

Các kỹ thuật chỉ số hóa và truy vấn dữ liệu ảnh

Nguyễn Đình Hóa
dinhhoa@gmail.com 0942807711

Tóm tắt nội dung bài 10

- ▶ Các mô hình truy vấn dữ liệu văn bản
 - ▶ Mô hình truy vấn Boolean trên tệp ngược
 - ▶ Mô hình truy vấn không gian vector
 - ▶ Mô hình truy vấn phân cụm
 - ▶ Các phương pháp phân cụm dữ liệu
 - ▶ Ứng dụng truy vấn dữ liệu trong lưu trữ phân cụm dữ liệu

Chỉ số hóa và truy vấn dữ liệu ảnh

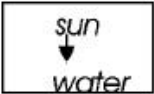












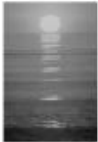













► Các kỹ thuật chính

- Dựa trên thuộc tính chung của ảnh và lưu trữ dưới dạng bảng và truy vấn bằng mô hình dữ liệu quan hệ.
 - VD: tên ảnh, loại ảnh, ngày tạo ảnh, chủ đề, tác giả,...
 - Nhược điểm: khó lưu trữ nội dung chính xác của ảnh
- Dựa trên sự kết hợp giữa việc trích xuất các thuộc tính của ảnh với nhận diện các thực thể trong ảnh. Nhược điểm: độ tính toán phức tạp.
- Sử dụng văn bản (text) để mô tả ảnh, sau đó áp dụng hệ thống IR để lưu trữ và tra cứu ảnh
- Sử dụng các thuộc tính bậc thấp (màu sắc, ngữ cảnh) để đánh chỉ số và truy vấn dữ liệu

Kỹ thuật thứ 2 và thứ 4 được gọi là truy vấn dữ liệu ảnh dựa trên nội dung.

Các thuộc tính bậc thấp của ảnh

- ▶ Màu sắc, hình dạng các vật thể, màu nền
- ▶ Các thông số diễn đạt kích thước vật thể
 - ▶ VD: Hình chữ nhật $XLB = 80$; $XUB = 120$; $YLB = 20$; $YUB = 55$.
- ▶ Các thông số diễn đạt giá trị màu của điểm ảnh
 - ▶ VD: tại điểm ảnh có tọa độ $(200, 150)$ có các giá trị màu sắc là $Red = 12$; $Green = 7$; $Blue = 8$.

	Example query	Example query result
exact	Spatial predicate 	  
	Image predicate <i>Amount of "sky" > 20% and amount of "sand" > 30%</i>	  
	Group predicate <i>Location = "Africa"</i>	  
approximate	Spatial example 	   
	Image example 	   
	Group example <i>pos</i>  <i>neg</i>  	   

Truy vấn dữ liệu ảnh dựa trên văn bản

- ▶ Ảnh được miêu tả bằng các nội dung văn bản tự do
- ▶ Câu truy vấn bao gồm các từ khóa hoặc văn bản tự do có kèm hoặc không kèm các phép toán logic.
- ▶ Kết quả truy vấn dựa trên so sánh sự tương đồng giữa câu truy vấn và nội dung lưu trữ (kỹ thuật IR)

Truy vấn dữ liệu ảnh dựa trên văn bản

- ▶ Văn bản dùng để miêu tả ảnh phải được nhập vào bằng tay (hoặc sử dụng phụ đề có sẵn)
 - ▶ Làm sao để nhập các nội dung miêu tả ảnh một cách chính xác và thống nhất
 - ▶ Cần sử dụng các thông tin về ngữ cảnh, hoặc nhiều từ đồng nghĩa để tăng thông tin về ảnh
 - ▶ Sự kết hợp các từ khác nhau cũng mang nhiều thông tin khác nhau
- ▶ Nội dung văn bản đôi khi không bao quát hết ý nghĩa của ảnh, hoặc đôi khi quá cụ thể.
 - ▶ Quá trình truy vấn cần có phản hồi.

