

MỘT PHƯƠNG PHÁP TRA CỨU DỮ LIỆU ẢNH DỰA TRÊN CÂY PHÂN CỤM ĐA NHÁNH CÂN BẰNG

Nguyễn Phương Hạc*, Văn Thế Thành

Trường Đại học Công nghiệp Thực phẩm TP.HCM

*Email: hacnp@hufi.edu.vn

Ngày nhận bài: 16/01/2019; Ngày chấp nhận đăng: 06/3/2019

TÓM TẮT

Bài báo tiếp cận phương pháp tra cứu ảnh tương tự dựa trên cây đa nhánh cân bằng lưu trữ các dữ liệu đặc trưng của hình ảnh dưới dạng chỉ mục nhằm mô tả hình dạng của đối tượng đặc trưng. Để trích xuất đặc trưng của hình ảnh, nhóm nghiên cứu đã thực hiện phân đoạn ảnh dựa trên đặc trưng thị giác cấp thấp gồm màu sắc và cấu trúc của hình ảnh. Các đặc trưng thị giác này được trích xuất trên mỗi khối của hình ảnh bằng phép biến đổi Wavelet và không gian màu CIE $L^*a^*b^*$. Tiếp theo, nhóm nghiên cứu thực hiện gom cụm các điểm ảnh để tạo thành vùng liên thông trên ảnh đồng thời loại bỏ các vùng có diện tích không vượt ngưỡng cho trước. Trên cơ sở các đặc trưng đã được trích xuất, nhóm nghiên cứu đã tạo ra độ tương phản của hình ảnh nhằm rút trích màu nền và màu chính của mỗi hình ảnh. Dựa trên ảnh đã được phân đoạn, chỉ mục đặc trưng được tạo ra nhằm thực hiện đối sánh các hình ảnh tương tự. Từ đó, nhóm nghiên cứu tạo ra cây đa nhánh cân bằng nhằm lưu trữ chỉ mục mô tả đặc trưng hình ảnh và thực hiện tìm kiếm ảnh. Để minh chứng tính đúng đắn cho phương pháp đề xuất, nhóm nghiên cứu xây dựng thực nghiệm và đánh giá kết quả trên các tập dữ liệu ảnh gồm Corel và CBIRimages. Kết quả thực nghiệm được so sánh với các công trình khác và cho thấy tính hiệu quả của phương pháp đề xuất của nhóm.

Từ khóa: Phương pháp phân đoạn, đặc trưng, chỉ mục hình ảnh.

1. GIỚI THIỆU

Sự xuất hiện các bộ dữ liệu đa phương tiện lớn đã dẫn đến yêu cầu phát triển các công cụ truy vấn thông tin thị giác. Vì vậy, việc tìm kiếm các hình ảnh liên quan đến yêu cầu người dùng là bài toán phù hợp với nhu cầu xã hội hiện đại. Tra cứu hình ảnh từ cơ sở dữ liệu ảnh là một bài toán quan trọng trong lĩnh vực thị giác máy tính và xử lý ảnh [1].

Hệ truy vấn ảnh bao gồm đối sánh đặc trưng để tìm ra ứng viên sau đó truy hồi để tìm ra các hình ảnh tương tự [2]. Hệ truy vấn ảnh dựa trên nội dung CBIR (*content-based image retrieval*) đã được giới thiệu vào khoảng thập niên 1980; trong hệ truy vấn này thực hiện các kỹ thuật tìm kiếm một tập hình ảnh liên quan từ cơ sở dữ liệu ảnh dựa trên trích xuất tự động đặc trưng nội dung hình ảnh [3].

Kiến trúc hệ CBIR gồm 2 phần: (1) Trích xuất đặc trưng thị giác để tạo chỉ mục cho hình ảnh; (2) Thực thi việc truy vấn ảnh tương tự dựa trên chỉ mục. Kết quả truy vấn ảnh sẽ trả về k -ảnh tương tự nhất theo một độ đo cho trước. Vì vậy, trong hệ truy vấn ảnh dựa trên nội dung là sự kết hợp của nhiều lĩnh vực như: xử lý ảnh, thị giác máy tính, nhận dạng, xử lý dữ liệu, truy hồi thông tin [2, 4].

Bước đầu tiên của bài toán truy vấn ảnh dựa trên nội dung sẽ thực hiện tiền xử lý ảnh như: phân tích màu, phân đoạn ảnh, lọc ảnh, khử nhiễu ảnh,... Với một hình ảnh đã được xử

lý, thực hiện trích lọc các đặc trưng thị giác để tạo ra chỉ mục mô tả hình ảnh [1, 2, 5], chữ ký này được dùng để phân lớp tự động và phân loại ngữ nghĩa theo nội dung thị giác của hình ảnh.

Vì vậy, trích xuất vùng đặc trưng là một tiền đề quan trọng để tìm ra các hình ảnh tương tự trong hệ truy vấn ảnh dựa trên nội dung CBIR. Có nhiều phương pháp trích xuất vùng đặc trưng đã được đề cập như trích xuất đối tượng trung tâm của hình ảnh, trích xuất thuộc tính đối tượng đặc trưng [6, 7].

Bài báo tiếp cận phương pháp phân đoạn hình ảnh để tạo chỉ mục hình ảnh và cây đa nhánh cân bằng trên không gian màu CIE $L^*a^*b^*$ và phép biến đổi Wavelet. Từ đó, nhóm nghiên cứu xây dựng phương pháp tra cứu ảnh tương tự dựa trên cấu trúc dữ liệu cây đa nhánh cân bằng này. Để giải quyết vấn đề đã đặt ra, nội dung tiếp theo của bài báo gồm những phần sau: Phần 2 đề cập đến các công trình liên quan để minh chứng tính khả thi và sự cải tiến của phương pháp đề nghị; Phần 3 đề cập các kiến thức cơ sở để làm nền tảng áp dụng cho phương pháp; Phần 4 trình bày phương pháp phân đoạn hình ảnh để trích xuất đối tượng đặc trưng làm cơ sở tạo chỉ mục hình ảnh; Phần 5 thực hiện truy vấn ảnh bao gồm việc tạo cây đa nhánh cân bằng và mô tả thực nghiệm; Phần 6 đưa ra kết luận và hướng phát triển tiếp theo của bài báo.

2. CÁC CÔNG TRÌNH LIÊN QUAN

Nhiều ứng dụng liên quan đến truy vấn ảnh đã được phát triển và ứng dụng trong nhiều lĩnh vực khác nhau như ứng dụng trong thư viện số: CIRES, C-BIRD, PhotoFile, iMATCH [5]. Ứng dụng truy vấn ảnh IRMA trong y học dựa trên SVM, hệ truy vấn ảnh y khoa CBMIR (Content-Based Medical Image Retrieval) trên ảnh CT, hệ truy vấn ảnh y khoa dựa trên phép biến đổi Wavelet, ứng dụng truy vấn ảnh trên hệ thống GIS (Geographic Information System) [8-11].

