

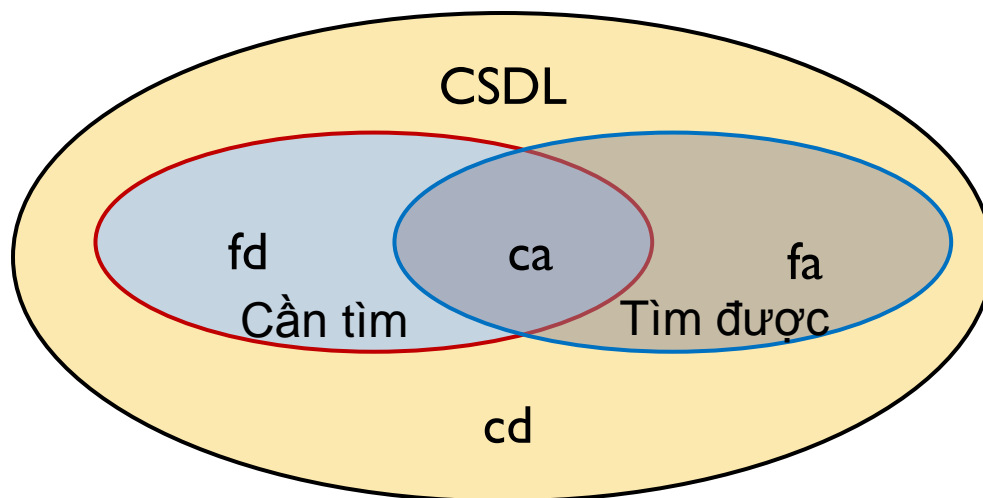
Các loại dữ liệu đa phương tiện

Nguyễn Đình Hóa

dinhhoa@gmail.com 0942807711

Tóm tắt nội dung bài 1

- ▶ Giới thiệu về hệ CSDL đa phương tiện
 - ▶ Đa phương tiện
 - ▶ Các phương tiện truyền dẫn thông tin
 - ▶ Các yêu cầu đối với hệ CSDL đa phương tiện
 - ▶ Ứng dụng của hệ CSDL đa phương tiện
 - ▶ Các thước đo hiệu quả của công cụ tra cứu dữ liệu
 - ▶ Precision
 - ▶ Recall



Các loại dữ liệu đa phương tiện

- ▶ Dữ liệu hình ảnh (image databases)
- ▶ Dữ liệu âm thanh (audio databases)
- ▶ Dữ liệu văn bản (text databases)
- ▶ Dữ liệu video (video databases)

Dữ liệu hình ảnh

Hãy mô tả các mẫu ảnh nền sau:



Dữ liệu hình ảnh

- ▶ Ví dụ cơ bản về những thuộc tính bậc thấp:
 - ▶ Màu sắc: màu của vật thể, màu nền,...
 - ▶ Hình dạng: kích thước, hình dạng của các vật thể,...
 - ▶ Phân bố màu nền: độ tương phản, sự bao phủ của màu nền,...

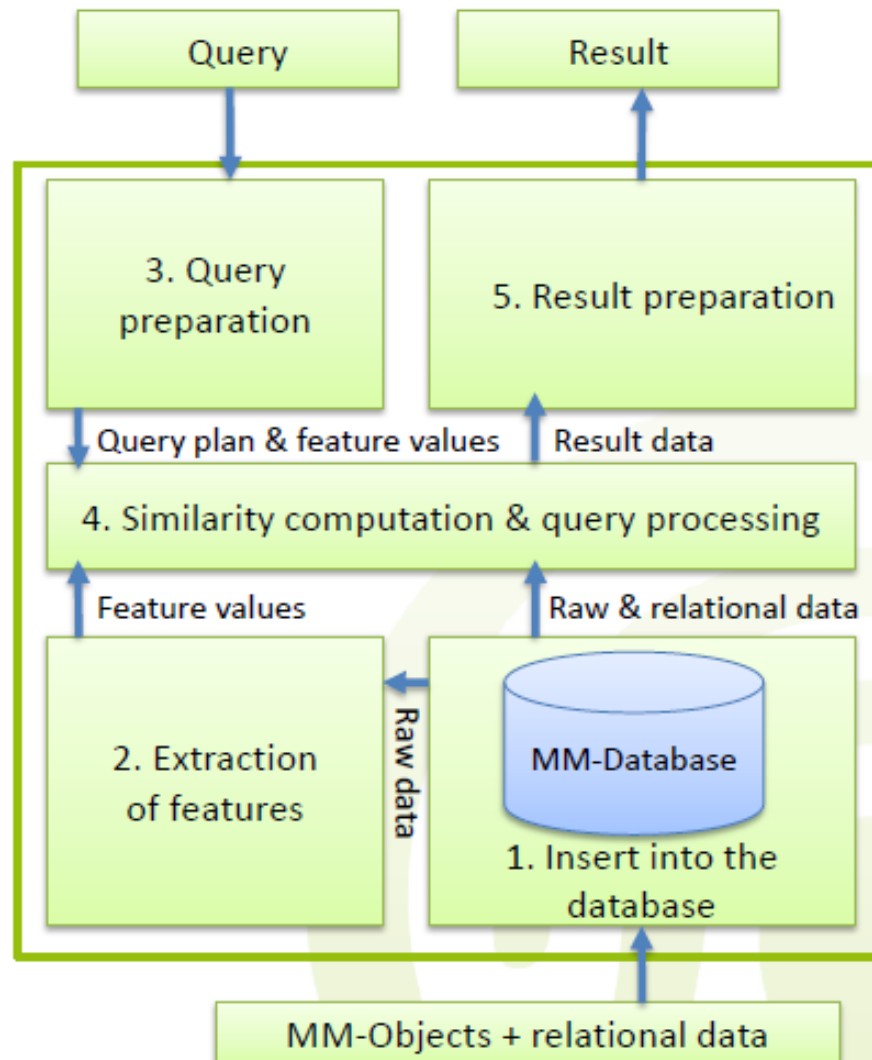


Dữ liệu hình ảnh

Đặc điểm của các thuộc tính nhận dạng:

