

Nghiên cứu cải tiến một số phương pháp tra cứu ảnh sử dụng đặc trưng ảnh : Luận án TS. Công nghệ thông tin: 62 48 01 01 / Nguyễn Hữu Quỳnh ; Nghđ. : PGS.TS. Ngô Quốc Tạo, PGS.TS. Đinh Mạnh Tường

1. Tính cấp thiết của luận án

Một lượng lớn thông tin ảnh đã được đưa lên Internet. Tuy nhiên, không thể truy cập hoặc sử dụng thông tin trong các tập ảnh khổng lồ này, nếu chúng không được tổ chức để tra cứu hiệu quả trên toàn bộ dữ liệu ảnh.

Các kỹ thuật dựa vào văn bản mô tả ảnh tốn nhiều thời gian, chi phí cao và phụ thuộc vào cảm nhận chủ quan của chuyên viên kỹ thuật. Hơn nữa, hệ thống dựa vào từ khoá rất khó thay đổi về sau này.

Để khắc phục các khó khăn này, tra cứu ảnh dựa vào đặc trưng thị giác của ảnh đã được đề xuất. Ý tưởng cơ bản của cách tiếp cận này là sử dụng kỹ thuật trích rút đặc trưng thị giác một cách tự động để cho ra các mô tả nội dung ảnh một cách trực tiếp từ chính bản thân ảnh.

Hầu hết các phương pháp đã được đề xuất sử dụng đặc trưng màu đều gặp phải vấn đề về chi phí không gian lưu trữ số các lược đồ màu biểu diễn ảnh lớn, độ chính xác tra cứu không cao, độ phức tạp tính toán lớn, nhạy cảm với quay và dịch chuyển.

Do đó, luận án chọn đề tài “**Nghiên cứu cải tiến một số phương pháp tra cứu ảnh sử dụng đặc trưng ảnh**” để góp phần giải quyết các vấn đề đặt ra.

2. Mục tiêu của luận án

Mục đích của luận án là nghiên cứu đề xuất một số phương pháp tra cứu ảnh sử dụng đặc trưng màu và thông tin không gian. Các phương pháp này sẽ hướng tới giải quyết các vấn đề về giảm không gian lưu trữ số các lược đồ màu biểu diễn ảnh, ít nhạy cảm với quay và dịch chuyển, giảm độ phức tạp tính toán và tăng độ chính xác tra cứu.

3. Các đóng góp của luận án

Đề xuất các kỹ thuật bao gồm: phương pháp HG, phương pháp IHG, phương pháp CSI, phương pháp CCS và hệ thống tra cứu ảnh dựa vào đặc trưng thị giác LVFIR.

4. Bố cục của luận án

Luận án này được bố cục thành bốn chương.

Chương 1 giới thiệu tổng quan về trích rút đặc trưng và tra cứu ảnh dựa vào đặc trưng.

Chương 2 trình bày kỹ thuật tra cứu ảnh dựa vào lược đồ màu khối.

Chương 3 trình bày phương pháp tra cứu dựa vào vùng ảnh.

Chương 4 trình bày thiết kế và thực hiện hệ thống tra cứu ảnh LVFIR, cùng với một số kết quả.

Cuối cùng, chúng tôi đưa ra một số kết luận và đề xuất các nghiên cứu tương lai.

Chương 1. TỔNG QUAN VỀ TRÍCH RÚT ĐẶC TRƯNG VÀ TRA CỨU ẢNH DỰA VÀO ĐẶC TRƯNG

1.1 Các đặc trưng

Các đặc trưng ảnh có thể được phân thành đặc trưng thị giác và đặc trưng ngữ nghĩa. Đặc trưng thị giác có thể được phân loại tiếp thành đặc trưng chung và đặc trưng theo lĩnh vực.

1.1.1 Các đặc trưng toàn cục và cục bộ

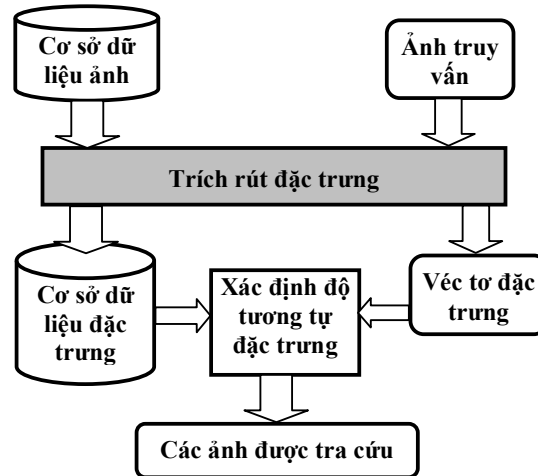
Các đặc trưng biểu diễn nội dung thị giác của toàn bộ ảnh được gọi là các đặc trưng toàn cục. Các đặc trưng biểu diễn nội dung thị giác của một phần của ảnh được gọi là đặc trưng cục bộ.

1.1.2 Các đặc trưng thị giác trong tra cứu ảnh dựa vào đặc trưng

Các đặc trưng thị giác bao gồm: Đặc trưng màu, đặc trưng kết cấu, đặc trưng hình dạng.

1.2 Kiến trúc của một hệ thống tra cứu ảnh dựa vào đặc trưng thị giác

Kiến trúc hệ thống tra cứu ảnh dựa vào đặc trưng thị giác được chỉ ra như Hình 1.1.



Hình 1.1. Kiến trúc hệ thống tra cứu ảnh dựa vào đặc trưng thị giác.

1.3 Trích rút đặc trưng

Trước khi đề cập đến đặc trưng màu, chúng tôi giới thiệu khái niệm về dải của lược đồ màu.

Định nghĩa 1.1 [Dải của lược đồ màu]:

Một dải của lược đồ màu là số điểm ảnh trong một diện tích ảnh được chỉ ra mà có chung màu.

Định nghĩa 1.2 [Khối ảnh]:

Một khối ảnh là một vùng ảnh hình chữ nhật trong ảnh.

1.3.1 Đặc trưng màu

Đặc trưng màu được sử dụng rất hiệu quả cho tra cứu các ảnh màu trong cơ sở dữ liệu ảnh. Các mô tả màu được trích rút và so sánh thuận lợi, do đó đặc trưng thích hợp cho tra cứu dựa vào đặc trưng thị giác.

1.3.2 Lượng hóa màu

Lượng hoá màu là quá trình giảm số các màu được sử dụng để biểu diễn một ảnh.

1.3.3 Biểu diễn màu

1.3.3.1 Lược đồ màu

Lược đồ màu biểu thị phân bố của số các điểm ảnh cho mỗi màu được lượng hóa. Lược đồ màu được tính toán dễ dàng và hiệu quả trong mô tả phân bố màu toàn cục và cục bộ trong ảnh.

1.3.3.2 Lược đồ màu toàn cục GCH

Sử dụng lược đồ màu toàn cục (GCH), một ảnh sẽ được mã hoá với lược đồ màu của nó, và khoảng cách giữa hai ảnh sẽ được xác định bởi khoảng cách giữa hai lược đồ màu này.

1.3.3.3 Lược đồ màu cục bộ LCH

Phương pháp LCH gồm thông tin liên quan đến phân bố màu của các vùng. Khi so sánh hai ảnh, chúng ta tính toán khoảng cách giữa lược đồ của một khối trong một ảnh và một khối ở cùng vị trí trong ảnh kia. Khoảng cách giữa hai ảnh sẽ được xác định bởi tổng tất cả các khoảng cách này.

1.3.3.4 Véc tơ gắn kết màu

Véc tơ gắn kết màu liên kết thông tin không gian vào lược đồ màu, mỗi dải của lược đồ màu được phân thành hai loại: gắn kết, nếu điểm ảnh thuộc về một vùng màu đồng nhất lớn và không gắn kết, nếu điểm ảnh không thuộc về một vùng màu đồng nhất lớn.

1.3.3.5 Tương quan màu

Tương quan màu mô tả phân bố màu của các điểm ảnh và chỉ ra tương quan không gian của các cặp màu.

1.3.3.6 Các màu trội

Các màu trội được sử dụng để mô tả đặc trưng màu của một ảnh. Phân cụm màu được thực hiện để thu các màu trội đại diện.

1.3.3.7 Các mô men màu

Mô men màu là các mô men thống kê của các phân bố xác suất của các màu.

1.3.4 Thông tin không gian

Thông tin không gian biểu thị vị trí không gian tuyệt đối và vị trí không gian tương đối của các vùng. Các vùng hoặc đối tượng với các đặc trưng màu tương tự có thể được phân biệt tốt hơn bằng việc tận dụng các thông tin không gian.

1.3.5 Phân vùng

Phân vùng là quá trình phân ảnh thành các vùng, trong trường hợp tốt nhất chúng ta sẽ thu được các đối tượng xuất hiện trong ảnh.

1.4 Các độ đo tương tự

Một số độ đo tương tự được sử dụng phổ biến nhất: Lược đồ giao, Khoảng cách L_1 , Khoảng cách dạng toàn phương, Khoảng cách EMD, Khoảng cách Kolmogorov-Smirnov,...

1.5 Đánh giá hiệu năng tra cứu

Để đánh giá một ứng dụng tra cứu ảnh, một cơ sở dữ liệu ảnh và một tập các truy vấn được yêu cầu. Các truy vấn được thực hiện với ứng dụng VFBIR để thu được các kết quả tra cứu. Sau đó phương pháp đánh giá hiệu năng được sử dụng để so sánh các kết quả được tra cứu này với các ảnh liên quan đến ảnh truy vấn trong cơ sở dữ liệu.

1.6 Các hệ thống VFBIR

Một số hệ thống tra cứu ảnh đã được xây dựng gồm: QBIC, Blobworld, RetrievalWare, VisualSeek và WebSeek, CIRES, Tìm kiếm ảnh của Google, ...

1.7 Kết luận và định hướng nghiên cứu

Trong chương này, chúng tôi đã giới thiệu một số khái niệm và kỹ thuật cơ bản về trích rút đặc trưng và tra cứu ảnh dựa vào đặc trưng thị giác. Đặc biệt chúng tôi tập trung vào trích rút và biểu diễn đặc trưng thị giác.

