

MỘT PHƯƠNG PHÁP TRÍCH XUẤT ĐẶC TRƯNG CHO BÀI TOÁN TÌM KIẾM ẢNH

Nguyễn Thị Uyên Nhi^{1,2}, Văn Thế Thành³

¹ Khoa Công nghệ thông tin, Trường Đại học Khoa học, Đại học Huế

² Khoa Thống kê – Tin học, Trường Đại học Kinh tế, Đại học Đà Nẵng

³ Phòng QLKH và ĐT SDH, Trường ĐH Công nghiệp Thực phẩm TP.HCM

Email: nhintu@due.edu.vn, thanhvt@hufi.edu.vn

Ngày nhận bài: 6/4/2021; ngày hoàn thành phản biện: 9/6/2021; ngày duyệt đăng: 16/6/2021

TÓM TẮT

Trong bài báo này, một phương pháp trích xuất đặc trưng cho bài toán tìm kiếm ảnh được đề xuất. Các kỹ thuật trích xuất đặc trưng bao gồm: phân vùng ảnh dựa trên độ tương phản; bộ mô tả màu chủ đạo MPEG-7; nhận dạng đối tượng dựa trên biên và làm mịn bề mặt dựa trên phép lọc Sobel; phát hiện biên đối tượng với phương pháp LoG (Laplace of Gaussian); nâng cao cường độ ảnh với phép lọc Gaussian... Từ đó, một phương pháp tìm kiếm ảnh dựa trên cây C-Tree [20] được thực hiện để tìm tập các ảnh tương tự với một ảnh truy vấn cho trước. Để minh chứng cho lý thuyết đề xuất thực nghiệm được xây dựng trên tập ảnh COREL-1K, WANG, đồng thời đánh giá độ chính xác của phương pháp đề xuất. Kết quả thực nghiệm cho thấy phương pháp trích xuất đặc trưng hình ảnh và tìm kiếm ảnh tương tự của chúng tôi là hiệu quả.

Từ khóa: C-Tree, phân cụm, tìm kiếm ảnh, trích xuất đặc trưng.

1. GIỚI THIỆU

Những năm gần đây, ảnh số đã tạo ra một lượng dữ liệu khổng lồ, ứng dụng trong nhiều lĩnh vực khác nhau [13]. Vì vậy, để truy cập cơ sở dữ liệu này, nhiều hệ thống truy xuất ảnh đã được phát triển với mục đích nâng cao hiệu quả của tìm kiếm. Hệ thống truy vấn hình ảnh dựa trên nội dung CBIR (Content-based Image Retrieval) [4][14][24] là phương pháp tìm kiếm phổ biến hiện nay. CBIR trích xuất các đặc trưng thị giác cấp thấp (màu sắc, kết cấu, hình dạng, bố cục không gian, v.v.) nhằm mô tả nội dung ảnh và tạo cơ sở cho quá trình tìm kiếm ảnh.

Trong bài báo này, chúng tôi trích xuất đặc trưng theo màu sắc, hình dạng, vị trí và bề mặt đối tượng nhằm thực hiện tìm kiếm ảnh tương tự. Việc trích xuất đặc trưng

này được dựa trên kỹ thuật phân vùng ảnh dựa trên độ tương phản [9]; kỹ thuật trích xuất đặc trưng màu MPEF-7 [9][22]; kỹ thuật nhận dạng đối tượng dựa trên biên và làm mịn bề mặt dựa trên phép lọc Sobel [10]; kỹ thuật phát hiện đường biên đối tượng dựa trên phương pháp LoG (Laplace of Gaussian) [2]; phương pháp trích xuất cấu trúc bề mặt ảnh dựa trên các phép lọc và nâng cao cường độ ảnh với phép lọc Gaussian [8].... Các vec-tơ đặc trưng được lưu trữ trên cấu trúc cây phân cụm C-Tree [20], làm cơ sở cho bài toán tìm kiếm ảnh. Thực nghiệm được xây dựng trên tập ảnh COREL, WANG để minh chứng cho lý thuyết đã đề xuất. Đồng thời, kết quả được so sánh với các phương pháp khác trên cùng tập ảnh để chứng tỏ tính hiệu quả của hệ truy vấn.

Đóng góp của bài báo gồm: (1) Trích xuất đặc trưng cấp thấp của hình ảnh bao gồm đặc trưng màu sắc, vị trí và cấu trúc bề mặt ảnh trên cơ sở cải tiến các thuật toán về phân vùng ảnh dựa trên độ tương phản, trích xuất đường biên đối tượng, trích xuất cấu trúc bề mặt, phân cụm các điểm ảnh theo cường độ màu sắc; (2) Đề xuất mô hình tìm kiếm ảnh tương tự dựa trên vec-tơ đa đặc trưng với cây phân cụm C-Tree; (3) Xây dựng thực nghiệm tìm kiếm ảnh tương tự và so sánh với một số công trình cùng bộ dữ liệu.