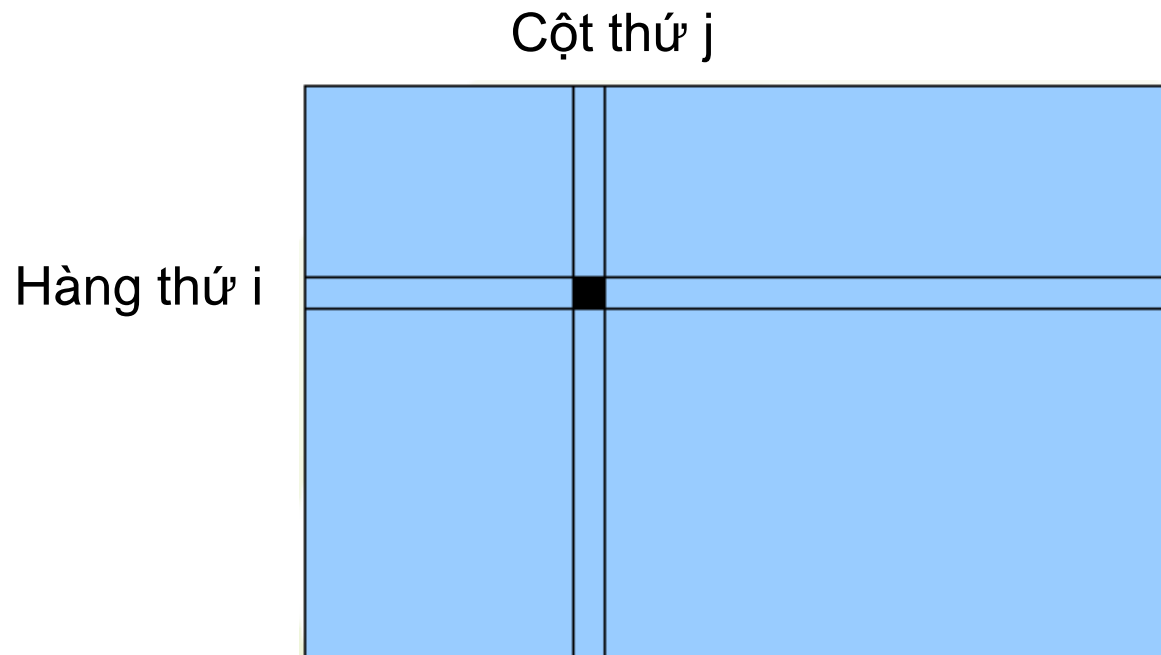
- ▶ Đánh giá sự khác nhau giữa các ảnh,
- ▶ Đôi khi khó diễn tả bằng ngôn ngữ
- ▶ Có thể trả về nhiều kết quả tra cứu tương tự nhau
- ▶ Có thể tạo ra nhiều câu truy vấn phức tạp bằng cách kết hợp nhiều thuộc tính với nhau.

Các bước truy vấn ảnh



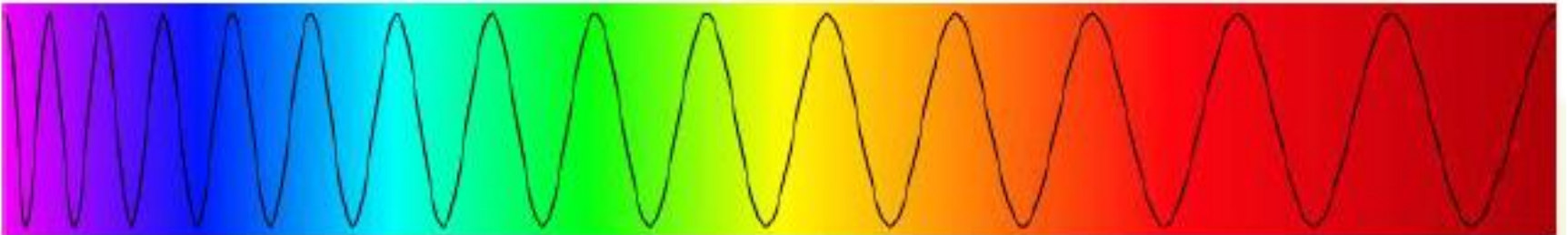
Dữ liệu hình ảnh

- ▶ Hình ảnh là một ma trận hai chiều
- ▶ Mỗi phần tử của ảnh được biểu diễn bởi
 - ▶ Tọa độ
 - ▶ Màu sắc



Dữ liệu hình ảnh

- ▶ Range of visible light: [380, 780] nm
- ▶ Blue: 435.8 nm, green: 546.1 nm, red: 700 nm



Dữ liệu hình ảnh

Giả thiết: nếu hai hình ảnh có màu sắc giống nhau thì nội dung của chúng cũng giống nhau

- ▶ Thông tin của ảnh sẽ bị mất nếu sử dụng thuộc tính cấp thấp để mô tả,
- ▶ Ví dụ: ảnh hoàng hôn (màu cam, vàng)