Đặc trưng thị giác được sử dụng phổ biến nhất là màu. Do màu cho phép cảm nhận và phân biệt ảnh rất hiệu quả. Hơn nữa, đặc trưng màu là tương đối ổn định với các biến dạng nhỏ và độc lập với hướng và cỡ của ảnh.

Thông tin màu thường được biểu diễn bởi lược đồ màu trong một không gian màu nào đó. Lược đồ màu có ưu điểm là được tính toán nhanh và không nhạy cảm với các thay đổi nhỏ về vị trí thu nhận ảnh. Tuy nhiên, lược đồ màu là một mô tả thô của ảnh nên hai ảnh rất

khác nhau có thể có các lược đồ màu tương tự. Hơn nữa, hai ảnh chỉ tương tự nếu chúng có các vùng màu tương tự tại những vị trí tương tự. Vì lý do này mà việc kết hợp đặc trưng màu với thông tin không gian để cải thiện hiệu năng tra cứu là cần thiết.

Trong luận án này chúng tôi sẽ tập trung vào vấn đề nâng cao hiệu năng hệ thống tra cứu ảnh dựa vào đặc trưng thị giác thông qua sử dụng đặc trưng của vùng ảnh:

Thứ nhất, chúng tôi sẽ đề xuất phương pháp sử dụng ít chi phí không gian lưu trữ các lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển.

Thứ hai, chúng tôi sẽ đề xuất phương pháp sử dụng đặc trưng của vùng ảnh vào trong quá trình tra cứu nhằm nâng cao hiệu năng tra cứu.

Chương 2. PHƯƠNG PHÁP TRA CỨU ẢNH DỰA VÀO LƯỢC ĐỒ MÀU KHỐI

2.1 Lược đồ màu khối

Dưới đây là mô tả cách tiếp cận lược đồ màu khối:

Với ảnh được lượng hoá thành C màu (trong không gian màu RGB) và ảnh được chia thành $m \times m$ khối ảnh có kích thước bằng nhau. Một lược đồ màu khối theo màu c ($0 < c \leq C$) là một tập $m \times m$ dải. Ở đây dải của lược đồ màu khối là số điểm ảnh trong một khối ảnh mà có chung màu và các giá trị dải được mô tả bởi hàm $p(b_k) = n_k / n$, với b_k là khối ảnh thứ k của ảnh ($0 < k \leq m \times m$), n_k là số các điểm ảnh có màu c trong khối b_k và n là tổng số các điểm ảnh trong ảnh.

2.2 Phương pháp tra cứu dựa vào lược đồ màu khối

2.2.1 Giới thiệu

GCH có ưu điểm là bất biến với quay và tỷ lệ và tính toán rất đơn giản. Tuy nhiên, GCH không bao gồm vị trí không gian của các màu trong ảnh.

Phương pháp LCH đưa thông tin không gian vào bản miêu tả ảnh. Tuy nhiên, phương pháp này sử dụng nhiều không gian để lưu trữ số các lược đồ màu biểu diễn ảnh và có độ phức tạp tính toán lớn.

Phương pháp CCH sử dụng ít không gian lưu trữ số các lược đồ màu biểu diễn ảnh. Tuy nhiên, phương pháp này không có khả năng xử lý đối với các biến đổi hình học như quay và dịch chuyển, do CCH chỉ so sánh mỗi khối ảnh của ảnh truy vấn với khối ảnh cùng màu và cùng vị trí trong ảnh cơ sở dữ liệu.

Để khắc phục nhược điểm trên, chúng tôi đề xuất phương pháp HG.

2.2.2 Phương pháp tra cứu đề xuất HG

Trong phần này chúng tôi trình bày phương pháp HG. Phương pháp này của chúng tôi đã được công bố trên tạp chí quốc tế IJCSSES.

2.2.2.1 Khái niệm về đồ thị hai phía

Định nghĩa 2.1 [Đồ thị]:

$G(N, E)$ được gọi là đồ thị vô hướng với N là tập đỉnh và E là tập cạnh. Nếu nó thỏa mãn: $E \subset N \times N$ (E là tập con của tích đề các $N \times N$)

Định nghĩa 2.2 [Đồ thị vô hướng có trọng số]:

$G(N, E)$ là đồ thị vô hướng mà mỗi cạnh của nó được gán một trọng số không âm.

Định nghĩa 2.3 [Đồ thị hai phía]: Đồ thị hai phía là đồ thị vô hướng $G(N, E)$ mà có thể tách N thành hai tập X và Y thỏa mãn các điều kiện sau:

- $N = X \cup Y$ và $X \cap Y = \emptyset$
- $X \times X \cap E = \emptyset$ và $Y \times Y \cap E = \emptyset$

Trong trường hợp đặc biệt ta ký hiệu $G(X, Y, E)$ là đồ thị hai phía.

Định nghĩa 2.4 [Đồ thị hai phía có trọng số]:

Đồ thị hai phía có trọng số $G(X, Y, E)$ là đồ thị hai phía mà mỗi cạnh của nó được gán một giá trị không âm.

Định nghĩa 2.5 [Đối sánh của đồ thị]:

Đối sánh M của đồ thị $G(X, Y, E)$ là một tập con các cạnh mà trong M không có hai cạnh nào có đỉnh chung.

Định nghĩa 2.6 [Giá trị của một đối sánh]:

Giá trị của một đối sánh trong đồ thị hai phía $G(X, Y, E)$ có trọng số được đánh giá bằng tổng các trọng số của các cạnh trong đối sánh.

Định nghĩa 2.7 [Giá trị đối sánh cực tiểu]:

Giá trị đối sánh cực tiểu là giá trị đối sánh nhỏ nhất trong tất cả các đối sánh có thể có của đồ thị hai phía có trọng số $G(X, Y, E)$.

2.2.2.2 Phương pháp HG

Ý tưởng của phương pháp HG:

Phương pháp tính lược đồ màu khối đối với mỗi màu của ảnh truy vấn và ảnh CSDL. Sau đó, tính khoảng cách của ảnh truy vấn và ảnh CSDL theo mỗi màu thông qua đồ thị hai phía có trọng số. Trong đồ thị này, mỗi đỉnh ở phía bên trái của đồ thị là một dải của lược đồ màu khối theo màu của ảnh truy vấn, mỗi đỉnh ở phía bên phải của đồ thị là một dải của lược đồ màu khối có màu tương ứng của ảnh CSDL. Cuối cùng, tính tổng khoảng cách của ảnh truy vấn và ảnh CSDL theo tất cả các màu và giá trị này được coi là khoảng cách giữa hai ảnh.

Nội dung của thuật toán HG:

Tiếp theo, chúng tôi mô tả chi tiết thuật toán HG trả lại khoảng cách của hai ảnh I_1 và I_2 .

Thuật toán HG(I_1, I_2, n)

Vào: ảnh I_1 và I_2 với cỡ $n \times n$ khối ảnh

Ra: D - khoảng cách giữa hai ảnh I_1 và I_2

1. For mỗi c_1 in C_1 do

 Tính $H(I_1, c_1, n)$

2. For mỗi c_2 in C_2 do

 2.1 Tính $H(I_2, c_2, n)$

3. For mỗi c in C do

 3.1 Xây dựng đồ thị $G(X, Y, E, c)$ gồm $2n^2$ đỉnh

 3.2 $D \leftarrow D + \text{MCM}(G(X, Y, E, c), n)$

4. Trả lại giá trị D

Trong thuật toán HG ở trên, tham số C_1 là số màu của ảnh I_1 , C_2 là số màu của ảnh I_2 và C là số màu của hai ảnh I_1 và I_2 . $H(I_1, c_1, n)$ là lược đồ màu khối theo màu c_1 của ảnh I_1 gồm $n \times n$ dải. $H(I_2, c_2, n)$ là lược đồ màu khối theo màu c_2 của ảnh I_2 gồm $n \times n$ dải. $G(X, Y, E, c)$ là đồ thị gồm $2n^2$ đỉnh, trong đó $n \times n$ dải của lược đồ màu khối $H(I_1, c, n)$ và $n \times n$ dải của lược đồ màu khối $H(I_2, c, n)$. Hàm $\text{MCM}(,)$ trả lại khoảng cách giữa hai ảnh theo màu c đã cho.

Trong thuật toán HG, chúng tôi có sử dụng hàm MCM. Hàm này được mô tả như sau:

Hàm MCM ($G(X, Y, E, c), n$)

Vào: $G(X, Y, E, c)$ - đồ thị theo màu c của ảnh I_1 và I_2 gồm $2n^2$ đỉnh

Ra: $costc$ - khoảng cách theo màu c giữa ảnh I_1 và I_2

1. For $i \leftarrow 1$ to $n \times n$ do
 For $j \leftarrow 1$ to $n \times n$ do
 $w(i,j) \leftarrow |h_{I_1}[c][i] - h_{I_2}[c][j]|$
2. $M \leftarrow$ Giá trị đối sánh cực tiểu của $G(X, Y, E, c)$
3. For mỗi $(i,j) \in M$ do
 $costc \leftarrow costc + w(i,j)$
4. Trả lại giá trị $costc$

Đầu tiên, hàm MCM tính trọng số của tất cả các cạnh trong đồ thị $G(X,Y,E,c)$. Sau đó, hàm tìm giá trị đối sánh cực tiểu của đồ thị $G(X, Y, E, c)$. Cuối cùng, hàm tính tổng giá trị các cạnh thuộc đối sánh và khoảng cách này là khoảng cách của ảnh I_1 và I_2 theo màu c .

Độ phức tạp của thuật toán HG:

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của hàm MCM thông qua Mệnh đề 2.1.

Mệnh đề 2.1 [Độ phức tạp của hàm MCM]: Độ phức tạp của hàm MCM là $O(n^4)$, với n^2 là số dải của lược đồ màu khối của ảnh,

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của thuật toán HG thông qua Mệnh đề 2.2.

Mệnh đề 2.2 [Độ phức tạp của thuật toán HG]: Độ phức tạp của thuật toán HG là $O(n^4)$ với n^2 là số dải của lược đồ màu khối của ảnh,

Dưới đây chúng tôi sẽ chỉ ra độ nhạy cảm với phép quay của phương pháp HG thông qua Mệnh đề 2.3.

