

LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các kết quả được viết chung với các tác giả khác đều được sự đồng ý của đồng tác giả trước khi đưa vào luận án. Các kết quả nêu trong luận án là trung thực và chưa từng được ai công bố trong các công trình nào khác.

Tác giả

Nguyễn Hữu Quỳnh

Lời cảm ơn

Thực hiện luận án tiến sĩ là một việc khó, nhưng là một nhiệm vụ đáng làm. Tôi rất hạnh phúc khi thực hiện xong luận án tiến sĩ, và quan trọng hơn là những gì tôi đã học được trong suốt ba năm qua. Bên cạnh kiến thức tôi thu được, tôi đã học được phương pháp nghiên cứu một cách độc lập. Sự thành công này không đơn thuần bởi sự nỗ lực của cá nhân tôi, mà còn có sự hỗ trợ và giúp đỡ của thầy giáo hướng dẫn và nhiều đồng nghiệp khác. Nhân cơ hội này, tôi muốn bày tỏ lời cảm ơn của tôi đến họ.

Đầu tiên, tôi muốn cảm ơn đến hai thầy giáo hướng dẫn của tôi, PGS TS Ngô Quốc Tạo và PGS TS Đinh Mạnh Tường, vì sự hướng dẫn tận tình và khoa học. Đó là một cơ hội lớn cho tôi để được nghiên cứu dưới sự hướng dẫn của hai thầy. Cảm ơn rất nhiều tới hai thầy vì sự hướng dẫn tôi cách đặt ra các câu hỏi nghiên cứu, hiểu các vấn đề, và viết các bài báo khoa học.

Tôi trân trọng cảm ơn Bộ môn Khoa học máy tính, Khoa Công nghệ thông tin, Phòng Đào tạo Sau Đại học - Nghiên cứu Khoa học, Ban giám hiệu trường Đại học Công nghệ đã tạo điều kiện thuận lợi cho tôi trong suốt quá trình thực hiện luận án.

Tôi bày tỏ sự cảm ơn đến PGS TS Vũ Đức Thi, PGS TS Lương Chi Mai, PGS TS Nguyễn Thanh Thủy vì sự giúp đỡ của họ cho các đề xuất và các trao đổi trong nghiên cứu của tôi. Tôi cũng bày tỏ sự cảm ơn đến PGS TS Đàm Xuân Hiệp – Hiệu trưởng trường Đại học Điện lực, người đã động viên và tạo điều kiện về thời gian và tài chính cho tôi trong việc công bố các bài báo trên các hội nghị và tạp chí quốc tế.

Tôi muốn cảm ơn đến các cán bộ, giảng viên trong khoa Công nghệ thông tin – Trường Đại học Điện lực đã cổ vũ động viên và sát cánh bên tôi trong quá trình nghiên cứu.

Tôi muốn cảm ơn những thành viên của đề tài nghiên cứu cơ bản NCCB200706 về sự tài trợ tài chính và các góp ý rất hữu ích về các bài báo được công bố trên các hội nghị và tạp chí quốc tế.

Tôi cảm ơn tất cả những người bạn của tôi. Những người luôn chia sẻ và cổ vũ tôi trong những lúc khó khăn và tôi luôn ghi nhớ điều đó.

Cuối cùng, tôi xin bày tỏ lòng biết ơn vô hạn đối với cha mẹ và gia đình đã luôn ủng hộ, giúp đỡ tôi.

MỤC LỤC

PHẦN MỞ ĐẦU	14
1. Tính cấp thiết của luận án	14
2. Mục tiêu của luận án	16
3. Các đóng góp của luận án	16
4. Bố cục của luận án	17
Chương 1. TỔNG QUAN VỀ TRÍCH RÚT ĐẶC TRUNG VÀ TRA CỨU ẢNH DỰA VÀO ĐẶC TRUNG	18
1.1 Các đặc trưng.....	18
1.1.1 Các đặc trưng toàn cục và cục bộ.....	18
1.1.2 Các đặc trưng thị giác trong tra cứu ảnh.....	19
1.2 Kiến trúc của một hệ thống tra cứu ảnh dựa vào đặc trưng thị giác.....	19
1.3 Trích rút đặc trưng	21
1.3.1 Đặc trưng màu	21
1.3.2 Lượng hóa màu.....	23
1.3.3 Biểu diễn màu.....	23
1.3.3.1 Lược đồ màu	23
1.3.3.2 Lược đồ màu toàn cục GCH.....	24
1.3.3.3 Lược đồ màu cục bộ LCH	26
1.3.3.4 Véc tơ gắn kết màu	28
1.3.3.5 Tương quan màu	28
1.3.3.6 Các màu trội.....	29
1.3.3.7 Mô men màu	29
1.3.4 Thông tin không gian.....	30