Không gian màu

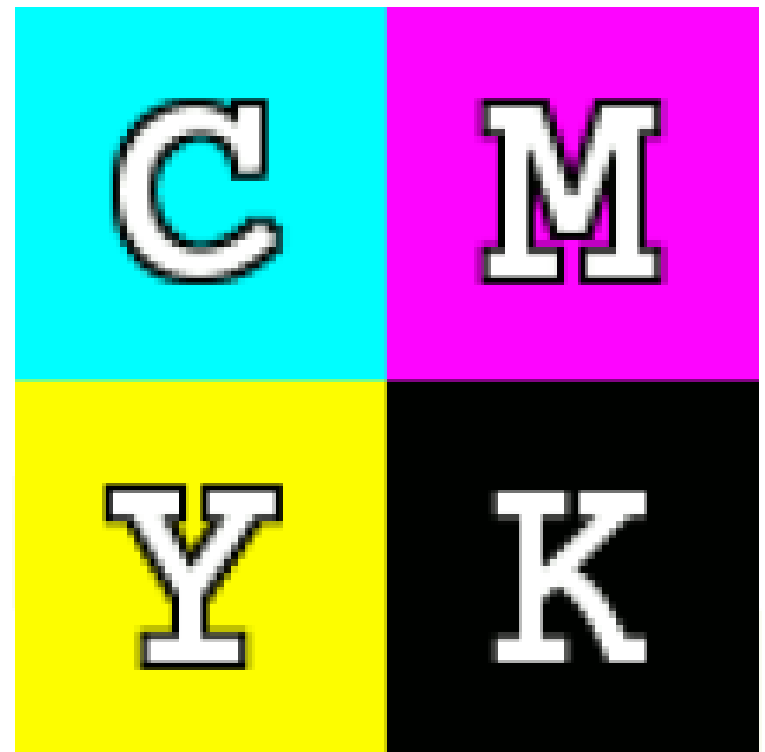
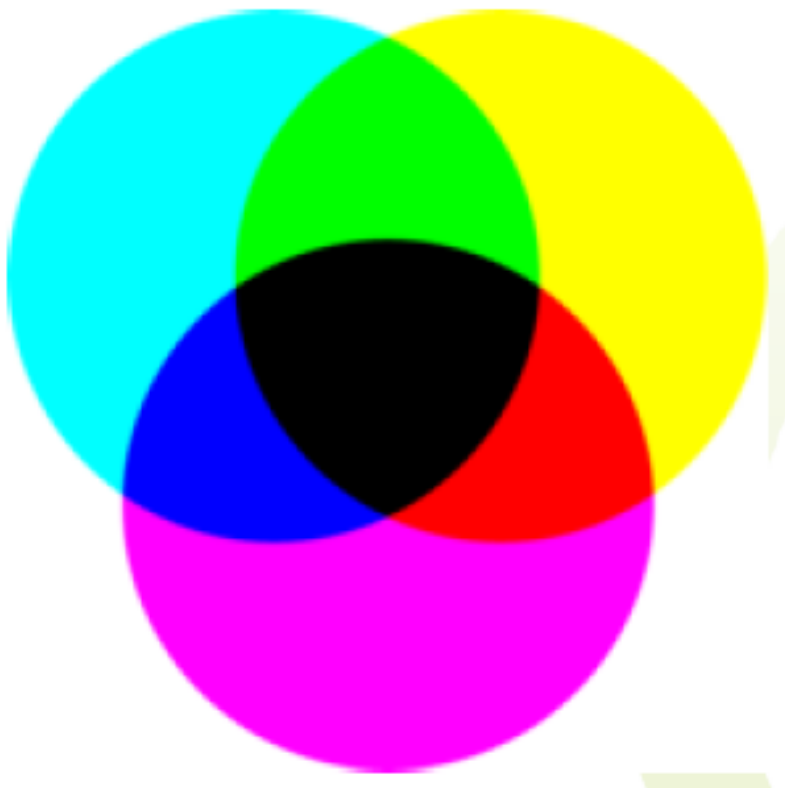
▶ Hệ RGB

- ▶ $(0,0,0)$ đen
- ▶ $(255,255,255)$ trắng
- ▶ $(255,0,0)$ đỏ
- ▶ $(0,255,0)$ lục
- ▶ $(0,0,255)$ lam
- ▶ $(255,255,0)$ vàng
- ▶ $(0,255,255)$ lục lam (cyan)
- ▶ $(255,0,255)$ đỏ tươi (magenta)



Không gian màu

- ▶ Hệ CMYK: Cyan, Magenta, Yellow, Black



Không gian màu

- ▶ Hệ CIE (Commission International de l'Eclairage – Standardization Commission on Illumination): chuyển đổi phi tuyến từ hệ RGB.
- ▶ Hệ HSV (Hue, Saturation, Value): chuyển đổi phi tuyến từ hệ RGB.

Dữ liệu hình ảnh

- ▶ Kết quả so sánh có thể tốt
 - ▶ Ảnh của con ếch xanh khác với ảnh hoàng hôn
- ▶ Kết quả so sánh có thể không tốt lắm
 - ▶ Ảnh của con ếch màu da cam không khác nhiều so với ảnh hoàng hôn.



Dữ liệu hình ảnh

- ▶ Kết hợp nhiều thuộc tính cấp thấp đôi khi mang lại kết quả tìm kiếm chính xác,
- ▶ Tuy nhiên, cùng một khung hình có thể mang nhiều nghĩa khác nhau
 - ▶ Hoàng hôn <-> bình minh bên bờ biển
 - ▶ Ảnh của một quả bóng màu da cam trên bờ biển.



Dữ liệu hình ảnh

- ▶ Các thuộc tính bậc cao

- ▶ Chuyển đổi Fourier

- ▶ Coi toàn bộ ảnh là một dãy tín hiệu
 - ▶ Chuyển đổi từ miền không gian sang miền tần số (mô tả hình ảnh bằng các dao động về mật độ)
 - ▶ Đảm bảo không mất thông tin
 - ▶ Khó diễn đạt bằng lời nói

- ▶ Chuyển đổi bằng wavelets

- ▶ Chuyển đổi Cosine

- ▶ ...