2. CÁC CÔNG TRÌNH LIÊN QUAN

Việc tìm kiếm ảnh bằng vec-tơ đặc trưng là một trong những phương pháp tìm kiếm nhằm giảm chi phí về bộ nhớ xử lý và thời gian đối sánh ảnh, trong đó mỗi hình ảnh được trích xuất một vec-tơ đặc trưng để đánh giá độ tương tự giữa các hình ảnh. Vec-tơ đặc trưng của hình ảnh được trích xuất dựa vào độ tương phản, màu sắc, kết cấu bề mặt, hình dạng, vị trí tương đối.... Những công trình công bố gần đây, nhiều kỹ thuật trích xuất đặc trưng cấp thấp nhằm mô tả nội dung ảnh như màu sắc, hình dạng, kết cấu được đề xuất cho bài toán tìm kiếm ảnh.

Shao H. và cộng sự [22] đề xuất sử dụng Bộ mô tả màu chủ đạo DCD (Dominant Color Descriptor) của MPEG-7 để trích xuất vec-tơ đặc trưng cho quá trình tìm kiếm ảnh. DCD mô tả sự phân bố màu sắc đại diện và các tính năng trong một hình ảnh hoặc một khu vực quan tâm thông qua một định dạng hiệu quả, nhỏ gọn và trực quan. A. Huneiti và cộng sự [18] đề xuất một phương pháp truy vấn ảnh theo nội dung bằng cách trích xuất cả các vectơ đặc trưng màu và kết cấu bằng cách sử dụng Biến đổi Wavelet rời rạc DWT (Discrete Wavelet Transform) và Bản đồ tự tổ chức (SOM). Nhóm nghiên cứu Ashraf R. [1] đề xuất một kỹ thuật biểu diễn hình ảnh dựa trên biến đổi dải màu nhằm xác định các đối tượng chính trong một hình ảnh, khắc phục cho các nhược điểm của kỹ thuật phân đoạn để xác định đối tượng như thời gian trích xuất chậm và kết quả không đáng tin cậy. Một phương pháp truy xuất ảnh dựa trên đặc trưng thông tin hợp nhất FIF-IRS (Fused Information Feature-based Image Retrieval System) của nhóm nghiên cứu Bella M. I. T. [5] đề xuất, bao gồm các đặc trưng ma trận đồng xuất hiện mức xám 8D-GLCM (8-Directional Gray Level Co-occurrence Matrix) và không gian màu HSV.

Phương pháp Fuzzy-NN được nhóm nghiên cứu Garg M. [17] đề xuất, kết hợp giữa Neuro-Fuzzy và Deep Neural Network nhằm trích xuất đặc trưng kết cấu ảnh cho bài toán phân loại và truy vấn ảnh. Nhóm nghiên cứu Alsmadi Mutasem K. [19] đề xuất phương pháp trích xuất các vec-tơ đặc trưng từ chữ ký màu, hình dạng và kết cấu. Theo đó, đánh giá cường độ sáng tương đồng của ảnh với thuật toán metaheuristic. Các đặc trưng màu trích xuất dựa trên không gian màu RGB với thuật toán phân cụm và phương pháp cạnh Canny để trích xuất các đặc điểm hình dạng, màu YCbCr với biến đổi wavelet rời rạc và ma trận đồng xuất hiện mức xám để trích xuất đặc điểm kết cấu. Sự kết hợp của các phương pháp này làm tăng hiệu suất của hệ truy vấn ảnh.

Từ phân tích các công trình liên quan cho thấy mô hình truy vấn ảnh dựa trên vec-tơ đặc trưng là một mô hình được đánh giá khả thi và hiệu quả. Trong bài báo này, chúng tôi tiếp cận theo phương pháp trích xuất vec-tơ đặc trưng theo nhiều kỹ thuật khác nhau để thực hiện tìm kiếm ảnh.

3. KỸ THUẬT TRÍCH XUẤT VEC-TƠ ĐẶC TRƯNG

Hình ảnh có cấu trúc không đồng nhất, kích thước các tập ảnh rất lớn, gây khó khăn trong việc lưu trữ trên bộ nhớ ngoài, và làm chậm quá trình tìm kiếm ảnh. Do đó, kỹ thuật lập chỉ mục [7] là một cách để quản lý ảnh số nhằm tiết kiệm bộ nhớ máy tính và lưu trữ dữ liệu đồng nhất, đồng thời tăng tốc độ hiển thị và tìm kiếm. Các nội dung ảnh được biểu diễn bằng các vec-tơ đặc trưng thị giác nhằm trở đến các tệp kê khai ảnh cụ thể, do đó tạo thành một cơ sở dữ liệu gồm các chỉ mục là các vec-tơ đặc trưng này. Mỗi đặc trưng ảnh được biểu diễn bằng cách sử dụng một hoặc nhiều bộ mô tả khác nhau. Trong bài báo này, các kỹ thuật trích xuất vec-tơ đặc trưng được đề xuất như sau:

3.1. Phân vùng ảnh

Việc phân đoạn ảnh màu được thực hiện bằng cách phân chia hình ảnh thành các vùng riêng biệt khác nhau để từ đó trích xuất đặc trưng trên mỗi vùng [8]. Để thực hiện được điều này, chúng tôi đề xuất phương pháp phân vùng dựa trên độ tương phản, nghĩa là vùng nào có độ tương phản thấp là hình nền và vùng nào có độ tương phản cao là hình đối tượng. **Hình 1** là kết quả ảnh phân đoạn tương ứng với hình đối tượng và hình nền gồm: (a) ảnh gốc, (b) ảnh mô tả độ tương phản, (c) ảnh đối tượng, (d) ảnh nền của đối tượng. Để làm giảm độ nhiễu các vùng quá sáng hoặc quá tối, một số điểm ảnh nằm trong lân cận của giá trị lớn nhất và giá trị nhỏ nhất của độ tương phản thì được quy về giá tương ứng.



Hình 1. Kết quả phân đoạn ảnh màu

3.2. Đặc trưng màu

Có rất nhiều kỹ thuật trích xuất đặc trưng màu sắc của ảnh, trong đó MPEG-7 với bộ mô tả màu chủ đạo DCD (Dominant Color Descriptor) [9] là sự thể hiện hiệu quả của các màu nổi bật trong một vùng ảnh [106][116]. Bộ mô tả màu chủ đạo DCD (Dominant Color Descriptor) được định nghĩa theo công thức (1):

$$DCD = \left\{ \left\{ CVal_j, Per_j, CVar_j \right\}, SC \right\}, (j = 1, 2, \dots, M) \quad (1)$$

Trong đó, M là số màu trội; $CVal_j$ là giá trị màu chủ đạo trong không gian màu tương ứng; ($Per_j, Per_j \in [0, 1], \sum_j Per_j = 1$) là phần trăm pixel tương đương với màu $CVal_j$ trong ảnh hoặc vùng ảnh; và phương sai màu $CVar_j$ xác định sự thay đổi của màu pixel trong một nhóm trên bộ màu có liên quan. Độ liên kết không gian SC là một số duy nhất đại diện cho sự đồng nhất hoàn toàn trong không gian của các màu chủ đạo của ảnh.

Quy trình trích xuất đặc trưng DCD bao gồm ba giai đoạn: chuyển đổi không gian màu, kỹ thuật gom cụm và tính toán tỷ lệ phần trăm của mỗi tâm. Giai đoạn đầu tiên là chuyển đổi không gian màu của MPEG-7 là chuyển từ RGB sang CIE-LUV nhằm thực hiện quá trình gom cụm các màu sắc chủ đạo. Sau khi ảnh đầu vào được chuyển đổi sang CIE-LUV, một thuật toán phân cụm k-Means được sử dụng để tìm màu sắc ảnh chủ đạo.

Như vậy các màu sắc ảnh hoặc một khu vực ảnh được trích xuất bởi DCD nhằm cung cấp một mô tả màu sắc hiệu quả, nhỏ gọn và đơn giản.

3.3. Kỹ thuật phát hiện biên đối tượng

Phương pháp phát hiện biên với phép lọc Sobel là phương pháp phát phổ biến dùng Gradient [10]. Sobel được sử dụng để tìm các đường biên theo chiều dọc và chiều ngang. Hai mặt nạ có kích thước $[3 \times 3]$ theo công thức (2) được dùng cho phép lọc Sobel. Tích chập giữa ảnh và các mặt nạ này, thu được kết quả là các gradient theo chiều dọc và chiều ngang G_x, G_y . Độ lớn Gradient được tính bằng cách sử dụng công thức (3).

$$Sobel_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad Sobel_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (2)$$

$$G[f(x, y)] = \sqrt{G_x^2 + G_y^2} \quad (3)$$



Hình 2. Một kết quả phát hiện biên đối tượng dựa vào phép lọc Sobel

Một kết quả phát hiện biên đối tượng dựa trên phép lọc Sobel được mô tả trong Hình 2. Từ hình 2 cho thấy, phép lọc Sobel cho kết quả là ảnh có đường biên dày và không sắc nét. Vì vậy phương pháp phát hiện biên đối tượng ảnh LoG (Laplacian of Guassian) [3][10] cho ảnh màu được sử dụng bổ sung để làm rõ nét đường biên của đối tượng. Laplacian ∇^2 của một ảnh $f(x, y)$ được tính theo công thức:

