

CẢI TIẾN TRA CỨU ẢNH THÔNG QUA TỔ HỢP ĐA PHÂN HẠNG

Nguyễn Hữu Quỳnh*

Trường Đại học Điện lực

TÓM TẮT

Các kỹ thuật tra cứu ảnh sử dụng SVM hiện nay xác định siêu phẳng tách dựa trên các mẫu luyện thu được từ lân cận của ảnh truy vấn, dẫn đến các ảnh thuộc cùng một chủ đề có thể bị phân lớp vào hai lớp khác nhau. Kết quả phân lớp này làm giảm độ chính xác của các phương pháp tra cứu. Trong bài báo này, chúng tôi đề xuất một phương pháp tra cứu ảnh - Combine Multiple Ranking (CMR), có thể lấy được nhiều ảnh thuộc cùng một chủ đề với ảnh truy vấn. Phương pháp của chúng tôi xây dựng siêu phẳng tách theo từng vùng của không gian đặc trưng để phân hạng các ảnh. Chúng tôi cũng cung cấp các kết quả thực nghiệm trên cơ sở dữ liệu gồm 10,800 ảnh để chỉ ra độ chính xác của phương pháp.

Từ khóa: Tra cứu ảnh dựa vào nội dung, không gian đặc trưng, đa siêu phẳng tối ưu, máy véc tơ hỗ trợ, tổ hợp đa phân hạng.

GIỚI THIỆU

Các cơ sở dữ liệu ảnh ngày càng trở nên phổ biến trong nhiều lĩnh vực ứng dụng đã thúc đẩy sự quan tâm lớn đến lĩnh vực tra cứu ảnh dựa vào nội dung. Các hệ thống CBIR (Content-based image retrieval) truyền thống thường chấp nhận độ đo khoảng cách Euclidean trong một không gian đặc trưng trực quan mức thấp nhiều chiều để đo độ tương tự giữa ảnh truy vấn và các ảnh trong cơ sở dữ liệu [1,11]. Tuy nhiên, độ đo khoảng cách Euclidean trong một không gian nhiều chiều thường rất không hiệu quả do khoảng trống ngữ nghĩa giữa các đặc trưng trực quan mức thấp và các khái niệm ngữ nghĩa mức cao. Phân hồi liên quan [13,16] là một công cụ quan trọng để thu hẹp khoảng cách ngữ nghĩa này và do đó cải tiến hiệu năng của tra cứu ảnh dựa vào nội dung. Trong những thập kỷ qua, nhiều cách tiếp cận RF(Relevance feedback) đã được thiết kế dựa vào các giả thiết khác nhau cho các mẫu phân hồi âm và dương [13]. Máy véc tơ hỗ trợ một lớp ước lượng mật độ của các mẫu phân hồi dương nhưng bỏ qua các mẫu phân hồi âm [14]. SVM hai lớp có thể xác định cả các mẫu phân hồi âm và dương nhưng coi hai nhóm khác nhau này có trọng số ngang bằng [10].

Ngoài cách tiếp cận RF truyền thống, một số lược đồ mới đã xuất hiện để giảm khoảng cách ngữ nghĩa trong CBIR [2,5,7,15]. Chẳng hạn, các nghiên cứu về tra cứu ảnh cộng tác (CIR) chia làm hai nhóm. Nhóm nghiên cứu thứ nhất có ý định cải tiến hiệu năng của RF truyền thống bằng việc sử dụng lại dữ liệu phản hồi lịch sử người dùng hoặc dữ liệu web cỡ lớn [2,5]. Nhóm nghiên cứu thứ hai cố gắng lựa chọn một tập các mẫu có chứa thông tin nhất từ cơ sở dữ liệu ảnh [7,12,3], có thể được gán nhãn bởi người dùng trong RF và được sử dụng làm dữ liệu huấn luyện để xác định một độ đo tương tự hiệu quả cho tra cứu ảnh.

