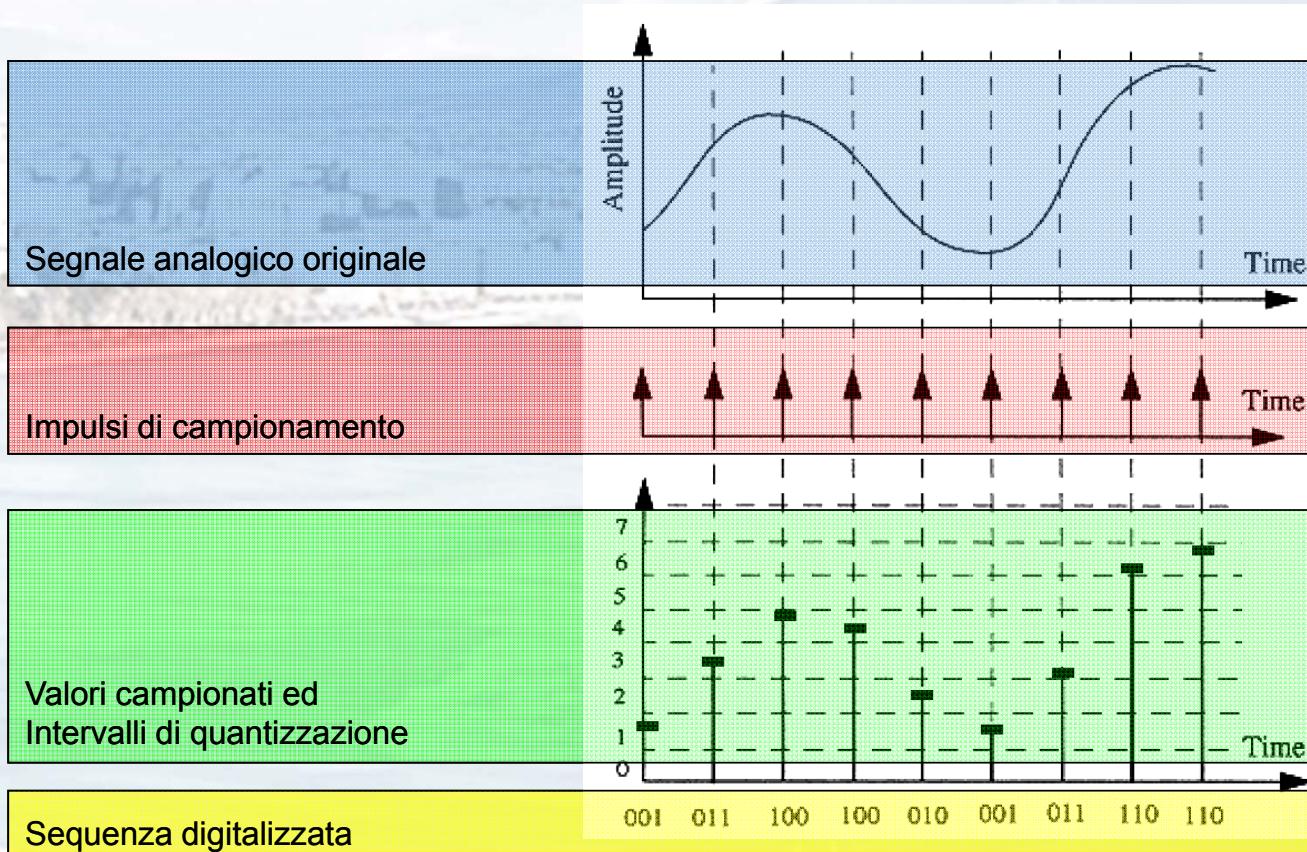
L' AUDIO



walter.balzano@gmail.com

Rappresentazione digitale dell'audio

Conversione **ANALOGICO → DIGITALE**



walter.balzano@gmail.com

ADC (Analog to Digital Conversion)

Le 3 fasi fondamentali di un processo ADC sono:

● CAMPIONAMENTO

Prelievo di valori assunti dal segnale analogico ad intervalli discreti di tempo

(gestiti da un clock). Il valore campionato resta costante durante il successivo intervallo di tempo. I campioni prelevati sono, in questa fase, ancora di tipo analogico assumendo pertanto un qualsiasi valore di un intervallo continuo.

● QUANTIZZAZIONE

Processo di conversione dei valori continui in valori discreti. L'intervallo del segnale viene suddiviso in un numero fisso di sotto-intervalli di uguale dimensione e viene assegnato un valore. Ciascun campione cade in uno specifico intervallo; quindi i valori possibili sono in numero limitato.

La grandezza del sotto-intervallo di quantizzazione è detto [passo di quantizzazione](#).

● CODIFICA

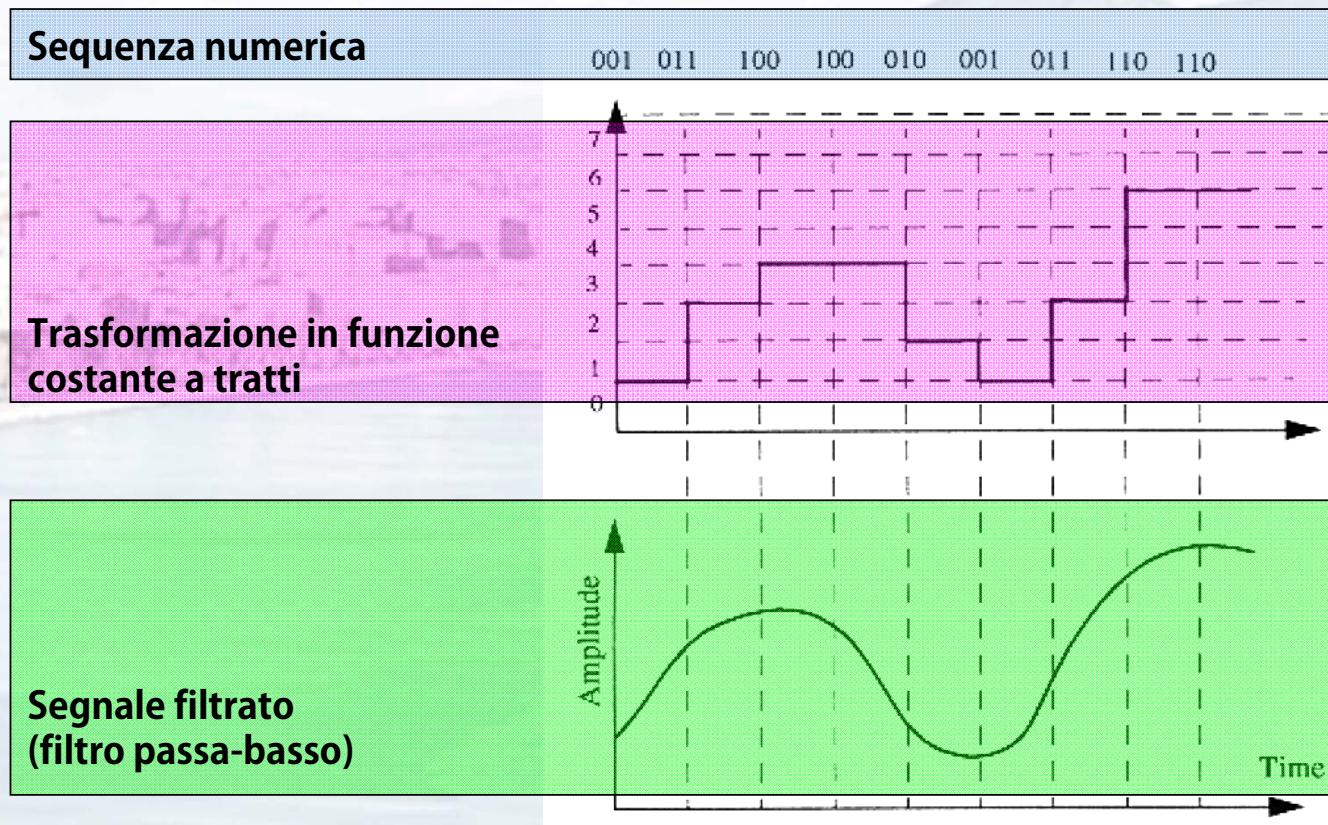
Processo di rappresentazione numerica dei valori quantizzati.

Quanto maggiore sarà sia la frequenza di campionamento sia il numero dei livelli di quantizzazione allora tanto maggiore sarà la fedeltà del segnale digitalizzato.



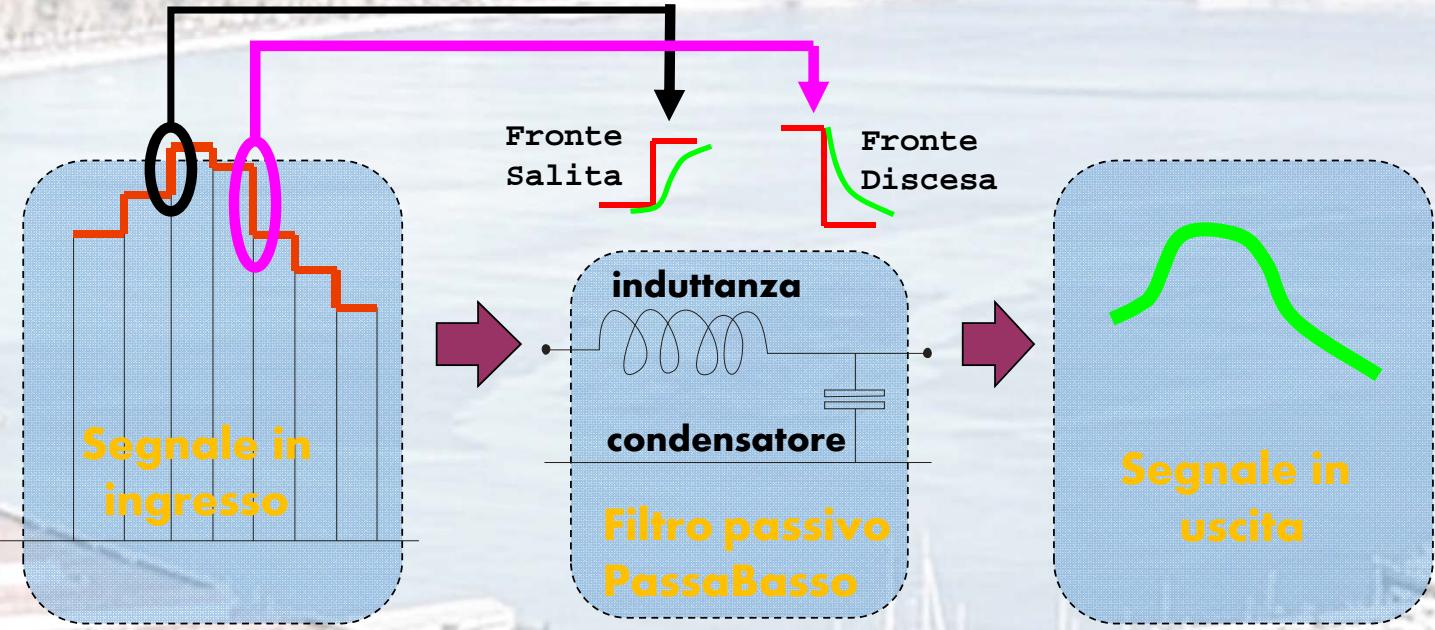
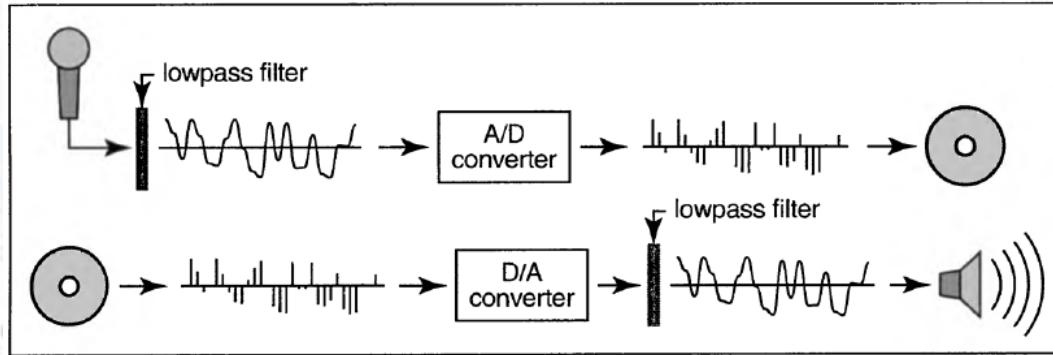
walter.balzano@gmail.com

DAC (Digital to Analog Conversion)



walter.balzano@gmail.com

ADC & DAC



walter.balzano@gmail.com

Scelta della Frequenza di Campionamento

- La *Frequenza di Campionamento* è strettamente dipendente dalla frequenza massima del segnale analogico da convertire;

infatti il **Teorema di Nyquist** afferma che:

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

- Campionamento Critico:

$$\text{Freq. Campionamento} = 2f \text{ Hz}$$



Scelta della Frequenza di Campionamento

Nella pratica le frequenze di campionamento sono di poco superiori al Campionamento critico.

Esempi:

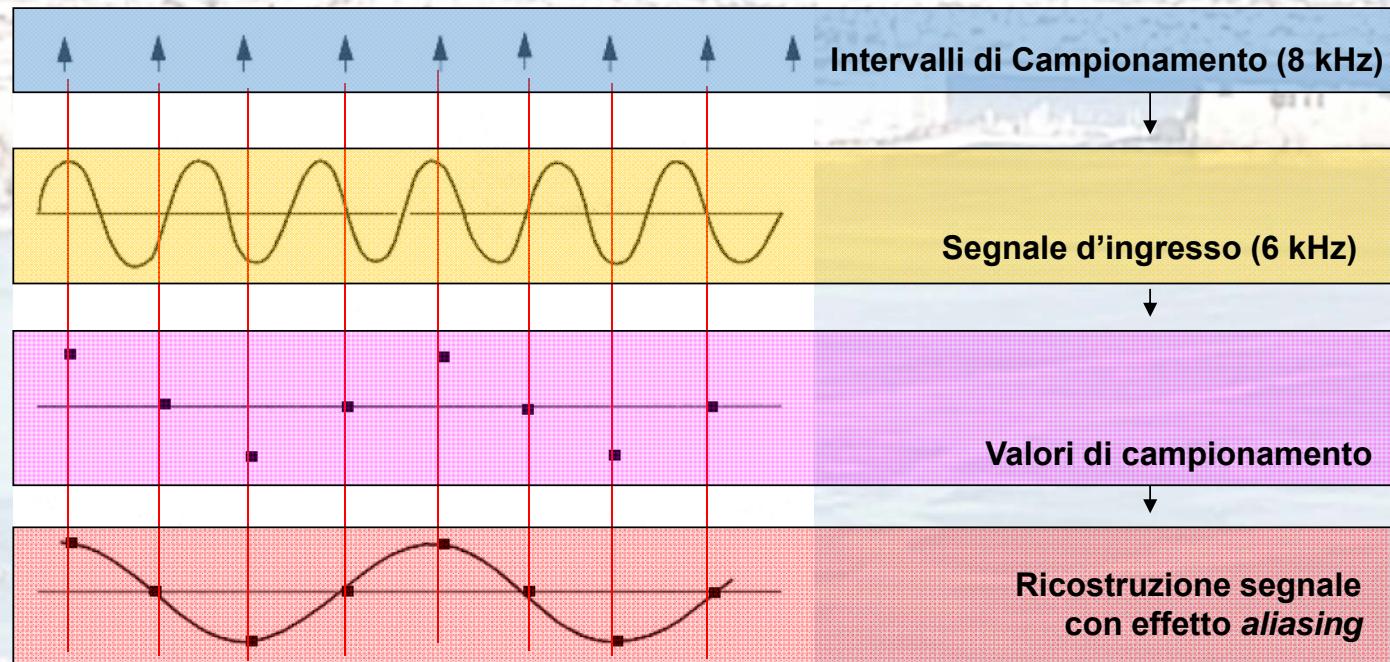
Dispositivo	Freq. Campionamento
CD-Audio	44.1 kHz
DAT (Digital Audio Tape)	48 kHz
Telefono	8 kHz



walter.balzano@gmail.com

Bassa Frequenza di campionamento

Esempio di **ricostruzione errata** per basso campionamento (effetto *aliasing*)



walter.balzano@gmail.com

Scelta numero dei livelli di Quantizzazione

Errore (o rumore) di quantizzazione:

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

- Il numero **Q** dei livelli di quantizzazione determina la quantità **b** di bit necessaria per rappresentare ciascun campione:

$$b = \log_2 Q$$

- La qualità del segnale digitale **SNR** (Signal Noise Ratio) viene misurata in decibel (db)

$$\text{SNR} = 20 \log_{10}(S/N) = 20b \log_{10}2 = 6b$$

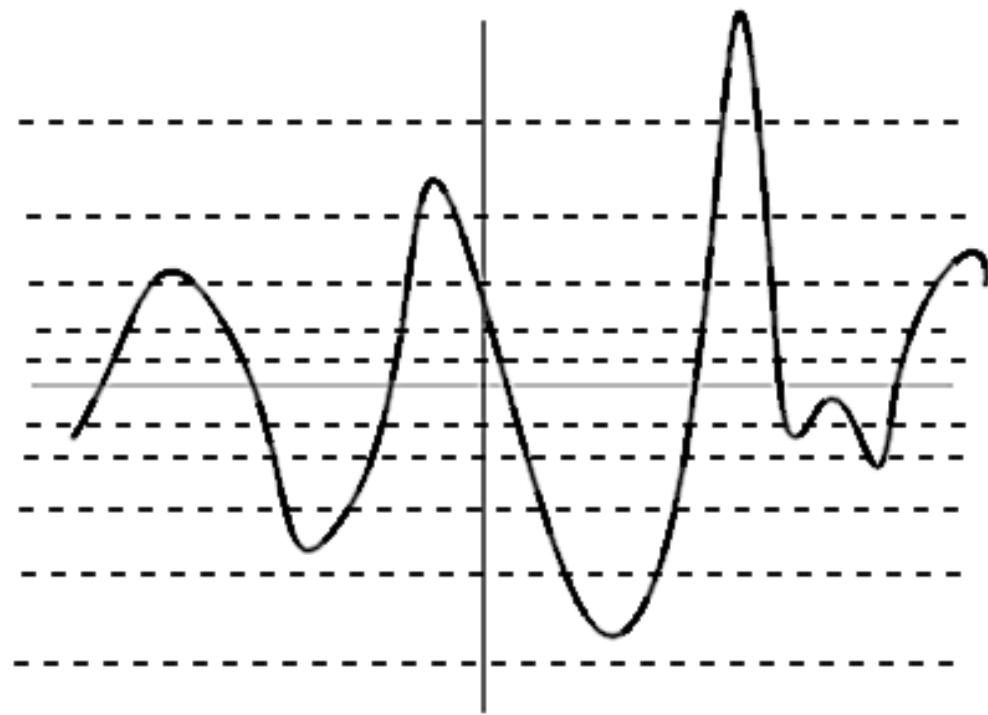
(S=max ampiezza segnale; N=Errore di quantizzazione; q=passo di quantizzazione; S=2^bq)

NOTA:

- Ogni bit in più usato per rappresentare il campione implica un aumento di 6 db del SNR
- Se l'errore di quantizzazione supera il valore della soglia uditiva allora viene avvertito.



Compressione Audio: Companding



walter.balzano@gmail.com

Compressione Audio: Companding

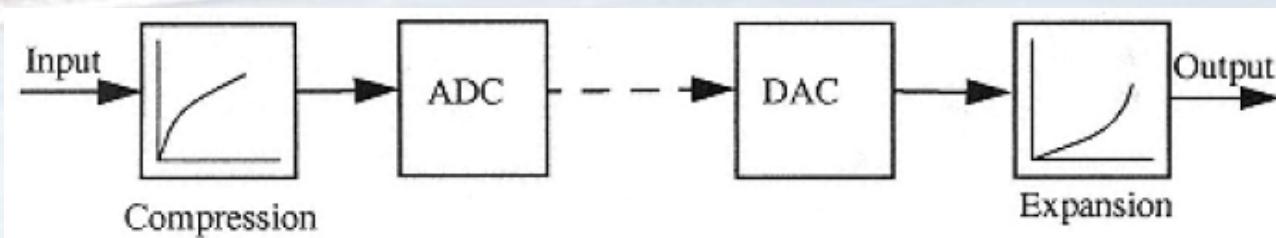
I metodi di compressione possono essere quindi basati su due approcci diversi:

- 1: Trasformazione non-lineare del segnale 2: Quantizzazione uniforme

- 1: Segnale Lineare 2: Quantizzazione non uniforme

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

- Companded PCM (*Log PCM*): Digitalizzazione uniforme di un segnale companded.



Schema della tecnica di companding

Mediante la tecnica del companding è possibile riprodurre un segnale ad 8-bit con la stessa qualità di un segnale a 12 bit e quindi il companding può quindi essere visto come una sorta di compressione analogica del segnale



walter.balzano@gmail.com

Predictive Coding

- Con il Predictive Coding, anziché codificare il valore del campione da trasmettere, si codifica **la differenza tra la predizione del valore del campione ed il valore del campione attuale** (*DPCM, differential pulse-coded modulation*)
- Il valore della predizione si ricava dai valori precedenti assunti dal segnale; tale valore è pertanto noto sia al codificatore che al decodificatore che applicano la medesima strategia.
- L'efficacia del Predictive Coding si basa sul fatto che:
 - **campioni vicini sono significativamente correlati.**
 - **per codificare una differenza occorre un numero inferiore di bit.**
- Se le differenze dovessero risultare molto grandi, vengono allora introdotte opportuni algoritmi correttivi



Mpeg Audio (Moving Picture Experts Group)

La tecnica di compressione Mpeg Audio è basata sul **mascheramento**: Suoni di maggiore intensità “coprono” i suoni a di bassa intensità che, pertanto, possono essere “ignorati” (compressi) senza influire sulla qualità udibile dal nostro orecchio.

Principali caratteristiche:

- E' una compressione a perdita (lossy)
- Frequenze di campionamento usate: 32 kHz, 44.1 kHz, 48 kHz.
- Supporto di 1 oppure 2 canali audio
- Bit rates del flusso compresso: da 32 a 244 kbps
- Rapporto di compressione: da 2.7:1 a 24:1 (dipende dalla frequenza di campionamento)
- Per rapporti di compressione fino a 6:1 la “perdita” non è percepibile neanche da “orecchie esperte”.

Lo standard Mpeg Audio rappresenta una famiglia di 3 schemi di compressione:

- Mpeg Layer 1 - Mpeg Layer 2 - Mpeg Layer 3
- La complessità dell'algoritmo impiegato è crescente dal Layer 1 al Layer 3
- Compatibilità verso il basso (Layer i è compatibile con Layer $i-1$)



Audio sintetico: Midi

- Con l'acronimo MIDI (Musical Instrument Digital Interface) si indica il protocollo standard per l'interazione degli strumenti musicali elettronici.
- **Il formato midi non contiene musica pre-registrata ma le direttive e le specifiche per sua riproduzione: uno spartito musicale non contiene la riproduzione del brano musicale ma solo le istruzioni per la riproduzione.** (*Un libro non contiene la narrazione diretta di una storia, bensì la storia stessa; la narrazione della storia dipende da fattori esterni quale il timbro vocale del narratore,...*)



walter.balzano@gmail.com

Audio sintetico: Midi

- Il file midi contiene quindi informazioni e direttive che vengono interpretate da sistemi speciali (sintetizzatori) Hardware o Software e ne realizzano l'esecuzione.
- Le direttive sono del tipo:

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

Possiamo quindi definire un file midi come una sorta di spartito musicale elettronico



- **Vantaggi:** Grandezze dei file molto ridotte (ottimo per essere trasmesso).
- **Svantaggi:** La riproduzione del suono non è univoca (l'esecuzione di uno stesso spartito musicale produce differenti interpretazioni in dipendenza degli strumenti utilizzati).



walter.balzano@gmail.com

Le immagini Digitali



**Percezione
Colori
Sintesi
Rappresentazioni
Modelli
Il video**



walter.balzano@gmail.com

La Percezione dei Colori

Immanuel Kant (Königsberg, 1724-1804) tratta il “fenomeno” (*Phainomenon*) percepito secondo quello che appare e non nella sua reale essenza

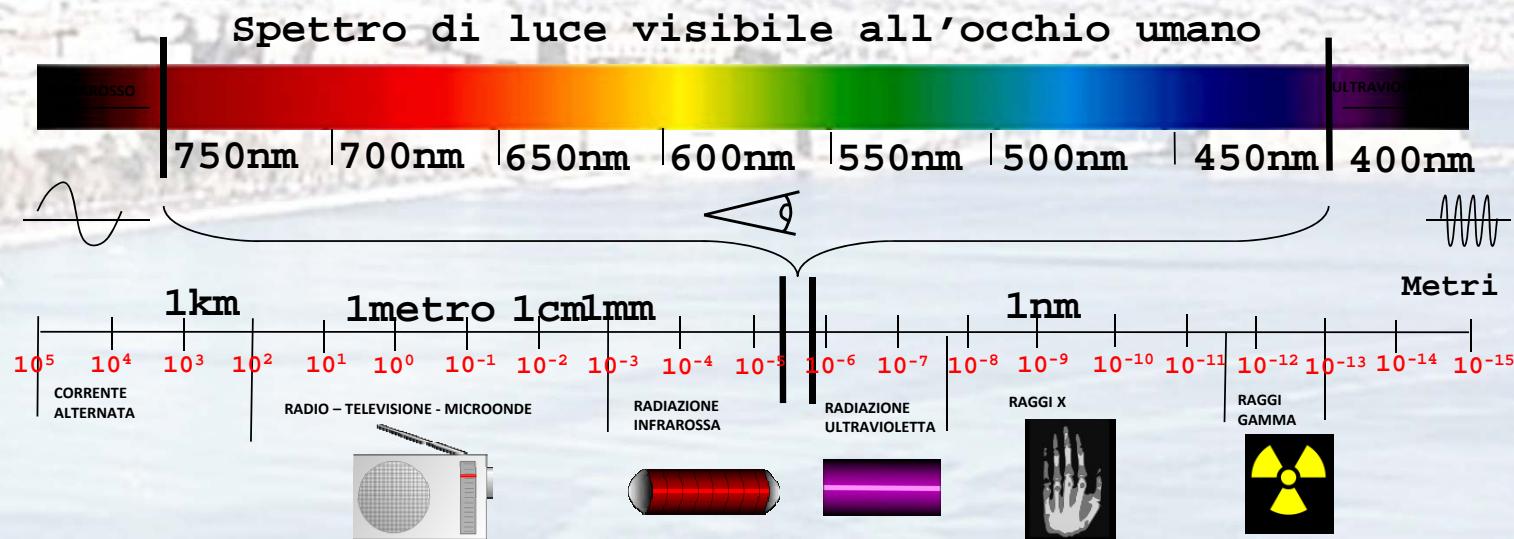


walter.balzano@gmail.com

Rappresentazione dei colori

Gli algoritmi di indicizzazione e di ricerca delle immagini devono essere basati su una corretta **Rappresentazione dei colori**.

La luce visibile è una radiazione elettromagnetica con lunghezza d'onda tra 400 e 780 nm.
Differenti lunghezze d'onda producono differenti sensazioni di colore

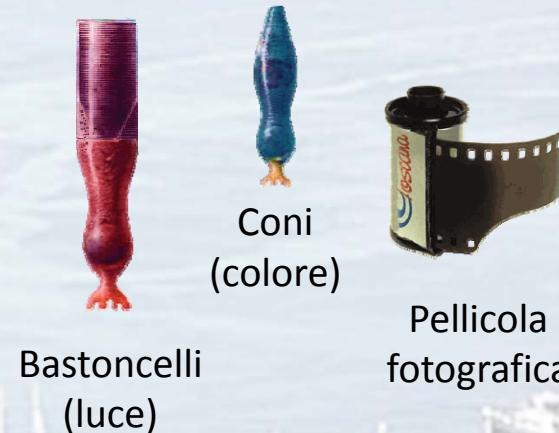
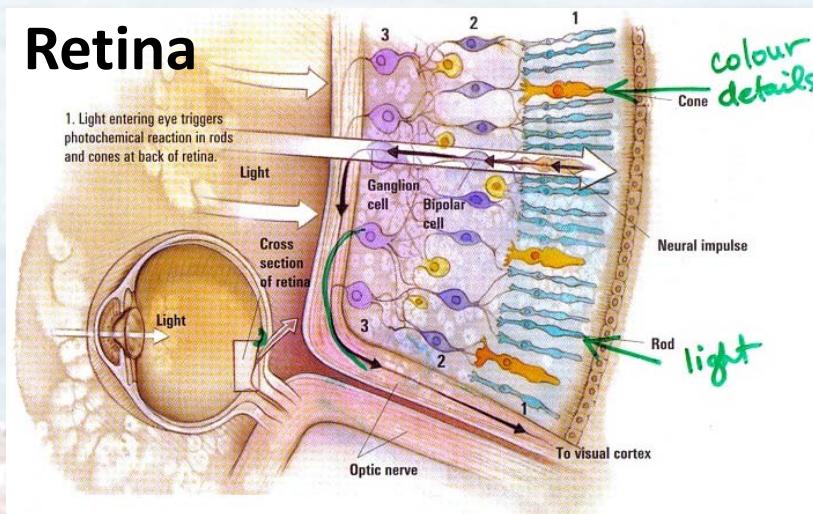
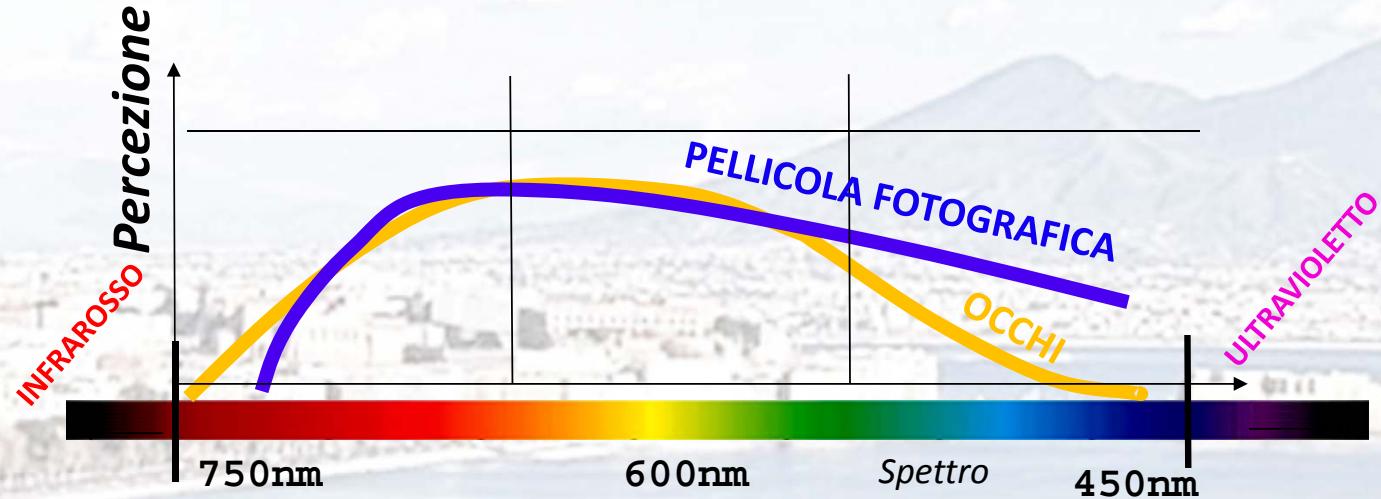


Le tre proprietà fisiche delle radiazioni di colore sono:

Luminanza (illuminazione) **Tinta** (il colore) **Saturazione** (la purezza)

walter.balzano@gmail.com

Sensibilità di percezione dei sistemi

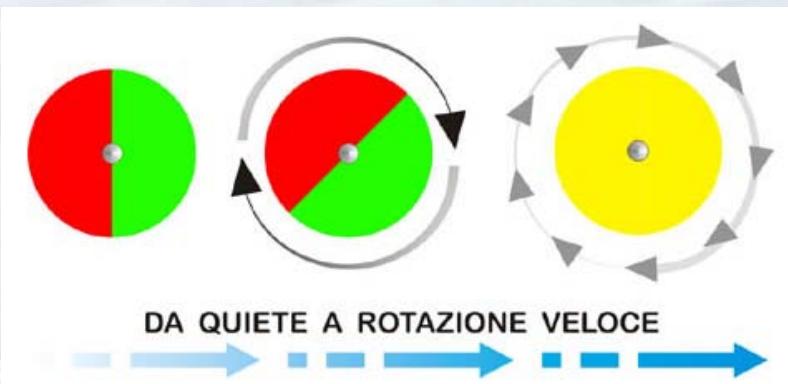


walter.balzano@gmail.com

Sintesi Additiva dei colori



Sintesi ADDITIVA: È possibile sperimentare che una opportuna **mescolanza di radiazioni di diversa lunghezza d'onda** produce la visione del bianco: tale risultato è opposto a quanto accade nella scomposizione della luce bianca solare nei colori dello spettro visibile ad opera di un prisma. Il verificarsi di tale fenomeno da origine a quella che comunemente viene detta **sintesi o mescolanza additiva**. Per convenzione si considerano 3 colori primari (una terna "naturale" non esiste) come quella RGB.

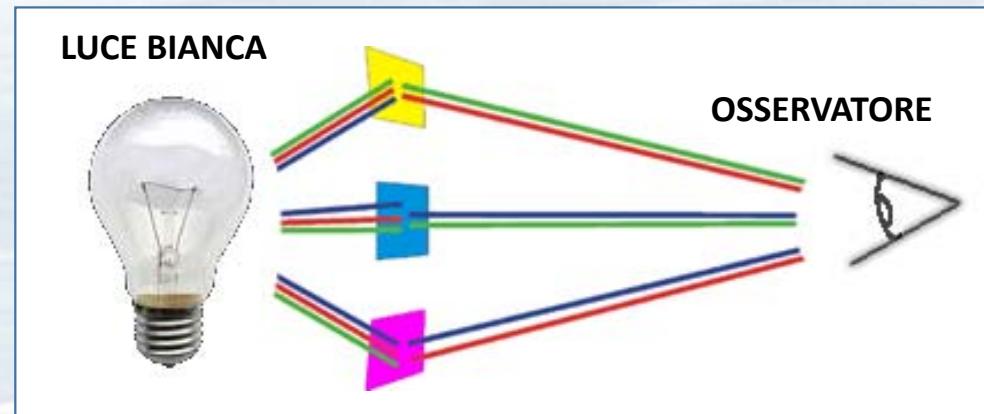
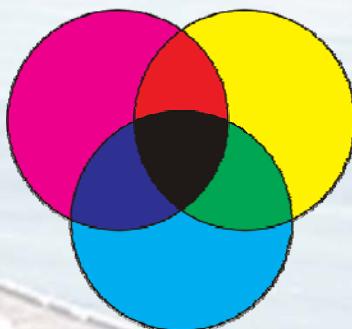


walter.balzano@gmail.com

Sintesi Sottrattiva dei colori



Sintesi SOTTRATTIVA: Considerando il fenomeno non dalla parte della radiazione riflessa, ma da quella della **radiazione assorbita**, allora le superfici che ci appaiono colorate (per es. **pittura e stampa**) sottraggono alla nostra visione una parte dello spettro visibile. Nella sintesi sottrattiva si analizza **l'effetto prodotto sulla visione della combinazione delle proprietà riflettenti dei colori** (cioè capacità di assorbimento della luce)

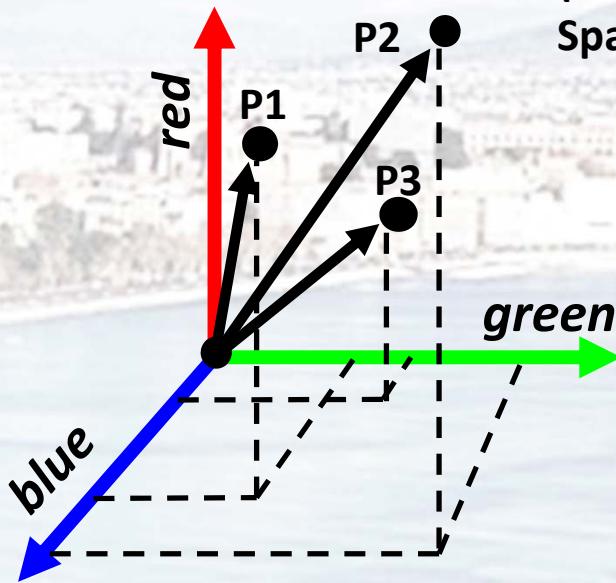


walter.balzano@gmail.com

Esempi di Modelli di Colore

Spazio RGB

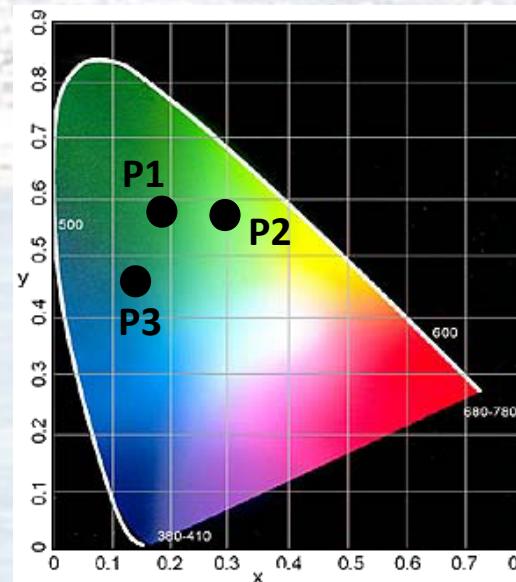
$$P_i(r_i, g_i, b_i)$$



- Spazio percettivamente NON uniforme
- Dipendente dal dispositivo.

Spazio CIE normalizzato (2D) $P_i(x_i, y_i)$

Commission Internationale de l'Eclairage - 1931
(Commissione Internazionale per l'Illuminazione)
Spazi (CIERGB, CIEXYZ, CIELUV, CIELAB)



Per la nostra capacità di percezione

$$\left(\overline{PP} \right)_{RGB} \neq \left(\overline{PP} \right)_{CIE}$$



walter.balzano@gmail.com

Sistemi di rappresentazione dei colori

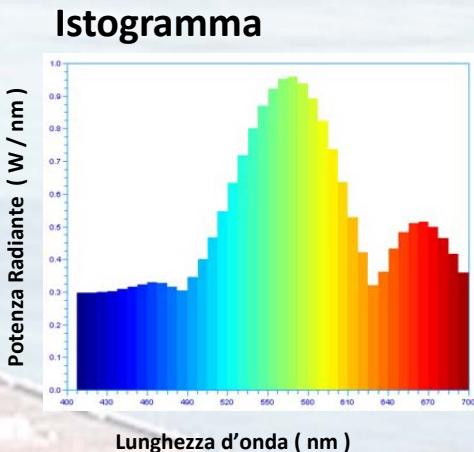
- Rappresentazione SPD (**Spectral Power Distribution**)
- Ogni colore “fisico” può essere rappresentato mediante la distribuzione dell’energia radiante alle varie lunghezze d’onda SPD

Vantaggi:

- è il modo più accurato per la specifica di un colore

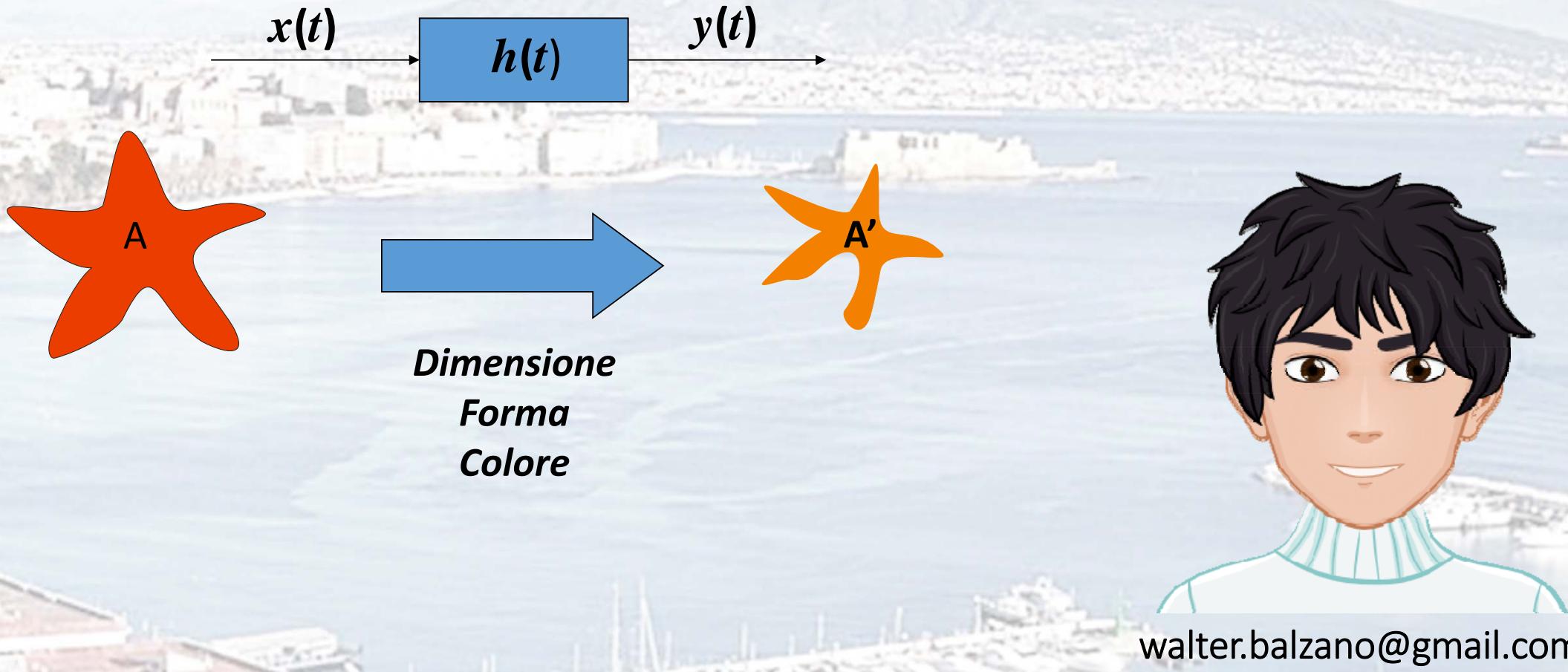
Svantaggi:

- Non descrive la relazione tra le proprietà fisiche del colore e la sua percezione visiva
- Complessa rappresentazioni numerica



walter.balzano@gmail.com

Funzioni di trasferimento



Correzione Gamma

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

$$\text{Luminanza} = V^\gamma$$

- Per neutralizzare la non linearità del dispositivo si usa la

$$\text{Correzione Gamma} = 1/\gamma$$

- La correzione di gamma è quindi **un'operazione non lineare usata per codificare e decodificare** la luminanza o i valori tristimolo in un sistema video o fotografico. Nel caso più semplice la correzioni di gamma è definita dalla seguente legge potenziale

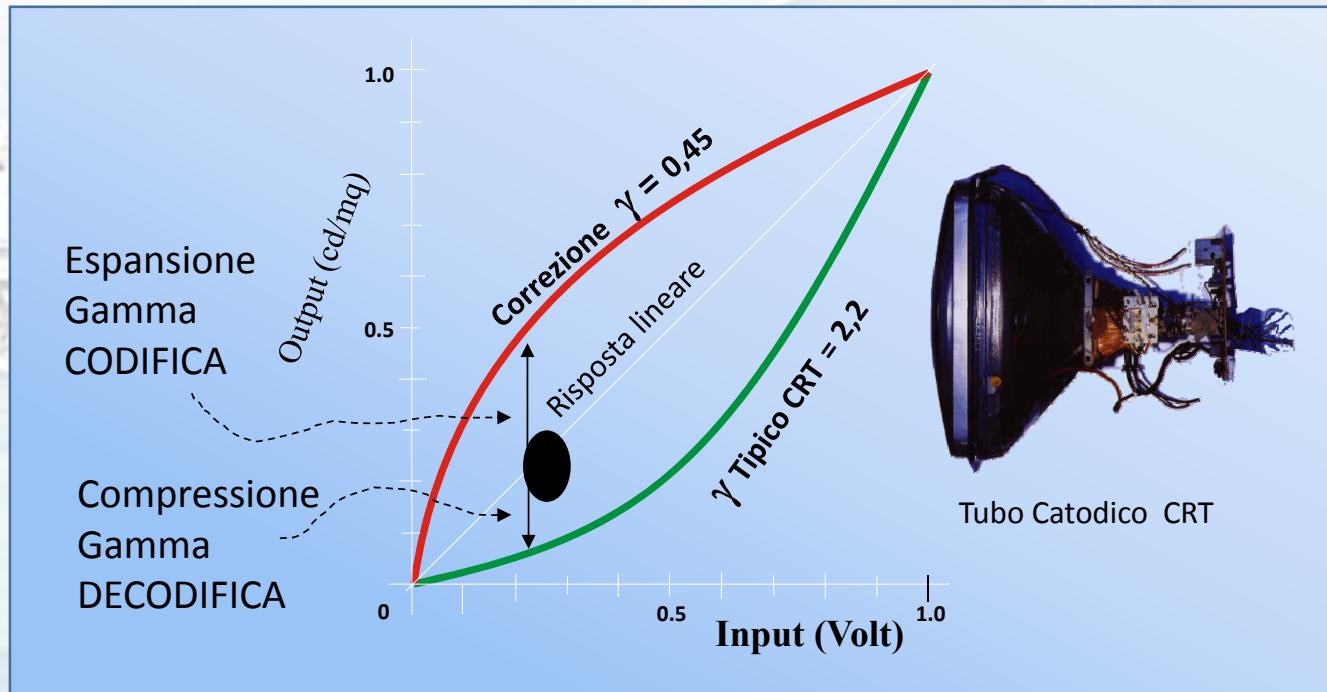
$$V_{out} = V_{in}^\gamma$$

- Tale funzione di trasferimento serve per **neutralizzare la non linearità del dispositivo**; **V** denota il segnale elettrico analogico (voltaggio) che il dispositivo preleva al proprio ingresso



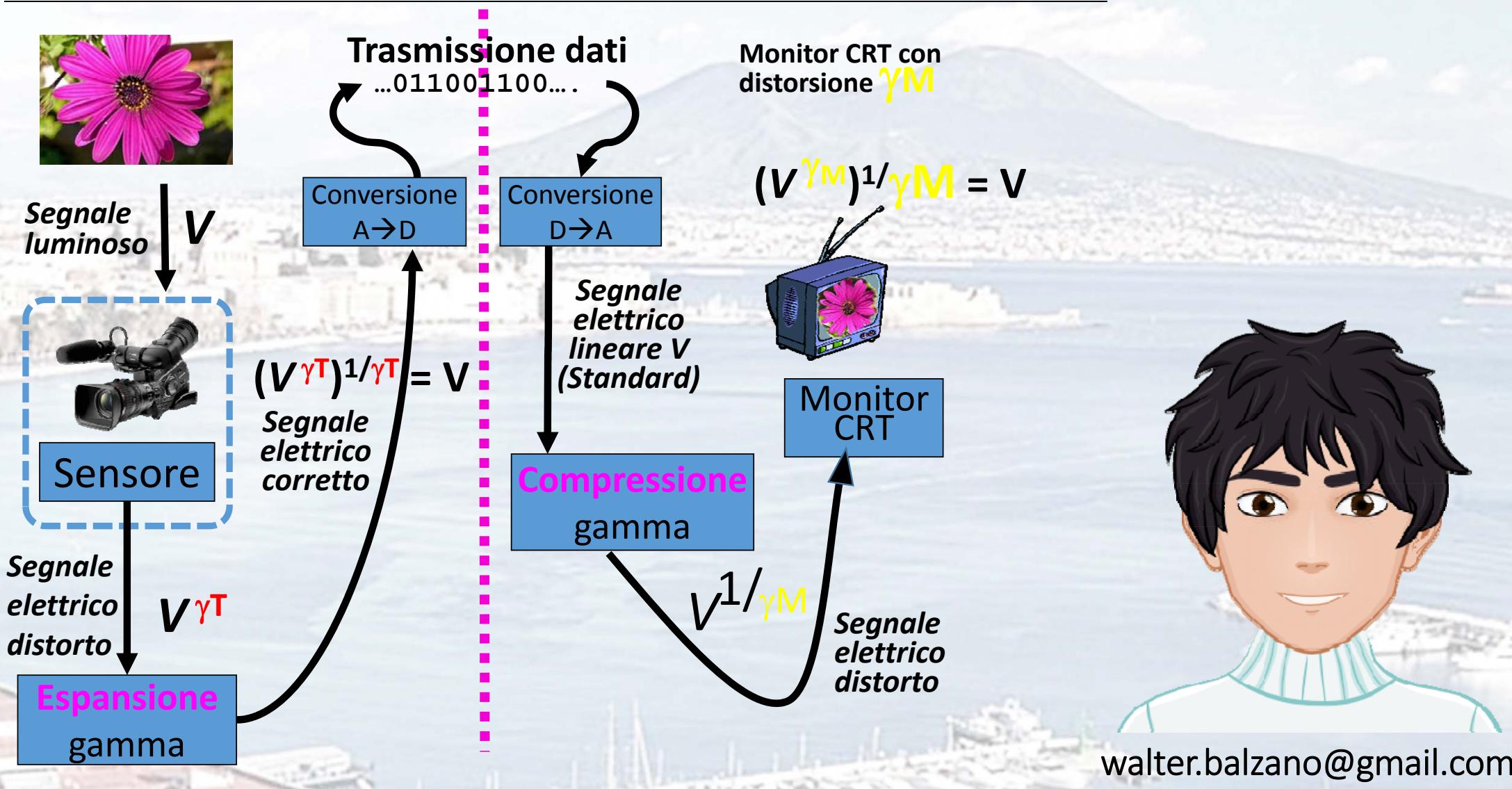
walter.balzano@gmail.com

Correzione Gamma (dispositivo)



walter.balzano@gmail.com

Correzione Gamma (schema generale)



Problemi di indicizzazione e ricerca legati allo spazio di colore

- Nella ricerca e nella indicizzazione delle immagini basate sul colore è necessario assumere che tutte le immagini siano rappresentate **nello stesso spazio di colore** e che il valore dei loro pixel rappresentino la stessa cosa.
- Data un'immagine nel formato RGB non è possibile interpretarla correttamente:
 - Non vengono indicate le definizioni delle primitive RGB e quindi non è possibile interpretare l'immagine in uno spazio comune
 - Non viene indicato il valore della correzione gamma applicata dal device che ha prodotto l'immagine
- Tutti i formati immagine più usati (GIF, JPEG, TIFF, ecc) non contengono informazioni sulle primitive dello spazio di colore e sulla correzione gamma usata.
- Il nuovo TIFF 6.0 contiene anche queste informazioni



walter.balzano@gmail.com

Le immagini digitali - [2]



**Le Origini
Grafica Raster e Vettoriale
Compressione
Analisi di Fourier
Frattali
Video**



walter.balzano@gmail.com

Le immagini digitali

Origini delle immagini digitali:

- Macchine fotografiche
- Scanner
- Fotogrammi di filmati
- Disegni elettronici

Schemi di classificazione delle immagini

● Raster (o bitmap)

- Raster (= griglia) definisce la griglia ortogonale di punti (scacchiera) che costituisce un'immagine. Ogni elemento è chiamato **pixel** a cui è associato un colore secondo uno specifico modello.
- I modelli di memorizzazione possono essere **LOSSY** o **LOSSLESS**

● Vettoriali

- Nella grafica vettoriale un'immagine è descritta mediante un insieme di primitive geometriche che descrivono punti, linee, curve e poligoni ai quali sono associabili svariati attributi.
- Grande qualità
- Elevata compressione



walter.balzano@gmail.com

Grafica Raster e Grafica Vettoriale

Grafica Raster:

Tecnica che descrive immagini mediante griglie di pixel colorati.

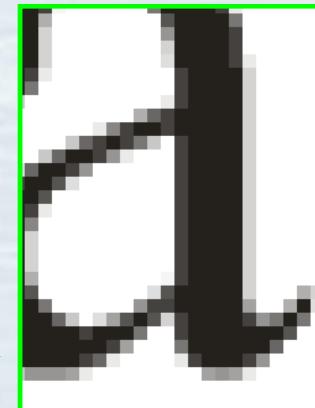
Grafica Vettoriale:

Tecnica che descrive immagini mediante un insieme di primitive geometriche quali punti, linee, curve e poligoni a cui possono essere attribuiti colori e sfumature.

Ciao mamma



Formato Vettoriale



Formato Raster



Grafica Raster e Grafica Vettoriale

Super-RResolution (SR): tecnica per migliorare la risoluzione delle immagini RASTER

Tecnica Single-Frame

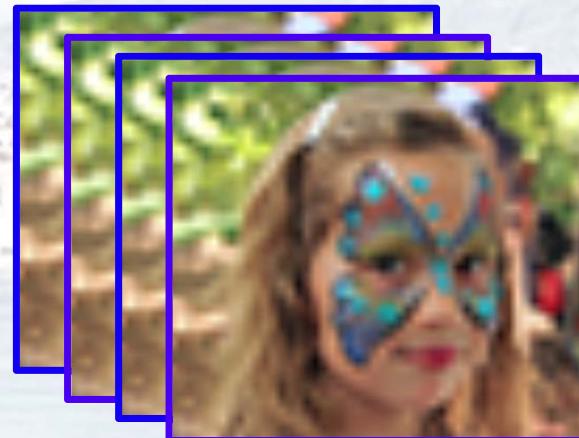
Input



Output



Tecnica Multi-Frame



Input

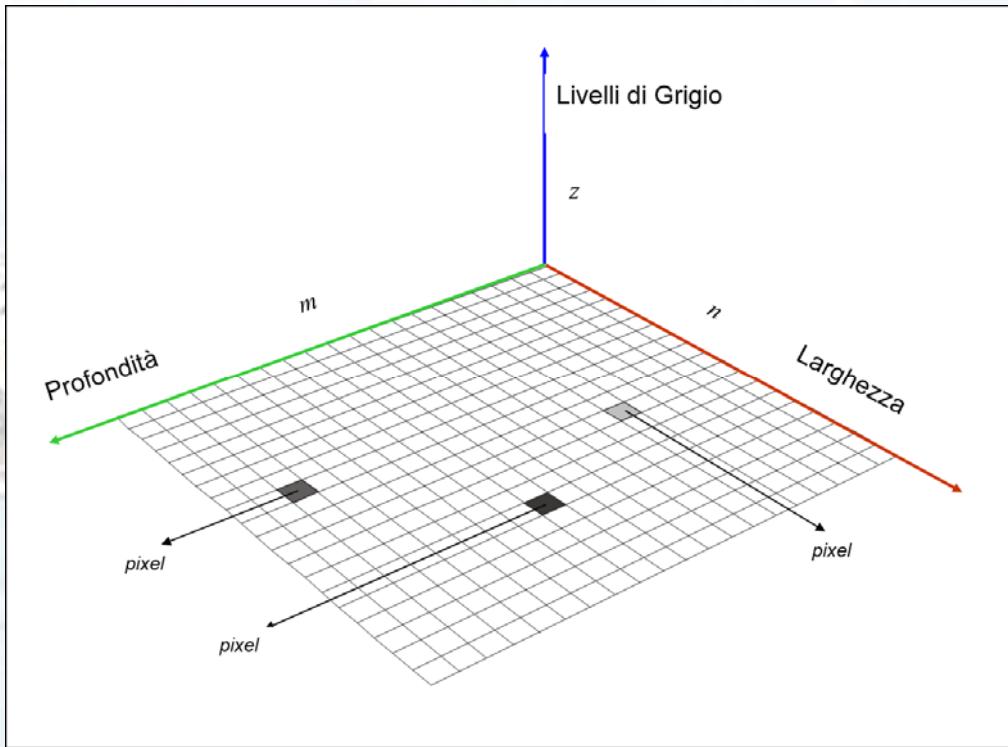


Output



walter.balzano@gmail.com

Immagini GrayScale (scala di grigio)



Grandezza immagine (misurata in bit): $m \cdot n \cdot \lceil \log_2 z \rceil$

Esempio:

Una immagine 640*480 con 256 toni di grigio
(cioè 1 byte per ogni pixel)
richiede $640 \times 480 = 307.200$ bytes



walter.balzano@gmail.com

Le immagini a colori

Un **modello di colore** è un modello matematico astratto che permette di rappresentare i colori delle immagini in forma numerica; esistono diversi modelli:

- **RGB** (Red, Green, Blue) (a sintesi additiva, per es. monitor, TV,...)
- **CMYK** (Cyan, Magenta, Yellow, Black) (a sintesi sottrattiva, per es. stampanti,...)
- **YUV** (opera in ambito analogico; nato per la compatibilità tv colori con tv B/N)
- **YCbCr** (l'equivalente digitale della YUV; **Y**=Luminanza; **Cb**=CromaBlu; **Cr**=CromaRed)
- **HSV** (Hue, Saturation, Value)

- La rappresentazione delle immagini a colori RGB è simile alla rappresentazione delle immagini a toni di grigio: **occorrono 3 valori per rappresentare ogni pixel** (3 matrici bi-dimensionali)

Trasformazione YCbCr \leftrightarrow RGB

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & -0.3313 & 0.5 \\ -0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



walter.balzano@gmail.com

Le immagini a colori: il modello RGB e YUV

Il Modello **RGB**



Il Modello **YUV**

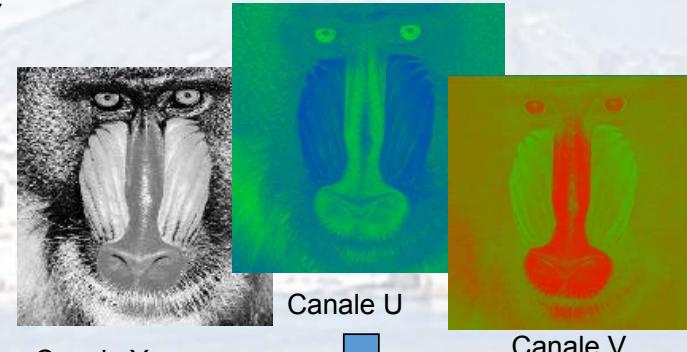


Immagine RGB

Immagine YUV



walter.balzano@gmail.com

Fondamenti della compressione:

- L'occhio umano è limitato poiché non percepisce tutto ciò che è realmente rappresentato.
- Generalmente, per ogni pixel di immagine, i punti adiacenti sono simili; la similitudine di tali punti ravvicinati costituisce la Ridondanza Spaziale.
- Differentemente dai testi, non è particolarmente significativa la perdita di informazioni; la compressione a perdita è quindi ammessa.



walter.balzano@gmail.com

Compressione immagine

Tipi di Compressione:

Senza perdita (**LossLess**)

permette di ricostruire perfettamente la rappresentazione dell'oggetto originale
(es. codifica di Huffman)

Con perdita (**Lossy**):

permette di ricostruire solo in parte la rappresentazione dell'oggetto originale.

Le tecniche “lossy” sono classificate in:

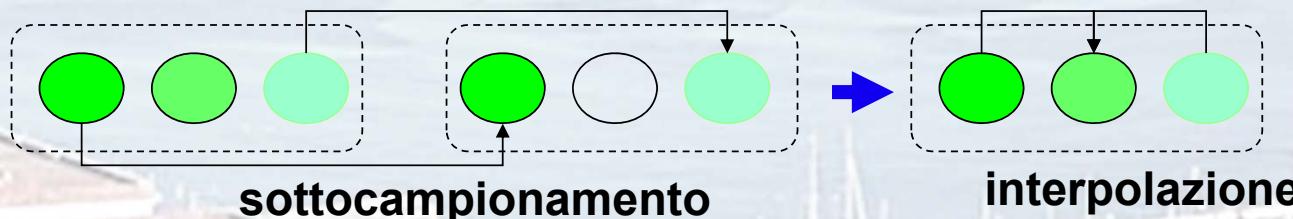
- Compressione con **sottocampionamento**
- Predictive Coding
- Codifica mediante trasformazione



walter.balzano@gmail.com

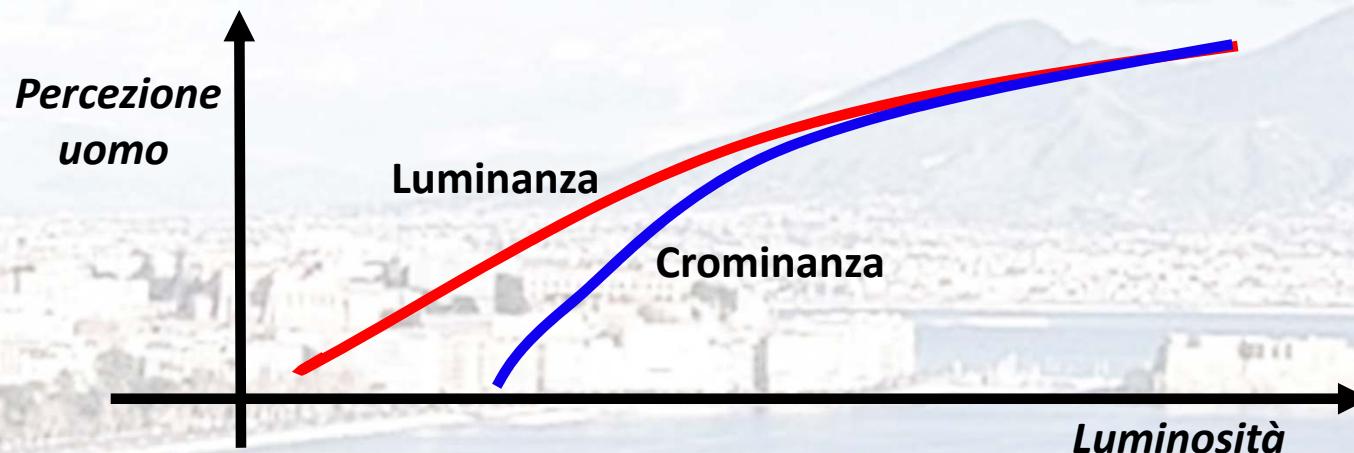
Compressione mediante sottocampionamento

- A causa della ridondanza spaziale (similitudine di pixel vicini), si considerano soltanto “alcuni” pixel; per esempio, si potrebbe considerare solo un pixel ogni due (sottocampionamento).
- In fase di decodifica i pixel “mancanti” vengono ricostruiti approssimando i loro valori mediante una interpolazione.
- E’ possibile raffinare questo metodo scegliendo di sottocampionare quelle componenti dell’immagine per le quali il nostro occhio è meno sensibile.
- Ogni immagine può essere decomposta mediante le componenti **Luminanza** e **Crominanza**.
- Il nostro occhio è molto più sensibile alla Luminanza che rispetto la Crominanza e quindi per la componente di Crominanza è possibile effettuare:
 - un livello di sottocampionamento maggiore.
 - Una quantizzazione meno raffinata.



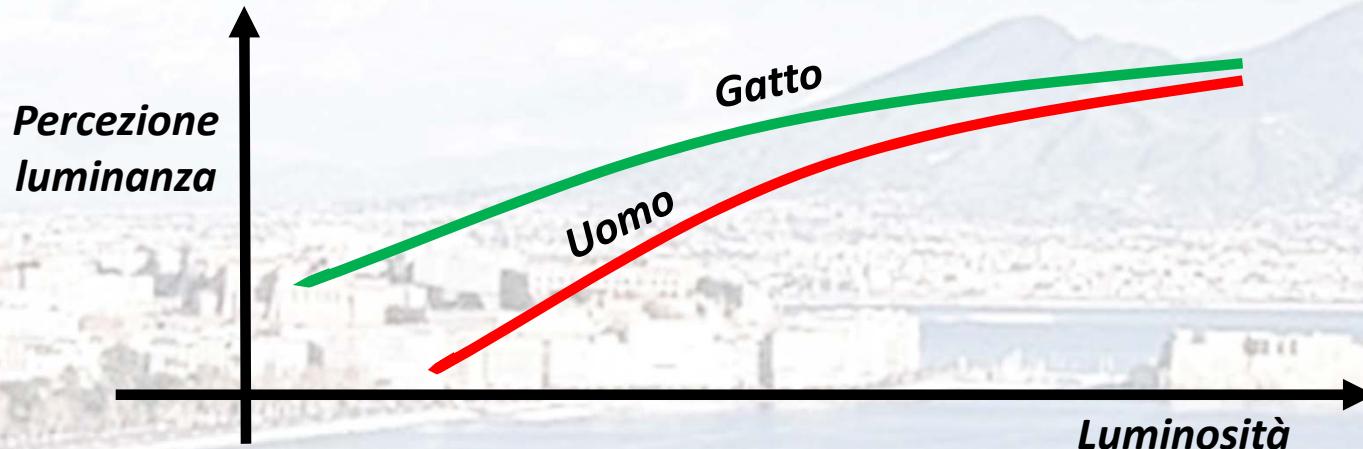
walter.balzano@gmail.com

Compressione mediante sottocampionamento



walter.balzano@gmail.com

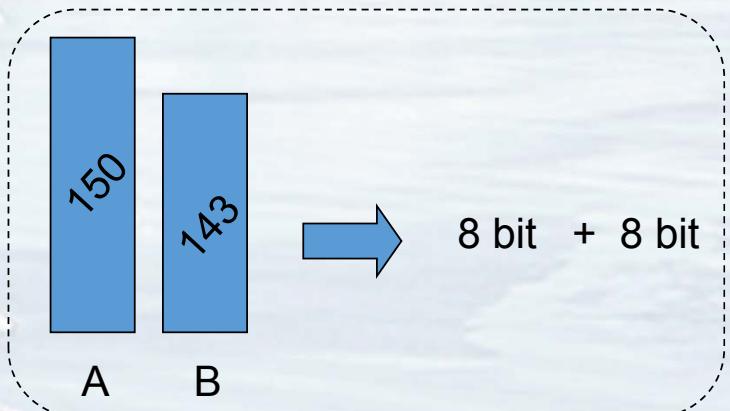
Compressione mediante sottocampionamento



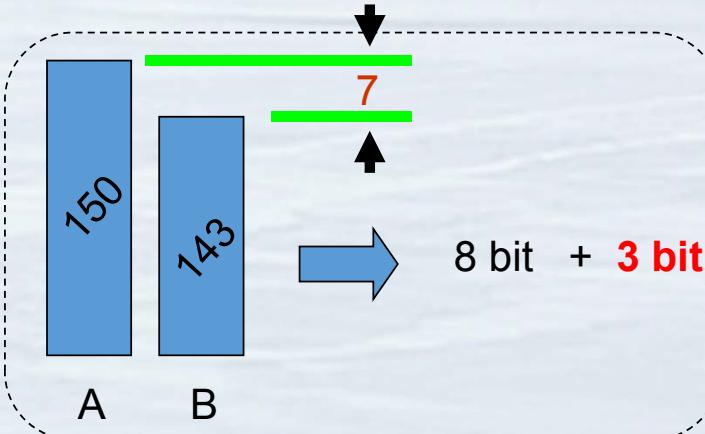
walter.balzano@gmail.com

Compressione mediante Predictive Coding

- E' analogo al predictive coding audio;
- Valori spazialmente vicini sono fortemente correlati;
- Se A e B sono due pixel vicini ed abbiamo già codificato A allora, **anziché codificare per intero anche B, è possibile codificare solo la differenza** tra i valori A e B. Tale differenza, essendo inferiore al valore di B, è codificabile con un minor numero di bit;
- La predizione può essere fatta sia tra le righe sia tra le colonne dell'immagine.



Codifica di pixel
A e B "contigui"



Codifica per differenza
di pixel A e B "contigui"

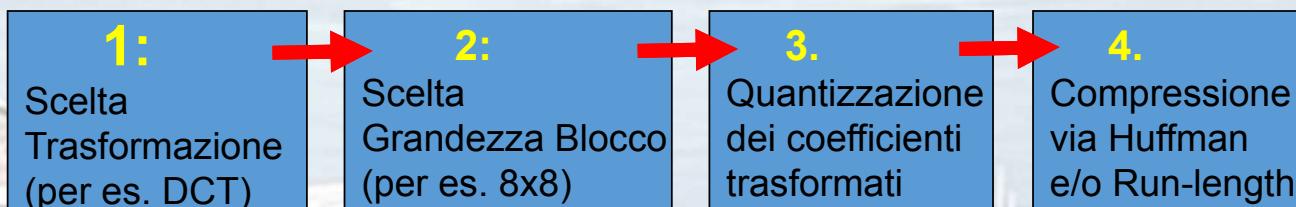


walter.balzano@gmail.com

Codifica mediante Trasformazione

- L'idea di base consiste nel considerare indipendenti i pixel dell'immagine convertendo gli elementi correlati dell'immagine in un **insieme minimale di coefficienti indipendenti** che implicano l'eliminazione della ridondanza.
- Una immagine viene suddivisa in sottoimmagini rettangolari su cui si applica una **trasformazione unitaria dal dominio spaziale al dominio delle frequenze**.
- Se nel dominio spaziale i dati sono fortemente correlati, allora i dati risultanti nel dominio delle frequenze sono adatti ad una fase di compressione (Huffman, Run-length,...)
- Le Trasformazioni maggiormente impiegate sono:
 - DFT (Discrete Fourier Transform)
 - DCT (Discrete Cosine Transform)

Implementazione del sistema di codifica:



walter.balzano@gmail.com

Nozioni (1): Serie di Fourier

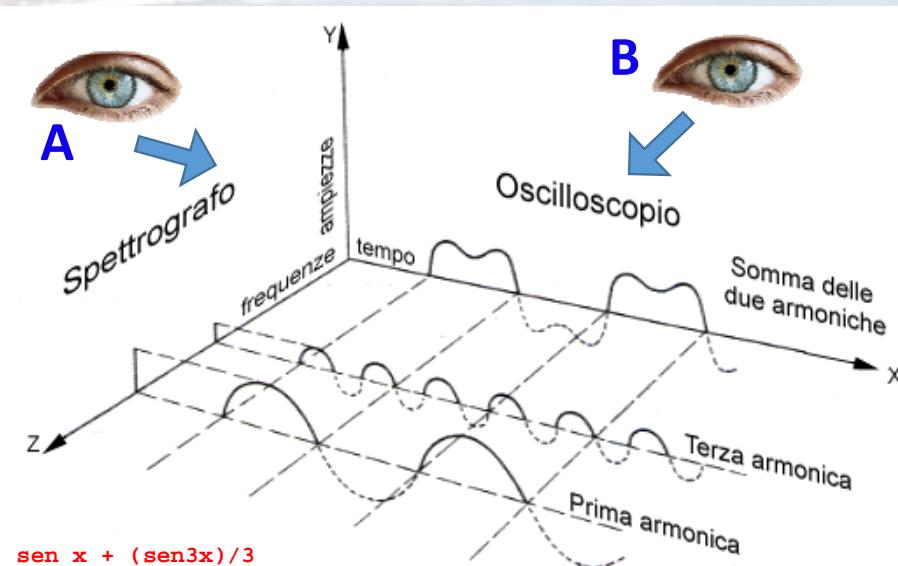
Fourier dimostrò che qualsiasi segnale periodico [$f(x)=f(x+T)$] può essere scomposto in una somma di (infiniti) segnali sinusoidali.

La serie di Fourier rappresenta un segnale periodico $x(t)$ di pulsazione ω_o mediante una somma pesata di sinusoidi di cui la prima (fondamentale) ha pulsazione ω_o e le successive (armoniche) hanno pulsazioni multiple di ω_o

$$x(t) = a_0 + a_1 \cos (\omega_o t + \theta_1) + a_2 \cos (2\omega_o t + \theta_2) + \dots + a_N \cos (N\omega_o t + \theta_N)$$

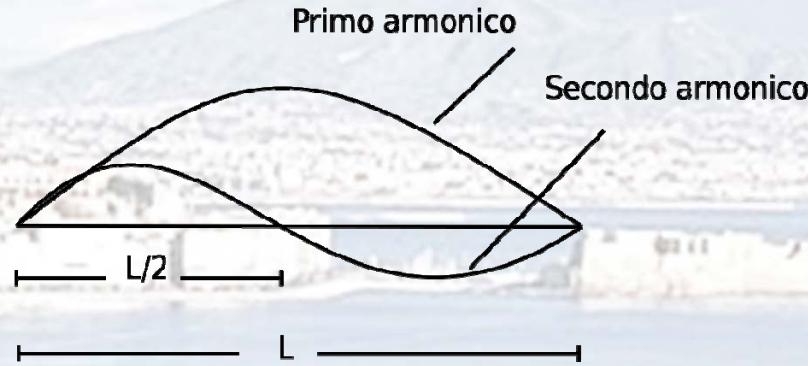


Jean Baptiste Joseph Fourier
1768-1830



walter.balzano@gmail.com

Nozioni (1): Serie di Fourier



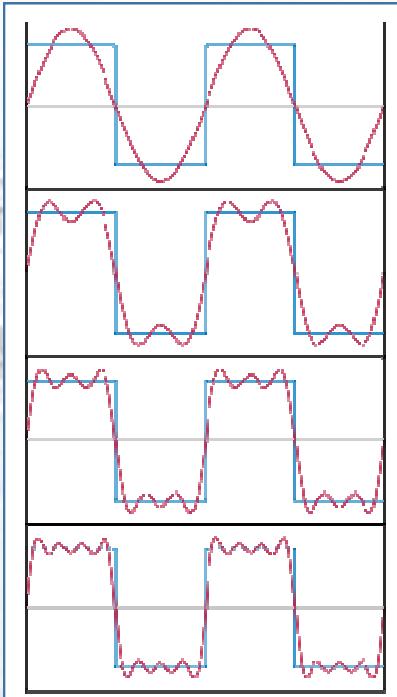
Gli armonici naturali sono una successione di suoni le cui frequenze sono multipli di una nota di base, chiamata fondamentale. Corrispondono alle frequenze naturali delle armoniche di una corda vibrante.

Un suono prodotto da un corpo vibrante non è mai puro (ovvero senza multipli in frequenza della nota di base) ma è costituito da più suoni, che si differenziano fra loro in intensità (volume) e frequenza (tono, alto o basso). Al suono fondamentale, quindi, se ne aggiungono altri: questi sono gli armonici, che hanno una importanza fondamentale sia nella determinazione del timbro di uno strumento che nella determinazione degli intervalli musicali.

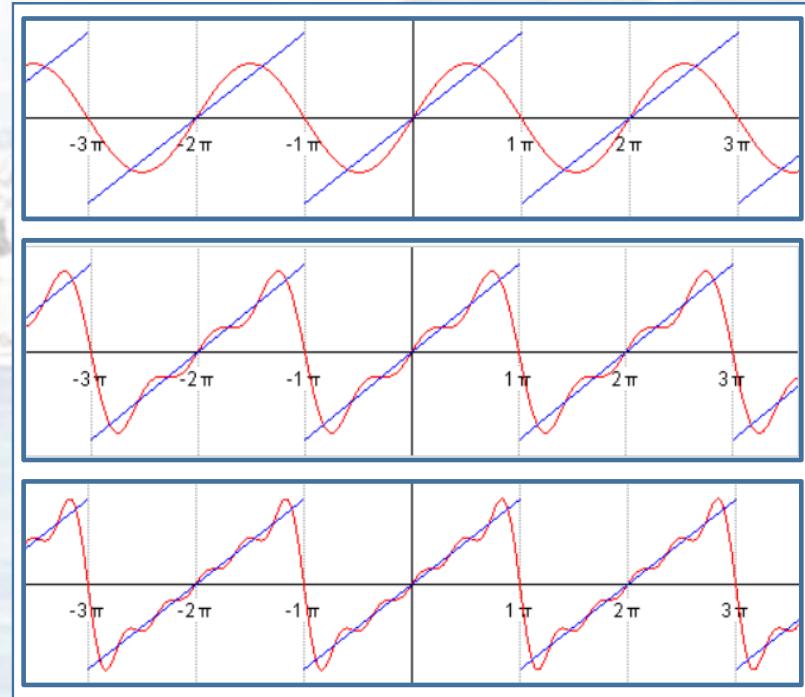


walter.balzano@gmail.com

Nozioni (2): Approssimazione con serie di Fourier



Approssimazioni per
l' «ONDA QUADRA»



Approssimazioni per il «DENTE di SEGA»



walter.balzano@gmail.com

Nozioni (3): Trasformata di Fourier

Data una funzione f , la sua trasformata di Fourier, $\mathcal{F}f$, è:

$$(\mathcal{F}f)(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

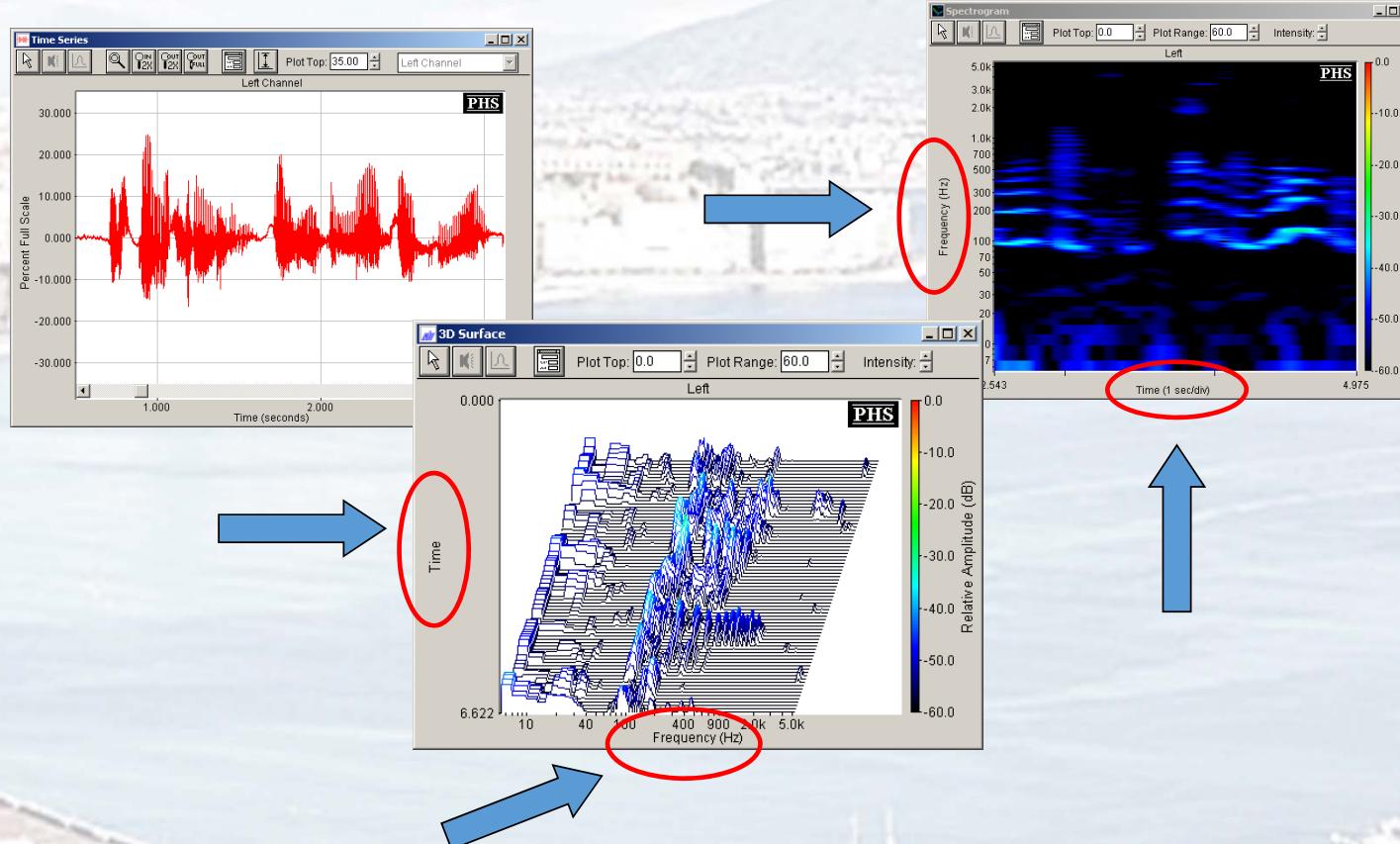
- La trasformata di Fourier decomponе un segnale nelle diverse frequenze che partecipano a costruirlo, comprese le loro fasi.
- Questa decomposizione è unica.
- DFT (Discrete Fourier Transform) DCT (Discrete Cosine Transform)
- Una semplificazione molto importante ed utilizza della Trasformata di Fourier è la FFT Fast Fourier Transform che è caratterizzata da una complessità di calcolo inferiore rispetto la Trasformata di Fourier.



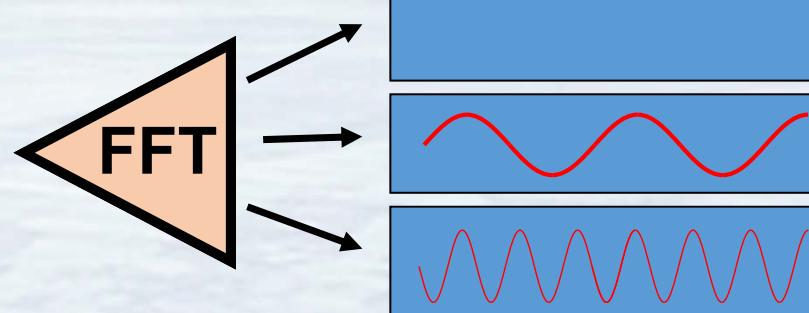
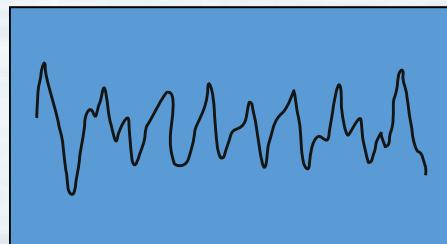
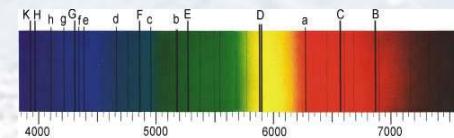
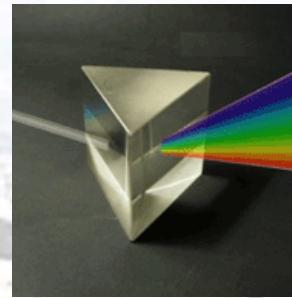
walter.balzano@gmail.com

Nozioni (4): Analisi spettrale

Con la Trasformata di Fourier passiamo al dominio delle frequenze



Nozioni (5): Analisi spettrale



walter.balzano@gmail.com

Immagini jpeg

- Standard JPEG (Joint Photographic Experts Group)
- Immagini a colori 24 bit ed immagini bianco e nero
- Modello LOSSY
- Modello Open Source



Fig.1 Qualità 10
1.443 byte



Fig.2 Qualità 30
2.166 byte



Fig.3 Qualità 60
3.826 byte



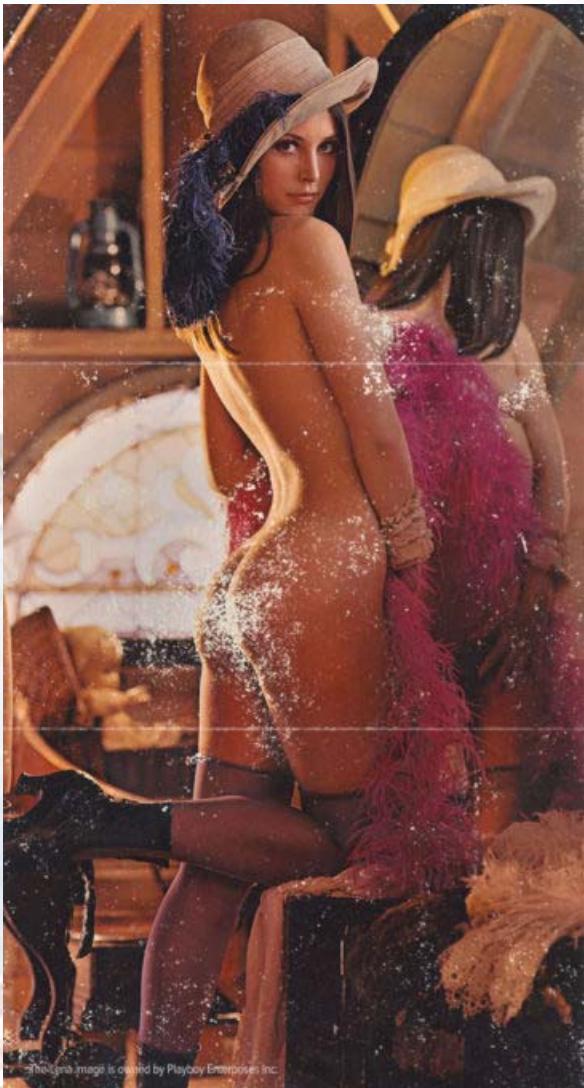
Fig.4 Qualità 80
6.216 byte



Idee di base:

- Limite percezione visiva.
- Analisi spettrale (DCT)

Immagini jpeg



Lena Forsén è una modella svedese nata (**Sjööblom** 31/3/1951), apparsa come Playmate della rivista Playboy (novembre 1972). È stata definita la "first lady di Internet".

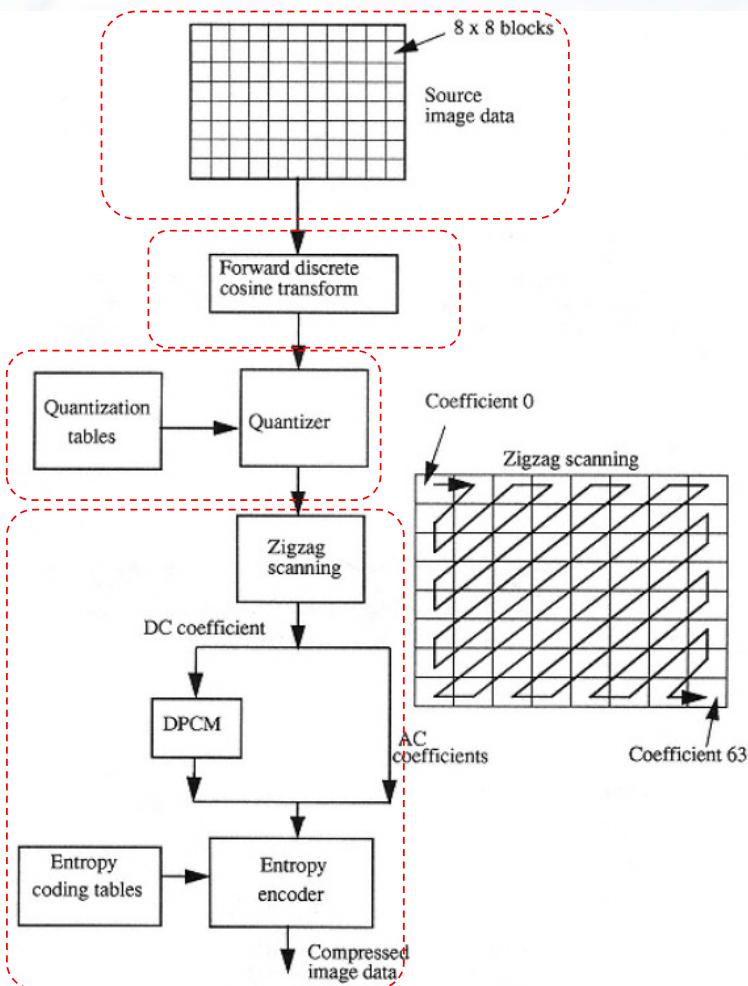


**Lena è il volto
dell'immagine di prova più
utilizzato al mondo**



walter.balzano@gmail.com

Jpeg: il modello



Fasi principali:

1. Preparazione in blocchi

RGB → YUV (matrici di Luminanza, Crominanza Blu, Crominanza rossa)

2. Passaggio al dominio frequenziale

Si applica la DTC alle 3 matrici Y Cr e Cb. Un'immagine può essere pensata composta dalla somma di tante variazioni di luminosità di forma cosinusoidale, di diversa frequenza (spaziale), fase e intensità. Le alte frequenze definiscono i dettagli dell'immagine mentre quelle basse le ombreggiature più estese.

3. Quantizzazione

Le frequenze della fase 2 vengono quantizzate in modo non lineare: le tabelle di quantizzazione utilizzano opportuni valori per arrotondare con meno precisione i coefficienti delle alte frequenze rispetto a quelle basse.

4. Codifica

Scansione ZIG-ZAG; Eliminazione delle ridondanze mediante codifica RLE e codici di Huffman



walter.balzano@gmail.com

Immagini Vettoriali

Grafica Vettoriale:

- Tecnica che descrive immagini mediante un insieme di primitive geometriche quali punti, linee, curve, poligoni,... a cui possono essere attribuiti colori e sfumature; è generalmente definita attraverso insiemi di equazioni matematiche.

Vantaggi:

- La sua rappresentazione è indipendente dalla risoluzione; una linea viene memorizzata mediante un'equazione, un punto iniziale ed un punto finale.
- Maggiore compressione dei dati → minor occupazione di RAM e HD rispetto una corrispondente immagine Raster: è particolarmente adatto per gestire grandi quantità di dati come quelli cartografici;
- L'ingrandimento o la riduzione delle misure e proporzioni del soggetto prodotto in vettoriale non incide in maniera significativa sul peso dell'immagine stessa.