1.3.5 Phân vùng.....	31
1.4 Các độ đo tương tự.....	32
1.5 Đánh giá hiệu năng tra cứu.....	37
1.6 Các hệ thống VFBIR.....	38
1.7 Kết luận và định hướng nghiên cứu.....	40
Chương 2. PHƯƠNG PHÁP TRA CỨU DỰA VÀO LƯỢC ĐỒ MÀU KHỐI....	42
2.1 Lược đồ màu khối.....	42
2.2 Phương pháp tra cứu dựa vào lược đồ màu khối.....	44
2.2.1 Giới thiệu	44
2.2.2 Phương pháp tra cứu đề xuất HG	47
2.2.2.1 Khái niệm về đồ thị hai phía.....	47
2.2.2.2. Phương pháp HG.....	48
2.3 Phương pháp cải tiến IHG.....	53
2.3.1 Khái niệm về sự tương tự lý tưởng giữa hai dải	53
2.3.2 Lý do đề xuất phương pháp IHG.....	54
2.3.3 Phương pháp IHG.....	54
2.4 Các thực nghiệm	60
2.4.1 Môi trường thực nghiệm	60
2.4.2 Các kết quả thực nghiệm.....	61
2.4.2.1 Kết quả thực nghiệm với phương pháp HG	61
2.4.2.2 Kết quả thực nghiệm với phương pháp IHG.....	65
2.5 Kết luận	69
Chương 3. PHƯƠNG PHÁP TRA CỨU DỰA VÀO VÙNG ẢNH.....	71
3.1 Biểu diễn ảnh sử dụng phương pháp cây tứ phân	71
3.2 Phương pháp tra cứu ảnh sử dụng đặc trưng của vùng ảnh.....	73
3.2.1 Giới thiệu	73

3.2.2 Trích rút đặc trưng	74
3.2.2.1 Trích rút màu và thông tin không gian.....	74
3.2.2.2 Trích rút các cụm màu thuần nhất.	82
3.2.3 Độ tương tự giữa hai ảnh	87
3.2.4 Các thực nghiệm.....	88
3.2.4.1 Môi trường thực nghiệm	88
3.2.4.2 Kết quả thực nghiệm	88
3.3 Kết luận	96
Chương 4. XÂY DỰNG ỨNG DỤNG TRA CỨU ẢNH DỰA VÀO NỘI DUNG	98
4.1 Thiết kế hệ thống tổng quát LVFIR.....	98
4.2 Module tra cứu group1	100
4.3 Module tra cứu group2.....	105
4.4 Một số kết quả	110
4.4.1 So sánh kỹ thuật LCH, CCH với HG và IHG	110
4.4.2 So sánh kỹ thuật QT, CBC và CCV với CSI và CCS	112
4.5 Kết luận.	116
KẾT LUẬN.....	117
DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ	119
TÀI LIỆU THAM KHẢO	120

DANH MỤC CÁC CHỮ VIẾT TẮT

Ký hiệu	Diễn giải
Black	Màu đen
CSDL	Cơ sở dữ liệu
CBC	Color Based Cluster
CCH	Color/Cell Histogram (Lược đồ màu khối)
CCS	Cluster of Colors and Space (Cụm màu và không gian)
CCV	Color Coherence Vectors (Véc tơ gắn kết màu)
CSI	Color and Spatial Information (Màu và thông tin không gian)
DistancebyColor	Khoảng cách theo màu
DRC	Distance by Region Comparing
EdgeDistance	Khoảng cách theo cạnh
EMD	Earth Mover Distance (Khoảng cách Earth Mover)
GCH	Global Color Histogram (Lược đồ màu toàn cục)
Gray	Màu xám
HG	Histogram Graph (Đồ thị lược đồ)
Hue	Sắc màu
IHG	Improving Histogram Graph method (Phương pháp cải tiến đồ thị lược đồ)
KLT	Karhunen–Loeve transform (Biến đổi Karhunen–Loeve)
LCH	Local Color Histogram (Lược đồ màu cục bộ)
LVFIR	Local Visual Feature-based Image Retrieval (Tra cứu ảnh dựa vào đặc trưng thị giác cục bộ)
MCM	Minimum Cost Matching (Giá trị đối sánh cực tiểu)
MTM	Mathematical Transform to Munsell (Biến đổi toán học sang hệ thống màu Munsell)