Một số công trình liên quan đến truy vấn nội dung ảnh đã công bố gần đây như: trích xuất đối tượng trên hình ảnh dựa trên sự biến đổi giá trị histogram [12], truy vấn ảnh tương tự dựa trên đối sánh các vùng đặc trưng và mối quan hệ tương tự của các vùng đặc trưng trên ảnh [13], truy vấn ảnh màu dựa trên dò tìm các vùng đặc trưng cục bộ bằng phương pháp Harris-Laplace [12], truy vấn ảnh màu dựa trên mặt phẳng bit và không gian màu $L^*a^*b^*$ [14], chuyển đổi không gian màu và xây dựng hàm băm nhằm truy vấn nội dung các ảnh màu [15],...

Có nhiều phương pháp dò tìm đặc trưng đã được giới thiệu, gồm: phương pháp dò góc và cạnh giới thiệu vào năm 1998 bởi Harris & M.Stephens, phương pháp dò tìm đặc trưng SIFT (*Scale Invariant Features Transform*) dựa trên phép lọc của mặt nạ tích chập giữa hình ảnh và đạo hàm DoG (*Difference of Gaussians*) để xấp xỉ toán tử Laplacian của hàm Gaussian được D.Lowe giới thiệu năm 2003, phương pháp dò tìm đặc trưng SURF (*Speeded Up Robust Features*) được Bay và cộng sự giới thiệu vào năm 2006, phương pháp dò điểm đặc trưng Harris Laplacian dựa trên toán tử Laplacian của hàm Gaussian được giới thiệu năm 2001 bởi Mikolajczyk & C.Schmid [16].

Năm 1973, Haralick *et al.* đã giới thiệu ma trận đồng xuất hiện (*co-occurrence matrix*) mô tả đặc tính cấu trúc [5]. Trong phương pháp này, ma trận đồng xuất hiện được xây dựng dựa trên hướng và khoảng cách giữa các điểm ảnh. Khi đó, đặc tính cấu trúc được trích xuất từ ma trận đồng xuất hiện bằng sự xuất hiện có tính lặp lại ứng với các mức xám.

Acharya *et al.* đề xuất việc tính toán xấp xỉ đối với đặc tính cấu trúc dựa trên các nghiên cứu tâm lý trong nhận thức thị giác về cấu trúc. Các thuộc tính cấu trúc này cho phân tích theo ngữ nghĩa trực quan và được sử dụng trong nhiều hệ truy vấn ảnh dựa trên nội dung. Phép biến đổi Wavelet đã được ứng dụng trong phân tích cấu trúc và phân lớp các hình ảnh

dựa trên phép phân tách đa phân giải của các hình ảnh và mô tả cấu trúc trên các tỷ lệ khác nhau [1].

Cũng theo Acharya *et al.*, năm 1962 Human đã xác định bảy moment chuẩn làm đặc trưng hình dạng và cũng bất biến đối với phép tỷ lệ [1].

Kumar *et al.* (2009) đã đề xuất phương pháp phân đoạn ảnh tự động dựa trên phép biến đổi Wavelet nhằm tạo ra phân đoạn nhanh và đơn giản. Trong bài báo đã cho thấy phương pháp hiệu quả tốt trên việc phân đoạn các hình ảnh lớn và thực thi dễ dàng hơn các phương pháp khác [17].

Tuy nhiên, nếu truy tìm hình ảnh dựa trên đối sánh trực tiếp nội dung của hình ảnh sẽ mất nhiều chi phí về thời gian cũng như không gian truy vấn. Do đó, cần có phương pháp mô tả nội dung hình ảnh dưới dạng dữ liệu mô tả (*metadata*) để từ đó thực hiện truy vấn hình ảnh tương tự qua dữ liệu mô tả này. Đối với nghiên cứu về ứng dụng chữ ký nhị phân tạo chỉ mục cho hình ảnh, Yannis Manolopoulos đã mô tả chữ ký nhị phân của hình ảnh và thực hiện phân cụm hình ảnh dựa trên cây S-Tree [18]. Trong thực nghiệm đã cho thấy tính hiệu quả khi áp dụng chữ ký nhị phân đối với dữ liệu hình ảnh.

Nascimento và Chitkara đã tiếp cận kỹ thuật truy vấn ảnh dựa trên chỉ mục nhị phân. Thực nghiệm đã cho thấy tính hiệu quả khi truy vấn trên các cơ sở dữ liệu ảnh lớn [19].

Năm 2013, Timothy Chappell và Shlomo Geva đã tiếp cận tìm kiếm ảnh tương tự dựa trên chỉ mục nhị phân, trong công trình đã đưa ra tính hiệu quả và gia tăng tốc độ truy vấn hình ảnh ứng dụng độ đo Hamming cho chữ ký nhị phân [20].

Gần đây, các công trình về truy vấn ảnh dựa trên chỉ mục nhị phân như: truy vấn ảnh dựa trên chỉ mục và cây chữ ký, truy vấn ảnh dựa trên độ đo EMD và cây S-Tree, truy vấn ảnh dựa trên chữ ký nhị phân, truy vấn ảnh dựa trên đồ thị chữ ký [21-25].

Việc truy vấn ảnh dựa trên chỉ mục mô tả là một hướng nghiên cứu khả thi và có tính hiệu quả. Do đó, bài báo tiếp cận phương pháp tạo ra chỉ mục cho hình ảnh dựa trên các đặc trưng cấp thấp như hình dạng, cấu trúc, màu sắc, vị trí của đối tượng đặc trưng. Trên cơ sở này, bài báo thực hiện tạo cây đa nhánh cân bằng và truy vấn nhanh các hình ảnh tương tự.

3. CÁC LÝ THUYẾT CƠ SỞ

3.1. Các định nghĩa cơ sở

3.1.1. Kỳ vọng của các giá trị rời rạc

Để trích xuất cấu trúc (*texture*) của các tín hiệu rời rạc, bước đầu tiên cần tìm kỳ vọng (bởi vì giá trị này phản ánh đặc trưng cho xu hướng trung tâm đối với các giá trị rời rạc [26, 27]).

