

KHOA CÔNG NGHỆ THÔNG TIN-ĐHKHTN

TRUY VẤN THÔNG TIN THỊ GIÁC

Giảng viên: PGS.TS. Lý Quốc Ngọc
TPHCM, 6-2022



KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

TRUY VẤN THÔNG TIN THỊ GIÁC

Bài giảng 6 Tổ chức và Truy vấn ảnh dựa vào vùng ảnh

Giảng viên: PGS.TS. Lý Quốc Ngọc



KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

Nội dung

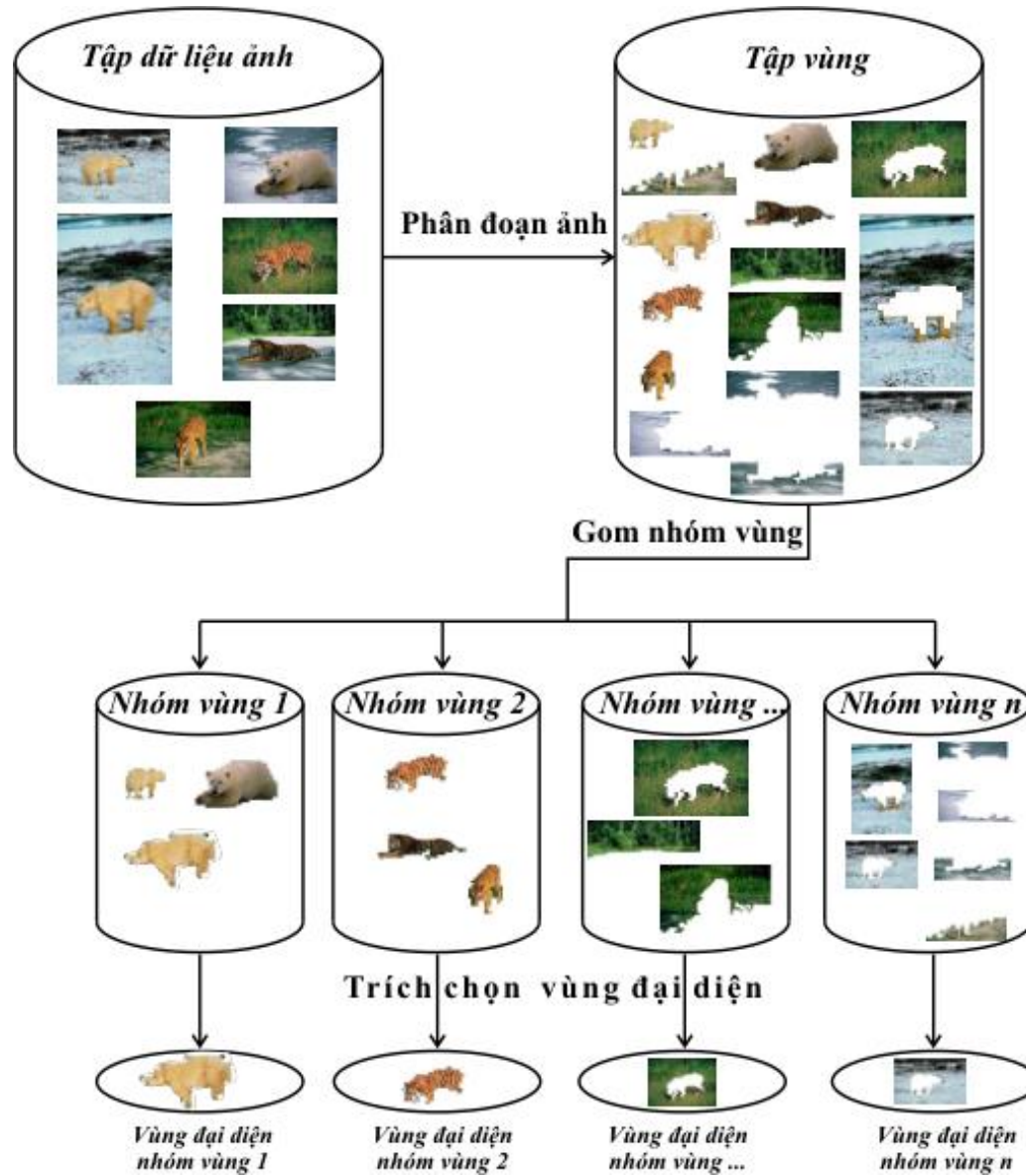
1. Mô hình tổ chức dữ liệu ảnh dưới dạng vùng.
2. Mô hình truy vấn dữ liệu ảnh dưới dạng vùng.