Truy vấn dữ liệu ảnh dựa trên văn bản

- ▶ Có khả năng chứa đựng thông tin về ngữ nghĩa (các thuộc tính bậc cao) một cách chính xác, chi tiết. VD: các thông tin về tâm trạng như “cười”, “hạnh phúc”...
- ▶ Nhược điểm: Một vài thuộc tính cấp thấp khó có thể diễn tả bằng văn bản như nền ảnh, hoặc các vật thể có hình dạng đặc biệt.
- ▶ Truy vấn bằng văn bản khó có thể áp dụng trong truy vấn ảnh trực tiếp. VD: “hãy tìm các hình ảnh giống với bức ảnh này...”

Truy vấn dữ liệu ảnh dựa trên màu sắc

► Kỹ thuật cơ bản

- Mỗi ảnh trong CSDL được diễn tả bằng các màu cơ bản (RGB)
- Mỗi dải màu cơ bản được chia nhỏ thành m mức. Tổng số sẽ có $n = m^3$ mức màu tương ứng với số ô màu
- Xây dựng biểu đồ tần suất màu
 $H(M) = [h_1, h_2, \dots, h_j, \dots, h_n]$ với mỗi h_j là số điểm ảnh mức màu j .
- $H(M)$ được lưu trữ và truy vấn sử dụng các kỹ thuật truy vấn dựa trên không gian vector, hoặc dựa trên phân cụm.

Truy vấn dữ liệu ảnh dựa trên màu sắc

- ▶ Các thước đo về độ tương tự, hoặc các thước đo khoảng cách (L1-norm, L2-norm) giữa các biểu đồ tần suất được sử dụng.
- ▶ VD: trong 3 bức ảnh xám 8x8 điểm, giả sử ta sử dụng 3 bit để mã hoá các mức xám:

$$H_1 = [8, 8, 8, 8, 8, 8, 8, 8]$$

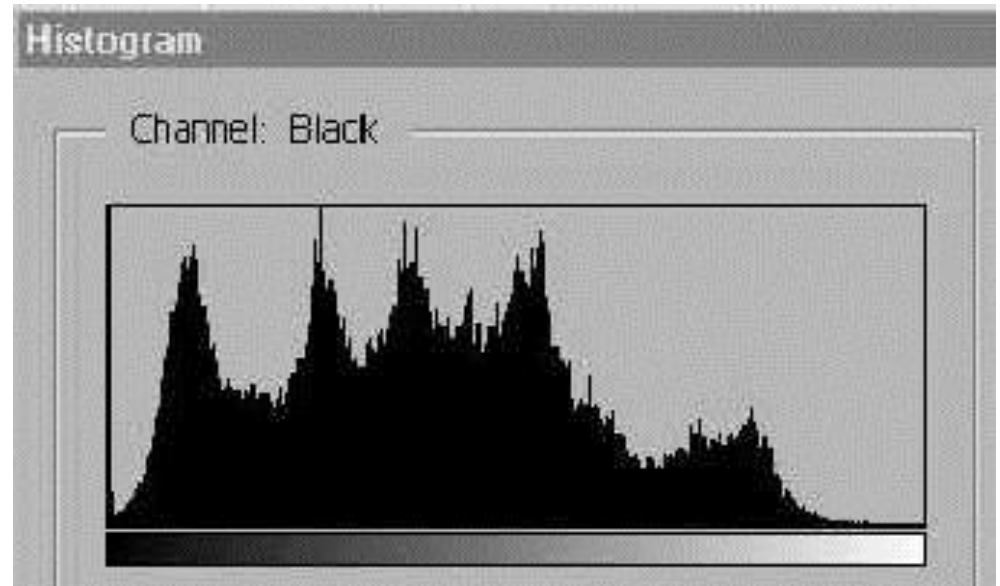
$$H_2 = [6, 6, 8, 8, 10, 10, 7, 9]$$

$$H_3 = [10, 10, 2, 2, 10, 10, 10, 10]$$

Hai bức ảnh nào gần nhau nhất theo công thức tính độ tương tự tổng quát, hoặc công thức tính khoảng cách L2-norm?

