Một phương pháp mới chuẩn hoá dữ liệu và hiệu chỉnh trọng số cho tổ hợp đặc trưng trong tra cứu ảnh theo nội dung

A Novel Method Normalized Data and Refine Weights for Combination Features in Content Based Image Retrieval

Vũ Văn Hiệu, Ngô Hoàng Huy, Ngô Quốc Tạo, Nguyễn Hữu Quỳnh

Abstract*: Relevance feedback as a bridge between high level semantic concepts and low features. It is important to improve the performance of content based image retrieval (CBIR) is preprocessing image features and refining distance measures for query based on user information needs. We propose a novel method* $3\sigma - FCM$ *to normilize features and distance for CBIR using combination features. In addition, we also use relevant feedback from users and learning from low features to update weights distance measures and refine query. Experimental results over the benchmark Corel dataset demonstrate the effectiveness of this propose method.*

Keywords*: Content Based Image Retrieval, Relevant Feedback, Normalized feature, Normalized distance, Fuzzy clustering c-means.*

I. GIỚI THIỆU

Với sự gia tăng nhanh chóng số lượng dữ liệu ảnh số, tra cứu ảnh dựa vào nội dung (Content based image retrieval - CBIR) trở thành lĩnh vực nghiên cứu tích cực trong những năm qua [6, 11, 17, 22, 24, 27, 37, 42 - 44, 53]. Các hệ thống này thường trích rút các biểu diễn trực quan của ảnh và định nghĩa các hàm tìm kiếm, so khớp mối liên quan để tra cứu theo sự quan tâm.

Tuy nhiên việc kết hợp các đặc trưng khác nhau rất phức tạp và phụ thuộc ứng dụng tra cứu [29, 31]. Bên cạnh việc sử dụng đơn đặc trưng không hiệu quả [5, 25, 27, 32, 47, 55], kết hợp nhiều đặc trưng nhưng sử dụng chung một độ đo khoảng cách cũng có một số hạn chế đáng kể [2, 12, 42]. Sự kết hợp các đặc trưng và độ đo khoảng cách phù hợp trong các ứng dụng tra cứu ảnh theo nội dung cụ thể là rất quan trọng.

Chúng tôi đã đề xuất một phương pháp mới chuẩn hóa dữ liệu, chuẩn hoá khoảng cách và cập nhật tự động trọng số của độ đo khoảng cách cho mỗi đặc trưng. Phương pháp này áp dụng cho hệ thống tra cứu ảnh theo nội dung sử dụng phản hồi liên quan và kết hợp nhiều đặc trưng trực quan mức thấp. Phương pháp này linh hoạt trong việc đánh chỉ số đặc trưng hoặc mở rộng thêm các đặc trưng khác mà không cần bất kì thay đổi thuật toán nào.

Cách tiếp cận của chúng tôi đề cập tới các nghiên cứu trước đó [1, 19, 35, 43, 46] trên chiến lược: chuẩn hoá đặc trưng kết hợp, chuẩn hoá khoảng cách, điều chỉnh trọng số dựa vào kiến thức người dùng và học từ dữ liệu. Những khác biệt này được trình bày trong phần ba và phần bốn.

Cấu trúc của bài báo được tổ chức như sau. Phần hai, một số nghiên cứu liên quan sử dụng kết hợp đặc trưng, chuẩn hoá đặc trưng, chuẩn hoá khoảng cách và phản hồi liên quan. Phần ba là đề xuất chuẩn hoá đặc trưng, chuẩn hoá khoảng cách và hiệu chỉnh trọng số dựa vào thông tin phản hồi từ người dùng và học từ dữ liệu. Các kết quả thực nghiệm đưa ra trong phần bốn. Kết luận và hướng nghiên cứu tương lai trong phần năm.

II. NGHIÊN CỨU LIÊN QUAN

II.1. Kết hợp đặc trưng trong CBIR

Trong tra cứu ảnh theo nội dung các đặc trưng trực quan thường được sử dụng kết hợp như là màu, kết cấu và hình dạng. Trong [14, 47] sử dụng kết hợp

màu, kết cấu, hình dạng. Các thành phần cho biểu diễn này gồm đặc trưng lược đồ màu, bất biến màu [51], kết cấu Tamura [52] và ma trận đồng mức [15, 33]. Trong [41] sử dụng kết hợp lược đồ màu, bất biến màu, Tamura, ma trận đồng mức, miêu tả Fourier, miêu tả hình dạng khối, và trong [33] sử dụng kết hợp lược đồ màu, bất biến màu, và ma trận đồng mức. Trong [36, 39] sử dụng véc tơ liên kết màu (lược đồ liên kết và lược đồ không liên kết). Trong [23] sử dụng kết hợp lược đồ màu, tương quan màu, bất biến màu, biến đổi Gabor và biến đổi wavelet.

II.2. Độ đo khoảng cách theo bộ đặc trưng

Để tính độ đo khoảng cách giữa truy vấn với mỗi ảnh trong cơ sở dữ liệu, Rahman và cộng sự [38] trích rút đặc trưng dựa trên khái niệm trực quan ở nhiều mức khác nhau, lược đồ biên (CLD) biểu diễn mức bán toàn cục và màu bề mặt (EHD) biểu diễn mức toàn cục. Độ đo khoảng cách có trọng số giữa hai véc tơ đặc trưng I_q và I_j được định nghĩa như sau:

$$Dis_{global}\ I_q, I_j = w_{cld} Dis_{cld}\ f^{cld}_{I_q}, f^{cld}_{I_j} + w_{ehd} Dis_{ehd}\ f^{ehd}_{I_q}, f^{ehd}_{I_j}, \quad (1)$$

với: $Dis_{cld}\ f^{cld}_{Iq}, f^{cld}_{Ij}$, $Dis_{ehd}\ f^{ehd}_{Iq}, f^{ehd}_{Ij}$ trên CLD và EHD tương ứng là L_2, $0 \le w_{cld}, w_{ehd}, w_{cld} + w_{ehd} = 1$.

Trong [9] đã cho một định nghĩa cụ thể độ đo khoảng cách có trọng số giữa hai ảnh (theo kiểu định nghĩa 2) :

$$S = \frac{w_C S_C + w_D S_D + w_A S_A}{w_C + w_D + w_A}, \quad (2)$$

trong đó S_c, S_D, và S_A là các độ đo khoảng cách giao cắt lược đồ theo màu, khoảng cách và hình dạng tương ứng.

Trong [10] đề xuất đánh giá độ đo khoảng cách giữa hai ảnh dựa trên đánh giá độ đo khoảng cách giữa các vùng ảnh. Trong đó miêu tả đặc trưng bao gồm biểu diễn các màu và phần trăm trong vùng.

II.3. Biểu diễn tổng quát đối tượng ảnh và độ đo khoảng cách

Hệ thống CBIR sử dụng nhiều bộ đặc trưng, trong mỗi bộ đặc trưng cũng có nhiều thành phần, các định nghĩa 2.1, 2.2 cho đối tượng ảnh và độ đo khoảng cách tổng quát.

Định nghĩa 2.1 ([43] Đối tượng ảnh):

$O \equiv O\ \mathrm{E}, F, \mathrm{T}$, trong đó:

E: dữ liệu thô của ảnh

$F = \{ f_i \}$: tập đặc trưng trong (màu, kết cấu, hình dạng).

$\mathrm{T} = t_{ij}$: tập biểu diễn cho đặc trưng f_i , $t_{ij} \equiv \left[t_{ij1}, ..., t_{ijk}, ..., t_{ijK} \right]$ véc tơ gồm nhiều thành phần.

Định nghĩa 2.2: Độ đo khoảng cách *D* giữa hai đối tượng $O^1(E^1, F^1, T^1)$ và $O^2(E^2, F^2, T^2)$ dạng đầy đủ xác định:

- Độ đo khoảng cách của một bộ đặc trưng:

$$D_{ij}\left(t^1_{ij}, t^2_{ij}\right) \overset{def}{=} D_{ij, w_{ijk}}\left(t^1_{ij}, t^2_{ij}\right) \quad (3)$$

- Độ đo khoảng cách của một kiểu đặc trưng:

$$D_i\left(f^1_i, f^2_i\right) \overset{def}{=} \sum_j W_{ij} D_{ij}\left(t^1_{ij}, t^2_{ij}\right) \quad (4)$$

- Độ đo khoảng cách toàn bộ:

$$D\left(O^1\left(E^1, F^1, T^1\right), O^2\left(E^2, F^2, T^2\right)\right) \overset{def}{=} \sum_i w_i D_i\left(f^1_i, f^2_i\right) \quad (5)$$

II.4. Một số phép chuẩn hoá hay được sử dụng

Mô hình đối tượng (định nghĩa 1) yêu cầu chuẩn hóa đặc trưng là cần thiết để bù đắp cho phạm vi khác biệt nhau giữa các thành phần đặc trưng được định nghĩa trong các miền khác nhau. Sau khi chuẩn hoá đặc trưng, chuẩn hoá hàm đo khoảng cách rất quan trọng, đảm bảo tính cân bằng giữa các đặc trưng khác nhau trên các hàm đo khoảng cách khác nhau.