Nhìn chung, các phương pháp tra cứu ảnh sử dụng các máy véc tơ hỗ trợ (SVM) hiện nay xây dựng siêu phẳng tách dựa trên các mẫu dương và mẫu âm được chọn ra từ các điểm (ảnh) thuộc lân cận của điểm ảnh truy vấn trong không gian đặc trưng trực quan [8]. Tuy nhiên, các ảnh (điểm) thuộc cùng một chủ đề có thể không thuộc lân cận của điểm truy vấn và do đó siêu phẳng tách thu được từ các phương pháp hiện nay có thể phân lớp các ảnh (điểm) này vào lớp âm. Khả năng phân lớp các ảnh thuộc cùng chủ đề vào lớp âm của các phương pháp tra cứu ảnh sử dụng SVM hiện có sẽ dẫn đến hiệu năng tra cứu nghèo nàn.

Các hạn chế ở trên cũng là động lực để chúng tôi đề xuất phương pháp tra cứu ảnh tổ hợp đa

* Email: quynhnh@epu.edu.vn

phân hạng CMR (Combine Multiple Ranking). Thay vì xác định một siêu phẳng tách trên toàn bộ không gian đặc trưng trực quan để phân hạng các kết quả, phương pháp của chúng tôi xác định siêu phẳng tách theo mỗi vùng (siêu phẳng tách vùng) và phân hạng kết quả theo siêu phẳng tách vùng này. Kết quả cuối cùng sẽ là tổ hợp kết quả của các phân hạng theo siêu phẳng tách vùng. Bằng thực nghiệm trên cơ sở dữ liệu đặc trưng gồm 10,800 ảnh, chúng tôi sẽ chỉ ra sự chính xác của phương pháp đề xuất.

Phần còn lại của bài báo được tổ chức như sau. Trong phần 2, chúng tôi trình bày phương pháp tra cứu ảnh tổ hợp đa phân hạng. Sau đó, phân tích các kết quả thực nghiệm của chúng tôi được đưa ra trong phần 3. Cuối cùng, chúng tôi đưa ra kết luận và hướng nghiên cứu tương lai trong phần 4.

PHƯƠNG PHÁP TRA CỨU ẢNH ĐỀ XUẤT

Trong phần này, đầu tiên chúng tôi trình bày tóm tắt cơ sở lý thuyết được sử dụng trong phương pháp đề xuất. Sau đó chúng tôi trình bày chi tiết phương pháp đề xuất.

Cơ sở lý thuyết

Với tập dữ liệu huấn luyện $D = (x_i, y_i)$, $1 \leq i \leq m$, x_i là véc tơ đặc trưng (một véc tơ m chiều của các đặc trưng đại diện cho một đối tượng) trong không gian R^m , m là số chiều của không gian, $y_i \in \{-1, 1\}$ là nhãn lớp của x_i tương ứng. Giả sử có vài siêu phẳng có thể phân chia tập D sao cho tất cả các điểm cùng nhãn nằm cùng một phía siêu phẳng gọi là siêu phẳng phân tách.

$$g(x) = w \cdot x + b = 0, w \in R^m, b \in R \tag{1}$$

ở đây x là một véc tơ đầu vào, w là trọng số véc tơ và b là độ lệch. SVM cố gắng đi tìm hai tham số véc tơ trọng số w và độ lệch b sao cho tất cả các điểm dữ liệu (x_i, y_i) đều thỏa mãn:

$$y_i(w \cdot x_i + b) \geq 1 \tag{2}$$

Các điểm véc tơ hỗ trợ là những điểm có khoảng cách gần nhất tới siêu phẳng, bằng $1/\|w\|$, đại lượng $2/\|w\|$ được gọi là lề. Như vậy việc tìm siêu phẳng tối ưu có lẽ cực đại tức là đi tìm siêu phẳng có giá trị $\|w\|^2$ nhỏ nhất thỏa mãn (2)

$$\min h(w) = \frac{1}{2} \|w\|^2 \tag{3}$$

thỏa mãn $y_i(w \cdot x_i + b) \geq 1, i = 1, \dots, n$

Với $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ là các nhân tử Lagrange khác không, hàm Lagrangian tổng quát của bài toán tối là hàm