Truy vấn dữ liệu ảnh dựa trên màu sắc

► Ví dụ:

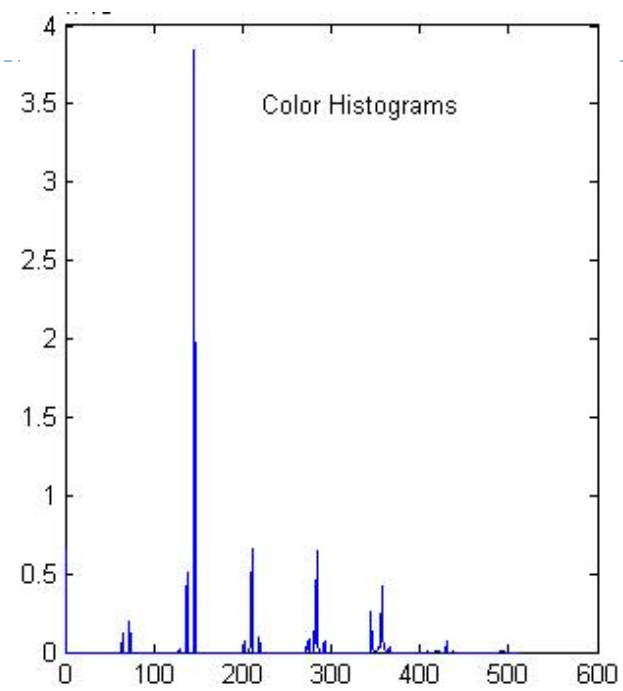
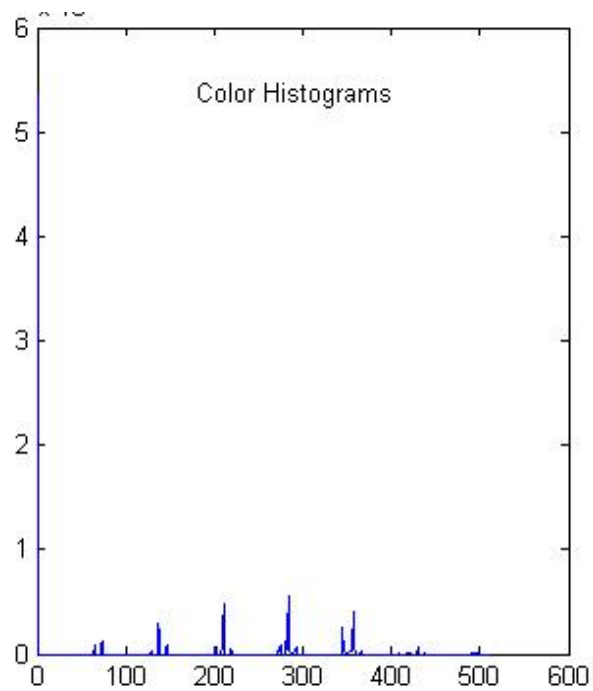


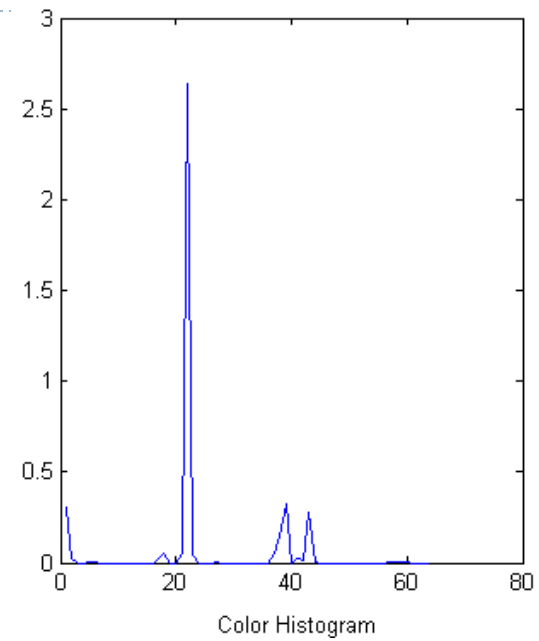
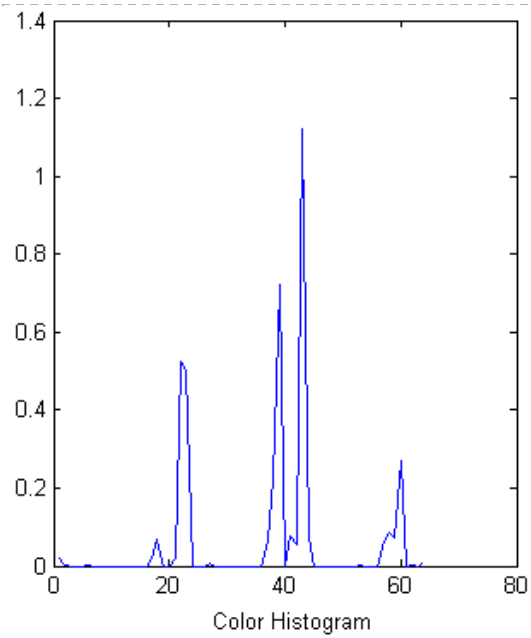
Truy vấn dữ liệu ảnh dựa trên màu sắc