- Chuẩn hóa min-max:

$$\{ f_i[j] \} \text{ a } \{ f_i'[j] \}, \forall j,\ f_i'[j] = \frac{f_i[j] - \min_j f_i[j]}{\max_j f_i[j] - \min_j f_i[j]}, \quad (6)$$

- Chuẩn hóa 3σ:

$$\{ f_i[j] \} \text{ a } \{ f_i'[j] \},\ f_i'[j] = \frac{f_i[j] - m_j}{3\sigma_j}, \quad (7)$$

trong đó $m_j \overset{\text{def}}{=} mean\left(\{f_i[j]\}\right), \sigma_j \overset{\text{def}}{=} \text{var}\left(\{f_i[j]\}\right)$

Tính chất : Chuẩn hóa theo min-max, 3σ bảo toàn thứ tự, $f_{1,i}[j] \geq f_{2,i}[j] \Rightarrow f'_{1,i}[j] \geq f'_{2,i}[j]$

Luật 3σ ([49]): x là $N(\mu,\sigma)$, thì xác suất

$$\text{P}\left(\left|\frac{x-\mu}{3\sigma}\right| \leq 1\right) \geq 0.99$$

Hạn chế : Chuẩn hóa theo min – max làm cho hầu hết thông tin hữu ích bị chuyển vào một phạm vi rất hẹp trong [0,1] nếu giá trị max lớn, 3σ rải đều trong [-1,1] nhưng yêu cầu dữ liệu là một chuỗi Gauss.

Phép chuẩn hóa 3σ đã được sử dụng trong [7, 28, 35, 42, 43] cho các đặc trưng dữ liệu (màu, kết cấu, hình dạng) và chuẩn hóa tập giá trị khoảng cách giữa hai mẫu dữ liệu. Trong [2] sử dụng phép chuẩn hoá min-max.

Chuẩn hoá giá trị khoảng cách sử dụng để ánh xạ giá trị khoảng cách của ảnh từ truy vấn dựa vào một véc tơ đặc trưng trong khoảng $[0,1]$. Trong [7] đã áp dụng phép chuẩn hóa min-max cho các giá trị khoảng cách, và xem xét phép chuẩn hoá sau:

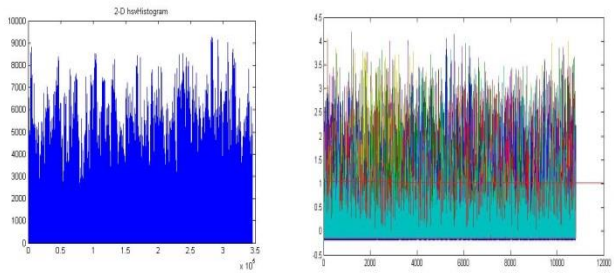
$X^k = \left\{X^k_{hs}, h = 1,...,p, s = 1,...,q\right\}$ là tập dữ liệu đặc trưng cơ sở dữ liệu ảnh, $k = \overline{1,n}$ và *h* là chỉ số của đặc trưng (chẳng hạn histogram), *s* là chỉ số của ảnh con mà đặc trưng đề cập đến.

Véc tơ khoảng cách chuẩn hoá giữa hai ảnh có chỉ số i và j tương ứng là:

$$D(X^i, X^j) = \left[\frac{D_{11}(X^i_{11}, X^j_{11})}{\mu_{11}}, \frac{D_{hs}(X^i_{hs}, X^j_{hs})}{\mu_{hs}}, ..., \frac{D_{pq}(X^i_{pq}, X^j_{pq})}{\mu_{pq}}\right]^T, \quad (8)$$

Trong đó các μ_{hs} là trung bình cộng các khoảng cách giữa các cặp ảnh trong cơ sở dữ liệu.

Trong lược đồ dữ liệu đặc trưng ở Hình 1, các thành phần của chuỗi đặc trưng thường có không ít hơn một đỉnh, tức là giả định phân bố chuẩn áp đặt là không hợp lý. Do đó khi chuẩn hóa theo 3σ, dữ liệu sau khi chuẩn hóa có khá nhiều thành phần rơi ra ngoài đoạn [-1,1]. Vì vậy sử dụng chuẩn hoá 3σ cho dữ liệu đặc trưng kết hợp là chưa đạt được mục tiêu của bước chuẩn hoá.



(a) Đặc trưng thô (b) Đặc trưng chuẩn hóa theo luật 3σ

Hình 1. Lược đồ đặc trưng HSV Histogram

II.5. Phản hồi liên quan và hiệu chỉnh truy vấn

Phản hồi liên quan trong tra cứu thông tin [46] sử dụng các mẫu tích cực và các mẫu tiêu cực thu được từ người dùng nhằm cải thiện hiệu năng của hệ thống. Nhiều nghiên cứu trong CBIR sử dụng phản hồi liên quan [8, 13, 26, 30, 50].

Hiệu chỉnh truy vấn là việc thay đổi véc tơ đặc trưng của truy vấn bằng một véc tơ mới. Truy vấn mới được xây dựng từ truy vấn hiện thời Q a Q_{new}, Q_{new} được sử dụng cho lần lặp sau.

Trong [16, 42, 48] hiệu chỉnh truy vấn theo công thức Rocchio [40] :

$$Q_{new} = Q + \beta \sum_{i=1}^{n_1} \frac{R_i}{n_1} - \gamma \sum_{i=1}^{n_2} \frac{S_i}{n_2}, \quad (9)$$

R_i là véc tơ cho tài liệu liên quan *i*, S_i là véc tơ cho tài liệu không liên quan *i*, n_1 là số các tài liệu liên quan, n_2 là số các tài liệu không liên quan, β và γ là tham số tuỳ biến.

Một nghiên cứu khác trong MARS [43], và trong [22] điều chỉnh truy vấn theo cách:

$$Q_{new} = \sum_{j=1}^{m} \frac{1}{\sigma_j}(C_j - Q_j)^2, \quad (10)$$

với σ_j là độ lệch chuẩn theo chiều thứ *j* của gía trị đặc trưng, *C* là trọng tâm của các đối tượng liên quan được đánh giá bởi người dùng.

III. KỸ THUẬT ĐỀ XUẤT

Phần này trình bày kỹ thuật đề xuất chuẩn hoá đặc trưng, chuẩn hoá khoảng cách và hiệu chỉnh trọng số. Bảng B.1 (xem phụ lục B) cho biết một số kí hiệu được sử dụng.

III.1. Chuẩn hoá đặc trưng dựa vào phân cụm mờ c-means (Fuzzy c-mean clustering (FCM))

Phân cụm mờ c-mean [3], sử dụng hiệu quả trong một số nghiên cứu CBIR [4, 54].

Để tối thiểu hóa các sai khác do dữ liệu được xem như các đại lượng ngẫu nhiên, có thể có nhiều đỉnh, chúng tôi đề xuất sử dụng phân cụm mờ cho từng bộ đặc trưng cụ thể. Sau khi phân cụm, việc chuẩn hoá được thực hiện theo luật $3\sigma - FCM$ xem như một mở rộng của chuẩn hóa theo luật 3σ.

Cho $\{E_{t,i}\}_{1\le i\le n}$ và các hằng số *p = p(t) > 1*, C=C(t) $\in N^+, C \ge 2$, $m_t = \dim(E_{t,i})$, $\forall 1 \le i \le n$. Thuật toán lặp FCM cực tiểu hóa hàm mục tiêu:

$$J(V,\eta) = \min_{V,\eta} \sum_{i=1}^{n}\sum_{c=1}^{C} \eta_{t,c,i}^{p} \left\| E_{t,i} - V_{t,c} \right\|^2, \quad (11)$$

với độ đo khoảng cách Ơcơlit, $\left\| E_{t,i} - V_{t,c} \right\|^2 = \sum_{j=1}^{m_t} \left(E_{t,i}[j] - V_{t,c}[j] \right)^2$ và các ràng buộc biến như sau:

$$\eta_{t,c,i} \in [0,1], \forall 1 \le i \le n, 1 \le c \le C, t \ge 1,$$

$$\sum_{c=1}^{C} \eta_{t,c,i} = 1, \forall 1 \le i \le n,$$

$$\sum_{i=1}^{n} \eta_{t,c,i} < n, \forall 1 \le c \le C,$$

Định nghĩa 3.1: Độ lệch chuẩn ở cụm *c (1≤c≤C)* trên dữ liệu $\{E_{t,i}\}_{1\le i\le n}$ có *C* cụm:

$$\forall 1 \le j \le m_t, \quad \sigma_{t,c,j} \overset{def}{=} \sqrt{\sum_{i=1}^{n} \eta_{t,c,i}^{p} \left(E_{t,i}[j] - V_{t,c}[j] \right)^2 / \sum_{i=1}^{n} \eta_{t,c,i}^{p}}, \quad (12)$$

Mệnh đề 3.1: Nếu $\{V_c\}_{c=1}^{C}$ tập *C* véc tơ tâm *m* chiều (đầu ra của thuật toán phân cụm sử dụng FCM), ∀ *c=1,2…,C* thì độ lệch chuẩn của cụm c tính theo công thức sau:

$$\sigma_{t,c,j}^2 = \left(\sum_{i=1}^{n} \eta_{t,c,i}^{p} \left(E_{t,i}[j] \right)^2 / \sum_{i=1}^{n} \eta_{t,c,i}^{p} \right) - V_{t,c}^2[j], \quad (13)$$

Chứng minh: xem phụ lục A.