Svantaggi:

- Meno intuitiva rispetto la grafica Raster;
- Richiede maggiore capacità di elaborazione e quindi Computer più potenti

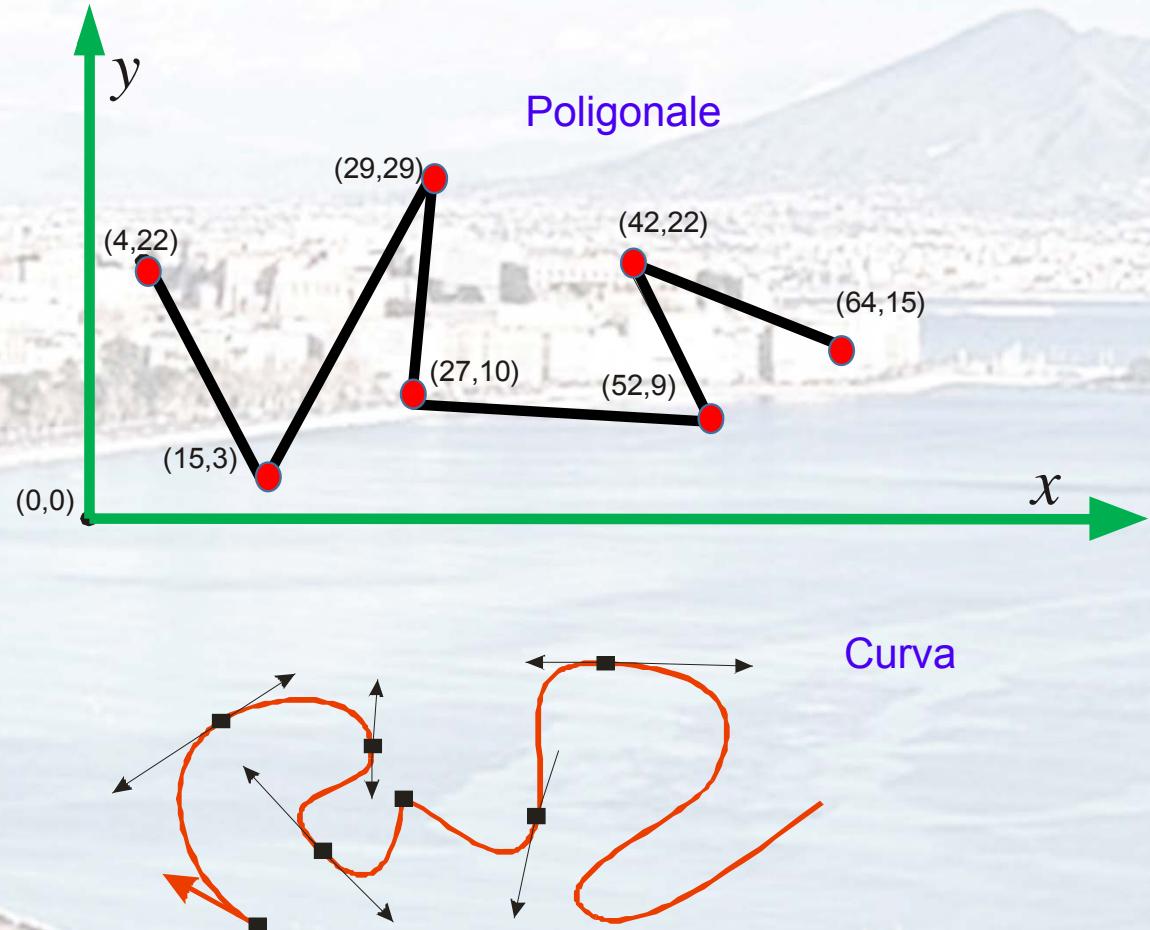
Applicazioni:

- Editoria, Architettura, Ingegneria, Computer Grafica,...
- *Altri esempi: Font TT (TrueType)*



walter.balzano@gmail.com

Immagini Vettoriali - 2



walter.balzano@gmail.com²⁵

Immagini Vettoriali – (2D e 3D)



da CorelDRAW (2D)

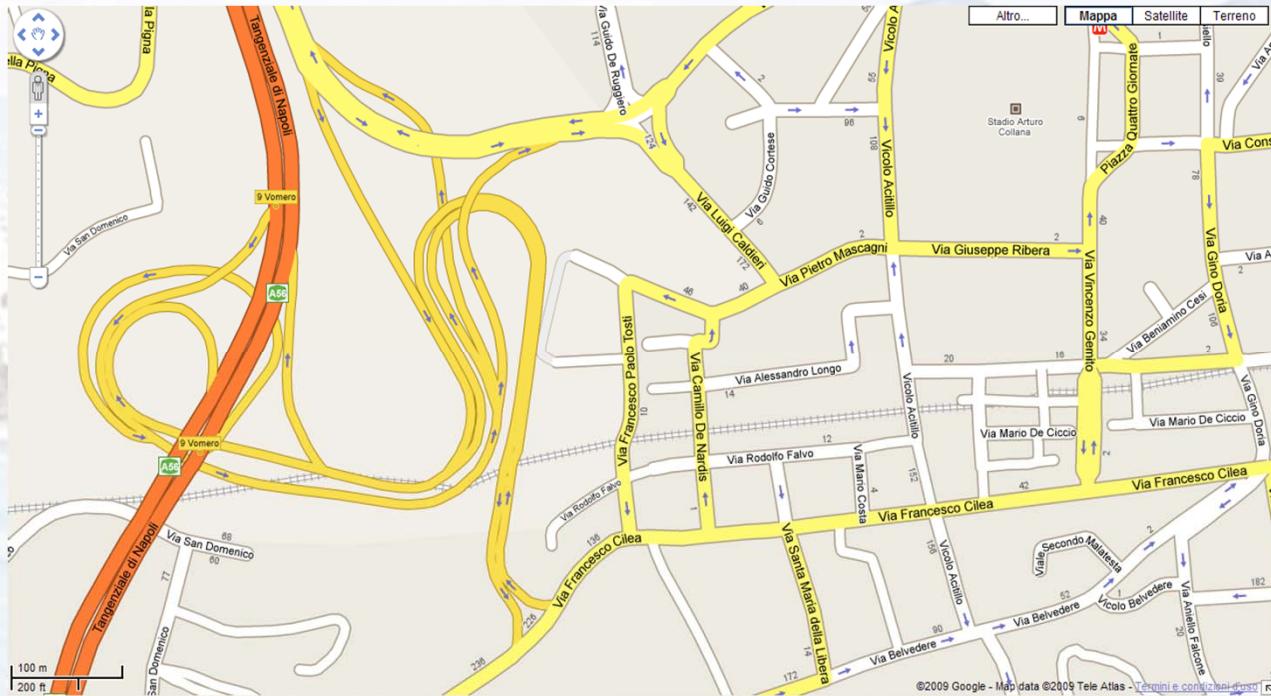


da Sketchup (3D)



walter.balzano@gmail.com

Immagini Vettoriali – (mappe)

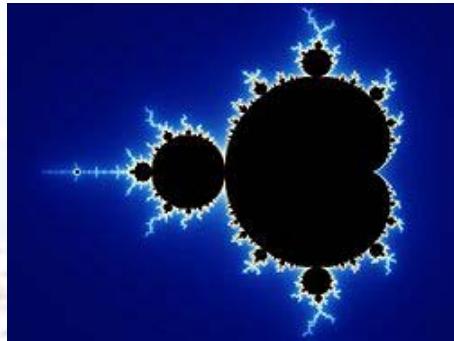


da Google MAPS



walter.balzano@gmail.com

Immagini Frattali



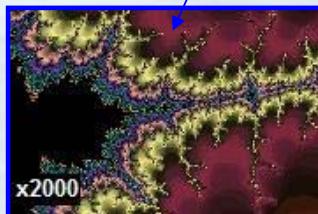
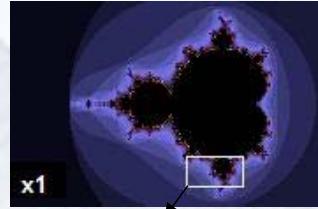
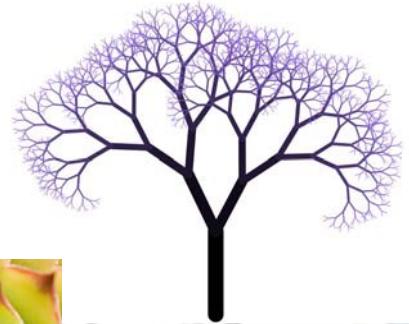
Le immagini cosiddette frattali per la loro somiglianza ai frattali, rappresentano, in un certo senso, una modo intermedio (tra Raster e Vettoriali) per la memorizzazione di immagini; infatti si cerca di **rappresentare la mappa di bit mediante una funzione matematica**. Tale tecnica è molto recente ed in evoluzione.

- A differenza della compressione jpeg che prende in esame le proprietà di frequenza e cromaticità dell'immagine, **il metodo frattale (che è un metodo LOSSY)** cerca di decomporre l'immagine in parti elementari che verranno memorizzate insieme alle regole che ne guideranno poi la ricomposizione.
- L'idea è basata sulla scomposizione dell'immagine in parti più piccole e cercando poi trasformazioni geometriche elementari (traslazioni, rotazioni, omotetie,...) che applicate iterativamente alle parti precedentemente decomposte permettano di ricostruire l'immagine originale.
- Il metodo è applicabile a qualunque “scala spaziale”: un ingrandimento molto forte di una immagine compressa con questo metodo non produce rilevante “effetto mosaico”; per lo stesso motivo, **in molti casi l'immagine compressa “si comporta meglio” della stessa immagine originale.**



walter.balzano@gmail.com

Immagini Frattali - esempi



Frattali in natura

Frattali di
Mandelbrot

walter.balzano@gmail.com



Il Video

- **Video digitale:** sequenza di fotogrammi o immagini visionate a frequenza costante.

- **Frame rate** (velocità di scorrimento)

- frame rate elevati → maggiore fluidità del filmato (25 frames al secondo)
- Frame rate elevati → maggiore banda per la trasmissione
- Per evitare il fenomeno del flickering, ogni fotogramma viene ripetuto più volte (refresh) finché diventa disponibile il fotogramma successivo (50 Hz)
- Con la tecnica dell'interlacciamento l'immagine viene composta a video componendo prima tutte le righe di immagine dispari e poi quelle pari.

- Il video è un media che richiede una grande quantità di dati per la sua rappresentazione; risulta quindi fondamentale l'applicazione di tecniche di compressione.

Esempio: un video di 10 minuti con fotogrammi 512*512, 24 bit/pixel, 30 Frame/secondo richiede $600*30*515*512*3 = 13.8 \text{ GB}$



walter.balzano@gmail.com

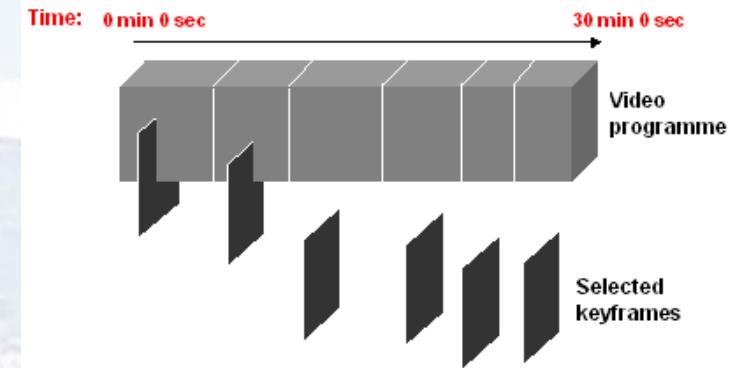
Codifica Intraframe / Interframe

Codifica Intraframe

- Codifica e decodifica di un flusso video descrivendo ogni singolo **fotogramma** che compone la sequenza video, rispettando quindi un approccio tradizionale alla **quantizzazione** video come sequenza di immagini statiche.
- Sequenze video particolarmente movimentate; descrizione dei singoli fotogrammi.
- Metodo preferito per la fase di authoring di un filmato.
- Semplice accesso diretto ad una scena generica.

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*

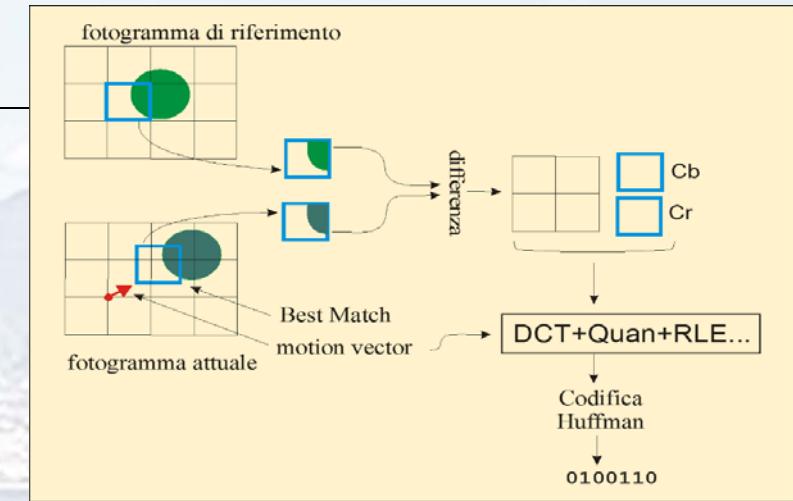


walter.balzano@gmail.com

Compressione Video

- Limite percezione visiva

- Riduzione della ridondanza
 - La ridondanza temporale viene trattata mediante la *similitudine dei fotogrammi adiacenti*.
 - Ogni fotogramma si divide in blocchi e si cerca la migliore corrispondenza tra blocchi di fotogrammi adiacenti. Per ogni coppia di blocchi “simili” si determina:
 1. Lo spostamento del blocco (*motion vector*)
 2. La differenza tra i 2 blocchi
 - Ciò che viene trasmesso è il *motion vector* e la relativa differenza



walter.balzano@gmail.com

MPEG

Obiettivi del progetto MPEG *Motion Picture Expert Group* (1988):

- Codifica congiunta audio video.
- Caratteristiche:

	Bitrate (Mbps)	Applicazioni	SIZE	Frame Rate (Frame/sec.)	Caratteristiche	Anno
MPEG-1*	1.5	VHS → VideoCD	360x280	30	Video a basso bitrate	1993
MPEG-2	10	Tv digitale	720x480	30	<ul style="list-style-type: none">• Multiplazione AudioVideo• Video interlacciato	1994
MPEG-3**	40	HDTV				-
MPEG-4					Gestione VO (Video Objects)	1998
MPEG-7					Organizza i contenuti Multimediali (XML)	
MPEG-21					<ul style="list-style-type: none">• Unifica MPEG-4 ed MPEG-7• Protezione digitale, diritti d'autore,...	

* MPEG-1 Layer 3 è meglio conosciuto come **MP3**

** Nel 1992 fu abbandonato poiché l'MPEG-2 riuscì a ricoprire anche le specifiche dell'MPEG-3

L'MPEG definisce le specifiche per gli standards di bitstream

Componenti MPEG: MPEG-Video, MPEG-Audio, MPEG-Systems

(gestione sincronizzazione e multiplexing diflussi di audio video compressi)



walter.balzano@gmail.com

Documenti multimediali compositi: l'SGML

SGML

- **Standard Generalized Markup Language**, standard per la descrizione logica dei documenti. L'organizzazione di un documento non è espressa usando la codifica dei sistemi di scrittura, che è finalizzata alla presentazione grafica, ma sono evidenziate le parti in cui è strutturato il documento (paragrafi, capitoli) ed altre particolarità del testo (come note, tabelle, intestazioni).
- Informazione è “**Marked up**”
- Fornisce una sintassi per la definizione del tipo di documento (**DTD**)
- In SGML, un DTD è necessario per la validazione del documento
- L'HTML e l'XML sono sottoinsiemi dell'SGML

Esempio di DTD:

```
<! ELEMENT persona (nome, cognome)>
<! ELEMENT nome (#PCDATA) >
<! ELEMENT cognome (#PCDATA)>
```



Struttura:

```
<persona>
  <nome>Mario</nome>
  <cognome>Rossi</cognome>
</persona>
```



walter.balzano@gmail.com

Documenti multimediali compositi: il PDF

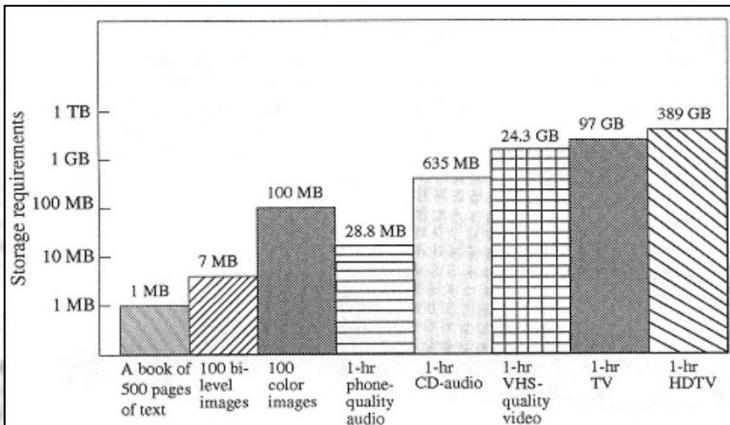
PDF

- Il **Portable Document Format**, comunemente abbreviato PDF, è un formato basato su un linguaggio di descrizione completa di una pagina. E' stato sviluppato da Adobe Systems nel 1993 per rappresentare documenti in modo **indipendente dall'hardware e dal software utilizzati** per generarli o per visualizzarli.
- Basato sul linguaggio Postscript di descrizione della pagina
- Permette una semplice
 - Estrazione dei dati
 - Annotazioni
 - Collegamenti ipertestuali
 - Bookmarks
 - Gestione della sicurezza



walter.balzano@gmail.com

Requisiti di memoria e Larghezza di banda



Requisiti di Memoria di alcuni media comuni

Applicazioni	Velocità (kbps)
CD-Audio	1.411
DAT	1.536
Telefonia Digitale	64
Radio Digitale, long-play DAT	1.024
Video a Qualità Televisiva	216.000
Video a qualità VHS	54.000
HDTV	864.000

Velocità di trasmissione di alcuni media comuni

● Requisiti di memoria

- Misurati in termini di **bytes** o **Megabytes**
- L'unità di base è il byte

● Larghezza di banda

- misurata in velocità di bits (**bit Rate**) o Megabits al secondo
- L'unità di base è il bit



walter.balzano@gmail.com

Progetto di Database MultiMediali

- 1. Architettura dei MIRS**
- 2. Modellizzazione dei dati di un MIRS**
- 3. Progetto dell'interfaccia utente in un MIRS**
- 4. Indicizzazione ed estrazione di dati da un MIRS**
- 5. Misure di efficienza di un MIRS**
- 6. Compressione, standardizzazione dei dati**



walter.balzano@gmail.com

Architettura dei MIRS

● Modularità

L'ampio range applicativo dei MIRS presuppone:

- Flessibilità ed adattività (librerie di funzioni)
- Gestione aggiornamenti



● Distribuzione

- Gestione dei dati multimediali (client-server)
- Accessi simultanei (librerie digitali, video-on demand).

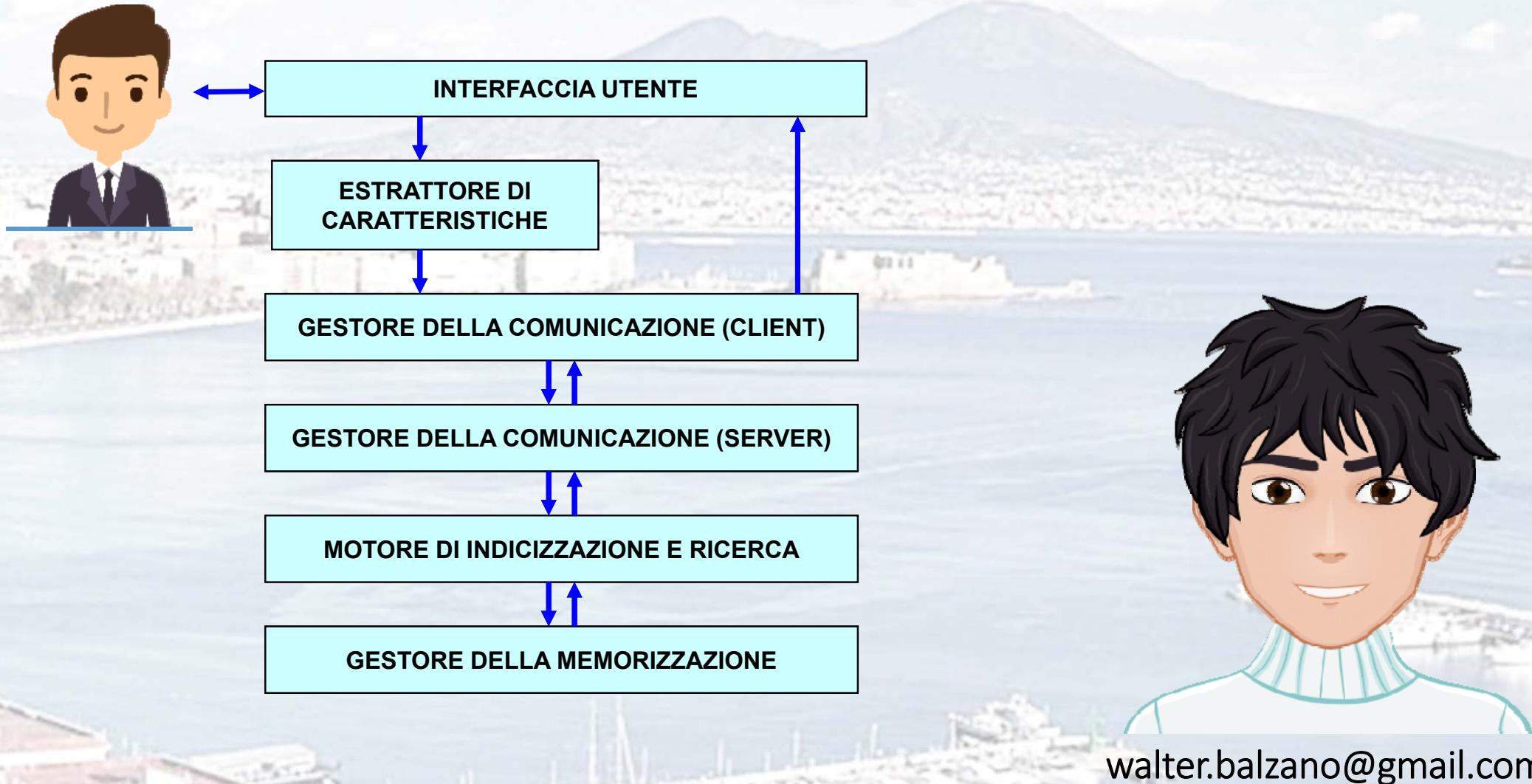
Moduli Opzionali

- **Thesaurus manager**: contiene sinonimi e altre relazioni tra le parole.
- **Integrity rule base**: testa l'integrità di una data applicazione.
- **Context manager**: tiene traccia del contesto dell'applicazione



walter.balzano@gmail.com

Architettura di base di un MIRS



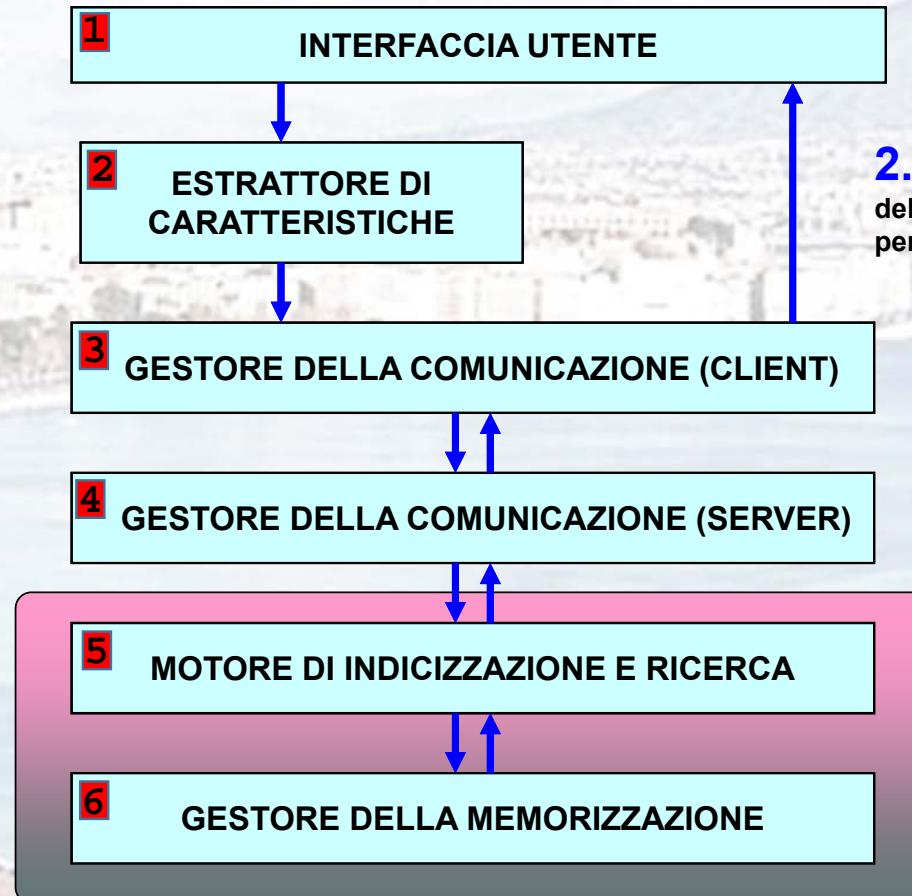
Architettura di un MIRS (FASE 1: inserimento)



1. L'utente specifica il tipo di input (microfono, videocamera, CD...).
2. I contenuti dei dati in input vengono estratti automaticamente o semi-automaticamente.
3. Gli oggetti multimediali originari con le relative caratteristiche estratte sono inviati al (o ai) server.
4. Gli oggetti multimediali in arrivo vengono organizzati e sono dati in input al blocco seguente.
5. Esegue l'indicizzazione delle informazioni.
6. Memorizza le informazioni di indicizzazione e gli oggetti multimediali originari.



Architettura di un MIRS (FASE 2: recupero)



1. L'utente esegue la query di un elemento non necessariamente già contenuto nel D.B.
2. Vengono estratti i contenuti della query, così come avviene per un inserimento.
3. Il client accoglie i dati dall'estrattore e li smista al motore di indicizzazione e ricerca.
4. Il server accoglie i dati dall'estrattore e li smista al motore di indicizzazione e ricerca.
5. Il motore di indicizzazione e ricerca reperisce l'elemento contenuto nel database che meglio si adatta alla query.
6. Recupero dell'oggetto richiesto; la destinazione di tale oggetto è l'interfaccia di visualizzazione.



walter.balzano@gmail.com

Modello dei dati

- In un DBMS, la finalità della modellizzazione è di specificare tipi e proprietà degli oggetti che dovrà contenere; In un MMDBMS o in un MIRS le finalità della modellizzazione comprendono anche una specifica dei diversi livelli di astrazione dei dati multimediali
- Un modello di dati per un MIRS deve comprendere la descrizione di proprietà statiche e dinamiche:
 - **PROPRIETA' STATICHE**: riguardanti gli oggetti stessi che costituiranno i dati multimediali, le loro relazioni e i loro attributi.
 - **PROPRIETA' DINAMICHE**: riguardanti le interazioni tra gli oggetti, operazioni disponibili sugli oggetti, interazione con gli utenti.
- L'usabilità di un MIRS è fortemente condizionata dal modello dei dati.
- L'indicizzazione di dati multimediali implica la considerazione di
 - spazi multidimensionali di caratteristiche
 - definizione di una metrica in tale spazio.



walter.balzano@gmail.com

Modello dei dati

Requisiti di un modello di dati per 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.

Il paradigma **Object Oriented (OO)** è il meglio adattabile alla modellizzazione dei dati multimediali:

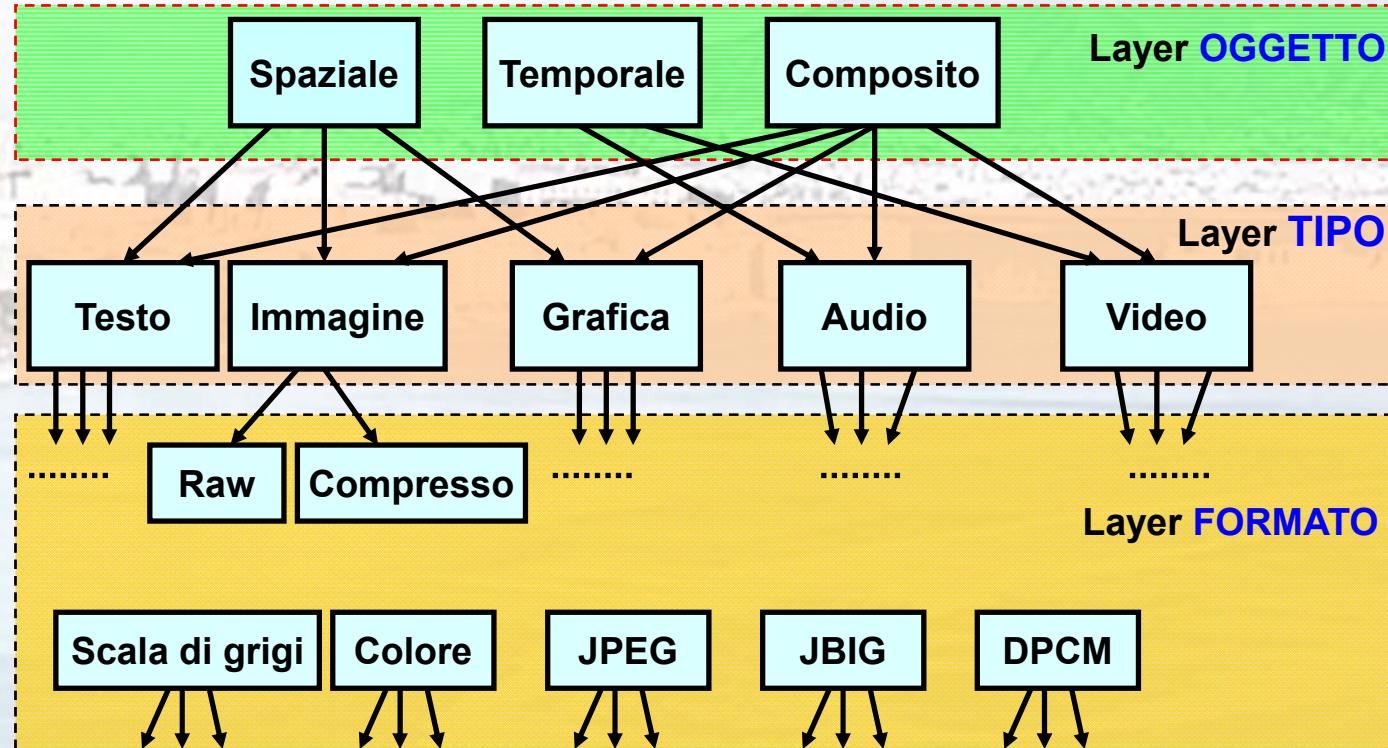
- **Incapsulamento** di codice e dati in una singola unità chiamata oggetto
- Il codice definisce le operazioni che possono essere effettuate sui dati

Struttura di tipo **Multilayer**: Permette l'indicizzazione, la ricerca e l'elaborazione a diversi livelli di astrazione.



walter.balzano@gmail.com

Modello dei dati (schema)



walter.balzano@gmail.com

Layer oggetto, tipo e formato

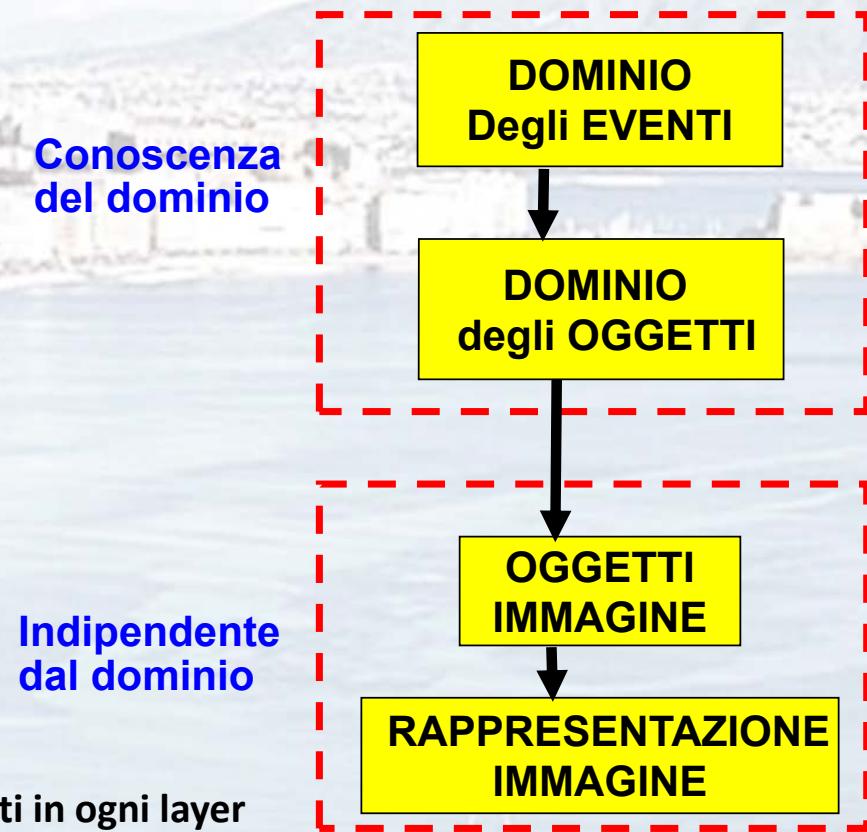
- **Layer OGGETTO:** Costituito da uno o più media con specifiche relazioni spaziali e temporali
 - Es.: *Una presentazione di immagini in successione accompagnate da un commento audio*
 - Relazioni spaziali (dimensione di ogni immagine e posizione di apparizione)
 - Relazioni temporali (tempo di inizio e tempo tra immagini; sincronizzazione con il contenuto audio)
- **Layer TIPO:** Contiene i tipi comuni di media
 - *Queste informazioni sono utilizzate nella fase di ricerca e di calcolo della similarità*
 - Ogni tipo è derivato da una comune classe media astratta
 - A questo livello sono specificati gli attributi del media;
 - per un'immagine:
 - Dimensione in pixel
 - Istogramma del colore (abbondanza relativa dei vari colori)
 - Oggetti contenuti nell'immagine (un'auto, una persona...)
- **Layer FORMATO:** Specifica il formato in cui il media è memorizzato
 - *Queste informazioni sono usate nelle fasi di decodifica, analisi e presentazione*
 - Raw Compresso Tipo di compressione usata
 - *NOTA: Il modello dei dati può dipendere anche dalla specificità e dalla particolarità dell'applicazione; Un'opportuna fase di progettazione del modello può contribuire alla definizione di un modello di base generico, sul quale possono essere aggiunte o ricavate nuove caratteristiche ed oggetti richiesti dalle particolari applicazioni.*



walter.balzano@gmail.com

Esempio 1 - VIMSYS

Il modello VIMSYS (Visual Information Management System) per la gestione di video ed immagini è formato da quattro layer:

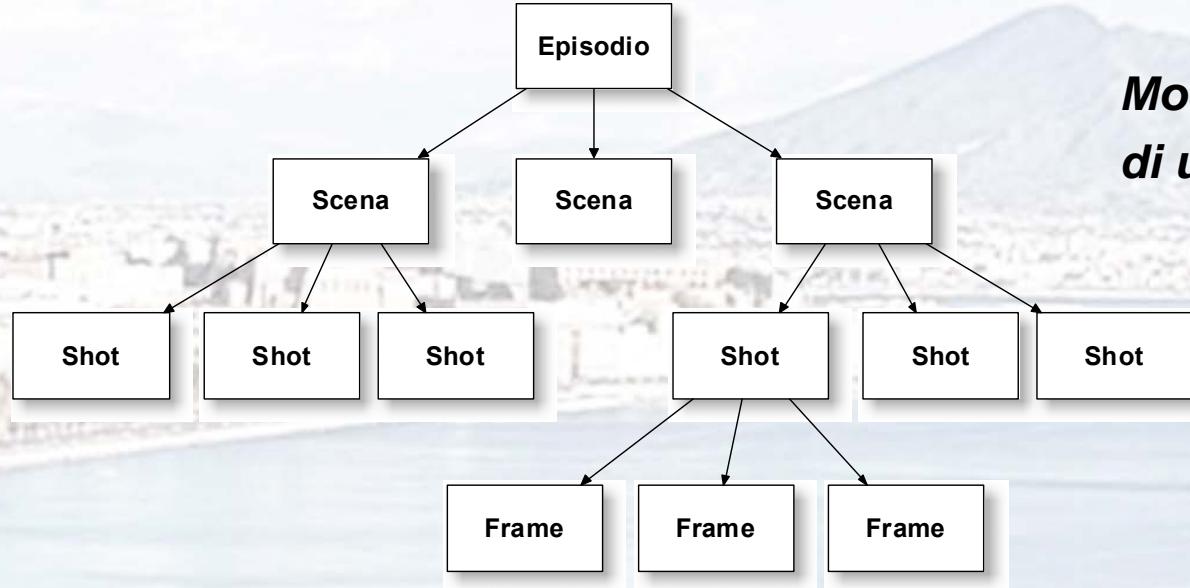


Tutti gli oggetti presenti in ogni layer hanno i propri attributi e metodi.



walter.balzano@gmail.com

Esempio 2: modello di un video



**Modello generale
di un video**

Ad ogni livello di dati vengono assegnati gli attributi relativi:

Episodio: autore, data di creazione, tipo di video, produttore, ecc...

Scena: semantica comune a tutti gli shot che contiene

Shot (ripresa): frame chiave, oggetto ripreso, data, luogo

Frame: statistiche sull'immagine, distribuzione del colore, ecc...



Interfaccia utente

Requisiti principali dell'interfaccia utente di un MIRS:

- fornire strumenti per inserire oggetti nel database in maniera semplice
- fornire strumenti per definire efficacemente le query e le esigenze di ricerca
- presentare i risultati delle ricerche in maniera efficiente ed efficace
- essere user-friendly

Popolazione del database

- A differenza dei DBMS tradizionali in un MIRS i dati sono costituiti da media diversi e non hanno struttura ed attributi prefissati → gestione complessa
- L'interfaccia deve consentire l'inserimento di dati multimediali semplici e composti e la specifica di tipologie di attributi che devono essere estratte ed indicizzate
- L'estrazione degli attributi può essere automatica o semiautomatica

Le query che possono essere inviate ad un MIRS sono:

- **Multiformi**: l'utente può inserirle utilizzando modi differenti e tipi di media differenti;
- **Incerte**: l'utente sa cosa vuole ma non sa descriverlo e riconosce il risultato corretto solo quando lo vede.



walter.balzano@gmail.com

Fase di ricerca

La ricerca è un compito fondamentale per i sistemi di gestione dei database.

Tipi di ricerca:

PER SPECIFICA:

Necessita di interpretazione per passare da testo a definizione di attributi e **feature** su cui ricercare:

Es. "Auto rossa" = range di colore dei pixel e forma di un oggetto contenuto nell'immagine

PER ESEMPI:

Necessita di strumenti di authoring multimediale per permettere l'inserimento dell'esempio con svariate modalità: **disegno, cattura di immagini e video**, ecc...



walter.balzano@gmail.com

Raffinamento di una query

Spesso per query “**incerte**” occorre una fase di browsing: l’utente sa riconoscere quello che cerca ma non sa bene come descriverlo:

- richiesta vaga e presentazione di un insieme ampio di risultati
- ricerca su una tassonomia che organizza le informazioni (categorie, ecc...)
- ricerca casuale su un insieme di risultati forniti a caso

- Data l’incertezza delle query sui dati multimediali, l’utente deve poter raffinare le sue richieste in base ai risultati ottenuti per la richiesta iniziale.
- Se l’utente si vede restituito dal sistema un elemento abbastanza simile a quello cercato, il sistema deve permettergli di riutilizzarlo per effettuare una nuova ricerca, più raffinata.
- Tale processo può essere ripetuto più volte.
- La conoscenza del dominio e il profilo utente possono essere utilizzati per raffinare una query.
- Un feedback sulla pertinenza è particolarmente utile nelle applicazioni multimediali.
- Nella pratica, il processo di individuazione di un’informazione multimediale è una combinazione delle fasi di ricerca, browsing e raffinamento.

L’interfaccia utente deve permettere la visualizzazione di **tutti i tipi** di dato multimediale (testo, immagini, video, audio...) e presentare gli aspetti essenziali degli oggetti o permettere all’utente di muoversi all’interno di esso (ad esempio una lunga sequenza video o audio). Il **tempo di elaborazione** deve essere basso (i risultati devono essere mostrati via via che vengono trovati)



walter.balzano@gmail.com

Estrazione delle Feature

- Gli oggetti multimediali gestiti dal database sono preprocessati per estrarne **feature** ed **attributi**.
- Il processo di ricerca si basa sulla **ricerca e comparazione** di tali feature (e non del dato originale): l'efficacia di estrazione delle feature è basilare per ottenere un sistema di buona qualità (se una feature non è stata identificata ed estratta dal dato, ad una query su tale feature il sistema non potrà rispondere in maniera adeguata)

I requisiti per l'estrazione delle feature sono:

- Le feature estratte devono essere **complete** e rappresentare il contenuto e l'informazione presente nel dato
- 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



Tipi di feature

METADATA:

- Catturano le informazioni di contesto che non descrivono o interpretano il contenuto del dato stesso (autore, data di creazione, titolo, ecc...)

ANNOTAZIONI TESTUALI:

- Sono descrizioni testuali del contenuto di un dato multimediale (soffrono del problema di essere soggettive e incomplete)

FEATURE DI BASSO LIVELLO:

- In genere possono essere estratte automaticamente
- Catturano dati e statistiche di un oggetto e le relazioni spazio temporali tra parti dell' oggetto
 - **Audio:** volume medio, distribuzione in frequenza
 - **Immagini:** distribuzione del colore, tessitura, forma degli oggetti,...
 - **Video:** struttura temporale e feature dei singoli frame

FEATURE DI ALTO LIVELLO:

- In genere l'estrazione richiede l'intervento umano.
- Cercano di riconoscere e capire gli oggetti: per esempio se in un file audio c'è musica o parlato, qual è il soggetto in primo piano in un video...



Tipi di feature

- Il retrieval basato sugli ultimi due tipi di feature è detto basato sul contenuto. Un sistema di gestione di dati multimediali deve supportare tutti e quattro i livelli di feature, che si completano a vicenda e rendono più completa la descrizione di un oggetto, adattandolo alle query.
- Ad esempio un'annotazione testuale si presta a catturare concetti astratti come i sentimenti, ma non è idonea a descrivere pattern complessi come forme o texture irregolari. Viceversa, le feature di basso livello riescono a catturarli, ma non a descrivere concetti astratti.
- Quando un oggetto multimediale comprende più tipi di media, le loro relazioni ed interazioni sono utili per l'estrazione di feature, l'interpretazione e il retrieval. Alcuni tipi di media sono più facilmente interpretabili e possono essere sfruttati per comprendere ed estrarre le feature di altri tipi di media.
- Ad esempio, se un oggetto è formato da uno streaming video e da una traccia audio, si può applicare lo speech recognition per acquisire conoscenza riguardo l'oggetto e poi usare questa conoscenza per sezionare e estrarre feature ed oggetti dallo streaming video. Si osservi che MPEG-7 è uno standard di descrizione di feature, ma non ne specifica le modalità di estrazione.



walter.balzano@gmail.com

Indicizzazione dei dati

- Dopo avere estratto le feature e gli attributi da un dato multimediale è necessario usare delle strutture di indicizzazione per organizzare la memorizzazione delle feature e fare in modo che la ricerca sia efficiente
- Data la natura diversa delle feature estratte, servono strategie di indicizzazione adeguate (ad esempio la distribuzione del colore in una immagine può essere memorizzata attraverso un istogramma)
- L'indicizzazione può essere gerarchica e avvenire a più livelli
- L'indicizzazione può prendere in considerazione le relazioni spazio-temporali tra gli oggetti
- Sono necessarie misure della similarità nello spazio delle feature estratte che simulino il giudizio umano (che è in genere soggettivo e dipendente dal contesto)

Misure di similarità

- Il retrieval multimediale è basato sulla similarità e non su un matching esatto tra query e elementi del database. La similarità è calcolata sulle feature estratte e sugli attributi ed è espressa come uno o più valori.
- La pertinenza dei risultati del retrieval è giudicata da esseri umani e quindi il maggior requisito delle misure di similarità e dei tipi di feature è che siano dei parametri adatti all'osservazione umana, circostanza che li rende comunque soggettivi e dipendenti dal contesto. Tutti questi fattori rendono complessa e cruciale la valutazione del retrieval.



walter.balzano@gmail.com

Garanzie sulla QoS (Quality of Service)

- I sistemi MIRS sono in genere distribuiti
- I dati multimediali impongono requisiti molto forti per una fruizione adeguata (larghezza di banda di trasmissione, spazio di memorizzazione, sincronizzazione spaziale e temporale)
- QoS specifica un insieme di parametri e requisiti richiesti in due gradi:
 - Qualità preferibile
 - Qualità accettabile
- La QoS è in genere negoziata tra client e server e sottoscritta tramite un “contratto” che garantisce tali 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



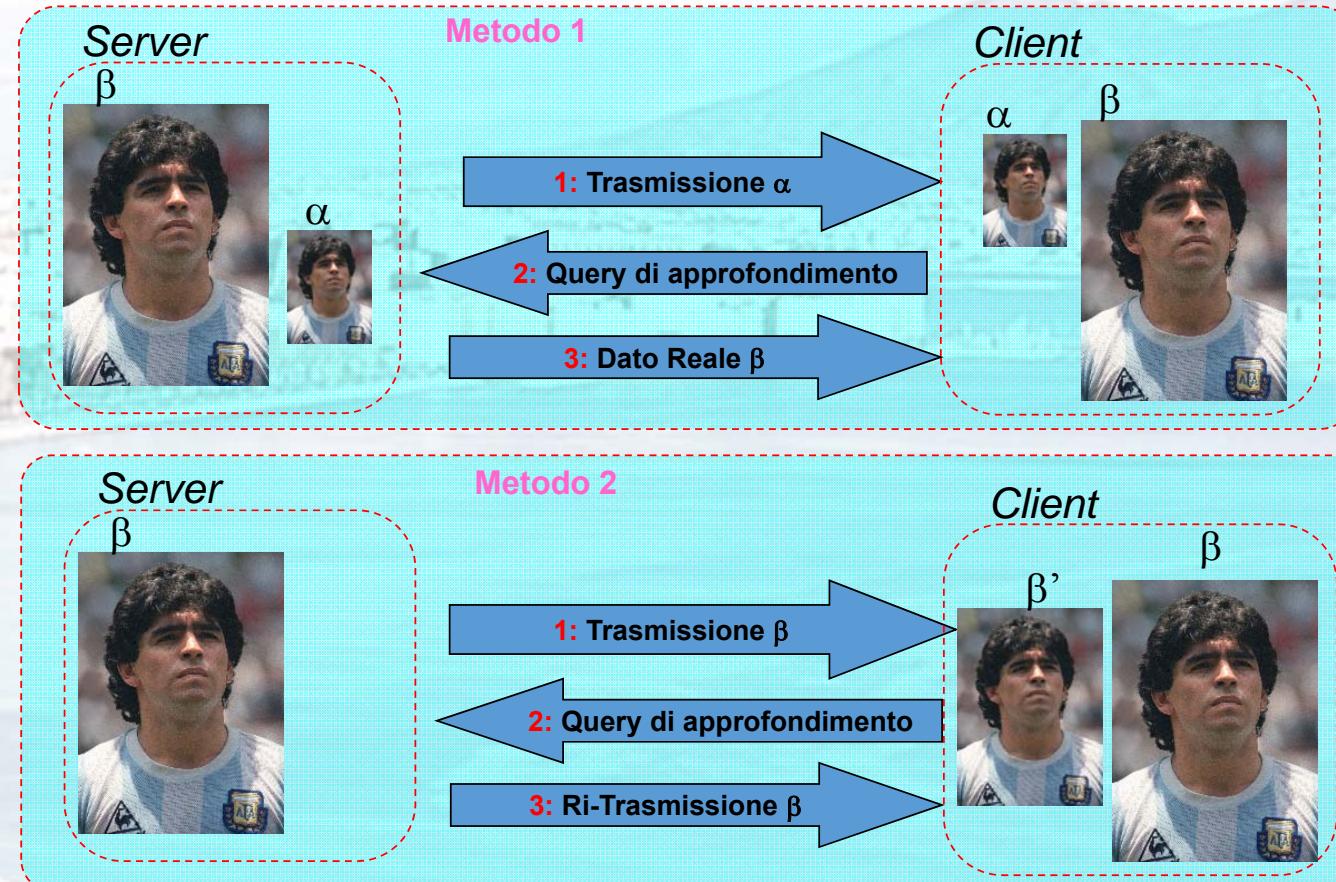
Multimedia Data Compression

- La gran parte dei dati Audio, Video, Immagini,... è salvata in formato compresso
 - L'obiettivo principale della compressione dei Dati Multimediali non considera gli aspetti relativi al retrieval e alla presentazione
 - Per l'estrazione delle caratteristiche dagli oggetti multimediali occorre prima effettuare una decompressione
 - L'operazione di estrazione delle caratteristiche è molto frequente → troppa elaborazione, occorre uno schema più efficiente (sia per la trasmissione che per la presentazione)
- **Metodo 1:**
- Sul server, per ogni grande immagine si salva anche una “copia ridotta”
 - La query dell'utente recupera sempre la copia ridotta
 - Se occorre anche il dettaglio allora oltre la copia ridotta viene recuperata anche l'immagine originale
 - Svantaggio: ridondanza dei dati sul server
- **Metodo 2:**
- Ogni query dell'utente recupera direttamente l'immagine originaria
 - L'immagine originaria viene ridotta per poter essere rappresentata sul client
 - Se occorre maggior dettaglio dell'immagine allora il server ritrasmette l'immagine originale
 - Svantaggio: Spreco di banda per la trasmissione
- **Metodo 3:**
- Si usano metodi di decompressione scalabili, progressivi e gerarchici (es.: gif e jpeg progressive)



walter.balzano@gmail.com

Multimedia Data Compression (2)

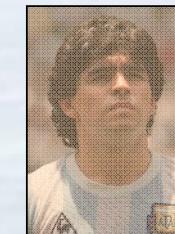


walter.balzano@gmail.com

Standard di Rappresentazione dei Dati

L'estrazione delle caratteristiche ed il processo di confronto presume che, per ogni media, il dato raw (grezzo o di base) sia lo stesso. Ciò non corrisponde al mondo reale, infatti, per esempio:

- Diversi brani audio possono essere registrati a diversi livelli di amplificazione → il loro confronto potrebbe perdere di significato
- Diverse immagini possono essere equalizzate (gamma correction) in modo completamente diverso → il loro confronto potrebbe perdere dei significato



walter.balzano@gmail.com

Indicizzazione e Recupero dei documenti di Testo

- 1. Introduzione**
- 2. Differenze IR e DBMS**
- 3. Indicizzazione automatica e modello booleano per il recupero**
- 4. Retrieval con modello spazio vettoriale**
- 5. Retrieval con modello probabilistico**
- 6. Recupero basato su cluster**
- 7. Metodi IR non tradizionali**
- 8. Misurazione delle prestazioni**
- 9. Tecniche di IR a confronto**
- 10. I motori di Ricerca WWW**



walter.balzano@gmail.com

Introduzione

IR: tecniche di recupero delle informazioni

- Rilevanza storica: grandezza del volume dei dati
(es: librerie digitali)
- Il testo è impiegato come strumento manuale di annotazione sfruttabile da un IR

Concetti di base:

- Rappresentazione delle query e dei documenti
- IR con diversi modelli di rappresentazione dei documenti ma simili processi di indicizzazione
- Tecniche di confronto tra documenti e query
- Modelli maggiormente usati:
 - Match esatto (modello booleano)
 - Spazio Vettoriale
 - Modello probabilistico
 - Modello su Cluster (raggruppamento)
- Query con linguaggio naturale
(problemi di lingua ed ambiguità linguistica)
- Tecniche di Intelligenza Artificiale.



walter.balzano@gmail.com

Differenze IR e DBMS

DBMS:

- Struttura omogenea dei record
- Componenti del record: gli attributi
- Un record è definito completamente ed univocamente dai propri attributi
- Retrieval con match esatto
(Query \leftrightarrow Valore dei campi dei records)

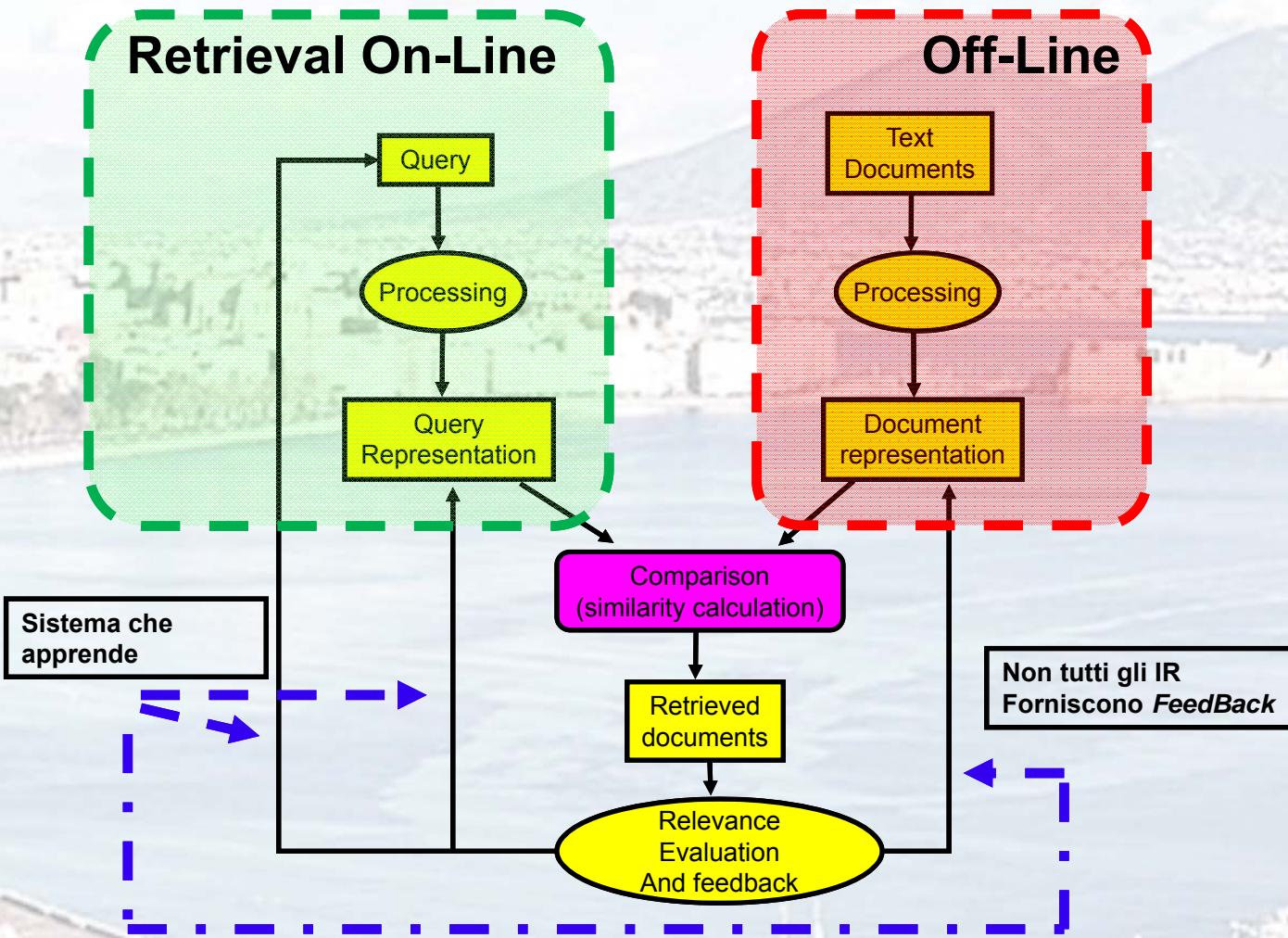
IR:

- Records non strutturati
- Attributi non prefissati
- Indicizzazione del documento:
 - keywords
 - descrittori
 - **Indici** (soggettivi ed equivoci)
- Dipendenza dal modello di rappresentazione dei documenti
- Retrieval con match **approssimato o parziale**



walter.balzano@gmail.com

Processo base del document Retrieval



walter.balzano@gmail.com

Indicizzazione automatica e modello booleano per il retrieve

Retrieve con modello booleano

● Text-pattern search systems

● Pattern

- Stringhe

- Espressioni regolari

● Strumenti

- **grep, egrep, fgrep, awk, ... (Unix).**

Esempio: home ~> grep root /etc/passwd

```
/[^a-zA-Z]a/  
| Mary had a little lamb.  
| And everywhere that Mary  
| went, the lamb was sure  
| to go.
```



Operatori Booleani OR AND NOT

Query con composizione booleana delle chiavi di ricerca

Esempio: termine_1 **ANT NOT** termine_2

walter.balzano@gmail.com

Struttura File

- Problema della scelta della rappresentazione della conoscenza:

- Flat file

- (ASCII, EBCDIC)
 - Ricerca lineare
 - grep, awk,...
 - poco efficiente

- Inverted files

- Signature files

- Generazione di “impronte” [signature]
 - metodo hash
 - Confronto tra le *signature* della query e del documento

- Alberi

- Grafi



walter.balzano@gmail.com

Inverted Files

- L'**inverted** file contiene un insieme di righe di testo
- Ogni riga contiene:
 - il *termine* che si vuole cercare
 - una sequenza di puntatori a documenti e/o records che contengono quel termine
- La parola “**inverted**” spiega quindi l’inversione del verso di ricerca: Prima la chiave e poi il documento che contiene la chiave.

Query:

(Term_1 AND Term_3)

Output: Record_3

Inverted file

Term_1	Record_1; Record_3
Term_2	Record_1; Record_2
Term_3	Record_2; Record_3; Record_4
Term_4	Record_1; Record_2; Record_3; Record_4



walter.balzano@gmail.com

Inverted Files (2)

Il processo di ricerca è più efficiente rispetto al flat-file: non si analizzano i documenti interi ma solo l'inverted file da cui si ricavano i collegamenti ai documenti che contengono la chiave (o soddisfano la query)

Query

Inverted file

Term_1	Record_1; Record_3
Term_2	Record_1; Record_2
Term_3	Record_2; Record_3; Record_4
Term_4	Record_1; Record_2; Record_3; Record_4

Output

Term_1 AND Term_3	Record_3
Term_1 OR Term_2	Record_1, Record_2, Record_3
Term_4 AND NOT Term_1	Record_2, Record_4
Term_1 AND NOT Term_4	nil



walter.balzano@gmail.com

Inverted File con operazioni estese

E' possibile raffinare la ricerca supponendo che sia significativa:

- La differenza di importanza (**peso**) tra un termine e un altro.
- La **posizione** in cui un termine compare.
- La **frequenza** in cui un termine compare all'interno di un documento.

Definiamo 2 nuovi **operatori di «prossimità»** quali

“WITHIN SENTENCE” e “ADJACENT”

con il seguente significato:

Term **i** **WITHIN SENTENCE** Term **j**

i Term **i** e **j** sono presenti nello stesso paragrafo
del record recuperato

Term **i** **ADJACENT** Term **j**

i Term **i** e **j** confinanti nel record recuperato



walter.balzano@gmail.com

Struttura generale dell'inverted file esteso:

Struttura generale dell'inverted file esteso:

Term *i*: Record_n°, Paragrafo_n°, Frase_n°, Parola_n°

Esempio di ricerca
di un Record «R»:

Inverted file:

information	R99, 10, 8, 3; R155, 15, 3, 6; R166, 2, 3, 1
retrieval	R77, 9, 7, 2; R99, 10, 8, 4; R166, 10, 2, 5

Compaiono nello
stesso paragrafo

Query:

Output:

information WITHIN SENTENCE retrieval

R99

Nota: anche R166 contiene entrambi i termini *information* e *retrieval*
ma in diversi paragrafi (paragrafo 2 e paragrafo 10) e quindi
non soddisfa la query



walter.balzano@gmail.com

Indicizzazione automatica

Il processo di indicizzazione del file di testo prevede diverse fasi; lo scopo principale è quello di filtrare il testo in modo da ottimizzare la significatività delle informazioni da considerare per le ricerche.



Fasi di filtraggio:

1. **Stop Words:** si escludono elementi insignificanti per la ricerca come “articoli”, “preposizioni”,...
2. **Stemming:** se nel testo compare più volte parole analoghe, per la ricerca si considera solo il **termine comune delle parole analoghe**. Esempio: se nel testo sono presenti i termini “pescare”, “pescatore”, “pesce” allora considero solo il loro prefisso comune, cioè “pesc”. Anche se tale operazione rende il file più compatto, d’altro canto occorre considerare che, per esempio, a causa della brevità del termine comune esso non sia più molto significativo.
3. **Thesaurus:** possibilità di sostituire diversi **termini simili** che compaiono nel testo con un unico termine (usando un vocabolario). Esempio: se nel testo sono presenti i termini “lavare”, “pulire”, “detergere” allora potrei considerare solo il termine “lavare”.
4. **Weighting:** I termini che compaiono nel testo hanno diversa importanza; la loro importanza può essere ricavata valutando le loro frequenze di occorrenza ed il risultato della ricerca viene mostrato in ordine di frequenza. NOTA: la frequenza di un termine ottenuto dopo un’operazione di Stemming o di Thesaurus è data dal numero delle occorrenze dei termini che il nuovo termine rappresenta.



walter.balzano@gmail.com

Indicizzazione automatica (STOP Words)

Esempi di STOP Words (dipendenza dal linguaggio)

A	BY	DELLO	K	OR	T	V
AD	C	DI	L	OTHER	TE	VI
AFTER	CI	E	LA	OUT	THAN	WE
AI	COL	ED	LE	OVER	THAT	WERE
AL	CON	ET	LO	P	THE	WHEN
ALL	D	F	M	PER	THEIR	WHERE
ALLA	DA	FOR	ME	Q	THEM	WHETHER
ALLE	DAGLI	FRA	MI	R	THERE	
ALLO	DAI	FROM	MORE	S	THESE	WHICH
ALSO	DAL	G	N	SAW	THEY	WHO
AN	DALL	GLI	NE	SI	THIS	WITH
AND	DALLA	H	NEGLI	SINCE	THOSE	WITHIN
AS	DALLE	HOWEVER	NEL	SU	TI	WITHOUT
AT	DALLO	I	NELL	SUCH	TO	
B	DE	IF	NELLA	SUGLI	TRA	X
BE	DEGLI	IL	NELLO	SUI	U	Y
BECAUSE	DEI	IN	NON	SUL	UN	Z
BEFORE	DEL	INTO	O	SULL	UNA	
BETWEEN	DELL	OPPURE	OF	SULLA	UNDER	
BIG	DELLA	J	OPPURE	SULLE	UNO	
BUT	DELLE				UPON	



walter.balzano@gmail.com

Indicizzazione automatica (pesi dei termini)

Esempio di rappresentazione dei pesi nell'inverted file:

Termine: Record1,Peso1; Record3,Peso3;

Term1: R1,0.3; R3,0.5; R6,0.8; R7,0.2; R11,1

Term2: R2,0.7; R3,0.6; R7,0.5; R9,0.5

Term3: R1,0.8; R2,0.4; R9,0.7

il risultato della ricerca viene mostrato in ordine di peso

Operazioni Booleane con i pesi dei termini

Operatore	Descrizione	esempio di Query	lista Output
OR	Si considera il peso maggiore dei records che contengono il termine della query	Term2 OR Term3	R1, R2, R9, R3, R7
AND	Si considera il peso minore dei records che contengono il termine della query	Term2 AND Term3	R9, R2
NOT	Si considera la differenza dei pesi dei termini che contengono il termine della query	Term2 AND NOT Term3	R3, R7



Indicizzazione automatica (calcolo dei pesi)

- Il peso di un termine deve considerare il numero di volte in cui il termine compare sia nel documento sia nell'insieme complessivo dei documenti.
- La formula comunemente utilizzata per il calcolo dei pesi è:

$$W_{ij} = tf_{ij} * \log(N/df_j)$$

In cui:

- W_{ij} = peso del termine j nel documento i
- tf_{ij} = frequenza del termine j nel documento i
- N = numero totale dei documenti del DataBase
- df_j = numero dei documenti del DataBase che contengono il termine j

NOTA:

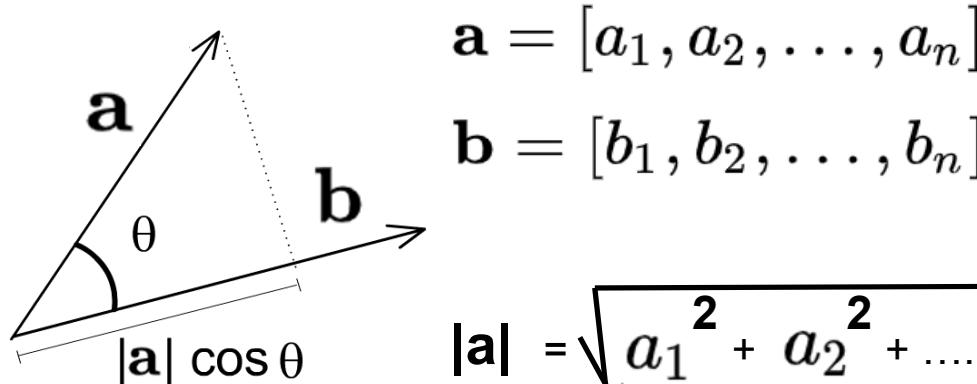
- se il termine è presente in tutti i documenti del DataBase allora $df_j = N$ da cui deriva che $W_{ij} = 0$ che significa che quel termine non ha un alto potere discriminante e quindi non è significativo per un buon filtraggio.
- Se il termine è parzialmente presente nei documenti del DataBase allora il suo potere discriminante è elevato e quindi è significativo per il filtraggio dei documenti



walter.balzano@gmail.com

Retrieval con modello spazio vettoriale

Prodotto Scalare


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

Possiamo calcolarlo in 2 modi diversi:

1. $|\mathbf{a}| |\mathbf{b}| \cos \theta$
2. $\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i$



walter.balzano@gmail.com

Retrieval con modello spazio vettoriale

- il modello spazio vettoriale presume l'esistenza di un determinato insieme di termini che rappresentino i Documenti e le Query: un documento D_i ed una Query Q_j sono definite nel seguente modo

■ $D_i = [T_{i1}, T_{i2}, \dots, T_{ik}, \dots, T_{iN}]$

■ $Q_j = [Q_{j1}, Q_{j2}, \dots, Q_{jk}, \dots, Q_{jN}]$ in cui:

- T_{ik} è il peso del k-esimo termine relativo al documento i
- Q_{jk} è il peso del k-esimo termine relativo alla query j
- N è il numero totale di termini usati nel documento e nella Query

Calcolo della similarità

$$S(D_i, Q_j) = \sum_{K=1}^N T_{ik} \cdot Q_{jk}$$

Calcolo della similarità normalizzata

$$S(D_i, Q_j) = \frac{\sum_{K=1}^N T_{ik} \cdot Q_{jk}}{\sqrt{\sum_{K=1}^N T_{ik}^2 \cdot \sum_{K=1}^N Q_{jk}^2}}$$

Prodotto
dei moduli dei
vettori
 T_{ik} e Q_{jk}

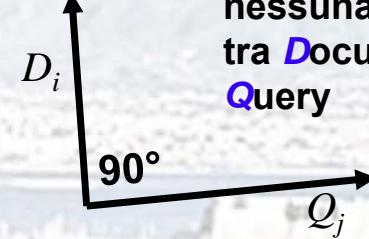
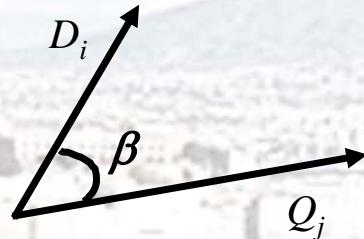
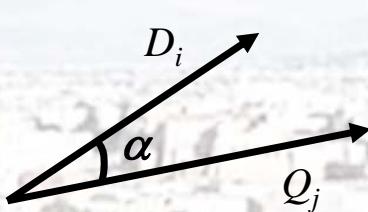
NOTA: lo scopo della normalizzazione è quello di ottenere un parametro di valutazione che sia indipendente dalle dimensioni N dei vettori.

Di fatto il valore di $S(D_i, Q_j)$ normalizzato corrisponde al valore del coseno dell'angolo che formano i due vettori D_i e Q_j



walter.balzano@gmail.com

Significato geometrico del modello spazio vettoriale



Vettori ortogonali → nessuna attinenza tra **Documento** e **Query**

$$S(D_i, Q_j) = \sum_{K=1}^N T_{ik} \cdot Q_{jk} = |D_i| \cdot |Q_j| \cdot \cos(\hat{D}_i Q_j)$$

- L'angolo che formano i due vettori influisce sul prodotto scalare
- Angolo maggiore → prodotto scalare minore, minore similitudine tra Query e Documento
- Angolo minore → prodotto scalare maggiore, maggiore similitudine tra Query e Documento



walter.balzano@gmail.com

Retrieval con modello spazio vettoriale – esempio

- Siano assegnati 4 documenti di cui i corrispondenti vettori sono:
 - $D_1 = [0.2, 0.1, 0.4, 0.5]$
 - $D_2 = [0.5, 0.6, 0.3, 0]$
 - $D_3 = [0.4, 0.5, 0.8, 0.3]$
 - $D_4 = [0.1, 0, 0.7, 0.8]$
- Sia assegnata una Query di cui il vettore corrispondente è
 - $Q = [0.5, 0.5, 0, 0]$

Allora le misure di similarità derivate sono:

$$S(D_1, Q) = 0.31$$

$$S(D_2, Q) = 0.93$$

$$S(D_3, Q) = 0.66$$

$$S(D_4, Q) = 0.07$$

Il sistema presenterà i documenti
in ordine decrescente e cioè: D_2, D_3, D_1, D_4



walter.balzano@gmail.com

Tecniche basate su Relevance feedback

L'idea di base di questo modello consiste nello sfruttare le valutazioni degli utenti (**feedback**) permettendo di raffinare i pesi dei termini sia della **query** sia dei **documenti**

Query Modification

- La modifica di una query beneficia solo l'utente che l'ha formulata
- I termini che sono stati ritenuti rilevanti vengono aggiunti alla query originale oppure viene aumentato il peso di quel termine
- I termini che non sono stati ritenuti rilevanti vengono cancellati dalla query oppure viene diminuito il peso di quel termine.
- La regola applicata è (**Formula di Rocchio**):

$$\overrightarrow{Q_m} = (a \cdot \overrightarrow{Q_o}) + \left(b \cdot \frac{1}{|D_r|} \cdot \sum_{D_j \in D_r} \overrightarrow{D_j} \right) - \left(c \cdot \frac{1}{|D_{nr}|} \cdot \sum_{D_k \in D_{nr}} \overrightarrow{D_k} \right)$$

Variable	Value
$\overrightarrow{Q_m}$	Modified Query Vector
$\overrightarrow{Q_o}$	Original Query Vector
$\overrightarrow{D_j}$	Related Document Vector
$\overrightarrow{D_k}$	Non-Related Document Vector

Variable	Value
a	Original Query Weight
b	Related Documents Weight
c	Non-Related Documents Weight
D_r	Set of Related Documents
D_{nr}	Set of Non-Related Documents

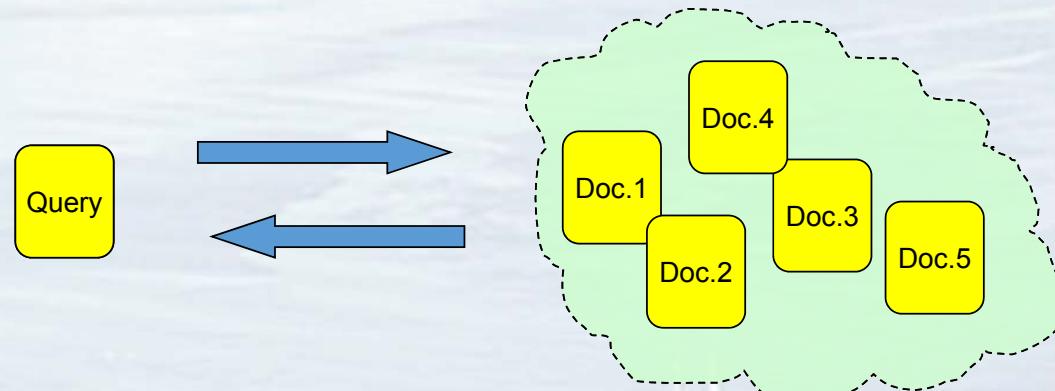


walter.balzano@gmail.com

Tecniche basate su Relevance feedback (2)

Document Modification

- La modifica di un documento beneficia anche altri utenti
- I valori dei pesi dei documenti vengono aggiornati considerando la query stessa (e non considerando la lista dei documenti giudicati rilevanti dall'utente)
- Vengono aumentati i pesi dei termini che si trovano sia nella query che nell'insieme dei documenti rilevanti
- Vengono diminuiti i pesi dei termini che si trovano nell'insieme dei documenti rilevanti ma non nella query



walter.balzano@gmail.com

Altri modelli per il Retrieval (1)

il modello probabilistico

- Considera la dipendenza dei termini e le loro principali relazioni
- Ha avuto un scarso successo a causa della difficoltà di calcolo delle probabilità su cui esso è fondato.

il modello basati su Cluster

- Il modello cluster è basato sull'idea di definire un insieme di gruppi in cui ciascun gruppo contiene elementi "simili". La query cercherà poi all'interno di un cluster



walter.balzano@gmail.com

Altri modelli per il Retrieval (2)

Generazione del Cluster

■ Similarità per coppie

- Ogni documento è rappresentato come un vettore
- Si calcola la similitudine per ogni coppia di documenti e si popola una matrice delle distanze
- Inizialmente ogni documento viene posizionato in un cluster (cluster con un sol documento)
- La coppia meno distante formerà un agglomerato; si rimuove la coppia dalla matrice delle distanze sostituendola con l'oggetto "congiunto" (il centroide del cluster)
- Si valutano tutte le distanze dei cluster rispetto il nuovo cluster
- Si ripete il processo finché la matrice di distanza sarà costituita da un solo elemento (i documenti saranno raggruppati in un unico super cluster).
- E' un metodo che ha una complessità temporale elevata ma la generazione prodotta è unica

■ Clustering Euristico

- Si considera il primo documento; esso costituirà il primo cluster
- Successivamente ogni documento viene confrontato con i cluster generati e viene quindi collocato nel cluster più "vicino"; in caso di distanza eccessiva si fonda un nuovo cluster.
- Si ripete il processo fino ad esaurimento dei documenti
- E' un metodo non molto complesso ma la generazione prodotta non gode di univocità (dipende infatti dall'ordine in cui i documenti sono considerati)

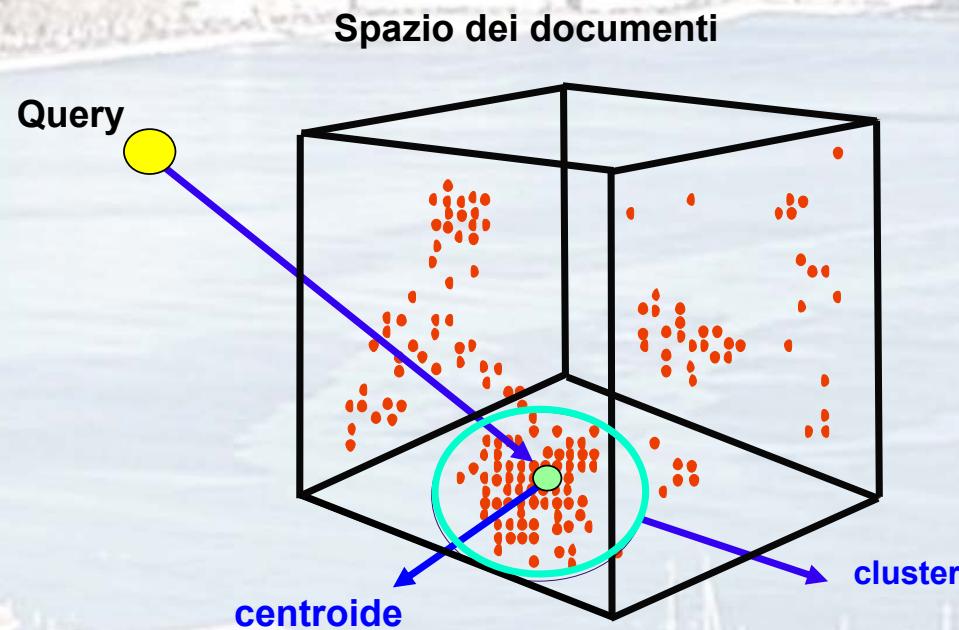


walter.balzano@gmail.com

Altri modelli per il Retrieval (3)

Cluster-Based Retrieval

- Dopo aver generato i cluster sia le ricerche che il recupero dei documenti risulteranno efficienti ed efficaci.
- Ogni Cluster è caratterizzato dal proprio centroide che è calcolato mediando tutti i vettori del cluster
- Il processo di retrieval confronta la query con i centroidi dei cluster e selezionerà il cluster più simile
- L'output sarà costituito dai documenti del cluster individuato (tutti i documenti del cluster, oppure i documenti del cluster con maggior grado di similitudine).



walter.balzano@gmail.com

Metodi IR non tradizionali

- I principali IR sono basati su misurazioni statistiche per la presenza dei termini.
- Numerose sono le evidenze che però confermano i limiti della metodologia statistica.
- Ambiguità del Linguaggio Naturale :
 - Query in linguaggio naturale (NLP, Natural Language Processing)
 - Il risultato di una query dipende dall'ordine in cui termini stessi compaiono nella query
 - I termini della query possono assumere significati diversi in relazione alla posizione che occupano.
 - Polisemia: uno stesso termine può assumere diversi significati
 - Sinonimia: termini diversi possono assumere lo stesso significato
 - Il significato di una Query può andar ben oltre i suoi “crudi” termini
- Metodi ad hoc
 - Introduzione di conoscenza aggiuntiva:
 - Conoscenza specifica del dominio (termini specifici,...)
 - Conoscenza specifica dell' utente che formula la query (le sue preferenze dovrebbero influenzare il processo di retrieving)



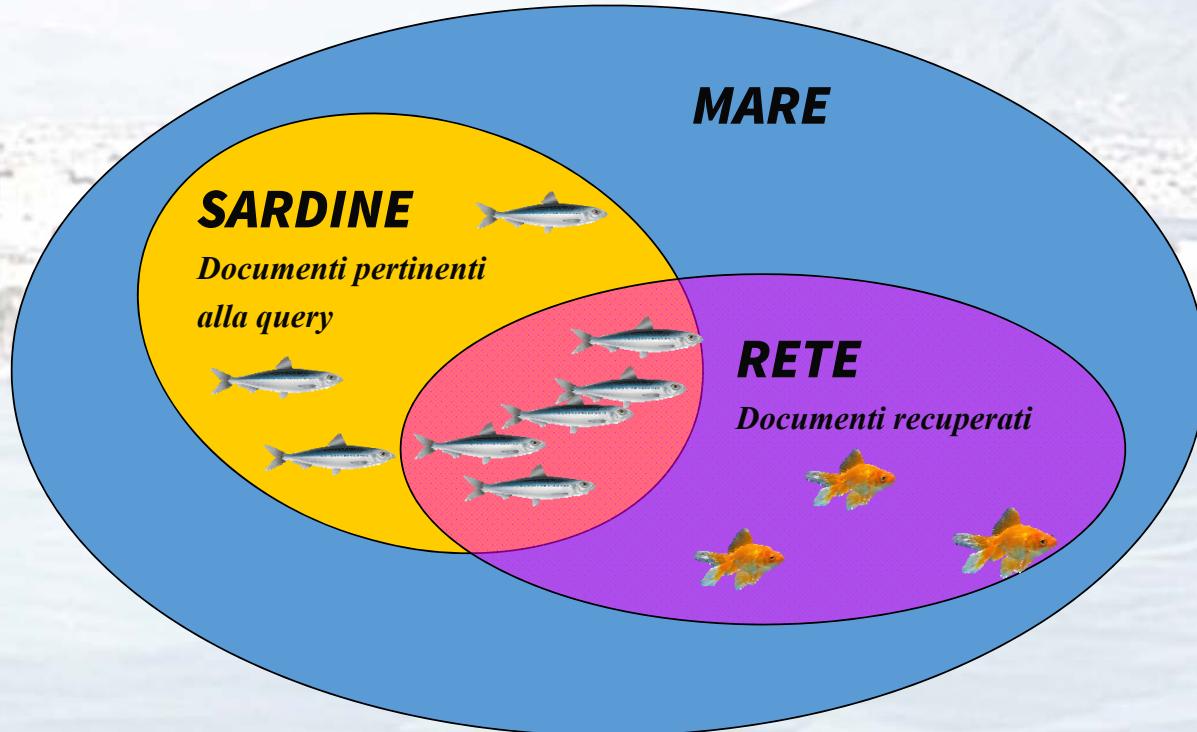
walter.balzano@gmail.com

Misurazione delle prestazioni (1)



walter.balzano@gmail.com

Misurazione delle prestazioni (2)



walter.balzano@gmail.com

Misurazione delle prestazioni (3)

- **Velocità di ricerca**

- **Recall:** capacità di recuperare informazioni rilevanti dal database. Si definisce come rapporto tra il numero dei documenti rilevanti recuperati ed il numero totale di elementi rilevanti nel database

- **Precisione:** misura l'accuratezza dei documenti recuperati. Si definisce come rapporto tra il numero di documenti rilevanti recuperati ed il numero totale di documenti recuperati

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

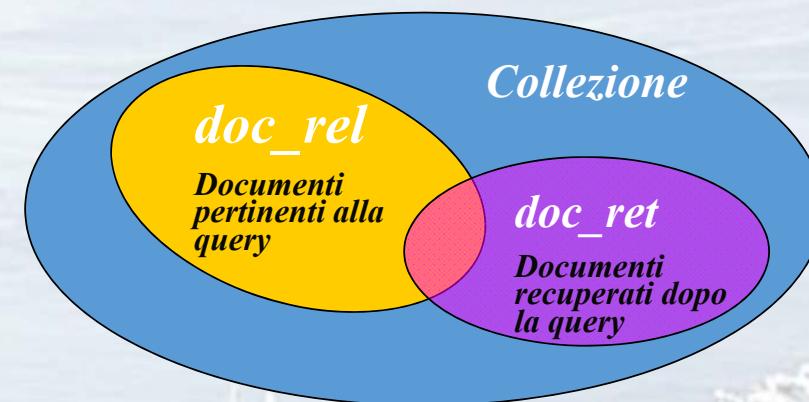
$$recall = \frac{|doc_ret \cap doc_rel|}{|doc_rel|}$$

(*retrieved*)

doc_ret = *documenti recuperati*

(*related*)

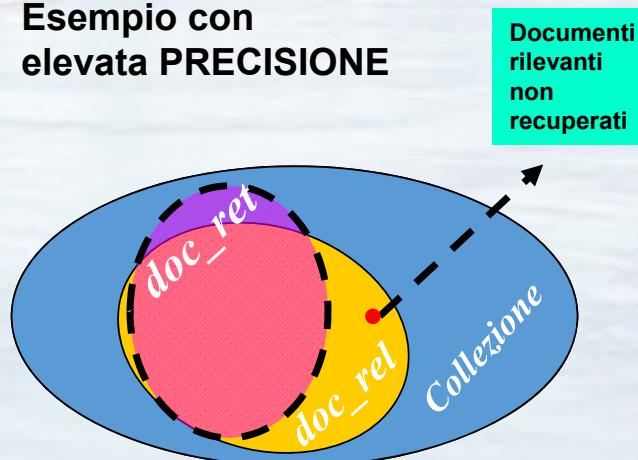
doc_rel = *documenti pertinenti*



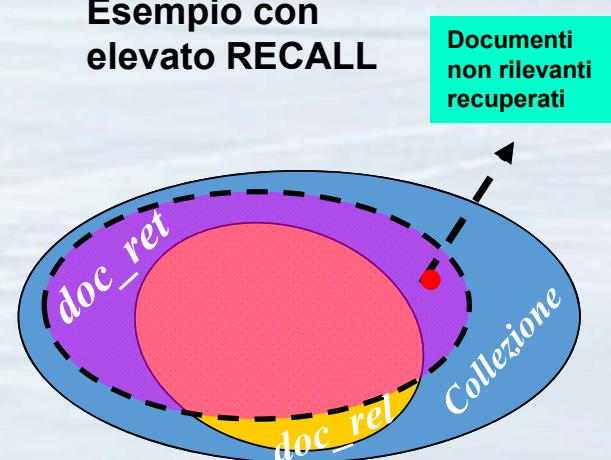
Misurazione delle prestazioni (4)

- Nella pratica, i parametri di valutazione recall e precisione vengono **valutati in modo congiunto**, ed una buona prestazione è indice di un oculato compromesso.
- Tipicamente quanto maggiore è la recall tanto minore risulta la precisione: infatti quanto maggiore è lo sforzo della query per prelevare tutti i documenti rilevanti, tanto maggiore sarà la probabilità che in output andranno anche documenti non rilevanti (e quindi imprecisione). Viceversa: quanto maggiore è la precisione tanto minore risulta la recall: per evitare di prendere elementi non rilevanti si finisce inevitabilmente per non prelevare qualche documento rilevante.

Esempio con elevata PRECISIONE

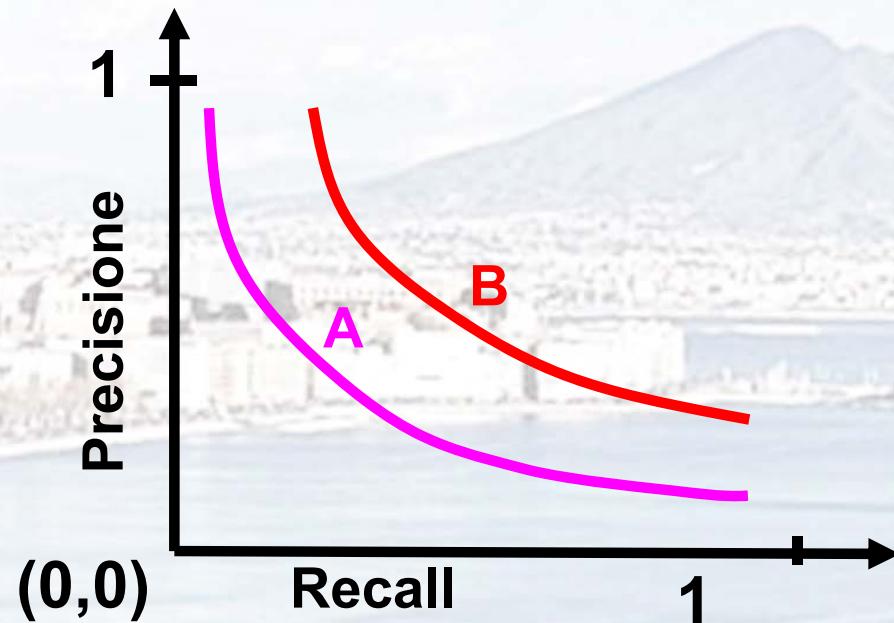


Esempio con elevato RECALL



walter.balzano@gmail.com

Misurazione delle prestazioni (5)



Esempio in cui
la prestazione del sistema B
è migliore
della prestazione del sistema A

B ha una maggiore distanza da (0,0) Rispetto ad A



walter.balzano@gmail.com

Tecniche di IR a confronto

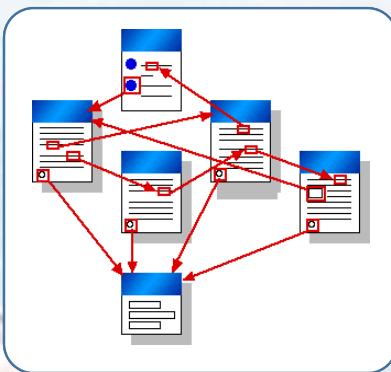
- L'indicizzazione automatica ha una prestazione simile a quella manuale ed un reale miglioramento si ottiene combinando entrambe le tecniche
- Impiegando un insieme di query simili la performance del recupero mediante confronti (match) parziali si è dimostrata essere migliore del match booleano esatto.
- Il modello probabilistico e spazio vettoriali hanno performance analoghe.
- Le tecniche di recupero basate su cluster e sul modello probabilistico hanno performance analoghe ma recuperano insiemi diversi di documenti.
- Se al primo tentativo non si è riusciti a recuperare documenti rilevanti allora l'uso di *Relevance feedback* consente un reale miglioramento della prestazione
- L'uso di uno specifico dominio di conoscenza e del profilo utente che effettua la query, produce un significativo miglioramento della prestazione.



walter.balzano@gmail.com

I motori di Ricerca WWW

- I motori di ricerca sul WWW costituiscono molto probabilmente il tipo di applicazione più utilizzato sul Web
- I documenti Web sono memorizzati come IPERTESTI in formati quali HTML
- I documenti sono strutturati in **NODI**, **LINK** ed **ANCHOR**.
 - Generalmente ogni nodo rappresenta un singolo concetto o un idea
 - I Link collegano diversi nodi
 - Le aree evidenziate “cliccabili” sono dette ANCHOR ed indicano l'esistenza di un LINK.
 - Selezionando un ANCHOR si attiva il LINK corrispondente e il NODO cui fa riferimento viene caricato e visualizzato

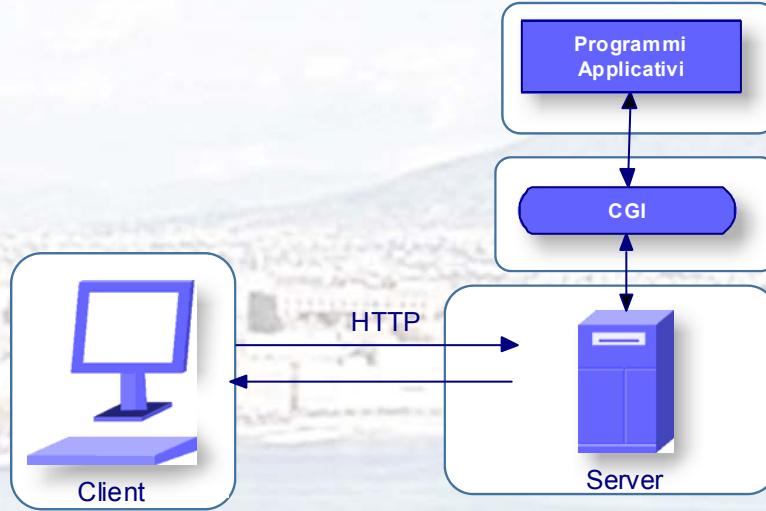


- Un **ipertesto** è un modo di organizzare l'informazione che permette un accesso non sequenziale
- L'**Ipermedia** è una stensione dell'ipertesto in cui le ancora e nodi possono essere di tipo qualsiasi
- Il **WWW** è una estensione geografica dell'ipermedia



walter.balzano@gmail.com

Modello di Comunicazione



- Diverse sono le regole di base che governano la comunicazione (i protocolli):
- HTTP, HTTPs, FTP, FTPs,
- Mediante l'interfaccia di un browser, il client effettua la richiesta di un documento al server che attraverso il versatile strato delle CGI (Common Gateway Interface) accede a svariate librerie di funzioni e/o servizi.
- Finita l'elaborazione sul server i risultati vengono reindirizzati al client che ne aveva fatto richiesta.



walter.balzano@gmail.com

Il web

- URL (Uniform Resource Locator): indirizzo che univocamente identifica ogni documento della rete:
 - *Protocol://Servername[:port]/Path/Document-name*
Es.: *http://www.na.infn.it/~walzano/index.html*
ftp://ftp.monash.edu.au/pub/internet/readme.txt

Crawler (detto anche spider o robot)

- Tipo di bot (programma o script che automatizza delle operazioni) che analizza i contenuti di una rete (o di un database) in un modo metodico e automatizzato, in genere per conto di un motore di ricerca nel Web.
- Solitamente acquisisce una copia testuale di tutti i documenti visitati e li inseriscono in un indice.
- Si basa su una lista di URL da visitare fornita dal motore di ricerca (il quale, inizialmente, si basa sugli indirizzi suggeriti dagli utenti o su una lista precompilata dai programmatore stessi).
- Durante l'analisi di un URL, identifica tutti gli hyperlink presenti nel documento e li aggiunge alla lista di URL da visitare. Il processo può essere concluso manualmente o dopo che un determinato numero di collegamenti è stato seguito.

Esempi di nomi degli spider dei principali motori di ricerca

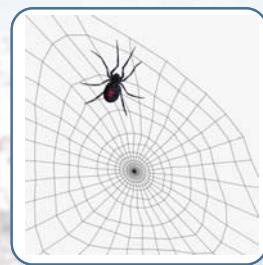
Nome Spider	Motore di ricerca
googlebot	Google
fast	Fast - Alltheweb
slurp	Inktomi – Yahoo!
scooter	Altavista
ask Jeeves	Ask Jeeves
teoma_agent	Teoma
ia_archiver	Alexa – Internet Archive
googlebot	Yahoo



walter.balzano@gmail.com

Il web (2)

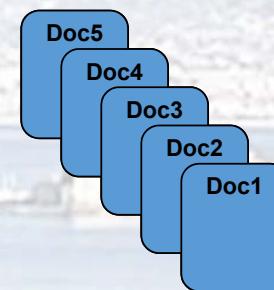
Motore di ricerca



Recupero
Dei documenti



Analisi ed
indicizzazione



Ranking

- Poiché i documenti selezionati sono comunque tanti, per poterli mostrare all'utente è necessario comunque organizzarli (Ranking)
- Poiché i documenti selezionati sono comunque tanti, la caratteristica di precisione del motore è più importante della caratteristica di recall (anche perché partendo dai "pochi" documenti selezionati, posso navigare e raggiungere gli altri documenti)
- Un modo per organizzare i documenti, assegnando loro dei pesi, è basato sulla formattazione del testo HTML



walter.balzano@gmail.com

Indicizzazione e Recupero dell'Audio

- 1. Introduzione**
- 2. Proprietà e caratteristiche principali**
- 3. Classificazione dell'audio**
- 4. Recupero e riconoscimento del parlato**
- 5. Indicizzazione e recupero della musica**
- 6. Indicizzazione e recupero
mediante l'uso di relazioni
tra l'audio ed altri Media.**



walter.balzano@gmail.com

Introduzione

Per gli esseri umani risulta semplice saper distinguere tra

- differenti tipi di audio (voce, musica, rumori,...)
- diverse “velocità di esecuzione” (lento, svelto,...)
- diversi toni (felice, arrabbiato, triste,...).