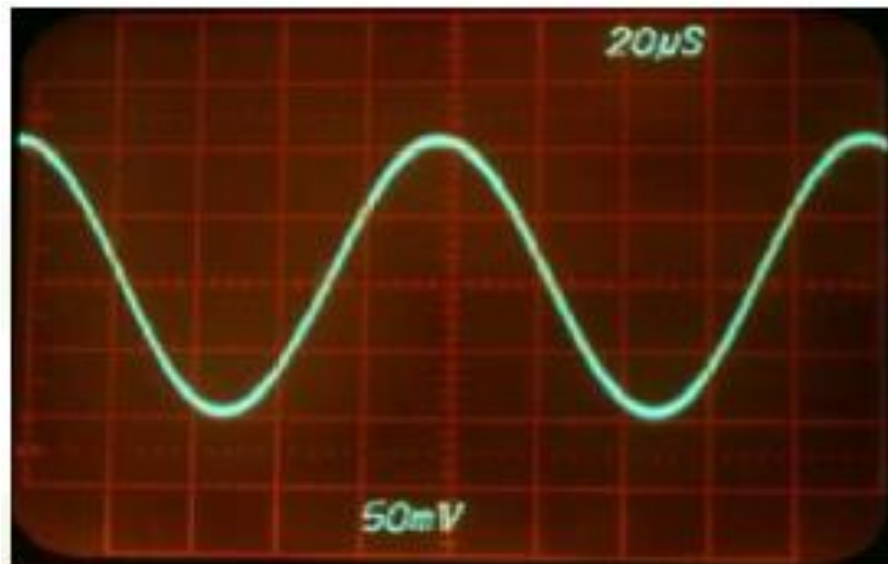
Dữ liệu âm thanh

- ▶ Thông tin được truyền tải bằng sóng âm
- ▶ Có 3 loại dữ liệu chính:
 - ▶ Âm nhạc
 - ▶ Lời nói
 - ▶ Tiếng ồn, nhiễu, ...



Dữ liệu âm thanh

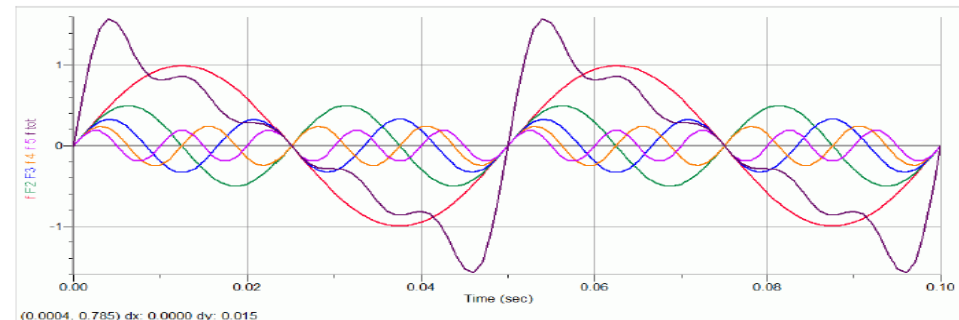
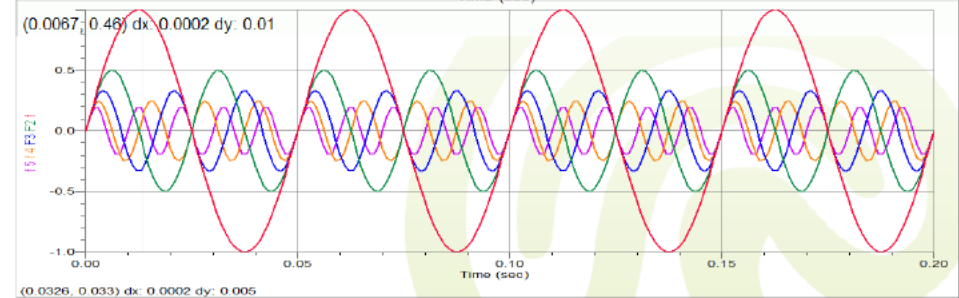
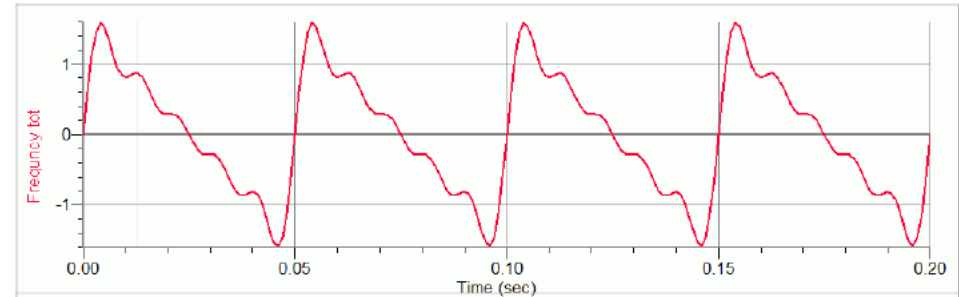
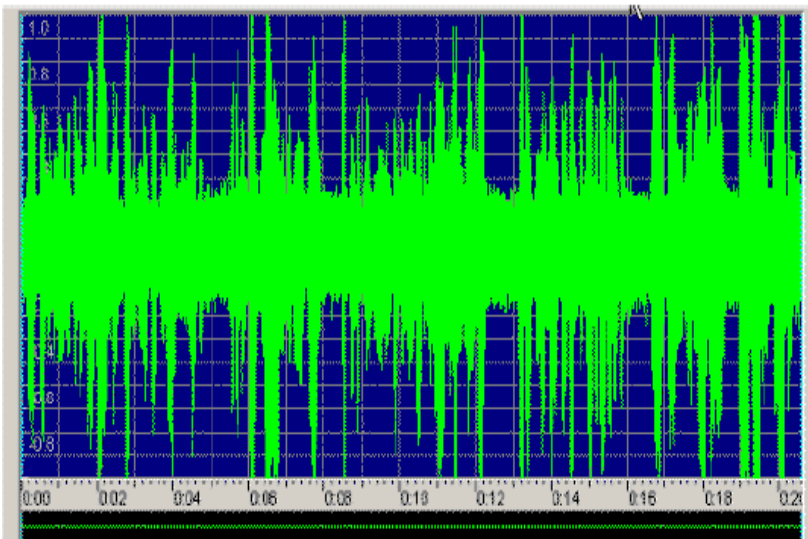
- ▶ Não người chỉ có thể nhận dạng hai thuộc tính cơ bản của âm thanh
 - ▶ Độ cao/thấp (tần số của sóng âm, từ 20Hz đến 20kHz)
 - ▶ Độ lớn/nhỏ (biên độ của sóng âm)