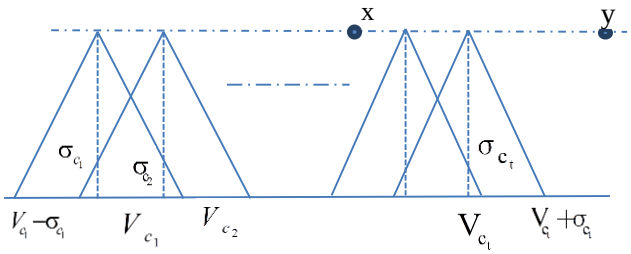
Định nghĩa 3.2: Phép chuẩn hóa $3\sigma - FCM$

$$x = \{x[j]\}_{j=1}^{m_t}, x_{norm} = \{x_{norm}[j]\}_{j=1}^{m_t}, \forall 1 \le j \le m_t, x_{norm}[j] \overset{def}{=} \frac{\min_{1\le c\le C}\left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\} + \max_{1\le c\le C}\left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\}}{C+1}, \quad (14)$$

Mệnh đề 3.2: $3\sigma - \mathrm{FCM}$ bảo toàn thứ tự.

Chứng minh: Xem phụ lục A.

Như vậy $3\sigma - \mathrm{FCM}$ có thể xem là mở rộng của phép chuẩn hóa 3σ (Khi không phân cụm, *C*=1). Hình 2 cho một minh họa của $3\sigma - \mathrm{FCM}$, ngoài ra các phản ví dụ 1-3 cho các phép biến đổi mở rộng 3σ sử dụng FCM không bảo toàn thứ tự cũng được cho dưới đây để so sánh.



Hình 2. Minh hoạ phép chuẩn hoá $3\sigma - \mathrm{FCM}$

Phản ví dụ 1:

$$F_{x[j]} = \frac{x[j] - V_{c_0}[j]}{3\sigma_{c_0,j}} \text{ với } c_0 = \arg\min_{1\le c\le C} \left| \frac{x[j] - V_c[j]}{3\sigma_{c,j}} \right|$$

Phản ví dụ 2:

$$F_{x[j]} = \frac{x[j] - V_{c_0}[j]}{3\sigma_{c_0,j}} \text{ với } c_0 = \arg\min_{1\le c\le C} \sum_{j=1}^{m_t} \left(x[j] - V_c[j] \right)^2$$

Phản ví dụ 3:

$$F_{x[j]} = \sum_{c=1}^{C} \eta_{c,x} \frac{x[j] - V_c[j]}{3\sigma_{c,j}}, \quad \eta_{c,x} = \frac{\left(\sum_{j=1}^{m_t} (x[j] - V_c[j]) \right)^{2\,\frac{1}{p-1}}}{\sum_{c'=1}^{C} \left(\sum_{j=1}^{m_t} (x[j] - V_{c'}[j]) \right)^{2\,\frac{1}{p-1}}}$$

Thuật toán 1 đề xuất chuẩn hoá $3\sigma - FCM$ cho dữ liệu đặc trưng.

Thuật toán 1. Chuẩn hoá đặc trưng $3\sigma - FCM$

Input: $\{E_{t,i}\}_{1\le i\le n;1\le t\le T}$, hằng số p = p(t) > 1, C = C(t) $\in N^+, C \ge 2$, $m_t = \dim(E^F_{t,i}), \forall i = \overline{1,n}$

Output: $\{E^{Norm}_{t,i}\}_{1\le i\le n}$ dữ liệu đã được chuẩn hoá, các tâm $\{V_{t,c}\}_{1\le c\le C_t}$, độ lệch chuẩn $\{\sigma_{t,c,j}\}_{1\le c\le C_t, 1\le j\le m_t}$

Bước 1: $FCM(C_t, p_t)\left(\{E_{t,i}\}_{1\le i\le n;1\le t\le T}\right)$ ta được $\{V_{t,c}\}_{c=1}^{C_t}$, $\{\eta_{t,c,i}\}_{1\le c\le C_t, 1\le i\le n}$ theo công thức (11)

Bước 2: Tính $\{\sigma_{t,c,j}\}_{1\le c\le C_t, 1\le j\le m_t}$ theo công thức (13)

Bước 3: For each $E_{t,i}$: $\forall j = \overline{1, m_t}$ tính $E^{norm}_{t,i}$ [j] theo công thức $3\sigma - FCM$ (3.4)

Return: $\{E^{norm}_{t,i}\}_{1\le i\le n}$, $\{V_{t,c}\}_{1\le c\le C_t}$, $\{\sigma_{t,c,j}\}_{1\le c\le C_t, 1\le j\le m_t}$

Thuật toán 1 có độ phức tạp $O(n * C_t * m_t)$.

III.2. Chuẩn hoá khoảng cách dựa vào phân cụm FCM

Giá trị khoảng cách ảnh truy vấn với mỗi ảnh cơ sở dữ liệu được chuẩn hoá theo thuật toán 2.

Thuật toán 2. Chuẩn hoá dữ liệu khoảng cách $3\sigma - FCM$ theo từng bộ

Input: $\{E^{norm}_{t,i}\}_{1\le i\le n}$, $\{V_{t,c}\}_{1\le c\le C_t}$, $\{\sigma_{t,c,j}\}_{1\le c\le C_t, 1\le j\le m_t}$, hằng số p=p(t) > 1, C= $C^{(D)}_t \in N^+, C \ge 2$

Output: Tập tâm $\{V^{(D)}_{t,c}\}_{1\le c\le C^{(D)}_t}$; độ lệch chuẩn $\{\sigma^{(D)}_{t,c}\}_{1\le c\le C^{(D)}_t}$

Bước 1: (1) Sinh ra 2 tập gồm K giá trị chỉ số ngẫu nhiên $RD_1 = \{RD_{1,i}\}_{1\le i\le K}$, $RD_2 = \{RD_{2,i}\}_{1\le i\le K}$ thỏa mãn:

(1.1) $RD_1 \cap RD_2 = \varnothing$,

$RD_{1,i} < RD_{1,j}, RD_{2,i} < RD_{2,j} \forall 1 \le i < j \le K$

(1.2) $1 \le RD_{1,i}, RD_{2,i} \le n \forall 1 \le i \le n$ (chọn K =[n/10]), M=K^2

(2) Xác định tập giá trị: $D_t(E^{norm}_{t,RD_{1,i}}, E^{norm}_{t,RD_{1,i}})$ được tập giá trị số dương $\{d_k\}_{1\le k\le M}$

Bước 2: $FCM(C^{(D)}_t, p_t)\left(\{d_k\}_{1\le k\le M}\right)$ ta được tập tâm $\{V^{(D)}_{t,c}\}_{c=1}^{C^{(D)}_t}$ và $\{\eta^{(D)}_{t,c,i}\}_{1\le c\le C^{(D)}_t, 1\le i\le M}$

Bước 3: Tính $\{\sigma^{(D)}_{t,c}\}_{1\le c\le C^{(D)}_t}$ theo công thức (13)

(Bước 2 và 3, sử dụng công thức (11), (13), (14) thay thế dữ liệu đặc trưng bằng giá trị khoảng cách là các số thực dương vô hướng.)

Return: $\{V^{(D)}_{t,c}\}_{c=1}^{C^{(D)}_t}$, $\{\sigma^{(D)}_{t,c}\}_{1\le c\le C^{(D)}_t}$

Thuật toán 2 có độ phức tạp $O(M * C^{(D)}_t)$. Qua thuật toán 2 xác định được các giá trị tâm của các cụm theo từng đặc trưng $\{V^{(D)}_{t,c}\}_{1\le c\le C, 1\le t\le T}$ và độ lệch chuẩn $\{\sigma^{(D)}_{t,c}\}_{1\le c\le C, 1\le t\le T}$, các giá trị này được lưu trong cơ sở dữ liệu để sử dụng trong chuẩn hoá lần sau.

III.3. Hiệu chỉnh trọng số và phản hồi liên quan

Trong kỹ thuật đề xuất này, chúng tôi coi đóng góp của mỗi đặc trưng là như nhau, không phân biệt số loại đặc trưng trong mỗi kiểu như [43]. Định nghĩa 3.3 về độ đo khoảng cách giữa hai đối tượng thể hiện điều này.