$$\nabla^2 = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \quad (4)$$

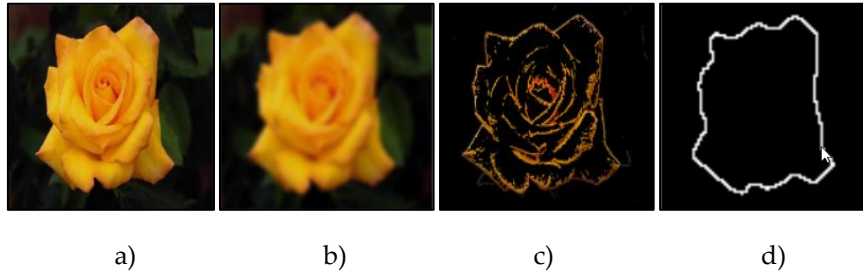
Kết hợp với bộ lọc Gaussian để làm mịn ảnh, nâng cao hiệu quả phát hiện biên. Hàm Gaussian được tính theo công thức:

$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5)$$

Trong đó σ là độ lệch chuẩn. Toán tử LoG được tính toán theo công thức:

$$\nabla^2 G(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (6)$$

Lúc này, ảnh sẽ bị mờ đi, mức độ mờ phụ thuộc vào độ lệch chuẩn σ . Trong phương pháp này, Gaussian cho phép nâng cao cường độ ảnh, khi kết hợp với Laplacian sẽ hiển thị vùng ảnh có cường độ thay đổi, do đó làm tăng hiệu quả phát hiện biên. Đồng thời, khi thực hiện phép lọc tần số cao với Gaussian, sẽ cho kết quả là ảnh đường nét của đối tượng. Hình 3 là một kết quả phát hiện biên đối tượng dựa trên phương pháp LoG. Hình 3 bao gồm: a) ảnh gốc; b) ảnh được làm mờ; c) đường nét ảnh theo phép lọc LoG; d) ảnh đường nét theo phép lọc tần suất cao. Từ kết quả này, các đặc trưng của đối tượng được trích xuất gồm chu vi, diện tích đối tượng, vị trí tương đối của các đường viền.



Hình 3. Một kết quả phát hiện biên dựa trên phương pháp LoG

3.4. Thực nghiệm trích xuất vec-tơ đa đặc trưng cho hình ảnh

Từ các lý thuyết đã đề xuất, một hệ trích xuất đặc trưng được xây dựng như trong Hình 4.



Hình 4. Hệ trích xuất đặc trưng hình ảnh

Mỗi ảnh được trích xuất một vec-tơ đa đặc trưng gồm 81 giá trị, cụ thể như sau:

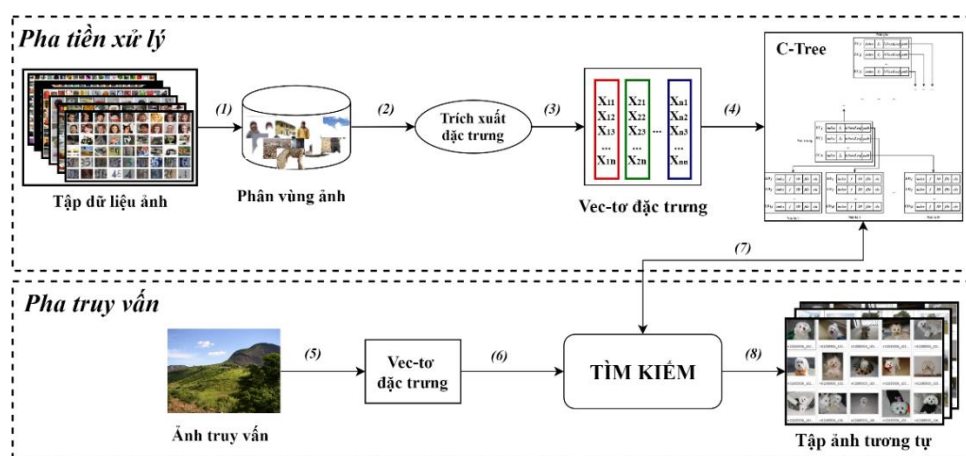
- Độ tương phản để thiết lập đối tượng và nền, và đặc trưng diện tích vùng ảnh phân đoạn (3 giá trị);
- Đặc trưng giá trị kỳ vọng theo trục X và Y của ảnh phân đoạn (2 giá trị);
- Tính đặc trưng độ lệch chuẩn theo trục X và Y của ảnh phân đoạn (2 giá trị);
- Lấy cấu trúc vân ảnh theo phép lọc Sobel và tính giá trị diện tích vùng theo vân ảnh (2 giá trị);
- Lấy đặc trưng cường độ các điểm ảnh theo lát giềng dựa vào phép lọc Sobel (1 giá trị);
- Tính giá trị kỳ vọng của vân ảnh theo trục X và Y (2 giá trị);
- Tính giá trị độ lệch chuẩn theo trục X và Y trên vân ảnh (2 giá trị);
- Lấy đặc trưng đường nét và tính chu vi đối tượng theo phương pháp LoG (2 giá trị);