Cho n tín hiệu được mô tả bởi vector $v = (v_1, v_2, \dots, v_n)$, theo tài liệu [26, 27] giá trị kỳ

vọng được định nghĩa:
$$\mu = \frac{1}{n} \sum_{i=1}^n v_i$$

3.1.2. Phương sai của các giá trị rời rạc

Phương sai phản ánh mức độ phân tán của các giá trị biến ngẫu nhiên xung quanh giá trị kỳ vọng μ . Từ đó, làm cơ sở trích xuất đặc trưng cấu trúc của các tín hiệu hình ảnh. Theo

tài liệu [26, 27], phương sai được định nghĩa:
$$\rho = \frac{1}{n-1} \sum_{i=1}^n |v_i - \mu|^2$$

3.1.3. Độ lệch chuẩn của các giá trị rời rạc

Khi đánh giá mức độ phân tán theo đơn vị đo của các tín hiệu ban đầu, cần tính độ lệch chuẩn. Từ mức độ phân tán này sẽ trích xuất giá trị cấu trúc của tín hiệu. Theo tài liệu [26, 27], độ lệch chuẩn được định nghĩa là: $\sigma = \sqrt{\rho}$.

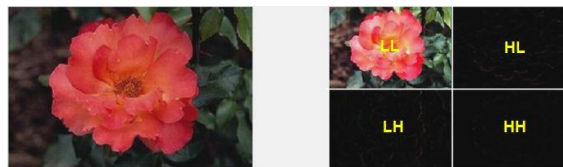
3.2. Phép biến đổi Wavelet

Phép biến đổi wavelet tạo ra sự phân tích đa tỷ lệ làm cho tín hiệu được phân rã và phân tích chi tiết hơn [28]. Do đó, phép biến đổi wavelet sẽ phân tích tín hiệu thành tổng các tín hiệu đồng dạng có tỷ lệ thời gian trễ khác nhau.

Vì vậy, các thành phần chi tiết mô tả cấu trúc của tín hiệu khi hàm wavelet được chọn đồng dạng với tín hiệu. Các hàm wavelet trực giao thường được sử dụng cho phép biến đổi wavelet rời rạc và rất tiện dụng cho việc tái tạo lại tín hiệu ban đầu sau quá trình nén dữ liệu [17].

Khi thực hiện phép biến đổi wavelet, mỗi một tín hiệu được phân tích thành hai thành phần: thành phần xấp xỉ tương ứng với thành phần tần số thấp và thành phần chi tiết tương ứng với thành phần tần số cao, thông qua 2 bộ lọc thông thấp và thông cao. Trong đó, bộ lọc thông cao sử dụng hàm wavelet $\psi(x)$ và bộ lọc thông thấp sử dụng hàm tỷ lệ $\Phi(x)$ [28].

Khi thực hiện phép biến đổi wavelet rời rạc cho một hình ảnh sẽ có được bốn dải tần con gồm LL, LH, HL, HH. Thành phần LL là phiên bản thô của hình ảnh gốc; các thành phần còn lại LH, HL, HH ứng với dải tần số cao chứa thành phần thông tin chi tiết của hình ảnh [1].



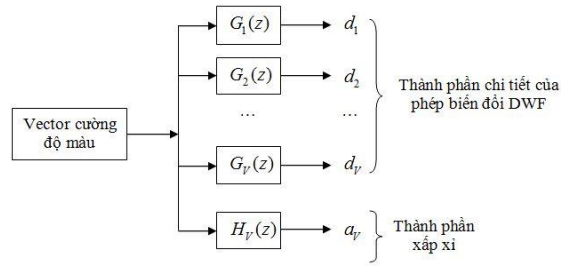
Hình 1. Minh họa phép biến đổi DFT

3.3. Phép biến đổi DWF

Để nhận diện và đưa ra các đặc tính cấu trúc của các điểm ảnh láng giềng cần sử dụng phép biến đổi DWF (*Discrete Wavelet Frames*) [28]. Đây là phương pháp tương tự với phép biến đổi wavelet rời rạc DWT (*Discrete Wavelet Transform*) dùng để biến đổi cường độ ảnh thành các dải tần con. Sự khác biệt chính của 2 phương pháp này đó là DWF sẽ đưa ra bộ lọc không có mẫu con.

Phép DWF thực thi trên băng tần lọc dựa trên phép lọc thông thấp $H(z)$ để phân giải mỗi thành phần cường độ của ảnh thành một tập các băng tần con. Độ lệch chuẩn của tất cả các thành phần chi tiết được tính trong láng giềng của điểm ảnh p để làm đặc trưng cấu trúc. Phép lọc thông cao $G(z) = zH(-z^{-1})$ được định nghĩa qua phép lọc thông thấp $H(z)$. Các phép lọc của dãy băng tần lọc $H_V(z)$, $G_i(z)$, $i = 1, \dots, V$ được tạo ra từ $H(z)$, $G(z)$ sao cho: $H_{k+1}(z) = H(z^{2^k})H_k(z)$, $G_{k+1}(z) = G(z^{2^k})H_k(z)$ với $k = 0, \dots, V - 1$, $H_0(z) = 1$.

Trong thực nghiệm, phép lọc thông thấp sử dụng phép biến đổi Haar Wavelet, tức là $H(z) = \frac{1}{2}(1 + z^{-1})$ với điều kiện thông thấp là $H(z)|_{z=1} = 1$. Khi đó, phép lọc thông cao được định nghĩa là $G(z) = zH(-z^{-1})$. Vector cấu trúc của điểm ảnh p được mô tả bằng độ lệch chuẩn của tất cả thành phần chi tiết và tính toán trên hình vuông láng giềng Φ của pixel p . Khi đó, vector đặc tính cấu trúc của điểm ảnh p sẽ là: $T(p) = [\sigma_1(p), \sigma_2(p), \dots, \sigma_{9 \times 9}(p)]$.



Hình 2. Mô hình phép biến đổi DFW

Ví dụ: cho cường độ hình chữ nhật láng giềng của điểm ảnh p như sau:

5	4	2
2	6	3
3	4	2

Vector cấu trúc $T(p)$ với $V = 2$ ứng với phép biến đổi Haar-Wavelet như sau:

$T(p) = (2.50, 1.25, 2.00, 1.00, 1.00, 0.50, 1.00, 0.50, 3.00, 1.50, 1.50, 0.75, 1.50, 0.75, 2.00, 1.00, 1.00, 0.50)$.

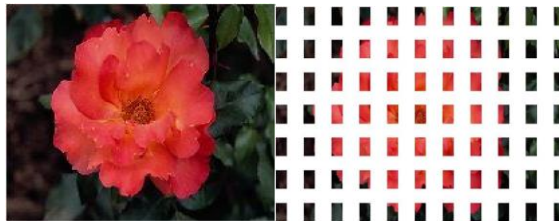
4. PHÂN ĐOẠN ẢNH

Để sử dụng hình dạng như là một đặc trưng của hình ảnh, bước cơ bản là phân đoạn hình ảnh để tìm đối tượng. Trong phương pháp này, sẽ gom cụm các điểm ảnh thuộc về các vùng liên thông dựa trên màu sắc và cấu trúc.