Định nghĩa 3.3: Độ đo khoảng cách D giữa hai đối tượng $O^1(E^1, F^1, T^1)$ và $O^2(E^2, F^2, T^2)$ dạng rút gọn xác định:

Độ đo khoảng cách của một bộ đặc trưng:

$$D_{ij}\left(t^1_{ij}, t^2_{ij}\right) \overset{def}{=} D_{ij,w_{ijk}}\left(t^1_{ij}, t^2_{ij}\right), \quad (15)$$

Độ đo khoảng cách toàn bộ:

$$D\left(O^1\left(E^1, F^1, T^1\right), O^2\left(E^2, F^2, T^2\right)\right) \overset{def}{=} \sum_{i,j} w_{ij} D_{ij}\left(f^1_i, f^2_i\right), \quad (16)$$

III.3.1. Truy vấn dựa trên thông tin phản hồi

Giả sử mỗi ảnh tương ứng là một mẫu trong không gian đặc trưng F^n và tập tất các mẫu là E^{Norm} có kích thước *n*. Giả định số các lớp *c* được biết, sau các tra cứu bởi các người dùng khác nhau, chúng ta có $E^{Norm} = NB^* \cup NB^+ \cup NB^-$, $NB = NB^+ \cup NB^-$, thông thường *#NB* là hằng số nhỏ thuộc [20, 40].

Định nghĩa 3.4: Tập đồng ý (Agreement) giữa độ đo toàn cục và độ đo theo bộ được định nghĩa:

$AGR_t(D, D_t, N) \overset{\text{def}}{=} NB \cap NB_t$, trong đó *NB*, NB_t tương ứng là tập *N* ảnh có độ đo khoảng cách cao nhất theo độ đo toàn cục *D*, và theo độ đo D_t của riêng bộ đặc trưng *t*.

Trong thực tế thường chọn *N*=20, và cho trước *D*, D_t nên chúng ta sẽ viết gọn là AGR_t.

Chúng tôi tiến hành thử nghiệm trên một số tập ảnh NB^+, NB^- và $NB^\sim$ ($NB^\sim \subset NB$) với một hàm độ đo khoảng cách toàn cục $D(Q^{Norm}, E_i^{Norm}) = \sum_{t=1}^{T} w_t D_t(Q_t^{Norm}, E_{t,i}^{Norm})$ và các hàm độ đo khoảng cách cục bộ $D_t(Q_t^{Norm}, E_{t,i}^{Norm})$.

Các ví dụ trong các hình 3.2.a và 3.2.b tính độ đo khoảng cách một số ảnh trong tập thử nghiệm (phần 4). Ký hiệu các cột (d1), (d2), (d3), (d4), (d5), (d6) tương ứng độ đo khoảng cách theo đặc trưng hsv Histogram [9, 43, 51], autoCorrelogram, Color moment, Gabor texture [21], Wavelet moment và Gist. Ký hiệu các hàm đo khoảng cách (f1): Histogram Intersection, (f2): L_2, (f3): L_1, (f4): Canberra.

Sử dụng truy vấn *Q* = {710.jpg}, theo nhận thức chủ quan chọn ra các tập NB^+ = {717.jpg, 704.jpg, 723.jpg, 700.jpg, 721.jpg}, NB^- = {100.jpg, 101.jpg, 102.jpg, 103.jpg, 104.jpg} và $NB^\sim$ ={676.jpg, 535.jpg, 509.jpg, 566.jpg, 551.jpg} (nằm trong tập thử nghiệm). Hình 3.2.a tính độ đo khoảng cách cho các tập NB^+, NB^-, $NB^\sim$ theo hàm khoảng cách tương ứng ở trên và hình 3.2.b sử dụng hàm khoảng cách L_2.

Qua các phép thử như Hình 3.a và 3.b chúng tôi nhận thấy sự phù hợp của các hàm khoảng cách (f1), (f2), (f3), (f4) và (f5) cho các bộ đặc trưng tương ứng (d1), (d2), (d3), (d4), (d5) và (d6). Một nhận định rút ra là: để hạn chế tối đa các ảnh nằm trong tập $NB^\sim \in NB$ thì tập AGR_t cần được sử dụng làm cơ sở hiệu chỉnh trọng số w_t.

Q	710.jpg	(d1) (f1)	(d2) (f2)	(d3) (f3)	(d4) (f4)	(d5) (f2)	(d6) (f2)	D
NB_1^+	717.jpg	0.024	0.368	0.349	0.571	0.232	0.322	1.867
NB_2^+	704.jpg	0.079	0.447	0.378	0.42	0.377	0.173	1.875
NB_3^+	723.jpg	0.023	0.416	0.318	0.45	0.295	0.395	1.898
NB_4^+	700.jpg	0.036	0.357	0.196	0.583	0.394	0.352	1.919
NB_5^+	721.jpg	0.074	0.441	0.295	0.43	0.34	0.365	1.945
NB_1^-	100.jpg	0.49	0.446	0.274	0.618	0.298	0.377	2.503
NB_2^-	101.jpg	0.51	0.382	0.517	0.61	0.442	0.278	2.738
NB_3^-	102.jpg	0.495	0.423	0.459	0.439	0.467	0.257	2.54
NB_4^-	103.jpg	0.543	0.448	0.374	0.502	0.471	0.405	2.743
NB_5^-	104.jpg	0.179	0.477	0.385	0.625	0.493	0.559	2.719
$NB_1^\sim$	676.jpg	0.01	0.355	0.353	0.442	0.354	0.309	1.823
$NB_2^\sim$	535.jpg	0.206	0.39	0.26	0.441	0.414	0.231	1.942
$NB_3^\sim$	509.jpg	0.246	0.445	0.286	0.434	0.323	0.237	1.97
$NB_4^\sim$	566.jpg	0.161	0.4	0.348	0.323	0.451	0.293	1.976
$NB_5^\sim$	551.jpg	0.254	0.418	0.217	0.427	0.446	0.256	2.018

Hình 3.a. Độ đo khoảng cách trên các tập NB^+, NB^-, $NB^\sim$

Q	710.jpg	(d1) (f2)	(d2) (f2)	(d3) (f2)	(d4) (f2)	(d5) (f2)	(d6) (f2)	D
NB_1^+	717.jpg	0.194	0.368	0.354	0.577	0.232	0.322	2.05
NB_2^+	704.jpg	0.312	0.447	0.4	0.422	0.377	0.173	2.13
NB_3^+	723.jpg	0.223	0.416	0.302	0.394	0.295	0.395	2.03
NB_4^+	700.jpg	0.254	0.357	0.19	0.514	0.394	0.352	2.06
NB_5^+	721.jpg	0.32	0.441	0.273	0.414	0.34	0.365	2.15
NB_1^-	100.jpg	0.518	0.446	0.258	0.681	0.298	0.377	2.58
NB_2^-	101.jpg	0.525	0.382	0.51	0.382	0.442	0.278	2.52
NB_3^-	102.jpg	0.478	0.423	0.462	0.372	0.467	0.257	2.46
NB_4^-	103.jpg	0.562	0.448	0.399	0.472	0.471	0.405	2.76
NB_5^-	104.jpg	0.347	0.477	0.405	0.402	0.493	0.559	2.68
$NB_1^\sim$	676.jpg	0.234	0.355	0.332	0.418	0.354	0.309	2
$NB_2^\sim$	535.jpg	0.312	0.39	0.234	0.443	0.414	0.231	2.03
$NB_3^\sim$	509.jpg	0.437	0.445	0.284	0.421	0.323	0.237	2.15
$NB_4^\sim$	566.jpg	0.361	0.4	0.337	0.341	0.451	0.293	2.18
$NB_5^\sim$	551.jpg	0.378	0.418	0.221	0.355	0.446	0.256	2.07

Hình 3.b. Độ đo khoảng cách của tập NB^+, NB^-, $NB^\sim$

Ba luật *R1*, *R2*, *R3* được rút ra khá phù hợp với trực giác như sau:

R1. $\forall E_i^{Norm} \in AGR_t$, tăng w_t nếu E_i^{Norm} là phản hồi dương, giảm w_t nếu ngược lại.

R2. Độ lệch chuẩn $\sigma_{l,t,D_t}^{NB^+}$ càng nhỏ thì trọng số w_t điều chỉnh tăng (giảm) càng nhiều.

R3. Độ lệch chuẩn $\sigma_{\sigma_{l,t,\|I_t\|}}^{NB+}$ càng nhỏ thì trọng số w_t điều chỉnh tăng (giảm) càng nhiều.

III.3.2. Cập nhật trọng số

Sử dụng hai thông tin quan trọng: (a) kiến thức chủ quan của người dùng đánh giá mức độ liên quan của các đối tượng nằm trong tập *NB*; (b) thông tin mức độ quan trọng của dữ liệu đặc trưng mức thấp được xác định nhờ tập AGR_t.