Dữ liệu âm thanh

Tín hiệu âm thanh là sự tổng hợp của nhiều sóng âm:

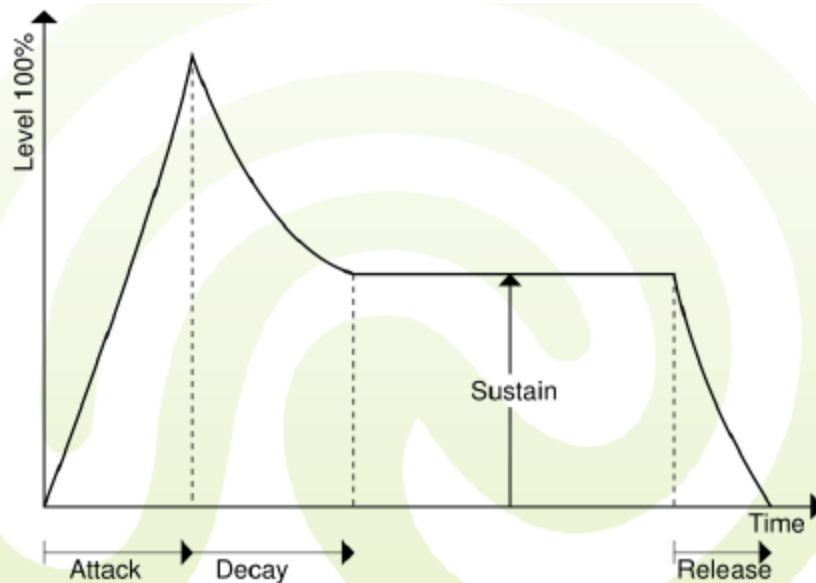
- ▶ Thông tin
- ▶ Nhiễu tương hỗ
- ▶ Nhiễu triệt tiêu



Dữ liệu âm thanh

Sự hình thành âm thanh

- ▶ Bộ phát dao động
- ▶ Bộ khuếch đại
- ▶ Biểu đồ về cường độ âm thanh theo thời gian (attack-decay-sustain-release: ADSR)

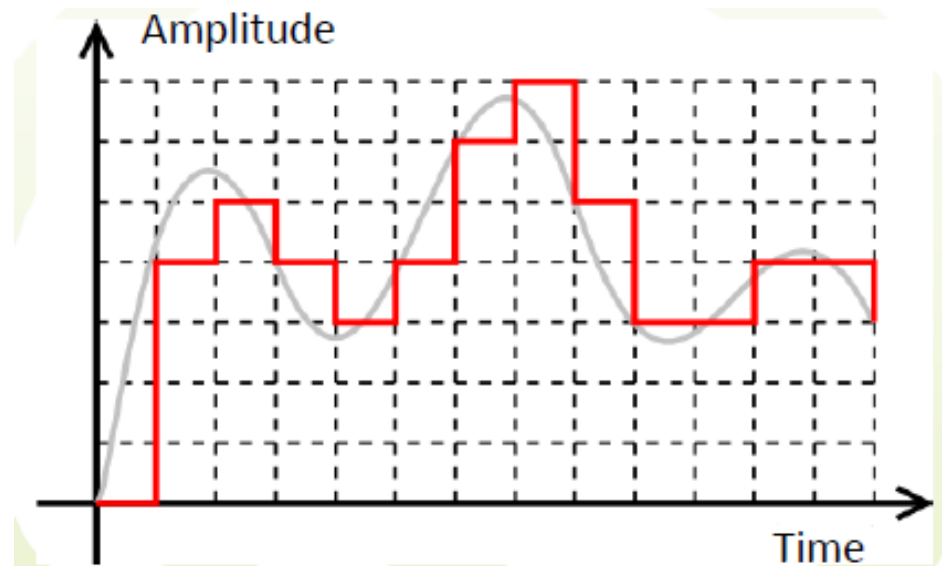


Dữ liệu âm thanh

Số hóa tín hiệu âm thanh:

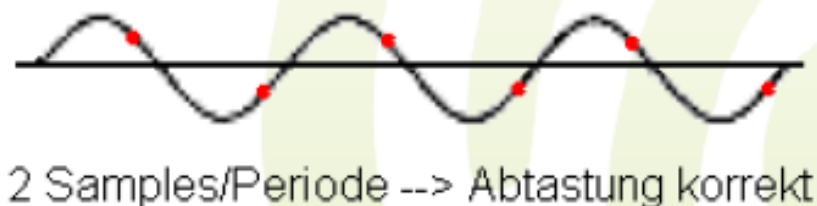
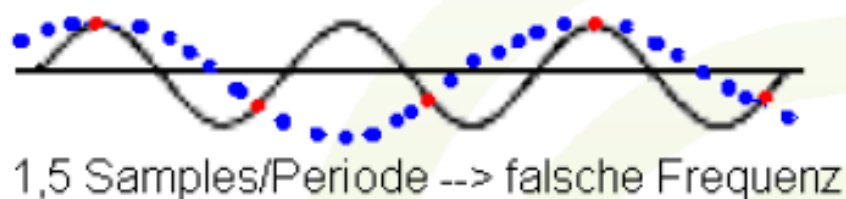
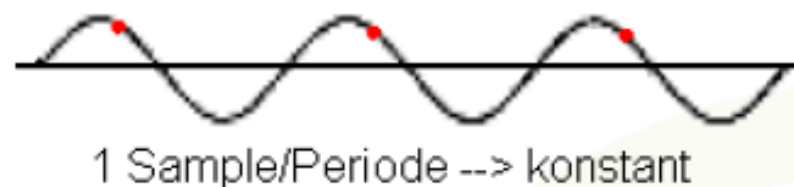
- ▶ Chuyển đổi tín hiệu từ liên tục sang gián đoạn
- ▶ Lấy mẫu: tần số lấy mẫu (audio CD: 44100Hz, Phone: 8000 Hz)
- ▶ Lượng tử hóa: độ phân giải (16 bits)

- Để truyền âm thanh audio CD stereo với mức lượng tử 16 bits thì cần tốc độ đường truyền là bao nhiêu?