Precision	Chính xác
Quantization	Lượng hóa
QT	Quad Tree (Cây tứ phân)
Recall	Hồi tưởng
RGB	Red (Đỏ), Green (Xanh lục), Blue (xanh lơ)
SR	Spatial Relationship (Quan hệ không gian)
Union	Hợp
VFBIR	Visual Feature Based Image Retrieval (Tra cứu ảnh dựa vào đặc trưng thị giác)
White	Màu trắng

DANH MỤC CÁC HÌNH

Hình 1.1. Kiến trúc hệ thống tra cứu ảnh dựa vào đặc trưng thị giác.....	20
Hình 1.2. Hai ảnh khác nhau nhưng có cùng lược đồ màu.	22
Hình 1.3. Từ trái sang: ảnh gốc sử dụng 256 màu, được lượng hoá trong 8 dải, và được lượng hoá trong 64 dải sử dụng không gian màu RGB.....	23
Hình 1.4. Ba ảnh I_1 , I_2 và I_3 và các lược đồ màu tương ứng của chúng.....	25
Hình 1.5. Tính khoảng cách giữa ảnh I_1 và I_2 sử dụng LCH, $d_{LCH}(I_1, I_2) = 1.319$, $d_{GCH}(I_1, I_2) = 0.088$	26
Hình 1.6. Tính khoảng cách giữa các ảnh I_1 và I_3 sử dụng LCH, $d_{LCH}(I_1, I_3) = 0.707$, $d_{GCH}(I_1, I_3) = 0.088$	27
Hình 1.7. Tính khoảng cách giữa các ảnh I_2 và I_3 sử dụng LCH $d_{LCH}(I_2, I_3) = 0.707$, $d_{GCH}(I_2, I_3) = 0$	27
Hình 1.8. Recall và Precision cho các kết quả truy vấn.....	38
Hình 2.1. Một ảnh được chia thành 9 khối ảnh và ba lược đồ màu khối của nó....	43
Hình 2.2. Ảnh I và ảnh I'	45
Hình 2.3. Lược đồ màu khối theo màu black và white biểu diễn ảnh I	45
Hình 2.4. Lược đồ màu khối theo màu black và white biểu diễn ảnh I'	45
Hình 2.5. Tính khoảng cách của ảnh I và I' theo màu black.....	46
Hình 2.6. Tính khoảng cách của ảnh I và I' theo màu white.....	46
Hình 2.7. Các khối ảnh của mỗi ảnh được đánh số từ trong ra và ngược chiều kim đồng hồ.	56
Hình 2.8. Lược đồ màu khối theo màu black của hai ảnh I_1 và I_2	56
Hình 2.9. Đồ thị hai phía biểu thị mối quan hệ của các dải của lược đồ màu khối của ảnh I_1 và I_2 theo màu black.	57

Hình 2.10. Các ảnh mẫu của các truy vấn từ 1 đến 6.....	61
Hình 2.11. So sánh LCH, CCH với HG theo các truy vấn 1, 2, 3 và 4 dưới dạng Recall - Precision.	63
Hình 2.12. So sánh LCH, CCH với HG theo các truy vấn 5 và 6 dưới dạng Recall - Precision.....	64
Hình 2.13. Các ảnh mẫu của các truy vấn từ 1 đến 6.....	65
Hình 2.14. So sánh HG với IHG theo các truy vấn 1 và 2 dưới dạng Recall – Precision.....	67
Hình 2.15. So sánh HG với IHG và SR theo các truy vấn 3, 4, 5 và 6 dưới dạng Recall-Precision.	68
Hình 2.16. Biểu đồ so sánh tốc độ của phương pháp HG và IHG.....	69
Hình 3.1 Ảnh gốc.	71
Hình 3.2. Cây tứ phân biểu diễn ảnh cho trong Hình 3.1.....	72
Hình 3.3 Cây biểu diễn ảnh cho trong Hình 3.1.	73
Hình 3.4. Ảnh I cỡ 10×10 điểm ảnh.	77
Hình 3.5. Ảnh I sau khi được tách ra thành hai vùng BR_1 và BR_2	78
Hình 3.6. Vùng BR_2 sau khi được tách ra thành hai vùng $BR_{2,1}$ và $BR_{2,2}$	80
Hình 3.7. Ảnh gồm 6×10 điểm ảnh.....	85
Hình 3.8. Các ảnh mẫu của các truy vấn từ 1 đến 6.....	89
Hình 3.9. So sánh CSI với QT và CBC theo các truy vấn 1 và 2 dưới dạng Recall-Precision.....	90
Hình 3.10. So sánh CSI với QT, CBC và SR theo các truy vấn 3, 4, 5 và 6 dưới dạng Recall – Precision.	92
Hình 3.11. Các ảnh mẫu của các truy vấn từ 1 đến 6.....	93
Hình 3.12. So sánh Recall – Precision theo các truy vấn 1,2 và 3 của CCS với CCV và CSI.	94