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i (y_i(w \cdot x_i + b) - 1) \tag{4}$$

$$\text{với } w = \sum_{i=1}^n \alpha_i y_i x_i, \quad \sum_{i=1}^n \alpha_i y_i = 0 \tag{5}$$

Thay thế các quan hệ này vào $L(w, b, \alpha)$ chúng ta có hàm mục tiêu

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j \tag{6}$$

$$\text{thỏa mãn } \sum_{i=1}^n \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, \dots, n$$

Hàm quyết định cho phép phân lớp một mẫu z được cho bởi công thức:

$$f(z) = \text{sign}(\sum_{i=1}^n \alpha_i y_i x_i \cdot z + b) \tag{7}$$

Khi dữ liệu là phân tách phi tuyến tính, dữ liệu đầu vào sẽ được ánh xạ sang một không gian mới nhiều chiều hơn bằng hàm φ thông qua hàm nhân Gaussian [4]:

$$K(x, y) = \exp(-\|x - y\|^2) \tag{8}$$

Hàm quyết định được biểu diễn thành:

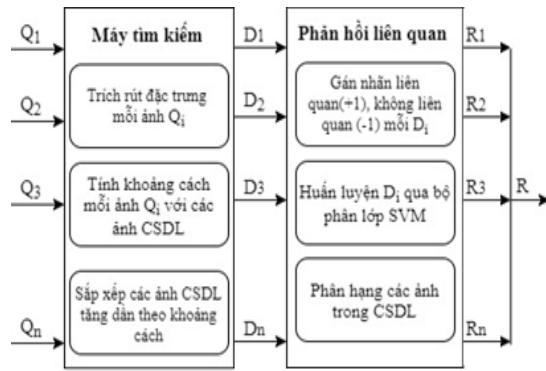
$$f(z) = \text{sign}(\sum_{i=1}^s \alpha_i y_i K(x_i, z) + b) \tag{9}$$

Trong đó $f(z)$ là đầu ra của hàm siêu phẳng quyết định của SVM và s là số lượng các điểm véc tơ hỗ trợ.

Phương pháp đề xuất

Đối với một nhiệm vụ tra cứu, người dùng có thể chọn các truy vấn khác nhau, những ảnh

này đều liên quan đến hình ảnh mà người dùng mong muốn. Những hình ảnh liên quan nhất sẽ là những hình ảnh biểu thị rõ ràng thông tin mong muốn của người dùng, còn những hình ảnh liên quan ít sẽ bổ sung thêm thông tin mong muốn của người dùng.



Hình 1. Mô hình hệ thống tra cứu ảnh đề xuất

Hệ thống tra cứu ảnh sử dụng đa truy vấn sử dụng một tập $Q = \{Q_1, Q_2, \dots, Q_m\}$, gồm m ảnh truy vấn. Với mỗi ảnh truy vấn Q_i đến Q_m sẽ được đưa vào làm ảnh truy vấn riêng biệt thu được tập D_i , $i = 1 \dots m$ gồm N_{Top} ảnh trả về tương ứng với mỗi ảnh truy vấn Q_i . Mỗi tập D_i người dùng lựa chọn gán nhãn các ảnh liên quan và không liên quan theo nhu cầu mong muốn của người dùng để thực hiện pha phân hồi liên quan. Xem xét D_i như là một tập mẫu huấn luyện cho bộ phân lớp SVM và xây dựng một bộ phân lớp phù hợp để thể hiện các khái niệm của truy vấn. Các truy vấn này sẽ tạo ra các kết quả tra cứu khác nhau bởi cùng một hệ thống tra cứu ảnh. Những danh sách kết quả khác nhau này trong bài báo này được gọi là “view – khung nhìn” có thể được xem như là những kênh, góc nhìn khác nhau, biểu diễn ngữ nghĩa khác nhau trong tâm trí người dùng. Điều này là rất có giá trị cho thấy việc kết hợp các kết quả tra cứu của đa truy vấn có tiềm năng đạt được độ chính xác cao. Kết quả tra cứu trong những khung nhìn này có khả năng được kết hợp thành một khung nhìn tổng quát hơn

Hình 2 dưới đây là thuật toán tra cứu trên một khung nhìn của người dùng **IRSV – Image Retrieval in Single View**.

