

## TRUY TÌM ẢNH DỰA VÀO NỘI DUNG VỚI PHƯƠNG PHÁP ĐÁNH VẤN ẢNH

Nguyễn Lâm<sup>(1)</sup>, Lý Quốc Ngọc<sup>(2)</sup>, Dương Anh Đức<sup>(2)</sup>

Nguyễn Bá Công<sup>(2)</sup>, Nguyễn Hữu Đức<sup>(2)</sup>

(1) Trường Đại Học Lương Thế Vinh, Nam Định

(2) Trường Đại Học Khoa Học Tự Nhiên-ĐHQG-HCM

(Bài nhận ngày 31 tháng 01 năm 2005, hoàn chỉnh sửa chữa ngày 18 tháng 7 năm 2005)

**TÓM TẮT:** Hiện nay trên thế giới, việc truy tìm ảnh đã bước sang thời kỳ mới, thời kỳ truy tìm ảnh dựa vào nội dung. Giải quyết bài toán truy tìm ảnh dựa vào nội dung là một vấn đề thách thức cần được vượt qua. Vấn đề cốt yếu để thiết kế hệ thống truy tìm ảnh hữu hiệu và thân thiện với người dùng là cần xây dựng phương pháp nhập liệu gắn gũi với cách nghĩ của con người, đồng thời đạt được mức độ chính xác cao trong kết quả truy tìm và mức độ tìm sót thấp. Chúng tôi đã giải quyết bài toán này bằng việc khám phá các đặc trưng cấp cao tiềm ẩn trong cơ sở dữ liệu ảnh. Các đặc trưng này được rút trích thông qua việc phân đoạn ảnh thành các vùng và phân lớp chúng thành các lớp với các phần tử đại diện mà chúng tôi gọi là vắn của ảnh. Chúng tôi đã dùng giải thuật phân lớp phân cấp trong phân đoạn ảnh, phân lớp vùng và phân lớp ảnh. Dữ liệu nhập có thể là toàn bộ ảnh, một số vùng của ảnh, các phần tử đại diện cho các lớp vùng. Việc truy tìm được tiến hành dựa trên cấu trúc cây phân cấp của các lớp vùng hoặc cấu trúc cây phân cấp của các lớp ảnh. Kết quả thực nghiệm cho thấy hệ thống có độ chính xác và độ trung thực cao, thân thiện với người dùng.

### 1. Giới thiệu

Khi sử dụng các hệ truy tìm ảnh dựa vào nội dung, con người luôn mong muốn có thể thể hiện được dễ dàng, tương đối đầy đủ các khái niệm về nội dung mà họ muốn tìm kiếm. Một cách truyền thống là họ chọn một ảnh truy vấn phù hợp với yêu cầu, tuy nhiên có thể họ không có sẵn bức ảnh truy vấn mà chỉ có một số khái niệm nào đó về bức ảnh (ví dụ như họ cần tìm các bức ảnh có dáng vẻ như : cỏ non xanh tận chân trời, cành lê trắng điểm một vài bông hoa), hoặc nhiều khi có được bức ảnh truy vấn nhưng họ chỉ muốn tìm các ảnh có chứa các khái niệm gần giống với một số đối tượng nào đó trong ảnh.

Chúng tôi đã thiết kế hệ thống truy tìm ảnh với mô hình dựa trên khái niệm, các khái niệm ở đây là các vắn của ảnh. Dữ liệu nhập có thể là toàn bộ ảnh (chứa đựng các đặc trưng cấp cao được tự động rút trích), một số vùng của ảnh truy vấn, hoặc một số phần tử đại diện của các lớp vùng. Việc truy tìm được thực hiện dựa trên cấu trúc cây phân cấp của các vùng (CTCPCV) hoặc có thể kết hợp với cấu trúc cây phân cấp ảnh (CTCPCA).

Hệ thống của chúng tôi được thiết kế gồm 4 giai đoạn chính sau :

Giai đoạn 1 : Phân đoạn ảnh.

Giai đoạn 2 : Phân lớp vùng.

Giai đoạn 3 : Phân lớp ảnh.

Giai đoạn 4 : Truy tìm ảnh dựa trên cấu trúc cây phân cấp vùng hoặc ảnh.