- Tính giá trị kỳ vọng theo trục X và Y của đường biên đối tượng (2 giá trị);
- Tính giá trị độ lệch chuẩn theo trục X và Y của đường biên đối tượng (2 giá trị);
- Lấy đặc trưng nâng cao cường độ ảnh với phép lọc Gaussian (9 giá trị);
- Lấy đường nét ảnh với phép lọc tần số cao (9 giá trị);
- Lấy đặc trưng cường độ của đối tượng (9 giá trị);
- Lấy đặc trưng cường độ của hình nền (9 giá trị);
- Lấy đặc trưng màu sắc chủ đạo MPEG-7 (25 giá trị)

4. HỆ TRUY VẤN ẢNH DỰA TRÊN VEC-TƠ ĐẶC TRƯNG

4.1. Mô hình hệ truy vấn ảnh

Các vec-tơ đặc trưng ảnh sau khi được trích xuất, tạo thành một cơ sở dữ liệu các đặc trưng nội dung ảnh. Trong bài báo này, các vec-tơ được lưu trữ trên cấu trúc cây phân cụm cân bằng C-Tree[20], làm cơ sở cho việc tìm kiếm tập ảnh tương tự dựa trên khoảng cách Euclide. Mô hình hệ truy vấn ảnh trên cây C-Tree dựa vào các vec-tơ đặc trưng, gọi là SBIR_CT, được mô tả trong **Hình 5**.



Hình 5. Mô hình hệ truy vấn ảnh trên cây C-Tree

Hệ truy vấn bao gồm 2 pha: Pha tiền xử lý và pha truy vấn. **Pha tiền xử lý** thực hiện các thao tác như sau:

Bước 1. Từ tập dữ liệu ảnh, thực hiện phân vùng ảnh (1) để tạo các đối tượng ảnh;

Bước 2. Thực hiện trích xuất đặc trưng cho các phân vùng đối tượng (2) và kết hợp các đặc trưng để tạo thành vec-tơ đa đặc trưng f cho mỗi ảnh trong tập dữ liệu (3);

Bước 3. Huấn luyện các mẫu dữ liệu với cấu trúc cây phân cụm cân bằng C-Tree để lưu trữ cơ sở dữ liệu các vec-tơ đặc trưng.

Pha truy vấn thực hiện các thao tác như sau:

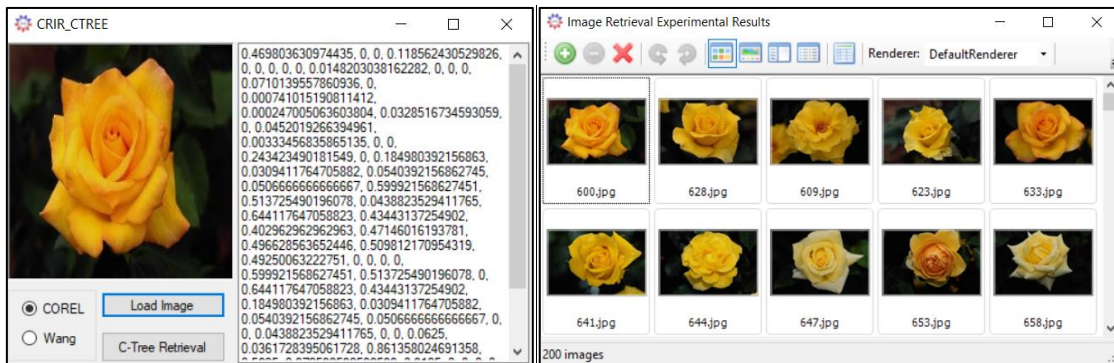
Bước 1. Mỗi ảnh truy vấn I_q đầu vào được trích xuất các vec-tơ đặc trưng f ;

Bước 2. Thực hiện so sánh f với cơ sở dữ liệu đặc trưng trên cây C-Tree để tìm nhánh có độ đo tương tự gần nhất và các nút lá phù hợp nhất dựa vào khoảng cách Euclid;

Bước 3. Kết quả là tập các ảnh tương tự được sắp xếp theo độ đo.