1. Mô hình tổ chức dữ liệu ảnh dưới dạng vùng

Gồm ba công đoạn chính sau :

- Phân đoạn ảnh
- Gom nhóm vùng
- Trích chọn vùng đại diện.

1. Mô hình tổ chức dữ liệu ảnh dưới dạng vùng



1.1. Phân đoạn ảnh

Mỗi ảnh $I_i \in SI$ ($i = 1..n_{SI}$, n_{SI} là số ảnh trong tập SI) được chia nhỏ thành các khối con K_j kích thước 4×4 (điểm ảnh), các đặc trưng của mỗi khối con gồm 3 đặc trưng màu, 3 đặc trưng vân và vị trí của trọng tâm khối.

Vector đặc trưng thị giác của các khối điểm ảnh

Định lượng màu

[] sử dụng không gian màu HSV và định lượng không gian màu này thành 12 thành phần H, 4 thành phần S và 4 thành phần V (192 màu).

Đặc trưng màu

Mỗi khối con K_j có 3 đặc trưng màu được biểu thị bởi $\{C_{j1}, C_{j2}, C_{j3}\}$,

$$C_{j1} = \left(\sum_{i=0}^{15} H[i] \right) / 16, C_{j2} = \left(\sum_{i=0}^{15} S[i] \right) / 16, C_{j3} = \left(\sum_{i=0}^{15} V[i] \right) / 16, \quad (3.4)$$

trong đó $H[i], S[i], V[i]$ là các thành phần H, S, V của điểm ảnh thứ i của khối con K_j .

1.1. Phân đoạn ảnh

Gọi $X = \{x_i, i = 1..N\}$ là tập các vector đặc trưng l chiều cần được gom nhóm,

N là số vector đặc trưng.

Gọi R là phép gom nhóm, $R = \{C_j, j = 1..M\}$, C_j là nhóm của các vector đặc trưng, $C_j \subset X$, M là số nhóm.

1.1. Phân đoạn ảnh

Gồm lược đồ tích lũy tổng quát và lược đồ tạo sinh nhóm.

Lược đồ tích lũy tổng quát (Generalized Agglomerative Scheme (GAS))

Bước 1. Giả sử $R_0 = \{C_i = \{x_i\}, i=1..N\}$ là phép gom nhóm ban đầu (khởi đầu, mỗi nhóm C_i gồm chỉ một vector đặc trưng x_i , N là số vector đặc trưng)

Bước 2. $t = 0$,

Bước 3. Lặp

Bước 4. $t = t + 1$

Bước 5. Chọn cặp nhóm (C_i, C_j) trong số các cặp nhóm (C_r, C_s) trong R_{t-1} sao cho

$d_C(C_i, C_j) = \min_{r,s=1,\dots,N, r \neq s} d_C(C_r, C_s)$, d_C là khoảng cách dị biệt giữa các nhóm.

Bước 5. Gọi C_q là nhóm được tạo do kết hợp cặp nhóm $\{C_i, C_j\}$, $C_q = C_i \cup C_j$,

Tạo phép gom nhóm mới $R_t = (R_{t-1} - \{C_i, C_j\}) \cup \{C_q\}$.

Bước 6. Đến khi $t = N - 1$ (tức là phép gom nhóm R_{N-1} được tạo và tất cả các vector đặc trưng thuộc cùng một nhóm).

1.1. Phân đoạn ảnh

Gồm lược đồ xây dựng cây phân hoạch và lược đồ tạo sinh nhóm.

Lược đồ xây dựng cây (được cảm sinh từ lược đồ tích lũy tổng quát)

Bước 1. Giả sử $T_0 = \{n_i = \{x_i\}, i = 1..N\}$ là cây phân hoạch ban đầu được tạo lập gồm N nút n_i (nút n_i ban đầu gồm vectơ đặc trưng x_i).

Bước 2. $t = 0$,

Bước 3. Lặp

Bước 4. $t = t + 1$

Bước 5. Chọn cặp nút (n_i, n_j) trong số các cặp nút (n_r, n_s) trong T_{t-1} sao cho

$d_N(n_i, n_j) = \min_{r, s=1, \dots, N, r \neq s} d_N(n_r, n_s)$, d_N là khoảng cách dị biệt giữa các nút.

Bước 5. Gọi n_q là nút được tạo do kết hợp cặp nút (n_i, n_j) , $n_q = n_i \cup n_j$,

Tạo cây mới $T_t = (T_{t-1} - \{n_i, n_j\}) \cup \{n_q\}$.

Bước 6. Đến khi $t = N - 1$ (tức là cây T_{N-1} được tạo).

1.1. Phân đoạn ảnh

Lược đồ tạo sinh nhóm

Bước 1. Rút gọn cây phân hoạch dựa vào độ đo dị biệt giữa nút cha và nút con.