(a) Dựa vào kiến thức thức người dùng:

Trên tập *NB*, người dùng đặt các mức độ liên quan cho các đối tượng. Chúng tôi thiết lập bảy mức độ liên quan phù hợp với nhận thức ngữ nghĩa của người dùng, đó là $score_l$ ={-3, -2, -1, 0, 1, 2, 3}, tương ứng với ý nghĩa "*không liên quan rất cao*", "*không liên quan cao*", "*không liên quan*", "*không ý kiến*", "*liên quan*", "*liên quan cao*", "*liên quan rất cao*", và sau đó tính giá trị trung bình và phân bố chuẩn trên tập NB^+ cho đại lượng độ đo khoảng cách và độ dài véc tơ như sau:

- Tính độ đo khoảng cách theo từng bộ đặc trưng t với truy vấn: $d_t = D_t(Q_t^{Norm}, E_{t,i}^{Norm})$.
- Tính giá trị $\sigma_{l,t,D_t}^{NB^+}$ trên tập dữ liệu khoảng cách của các phản hồi dương.
- Tính giá trị $\sigma_{l,t,\|I_t\|}^{NB^+}$ trên tập dữ liệu độ dài véc tơ của các phản hồi dương.

(b) Học từ dữ liệu đặc trưng mức thấp kết hợp kiến thức thu được từ người dùng:

Sử dụng tập AGR_t tính điều chỉnh tăng hoặc giảm trọng số w_t theo từng bộ đặc trưng *t*:

For each *l* của phản hồi

For each bộ *t*

For each $I \in AGR_t$

$$w_t^{(l+1)} = max\left(w_t^{(l)} + \Delta w_t^{(l)}, 0\right), \quad (17)$$

$\forall t = \overline{1,T}$, gán lại $w_t^{(l+1)} = w_t^{(l+1)} / \sum_{t'=1}^{T} w_{t'}^{(l+1)}$,

Ở đó $\Delta w_t^l = f(score(E_{t,i}^{Norm}), \sigma_{l,t,D_t}^{NB^+}, \sigma_{l,t,\|I_t\|}^{NB^+}) \in (-\infty, +\infty)$,

hàm *f* có thể chọn chẳng hạn:

$$f(score(E_{t,i}^{Norm}), \sigma_{l,t,D_t}^{NB^+}, \sigma_{\sigma_{l,t,\|I_t\|}}^{NB+}) = \frac{score(E_{t,i}^{Norm})}{3} / (\sigma_{l,t,D_t}^{NB^+} * \sigma_{\sigma_{l,t,\|I_t\|}}^{NB+}), \quad (18)$$

Thuật toán 3. IR-FCM (Hiệu chỉnh trọng số độ đo khoảng cách toàn cục)

Input: $\{E_{t,i}^{Norm}\}_{1\le i\le n}$, $\{V_{t,c}\}_{1\le c\le C_t}$, $\{\sigma_{t,c,j}\}_{1\le c\le C_t, 1\le j\le m_t}$, $\{V_{t,c}^{(D)}\}_{1\le t\le T, 1\le c\le C_t^{(D)}}$, $\{\sigma_{t,c}^{(D)}\}_{1\le c\le C_t^{(D)}}$

/ Véc tơ đặc trưng mức thấp Q của ảnh truy vấn. Lmax số lần lặp phản hồi, K: số ảnh có độ đo khoảng cách toàn cục so với Q là bé nhất. */*

Output: Tập trọng số $\{w_t\}_{1\le t\le T}$

Bước 1: Khởi tạo, $\forall t = \overline{1,T}$: $w_t = \frac{1}{T}$, chuẩn hóa từng bộ *t* của *Q* theo $3\sigma - FCM$

Bước 2: Lặp lại cho mỗi *l* từ 1 đến L_{max}

Bước 3: Bước lặp phản hồi

3.1: For each $E_i^{norm} \in \{E_i^{norm}\}_{1\le i\le n}$: $\forall t = \overline{1,T}$

3.1.1: Chuẩn hoá độ đo khoảng cách cho bộ *t* bằng cách: $D_t(E_{t,i}^{Norm}) = 3\sigma - FCM\ D_t(Q_t^{Norm}, E_{t,i}^{Norm})$

3.1.2: Chuyển $D_t(E_{t,i}^{Norm})$ về [0,1] : $D_t(E_{t,i}^{Norm}) = \min\left(\max\left(\frac{D_t(E_{t,i}^{Norm})+1}{2}, 0\right), 1\right)$

3.1.3: Tính độ đo khoảng cách $D(E_i^{norm}) = \sum_{t=1}^{T} w_t D_t(E_{t,i}^{norm})$

3.1.4: Hiệu chỉnh trọng số dựa trên AGR_t

For each $I \in NB^{(l)}$, $\forall t = \overline{1,T}$ nếu $I \in AGR_t^{(l)}$ thì hiệu chỉnh w_t theo công thức (3.7) và (3.8)

3.1.5: Chuẩn hóa lại trọng số w_t, $\forall t = \overline{1,T}$, $w_t = w_t / \sum_{t=1}^{T} w_t$

Return: Kết thúc bước 2, thu được $\{w_t\}_{1\le t\le T}$

Thuật toán 3 đề xuất hiệu chỉnh độ đo khoảng cách toàn cục dựa trên độ đo khoảng cách theo bộ và phép chuẩn hóa $3\sigma - FCM$, thuật toán có độ phức tạp $O\left(n * T * max\{C_t, C_t^{(D)}\}\right)$.

III.3.3. Hiệu chỉnh truy vấn

Đặt $\overline{Q}_t \overset{\text{def}}{=} \dfrac{1}{\#\text{AGR}_t} \sum\limits_{E_{t,i}^{norm} \in AGR_t} E_{t,i}^{norm}$,

$$\overline{\sigma}_t \overset{\text{def}}{=} \sqrt{\frac{1}{\#AGR_t} \sum_{E_i^{norm} \in AGR_t} \left(E_{t,i}^{norm}\right)^2 - \overline{Q}_t^{\,2}} \tag{19}$$

$$AGR_t^* \overset{\text{def}}{=} \left\{ E_{t,i}^{norm} \in AGR_t \,\middle|\, \left| E_{t,i}^{norm} - \overline{Q}_t \right| \le 3\overline{\sigma}_t \right\} \tag{20}$$

(AGR_t có sự đồng nhất cao về đặc trưng, nên ta dùng biến đổi 3σ). Khi $\#AGR_t \ge 2$ và $\#AGR_t^* \ge 1$, truy vấn mới ứng với bộ đặc trưng *t* được lập lại như sau:

$$Q_{t,new} \overset{\text{def}}{=} \frac{1}{\#AGR_t^*} \sum_{E_{t,i}^{norm} \in AGR^*} E_{t,i}^{norm} \tag{21}$$

Truy vấn mới Q_{new} là biểu diễn tốt, mang thông tin của từng bộ đặc trưng mà người dùng quan tâm.

IV. THỰC NGHIỆM

IV.1. Cơ sở dữ liệu ảnh

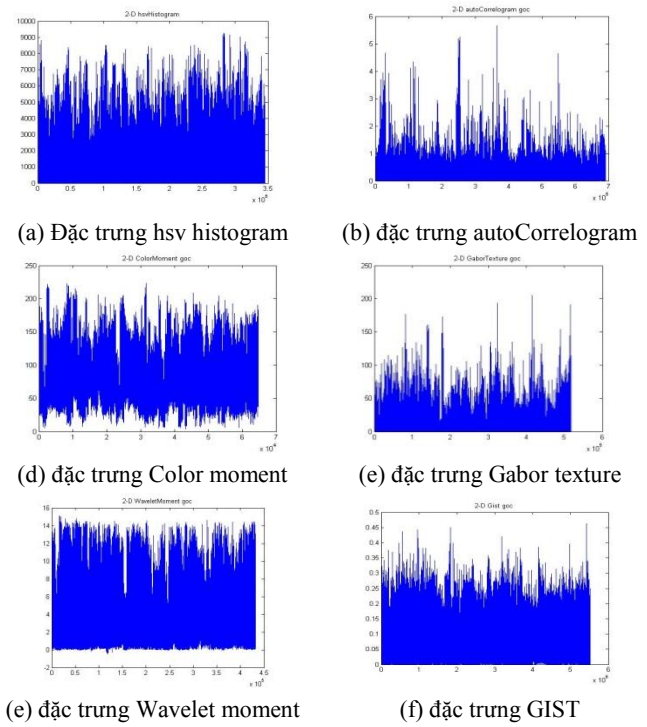
Nhiều nghiên cứu sử dụng cơ sở dữ liệu chuẩn Corel như trong [24, 32, 43, 53]. Cơ sở dữ liệu tiêu chuẩn khác như ImageCLEF được sử dụng ở [38], MIT VisTex sử dụng trong thực nghiệm [2, 11].

Tập ảnh Corel (http://corel.digitalriver.com) bao gồm khoảng 68040 ảnh màu gồm nhiều chủ đề. Có khoảng 100 ảnh trong mỗi chủ đề. Thực nghiệm tra cứu sử dụng tập Wang[1] là tập con của tập Corel với 10 chủ đề đó là: Biển, Châu Phi, hoa hồng, ngựa, núi, thức ăn, xe buýt, khủng long, lâu đài, voi. Các ảnh cùng chủ đề được xem như là liên quan cao nhất, và các ảnh thuộc chủ đề gần gũi được xem như là có liên quan (ví dụ chủ đề ngựa và voi) được xem như là liên quan.