- ...inoltre, per un essere umano, risulta semplice determinare la similitudine tra diversi pezzi audio.
- Attualmente, il modo più utilizzato per classificare vari brani audio è basato sul titolo oppure sul nome del file: è evidente sia la **soggettività** di tale metodologia e sia l'incapacità di poter supportare query come “cercare un brano audio simile a quello che abbiamo appena ascoltato (query per esempio).



walter.balzano@gmail.com

Introduzione

- Ulteriore complicazione nasce dal fatto che **non esiste uno standard di memorizzazione** del file audio (diverse frequenze di campionamento, ampiezze,...) e ciò comporta ovviamente grosse problematiche per il loro confronto.
- Risulta pertanto necessario sviluppare tecniche e metodi di retrieving.



walter.balzano@gmail.com

Approccio generale

- L'approccio che viene generalmente seguito per lo studio dei vari brani audio consiste nel realizzare un primo livello di distinzione:
 - Parlato
 - Musica
 - Rumori
- La gestione è quindi diversificata per tipologie: per esempio, lo studio del parlato consiste principalmente nel convertire il file audio in parole di testo (*speech recognition*) su cui si potrà poi effettuare una query tradizionale (tramite un IR)
- Le query audio vengono gestite su sottoinsiemi simili di brani audio.



walter.balzano@gmail.com

Approccio generale

La **distinzione tra diverse tipologie audio è importante** per diverse ragioni:

- Diversi tipi di audio richiedono **differenti tecniche di indicizzazione e recupero**
- Diversi tipi di audio assumono significati diversi per le applicazioni
- Per il parlato esistono già numerosi applicativi per il riconoscimento
- La classificazione è di per se utile per molte applicazioni
- La classificazione comporta poi una riduzione dello spazio di ricerca



Proprietà e caratteristiche principali dell'audio

- I segnali audio vengono rappresentati:
 - Nel **dominio temporale** [*time domain*] (rappresentazione Tempo/Aampiezza)
 - Nel **dominio delle frequenze** (rappresentazione Frequenza/Magnitudine)
- Ciascun tipo di rappresentazione è particolarmente idonea per l'estrazione di determinate caratteristiche
- Oltre alle caratteristiche estraibili dall'audio rappresentato nei due precedenti domini, è possibile estrarre **caratteristiche che possono essere soggettive (come il timbro,...)**



walter.balzano@gmail.com

Caratteristiche derivabili dal Time Domain

- 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.
- Si assume che ogni campione audio sia rappresentato mediante un insieme di 16 bit; ciò comporta che l'escursione dei valori varia del segnale varia da $32767 (2^{15}-1)$ a -32767

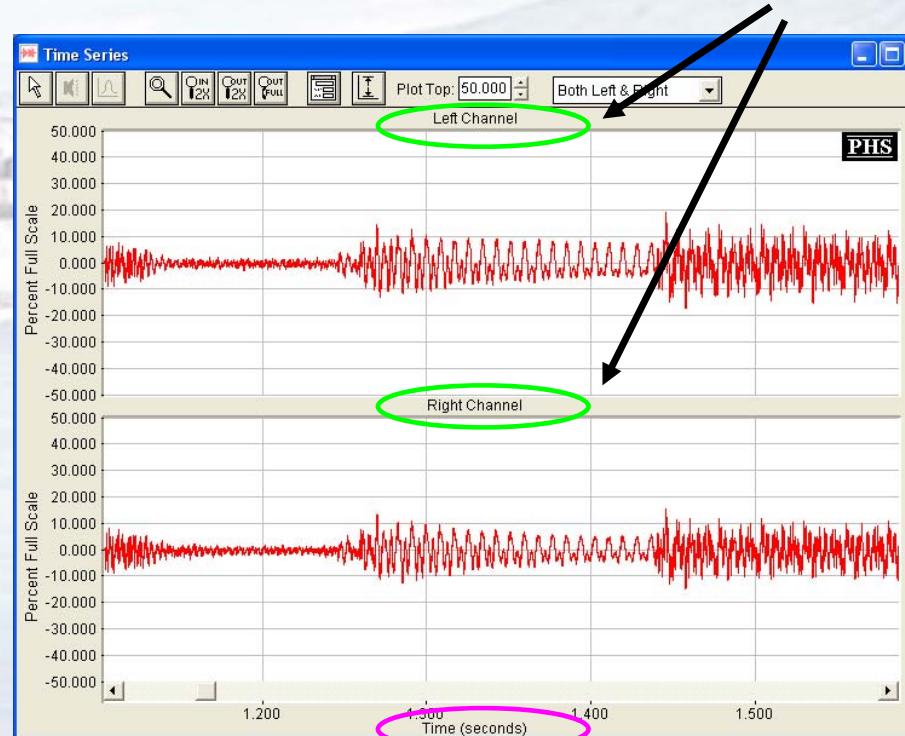


walter.balzano@gmail.com

Esempio di Time Domain

Canale sinistro e destro

Ampiezza



Tempo



walter.balzano@gmail.com

Caratteristiche derivanti dal Time Domain

Average Energy (*Energia Media*)

Indica la “rumorosità” del segnale audio ed è calcolabile mediante la relazione:

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

In cui:

E = energia media del brano audio

N = numero totale dei campioni valutati

$x(n)$ = valore del campione n -esimo



Caratteristiche derivanti dal Time Domain

Zero Crossing Rate

(*Frequenza di passaggio per lo zero*)

Indica con quale frequenza cambia segno
l'ampiezza del segnale ed è calcolabile
mediante la relazione:

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

$\operatorname{sgn} x(n)$ = segno di $x(n)$ =  $1 \text{ se } x(n) > 0$
 $-1 \text{ se } x(n) < 0$



Caratteristiche derivanti dal Time Domain

Silence Ratio (*Quantità di silenzio*)

Indica la proporzione di silenzio nel brano musicale; è il periodo entro il quale i valori assoluti di ampiezza di un certo numero di campioni (e non solo un singolo valore) e per un “certo” tempo siano prossimi ad una soglia specifica.

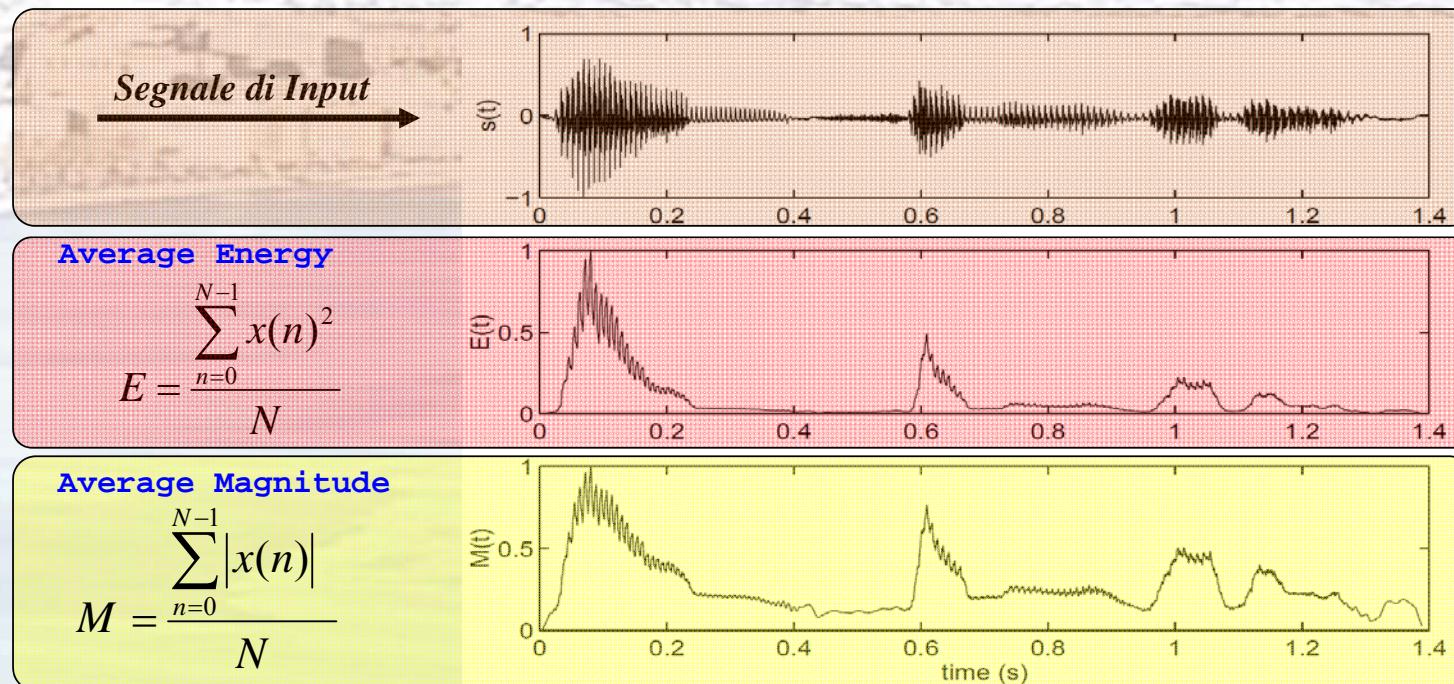
$$\text{Silence Ratio} = \frac{\text{Somma dei periodi di silenzio}}{\text{Lunghezza totale del brano}}$$



walter.balzano@gmail.com

Caratteristiche derivanti dal Time Domain (Energy)

Poiché l' *Average Energy* aumenta fortemente per grandi ampiezze di segnale (i campioni compaiono elevati al quadrato) si introduce un'ulteriore grandezza, l'*Average Magnitude*.



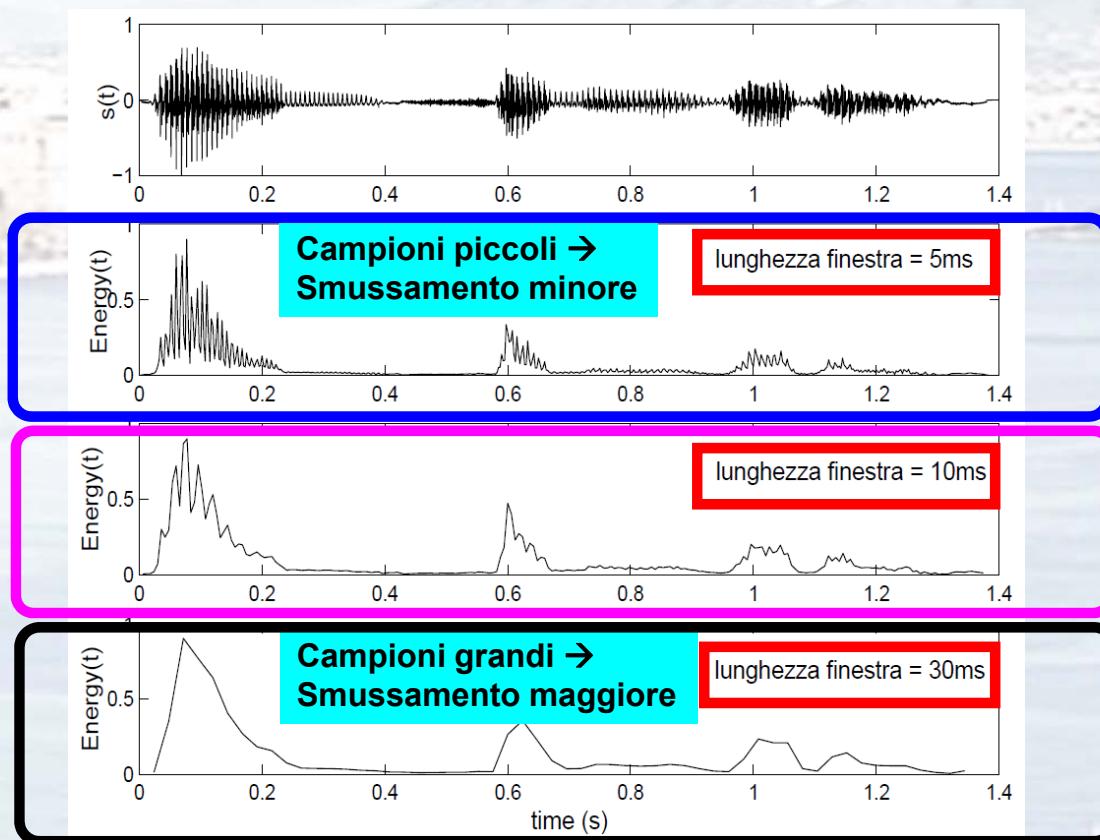
Nota: nell'esempio E ed M sono normalizzate rispetto il valore massimo (con 100 campioni e con stesso frame Rate di campionamento del segnale).



walter.balzano@gmail.com

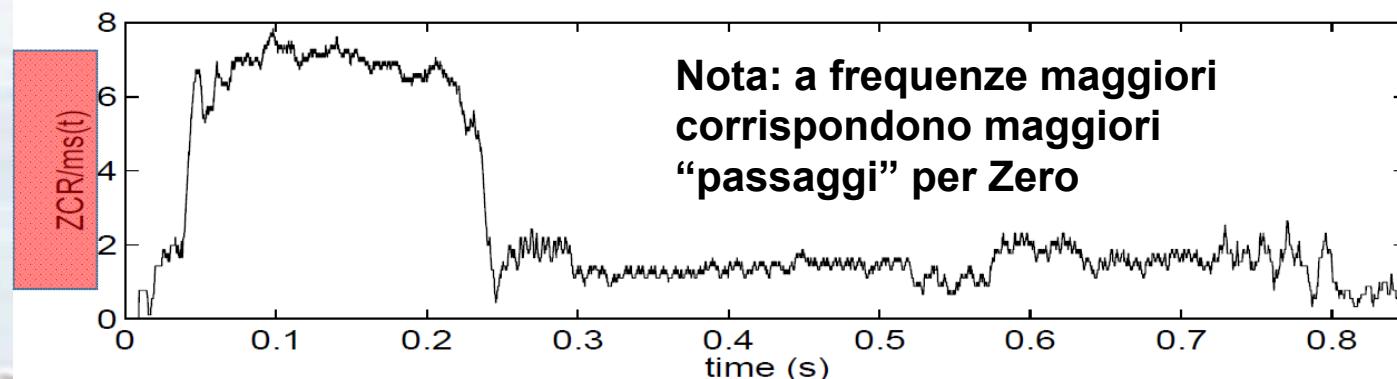
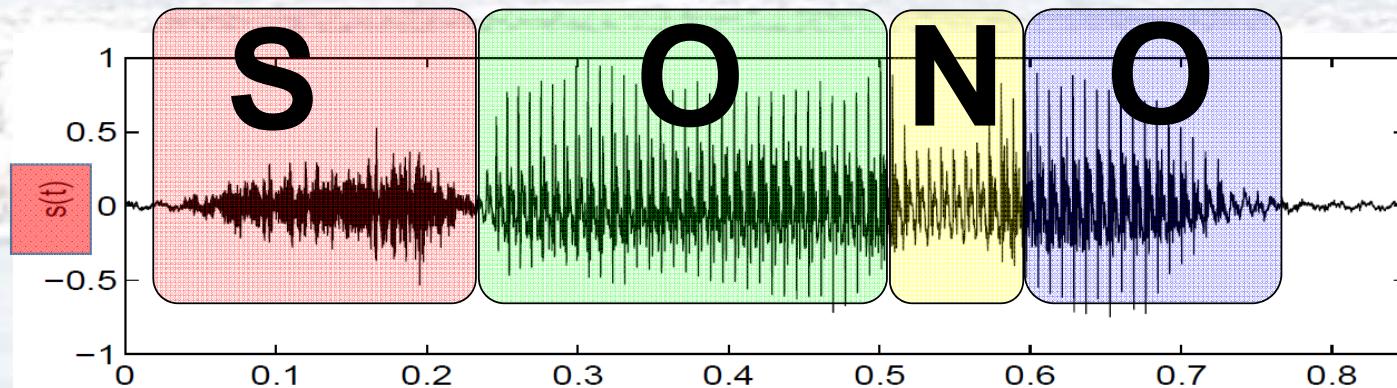
Caratteristiche derivanti dal Time Domain (Energy - 2)

Si noti che fissata la frequenza di Campionamento e facendo invece variare la grandezza del campione (cioè la lunghezza della finestra), la funzione Energia assume un andamento più “smussato”



Caratteristiche derivanti dal Time Domain (ZCR)

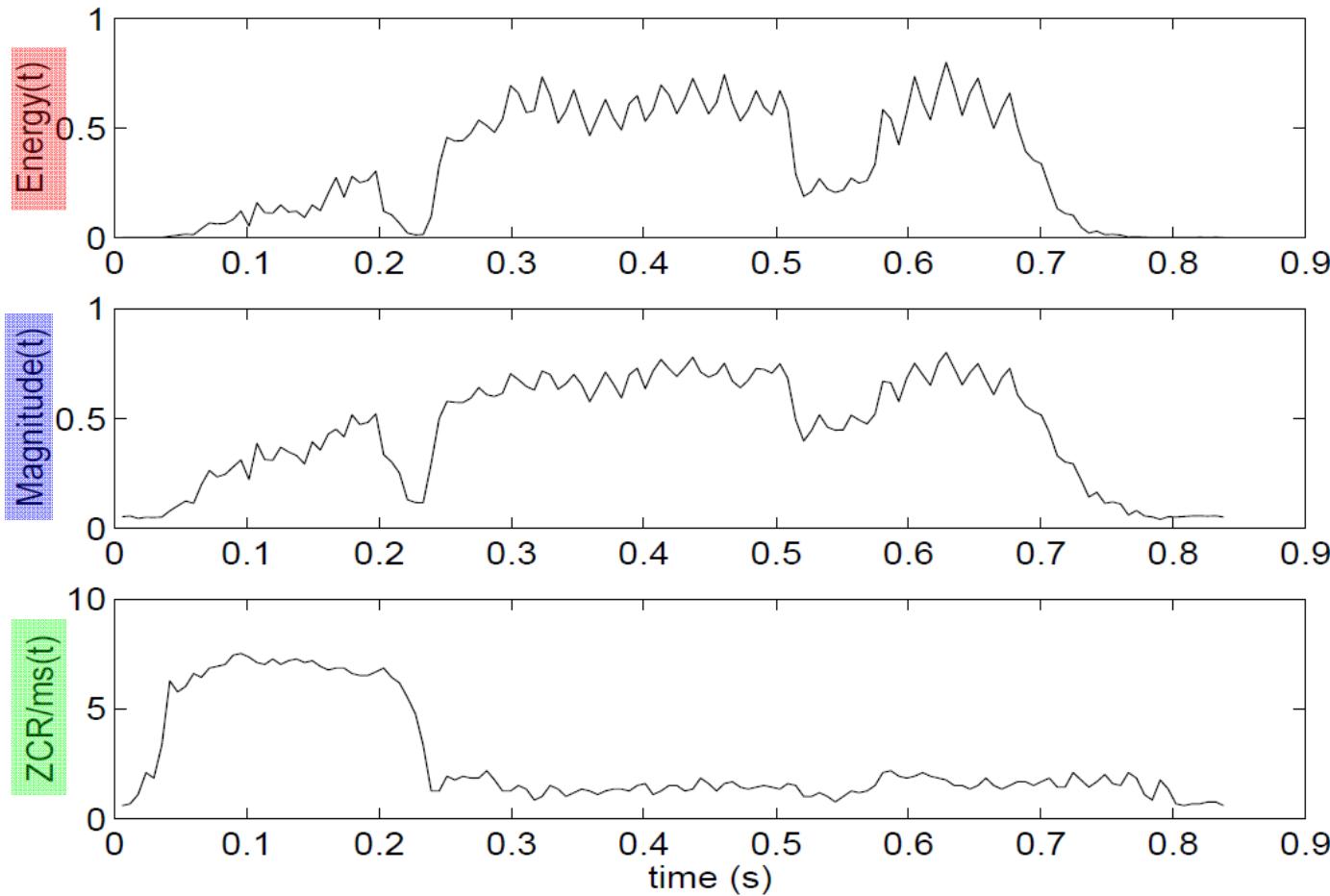
I grafici nell'esempio sottostante sono relativi all'espressione "SONO" calcolata con una finestra rettangolare di N=100 campioni e frame rate pari a quella del segnale (11KkHz); si noti come si riesca a discriminare la "S" iniziale (suono non vocalizzato) dal resto della parola (suono vocalizzato)



walter.balzano@gmail.com

Caratteristiche derivanti dal Time Domain (confronti)

ZCR, Magnitudine ed Energia a confronto



walter.balzano@gmail.com

Caratteristiche derivanti dal Dominio delle Frequenze

- La rappresentazione nel dominio delle frequenze deriva dalla rappresentazione nel dominio temporale applicando la trasformata di Fourier



- La trasformata di Fourier descrive in che modo un segnale possa essere decomposto nelle proprie frequenze componenti
- Nel Dominio delle Frequenze il segnale viene rappresentato come ampiezza che varia in dipendenza della frequenza; in altre parole tale rappresentazione mostra in che modo è distribuita l'energia alle varie frequenze.

La rappresentazione nel dominio delle Frequenze è comunemente detta **Spettro del segnale**



walter.balzano@gmail.com

Trasformata di Fourier

La Trasformata di Fourier **DFT** è data dalla formula:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-jn\omega_k}$$

In cui:

$$\omega_k = 2\pi k/N$$

$x(n)$ = segnale discreto con N campioni

k = valore binario della DFT

Se la frequenza di campionamento
del segnale è f_s Hz allora la
frequenza f_k del binario k espressa
in Herz è data da:

$$f_k = f_s \frac{\omega_k}{2\pi} = f_s \frac{k}{N}$$

Se $x(n)$ è temporalmente
limitato da N
allora può essere completamente
ricostruito usando la **Trasformata di Fourier Discreta
Inversa IDFT** per la frequenza di N campioni

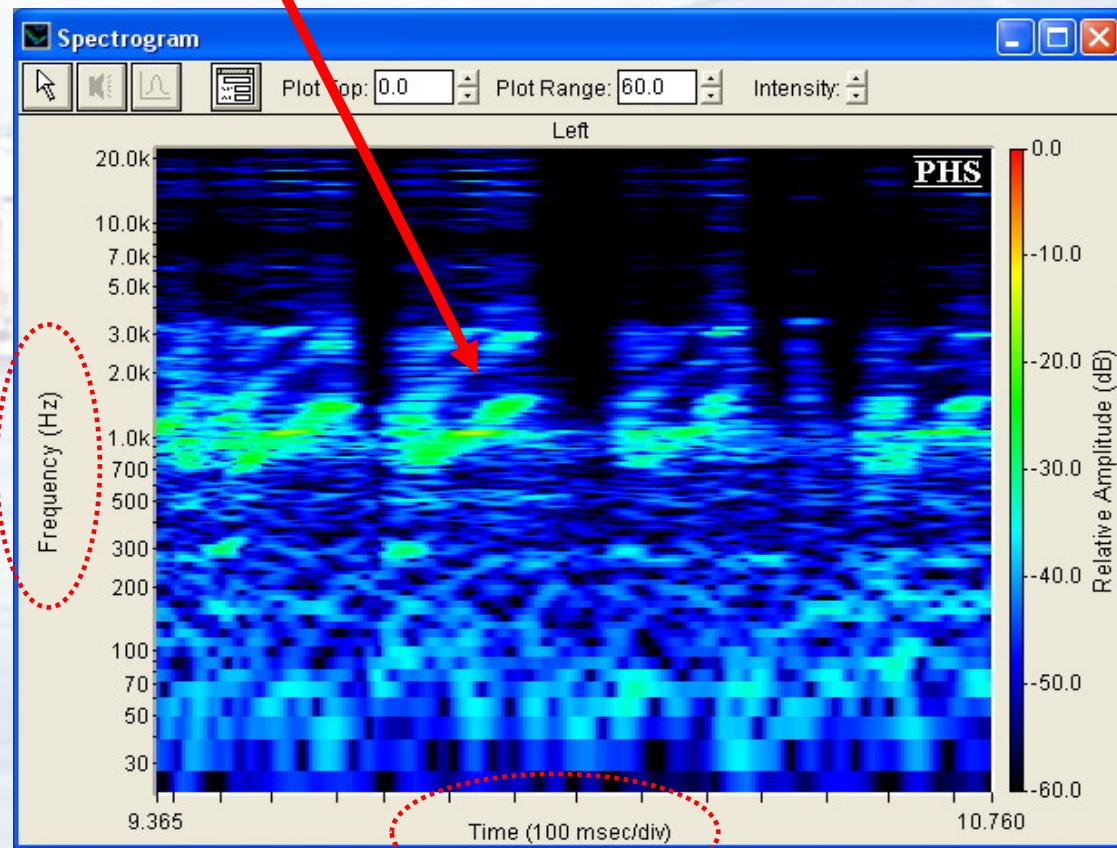
La DFT e la IDFT sono calcolate in modo efficiente
tramite la Fast Fourier Transforms **FFT**



walter.balzano@gmail.com

Spettro di un segnale audio

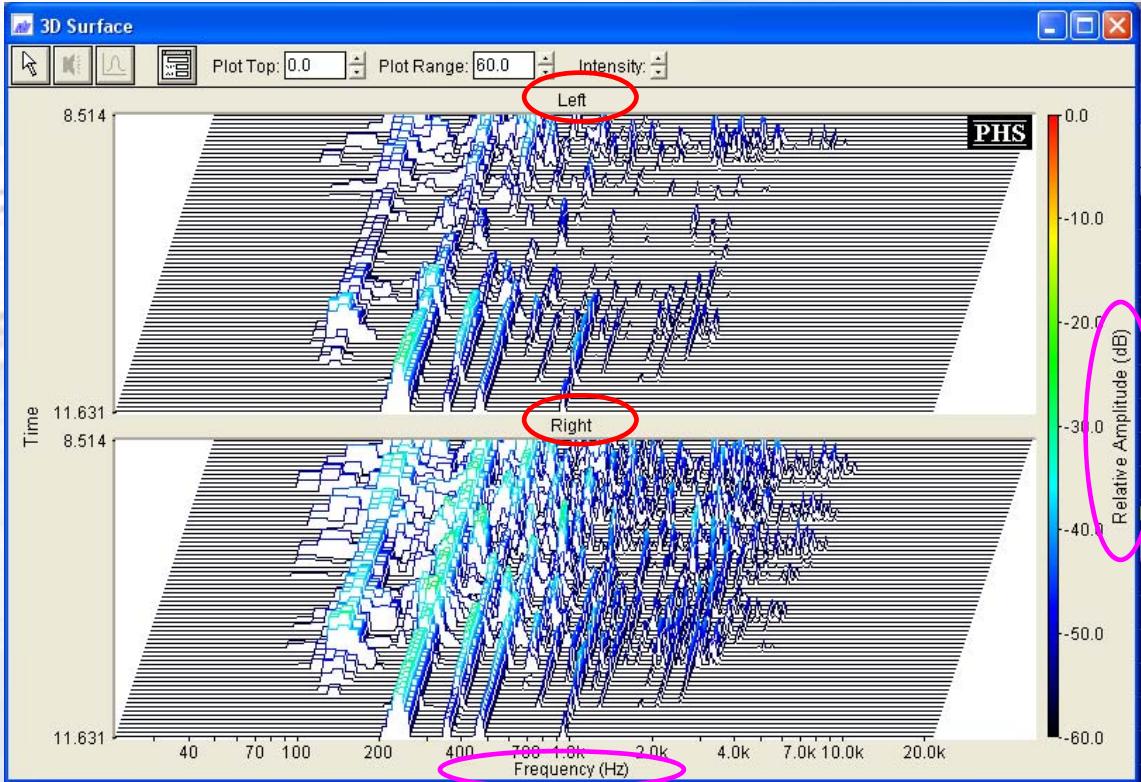
Il colore rappresenta l'ampiezza



walter.balzano@gmail.com

Spettrogramma (2)

Ampiezza



Frequenza

Tempo



walter.balzano@gmail.com

Caratteristiche derivanti dal Dominio delle Frequenze (2)

- **Bandwidth** (Larghezza di Banda) = gamma (o *range*) delle frequenze di un suono
- Mediamente la musica ha un range molto più ampio del parlato
- Il metodo più semplice adottato per il calcolo del range effettua una differenza:

Range_frequenze =
Massima_Frequenza – Minima_Frequenza

Con:

Massima_Frequenza, Minima_Frequenza ∈ Spettro dei componenti

Massima_Frequenza, Minima_Frequenza = Componenti non nulle

Nota: in alcuni casi si definisce una componente non nulla se essa supera una soglia di **3 dB** oltre il silenzio



Caratteristiche derivanti dal Dominio delle Frequenze (3)

- Lo spettro del segnale facilita la valutazione della distribuzione delle frequenze componenti; ad esempio è possibile rapidamente valutare se il segnale possiede componenti con frequenze elevate
- La presenza di alte frequenze nel segnale audio comporta che con alta probabilità il segnale in oggetto contiene musica
(la musica ha componenti molto più elevate rispetto il parlato)
- **7 kHz** rappresenta un buon valore nello spettro per la soglia che determina se un file audio contiene parlato oppure musica; frequenze fino a 7 kHz → banda bassa; frequenze oltre i 7 kHz → banda alta
- La classificazione di frequenze “alte” o “basse” dipendono dall’applicazione che dovrà trattare tali segnali
- L’energia complessiva di ciascuna banda è la somma dei suoi componenti
- Il **Centroide Spettrale** individua il punto medio della distribuzione di energia di un suono
- Il centroide del parlato è inferiore al centroide della musica



walter.balzano@gmail.com

Caratteristiche derivanti dal Dominio delle Frequenze (4)

Spettro del parlato



Spettro musica



walter.balzano@gmail.com

Caratteristiche derivanti dal Dominio delle Frequenze (5)

Armoniche

- **Un suono prodotto da un corpo vibrante non è mai puro**, ma è costituito da un amalgama in cui al suono fondamentale se ne aggiungono altri più acuti e meno intensi: questi sono gli armonici, che hanno una importanza fondamentale nella determinazione del timbro di uno strumento e nella determinazione degli intervalli musicali
- Le armoniche di un suono sono multiple in frequenza rispetto una frequenza più bassa detta **frequenza fondamentale**
- In genere **la musica contiene molte più armoniche di un semplice suono**
- **Il test che stabilisce se un suono contiene armoniche controlla che le frequenze di componenti dominanti siano multiple di una frequenza fondamentale**

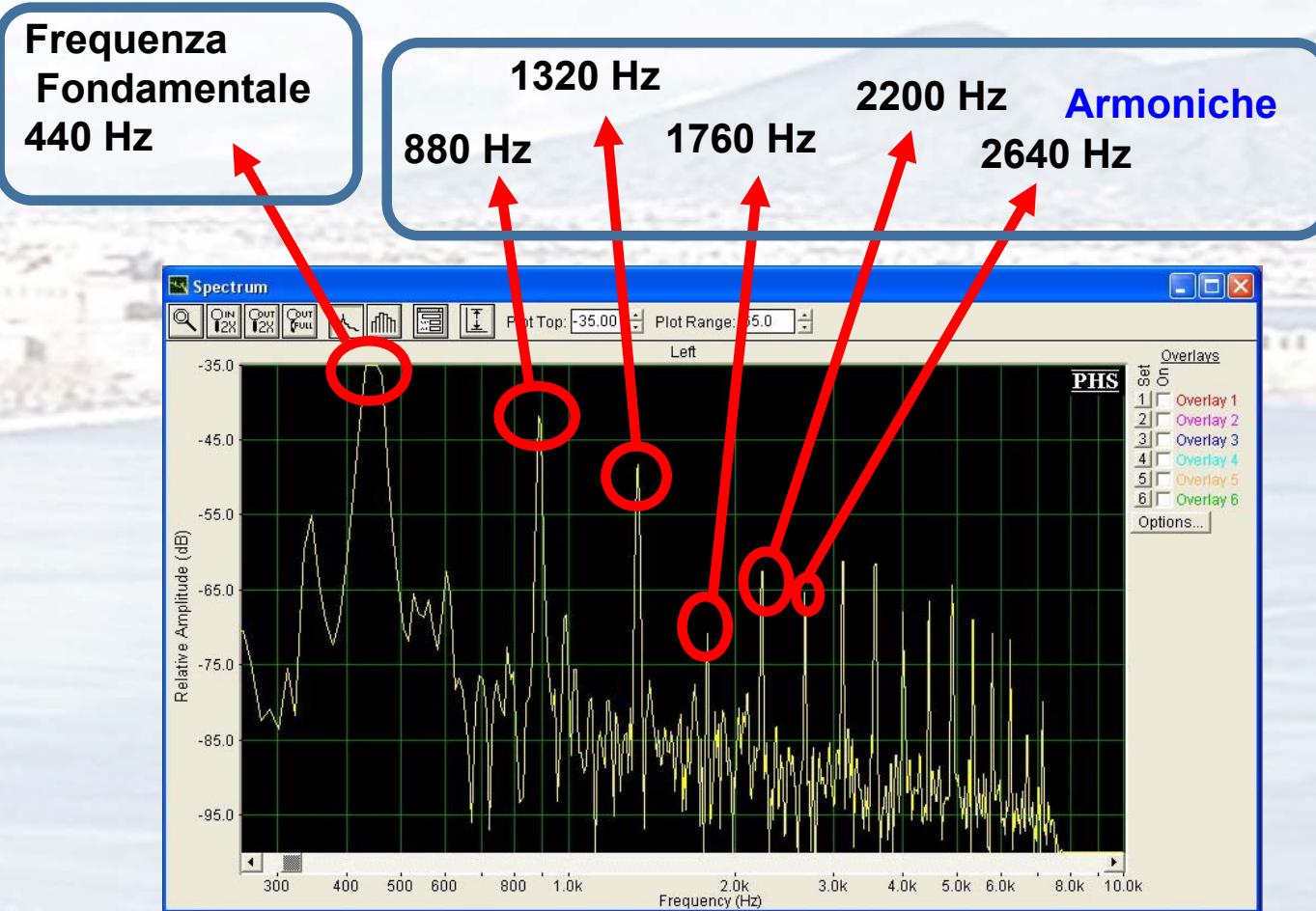
Esempio:

- Se lo spettro di un flauto contiene la serie di picchi di frequenza:
- 400 Hz, 800 Hz, 1200 Hz, 1600 Hz,
- La serie è del tipo $f, 2f, 3f, 4f, \dots$ in cui $f=400\text{Hz}$ costituisce la frequenza fondamentale del suono e le frequenze $2f, 3f, 4f, \dots$ sono le armoniche



walter.balzano@gmail.com

Caratteristiche derivanti dal Dominio delle Frequenze (6)



walter.balzano@gmail.com

Classificazione dei segnali audio

- Le principali categorie oggetto di studi di classificazione dei suoni sono la Musica ed il Parlato
- Ovviamente è possibile effettuare una classificazione molto più raffinata distinguendo, ad esempio, la voce maschile da quella femminile
- Le principali caratteristiche/differenze sono riassunte nella seguente tabella:

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



walter.balzano@gmail.com

Metodi di classificazione dell'audio

I principali metodi di classificazione audio sono:

● Classificazione Step-by-Step (passo passo)

- *Ciascuna caratteristica audio viene valutata singolarmente* per poter determinare se si tratta di un brano musicale o di parlato;
- Ogni caratteristica è vista come filtro o criterio di selezione che può essere in grado di stabilire la tipologia di un determinato pezzo audio
- Il primo filtro valuta la **posizione del centroide**: se il valore supera una soglia prefissata allora si presume che si tratti di un brano musicale
- Successivamente si valuta la quantità di silenzio (**basso silenzio → musica**)
- L'ordine con cui si serializzano i filtri l'uno dopo l'altro qualificano un algoritmo, compresa la complessità totale di elaborazione

● Classificazione basata su caratteristiche vettoriali

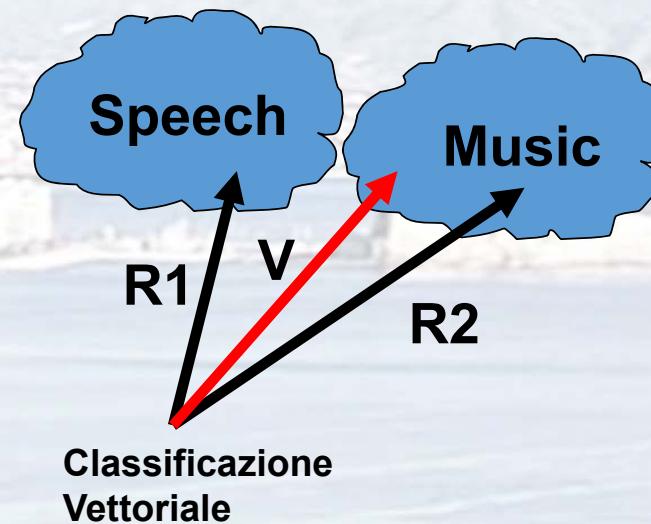
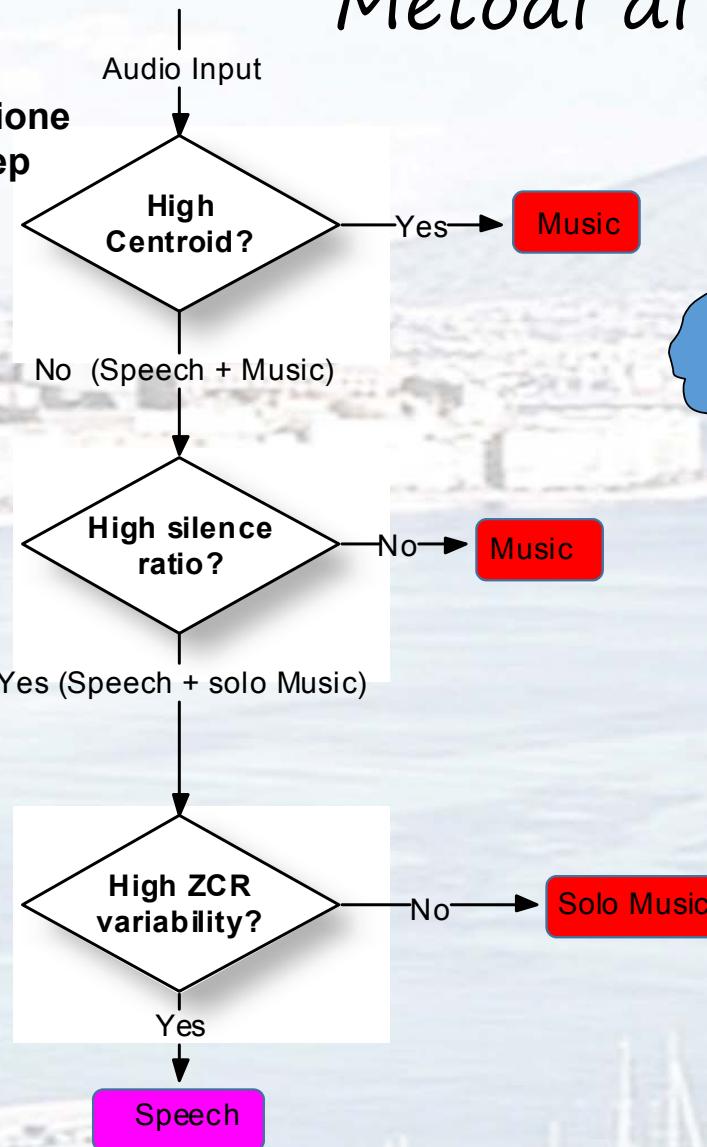
- I valori di un insieme di caratteristiche del suono considerato costituiscono le componenti di un **vettore V** che verrà confrontato (tramite **distanza Euclidea**) con un altro vettore di caratteristiche **R** che rappresenta il vettore di riferimento di ciascuna classe di pezzi audio



walter.balzano@gmail.com

Metodi di classificazione dell'audio

Classificazione Step-by-Step



walter.balzano@gmail.com

Riconoscimento del parlato

- L'approccio fondamentale per l'indicizzazione ed il recupero del parlato è basato sulla conversione dei segnali audio vocali in testo su cui successivamente applicare tecniche di IR
- Da un punto di vista generale, risulta informativo (e quindi possibile oggetto di query) anche l'identità dello speaker, la sua inflessione, ed altre caratteristiche "secondarie".



Riconoscimento del parlato

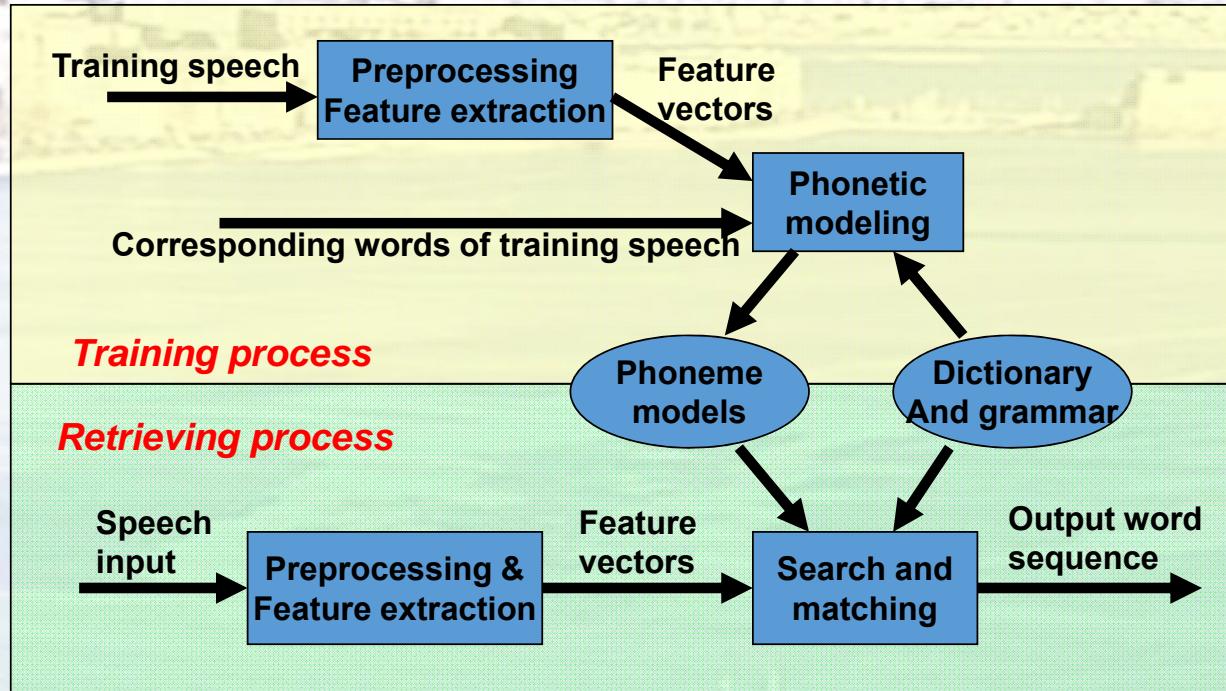
- In generale il problema del riconoscimento del parlato (**ASR**, Automatic Speech Recognition) viene ricondotto ad un problema di pattern matching
 - Al sistema vengono fornite le componenti di base del parlato (i **fonemi**, le **parole** e le **frasi**)
 - Il sistema le rappresenta utilizzando vettori di feature (bandwidth, ZCR, Silence ratio, Spectral centroid, ecc...).
- In genere si suddivide il parlato in singole unità ognuna delle quali viene confrontata con i vettori di feature raccolti nella fase di training: in tal modo viene trovato il matching migliore utilizzando la distanza euclidea tra i vettori di feature.
- Dal punto di vista applicativo, il problema del riconoscimento del parlato si complica perché:
 - fonemi prodotti da persone diverse hanno grandi differenze in durata, ampiezza, componenti in frequenza la presenza di rumore ambientale o di sottofondo
 - la separazione del brano in fonemi è critica (ogni fonema ha una durata diversa)
 - I fonemi variano a seconda del contesto (es: vocale in mezzo a consonanti)



walter.balzano@gmail.com

Concetti di base dell'ASR (Automatic Speech Recognition)

- Il processo di riconoscimento è quindi un processo **STATISTICO** che può essere migliorato fornendo una conoscenza del linguaggio utilizzato (dizionario dei termini e regole grammaticali che determinano le sequenze di parole ammesse)



walter.balzano@gmail.com

Tecniche basate sul Dynamic Time Warping

- Ogni “pezzetto” di parlato (frame temporale per esempio di 10 ms) viene rappresentato da un vettore P di caratteristiche
- Il processo di riconoscimento consiste nel considerare più piccola delle differenze tra il vettore P e ciascun vettore memorizzato nella precedente fase di Training

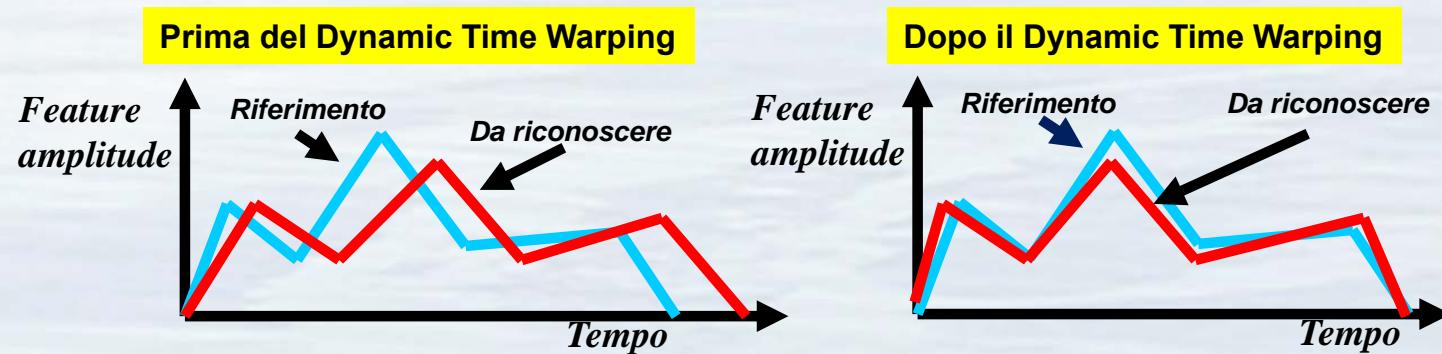
Il problema è abbastanza complesso poiché:

- Persone diverse possono impiegare tempi diversi per pronunciare la stesso fonema.
- La stessa persona può pronunciare lo stesso fonema in modo differente



Tecniche basate sul Dynamic Time Warping

La tecnica del *Time Warping* tenta di normalizzare la durata dei frame del parlato da riconoscere con quella dei frame memorizzati durante la fase di training; tale tecnica è basata sul considerare le variazioni temporali in modo non lineare: si “dilata” o si “contrae” l’asse dei tempi in modo da far coincidere picchi di segnale:



Tecniche basate sul Dynamic Time Warping

L'algoritmo Dynamic Time Warping algorithm (DTW) è un noto algoritmo introdotto negli anni'60 ed ampiamente esplorato negli anni '70 in molte aree tra cui:

- **Riconoscimento del parlato
(speech recognition)**
- **Riconoscimento della scrittura
(handwriting and online signature matching)**
- **Riconoscimento dei gesti (gesture recognition)**
- **Datamining e Time series clustering**
- **Computer Vision and Animation**
- **Allineamento delle sequenza di proteine**
- **Musica ed elaborazione di segnali**



walter.balzano@gmail.com

Tecniche basate sul Dynamic Time Warping

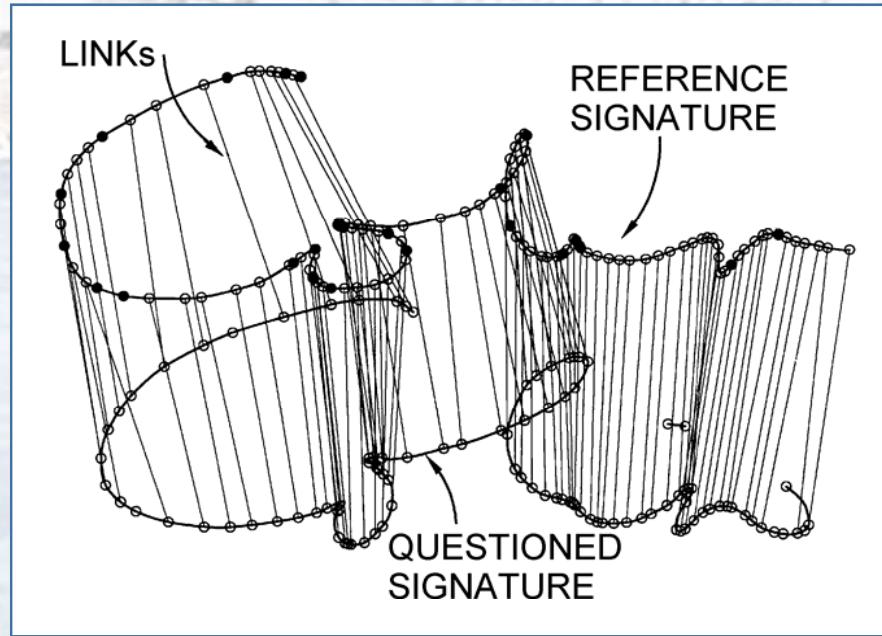
Allineamento delle sequenze di due campioni di proteine mediante la ricerca di sottopattern comuni.

-----PKKAVQLVLQMRD-----	-----AEKIANGLLNEARAMR-----
-----SSDVVSYVRDQLR-----	-----VQLACESLAQVALDRR-----
-----RQDVVRIVGEYLT-----	-----DQNAATHLIRHAVGNN-----
-----NEEIASLVIRWMD-----	-----DKNVATHLIRNALS-----
-----DQEVVVDLITSTVN-----	-----LKKATPQFVAEETIKF-----
-----VEQYWQDDFLPVNLQ-----	-----VAEAAPRLIELANQVN-----
-----KELEFFEITNSK-----	-----LKQAVFNLRYRSIENA-----
-----QLLSQKFVDNVVSLSK-----	-----PSVFAQEISKLTGKNY-----
-----YETVCDISRENKS-----	-----PMSAAEKMKDYAISYG-----
-----VDTVCDIARENST-----	-----PLRAAAELKDHAMAYG-----
-----YQTAVDIARTQRN-----	-----PMIAAQQLRDFAIISYG-----
-----PELIVDVARECRS-----	-----LMKASQQLRDLAIAYG-----
-----KRTVIDVVANRH-----	-----PLLASTKLRDYAIAYG-----
-----IDDLNSTFHNNRS-----	-----PIVVAKKIQDQLQSYD-----
-----HVENYEMWMKRKIG-----	-----PQEIAIDLIMEEVIRTK-----



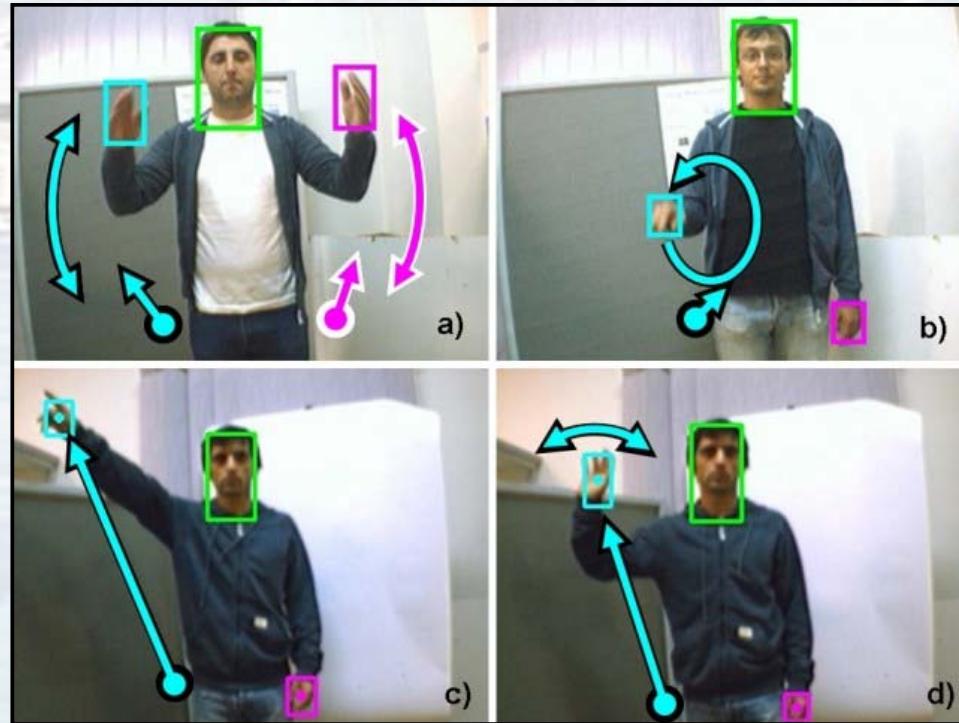
Tecniche basate sul Dynamic Time Warping

Riconoscimento di una firma basata sull'abbinamento e sul confronto di punti significativi della firma di riferimento e quella da verificare



Tecniche basate sul Dynamic Time Warping

Riconoscimento dei gesti mediante il confronto di traiettorie di punti osservati con traiettorie di riferimento



walter.balzano@gmail.com

Tecniche basate sul Dynamic Time Warping

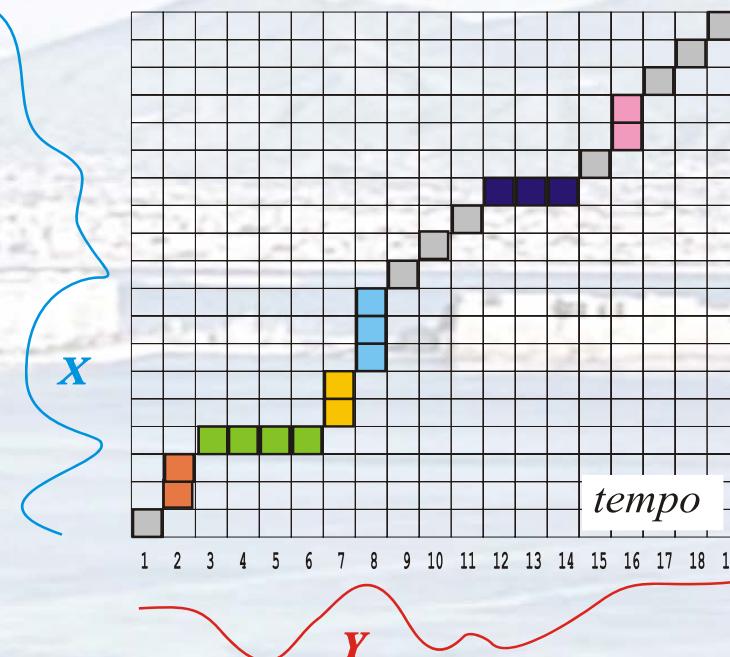
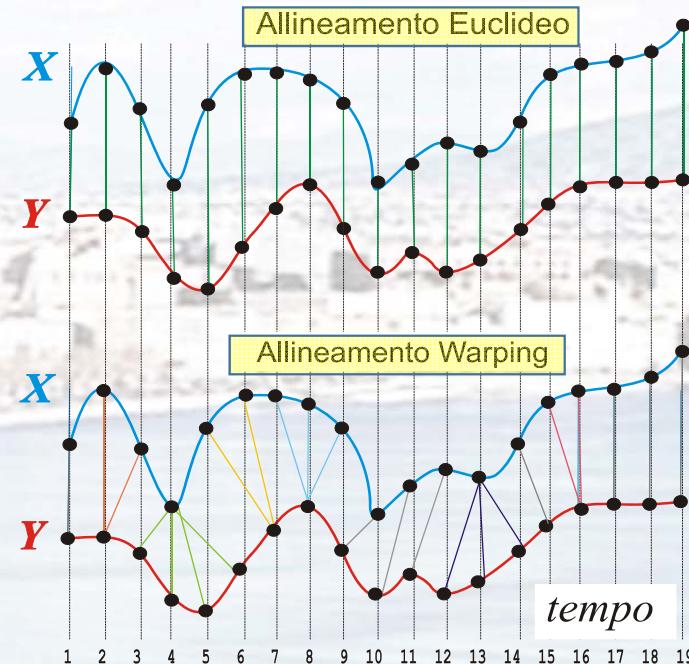
L'obiettivo primario del DTW consiste nel confrontare 2 sequenze temporali $X=(x_1, x_2, \dots, x_N)$ ed $Y=(y_1, y_2, \dots, y_M)$ e cercare il **MINIMO COSTO COMPLESSIVO** di una serie scelta di coppie (x_i, y_j)

Matrice delle distanze

y_7	0.51	0.51	0.49	0.49	0.35	0.17	0.21	0.33	0.41	0.49
y_6	0.27	0.27	0.26	0.25	0.16	0.18	0.23	0.25	0.31	0.68
y_5	0.13	0.13	0.13	0.12	0.08	0.26	0.40	0.47	0.49	0.49
y_4	0.08	0.08	0.08	0.08	0.10	0.31	0.47	0.57	0.62	0.65
y_3	0.06	0.06	0.06	0.07	0.11	0.32	0.50	0.60	0.65	0.68
y_2	0.04	0.04	0.06	0.08	0.11	0.32	0.49	0.59	0.64	0.66
y_1	0.02	0.05	0.08	0.11	0.13	0.34	0.49	0.58	0.63	0.66
	x1	x2	x3	x4	x4	x5	x6	x7	x8	x9



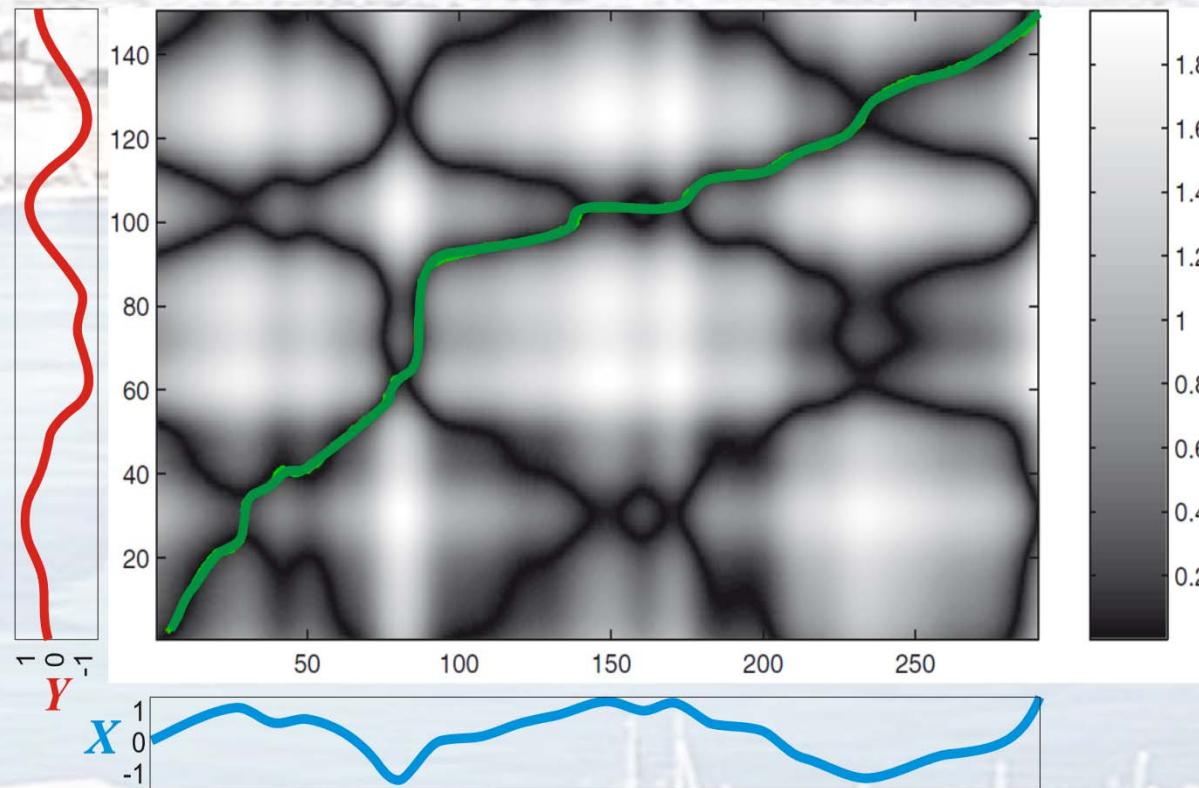
Tecniche basate sul Dynamic Time Warping



walter.balzano@gmail.com

Tecniche basate sul Dynamic Time Warping

Da un punto di vista intuitivo, un allineamento ottimale tra le sequenze X ed Y percorre **un sentiero di avallamenti più profondi possibile (valori bassi)** all'interno della matrice C :



walter.balzano@gmail.com

Tecniche basate su Hidden Markov Models (HMM)

- Una tecnica molto usata e con buoni risultati (sia per il riconoscimento dello scritto che del parlato ed i suoi fonemi) è quella dei modelli di Markov nascosti (HMM, Hidden Markov Models). La parola “nascosto” indica che per un osservatore non è visibile la sequenza di *stati* ma solo la sequenza di output dei *simboli*

Una HMM è costituita da:

- *Stati*
- *Probabilità di transizione*
- *Probabilità di generazione dei simboli*



Tecniche basate su Hidden Markov Models (HMM)

- Durante la fase di **TRAINING** viene costruita una HMM che per ogni fonema considera (sotto forma di probabilità) la variabilità degli speaker, il rumore di fondo e le differenze temporali
- Durante la **RICONOSCIMENTO** occorre trovare quale HMM è più probabile che generi la sequenza di feature vector di input

In cui:

simboli:

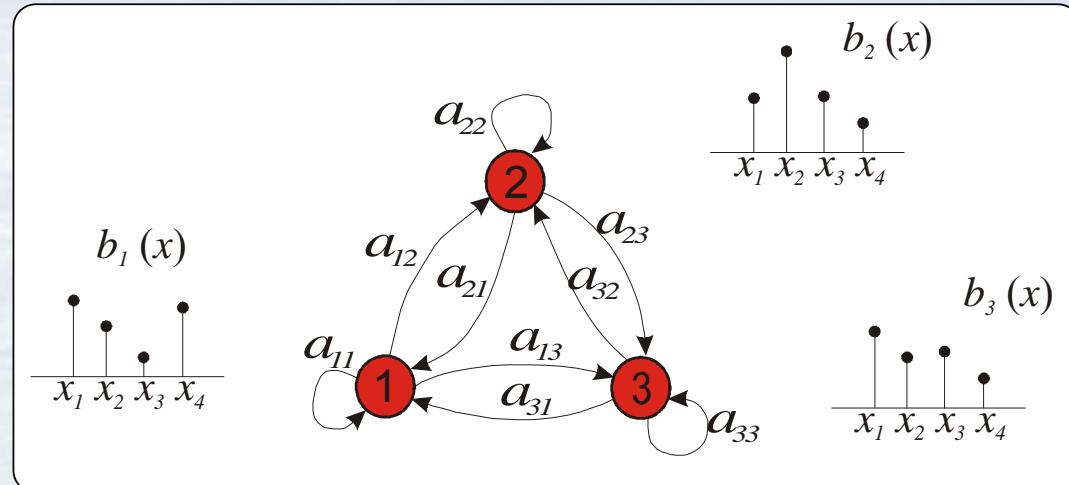
x_1, x_2, x_3, x_4

probabilità dei simboli:

$b_i(x)$

probabilità di transizione:

a_{ij}



Tecniche basate Reti Neurali Artificiali ANN (1)

- Le **ANN** (**A**rtificial **N**eural **N**etworks), largamente impiegate per il riconoscimento, simulano i processi cognitivi, in particolare l'apprendimento, del cervello umano
- Una ANN simula una rete interconnessa da link con peso.

- Una ANN è costituita da due fasi:

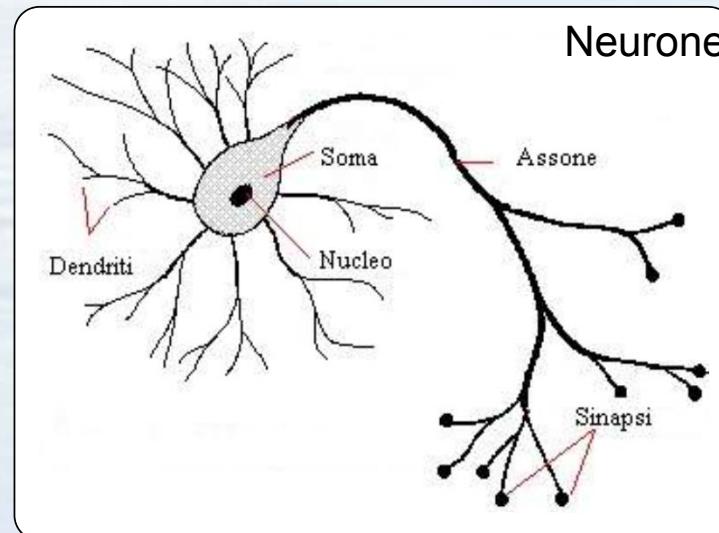
- **Training** (addestramento): i vettori di caratteristiche ottenuti durante l' addestramento di parlato servono per tarare i pesi dei link della rete
- **Recognition**: l' ANN seleziona il fonema più verosimile basandosi sulle caratteristiche dei vettori

Lo scopo della fase di addestramento consiste nel fornire un risultato più probabile per un input assegnato



Tecniche basate Reti Neurali Artificiali ANN (2)

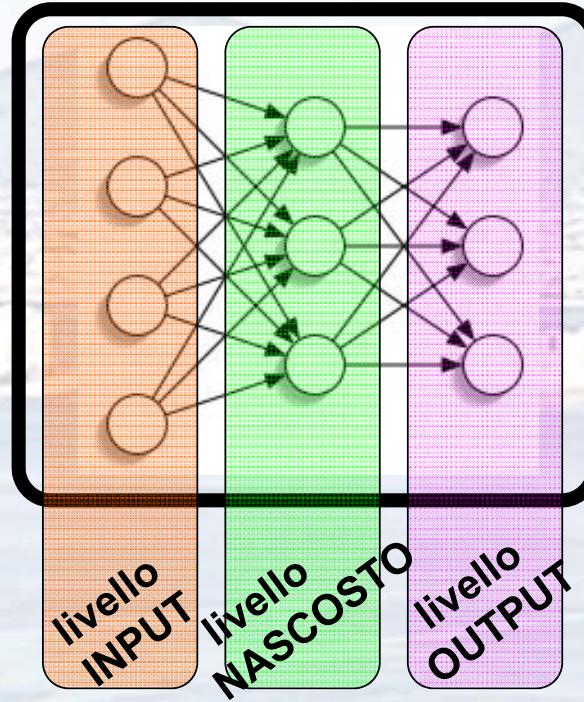
- Le ANN sono basate su implementazioni che prevedono l'utilizzo di un grande numero di elementi di calcolo (**PE – Processing Elements**) molto semplici e interconnessi tra di loro.
- Ogni PE implementa una semplice funzione matematica di tipo non lineare e rappresenta un **NEURONE** (unità elementare del cervello)
- 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)



walter.balzano@gmail.com

Tecniche basate Reti Neurali Artificiali ANN (3)

- 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.
- La rete può essere addestrata a implementare un determinato compito
- L'addestramento consiste nel variare i pesi delle connessioni secondo determinate leggi
- 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.

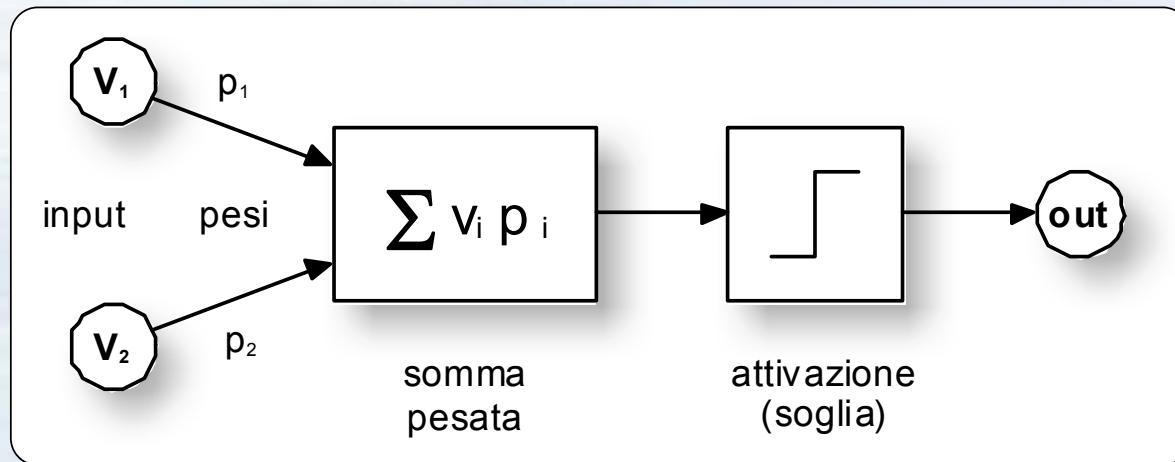


walter.balzano@gmail.com

Tecniche basate Reti Neurali Artificiali ANN (4)

Elaborazione di un PE:

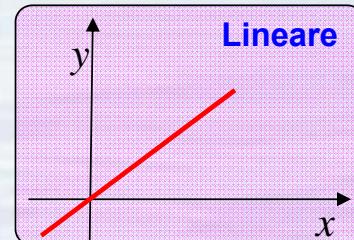
- Ogni neurone della rete (PE) effettua una somma pesata **INTEGRAZIONE** degli input derivanti dalle connessioni con gli altri neuroni
- L'input pesato viene poi valutato da una funzione detta di **TRASFORMAZIONE** che determina l'output del singolo PE
- Normalmente le funzioni di trasformazione sono funzioni molto semplici e non lineari (**FUNZIONI SOGLIA**)



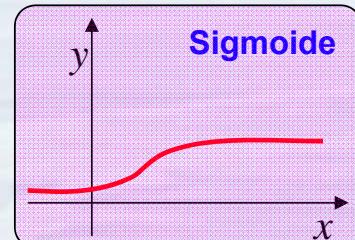
Tecniche basate Reti Neurali Artificiali ANN (5)

Funzioni di trasformazione

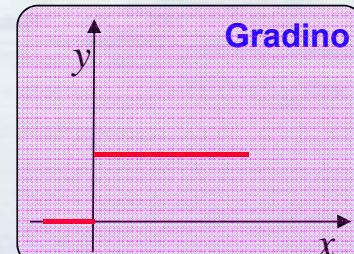
- Le funzioni di trasformazione hanno il compito di restituire l'output in relazione all'input totale ricevuto sul neurone
- Si utilizzano delle funzioni a soglia che danno luogo ad un'attivazione del neurone solo nel caso in cui l'input su tale neurone superi una soglia pre-determinata
- Questo simula il comportamento dei neuroni reali i quali reagiscono solo se stimolati sopra una determinata soglia
- Le funzioni di trasformazione più usate sono le seguenti:



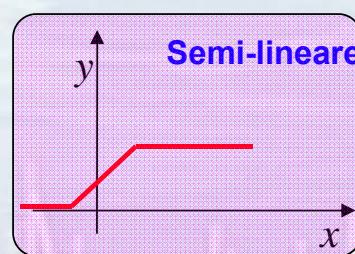
Lineare



Sigmoide



Gradino



Semi-lineare



walter.balzano@gmail.com

Prestazione dei Sistemi di Speech Recognition

Le prestazioni dei Sistemi di Speech Recognition sono influenzate dai seguenti fattori:

- **Soggetto** del parlato: articolo di giornale, libro tecnico, ecc...
- **Tipo** di parlato: letto o conversazione spontanea
- **Dimensione** del vocabolario utilizzato

Esempio di prestazione della tecnica HMM:

Soggetto	Tipo	Dimensione Vocabolario	Prestazione (Error %)
Cifre connesse	Lettura	10	<0.3
Speaker aeroporto	Spontaneo	2500	2
Giornale	Lettura	64000	7
Radio News	Lettura/Spontaneo	64000	30
Telefono	Conversazione	10000	50

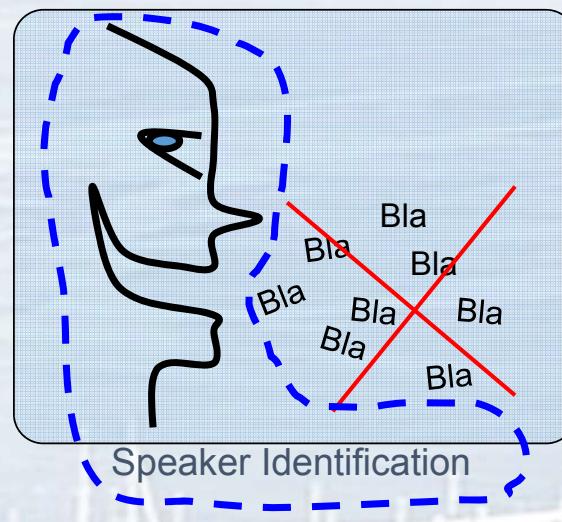
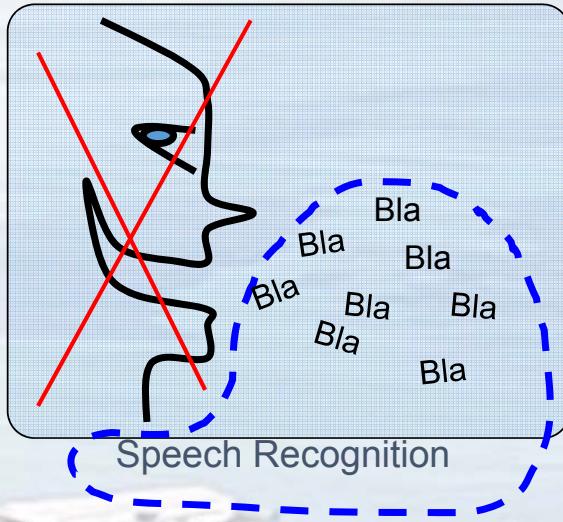


walter.balzano@gmail.com

Tecniche di identificazione dello speaker

Cercano di estrarre informazioni su chi sta parlando ed alcuni degli obiettivi sono:

- Identificazione del numero di speaker che stanno parlando
- Identificazione del sesso o dell'età dello speaker
- Identificazione dello stato emotivo o attitudinale (allegro, triste, ecc...)
- Riconoscimento della persona che sta parlando (Voice Recognition)
- Utilizzano un approccio contrario ai metodi di Speech Recognition (i quali devono essere speaker-independent e quindi cercano di eliminare le caratteristiche peculiari del parlato di ognuno degli speaker) cercando di enfatizzare le differenze di pronuncia, linguistiche e temporali tra i vari speaker



walter.balzano@gmail.com

Indicizzazione e ricerca di brani musicali

UTILIZZO DI FEATURE

- Un brano musicale viene suddiviso in frame e ogni frame viene rappresentato da un vettore di feature (ad esempio: volume, tono, centroide spettrale, bandwidth, armonicità) che viene utilizzato per il calcolo della similarità

UTILIZZO DEL TONO

- Un brano musicale viene suddiviso in sequenze di toni (frequenza fondamentale del suono). Si può utilizzare una scala di riferimento oppure classificare una sequenza musicale come una sequenza di U(p), D(own) e S(imilar).
- Occorrono tecniche di **pattern matching** per ritrovare la sequenza che comporta un numero di errori minimo rispetto al brano cercato dall'utente.
- In questi sistemi si ottengono alte performance solo se ogni suono è intervallato da una pausa.



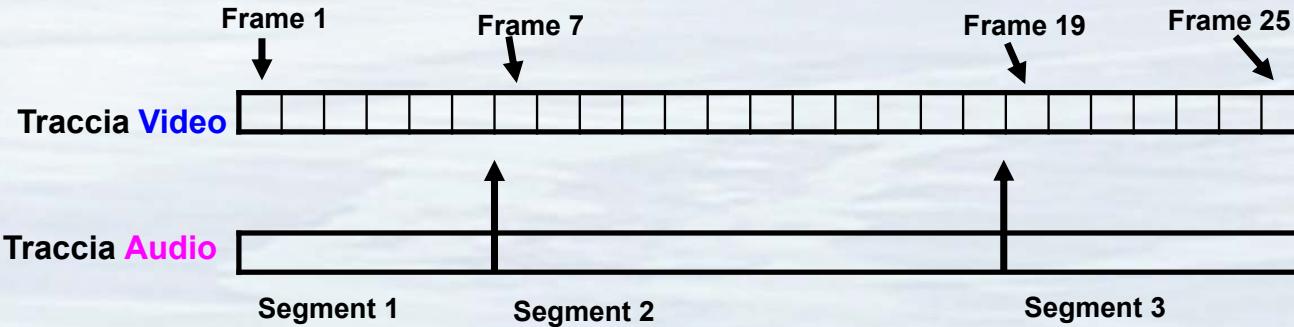
walter.balzano@gmail.com

Relazioni tra audio ed altri media

- In molte applicazioni l'audio è parte di un oggetto multimediale composito (ad esempio un film) dove esistono delle forti relazioni temporali tra video ed audio.
- Possiamo utilizzare la conoscenza su uno dei media per migliorare l'indicizzazione e la comprensione del contenuto dell'altro media.

Esempio:

Possiamo utilizzare la classificazione audio (parlato, musica, parlato) per segmentare il video in maniera analoga, ed usare il riconoscimento del parlato per comprendere il contenuto del video.



Indicizzazione e Recupero delle Immagini

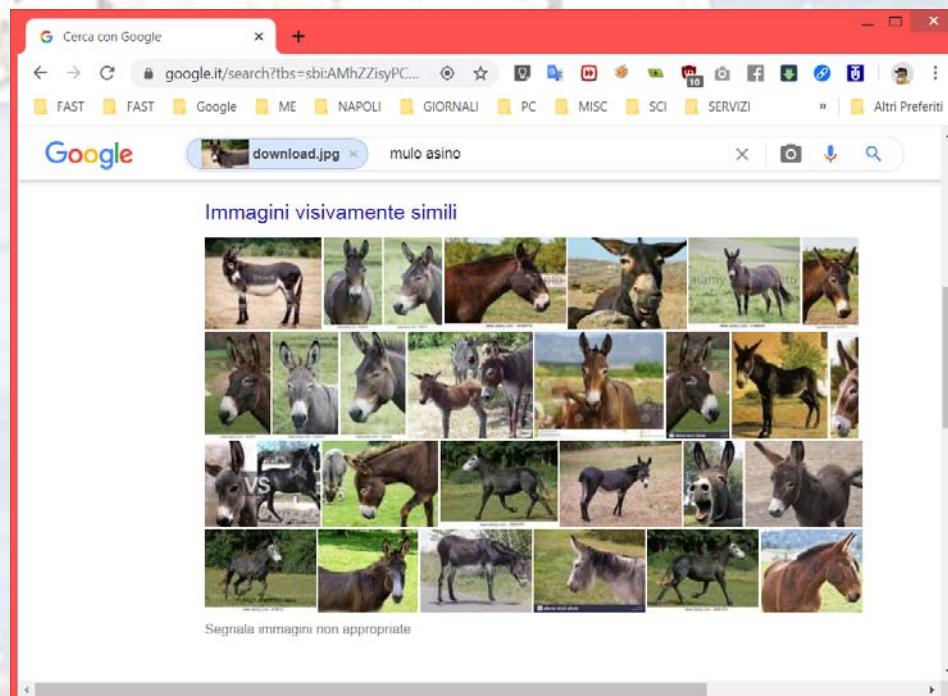
- **Introduzione**
- **Approcci di Indicizzazione e Recupero**
- **Tecniche Text-Based**
- **Tecniche Color-Based**
- **Miglioramenti delle tecniche di base**
- **Retrieving basato sulla forma**



walter.balzano@gmail.com

Introduzione

- L'indicizzazione ed il recupero delle immagini costituisce un settore di ricerca in cui si sono raggiunti risultati molto importanti e avanzati.
- I risultati ottenuti in questa di ricerca sono spesso sfruttati nei database relazionali di tipo commerciale.



walter.balzano@gmail.com

Introduzione



Google Photos



Amazon Photos



Flickr



walter.balzano@gmail.com

Introduzione

I 4 approcci principali dell'indicizzazione e ricerca delle immagini:

Tradizionali

Basati su contenuti

Attributi di Struttura	Testo	Caratteristiche di basso livello	Riconoscimento degli oggetti
<ul style="list-style-type: none">• Nome_file• CATEGORIA• Data di creazione• Soggetto• Autore....	<ul style="list-style-type: none">• Annotazioni	<ul style="list-style-type: none">• Colori• Texture.....	<ul style="list-style-type: none">• Computer vision• OCR.....
DBMS	<i>Incompletezza e soggettività</i>		



walter.balzano@gmail.com

Memorizzazione di attributi strutturati

- Questa strategia è fondata sulla memorizzazione di informazioni per ognuna delle immagini del D.B.M.S.:
 - Nome del file,
 - Data di creazione,
 - Autore,
 - Categoria dell'immagine
 - Soggetto
 -
- L'indicizzazione e la ricerca possono quindi essere effettuate con **tecniche standard dei DBMS**
- I limiti più evidenti di questa metodologia sono:
 - Gli attributi possono non descrivere in maniera completa le immagini
 - Le query sono limitate ai soli attributi memorizzati



walter.balzano@gmail.com

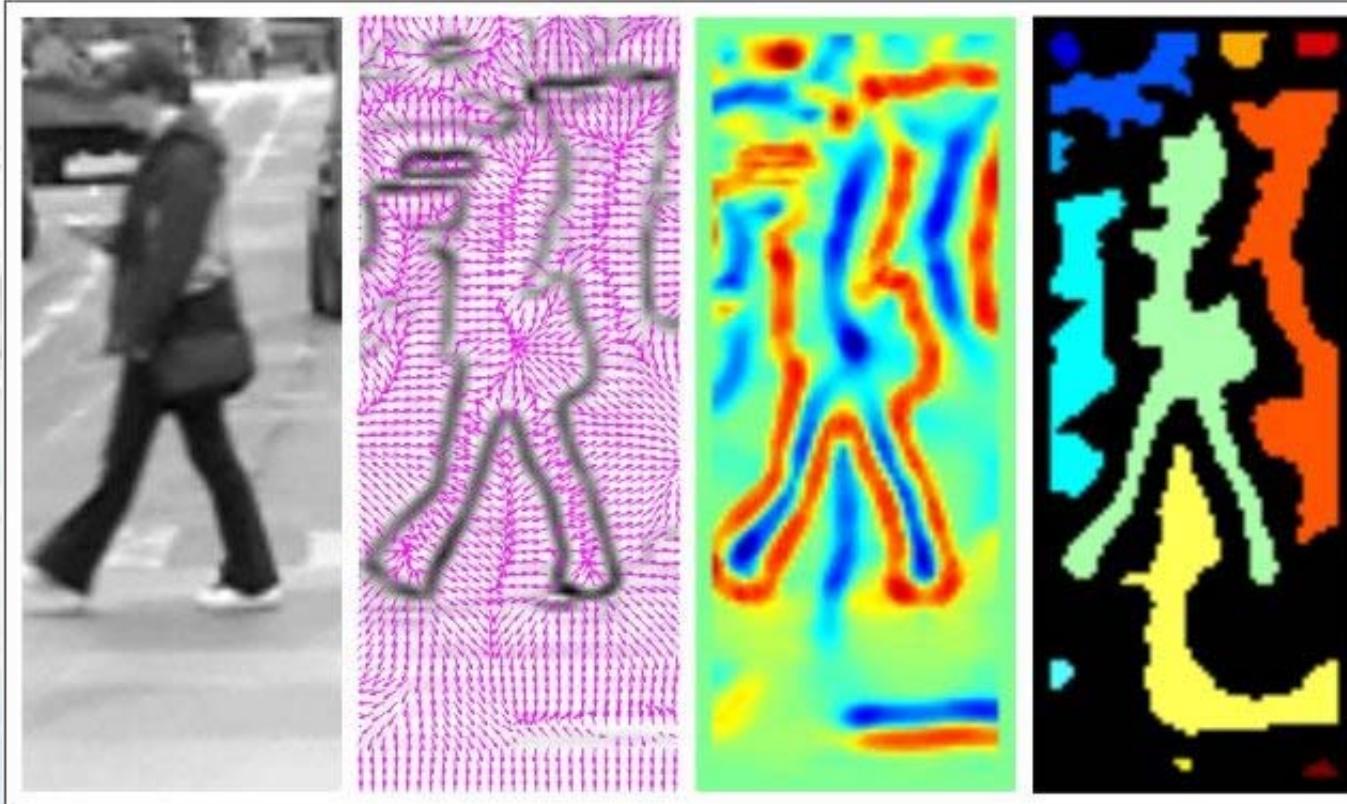
Riconoscimento degli oggetti

- Questa strategia è fondata su sofisticate **tecniche di estrazione di feature** e di **algoritmi di riconoscimento** degli oggetti presenti nella scena.
- Caratteristiche principali:
 - Sono computazionalmente molto pesanti
 - Richiedono algoritmi estremamente sofisticati
 - Godono di una buona prestazione solo per domini applicativi ristretti
 - Sono argomento di ricerca avanzata (**Computer Vision**) tecniche



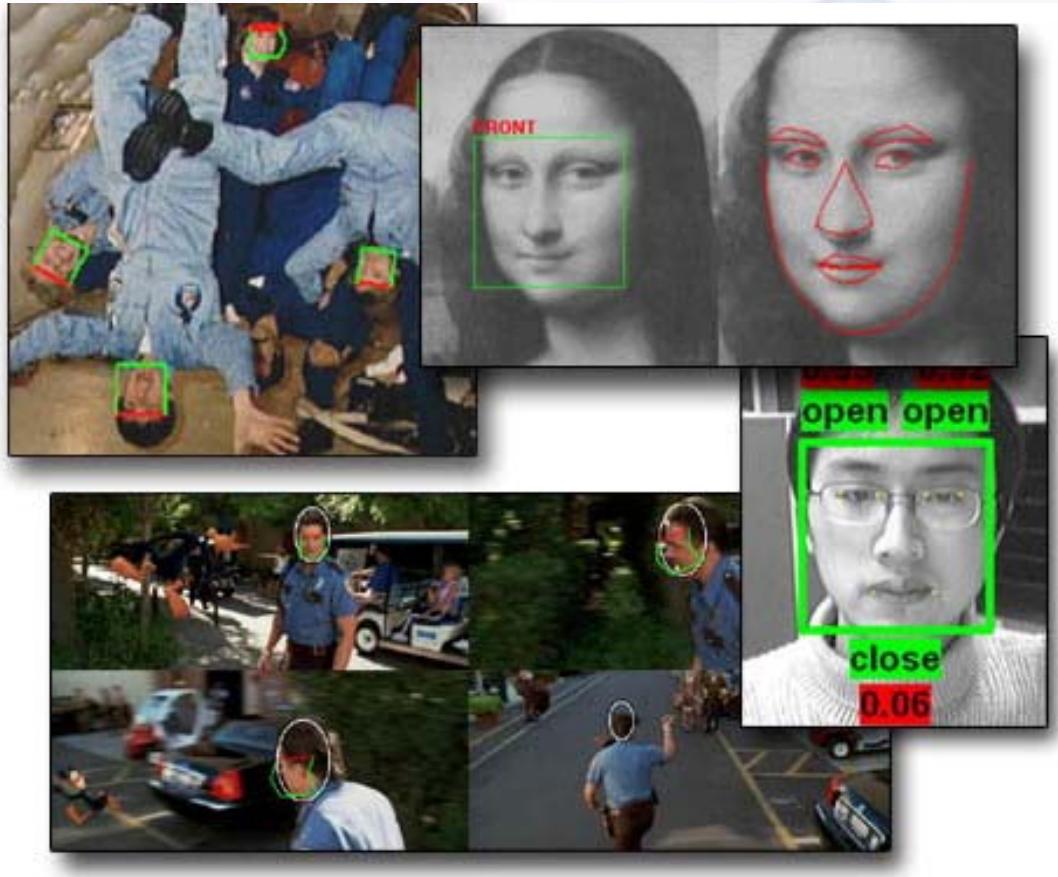
walter.balzano@gmail.com

Riconoscimento degli oggetti



walter.balzano@gmail.com

Riconoscimento degli oggetti



walter.balzano@gmail.com

Riconoscimento degli oggetti



walter.balzano@gmail.com

Annotazioni libere (Text-based image retrieval)

- Le immagini sono descritte con testo libero (non controllato).
- Le query sono nella forma di “**parole chiave**” o testo libero (con eventuale utilizzo di operatori booleani)
- La ricerca usa gli **algoritmi convenzionali di IR** (Information retrieval) basati sulla ricerca di similarità tra query e testo descrittivo delle immagini
- Essendo una tecnica “**manuale**”, occorre notare che:
 - Il testo deve essere il più possibile completo e consistente
 - E’ necessario l’utilizzo di un dizionario per ricercare anche su sinonimi (uomo/donna/bambino = persona)
 - Il testo introdotto potrebbe essere soggettivo → occorre utilizzare tecniche di *relevance feedback*
- **Vantaggi:**
- Si possono “catturare” anche concetti di alto livello (cioè astratti) presenti nell’immagine (per es. una persona che sorride o che è triste)

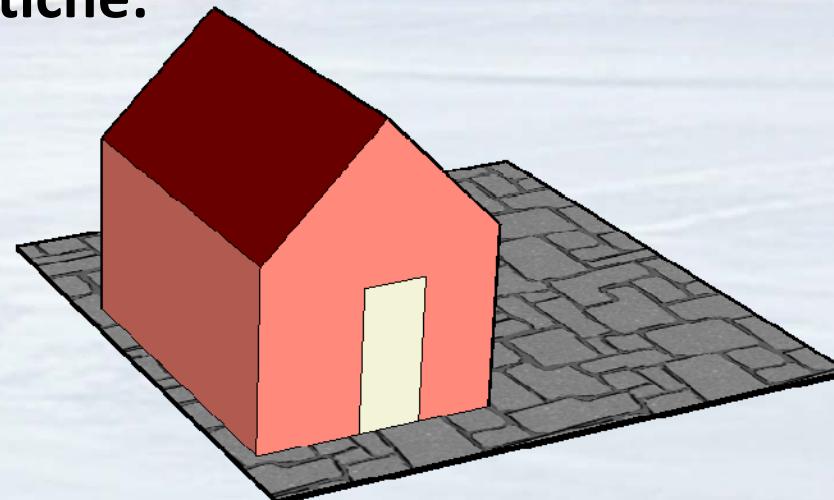


walter.balzano@gmail.com

Estrazione di feature di basso livello

- L'estrazione di feature di basso livello è definita su tecniche di indicizzazione e ricerca basate sul contenuto (*content-based*)
- Possono prendere in considerazione una o più delle seguenti caratteristiche:

- COLORE
- FORMA
- TEXTURE



- Molti sistemi reali sono basati su tali metodi

Algoritmi basati sull'analisi del colore

- Ogni immagine è memorizzata assegnando ai pixel tre valori numerici ad ognuno dei canali di colore (ad esempio RGB)
- Ogni canale è discretizzato in m intervalli (**quantizzazione dei colori**): il numero totale di combinazioni diverse (bins) è m^3
(es: 16 intervalli di colore → 4096 bins)

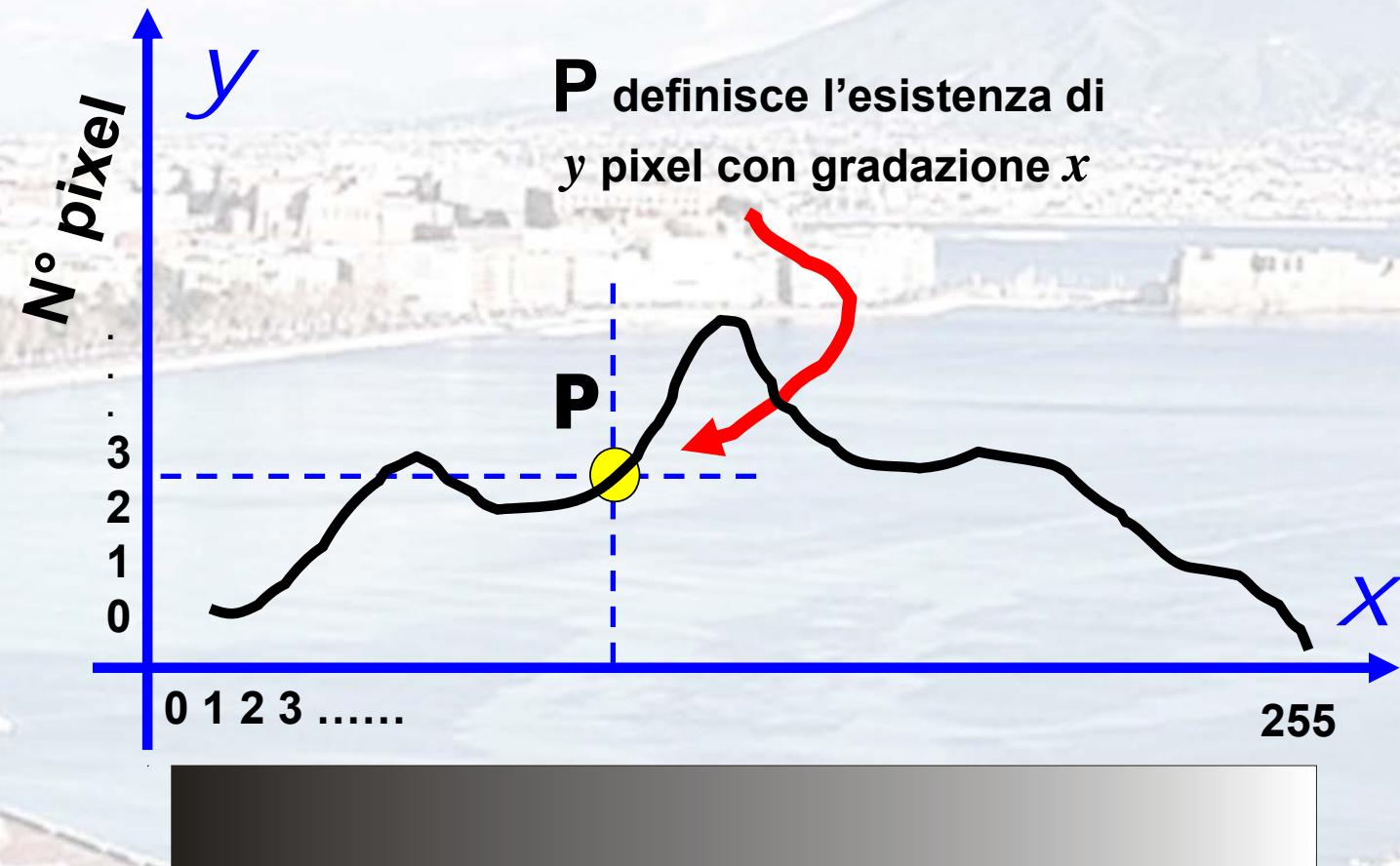
- Si definisce **ISTOGRAMMA di COLORE**:

il vettore $H(M) = (h_1, h_2, \dots, h_j, \dots h_n)$

in cui h_j rappresenta il numero di pixel dell'immagine M che ricadono nel bin j.