Thuật toán **IRSV** trên Hình 2 thực hiện như sau:

Thuật toán IRSV (Q, N_{Top}, DB);

Input:

- Q Ảnh truy vấn
- N_{Top} Số lượng ảnh
- DB Tập các ảnh toàn bộ cơ sở dữ liệu

Output:

- R Tập gồm N_{Top} ảnh trả về.

$D \leftarrow \text{BasicIR}(Q, N_{Top}, DB)$;
 $D^+ \leftarrow \text{MarkRelevance}(D)$;
 $D^+_{label} \leftarrow \text{SetLabel}(D^+, 1)$;
 $D^- \leftarrow \text{MarkIrrelevance}(D)$;
 $D^-_{label} \leftarrow \text{SetLabel}(D^-, -1)$;
 $X \leftarrow D^+ \cup D^-$; $X_{label} \leftarrow D^+_{label} \cup D^-_{label}$;
 $\langle \alpha, b, SV, SV^{label} \rangle \leftarrow \text{Classifier}(X, X_{label})$;
for $k \leftarrow 1$ **to** $DB.Count$ **do**
 $\psi(DB_k) \leftarrow \sum_{j=1}^{SV.Count} \alpha_j SV^{label}_j SV_j * DB_k + b$
 $R \leftarrow \text{Sort}(DB, \psi, N_{Top})$;
Return R ;

Hình 2. Thuật toán tra cứu ảnh trên một khung nhìn IRSV

Với đầu vào là một ảnh truy vấn Q thuật toán sẽ lấy được tập ảnh D gồm N_{Top} ảnh đầu tiên được phân hạng đầu tiên trong toàn bộ cơ sở dữ liệu DB thông qua **BasicIR()**. Sau đó người dùng dựa vào nhu cầu thông tin mong muốn của bản thân gán nhãn liên quan +1 được tập D^+ , gán nhãn không liên quan -1 được tập D^- thông qua các hàm **MarkRelevance**, **MarkIrrelevance** và **SetLabel**. Gộp cả hai tập liên quan và không liên quan này thu được tập huấn luyện X cùng nhãn tương ứng X_{label} của chúng đưa vào tiến hành huấn luyện phân lớp sử dụng hàm **Classifier()** dùng SVM được mô tả ở phần đầu nội dung 2. Dựa vào các ảnh véc tơ hỗ trợ SV cùng các tham số α, b thuật toán duyệt từng ảnh trong toàn bộ cơ sở dữ liệu DB để tính khoảng cách ψ từng ảnh tới siêu phẳng phân tách. Sắp xếp các ảnh trong cơ sở dữ liệu theo thứ tự giảm dần theo khoảng cách ψ vừa tính được, lấy ra tập ảnh R gồm N_{Top} ảnh đầu tiên sau khi sắp xếp.

Thuật toán tra cứu ảnh CMR, cho phép thu được tập ảnh kết quả kết hợp từ những vùng không gian ngữ nghĩa khác nhau thông qua

việc tận dụng hiệu quả phản hồi thông tin của người dùng dựa trên mỗi khung nhìn được mô tả như hình 3 dưới đây:

```

Thuật toán CMR ( $Q, N_{Top}, DB$ );
Input:
-  $Q$  Tập ảnh truy vấn
-  $N_{Top}$  Số lượng ảnh
-  $DB$  Tập các ảnh toàn bộ cơ sở dữ liệu
Output:
-  $R$  Tập gồm  $N_{Top}$  ảnh trả về.
for  $i \leftarrow 1$  to  $Q.Count$  do
   $R_i \leftarrow \mathbf{IRSV}(Q_i, N_{Top}, DB)$ ;
 $IR \leftarrow \emptyset$ ;
for  $i \leftarrow 1$  to  $N_{Top}$  do
  for  $j \leftarrow 1$  to  $Q.Count$  do
    if( $\mathbf{isMember}(R_{ij}, IR) == 0$ )
       $IR \leftarrow IR \cup R_{ij}$ ;
 $R \leftarrow \mathbf{getTop}(IR, N_{Top})$ ;
Return  $R$ ;