Giải thuật chính được dùng trong giai đoạn phân đoạn và phân lớp là giải thuật phân lớp phân cấp. Mỗi ảnh được phân đoạn thành cấu trúc cây phân cấp của vùng từ thô đến

mịn. Các vùng của cơ sở dữ liệu ảnh được phân lớp theo đặc trưng vùng thành cấu trúc cây phân cấp của các lớp vùng. Các ảnh gồm các vùng, các vùng này có thể được thay thế bởi phần tử đại diện của lớp vùng, và các ảnh được phân lớp thành cấu trúc cây phân cấp của lớp ảnh.

Cách tiếp cận này có một số ưu điểm sau :

Người dùng chưa có ý niệm rõ lắm về ảnh cần tìm mà chỉ có một số yêu cầu nào đó về các vẫn cấu thành nên nó hoặc họ chỉ cần tìm các ảnh tương tự với các vẫn chính của ảnh truy vấn chứ không cần sự tương tự toàn cục.

Giải quyết bài toán phân lớp ảnh với nội dung phong phú. Bài toán phân lớp ảnh được qui về bài toán phân lớp các vẫn cấu thành nên ảnh, hoặc phân lớp ảnh gồm các khái niệm của vùng .

Tránh được việc truy tìm tuần tự trên cơ sở dữ liệu lớn. Các ảnh được xét chỉ là lớp các ảnh có các vẫn thuộc lớp các vẫn truy vấn.

Tăng độ chính xác và độ trung thực trong kết quả truy tìm.

## 2. Phân đoạn ảnh

Mỗi ảnh được phân đoạn dựa trên giải thuật phân lớp phân cấp (PLPC) với các vector đặc trưng của các khối điểm ảnh kích thước 4x4 và độ đo tính dị biệt giữa các khối điểm ảnh.

Mỗi ảnh A được chia nhỏ thành các khối con  $K_i$  kích thước 4x4 (điểm ảnh), các đặc trưng trong mỗi khối con gồm 3 đặc trưng màu và 3 đặc trưng về vẫn.

### Đặc trưng màu

Chúng tôi sử dụng hệ màu HSI (Hue,Saturation,Intensity) cùng với việc định lượng không gian màu gồm 12 H, 3 S, 3 I. Mỗi khối con gồm 3 đặc trưng màu :  $M_1, M_2, M_3$ .

$$M_1 = \frac{1}{16} \sum_{i=0}^{15} H[i], \quad M_2 = \frac{1}{16} \sum_{i=0}^{15} S[i], \quad M_3 = \frac{1}{16} \sum_{i=0}^{15} I[i],$$

trong đó  $H[i], S[i], I[i]$  là thành phần  $H, S, I$  của điểm ảnh thứ  $i$

### Đặc trưng vẫn

Mỗi khối con gồm 3 đặc trưng vẫn :  $V_1, V_2, V_3$ , thực hiện phép biến đổi Haar wavelet [8] trên khối con 4x4 đối với thành phần màu  $I$ , ta được 4 băng tần số, mỗi băng chứa 2\*2 hệ số, giữ lại băng tần số HL, LH và HH.

Gọi 4 hệ số của khối qua phép biến đổi thành băng tần số HL là  $\{hl_0, hl_1, hl_2, hl_3\}$ ,

$$\text{Đặc trưng } V_1 \text{ trong băng tần số HL là } V_1 = \left(\frac{1}{4} \sum_{i=0}^3 hl_i\right)^{1/2} .$$

Gọi 4 hệ số của khối qua phép biến đổi thành băng tần số LH là  $\{lh_0, lh_1, lh_2, lh_3\}$ ,

$$\text{Đặc trưng } V_2 \text{ trong băng tần số LH là } V_2 = \left(\frac{1}{4} \sum_{i=0}^3 lh_i\right)^{1/2} .$$

Gọi 4 hệ số của khối qua phép biến đổi thành băng tần số HH là  $\{hh_0, hh_1, hh_2, hh_3\}$ ,

$$\text{Đặc trưng } V_3 \text{ trong băng tần số HH là } V_3 = \left(\frac{1}{4} \sum_{i=0}^3 hh_i\right)^{1/2} ,$$

### Đặc trưng vị trí

Mỗi khối con  $i$  có tâm  $C_i$  tọa độ  $x_i, y_i$ .

Việc đưa thêm đặc trưng vị trí giúp phân đoạn được các vùng có tính liên thông.