Bài báo sẽ tiếp cận phân đoạn hình ảnh sao cho mỗi hình ảnh được phân đoạn thành các vùng đặc trưng để từ đó làm cơ sở xây dựng chữ ký nhị phân nhằm mô tả nội dung hình ảnh.

Ảnh phân đoạn được tạo ra từ việc nhóm các điểm ảnh trở thành một vùng tương tự. Bài báo tiếp cận phương pháp phân đoạn ảnh tự động dựa trên các thông tin cấp thấp gồm màu sắc, cấu trúc và vị trí các điểm ảnh.

Để tính toán các giá trị này, hình ảnh được chia ra thành các khối vuông $f \times f$ không giao nhau. Do đó, hình ảnh được chia thành L khối b_l , $l = 1, \dots, L$. Sau đó tính vector cấu trúc và vector màu sắc của mỗi khối bằng giá trị trung bình của các điểm ảnh trong khối đó.



Hình 3. Ví dụ hình ảnh tách thành 7×11 khối

Kết quả phân đoạn là mặt nạ phân đoạn, tức là một ảnh xám mô tả đối tượng đặc trưng của hình ảnh. Dựa trên mặt nạ phân đoạn, thực hiện tính toán vùng liên thông và loại bỏ các vùng liên thông có diện tích không vượt ngưỡng (trong thực nghiệm sẽ loại bỏ các vùng liên thông có diện tích nhỏ hơn 5% diện tích ảnh).

Thuật toán 1: Phân đoạn ảnh

Đầu vào: Ảnh màu I

Đầu ra: Mặt nạ phân đoạn M

Bước 1: Trích xuất vector cấu trúc và vector cường độ cho mỗi điểm ảnh.

Bước 2: Tính tâm các khối bằng cách lấy giá trị trung bình vector cấu trúc và vector màu sắc của tất cả các điểm ảnh trong khối.

Bước 3: Tính độ tương phản C của hình ảnh để tạo thành đối tượng nền và đối tượng đặc trưng.

Bước 4: Tìm các tâm cụm cho các đối tượng đặc trưng hỗ trợ dựa trên độ tương phản.

Bước 5: Dựa trên tập các tâm cụm của các đối tượng đặc trưng, thực hiện gom cụm các điểm ảnh.

Bước 6: Tạo mặt nạ phân đoạn M ứng với các điểm ảnh đã phân cụm.

Bước 7: Loại bỏ các đối tượng có diện tích nhỏ dựa trên mặt nạ phân đoạn M

Bước 8: Trả về mặt nạ phân đoạn M .

Đối với vector đặc trưng màu sắc của mỗi điểm ảnh $I(p) = (I_L(p), I_a(p), I_b(p))$ được xây dựng dựa trên không gian màu $CIE L^*a^*b^*$. Không gian màu $CIE L^*a^*b^*$ được công nhận là chuẩn quốc tế vào thập niên 1970 bởi tổ chức CIE và đồng nhất với nhận thức con người. Khoảng cách Euclidean giữa hai điểm trong không gian màu $CIE L^*a^*b^*$ tương đương với khoảng cách nhận thức giữa 2 màu theo hệ thống thị giác của con người [1].

Sau khi thực hiện trích xuất cấu trúc và màu sắc của hình ảnh, thực hiện quá trình gom cụm các điểm ảnh bằng phương pháp gom cụm K-Means. Bước đầu tiên của quá trình gom cụm này đó là chọn ra tâm cụm dựa trên độ tương phản C của hình ảnh. Để thực hiện nhanh quá trình này, hình ảnh được chia thành L khối ảnh không giao nhau, mỗi khối ảnh này được xem là một điểm ảnh lớn (*super pixel*). Do đó, vector cấu trúc $T^b(b_l)$ và vector màu sắc $I^b(b_l)$ của mỗi khối b_l tương ứng với các giá trị trung bình của vector cấu trúc và vector màu sắc của tất cả các điểm ảnh trong khối. Với b_l, b_n là 2 khối bất kỳ, độ tương phản của hình ảnh sẽ là: $C = \max\{\alpha \|I^b(b_l) - I^b(b_n)\| + \beta \|T^b(b_l) - T^b(b_n)\|\}$ (trong thực nghiệm $\alpha = \beta = 0.5$). Sau khi tìm được độ tương phản C của hình ảnh, khối có năng lượng thấp hơn sẽ là tâm cụm nền (*background*) và khối có năng lượng cao sẽ là tâm cụm của đối tượng đặc trưng (*foreground*).

Bước tiếp theo sẽ thực hiện tìm tâm các cụm của các đối tượng đặc trưng bổ sung, tức là tìm các khối có độ đo d gần với đối tượng đặc trưng nhất (tương ứng sẽ là xa nhất đối với màu nền). Trong thực nghiệm sẽ tìm các tâm cụm có độ đo $d > \mu C$ (với $\mu = 0.4$). Sau đó, thực hiện phép gom cụm cho tất cả các điểm ảnh.

Bước cuối cùng là loại bỏ các vùng liên thông có diện tích không vượt ngưỡng θ . Việc tính diện tích vùng dựa trên thuật toán loang 4-liên thông được tóm tắt như sau:

Thuật toán 2: Tính diện tích vùng

Đầu vào: Mặt nạ phân đoạn M và vị trí (r,c)

Đầu ra: Giá trị diện tích S

Bước 1: Khởi tạo $S = 0$;

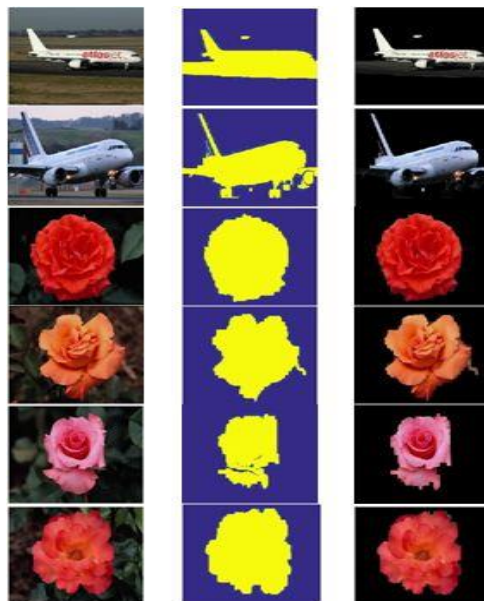
Stack = \emptyset ;

Bước 2: Push(Stack, r, c);

Bước 3: **while** Stack $\neq \emptyset$ **do**

```

(r,c) = Pop(Stack);
S = S + 1;
If (r > 1)&& (M(r,c) = M(r-1,c) then
    Push(Stack,r-1,c);
If (r < rows)&&(M(r,c)= M (r+1,c) then
Push(Stack,r+1,c);
If (c > 1)&&M(r,c) == M (r,c-1) then
    Push(Stack,r,c-1);
If (c < columns)&&M(r,c) = M(r,c+1) then
    Push(Stack,r,c+1);
Bước 4: Return S;
    
```

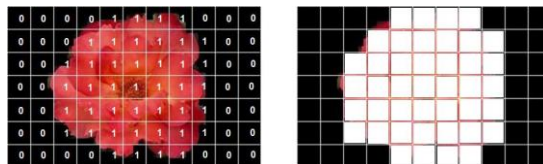


Hình 4. Một số kết quả mẫu phân đoạn ảnh

5. TRUY VẤN ẢNH

5.1. Tạo chỉ mục nhị phân

Đặc trưng của hình ảnh có thể được mô tả bằng một vector đặc tính đa chiều gọi là chữ ký của hình ảnh (*image signature*).