4.2. Thực nghiệm hệ truy vấn ảnh

Hệ truy vấn ảnh SBIR_CT thực nghiệm trên tập ảnh COREL (1000 ảnh) và WANG (10.800 ảnh). Thực nghiệm được xây dựng trên nền tảng dotNET Framework 4.8, ngôn ngữ lập trình C#. Các đồ thị được biểu diễn trên Matlab 2015. Cấu hình máy tính của thực nghiệm: Intel(R) CoreTM i7-8750H, CPU 2,70GHz, RAM 8GB và hệ điều hành Windows 10 Professional. Trong thực nghiệm này, việc truy vấn ảnh được thực hiện so sánh độ đo tương tự giữa hai hình ảnh dựa trên khoảng cách Euclide của vec-tơ đặc trưng. **Hình 6** là một kết quả tìm kiếm hệ truy vấn SBIR_CT, bao gồm: giao diện ảnh truy vấn cùng với vec-tơ của ảnh đó và tập các hình ảnh tương tự với ảnh truy vấn theo độ đo.



Hình 6. Một kết quả tìm kiếm của hệ truy vấn SBIR_CT

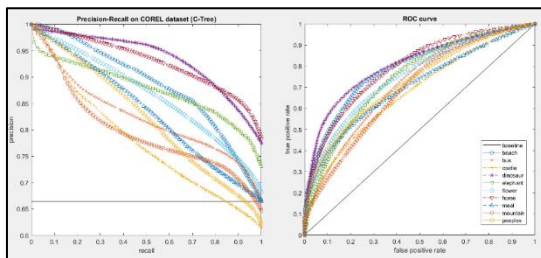
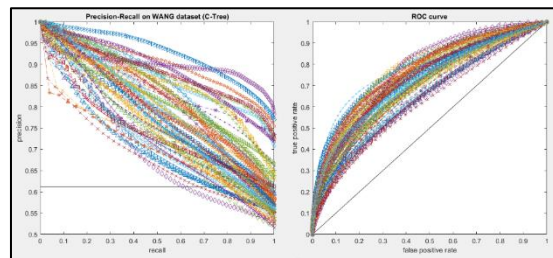
4.3. Đánh giá thực nghiệm

Để đánh giá hiệu quả tìm kiếm ảnh, luận án sử dụng các yếu tố để đánh giá bao gồm: precision, recall và F-measure, thời gian truy vấn (milli seconds). Trên cơ sở thực nghiệm, các giá trị hiệu suất và thời gian tìm kiếm trung bình của các thư mục ảnh trên tập ảnh COREL, WANG được tổng hợp trong **Bảng 1**.

Bảng 1. Hiệu suất tìm kiếm của hệ truy vấn SBIR_CT trên tập ảnh COREL, WANG

Giá trị hiệu suất	COREL	WANG
Avg. Precision	0.677655	0.6072217
Avg. Recall	0.699885	0.4891839
Avg. F-measure	0.688521	0.5450113
Avg. Query time (ms)	19.91437	39.746901

Từ Bảng 1 cho thấy, độ chính xác của hệ truy vấn SBIR_CT trên bộ ảnh COREL và WANG là khá cao, thời gian truy vấn trên cây C-Tree nhanh. Điều này chứng tỏ, các lý thuyết đề xuất trong bài báo về trích xuất đặc trưng thị giác là hiệu quả. Tập COREL với số lượng ít (1000 ảnh), nên cho độ chính xác, độ phủ cao hơn so với tập WANG (10.800 ảnh). Dựa trên các số liệu thực nghiệm, các đồ thị Precision-Recall curve và ROC curve được thực hiện để đánh giá độ chính xác của hệ truy vấn SBIR_CT. Hình 7, Hình 8 lần lượt là hiệu suất trên tập COREL và WANG. Trong đồ thị đường cong PR, mỗi đường cong là một thư mục của tập ảnh. Mỗi đường cong này cho thấy tỷ lệ giữa độ chính xác (precision) và độ phủ (recall) của một chủ đề ảnh; Đồng thời, đường cong tương ứng trong đồ thị ROC cho biết tỷ lệ kết quả truy vấn đúng và sai, nghĩa là diện tích dưới đường cong này đánh giá được tính đúng đắn của các kết quả truy vấn.

**Hình 7.** Hiệu suất tìm kiếm dựa trên cây C-Tree của tập ảnh COREL**Hình 8.** Hiệu suất tìm kiếm dựa trên cây C-Tree của tập ảnh WANG**Bảng 2.** So sánh độ chính xác giữa các phương pháp trên tập ảnh COREL

Phương pháp	Độ chính xác trung bình
A. Huneiti, 2015 [18]	0.559
Garg, M., 2019 [17]	0.602
Bella M. I. T., 2019 [5]	0.658
SBIR_CT	0.6777

Bảng 3. So sánh độ chính xác giữa các phương pháp trên tập ảnh WANG

Phương pháp	Độ chính xác trung bình
Dos Santos, 2015 [15]	0.570
R. Das, 2017 [12]	0.559
P. Chhabra, 2018 [11]	0.577
SBIR_CT	0.6072