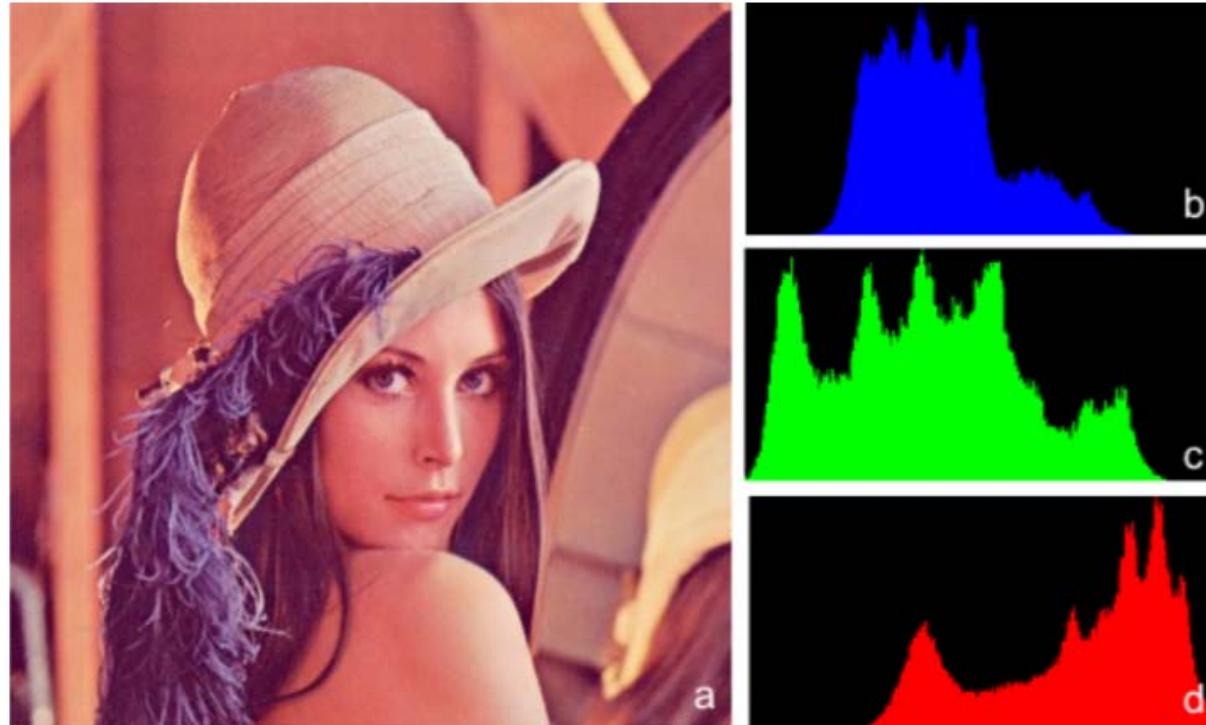


Algoritmi basati sull'analisi del colore



walter.balzano@gmail.com

Algoritmi basati sull'analisi del colore



walter.balzano@gmail.com

Confronto degli Istogrammi di immagini

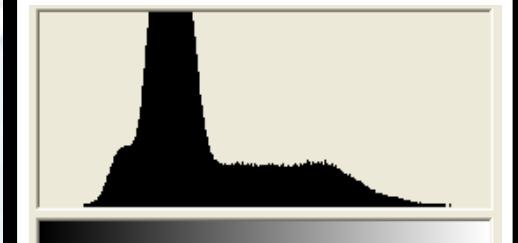
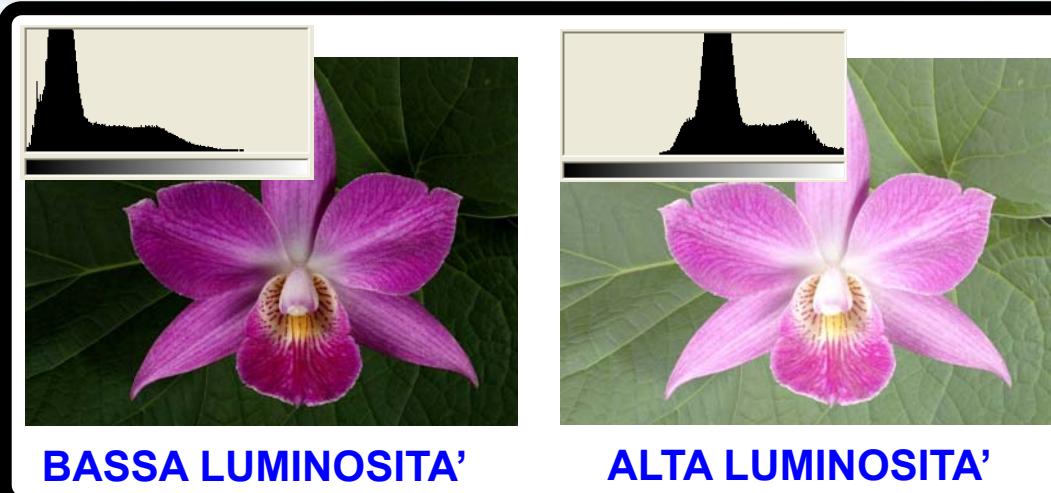
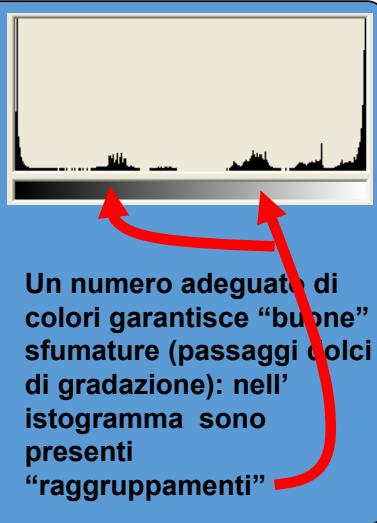
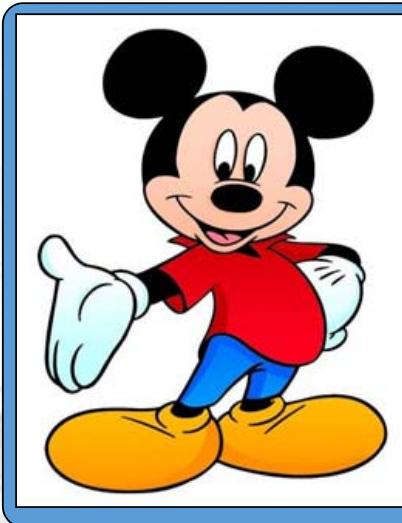


Foto ORIGINALE



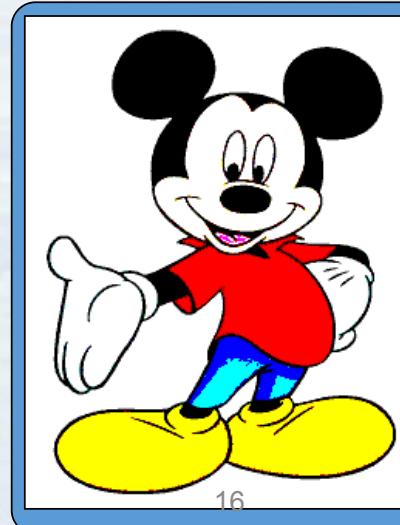
walter.balzano@gmail.com

Confronto degli Istogrammi di immagini (2)



**Disegno con
8 bit/pixel
(max colori = $2^8=256$)**

**Disegno con
4 bit/pixel
(max colori = $2^4=16$)**



walter.balzano@gmail.com

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:

in cui:

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

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



Esempio di calcolo della distanza

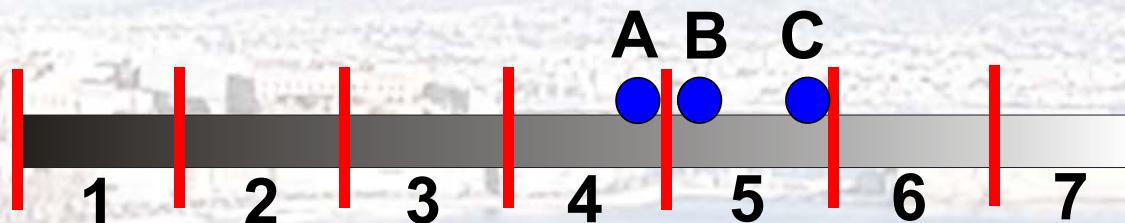
- Si supponga di avere 3 immagini 8x8 pixels aventi i seguenti istogrammi:
 - $H_1 = (8, 8, 8, 8, 8, 8, 8, 8)$
 - $H_2 = (7, 7, 7, 7, 9, 9, 9, 9)$
 - $H_3 = (2, 2, 10, 10, 10, 10, 10, 10)$
- Quindi, le distanze tra le immagini sono rispettivamente:
 - $d(H_1, H_2) = 1+1+1+1+1+1+1+1 = 8$
 - $d(H_1, H_3) = 6+6+2+2+2+2+2+2 = 24$
 - $d(H_2, H_3) = 5+5+3+3+3+1+1+1+1 = 23$
- Pertanto le immagini 1 e 2 sono più simili che 1 e 3 oppure 2 e 3
- Tale metodo di base è soggetto ad una serie di problemi e può fornire risultati fortemente incorretti se non si applicano delle varianti necessarie per trattare in maniera migliore l'informazione sul “colore”.



walter.balzano@gmail.com

Problemi relativi all'indicizzazione tramite istogramma del colore

- La **discretizzazione dello spazio dei colori in classi (bins)**
non tiene conto della similarità dei colori:



- Due bins adiacenti vengono considerati totalmente diversi.
- Il posizionamento della linea di demarcazione tra i bins influisce fortemente sull'istogramma e sul calcolo della distanza.

La rappresentazione dei colori non è univoca:

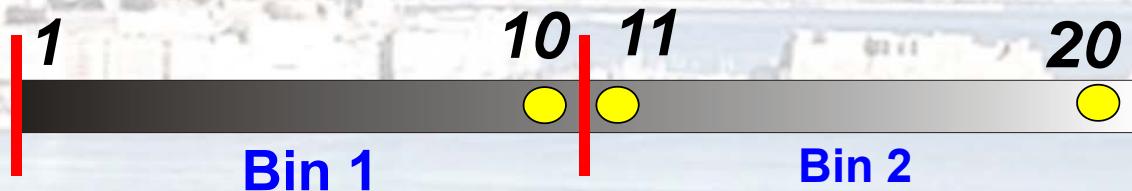
- Dipendenza dal sistema di rappresentazione
- Dipendenza dal «device»



walter.balzano@gmail.com

Istogrammi del colore e similarità

- L'esempio precedente non considera la similarità tra i colori dei bins:
- due immagini con colori percepiti come simili ma numericamente diversi risultano molto distanti tra di loro.



- *Esempio:*
- Si supponga di avere due bin che rappresentano gli insiemi di colori 1-10 e 11-20 (in cui i colori con numeri vicini sono simili tra di loro). Quindi il colore 10 viene assegnato al bin 1, ed i colori 11 e 20 vengono assegnati al bin 2
- Per paradosso il colore 11 viene considerato uguale al 20 ma diverso dal 10



Istogrammi del colore e similarità

Sono stati proposti vari metodi per superare tale problema:

- Distanza tra i bins
- Istogramma cumulativo
- Istogramma pesato percettivamente



Distanza tra i bins

- La distanza tra bins si definisce come misura di similarità (inverso della distanza) calcolata bin-per-bin. E' quindi possibile definire l'istogramma Z di similarità tra due histogrammi X ed Y nel seguente modo:

$$\|Z\| = Z^T AZ$$

in cui:

- A è la matrice simmetrica di similarità dei colori in cui $a(i,j) = a(j,i) = 1 - d(c_i, c_j)/d_{max}$
- $d(c_i, c_j)$ è la distanza tra i colori in uno spazio di uniforme
- d_{max} è la distanza massima tra due colori
- Se due colori sono molto diversi $d(c_i, c_j)$ è prossima a d_{max} e $a(i, j)$ è vicino allo 0 contribuendo al calcolo della similarità

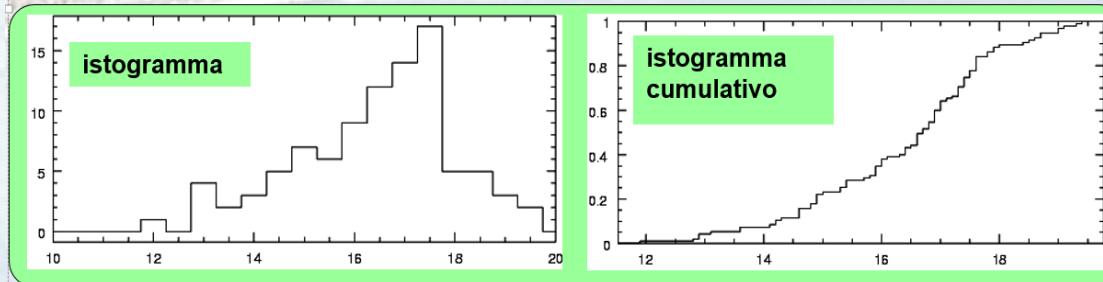


Iistogramma cumulativo

- L'istogramma cumulativo non considera la distanza tra i bin ma risolve parzialmente il problema creando delle classi cumulative.
- Un istogramma cumulativo (Cumulative Histogram) per un'immagine M è definito da:

$\text{CH}(M) = (ch_1, ch_2, \dots, ch_n)$ in cui

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



- Per il calcolo della distanza si usa la normale **distanza euclidea**.
- Prestazione: buona per valori bassi di i .
- Per valori alti non riflette la similarità tra i colori



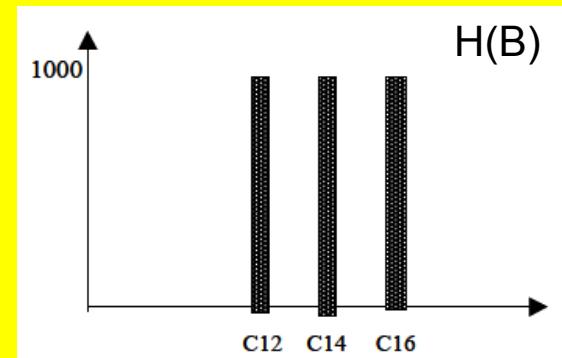
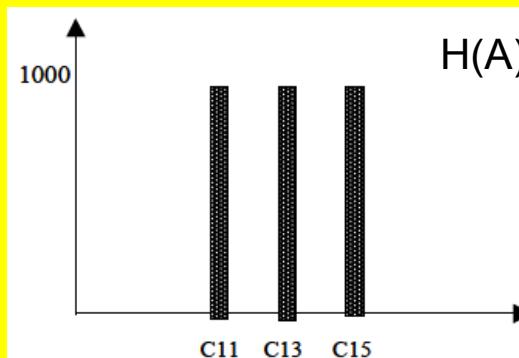
walter.balzano@gmail.com

Istogramma pesato percettivamente (PWH - Percentually Weighted Histogram)

A causa del processo di quantizzazione dei colori e di “mappatura” sui bin corrispondenti, **un singolo colore originario potrebbe essere simile a colori appartenenti a più di un unico bin** \Rightarrow immagini costituite da colori simili potrebbero avere istogrammi la cui misurazione di distanza genera un valore elevato

Se si assume che i bin siano ordinati da C1 a Cn, l'esempio di seguito illustrato mostra gli istogramma di due immagini A e B che non avendo alcun colore comune avranno una distanza pari a 6000

RGB



walter.balzano@gmail.com

Istogramma pesato percettivamente (PWH - Perceptually Weighted Histogram)

Una soluzione consiste nel **cambiare lo spazio di rappresentazione dei colori**, da **RGB** a **CIELuv**. Il cambiamento di rappresentazione dei colori comporta che ogni singolo colore in RGB può avere 10 diverse **corrispondenze “pesate”** in CIELuv (si assume che per ogni pixel in RGB, la somme dei pesi relativi alle corrispondenze in CIELuv sia uguale a 1).

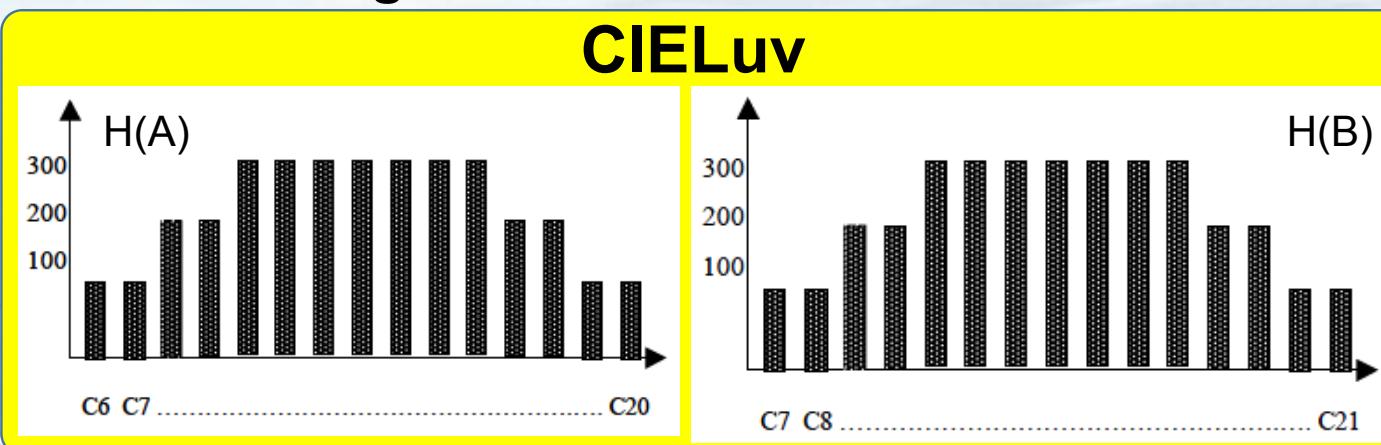
Infine si costruisce l'istogramma dell'immagine rappresentata nel nuovo dominio CIELuv



Iistogramma pesato percettivamente (PWH - Perceptually Weighted Histogram)

Ora la nuova “distanza” è 600 (cioè molto inferiore alla precedente cioè 6000)
⇒ maggiore similitudine

- la similarità tra i colori viene considerata direttamente nella **costruzione dell’istogramma** e non nel calcolo della distanza tra istogrammi (**ogni pixel contribuisce con un peso diverso a seconda della similarità a tutti i bin di colori**)
- In genere il PWH è il metodo che fornisce i risultati migliori

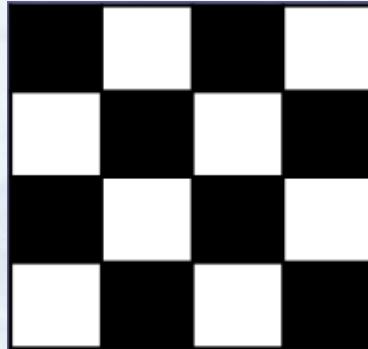


walter.balzano@gmail.com

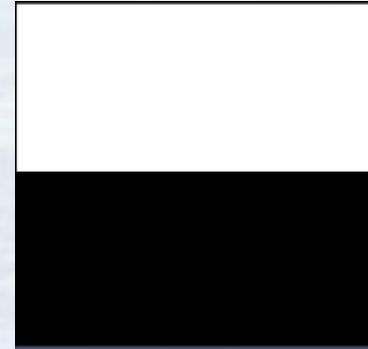
Limiti dell'approccio Color-Based: relazioni spaziali

- Una limitazione dei sistemi di indicizzazione di immagini basate sui colori consiste nel fatto che generalmente **vengono ignorate le relazioni spaziali tra i pixel**. Ad esempio le due immagini seguenti verrebbero sempre considerate uguali:

1

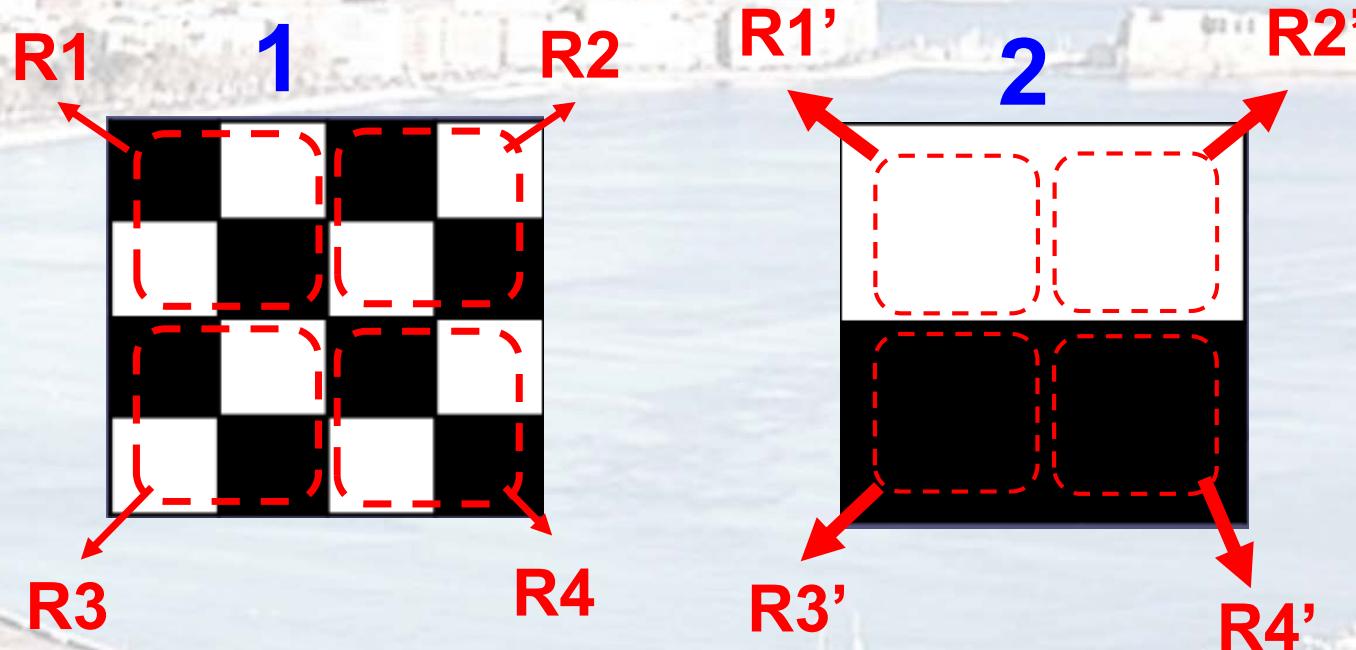


2



Limiti dell'approccio Color-Based: relazioni spaziali

Per risolvere il problema si possono utilizzare tecniche che prevedono la **suddivisione delle immagini in regioni più piccole** e applicare la tecnica basandosi sugli istogrammi delle sotto-regioni corrispondenti delle immagini



Limiti dell'approccio Color-Based: background

- **Mascheramento ad opera dello sfondo**
- **Gli istogrammi sono fortemente condizionati dalla presenza di grandi blocchi di colore omogeneo e risulta che le immagini che contengono un oggetto in primo piano (**foreground**) siano trattate in maniera errata in quanto nell'istogramma i pixel dello sfondo (**background**) sono maggiormente rappresentati e “mascherano” l’oggetto in primo piano.**

Esempio:

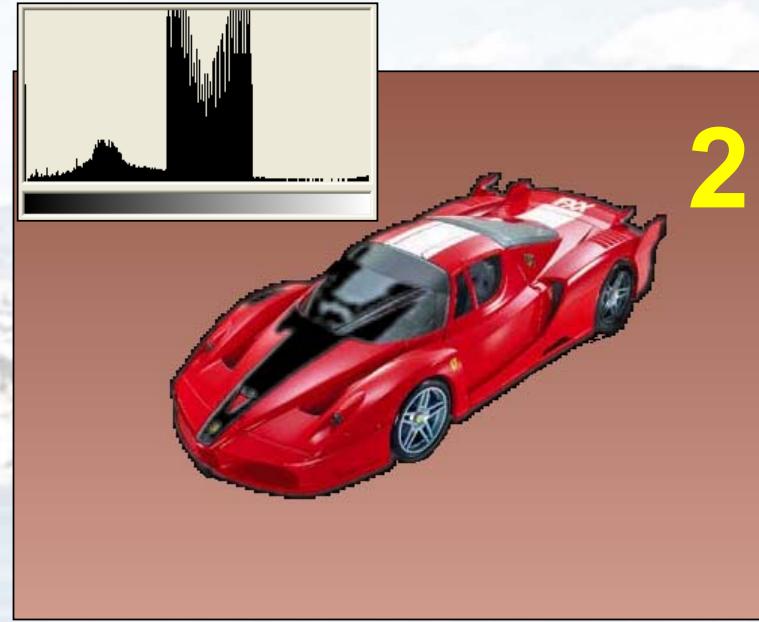
- Immagine¹: macchina rossa su sfondo nero
- Immagine 2: macchina rossa su sfondo marrone
- Immagine 3: macchina blu su sfondo marrone



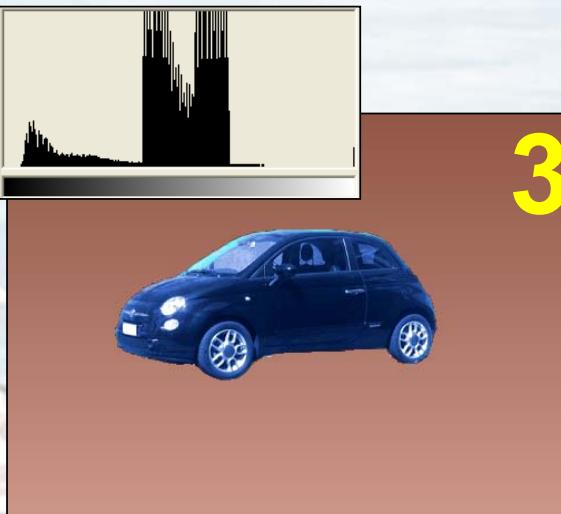
Limiti dell'approccio Color-Based: background



1



2



3

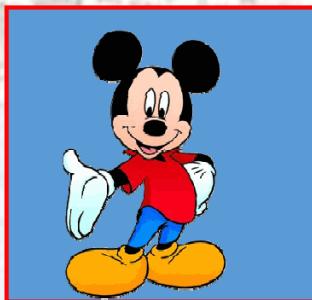
- 2 e 3 sono considerate più vicine rispetto a 1 e 2
- La possibile soluzione può consistere nell'utilizzare **tecniche di segmentazione** (automatiche o semi-automatiche) per suddividere le immagini in soggetto e sfondo e, quindi, calcolare separatamente i due istogrammi usandoli singolarmente



walter.balzano@gmail.com

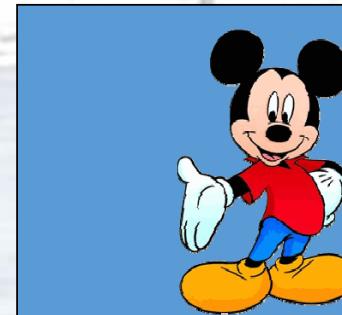
Indicizzazione e Ricerca basate sulla forma

- Si basa su algoritmi di segmentazione delle immagini che sono in grado di suddividere una immagine in singoli oggetti attraverso metodi automatici o semiautomatici.
- I sistemi efficienti devono avere una rappresentazione unica, invariante rispetto a **traslazione, rotazione o scala**

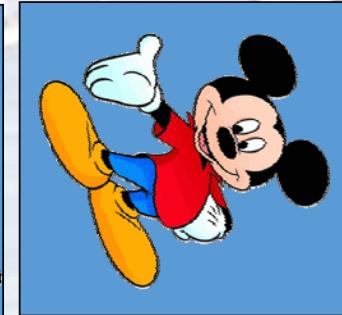


QUERY

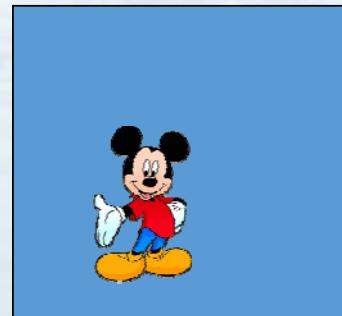
1: Traslata



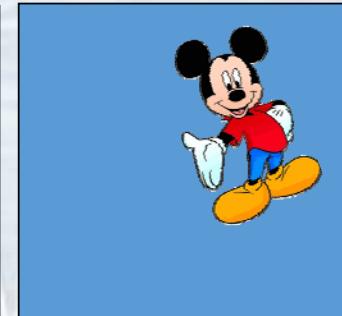
2: Ruotata



3: Scala



1+2+3



walter.balzano@gmail.com

Indicizzazione e Ricerca basate sulla forma

Il processo di ricerca è spesso realizzato in modalità “query by example” e il sistema ritorna come risposta la forma che meglio approssima quella disegnata dall’utente

- **QBIC(TM)**

IBM's Query By Image Content

The State Hermitage Museum <http://www.heritagemuseum.org/>

- **CIRES**

Content based Image REtrieval System

- **SIMPLIcity -- Semantics-sensitive**

Integrated Matching for Picture LIBraries

<http://wang14.ist.psu.edu/>



walter.balzano@gmail.com

Indicizzazione e Ricerca basate sulla forma

The State Hermitage Museum

The screenshot shows the QBIC Layout Search interface. At the top, there is a color palette and a preview window showing a painting with a red background, a blue rectangle at the top, and a white figure. Below the preview are three sliders for Red (R), Green (G), and Blue (B) with values 22, 255, and 29 respectively. To the left of the search area are links for BROWSE, QBIC SEARCHES, COLOUR SEARCH, LAYOUT SEARCH, and ADVANCED SEARCH. On the right are buttons for delete, clear all, and search.

Query per Disposizione



The screenshot shows a list of eight search results, each with a thumbnail image and a link to the original painting. The results are:

- 3) [Girl with Tulips](#)
Matisse, Henri
1910
- 4) [Bridge in Alcantara](#)
Premazzi, Luigi
1887
- 5) [Port](#)
Derain, Andre
Circa 1905
- 6) [Portrait of Suzanne Dufy, the Artist's Sister](#)
Dufy, Raoul
1904
- 7) [Glass Vessels](#)
- 8) [Still Life with The Dancer](#)

Risultati



walter.balzano@gmail.com

Indicizzazione e Ricerca basate sulla forma

The State Hermitage Museum: Digital Collection -- Powered by IBM - Windows Internet Explorer

INFORMATION COLLECTION HIGHLIGHTS EXHIBITIONS HERMITAGE HISTORY CHILDREN & EDUCATION DIGITAL COLLECTION

QUICK SEARCH go

BROWSE •
QBIC SEARCHES •
COLOUR SEARCH •
LAYOUT SEARCH •
ADVANCED SEARCH •

QBIC COLOUR SEARCH

1. Use your mouse to select a colour from the palette.
2. Click the arrow button to add the colour to the bucket.
3. Slide the triangular handles on the bucket to adjust the percentage of this colour.
4. You may repeat this process until the bucket is full. When you are ready, click Search.

You may also use the Colour Mixer to adjust RGB

R [255] G [247] B [118]

delete clear all search

The State Hermitage Museum

Query per Colore



IBM - Windows Internet Explorer

TICKETS SHOP

Digital Collection

INFORMATION COLLECTION HIGHLIGHTS EXHIBITIONS HERMITAGE HISTORY CHILDREN & EDUCATION DIGITAL COLLECTION

QUICK SEARCH go

SEARCH RESULTS

modify search
new search

BROWSE •
QBIC SEARCHES •
COLOUR SEARCH •
LAYOUT SEARCH •
ADVANCED SEARCH •

	1) Seville Still Life Matisse, Henri Circa 1910/1911		2) Icon: St Nicholas UNKNOWN 13th - 14th century
	3) Avenue in a Park Watteau, Antoine Circa 1715		4) Portrait of S. I. Shchukin Krohn, Christian Cornelius (Xan) 1915

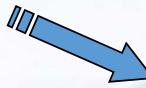
Risultati



walter.balzano@gmail.com

Indicizzazione e Ricerca basate sulla forma

Query per Grouping, Color, Texture



Cires: Results - Windows Internet Explorer

Query Image

Retrieved Images

Relevance feedback type: Cluster. Weights: Perceptual Grouping = 0.33, Color = 0.33, Texture = 0.33, L, A, B channels.
For relevance feedback, please select the check boxes below each image, and then select the feedback type. Note that NS = "Not Sure".

Image 1	Image 2	Image 3	Image 4	Image 5
<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No
<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input checked="" type="radio"/> Yes <input type="radio"/> NS <input type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No	<input type="radio"/> Yes <input type="radio"/> NS <input checked="" type="radio"/> No

Cires: Content based Image RETrieval System - Windows Internet Explorer

Manmade: Bridges

Weights (should sum to 1):

Perceptual grouping: 0.33 Color: 0.33 Texture: 0.33

Texture:

L, A and B channels
 L channel only (-Grayscale texture)

Cires

Risultati

Relevance
FeedBack



walter.balzano@gmail.com

Definizioni

ASSE MAGGIORE:

segmento che congiunge i due punti della forma
che sono più distanti fra di loro

ASSE MINORE:

segmento perpendicolare all'asse maggiore e tale
che il rettangolo il cui lati sono paralleli ai due
assi racchiuda completamente l'intera forma

RETTOANGOLO DI BASE:

il rettangolo descritto nella precedente
(esso coincide con il più piccolo rettangolo che
contiene l'intera figura)

ECCENTRICITA':

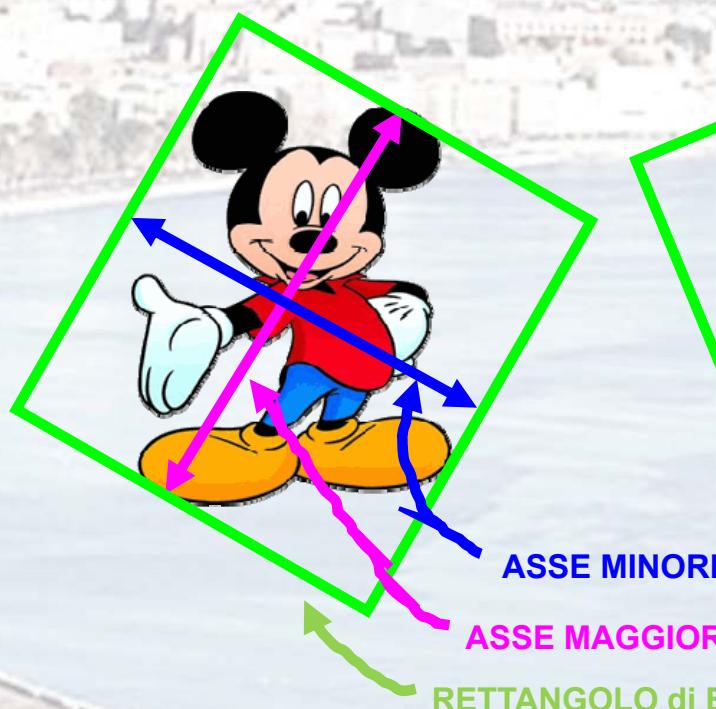
rapporto tra la lunghezza dell'asse maggiore e la
lunghezza dell'asse minore



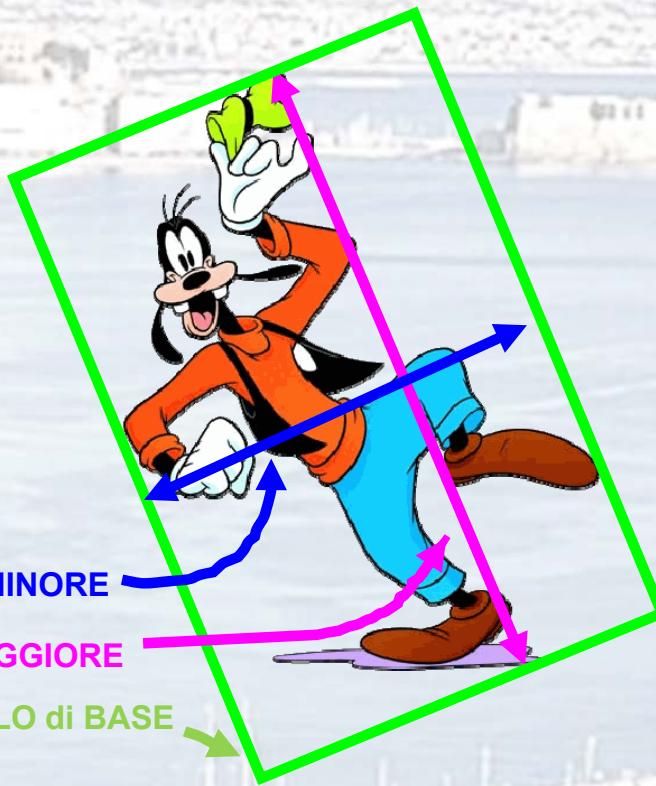
Definizioni

Le caratteristiche elencate possono essere considerate basilari in quanto forniscono una prima caratterizzazione delle forme e sono utilizzati durante l'indicizzazione e la ricerca

Bassa Eccentricità



Alta Eccentricità



Descrizioni numeriche delle forme

Per la descrizione numerica delle forme sono state adottate molteplici strategie tra cui:

- Metodi basati su **formulazioni matematiche** che descrivono un'immagine come funzione $f(x,y)$ dove x e y sono le coordinate di un pixel:

- Momenti invarianti
- Momenti centrali normalizzati

- Metodi che cercano di descrivere le forme mediante **le coordinate dei punti** che le definiscono:

- Metodo dei descrittori di Fourier
- Istogramma dei lati significativi
- Lista ordinata dei punti interessanti
- Elastic template matching

Una limitazione “classica” delle precedenti metodologie è dovuta al fatto che trattano con difficoltà l’invarianza rispetto alle traslazioni, rotazioni e cambiamenti di scala necessari per ottenere un buon sistema di indicizzazione e ricerca basato sulle forme.



walter.balzano@gmail.com

Rappresentazione delle forme basata su REGIONI

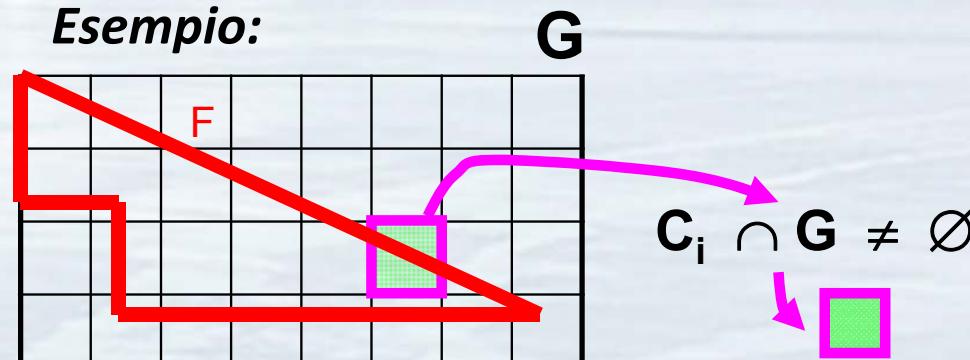
La rappresentazione delle forme basata su regioni è una metodica semplice i cui risultati sono molto promettenti. Data una forma F, la metodologia adottata è così descritta:

Si sovrappone una griglia G che contenga completamente F.

Per ogni cella C_i di G, se $C_i \cap G = \emptyset$ allora $C_i = 0$ altrimenti $C_i = 1$

Si ottiene una stringa binaria che rappresenta la forma leggendo gli 1 o 0 da sinistra a destra e dall'alto al basso.

Esempio:



Rappresentazione:

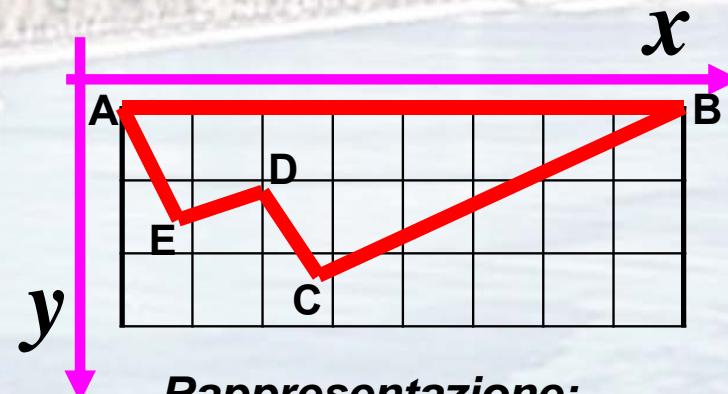
11100000111110000100111001111111



walter.balzano@gmail.com

Normalizzazione rispetto alle rotazioni

- Per ottenere l'invarianza rispetto alla rotazione, la forma viene ruotata in modo tale che il suo asse maggiore sia allineato con l'asse x (orizzontale): Il suddetto allineamento è ugualmente possibile ruotando tutto di 180° ed andrebbero quindi esaminati 2 casi.
- Per evitare di raddoppiare le rappresentazioni di stringhe binarie, generalmente si preferisce, invece, di "raddoppiare" la query e cioè, in altre parole, si effettua una trasformazione sulla query stessa che viene riproposta quindi per la seconda volta.



Query1 = Riga1, Riga2, Riga3

Query2 = Riga3_{rev}, Riga2_{rev}, Riga1_{rev}



walter.balzano@gmail.com

Normalizzazione rispetto ai cambiamenti di scala



- E' possibile effettuare una normalizzazione di scala ridimensionando tutte le forme presenti nel DataBase in modo tale che abbiano tutte lo stesso asse maggiore (di numero di pixel). In altre parole, fissata la larghezza x di una forma è possibile che possa variare solo la sua altezza y (finché non se ne conosce anche l'eccentricità)
- Il criterio di **univocità di rappresentazione di una forma** è pienamente determinato se:
 - Viene fissata la **CellSize** (per es. 10×10 pixel)
 - Si presume che l'asse maggiore sia univoco
 - Si effettuano le operazioni di normalizzazione di Rotazione
 - Si effettuano le operazioni di normalizzazione di Scala
 - Per ogni forma si conosce la sua eccentricità
- La coppia (**Stringa_binaria; Eccentricità**) determina univocamente la forma



walter.balzano@gmail.com

Calcolo della similarità

Il calcolo della similarità viene generalmente compiuto in 2 fasi:

FASE 1: Eccentricità molto diverse

- Se due forme hanno eccentricità molto diversa tra di loro, allora non c'è alcun bisogno di calcolare la similarità in quanto possiamo assumere che, di conseguenza, le figure siano molto diverse.

FASE 2: Eccentricità uguale o simile

- Se due forme hanno **uguale** eccentricità (e quindi la loro stringa binaria ha uguale lunghezza), la distanza tra di loro si calcola come numero di posizioni delle stringhe che hanno valore diverso



Calcolo della similarità

170x60

ECCENTRICITA' = 2,6

190x30

ECCENTRICITA' = 6,3

120x90

ECCENTRICITA' = 1,3

100x20

ECCENTRICITA' = 5

Esempio:

figura1: **11111111111100000**

figura2: **11111111111111100**

la distanza tra figura1 e figura2 è 3

Se due forme hanno **eccentricità diversa di poco**, si possono aggiungere zeri alla stringa che descrive la forma più "bassa" e quindi applicare la distanza come sopra.

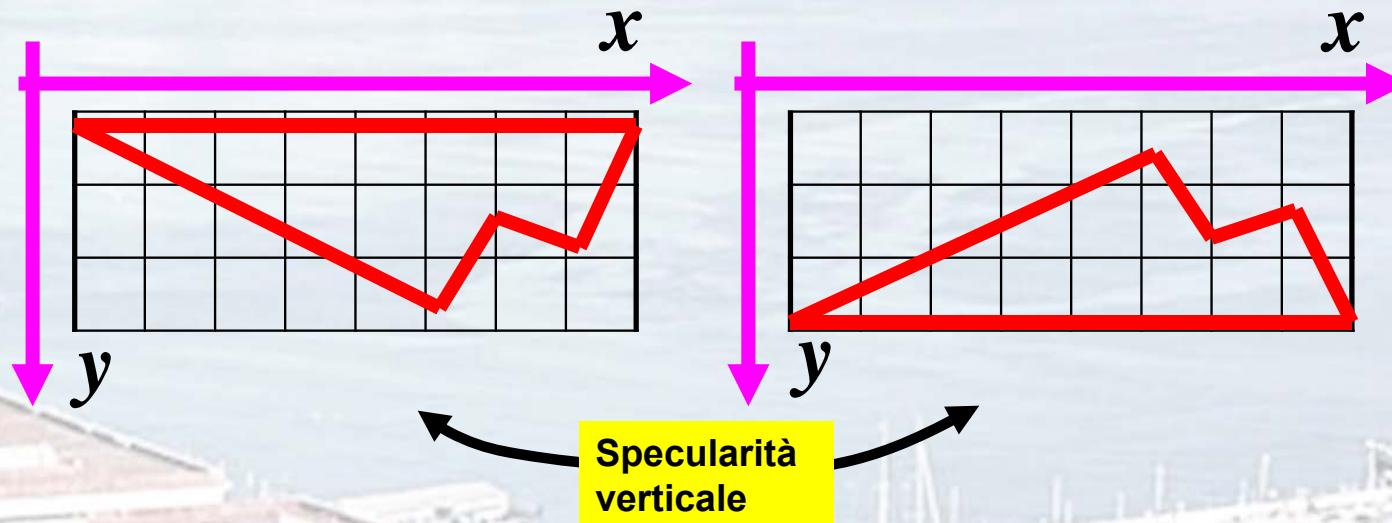


walter.balzano@gmail.com

Forme speculari

- Un tipo di normalizzazione “estesa” considera anche le ulteriori trasformazioni che potrebbero riguardare una figura: una buona strategia di riconoscimento dovrebbe essere resistente anche rispetto alle **trasformazioni di specularità orizzontale e verticale**:
- Anche in questo caso si preferisce non appesantire il database memorizzando 2 sequenze come indice della forma ma, piuttosto, è preferibile effettuare un maggior numero di ricerche **applicando tali trasformazioni alla forma da ricercare**

(considerando anche la rotazione di 180°,
ogni ricerca dovrà essere realizzata con 4 stringhe diverse)



walter.balzano@gmail.com

Riepilogo

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

1. Si determinano gli **assi maggiore e minore e l'eccentricità**. Si **ruota** la forma in modo che l'asse maggiore coincida con l'asse X.
2. Si **scala** la forma in modo che l'asse maggiore abbia una lunghezza prefissata.
3. Si sovrappone una **griglia regolare** alla forma.
4. Si assegnano 1 e 0 alle celle e **si calcola la stringa binaria**. scorrendo le celle della griglia da SX a DX e dall'alto in basso
5. L'indice della forma è costituito dalla stringa binaria e dal numero di celle che la forma occupa.

Durante la ricerca la forma da trovare è elaborata in modo analogo ottenendo 4 stringhe diverse (rotazione di 180° e 2 mirroring) e per ogni forma presente nel database che ha una eccentricità simile si calcola la distanza (si considera la distanza minima tra le quattro rappresentazioni della forma)



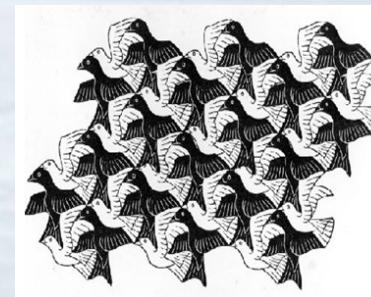
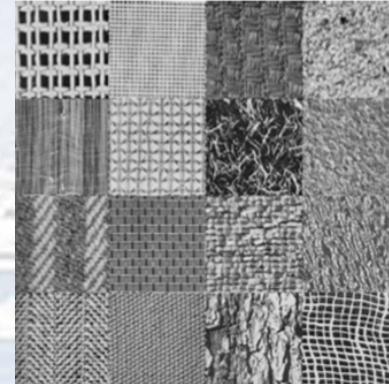
walter.balzano@gmail.com

Metodi basati sulle texture

Texture Analysis

- La Texture (detta anche tessellatura) descrive una “percezione” dell’immagine che è difficilmente descrivibile e riguarda aree caratterizzate da comuni caratteristiche di intensità e struttura

- Granularità (fine o grossa)
- Contrasto
- Direzionalità (direzione dominante della texture)
- Regolarità
- ecc...



- Ci sono esempi di sistemi reali che utilizzano una misurazione matematica (più o meno sofisticata) di tali concetti e la utilizzano per indicizzare e ricercare immagini.



walter.balzano@gmail.com

Altre metodologie

La compressione di immagine sfrutta già **caratteriste “intime” dell’immagine** stessa. Proprio questi parametri di “sintesi” possono essere sfruttati per indicizzare l’immagine

esempio:

- i coefficienti DCT nella compressione JPEG
- i coefficienti Wavelet nella compressione Wavelet

- **Memorizzazione basata su un modello descrittivo dell’immagine**

esempio:

utilizzo delle tecniche di compressione **FRATTALE**

Memorizzazione delle **relazioni tra gli oggetti presenti in una immagine**. E’ simile all’utilizzo della informazione topologica nei GIS: tale relazione consente di inviare al sistema richieste quali:

- **“trova tutte le immagini contenenti un lago vicino ad una montagna”**
- **“trova le immagini contenenti una autostrada sulla sinistra di una foresta”**



walter.balzano@gmail.com

Indicizzazione e Recupero del Video

- **Introduzione**
- **Approcci di Indicizzazione
e Recupero basati su SHOT**
- **Segmentazione**
- **Indicizzazione e Recupero**
- **Rappresentazione ed
Astrazione**



walter.balzano@gmail.com

Introduzione

- La grande quantità di informazione è sicuramente la proprietà più evidente di un flusso dati di tipologia video.
- Il Video è ricco di informazioni in quanto è spesso corredata di:
 - **Sottotitoli** (testo)
 - **Colonna sonora** (audio parlato e/o musica)
 - **Frame** (immagini catturate e riprodotte a velocità costante),
 - **Metadati** (titolo, autore, produttore, ecc...).



walter.balzano@gmail.com

Indicizzazione e ricerca video

Le principali metodologie impiegate sono basate su:

- **Metadati**: informazioni strutturate memorizzate in RDBMS tradizionali.
- **Testo**: sottotitoli o trascrizioni dei dialoghi trattati con tecniche di IR standard per l'indicizzazione e la ricerca.
- **Audio**: segmentazione in parlato e non parlato; tecniche di **speech recognition** per ricavare il testo su cui poi applicare tecniche di IR.
- **Contenuto**:
 - *Metodo 1 (singolo fotogramma)*: il video viene considerato come una serie di immagini indipendenti su cui è possibile applicare tecniche per immagini. Questo approccio comporta evidenti svantaggi:
 - perdita di informazioni temporali
 - la mole di dati da elaborare è molto grande
 - *Metodo 2 (gruppi di fotogrammi)*: i frame vengono suddivisi in gruppi (**shots**) e l'indicizzazione è basata sui frame rappresentativi di ogni shot
- **I migliori risultati derivano combinando due o più metodi.**



walter.balzano@gmail.com

Indicizzazione e ricerca basata sugli SHOT

- **SHOT:** gruppo di frame contigui che hanno uno o più delle seguenti caratteristiche:

- Frame fanno parte della **stessa scena**
- **Frame non sono interrotti** da uno “stacco” della telecamera
- Frame contengono lo **stesso evento / azione**

- *Esempi:*

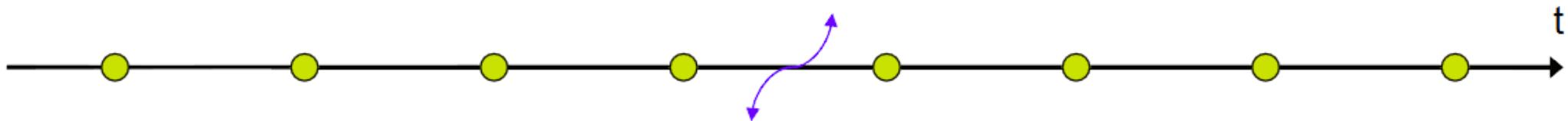
- in un video di un telegiornale, ogni notizia rappresenta uno shot diverso.
- In un film, una serie di immagini con gli stessi protagonisti presenti nello stesso ambiente è uno shot

- In molte applicazioni è necessario non soltanto individuare quale video contiene una determinata informazione, ma anche quale parte del video la contiene.
- Scorrere un video alla ricerca di una particolare sequenza è una operazione che richiede molto tempo:
- Sono necessarie tecniche di **SEGMENTAZIONE** dei video per l'identificazione automatica degli shot



walter.balzano@gmail.com

Segmentazione automatica



- I frame consecutivi che sono ubicati prima e dopo un “cambio di camera” generalmente mostrano un elevato cambiamento quantitativo nel loro contenuto.
- Per individuare gli shot occorre definire una misura quantitativa che catturi la differenza tra una coppia di frame:
 - Data una certa soglia, se la differenza tra un frame e il successivo supera tale valore, il punto viene considerato una interruzione tra shot diversi.



walter.balzano@gmail.com

Segmentazione automatica(2)

Per verificare eventuali passaggi tra shot occorre misurare la differenza tra 2 frame consecutivi.

- **Metodo 1:** si calcola la somma, pixel per pixel, delle differenze tra i 2 frame consecutivi:

■ questo metodo non garantisce buoni risultati poiché la presenza di oggetti in movimento causa grandi differenze anche all'interno della stessa scena e porta a ottenere molte false interruzioni di shot.

- **Metodo 2:** si calcola la differenza tra gli istogrammi di colore dei 2 frame consecutivi:

■ Il metodo è basato sul fatto che il movimento di oggetti all'interno della stessa scena causa piccole variazioni nell'istogramma dell'immagine

- Se le metodologie adottate “scoprono” grosse variazioni allora probabilmente si tratta di un cambio di scena.



walter.balzano@gmail.com

Segmentazione automatica(3)

In particolare definiamo la distanza tra frame:

$$SD_i = \sum_j |H_i(j) - H_{i+1}(j)|$$

In cui:

$H_i(j)$ = valore del j -esimo bin
dell'istogramma dell' i -esimo frame

Se SD_i supera una soglia prefissata allora assumiamo che si sia verificato un cambio di shot.

Ovviamente, la scelta della soglia per l'individuazione degli shot è molto critica per l'ottenimento di risultati corretti: deve tollerare le variazioni presenti normalmente tra singoli frame ma individuare tutti i cambi di shot. Un buon compromesso consiste nel considerare un valore di soglia leggermente più grande della media di tutte le differenze tra un frame ed il successivo.



Shot e transizioni

- Quando tra una scena e l'altra di un video sono utilizzate le transizioni (dissolvenza incrociata, fade-in, fade-out, ecc...) la differenza frame per frame è sempre molto bassa.
- Sarebbe necessario introdurre una soglia molto più bassa ma con il rischio di trovare molti cambiamenti di shot che non corrispondono a effettivi cambi di scena.

Esempio di fade-in e fade-out



Cambio
SHOT “non
evidente”



walter.balzano@gmail.com

Shot e transizioni (2)

Esempio di dissolvenza incrociata



Cambio SHOT “non evidente”

Cambio SHOT “evidente”



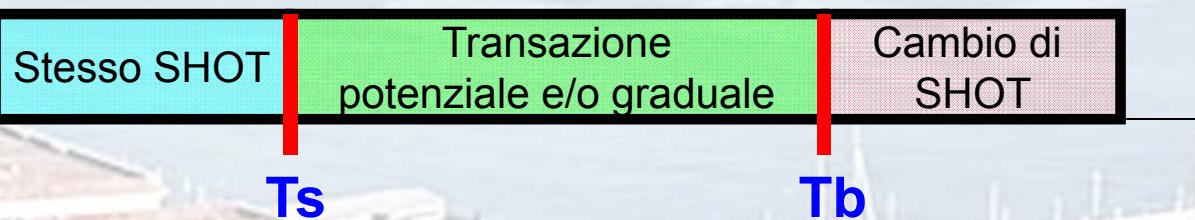
walter.balzano@gmail.com

Segmentazione a due soglie

- Un meccanismo utilizzato consiste nell'introdurre **due soglie diverse**: una più alta **T_b** utilizzata per determinare i cambi di camera e una più bassa **T_s** per determinare i frame nei quali avviene una transizione graduale:

- Confrontando un frame con il successivo, se la differenza è maggiore di T_b si introduce un cambio di shot, se la differenza è minore di T_b ma superiore a T_s il frame è dichiarato come un potenziale frame di transizione.
- Fino a quando si hanno potenziali frame di transizione consecutivi, la loro distanza viene sommata e quando raggiunge la soglia T_b si genera un cambio di shot (se la sequenza di potenziali frame di transizione è interrotta da una differenza inferiore alla soglia T_s la somma riparte da zero)

- In generale il compito di individuazione degli shot è molto più complesso in presenza di transizioni e le percentuali di riconoscimento corretto degli shot sono basse.



walter.balzano@gmail.com

Il problema del panning e dello zoom



Le operazioni di **panning** e **zoom** effettuate con la telecamera generano dei cambiamenti graduali (simili alle transizioni) che possono portare all'individuazione di falsi shot.

- Per limitare questo fenomeno si possono introdurre tecniche di analisi del movimento di blocchi di pixel tra frame successivi.
- Nel caso di panning e zoom i vettori di movimento hanno un preciso comportamento:
 - Nel panning, vanno tutti nella stessa direzione
 - Nello zoom convergono o divergono dal punto centrale dell'immagine.



Cambiamenti di illuminazione

- Un secondo problema è dovuto alla determinazione di falsi shot a causa di cambiamenti di illuminazione.
- Per evitarli, usando diversi step di elaborazione, si possono applicare operazioni di normalizzazione ai colori.
- Normalizzazione dei valori dei pixel (N) dell'immagine:

$$R'_i = \frac{R_i}{\sqrt{\sum_{i=1}^N R_i^2}} \quad G'_i = \frac{G_i}{\sqrt{\sum_{i=1}^N G_i^2}} \quad B'_i = \frac{B_i}{\sqrt{\sum_{i=1}^N B_i^2}}$$

Es: aumentando la luminosità dell'immagine di k volte, il pixel $P1(R_1, G_1, B_1)$ si trasforma in $P2(kR_1, kG_1, kB_1)$ ma $P1$ e $P2$ normalizzati coincidono

Conversione cromatica

$$r_i = \frac{R'_i}{R'_i + G'_i + B'_i} \quad g_i = \frac{G'_i}{R'_i + G'_i + B'_i}$$



Utilizzo degli SHOT

- Dopo aver individuato gli shot, occorre rappresentarli ed indicizzarli in modo da essere efficacemente gestiti dalla fase di ricerca.
- Per ogni shot si determinano uno o più frame rappresentativi (**Representative Frame r-frame**) che vengono usati per l'indicizzazione e la ricerca con tecniche simili a quelle viste per le immagini statiche (*colore, forma o texture*)
- Quanti r-frame occorre utilizzare per ogni shot?
■ 1 r-frame per ogni shot.

Svantaggi:

- Non vengono considerate correttamente le differenze di lunghezza tra i vari shot
- Non è “catturato” completamente il contenuto di uno shot



Utilizzo degli SHOT

■ N r-frame al secondo

Svantaggi:

- Gestisco la lunghezza ma non il contenuto
- Potrebbero anche essere sovrabbondanti nel caso di uno shot molto uniforme

■ Suddivisione in sotto-shot

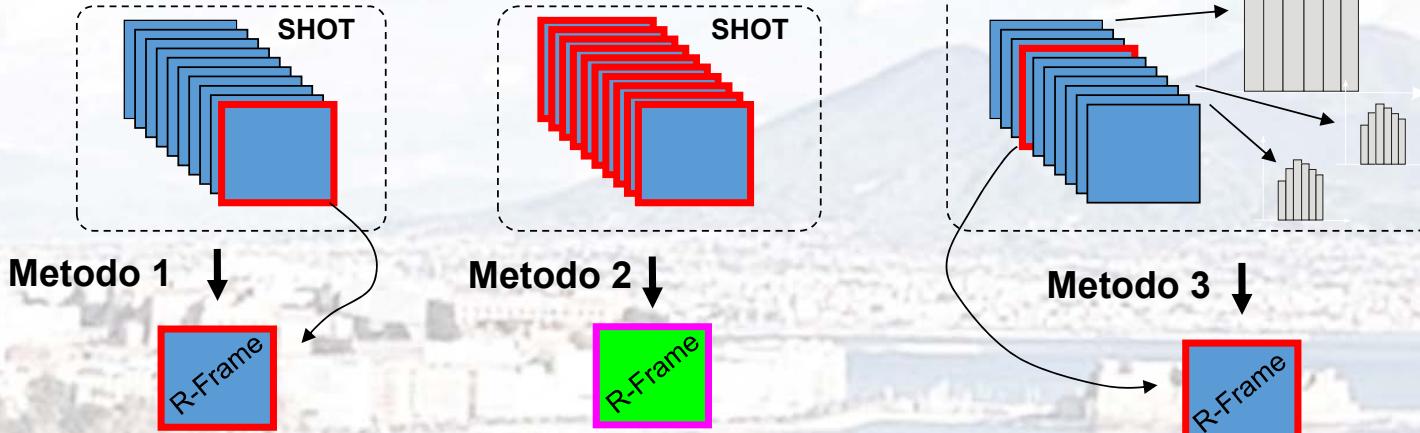
Metodologia:

- Mediante le stesse tecniche utilizzate per la determinazione degli shot, si introduce un nuovo livello e si genera un r-frame per ogni sotto-shot



walter.balzano@gmail.com

Individuazione degli r-frame



- Se il video è abbastanza statico, ognuno dei frame di uno shot può essere scelto come r-frame
- Se ci sono movimenti di oggetti e/o della camera all'interno dello shot, la scelta dell'r-frame riveste una importanza fondamentale per ottenere buone performance di ricerca.

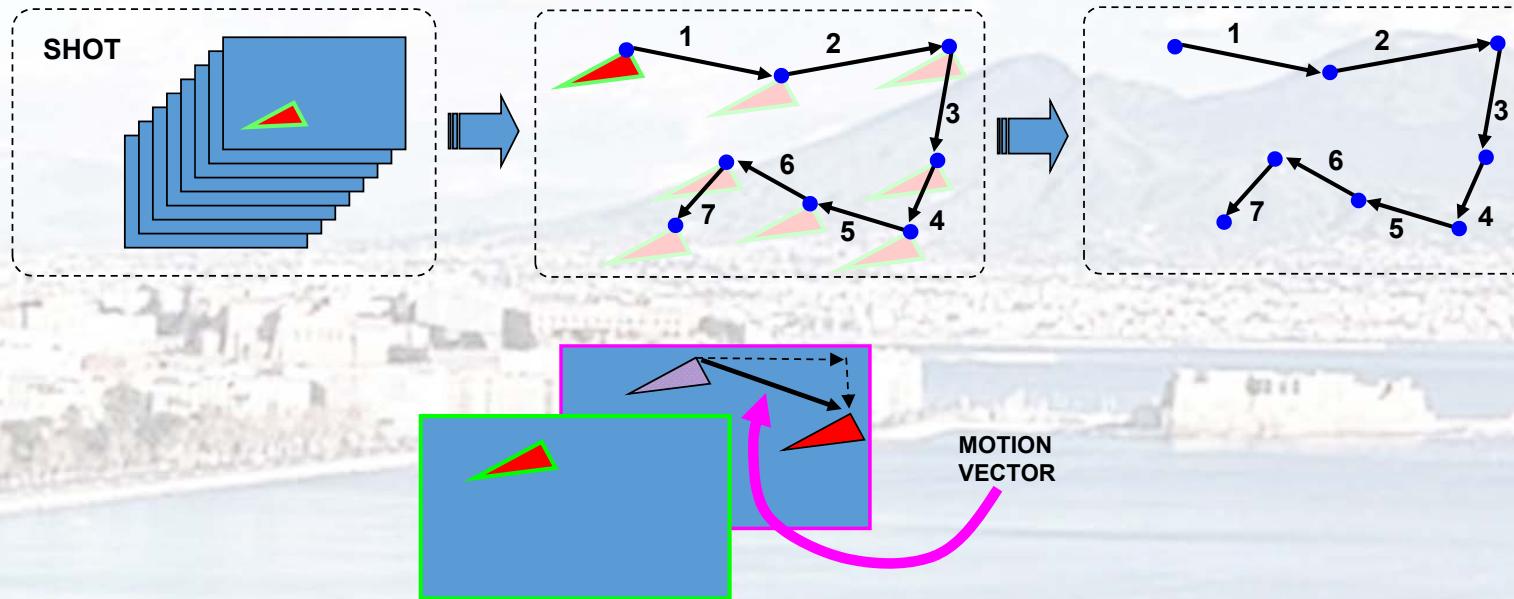
- Metodi possibili:

1. Si considera come r-frame **il primo frame** dello shot (o del sotto-shot)
2. Si calcola un **frame “medio”** facendo la media dei colori di tutti frame dello shot pixel per pixel
3. Si calcola la media di tutti gli istogrammi dei frame dello shot e scelgo come r-frame **il frame il cui istogramma è più vicino all'istogramma medio**



walter.balzano@gmail.com

Tecniche basate sulle informazioni di movimento



- Le tecniche basate sulle informazioni di movimento memorizzano, per ogni frame , le informazioni derivanti dai “**motion vectors**”:
 - direzione prevalente di movimento
 - uniformità del movimento
 - Panning
 - Zoom
- Ogni shot viene indicizzato attraverso i motion vectors che lo caratterizzano



walter.balzano@gmail.com

Tecniche basate sul riconoscimento degli oggetti

- Le tecniche di riconoscimento sono di difficile applicazione nel caso di immagini statiche.
- Contrariamente a quanto può essere intuitivamente pensato, nel caso di dati video tali tecniche funzionano meglio in quanto il movimento degli oggetti tra frame successivi fornisce informazioni utili per segmentare il video e riconoscerne i singoli elementi.



- Gli oggetti riconosciuti sono quindi utilizzati per l'indicizzazione degli shot e la successiva ricerca.



walter.balzano@gmail.com

Rappresentazione dati video

- Ricchi dal punto di vista informativo
- Contengono una **dimensione temporale**
- Metodi per una rappresentazione compatta che catturino le informazioni essenziali da presentare all'utente.

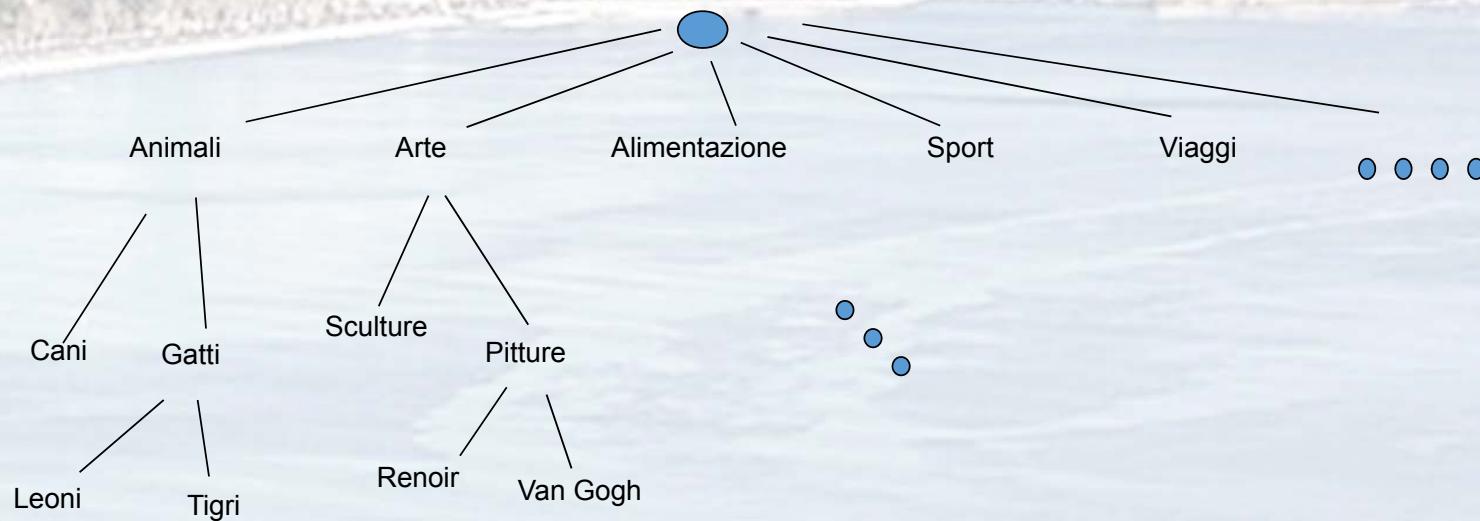
Tali tecniche hanno 3 applicazioni diverse:

- **Video browsing** (scorrere un video ricercando una porzione di interesse): Una rappresentazione compatta velocizzerebbe di molto tali operazioni su video di una certa lunghezza
- **Presentazione dei risultati di una ricerca**: Un sistema di ricerca ritorna video o shot a seguito di una query dell'utente. Una rappresentazione compatta occupa una limitata porzione di schermo
- **Limitazione dei requisiti sulla banda di trasmissione**: Un utente potrebbe consultare la rappresentazione compatta che occupa molti meno byte dell'originale



Classificazione dei video per argomenti

- E' un ausilio alla ricerca in quanto consente una maggiore focalizzazione sugli interessi informativi dell'utente (riduce lo spazio di ricerca)
- Può essere utilizzata sia a livello di intero video che dei suoi singoli shot (e quindi realizzare una rappresentazione compatta del video stesso)



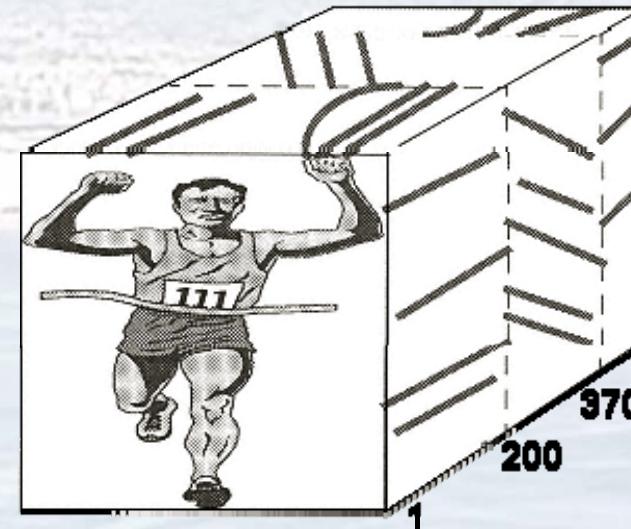
walter.balzano@gmail.com

Presentazione video tramite MICON

MICON (Motion Icon):

rappresentazioni compatta degli shot (l'analogo delle miniatura per immagini fisse)

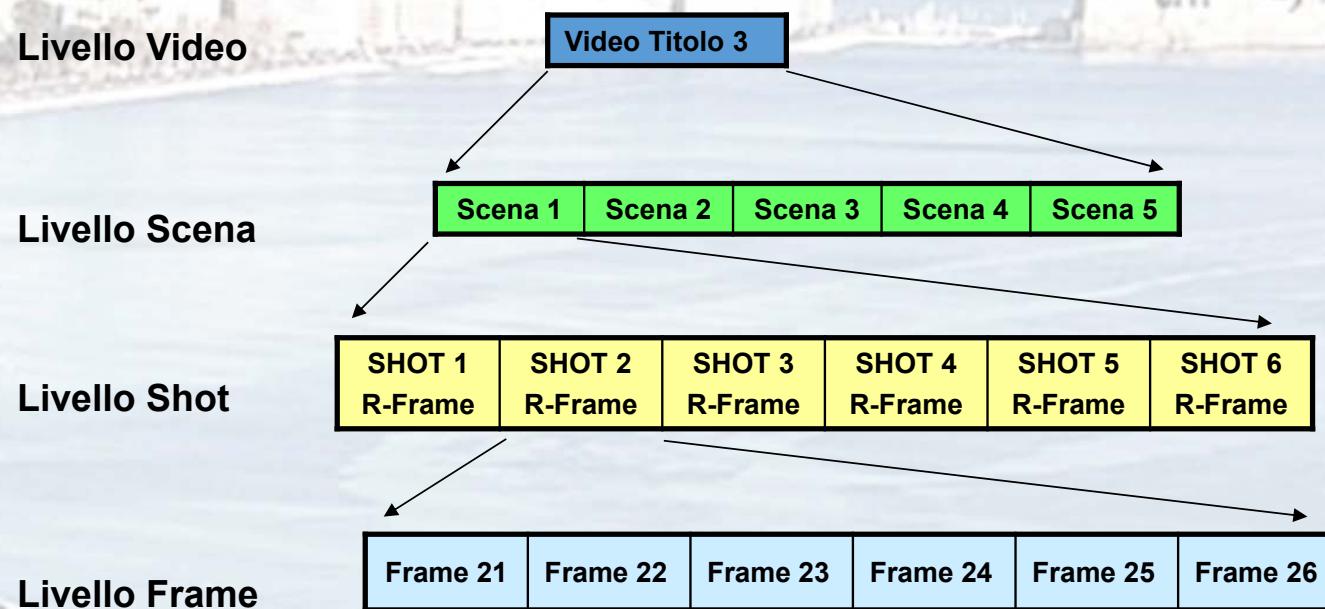
- Una MICON è caratterizzata da tre dimensioni:
 - Un r-frame
 - La profondità (indica la durata temporale)
 - I pixel presenti sui bordi orizzontali e verticali → danno una idea del movimento all'interno dello shot
- Si possono implementare alcune operazioni sulle MICON:
 - Browsing: posso portare in primo piano qualunque frame dello shot
 - Taglio: posso tagliare orizzontalmente o verticalmente il "cubo" per ottenere informazioni aggiuntive



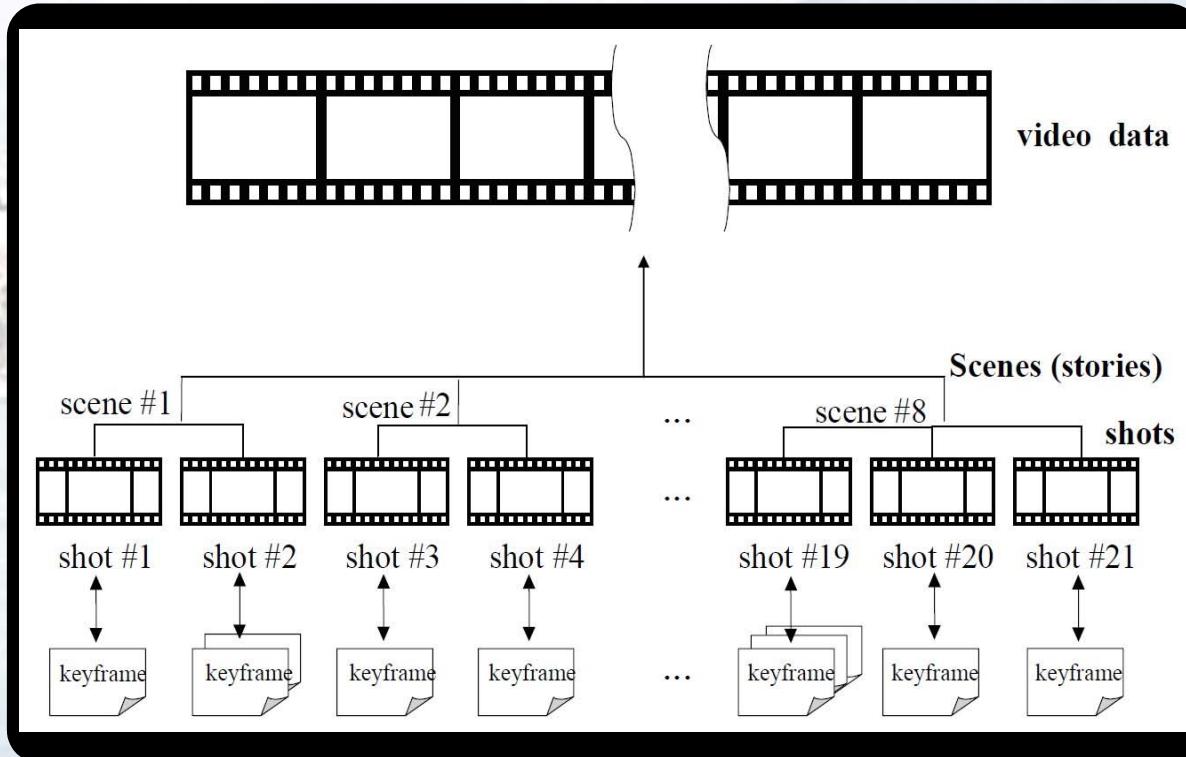
walter.balzano@gmail.com

Browser video gerarchico

- Il browser video gerarchico è uno strumento che consente di ispezionare in maniera gerarchica i vari elementi di un video.
- L'utente si muove sulla gerarchia alla ricerca delle informazioni che gli interessano



Browser video gerarchico (2)



walter.balzano@gmail.com

Altre tecniche di rappresentazione compatta

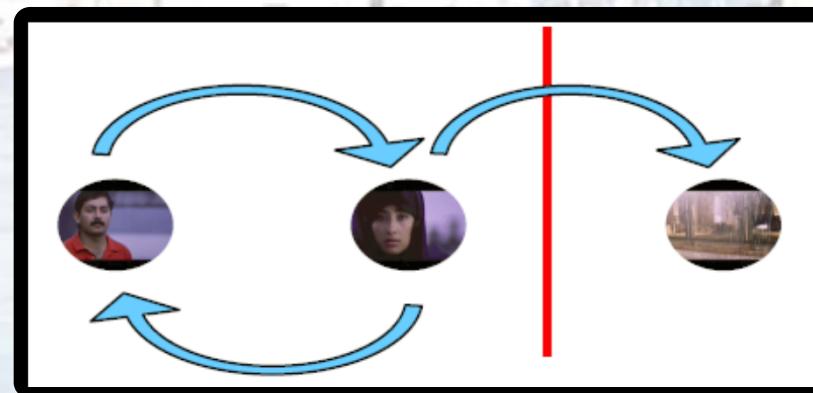
- **Storyboard:**
- Collezione di r-frame che rappresentano le porzioni più importanti di un video
- **Mosaicatura:**
- Unione di frame diversi che nel loro insieme descrivono un elemento complesso (ad esempio un edificio)
- **Scene Transition Graph (STG):**
- E' una struttura a grafo che cattura in modo compatto sia il contenuto che il flusso temporale di un video (utilizza un concetto simile a quello degli ipertesti per la rappresentazione dei documenti)



walter.balzano@gmail.com

Scene Transition Graph (STG)

- Uno Scene Transition Graph (STG) consiste in una rappresentazione di un video mediante un grafo orientato $G=(V,E)$ in cui:
 - V : Vertici. Ciascun Vertice rappresenta un cluster di Shot simili
 - E : Archi. Una coppia di Vertici i e j sono connessi da un arco se esiste uno shot in un cluster i che precede uno shot in un cluster j

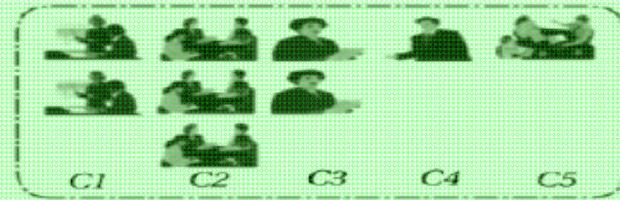
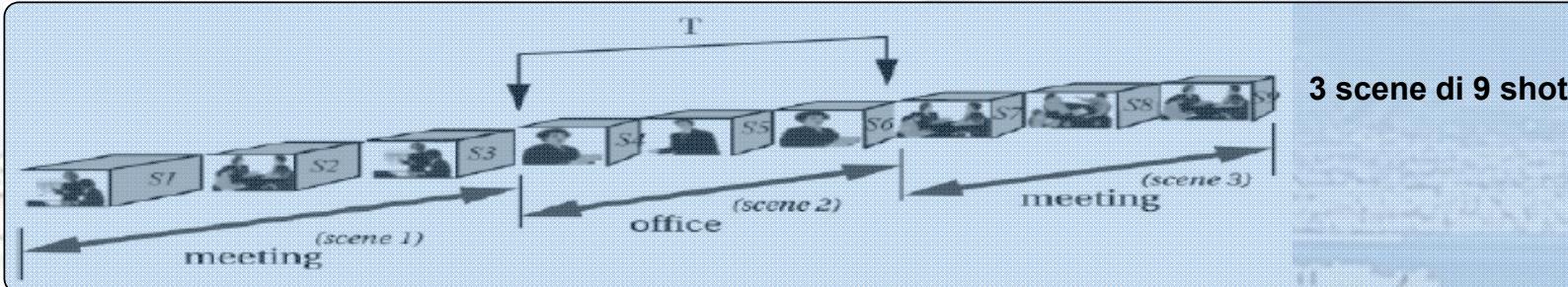


La segmentazione delle scene consiste nel trovare tagli nel grafo la cui applicazione genererebbe due grafi disconnessi

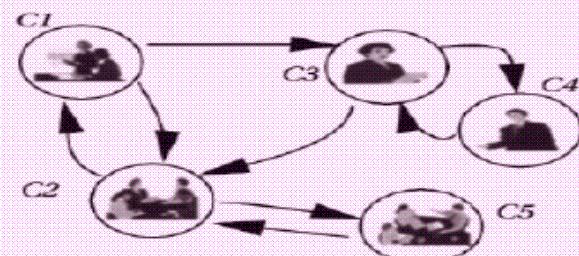


Dal Video all' STG

Esempio di video sequenza con ripetizione di scene



Risultato di una
fase di
raggruppamento



L' STG
Scene
Transition
Graph



walter.balzano@gmail.com

Tecniche e strutture dati per la ricerca efficiente di similarità nei dati multimediali

- Le feature, una volta estratte, vengono rappresentate attraverso **vettori multidimensionali**:
 - Zero crossing rate
 - Silence ratio
 - Iistogrammi di colore
 - ,.. ecc
- La similarità o la distanza sono calcolate come **distanza tra i vettori** che rappresentano le feature estratte.
- Il calcolo della similarità è un'operazione computazionalmente dispendiosa.
 - Se la dimensionalità del vettore è grande, i tempi di ricerca per confrontare tutti i vettori dei dati memorizzati con la query diventano troppo lunghi
 - Occorrono tecniche e strutture dati che permettano di organizzare i vettori multidimensionali e di effettuare le ricerche in modo veloce.
- Lo scopo fondamentale di tali tecniche è di **suddividere lo spazio multidimensionale in sottospazi** in modo tale che solo alcuni di questi debbano essere considerati per rispondere ad una query



walter.balzano@gmail.com

Tipi di query e vettori

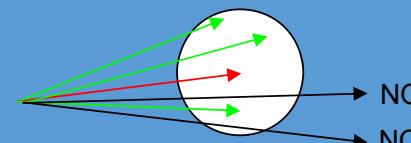
- **POINT QUERY:**

- Una query utente è rappresentata da un vettore e la ricerca ha come obiettivo quello di trovare gli oggetti memorizzati nel database il cui vettore delle feature risulta identico a quello della query (exact match)



- **RANGE QUERY:**

- La query è rappresentata da un vettore e da un range di distanza. Tutti gli oggetti la cui distanza dalla query è inferiore o uguale a quella specificata sono ritrovati (in alcuni casi si possono dare range di distanze per ogni dimensione del vettore)



- **K-NEAREST NEIGHBOR:**

- La query è rappresentata da un vettore e da un intero k: i k oggetti con le distanze più piccole dal vettore di query sono ritrovati

- Es. per k=3



walter.balzano@gmail.com

Tipi di query e vettori

- **Le tecniche e le strutture dati che vedremo potranno permettere una o più delle tipologie di query.**
- Oltre alla ricerca ed al calcolo della distanza si devono permettere **INSERIMENTI ed ELIMINAZIONI**. Tali operazioni non risultano critiche poiché sono possibili off-line)



walter.balzano@gmail.com

Riduzione dello spazio di ricerca tramite filtraggio

- In alcune applicazioni è possibile utilizzare alcuni criteri per focalizzare la ricerca solo in alcune “zone” dello spazio multidimensionale: un filtraggio efficiente serve ad escludere dal confronto una gran parte dei vettori multidimensionali.
 - Filtraggio tramite classificazione
 - Filtraggio basato sulla diseguaglianza triangolare
 - Filtraggio per histogrammi di colore
 - Filtraggio basato sul colore medio



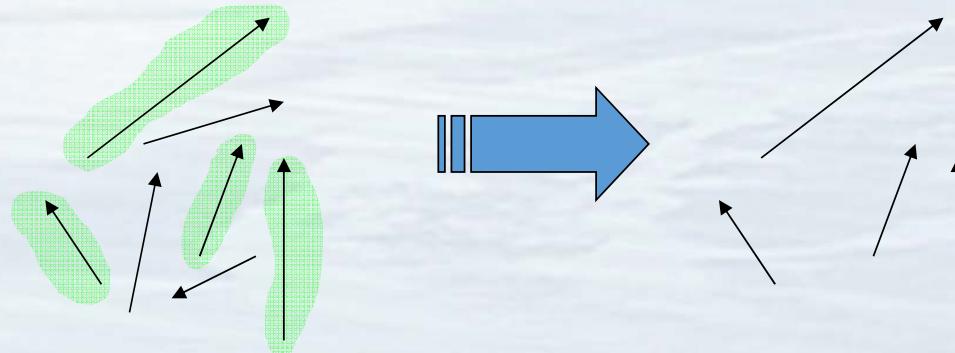
walter.balzano@gmail.com

Filtraggio tramite classificazione

- Filtraggio tramite classificazione, attributi strutturati o parole chiave

- Utilizza informazioni associate ai dati multimediali per ridurre lo spazio di ricerca.

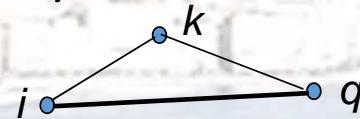
- Esempio: l'utente può essere interessato ad un intervallo di date di produzione o a prodotti di un autore/regista, ecc...)
 - (*Nell'indicizzazione delle immagini "region-based" basato sulla forma abbiamo utilizzato l'eccentricità delle forme come criterio di filtraggio: solo le forme con una eccentricità simile sono confrontate*)



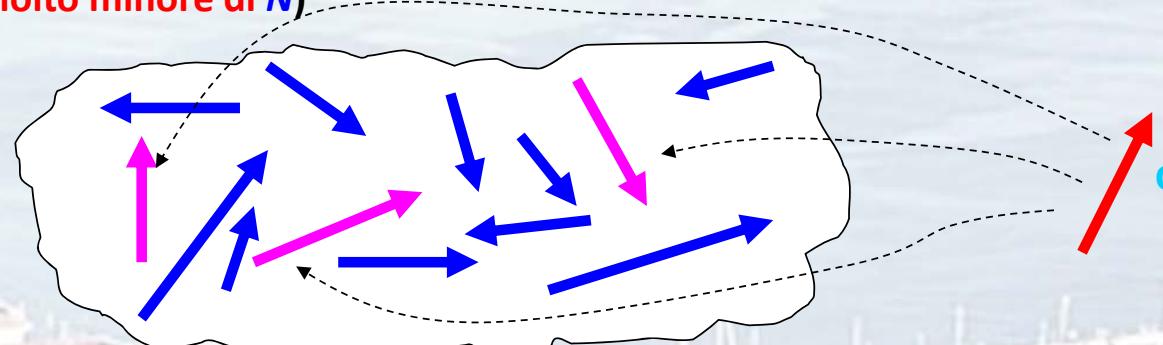
Filtraggio basato sulla diseguaglianza triangolare (definizione)

- Per molte metriche di distanza (es. distanza euclidea) vale la “diseguaglianza triangolare”:
- La distanza d tra due oggetti i e q non può essere inferiore alla differenza tra le loro distanze con un terzo oggetto k , cioè

- $d(i, q) \geq | d(i, k) - d(q, k) | \Rightarrow d(i, q) \geq \max_{1 \leq j \leq m} | d(i, k_j) - d(q, k_j) |$
• (i, q, k sono vettori di feature)
($m = \text{numero di vettori di feature}$)



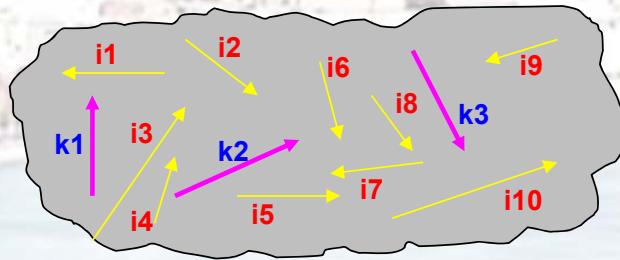
- Obiettivo fondamentale del filtraggio è la riduzione del numero dei confronti (tra la query q e gli N vettori di feature del database): partendo dagli N vettori presenti nel database, solo un sottoinsieme di m vettori è scelto come base di confronto (con m molto minore di N)



Filtraggio basato sulla diseguaglianza triangolare (algoritmo)

- **FASE 1 OFF-line (indicizzazione):**

- Per ogni vettore i -esimo presente nel database e per ogni vettore di confronto K_j calcoliamo la distanza $d(i, K_j)$ e la memorizziamo nel database.