Trong chuẩn hoá khoảng cách với tập ảnh lớn chúng ta nên chọn K ở khoảng 10% tới 50% số lượng ảnh của tập. Sử dụng 100 ảnh ngẫu nhiên làm truy vấn và đánh giá chất lượng tra cứu trên các lần lặp với các truy vấn khởi tạo này. Mỗi truy vấn thực hiện 10 lần lặp phản hồi và độ đo hiệu năng được đánh giá bằng trung bình của toàn bộ 100 truy vấn này.

IV.2. Trích rút bộ đặc trưng kết hợp

Như đã trình trong phần hai, chúng tôi lựa chọn bộ đặc trưng gồm lược đồ màu hsv (hsv histogram) [51], gắn kết màu (autoCorrelogram) [20], bất biến màu (Colormoment) [49], kết cấu Gabor (Gabor Texture) [55], bất biến Wavelet (Wavelet Moment) [18], hình dạng GIST [34]. Hình 4 là lược đồ phân bố dữ liệu theo bộ đặc trưng trên tập dữ liệu tiêu chuẩn này.



(a) Đặc trưng hsv histogram (b) đặc trưng autoCorrelogram

(d) đặc trưng Color moment (e) đặc trưng Gabor texture

(e) đặc trưng Wavelet moment (f) đặc trưng GIST

Hình.4. Lược đồ đặc trưng thô

Các đặc trưng trên được chuẩn hoá theo luật $3\sigma - FCM$ sử dụng các tham số phân cụm FCM trong Bảng 1. Các tham số được lựa chọn đảm bảo dữ liệu đặc trưng trong khoảng [-1,1] nhiều nhất.

Bảng 1. Tham số phân cụm FCM

hsv Histogram		auto Correlogram		Color Moment		Gabor Texture		Wavelet Moment		Gist	
C	p	C	p	C	p	C	p	C	p	C	p
8	1.2	13	1.8	5	2	5	2	5	2	5	2

[1] http://wang.ist.psu.edu/docs/related.shtml

IV.3. Các kết quả và luận giải

Hai chỉ số khách quan, độ chính xác (Precision, ký hiệu là *Pr*) và độ triệu hồi (Recall, ký hiệu là Re) [45] được sử dụng trong thực nghiệm để đánh giá hiệu năng của đề xuất.

Recall được định nghĩa là tỉ số của số ảnh liên quan được tra cứu (ký hiệu là *NR*) với toàn bộ ảnh liên quan trong cơ sở dữ liệu (ký hiệu là *NT*).

$Pr = NR/N$, $Re = NR/NT$

Các đề xuất đã được so sánh với nghiên cứu liên quan trước ở các khía cạnh:

- Độ đo khoảng cách giữa đối tượng ảnh chúng tôi sử dụng định nghĩa 3.3 theo từng bộ đặc trưng nên đơn giản hơn, chính xác hơn. Trong khi [43] sử dụng định nghĩa 2.2 theo kiểu đặc trưng.
- Sử dụng chuẩn hoá $3\sigma - FCM$ rất phù hợp dữ liệu kết hợp nhiều đặc trưng, đảm bảo 99% dữ liệu sau chuẩn hoá thuộc [-1,1]. Trong khi đó nhiều nghiên cứu [7, 28, 35, 42, 43] sử dụng chuẩn hoá 3σ và trong [2] sử dụng chuẩn hoá min-max có nhiều hạn chế (đã chứng minh trong mục 2 và 3). Chúng tôi sử dụng hàm chuẩn hoá khoảng cách d=min(max(d+1/2,0),1) nên độ mất thông tin ít hơn của các phương pháp khác khi quy khoảng cách về [0,1].
- Trong hiệu chỉnh trọng số khoảng cách, [43] chỉ sử dụng độ lệch chuẩn của độ dài véc tơ đặc trưng, phương pháp của chúng tôi sử dụng kết hợp cả hai độ lệch trên độ dài véc tơ và độ dài theo bộ đặc trưng (phương trình 3.8). Sự khác nhau ở đây là chúng tôi sử dụng tập AGR_t theo bộ đặc trưng, trong [43] sử dụng tập AGR_t theo kiểu đặc trưng.

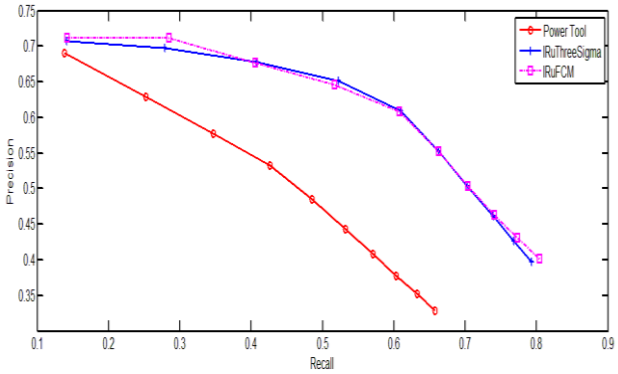
Hiệu chỉnh truy vấn trong phương pháp của chúng tôi sử dụng tập AGR_t theo bộ đặc trưng có kết hợp thông tin phản hồi và thông tin dữ liệu. Tập AGR_t thay đổi sau mỗi phản hồi.

Chúng tôi thực hiện tra cứu theo kỹ thuật Power Tool [43] và kỹ thuật đề xuất trên đặc trưng chuẩn hoá theo 3σ (*IR-3Sigma*) và đặc trưng chuẩn hoá theo $3\sigma - FCM$ (IR-FCM). Bảng 2 cho biết độ chính xác tra cứu với số lần lặp *l*.

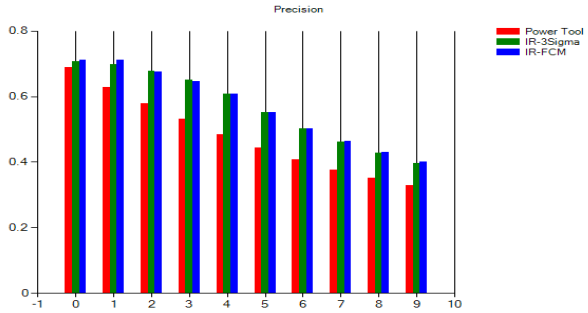
Bảng 2. Độ chính xác trên l lần lặp

Kỹ thuật	*l=1*	*l=2*	*l=3*	*l=4*	*l=5*	Trung bình
Power Tool	0.69	0.629	0.578	0.532	0.485	58.2%
IR-3Sigma	0.707	0.698	0.678	0.652	0.609	66.9%
IR-FCM	0.712	0.712	0.676	0.646	0.608	67.1%

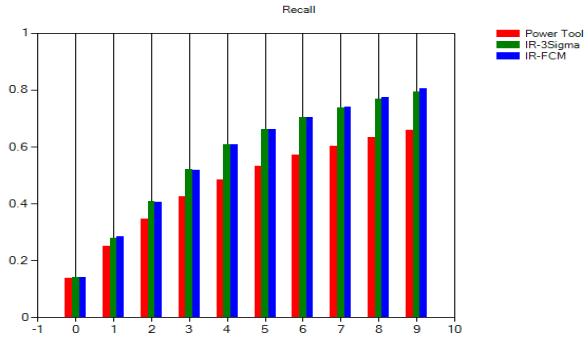
Các Hình 5, Hình 6 và Hình 7 là biểu đồ của độ chính xác và triệu hồi, độ chính xác, triệu hồi tương ứng trong 10 lần lặp thể hiện tính hiệu quả của kĩ thuật đề xuất.



Hình 5. Biểu đồ độ chính xác và triệu hồi



Hình 6. Biểu đồ độ chính xác từng lần lặp



Hình 7. Biểu đồ triệu hồi từng lần lặp

V. KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TIẾP THEO

Phương pháp chuẩn hoá $3\sigma - FCM$ được đề xuất cho dữ liệu không phải là một phân bố Gauss, đảm bảo 99% nằm trong phạm vi [-1,1]. Hệ thống tra cứu ảnh theo nội dung cải thiện độ chính xác đạt đến 67.1% do chỉnh trọng số khoảng cách bằng thông tin phản hồi và tập ARG_t.

Trong tương lai chúng tôi tiếp tục sử dụng phương pháp chuẩn hoá này và cải tiến phương pháp hiệu chỉnh trọng số dựa vào thông tin phản hồi liên quan để tăng hiệu năng hệ thống tra cứu.

LỜI CẢM ƠN

Chúng tôi xin cám ơn sự hỗ trợ từ đề tài CS'16.03 "Phát triển một số phương pháp tra cứu ảnh dựa vào nhận thức trực quan và đa truy vấn" của Viện CNTT, Viện Hàn lâm KH&CN Việt Nam.