Bảng 4. So sánh thời gian tìm kiếm trên tập ảnh COREL

Phương pháp	Thời gian tìm kiếm trung bình (ms)
Bibi R., 2019 [6]	512.675
Pavithra L. K., 2018 [21]	657.568
Zhou J., 2019 [1]	783.712
SBIR_CT	19.91437

Để đánh giá độ chính xác và hiệu quả của hệ truy vấn SBIR_CT, chúng tôi so sánh hiệu suất thu được từ thực nghiệm với các công trình nghiên cứu khác trên cùng tập dữ liệu ảnh. **Bảng 2**, **Bảng 3** lần lượt là kết quả so sánh độ chính xác trung bình của tập ảnh COREL và WANG với kết quả của các phương pháp khác. Đồng thời, **Bảng 4** là kết quả so sánh về thời gian tìm kiếm ảnh (milli giây) trên tập ảnh COREL.

Qua số liệu của các bảng trên, cho thấy phương pháp tìm kiếm ảnh dựa trên cây C-Tree với vec-tơ đa đặc trưng được đề xuất trong bài báo có độ chính xác vượt trội so với các phương pháp khác trên cùng tập ảnh, thời gian tìm kiếm nhanh (bộ COREL). Điều này chứng tỏ, phương pháp đề xuất của chúng tôi là đúng đắn và hiệu quả trong bài toán về tìm kiếm ảnh.

5. KẾT LUẬN

Trong bài báo này, phương pháp trích xuất vec-tơ đa đặc trưng của hình ảnh được đề xuất với các kỹ thuật trích xuất đặc trưng màu trội MPEG-7, kỹ thuật phát hiện biên với LoG, phép lọc Sobel, nâng cao cường độ ảnh với Gaussian... Mỗi hình ảnh trong tập dữ liệu được trích xuất thành một vec-tơ đặc trưng, tạo thành cơ sở dữ liệu đặc trưng, và lưu trữ trên cây C-Tree cho bài toán tìm kiếm ảnh. Thực nghiệm về trích xuất đặc trưng và tìm kiếm ảnh được thực hiện trên tập ảnh COREL, WANG. Kết quả thực nghiệm được so sánh với các phương pháp khác trên cùng tập ảnh cho thấy, các đề xuất trong bài báo về trích xuất đặc trưng áp dụng cho bài toán tìm kiếm ảnh trên cây C-Tree là hiệu quả: có độ chính xác cao, thời gian tìm kiếm nhanh. Hướng phát triển tiếp theo

của bài báo này là từ các đặc trưng cấp thấp, xác định ngữ nghĩa của hình ảnh nhằm đáp ứng được yêu cầu của người dùng.

TÀI LIỆU THAM KHẢO

- [1] Ashraf, R. B. (2015). Content based image retrieval using embedded neural networks with bandletized regions. *Entropy*, 3552-3580.
- [2] Babu, S. A. (2014). Improving Quality of Content Based Image Retrieval with Graph Based Ranking. *International Journal of Research and Applications*, 1(1), 2349-0020.
- [3] Bagri, N. &. (2015). A comparative study on feature extraction using texture and shape for content based image retrieval. *International Journal of Advanced Science and Technology*, 41-52.
- [4] Barrios, J. M.-E. (2009). Text-based and content-based image retrieval on Flickr. *Second International Workshop on Similarity Search and Applications*. IEEE.
- [5] Bella, M. I. (2019). An efficient image retrieval framework using fused information feature. *Computers & Electrical Engineering*, 75, 46-60.
- [6] Bibi, R. M. (2020). Query-by-visual-search: multimodal framework for content-based image retrieval. *Journal of Ambient Intelligence and Humanized Computing*, 11(11), 5629-5648.
- [7] Bora, D. J. (2017). A novel approach for color image edge detection using multidirectional Sobel filter on HSV color space. *Int. J. Comput. Sci. Eng*, 5(2), 154-159.
- [8] Chaki, J. &. (2018). *A Beginner's Guide to Image Preprocessing Techniques*. CRC Press.
- [9] Chaki, J. &. (2021). *Image Color Feature Extraction Techniques: Fundamentals and Applications*. Singapore: Springer.
- [10] Chaki, J. &. (n.d.). *A beginner's guide to image shape feature extraction techniques*. 2019: CRC Press.
- [11] Chhabra, P. G. (2020). Content-based image retrieval system using ORB and SIFT features. *Neural Computing and Applications*, 32(7), 2725-2733.
- [12] Das, R. T. (2017). Novel feature extraction technique for content-based image recognition with query classification. *International Journal of Computational Vision and Robotics*, 7(1-2), 123-147.
- [13] Deloitte. (2016). *Photo sharing: trillions and rising*. Deloitte Touche Tohmatsu Limited.
- [14] Dharani, T. &. (2013). A survey on content based image retrieval. *International Conference on Pattern Recognition, Informatics and Mobile Engineering*. IEEE.
- [15] Dos, S. J. (2015). A signature-based bag of visual words method for image indexing and search. *Pattern Recognition Letters*, 65, 1-7.
- [16] Erwin, M. F. (2017). Content Based Image Retrieval for Multi-Objects Fruits Recognition using k-Means and k-Nearest Neighbor. *International Conference on Data and Software Engineering*.
- [17] Garg, M. S. (2019). Fuzzy-NN approach with statistical features for description and classification of efficient image retrieval. *Modern Physics Letters A*, 34(03)(1950022).