### Chuẩn hóa các đặc trưng màu, vân, vị trí

Gọi  $f_i$  là vector đặc trưng của khối con  $K_i$  của ảnh,

$$f_i = \{M_{i1}, M_{i2}, M_{i3}, V_{i1}, V_{i2}, V_{i3}, x_i, y_i\},$$

$M_{ij}, j \in [1..3]$  là các đặc trưng màu đã được chuẩn hóa của khối con  $K_i$ ,

$V_{ij}, j \in [1..3]$  là các đặc trưng vân đã được chuẩn hóa của khối con  $K_i$ ,

$x_i, y_i$  là các đặc trưng vị trí đã được chuẩn hóa của khối con  $K_i$ ,

### Độ đo tính dị biệt giữa các khối con

$$d_K(K_i, K_j) = (w_m \sum_{l=1}^3 (M_{il} - M_{jl})^2 + w_v \sum_{l=1}^3 (V_{il} - V_{jl})^2 + w_t ((x_i - x_j)^2 + (y_i - y_j)^2))^{1/2},$$

trong đó:

$M_{il}$  là đặc trưng màu thứ  $l$  của khối  $K_i$ ,

$V_{il}$  là đặc trưng vân thứ  $l$  của khối  $K_i$ ,

$x_i, y_i$  là tọa độ trọng tâm của khối  $K_i$ ,

$w_m, w_v, w_t$  là trọng số xác định mức độ quan trọng của đặc trưng màu, vân và vị trí

### Giải thuật phân lớp phân cấp

Giả sử  $A = \{K_i, i = 1, 2, \dots, N\}$ , với  $A$  là ảnh,  $K_i$  là các khối con thứ  $i$  của ảnh.

Giai đoạn khởi động

B1. Tạo phân cấp  $R_0 = \{K_i, i = 1, 2, \dots, N\}$ , với  $C_i = K_i$

B2.  $P_0 = P(K)$

Trong đó  $P(K)$  là ma trận kích thước  $N \times N$  với  $P(i, j) = d(K_i, K_j)$ .

B3.  $t=0$

Giai đoạn phân lớp

B4.  $t=t+1$

B5. Chọn cặp  $C_i, C_j$  sao cho  $d(C_i, C_j) = \min d(C_r, C_s), r, s = 1, 2, \dots, N, r \neq s$

B6. Kết nạp  $C_i, C_j$  vào lớp  $C_q$  và tạo một phân cấp  $R_t$ ,

$$R_t = (R_{t-1} - \{C_i, C_j\}) \cup \{C_q\}$$

B7. Cập nhật  $P(t)$  từ  $P(t-1)$  bằng 2 bước:

Xoá các dòng và cột tương ứng với 2 lớp vừa được kết nạp.

Thêm dòng mới và cột mới chứa khoảng cách giữa lớp mới tạo và các lớp cũ.

B8. Lặp lại B4 khi nào các  $K_i$  đều chưa thuộc cùng một lớp.

B9. Rút gọn cây dựa vào độ đo dị biệt giữa nút con và nút cha.

B10. Các nút lá cùng một cha sẽ tạo thành một vùng.

B11. Tiến hành trộn các vùng kích thước bé với các vùng lân cận.

Phân đoạn ảnh thực chất là phân lớp các điểm ảnh. Chúng tôi chọn phương pháp phân lớp phân cấp vì nó có khả năng tổ chức vùng theo cấu trúc phân cấp. Kết quả của giai đoạn này giúp rút trích các đặc trưng cấp cao chứa đựng ngữ nghĩa của ảnh.

### 3. Phân lớp vùng

Sau khi phân đoạn ảnh thành các vùng thì các vùng chính là các đặc trưng cấp cao đóng vai trò rất quan trọng giúp chúng ta lấp một phần hố ngăn cách giữa các đặc trưng cấp thấp và ngữ nghĩa của ảnh. Chúng tôi tiến hành phân lớp vùng nhằm rút trích được các khái niệm chung nhất giữa các vùng góp phần vào việc truy tìm được hiệu quả hơn. Chúng tôi phân lớp vùng dựa vào đặc trưng về màu, vân, hình dạng của vùng, độ đo dị biệt giữa các vùng và giải thuật phân lớp phân cấp.