Bước 2. Các nút lá cùng một cha sẽ được tạo thành một nhóm.

Trong giai đoạn này, luận án dùng phương pháp rút gọn cây phân hoạch [19] để tạo sinh nhóm. Nút cần loại bỏ dựa trên tiêu chuẩn xét độ dị biệt giữa nút con với nút cha.

Gọi a, b là 2 nút khác nút lá,

$d_{\min}(a, b) = \min\{d(l_a, l_b), \forall l_a \in a, \forall l_b \in b\}$, là độ dị biệt bé nhất giữa a và b ,

$d_{\max}(a, b) = \max\{d(l_a, l_b), \forall l_a \in a, \forall l_b \in b\}$, là độ dị biệt lớn nhất giữa a và b ,

1.1. Phân đoạn ảnh

Trong đó l_a, l_b là các nút lá.

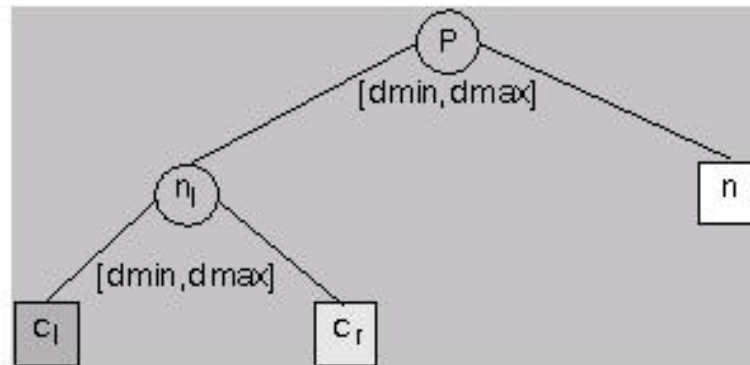
Trong quá trình tạo cấu trúc cây phân cấp, ta có :

Nút p_l được tạo lập và được liên kết 2 giá trị :

$d_{\min}(n_l, n_r), d_{\max}(n_l, n_r), n_l, n_r$ là nút con trái và phải của nút p_l

Nút n_l được tạo lập và được liên kết 2 giá trị :

$d_{\min}(c_l, c_r), d_{\max}(c_l, c_r), c_l, c_r$ là nút con trái và phải của nút n_l



1.1. Phân đoạn ảnh

Giả sử nút n là con của nút p , xét 2 đại lượng :

$$\Delta_{\min}(p, n) = \frac{d_{\min}^p - d_{\min}^n}{d_{\min}^n}, \text{ thể hiện độ dị biệt giữa độ dị biệt bé nhất của nút } n$$

và nút p .

$$\Delta_{\max}(p, n) = \frac{d_{\max}^p - d_{\max}^n}{d_{\max}^n}, \text{ thể hiện độ dị biệt giữa độ dị biệt lớn nhất của nút}$$

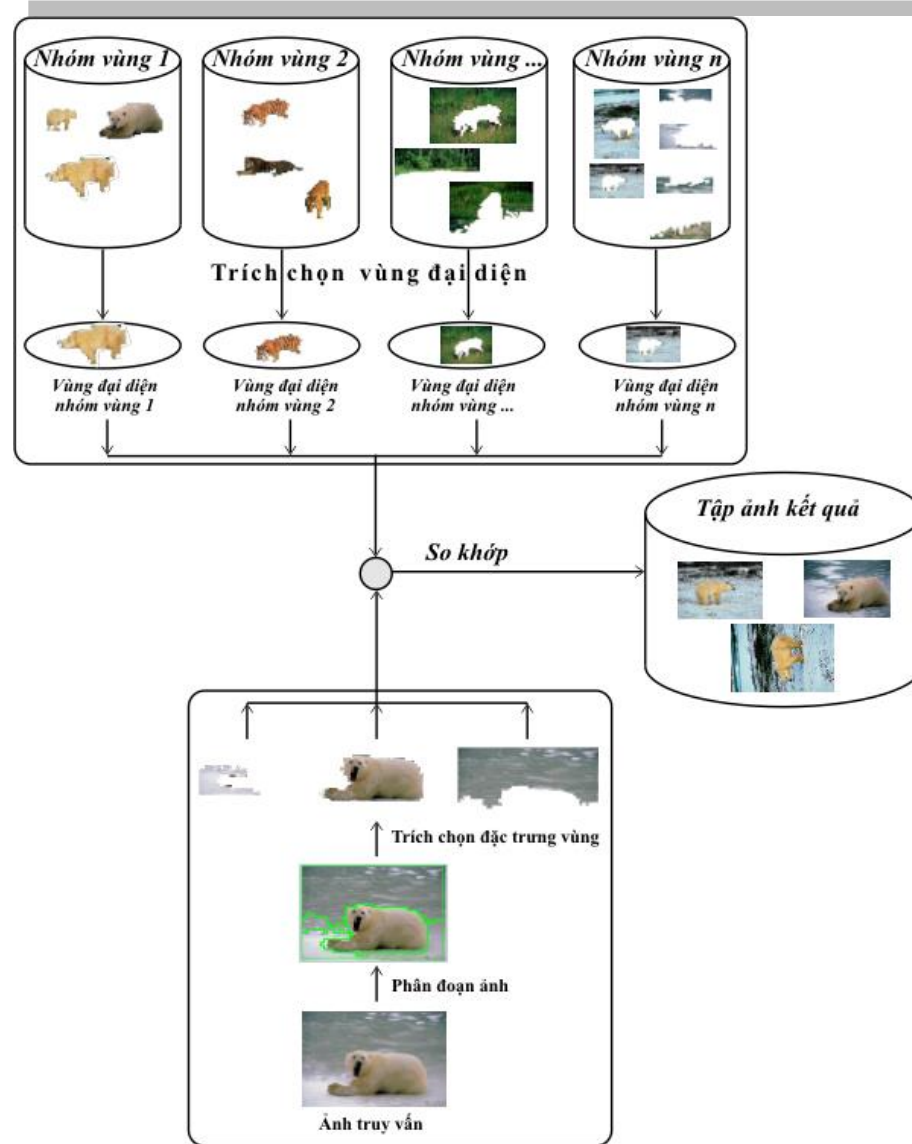
n và nút p .