ITEM	$d(i, k1)$	$d(i, k2)$	$d(i, k3)$	$d(i, k4)$	$d(i, k5)$
i1	Δ	Δ	Δ		
i2	Δ	Δ	Δ		
⋮	⋮	⋮	⋮	⋮	⋮

ITEM	$d(i, k1)$	$d(i, k2)$	$d(i, k3)$	$ d(i, k1) - d(q, k1) $	$ d(i, k2) - d(q, k2) $	$ d(i, k3) - d(q, k3) $	$L(i)$
i1				Δ	Δ	Δ	Δ
i2				Δ	Δ	Δ	Δ
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮



Filtraggio basato sulla disuguaglianza triangolare (algoritmo)

FASE 2 ON-line (ricerca):

- Si calcola la distanza $d(\mathbf{q}, \mathbf{K}_j)$ tra la query \mathbf{q} ed ognuno dei vettori di confronto \mathbf{K}_j .
- Si calcola $L(i) = \max_{1 \leq j \leq m} |d(i, k_j) - d(q, k_j)|$ per ogni vettore i presente nel database
- Soltanto i vettori i con $L(i)$ inferiore ad una determinata soglia T sono selezionati per calcolare la loro distanza $d(q, i)$.
- Per gli altri vettori non calcoliamo la distanza da \mathbf{q} in quanto, per la disuguaglianza triangolare, tale distanza è maggiore della soglia T



walter.balzano@gmail.com

Filtraggio basato sulla disuguaglianza triangolare (esempio)

• Si supponga di avere:

- un database multimediale con 8 oggetti memorizzati da i_1 a i_8
- due vettori di confronto (k_1 e k_2).

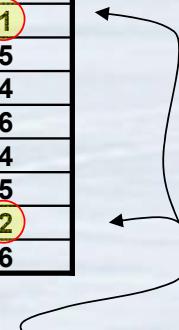
• In modalità offline si calcolano le distanze tra ogni oggetto del database e i due vettori di confronto.

• **Ricerca:** si supponga poi di voler trovare tutti gli oggetti del database che distano meno di 3 dalla query q e che le distanze tra q e i due vettori di confronto siano 3 e 4;

in altre parole: $T < 3$ e $d(q, k_1) = 3$ e $d(q, k_2) = 4$

ITEM	$d(i, k_1)$	$d(i, k_2)$	$ d(i, k_1) - d(q, k_1) $	$ d(i, k_2) - d(q, k_2) $	$L(i)$
i_1	2	5	1	1	1
i_2	4	9	1	5	5
i_3	7	2	4	2	4
i_4	9	3	6	1	6
i_5	3	8	0	4	4
i_6	2	9	1	5	5
i_7	1	4	2	2	2
i_8	4	10	1	6	6

Conclusione: Solo gli oggetti i_1 e i_7 potrebbero avere distanza da q inferiore a 3 e quindi solo 2 oggetti su 8 devono essere confrontati con q .



Filtraggio per histogrammi di colore

- Un istogramma è composto da n bin e la distanza tra due histogrammi è calcolata sommando le differenze tra i valori dei bin corrispondenti.
- Per ottenere una elevata velocità di ricerca si possono:
 - Ridurre il numero dei bin (diminuendo l'accuratezza in quanto colori diversi vengono considerati uguali)
 - Effettuare il calcolo della distanza solo per un sottoinsieme delle immagini presenti nel database
- Le due strategie si possono combinare nel seguente modo:
 - Si utilizzano histogrammi con pochi bin per selezionare i possibili candidati tra tutte le immagini presenti nel database
 - Di tale sottinsieme si calcola la differenza con la query utilizzando gli histogrammi completi



walter.balzano@gmail.com

Filtraggio basato sul colore medio

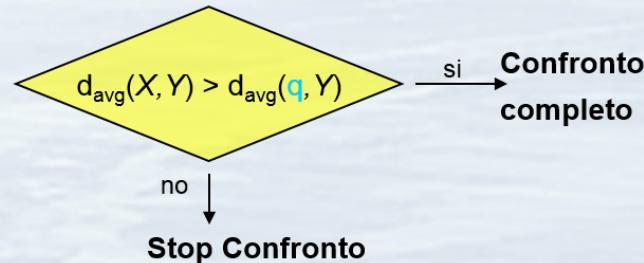
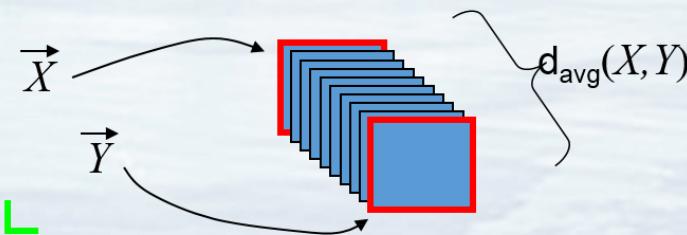
- Il filtraggio può sfruttare il colore medio delle immagini.
- Si definisce il **vettore dei colori medi** $M(R,G,B) = M(R_{avg},G_{avg},B_{avg})$ le cui componenti sono definite da:

$$R_{avg} = \frac{\sum_{p=1}^P R(p)}{P} \quad G_{avg} = \frac{\sum_{p=1}^P G(p)}{P} \quad B_{avg} = \frac{\sum_{p=1}^P B(p)}{P}$$

In cui:
 P = numero di pixel dell'immagine
 $R(p), G(p)$ e $B(p)$ = componenti di colore del pixel p .

Si definisce la distanza euclidea tra i vettori X e Y dei colori medi di 2 immagini:

$$d_{avg}(X,Y) = \sqrt{\sum_{i=1}^3 (x_i - y_i)^2}$$



- Usando il “colore medio”, il filtro riduce il numero dei confronti della query q con tutti gli elementi del database”: si escludono dal confronto completo gli elementi che superano il valore del colore medio precedentemente calcolato. Viene invece effettuato il confronto tra q e gli istogrammi completi che “rientrano” nel valore medio



Strutture Dati Efficienti per la Ricerca della Similarità

- **Introduzione**
- **Alberi B e B+**
- **Clustering**
- **Alberi B+ Multidimensionali**
- **Alberi K-d**
- **Grid Files**
- **R Tree**



walter.balzano@gmail.com

Introduzione

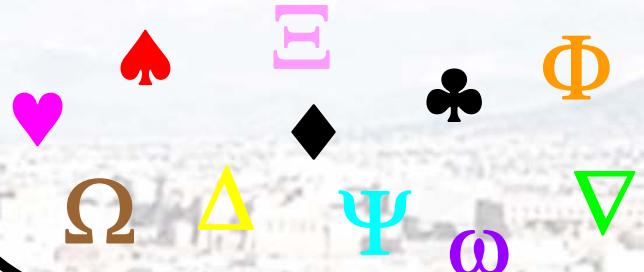
- Per un insieme di oggetti multimediali, la fase di indicizzazione produce un insieme di vettori di caratteristiche (**feature vector**)
- Ciascun vettore V, a seconda della complessità dell'oggetto indicizzato, può contenere un **numero grande di componenti** (caratteristiche testo, audio, immagini, video,...)
- La **fase di Retrieving** è, pertanto, fondamentalmente caratterizzata da un gran numero di confronti di caratteristiche tra la query Q e le caratteristiche degli oggetti precedentemente memorizzate.
- Non è quindi pensabile che questi confronti siano eseguiti in modo “lineare”
- Le strategie proposte sono basate sulla **suddivisione dello spazio multidimensionale** delle caratteristiche in sottospazi.



walter.balzano@gmail.com

Introduzione

dati multimediali



Query

Q

Confronto

Feature Vector

$V(v_1, v_2, v_3, \dots, v_n)$

$Z(z_1, z_2, z_3, \dots, z_n)$

$T(t_1, t_2, t_3, \dots, t_n)$

$R(r_1, r_2, r_3, \dots, r_n)$

.....



walter.balzano@gmail.com

Alberi B

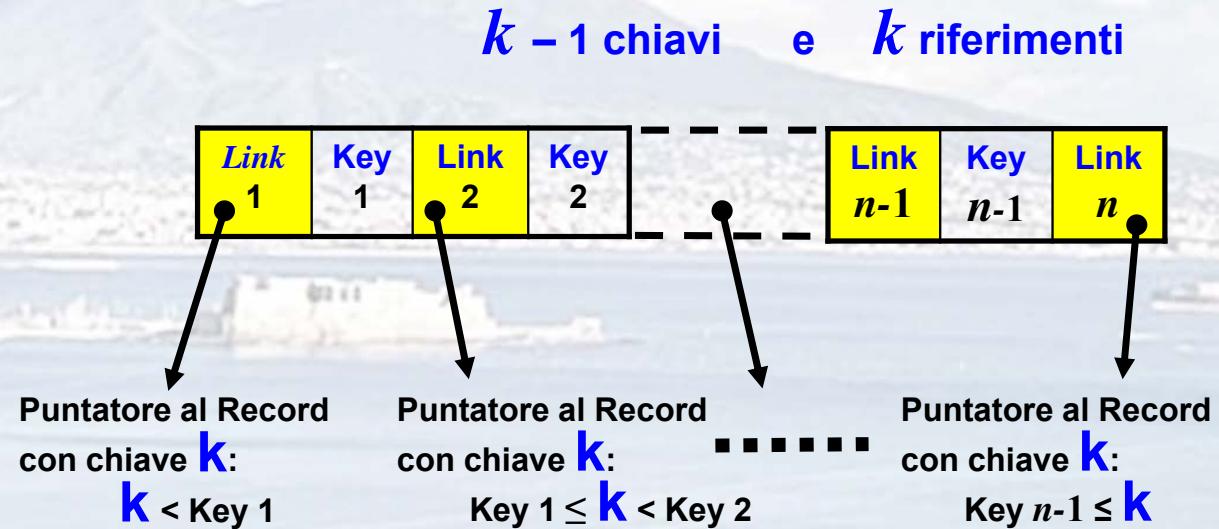
- Un albero B di ordine m (in cui m rappresenta il massimo numero di figli che ogni nodo può avere) è un albero di ricerca generico con le seguenti proprietà:
- La radice ha almeno 2 sottoalberi, a meno che non sia una foglia
- Ogni nodo che non sia la radice e che non sia una foglia contiene $k - 1$ chiavi e k riferimenti a sottoalberi, 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
- Secondo tali definizioni un albero B è sempre almeno pieno per metà, ha pochi livelli ed è perfettamente bilanciato



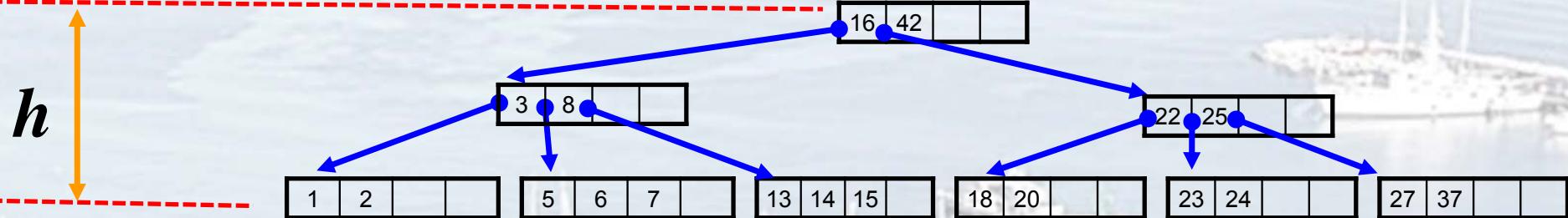
Alberi B (struttura nodo)



walter.balzano@gmail.com



Esempio:



Alberi B (Operazioni e Proprietà)

- Le operazioni fondamentali per gli alberi B sono:

- Creazione
- Inserimento
- Cancellazione
- Ricerca

- Principali proprietà strutturali di un albero B:

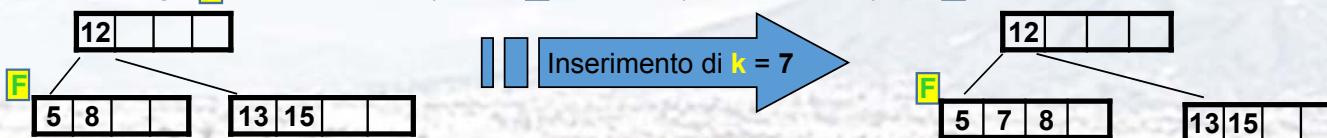
- E' un albero bilanciato
- Prevedibilità per la complessità delle operazioni di ricerca e/o attraversamento dell'albero



Alberi B (Inserimento)

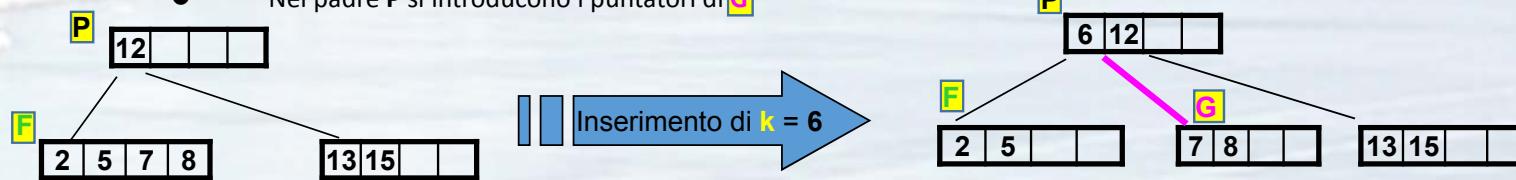
- Per l'inserimento di una nuova chiave **K** nell'albero B si hanno 3 casi:

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



- 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**



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

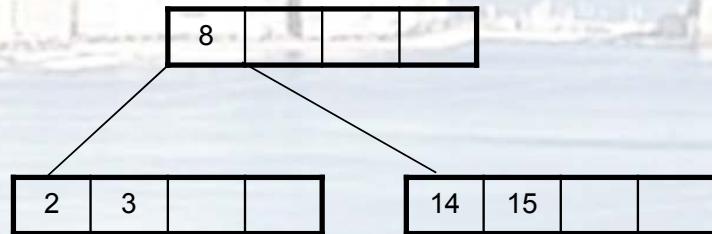


Alberi B (esempio di costruzione ed inserimento)

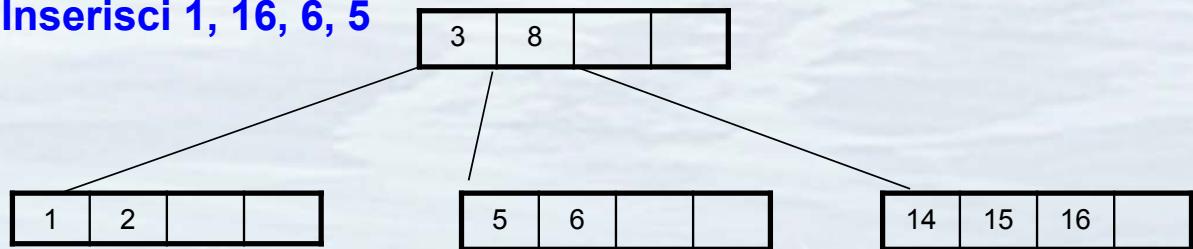
Inserisci 8, 14, 2, 15



Inserisci 3

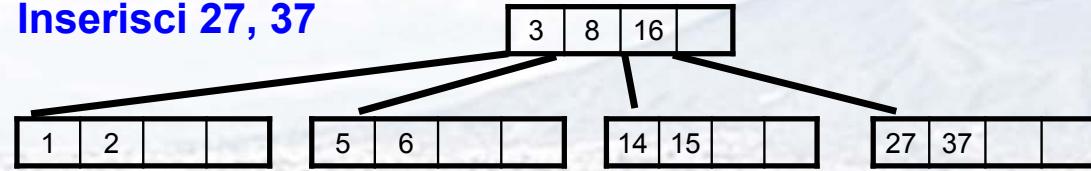


Inserisci 1, 16, 6, 5

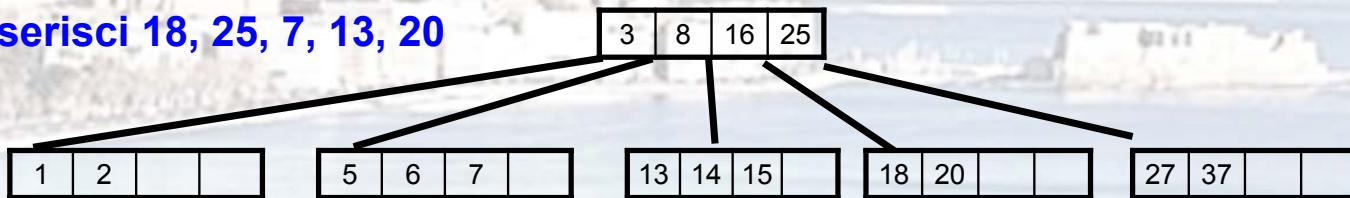


Alberi B (esempio di costruzione ed inserimento)

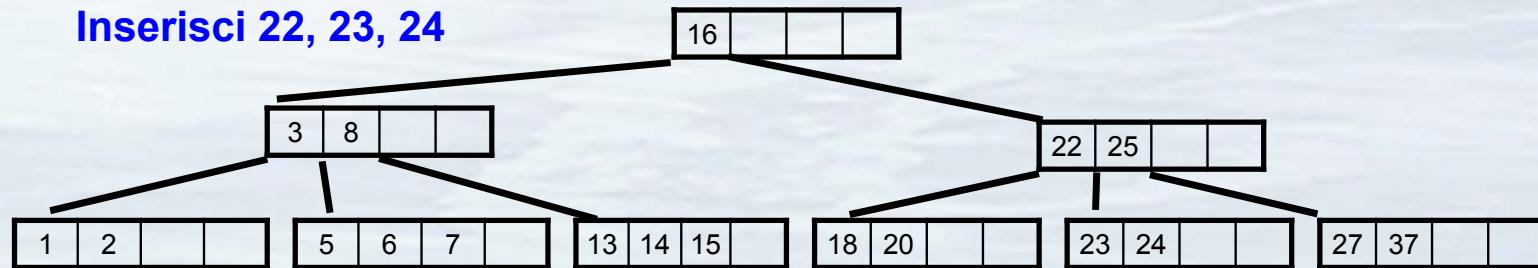
Inserisci 27, 37



Inserisci 18, 25, 7, 13, 20



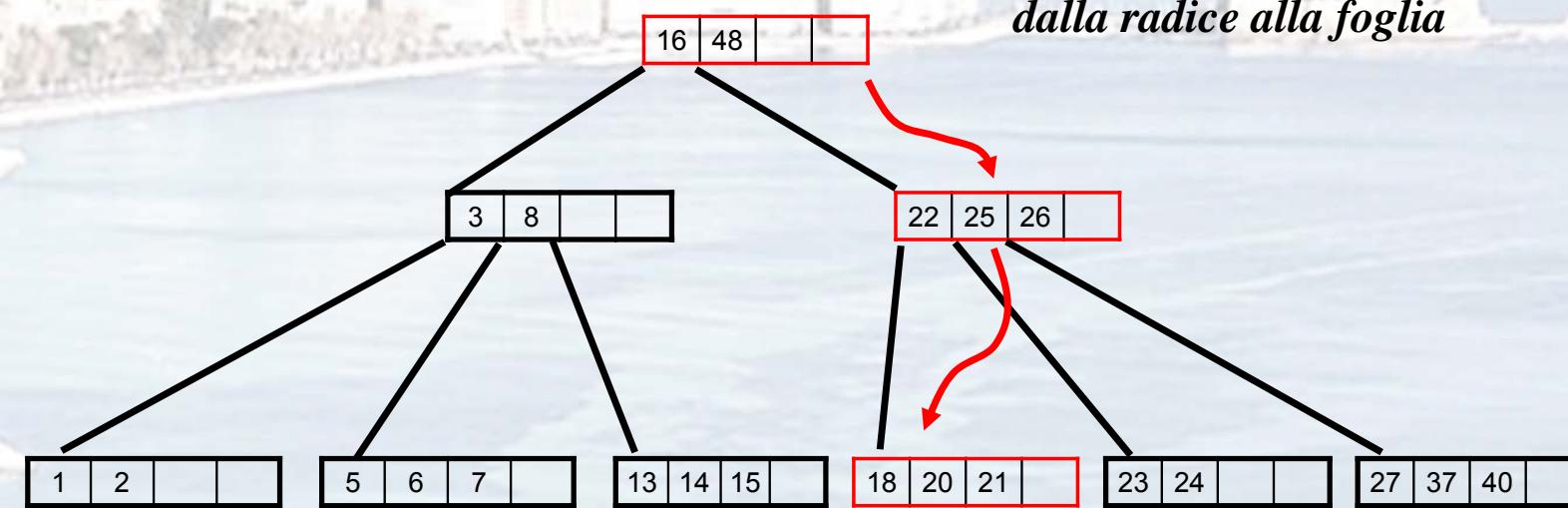
Inserisci 22, 23, 24



Alberi B (Ricerca)

La ricerca sugli alberi B è particolarmente **efficiente** in quanto essi sono intrinsecamente bilanciati

*Esempio di
attraversamento:
dalla radice alla foglia*

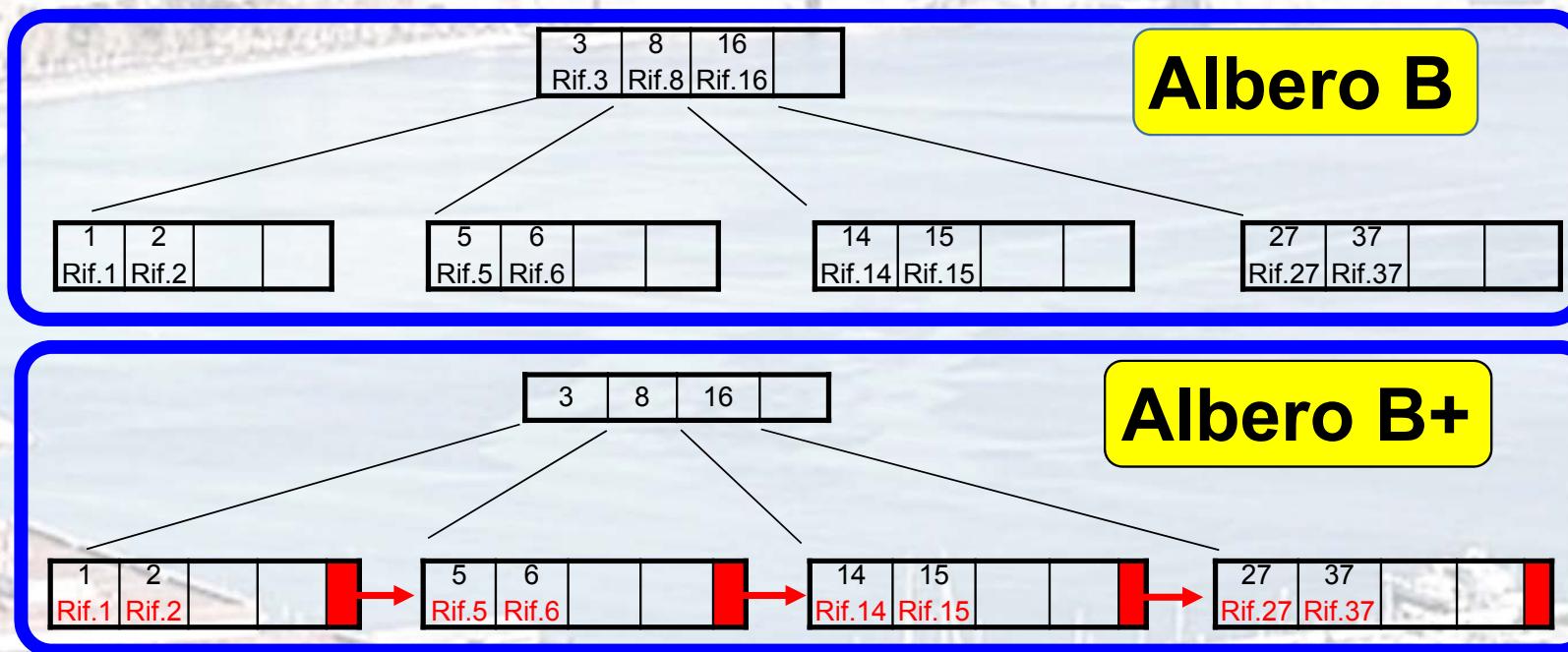


Alberi B+

Gli Alberi B+ possono essere considerati come una variante degli alberi B.

In particolare, le caratteristiche salienti di un Albero B+ sono:

- In un albero B+ i riferimenti ai dati sono contenuti solo nelle foglie (invece in un albero B i riferimenti ai dati sono contenuti in un nodo qualsiasi dell'albero)
- Le foglie di un albero B+ contengono anche un campo puntatore aggiuntivo che permette di “navigare” tra le foglie come una linked list.



walter.balzano@gmail.com

Clustering

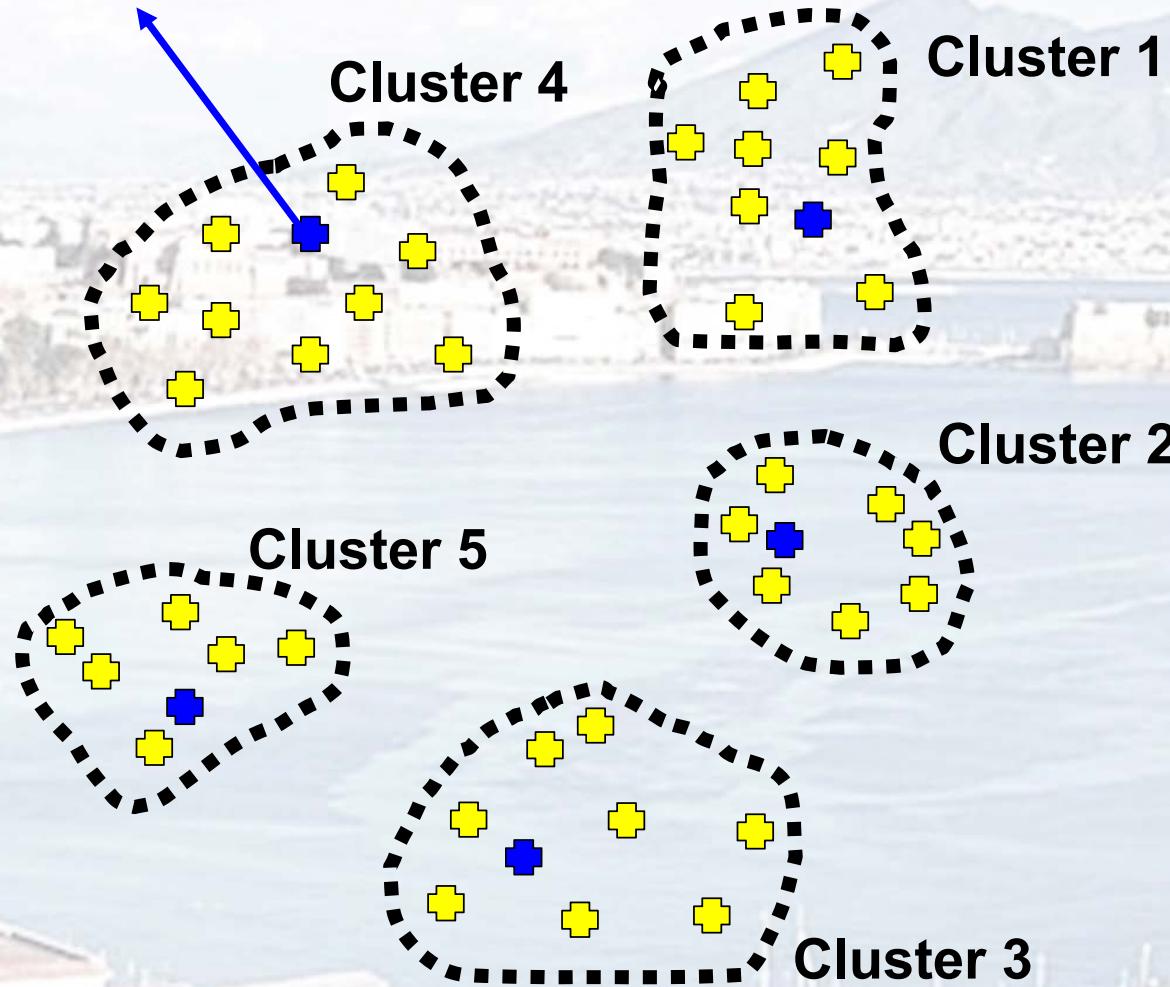
- Tecnica per ottimizzare i tempi di ricerca nello spazio di feature **n-dimensionale**
- Vettori di features simili vengono raggruppati in **cluster** in base a **misure di similarità**
- Ogni cluster è rappresentato dal proprio **centroide**
- Il calcolo della similarità avviene tra la **query** ed il **centroide** di ogni cluster
- I cluster il cui centroide è più simile alla query vengono utilizzati per la **ricerca completa** sui vettori di features che contengono.



walter.balzano@gmail.com

Clustering

CENTROIDE

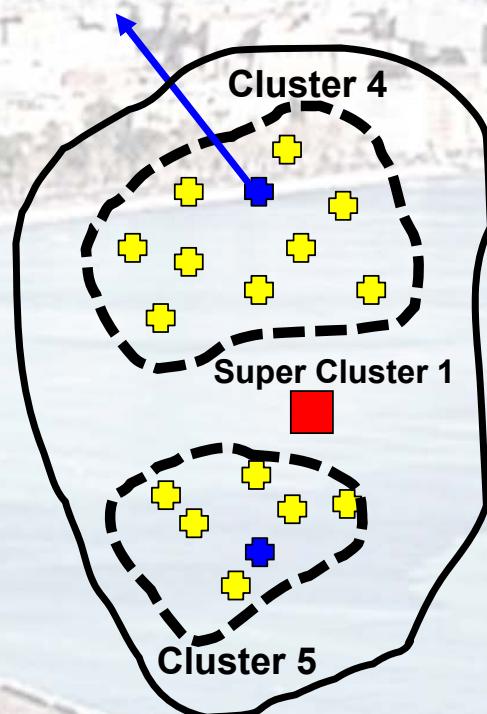


walter.balzano@gmail.com

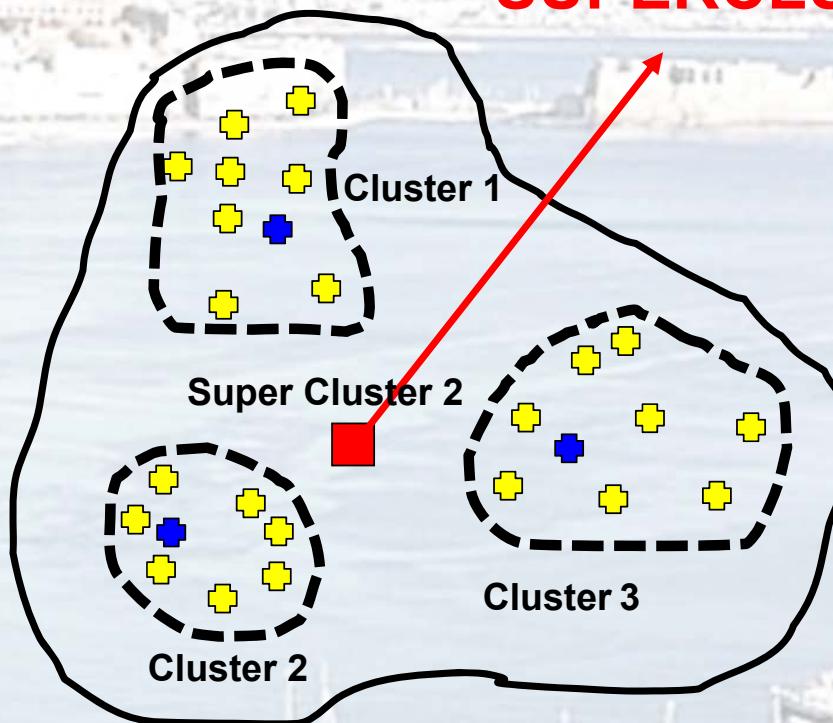
Clustering a più livelli

- Quando il numero di cluster è comunque alto si utilizzano cluster a livelli multipli per ridurre il numero di calcoli di similarità

CENTROIDE



CENTROIDE DEL
SUPERCLUSTER 2



B+ tree MultiDimensionali (MB+ tree)

- Il B+ tree Multidimensionale (**MB+ tree**) è una estensione del B+ Standard (da MonoDimensionale a MultiDimensionale)
- L'MB+ tree supporta le “Similarity Query” (Query per intervalli e per prossimità)

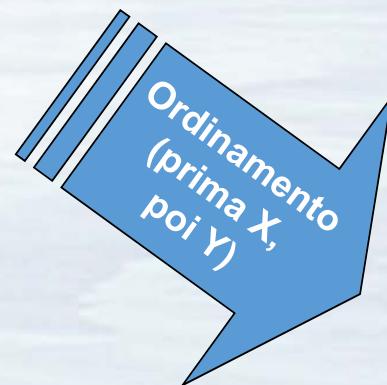
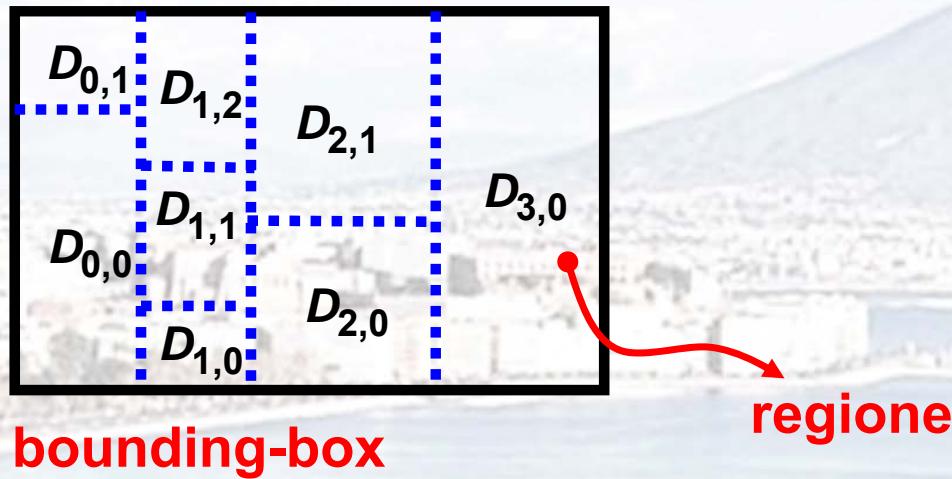
Esempio in 2D:

- Ogni feature vector è un punto nello spazio 2D.
- L'intero spazio delle feature “**bounding-box**” contenente tutti i punti è il rettangolo identificato dallo spigolo in basso a sinistra (x_{\min}, y_{\min}) e dallo spigolo in alto a destra (x_{\max}, y_{\max}).
- Dividiamo il bounding-box in regioni contenenti feature simili.
- **Ordiniamo le regioni secondo un criterio (Prima X, poi Y).**
Ogni regione contiene i puntatori ai feature-vector che ricadono all'interno della regione
- Ogni feature-vector ha un link con il dato multimediale di cui è una rappresentazione



walter.balzano@gmail.com

$B+$ tree MultiDimensionali (MB+ tree)



$D_{0,0}, D_{0,1}, D_{1,0}, D_{1,1}, D_{1,2}, D_{2,0}, D_{2,1}, D_{3,0}$



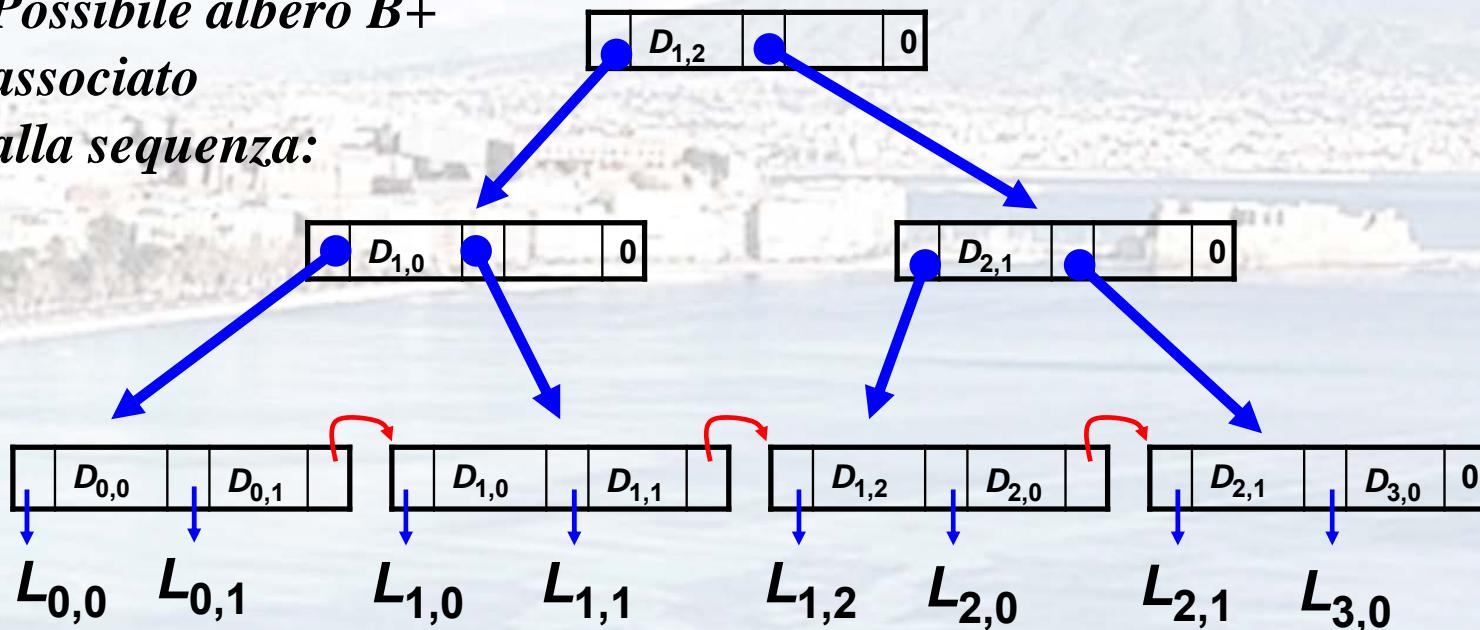
walter.balzano@gmail.com

$B+$ tree MultiDimensionali (MB+ tree)

Ordinamento (prima X, poi Y)

$D_{0,0}, D_{0,1}, D_{1,0}, D_{1,1}, D_{1,2}, D_{2,0}, D_{2,1}, D_{3,0}$

Possibile albero $B+$
associato
alla sequenza:



In cui $L_{i,j}$ rappresenta la lista degli elementi relativi alla regione $D_{i,j}$

Differenze con $B+$ standard:

- L'albero ottenuto dipende dall'ordinamento delle regioni (e non dalle chiavi)
- Ogni regione corrisponde alla lista di vettori di caratteristiche (anziché i record dei dati)



walter.balzano@gmail.com

Ricerca su MB trees MultiDimensionali

Point query

- Ricerca di un vettore dato (x,y)
- Si parte dalla root e si trova la regione che contiene il vettore da ricercare
- Si scorre la lista di feature-vector associata alla regione

Range query

- Ricerca di tutti i vettori che ricadono in un rettangolo
- Partendo dalla root troviamo tutte le regioni che si sovrappongono al rettangolo di ricerca
- Si scorre la lista di feature-vector associata alla regione K

Nearest-Neighbor query

- Ricerca dei k vettori più vicini ad un vettore dato
- Si usa un procedimento iterativo, basato sulla ripetizione di Range query finché non si trova un numero sufficiente di vettori candidati.
- Si usa il calcolo della distanza euclidea tra il vettore da ricercare ed i vettori candidati



walter.balzano@gmail.com

K-d trees

I K-d trees sono considerati una estensione degli ALBERI BINARI:

ALBERO BINARIO

- ogni nodo ha tre elementi:
 - un valore della chiave X
 - un puntatore ai record con chiave < x
 - un puntatore ai record con chiave > x
- È normalmente non bilanciato
- in fase di inserimento si applicano **metodi di bilanciamento** per mantenere i tempi di ricerca in $O(\log n)$



ALBERO K-D TREE

- ogni chiave è costituita dal vettore K-dimensionale invece che da un solo valore
- per generare l'albero occorre regolamentare la modalità di inserimento di un nuovo elemento (sinistra oppure destra):
 - Al primo livello si decide basandosi sulla prima componente del vettore
 - Al secondo livello si decide basandosi sulla seconda componente del vettore
 - ecc...
 - Esaurite le K dimensioni si ricomincia dalla prima componente del vettore



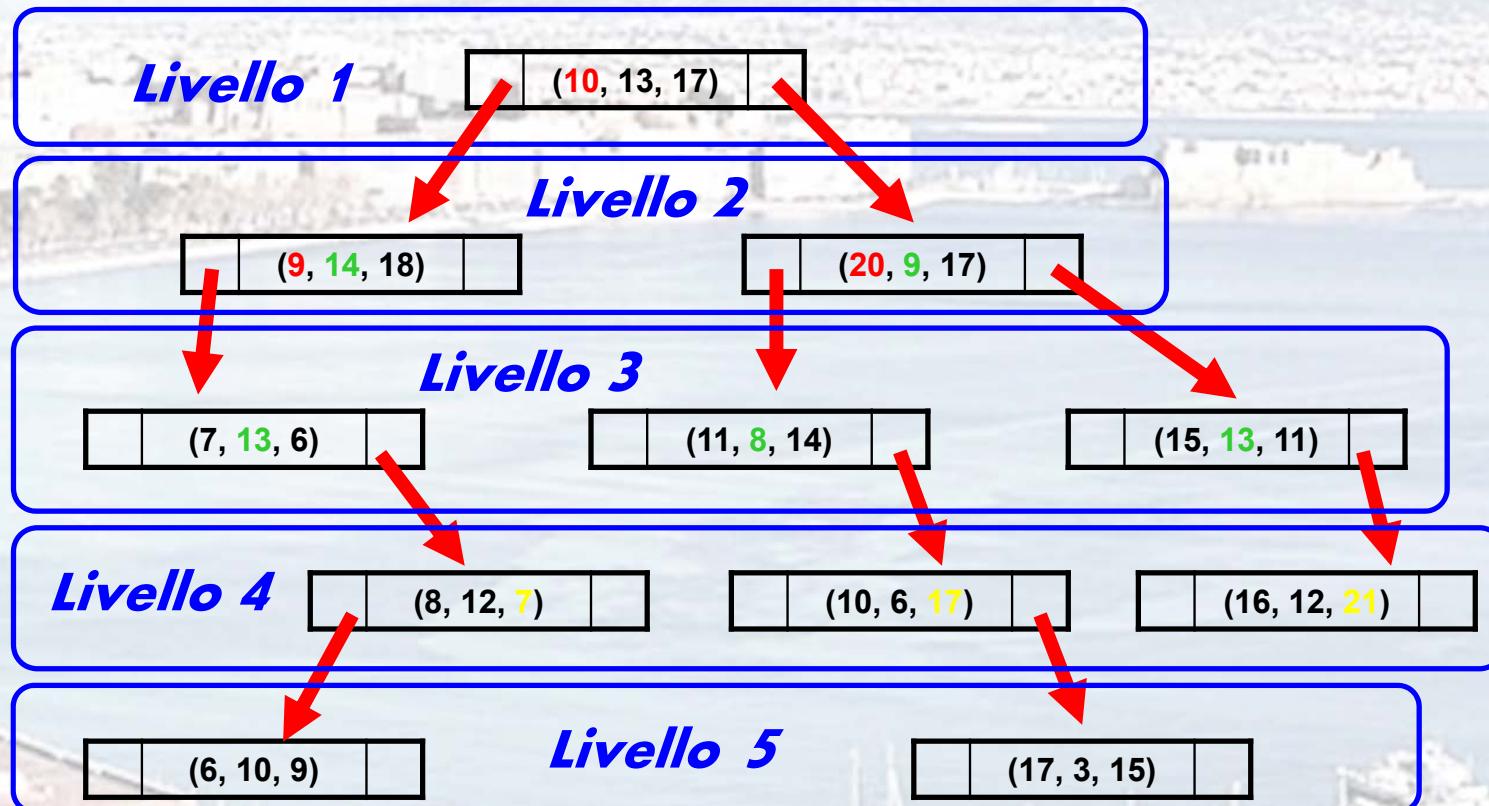
walter.balzano@gmail.com

K-d trees (esempio)

Siano assegnati i seguenti vettori in uno spazio 3D:

(10,13,7) (9,14,8) (20,9,17) (7,13,6) (8,12,7) (6,10,9) (11,8,14) (15,13,11) (10,6,17) (16,12,21) (17,3,15)

• E' possibile costruire il seguente albero:



K-d trees (operazioni)

● Inserimento:

- Per ciascun livello vale l'ordinamento relativo alla componente corrispondente del feature vector.
- Problema: la struttura dell'albero dipende dall'ordine di inserimento dei record.
- L'albero può diventare sbilanciato e richiedere operazioni di ri-bilanciamento

● Ricerca:

- E' simile al processo di inserimento
- Per ogni livello la scelta dipende solo dal valore della relativa componente del vettore

● Eliminazione:

- Può risultare complicata quando occorre eliminare un nodo intermedio dell'albero
- Si possono effettuare delle cancellazioni logiche senza modificare la struttura dell'albero

● Range query:

- Facile da implementare e comporta la riduzione del numero dei nodi da visitare



walter.balzano@gmail.com

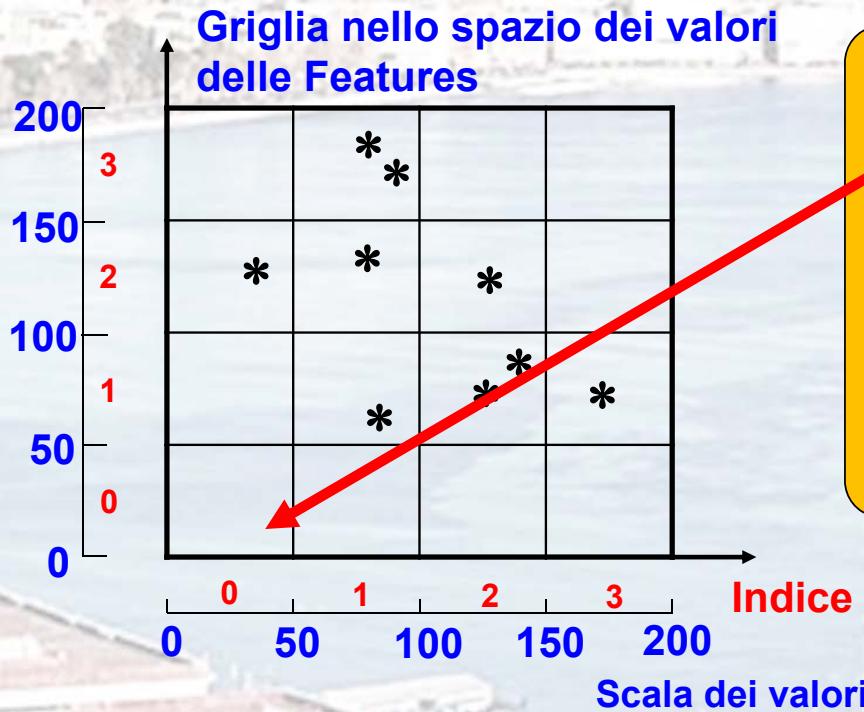
Grid files

I Grid files costituiscono una modalità di indicizzazione e ricerca semplice, spesso impiegata in implementazioni reali.

Consistono nella suddivisione dello spazio n-dimensionale in ipercubi aventi tutti la stessa dimensione.

- Ogni ipercubo contiene zero o più feature-vector

Esempio in 2D:



Array 2D di puntatori alle griglie

$P_{0,0}$	$P_{0,1}$	$P_{0,2}$	$P_{0,3}$
$P_{1,0}$	$P_{1,1}$	$P_{1,2}$	$P_{1,3}$
$P_{2,0}$	$P_{2,1}$	$P_{2,2}$	$P_{2,3}$
$P_{3,0}$	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$



Grid Files (operazioni)

- L'inserimento di un valore nella struttura è molto semplice:

- **Esempio:** se nella griglia 2D precedente si deve inserire il vettore di feature {80, 70} allora esso dovrà essere inserito nella griglia con indice (1,1) e quindi sarà puntato dal puntatore $P_{1,1}$

- La ricerca di un valore avviene in un modo simile:

- **Esempio di point query:** per trovare il vettore {40, 125} basterà puntare alla cella (0,2) della griglia e quindi scorrere la lista di vettori puntati da questo (tutti i vettori di feature che si trovano nella stessa griglia sono puntati dallo stesso puntatore)

- **Esempio di range query:** è necessario trovare tutte le liste di vettori puntate dai puntatori le cui griglie vengono intersecate dal rettangolo descritto dalla query



Grid files (considerazioni)

SE i vettori di features sono distribuiti abbastanza uniformemente all'interno dello spazio dei valori

ALLORA tale metodo da buoni risultati

ALTRIMENTI alcune griglie risultano vuote o quasi e altre sovraffollate

- Se una griglia è sovraffollata il puntatore alla griglia individuerà una lista di vettori molto lunga il cui scorrimento e calcolo della similarità comporta perdita di tempo elevata
- Per far fronte a questo problema (nel caso di distribuzione non uniforme) invece di utilizzare griglie fisse della stessa dimensione si crea una suddivisione adattativa cercando di bilanciare il contenuto delle diverse griglie
- Nelle zone dello spazio densamente popolate si utilizzano griglie di piccole dimensioni mentre nelle zone scarsamente popolate si utilizzano griglie di dimensioni più grandi



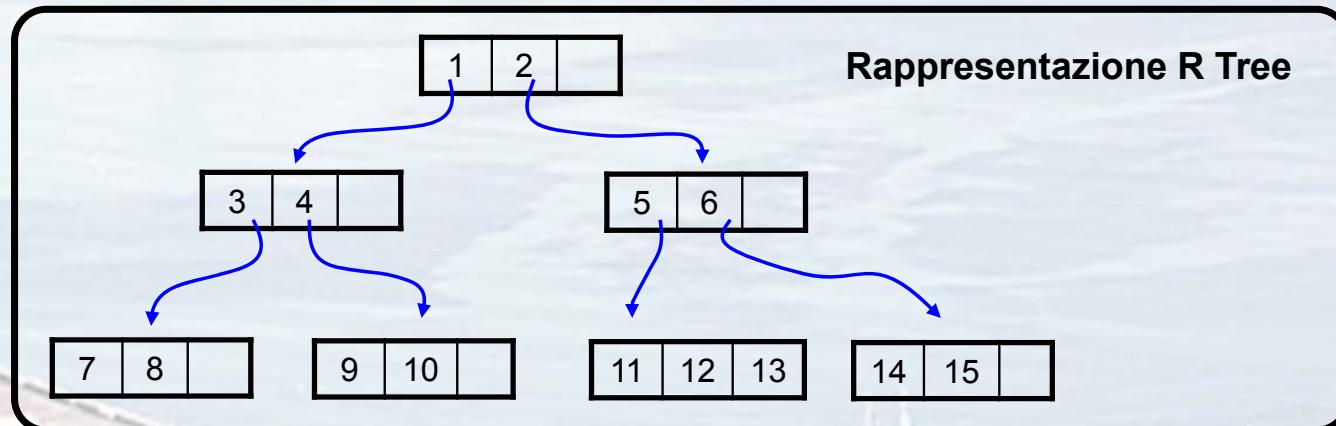
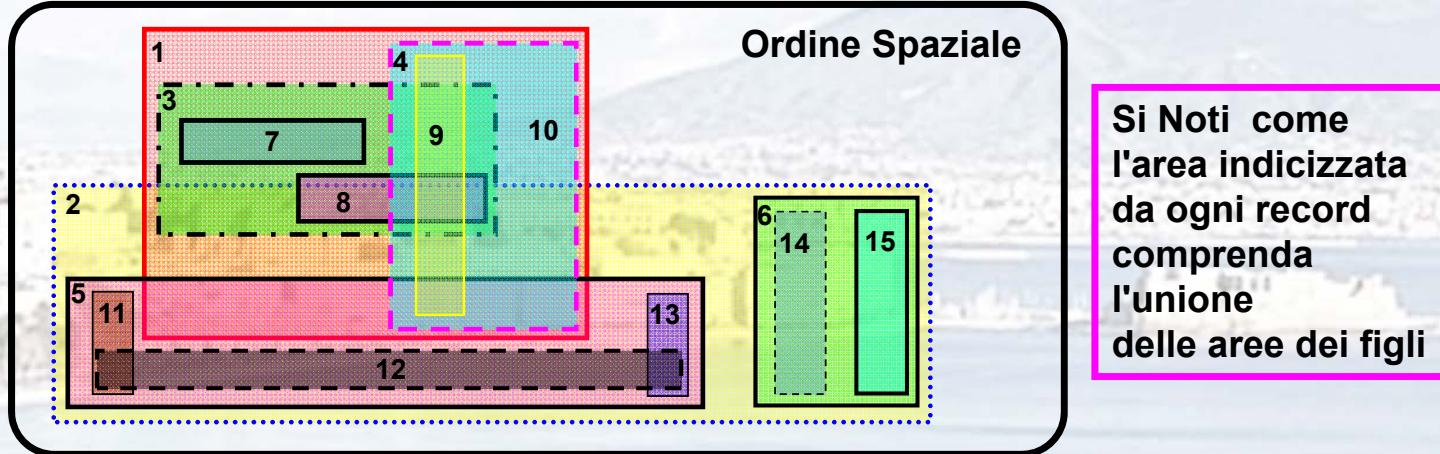
R Tree (Rectangle)

- L' R Tree (**Rectangle**) è una generalizzazione dell' MB+ Tree ed identifica una famiglia di strutture indicizzate molto utilizzate per l'organizzazione dei **dati multidimensionali** (per es. dati spaziali e geografici).
- La struttura dati divide lo spazio in **MBR** (**Minimum Bounding Rectangles**)
- Ogni nodo dell'R-tree ha un numero variabile di entry (fino ad un massimo predeterminato).
- Ogni entry che non sia un nodo foglia contiene due entità:
(A) una entità identifica il nodo figlio, e nell'altra entità c'è
(B) l'MBR che contiene tutte le entry del nodo figlio.
- **(A)** In ogni nodo non foglia viene memorizzato un puntatore che punta ad un nodo di livello più basso nell'albero e un rettangolo che copre tutti i rettangoli associati ai discendenti del nodo.
- **(B)** Nei nodi foglia viene memorizzata la lista dei vettori che ricadono dentro al singolo rettangolo di livello più basso.
- Sono strutture dati utilizzate sia per memorizzare dati che hanno un "bounding-box" che dati di tipo puntuale



walter.balzano@gmail.com

R Tree (Esempio)



walter.balzano@gmail.com

R Tree (Operazioni)

Query:

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

Insert:

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

Delete:

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



walter.balzano@gmail.com

R tree (Dati puntuali ed efficienza di ricerca)

Dati Puntuali

- Per la **ricerca, l'inserimento e l'eliminazione** di dati puntuali in un R TREE gli algoritmi sono simili a quelli precedenti (dati multidimesionali) e l'unica differenza consiste nel fatto che **ogni nodo foglia contiene più di un elemento puntuale di cui viene memorizzato l'MBR**
- E' possibile implementare la **ricerca k nearest neighbor** attraverso la stima di un rettangolo che contiene sicuramente tutti i punti da ricercare e utilizzando la ricerca su oggetti rettangolari descritta in precedenza
- L'inserimento di un punto avviene ricercando il rettangolo dell'albero che deve essere **ampliato di meno** per contenerlo
- In modo analogo al caso di oggetti rettangolari vengono trattati i casi di **splitting** (quando la lista di punti di un nodo
- conterebbe troppi elementi) o di eliminazione di un nodo a seguito della eliminazione di uno o più punti
- **Efficienza di ricerca**
- Dipende da 2 concetti (definiti per ognuno dei livelli dell'albero):
 - **COVERAGE**: E' l'area totale di tutti i rettangoli associati ai nodi del livello
 - **OVERLAP**: E' l'area totale coperta da due o più nodi
- Un R tree è efficiente se **sia la COVERAGE che l'OVERLAP sono minimizzati**; in particolare l'overlap comporta problemi in fase di ricerca.
- Inoltre è cruciale l'ordine di inserimento per ottenere un albero maggiormente bilanciato



walter.balzano@gmail.com

Multimedia WaterMarking

- **Introduzione**
- **Obiettivi**
- **Classificazione**
- **WaterMarking di Testi**
- **WaterMarking di Immagini**
- **WaterMarking di Audio**
- **WaterMarking di Filmati**



walter.balzano@gmail.com

Introduzione

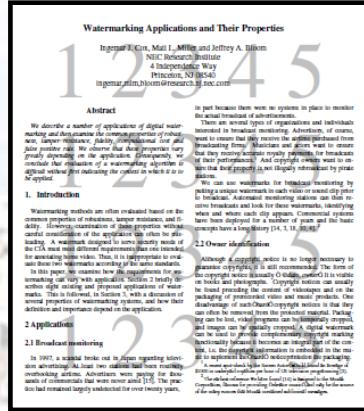
- Con il termine **watermarking** si definiscono tecniche e metodi per l'inclusione di informazioni all'interno di un file multimediale, che possono essere successivamente rilevate o estratte per trarre informazioni sulla sua origine e/o provenienza del file considerato.
- Tali indicazioni, dette **watermark**, possono essere:
 - **evidenti** per l'utente del file (per esempio nel caso di una indicazione di copyright applicata, per esempio, in sovrapposizione su una immagine digitale)
 - **latenti** (nascoste all'interno del file); in tal caso il watermarking può essere considerato una forma di **steganografia** (filigrana digitale).
- Per mezzo del watermark il documento è ancora accessibile, ma contrassegnato in modo permanente



walter.balzano@gmail.com

WaterMark Evidenti vs WaterMark Latenti

WaterMark Evidenti



Esempio di
Articolo Marcato



Crop dell'immagine originale di
Lena (a destra compare marchio)

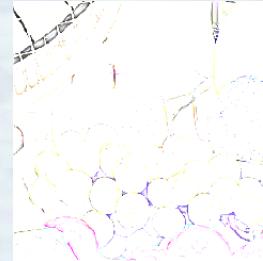
WaterMark Latenti



Originale



Marcata



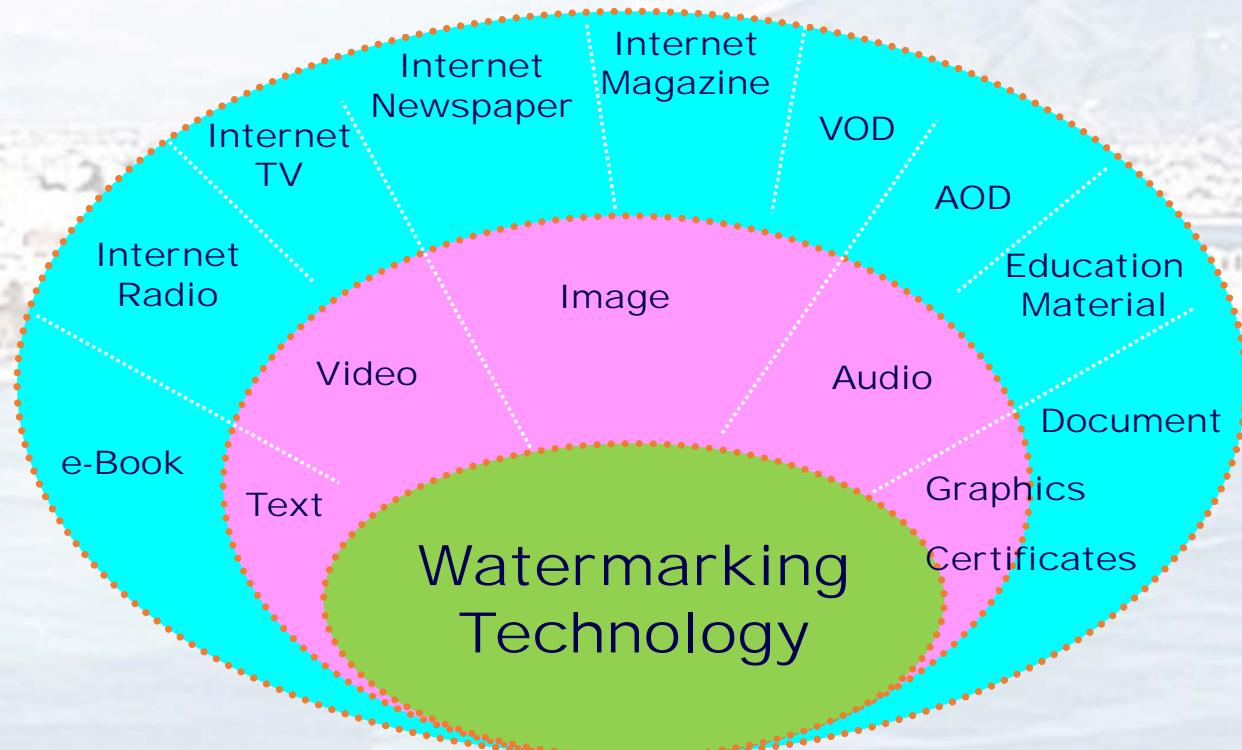
Differenza

Ciao, questo è un
messaggio segreto

WaterMark:
Differenza codifica

walter.balzano@gmail.com

Ambiti di applicazione



walter.balzano@gmail.com

Obiettivi del WaterMarking

Numerosi possono essere gli obiettivi del Watermarking tra cui quelli più comuni sono:

- Rendere manifesto a tutti gli utenti chi sia il **legittimo proprietario** del documento (nel caso in cui il marchio sia visibile);
- **Dimostrare l'originalità** di un documento non contraffatto;
- **Evitare la distribuzione** di copie non autorizzate;
- Marcare alcune **caratteristiche specifiche**;
- Segnare il **percorso di vendita** del **documento**, utilizzando un marchio differente per ciascun acquirente.



Classificazione (1)

I WaterMark possono essere classificati a seconda di alcune loro proprietà, che dipendono dallo scopo con cui sono stati inseriti all'interno del documento. Innanzitutto un watermark può essere visibile o invisibile.

Visibilità:

- WaterMark **visibile**: utilizzato per codificare informazioni che devono essere rese pubbliche all'utente finale.
- Watermark **invisibile**: utilizzato in quei contesti in cui il proprietario legittimo vuole garantirsi i diritti d'autore, nascondendo quindi il marchio nel documento. In pratica la copia marcata è quasi identica all'originale, a meno di alcune **differenze non riscontrabili dalle percezioni umane**.



Classificazione (2)

Resistenza agli attacchi:

- Un watermark **FRAGILE** può essere facilmente attaccato, distrutto e reso irriconoscibile da quasi ogni tipo di manipolazione dei dati.
- Un watermark **SEMIFRAGILE** subisce la stessa sorte di quello fragile se i cambiamenti inflitti sono superiori ad una certa soglia definita dall'utente.
- Un watermark **ROBUSTO** deve resistere alle più comuni operazioni e trasformazioni sui dati, in quanto è utilizzato quando la proprietà del documento deve essere provata o garantita. L'informazione che trasporta non deve perdersi e deve potersi recuperare, anche se il documento è stato modificato. Inoltre si tende a considerare un watermark robusto quando è in grado di resistere anche ad attacchi intenzionali volti alla sua rimozione.



walter.balzano@gmail.com

WaterMark Robusto vs WaterMark Fragile

- Il **watermark robusto** si applica quando occorre dimostrare l'origine di una informazione, anche quando essa viene fortemente distorta e/o manipolata
- Il **watermark fragile** si applica quando occorre dimostrare che una informazione non è più quella originale anche se alterata in minima parte.



Esempio:

Usando uno strumento grafico come Photoshop in pochi minuti
è possibile modificare una immagine:

- Se occorre dimostrare che l'immagine, anche se modificata,
è la stessa allora uso un **watermark robusto**.
- Se occorre dimostrare che l'immagine, anche se minimamente
modificata, non è più autentica allora uso un **watermark fragile**.

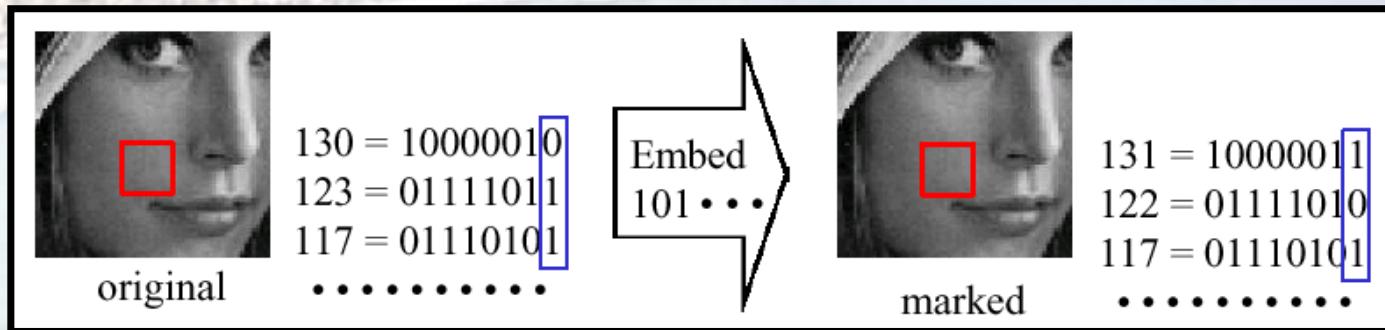


walter.balzano@gmail.com

Esempio

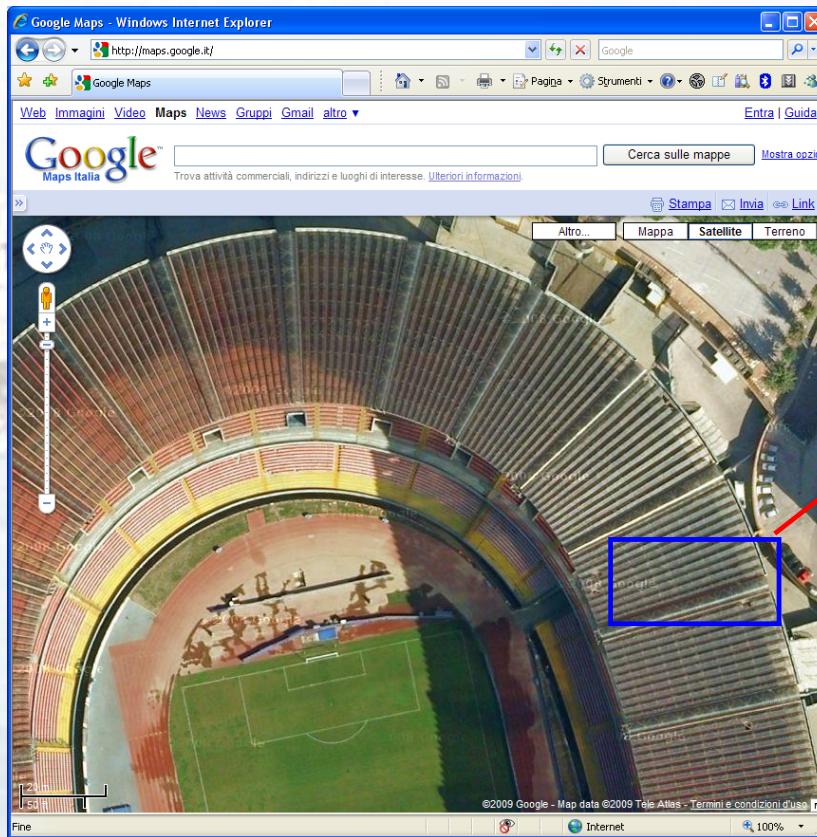
Esempio di

WaterMark non visibile

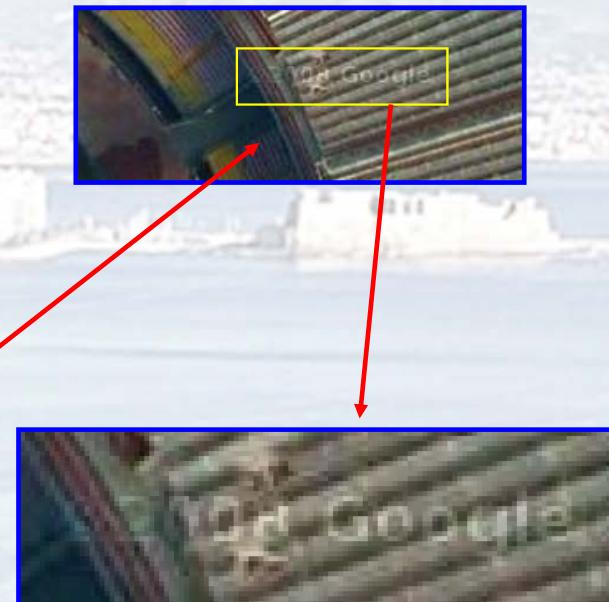


walter.balzano@gmail.com

Esempio



Google Maps Watermark



walter.balzano@gmail.com

Classificazione (2)

Autonomia:

- **Ciechi**: i watermark si definiscono ciechi se per verificare la loro presenza **non è necessario il documento originale**;
- **Non ciechi**: **è necessario il documento originale**;
 - Più robusti, ma non sempre è possibile avere a disposizione l'originale;
 - Solo il proprietario può dimostrare la presenza del marchio.

Dominio Pubblico o privato:

- **WaterMark privati**: possono essere estratti solo quando si conosca a priori il loro contenuto e si possegga il documento originale non marchiato;
- **WaterMark pubblici**: quelli rilevabili anche se non se ne conosce il contenuto, senza l'ausilio del documento originale. Nel caso di watermark pubblici è più semplice identificare e alterare (o rimuovere) il marchio, ma in molti contesti essi sono più utili, in quanto permettono a chiunque di individuare il proprietario del documento



walter.balzano@gmail.com

Proprietà (1)

I watermark possono essere utilizzati per diversi scopi e, quindi, devono soddisfare a ben determinate esigenze. Esistono però delle **caratteristiche comuni** a tutti i watermark:

- il legittimo proprietario o un'autorità indipendente di controllo devono poter facilmente **estrarre le informazioni del watermark**;
- il recupero del watermark **deve provare in maniera non ambigua l'identità del proprietario**;
- deve essere possibile **sovrapporre più watermark sul documento**, senza che i precedenti siano distrutti;
- il watermark deve essere inserito all'interno del segnale da proteggere per maggiore sicurezza e portabilità.



Proprietà (2)

Per soddisfare le precedenti richieste i watermark devono essere:

- **Invisibili**: inserire un segnale di watermark comporta necessariamente un seppur piccolo degrado dell'immagine; questo degrado deve essere il più lieve possibile in modo da non alterare la percezione del documento. Il grado d'alterazione deve essere deciso dal proprietario del documento, il quale può scegliere tra forti alterazioni, che danno una garanzia di robustezza ad eventuali attacchi, e deboli alterazioni, che non degradano il prodotto.
- **Codificati a chiave**: ogni segnale di watermark è associato ad una particolare sequenza di bit detta chiave (key). La chiave serve sia per produrre il segnale di watermark che per riconoscerlo all'interno di un documento. La chiave è privata e caratterizza univocamente il legittimo proprietario del documento. Solo chi è in possesso della chiave è in grado di dimostrare la presenza del watermark nel prodotto digitale. Il numero di chiavi possibili deve essere enorme.



Proprietà (3)

- **Efficienti statisticamente:** un documento firmato con un segnale di watermark deve essere facilmente riconoscibile se si conosce la giusta chiave. La probabilità che la chiave (nella fase di riconoscimento) venga rifiutata, pur essendo corretta, deve essere sufficientemente bassa.
- **Invisibili statisticamente:** possedere un gran numero di documenti digitali, tutti firmati con la stessa chiave, non deve rendere riconoscibile (e quindi eliminabile) la firma. Diversi prodotti firmati con la stessa chiave devono generare segnali di watermark differenti. Dobbiamo essere sicuri che il riconoscimento della chiave all'interno dell'immagine da parte di terzi sia impossibile.



walter.balzano@gmail.com

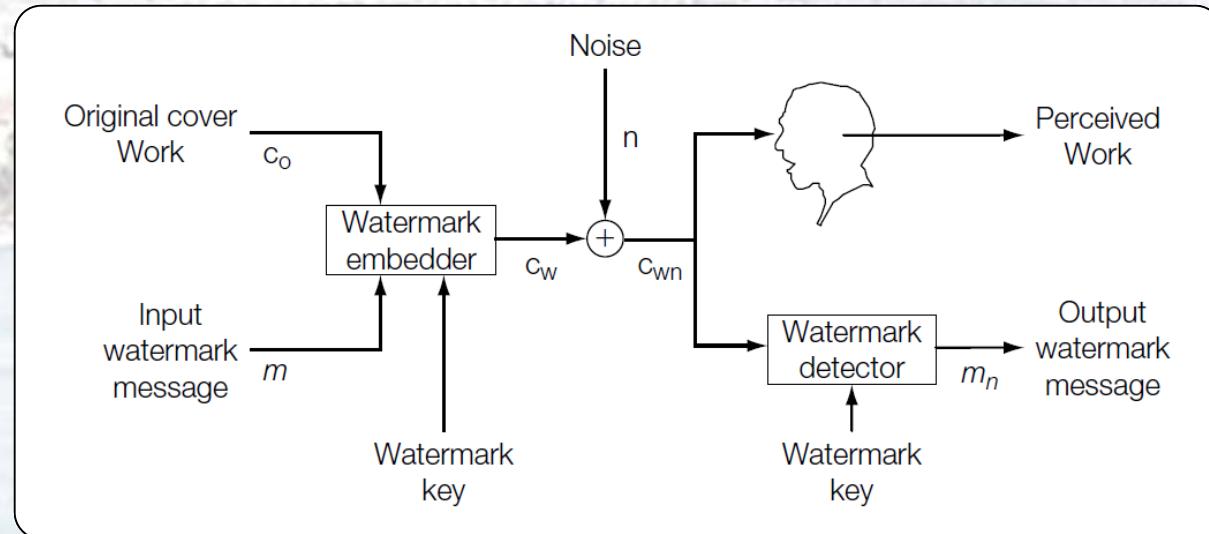
Proprietà (4)

- **Multipli:** deve essere possibile inserire un elevato numero di segnali di watermark all'interno dello stesso documento; ognuno di questi segnali può essere riconosciuto mediante la corrispondente chiave.
- **Robusti:** sui documenti digitali possono venir fatte numerose operazioni per migliorare la loro qualità o per comprimere la loro dimensione. I segnali di watermark devono essere tali da non venire eliminati da questo tipo di operazioni, né da operazioni mirate ad alterare o cancellare il marchio stesso.
- **Invertibili:** il legittimo proprietario del documento deve poter rimuovere il watermark. In realtà questa proprietà non si può ottenere se devono essere garantite la robustezza e la resistenza alle aggressioni



walter.balzano@gmail.com

Schema generale di un sistema stenografico



walter.balzano@gmail.com

Metodologia

L'aggiunta di un watermark all'immagine può essere vista come l'inserimento di una componente di "rumore" nell'immagine stessa. Un qualunque schema di watermarking è realizzato attraverso l'implementazione di due ben specifici algoritmi di codifica e decodifica:

- **Codifica:** la codifica del marchio prende in input un' immagine originale e ne restituisce in output la corrispondente immagine opportunamente marcata e il marchio vero e proprio;
 - Sia V l'immagine originale e sia $W = \{w_1, w_2, \dots, w_n\}$ il watermark da inserire (che potrebbe dipendere da V) e V_w l'immagine marchiata.
 - V_w si può ottenere da V e W tramite un'opportuna funzione di codifica E : $E(V, W) = V_w$
- **Decodifica:** la decodifica D prende in input l'immagine marchiata e l'immagine originale (se il watermark è non cieco), restituisce il marchio associato.
 - **watermark cieco:** vuole in ingresso un'immagine marchiata V_w e restituisce il watermark W' : $D(V_w) = W'$
 - **watermark non cieco:** vuole in ingresso l'immagine originale V e quella marchiata V_w e restituisce il watermark W' : $D(V, V_w) = W'$

I watermark W e W' non sono necessariamente identici, in quanto l'immagine può essere stata modificata tra la fase di codifica e di decodifica. E' quindi necessaria una funzione di confronto $C_\delta(W, W')$ ed una soglia S : il confronto fra $C_\delta(W, W')$ ed S permette di stabilire se i due watermark corrispondono:

IF $C_\delta(W, W') < S$ **THEN** W e W' corrispondono **ELSE** W e W' non corrispondono



walter.balzano@gmail.com

Tecniche di WaterMarking su Testi

- Metodi Spaziali
- Metodi Sintattici
- Metodi Semantici

Testo Normale

The	q	u	i	c	k	b	r	o	w	n	f	o	x		
j	u	m	p	s	o	v	e	r	t	h	e	l	a	z	y
d	o	g	.												

*Testo codificato
con aggiunta
di spazi*

The	q	u	i	c	k	b	r	o	w	n	f	o	x		
j	u	m	p	s	o	v	e	r	t	h	e	l	a	z	y
d	o	g													

*Esempio di
altra
codifica*

This distressed the monks and terrified them. They were not used to hearing these awful beings called names, and they did not know what might be the consequence. There was a dead silence now; superstitious bodings were in every mind. The magician began to pull his wits together, and when he presently smiled an easy, nonchalant smile, it spread a mighty relief around; for it indicated that his mood was not destructive.



Tecniche di WaterMarking su Immagini

- Least Significant Bit
- Tecniche Correlation-Based
- Tecniche Frequency Domain
- Tecniche di Wavelet Watermarking
- Tecniche Spread Spectrum

Least Significant Bit



Senza WaterMark

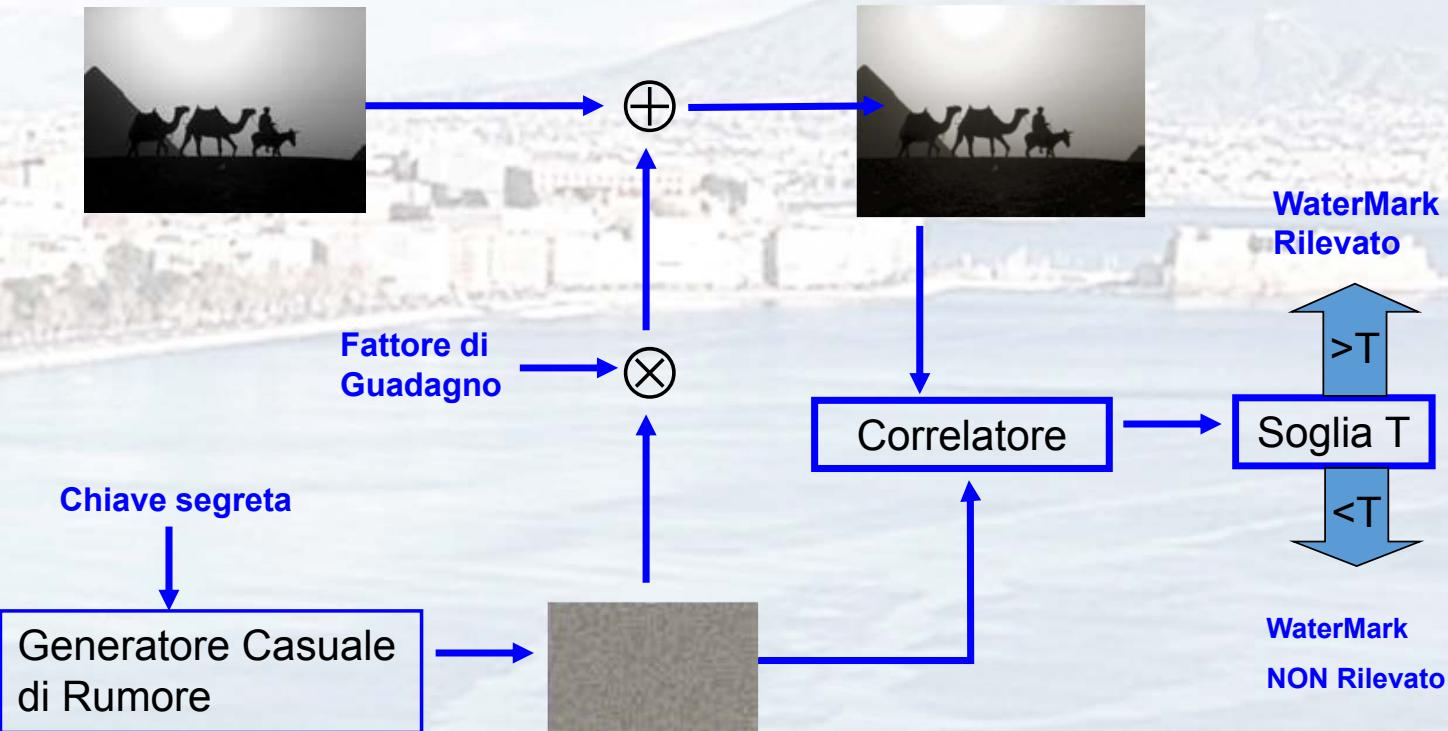


Con WaterMark



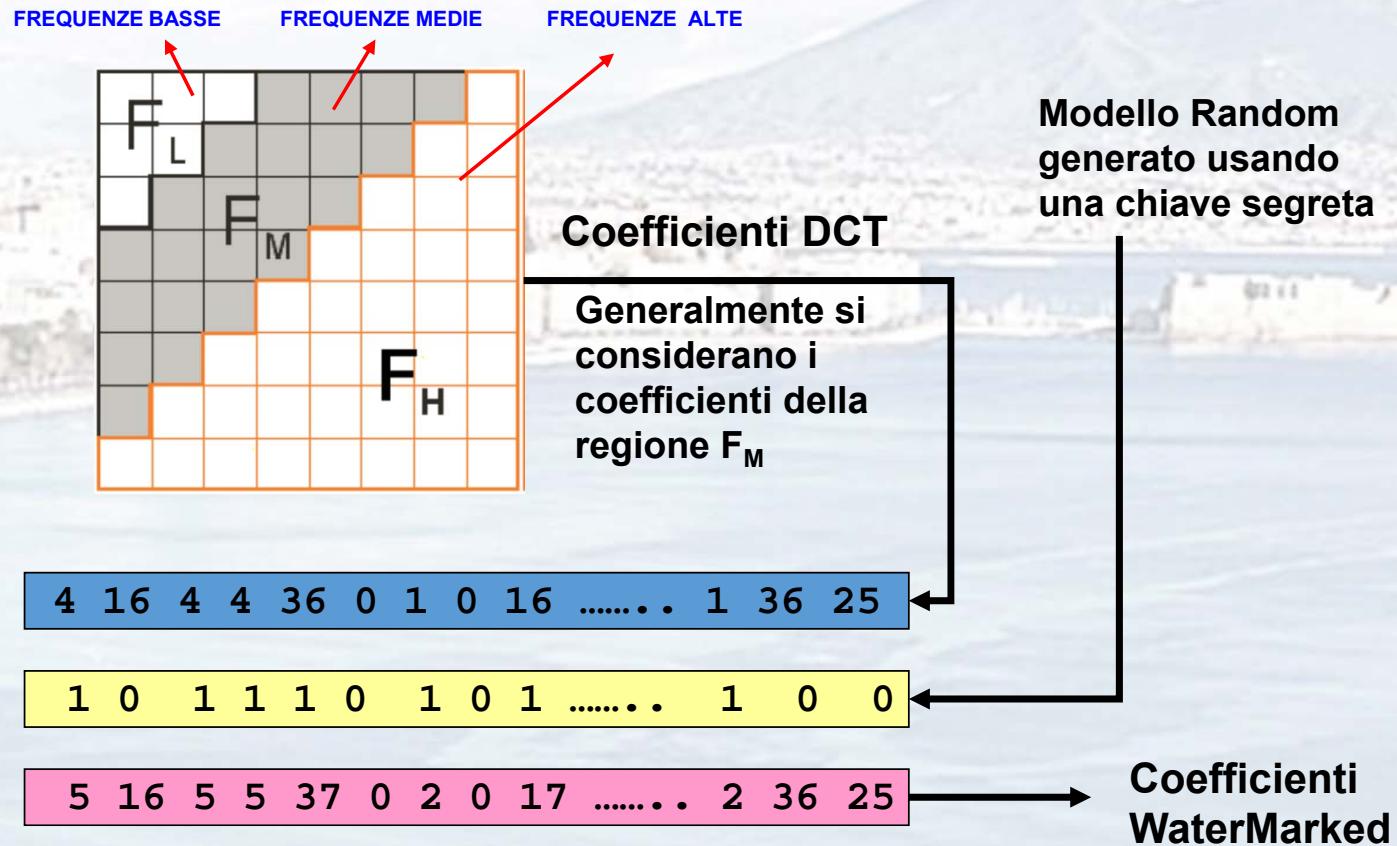
walter.balzano@gmail.com

Tecniche Correlation-Based



walter.balzano@gmail.com

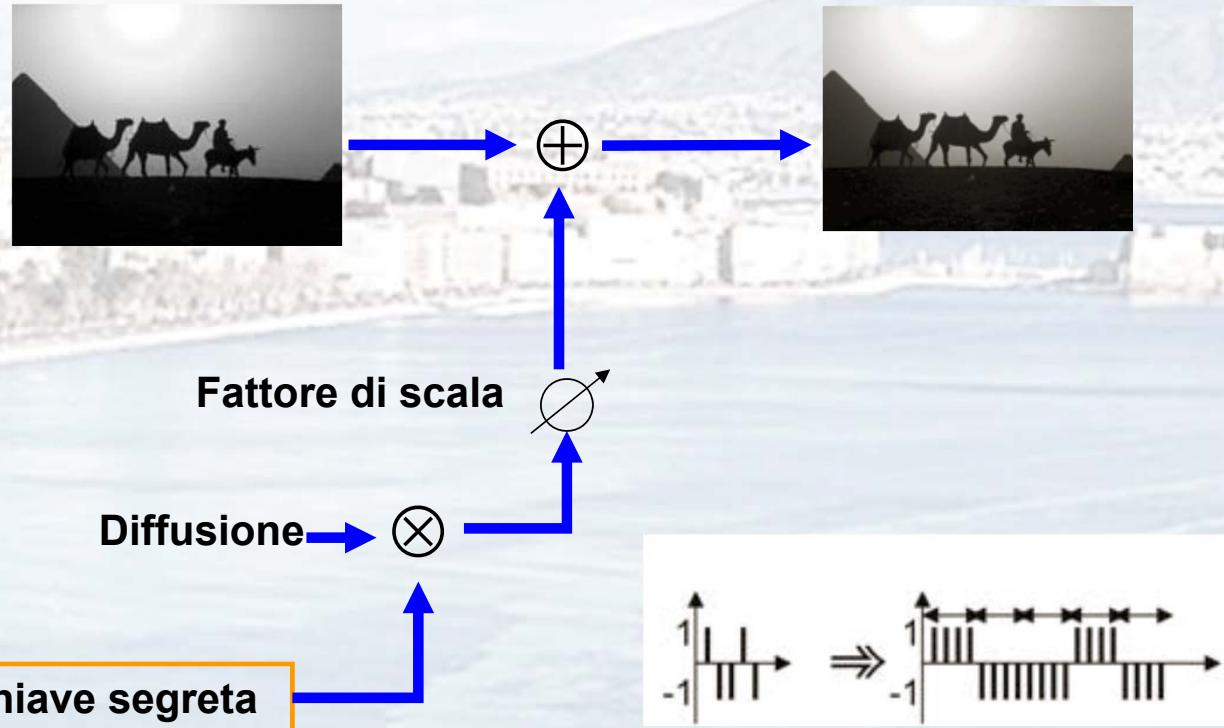
Tecniche basate su Frequenza



Modello Random generato usando una chiave segreta



Tecniche basate su spettro



walter.balzano@gmail.com

Video WaterMarking

Scopi

- Copyright protection
 - Proprietà dei contenuti
 - Dimostrazione di acquisto
 - Accesso condizionato
- Annotazioni Nascoste
 - Etichettatura
 - Indicizzazione
 - Livelli di controllo

Tecniche

Raw Video

- Codifica su singolo Frame
- Codifica sul 1° frame di ogni Shot
- Codifica dell'R-Frame di ogni Shot
- Aggiunta di un segnale sotto la soglia percepibile

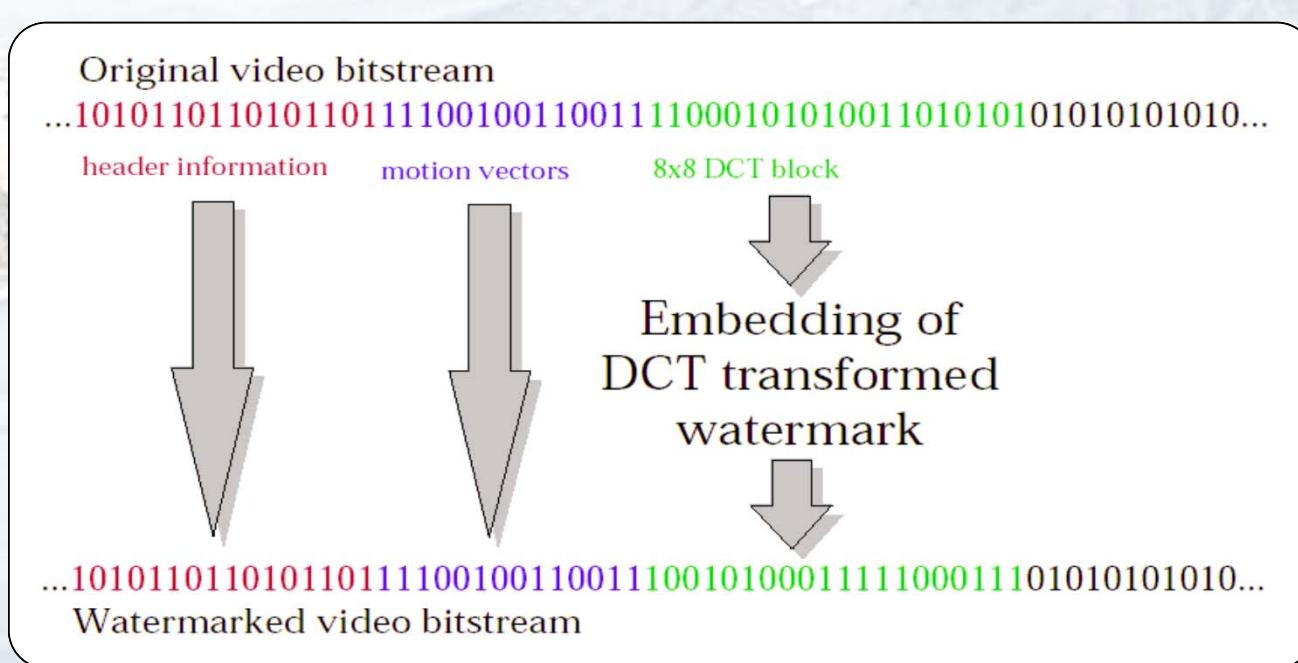
Video Compresso

- Modifica dei Coefficienti della DCT



walter.balzano@gmail.com

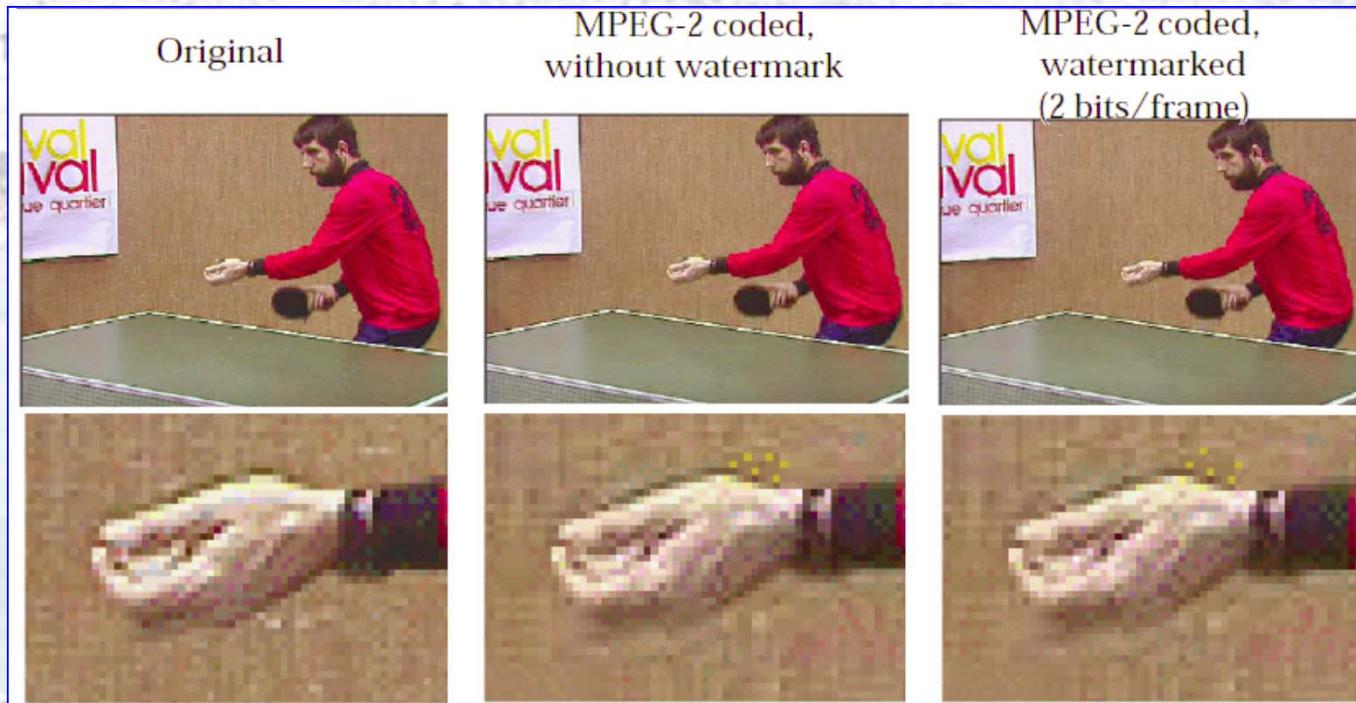
WaterMarking di video Compresso (1)



walter.balzano@gmail.com

WaterMarking di video Compresso (2)

Esempio



walter.balzano@gmail.com

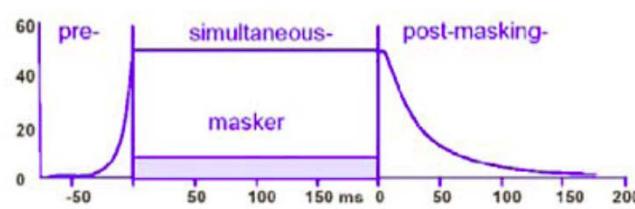
Audio WaterMarking

Utilizzo

- Creazione di tracce audio, Manipolazione e Versioning
- Fornire differenti livelli di accesso
- “Ritagliare” un segnale audio secondo le necessità dell’utente

Tipologie

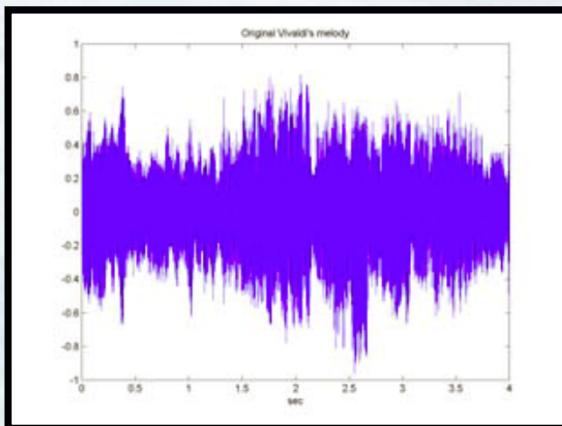
- **Mascheramento delle frequenze**
 - Mascheramento tra frequenze componenti
 - Se 2 segnali hanno frequenze molto simili, il più forte “maschera” il più debole
- **Mascheramento Temporale**
 - Pre-masking: un segnale debole (seguito da uno forte) viene “spento”, dovrà apponendogli quello più forte che viene anticipato di 5-20ms)
 - Post-masking: 50-200ms dopo il segnale forte viene disattivato il masker



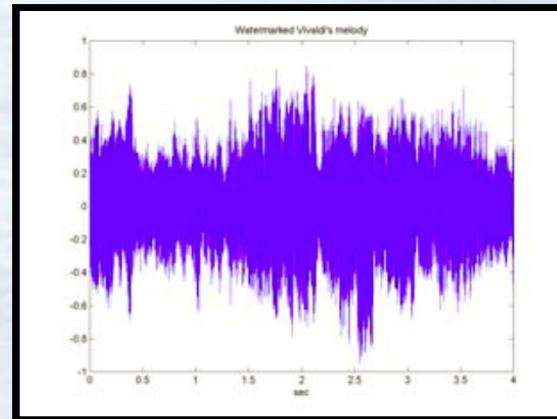
walter.balzano@gmail.com

Tecniche di Audio WaterMarking

- Low-Bit Coding
- Phase Coding
- Direct Sequence Spread Spectrum (DSSS)
- Frequency Hopped Spread Spectrum (FHSS)
- Echo Coding



Originale

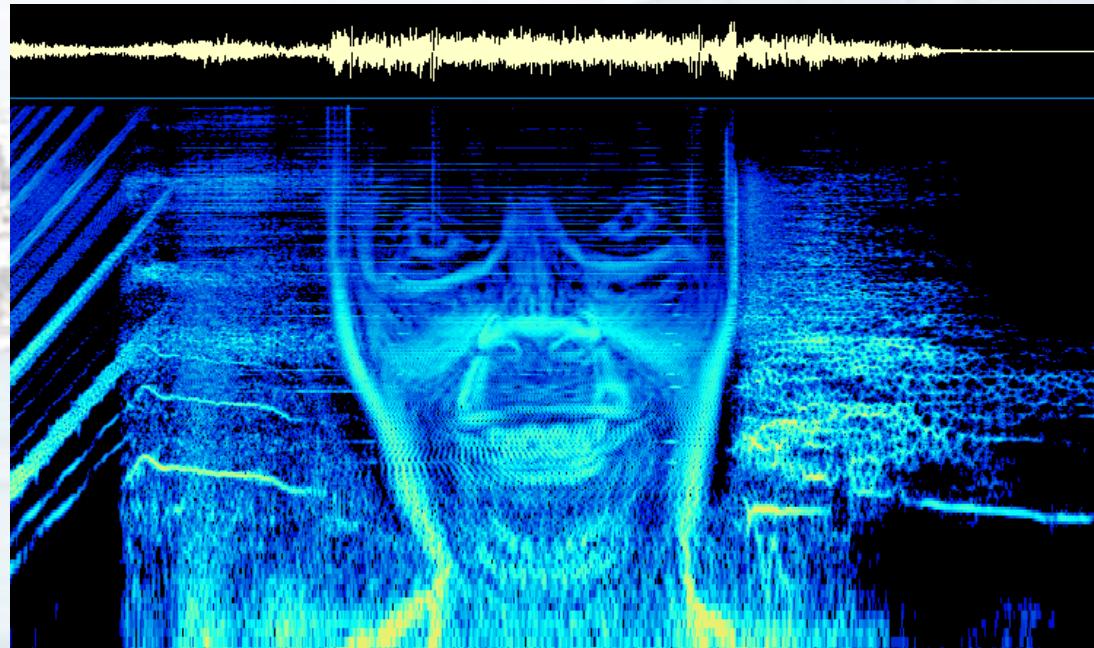


WaterMarked



walter.balzano@gmail.com

Tecniche di Audio WaterMarking



walter.balzano@gmail.com

Introduzione ai Gis

- **Introduzione**
- **Struttura Generale**
- **Caratteristiche**
- **Il modello spaziale
del mondo reale**
- **Oggetti del GIS**
- **Georeferenziazione**
- **Modelli della Terra**
- **Datum Geodetici**
- **Coordinate**
- **Proiezioni**
- **Mappe**



walter.balzano@gmail.com

Definizioni

- **GIS: Geographical Information System**

Un Gis può essere definito come un **sistema informatico per la gestione di dati geografici**

- **Geographical:**

si basano su rappresentazioni del territorio realizzate attraverso la memorizzazione della localizzazione geografica degli elementi in esso contenuti

- **Information:**

Indica che nei GIS i dati portano informazione in vari forme (**mappe, immagini, tavole, relazioni, interrogazioni, statistiche, incroci, ecc...**)

- **System:**

Un GIS è un “sistema” fatto di più componenti: hardware, software, dati geografici, risorse umane, ecc...