Mệnh đề 2.3 [Độ nhạy cảm với phép quay của phương pháp HG]: Phương pháp HG ít nhạy cảm với phép quay của ảnh hơn phương pháp CCH.

Dưới đây chúng tôi sẽ chỉ ra độ nhạy cảm với phép dịch chuyển của phương pháp HG thông qua Mệnh đề 2.4.

Mệnh đề 2.4 [Độ nhạy cảm với phép dịch chuyển của phương pháp HG]: Phương pháp HG ít nhạy cảm với phép dịch chuyển của ảnh hơn phương pháp CCH.

2.3 Phương pháp cải tiến IHG

Trong phần này chúng tôi trình bày phương pháp cải tiến IHG. Phương pháp này của chúng tôi đã được công bố trên tạp chí quốc tế IJCSSES.

2.3.1 Khái niệm về sự tương tự lý tưởng giữa hai dải

Định nghĩa 2.8 [Sự tương tự lý tưởng giữa hai dải của hai lược đồ màu khối]: Hai dải của hai lược đồ màu khối được gọi là tương tự lý tưởng nếu chúng thỏa mãn cả hai điều kiện:

- Khoảng cách chấp nhận được: Khoảng cách giữa một dải của lược đồ màu khối của ảnh truy vấn với một dải của lược đồ màu khối của ảnh CSDL phải không quá lớn.
- Vị trí thích hợp: mỗi dải của lược đồ màu khối tương ứng với mỗi khối ảnh thuộc cạnh hình vuông trong ảnh truy vấn chỉ được so sánh với mỗi dải của lược đồ màu khối tương ứng với mỗi khối ảnh thuộc cạnh hình vuông tương ứng của ảnh CSDL.

2.3.2 Lý do đề xuất phương pháp IHG

Khi phương pháp HG sử dụng giá trị đối sánh cực tiểu để tính toán khoảng cách giữa hai ảnh theo màu c sẽ gặp phải các vấn đề sau:

- ❖ Ảnh hưởng nhiều: Các giá trị đối sánh có thể chứa các cạnh có giá trị rất lớn. Trong trường hợp này, các dải của lược đồ màu khối tương ứng sẽ rất khác nhau. Điều đó sẽ làm tăng các giá trị nhiễu và ảnh hưởng đến khoảng cách cuối cùng giữa hai ảnh.
- ❖ Tốn thời gian: Các dải của lược đồ màu khối có vị trí không thích hợp xuất hiện với tần xuất nhiễu sẽ ảnh hưởng đáng kể đến thời gian so sánh của phương pháp cũng như tăng giá trị nhiễu vào khoảng cách cuối cùng.

Vì lý do này, một vấn đề đã xuất hiện *các dải của hai lược đồ màu khối có điều kiện như thế nào có thể được sử dụng hiệu quả cho so sánh của phương pháp HG?*

Qua quan sát ảnh bị quay hoặc dịch chuyển, chúng tôi nhận thấy rằng, vị trí của các khối ảnh sẽ thay đổi, nhưng các khối ảnh này vẫn nằm trên cạnh hình vuông xác định. Điều đó nói lên rằng chúng ta chỉ cần so sánh một dải của lược đồ màu khối của ảnh truy vấn với các dải của lược đồ màu khối tương ứng với các khối thuộc cạnh hình vuông tương ứng của ảnh CSDL. Nói cách khác, chúng ta chỉ so sánh một dải của lược đồ màu khối của ảnh truy vấn với mỗi dải của lược đồ màu khối của ảnh CSDL nếu chúng là *tương tự lý tưởng*.

Việc xác định các dải của hai lược đồ màu khối *tương tự lý tưởng* sẽ ảnh hưởng đến độ chính xác và thời gian so sánh của phương pháp HG. Chúng tôi đề xuất phương pháp HG cải tiến, có tên là IHG [44].

2.3.3 Phương pháp IHG

Ý tưởng cơ bản của phương pháp IHG:

Chỉ so sánh mỗi dải của lược đồ màu khối tương ứng với các khối thuộc cạnh hình vuông của ảnh truy vấn với mỗi dải của lược đồ màu khối tương ứng với các khối thuộc cạnh hình vuông tương ứng có cùng màu của ảnh CSDL.

Nội dung của thuật toán IHG:

Tiếp theo, chúng tôi trình bày chi tiết thuật toán IHG(). Thuật toán này trả lại khoảng cách giữa hai ảnh I_1 và I_2 với kích thước $n \times n$ khối ảnh.

Thuật toán IHG(I_1, I_2, n)

Vào : I_1, I_2 – hai ảnh với cỡ $n \times n$ khối ảnh

Ra : dis - khoảng cách giữa ảnh I_1 và I_2

1. $dis \leftarrow 0$
2. $C \leftarrow \text{Quantization}(I_1, I_2)$
3. For mỗi c in C do
 - 3.1 $dis \leftarrow dis + \text{DistancebyColor}(I_1, I_2, n, c)$
4. Trả lại giá trị dis

Trong thuật toán IHG(), chúng tôi có sử dụng hàm DistancebyColor(). Hàm này có đầu vào là hai ảnh I_1 và I_2 , kích thước $n \times n$ khối ảnh của ảnh và màu c . Sau đó hàm trả về khoảng cách giữa hai ảnh I_1 và I_2 theo màu c .

Hàm DistancebyColor (I_1, I_2, n, c)

Vào: hai ảnh I_1 và I_2 với cỡ $n \times n$ khối ảnh, màu c .

Ra: $distc$ - khoảng cách giữa hai ảnh I_1 và I_2 theo màu c .

1. $r \leftarrow 0$
2. while ($r < n$)
 - 2.1 $distc \leftarrow 0$
 - 2.2 $r \leftarrow r + 2$
 - 2.3 for $i \leftarrow \text{sqr}(r-2)+1$ to $\text{sqr}(r)$ do
 - 2.3.1 for $j \leftarrow \text{sqr}(r-2)+1$ to $\text{sqr}(r)$ do
$$w(i,j) \leftarrow |h_{I_1}[c][i] - h_{I_2}[c][j]|$$
 - 2.4 $distc \leftarrow distc + \text{EdgeDistance}(G(X, Y, E, c), r)$
3. Trả lại giá trị $distc$

Trong hàm DistancebyColor(), chúng tôi sử dụng hàm EdgeDistance(). Hàm EdgeDistance() tính khoảng cách các dải của lược đồ màu khối tương ứng với các khối ảnh thuộc cạnh hình vuông kích thước r của ảnh I_1 và các dải của lược đồ màu khối tương ứng với các khối ảnh thuộc cạnh hình vuông kích thước r của ảnh I_2 theo màu c .

Hàm EdgeDistance($G(X, Y, E, c), r$)

Vào:

- r - cỡ của hình vuông
- $G(X, Y, E, c)$ - đồ thị theo màu c của ảnh I_1 và I_2 gồm $(4r-5)^2$ đỉnh

Ra: $costr$ - khoảng cách giữa hai cạnh của hình vuông kích thước r .

1. $M \leftarrow$ Giá trị đối sánh cực tiểu của $G(X, Y, E, c)$ gồm $(4r-5)^2$ đỉnh
2. For mỗi cạnh $(i,j) \in M$ do
$$costr \leftarrow Constr + w(i,j)$$
3. Trả lại giá trị $costr$

Độ phức tạp của thuật toán IHG:

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của hàm EdgeDistance thông qua Mệnh đề 2.5.

Mệnh đề 2.5 [Độ phức tạp của hàm EdgeDistance]: Độ phức tạp của hàm EdgeDistance($G(X, Y, E, c), r$) là $O(r^2)$, với $(4r-5)$ là số dải của lược đồ màu khối tương ứng với các khối ảnh thuộc cạnh hình vuông kích thước r của ảnh,

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của hàm DistancebyColor() thông qua Mệnh đề 2.6.

Mệnh đề 2.6 [Độ phức tạp của hàm DistancebyColor(I_1, I_2, n, c)]: Độ phức tạp của hàm DistancebyColor (I_1, I_2, n, c) là $O(n^3)$, với $n \times n$ là số dải của lược đồ màu khối,

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của thuật toán IHG thông qua Mệnh đề 2.7.

Mệnh đề 2.7 [Độ phức tạp của thuật toán IHG]: Độ phức tạp của thuật toán IHG(I_1, I_2, n) là $O(n^3)$, với n là số dải của lược đồ màu khối,

2.4 Các thực nghiệm

2.4.1 Môi trường thực nghiệm

Cơ sở dữ liệu gồm 7,812 ảnh jpeg. Cơ sở dữ liệu ảnh này là tập con của tập ảnh của GS Wang và chúng tôi tập hợp từ Internet được sử dụng để đánh giá hiệu năng tra cứu.

2.4.2 Các kết quả thực nghiệm

2.4.2.1 Kết quả thực nghiệm với phương pháp HG

Để kiểm tra độ chính xác của kỹ thuật tra cứu được phát triển, sáu truy vấn được chọn cho mục tiêu đánh giá. Tập truy vấn của chúng tôi gồm 6 ảnh truy vấn. Các truy vấn từ 1 đến 5 cùng với tập ảnh liên quan được tạo ra từ cơ sở dữ liệu “Wang 1000”, truy vấn 6 cùng tập ảnh liên quan được chúng tôi tập hợp từ Internet. Bảng 2.1 chỉ ra các loại của ảnh truy vấn và tập ảnh liên quan.

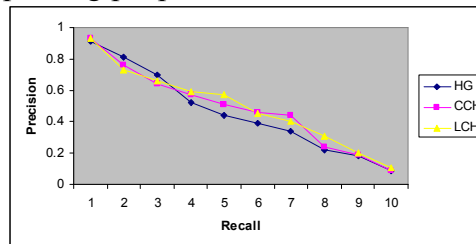
Bảng 2.1. Các loại của ảnh truy vấn và các ảnh liên quan.