Hình 3.13. So sánh Recall-Precision theo các truy vấn 4, 5 và 6 của CCS với CCV, CSI và SR.....	96
Hình 4.1. Kiến trúc của hệ thống LVFIR.	99
Hình 4.2. Kiến trúc của Module tra cứu group1.	100
Hình 4.3. Màn hình chính của module tra cứu group1.	102
Hình 4.4. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng LCH.....	102
Hình 4.5. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng CCH.....	103
Hình 4.6. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng HG.....	103
Hình 4.7. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng IHG.....	104
Hình 4.8. Kiến trúc của Module tra cứu group2.	105
Hình 4.9. Giao diện sử dụng kỹ thuật QT, CBC và CCV của module tra cứu group2.	106
Hình 4.10. Giao diện sử dụng kỹ thuật CSI và CCS của module tra cứu group2.....	107
Hình 4.11. Giao diện tra cứu khi sử dụng phương pháp QT với ảnh truy vấn.....	107
Hình 4.12. Giao diện tra cứu khi sử dụng phương pháp CBC với ảnh truy vấn... ..	108
Hình 4.13. Giao diện tra cứu khi sử dụng phương pháp CCV với ảnh truy vấn... ..	108
Hình 4.14. Giao diện tra cứu khi sử dụng phương pháp CSI với ảnh truy vấn.....	109
Hình 4.15. Giao diện tra cứu khi sử dụng phương pháp CCS với ảnh truy vấn....	109
Hình 4.16. Kết quả thực hiện truy vấn 1.	110
Hình 4.17. Kết quả thực hiện truy vấn 2.	111
Hình 4.18. Kết quả thực hiện truy vấn 3.	112
Hình 4.19. Kết quả thực hiện truy vấn 1.	113
Hình 4.20. Kết quả thực hiện truy vấn 2.	114
Hình 4.21. Kết quả thực hiện truy vấn 3.	115

DANH MỤC CÁC BẢNG

Bảng 2.1. Các loại của ảnh truy vấn và các ảnh liên quan.	61
Bảng 2.8. Các loại của ảnh truy vấn và các ảnh liên quan.	65
Bảng 3.1. Tính độ lệch $DX_{selectedrow}$ cho phân hoạch theo dòng của ảnh I	78
Bảng 3.2. Tính độ lệch $DX_{selectedcol}$ cho phân hoạch theo cột của ảnh I	79
Bảng 3.3. Tính độ lệch $DX_{selectedrow}$ cho phân hoạch theo dòng của vùng BR_2	80
Bảng 3.4. Tính độ lệch $DX_{selectedcol}$ cho phân hoạch theo cột của vùng BR_2	81
Bảng 3.5. Tính toán giá trị của v_i	85
Bảng 3.6. Tính toán giá trị của h_j	86
Bảng 3.7. Các loại của ảnh truy vấn và các ảnh liên quan.	89
Bảng 3.14. Các loại của ảnh truy vấn và tập ảnh liên quan.....	92
Bảng 3.17. Các kết quả của truy vấn 3.....	94
Bảng 3.20. Các kết quả của truy vấn 3.....	95

PHẦN MỞ ĐẦU

1. Tính cấp thiết của luận án

Những năm gần đây, chúng ta đã chứng kiến sự tăng nhanh kích cỡ của các tập hợp ảnh số cùng với sự phát triển bùng nổ của các ứng dụng Internet. Hàng ngày, việc sử dụng các thiết bị thu nhận ảnh sinh ra nhiều giga-bytes dữ liệu ảnh. Một lượng lớn thông tin ảnh, khoảng hàng trăm triệu ảnh [12, 59, 70, 79], 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. Quản trị cơ sở dữ liệu (CSDL) và thị giác máy là hai cộng đồng có đóng góp chính cho lĩnh vực tra cứu ảnh. Hai cộng đồng này tiếp cận tra cứu ảnh từ hai góc độ khác nhau, dựa vào văn bản mô tả ảnh và dựa vào đặc trưng thị giác của bản thân ảnh.