#### Đặc trưng màu

Vì vùng  $C_i$  bây giờ gồm nhiều khối con  $K_i$  có đặc trưng màu trung bình tương tự nhau nên chúng tôi không chọn đặc trưng màu cho vùng là đặc trưng màu trung bình vì sẽ không thể hiện được chính xác phân bố màu trong vùng, vì vậy chúng tôi chọn đặc trưng lược đồ màu ứng với mỗi vùng, với hệ màu HSI(Hue,Saturation,Intensity) cùng việc định lượng không gian màu gồm 12 H, 3 S, 3 I, ta có đặc trưng lược đồ màu :

$\{h_{C_i}[j], j \in [0..107]$ , trong đó  $h_{C_i}[j]$  là phần tử thứ  $j$  của lược đồ màu của vùng  $C_i$ ,

#### Đặc trưng vân

Giả sử mỗi vùng có  $N$  khối con.

Đặc trưng vân của mỗi khối con  $K_i$  là  $\{V_{i1}, V_{i2}, V_{i3}\}$ ,

Đặc trưng vân của vùng  $C_i$  là :

$$V_{C_i,1} = \frac{1}{N} \left( \sum_{i=1}^N V_{i1} \right), V_{C_i,2} = \frac{1}{N} \left( \sum_{i=1}^N V_{i2} \right), V_{C_i,3} = \frac{1}{N} \left( \sum_{i=1}^N V_{i3} \right),$$

#### Đặc trưng hình dáng

Là đặc trưng cấp cao góp phần quan trọng trong việc truy tìm ảnh phù hợp hơn về mặt ngữ nghĩa.

Đặc trưng thứ nhất đặc tả diện tích của vùng :

$$DT_{C_i} = \frac{|C_i|}{|A|},$$

$|C_i| = \sum_{(x,y) \in C_i} 1$ , là số điểm ảnh trong vùng  $C_i$ ,

$|A| = \sum_{(x,y) \in A} 1$ , là số điểm ảnh trong ảnh  $A$

Đặc trưng thứ hai đặc tả vị trí của vùng :

Vị trí này chính là trọng tâm của vùng, điểm này có thể không nằm trong vùng, tuy nhiên điều này không quan trọng vì khoảng cách so sánh giữa các vùng không phụ thuộc điều này.

$$VT_{C_i} = \left( \sum_{(x,y) \in C_i} x / |C_i|, \sum_{(x,y) \in C_i} y / |C_i| \right),$$

$x, y$  là vị trí của điểm ảnh trong vùng  $C_i$ ,

$|C_i|$  là số điểm ảnh trong vùng  $C_i$ ,

Đặc trưng thứ ba đặc tả tính Compact của vùng :

$$Com_{C_i} = CV_o / CV_{C_i},$$

$CV_o$  là chu vi hình tròn có cùng diện tích với  $C_i$ ,

$CV_{C_i}$  là chu vi của  $C_i$

### Độ đo đối sánh vùng

Phân lớp vùng với đặc trưng diện tích, sử dụng độ đo tính dị biệt :

$$d_v(C_i, C_j) = |DT_{C_i} - DT_{C_j}|$$

Phân lớp vùng với đặc trưng tính compact, sử dụng độ đo tính dị biệt :

$$d_v(C_i, C_j) = |Com_{C_i} - Com_{C_j}|$$

Phân lớp vùng với đặc trưng màu, sử dụng độ đo tính dị biệt :

$$d_v(C_i, C_j) = d(h_{C_i}, h_{C_j})$$

trong đó :

$$d(h_{C_i}, h_{C_j}) = \sum_{k=1}^N \sum_{l=1}^N (h_{C_i}(k) - h_{C_j}(k))(h_{C_i}(l) - h_{C_j}(l))a_{kl},$$

$h_{C_i}(k)$  là phần tử thứ  $k$  của lược đồ màu của vùng  $C_i$ ,  $N$  là số ô màu của lược đồ màu,

$a_{kl} = \|m_k - m_l\|_{L^2}$ , thể hiện mức độ tương quan giữa các màu trong không gian màu HSI

$$a_{kl} = \frac{1}{\sqrt{5}} [(I_k - I_l)^2 + (S_k \cos H_k - S_l \cos H_l)^2 + (S_k \sin H_k - S_l \sin H_l)^2]^{1/2},$$

$H_k, S_k, I_k$  là các thành phần  $H, S, I$  của màu  $m_k$ .