- [18] Huneiti, A. &. (2015). Content-based image retrieval using SOM and DWT. *Journal of software Engineering and Applications*, 8(02)(51).
- [19] K., A. M. (2020). Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features. *Arabian Journal for Science and Engineering*, 82-94.
- [20] Nhi, N. T. (2020). A SELF-BALANCED CLUSTERING TREE FOR SEMANTIC-BASED IMAGE RETRIEVAL. *Journal of Computer Science and Cybernetics*, 49-67.
- [21] Pavithra, L. K. (2019). An efficient seed points selection approach in dominant color descriptors (DCD). *Cluster Computing*, 22(4), 788–795.
- [22] Shao, H. W. (2008). Image retrieval based on MPEG-7 dominant color descriptor. *The 9th International Conference for Young Computer Scientists* (pp. 753-757). IEEE.
- [23] Singha, M. &. (2012). Content based image retrieval using color and texture. *Signal & Image Processing*.
- [24] Van, T. T. (2018). Content-based image retrieval based on binary signatures cluster graph. *Expert Systems*, 35(1), e12220.
- [25] Vinayak, V. &. (2017). CBIR system using color moment and color auto-Correlogram with block truncation coding. *International Journal of Computer Applications*, 161(9), 1-7.
- [26] Zhou, J. L. (2019). Image retrieval based on effective feature extraction and diffusion process. *Multimedia Tools and Applications*, 78(5), 6163-6190.

A FEATURE EXTRACTION METHOD FOR IMAGE RETRIEVAL

Nguyen Thi Uyen Nhi^{1,2}, Van The Thanh³

¹Faculty of Information Technology, University of Sciences, Hue University

²Faculty of Statistics and Informatics, University of Economics, The University of Danang

³HCMC University of Food Industry

Email: nhintu@due.edu.vn, thanhvt@hufi.edu.vn

ABSTRACT

In this paper, a feature extraction method for image retrieval is proposed. Feature extraction techniques include: segmentation of an image based on contrast; the MPEG-7 dominant color descriptor; edge-based object recognition and surface smoothing based on Sobel filtering; detecting object boundary with LoG (Laplace of Gaussian) method; enhancing image intensity with Gaussian filtering ... Hence, a method for image retrieval based on C-Tree [20] was implemented to find a set of similar images. To demonstrate the effectiveness of the proposals, the experiment was built on the COREL-1K, WANG image set. At the same time, the experimental results are used to evaluate the effectiveness of the proposed method. The experimental results show that our method of feature extraction image and similar image retrieval on C-Tree is effective.

Keywords: C-Tree, feature extraction, image retrieval.



Nguyễn Thị Uyên Nhi sinh năm 1985. Bà nhận bằng cử nhân và thạc sĩ chuyên ngành Khoa học máy tính và kỹ thuật tính toán, tại trường Đại học tổng hợp kỹ thuật Volgograd, Liên bang Nga, lần lượt vào các năm 2008, 2010. Từ năm 2017, bà học NCS ngành Khoa học máy tính tại Trường Đại học Khoa học, Đại học Huế. Hiện nay, bà công tác tại khoa Thống kê – Tin học, Trường Đại học Kinh Tế (DUE), Đại học Đà Nẵng.

Lĩnh vực nghiên cứu: xử lý ảnh, tìm kiếm ảnh, cơ sở dữ liệu.



Văn Thế Thành sinh năm 1979. Ông tốt nghiệp đại học chuyên ngành Toán tin tại Đại học Khoa học Tự nhiên - Đại học Quốc gia TP.HCM vào năm 2001, nhận bằng Thạc sĩ Khoa học Máy tính tại Đại học Quốc gia TP.HCM vào năm 2008. Năm 2016, ông nhận bằng Tiến sĩ Khoa học Máy tính tại Trường Đại học Khoa học, Đại học Huế. Hiện nay, ông công tác tại Trường ĐH Công nghiệp Thực phẩm TP.HCM.

Lĩnh vực nghiên cứu: xử lý ảnh, khai thác dữ liệu ảnh và tìm kiếm ảnh.