TT	Tên truy vấn	Số ảnh liên quan	Tính chất của các ảnh liên quan
1	Ngựa	12	không quay hoặc dịch chuyển
2	Voi	15	không quay hoặc dịch chuyển
3	Hoa	9	không quay hoặc dịch chuyển
4	Người châu phi	14	không quay hoặc dịch chuyển
5	Xe buýt	10	gồm các ảnh quay và dịch chuyển
6	Cầu Thê Húc	17	gồm các ảnh quay và dịch chuyển



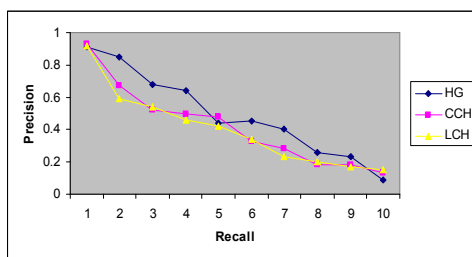
Hình 2.10. Các ảnh mẫu của các truy vấn từ 1 đến 6.

Từ đồ thị Precision-Recall trong Hình 2.11, chúng tôi nhận thấy rằng phương pháp CCH và LCH cho độ chính xác tra cứu tương đương nhau và độ chính xác của phương pháp HG tương đương với hai phương pháp CCH và LCH. Điều đó nói lên rằng, đối với các ảnh truy vấn có tập ảnh liên quan không bị quay và dịch chuyển thì phương pháp CCH và LCH làm việc hiệu quả tương đương phương pháp HG.



Hình 2.11. So sánh LCH, CCH với HG theo các truy vấn 1, 2, 3 và 4 dưới dạng Recall - Precision.

Từ đồ thị Precision-Recall trong Hình 2.12, chúng tôi nhận thấy rằng phương pháp CCH và LCH cho độ chính xác tra cứu thấp hơn hẳn phương pháp HG. Điều đó nói lên rằng, đối với các ảnh truy vấn có tập ảnh liên quan bị quay và dịch chuyển thì phương pháp HG cho kết quả chính xác hơn phương pháp CCH và LCH.



Hình 2.12. So sánh LCH, CCH với HG theo các truy vấn 5 và 6 dưới dạng Recall - Precision.

Nhận xét về thực nghiệm đối với phương pháp HG: Với tra cứu tập các ảnh quay và dịch chuyển, phương pháp HG cho kết quả chính xác hơn hẳn LCH và CCH.

2.4.2.2 Kết quả thực nghiệm với phương pháp IHG

Để kiểm tra độ chính xác của kỹ thuật tra cứu IHG, sáu truy vấn được chọn cho mục tiêu đánh giá. Tập truy vấn của chúng tôi gồm 6 ảnh truy vấn. Các truy vấn 1, 3, 4 và 5 cùng với tập ảnh liên quan được tạo ra từ cơ sở dữ liệu “Wang 1000”, truy vấn 2 cùng tập ảnh liên quan được chúng tôi tập hợp từ Internet. Tập ảnh truy vấn được chia ra làm hai loại: Loại 1 gồm các ảnh truy vấn và tập ảnh liên quan của nó được điều chỉnh quay, loại 2 gồm các ảnh truy vấn và tập ảnh liên quan của nó được điều chỉnh dịch chuyển. Bảng 2.8 chỉ ra các loại của ảnh truy vấn và tập ảnh liên quan.

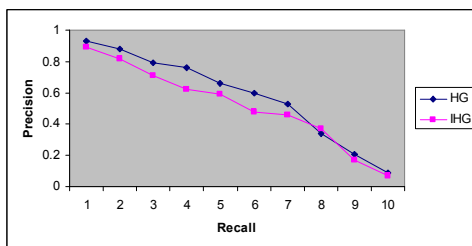
Bảng 2.8. Các loại của ảnh truy vấn và các ảnh liên quan.

TT	Tên truy vấn	Số ảnh liên quan	Tính chất của các ảnh liên quan
1	Xe buýt	10	Điều chỉnh dịch chuyển
2	Cầu Thê Húc	17	Điều chỉnh dịch chuyển
3	Khủng long	14	Điều chỉnh quay
4	Ngựa	12	Điều chỉnh quay
5	Voi	15	Điều chỉnh quay
6	Hoa	9	Điều chỉnh quay



Hình 2.13. Các ảnh mẫu của các truy vấn từ 1 đến 6.

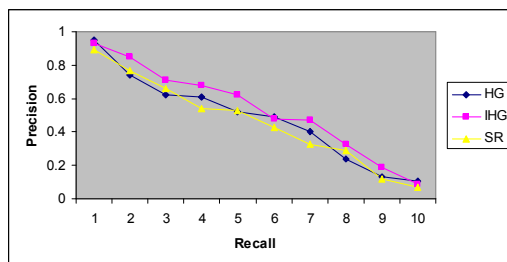
Từ đồ thị Precision-Recall trong Hình 2.14, chúng tôi nhận thấy rằng phương pháp IHG cho độ chính xác tra cứu có phần thấp hơn phương pháp HG. Điều đó nói lên rằng, đối với các ảnh truy vấn có tập ảnh liên quan bị dịch chuyển thì phương pháp HG cho kết quả chính xác hơn phương pháp IHG.



Hình 2.14. So sánh HG với IHG theo các truy vấn 1 và 2 dưới dạng Recall – Precision.

Từ đồ thị Precision-Recall trong Hình 2.15, chúng tôi nhận thấy rằng phương pháp IHG cho kết quả tra cứu tốt hơn phương pháp HG và SR. Điều đó nói lên rằng, đối với các ảnh

truy vấn có tập ảnh liên quan bị quay thì phương pháp IHG cho kết quả chính xác hơn HG và SR.



Hình 2.15. So sánh HG với IHG và SR theo các truy vấn 3, 4, 5 và 6 dưới dạng Recall-Precision.

Nhận xét về thực nghiệm đối với IHG: Với tra cứu tập các ảnh quay, phương pháp IHG cho kết quả chính xác hơn phương pháp HG và SR.

2.5 Kết luận

Chương này đã trình bày phương pháp tra cứu ảnh dựa vào lược đồ màu khối. Trên cơ sở phương pháp này chúng tôi đã trình bày kỹ thuật tra cứu ảnh sử dụng ít không gian lưu trữ số các lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển, được gọi là HG. Sử dụng các mệnh đề đã được chứng minh và so sánh các kết quả thực nghiệm của các phương pháp LCH, CCH và HG, chúng tôi đã chỉ ra, đối với tập ảnh được điều chỉnh quay và dịch chuyển, HG cho kết quả tra cứu chính xác hơn LCH và CCH.

Tiếp theo đó, chúng tôi đã đề xuất một phương pháp đối sánh ảnh cải tiến, gọi là IHG. Các mệnh đề đã được chứng minh và so sánh các kết quả áp dụng ở phương pháp HG và phương pháp cải tiến IHG, chúng tôi đã chỉ ra rằng, đối với tập các ảnh quay, độ chính xác và tốc độ của IHG cao hơn phương pháp HG. Phương pháp IHG làm việc hiệu quả với các ảnh khi chúng được tách thành nhiều hơn bốn khối ảnh.

Chương 3. PHƯƠNG PHÁP TRA CỨU DỰA VÀO VÙNG ẢNH

3.1 Biểu diễn ảnh sử dụng phương pháp cây tứ phân

Với phương pháp cây tứ phân, một vùng trong ảnh được tách thành bốn vùng có kích cỡ bằng nhau. Với mỗi vùng, nếu vùng là thuần nhất, thì vùng này không được xem xét là ứng cử viên cho lần tách tiếp theo. Nếu vùng không thuần nhất, nó sẽ được tách thành bốn vùng con cho đến khi tất cả các vùng đều đồng nhất. Ảnh thu được sau khi sử dụng phương pháp này được biểu diễn bằng cấu trúc cây. Mỗi nút trong cấu trúc này hoặc là nút lá hoặc có bốn nút con.

Tuy nhiên, cách tiếp cận dựa vào các khối có kích cỡ đều này khó có thể cung cấp thông tin màu cục bộ chính xác do các đối tượng trong ảnh khó có thể ép vào các khối có kích thước đều.

3.2 Phương pháp tra cứu ảnh sử dụng đặc trưng của vùng ảnh

3.2.1 Giới thiệu

Để đưa thông tin không gian vào bản mô tả ảnh, cách tiếp cận thông thường là chia ảnh thành các khối và trích rút đặc trưng màu từ mỗi khối. Một biến thể của cách tiếp cận này là cách tiếp cận dựa vào cây tứ phân, trong đó ảnh được tách ra thành cấu trúc cây tứ phân và mỗi nhánh của cây có một lược đồ riêng để mô tả nội dung màu của nhánh.

Tuy nhiên, cách tiếp cận dựa vào các khối có kích cỡ đều này khó có thể cung cấp thông tin màu cục bộ chính xác do các đối tượng trong ảnh rất khó để ép vào các khối có kích thước đều. Để khắc phục hạn chế này, chúng tôi đề xuất kỹ thuật CSI và CCS.

3.2.2 Trích rút thông tin không gian

3.2.2.1 Trích rút thông tin không gian

Thuật toán CSI được mô tả như sau:

Đầu tiên, toàn bộ ảnh đã cho Img được coi như một vùng. Trong bước đầu tiên, ảnh có thể được tách thành hai vùng phụ thuộc vào hàm giá trị $Cost(BR_i)$ và màu phân cụm sử dụng kỹ thuật cân bằng lược đồ. Với mỗi vùng, một điều kiện tách được sử dụng để xác định một vùng có được phân hoạch hay không. Nếu các mẫu quan sát rơi vào vùng lệch đáng kể so với tần suất kỳ vọng, vùng cần phân hoạch tiếp, với mỗi vùng cần xác định hàm giá trị $Cost(BR_i)$, để xác định vùng BR_i hoặc được phân hoạch theo chiều ngang hoặc phân hoạch theo chiều dọc.

Tần suất kỳ vọng có thể được tính toán theo kinh nghiệm về phân bố mẫu. Nếu các mẫu quan sát quá xa tần suất kỳ vọng, phân hoạch vùng cần được tiếp tục và với mỗi vùng cần xác định giá trị hàm giá, để xác định BR_i sẽ được phân hoạch theo chiều ngang hay dọc. Phương pháp dựa trên tri thức về phân bố phụ thuộc vào kinh nghiệm chuyên gia sẽ đánh giá tần suất kỳ vọng được mô tả như dưới.