Phân lớp vùng với đặc trưng vân, sử dụng độ đo tính dị biệt :

$$d_v(C_i, C_j) = |V_{C_i} - V_{C_j}|_{L_1}$$

Chúng tôi chọn phương pháp phân lớp phân cấp để phân lớp các vùng. Chúng tôi cải tiến việc phân lớp các vùng theo đặc trưng hình học ( $DT_{C_i}$ ) rồi đến đặc trưng compact ( $Com_{C_i}$ ), đặc trưng màu ( $h_{C_i}[j], j = 0, 2, \dots, 107$ ) và cuối cùng là đặc trưng vân ( $V_{C_i}$ ).

Mỗi lớp vùng sẽ được trích chọn một phần tử đại diện.

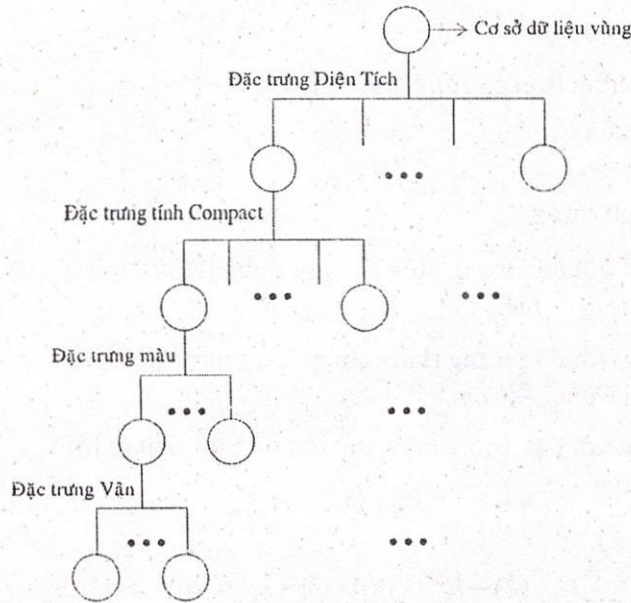
Việc phân lớp này nhằm chuẩn bị cho giai đoạn truy tìm. Nó giúp rút ngắn thời gian tìm kiếm nhờ vào việc tránh phải duyệt toàn bộ cơ sở dữ liệu vùng với độ đo đối sánh trên toàn bộ các đặc trưng của vùng.

### 4. Phân lớp ảnh

Chúng tôi thay các vùng trong ảnh bởi phần tử đại diện của lớp vùng, đặc trưng cho khái niệm về vùng, như vậy mỗi ảnh sẽ chứa các khái niệm về vùng. Chúng tôi tiến hành phân lớp ảnh dựa trên các đặc trưng của các khái niệm vùng, độ đo dị biệt giữa 2 ảnh và giải thuật phân lớp phân cấp.

Trong giai đoạn này, việc quan trọng là xác định độ đo dị biệt giữa 2 ảnh.

Giả sử ảnh A gồm các vùng  $\{C_i, i \in [1..Nc]\}$ , sau khi thay các vùng này bởi các phần tử đại diện của lớp vùng, ảnh A gồm các khái niệm của vùng  $\{Re p_j, j \in [1..Nrep]\}$ .



Hình 1. Sơ đồ phân lớp phân cấp các vùng

Giả sử xét 2 ảnh  $A_i, A_j$  :

$A_i = \{Re p_{i0}, Re p_{i1}, \dots, Re p_{iN-1}\}$ , với  $Re p_{ik}, k \in [0..N-1]$  là khái niệm vùng của ảnh  $A_i$

$A_j = \{Re p_{j0}, Re p_{j1}, \dots, Re p_{jM-1}\}$ , với  $Re p_{jl}, l \in [0..M-1]$  là khái niệm vùng của ảnh  $A_j$

Chúng tôi đề xuất một độ đo tính dị biệt giữa 2 ảnh, thể hiện được sự tham gia của các vùng trong cả 2 ảnh.