```

Hình 3. Thuật toán CMR

Thuật toán CMR trên Hình 3 được thực hiện như sau:

Khi người dùng gửi vào một tập các ảnh truy vấn Q gồm Q_1, Q_2, \dots, Q_m , thuật toán sẽ đưa từng ảnh truy vấn Q_i thông qua thuật toán **IRSV** (xem thuật toán **IRSV** trên hình 2) thu được m tập R gồm N_{Top} ảnh tương ứng. Sau đó thuật toán sẽ kết hợp m tập ảnh R này lại thành một tập ảnh duy nhất IR bằng cách duyệt từng ảnh có trong mỗi tập R sẽ được thêm vào tập IR cuối cùng nếu như tập IR chưa chứa ảnh đó. Cuối cùng thuật toán thu được tập ảnh R trả về cuối cùng gồm N_{Top} ảnh phân hạng đầu tiên thông qua hàm **getTop()**

THỰC NGHIỆM

Trong thực nghiệm, chúng tôi sử dụng một tập con của Corel Photo Gallery làm cơ sở dữ liệu thử nghiệm để đánh giá hiệu quả của phương pháp. Tập này gồm 80 loại, ví dụ như là: mùa thu, cây cảnh, lâu đài, đám mây, chó, voi, hổ, tàu hỏa, thác nước,.... Tất cả các ảnh trong tập ảnh này có tính chất là đều chứa đối tượng tiền cảnh nổi bật. Đa số nhóm đều gồm 100 ảnh, có một vài nhóm có hơn 100 hình

ảnh. Cỡ của các ảnh có $\max(\text{chiều rộng, chiều cao})=120$ và $\min(\text{chiều rộng, chiều cao})=80$.

Chúng tôi thiết kế một kịch bản mô hình hóa quá trình tra cứu ảnh để đánh giá độ chính xác¹ của phương pháp đề xuất. Trong các hệ thống tra cứu ảnh thế giới thực, một ảnh truy vấn thường không tồn tại trong cơ sở dữ liệu. Để mô phỏng điều này, chúng tôi chia tập ảnh cơ sở dữ liệu ảnh ban đầu thành 5 tập ảnh con bằng nhau và chứa 20% của mỗi loại. Tại mỗi lần chạy kiểm chứng chéo, một tập con sẽ được đưa lần lượt vào làm tập truy vấn và bốn tập ảnh còn lại được sử dụng như là tập ảnh cơ sở dữ liệu cho việc tra cứu ảnh. Sau đó 500 ảnh được lấy ngẫu nhiên trong tập ảnh truy vấn được tra cứu. Trong thực nghiệm, phương pháp CMR thực hiện mỗi lần tra cứu sử dụng một tập gồm ba ảnh bằng cách với mỗi ảnh trong 500 ảnh sẽ lấy ngẫu nhiên thêm hai ảnh truy vấn cùng loại trong tập ảnh truy vấn.

Chúng tôi tính độ chính xác trung bình kiểm chứng chéo 5 tập con với tập ảnh kết quả gồm 60 ảnh được phân hạng đầu tiên (xem phụ lục A). Để thấy được sự hiệu quả về độ chính xác của phương pháp đề xuất CMR (sử dụng SVM trong thư viện matlab bản R2016a), chúng tôi so sánh CMR với một số phương pháp khác như DSSA [9], GOED [7], và GBDA [6] trên cùng một tập cơ sở dữ liệu ảnh Corel. GBDA (Generalized Biased Discriminant Analysis) sử dụng cách tiếp cận phản hồi liên quan với phân tích phân biệt để khắc phục mất cân bằng mẫu phản hồi trong tra cứu ảnh. DSSA (Discriminative Semantic Subspace Analysis) học trực tiếp một không gian con ngữ nghĩa từ các ràng buộc cặp tương tự và không tương tự mà không sử dụng thông tin nhãn lớp để nâng cao hiệu độ chính xác tra cứu ảnh. GOED (geometric optimum experimental design) tận dụng cấu

¹Độ chính xác là tỉ số giữa số các ảnh liên quan với ảnh truy vấn trong tập kết quả trả về trên tổng số các ảnh trả về.

trúc hình học của các mẫu chưa có nhãn để giảm nhẹ vấn đề cỡ nhỏ của tập huấn luyện trong SVM.