Độ lệch giữa một quan sát i và tần suất kỳ vọng được tính theo công thức:

$$DX = \frac{obs(i) - E(i)}{\sqrt{E(i)}} \quad (3-1)$$

Nếu DX vượt quá một ngưỡng T , vùng bị bỏ qua đối với phân hoạch tiếp theo. Trong trường hợp ngược lại, vùng hiện tại sẽ được thêm vào Stack cho phân hoạch tiếp theo. Quá trình phân hoạch được lặp cho đến khi một trong các điều kiện sau được thỏa mãn: tất cả các vùng là thuần nhất hoặc số các mẫu trong một vùng nhỏ hơn một ngưỡng đã cho.

Giá trị $Cost(BR_i)$ được tính như sau:

$$Cost(BR_i) = \text{Max}(DX_{selectedrow}, DX_{selectedcol}) \quad (3-2)$$

Ký hiệu $DX_{selectedrow}$ là tần suất kỳ vọng trong dòng $selectedrow$, $DX_{selectedcol}$ là tần suất kỳ vọng trong cột $selectedcol$.

$$Ở đây \quad DX_{selectedrow} = \text{Max}(DX_{toprow}, DX_{bottomrow}) \quad (3-3)$$

Ký hiệu DX_{toprow} và $DX_{bottomrow}$ là tần suất kỳ vọng trong dòng $toprow$ và $bottomrow$ theo các hướng trên xuống/dưới lên và các giá trị này được tính toán theo công thức ở dưới:

$$DX_{toprow} = \frac{obs_{toprow}(i) - E_{toprow}(i)}{\sqrt{E_{toprow}(i)}} \quad (3-4)$$

$$DX_{bottomrow} = \frac{obs_{bottomrow}(i) - E_{bottomrow}(i)}{\sqrt{E_{bottomrow}(i)}} \quad (3-5)$$

$$\text{và } DX_{selectedcol} = \text{Max}(DX_{leftcol}, DX_{rightcol}) \quad (3-6)$$

Tương tự, ký hiệu $DX_{leftcol}$, $DX_{rightcol}$ là tần suất kỳ vọng trong dòng $leftcol/rightcol$ theo các hướng trái sang phải/phải sang trái và các giá trị này được tính theo công thức sau:

$$DX_{leftcol} = \frac{obs_{leftcol}(i) - E_{leftcol}(i)}{\sqrt{E_{leftcol}(i)}} \quad (3-7)$$

$$DX_{rightcol} = \frac{obs_{rightcol}(i) - E_{rightcol}(i)}{\sqrt{E_{rightcol}(i)}} \quad (3-8)$$

Nếu $Cost(BR_i)$ bằng $DX_{selectedrow}$, thì BR_i được phân hoạch theo chiều dọc. Trong trường hợp ngược lại, nếu $Cost(BR_i)$ bằng $DX_{selectedcol}$, BR_i được phân hoạch theo chiều ngang.

Dưới đây, chúng tôi trình bày thuật toán trích rút màu và thông tin không gian CSI.

Thuật toán CSI:

Vào: - *Img* - ảnh
 - *minarea* – ngưỡng diện tích của vùng
 - *E* – ngưỡng nhiễu

Ra: - *C*- thông tin màu
 - *BR* – thông tin không gian của vùng màu

1. $Stack \leftarrow Img$
2. do
 - 2.1 $BR \leftarrow Stack$
 - 2.2 If ($area(BR) > minarea$)
 - 2.2.1 Tách BR thành 2 vùng $\{BR_j\}, j=1,2$ theo chiều ngang hoặc dọc theo giá trị hàm giá $Cost(BR)=Max(DX_{selectedrow}, DX_{selectedcol})$.
 - 2.2.2 for mỗi BR_j
 - tính $DX = \frac{obs(j) - E(j)}{\sqrt{E(j)}}$
 - If ($DX < T$) $Stack \leftarrow BR_j$
 - else xuất (C, BR_j)
 - 2.3 else xóa (BR)
3. while ($Stack \neq \emptyset$)

Trong thuật toán **CSI**, có ba tham số *minarea*, $Cost(BR_j)$ và *T*. *minarea* là diện tích tối thiểu của một vùng. Nếu diện tích của một vùng nhỏ hơn *minarea*, vùng không được sử dụng cho phân hoạch tiếp theo. Nếu $Cost(BR_j)$ bằng $DX_{selectedrow}$, vùng này được phân hoạch theo chiều dọc. Trong trường hợp ngược lại, nếu $Cost(BR_j)$ bằng $DX_{selectedcol}$ vùng được phân hoạch theo chiều ngang. *T* là ngưỡng nhiễu được chấp nhận của mỗi vùng.

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của thuật toán CSI thông qua Mệnh đề 3.1.

Mệnh đề 3.1 [Độ phức tạp của thuật toán CSI]:

Độ phức tạp của thuật toán CSI là $O(n^2)$ với *n* là số điểm ảnh của ảnh,

3.2.2.2 Trích rút các cụm màu thuần nhất.

Đầu tiên thuật toán coi một ảnh đã cho *I* như một vùng. Nếu diện tích của vùng này nhỏ hơn một ngưỡng đã cho thì thuật toán sẽ loại bỏ vùng này. Nếu vùng là thuần nhất, CCS sẽ xuất vùng này và màu của nó, và dừng. Ngược lại nó gọi thủ tục Split() để phân hoạch vùng *Rec* thành hai vùng Rec_1 và Rec_2 và đẩy chúng vào Stack. Quá trình này sẽ lặp đối với mỗi vùng trong Stack cho đến khi Stack rỗng.

Trong thuật toán CCS, có các tham số *minsize* và *tolerance*. Ở đây *minsize* là diện tích nhỏ nhất của một vùng, *tolerance* chỉ ra mức nhiễu cho phép trong mỗi vùng. Nếu diện tích của một vùng nhỏ hơn *minsize*, vùng sẽ không được phân hoạch tiếp.

Thuật toán CCS có thể được viết như sau.

Thuật toán CCS:

Vào: I – ảnh ,

$minsize$ –ngưỡng diện tích của một vùng.

Ra: Các cụm màu thuần nhất trong ảnh.

1. $Stack \leftarrow I$
2. do
 - 2.1 $REC \leftarrow Stack$
 - 2.2 If ($size(REC) > minsize$)
 - 2.2.1 if ($deviation(REC) > tolerance$)
 - 2.2.1.1 **Split** ($REC, c_1, Rec_1, c_2, Rec_2$)
 - 2.2.1.2 If ($size(Rec_1) > 0$ and $size(Rec_2) > 0$)
 - $Stack \leftarrow Rec_1$
 - $Stack \leftarrow Rec_2$
 - 2.2.2 else xuất (c, REC)
 - 2.3 else xóa (REC)
3. while ($Stack \neq \emptyset$)

Dưới đây, chúng ta sẽ mô tả chi tiết thủ tục Split().

Thủ tục **Split** ($REC, c_1, Rec_1, c_2, Rec_2$)

Vào: Cụm REC với cỡ $n \times n$,

Ra: Các cụm và các màu của nó (c_1, Rec_1), (c_2, Rec_2)

1. for $i \leftarrow 0$ to $n-1$ do
 - 1.1 for $j \leftarrow 0$ to $n-1$ do

$$\{ row \leftarrow \sum_j p_{i,j} ; afterrow \leftarrow \sum_j p_{i+1,j} \}$$
 - 1.2 $v_i \leftarrow |(\frac{min(i, n-i)}{n} * (row - afterrow))|$
 - 1.3 calculate $maxv(k) \leftarrow \max(<v_1, v_2, \dots, v_n>)$
2. for $j \leftarrow 0$ to $n-1$ do
 - 2.1 for $i \leftarrow 0$ to $n-1$ do

$$\{ col \leftarrow \sum_i p_{i,j} ; aftercol \leftarrow \sum_i p_{i,j+1} \}$$
 - 2.2 $h_j \leftarrow |(\frac{min(j, n-j)}{n} * (col - aftercol))|$
 - 2.3 calculate $maxh(l) \leftarrow \max(<h_1, h_2, \dots, h_n>)$
3. if ($maxv(k) > maxh(l)$) then
 - 3.1 Split REC by vertical at row k
 - 3.2 $Rec_1 \leftarrow size((0,0);(k,n-1)); c_1 \leftarrow color(Rec_1)$
 - 3.3 $Rec_2 \leftarrow size((k,0);(m-1,n-1)); c_2 \leftarrow color(Rec_2)$
4. else if ($maxv(k) < maxh(l)$) then
 - 4.1 Split REC by horizontal at column l
 - 4.2 $Rec_1 \leftarrow size((0,0);(n-1,l)); c_1 \leftarrow color(Rec_1)$
 - 4.3 $Rec_2 \leftarrow size((0,l);(n-1,n-1)); c_2 \leftarrow color(Rec_2)$
5. else if ($v_1 = v_2 = \dots = v_n = h_1 = h_2 = \dots = h_n$)

$$\{ Rec_1 \leftarrow 0; c_1 \leftarrow Null ; Rec_2 \leftarrow 0; c_2 \leftarrow Null \}$$
6. Trả lại $<(c_1, Rec_1), (c_2, Rec_2)>$

Thủ tục Split() xuất ra các vùng Rec_1 , Rec_2 và các màu c_1 , c_2 tương ứng của nó. Trong thủ tục này, tham số k giữ chỉ số dòng sẽ được sử dụng để tách theo chiều đứng và l giữ lại chỉ số cột sẽ được sử dụng để tách theo chiều ngang. Hàm maxv() tính giá trị lớn nhất của danh sách $\langle v_1, v_2, \dots, v_n \rangle$ và hàm maxh() tính giá trị lớn nhất của danh sách $\langle h_1, h_2, \dots, h_n \rangle$.