Sử dụng các kỹ thuật dựa vào văn bản mô tả ảnh hoặc từ khoá mô tả ảnh để quản lý CSDL ảnh là cách đơn giản thường được sử dụng. Các từ khoá mô tả ảnh cung cấp thông tin nội dung mô tả ảnh trong một CSDL ảnh đã cho, nhưng để mô tả các ảnh đủ chi tiết, cần một tập từ khoá rất lớn và phức tạp. Một hạn chế nữa của cách tiếp cận này là cần nhân lực được đào tạo kỹ lưỡng để xây dựng các từ khoá đối với mỗi ảnh và chọn các từ khoá phù hợp cho tra cứu các ảnh hiệu quả. Công việc mô tả nội dung ảnh thủ công này tốn nhiều thời gian, chi phí cao và phụ thuộc vào cảm nhận chủ quan của chuyên viên kỹ thuật theo nghĩa cùng một nội dung ảnh, những người khác nhau có thể đưa ra cảm nhận về ảnh khác nhau. Cảm nhận chủ quan và mô tả nội dung ảnh không chính xác là nguyên nhân làm cho so sánh sai trong lúc tra cứu. Hơn nữa, hệ thống dựa vào từ khoá rất khó thay đổi về sau. Do đó, cần có cách tiếp cận mới để khắc phục các hạn chế này.

Để khắc phục các khó khăn ở trên, 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.

Hệ thống tra cứu ảnh dựa vào đặc trưng thị giác sẽ xác định các ảnh trong CSDL ảnh có đặc trưng thị giác tương tự với ảnh truy vấn theo hai pha: Pha 1, tất cả các ảnh trong CSDL được xử lý, được trích chọn đặc trưng thị giác. Quá trình xử lý và trích chọn đặc trưng thị giác được thực hiện một cách tự động ngay khi các ảnh được nhập vào CSDL. Quá trình này gán cho mỗi ảnh một tập các ký hiệu mô tả, các ký hiệu mô tả ảnh này sẽ được lưu trữ trong CSDL và được sử dụng trong pha tiếp theo. Pha 2, trích rút các đặc trưng thị giác của ảnh truy vấn và so sánh các đặc trưng này với các đặc trưng thị giác của ảnh trong CSDL theo một độ đo tương tự nào đó. Các ảnh trong CSDL được phân hạng theo mức độ tương tự của nó với ảnh truy vấn. Ảnh có hạng cao nhất được truy xuất. Trích rút nội dung thị giác của các ảnh hiệu quả và đo độ tương tự giữa các ảnh dựa trên đặc trưng thị giác là hai phần quan trọng trong tra cứu ảnh dựa vào đặc trưng thị giác.

Các nghiên cứu gần đây trong tra cứu ảnh tập trung vào trích chọn đặc trưng thị giác gồm màu, kết cấu, hình dạng và thông tin không gian. Màu là đặc trưng được sử dụng rộng rãi nhất cho tra cứu ảnh do tính toán nhanh, tương đối ổn định với các biến dạng nhỏ, thay đổi về kích thước và hướng. Một số phương pháp đã được đề xuất như: Phương pháp lược đồ màu toàn cục và lược đồ màu cục bộ [51], phương pháp véc tơ gắn kết màu [16], phương pháp tương quan màu [30], phương pháp lược đồ màu khối [54],... Tuy nhiên, hầu hết các phương pháp này đều gặp phải vấn đề sử dụng nhiều không gian để lưu trữ các lược đồ màu biểu diễn ảnh, độ chính xác tra cứu không cao, độ phức tạp tính toán lớn, nhạy cảm với quay và dịch chuyển, không cho phép nhận biết các đối tượng tương tự có màu khác nhau [7, 40, 52, 69].

Do đó, việc đề xuất các giải pháp tra cứu ảnh dựa vào đặc trưng thị giác để khắc phục được các hạn chế ở trên là một nhu cầu cấp thiết. Đó cũng là lý do mà

luận án chọn đề tài “**Nghiên cứu cải tiến một số phương pháp tra cứu ảnh sử dụng đặc trưng ảnh**”.

2. Mục tiêu của luận án

Mục tiêu 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 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

Trong luận án này, tác giả nghiên cứu đề xuất các kỹ thuật tra cứu ảnh dựa vào đặc trưng của vùng ảnh gồm: phương pháp HG (Histogram Graph) [42], phương pháp IHG (Improving Histogram Graph) [43], phương pháp CSI (Color and Spatial Information) [45] và phương pháp CCS (Cluster of Colors and Space) [46]:

- Phương pháp tra cứu ảnh dựa vào đặc trưng màu, có tên là HG [42]. Đặc điểm của phương pháp này là sử dụng ít không gian lưu trữ các lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển.