► Các kỹ thuật nâng cao

► Sử dụng sự tương đồng màu sắc:

- Giả sử mỗi ảnh bao gồm N điểm ảnh, nếu sử dụng L1-norm, khoảng cách tối đa giữa 2 ảnh sẽ là $2N$, xảy ra khi hai ảnh đó không có bất kỳ điểm chung về giá trị màu nào.
- Trên thực tế, mắt người vẫn cảm nhận được sự tương đồng về màu sắc mặc dù giá trị màu của chúng khác nhau.
- Đôi khi giá trị màu sắc bị thay đổi do nhiễu, dao động về độ sáng,... dẫn đến hai màu tương tự nhau có thể thuộc hai mức màu khác nhau. Hoặc đôi khi trong quá trình truy vấn, ta cần tìm các ảnh có màu sắc tương tự nhau chứ không nhất thiết phải có các giá trị màu bằng nhau.





Truy vấn dữ liệu ảnh dựa trên màu sắc

- ▶ Sử dụng biểu đồ tần suất lũy tích $CH(M) = [ch_1, ch_2, \dots, ch_n]$ trong đó $ch_i = \sum_{j \leq i} h_j$
 - Nhược điểm: không diễn đạt được chính xác độ tương tự giữa các màu sắc
- ▶ Sử dụng công thức tính sự tương tự dựa trên màu sắc
 - Giả sử X và Y là biểu đồ tần suất của 2 ảnh. Z là vector hiệu số giữa X và Y. Độ tương đồng giữa X và Y được tính bằng $\|Z\| = Z^T A Z$, trong đó: $a_{i,j} = 1 - d(c_i, c_j) / d_{\max}$, là khoảng cách màu trong không gian màu MTM.
- ▶ Sử dụng biểu đồ tần suất theo trọng số (PWH)
 - Ứng với từng ô màu trong cách tính biểu đồ tần suất, người ta chọn ra một màu đại diện.
 - 10 màu đại diện gần nhất được chọn làm tâm.
 - Tính khoảng cách giữa màu của từng pixel đến các màu đại diện đó.
 - Trọng số của 10 màu đại diện cho mỗi điểm ảnh sẽ được tính bằng nghịch đảo của các khoảng cách đó.

Truy vấn dữ liệu ảnh dựa trên màu sắc

- ▶ Nhược điểm của phương pháp truy vấn ảnh dựa trên màu sắc là không tính đến mối liên hệ về không gian giữa các điểm ảnh.
- ▶ Sử dụng mối liên hệ về không gian giữa các điểm ảnh
 - ▶ Ảnh được chia thành nhiều vùng nhỏ cố định
 - ▶ Tại mỗi vùng ảnh, biểu đồ tần suất màu được thiết lập
 - ▶ Trong quá trình truy vấn, biểu đồ tần suất màu tại các vùng tương ứng sẽ được so sánh với nhau.



Truy vấn dữ liệu ảnh dựa trên màu sắc

- ▶ Nhược điểm của phương pháp này là sự ảnh hưởng của hiệu ứng che mặt.
 - Các ô màu lớn sẽ lấn át nội dung của ô màu nhỏ.
 - Các ảnh có màu nền giống nhau sẽ mang nội dung giống nhau, mặc dù các vật thể trong đó là khác hẳn nhau.
- ▶ Có thể sử dụng biểu đồ tần suất màu sắc riêng biệt cho hình nền và các vật thể: phân chia ảnh thành nhiều khung chữ nhật, mỗi khung chứa riêng vật thể hoặc hình nền.
- ▶ Việc xác định vật thể hay hình nền được thực hiện dựa trên sự biến thiên giá trị màu theo trục ngang hay trục dọc
- ▶ Mỗi ảnh được đại diện bởi hai biểu đồ tần suất.
- ▶ Quá trình truy vấn có thể thực hiện theo một trong các cách:
 - Dựa vào biểu đồ tần suất màu tổng thể của ảnh
 - Dựa vào biểu đồ tần suất của các vật thể trong ảnh
 - Dựa vào biểu đồ tần suất màu của hình nền
 - Dựa vào biểu đồ tần suất màu của cả hình nền lẫn vật thể (có trọng số)