Hình 5. Minh họa cách tạo chữ ký nhị phân

Sau khi tạo chỉ mục cho hình ảnh trong cơ sở dữ liệu, điều quan trọng là sử dụng một độ đo tương tự nhằm truy vấn trong cơ sở dữ liệu.

5.2. Độ đo tương tự

Gọi sig^I và sig^J lần lượt là 2 chữ ký nhị phân của hai hình ảnh I và J . Độ trùng khớp d_i được đối sánh trên mỗi phần tử của 2 chữ ký và được định nghĩa như sau:

$$d_i = \begin{cases} 1 & \text{if } (sig_i^I = sig_i^J) \\ 0 & \text{if } (sig_i^I \neq sig_i^J) \end{cases}$$

Độ đo tương tự của 2 hình ảnh I và J được định nghĩa là: $\phi = \frac{1}{n} \sum_{i=1}^n d_i$. Dễ dàng chứng minh ϕ thỏa các tính chất của một chuẩn, gồm:

(1) Không âm: $\phi(sig_I, sig_J) \geq 0$

Nếu $\phi(sig_I, sig_J) = 0 \Leftrightarrow sig_I = sig_J$

(2) Đối xứng: $\phi(sig_I, sig_J) = \phi(sig_J, sig_I)$

(3) Bất đẳng thức tam giác:

$$\phi(sig_I, sig_J) + \phi(sig_J, sig_K) \geq \phi(sig_I, sig_K)$$

5.3. Cây đa nhánh cân bằng

Nhằm giảm không gian và tăng tốc độ truy vấn, nhóm nghiên cứu xây dựng cây đa nhánh cân bằng lưu trữ các chỉ mục mô tả hình ảnh. Mỗi một nút trong cây lưu trữ tập các phần tử $\langle \langle SIG, p \rangle \rangle$, với SIG là chữ ký và p là con trỏ tham chiếu đến nút con. Các nút lá sẽ lưu trữ các phần tử $\langle \langle sig, oid \rangle \rangle$, với sig là chỉ mục của mỗi hình ảnh và oid là định danh của hình ảnh tương ứng. Quá trình tạo cây dựa trên thao tác chèn và tách nút trong cây. Thuật toán tạo cây chữ ký được đề xuất như sau:

Input: tập các chỉ mục $S = \{sig_1, \dots, sig_n\}$

Output: cây đa nhánh T

Algorithm1. Gen-Tree(S , Root)

Begin

Bước 1.

$v = \text{Root}$;

If $S = \emptyset$ **then** STOP;

Else Chọn $\langle sig, oid \rangle \in S$ và $S = S \setminus \langle sig, oid \rangle$;

Qua bước 2;

Bước 2.

If (v là nút lá) **then**

begin

If $v.\text{count} < M$ **then** $v = v \cup \langle sig, oid \rangle$;

Else SplitNode(v , sig);

 Quay lại bước 1;

end

Else

begin

$W = \{SIG_i \mid SIG_i \in v \text{ và } SIG_i \rightarrow sig \wedge sig = sig\};$

$EMD(SIG_0 \rightarrow sig, sig) = \min\{EMD(SIG_i \rightarrow sig, sig) \mid SIG_i \in W\};$

$v = SIG_0 \rightarrow p;$

Quay lại bước 2;

end

End.

Thuật toán **Algorithm1** lần lượt đưa các chỉ mục sig từ tập chữ ký S vào trong cây. Với mỗi chỉ mục sig sẽ được chèn vào nút lá phù hợp, nếu nút lá đầy thì quá trình tách nút sẽ được thực hiện và cây đa nhánh tăng trưởng chiều cao theo hướng gốc của cây. Tại mỗi nút trong của cây, sẽ ưu tiên đi theo hướng có độ tương tự nhiều hơn, quá trình này sẽ được duyệt cho đến khi tìm ra được nút lá phù hợp. Quá trình duyệt cây không nhất thiết phải duyệt qua tất cả các hướng có chữ ký phù hợp với chữ ký hình ảnh cần chèn, điều này sẽ giảm một khoảng chi phí đáng kể trong quá trình tìm ra các nút lá phù hợp để chèn chữ ký vào cây. Do đó, ứng với mỗi chữ ký cần chèn sẽ duyệt qua đường đi có chiều cao $h = \lceil \log_m n - 1 \rceil$, với m là số chữ ký tối thiểu của một nút trong cây. Gọi k là chiều dài của mỗi chữ ký, mỗi một nút trong của cây sẽ có tối đa là M chữ ký, vì vậy quá trình duyệt cây để tìm ra nút lá phù hợp sẽ có chi phí tối đa là $k \times M \times \lceil \log_m n - 1 \rceil$. Tuy nhiên, khi tìm ra nút lá phù hợp nhưng đã bị đầy, cần phải thực hiện quá trình tách nút. Việc tách nút dựa trên cơ sở phép toán α -seed, β -seed, được thực hiện theo thuật toán đề xuất như sau:

Input: Nút cần tách v

Output: Cây T sau khi thực hiện phép tách nút

Algorithm2. SplitNode(v)

Begin

Tạo nút v_α và v_β lần lượt chứa chữ ký α -seed và β -seed ;

$v = v \setminus \{ \alpha\text{-seed}, \beta\text{-seed} \};$

For ($SIG_i \in v$)

begin

If($EMD(SIG_i \rightarrow sig, \alpha\text{-seed}) < EMD(SIG_i \rightarrow sig, \beta\text{-seed})$)**then**

$v_\alpha = v_\alpha \cup SIG_i;$

Else

$v_\beta = v_\beta \cup SIG_i;$

end

$s_\alpha = \bigcup SIG_i^\alpha$, với $SIG_i^\alpha \in v_\alpha$

$s_\beta = \bigcup SIG_i^\beta$, với $SIG_i^\beta \in v_\beta$

If ($v_{parent} \neq \text{null}$) **then** $v_{parent} = v_{parent} \cup s_\alpha$; $v_{parent} = v_{parent} \cup s_\beta$;

If ($v_{parent}.count > M$) **then** SplitNode(v_{parent});

If ($v_{parent} = \text{null}$) **then** Root = $\{ s_\alpha, s_\beta \};$

End.