PHỤ LỤC A

Mệnh đề 3.1: (Chứng minh).

Theo công thức tính tâm dựa vào FCM ta có:

$$\forall 1 \le j \le m_t, V_{t,c,j} = \sum_{i=1}^{n} \eta_{t,c,i}^{p} E_{t,i}[j] / \sum_{i=1}^{n} \eta_{t,c,i}^{p},$$

$$\forall 1 \le j \le m_t, \sigma_{t,c,j}^2 = \sum_{i=1}^{n} \eta_{c,i}^{p} \left(E_{t,i}[j] - V_{t,c}[j]\right)^2 / \sum_{i=1}^{n} \eta_{c,i}^{p} =$$

$$= \left(\sum_{i=1}^{n} \eta_{c,i}^{p} \left(E_{t,i}[j]\right)^2 - 2V_{t,c}[j] \sum_{i=1}^{n} \eta_{c,i}^{p} E_{t,i}[j]\right) / \sum_{i=1}^{n} \eta_{c,i}^{p} + V_{t,c}^2$$

$$\Rightarrow \sigma_{t,c,j}^2 = \sum_{i=1}^{n} \eta_{c,i}^{p} \left(E_{t,i}[j]\right)^2 / \sum_{i=1}^{n} \eta_{c,i}^{p} - 2V_{t,c}^2 + V_{t,c}^2$$

$$= \sum_{i=1}^{n} \eta_{c,i}^{p} \left(E_{t,i}[j]\right)^2 / \sum_{i=1}^{n} \eta_{c,i}^{p} - V_{t,c}^2.$$

Mệnh đề 3.2: (Chứng minh).

$$\forall 1 \le c \le C, \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \ge \frac{y[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \Rightarrow \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \ge \min_{1 \le c \le C} \left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\}$$

$$\Rightarrow \min_{1 \le c \le C} \left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\} \ge \min_{1 \le c \le C} \left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\}$$

$$\forall 1 \le c \le C, \frac{y[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \le \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \Rightarrow \frac{y[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \le \max_{1 \le c \le C} \left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\}$$

$$\Rightarrow \max_{1 \le c \le C} \left\{ \frac{x[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\} \ge \max_{1 \le c \le C} \left\{ \frac{y[j] - V_{t,c}[j]}{3\sigma_{t,c,j}} \right\}$$

PHỤ LỤC B

***Bảng B.1**. Các ký hiệu và các định nghĩa*

Kí hiệu	Định nghĩa
I, Q	Ảnh cơ sở dữ liệu và ảnh truy vấn
n	Kích thước của cơ sở dữ liệu ảnh
F	Không gian đặc trưng
m	Độ dài của một véc tơ đặc trưng
T={t1,..,tk}	Tập kết hợp nhiều đặc trưng
E	Dữ liệu đặc trưng
Ei (E_i^{Norm})	Đặc trưng thô (chuẩn hoá) của ảnh thứ i
$E_{t,i}$ ($E_{t,i}^{Norm}$)	Đặc trưng thô (chuẩn hoá) bộ t của ảnh thứ i
Qt(Q_t^{Norm})	Đặc trưng thô (chuẩn hoá) bộ đặc trưng t ảnh truy vấn
D (D_t)	Hàm khoảng cách toàn bộ (bộ đặc trưng t)
wt	Trọng số khoảng cách của bộ đặc trưng t
NB	Tập đối tượng ảnh có độ đo khoảng cách nhỏ nhất theo đặc trưng toàn cục
NBt	Tập ảnh có độ đo khoảng cách cao theo đặc trưng t
NB+	Tập ảnh xác nhận tích cực
NB-	Tập ảnh xác nhận tiêu cực
NB~	Tập ảnh có độ đo khoảng cách cao và thuộc tập NB-
NB*	Tập ảnh chưa được tra cứu
$V_t^{(D)}$	Tập tâm cụm theo đặc trưng t
$V_{t,c,j}$	Tâm cụm c của thành phần đặc trưng j ở bộ đặc trưng t (theo FCM)
$V_{t,c}^{(D)}$	Tâm cụm c giá trị khoảng cách Dt (theo FCM)
$\eta_{t,c,i}^{p}$	Giá trị độ thuộc của phần tử đặc trưng i ở bộ t so với tâm c, p là hệ số FCM
$\sigma_{t,c,j}$	Độ lệch chuẩn thành phần j của đặc trưng t theo cụm c
$\sigma_{t,c}^{(D)}$	Độ lệch chuẩn khoảng cách đặc trưng t với tâm c
$\sigma_{l,t,D_t}^{NB^+}$	Độ lệch chuẩn khoảng cách bộ t trong lần lặp l đối với các ảnh ở tập NB+
$\sigma_{\sigma_{l,t,\|I_t\|}}^{NB+}$	Độ lệch chuẩn theo độ dài đặc trưng bộ t trong lần lặp l cho ảnh I trong tập NB+

TÀI LIỆU THAM KHẢO

[1] ANDROUTSOS, PANAGIOTIS, et al, *Aggregation of color and shape features for hybrid query generation in content based visual information retrieval*, Signal Processing 85.2 (2005): 385-393.

[2] BAI, CONG, KIDIYO KPALMA, and JOSEPH RONSIN, *Color textured image retrieval by combining texture and color features*, Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European. IEEE, 2012.

[3] BEZDEK, JAMES C, *Pattern recognition with fuzzy objective function algorithms*, Springer Science & Business Media, 2013.

[4] BHANU, BIR, AND ANLEI DONG, *Concepts learning with fuzzy clustering and relevance feedback*, Engineering Applications of Artificial Intelligence 15.2 (2002): 123-138.

[5] BINAGHI, E., et al, *Fuzzy reasoning approach to similarity evaluation in image analysis*, International Journal of Intelligent Systems 8.7 (1993): 749-769.

[6] CHANG, SHIH-FU, et al, *Visual information retrieval from large distributed online repositories*, Communications of the ACM 40.12 (1997): 63-71.

[7] CIOCCA, GIANLUIGI, and RAIMONDO SCHETTINI, *A relevance feedback mechanism for content-based image retrieval*, Information processing & management 35.5 (1999): 605-632.

[8] COX, INGEMAR J., et al. *"The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments."* Image Processing, IEEE Transactions on 9.1 (2000): 20-37.

[9] ŞAYKOL, EDIZ, UĞUR GÜDÜKBAY, and ÖZGÜR ULUSOY, *A histogram-based approach for object-based query-by-shape-and-color in image and video databases*, Image and Vision Computing 23.13 (2005): 1170-1180.

[10] DENG, YINING, et al, *An efficient color representation for image retrieval*, Image Processing, IEEE Transactions on 10.1 (2001): 140-147.

[11] DO, MINH N., and MARTIN VETTERLI, *Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance*, Image Processing, IEEE Transactions on 11.2 (2002): 146-158.

[12] DUBEY, RAJSHREE S., RAJNISH CHOUBEY, and JOY BHATTACHARJEE, *Multi feature content based image retrieval*, International Journal on Computer Science and Engineering 2.6 (2010): 2145-2149.

[13] GRIGOROVA, ANELia, et al, *Content-based image retrieval by feature adaptation and relevance feedback*, Multimedia, IEEE Transactions on 9.6 (2007): 1183-1192.

[14] Gudivada, Venkat N., and Vijay V. Raghavan, *Content based image retrieval systems*, Computer 28.9 (1995): 18-22.

[15] HARALICK, ROBERT M., KARTHIKEYAN SHANMUGAM, and ITS' HAK DINSTEIN, *Textural features for image classification*, Systems, Man and Cybernetics, IEEE Transactions on 6 (1973): 610-621.

[16] Harman, Donna, *Relevance Feedback and Other Query Modification Techniques*, (1992): 241-263.

[17] HIEU VU VAN, QUYNH NGUYEN HUU, and HA NGUYEN THI THU, *Content based image retrieval with bin of color histogram*, Audio, Language and Image Processing (ICALIP), 2012 International Conference on, IEEE, 2012.

[18] HIREMATH, P. S., S. SHIVASHANKAR, and JAGADEESH PUJARI, *Wavelet based features for color texture classification with application to CBIR*, International Journal of Computer Science and Network Security 6.9A (2006): 124-133.

[19] HUANG, JING, S. RAVI KUMAR, and MANDAR MITRA, *Combining supervised learning with color correlograms for content-based image retrieval*, Proceedings of the fifth ACM international conference on Multimedia, ACM, 1997.

[20] HUANG, JING, et al, *Image indexing using color correlograms*, Computer Vision and Pattern Recognition, 1997, Proceedings., 1997 IEEE Computer Society Conference on, IEEE, 1997.

[21] HUANG, ZHI-CHUN, et al, *Content-based image retrieval using color moment and Gabor texture feature*, Machine Learning and Cybernetics (ICMLC), 2010 International Conference on, Vol. 2, IEEE, 2010.