- **Obiettivo ultimo di un GIS è di fornire supporto decisionale basandosi su dati spaziali**



Struttura di un GIS

un GIS può essere concepito in modo gerarchico
identificando 6 **componenti** distinte
corrispondenti ad altrettante **attività** :



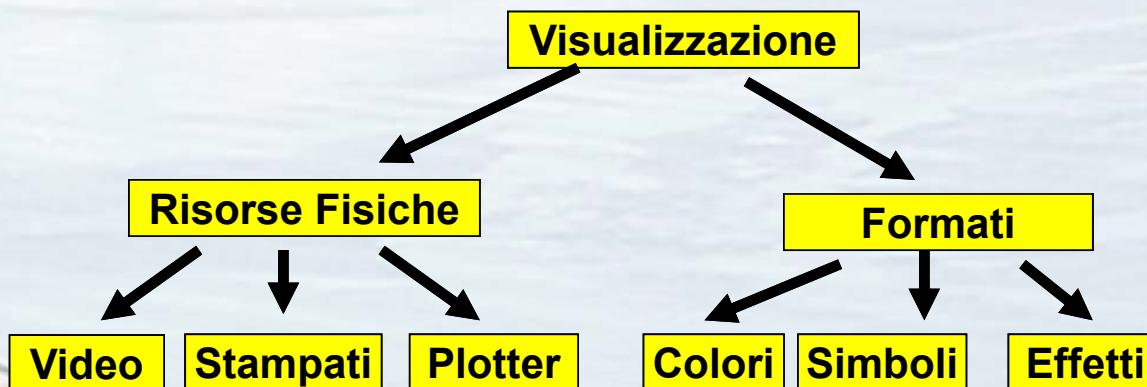
1: Organizzazione

- Un Gis deve collezionare grandi quantità di dati; oltre che dalla **quantità** un GIS è caratterizzato anche dalla **eterogeneità** delle sorgenti informative da cui provengono i dati
- Affinché il “Collage Informativo” di un GIS aggiunga di fatto una ulteriore dimensione informativa al materiale raccolto occorre che lo stesso Gis realizzi una Consistente **Organizzazione dei dati**
- Data models: modalità di organizzazione delle informazioni
- Le tipologie di informazioni contenute in un GIS possono essere, principalmente, di 2 tipi fortemente legati tra loro:
 - SPAZIALE (geografico)
 - ALFANUMERICO (attributi)



2: Visualizzazione

- Un GIS sfrutta le capacità umane di comprendere situazioni complesse in maniera migliore attraverso un approccio **VISUALE**
- Per tale motivo una delle prime componenti di un GIS è la capacità di **presentare graficamente** ed in maniera adeguata le informazioni raccolte e in esso organizzate (creazione di **MAPPE**)



3: Interrogazione

- In un Gis la realtà è memorizzata attraverso oggetti grafici elementari (**FEATURES**) che costituiscono una banca dati
- Ogni feature ha indissolubilmente collegate delle informazioni alfanumeriche associate (Esempi: edificio X, via, n. civico, ecc...)
- Il Gis deve permettere di richiedere informazioni sui dati da esso gestiti. Tali richieste posso essere di 3 tipi:
 - **Interrogazione degli attributi:** conoscere le caratteristiche di una feature memorizzata nella banca dati
 - **Interrogazione spaziale:** richieste complesse che intendono ottenere informazioni sui punti geografici dove si verificano una serie di condizioni specificate dall'utente;
 - **Interrogazione topologica:** domande su distanza, adiacenza, orientamento, contenimento,...



walter.balzano@gmail.com

4: Combinazione

- Un Gis deve permettere di **combinare** tra loro **informazioni** provenienti da fonti diverse
- Ciò permette di scoprire **correlazioni** o di generare dati derivati
- Si possono avere **2 modalità di combinazione**:
 - Visualizzazione congiunta di tematismi differenti
 - Costruzione “fisica” di temi ottenuti combinandone altri.

Esempi:

- **Visualizzazione** di una carta geologica sovrapposta ad una immagine da satellite (può permettere di scoprire relazioni tra la litologia ed il suo “effetto” sull’immagine)
- **Intersezione** della carta dei vincoli paesaggistici con la carta di uso del suolo (può permettere di derivare alcune informazioni su “abusì edilizi”)



walter.balzano@gmail.com

5: Analisi

- Si riferisce alla possibilità di utilizzare un Gis per ottenere un **SIGNIFICATO** a partire dai dati grezzi (raw)
- Oltre a poter essere effettuata in maniera visuale, in genere il Gis mette a disposizione strumenti **STATISTICI** e di **MODELLAZIONE**

Esempi:

- A partire da valori di una variabile rilevati in una serie di punti (x,y) ottenere una superficie tridimensionale attraverso metodi di interpolazione.
- Data una carta di uso del suolo, ottenere informazioni sull'area complessiva occupata da ognuna delle classi
- **Ottenerne statistiche sulla correlazione tra la litologia (composizione chimica e mineralogica) e la concentrazione al suolo di un inquinante**



walter.balzano@gmail.com

6: Predizione

- E' l'aspetto più avanzato di un Gis, quello in cui la componente umana del progettista è determinante
- E' strettamente legato al punto precedente ma si spinge anche a fare previsioni di scenari evolutivi futuri. Lo scopo è quello di rispondere a domande quali:
 - Che cosa succede se prendo questa decisione? (analisi "What-If").
 - Esempio: pianificazione del traffico.
- E' una componente fondamentale in campi quali:
 - Pianificazione territoriale e urbanistica
 - Analisi ambientale
 - valutazione di impatto
 - ecc...
- E' un aspetto complesso e legato ad altre discipline:
 - Simulazione
 - multimedialità,
 - ecc...



walter.balzano@gmail.com

Caratteristiche

I settori più “vicini” al Gis sono il **CAD** e **l’Image processing**: oggi tutti i software Gis di buon livello contengono al loro interno funzionalità CAD e di elaborazione delle immagini

● **CAD:**

- Strumenti computerizzati di **supporto al disegno tecnico**. Sono nati in ambiente ingegneristico e architettonurale prima in bidimensionale, ed in seguito si sono estesi alla gestione della terza dimensione.
- Strumenti di tipo **VETTORIALE** (basati su vettori). Ciò significa che ogni oggetto del disegno viene memorizzato e gestito attraverso le sue coordinate.
- Strumenti con ottime funzionalità di visualizzazione
- Poco adatti alla “gestione”: non possiedono il concetto di **ATTRIBUTO** inteso come informazione alfanumerica collegata agli oggetti grafici

● **Image Processing:**

- Strumenti di visualizzazione, gestione ed elaborazione di immagini **RASTER** (matrice o grigliato di pixel).
- **Esempi:** Immagini da satellite, Fotografie aeree,...



walter.balzano@gmail.com

Esempio di Cartografia Tradizionale

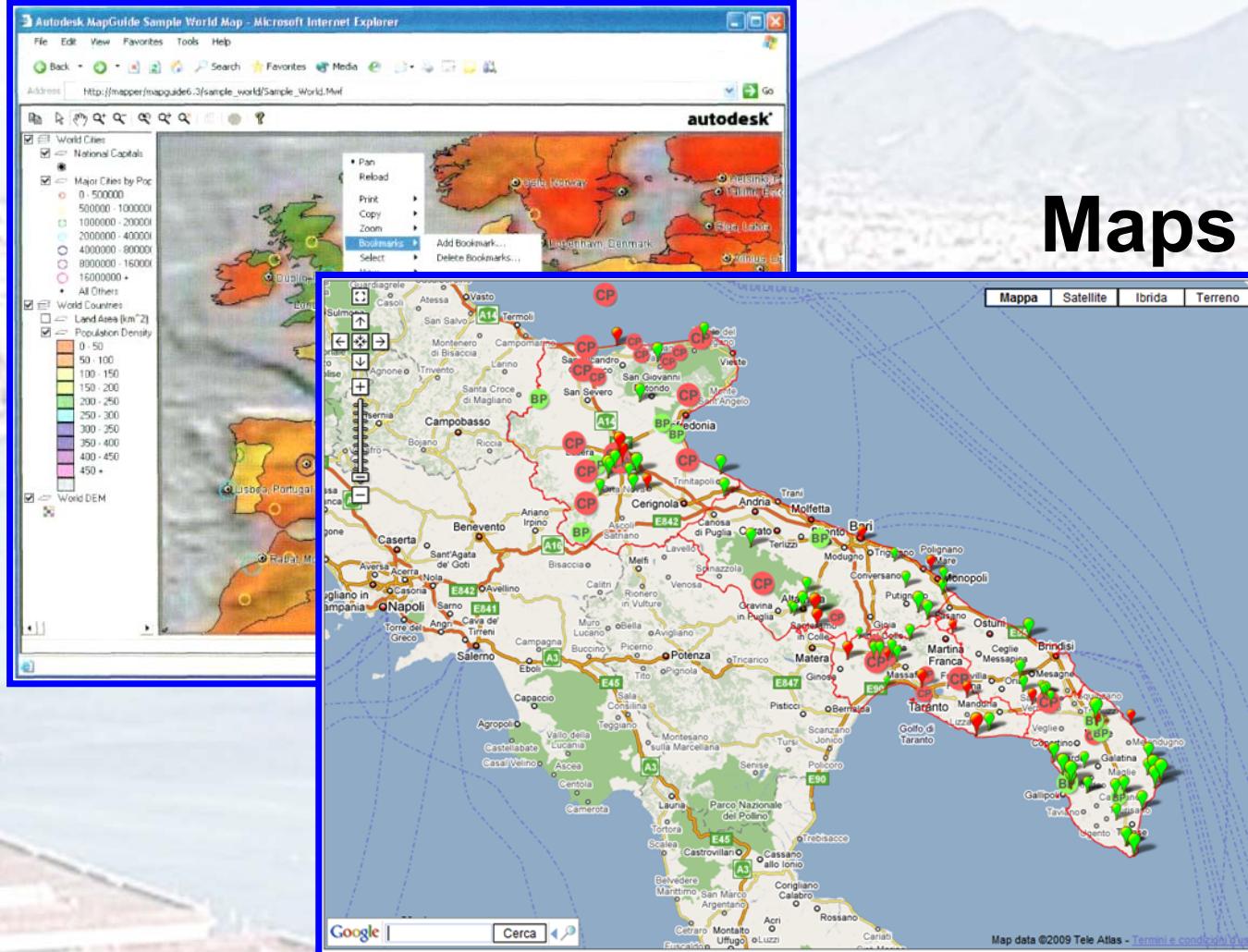


walter.balzano@gmail.com

MapGuide

Esempio di Cartografia Digitale

Maps



walter.balzano@gmail.com

Modello spaziale del mondo reale

- La peculiarità dei GIS consiste nell'introdurre la dimensione dello **SPAZIO** in un database tradizionale
- Un database tradizionale può contenere informazioni territoriali e informazioni relative allo spazio:
 - Localizzazione di persone
 - Localizzazione di eventi e attività
 - Dati sulle distanze
- Un database tradizionale non è in grado di gestire implicitamente né di derivare informazioni originali di tipo spaziale
- Per fare ciò è necessario introdurre un **modello spaziale del mondo reale** che gestisca la memorizzazione delle informazioni nel DB



walter.balzano@gmail.com

Modello spaziale del mondo reale

- La **modellazione** del mondo reale e **l'organizzazione** delle informazioni sotto forma di dati che lo descrivono presuppongono un processo diviso in diversi stadi:
 - Identificazione delle **entità** che definiscono i fenomeni di nostro interesse
 - 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**



walter.balzano@gmail.com

L' oggetto GIS

- Un oggetto GIS ha diverse caratteristiche che lo contraddistinguono:
 - TIPO
 - ATTRIBUTI
 - RELAZIONI SPAZIALI
 - GEOMETRIA
 - QUALITA' DEL DATO
- Gli elementi all'interno di un GIS sono costituiti da una **collezione di oggetti** rappresentati con figure geometriche
- Tali oggetti rappresentano in **forma semplificata** fenomeni del mondo reale conservandone le proprietà spaziali di volta in volta giudicate rilevanti



walter.balzano@gmail.com

● PUNTO:

■ Punto isolato

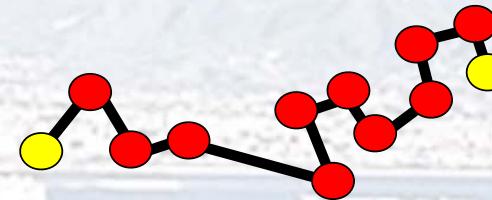
■ Se fa parte di una linea è detto
VERTICE o **NODO**



● ARCO:

■ Sequenza ordinata di VERTICI.

■ Il primo è detto **NODO INIZIALE**,
l'ultimo **NODO FINALE**



● ANELLO:

■ Insieme di uno o più archi che costituiscono un anello chiuso

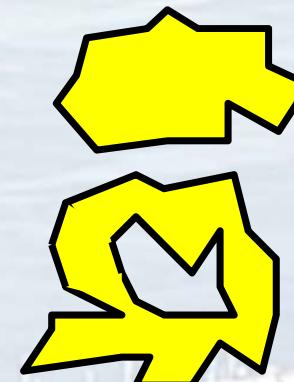


● POLIGONO:

■ Insieme di uno o più anelli che delimitano una area chiusa.

■ Il primo anello è detto **ESTERNO**, gli altri **INTERNI**.

■ Se ha 1 solo anello è detto **SEMPLICE** altrimenti **COMPLESSO**



L'oggetto punto

- Entità **ADIMENSIONALE** che specifica la localizzazione di un fenomeno nello spazio (codifica la posizione)

Diversi tipi di punti:

- **Punto entità**: identifica la localizzazione di un fenomeno nello spazio senza tenere conto dell'eventuale forma o dimensione del fenomeno
- **Punto area**: denota una posizione all'interno di un'area e ne rappresenta una forma semplificata
- (*Esempio*: baricentro, centro urbano per i comuni)
- **Nodo**: localizzazione puntuale con più proprietà topologiche (collegamento ad un altro punto tramite segmento)



walter.balzano@gmail.com

L'oggetto arco o linea

- L'arco è inteso come una entità **UNIDIMENSIONALE** che ha la capacità di rappresentare un fenomeno reale evidenziandone:
 - Posizione
 - Direzione
 - Lunghezza
- Può essere rappresentata con diversi spessori che identificano **attributi metrici** diversi dalla lunghezza (rimane un oggetto unidimensionale)
- Può avere un attributo che ne specifica l'orientamento

Esempi:

- Strade, fiumi, confini, flussi di comunicazione, ecc.



walter.balzano@gmail.com

L'oggetto poligono o area

- Il poligono è inteso come una entità **BIDIMENSIONALE** che ha la capacità di rappresentare un fenomeno reale evidenziandone:
 - Posizione
 - Morfologia
 - Superficie (dimensione quadrata)
- Viene usato tipicamente per rappresentare strutture territoriali come:
 - Comuni
 - Quartieri
 - Regioni
 - Tutte le volte che superficie e topologia sono importanti nell'analisi



Rappresentazione spaziale

- La scelta degli oggetti nella rappresentazione spaziale dipende dall'**analisi** che dovrà essere fatta del fenomeno:
- Dalle **proprietà** dei singoli oggetti dipende la possibilità di effettuare determinate analisi piuttosto che altre
- *Esempio:*
 - Se ci interessa la valutazione della distanza tra le città potremmo rappresentare queste con dei punti ma se volessimo fare un'analisi di distribuzione demografica dovremmo usare dei poligoni magari con sottopoligoni per la rappresentazione dei quartieri



walter.balzano@gmail.com

Georeferenziazione

- La rappresentazione degli **oggetti nello spazio** richiede un sistema di riferimento spaziale
- Gli oggetti devono mantenere:
 - **Posizioni relative corrette**
 - **Dimensioni proporzionali agli oggetti reali corrispondenti**
 - **Relazioni spaziali corrette**
- **GEOREFERENZIAZIONE:**
 - Tecnica di localizzazione territoriale che permette di associare un oggetto ad un particolare punto nello spazio reale



walter.balzano@gmail.com

Georeferenziazione (2)

- Tecniche di georeferenziazione **CONTINUE**:

- La misura della posizione di un fenomeno è ottenuta rispetto ad un sistema di **riferimento assoluto**
- Normalmente si usano i sistemi di **coordinate terrestri** cartografiche
- La posizione reale di un oggetto viene rilevata attraverso **sistemi di posizionamento globale (GPS)**

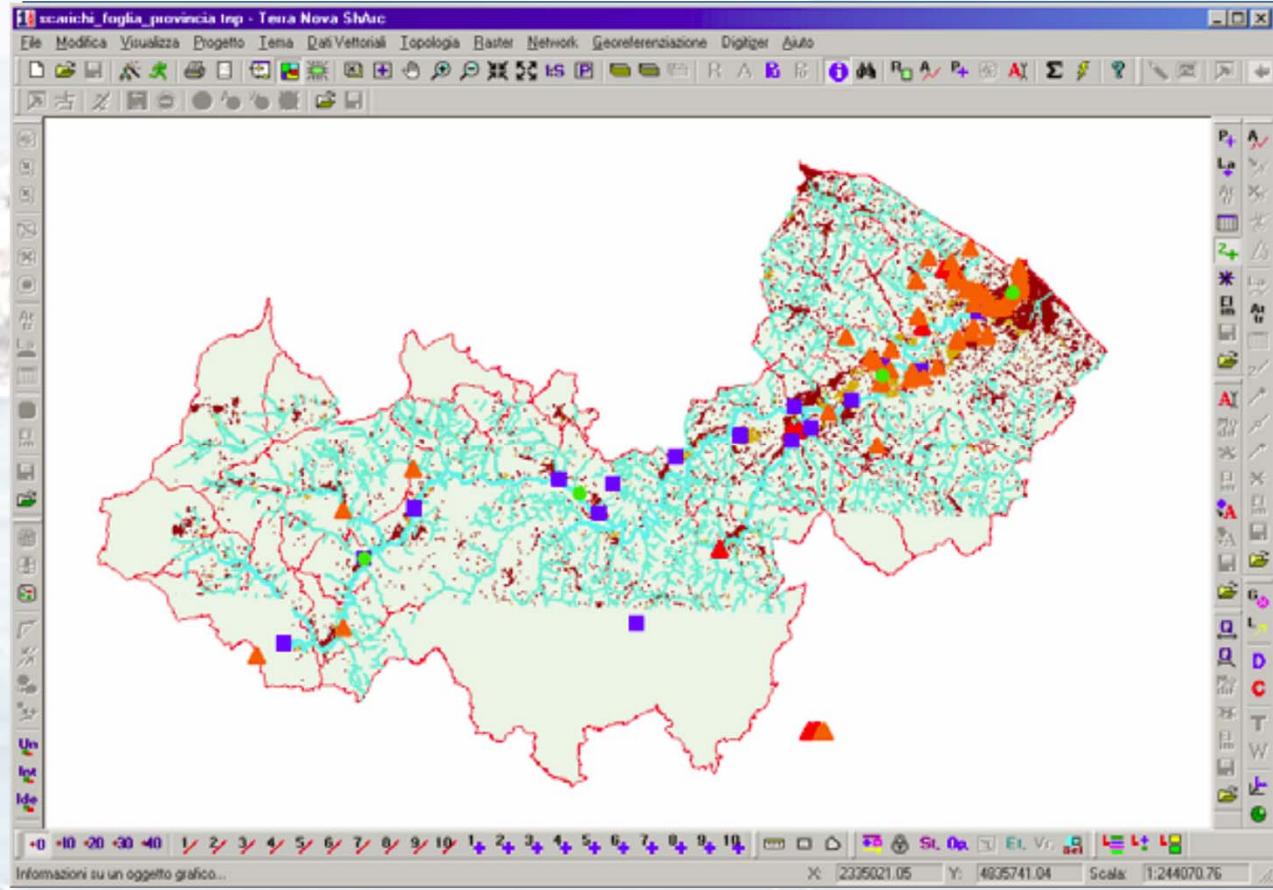
- Tecniche di georeferenziazione **DISCRETE**:

- La misura della posizione di un fenomeno è ottenuta **indirettamente** rispetto ad unità territoriali di riferimento **già georeferenziate**
- La posizione di un oggetto viene rilevata valutando la distanza da un **oggetto territoriale di riferimento**



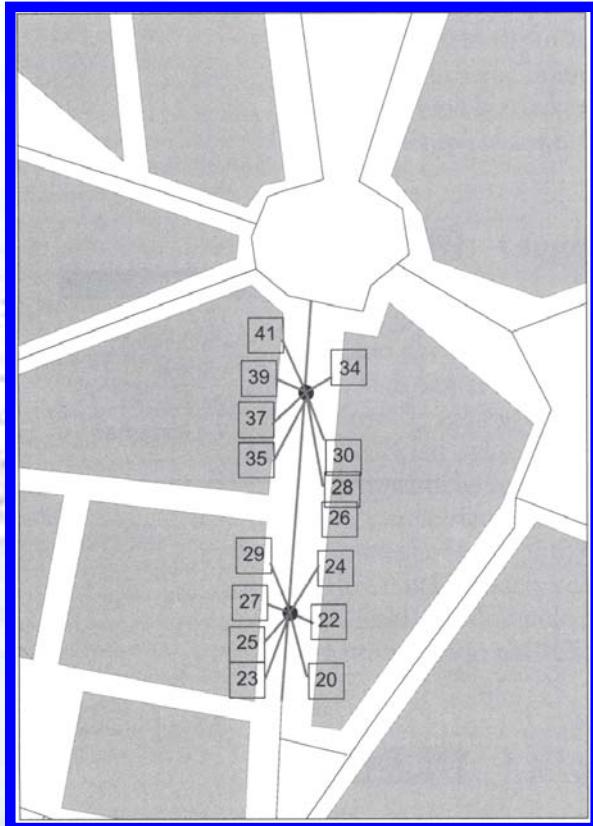
walter.balzano@gmail.com

Georeferenziazione



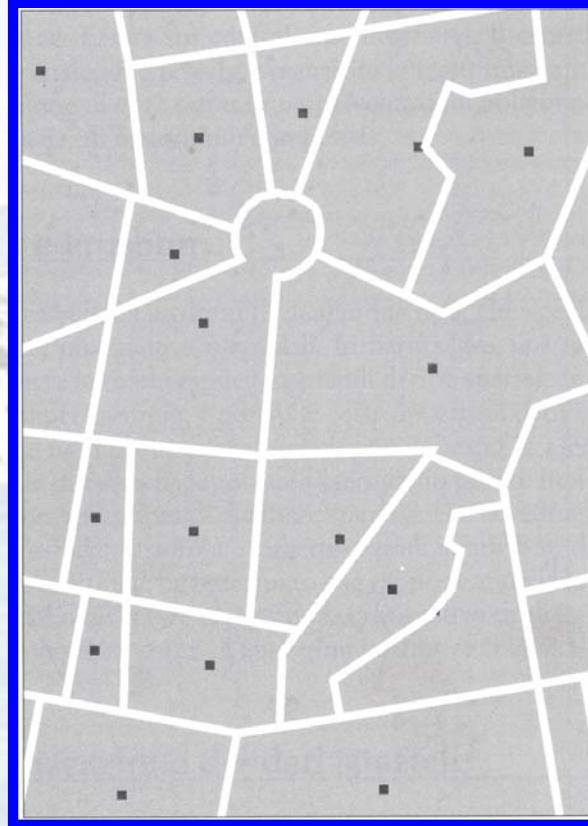
walter.balzano@gmail.com

Georeferenziazione



GeoReferenziazione discreta toponomastica.

I fenomeni sono localizzati approssimativamente rispetto ai numeri civici degli edifici



GeoReferenziazione discreta areale

I fenomeni sono localizzati in modo approssimato nel **centroide** delle zone statistiche



walter.balzano@gmail.com

Modelli della Terra

Modello della “Terra piatta”

- La superficie della Terra viene rappresentata attraverso un piano orientato a nord con l'asse ortogonale coincidente con la direzione della gravità
- Gli oggetti vengono rappresentati sul piano e posizionati relativamente alla piano stesso
- **Va bene per mappe di dimensioni non superiori a 2-3 Km e pianeggianti**

Modello della “Terra curva”

- Tiene conto della curvatura della superficie terrestre
- Modello complesso e non univoco
- Cerca di limitare gli errori nelle relazioni spaziali degli oggetti rappresentati sulla Terra

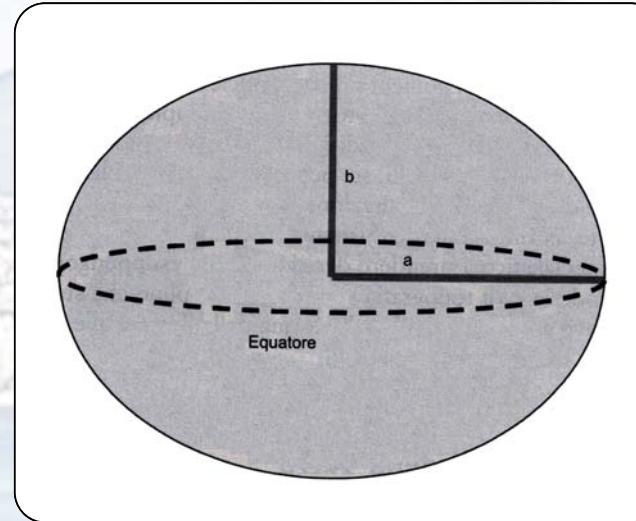


walter.balzano@gmail.com

Ellissoide o Sferoide (Terra curva)

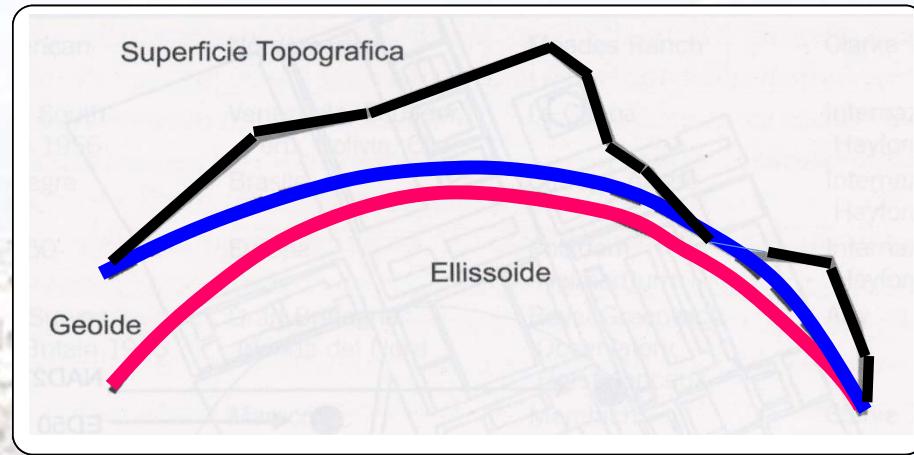
Modello matematico della Terra

- Generato dalla rotazione di un ellisse attorno al suo asse minore
- Riproduce lo schiacciamento ai poli e il rigonfiamento all'equatore
- Non tiene conto della distribuzione non uniforme delle masse nella Terra e delle irregolarità della superficie
- Utilizzato per la valutazione della posizione geografica di un punto sulla superficie



walter.balzano@gmail.com

Geoide (Terra curva)



Modello fisico della Terra

- Generato dalla **superficie perpendicolare** alla direzione della **forza di gravità** in ogni punto della Terra
(simile alla superficie del mare)
- Tiene conto delle **irregolarità locali** della superficie terrestre
- Viene utilizzato per la valutazione **dell'altitudine** di un punto



walter.balzano@gmail.com

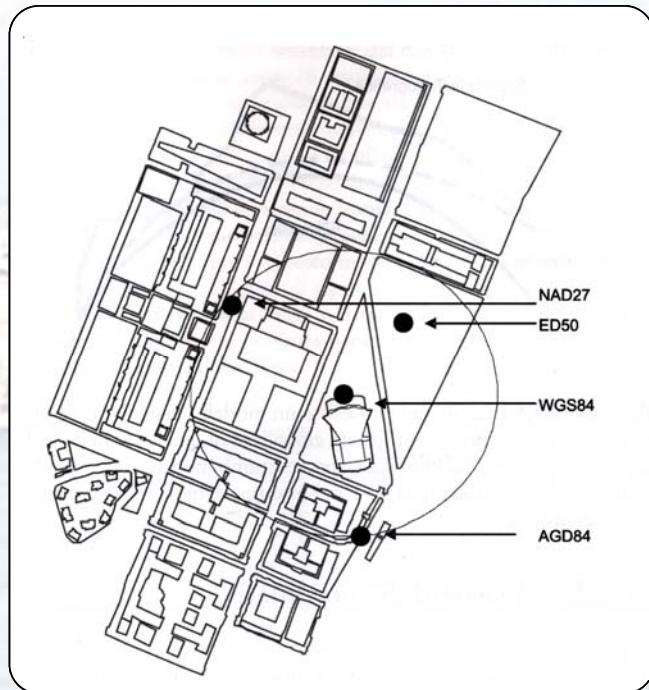
DATUM GEODETICI

- I sistemi di coordinate attualmente utilizzati si basano sul **DATUM GEODETICO** (modello matematico-fisico) che adatta localmente la superficie dell'ellissoide a quella del **geoide**
- Diversi tipi di **DATUM**:
 - Ogni paese ha definito un proprio **DATUM** al fine di approssimare al meglio la superficie terrestre all'interno della **propria area** di interesse (Es: *NAD27 North American Datum 1927*, *ED50 European Datum 1950*)
 - La diffusione dei sistemi GPS ha portato alla **definizione di DATUM GLOBALI** che non si concentrano su nessuna area in particolare ma cercano di minimizzare gli errori in modo globale (Es: *WGS84 World Geodetic System 1984*) standard dei GPS



walter.balzano@gmail.com

Posizionamento con DATUM diversi



- La rilevazione con diversi DATUM produce localizzazioni non coincidenti
- Nell'esempio la localizzazione dello stesso punto porta ad un errore massimo di 400 metri



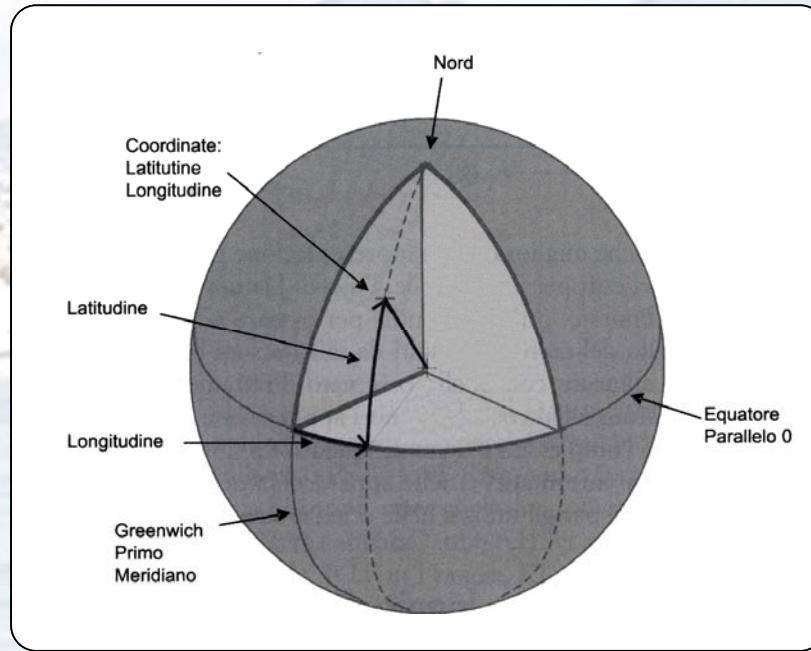
walter.balzano@gmail.com

Latitudine e Longitudine

- Sul modello della Terra (DATUM) viene generata una griglia immaginaria al fine di effettuare misurazioni con un sistema di coordinate terrestri

- La griglia è costituita da due insiemi di linee immaginarie che si intersecano:

- Paralleli o linee di latitudine
- Meridiani o linee di longitudine



- L'origine del sistema di coordinate viene collocato all'interno del modello della Terra e i punti sulla superficie vengono calcolati con gli strumenti della geometria sferica



walter.balzano@gmail.com

Parallel e Meridiani

Paralleli

- Cerchi paralleli all'equatore dove l'equatore è definito come il cerchio massimo che divide la terra nei due emisferi NORD e SUD
- All'equatore viene assegnato il valore convenzionale di **PARALLELO ZERO**
- La distanza tra i parallelviene misurata in **gradi sessualiimali** corrispondenti a circa 111 Km

- I gradi di latitudine variano tra 0 e 90 a nord e a sud dell'equatore

Meridiani

- Cerchi tracciati in **verticale** attraverso il NORD ed il SUD geografici
- La **longitudine** misura lo **scostamento** rispetto ad un cerchio di meridiano assunto come origine
- Per convenzione venne definito come **MERIDIANO ZERO** il meridiano che passa per l'osservatorio di **Greenwich**
- I meridiani si estendono a est di gradi con segno positivo e a ovest con segno negativo per 180 gradi
- I meridiani si incontrano nell'oceano pacifico (180 gradi) formando la *Linea della data internazionale*



walter.balzano@gmail.com

Posizionamento

- Ad ogni punto sulla superficie Terrestre viene assegnata una posizione espressa in GRADI, MINUTI e SECONDI di latitudine e longitudine
 - *Esempio:* Napoli = **40° 51' 46,80" NORD**
14° 16' 36,12" EST
- La natura sferica delle coordinate porta a far sì che la linea più breve che collega due punti sul modello curvo della Terra non è una linea ma è una curva
 - Tale curva può essere considerata come una **sezione di un circolo massimo** che passi per il centro della Terra
 - Tra due punti sulla superficie della Terra si può tracciare **un solo circolo massimo**
 - La distanza minore fra due punti sarà l'arco di circolo massimo che passa per i due punti



walter.balzano@gmail.com

Proiezioni

- Compito principale della **cartografia** è quello di rappresentare su carta la realtà tridimensionale:
per poter analizzare e rappresentare i dati territoriali bisogna trasferire le coordinate riferite alla superficie della Terra (modello Terra curva) su di un **foglio piano**: Questo processo è detto **PROIEZIONE** e porta alla costruzione di una **MAPP**A
- La costruzione di una mappa ovviamente comporta anche la riduzione delle dimensioni rispetto al mondo reale
- Il rapporto tra le dimensioni degli oggetti riportati sulla mappa e gli oggetti nel mondo reale è detto **SCALA** ed è rappresentato mediante una frazione numerica in cui il numeratore è uguale a 1
 - **Esempio:** mappe 1:10 000 → 1 cm sulla mappa equivale a 10 000 cm nel mondo reale
- Distorsioni delle proiezioni
- Non esistono proiezioni o trasformazioni matematiche di qualsiasi tipo che permettano di “sviluppare” fedelmente una superficie curva su di una superficie piana senza introdurre una o più distorsioni; in particolare la distorsione può ripercuotersi su:

- Distanze
- Direzioni
- Forme
- Superfici
- Scala



walter.balzano@gmail.com

Classificazione delle Proiezioni

E' possibile classificare le proiezioni in base a:

PROPRIETA' FISICHE (conservate senza errori):

Conformi o isogone

Conservano la forma degli oggetti su scala locale (piccole aree); i meridiani e i paralleli si intersecano con angoli retti formando una griglia regolare di riferimento. Invece per grandi aree, come quella di un continente, la forma sarà distorta. Le proiezioni Conica Conforme di Lambert e la Cilindrica Conforme di Mercatore sono le proiezioni Conformi più comuni

Ad aree equivalenti

Riproducono correttamente le superfici in modo corrispondente (proporzionale) alla realtà, a scapito della forma degli angoli e della scala

Equidistanti

Conservano le distanze fra determinati punti, anche se la scala non è riprodotta correttamente in tutti i punti della mappa (Nota: nessuna proiezione è "equidistante" per tutti i punti della mappa). E' ragionevole usarla se occorre valutare la distanza tra due punti non troppo distanti.

Afilattiche (non conservano nessuna proprietà)

METODI GEOMETRICI (usati per generare la proiezione stessa):

Cilindriche (proiezione superficie sferica su cilindro)

Tangenti, Secanti, Traverse, Oblique

Coniche (proiezione superficie sferica su cono)

Azimutali o piane (la superficie sferica viene proiettata direttamente su un piano)



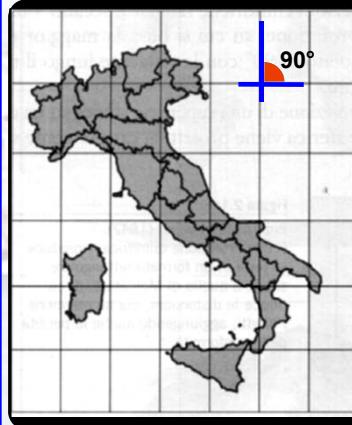
walter.balzano@gmail.com

Proiezioni: distorsioni

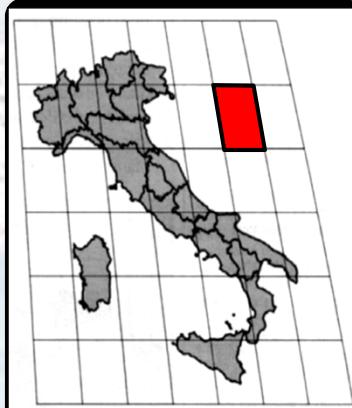
Modello sferico



Conformi o isogone



Ad aree equivalenti



Equidistanti



Area km^2

Lombardia	24.230
Lazio	17.310
Calabria	15.230

49.486
31.282
25.230

24.285
17.348
15.267

24.368
17.337
15.238

Distanza km

Milano-Roma 477
Torino-Venezia 367

660
520

510
360

477
368

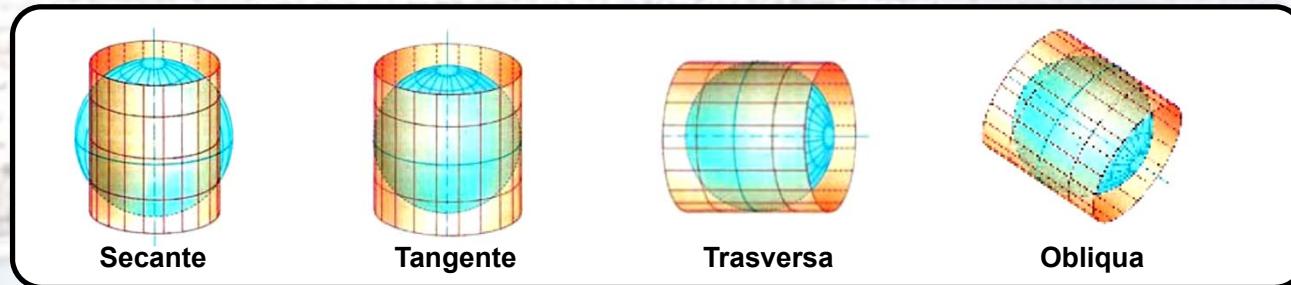


walter.balzano@gmail.com

Metodo delle proiezioni

Cilindriche

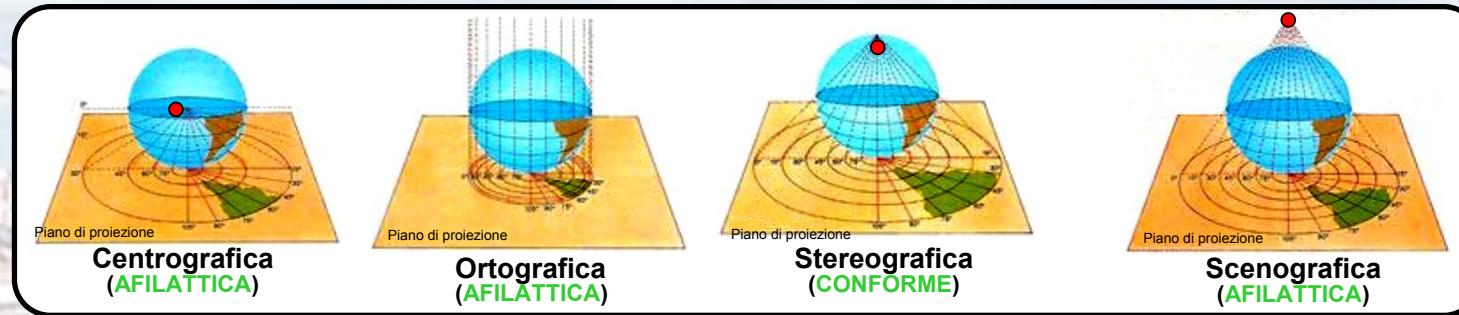
- Generate mediante proiezione di una superficie sferica su un cilindro
- A seconda della posizione relativa del cilindro rispetto alla sfera si hanno proiezioni cilindriche:



Azimutali

Generate mediante proiezione di una superficie sferica direttamente su di un piano.

A seconda della posizione del centro di proiezione si dividono in:



walter.balzano@gmail.com

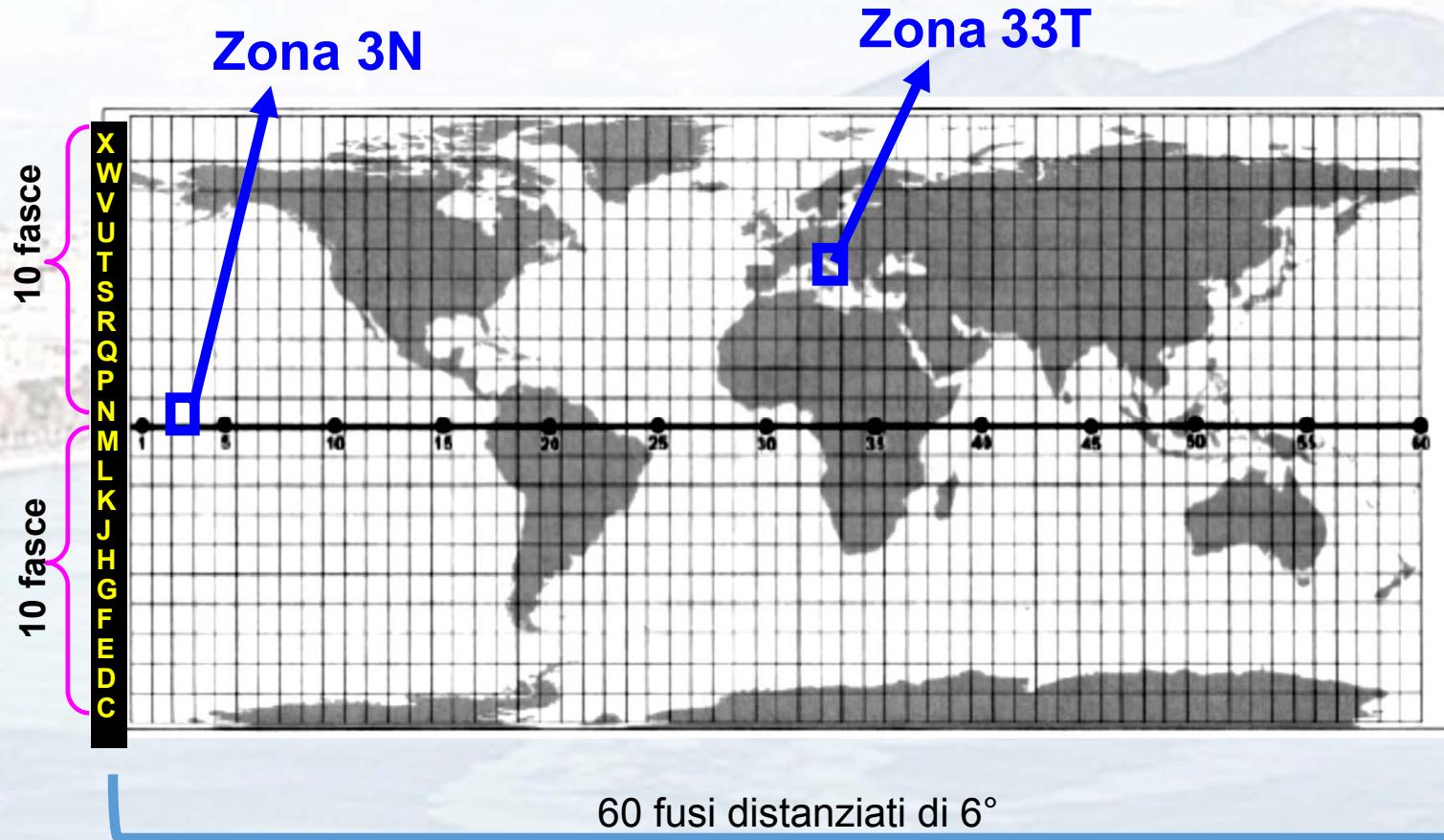
Proiezione UTM (Universale Trasversa di Mercatore)

- La proiezione UTM (Cassini-Gauss) è una modifica della proiezione di Mercatore (a sua volta definibile come cilindrica modificata); essa UTM rappresenta in modo corretto gran parte delle regioni abitate sulla Terra
- Crea un mosaico di proiezioni centrate sul meridiano centrale della zona di interesse di dimensione contenute
- Ogni elemento del mosaico minimizza gli errori lineari, angolari e superficiali
- La deformazione cresce verso i lati di ogni elemento del mosaico e ne rende difficile l'accostamento
- Il mosaico di proiezioni su cui si basa la cartografia UTM è costituito da strisce verticali che si allungano allontanandosi dall'equatore e strisce orizzontali che suddividono ogni emisfero in 10 sottozone
- In dettaglio, presi come riferimento il meridiano di Greenwich e dall'equatore, per costruire il reticolo si sceglie un meridiano ogni 6° di longitudine ed un parallelo ogni 8° di latitudine.
- Il reticolo non copre l'intera superficie terrestre ma lascia fuori i poli (ci si ferma a circa 80°). Questo porta ad ottenere un reticolo di 60x20 elementi, ciascuno "largo" 6° ed "alto" 8° .



walter.balzano@gmail.com

Proiezione UTM (Universale Trasversa di Mercatore)



Emisfero Nord

Emisfero Sud



walter.balzano@gmail.com

Coordinate UTM

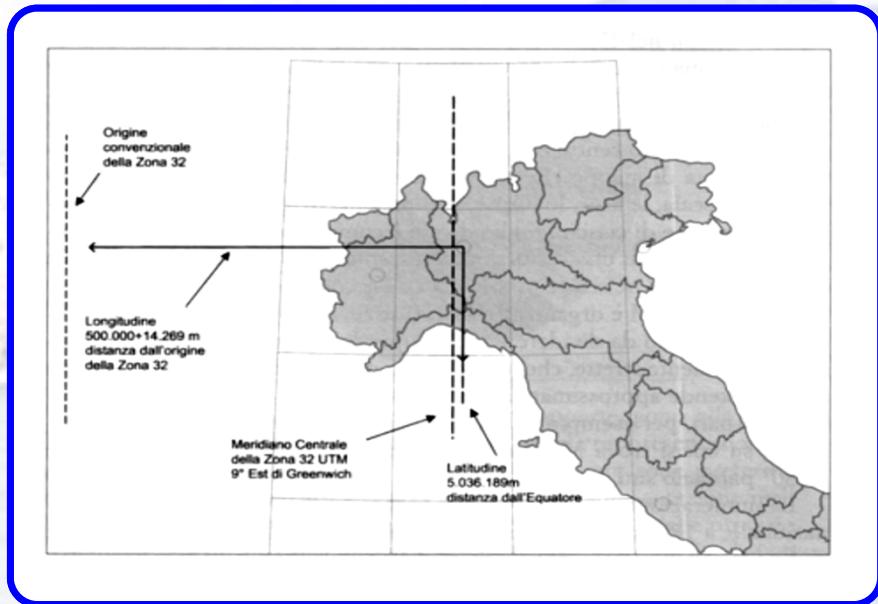
- All'interno di ogni zona, le coordinate (esprese in metri) sono descritte in forma cartesiana in cui la y è rappresentata sulla direzione Nord-Sud
- Nell'emisfero nord l'equatore rappresenta l'origine della coordinata y e la latitudine si misura come distanza positiva dall'equatore
- Nell'emisfero sud il polo Sud rappresenta l'origine delle y e la latitudine si misura come distanza positiva dal polo
 - In questo modo anche nell'emisfero Sud abbiamo latitudini positive
- Come origine dell'asse x si assume il meridiano centrale della zona stessa spostato convenzionalmente a ovest di 500 Km
 - In questo modo si lavora sempre con valori positivi delle x all'interno di ciascuna zona



walter.balzano@gmail.com

Coordinate UTM

L'Italia è coperta dalla zona 32 e 33 delle coordinate UTM



Esempio:

La cima del monte Galero ha

Coordinate geografiche: 44°09'22"N - 8°00'56"E

Coordinate UTM: 32T , 0421274 east , 4889679 north

mappa 32T; distanza 78726 metri dal meridiano centrale di mappa
(ovvero 500000 - 421274); distanza 4889679 metri dall'equatore.



walter.balzano@gmail.com

Le Mappe nei GIS

- In un GIS una mappa è il risultato della sovrapposizione di più mappe elementari dette **strati informativi**
- I vari strati informativi hanno la caratteristica di essere **georeferenziati** attribuendo informazioni diverse ai punti del territorio
 - Modello di mappa **VETTORIALE**
 - Modello di mappa **RASTER**
 - Modelli di mappe **TRIDIMENSIONALI**

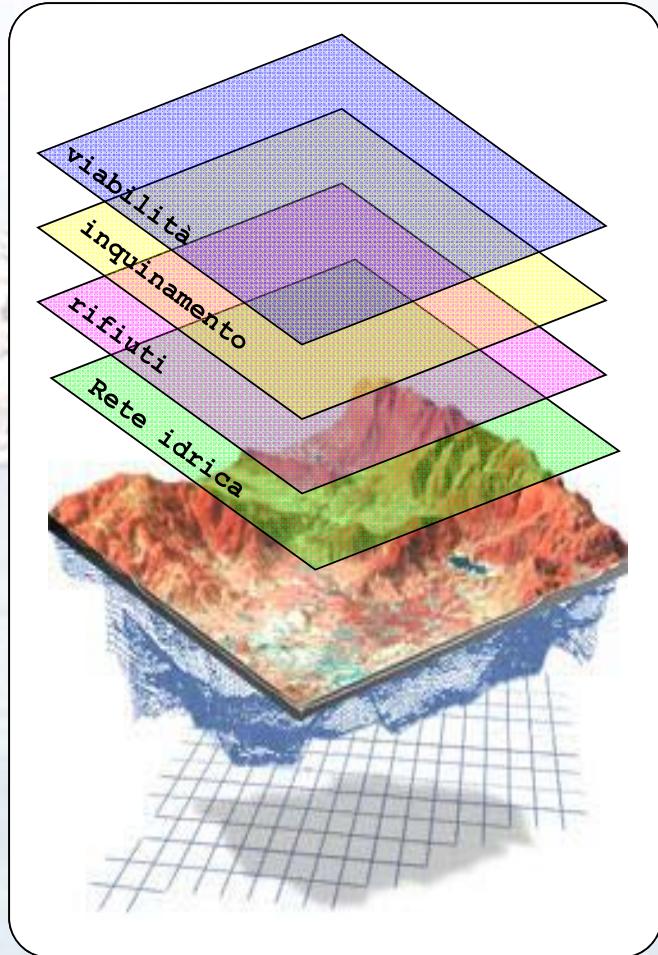
MAPPE RASTER

- E' adatto per dati che cambiano con continuità nello spazio
- E' rappresentato come una matrice rettangolare di numeri (un valore per ogni cella o pixel)
- Grande facilità di gestione e di manipolazione: possibilità di effettuare sovrapposizioni tra immagini raster diverse
- **Facilità di acquisizione e riproduzione**
- **SVANTAGGI:**
 - Per avere una buona definizione delle informazioni occorre creare raster con molti pixel, con **grande occupazione di memoria** (esempio: cella con valori da 1 a 255, 1 pixel = 1byte, Area di 100x100km con risoluzione di 10m = $10.000 \times 10.000 \text{pixel} = 100 \text{ MB}$). Si usano tecniche di compressione quali RLE, JPEG, ecc...)
 - Limitata scalabilità (non posso cambiare scala o zoomare a piacere in quanto: il dettaglio dell'informazione contenuta è limitato)
 - Lentezza di elaborazione
 - Memorizzazione di una sola variabile (o attributo)



walter.balzano@gmail.com

Le Mappe nei GIS

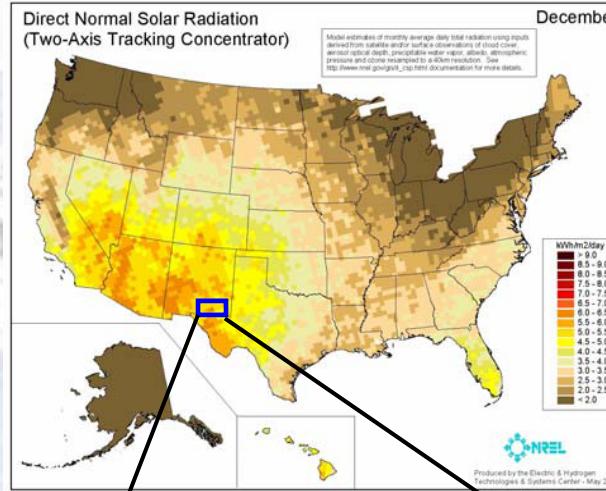


walter.balzano@gmail.com

Esempi di Mappa Raster



Cell.1	Cell.2	Cell.3	Cell.4	Cell.5
10	5	168	28	211
Cell.6
71				
....
....	Cell.20 21



Cell.1	Cell.2	Cell.3	Cell.4	Cell.5
100	145	23	178	1
Cell.6
44				
....
....	Cell.20 267



walter.balzano@gmail.com

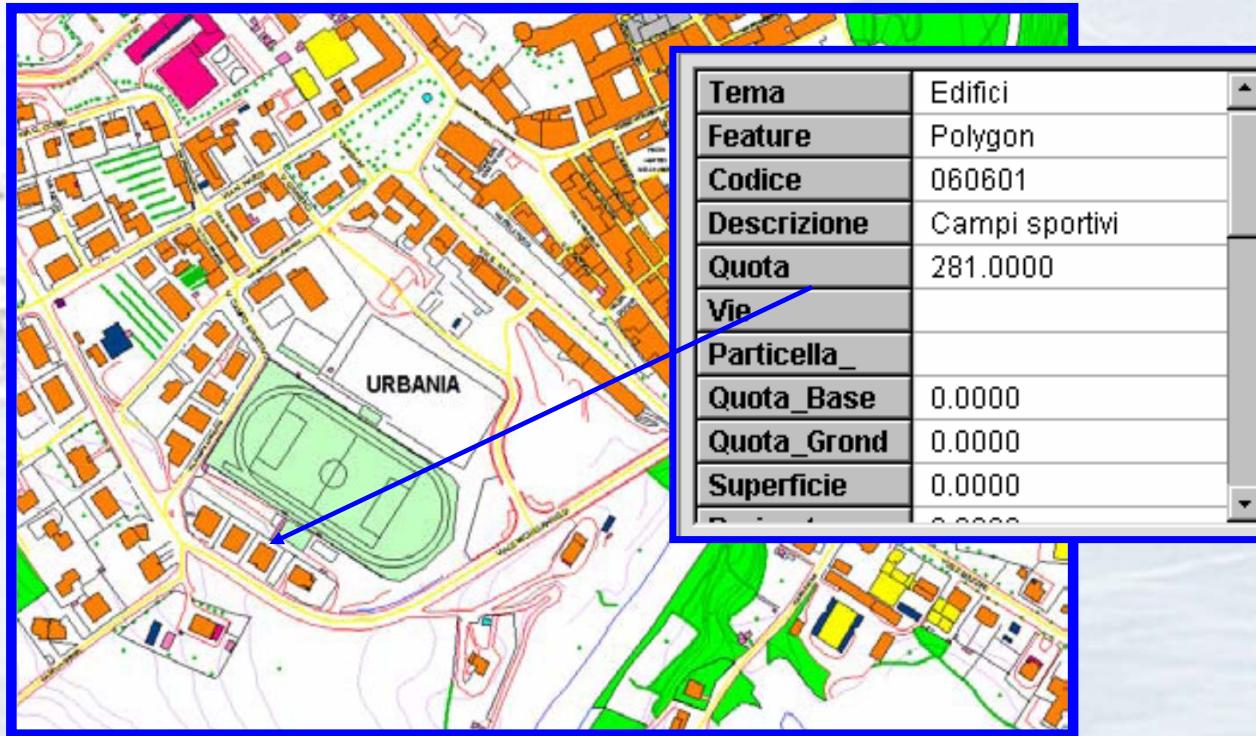
Mappa Vettoriale

- Rappresenta le informazioni territoriali con i 3 elementi: **PUNTI, LINEE, POLIGONI.**
- Ogni elemento (o FEATURE) è caratterizzato dalle sue coordinate geografiche
 - **PUNTO = (x,y)**
 - **LINEA (o ARCO) = (x₁,y₁), (x₂,y₂), ... , (x_n,y_n)**
 - **POLIGONO = ARCO che si chiude su se stesso e definisce un'area chiusa**
- Permette una **scalabilità completa** senza perdita delle informazioni
- Permette la gestione degli **ATTRIBUTI** (=informazioni alfanumeriche):
 - 1 record associato ad ogni feature
- Permette il **rendering** (colori, simboli, spessori, riempimenti,...) in funzione dei valori degli attributi



walter.balzano@gmail.com

Mappa Vettoriale



walter.balzano@gmail.com

Mappa TRIDIMENSIONALE

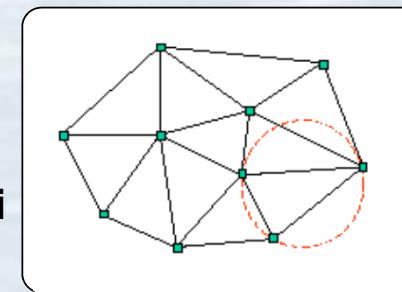
● TIN (Triangulated Irregular Network)

- E' un modello che serve per la rappresentazione di superfici 3D. In particolare viene usato per la generazione dei DEM o DTM (**Digital Elevation Model**), strutture che modellano l'elevazione e l'orografia del territorio.
- Viene costruito a partire da una serie di punti (x,y) dei quali si conosce la **coordinata Z**.
- Viene costruita una **TRIANGOLAZIONE** (=insieme di triangoli che coprono l'intera area di interesse senza sovrapporsi) che rappresenta in maniera compatta l'andamento della superficie 3D

Il TIN viene costruito rispettando la **REGOLA DI TRIANGOLAZIONE DI DELAUNAY**:

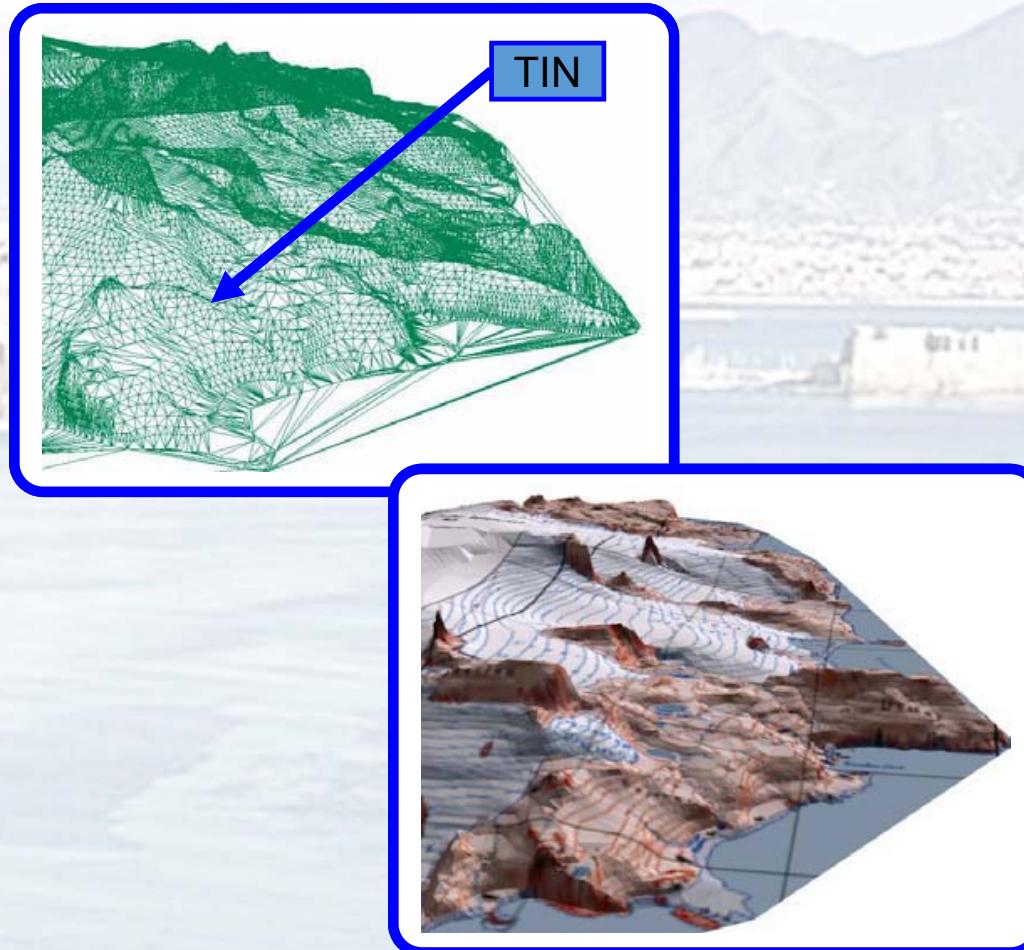
il cerchio che passa per i 3 vertici di un triangolo **NON** contiene alcun altro vertice (significa che i triangoli sono il più possibile "vicini" ad essere equilateri).

C'è un teorema che dimostra che: comunque siano dati un insieme di punti nel piano **esiste sempre almeno una triangolazione che verifica la regola di Delaunay**.



walter.balzano@gmail.com

Mappa TRIDIMENSIONALE



walter.balzano@gmail.com

Mappa TEMATICHE

- Con il termine **MAPPE TEMATICHE** si identificano diverse **tecniche di rappresentazione** in mappa dei dati
 - Hanno lo scopo di **concentrare l'attenzione sulla distribuzione di una particolare variabile**
 - Facilitare l'analisi comparativa individuando omogeneità o concentrazione di un fenomeno sul territorio
- Si possono classificare a seconda della **natura dei dati utilizzati**, della **tecnica di visualizzazione** e del **modello di mappa** in 5 tipi:
 - 1.COROPLTE
 - 2.A DENSITA' DI PUNTI
 - 3.SCALARI
 - 4.CATEGORIALI
 - 5.A ISOLINEE



walter.balzano@gmail.com

1 - Coroplete

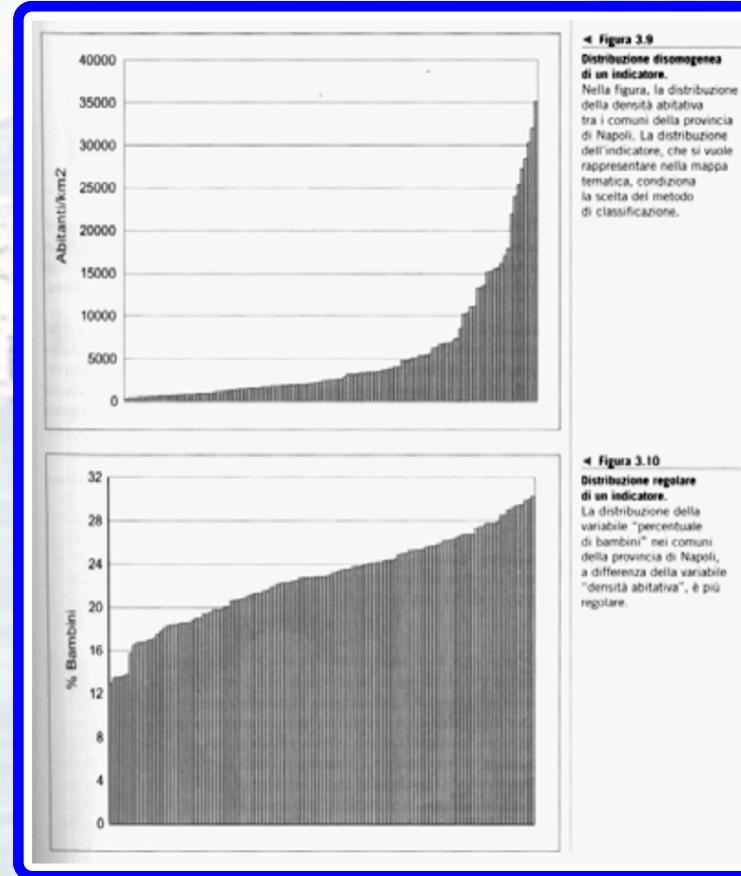
- Sono molto usate e **visualizzano la distribuzione** di un attributo nello spazio in forma classificata attraverso l'uso di scale cromatiche
- Vengono costruite attraverso un processo di **classificazione della variabile quantitativa**
 - Decidere il numero e l'ampiezza delle classi non è un problema banale
- I metodi di classificazione si basano su diversi schemi:
 - Metodi basati su **caratteristiche statistiche** della variabile in esame (metodo a intervalli costanti, metodi delle progressioni o dei quantili)
 - **Metodi idiografici** (le classi sono basate su punti caratteristici delle distribuzioni)
 - **Metodi esogeni** (utilizzano valori esterni alla distribuzione della variabile per creare le classi)



walter.balzano@gmail.com

Esempio visualizzazione indicatori

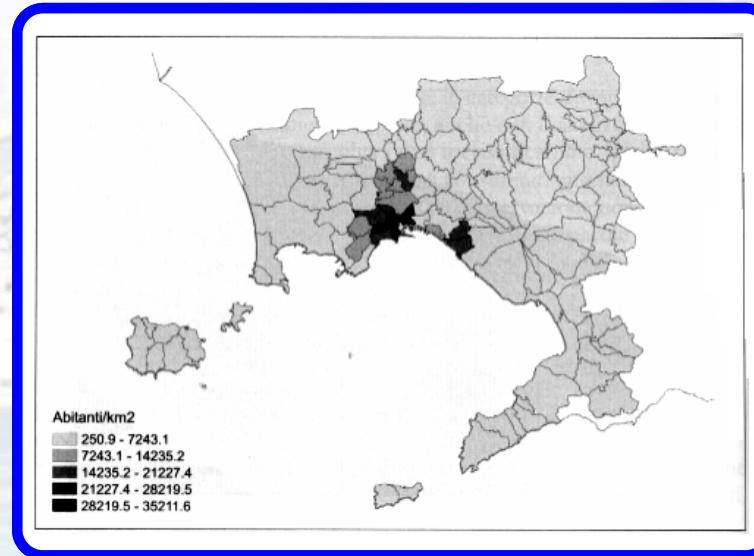
- Nella prima figura viene rappresentata la **distribuzione della densità abitativa** tra i comuni della provincia di Napoli
 - Si può notare la forte **disomogeneità** dell'indicatore tra i diversi comuni
- La seconda figura rappresenta la **distribuzione della percentuale di bambini** sempre all'interno della provincia di Napoli
 - Si può notare la **maggior regolarità** della distribuzione
- L'andamento della distribuzione dell'indicatore **influerà la scelta del metodo** di classificazione della variabile da visualizzare



walter.balzano@gmail.com

Metodo delle classi di uguale ampiezza

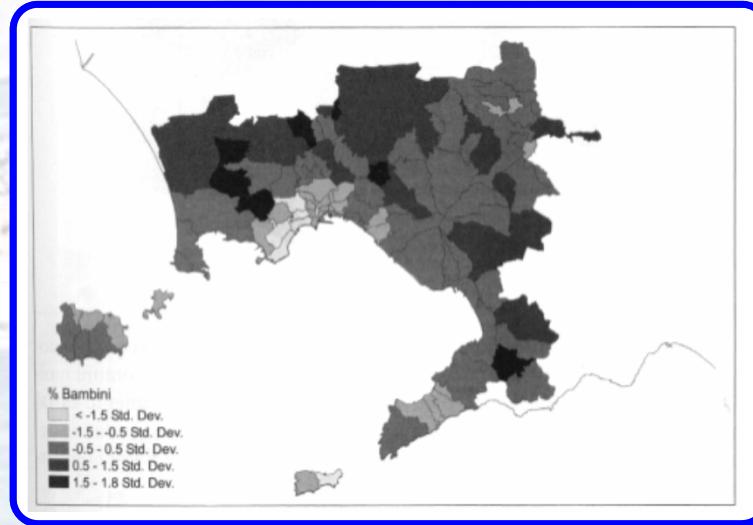
- Si suddivide la distribuzione della variabile in intervalli della stessa ampiezza
- Si divide il **range** della variabile (distanza tra il min ed il max) per il **numero delle classi** che si vuole utilizzare e si ottiene **l'ampiezza della classe**
- **NON** funziona per variabili distribuite **DISOMOGENEAMENTE**
- Produce una mappa appiattita dove spiccano solo le unità con i valori estremi



walter.balzano@gmail.com

Metodo delle deviazioni standard

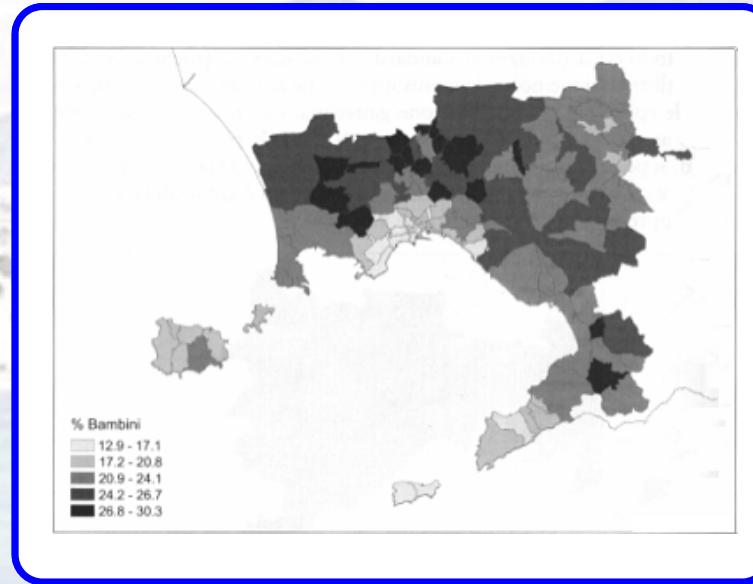
- Presuppone che la distribuzione sia uniformemente distribuita
- Le classi sono create come la media della distribuzione + multipli della deviazione standard
- Offre una buona analisi comparativa del fenomeno



walter.balzano@gmail.com

Metodo dei quantili

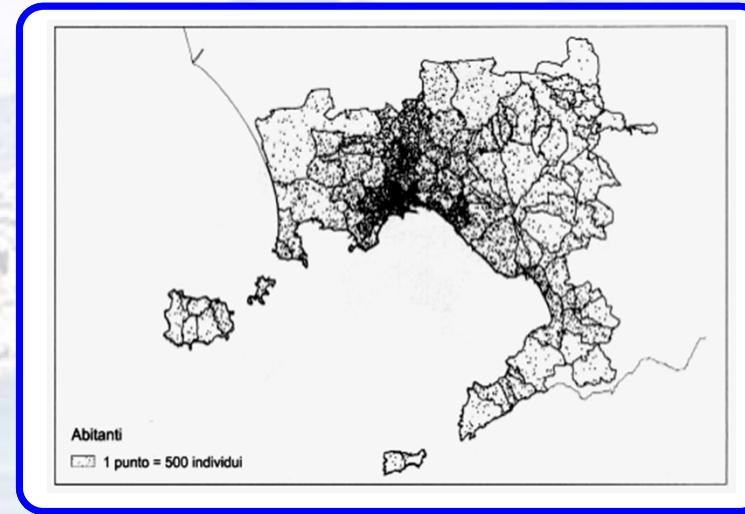
- Detto anche di *equinumerosità delle osservazioni*
- Crea delle classi di **ampiezza variabile** e tale da avere **lo stesso numero di osservazioni** su ogni classe
- Le classi possono avere un range dell'indicatore anche molto grande
- Per distribuzioni fortemente **disomogenee** sono **fuorvianti**



walter.balzano@gmail.com

2 - Mappe a densità di punti

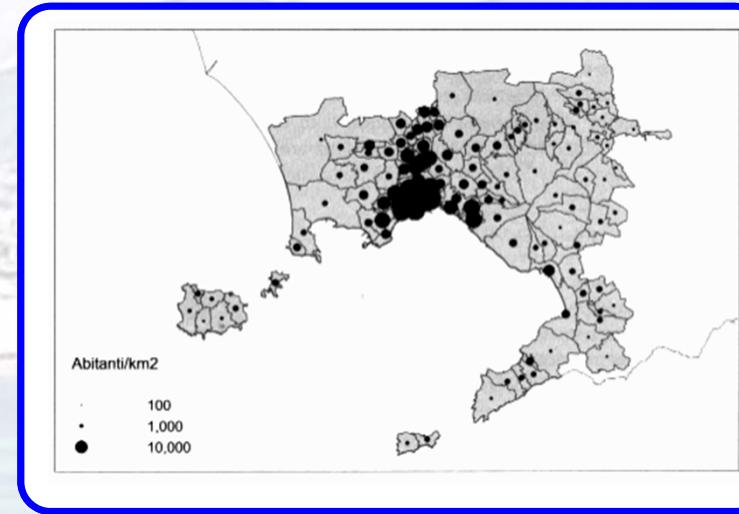
- Visualizzano la distribuzione nello spazio di un attributo nello spazio sotto forma di simboli grafici in numero proporzionale al valore dell'attributo
- All'interno di un comune ad esempio si può visualizzare il numero di abitanti creando un punto disposto casualmente sul territorio ogni un determinato numero di abitanti (al limite un punto per abitante)
- Forniscono una visualizzazione della densità molto efficace



walter.balzano@gmail.com

3 - Mappe a simboli scalari

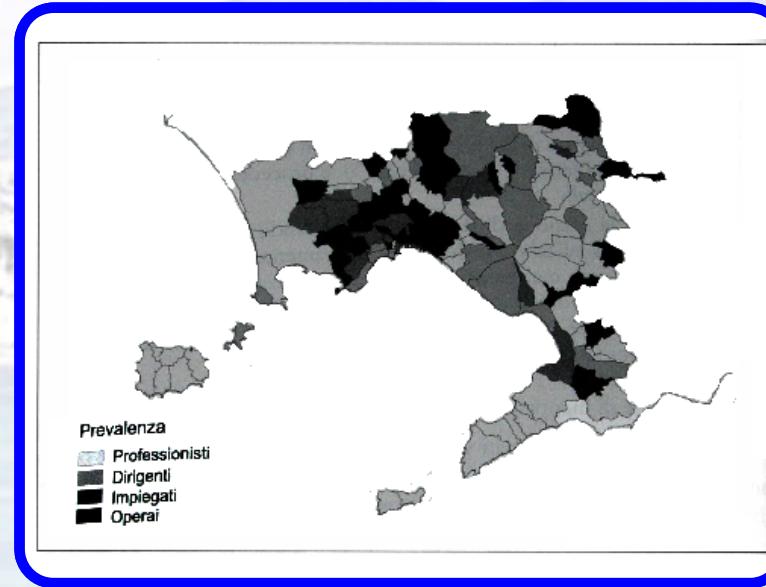
- I simboli assumono una dimensione proporzionale all'entità della variabile che rappresentano
- La riproduzione proporzionale viene solitamente fatta in modo tale che il raggio del cerchio (simbolo) sia proporzionale al dato e non alla sua superficie come sarebbe matematicamente corretto per limitare l'errore visuale di sottostima dei simboli di maggiore dimensione



walter.balzano@gmail.com

4 - Mappe categoriali

- Le variabili categoriali sono visualizzate con simboli o colori che identificano individualmente i diversi valori
- Le mappe categoriali evidenziano intuitivamente il grado di dispersione o di concentrazione dei fenomeni nello spazio

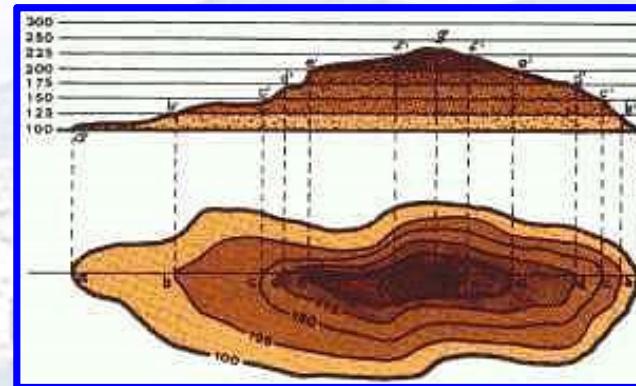


walter.balzano@gmail.com

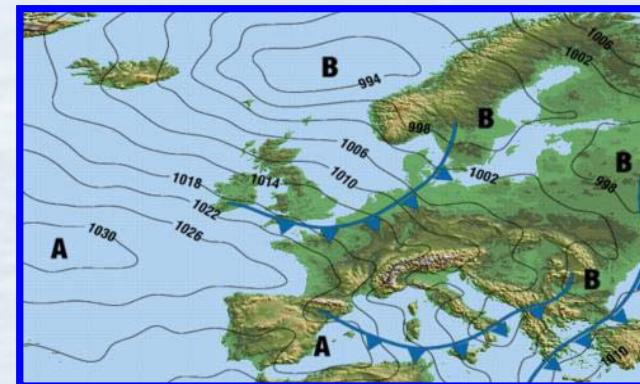
5 - Mappe a isolinee

- Vengono disegnate **linee** che congiungono **zone** a ugual **valore della variabile** formando degli anelli
- In relazione alla grandezza che descrivono assumono diversi nomi:
 - Isobare → uguale pressione
 - Isobate → stessa profondità del mare
 - Isoipse → stessa altitudine
 - Isocline → stessa acclività
- Si utilizzano serie di isolinee ottenute dividendo la variabile in **intervalli regolari**

isoipse



isobare



walter.balzano@gmail.com

Analisi Spaziale

- **Introduzione**
- **L'esplorazione dello spazio**
- **Modelli di ripartizione territoriale**
- **Indici statistici geospaziali**
- **Modellazione delle superfici**
- **Metodi di Classificazione**



walter.balzano@gmail.com

Esplorazione dello spazio

- Obiettivo della **misurazione** degli attributi spaziali e non spaziali è quello di **individuare**, **contare** e **localizzare** gli oggetti presenti in una regione
- La **descrizione** accurata delle proprietà geometriche degli oggetti, la **misurazione** delle loro relazioni, **l'enumerazione** secondo diversi criteri topologici aiuta a formare una **tassonomia dello spazio** e fornisce una prima comprensione dei fenomeni rappresentati nella mappa.
- **Nell'esplorazione dello spazio occorre considerare:**
 - **Gli attributi spaziali**
 - **L'analisi multistrato**



walter.balzano@gmail.com

Gli attributi spaziali

● Geometria degli oggetti derivabili dalla forma degli oggetti presenti nella mappa.