Dữ liệu âm thanh

- Lý thuyết lấy mẫu:
(Nyquist, 1928) tần số lấy mẫu ít nhất phải bằng 2 lần tần số thực của tín hiệu.



Dữ liệu âm thanh

Nén tín hiệu:

- ▶ Tín hiệu chưa nén: AIFF (*.aif), Wave (*.wav), IRCAM (*.snd), AU (*.au),...
- ▶ Nén có suy hao: chuyển đổi Cosine rời rạc (MDCT), wavelets,...
- ▶ Mã hóa: chuyển sang miền tần số (lấy mẫu)
- ▶ Giải mã: khôi phục tín hiệu
- ▶ Lược bỏ thông tin:
 - ▶ Bỏ tần số cao/thấp
 - ▶ Giảm độ chính xác của thông tin ở tần số giữ lại
 - ▶ Sử dụng các hiệu ứng âm thanh khác để lược bỏ thông tin.

Dữ liệu âm thanh

Thông tin về âm thanh chứa trong CSDL

- ▶ Dữ liệu âm thanh
 - ▶ Nhạc, CD...
 - ▶ Các hiệu ứng âm thanh,...
- ▶ Tài liệu dạng âm thanh
 - ▶ Các bài diễn thuyết
 - ▶ Các bản thu âm hội thoại, cuộc gọi điện thoại, đàm phán.

Dữ liệu âm thanh

Ứng dụng trong hệ CSDL ĐPT

- ▶ Nhận diện các tín hiệu âm thanh (truy vấn)
- ▶ Phân loại / tìm kiếm các tín hiệu giống nhau
- ▶ Đồng bộ tín hiệu âm thanh

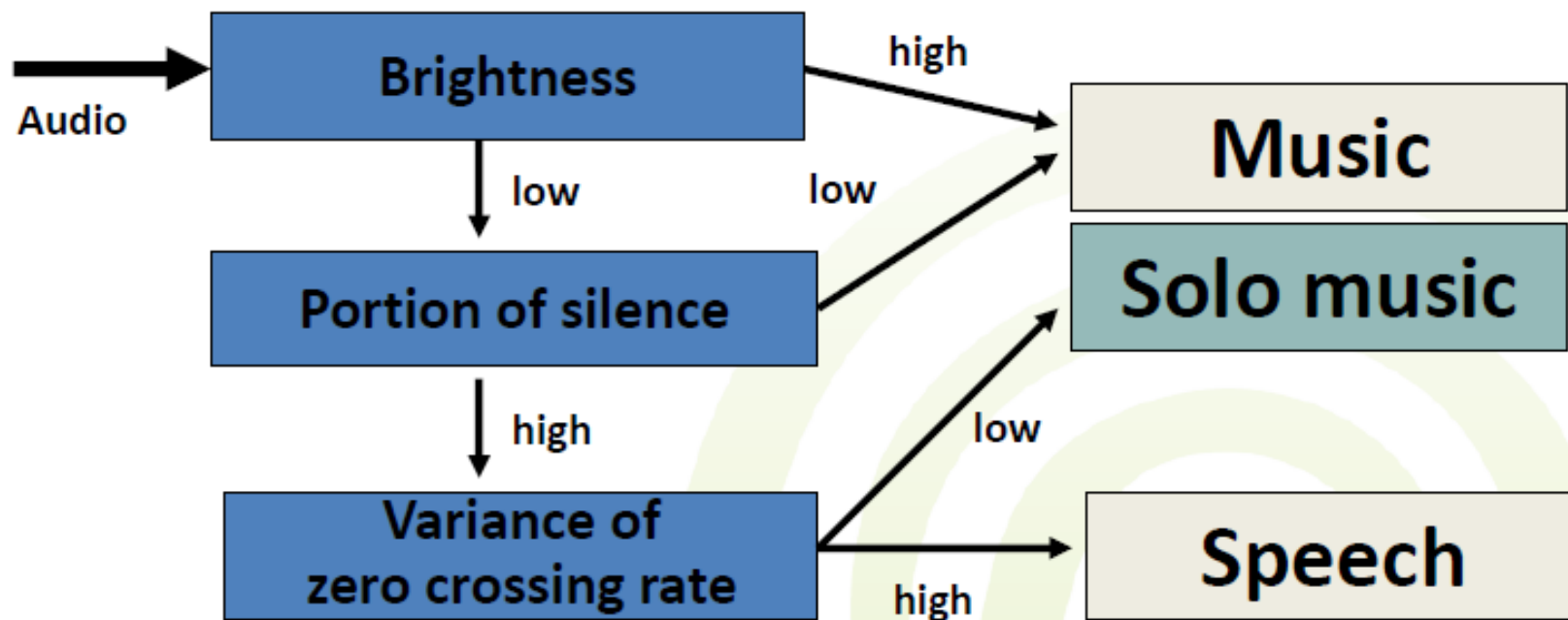
Dữ liệu âm thanh

Các thuộc tính bậc thấp:

- ▶ Cường độ âm thanh
- ▶ Phân bố tần số âm thanh
- ▶ Độ cao/thấp
- ▶ Độ rõ nét
- ▶ Năng lượng trung bình
- ▶
- ▶ Các thuộc tính bậc thấp được đo trong
 - ▶ Miền thời gian (cường độ tại từng thời điểm)
 - ▶ Miền tần số (năng lượng âm thanh tại từng tần số)