- Để tăng cường phương pháp HG, chúng tôi đã đề xuất phương pháp IHG [43], nhằm giảm thời gian và tăng độ chính xác tra cứu của phương pháp HG nhưng vẫn sử dụng ít không gian lưu trữ các lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển.

- Phương pháp CSI [45] trích rút đặc trưng màu và thông tin không gian của các vùng ảnh và sử dụng trong quá trình tra cứu để nâng cao hiệu năng tra cứu.

- Phương pháp CCS [46] phân hoạch ảnh thành các cụm màu thuần nhất (các cụm màu này có thể có kích cỡ khác nhau) và trích rút thông tin màu và không gian của mỗi vùng phục vụ quá trình tra cứu.

- Xây dựng hệ thống tra cứu ảnh dựa vào đặc trưng thị giác có tên là LVFIR (Local Visual Feature-based Image Retrieval) trên cơ sở các kỹ thuật đề xuất của tác giả. Hệ thống này gồm hai module chính là module tiền xử lý và module tra cứu.

4. Bố cục của luận án

Luận án này được bố cục thành bốn chương, gồm 125 trang.

Chương 1 giới thiệu tổng quan về trích rút đặc trưng và tra cứu ảnh dựa vào đặc trưng thị giác và đưa ra một số kết luận và định hướng cho nghiên cứu.

Chương 2 trình bày kỹ thuật tra cứu ảnh dựa vào lược đồ màu khối, có tên là HG [42] và cải tiến của nó, có tên là IHG [43].

Chương 3 trình bày kỹ thuật trích rút đặc trưng của vùng ảnh sử dụng trong quá trình tra cứu ảnh, có tên là CSI [45] và CCS [46].

Chương 4 trình bày thiết kế và thực hiện hệ thống thực nghiệm tra cứu ảnh dựa vào đặc trưng thị giác (sử dụng các kỹ thuật được đề xuất trong Chương 2 và Chương 3) LVFIR, cùng với một số kết quả.

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

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

Trong chương này, chúng tôi sẽ 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 gồm: các đặc trưng, kiến trúc của hệ thống tra cứu ảnh dựa vào đặc trưng thị giác, trích rút đặc trưng, các độ đo tương tự, đánh giá hiệu năng tra cứu và giới thiệu một số hệ thống tra cứu ảnh dựa vào đặc trưng thị giác. Đặc biệt chúng tôi nhấn mạnh vào đặc trưng màu. Cuối cùng chúng tôi sẽ đưa ra một số kết luận và định hướng cho nghiên cứu.

1.1 Các đặc trưng

Dữ liệu ảnh thô không được sử dụng trực tiếp trong hầu hết các hệ thống thị giác máy vì hai lý do: Thứ nhất, tốn nhiều không gian để lưu trữ ảnh và độ phức tạp tính toán lớn. Thứ hai, nhiều thông tin của ảnh là dư thừa và/ hoặc không hữu ích. Thay vì sử dụng toàn bộ ảnh, chúng ta chỉ cần sử dụng một biểu diễn quan trọng nhất. Bước tìm biểu diễn được gọi là *trích rút đặc trưng* và kết quả của biểu diễn là véc tơ đặc trưng. Trích rút đặc trưng có thể xem như việc ánh xạ ảnh từ không gian ảnh sang không gian đặc trưng.

Liên quan đến nội dung ảnh, 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 [12, 70, 79]. Các đặc trưng thị giác chung gồm màu, kết cấu, hình dạng và quan hệ không gian. Các đặc trưng theo lĩnh vực bao gồm tri thức về lĩnh vực như mặt người, vân tay,... Đặc trưng ngữ nghĩa không dễ dàng được trích rút và thường được suy diễn từ các đặc trưng mức thấp hoặc sử dụng văn bản mô tả ảnh.

1.1.1 Các đặc trưng toàn cục và cục bộ

Các đặc trưng ảnh có thể là toàn cục hoặc cục bộ. Nếu các đặc trưng biểu diễn nội dung thị giác của toàn bộ ảnh, các đặc trưng này được gọi là các đặc trưng toàn

cục. Ngược lại, các đặc trưng biểu diễn nội dung thị giác của một phần ảnh thì được gọi là đặc trưng cục bộ.

1.1.2 Các đặc trưng thị giác trong tra cứu ảnh

Đặc trưng màu: Màu có vai trò quan trọng trong tra cứu ảnh dựa vào đặc trưng thị giác. Các màu có thể được biểu diễn trong các không gian màu khác nhau như RGB, HSV,...