5.4. Thuật toán truy vấn ảnh dựa trên cây S-tree

Sau khi lưu trữ chỉ mục và định danh của hình ảnh tương ứng trên cây, quá trình truy vấn sẽ tìm ra các chữ ký tương tự của hình ảnh dựa trên việc duyệt cây. Sau khi tìm ra các chữ ký hình ảnh, dựa vào định danh của các hình ảnh sẽ tìm ra cụ thể các hình ảnh tương tự với hình ảnh truy vấn. Do đó, bài toán cần thực hiện là tìm ra chữ ký của hình ảnh và định danh tương ứng, quá trình truy vấn này được thực hiện theo thuật toán đề xuất như sau:

Input: chữ ký truy vấn sig và T

Output: Tập các chữ ký ảnh và các định danh tham chiếu đến hình ảnh tương ứng

Algorithm3. Search-Image-Sig(sig, S-tree)

Begin

$v = \text{root}; \text{SIGOUT} = \emptyset; \text{Stack} = \emptyset; \text{Push}(\text{Stack}, v);$

while(not Empty(Stack)) **do**

begin

$v = \text{Pop}(\text{Stack});$

If(v is not Leaf) **then**

begin

For($\text{SIG}_i \in v$ and $\text{SIG}_i \rightarrow \text{sig} \wedge \text{sig} = \text{sig}$) **do**

$\text{EMD}(\text{SIG}_0 \rightarrow \text{sig}, \text{sig}) = \min\{\text{EMD}(\text{SIG}_i \rightarrow \text{sig}, \text{sig}) | \text{SIG}_i \in v\};$

$\text{Push}(\text{Stack}, \text{SIG}_0 \rightarrow \text{next});$

end

Else $\text{SIGOUT} = \text{SIGOUT} \cup \{ \langle \text{SIG}_i \rightarrow \text{sig}, \text{oid}_i \rangle | \text{SIG}_i \in v \};$

end

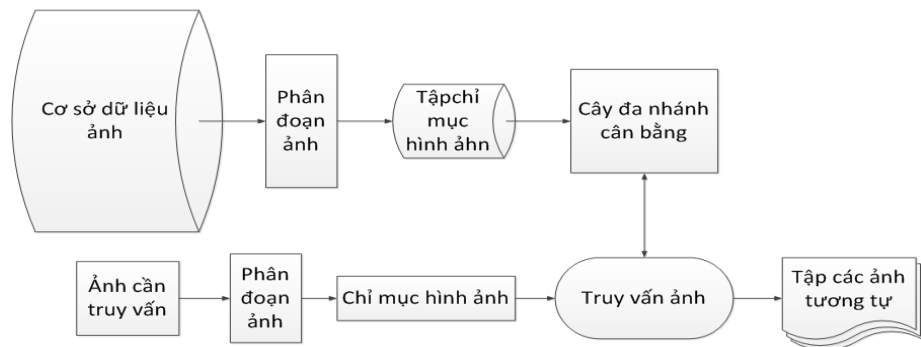
return SIGOUT;

End.

5.5. Thực nghiệm truy vấn ảnh

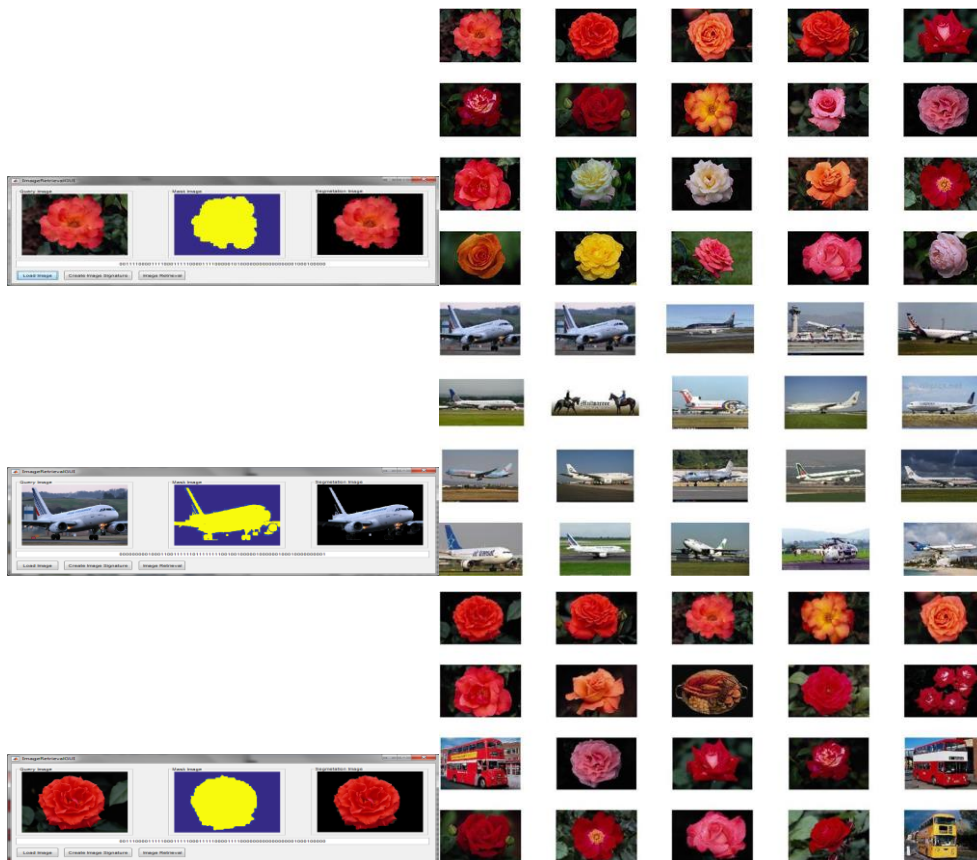
Bài báo thực hiện quá trình truy vấn ảnh tương tự trên dữ liệu thực nghiệm gồm Corel và Image Object (MSRC). Các hình ảnh trong các tập dữ liệu này được chia thành các chủ đề khác nhau. Ứng với mỗi hình ảnh truy vấn sẽ tìm ra các hình ảnh tương tự trong tập dữ liệu ảnh này.