Dữ liệu âm thanh

- ▶ Ví dụ về phân loại dữ liệu âm thanh dựa trên các thuộc tính bậc thấp



Dữ liệu văn bản

- ▶ Văn bản đơn thuần: tổ hợp các ký tự.
 - ▶ Các bảng mã: ASCII code (8 bits)
 - ▶ Cần lưu trữ 300 trang văn bản, mỗi trang chứa 3000 ký tự thì cần dung lượng bao nhiêu?
- ▶ Văn bản có cấu trúc:
 - ▶ HyperText Markup Language (HTML)
 - ▶ Standard General Markup Language (SGML)
 - ▶ LaTeX
 - ▶ Office Document Architecture (ODA)
 - ▶ Portable Document Format (PDF)

Bảng mã ASCII

Dec	Hx	Oct	Char	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr
0	0	000	NUL (null)	32	20	040	 	Space	64	40	100	@	@	96	60	140	`	`
1	1	001	SOH (start of heading)	33	21	041	!	!	65	41	101	A	A	97	61	141	a	a
2	2	002	STX (start of text)	34	22	042	"	"	66	42	102	B	B	98	62	142	b	b
3	3	003	ETX (end of text)	35	23	043	#	#	67	43	103	C	C	99	63	143	c	c
4	4	004	EOT (end of transmission)	36	24	044	$	\$	68	44	104	D	D	100	64	144	d	d
5	5	005	ENQ (enquiry)	37	25	045	%	%	69	45	105	E	E	101	65	145	e	e
6	6	006	ACK (acknowledge)	38	26	046	&	&	70	46	106	F	F	102	66	146	f	f
7	7	007	BEL (bell)	39	27	047	'	'	71	47	107	G	G	103	67	147	g	g
8	8	010	BS (backspace)	40	28	050	((72	48	110	H	H	104	68	150	h	h
9	9	011	TAB (horizontal tab)	41	29	051))	73	49	111	I	I	105	69	151	i	i
10	A	012	LF (NL line feed, new line)	42	2A	052	*	*	74	4A	112	J	J	106	6A	152	j	j
11	B	013	VT (vertical tab)	43	2B	053	+	+	75	4B	113	K	K	107	6B	153	k	k
12	C	014	FF (NP form feed, new page)	44	2C	054	,	,	76	4C	114	L	L	108	6C	154	l	l
13	D	015	CR (carriage return)	45	2D	055	-	-	77	4D	115	M	M	109	6D	155	m	m
14	E	016	SO (shift out)	46	2E	056	.	.	78	4E	116	N	N	110	6E	156	n	n
15	F	017	SI (shift in)	47	2F	057	/	/	79	4F	117	O	O	111	6F	157	o	o
16	10	020	DLE (data link escape)	48	30	060	0	0	80	50	120	P	P	112	70	160	p	p
17	11	021	DC1 (device control 1)	49	31	061	1	1	81	51	121	Q	Q	113	71	161	q	q
18	12	022	DC2 (device control 2)	50	32	062	2	2	82	52	122	R	R	114	72	162	r	r
19	13	023	DC3 (device control 3)	51	33	063	3	3	83	53	123	S	S	115	73	163	s	s
20	14	024	DC4 (device control 4)	52	34	064	4	4	84	54	124	T	T	116	74	164	t	t
21	15	025	NAK (negative acknowledge)	53	35	065	5	5	85	55	125	U	U	117	75	165	u	u
22	16	026	SYN (synchronous idle)	54	36	066	6	6	86	56	126	V	V	118	76	166	v	v
23	17	027	ETB (end of trans. block)	55	37	067	7	7	87	57	127	W	W	119	77	167	w	w
24	18	030	CAN (cancel)	56	38	070	8	8	88	58	130	X	X	120	78	170	x	x
25	19	031	EM (end of medium)	57	39	071	9	9	89	59	131	Y	Y	121	79	171	y	y
26	1A	032	SUB (substitute)	58	3A	072	:	:	90	5A	132	Z	Z	122	7A	172	z	z
27	1B	033	ESC (escape)	59	3B	073	;	;	91	5B	133	[[123	7B	173	{	{
28	1C	034	FS (file separator)	60	3C	074	<	<	92	5C	134	\	\	124	7C	174	|	
29	1D	035	GS (group separator)	61	3D	075	=	=	93	5D	135]]	125	7D	175	}	}
30	1E	036	RS (record separator)	62	3E	076	>	>	94	5E	136	^	^	126	7E	176	~	~
31	1F	037	US (unit separator)	63	3F	077	?	?	95	5F	137	_	_	127	7F	177		DEL

Source: www.LookupTables.com

Bảng mã ASCII

128	Ç	144	É	160	á	176	░	192	Ł	208	⋈	224	α	240	≡
129	ü	145	æ	161	í	177	▒	193	⌈	209	⌋	225	β	241	±
130	é	146	Æ	162	ó	178	▓	194	⌊	210	⌌	226	Γ	242	≥
131	â	147	ô	163	ú	179		195	┐	211	⋈	227	π	243	≤
132	ä	148	ö	164	ñ	180	└	196	━	212	↳	228	Σ	244	∫
133	à	149	ò	165	Ñ	181	┌	197	┘	213	℞	229	σ	245	∫
134	â	150	û	166	²	182	▯	198	┐	214	⌈	230	μ	246	÷
135	ç	151	ù	167	°	183	▯	199	┐	215	⌈	231	τ	247	≈
136	ê	152	ÿ	168	¿	184	┐	200	⋈	216	⌈	232	Φ	248	°
137	ë	153	Ö	169	┐	185	▯	201	℞	217	┐	233	⊙	249	·
138	è	154	Ü	170	┐	186	▯	202	⋈	218	┐	234	Ω	250	·
139	ï	155	◊	171	½	187	┐	203	⌋	219	■	235	δ	251	√
140	î	156	£	172	¼	188	┐	204	┐	220	■	236	∞	252	π
141	ì	157	¥	173	¡	189	┐	205	=	221	■	237	φ	253	²
142	Ä	158	£	174	«	190	┐	206	┐	222	■	238	ε	254	■
143	Å	159	f	175	»	191	┐	207	⋈	223	■	239	∩	255	