Đặc trưng kết cấu: Kết cấu là tập các điểm trong một vùng thỏa mãn ràng buộc hay qui luật nào đó. Đặc trưng này khá quan trọng cho tra cứu ảnh.

Về cơ bản, các phương pháp biểu diễn kết cấu có thể được chia thành hai loại: các phương pháp *cấu trúc* và các phương pháp *thống kê*. Các toán tử được dùng phát hiện cấu trúc bao gồm *các toán tử hình thái* và *đồ thị liên kề* xác định các kết cấu cơ sở và luật phân bố của chúng. Các phương pháp thống kê bao gồm: Phương pháp phổ năng lượng Fourier, Tamura, trường ngẫu nhiên Markov, mô hình fractal, các bộ lọc đa phân giải như biến đổi Gabor và biến đổi dạng sóng... thể hiện kết cấu bằng sự phân bố thống kê của độ sáng của các điểm ảnh.

Đặc trưng hình dạng: Các đặc trưng hình dạng có quan hệ chặt chẽ với mô tả vùng hoặc các đối tượng được phân đoạn. Đặc trưng hình dạng được trích rút từ các đường bao đối tượng hoặc vùng chứa đối tượng.

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

Quá trình thực hiện của hệ thống tra cứu ảnh dựa vào đặc trưng thị giác được chia thành hai giai đoạn:

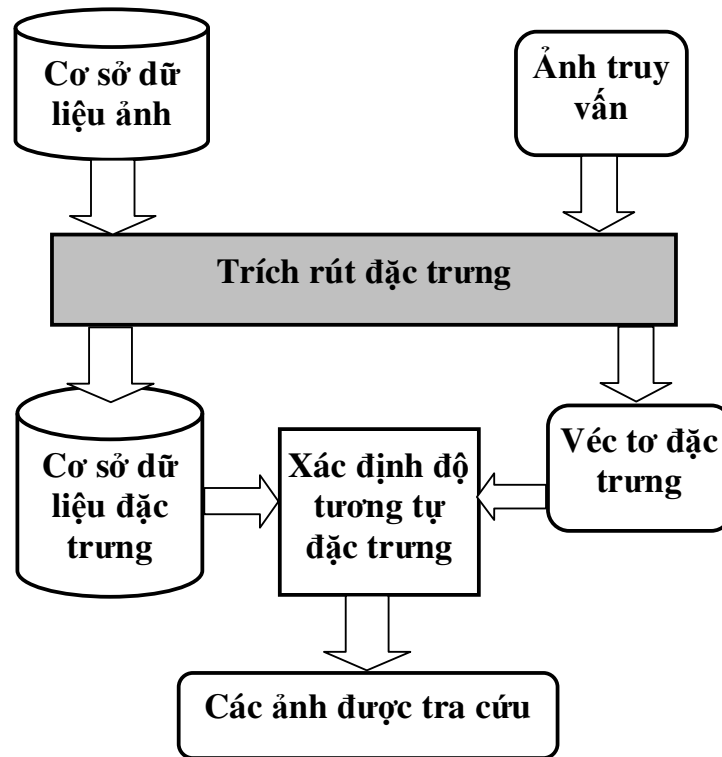
Giai đoạn 1: Tạo lập CSDL ảnh cùng với thông tin đặc trưng (ngoại tuyến)

Trích rút đặc trưng của ảnh trong CSDL ảnh. Quá trình xử lý gồm lọc, chuẩn hóa, phân đoạn và nhận dạng đối tượng. Đầu ra của bước này là một tập các mô tả nội dung các ảnh trong CSDL.

Giai đoạn 2: Tra cứu ảnh (trực tuyến)

1. Tạo lập truy vấn: trích rút đặc trưng thị giác của ảnh truy vấn.
2. So sánh: các đặc trưng thị giác của ảnh truy vấn được so sánh với các đặc trưng thị giác của các ảnh trong CSDL ảnh. Các kỹ thuật đánh chỉ số có thể được sử dụng nhằm tăng tốc quá trình tra cứu.

Dưới đây là kiến trúc hệ thống tra cứu ảnh dựa vào đặc trưng thị giác.



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

Hình 1.1 chỉ ra kiến trúc hệ thống tra cứu ảnh dựa vào đặc trưng thị giác. Các đặc trưng thị giác của ảnh trong CSDL ảnh được trích rút và được biểu diễn bằng các véc tơ đặc trưng nhiều chiều. Các véc tơ đặc trưng của các ảnh trong CSDL ảnh tạo thành CSDL đặc trưng. Khi thực hiện tra cứu, người sử dụng cung cấp cho hệ thống ảnh truy vấn, sau đó hệ thống trích rút các véc tơ đặc trưng của ảnh truy vấn này. Xác định độ tương tự giữa các véc tơ đặc trưng của ảnh truy vấn và các véc tơ đặc trưng của các ảnh trong CSDL đặc trưng. Trên cơ sở độ tương tự xác định được,

hệ thống cho ra kết quả tra cứu gồm một danh sách các ảnh có độ tương tự với ảnh truy vấn nhất.