Sự tương tự giữa 2 ảnh càng cao khi có sự tương tự cao giữa các vùng trong 2 ảnh.

$$d_A(A_i, A_j) = \frac{\sum_{k=0}^{N-1} w_{ik} d(Re p_{ik}, A_j) + \sum_{l=0}^{M-1} w_{jl} d(Re p_{jl}, A_i)}{2}$$

$N$  là số vùng của ảnh  $A_i$ ,  $M$  là số vùng của ảnh  $A_j$ ,

$d(Re p_{ik}, A_j) = \min\{d_V(Re p_{ik}, Re p_{jl}), l = 0..M-1,$

là khoảng cách từ vùng  $Re p_{ik}$  của ảnh  $A_i$  đến ảnh  $A_j$ ,

$d(Re p_{jl}, A_i) = \min\{d_V(Re p_{jl}, Re p_{ik}), k = 0..N-1,$

là khoảng cách từ vùng  $Re p_{jl}$  của ảnh  $A_j$  đến ảnh  $A_i$ ,

$w_{ik} = N_{ik} / NK_i$ ,  $N_{ik}$  là số khối trong vùng  $Re p_{ik}$ ,  $NK_i$  là số khối con của ảnh  $A_i$

$w_{jl} = N_{jl} / NK_j$ ,  $N_{jl}$  là số khối trong vùng  $Re p_{jl}$ ,  $NK_j$  là số khối con của ảnh  $A_j$

$w_{ik}$  và  $w_{jl}$  thể hiện vai trò quan trọng của vùng có diện tích lớn trong độ đo đối sánh.

$$d_V(C_i, C_j) = w_m d(h_{C_i}, h_{C_j}) + w_v |V_{C_i} - V_{C_j}|_{L_1} +$$

$$w_{vt} |VT_{C_i} - VT_{C_j}|_{L_2} + w_{dt} |DT_{C_i} - DT_{C_j}| + w_{com} |Com_{C_i} - Com_{C_j}|$$

$w_m, w_v, w_{vt}, w_{dt}, w_{com}$  là các trọng số xác định mức độ quan trọng của các đặc trưng màu, vân, vùng

### 5. Truy tìm ảnh dựa trên cấu trúc cây phân cấp vùng

Trong quá trình tìm ảnh dựa trên cấu trúc cây phân cấp vùng, hệ thống của chúng tôi được tổ chức thành 3 giai đoạn chính là giai đoạn nhập dữ liệu, truy tìm và kết xuất.

### 5.1. Giai đoạn nhập dữ liệu

Trong giai đoạn này người sử dụng có thể nhập liệu bằng 3 cách :

#### Cách thứ nhất

Người dùng chọn ảnh truy vấn tùy ý, sau đó hệ thống sẽ tự động phân đoạn ảnh này thành các vùng.

#### Cách thứ hai

Người dùng chọn một số ảnh truy vấn tùy ý, sau đó hệ thống sẽ tự động phân đoạn các ảnh này, cuối cùng người dùng chọn các vùng cần truy tìm và các vùng không cần tìm trong các ảnh này theo luận lý :

$$(Sel(C_i) OR Sel(C_j)) AND (Sel(C_k) OR Sel(C_l)) AND NOT (Sel(C_m) OR Sel(C_n)),$$

$$Sel(C) = \begin{cases} \text{True nếu } C \text{ được chọn} \\ \text{False nếu } C \text{ không được chọn} \end{cases}$$

#### Cách thứ ba

Người dùng chọn các mẫu truy vấn đã có sẵn trong cơ sở dữ liệu vùng theo luận lý :

$$(Sel(C_i) OR Sel(C_j)) AND (Sel(C_k) OR Sel(C_l)) AND NOT (Sel(C_m) OR Sel(C_n)),$$

### 5.2. Giai đoạn truy tìm

Với vùng truy vấn  $C_q$ , so sánh với các phần tử đại diện trong cây phân lớp vùng để chọn được mẫu đại diện gần nhất với mẫu truy vấn.

Gọi mẫu đại diện gần nhất với mẫu truy vấn là  $C_r$ , lớp có mẫu đại diện  $C_r$  là  $B_r$ .

Gọi tập các ảnh có ít nhất một vùng thuộc về lớp  $B_r$  là  $TA(B_r)$ .