Phương pháp CMR mà chúng tôi đề xuất xây dựng siêu phẳng tách theo từng vùng của không gian đặc trưng để phân hạng các ảnh. Do đó CRM có thể lấy được nhiều ảnh thuộc cùng một chủ đề với ảnh truy vấn trên từng vùng không gian đặc trưng. Các kết quả độ chính xác được thể hiện bằng số liệu trong Bảng 1.

Bảng 1. Bảng kết quả của 4 phương pháp

Phương pháp	GBDA	GOED	DSSA	CMR
Độ chính xác (%)	29.3	33	35	46.32

KẾT LUẬN

Chúng tôi đã đề xuất phương pháp tra cứu ảnh dựa vào nội dung cho cải tiến độ chính xác tra cứu của các hệ thống tra cứu phản hồi liên quan sử dụng SVM truyền thống. Phương pháp của chúng tôi quan tâm đến việc thu được các ảnh liên quan ngữ nghĩa với mong muốn người dùng ở những vùng khác nhau trong cơ sở dữ liệu. Để tạo ra vùng có chứa các ảnh liên quan ngữ nghĩa khác nhau, chúng tôi tận dụng thông tin phản hồi của người dùng trên các tập ảnh kết quả trả về của nhiều truy vấn khác nhau.

Kết quả thực nghiệm của chúng tôi trên cơ sở dữ liệu đặc trưng gồm 10800 ảnh đã chỉ ra rằng phương pháp được đề xuất CMR cung cấp một độ chính xác cao hơn hẳn so với các phương pháp DSSA, GOED và GBDA.

TÀI LIỆU THAM KHẢO

1. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain (2000), "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no.12, pp. 1349–1380.
2. C. H. Hoi, M. R. Lyu, and R. Jin (2006), "A unified log-based relevance feedback scheme for image retrieval," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 4, pp. 509–524.
3. C. H. Hoi, R. Jin, J. Zhu, and M. R. Lyu (2009), "Semisupervised svm batch mode active learning

with applications to image retrieval," *ACM Transactions on Information System*, vol. 27, no. 3, pp. 16:1–16:29.

4. J.C. Burges (1998), "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167.

5. L. Zhang, L. Wang, and W. Lin (2012), "Conjunctive patches subspace learning with side information for collaborative image retrieval" *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3707–3720.

6. L. Zhang, L. Wang, and W. Lin (2012), "Generalized biased discriminant analysis for content-based image retrieval," *IEEE Transactions on Cybernetics*, vol. 42, no. 1, pp.282–290.

7. L. Zhang, L. Wang, W. Lin, S. Yan (2014), "Geometric optimum experimental design for collaborative image retrieval," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 24, no. 2, pp. 346–359.

8. Lei. Zhang, F. Liu, B. Zhang (2001), Support vector machine learning for image retrieval. In: International Conference on Imag. Processing, pp. 7–10.

9. Lining Zhang, Hubert P. H. Shum and Ling Shao (2016), "Discriminative Semantic Subspace Analysis for Relevance Feedback," *IEEE Transaction on Image Processing*, vol. 25, no. 3.

10. P. Hong, Q. Tian, and T. S. Huang (2000), "Incorporate support vector machines to content-based image retrieval with relevance feedback," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 750–753.

11. R. Datta, D. Joshi, J. Li, and J. Z. Wang (2008), "Image retrieval: ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, no. 2, pp. 1–60.

12. S. Tong and E. Chang (2001), "Support vector machine active learning for image retrieval," in *Proceedings of the 9th ACM International Conference on Multimedia*, pp. 107–118.

13. X. S. Zhou and T. S. Huang (2003), "Relevance feedback in image retrieval: A comprehensive review," *Multimedia Systems*, vol. 8, no. 6, pp. 536–544.