Truy vấn dữ liệu ảnh dựa trên màu sắc

- ▶ Sử dụng thông tin về thống kê của sự phân bố màu sắc (giống lượng tử hoá phi tuyến): dải màu có phân bố nhiều điểm ảnh hơn sẽ được phân chia nhỏ hơn.
- ▶ Sử dụng các cách diễn đạt màu sắc rõ ràng hơn
 - ▶ Hệ RGB có một số nhược điểm:
 - Phụ thuộc vào thiết bị định dạng màu (nguyên tắc phối màu)
 - Không nhất quán trong biểu diễn màu giống nhau.
 - ▶ Thay đổi không gian màu sắc (color space) từ RGB sang CIE Luv, CIE Lab, HSI (hue, saturation, intensity), HVC (hue, value, chroma)
 - ▶ Thay đổi cách diễn đạt ảnh: các thiết bị chụp ảnh khác nhau sẽ có các thông số điều chỉnh khác nhau đối với ảnh. Cùng một ảnh thu từ nhiều nguồn khác nhau sẽ có các thông số màu khác nhau

Truy vấn dữ liệu ảnh dựa trên hình dạng

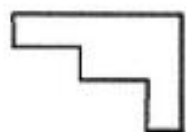
- ▶ Ảnh cần được chia thành các mảnh chứa vật thể
 - ▶ Mỗi hình ảnh vật thể cần phải được diễn tả thống nhất, không phụ thuộc vào sự biến đổi của sự vật, vị trí, kích thước,...
 - ▶ Các hình giống nhau cần phải được diễn đạt giống nhau
- ▶ Trong quá trình truy vấn, người dùng sử dụng một hình ảnh mẫu hoặc tự vẽ ra một hình mẫu để hệ thống tìm kiếm các ảnh có hình giống như vậy.
- ▶ Kết quả trả về thường là nhiều ảnh, sắp xếp theo thứ tự giảm dần về sự giống nhau với hình mẫu.

Truy vấn dữ liệu ảnh dựa trên hình dạng

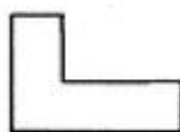
- ▶ Các khái niệm chung và một số thước đo về hình dạng đơn giản.
 - ▶ Trục chính: đường thẳng nối hai điểm xa nhất trong hình
 - ▶ Trục phụ: đường thẳng vuông góc với trục chính, có độ dài sao cho một hình chữ nhật với các cạnh song song và dài bằng trục chính và trục phụ chứa được toàn bộ đường bao của hình.
 - ▶ Hình chữ nhật cơ bản: được hình thành như mô tả ở trên
 - ▶ Tâm sai / độ lệch tâm: tỉ lệ về độ dài giữa trục chính và trục phụ
 - ▶ Những thông tin cơ bản nói trên có thể được sử dụng để truy vấn hình ảnh.

Truy vấn dữ liệu ảnh dựa trên hình dạng

- ▶ Diễn tả hình dạng theo khu vực và các tiêu chí so sánh tương tự
 - ▶ Khái niệm cơ bản của diễn tả hình dạng theo khu vực:
 - ▶ Ứng với mỗi hình thể nào đó, ta đặt một lưới (gồm các hình vuông có kích thước cố định) lên đó
 - ▶ Ứng với mỗi hình vuông có trên 15% là diện tích vật thể, ta gán giá trị 1
 - ▶ Các hình vuông còn lại có giá trị là 0.
 - ▶ Đọc giá trị của các hình vuông từ trái sang phải, từ trên xuống dưới để có một chuỗi nhị phân đại diện cho hình thể đó.
 - ▶ Chuẩn hóa xoay vòng
 - ▶ Chuẩn hóa thước đo
 - ▶ Diễn tả hình dạng thống nhất – đánh chỉ số hình dạng
 - ▶ Các thước đo về sự tương tự.
 - ▶ Một vài kỹ thuật khác thực hiện trên hình dạng