Ứng dụng thực nghiệm được xây dựng trên nền tảng công cụ IPT (Image Processing Toolbox) của Matlab 2015. Thực nghiệm được thực thi trên máy tính với bộ xử lý Intel(R) CoreTM i7-2620M, CPU 2.70GHz, RAM 4GB, Hệ điều hành Windows 7 Professional 64 bit.

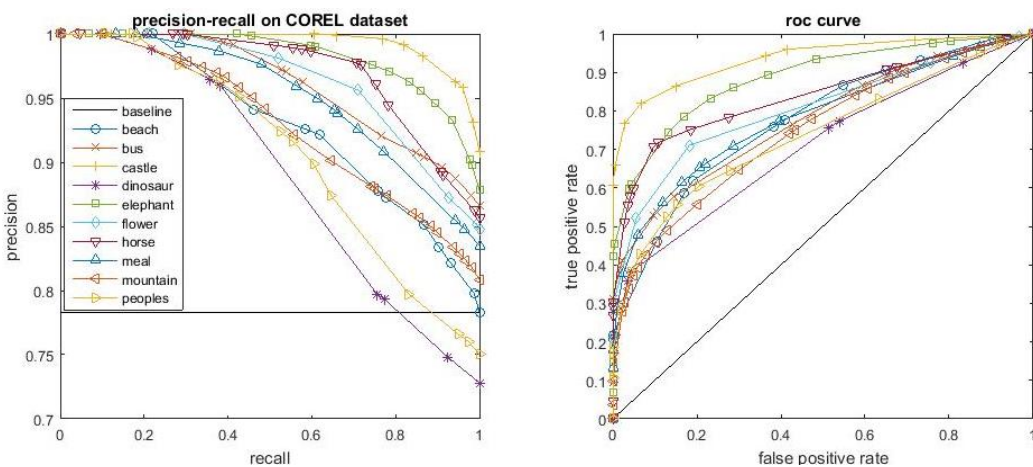


Hình 6. Mô hình truy vấn ảnh

Quá trình truy vấn hình ảnh được chia làm 2 giai đoạn gồm: Giai đoạn tiền xử lý gồm các bước: (1) Phân đoạn hình ảnh ứng; (2) Tạo chỉ mục để tạo thành tập chữ ký ảnh; (3) Tạo cây đa nhánh cân bằng. Giai đoạn truy vấn ảnh thực hiện: (1) Phân đoạn ảnh truy vấn; (2) Tạo chỉ mục cho ảnh truy vấn; (3) Thực hiện truy vấn ảnh để tìm các hình ảnh tương tự.



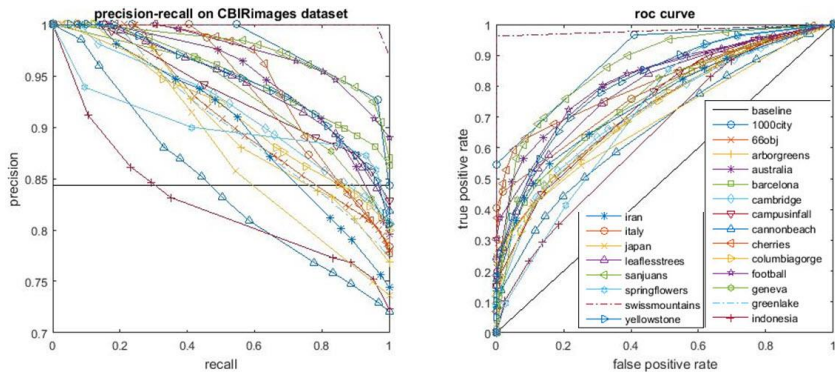
Hình 7. Một số kết quả truy vấn ảnh



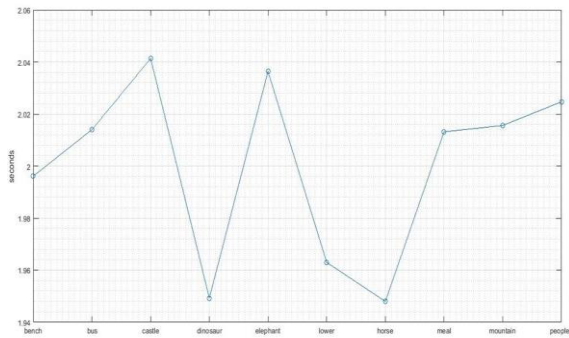
Hình 8. Độ phủ và độ chính xác khi truy vấn ảnh trên tập dữ liệu ảnh COREL

Độ phủ (*recall*) = (số ảnh truy vấn có liên quan)/(Tổng số ảnh có liên quan trong tập dữ liệu ảnh)

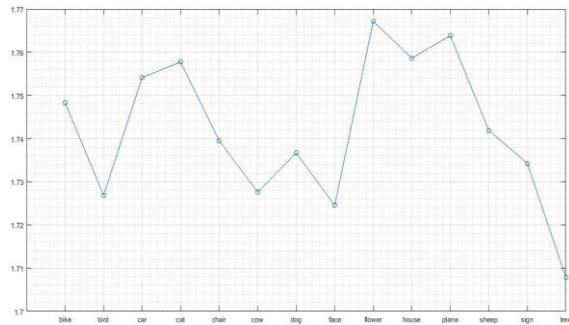
Độ chính xác (*precision*) = (số ảnh truy vấn có liên quan)/(Ngưỡng xác định số ảnh truy vấn)



Hình 9. Độ phủ và độ chính xác khi truy vấn ảnh trên tập dữ liệu ảnh CBIRimages



Hình 10. Thời gian truy vấn ảnh trung bình trên tập dữ liệu ảnh COREL



Hình 11. Thời gian truy vấn ảnh trên tập dữ liệu ảnh CBIRimages

Bảng 1. So sánh độ chính xác truy vấn giữa các phương pháp

Phương pháp	Độ chính xác trung bình	Độ phủ trung bình	Độ đo F-measure trung bình	Thời gian truy vấn trung bình
S-Tree	0,42	0,55	0,476289	186,25 I/Os
Fuzzy Signatures	N/A	N/A	N/A	20-50 I/Os
Fuzzy color histogram	0,50688	0,61625	0,55624	4,41863 sec
Interest region	0,85200	0,78375	0,81645	4,78516 sec
Phương pháp đề xuất	0,687480469	0,687487535	0,687484002	ec

6. KẾT LUẬN

Bài báo đã tiếp cận quá trình tạo chỉ mục dựa trên phân đoạn hình ảnh để tạo thành chữ ký nhị phân và cây đa nhánh cân bằng để từ đó thực hiện bài toán truy vấn ảnh tương tự dựa trên nội dung. Bài báo đã mô tả quá trình phân đoạn hình ảnh dựa trên không gian màu CIE