Nút n được loại bỏ nếu :

$$\Delta_{\min} < \varepsilon \quad \text{và} \quad |\Delta_{\max}| < \varepsilon$$

Khi nút n bị loại bỏ, nút con của nút n là nút c không bị mất đi mà được liên kết với trở lại với nút p .

2. Mô hình truy vấn ảnh dưới dạng vùng



Để sắp hạng kết quả truy vấn, cần xác định độ đo đối sánh giữa ảnh truy vấn và ảnh kết quả, độ đo này được xác định như sau.

Luận án thay các vùng ảnh bởi vùng đại diện, đặc trưng cho nhóm các vùng, như vậy mỗi ảnh sẽ chứa các vùng đại diện.

Giả sử $RI \in SRI^R$, ảnh RI gồm các vùng $\{r_j, j = 1..n_{RI}\}$,

Xét hai ảnh QI và RI :

Luận án chọn khoảng cách dị biệt giữa hai ảnh chịu ảnh hưởng của các vùng trong cả hai ảnh. Sự dị biệt giữa hai ảnh càng lớn khi có sự dị biệt lớn giữa các vùng trong hai ảnh.

2. Mô hình truy vấn ảnh dưới dạng vùng

$$d(RI, QI) = \frac{\sum_{l=1}^{n_{RI}} w_l d(r_l, QI) + \sum_{k=1}^{n_{QI}} w_k d(rq_k, RI)}{2},$$

n_{RI} là số vùng của ảnh RI , n_{QI} là số vùng của ảnh QI ,

$d(r_l, QI) = \min_{k=1..n_{QI}} \{d_{Reg}(r_l, rq_k)\}$ là độ dị biệt của vùng r_l (ảnh RI) đối với ảnh QI ,

$d(rq_k, RI) = \min_{l=1..n_{RI}} \{d_{Reg}(rq_k, r_l)\}$ là độ dị biệt của vùng rq_k (ảnh QI) đối với ảnh RI ,

$w_l = N_l / NK_{RI}$, N_l là số khối trong vùng r_l , NK_{RI} là số khối con của ảnh RI ,

$w_k = N_k / NK_{QI}$, N_k là số khối trong vùng rq_k , NK_{QI} là số khối con của ảnh QI ,

w_k và w_l thể hiện vai trò quan trọng của vùng có diện tích lớn trong độ đo đối sánh.

Gọi SRI_ε là tập ảnh kết quả ứng với sai số ε , có thể biểu diễn SRI_ε như sau :

$$SRI_\varepsilon = \{RI_i | d(RI_i, QI) \leq d(RI_{i+1}, QI) < \varepsilon, RI_i \in SI, i = 1..n(\varepsilon)-1\},$$

trong đó :

RI_i là ảnh kết quả thứ i (sắp theo thứ tự tăng dần dựa vào khoảng cách dị biệt d),

$d(RI_i, QI)$ là khoảng cách dị biệt giữa ảnh RI_i và ảnh truy vấn QI ,