Tập ảnh ứng với câu truy vấn :

$$(Sel(C_i) OR Sel(C_j)) AND (Sel(C_k) OR Sel(C_l)) AND NOT (Sel(C_m) OR Sel(C_n))$$

sẽ là tập :

$$((TA(B_i) \cup TA(B_j)) \cap ((TA(B_k) \cup TA(B_l)) \setminus ((TA(B_m) \cup TA(B_n))))$$

Với cách làm trên, hệ thống sẽ đạt độ chính xác khá cao, mức độ tìm sót thấp, đồng thời tốc độ truy tìm khá nhanh vì không phải duyệt qua toàn bộ cơ sở dữ liệu ảnh và sử dụng toàn bộ các đặc trưng trong đối sánh ảnh. Tuy nhiên tập kết quả có thể là tập rỗng.

### 5.3. Giai đoạn kết xuất

Trong giai đoạn này, chúng tôi tiến hành xếp hạng kết quả có được từ giai đoạn truy tìm dựa trên độ đo đối sánh giữa 2 ảnh.

## 6. Truy tìm ảnh dựa trên cấu trúc cây phân cấp vùng và ảnh

Trong quá trình tìm ảnh dựa trên cấu trúc cây phân cấp vùng và ảnh, hệ thống của chúng tôi được tổ chức thành 3 giai đoạn chính là giai đoạn nhập dữ liệu, truy tìm và kết xuất.

### 6.1. Giai đoạn nhập liệu

Trong giai đoạn này, người sử dụng có thể nhập liệu bằng 3 cách như đã trình bày ở mục 5.1. Sau đó các vùng sẽ được thay thế bởi các phần tử đại diện gần nhất trong cây phân cấp vùng.

**6.2. Giai đoạn truy tìm**

Kết quả có được từ bước nhập liệu là ảnh truy vấn chứa một số khái niệm là các phần tử đại diện vùng. So sánh ảnh truy vấn với các phần tử đại diện trên cây phân lớp ảnh để chọn phần tử đại diện gần nhất, tập các ảnh trong lớp có phần tử đại diện này là kết quả cần tìm. Dùng phương pháp này, tập ảnh kết quả sẽ khác tập rỗng.

**6.3. Giai đoạn kết xuất**

Trong giai đoạn này, chúng tôi tiến hành xếp hạng kết quả có được từ giai đoạn truy tìm dựa trên độ đo đối sánh giữa 2 ảnh .

**7. Thực nghiệm**

Chúng tôi đã thử nghiệm trên cơ sở dữ liệu gồm 9000 ảnh bao gồm đủ loại :

Thiên nhiên : Phong cảnh, hoa, quả, động vật

Nhân tạo : Thành phố, xe hơi, xe lửa, tàu thuyền, máy bay, áo quần, logo

Hoạt động có người : Thể thao

Từ 9000 ảnh chúng tôi rút trích được 54624 vùng, và phân thành 325 lớp.

Chúng tôi sử dụng 2 đại lượng :

Độ chính xác và độ trung thực để đánh giá tính hiệu quả của hệ thống.

Độ chính xác = Số ảnh tìm được đúng / Số ảnh tìm được.

Độ trung thực = Số ảnh tìm được đúng / Tổng số ảnh đúng thực có

Phương pháp nhập liệu Phương pháp truy tìm	Số mẫu truy vấn	Số mẫu kết quả tốt nhất được chọn	Độ chính xác trung bình	Độ trung thực trung bình
Chọn ảnh truy vấn Truy tìm dựa vào lược đồ màu toàn cục	50	10	72	74
		20	70	78
		30	67	81
Chọn ảnh truy vấn Truy tìm dựa vào độ đo đối sánh ảnh	50	10	88	85
		20	86	86
		30	82	88
Chọn vùng của ảnh truy vấn Truy tìm dựa vào độ đo đối sánh vùng và ảnh	50	10	90	86
		20	87	87
		30	83	89
Chọn vùng truy vấn trong cơ sở dữ liệu vùng Truy tìm dựa vào độ đo đối sánh vùng và ảnh	50	10	92	89
		20	90	90
		30	89	91

**Bảng 1. Bảng thống kê kết quả truy tìm theo phương pháp đánh vân ảnh**

Kết quả thực nghiệm cho thấy phương pháp truy tìm ảnh dựa trên phương pháp đánh vân ảnh khá hiệu quả, giúp lấp bớt được hố ngăn cách giữa đặc trưng cấp thấp và ngữ nghĩa ảnh.



**8. Kết luận và hướng phát triển**

Chúng tôi đã tiếp cận bài toán truy tìm ảnh dựa vào nội dung với các phương pháp nhằm giúp giảm bớt hồ ngấn cách giữa đặc trưng cấp thấp và ngữ nghĩa của ảnh.