$L^*a^*b^*$ và phép biến đổi Wavelet để làm tiền đề tạo chữ ký nhị phân. Theo thực nghiệm cho thấy việc truy vấn ảnh dựa trên chữ ký nhị phân mang lại nhiều hiệu quả đồng thời giúp cho việc tìm kiếm nhanh và giảm đáng kể không gian truy vấn trong quá trình đối sánh ảnh tương tự. Tuy nhiên, đối với phương pháp tạo chỉ mục nhị phân như trên sẽ cho kết quả chính xác thấp nếu ảnh bị phân tán màu sắc và vị trí. Hơn nữa, nếu thực hiện truy vấn trực tiếp trên dữ liệu mô tả này có thể tốn kém thời gian nếu như dữ liệu chữ ký ảnh lớn. Vì vậy, hướng phát triển của bài báo sẽ tạo ra chữ ký nhị phân của hình ảnh vừa mô tả hình dạng, vị trí và màu sắc. Đồng thời thực hiện tiền xử lý gom cụm chữ ký tương tự để giúp cho quá trình truy vấn nhanh và hiệu quả.

Lời cảm ơn: Nghiên cứu này được Trường Đại học Công nghiệp Thực phẩm TP.HCM tài trợ và được nhóm nghiên cứu SBIR-HCM, Trường Đại học Sư phạm TP.HCM hỗ trợ về chuyên môn và cơ sở vật chất.

TÀI LIỆU THAM KHẢO

1. Acharya T., Ray A.K - Image processing: principles and applications, John Wiley & Sons Inc. (2005).
2. Oge Marques Borko Furht - Content-based image and video retrieval, Springer Science + Business Media New York, 2002.
3. James Z. Wang - Integrated Region-Based Image Retrieval, Springer Science Business Media New York, 2001.
4. Michael. S. Lew - Principles of Visual Information Retrieval, Springer-Verlag London, 2001.
5. Muneesawang P. - Multimedia Database Retrieval: Technology and Applications, Springer Heidelberg New York, 2014.
6. Kim S. - Central Object Extraction for Object-Based Image Retrieval, Springer-Verlag, LNCS, 2728, (2003) 39-49.
7. Jianru Xue - Automatic salient object extraction with contextual cue and its app
8. Huang Y., Zhang J., Zhao Y., Ma D. - Medical image retrieval with query-dependent feature fusion based on one-class SVM, in ICCES: 13th IEEE International Conference on Computational Science and Engineering (2010) 176-183.
9. Jin L., Hong L. Lianzhi T. - A mapping modelling of visual feature and knowledge representation approach for Medical Image Retrieval, in: 2009 International Conference on Mechatronics and Automation (2009) 1778-1783.
10. Rajakumar K., Muttan S. - Medical image retrieval using energy efficient wavelet transform, in: 2010 Second International conference on Computing, Communication and Networking Technologies (2010) 1-5.
11. Shea G. Y. K., Cao J. - Geo-Planar Indexing (GPI) – An efficient indexing scheme for fast retrieval of raster-based geospatial data in mobile GIS application, Proc. of IEEE CISP (2012) 1047-1052.
12. Wang X. Y. - Robust Image Retrieval Based on Color Histogram of Local Feature Regions, Springer Science, Multimed Tools Appl. **49** (2010) 323-345.
13. Bartolini I., Ciaccia P., Patel M. - Query processing issues in region-based image databases, Knowledge and Information Systems **25** (2) (2010) 389-420.
14. Wang X.Y. - Robust color image retrieval using visual interest point feature of significant bit-planes, DSP **23** (4) (2013) 1136-1153.
15. Tang Z. - Robust Image Hash Function Using Local Color Features, Int. J. Electron. Commun. **67** (2013) 717-722.
16. Wang X.Y. - Robust color image retrieval using visual interest point feature of significant bit-planes, Digital Signal Processing **23** (4) (2013) 1136-1153.

17. Kumar H.C.S., Raja K.B., Venugopal K.R., Patnaik L.M. - Automatic Image Segmentation using Wavelets, IJCSNS International Journal of Computer Science and Network Security **9** (2) (2009) 305-313.
18. Manolopoulos Y. - Advanced Signature Indexing for Multimedia and Web Applications, Springer Science New York, 2003.
19. Nascimento M. A., Chitkara V. - Color-based image retrieval using binary signatures, SAC Madrid (2002) 687-692.
20. Timothy Chappell, Shlomo Geva - Efficient Top-K retrieval with signatures, ADCS'13, Brisbane (2013) 10-17.
21. Nascimento M. A. - Image indexing and retrieval using signature trees, Data & Knowl. Engineering **43** (2002) 57-77.
22. Thanh Manh Le, Thanh The Van - Image retrieval system based on emd similarity measure and S-Tree, ICITES-2012, Springer Verlag, LNEE 234 (2013) 139-146.
23. Huang P.W. - Indexing pictures by key objects for large-scale image database, Pattern Recognition **30** (7) (1997) 1229-1237.
24. Thanh T.V., Thanh M. Le. - Image retrieval Based on Binary signature and S-kGraph, Jour. of Annales Univ. Sci. Budapest., Sect. Comp. **43** (2014) 105-122.
25. Thanh The Van, Thanh Manh Le - RBIR Based on Signature Graph, ICCCI-2014, IEEE Xplore (2014) 1-4.
26. Lê Sĩ Đồng - Xác suất và thống kê và ứng dụng, NXB Giáo dục, 2004.
27. Feller W. - An introduction to probability theory and its applications, Vol. 1, 3rd Ed., John Wiley & Sons Inc., New York (1968).
28. Unser M. - Texture classification and segmentation using wavelet frames, IEEE Trans. on Im. Processing **4** (11) (1995) 1549-1560.

ABSTRACT

A METHOD OF IMAGE RETRIEVAL USING BALANCE TREE

Nguyen Phuong Hac*, Van The Thanh
Ho Chi Minh City University of Food Industry
*Email: hacnp@hufi.edu.vn

The paper approaches a method of image retrieval using balance tree to store meta-data of image features by index to describe the shape of interest objects of image. In order to extract interest objects, we propose the image segmentation method on the base of low-level visual features including color and texture of image. The features are extracted at each block of image by Wavelet transformation and CIE L*a*b* color space. From there, we cluster all pixels to give connected regions; at the same time, we remove regions that have area less than threshold. Based on extracted features, we find out the contrast of image in order to extract background and foreground. On the base of segmented image, the feature indexes are created in order to match similar images. Then, we create the balance tree to store the feature indexes and search the set of similar images based on this structure. To illustrate the proposed method, the paper builds application and assesses experimental results on image databases including Corel and Image Object (MSRC). The results of our method are compared with others and show the effective of our proposed method.

Keywords: Segmentation method, interest object, feature index.