Source: www.LookupTables.com

Dữ liệu văn bản

- ▶ Nén văn bản: thu gọn kích thước lưu trữ mà không làm mất thông tin
 - ▶ Mã hóa Huffman
 - ▶ Dúng ít bits để mã hóa các ký tự thường xuyên xuất hiện,
 - ▶ Dùng nhiều bit để mã hóa các ký tự ít xuất hiện,
 - ▶ Cần có codebook

<i>Symbol</i>	<i>Probability</i>	<i>Code</i>
<i>e</i>	0.8	1
<i>t</i>	0.16	01
<i>x</i>	0.02	001
<i>z</i>	0.02	000

Dữ liệu văn bản

- ▶ Mã hóa run-length

- ▶ Thay thế một chuỗi ký tự trùng lặp bằng một ký tự cùng với số lần trùng lặp.

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

- ▶ Cần thiết lập ngưỡng số lần lặp.
 - ▶ Phát sinh ký tự mới.
 - ▶ VD: “MMMMannnnnnnnn” (13 ký tự) -> “@M4a@n8” (7 ký tự)

Dữ liệu văn bản

- ▶ Mã hóa LZW: sử dụng từ điển của các cụm ký tự. Lưu trữ mã của các cụm ký tự thay cho mã của từng ký tự.
 - ▶ VD: văn bản có 10,000 ký tự bao gồm 2000 từ, trong đó có 500 từ khác nhau: ta mã hóa 500 từ thay cho mã hóa 10,000 ký tự.

Dữ liệu video

- ▶ Yêu cầu đối với CSDL video:
 - ▶ Lưu trữ
 - ▶ Có thể truy nhập được dữ liệu
 - ▶ Có thể khôi phục / trình chiếu video
- ▶ Những CSDL video phổ biến hiện nay
 - ▶ Blobs (binary large objects)
 - ▶ Tra cứu bằng siêu dữ liệu
 - ▶ Cắt thành các khung / đoạn



Dữ liệu video

- ▶ Các công cụ hỗ trợ
 - ▶ IBM AIV Extenders dành cho IBM DB2 UDB: tích hợp dữ liệu video vào các dữ liệu thương mại trên cùng một câu truy vấn.
 - ▶ VD: tra cứu tất cả các đoạn video có cùng một chủ đề và sắp xếp theo thời gian quay video.
 - ▶ Tích hợp cấu hình QBIC vào CSDL thương mại

Dữ liệu video

- ▶ Dữ liệu video được lưu trữ ra sao?
 - ▶ Các thuộc tính của video được đưa vào/ lấy ra khỏi CSDL
 - ▶ CSDL chỉ lưu trữ và bảo quản các thuộc tính của video: tốc độ khung hình, khuôn mẫu nén hình ảnh, số lượng các điểm dừng khung hình,...
 - ▶ Video được tra cứu dựa trên các thuộc tính đã lưu trữ: tên, số thứ tự, nội dung, hoặc các thuộc tính định dạng.
 - ▶ Công cụ hỗ trợ quay video

Dữ liệu video

- ▶ Các vấn đề cần giải quyết:
 - ▶ Phương tiện truyền dẫn liên tục
 - ▶ Bao gồm một vài đường truyền
 - ▶ Truyền các chuỗi hình ảnh
 - ▶ Truyền cả âm thanh
 - ▶ Truyền cả văn bản

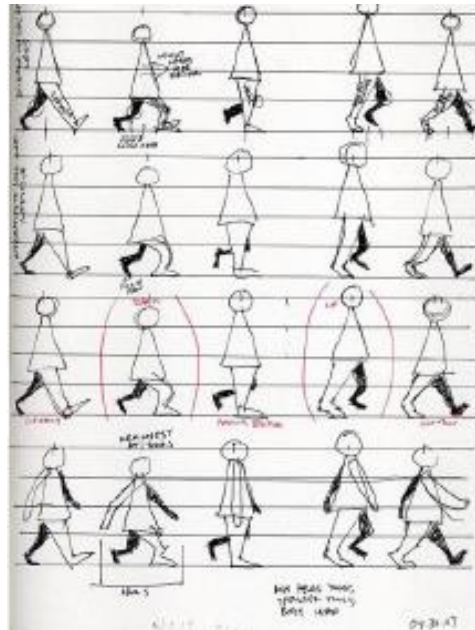


Dữ liệu video

► Các vấn đề cần giải quyết

► Tổ chức thông tin

- Video truyền thông tin có cấu trúc cả về không gian và thời gian
- Video là một dạng tài liệu
- Video bao gồm nhiều phần có cấu trúc khác nhau



Dữ liệu video

- ▶ Tra cứu video là sự tổng hợp của tất cả các công cụ tra cứu:
 - ▶ Hình ảnh
 - ▶ Âm thanh
 - ▶ Văn bản

Dữ liệu video

- ▶ Các thuộc tính sử dụng để tra cứu video
 - ▶ Nhận dạng vật thể dựa trên phân đoạn và phụ đề
 - ▶ Giọng nói, âm thanh
 - ▶ Hình dáng và nội dung thoại
 - ▶ Các ngữ cảnh cụ thể (khán giả, không gian,...)
 - ▶ Hướng/tốc độ di chuyển của vật thể
 - ▶ Sự dịch chuyển vị trí/góc chiếu máy quay
 - ▶ Quan hệ không gian, thời gian
 - ▶ ...