Cách tiếp cận của chúng tôi là dùng đặc trưng cấp cao là vùng mà chúng tôi còn gọi là vắn ảnh để bắc cầu nối giữa đặc trưng cấp thấp và ngữ nghĩa của ảnh. Chúng tôi đã dùng giải thuật phân lớp phân cấp để phân đoạn ảnh, phân lớp vùng và phân lớp ảnh. Việc truy tìm được tiến hành trên cây phân lớp vùng hoặc trên cây phân lớp vùng và ảnh. Hệ thống truy tìm của chúng tôi đạt được độ chính xác và trung thực khá cao, và quan trọng hơn là người sử dụng đã có thể diễn đạt ảnh truy vấn gần với suy nghĩ của họ hơn.

Với hệ thống truy tìm này, chúng tôi đã xây dựng được một khung nền để có thể tiến xa hơn trong việc xây dựng mô hình truy tìm ảnh dựa vào khái niệm. Công việc chúng tôi sẽ tiến hành trong tương lai là dịch tự động từ ảnh sang từ nhằm giúp người sử dụng có thể diễn đạt bằng ngôn từ trong truy tìm ảnh.

## CONTENT-BASED IMAGE RETRIEVAL BASED ON SPELLING IMAGES

Nguyen Lam<sup>(1)</sup>, Ly Quoc Ngoc<sup>(2)</sup>, Duong Anh Đức<sup>(2)</sup>

Nguyen Ba Cong<sup>(2)</sup>, Nguyen Huu Đức<sup>(2)</sup>

(1) Luong The Vinh University, Nam Dinh City

(2) University of Natural Sciences, HCM city

**ABSTRACT:** *Nowadays, image retrieval have a new age, content-based image retrieval age. Resolving the problem of content-based image retrieval is the challenge must been overcome in our life. The key problem in achieving an efficient and user-friendly system is the development the description of query data close to human beings and athering the minimal irrelevant information together the relevant information is not overlooked. We have solved this problem by discovering the implicit high-level features in image database. These features are extracted by segmenting the image into the regions and clustering them into the clusters with their representatives that we called the words of image. We have used the Hierarchical Agglomerative Clustering algorithm to segment the images, cluster the regions and cluster the images. The query data can be the whole image, or some regions of query images or some region cluster's representatives. Retrieving is performed on the region clusters's hierarchical tree structure or the image cluster's hierarchical tree structure. Our experiment results have shown that our retrieval system is efficient and user-friendly.*

## TÀI LIỆU THAM KHẢO

[1] Al Bovik. *Handbook of Image and Video Processing*. Academic Press, 2000

[2] D. Bimbo. *Visual Information Retrieval*. Morgan Kaufmann, 1999

- [3] G. Sheikholeslami, W. Chang, A. Zhang. *SemQuery: Semantic Clustering and Querying on Heterogeneous Features for Visual Data*. IEEE Trans. Knowledge and Data Engineering, vol. 14, no. 5, pp. 988-1002, 2002
- [4] I. Kompatsiaris, E. Triantafillou, M. G. Strintzis. *Region-based color image indexing and retrieval*. IEEE International Conference on Image Processing (ICIP'01), 2001
- [5] Lý Quốc Ngọc. *Truy tìm thông tin thị giác dựa vào nội dung bằng phương pháp gán nhãn ngữ nghĩa cho ảnh*. Tạp chí phát triển khoa học công nghệ , tập 7, tháng 4&5, năm 2004.
- [6] Quoc Ngoc Ly, Anh Duc Duong. *Hierarchical Agglomerative Clustering in Content-based Image Retrieval*. Proceedings of The International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC 2005), July 4-7, 2005, Korea.
- [7] Quoc Ngoc Ly, Anh Duc Duong. *Hierarchical Data Model in Content-based Image Retrieval*. International Journal of Information technology, International Conference on Intelligent Computing (ICIC 2005), August 23-28, 2005, China
- [8] Rafael C. Gonzalez, Richard E. Woods. *Digital Image Processing. Second Edition*, Prentice-Hall, Inc. 2002.
- [9] Sergios Theodoridis, Konstantinos Koutroumbas. *Pattern Recognition*, Academic Press, 1999