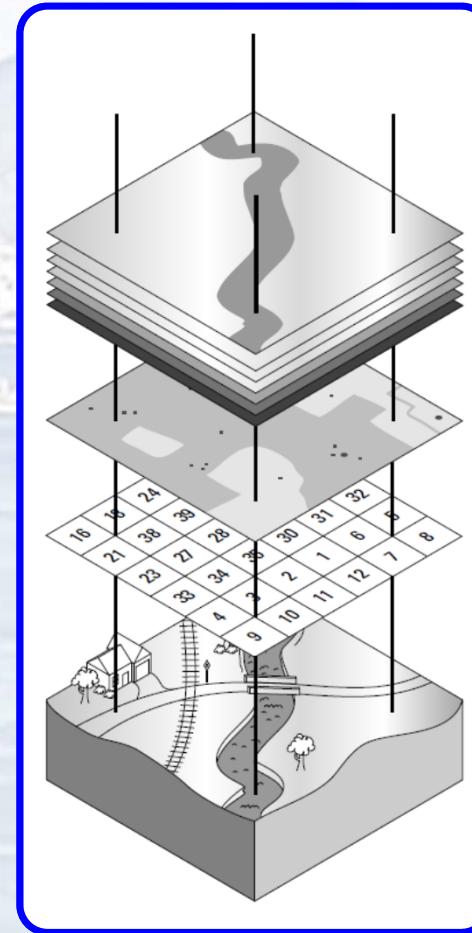
- (un sistema GIS è in grado di lavorare direttamente sul modello sferico per evitare le distorsioni introdotte dalle proiezioni)
- Superficie e perimetro dei poligoni
- Lunghezza linee
- Compattezza dei poligoni
- Frastagliatura linee di confine
- Centroidi

● Rapporti nello spazio (proprietà prossimali e topologiche)

- Distanze (assolute o relative)
 - Distanza euclidea (o linea d'aria)
 - Distanza funzionale (vincolata a direttrici di spostamento)
 - Accessibilità temporale (tempo necessario per spostarsi tra due punti)
 - Costi del viaggio

● Proprietà topologiche

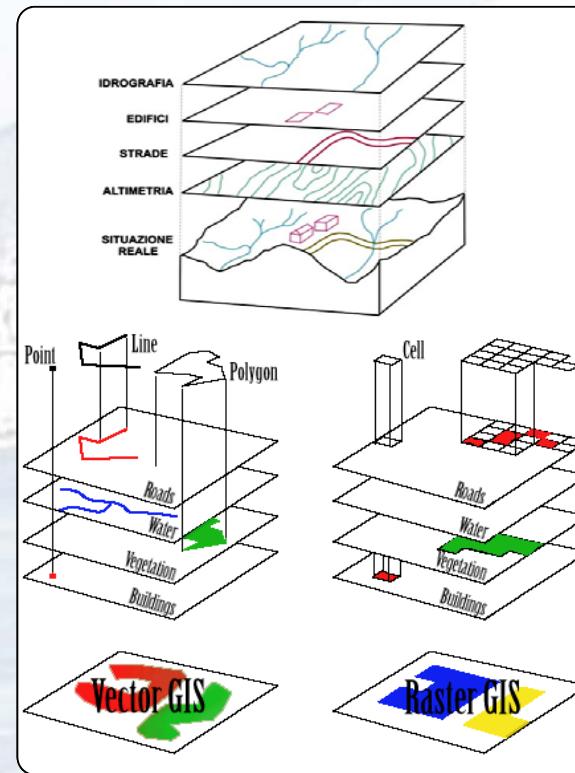
- Contiguità spaziale tra oggetti (intersezione tra linee, poligoni,...)
- Le entità descritte sono trattate nei GIS mediante Query



walter.balzano@gmail.com

L'Analisi Multistrato

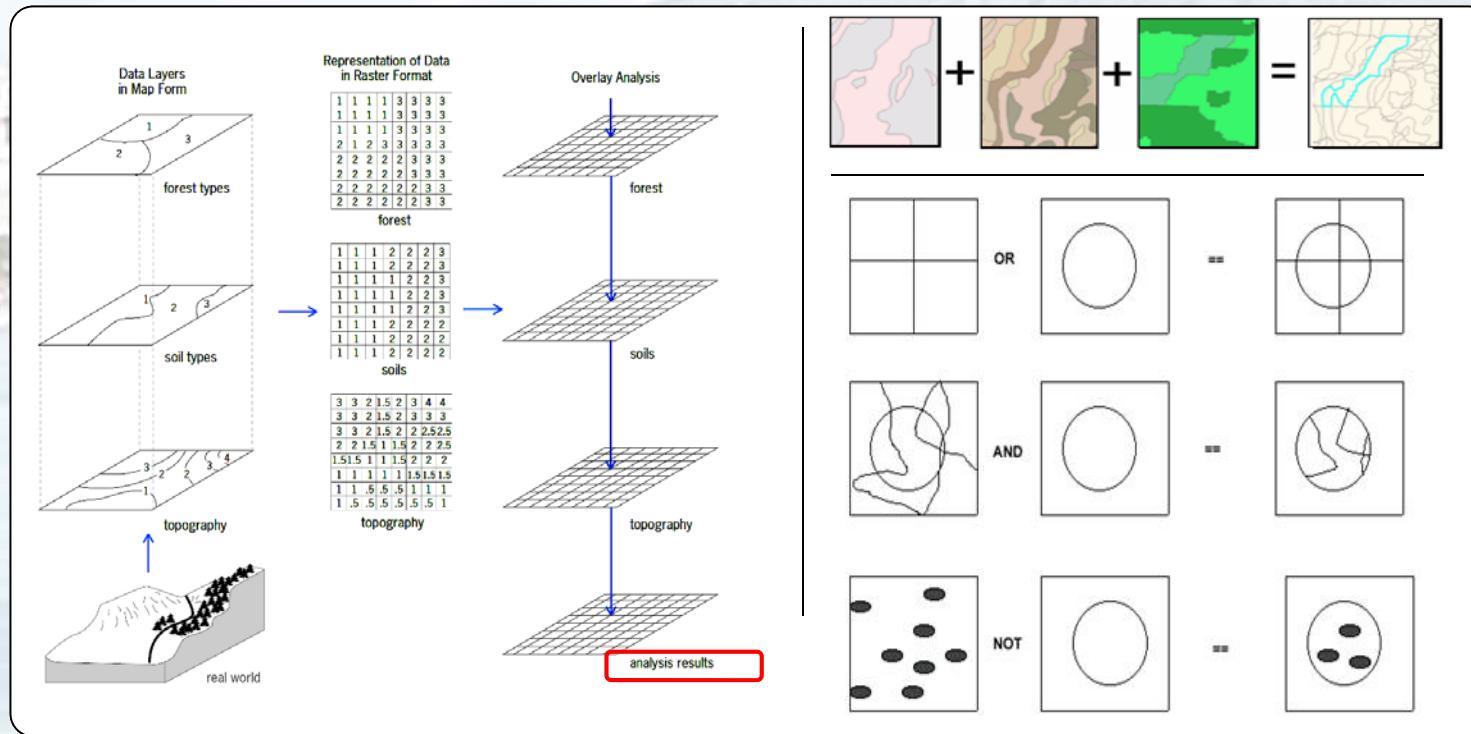
- **Analisi multistrato:** efficace tecnica di analisi territoriale il cui obiettivo è quello di **confrontare informazioni contenute in più strati informativi** riferiti a diverse unità di osservazione su un'**area comune**.
- Permette di individuare aree caratterizzate dalla **compresenza di più fenomeni**.
 - *Esempi:*
 - Studio della relazione tra l'immigrazione straniera ed il degrado edilizio
 - Confronto della densità abitativa in diversi periodi di tempo
- Le più comuni modalità di confronto tra differenti layer informativi sono realizzate mediante tecniche di **OVERLAY**:
 - Sovrapposizione tra mappe raster
 - Sovrapposizione tra mappe vettoriali



walter.balzano@gmail.com

Sovrapposizione tra mappe raster

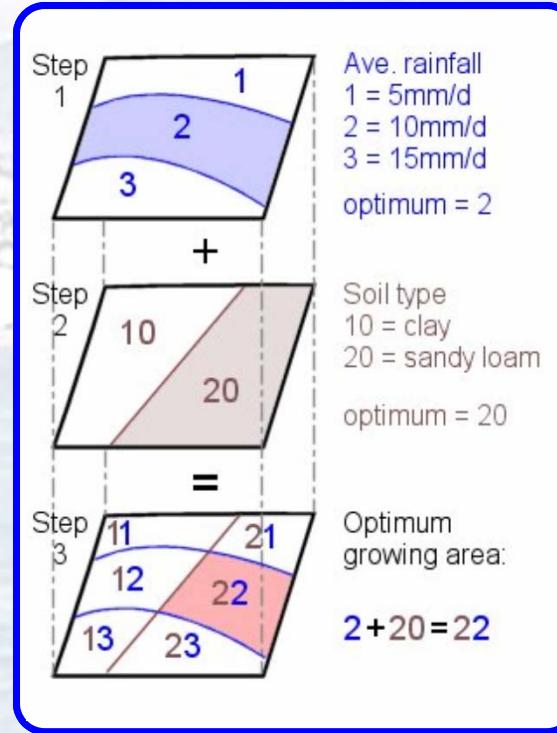
- Pre-requisito: le mappe da confrontare devono essere caratterizzate dallo **stesso livello di precisione**
- Le cellette possono essere confrontate pixel per pixel usando operatori logici e/o operatori aritmetici



walter.balzano@gmail.com

Sovrapposizione tra mappe vettoriali

- Dati gli oggetti vettoriali rappresentati nei diversi strati informativi ne viene calcolata l'intersezione
- Viene generato un nuovo strato informativo che mette in evidenza una possibile combinazione delle due informazioni
- Le combinazioni possono essere di tipo logico (booleano) o matematico
- Il nuovo strato fornisce all'utente informazioni per la scelta di siti per opportune destinazioni d'uso o correlazioni tipo causa-effetto di alcuni fenomeni socio-economici



walter.balzano@gmail.com

Modelli di ripartizione territoriale

- Sono modelli di analisi spaziale che consistono nella creazione di **nuove unità territoriali** per:
 - Trasformazioni di entità esistenti
 - **Ripartizione dello spazio secondo diversi modelli**
- Sono rappresentati da famiglie di metodi che hanno in comune l'obiettivo di **riorganizzare lo spazio secondo opportuni criteri**
- Tali criteri possono essere di tipo sociologico, funzionale, urbanistico, fisico, ecc
- **Principali modelli di ripartizione territoriale:**
 - Riclassificazione
 - Buffer
 - Vicinato di Voronoi
 - Gravitazione



walter.balzano@gmail.com

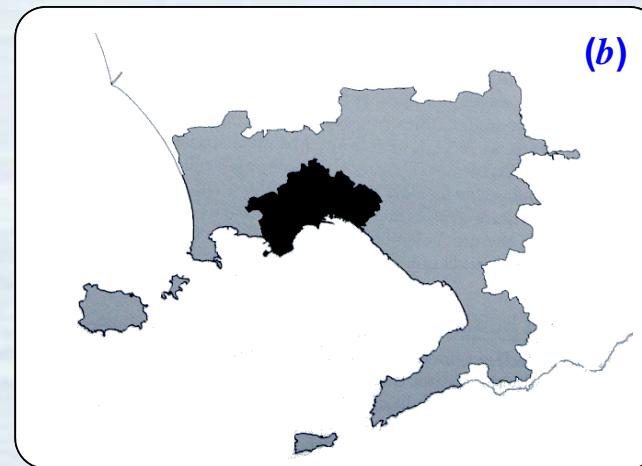
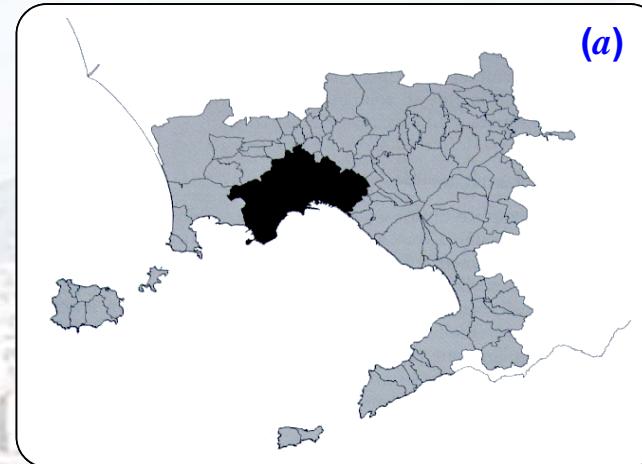
Riclassificazione

- Riclassificazione: riorganizzazione degli oggetti preesistenti dello spazio
- Mediante la riclassificazione vengono trasformate aree omogenee visualizzate nella mappa tematica in nuove aree che risultano dalla aggregazione di aree simili
- I criteri di aggregazione possono essere basati su attributi oppure su proprietà topologiche/prossimali
- Gli attributi sono aggregabili con operazioni e funzione tipiche dei DataBase (somma, medio, minimo,...) con possibilità di creare nuovi indicatori ottenuti pesando il contributo di ciascun oggetto componente
- Dal punto di vista geometrico la riclassificazione cancella i confini che delimitano le aree originali e ridisegna nuove figure geometriche prive di divisioni interne.

Esempio di riclassificazione dei quartieri urbani

(a) comune di Napoli e dei comuni della sua provincia

(b) il capoluogo ed il resto della provincia



walter.balzano@gmail.com

- Con il termine **BUFFER** si identifica un'area che si estende intorno ad un oggetto individuando una **SUPERFICIE DI PERTINENZA** dell'oggetto stesso

- *Esempio:*

- Un'area circolare intorno ad una scuola ne identifica il bacino d'utenza
- Un buffer attorno ad un'autostrada rappresenta la fascia soggetta ad inquinamento acustico o chimico dovuto al traffico veicolare

- I buffer identificano aree **EQUIDISTANTI** da un particolare fenomeno

- Possono essere creati generando **aree di ampiezza dipendente** da un attributo specifico dell'oggetto in questione.

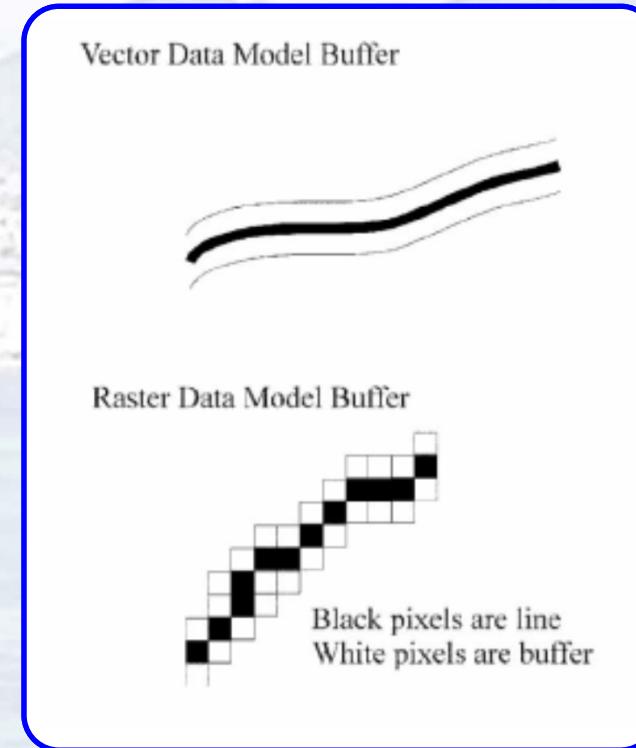
- *Esempio:*

- Un buffer attorno ad una sorgente inquinante potrebbe avere un'area proporzionale all'entità di emissione in quanto l'estensione della zona inquinata dipende dall'entità della sorgente

- L'analisi multistrato tra buffer e altri strati informativi ci può dare diverse informazioni utili

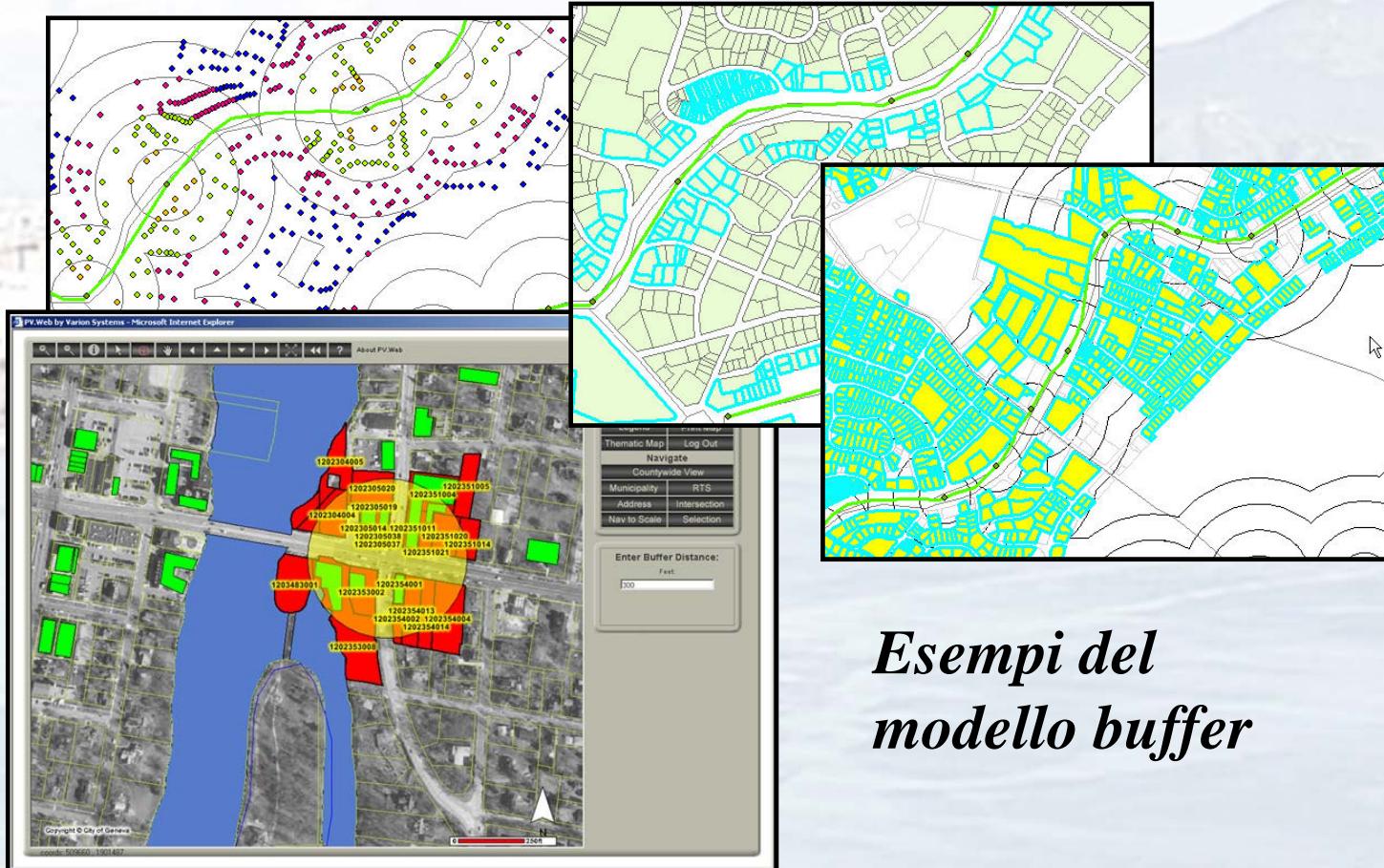
- La produzione di un buffer risulta computazionalmente abbastanza semplice sia nel caso di oggetto vettoriale sia nel caso di una mappa raster

Modello a buffer



walter.balzano@gmail.com

Modello a buffer (2)



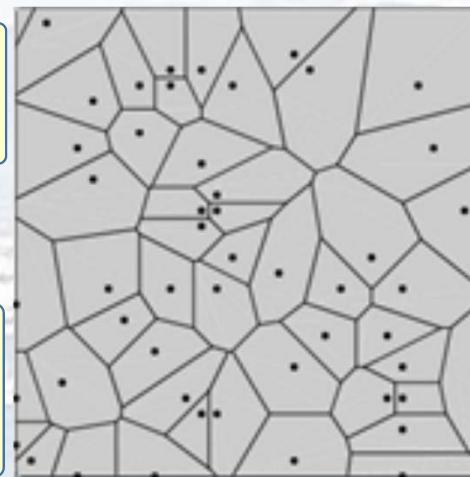
*Esempi del
modello buffer*



walter.balzano@gmail.com

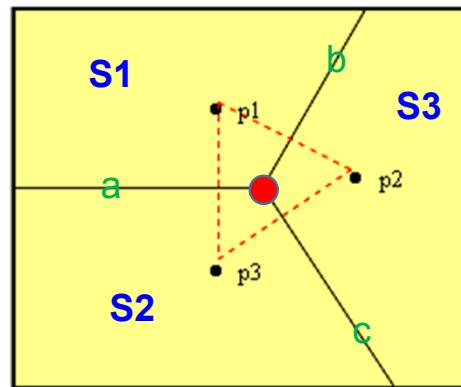
Modello di vicinato di Voronoi

- Attraverso questo modello viene generata una ripartizione territoriale, basata su un criterio detto dell'AREA DEL VICINATO, a partire da una distribuzione di punti sul territorio
- Il modello di Voronoi genera su un insieme di punti una collezione di regioni che ripartiscono il piano in maniera continua
- Il modello gode della seguente proprietà:
 - Ogni regione contiene esattamente un punto del set dato e ha la proprietà di contenere tutti i punti del piano che sono più vicini a quel punto piuttosto che ad un altro



Generazione dei POLIGONI di VORONOI

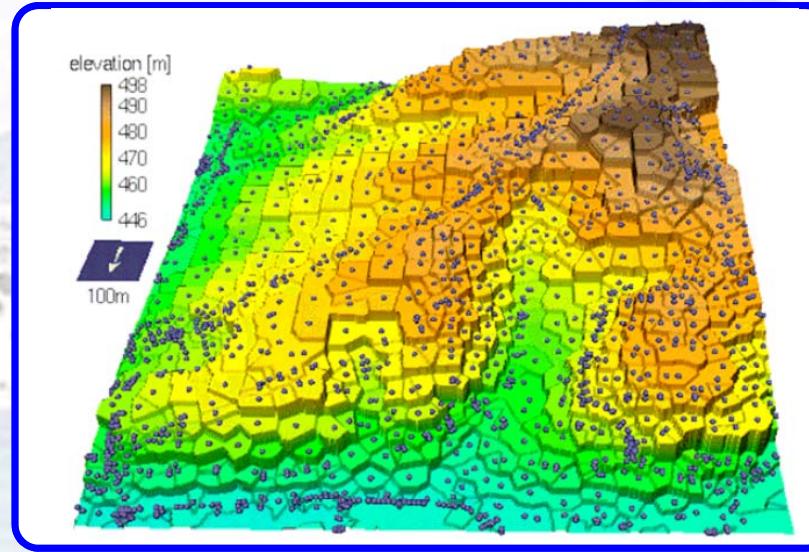
- Dati tre punti nel piano tracciamo le congiungenti dei tre punti
- Su ogni congiungente tracciamo la linea bisecante (divide in due la congiungente ed è perpendicolare)
- L'intersezione delle tre bisecanti individua un punto nel piano, chiamato punto di Voronoi, che è equidistante ai tre punti del piano (p_1, p_2, p_3).
- In questo caso particolare vengono quindi identificate tre regioni sul piano, chiamate regioni di Voronoi o celle di Voronoi, ciascuna contenente un punto dato.



walter.balzano@gmail.com

Modello di Voronoi nel GIS

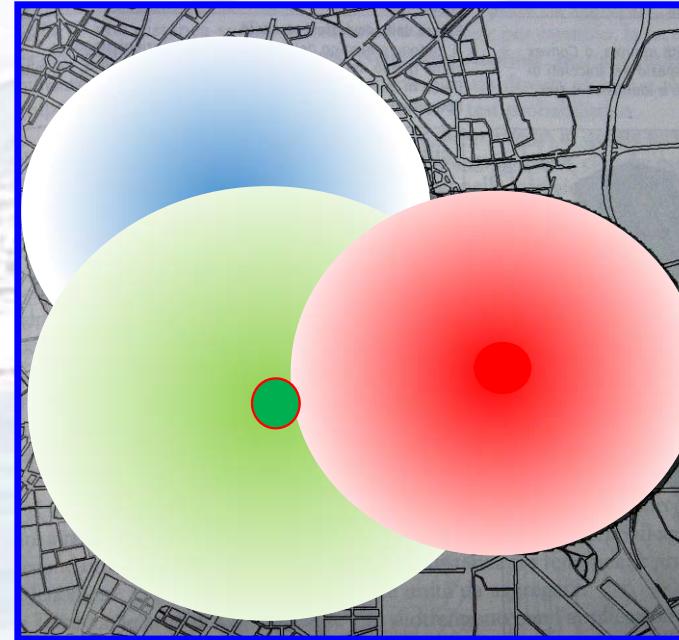
- Le regioni definite dalla poligonazione di Voronoi sono chiamate anche aree di prossimità o bacini di attrazione
- Tali poligoni costituiscono un buon modello per la definizione dei bacini di utenza o attrazione:
 - Contengono tutti i punti del piano più vicini al punto generatore che non ad altri punti
- Definita l'attrazione di un punto sulla superficie (Esempio: centro commerciale) come una grandezza proporzionale alla distanza dal punto, il poligono di Voronoi identifica l'area contenente tutti i punti (le località geografiche) più vicini al punto generatore (centro commerciale)
- Servono anche alla ricostruzione della distribuzione di un fenomeno nello spazio disponendo di osservazioni campionarie
 - Si divide lo spazio con poligoni di Voronoi ed ad ogni poligono si associa il valore della grandezza rilevato nel punto generatore
 - Ciò evita di associare un valore campionario a punti del territorio più vicini ad altri campioni



walter.balzano@gmail.com

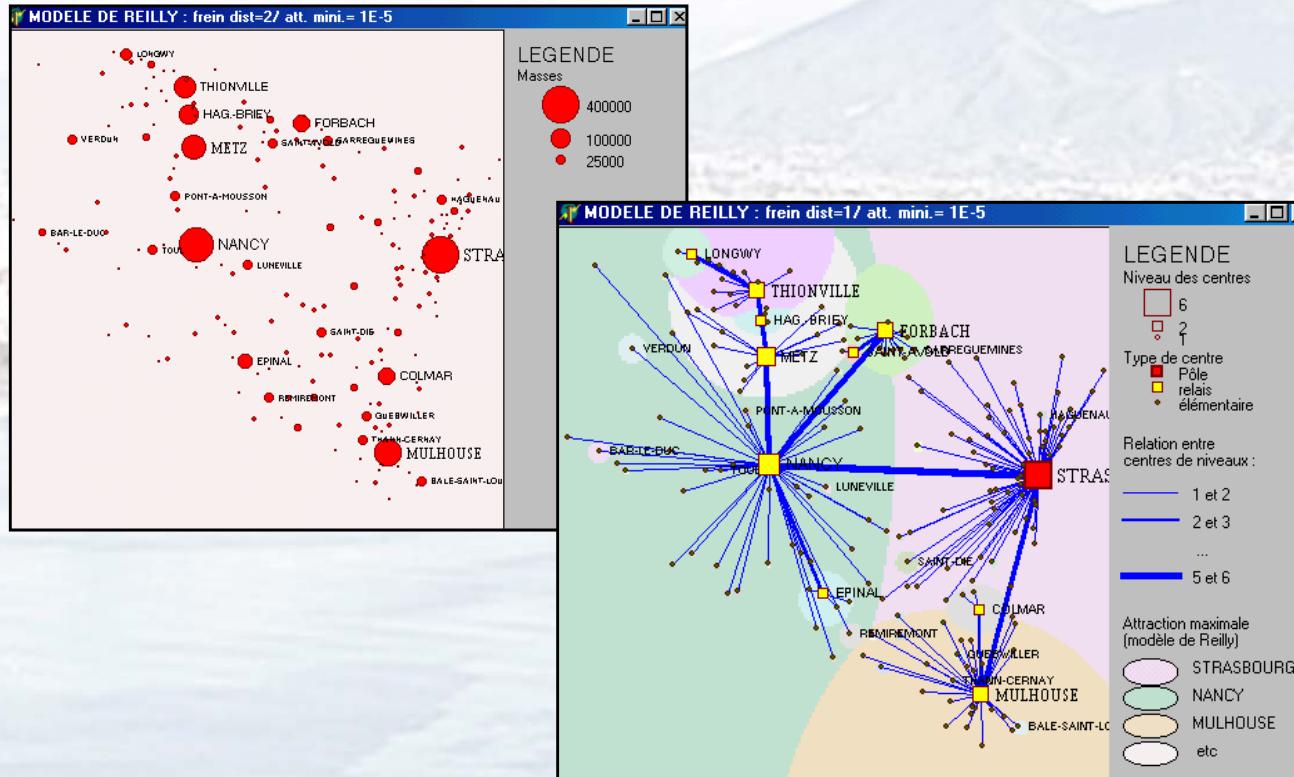
Modelli di gravitazione

- Sono modelli di ripartizione territoriale basati sull'analogia con il modello di attrazione gravitazionale proposto da Newton secondo il quale **due corpi si attraggono** in modo dipendente dalla loro massa e dalla **distanza** che li separa
 - Oggetti più vicini esercitano un'attrazione maggiore
 - Oggetti più grandi esercitano un'attrazione maggiore
- Nei sistemi GIS questi modelli definiscono l'attrazione tra i punti in base alla distanza e al valore di un attributo associato al punto
 - **Esempio:** l'attrazione commerciale tra due città è proporzionale alla loro popolazione ed è inversamente proporzionale al quadrato della loro distanza
- Il Modello di gravitazione genera una ripartizione completa del territorio in aree di attrattività



walter.balzano@gmail.com

Modelli di gravitazione (2)



walter.balzano@gmail.com

Indici statistici geospaziali

- Sono sistemi di analisi spaziale del territorio che forniscono misure sintetiche della **disposizione spaziale** di un fenomeno
- Tali indici danno informazioni riguardo:
 - Il **baricentro** del fenomeno
 - La **dispersione** territoriale
 - La **direzione** prevalente
 - La **regolarità** della distribuzione
- Con il termine **disposizione spaziale** o **pattern** si intende una determinata configurazione degli oggetti nello spazio
 - a) **Pattern casuale**: non esiste nessuna regolarità o motivo
 - b) **Pattern uniforme**: la distanza di ciascun fenomeno da un altro è costante
 - c) **Pattern a cluster**: i fenomeni si trovano disposti in addensamenti staccati da zone vuote

Pattern casuale



Pattern uniforme



Pattern a cluster



walter.balzano@gmail.com

Statistiche centrografiche

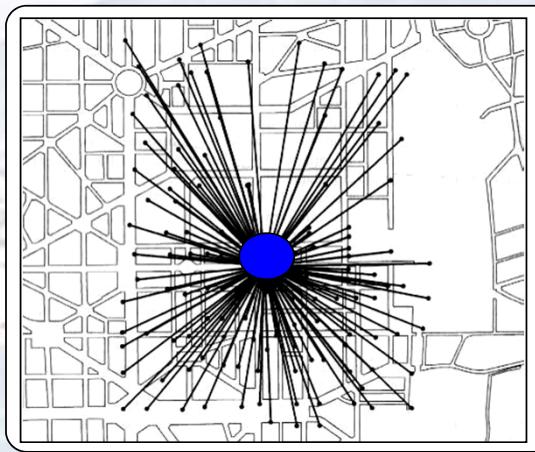
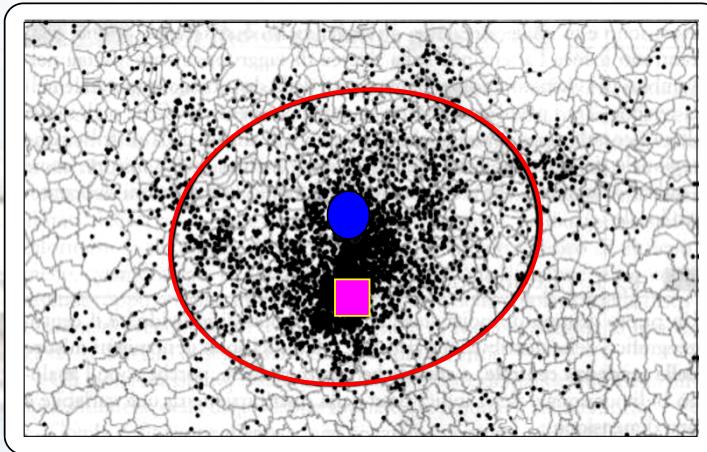
- Obiettivo primario delle statistiche centrografiche è l'individuazione del centro geografico della distribuzione dei dati:
 - **CENTRO MEDIO:** media delle coordinate di latitudine e longitudine delle singole osservazioni (detto anche centro di gravità o punto di equilibrio delle osservazioni)
 - **CENTRO DELLA DISTANZA MINIMA:** rappresenta il punto che minimizza le distanze da tutti gli altri punti della distribuzione detto anche punto di "viaggio minimo"
 - **DEVIAZIONE STANDARD DELLA DISTANZA:** misura la dispersione dei fenomeni sul territorio in termini di distanza rispetto al centro medio
 - **CERCHIO della deviazione standard:** l'area viene evidenziata da un cerchio centrato sul CENTRO MEDIO e il cui raggio è multiplo della deviazione standard
 - **ELLISSE delle deviazioni standard:** Si calcolano le deviazioni standard lungo l'asse della latitudine e della longitudine definendo gli assi di un ellissi; oltre ad una misura delle dispersione



walter.balzano@gmail.com

Statistiche centrografiche

Esempi



- **Punti:** residenze studenti iscritti Università di Milano Bicocca
- **Cerchio centrale:** Università di Milano Bicocca
- **Quadrato:** centro medio
- **Ellisse:** area di maggiore concentrazione

- **Cerchio centrale:** centro della distanza minima da tutti i punti rappresentati nell'area.
In questo esempio si misurava l'accessibilità pedonale da un servizio di trasporto pubblico



walter.balzano@gmail.com

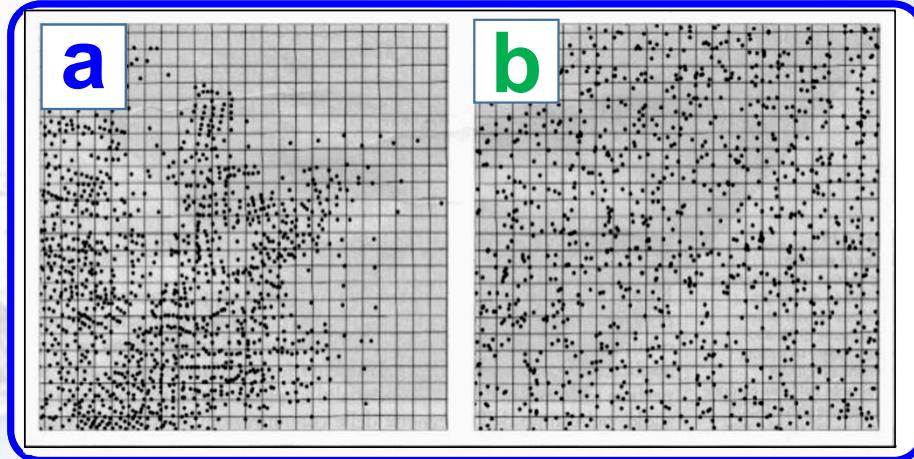
Analisi dei quadranti

- L'Analisi dei quadranti è uno strumento per l'analisi del **pattern** (distribuzione di una variabile sul territorio)
- E' largamente usato in campo biologico per lo studio delle immagini derivanti dalla microscopia
- Si basa sul conteggio delle osservazioni su di una griglia regolare a maglie quadrate che suddivide l'immagine
 - Per ciascun quadrato viene valutata la frequenza delle osservazioni (**OCCORRENZA**)
 - Si calcola la varianza delle frequenze (misura di dispersione del fenomeno)
 - Dal confronto tra la distribuzione osservata e una distribuzione casuale si deriva una **misura della presenza di pattern** (raggruppamenti) della variabile sul territorio
- **PUNTI DEBOLI:**
 - La **dimensione della griglia** influenza notevolmente il risultato ma non esiste una regola per determinarne la dimensione ottima
 - Non tiene conto delle **variazioni locali** tra le celle confinati



walter.balzano@gmail.com

Analisi dei quadranti



La mappa (a) rappresenta la popolazione residente nel quadrante Nord-est del comune di Milano
In (b) viene rappresentata una distribuzione casuale di un ugual numero di punti rispetto a (a)
Il confronto tra la distribuzione dei punti nelle due griglie fornisce una misura statistica della regolarità del pattern



walter.balzano@gmail.com

L'analisi di vicinato (Nearest Neighbour Analysis)

- L'Indice Nearest Neighbour (NNI) misura il grado di dispersione spaziale di una distribuzione di punti. Il calcolo è basato sulla misurazione delle distanze di punti adiacenti: la distanza tra punti raggruppati (clustered) è inferiore a quella in cui i punti sono distribuiti in un'area in modo casuale o uniforme.
- Le formule per il calcolo dell' NNI sono:

$$Ed = 0.5\sqrt{A/N}$$

$$Ad = \left(\sum_{i=1}^N d_i \right) / N$$

$$NNI = Ad / Ed$$

In cui:

Ed = distanza attesa

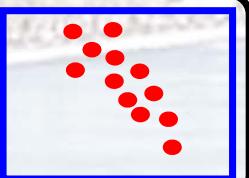
Ad = distanza di vicinato

d_i = distanza del punto i al punto più vicino

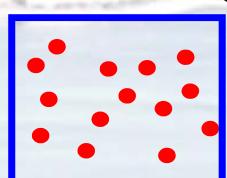
A = Area della mappa

N = numero totale dei punti

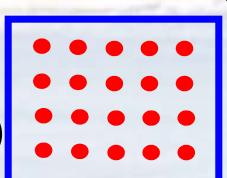
$NNI \rightarrow 0$
(Cluster)



$NNI \rightarrow 1.0$
(Random)



$NNI \rightarrow 2.14$
(Uniform)



Valori di NNI

0

0.23

0.5

1.0

1.5

2.0

2.1491

Tendenza al
raggruppamento

RANDOM

Tendenza alla
regolarità

- Se i punti cadono tutti nel medesimo luogo $\rightarrow Ad = 0 \rightarrow NNI = 0$
- Se i punti sono molto raggruppati $\rightarrow NNI$ tende a 0 poiché è piccola la loro distanza media
- Se i punti hanno una distribuzione casuale $\rightarrow NNI$ tende a 1
- Se i punti sono distribuiti in modo perfettamente uniforme $\rightarrow NNI$ tende a 2.1491

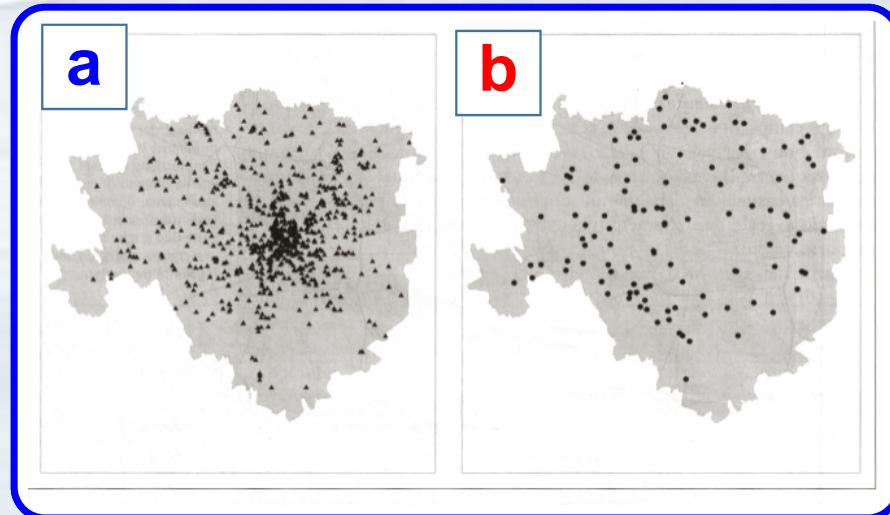


walter.balzano@gmail.com

• Esempio:

L'analisi di vicinato (2)

- La mappa **(a)** mostra i luoghi di cultura a Milano; dal calcolo si ottiene un indice di vicinato $NNI=0.53$
- La mappa **(b)** mostra i centri sportivi sempre a Milano; si ottiene un indice di vicinato $NNI = 0.27$
- In questo esempio l'analisi di vicinato mostra che i centri sportivi sono maggiormente organizzati a cluster (NNI minore) anche se all'apparenza sembrerebbe il contrario



walter.balzano@gmail.com

Metodi di Classificazione

- Metodi **LOCALI**: calcolano il valore sconosciuto **basandosi sui valori noti nell'immediato intorno**; riproducono fedelmente le variazioni locali della variabile. Sono basati su singole funzioni matematiche applicate ad una parte della totalità dei punti campionati
- Metodi **GLOBALE**: stimano i valori nei punti sconosciuti **utilizzando tutti i dati disponibili** e cercando di valutarne la distribuzione e l'**ANDAMENTO GLOBALE**; dato l'andamento globale si estrapolano i valori in qualunque zona d'interesse. Sono basati su singole funzioni matematiche applicate a tutti i punti.
- Metodi **ESATTI**: **(interpolazione)**
 - La superficie risultante passa esattamente in tutti i punti del data set
 - Valida con dati molto accurati
- Metodi **APPROSSIMATI**: **(smoothing)**
 - La superficie risultante NON passa esattamente in tutti i punti del data set
 - Valida con dati con alto grado di incertezza
- Metodi **GRADUALI**:
 - produce superfici lisce (smussate) che passano per tutti i punti
 - Appropriata per l'interpolazione di dati con piccola variabilità locale
- Metodi **BRUSCHI**:
 - produce superfici lisce a gradini
 - Appropriata per l'interpolazione di dati con grandi variabilità locali o con forti discontinuità
- Metodi **DETERMINISTICI**:
 - Usati quando sussistono sufficienti conoscenze circa la superficie da modellare
 - Permette di utilizzare modelli matematici specifici
- Metodi **STOCASTICI**:
 - Utilizzati per incorporare variabili random nelle superfici da interpolare



walter.balzano@gmail.com

Interpolazione

L' **INTERPOLAZIONE SPAZIALE** permette di stimare il **valore di una variabile in zone non “coperte” da osservazioni basandosi sulla conoscenza della variabile in punti noti.**

Molti GIS offrono alcuni sistemi di interpolazione; tipici metodi sono:

- Poligoni di Thiessen o Voronoi
- Triangulated Irregular Networks (TINs)
- Spatial moving average o altre tecniche matematiche
- Trend Surfaces
- Kriging

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

- Il fenomeno d'interesse si assume abbia natura continua e sia quindi misurabile con una variabile metrica
- I fenomeni osservati siano spazialmente dipendenti (è possibile derivare il valore della variabile in un punto dati i punti circostanti (LOCALE) o data la caratterizzazione di una certa regione (GLOBALE))

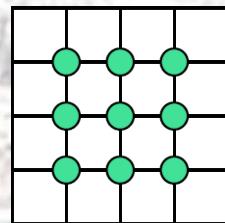
Il metodo di interpolazione deve essere scelto anche in base alla distribuzione dei punti noti (campioni)



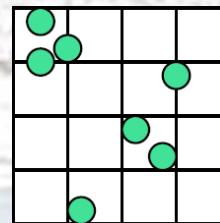
walter.balzano@gmail.com

Metodi di campionamento

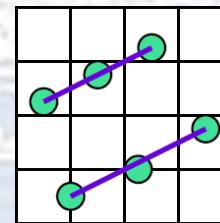
Il metodo di campionamento è fondamentale per la scelta dei successivi metodi di interpolazione:



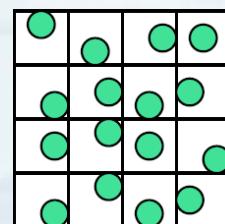
Regolare



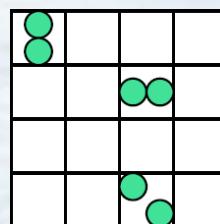
Random



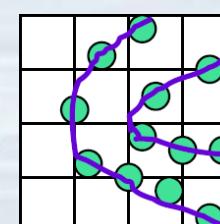
Transetti



Stratified Random



Cluster

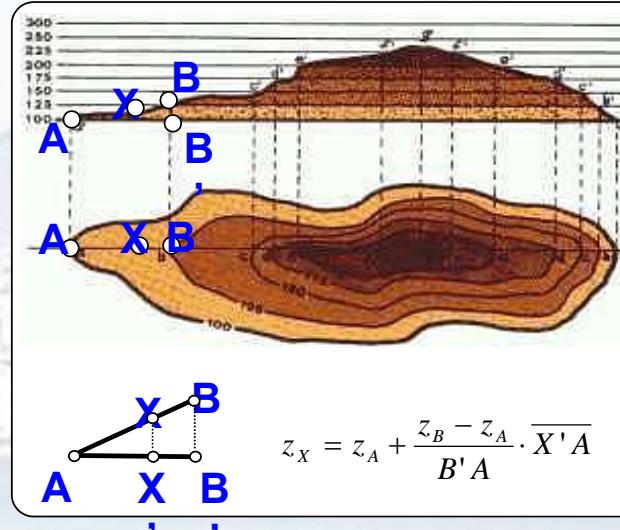


Contour



Interpolazione lineare da isolinee (stima locale)

- Un metodo molto comune per rappresentare l'andamento di un valore nello spazio è quello che utilizza le **isolinee**
- Il procedimento di interpolazione applicato a dati rappresentati sotto forma di isolinee permette di avere una **stima della variabile in qualsiasi punto** derivandolo dal valore nei punti che appartengono alle isolinee
- Il procedimento più semplice è quello del metodo di interpolazione lineare:
 - Si traccia il segmento più breve che unisce due isolinee adiacenti e passante per il punto di interesse.
 - Sul segmento generato si applica l'interpolazione lineare.



Interpolazione di Lagrange:

$$f(x) \approx r(x) = f(a) + \frac{f(b) - f(a)}{b-a}(x-a)$$
$$f(x) = r(x) + \omega(x-a)(x-b)$$

↑ ↓ ↓

Valore vero = valore approssimato + errore



walter.balzano@gmail.com

Metodo di KRIGING (stima globale)

- Prende il nome dall'ingegnere minerario sudafricano **D.G. Krige** che sviluppò una tecnica per ottenere una mappa dei filoni d'oro a partire da dati campionari
- Il processo di interpolazione di KRIGING è costituito da due fasi distinte:
 - Analisi statistica delle **TENDENZE DI FONDO** della distribuzione dei punti campionari attraverso la costruzione del **VARIOGRAMMA**
 - **Interpolazione** vera e propria guidata dal **modello** della regione



walter.balzano@gmail.com

Variogramma

- **Definizione:**
- Il Variogramma descrive in modo statistico l'andamento della variabile di interesse sul territorio misurando il grado di cambiamento dei dati campionari nello spazio
- Si basa su di un concetto di autocorrelazione spaziale dei dati ed è in grado di tenere conto di eventuali andamenti non lineari del flusso del valore (direzioni preferenziali, filoni, ecc)
- Tecnicamente è un modello statistico della morfologia dello spazio costruito sommando tre componenti:
 - Una componente strutturale *drift* (corrente, flusso) che rappresenta una costante di trend della regione
 - Una componente casuale correlata spazialmente
 - Un errore casuale che rappresenta l'errore residuo

• Costruzione:

- È una funzione che interpola la varianza dei valori osservati in gruppi di coppie di punti a determinate distanze

- Il variogramma è definito come:

$$\gamma(h) = \sum_{i=1}^{n(h)} \frac{(z(x+h) - z(x))^2}{n(h)}$$

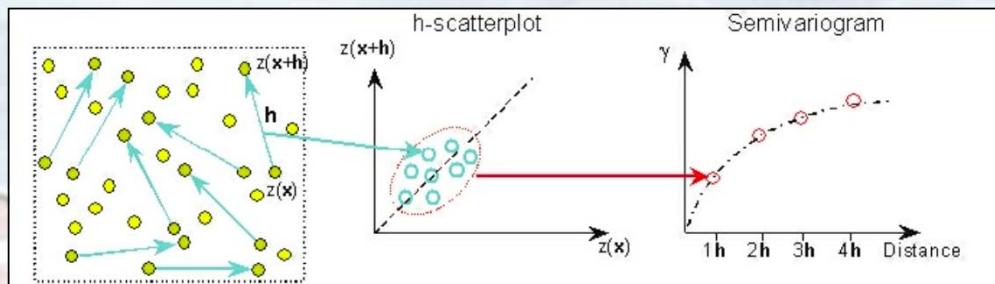
in cui:

z = valore della misura nel relativo punto

h = classe di distanza tra punti di misura

$n(h)$ = conteggio del numero di coppie di osservazioni effettuate alla distanza h

• Esempio:

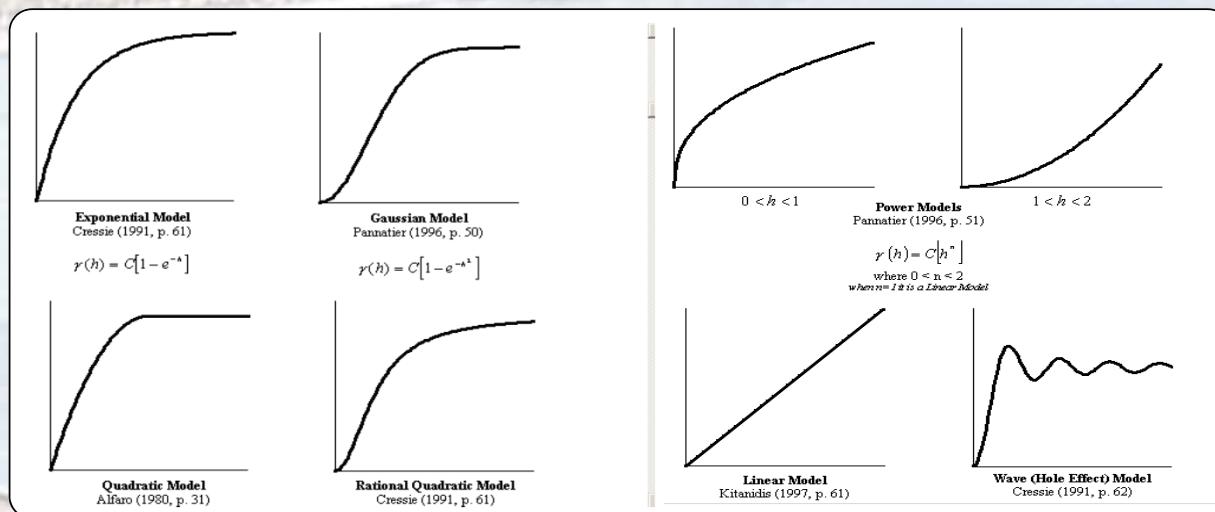


walter.balzano@gmail.com

Interpolazione del variogramma

Una volta costruito il variogramma empirico si cerca il modello che meglio lo approssima tra i modelli standard

In questo modo otteniamo una legge che descrive l'andamento della nostra variabile sul territorio a scala globale (in tutta la zona coperta dai campioni)



walter.balzano@gmail.com

GPS

- Introduzione
- Struttura
- Funzionamento
- Coordinate spaziali
- Errori di misurazione
- Sincronizzazione
- A-GPS e D-GPS
- Riferimenti Spazio/Tempo
- Logger / Tracker
- FingerPrint
- Sperimentazione



walter.balzano@gmail.com

Introduzione

- 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

Sistemi GPS

USA



NAVSTAR

RUSSIA



GLONASS

EUROPA



GALILEO

CINA



BEIDOU

INDIA



IRNSS



walter.balzano@gmail.com

NAVSTAR

Navstar, (**N**avigation **S**ystem with **T**iming **A**nd **R**anging **G**lobal **P**ositioning **S**ystem) 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.
-
- Le applicazioni del sistema GPS non sono limitate al campo militare, ma sono disponibili a tutti anche per uso civile, seppure con qualche limitazione nella precisione ottenibile nelle misure. La **disponibilità del segnale GPS 24 ore su 24 in ogni angolo del globo** e la progressiva riduzione dei costi dei ricevitori hanno trasformato il sistema GPS addirittura in un fenomeno di massa.
 - Inizialmente, per motivi di sicurezza, il sistema era affetto da un errore indotto ma poi reso libero



walter.balzano@gmail.com

Struttura NAVSTAR

- La costellazione di 24 satelliti che costituisce il sistema GPS fu completata nel 1993, prevede 21 satelliti operativi e 3 satelliti di scorta pronti a intervenire in caso di guasto.
- 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.
- Il centro di controllo del sistema GPS si trova nei pressi di Colorado Springs ed ha il compito di eseguire tutte le misure necessarie per "correggere" le informazioni inviate dai satelliti GPS.



walter.balzano@gmail.com

Sistema GPS

Il sistema GPS si compone di 3 parti:

- **Il Segmento spaziale**
(24 satelliti)
- **Il Segmento di controllo**
(5 centri di controllo)
- **Il Segmento d'utilizzo**
(i ricevitori)



walter.balzano@gmail.com

Il segmento spaziale

- I 24 satelliti Sono disposti su 6 piani orbitali diversi ma ugualmente spaziati tra di loro, inclinati di 55° rispetto al piano equatoriale (quindi non coprono le zone polari) a forma di ellissi a bassa eccentricità.
- Ogni piano orbitale ha 3 o 4 satelliti, e i piani sono disposti in modo tale che ogni utilizzatore sulla terra possa ricevere i segnali di almeno 5 satelliti. La loro quota è di 20 200 km e compiono due orbite complete in un giorno siderale.
- Da un punto del globo terrestre il ricevitore riesce a vedere solo la metà di essi, quindi 12. Ma non li vedrà mai tutti e 12 per via della loro inclinazione rispetto all'equatore. In più il ricevitore GPS stesso fa una discriminazione dei satelliti in base alla loro geometria e alla stima degli errori su ciascuno privilegiando quelli che forniscono maggior precisione (basandosi, per esempio, sulla loro elevazione rispetto l'orizzonte).
- Ogni satellite è dotato di OROLOGI ATOMICI basati su 4 oscillatori ad altissima precisione, di cui 2 al cesio e 2 al rubidio con una precisione di $3 \cdot 10^{-9}$ secondi (possibilità di errore è di un secondo ogni 30000 anni. Il Costo di ogni orologio atomico è di circa 160.000 Euro); ha dei razzi per effettuare le correzioni di orbita. Ha due pannelli solari di area pari a 7,25 m² per la produzione di energia. Ha infine batterie di emergenza per garantire l'apporto energetico nei periodi in cui il sole è eclissato. Pesa circa 845 kg ed ha una vita di progetto di 7,5 anni.

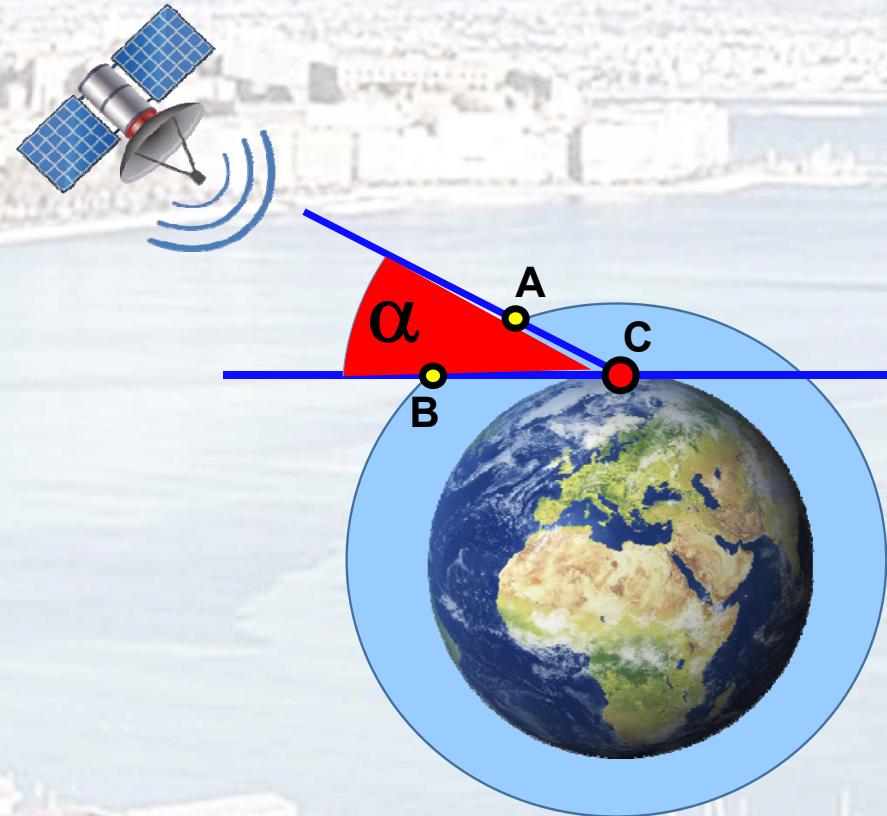


walter.balzano@gmail.com

Il segmento spaziale (2)

Elevazione dei satelliti:

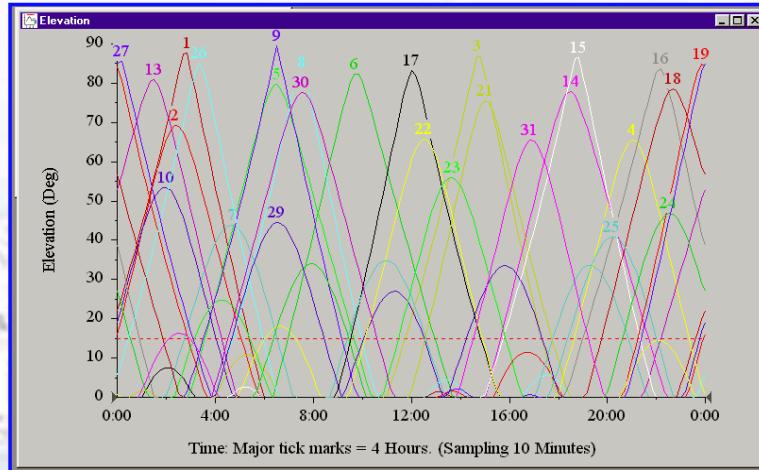
Angolo α individuato fra la direzione del segnale ed il piano tangente all'ellissoide nel punto occupato dal ricevitore. In genere si utilizzano solo segnali provenienti da satelliti con $\alpha > 10^\circ$ o 15° , in modo da attenuare i problemi legati al disturbo atmosferico



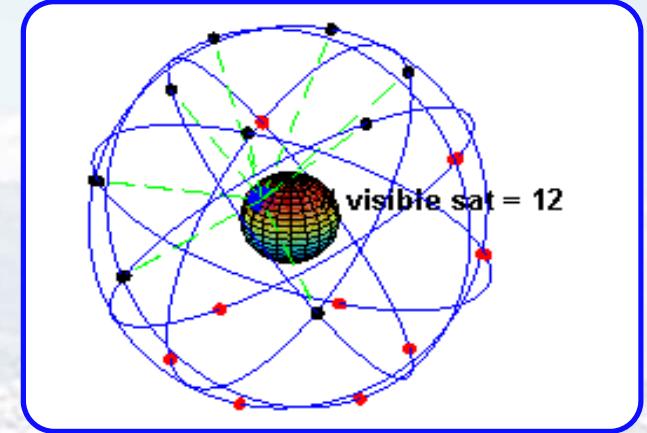
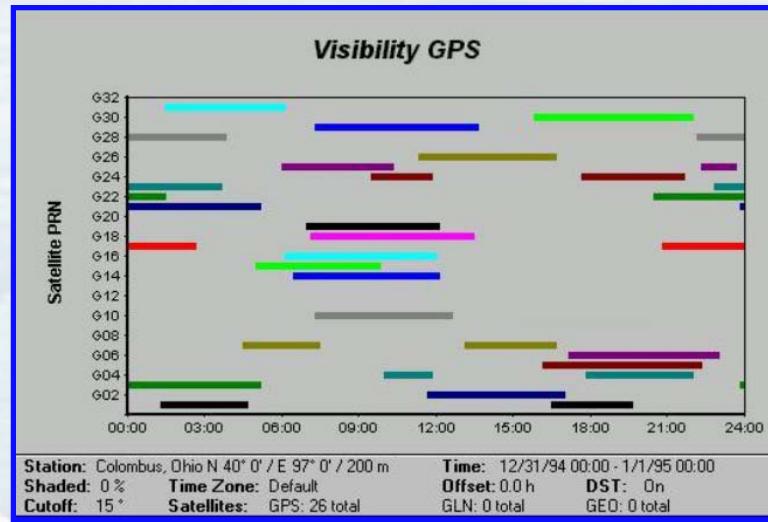
walter.balzano@gmail.com

Il segmento spaziale (2)

Tabella tipica
delle elevazioni
su 24 ore



Visibilità dei
Satelliti da un
punto fisso
della terra.
Ciascuno è
visibile per
circa 4 ore



walter.balzano@gmail.com

Vista da dispositivi



walter.balzano@gmail.com

Segmento di controllo

Il Segmento di controllo è costituito da un gruppo di stazioni di monitoraggio a terra che hanno la loro **sede centrale in Colorado**: in particolare, i centri di controllo terrestre sono 5 e sono disposti in posizione pressoché equatoriale intorno al globo. I loro nomi sono:

1. Colorado Springs,
 - Raccoglie i dati delle altre stazioni
 - Compensazione dei dati →
 - errori degli orologi sui satelliti,
 - effemeridi satelliti,
 - scostamento delle orbite
 - Invia ai satelliti dati di correzione per Orologi, Orbite, Effemeridi
2. Hawaii,
3. Ascension Island,
4. Diego Garcia,
5. Kwajalein



walter.balzano@gmail.com

Segmento di controllo

Le stazioni a terra hanno il compito di:

- "tracciare", ovvero **seguire in maniera continua i satelliti** ed elaborare i dati ricevuti per calcolarne la posizione spazio-temporale (effemeridi);
- **imporre correzioni d'orbita;**
- **sincronizzare gli orologi atomici a bordo dei satelliti** usando un orologio Maser all'idrogeno
- **scaricare i dati per la trasmissione tramite i satelliti** (memorizzare nuovi dati sui satelliti: fra i più importanti ci sono le "effemeridi" previste per le successive 12 o 24 ore, che vengono trasmesse agli utenti).



Il Segmento d'utilizzo

- **Il segmento d'utilizzo è definito da ogni ricevitore**, come quelli a bordo delle automobili, degli aerei e delle navi. 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**.
- La precisione, a seconda della applicazione, può variare da **±100 m** a **±1 cm**



walter.balzano@gmail.com

Interazione dei Sottosistemi GPS

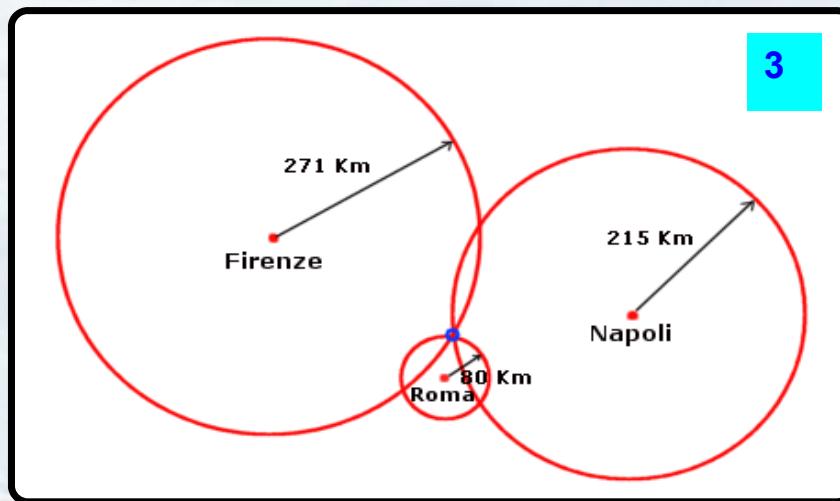
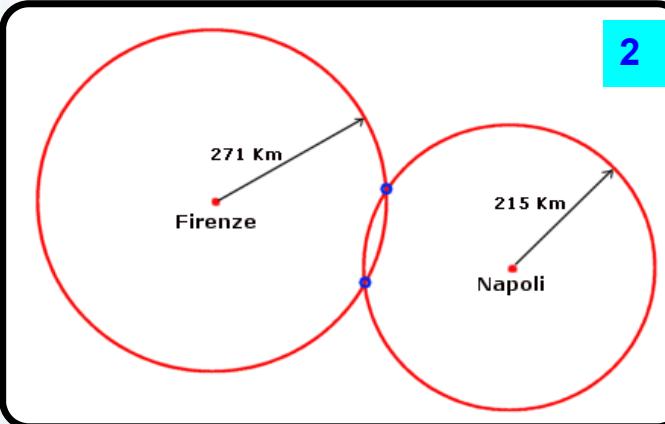
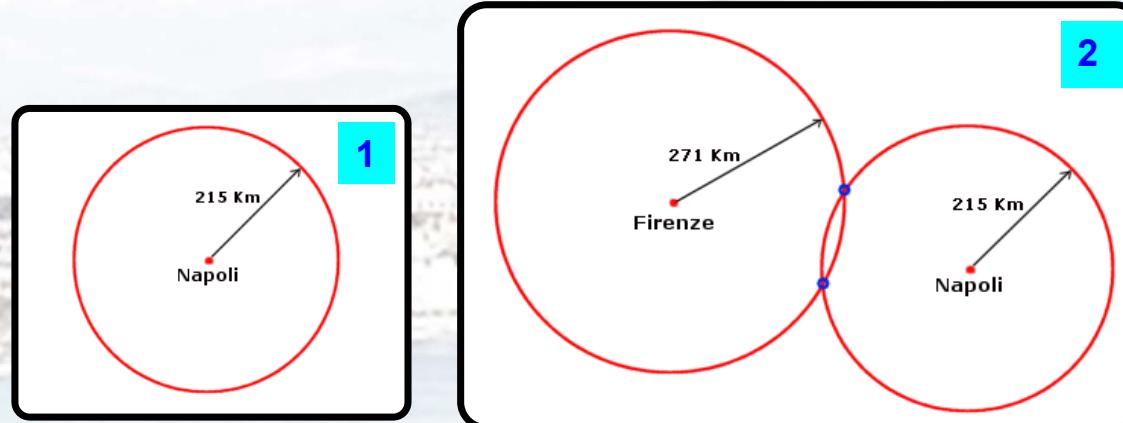


walter.balzano@gmail.com

Funzionamento GPS: la trilaterazione

La **trilaterazione** è il metodo usato per il calcolo effettivo della posizione.

Esempio su un spazio bidimensionale (quello reale lavora sullo spazio 3D ed usa lo stesso concetto).



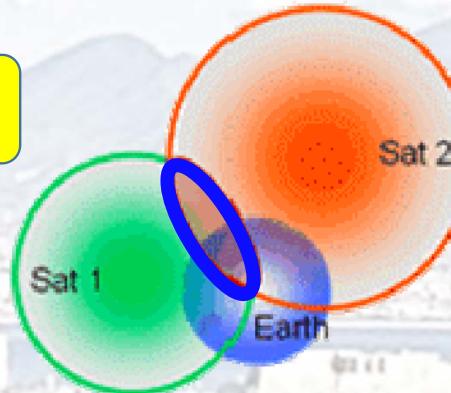
walter.balzano@gmail.com

Funzionamento GPS: la trilaterazione (2) – 3D

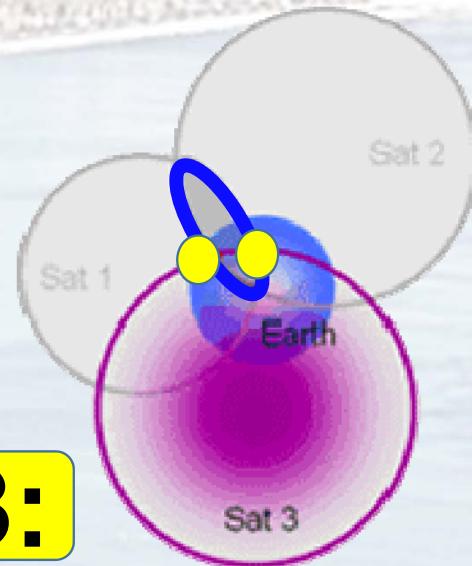
1:



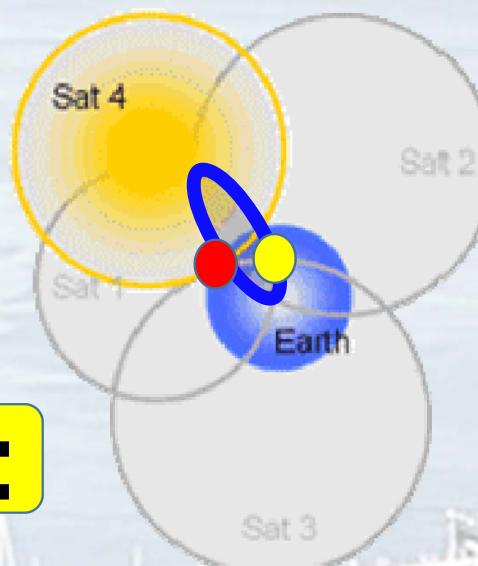
2:



3:



4:

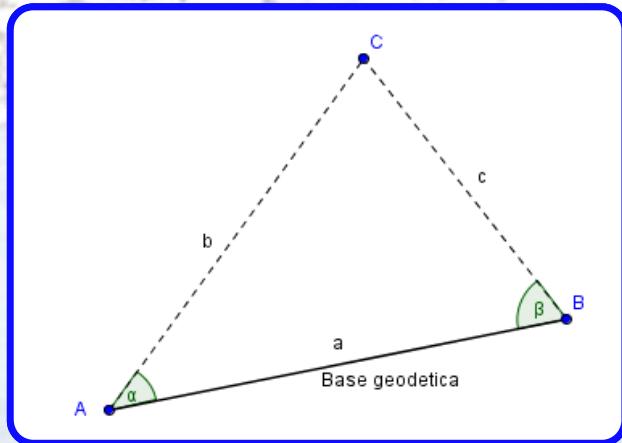


walter.balzano@gmail.com

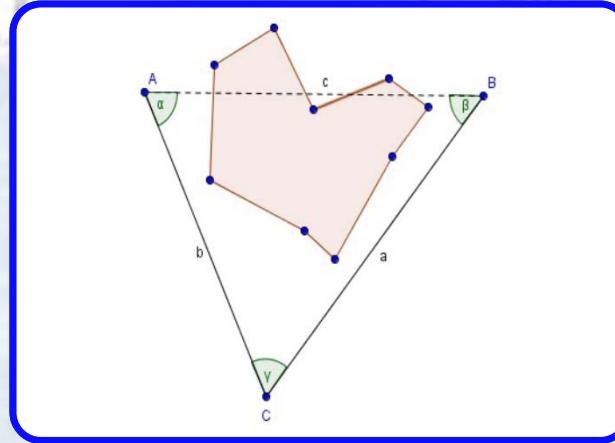
Trilaterazione vs Triangolazione

- La **Triangolazione**, in Geodesia e Topografia, è quel procedimento che permette la determinazione indiretta di distanze tra punti del terreno e, quindi, la determinazione delle loro coordinate geografiche. Per effettuare una triangolazione si scelgono tre punti opportuni sul terreno, considerati vertici di un triangolo, **uno di questi lati viene misurato direttamente** ed è detto base geodetica misurata, la misura va effettuata con altissima precisione. Da ognuno degli estremi della base **si misurano gli angoli** sotto i quali viene visto l'altro punto e con semplici calcoli trigonometrici se ne determina la distanza e, quindi, la posizione.

ESEMPIO: metodo delle parallasse per il calcolo della distanza delle stelle



Triangolazione



Triangolazione per il calcolo della distanza fra due punti accessibili ma non visibili l'uno all'altro

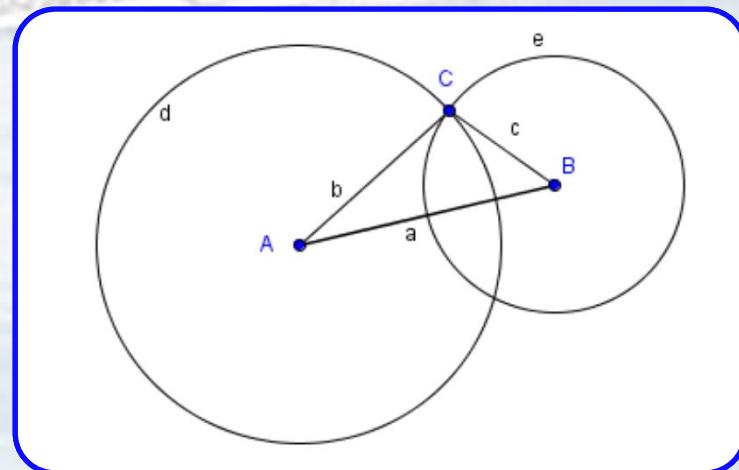


walter.balzano@gmail.com

Trilaterazione vs Triangolazione

La **Trilaterazione** è quel procedimento che consente la determinazione della posizione di un punto **in base a misure di distanze da altri di coordinate note, piuttosto che misure di angoli**. Un tempo era difficile ottenere precise misure di distanze e per questo si preferivano metodi di triangolazione, ma oggi con i moderni strumenti ad emissione elettromagnetica si è superato questo problema e la trilaterazione si affianca spesso o sostituisce la triangolazione.

ESEMPIO: metodo impiegato dai rilevamenti satellitari (GPS).



Trilaterazione



walter.balzano@gmail.com

Funzionamento GPS (3)

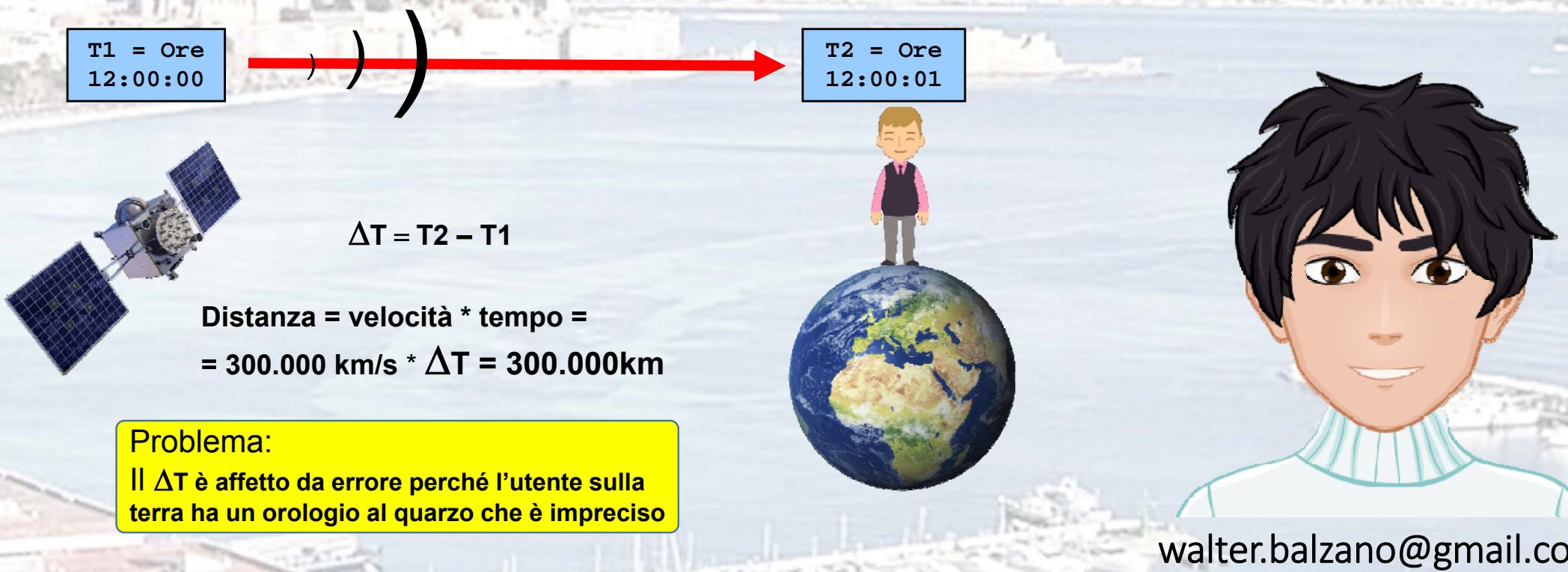
- Per capire la distanza tra il receiver ed il satellite viene misurato il tempo che un segnale impiega per arrivare a terra.
- Il calcolo matematico è semplice ma il livello di precisione dipende fortemente dalla accuratezza della lettura degli orologi (un solo millesimo di secondo di differenza potrebbe penalizzare la rilevazione con un errore nell'ordine dei 300 Km!)
- Considerando che un ricevitore non può montare orologi atomici da 160.000 Euro, si è pensato di usare orologi capaci di mantenere un'estrema precisione per brevi periodi che però nel tempo vanno spesso corretti sfruttando direttamente i segnali dei satelliti.



Funzionamento GPS (3)

Esempio:

- Ad un ora prestabilita (supponiamo le 12:00) il satellite genera un codice (detto pseudo random code) e lo invia sulla terra.
- Sempre alle 12:00 anche il receiver GPS genera lo stesso identico codice per cui, quando il segnale dal satellite arriva a terra e viene letto dal receiver, questo lo riconosce ed è in grado di misurare quanto tempo ha impiegato il segnale per arrivare.
- Moltiplicando il tempo per la velocità della luce (300.000 km/s) si ottiene la distanza tra il satellite ed il receiver GPS.



Funzionamento GPS (4)

- Il funzionamento del sistema GPS si basa sulla **misura del tempo di percorrenza del segnale dall'antenna del satellite all'antenna del ricevitore**
- Ogni satellite ha il suo nome (un numero da 1 a 32) ed invia un messaggio codificato tramite il suo nome.

MESSAGGIO:

- **Formato:** Il messaggio viene trasmesso con continuità e contemporaneamente su 2 frequenze (1,2 e 1,5 GHz) in modo da eliminare l'errore dovuto alla rifrazione atmosferica.
- **Contenuto:** Il messaggio inviato da ogni satellite contiene:
 - l'almanacco (parametri orbitali approssimati) dell'intera costellazione
 - le effemeridi relative a se stesso
 - Dati relativi al satellite (orario, stato del sistema,...)



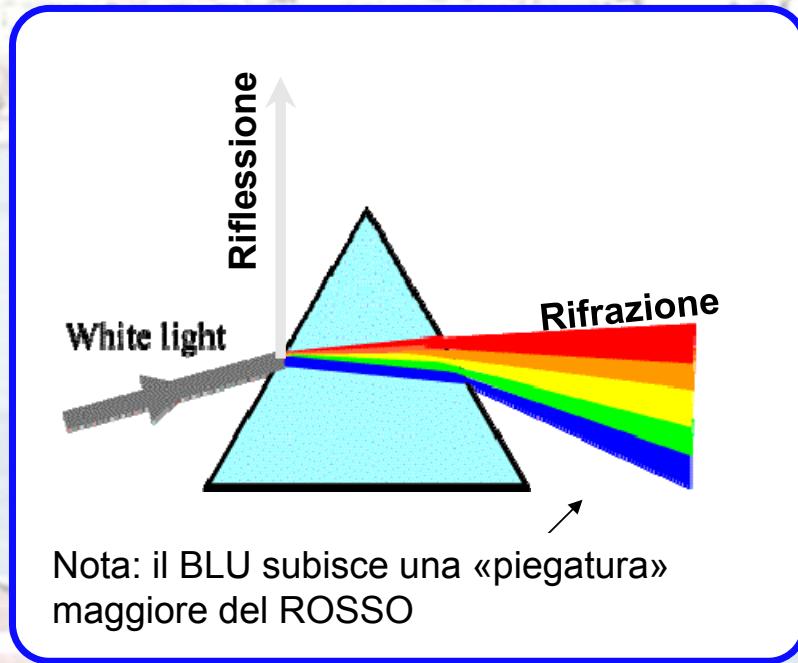
walter.balzano@gmail.com

La Rifrazione e Riflessione

La rifrazione avviene ogni qualvolta la luce attraversa uno spazio con differente densità nel quale cambia la sua velocità di propagazione e le “piegature” dipendono dal valore della frequenza considerata.

L'angolo di rifrazione dipende da:

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

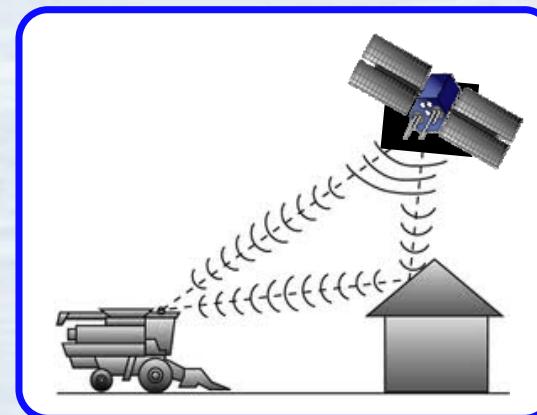
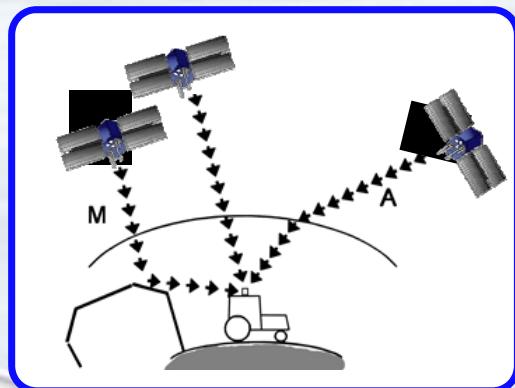
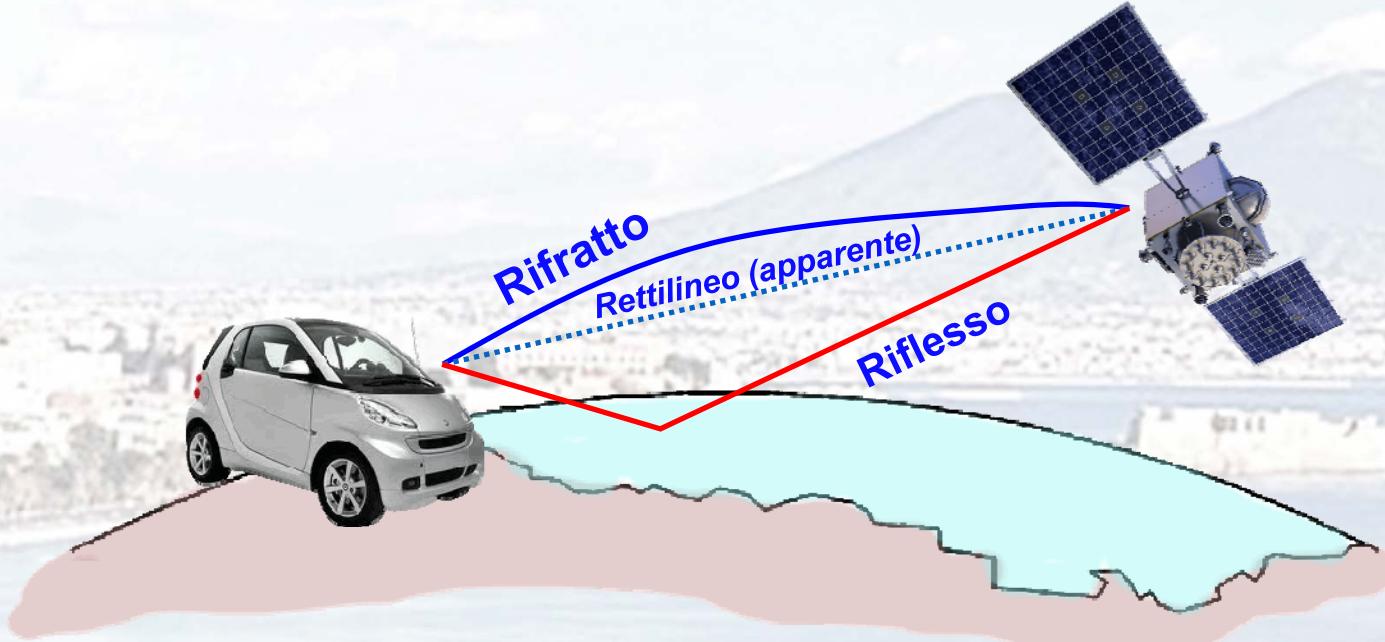


Nota: il BLU subisce una «piegatura» maggiore del ROSSO



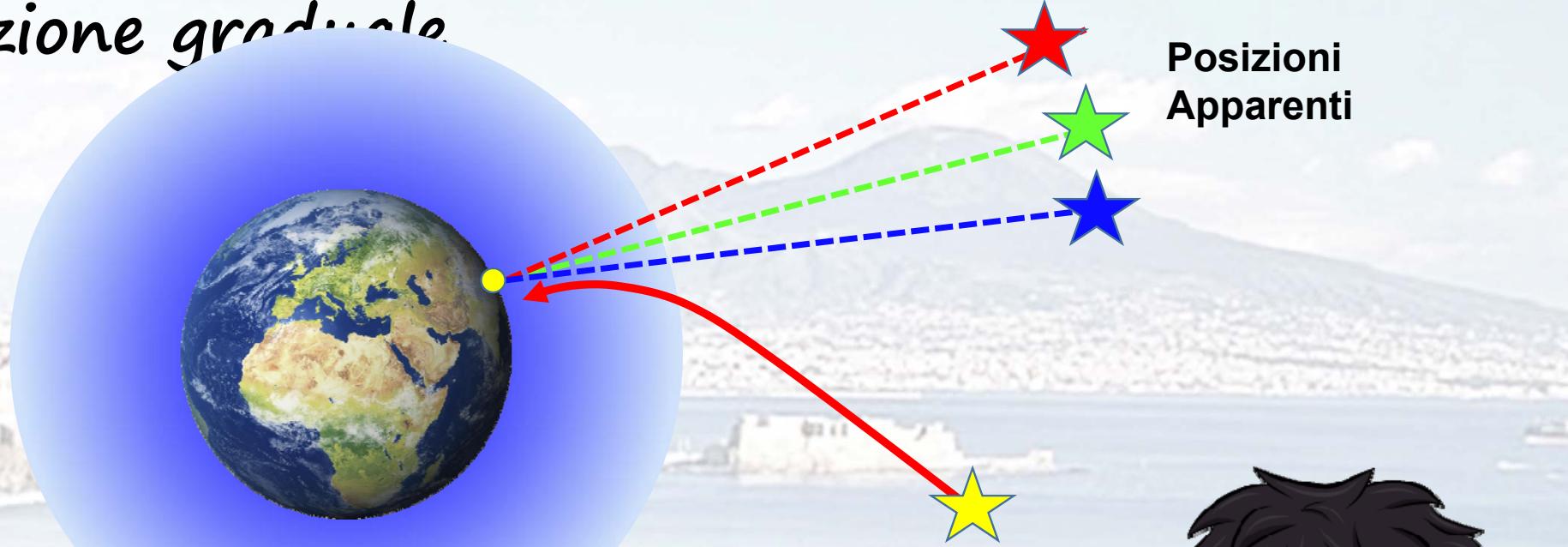
walter.balzano@gmail.com

La Rifrazione e Riflessione nel GPS



walter.balzano@gmail.com

Rifrazione graduale



- A causa della gravità, l'atmosfera terrestre è più densa alla superficie che ad alta quota e quindi la "piegatura" che subisce il segnale che attraversa l'atmosfera varia da un minimo (alta quota) ad un massimo (bassa quota).

Un raggio di luce di una stella che arriva sulla terra subisce una deviazione ed i colori di cui il raggio si compone vengono rifratti con angoli diversi. Ciò condurrebbe a percepire una stella non come un puntino luminoso bensì come un piccolo segmento tipo arcobaleno. Nella pratica tali differenze di rifrazione non sono distinguibili dall'occhio umano per cui continuiamo a percepire la loro mescolanza (cioè vediamo la stella bianca).

walter.balzano@gmail.com

Funzionamento GPS (4)

L1 = 1.57542 GHz
L2 = 1.2276 GHz

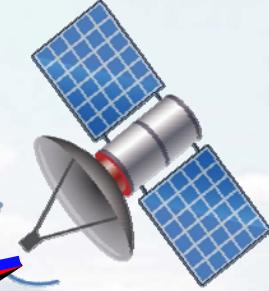
Rifrazione

L1

L2

Rettilineo apparente

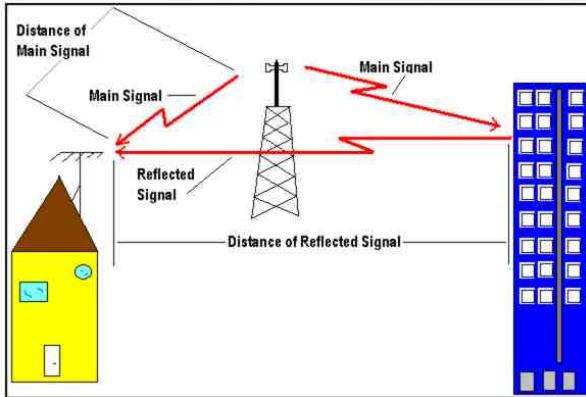
ionosfera



walter.balzano@gmail.com

Esempi di Riflessione nel mondo reale

Il fenomeno della rifrazione di segnali è riscontrabile in molte circostanze del mondo reale



TV analogica:
**Sdoppiamento
dell'immagine
(ghosting)** in
presenza di
superfici riflettenti
(mare, aeroporti,...)



walter.balzano@gmail.com

Le effemeridi

Le effemeridi (dalla parola greca *ephemeros* = giornaliero) sono tabelle che contengono un insieme di parametri sintetici necessari e sufficienti per calcolare la posizione del satellite. Nel caso dei satelliti GPS occorre distinguere:

- **effemeridi trasmesse** (trasmesse dal satellite);
- **effemeridi precise**: calcolate a posteriori da diverse organizzazioni governative (ad es. NIMA) e di ricerca (ad es. IGS) e vengono poi distribuite via web da diversi enti.
 - effemeridi **rapide**: disponibili con un ritardo di 1 giorno;
 - effemeridi **finali**: disponibili con un ritardo di 14 giorni.

Ephemeris Data Set Used in Pseudo-Range Navigation Example (GPS Time = 150000 seconds)				
Ephemeris Data Parameter	Value	Value	Value	Value
SV	15	27	31	7
Issue of Data Ephemeris	196	200	125	125
Cosine Correction to Inclination	-9.313225746E-08	1.136213541E-07	2.793967724E-08	-1.285225153E-07
Sine Correction to Inclination	-3.725290298E-09	-1.061707735E-07	9.126961231E-08	-1.322478056E-07
Cosine Correction to Radius	146.09375	148.84375	306.28125	322
Sine Correction to Radius	-69.9375	79.09375	-130.71875	-128.5
Cosine Correction to Latitude	-3.630295396E-06	4.122033715E-06	-6.921589375E-06	-6.720423698E-06
Sine Correction to Latitude	1.228414476E-05	1.15185976E-05	3.74391675E-06	2.983957529E-06
Mean Motion Difference	4.023024718E-09	4.513045129E-09	4.656622538E-09	4.650550857E-09
Eccentricity	0.006778693292	0.01127019501	0.005836840719	0.006999379606
Rate of Inclination Angle	1.817932867E-10	-5.928818388E-11	-5.418082828E-10	-4.207318109E-10
Orbital Inclination	0.9721164968	0.9459886628	0.9633626261	0.963950905
Mean Anomaly at Reference Time	-0.8856059028	0.1225249	-0.6775731485	3.019737078
Argument of Perigee	1.738558535	2.601538834	0.6715504011	-2.568758665
Rate of Right Ascension	-7.783538501E-09	-8.143553497E-09	-8.411421798E-09	-8.25355808E-09
Longitude of Ascending Node	-2.8654714	0.2200327977	2.320031302	2.317137898
Square Root of Semi-Major Axis	5153.618444	5153.653282	5153.789852	5153.644896
Reference Time Ephemeris	151200	151200	136800	151200



walter.balzano@gmail.com

Fonti di errore nella determinazione della posizione

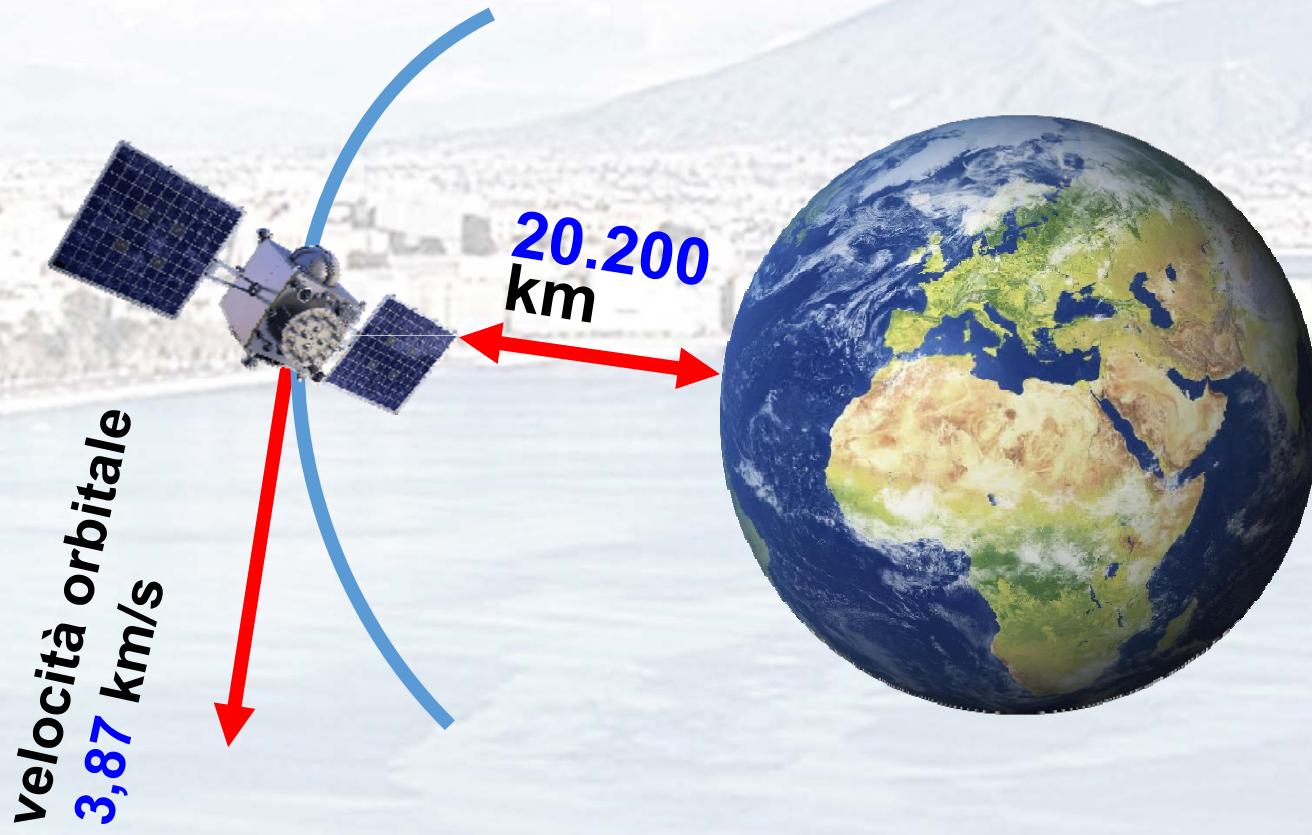
- Per ottenere una sincronizzazione perfetta, ogni orologio a bordo dei satelliti deve essere sincronizzato con tutti gli altri e con quelli sulla Terra
- Tuttavia occorre considerare alcuni fattori che costituiscono fonte di errore e sono di diversa natura:
 - **Fattori atmosferici** (ritardo da 1 ns 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
 - tempo di elaborazione nei computer del GPS
 - **Fattori RELATIVISTICI** (Anticipo/Ritardo del tempo) degli stessi orologi atomici

NOTA: Gli errori relativistici sono di quasi tre ordini di grandezza più grandi rispetto alle altre sorgenti di ritardo. La loro correzione è quindi indispensabile per il funzionamento dei sistemi di navigazione satellitare.



walter.balzano@gmail.com

Combinazione errori: Velocità + Gravità

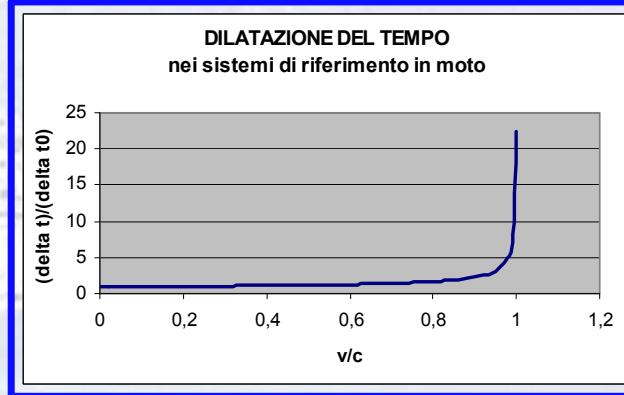


walter.balzano@gmail.com

Errori relativistici

Secondo la teoria della Relatività, il tempo proprio di ogni corpo in movimento viene deformato secondo questa relazione:

$$\Delta t = \frac{\Delta t_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$



In cui Δt_0 è il tempo proprio, cioè la differenza tra i tempi di due eventi che accadono nello stesso luogo.