1



2



3



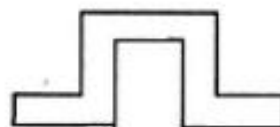
4



5



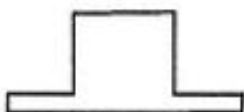
6



7



8



9



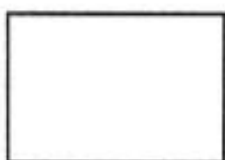
10



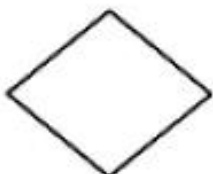
11



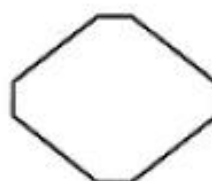
12



13



14



15



16



17



18

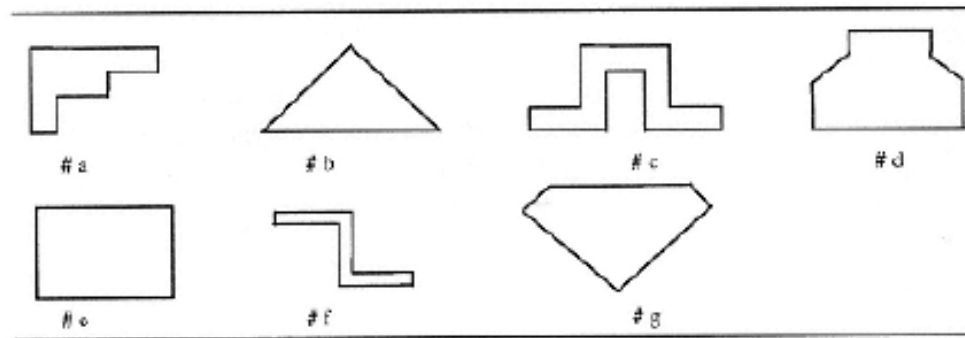


19



20

Hình ảnh truy vấn



Kết quả so sánh khoảng cách

	a	b	c	d	e	f	g
1	0	104	46	76	37	44	65
2	33	65	69	80	38	61	56
3	17	81	33	96	48	37	60
4	104	0	87	76	71	100	37
5	72	62	75	41	53	78	61
6	98	55	104	50	110	103	77
7	46	87	0	86	79	44	60
8	54	88	32	117	89	40	69
9	32	84	34	101	55	52	65
10	76	76	86	0	72	93	76
11	40	76	50	89	47	64	59
12	105	68	121	65	69	90	80
13	37	71	79	72	0	71	66
14	108	55	116	64	74	91	69
15	123	71	132	62	88	137	62
16	44	100	44	93	71	0	63
17	54	92	71	109	49	50	91
18	54	51	52	54	62	69	36
19	65	37	60	76	66	63	0
20	91	38	109	65	81	86	70

Truy vấn dữ liệu ảnh dựa trên hình dạng

- ▶ Các moment bất biến (invariant moments)

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y)$$

- ▶ Hệ số Fourier
- ▶ Biểu đồ tần suất màu của các nét chính
- ▶ Danh mục các điểm ảnh trọng tâm
- ▶ So sánh hình mẫu mềm dẻo

Truy vấn ảnh dựa trên bố cục

- ▶ Các thuộc tính chủ yếu liên quan đến bố cục ảnh
 - ▶ Độ thô của ảnh (ngược với độ mịn): sự khác biệt giữa các phần tử của ảnh
 - ▶ Độ tương phản: khoảng động của dải màu xám, sự phân cực giữa hai màu sáng và tối trong biểu đồ tần suất, độ sắc của các nét, tần suất lặp lại của các khối màu
 - ▶ Hình dạng và chiều sắp xếp của các hình trong ảnh
 - ▶ Sự tương đồng của các đường viền
 - ▶ Tính trật tự của ảnh
 - ▶ Bố trí màu sắc trong ảnh (liên tục hay gián đoạn)

Truy vấn dựa trên ảnh nén

- ▶ Đánh chỉ số và truy vấn dựa trên các hệ số DCT
- ▶ Đánh chỉ số và truy vấn dựa trên các hệ số FFT
- ▶ Đánh chỉ số và truy vấn dựa trên các hệ số Wavelet
- ▶ Đánh chỉ số và truy vấn dựa trên dữ liệu nén VQ