[22] ISHIKAWA, YOSHIHARU, RAVISHANKAR SUBRAMANYA, and CHRISTOS FALOUTSOS, *MindReader: Querying databases through multiple examples*, Computer Science Department (1998): 551.

[23] JOSE, SEBIN, and PHILUMON JOSEPH, *Content based Image Retrieval System with Watermarks and Relevance Feedback*, International Journal of Computer Applications 99.11 (2014): 1-6.

[24] KIM, DEOK-HWAN, CHIN-WAN CHUNG, and KOBUS BARNARD, *Relevance feedback using adaptive clustering for image similarity retrieval*, Journal of Systems and Software 78.1 (2005): 9-23.

[25] KOKARE, MANESH, PRABIR K. BISWAS, and BISWANATH N. CHATTERJI, *Texture image retrieval using new rotated complex wavelet filters*, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 35.6 (2005): 1168-1178.

[26] LU, YE, et al, *A unified framework for semantics and feature based relevance feedback in image retrieval*

systems, Proceedings of the eighth ACM international conference on Multimedia, ACM, 2000.

[27] MANJUNATH, BANGALORE S., and WEI-YING MA, *Texture features for browsing and retrieval of image data,* Pattern Analysis and Machine Intelligence, IEEE Transactions on 18.8 (1996): 837-842.

[28] MEHROTRA, SHARAD, et al, *Multimedia analysis and retrieval system,* Proc. of The 3rd Int, Workshop on Information Retrieval Systems, 1997.

[29] MEHTRE, BABU M., MOHAN S. KANKANHALLI, and WING FOON LEE, *Content-based image retrieval using a composite color-shape approach,* Information Processing & Management 34.1 (1998): 109-120.

[30] MEILHAC, CHRISTOPHE, and CHAHAB NASTAR, *Relevance feedback and category search in image databases,* Multimedia Computing and Systems, 1999, IEEE International Conference on, Vol. 1, IEEE, 1999.

[31] MINKA, THOMAS P., and ROSALIND W. PICARD, *Interactive learning with a "society of models",* Computer Vision and Pattern Recognition, 1996, Proceedings CVPR'96, 1996 IEEE Computer Society Conference on, IEEE, 1996.

[32] MOGHADDAM, H. ABRISHAMI, and M. SAADATMAND-TARZJAN, *Gabor wavelet correlogram algorithm for image indexing and retrieval,* Pattern Recognition, 2006, ICPR 2006, 18th International Conference on, Vol. 2, IEEE, 2006.

[33] OHANIAN, PHILIPPE P., and RICHARD C. DUBES, *Performance evaluation for four classes of textural features,* Pattern recognition 25.8 (1992): 819-833.

[34] OLIVA, AUDE, and ANTONIO TORRALBA, *Modeling the shape of the scene: A holistic representation of the spatial envelope,* International journal of computer vision 42.3 (2001): 145-175.

[35] ORTEGA, MICHAEL, et al, *Supporting similarity queries in MARS,* Proceedings of the fifth ACM international conference on Multimedia, ACM, 1997.

[36] PASS, GREG, RAMIN ZABIH, and JUSTIN MILLER, *Comparing images using color coherence vectors,* Proceedings of the fourth ACM international conference on Multimedia, ACM, 1997.

[37] PENTLAND, ALEX, ROSALIND W. PICARD, and STAN SCLAROFF, *Photobook: Content-based manipulation of image databases,* International journal of computer vision 18.3 (1996): 233-254.

[38] RAHMAN, M. M., BIPIN C. DESAI, and PRABIR BHATTACHARYA, *Multi–modal interactive approach to ImageCLEF 2007 photographic and medical retrieval tasks by CINDI,* Working Notes of CLEF 7 (2007).

[39] RAVANI, REZA, MOHAMAD REZA MIRALI, and MARYAM BANIASADI, *Parallel CBIR system based on color coherence vector,* 17th International Conference on Systems, Signals and Image Processing, 2010.

[40] ROCCHIO, JJ, *Relevance Feedback in Information Retrieval,* SMART Retrieval System Experimens in Automatic Document Processing (1971).

[41] RUI, YONG, et al, *Automatic matching tool selection using relevance feedback in MARS,* Proc. of 2nd Int. Conf. on Visual Information Systems, 1997.

[42] RUI, YONG, THOMAS S. HUANG, and SHARAD MEHROTRA, *Content-based image retrieval with relevance feedback in MARS,* Image Processing, 1997, Proceedings., International Conference on, Vol. 2, IEEE, 1997.

[43] RUI, YONG, et al, *Relevance feedback: a power tool for interactive content-based image retrieval,* Circuits and Systems for Video Technology, IEEE Transactions on 8.5 (1998): 644-655.

[44] RUI, YONG, THOMAS S. HUANG, and SHIH-FU CHANG, *Image retrieval: Current techniques, promising directions, and open issues,* Journal of visual communication and image representation 10.1 (1999): 39-62.

[45] SALTON, GERARD, and MICHAEL J. MCGILL, *Introduction to modern information retrieval,* (1986).

[46] SCLAROFF, STAN, LEONID TAYCHER, and MARCO LA CASCIA, *Imagerover: A content-based image browser for the world wide web,* Content-Based Access of Image and Video Libraries, 1997, Proceedings, IEEE Workshop on, IEEE, 1997.

[47] SMITH, JOHN R., and SHIH-FU CHANG, *VisualSEEk: a fully automated content-based image query system,* Proceedings of the fourth ACM international conference on Multimedia, ACM, 1997.

[48] SPINK, AMANDA, and ROBERT M. LOSEE, *Feedback in information retrieval,* Annual review of information science and technology 31 (1996): 33-78.

[49] STRICKER, MARKUS A., and MARKUS ORENGO, *Similarity of color images,* IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology. International Society for Optics and Photonics, 1995.

[50] `SU, ZHONG, et al, *Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces, and progressive learning,* Image Processing, IEEE Transactions on 12.8 (2003): 924-937.

[51] SWAIN, MICHAEL J., and DANA H. BALLARD, *Color indexin,* International journal of computer vision 7.1 (1991): 11-32.

[52] TAMURA, HIDEYUKI, SHUNJI MORI, and TAKASHI YAMAWAKI, *Textural features corresponding to visual perception,* Systems, Man and Cybernetics, IEEE Transactions on 8.6 (1978): 460-473.

[53] WANG, TAO, YONG RUI, and SHI-MIN HU, *Optimal adaptive learning for image retrieva*, Computer Vision and Pattern Recognition, 2001, CVPR 2001, Proceedings of the 2001 IEEE Computer Society Conference on, Vol. 1. IEEE, 2001.

[54] YANG, MIIN-SHEN, PEI-YUAN HWANG, and DE-HUA CHEN, *Fuzzy clustering algorithms for mixed feature variables*, Fuzzy Sets and Systems 141.2 (2004): 301-317.

[55] ZHANG, DENGSHENG, et al, *Content-based image retrieval using Gabor texture features*, IEEE Pacific-Rim Conference on Multimedia, University of Sydney, Australia, 2000.

Nhận bài ngày: 28/10/2015

SƠ LƯỢC VỀ TÁC GIẢ

VŨ VĂN HIỆU

Sinh năm 1976 tại Kiến Thuỵ, Hải Phòng.

Tốt nghiệp thạc sĩ Khoa học máy tính năm 2007 tại Học viện Kỹ thuật Quân sự; Đang là nghiên cứu sinh năm thứ 3 tại Viện CNTT, Viện Hàn lâm KH&CN Việt Nam. Chuyên ngành cơ sở toán cho tin học.

Hiện công tác tại Khoa CNTT, Trường ĐH Hải Phòng. Lĩnh vực nghiên cứu: Nhận dạng, Xử lý ảnh, Trí tuệ nhân tạo, Khai phá dữ liệu..

Email: hieuvv@dhhp.edu.vn. ĐT: 0975482486

NGÔ HOÀNG HUY

Sinh năm 1969 tại Hà Nội.

Tốt nghiệp chuyên ngành Toán tin tại Trường ĐH Sư phạm Hà Nội.

Hiện đang là nghiên cứu viên Viện CNTT, Viện Hàn lâm KH&CN Việt Nam.

Email: huyngo3i@gmail.com

NGÔ QUỐC TẠO

Tốt nghiệp chuyên ngành Toán Máy tính Trường ĐH Bách Khoa Hà Nội năm 1982, tiến sỹ Toán lý năm 1997 chuyên ngành đảm báo toán học cho các hệ thống tính toán. Được phong Phó Giáo sư năm 2002. Công tác tại Viện CNTT, Viện Hàn lâm KH&CN Việt Nam.

Email: nqtao@ioit.ac.vn

NGUYỄN HỮU QUỲNH

Tốt nghiệp ĐH, Cao học và Tiến sĩ tại ĐH Quốc Gia Hà Nội vào các năm 1998, 2004 và 2010. Công tác tại Khoa CNTT, Trường ĐH Điện Lực, Hà Nội.

Email: quynhnh@epu.edu.vn