Với mỗi dòng i ($i=0, 1, \dots, n-1$), Split() tính tổng số các điểm ảnh của dòng i và tổng các điểm ảnh của dòng $i+1$. Sau đó, nó tính giá trị độ lệch theo chiều đứng v_i của hai tổng này. Tương tự, với mỗi cột j ($j=0, 1, \dots, n-1$), Split() cũng tính toán giá trị độ lệch theo chiều ngang h_j giữa tổng số các điểm ảnh trong cột j và cột $j+1$.

Dựa trên các giá trị $|v_i|$ và $|h_j|$ tính được, thủ tục Split() sẽ phân hoạch vùng REC thành hai vùng Rec_1 và Rec_2 theo chiều ngang hoặc chiều đứng.

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của thủ tục Split() thông qua Mệnh đề 3.2.

Mệnh đề 3.2 [Độ phức tạp của thủ tục Split]:

Độ phức tạp của thủ tục Split() là $O(n)$ với n là số điểm ảnh của ảnh,

Dưới đây chúng tôi sẽ đánh giá độ phức tạp của thuật toán CCS thông qua Mệnh đề 3.3.

Mệnh đề 3.3 [Độ phức tạp của thuật toán CCS]:

Độ phức tạp của thuật toán CCS là $O(n^2)$ với n là số điểm ảnh của ảnh,

3.2.3 Khoảng cách giữa hai ảnh

Sau khi sử dụng kỹ thuật CSI hoặc CCS để chia ảnh Img_1 và Img_2 thành dãy các vùng, chúng tôi sẽ sử dụng thuật toán DRC để tính khoảng cách giữa ảnh Img_1 và Img_2 .

Hàm **DRC**:

Vào: T_c - tổng số các màu của tập màu

T_k^{s1} - các vùng của màu k của ảnh Img_1

T_k^{s2} - các vùng của màu k của ảnh Img_2

Ra : dist - khoảng cách giữa ảnh Img_1 và Img_2

1. dist ← 0;

2. For $k \leftarrow 1$ to T_c do

2.1 For $i \leftarrow 1$ to T_k^{s1} do

2.1.1 For $j \leftarrow 1$ to T_k^{s2} do

If $(R_{Img_1}(i, k) \cap R_{Img_2}(j, k))$ then

dist ← dist + $|R_{Img_1}(i, k) \cap R_{Img_2}(j, k)|$

3. Trả lại giá trị dist

3.2.4 Các thực nghiệm

3.2.4.1 Môi trường thực nghiệm

Hiệu năng tra cứu được đánh giá sử dụng một cơ sở dữ liệu gồm 7,812 ảnh jpeg. Cơ sở dữ liệu này là tập con của tập ảnh của GS WANG và chúng tôi tập hợp qua Internet.

3.2.4.2 Kết quả thực nghiệm

Phương pháp CSI:

Để kiểm tra độ chính xác của phương pháp tra cứu CSI, sáu truy vấn được thực hiện và các truy vấn 1 và 2 được sử dụng ba phương pháp CSI, QT và CBC, các truy vấn từ 3 đến 6 sử dụng thêm phương pháp SR. Các truy vấn từ 1 đến 6 cùng với tập ảnh liên quan được tạo ra từ cơ sở dữ liệu “WANG 1000”. Thực nghiệm của chúng tôi đã sử dụng các tham số

$minarea$ và E , ở đây $minarea=36$ và $E=0.42$. Bảng 3.7 chỉ ra các loại ảnh truy vấn và tập ảnh liên quan tương ứng.

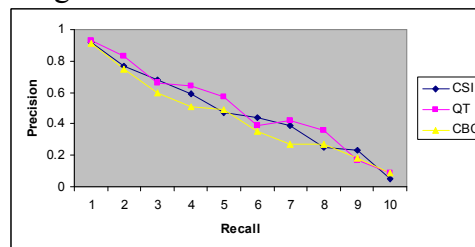
Bảng 3.7. Các loại của ảnh truy vấn và các ảnh liên quan.

TT	Tên truy vấn	Số ảnh liên quan	Tính chất của các ảnh liên quan
1	Ngựa	15	Tương đối hỗn tạp
2	Voi	18	Tương đối hỗn tạp
3	Hoa	12	Có độ thuần nhất cao
4	Bãi biển	21	Có độ thuần nhất cao
5	Núi	14	Có độ thuần nhất cao
6	Di tích cổ	18	Có độ thuần nhất cao



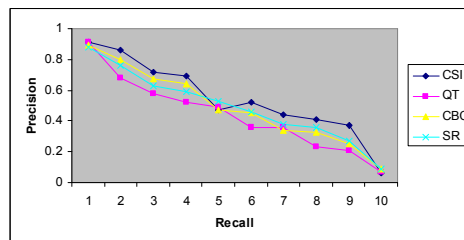
Hình 3.8. Các ảnh mẫu của các truy vấn từ 1 đến 6.

Từ đồ thị Precision-Recall trong Hình 3.9, chúng tôi nhận thấy rằng, đối với các truy vấn có ảnh mẫu và tập ảnh liên quan tương đối hỗn tạp, các phương pháp CSI, QT và CBC cho kết quả tra cứu tương đương nhau.



Hình 3.9. So sánh CSI với QT và CBC theo các truy vấn 1 và 2 dưới dạng Recall-Precision.

Từ đồ thị Precision-Recall trong Hình 3.10, chúng tôi nhận thấy rằng, đối với các ảnh truy vấn có tập ảnh liên quan tương đối thuần nhất thì phương pháp CSI làm việc hiệu quả hơn phương pháp QT, CBC và SR.



Hình 3.10. So sánh CSI với QT, CBC và SR theo các truy vấn 3, 4, 5 và 6 dưới dạng Recall – Precision.

Nhận xét đối với thực nghiệm của phương pháp CSI: Với tra cứu tập các ảnh có độ thuần nhất cao, phương pháp CSI làm việc hiệu quả hơn phương pháp QT, CBC và SR.

Phương pháp CCS:

Để kiểm tra độ chính xác của phương pháp tra cứu CCS, sáu truy vấn được thực hiện và các truy vấn 1, 2 và 3 được sử dụng ba phương pháp CCS, CCV và CSI, các truy vấn từ 4 đến 6 sử dụng thêm phương pháp SR. Các truy vấn từ 1 đến 6 cùng với tập ảnh liên quan được tạo ra từ cơ sở dữ liệu “WANG 1000”. Thực nghiệm của chúng tôi đã sử dụng các

tham số *minsize* và *tolerance*, ở đây *minsize*=64 và *tolerance*= 0.31. Bảng 3.14 chỉ ra các loại của ảnh truy vấn và tập ảnh liên quan.

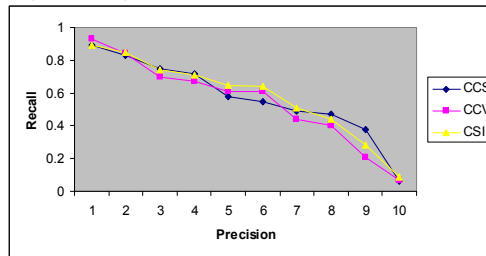
Bảng 3.14. Các loại của ảnh truy vấn và tập ảnh liên quan.

TT	Tên truy vấn	Số ảnh liên quan	Tính chất của các ảnh liên quan
1	Thức ăn	15	Hỗn tạp
2	Ngựa	15	Hỗn tạp
3	Voi	18	Hỗn tạp
4	Bãi biển	21	Tương đối thuần nhất
5	Núi	14	Tương đối thuần nhất
6	Di tích cổ	18	Tương đối thuần nhất



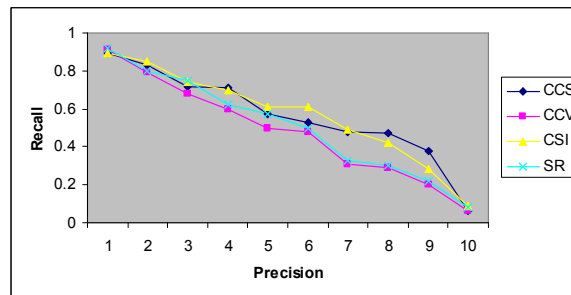
Hình 3.11. Các ảnh mẫu của các truy vấn từ 1 đến 6.

Từ đồ thị Precision-Recall trong Hình 3.12, chúng tôi nhận thấy rằng, đối với các truy vấn có ảnh mẫu và tập ảnh liên quan có độ thuần nhất thấp, các phương pháp CCS, CCV và CSI cho kết quả tra cứu tương đương nhau.



Hình 3.12. So sánh Recall – Precision theo các truy vấn 1,2 và 3 của CCS với CCV và CSI.

Từ đồ thị Precision-Recall trong Hình 3.13, chúng tôi nhận thấy rằng, đối với các truy vấn có ảnh mẫu và tập ảnh liên quan có độ thuần nhất cao, hai phương pháp CCS và CSI cho kết quả tra cứu tương đương nhau. Hai phương pháp này cho kết quả tra cứu tốt hơn phương pháp CCV và SR.



Hình 3.13. So sánh Recall-Precision theo các truy vấn 4, 5 và 6 của CCS với CCV, CSI và SR.

Nhận xét đối với thực nghiệm của phương pháp CCS: Với tra cứu tập các ảnh có độ thuần nhất cao, phương pháp CCS làm việc hiệu quả hơn phương pháp CCV và SR.

3.3 Kết luận

Chúng tôi đã trình bày phương pháp biểu diễn ảnh sử dụng cây tứ phân. Trên cơ sở phân tích hạn chế của phương pháp này, chúng tôi đề xuất kỹ thuật tra cứu ảnh dựa vào màu và không gian CSI. Kỹ thuật bao gồm ba giai đoạn:

- Sắp xếp lược đồ cấp xám theo thứ tự giảm dần của tần số xuất hiện, sử dụng phương pháp cân bằng lược đồ để giảm số các màu của ảnh.
- Chia ảnh thành dãy các hình chữ nhật theo thủ tục tách chiều ngang và dọc.
- Sử dụng thông tin không gian để tra cứu các ảnh liên quan từ cơ sở dữ liệu.