14. Y. Chen, X. S. Zhou, and T.S. Huang (2001), "One-class svm for learning in image retrieval," in *Proceedings of IEEE International Conference on Image Processing*, pp. 34–37.

15. Y. Liu, D. Xu, I.W. Tsang, and J. Luo (2011), "Textual query of personal photos facilitated by large-scale web data," *IEEE Transactions on*

Pattern Analysis and Machine Intelligence, vol. 33, pp. 1022–1036.

16. Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra (1998), “Relevance feedback: a power

tool for interactive content-based image retrieval,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no.5, pp. 644 –655.

ABSTRACT

IMPROVING IMAGE RETRIEVAL ACCURACY BY COMBINING MULTIPLE RANKING

Nguyen Huu Quynh*

University of Electricity

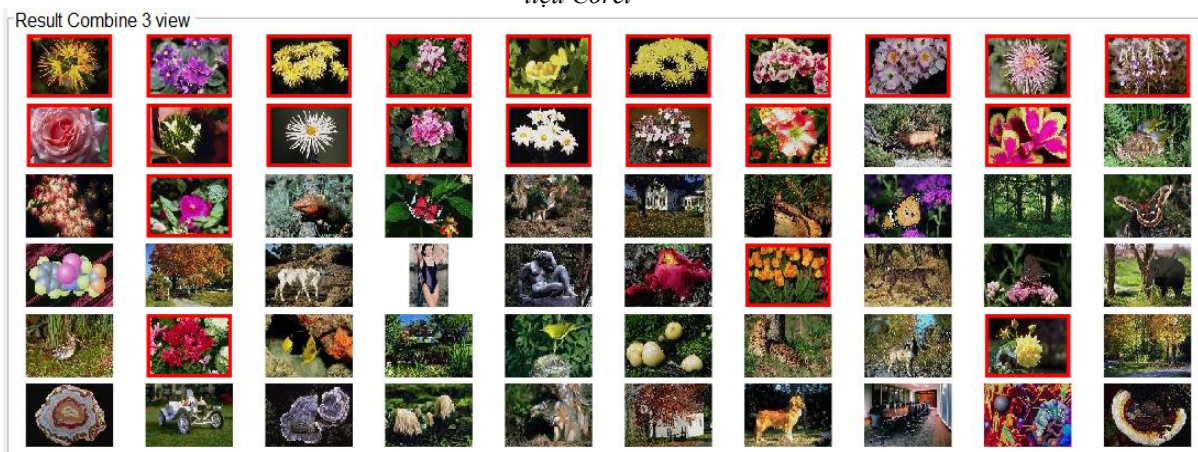
Exist image retrieval techniques using SVM define the optimal hyperplane based on samples obtained from neighboring images of query image, resulting in images of the same topic that can be classed into two different classes. This classification result reduces the accuracy of the retrieval methods. In this paper, we propose a image retrieval method - Combine Multiple Ranking (CMR), which can capture many images of the same topic with the query image. Our method of building the optimal hyperplane by each area of the feature space in order to rank images. We also provided empirical results on a database of 10,800 images to show the accuracy of the method.

Keywords: Content-based image retrieval, feature space, multiple optimal hyper plane, support vector machine (SVM), combine Multiple Ranking

PHỤ LỤC A



Hình A1. Tập ảnh truy vấn gồm các ID là 124050, 124053, 476089 thuộc loại *pl_flower* trong cơ sở dữ liệu Corel



Hình A2. Kết quả 60 ảnh được phân hạng đầu tiên của tập ảnh truy vấn gồm các ID là 124050, 124053, 476089 thuộc loại *pl_flower*³. Độ chính xác là 36.67%(Ảnh viền đỏ là biểu thị các ảnh trong cơ sở dữ liệu thuộc cùng loại *pl_flower* ảnh truy vấn.)

Ngày nhận bài: 12/9/2017; Ngày phản biện: 17/10/2017; Ngày duyệt đăng: 30/11/2017

* Email: quynhnh@epu.edu.vn