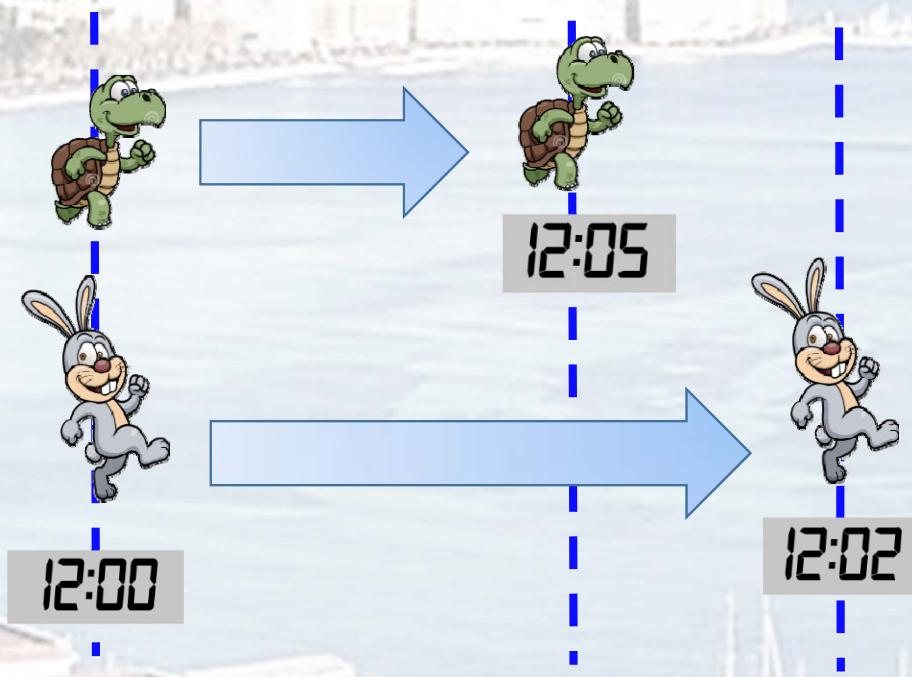
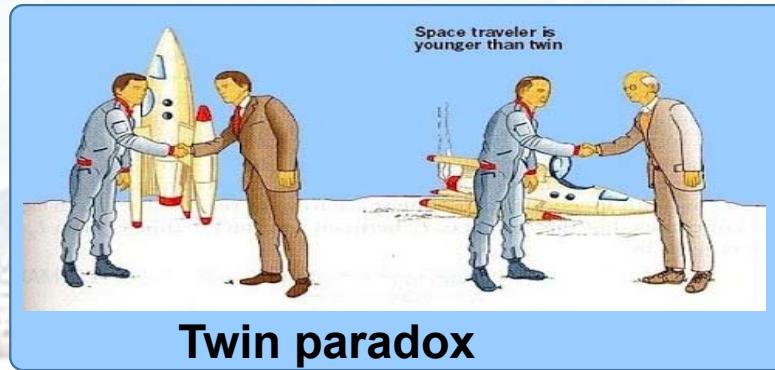
Come possiamo vedere nel grafico, maggiore è la velocità del sistema di riferimento, maggiore è l'intervallo di tempo Δt , cioè l'orologio rallenta fino a fermarsi quando $v=c$.

Esempio: Ci sono due orologi sincronizzati sulla superficie della Terra; il primo orologio rimane sul nostro pianeta ed il secondo è lanciato nella profondità dello spazio a bordo di un'astronave che viaggia vicina alla velocità della luce ($c = 300000 \text{ Km/sec}$). Quando torna a casa, comparando i due orologi si può osservare che l'orologio che ha viaggiato è in ritardo rispetto a quello che è rimasto sulla Terra!



walter.balzano@gmail.com

Relatività: tempo e velocità

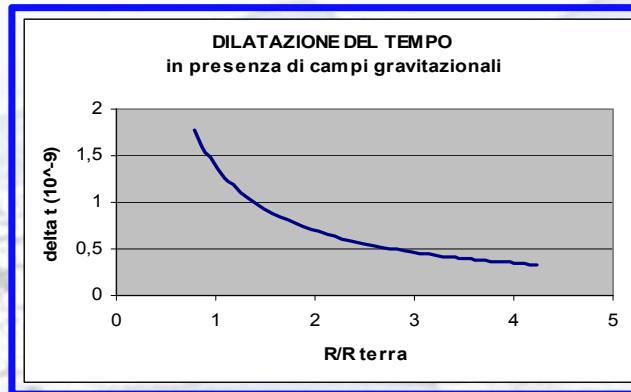


walter.balzano@gmail.com

Errori relativistici (2)

Allo stesso modo qualsiasi corpo soggetto alla forza di gravità ha una deformazione del tempo secondo questa relazione:

$$\Delta t = \frac{\Delta t_0}{1 + \frac{2\Phi}{c^2}}$$



In cui $\Phi = -GM/R$ indica il potenziale gravitazionale scalare pseudo newtoniano con G che è la costante di gravitazione universale, M la massa del corpo generatore del campo gravitazionale e R la distanza radiale non euclidea fra i corpi (cioè soggetta alla legge relativistica dell'allungamento del raggio: $DR = GM / 3c^2$). Δt_0 è ancora il tempo proprio ossia l'intervallo di tempo rispetto ad un sistema di riferimento in cui non vi sono campi gravitazionali.

Consegue che: gli orologi immersi in campi gravitazionali intensi rallentano la loro misurazione del tempo rispetto agli orologi soggetti a campi gravitazionali minori o nulli.

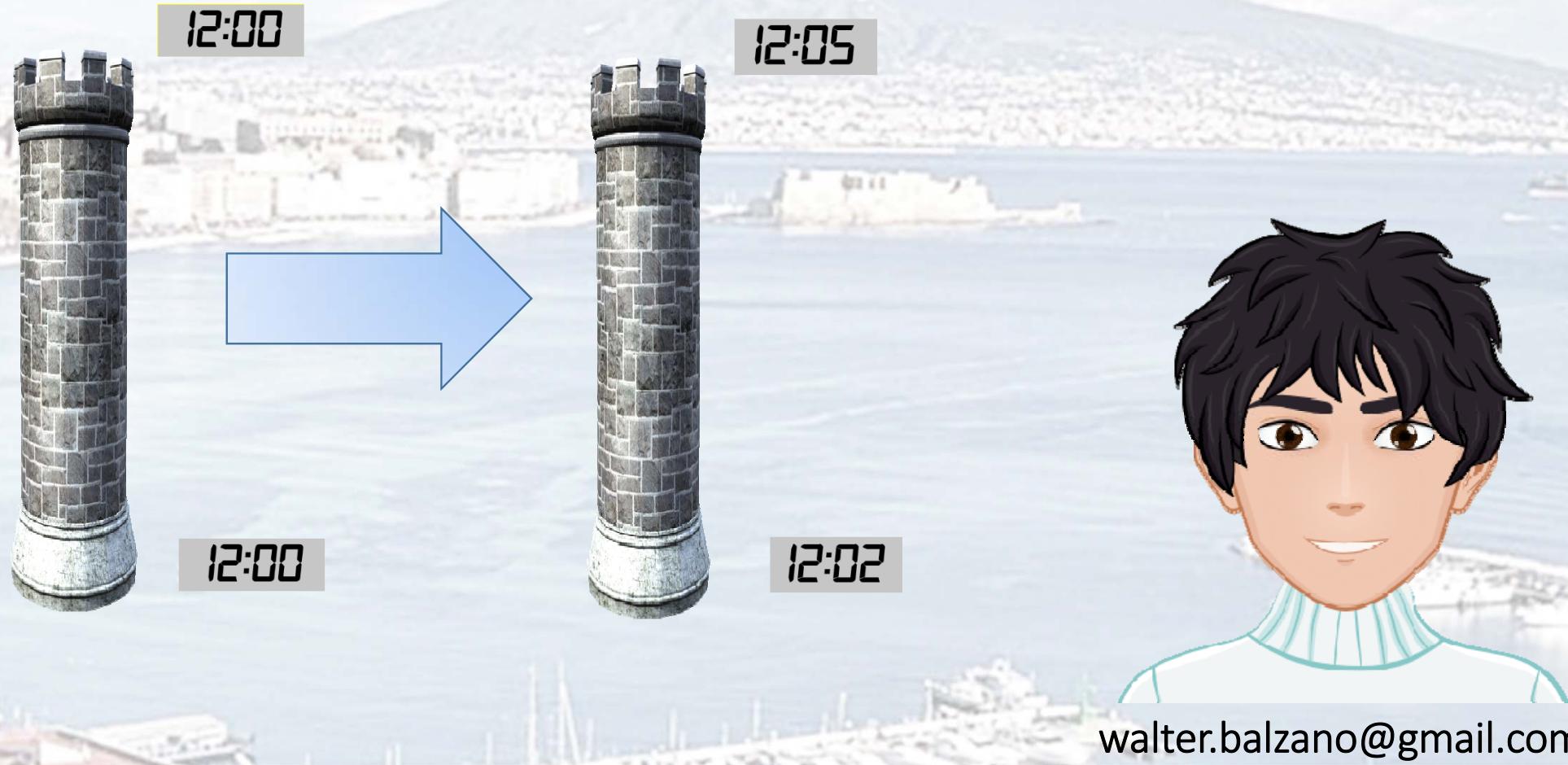
Esempio: utilizziamo gli stessi due orologi sincronizzati sulla superficie della Terra.

Questa volta mandiamo un orologio in una stazione spaziale orbitante intorno alla Terra, dove la gravità è minore che sulla superficie terrestre. Quando torna indietro possiamo osservare che: L'orologio che è stato spedito nella stazione spaziale è in anticipo rispetto a quello che è rimasto sulla superficie terrestre!!!



walter.balzano@gmail.com

Relatività: tempo e gravità



walter.balzano@gmail.com

Effetti relativistici sul GPS

Quanto sono importanti questi effetti per il corretto funzionamento degli orologi a bordo dei satelliti del GPS?

- **Effetto dovuto alla velocità:** la formula $t = t_0(1 - (v/c)^2)^{-1/2}$ applicata ad un satellite la cui velocità orbitale è 3,87 km/s mostra un **ritardo di 7 μ -sec al giorno** (la velocità, ricordiamo, dilata il tempo di un corpo relativamente ad uno più lento) rispetto agli orologi a terra.
- **Effetto dovuto alla quota:** l'influenza del campo gravitazionale terrestre sui satelliti è **quattro volte** minore. Applicando la formula: $t = t_0(1 + 2\Phi/c^2)^{-1}$, che lega il tempo relativo e la gravità, ne deduciamo che ogni satellite **“guadagna” 45 μ -sec al giorno** (μ -sec = microsecondi = $1 \cdot 10^{-6}$ secondi)

La somma algebrica delle variazioni dei tempi così ottenuti (7 μ -sec – 45 μ -sec) ci porta a dire che un satellite del GPS anticipa il tempo segnato dall'orologio sulla terra di **38 μ -sec al giorno**.

Nella localizzazione del ricevitore, che errore genera un anticipo di 38 μ -sec al giorno? Per calcolarlo basta moltiplicare questo tempo per la velocità della luce:

$$38 \mu s \times 3 \cdot 10^8 \frac{m}{s} = 11400 m = 11,4 km$$

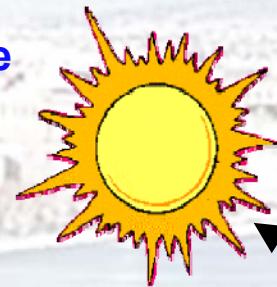


walter.balzano@gmail.com

Fonti di errori minori

Esistono numerose componenti fonti di errori che rendono complessa l'elaborazione anche perché tali fonti sono variabili nel tempo (basti pensare al diverso scenario gravitazionale della terra a seconda se si trovi in apogeo (più lontana dalla terra) o perigeo (più vicina alla terra))

Attrazione
gravitazione
SOLE



Pressione
radiazione solare
(pannelli solari)

Attrazione
gravitazione
LUNA

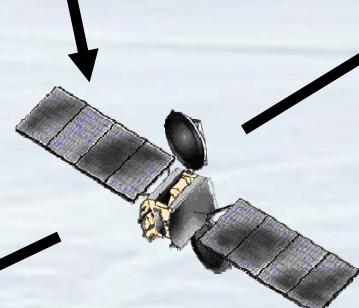


Attrazione terra

- Gravitazione centro terra
- Perturbazioni oceaniche
- maree

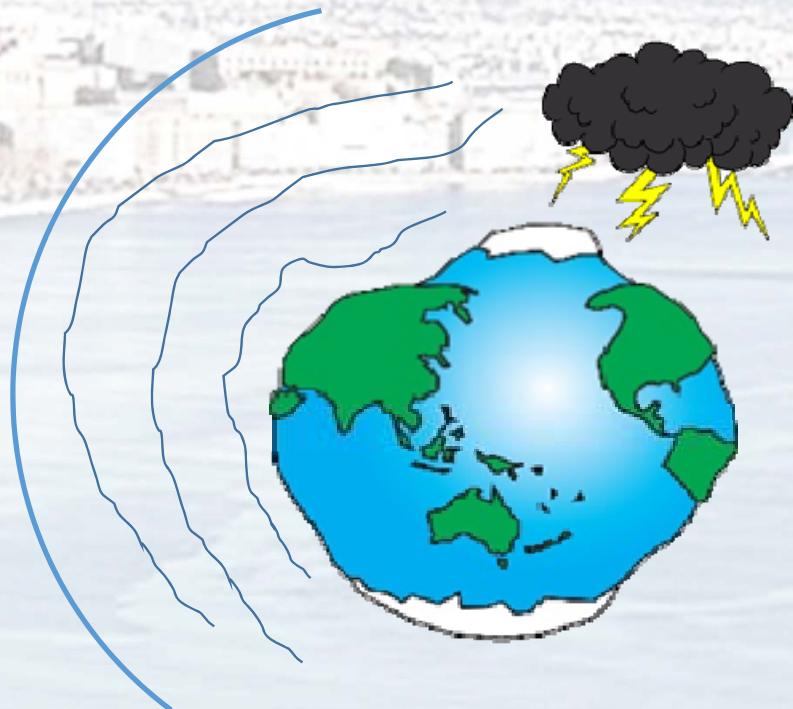


Attrazione
gravitazione
altri corpi celesti



Sorgenti di errori

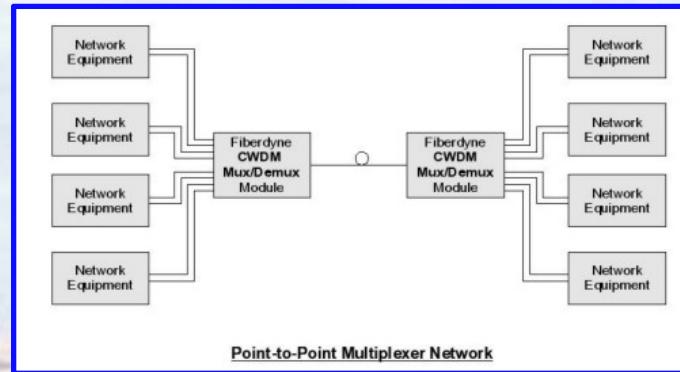
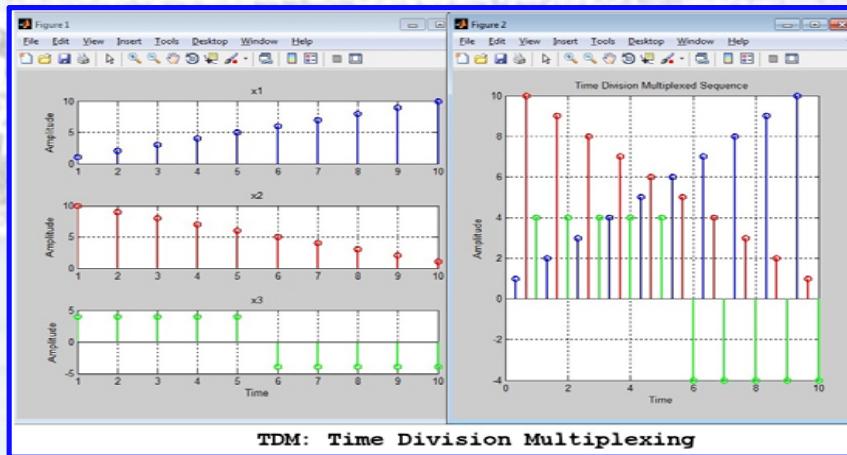
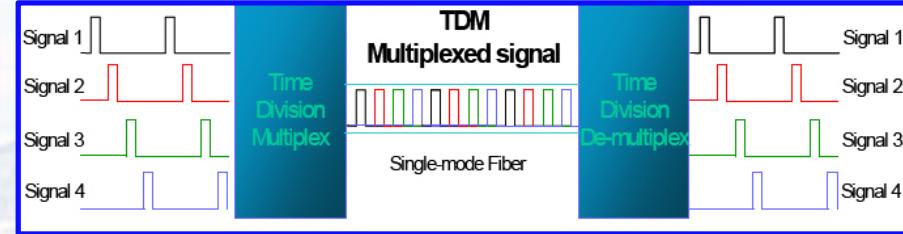
**Regolarità orbita satellite
proporzionata alla distanza dalla terra**



walter.balzano@gmail.com

Trasmissioni GPS ortogonali: Time Division Multiplexing

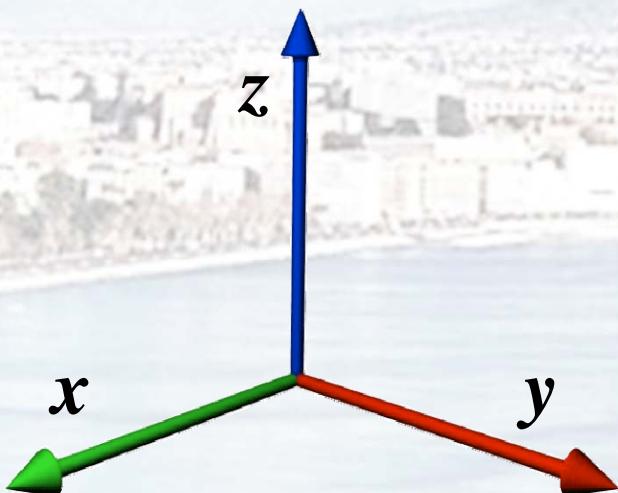
I segnali trasmessi dai satelliti sono tra loro ortogonali: per ogni coppia S1 ed S2 di segnali, il loro prodotto scalare è nullo



walter.balzano@gmail.com

Spazio/Tempo

Alcune questioni di Spazio / Tempo...



UTM-WGS84
Universal Transverse
Mercator World
Geodetic System 1984



UTC
Coordinated
Universal
Time



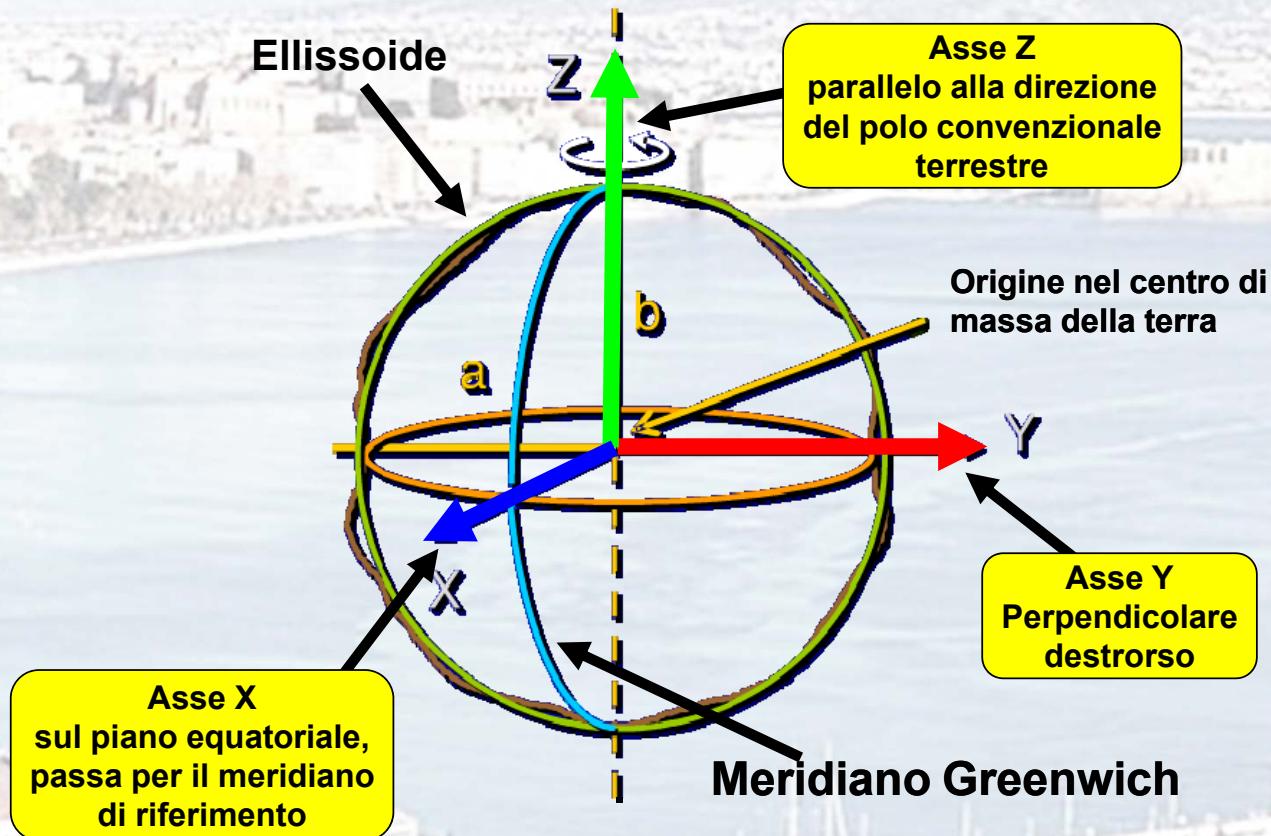
walter.balzano@gmail.com

Spazio: Sistema di Riferimento

WGS84 World Geodetic System 1984

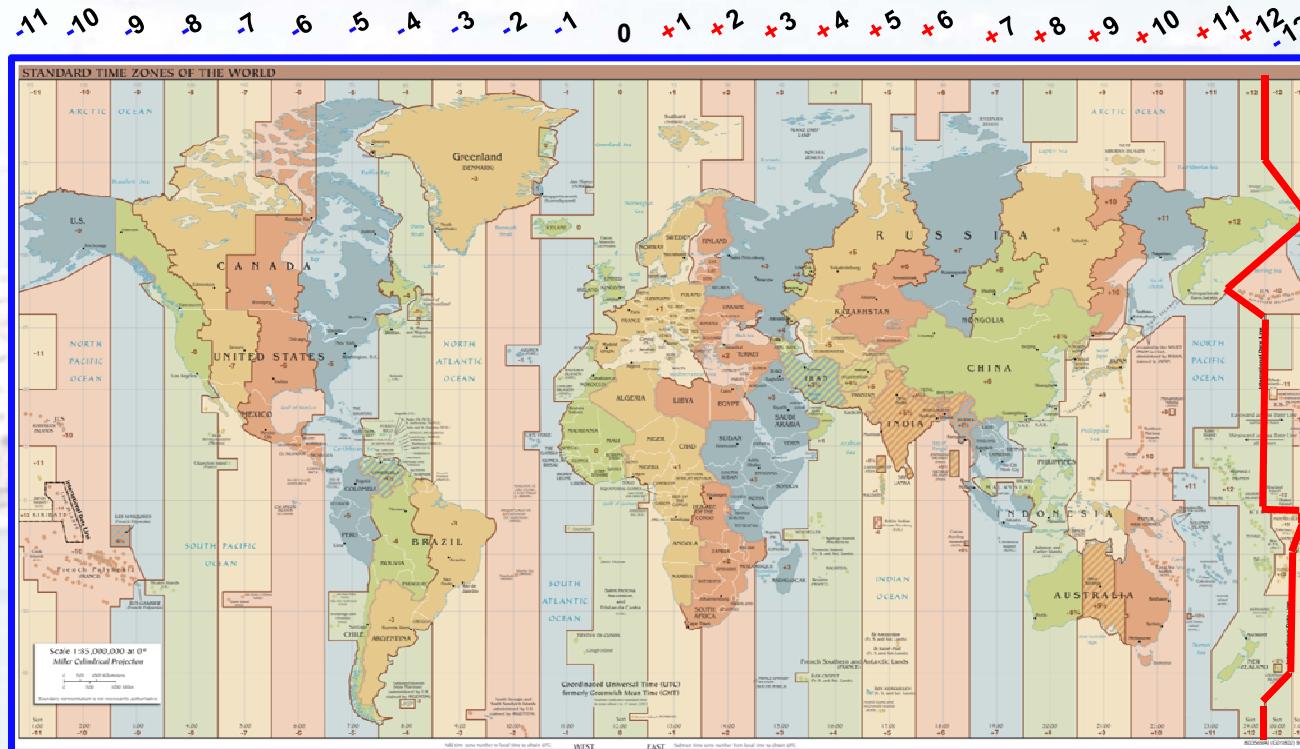
(sistema geodetico mondiale, riferito al 1984)

- La posizione del satellite è nota attraverso le effemeridi. Sono cioè note X, Y, Z del satellite nel sistema WGS84



walter.balzano@gmail.com

Time Zones



Un fuso orario è una regione del globo terrestre che osserva un orario standard uniforme per scopi legali , commerciali e sociali . I fusi orari tendono a seguire i confini dei paesi e le loro suddivisioni invece di seguire rigorosamente la longitudine, poiché è conveniente che le aree in stretta comunicazione commerciale o di altro tipo mantengano lo stesso orario.



walter.balzano@gmail.com

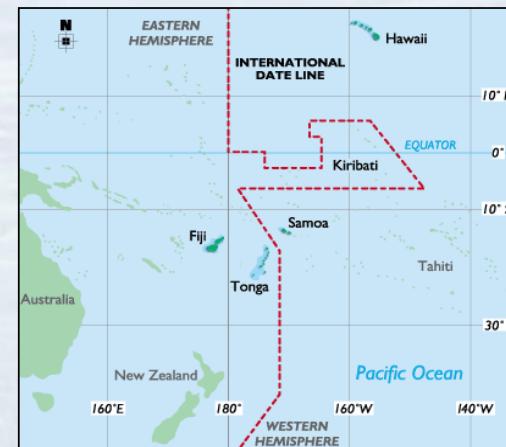
La Linea del cambiamento di data



TONGA = Giovedì 16 giugno 2011
ore 7:44:23

Samoa = Mercoledì 15 giugno 2011
ore 7:44:23

- **Subito a sinistra della “International Data Line” la data è sempre un giorno avanti rispetto alla data subito a destra.**
- **Tonga e Samoa hanno lo stesso orario ma hanno 1 giorno di differenza.**



walter.balzano@gmail.com

UTC (Tempo Coordinato Universale)

- **Motivazione:** il termine UTC è stato coniato per non dover menzionare una specifica località in uno standard internazionale.
- **Riferimento:** basa su misurazioni condotte da orologi atomici invece che su fenomeni celesti come il GMT. A causa delle oscillazioni nella velocità di rotazione della Terra (dovute alle influenze gravitazionali della Luna, del Sole e, in misura minore, degli altri pianeti), il GMT ritarda costantemente rispetto al "tempo atomico" UTC.
- **Rappresentazione:** L'UTC nei sistemi informatici tipicamente Unix memorizzano la data come il numero di secondi passati dal 1° gennaio 1970.

[https://it.wikipedia.org/wiki/Tempo_\(Unix\)](https://it.wikipedia.org/wiki/Tempo_(Unix))
<http://www.epochconverter.com/clock>



walter.balzano@gmail.com

UTC: Rappresentazione

EPOCA: **1º gennaio 1970**. Sulla maggior parte dei sistemi a 32 bit, il valore del dato `time.h` usato per questo calcolo è un numero intero a **32 bit di tipo signed**

Usando questo sistema, l'istante più lontano rappresentabile scoccherà alle ore:

03: 14: 07 del 19 gennai o 2038 (UTC).

Dopo questo momento, il contatore supererebbe il valore massimo, e verrebbe considerato come un numero negativo. I computer leggeranno la data non come 2038 ma come **1091** (precisamente, le 20:45:52 UTC di **venerdì 13 dicembre 1901**), causando errori di calcolo

03:14:07 del
19 gennaio 2038

+ 1 secondo

20:45:52 del
13 DICEMBRE 1901

Esempio:

Binary:	01111111 11111111 11111111 11111110
Decimal:	2.147.483.646
Date:	2038-01-19 03:14:06 (UTC)



walter.balzano@gmail.com

Rappresentazioni possibili di 32 bit signed

-2.147.483.648

Data MINIMA :

20:45:52 UTC
di venerdì
13 dicembre 1901

0

EPOCH:

00:00:00 UTC
di giovedì
1 gennaio 1970

1.000.000.000

1.111.111.111

.....

1.234.567.890

.....

Data MASSIMA:

03:14:07 UTC
di martedì
19 gennaio 2038

2.147.483.647

Eventi degni di nota:

9 settembre 2001 01:46:40 AM (UTC).

: Unix millennium, l'orologio segnava 1.000.000.000 secondi: L'evento celebrato a Copenaghen durante una festa organizzata dal DKUUG (Danish UNIX User Group).

18 marzo 2005: L'orologio segnava

1.111.111.111 secondi.

12 maggio 2006 alle 21:27:28 CRASH

Web Server AOLserver: Un miliardo di secondi prima del 2038 il server superò il limite critico restituendo una data del passato, causando crash del sistema.

Il 13 febbraio 2009 alle ore 23:31:30

(UTC), il Tempo Unix ha raggiunto la cifra '1234567890'. Si sono tenuti festeggiamenti vari in tutto il mondo per celebrare l'evento

26 gennaio 2011: è stato il 15.000°

giorno di UNIX.



walter.balzano@gmail.com

Epoca

<http://www.epochconverter.com/>



Epoch & Unix Timestamp Conversion Tools

The current Unix epoch time is **1462649527**

Convert epoch to human readable date and vice versa

[Timestamp to Human date](#) [\[batch convert timestamps to human dates\]](#)

GMT: Thu, 01 Jan 1970 00:00:00 GMT

Your time zone: 1/1/1970, 01:00:00 GMT+1:00

: : [Human date to Timestamp](#)

[Human date to Timestamp](#) [\[batch convert\]](#)

Input format: RFC 2822, D-M-Y, M/D/Y, Y-M-D, etc.

Strip 'GMT' to convert to local time, no support for other time zones.

Press or [click here](#) to clear all forms.

Epoch dates for the start and end of the year/month/day

Show start & end of year month day

[\[Epoch List By Month & Year\]](#)



walter.balzano@gmail.com

The end of Epoch: Soluzioni

Non è semplice risolvere il problema a causa della grande diversità dei processori, Sistemi Operativi e file system. Molti sistemi operativi per sistemi a 64 bit usano già dei numeri interi a 64 bit per il time.h. Il passaggio a questo tipo di architetture è in corso, e ci si aspetta che sia completo prima del 2038. Tuttavia, ancora oggi esistono centinaia di milioni di sistemi a 32 bit sul mercato, di cui molti in sistemi integrati.

Soluzioni:

- Cambiare time.h in un intero unsigned. Ciò permetterà di rimandare il problema al 7 febbraio 2106, ma ciò causerebbe comunque problemi a molti programmi.
- L'uso di un valore di tipo signed a 64 bit sposterebbe l'emergere del problema in avanti nel tempo di circa 290 miliardi di anni, una data che è addirittura al di là della previsione di vita del sistema solare: spostamento eccessivo!!!
 - Sono state avanzate anche varie proposte alternative, alcune delle quali in uso, per sfruttare questo spostamento eccessivo della data massima calcolabile: tra queste, includere nel calcolo delle ore i millisecondi o i microsecondi, abbreviando la vita utile delle macchine a 300.000 anni



walter.balzano@gmail.com

Esempio:

UTC

547667220000

LAT

40.827988, 14.192942

LON

10 maggio del 1987
ore 17:47:00,000

Centro Campo dello
Stadio S.Paolo



walter.balzano@gmail.com

Sincronizzazione Orologi

- Come più volte detto, l'efficacia e la precisione dell'intero sistema GPS dipende fortemente dalla precisione (e quindi dalla sincronizzazione) degli orologi sia terrestri che satellitari
- Gli orologi atomici sono basati su Cesio e Rubidio e sono troppo costosi
- Gli orologi dei ricevitori sono basati su quarzo e sono economici; essi sono continuamente “resettati”

SINCRONIZZAZIONE:

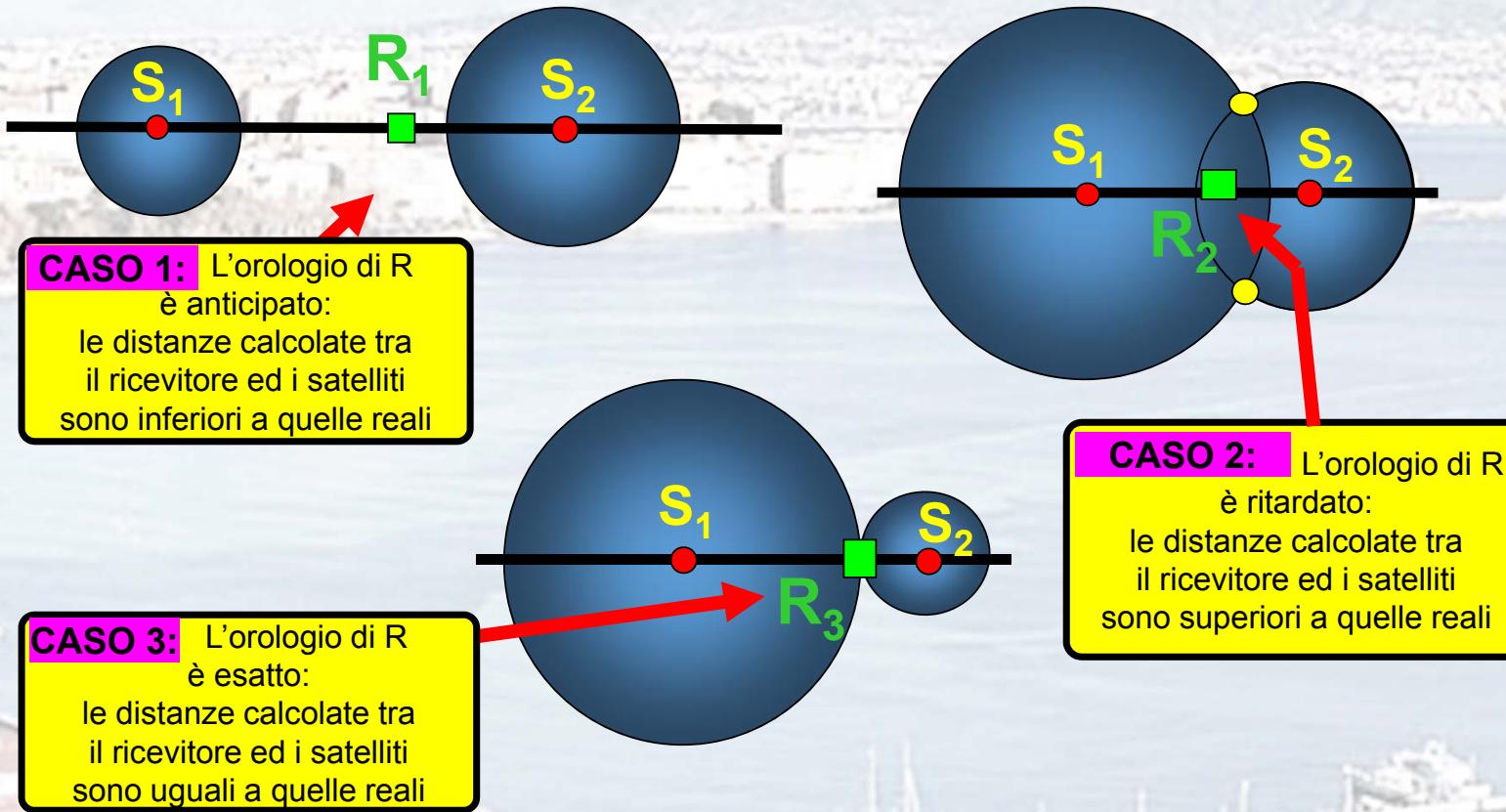
- Il ricevitore R, esaminando l'orario di 4 o più satelliti può misurare la propria imprecisione: in altre parole, **solo e soltanto un orario T può soddisfare le equazioni geometriche per poter individuare un unico punto nello spazio: le 4 sfere di distanza individuate dai 4 satelliti e dal ricevitore si intersecano in un solo punto soltanto se le misurazioni ottenute sono precise**: T è proprio l'orario degli orologi atomici dei satelliti. Il ricevitore R quindi calcola ed imposta il proprio orario a T semplicemente imponendo che l'intersezione delle 4 sfere coincida in un unico punto.
- Il sistema GPS può quindi fornire gratuitamente l'accuratezza di un orologio atomico



walter.balzano@gmail.com

Sincronizzazione (2)

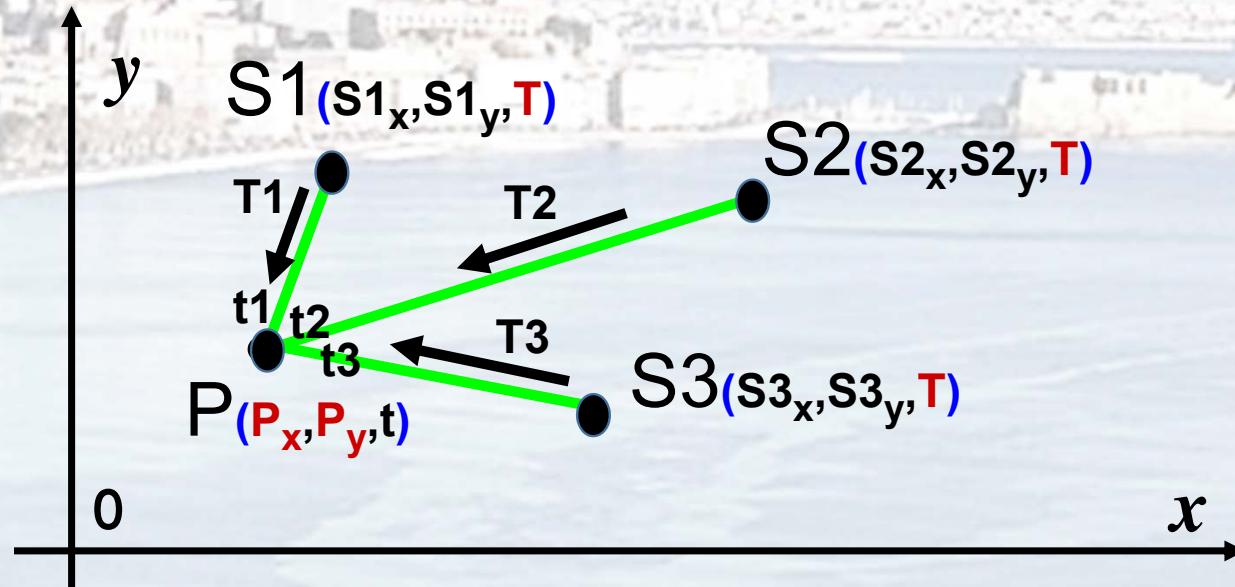
- **Esempio in 1D** (in 3D il concetto è equivalente ma esteso a 4 sfere)
- **Le coordinate "spaziali" dei satelliti S_1 ed S_2 sono note al Ricevitore R** (R infatti possiede l'Almanacco)
- A seconda dell'ora dell'orologio di R (orario R_1 , R_2 ed R_3) posso quindi ricavare diverse grandezze delle sfere di distanza.
- Nell'esempio si suppongono 3 casi relativi a 3 orari differenti (orario R_1 , R_2 ed R_3) dell'orologio del ricevitore R



Sincronizzazione orologi - Caso 2D

Si supponga che:

- In un piano C siano note le coordinate dei 3 punti S1, S2 ed S3 e siano invece ignote le coordinate del punto P.
- Gli orologi di S1, S2 ed S3 siano precisi e perfettamente sincronizzati con tempo T.
- L'orologio di P sia impreciso con tempo t ($t + e = T$; e rappresenta l'errore, positivo o negativo, dell'orologio di P).



Noto: $S1_x, S1_y, S2_x, S2_y, S3_x, S3_y, T, T1, T2, T3, t$ Incognite: P_x, P_y, T

Altre relazioni:

- $t + e = T$ (in cui e = errore orologio di P); Spazio= $V*T$; $c=300.000 \text{ km/s}$



walter.balzano@gmail.com

Sincronizzazione orologi - Caso 2D

Si supponga che P misuri le distanze dai 3 punti S1, S2 ed S3 usando il proprio orologio t cioè P legge il proprio orologio ogni volta che riceve un messaggio: quindi P legge i 3 orari t_1 , t_2 e t_3 relativi all'arrivo dei messaggi di S1, S2 ed S3.

Con il teorema di Pitagora si calcolano le distanze tra P ed i punti S1, S2 ed S3:

$$\left\{ \begin{array}{l} \text{distanza}(P,S1) \rightarrow \sqrt{(S1_x - P_x)^2 + (S1_y - P_y)^2} = v \cdot (t_1 + e - T_1) \\ \text{distanza}(P,S2) \rightarrow \sqrt{(S2_x - P_x)^2 + (S2_y - P_y)^2} = v \cdot (t_2 + e - T_2) \\ \text{distanza}(P,S3) \rightarrow \sqrt{(S3_x - P_x)^2 + (S3_y - P_y)^2} = v \cdot (t_3 + e - T_3) \end{array} \right.$$

in cui:

T_i = istante di partenza del messaggio da S_i (incluso nel messaggio da S_i)

t_i = istante d'arrivo del messaggio di S_i in P (contenente il tempo T_i di S_i)

$t_i - T_i$ = tempo approssimato di percorrenza da S_i a P (calcolato in P)

$t_i + e - T_i$ = tempo esatto di percorrenza da S_i a P.

v = velocità della luce = 300.000 km/s

Con la risoluzione di tale sistema di 3 equazioni con 3 incognite (P_x , P_y , e) si ricava anche il tempo T (infatti $T = t + e$)



walter.balzano@gmail.com

A-GPS: Assisted GPS

● L' A-GPS, è un sistema che consente di abbattere i tempi necessari alla prima localizzazione durante l'uso di un terminale GPS. 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. Questo sistema sta mostrando una notevole diffusione, ed è normalmente associato ai sistemi di localizzazione (LBS, Location Based Service) basati su telefonia cellulare.

● Uno dei principali problemi dei terminali GPS classici è relativo alla prima localizzazione (il Fixing), in quanto un terminale, all'accensione, deve ricavare la lista dei satelliti in vista in quel momento, per potervisi agganciare e cercare di rilevare la propria posizione. Tale processo è in genere alquanto dispendioso in termini di tempo e risorse, ed il sistema A-GPS è stato studiato al fine di abbattere tali costi, anche in previsione di un utilizzo su terminali con basse capacità di elaborazione o con risorse energetiche limitate, quali i telefoni cellulari.



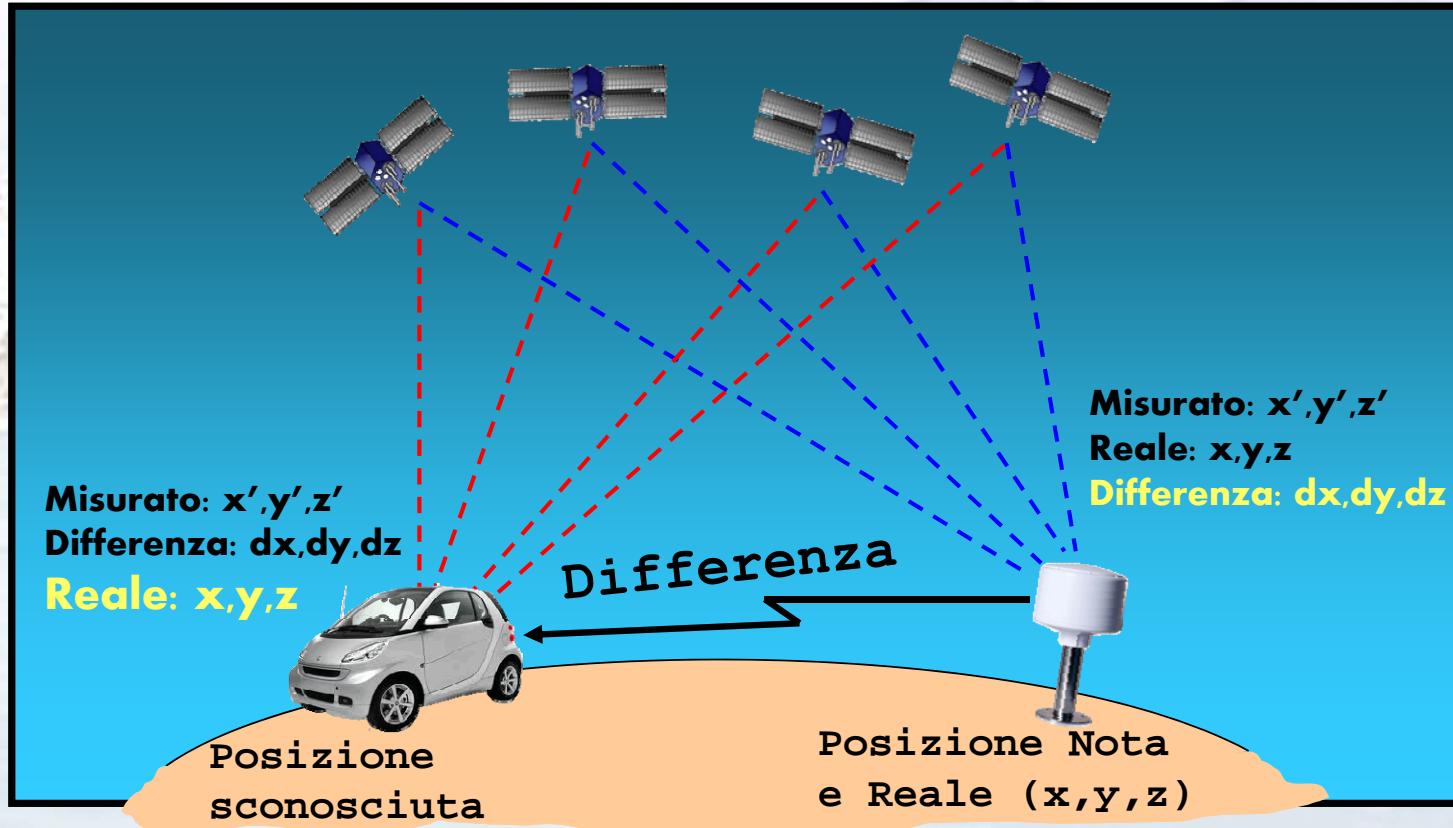
walter.balzano@gmail.com

A-GPS: Assisted GPS - 2

- Lo scopo principale di questo sistema è quello di "assistere" il ricevitore GPS nel calcolo della posizione, fornendogli informazioni sui satelliti in vista. Tale metodologia richiede il supporto dell'operatore di telefonia mobile.
- L'idea di fondo è la seguente: dato che ogni cella di telefonia mobile presente sul territorio ha una posizione fissa, si fa in modo che sia la cella a ricavare quali siano i satelliti GPS ad essa in vista, istante per istante. Quando un terminale A-GPS vuole conoscere la sua posizione, si collega tramite la rete cellulare ad un Assistance Server (che può anche essere gestito dall'operatore stesso), al quale viene inviata anche l'informazione sulla cella a cui l'utente è agganciato. Dato che sono noti i satelliti in vista alla cella, si può assumere ragionevolmente che anche il terminale A-GPS veda i medesimi satelliti. Pertanto il server elabora una lista con i satelliti in vista, e la invia attraverso la rete cellulare al terminale, che in questo modo può ricavare immediatamente la propria posizione.
- La ridotta quantità di dati da inviare ha suggerito l'impiego di messaggi di tipo cell broadcast, già ampiamente utilizzati, ad esempio, per effettuare tariffazioni diverse a seconda dell'area in cui l'utente si trova, e che compaiono sullo schermo dell'utente sotto forma di indicazione della provincia.



D-GPS: Differential GPS

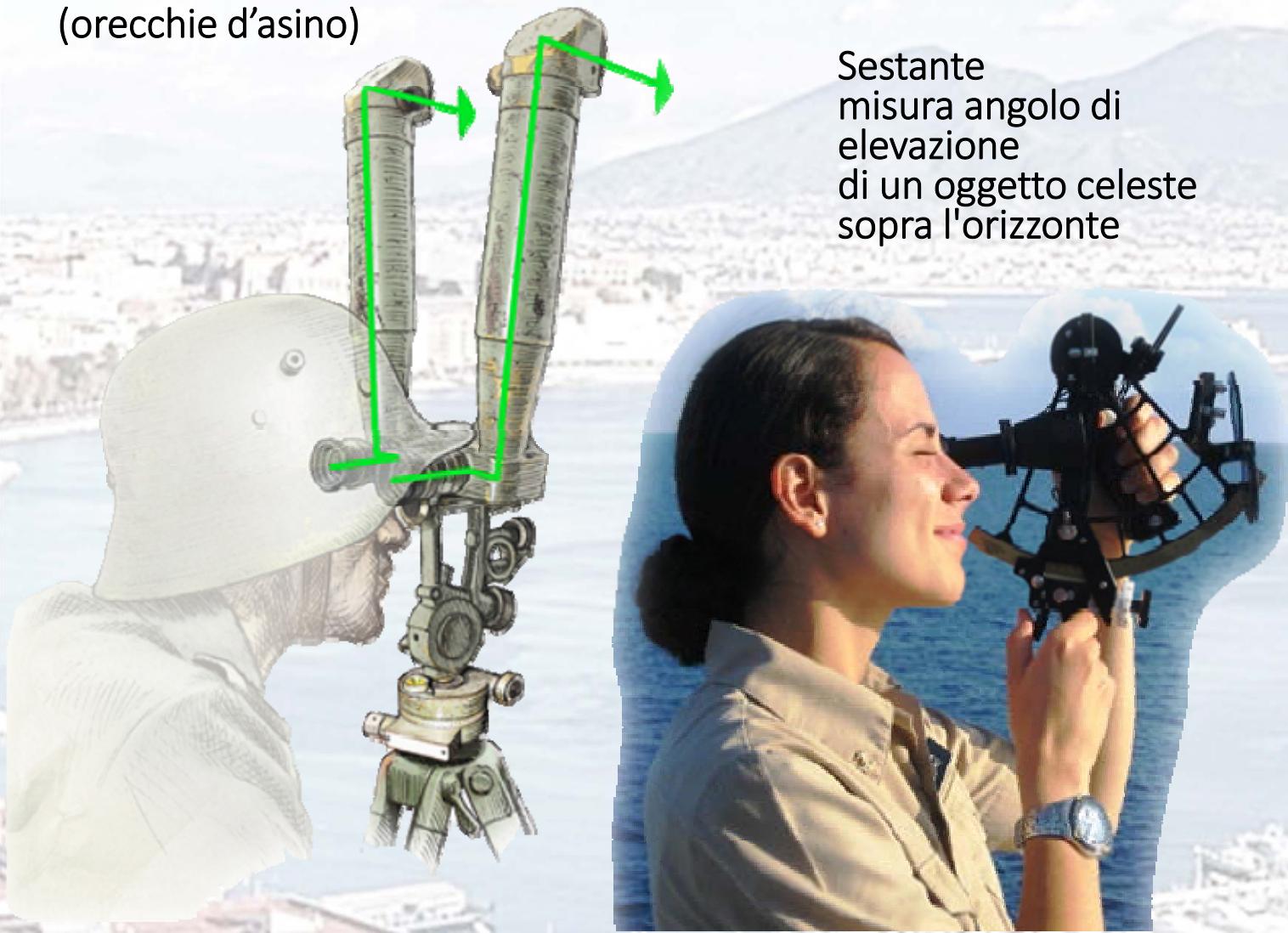


walter.balzano@gmail.com

trench
binoculars
(orecchie d'asino)

...Misurazioni

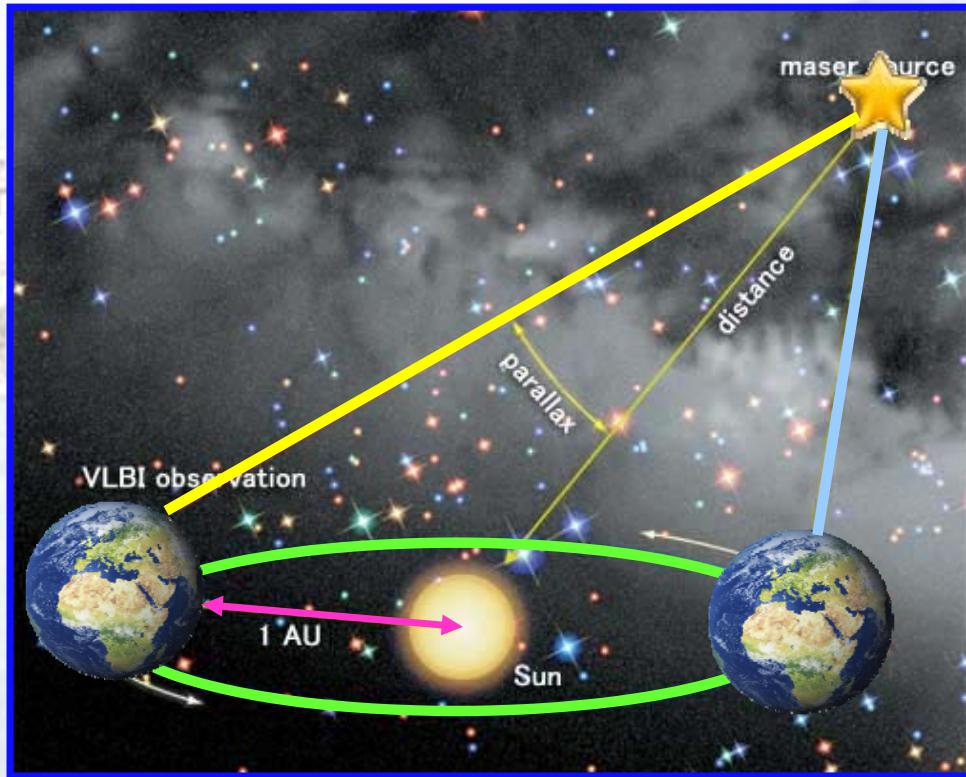
Sestante
misura angolo di
elevazione
di un oggetto celeste
sopra l'orizzonte



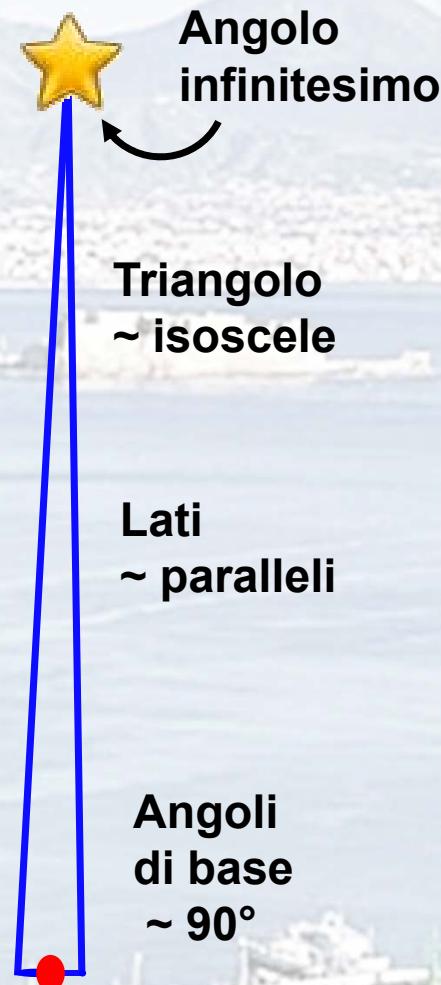
walter.balzano@gmail.com

...Misurazioni (2)

Importanza dell'angolo:
il Parallax Method



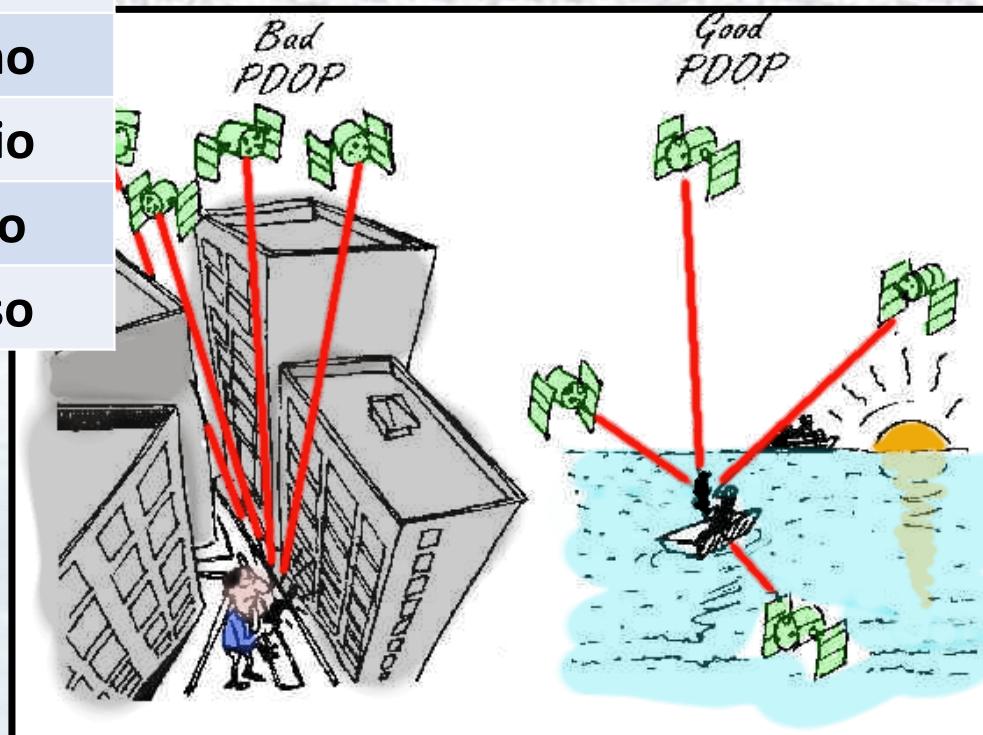
Base Triangolo =
Diametro orbita Terra-Sole



walter.balzano@gmail.com

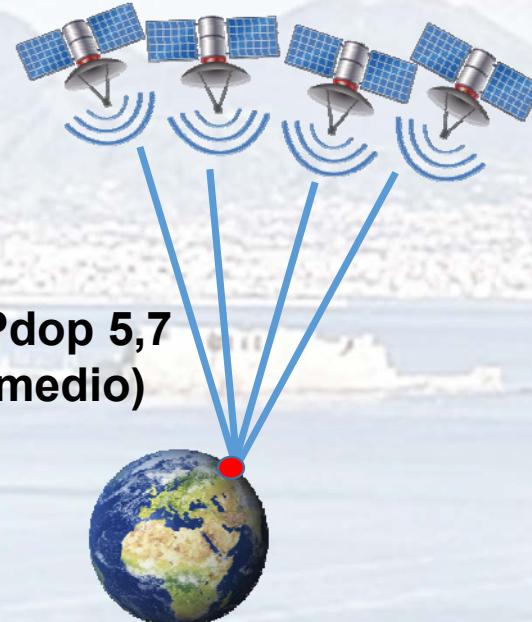
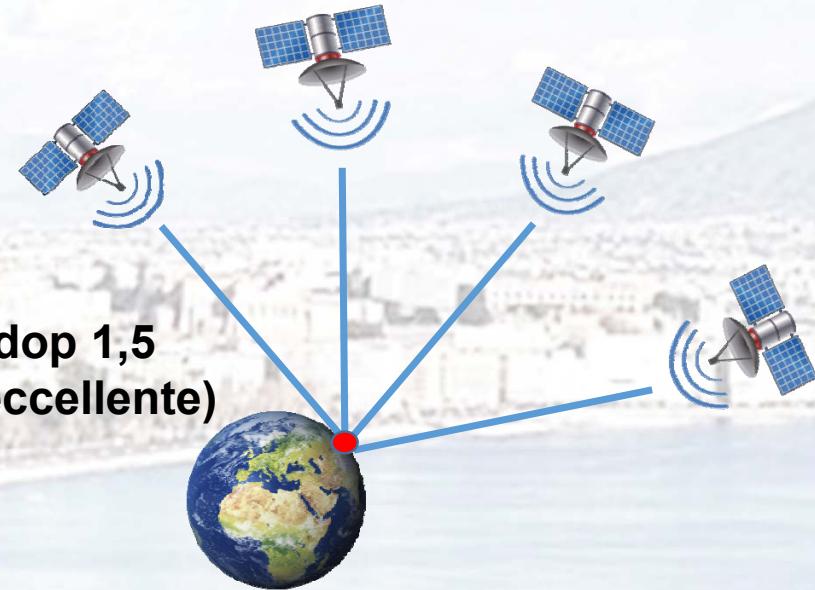
DOP - Diluition Of Precision

DOP	Valutazione
1	ideale
1-2	Eccellente
2-5	Buono
5-10	Medio
10-20	Basso
>20	Scarso



walter.balzano@gmail.com

DOP - Dilution Of Precision - 2



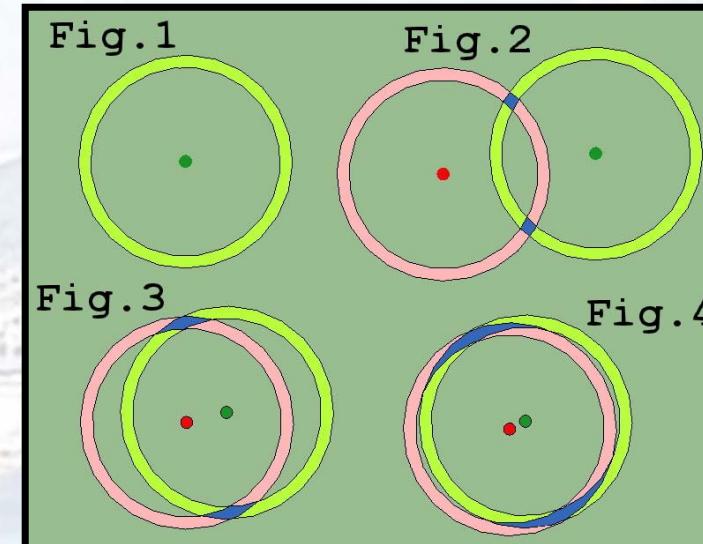
Abbiamo visto che per il calcolo effettivo della posizione del ricevitore GPS viene utilizzato il metodo della **Trilaterazione**, basata sulle distanze e non sugli angoli... Allora perché è comunque desiderabile che la distanza per esempio tra due satelliti sia maggiore possibile in modo che essi formino un angolo maggiore possibile?

walter.balzano@gmail.com

DOP - Dilution Of Precision - Esempio in 2D

- Natura geometrica
 - I dispositivi hanno una precisione limitata
- Esempio:**

• **Figura 1:** se conosco la mia distanza dal satellite “●“ e suppongo che esista un errore allora posso ragionevolmente supporre di trovarmi in un punto qualsiasi della relativa corona circolare (zona verde), cioè la mia distanza dal satellite è compresa tra un valore minimo ad uno massimo...



• **Figura 2, 3 e 4:** se conosco la mia distanza sia dal satellite1 “●“ che dal satellite2 “●“ allora, sapendo di dover considerare un errore, potrò trovarmi nell'intersezione (zona blu) delle rispettive 2 corone circolari (rosa e verde)

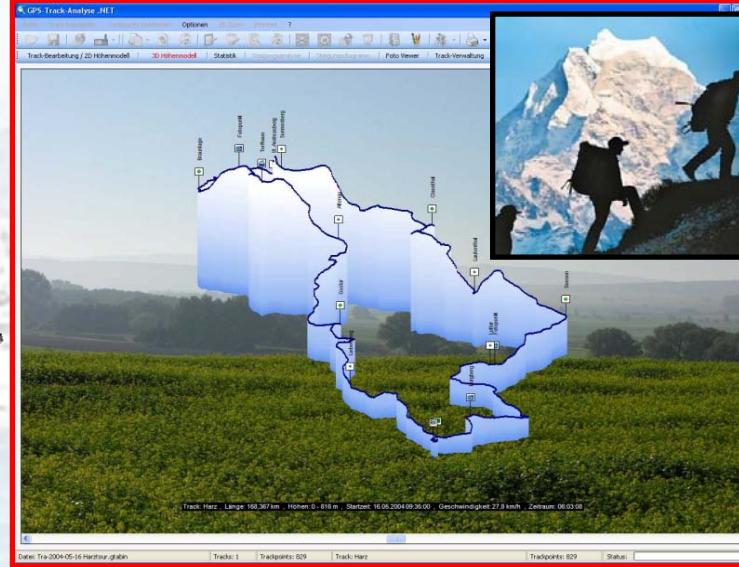
• **Si nota facilmente che:** quanto più sono vicini i satelliti “●“ e “●“, tanto maggiore è l'area di intersezione (zona blu) e, di conseguenza, la mia posizione è più 'incerta'... **viceversa:** quanto più sono distanti i satelliti allora l'area di intersezione (zona blu) tende ad essere minore...!

Se, per assurdo, il mio dispositivo avesse una precisione infinita allora anziché corone circolari avrei solo semplici circonferenze e le loro intersezioni sarebbero solo punti e non aree e quindi la distanza tra i satelliti di riferimento non sarebbe poi così importante.



walter.balzano@gmail.com

GPS LOGGER



walter.balzano@gmail.com

GPS Tracker



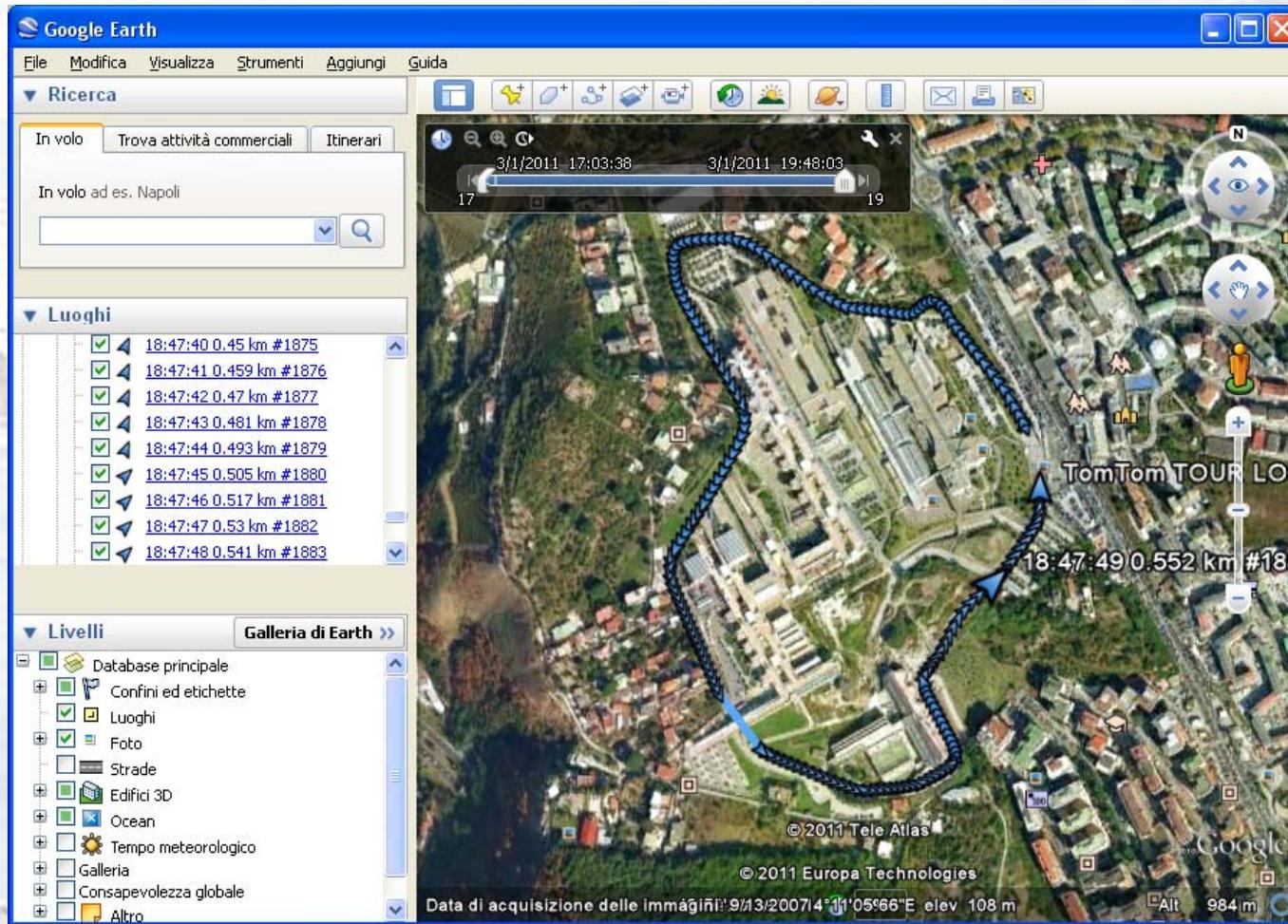
walter.balzano@gmail.com

GPS Tracker



walter.balzano@gmail.com

GPS Tracker



walter.balzano@gmail.com

Formato File GPX

```
<gpx version='1.1' creator='TTTracklog v.1.13' .....>
<trk> <name>TomTom TOUR LOG 29.783 km (1551/1802)</name>
<number>0</number>
<trkseg>
  <trkpt lat='40.83995687' lon='14.18720898'>
    <ele>93.3</ele>
    <time>2011-03-01T15:03:38Z</time>
    <name>16:03:38 28.574 km #1495 </name>
    <course>332.3</course>
    <speed>8.12822</speed>
    <pdop>1.9</pdop>
    <hdop>1.1</hdop>
    <vdop>1.6</vdop>
    <sat>8</sat>
  </trkpt>
  .....
</trkseg>
</trk>
</gpx>
```



walter.balzano@gmail.com

Riferimenti

centrini topografici
(punti fiduciali)



Via cintia

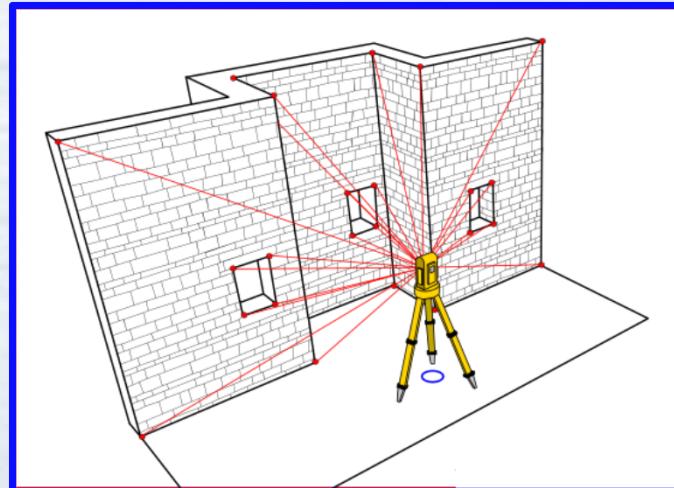
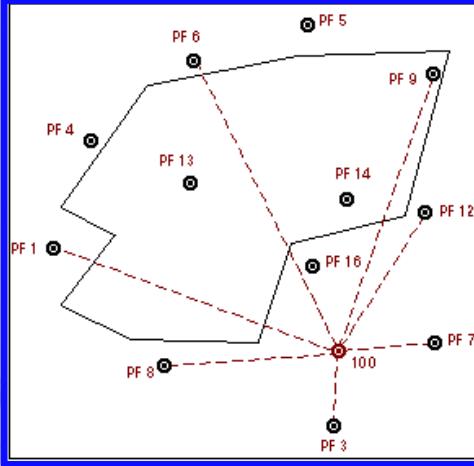
GPS = 40.841074, 14.186843

vedi su Maps: <http://bit.do/puntofiduciale>



walter.balzano@gmail.com

Riferimenti



walter.balzano@gmail.com

Metodi di localizzazione alternativi / supporto al GPS

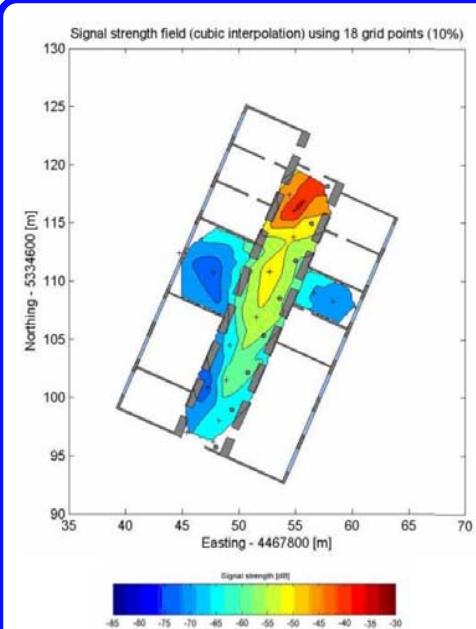
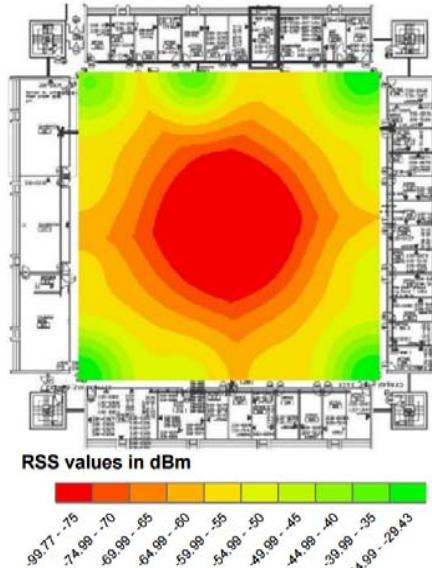
Fingerprints



walter.balzano@gmail.com

Localizzazione Indoor

RadioMappe per il FingerPrints



MAC	Avg RSS
00:40:96:A1:0D:3A	-33.15
00:40:96:A1:0C:B1	-56.62
00:40:96:A1:0C:AD	-59.69
00:02:8A:9E:4F:31	-62.77
00:40:96:A1:0D:08	-64.77
00:40:96:A1:0D:45	-66
00:40:96:A1:0D:7C	-67
00:09:B7:7B:9F:60	-86
00:40:96:A1:0C:A3	-86

List of signal strengths (dBm) measured at one unknown location from user

Received Signal Strength (**RSS**) radio map



walter.balzano@gmail.com

Localizzazione Indoor «casalinga»

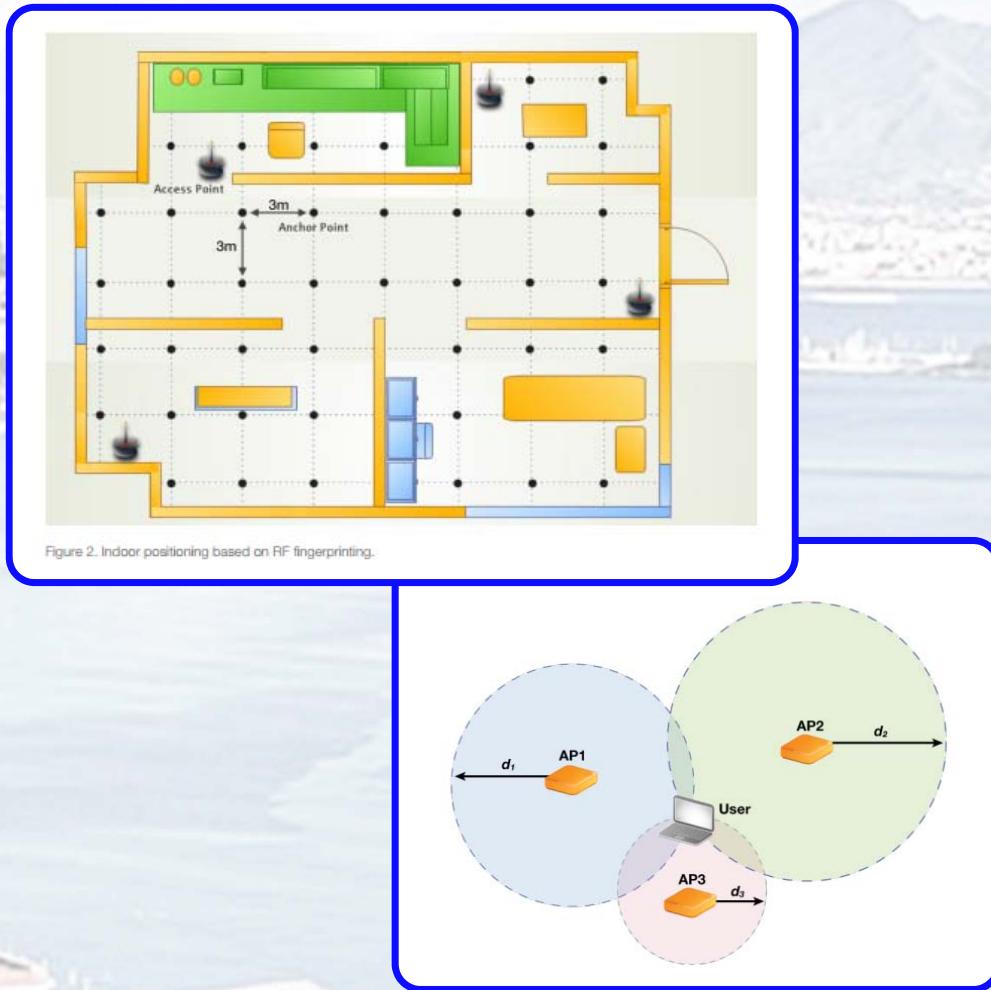
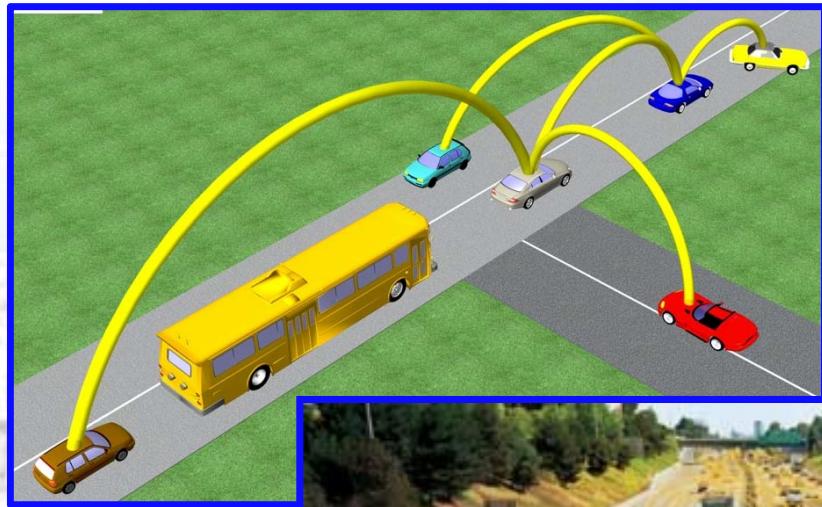


Figure 2. Indoor positioning based on RF fingerprinting.

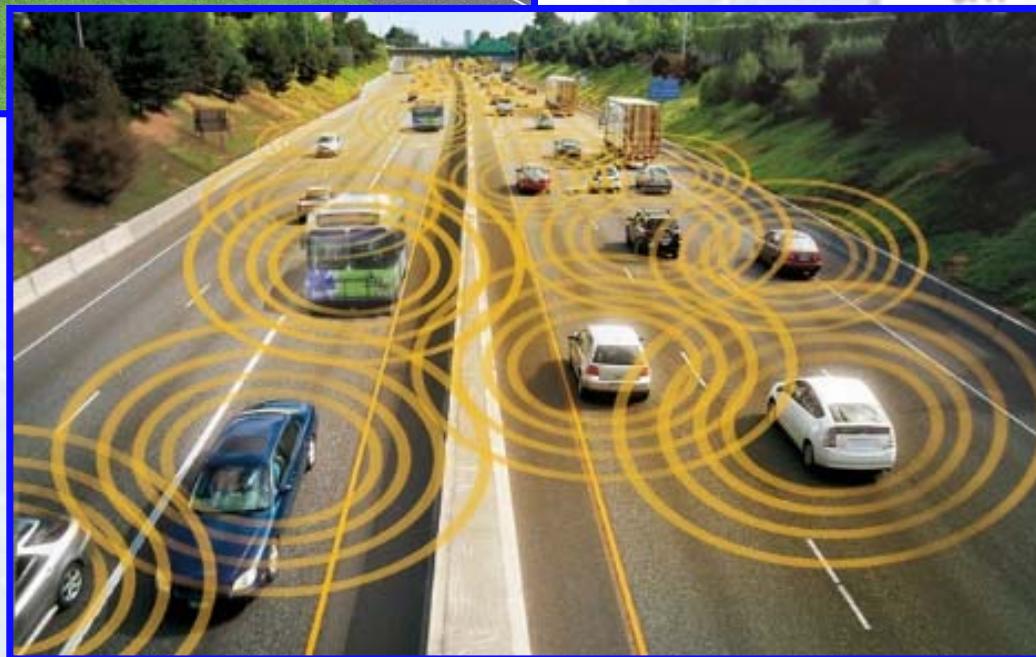


walter.balzano@gmail.com

Localizzazione OUTdoor «Stradale»



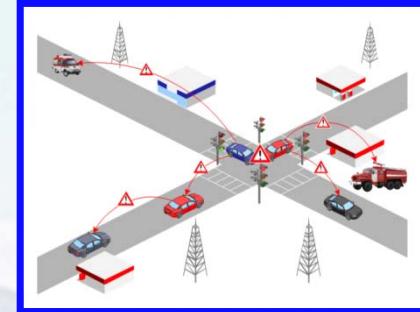
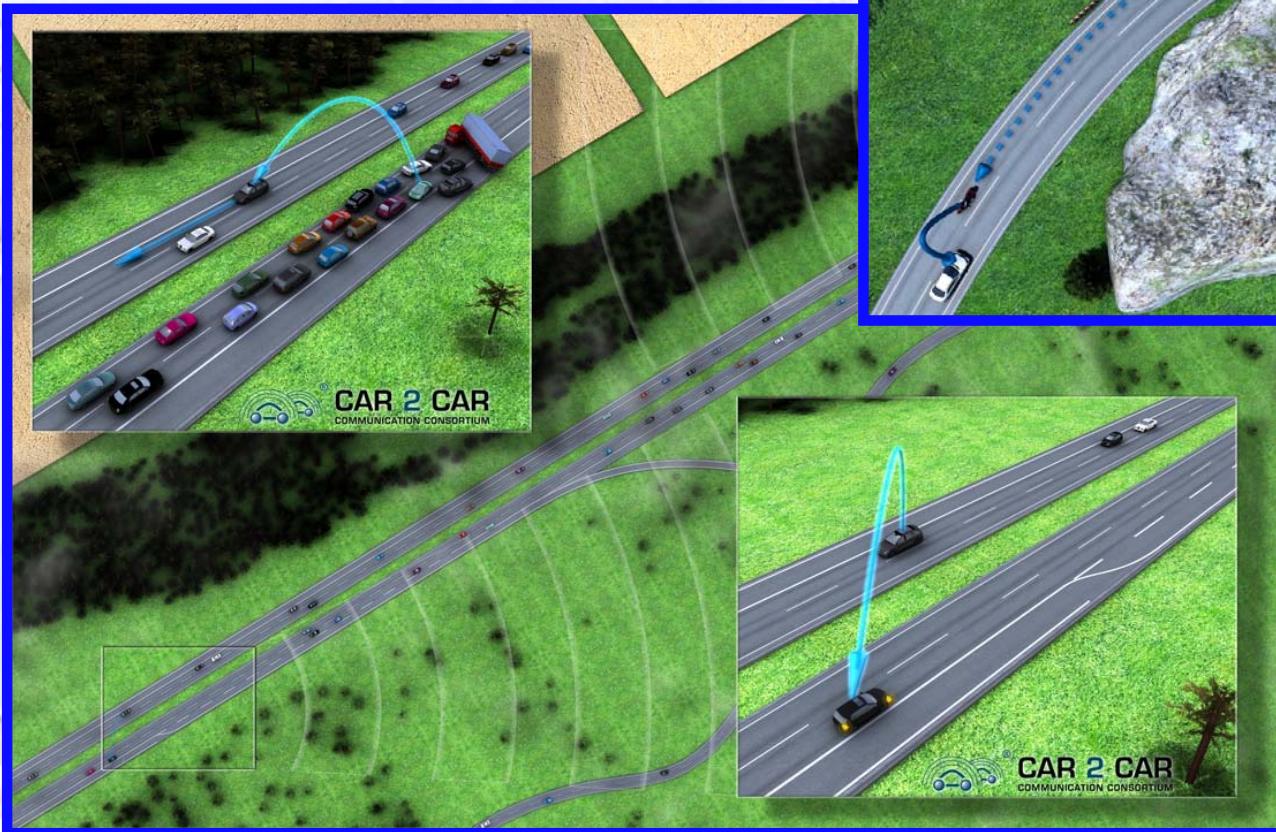
Car2Car



walter.balzano@gmail.com

Localizzazione OUTdoor «Stradale»

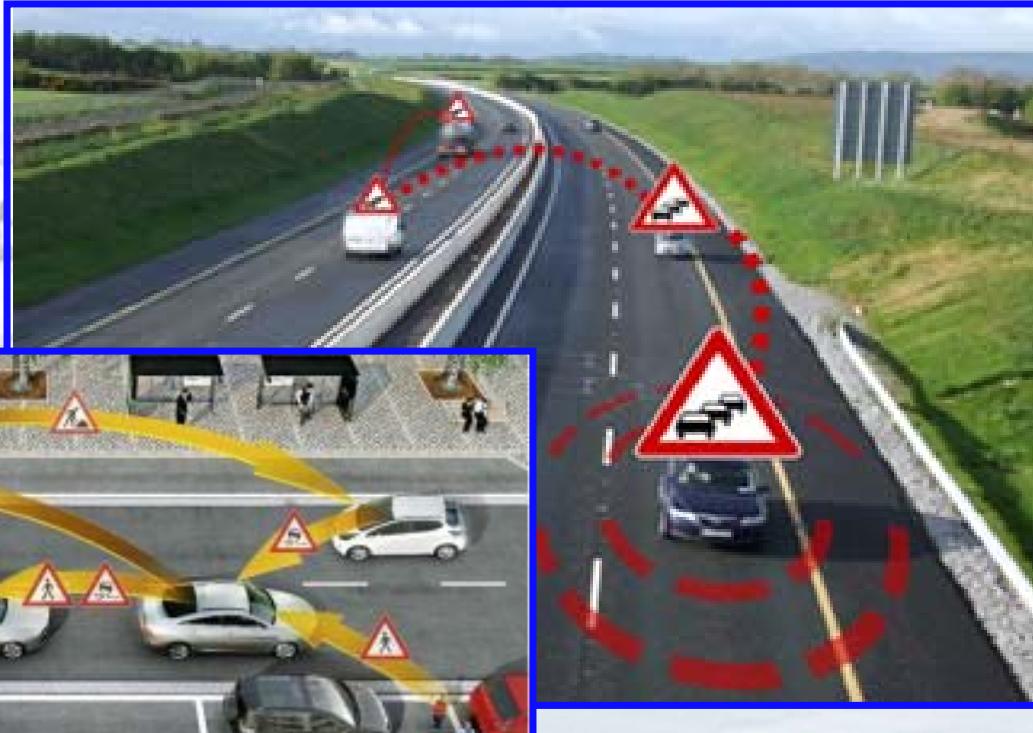
- car2car
- V2V! Vehicle2Vehicle
- V2I! Vehicle2Infrastructure



walter.balzano@gmail.com

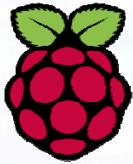
Localizzazione OUTdoor «Stradale»

Car2Car

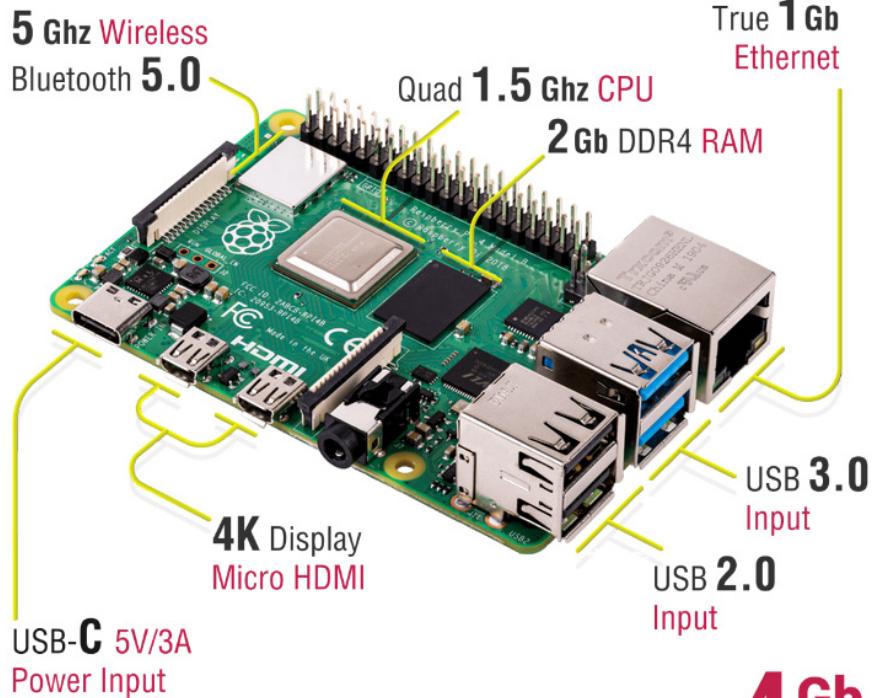


walter.balzano@gmail.com

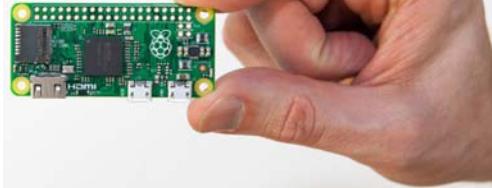
Prototipi per la Sperimentazione



Raspberry Pi 4 b



Raspberry Pi Zero



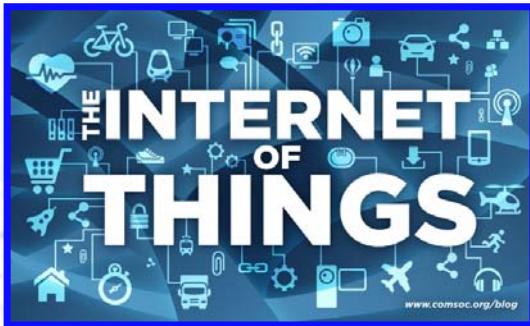
Berry IMU
Inertial
Measurement
Unit

- Giroscopio
- Accelerometro
- Pressione
- Altitudine
- Temperatura



walter.balzano@gmail.com

Localizzazione ed iot (Internet of Things)



walter.balzano@gmail.com