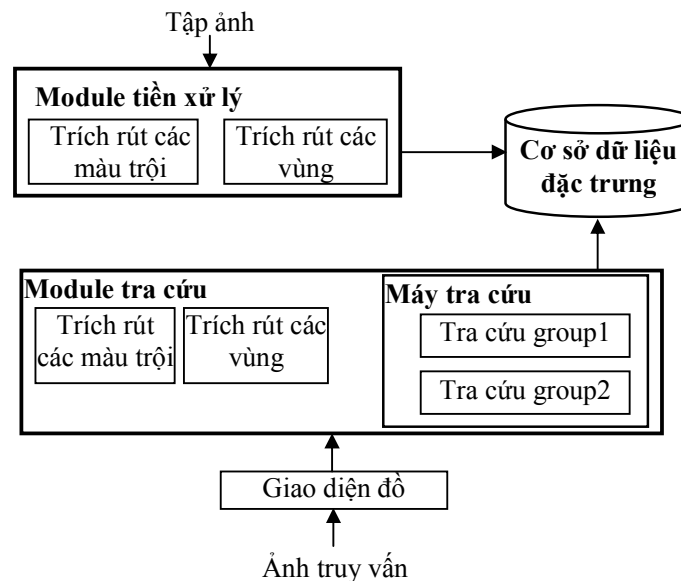
Hơn nữa, chúng tôi cũng đề xuất một kỹ thuật tra cứu khác sử dụng màu và các cụm màu thuần nhất của nó để phục vụ quá trình tra cứu, gọi là CCS. Các mệnh đề đã được chứng minh và các kết quả thực nghiệm đã minh chứng độ chính xác của kỹ thuật đề xuất.

Cả hai kỹ thuật CSI và CCS đều có khả năng tự động chia ảnh thành các vùng có kích cỡ khác nhau và sử dụng các vùng này vào trong quá trình tra cứu ảnh.

Chương 4. XÂY DỰNG ỨNG DỤNG TRA CỨU ẢNH DỰA VÀO NỘI DUNG

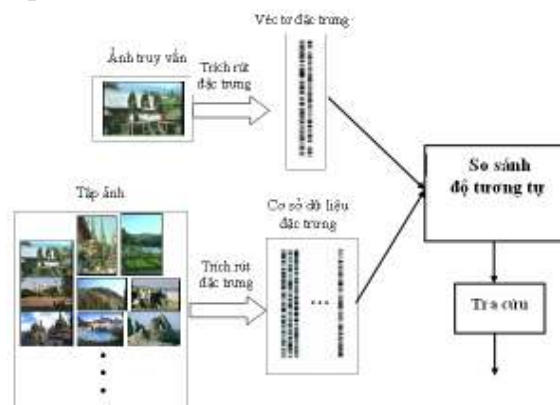
4.1 Thiết kế hệ thống tổng quát LVFIR

Kiến trúc toàn bộ hệ thống được chỉ ra trong hình 4.1.



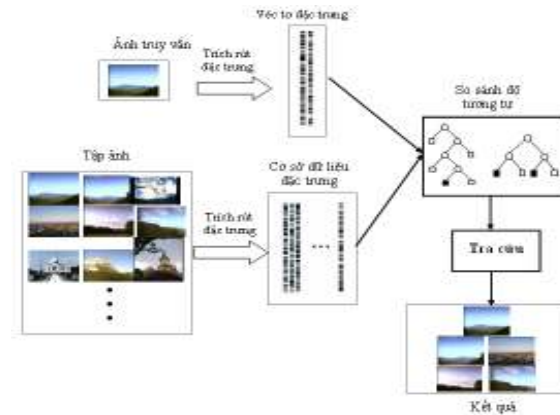
Hình 4.1. Kiến trúc của hệ thống LVFIR.

4.2 Module tra cứu group1



Hình 4.2. Kiến trúc của Module tra cứu group1.

4.3 Module tra cứu group2

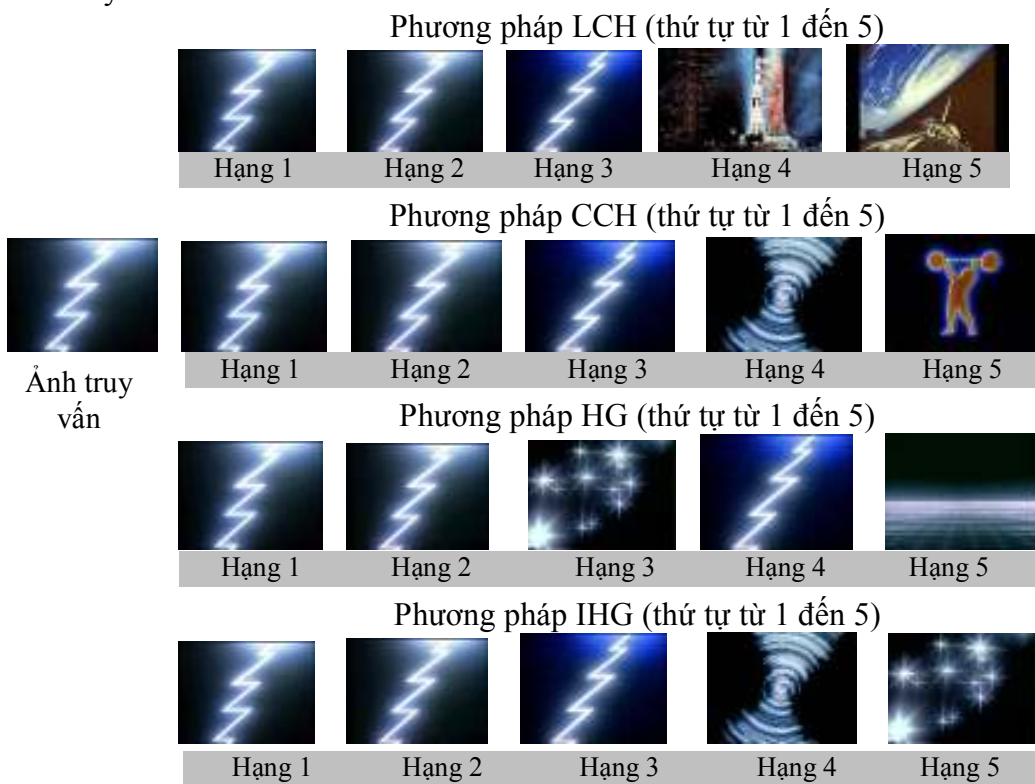


Hình 4.8. Kiến trúc của Module tra cứu group2.

4.4 Một số kết quả

4.4.1 So sánh kỹ thuật LCH, CCH với HG và IHG

Truy vấn 1:



Hình 4.16. Kết quả thực hiện truy vấn 1.

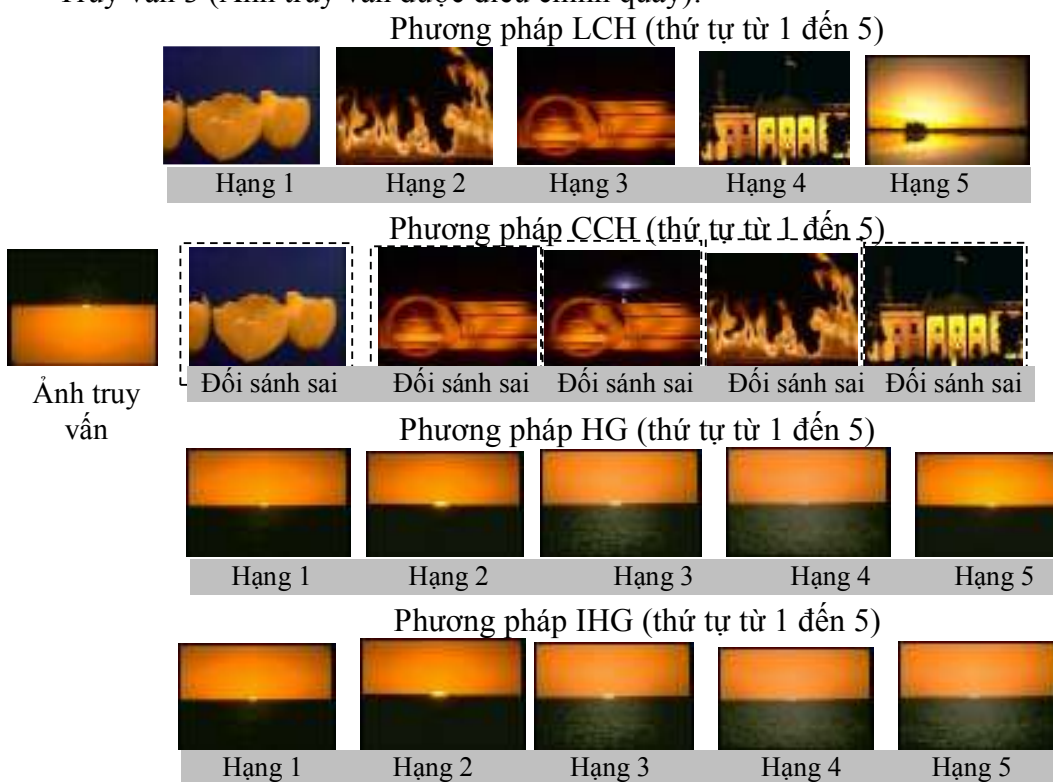
Truy vấn 2 (Ảnh truy vấn được điều chỉnh dịch chuyển):





Hình 4.17. Kết quả thực hiện truy vấn 2.

Truy vấn 3 (Ảnh truy vấn được điều chỉnh quay):



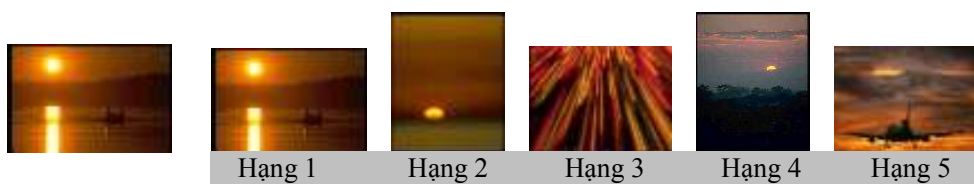
Hình 4.18. Kết quả thực hiện truy vấn 3.

Từ các truy vấn 1, 2 chúng tôi nhận thấy phương pháp HG và IHG cho kết quả xấp xỉ phương pháp LCH và CCH. Tuy nhiên, trong trường hợp ảnh truy vấn được điều chỉnh quay hoặc dịch chuyển (truy vấn 2 và 3), phương pháp HG và IHG thực hiện tốt hơn hẳn phương pháp LCH và CCH.

4.4.2 So sánh kỹ thuật QT, CBC và CCV với CSI và CCS

Truy vấn 1:

Phương pháp QT



Ảnh truy
vấn

Phương pháp CCV



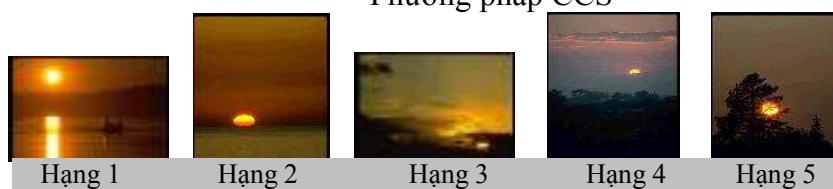
Phương pháp CBC



Phương pháp CSI



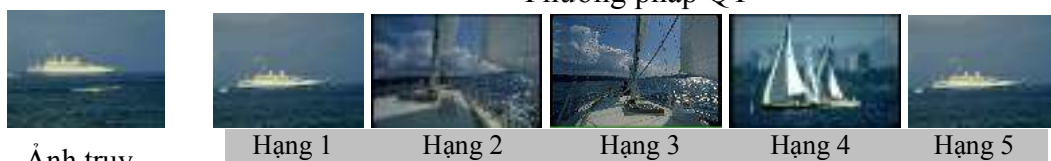
Phương pháp CCS



Hình 4.19. Kết quả thực hiện truy vấn 1.

Truy vấn 2:

Phương pháp QT

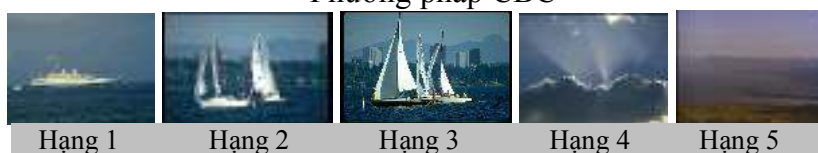


Ảnh truy
vấn

Phương pháp CCV



Phương pháp CBC



Phương pháp CSI



Phương pháp CCS



Hình 4.20. Kết quả thực hiện truy vấn 2.

Truy vấn 3:



Ảnh truy
vấn

Phương pháp QT



Phương pháp CCV



Phương pháp CBC



Phương pháp CSI



Phương pháp CCS



Hình 4.21. Kết quả thực hiện truy vấn 3.

Từ các truy vấn 1, 2 và 3 chúng tôi nhận thấy phương pháp CSI và CCS thực hiện tốt hơn phương pháp QT, CBC và phương pháp CCV.

4.5 Kết luận.

Ứng dụng được phát triển thử nghiệm bằng C# trên hệ điều hành Windows XP, bộ xử lý Pentium 1.73 GHz, 512 MB bộ nhớ với cơ sở dữ liệu ảnh gồm 7,812 ảnh. Đối với cơ sở dữ liệu ảnh này, kết quả cho thấy phương pháp HG và IHG cho kết quả tốt hơn phương pháp LCH và CCH, đặc biệt là khi ảnh được điều chỉnh quay hoặc dịch chuyển. Cũng trên cơ sở dữ liệu ảnh này, phương pháp CSI và CCS cho độ chính xác cao hơn phương pháp QT, CCV và CBC.

KẾT LUẬN

Sử dụng các đặc trưng thị giác trích rút được, đặc biệt là đặc trưng của vùng ảnh, trong tra cứu ảnh dựa vào đặc trưng thị giác là chủ đề nghiên cứu được nhiều người quan tâm. Nhiều kỹ thuật đã được đề xuất để đáp ứng các yêu cầu khác nhau. Hầu hết các kỹ thuật đều cố gắng nâng cao hiệu năng tra cứu theo hướng tra cứu nhanh và chính xác. Trong luận án này, ngoài việc tập trung vào giải quyết bài toán tra cứu theo hướng nhanh và chính xác. Tác giả còn hướng đến giải quyết vấn đề giảm không gian lưu trữ số các lược đồ màu biểu diễn ảnh.

Để giải quyết vấn đề giảm không gian lưu trữ số các lược đồ màu biểu diễn ảnh, tăng tốc độ và độ chính xác tra cứu trong trường hợp ảnh quay và dịch chuyển. Chúng tôi đã nghiên cứu một số kỹ thuật khác nhau. Trong đó đã phân tích các kỹ thuật GCH, LCH và lược đồ màu khối CCH. Trên cơ sở phân tích ưu và nhược điểm của các kỹ thuật này, chúng tôi đã đề xuất phương pháp tra cứu ảnh dựa vào đặc trưng thị giác sử dụng ít không gian lưu trữ số các lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển có tên là HG và phương pháp HG cải tiến. Các mệnh đề đã được chứng minh và các kết quả thực nghiệm đã chỉ ra tốc độ và độ chính xác của kỹ thuật tra cứu.

Để giải quyết vấn đề tăng độ chính xác tra cứu thông qua sử dụng các đặc trưng cục bộ, chúng tôi đã phân tích ưu điểm và hạn chế của kỹ thuật biểu diễn ảnh sử dụng cây tứ phân. Trên cơ sở phân tích này, chúng tôi đã đề xuất phương pháp tra cứu ảnh dựa vào đặc trưng thị giác CSI và CCS. Hai phương pháp này sử dụng đặc trưng của vùng ảnh vào trong quá trình tra cứu. Từ các mệnh đề đã được chứng minh và từ các kết quả thực nghiệm đã chỉ ra độ chính xác của kỹ thuật tra cứu được đề xuất là hiệu quả.

Tóm lại, đóng góp chính của luận án đó là:

Thứ nhất, luận án đã đề xuất được phương pháp, có tên là HG, để giải quyết bài toán tra cứu ảnh dựa vào đặc trưng thị giác trong trường hợp ảnh bị quay và dịch chuyển và giảm chi phí không gian lưu trữ các lược đồ màu. Phương pháp này đã được công bố trên tạp chí quốc tế IJCSSES.

Thứ hai, trên cơ sở phương pháp HG luận án cũng đã đưa ra phương pháp HG cải tiến, có tên là IHG, phương pháp này cải tiến độ chính xác và tốc độ của phương pháp HG. Phương pháp này đã được công bố trên tạp chí quốc tế IJCSSES.

Thứ ba, luận án đã đề xuất được kỹ thuật tra cứu ảnh CSI dựa vào đặc trưng của vùng ảnh. Kỹ thuật này có khả năng tự động chia ảnh thành các vùng có kích cỡ khác nhau và sử dụng các vùng này trong quá trình tra cứu ảnh. Kỹ thuật này đã được công bố tại hội nghị quốc tế về xử lý ảnh CISP08.

Thứ tư, bên cạnh kỹ thuật CSI tác giả cũng đã đề xuất được kỹ thuật có tên là CCS. Kỹ thuật trích rút màu và các cụm màu thuần nhất để phục vụ quá trình tra cứu. Kỹ thuật này cũng có khả năng tự động chia ảnh thành các vùng có kích cỡ khác nhau và sử dụng các vùng này trong quá trình tra cứu ảnh. Kỹ thuật này đã được công bố trên tạp chí Công nghệ thông tin và Truyền thông *PTITJ*.

Cuối cùng, trên cơ sở các kỹ thuật đã được đề xuất, chúng tôi đã xây dựng được hệ thống tra cứu ảnh dựa vào đặc trưng thị giác có tên là LVFIR. Hệ thống này gồm hai module chính là module tiền xử lý và module tra cứu.

Một số vấn đề cần được nghiên cứu tiếp trong tương lai:

- Kết hợp đặc trưng kết cấu và đặc trưng hình vào quá trình tra cứu.
- Xây dựng cơ chế đánh chỉ số cơ sở dữ liệu ảnh để tăng tốc độ quá trình tra cứu ảnh.
- Thực nghiệm trên cơ sở dữ liệu ảnh có kích thước lớn hơn và đa dạng hơn.

DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ

- [1] Quynh, N. H. and Tao, N. Q (2010), “A Novel Content Based Image Retrieval Method Based on Splitting the Image into Homogeneous Regions”, *International Journal of Innovative Computing, Information and Control*, Vol.6, No.10 (Accepted).
- [2] Quynh, N. H. and Tao, N. Q (2009), “A novel method for content based image retrieval using color features”, *International Journal of Computer Sciences and Engineering Systems*, Vol.3, No.1, pp. 1-6.
- [3] Quynh, N. H. and Tao, N. Q (2009), “Improving HG Method for Content based Landscape Image Retrieval”, *International Journal of Computer Sciences and Engineering Systems*, Vol.3, No.1, pp. 47-51.
- [4] Quynh, N. H. and Tao, N. Q., Giang, N. T. (2008), “A efficient method for content based image retrieval using histogram graph”, *Proc. of IEEE on Control, Automation, Robotics and Vision*, pp. 874-878.
- [5] Quynh, N. H. and Tao, N. Q., Giang, N. T. (2008), “Efficient content based image retrieval through sector histogram”, *Proc. of IEEE on Circuits and Systems*, pp. 1814-1817.
- [6] Quynh, N. H. and Tao, N. Q. (2008), “Combining Color and Spatial Information for Retrieving Landscape Images” *In Proc. of IEEE on Image and Signal Processing, Vol. 2 - Volume 02*, IEEE Computer Society, Washington, DC, pp. 480-484.
- [7] Quynh, N. H. and Tao, N. Q (2008), “Segmenting the images into homogeneous clusters for retrieving landscape images”, *Posts, Telecommunications and Information Technology Journal (PTITJ)*, Issue 3, pp. 54-59.
- [8] Nguyễn Hữu Quỳnh, Ngô Quốc Tạo (2007), “Sử dụng đặc tính cục bộ của vùng phục vụ tra cứu ảnh phong cảnh”, *Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông*, tr. 608-617, Đại Lải.