1.3 Trích rút đặc trưng

Đặc trưng màu là một trong những đặc trưng thị giác quan trọng và được sử dụng rộng rãi nhất trong tra cứu ảnh. Do đó trong phần này, chúng tôi sẽ đề cập đến đặc trưng màu.

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

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.

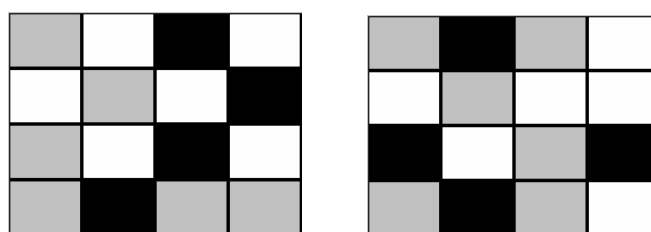
1.3.1 Đặc trưng màu

Đặc trưng màu được sử dụng rất hiệu quả cho tra cứu các ảnh màu trong CSDL ảnh [66, 69, 80]. Các mô tả màu được trích rút và so sánh tương đối thuận lợi và do đó nó thích hợp cho tra cứu dựa vào đặc trưng thị giác.

Ảnh được thu thập từ camera số, hoặc được tải xuống từ Internet thường có ba kênh màu (các ảnh đa cấp xám chỉ có một kênh, các ảnh đa phổ có thể có nhiều hơn ba kênh).

Các tín hiệu màu một hoặc hai chiều cũng được sử dụng rộng rãi trong tra cứu ảnh dựa vào đặc trưng thị giác (VFBIR) đặc biệt trong các ứng dụng với điều kiện thu nhận ảnh tương phản là quan trọng. Trong [19, 20] đã chỉ ra rằng màu là bất biến dưới ánh sáng, bóng và sự thay đổi hình học của người quan sát và các góc chiếu sáng.

Lược đồ màu được dùng để miêu tả đặc trưng màu của một ảnh, đếm số lần xuất hiện của mỗi màu trong một ảnh [37]. Từ lược đồ màu ta suy ra phân bố xác suất của màu trong ảnh. Phân bố này bất biến với quay, dịch chuyển và tỷ lệ; do đó, lược đồ màu rất thích hợp cho tra cứu ảnh dựa vào đặc trưng thị giác. Tuy nhiên, hạn chế chính của lược đồ màu là chưa tận dụng được thông tin không gian của các vùng ảnh. Điều này có thể dẫn đến các sai số không mong muốn; thí dụ, trong tra cứu ảnh sử dụng một lược đồ màu là không thể phân biệt giữa một quả bóng màu xanh và một bụi cỏ màu xanh. Chẳng hạn, Hình 1.2 chỉ ra nhược điểm này.



Hình 1.2. Hai ảnh khác nhau nhưng có cùng lược đồ màu.

Nhiều phương pháp khác đã được đề xuất, bao gồm: mô men màu [32, 34], các dấu hiệu màu [31], các lược đồ màu cải tiến [51], các véc tơ gắn kết màu [16], tra cứu các ảnh dựa vào phân cụm [80], các tương quan màu [30], các vùng màu cục bộ [2, 31], phương pháp Harbin [63] và các đốm màu [64]. Các phương pháp này đề cập đến các kỹ thuật tra cứu theo màu ở mức không gian. Các kỹ thuật này đều có xuất phát điểm sử dụng một trong hai cách tiếp cận, cách tiếp cận thứ nhất theo hướng cố gắng liên kết thông tin không gian vào lược đồ màu toàn cục, cách tiếp cận thứ hai cố gắng tăng thông tin không gian thông qua chia ảnh thành các khối đều. Nhóm thứ nhất có hạn chế là rất khó thu nhận được thông tin không gian của các vùng (đối tượng) trong ảnh, do bản chất của lược đồ màu toàn cục biểu thị phân bố xác suất của toàn bộ ảnh. Tuy nhóm thứ hai có thể thu nhận được thông tin không gian tốt hơn nhóm thứ nhất nhưng vẫn có các hạn chế: Hạn chế thứ nhất, chúng ta muốn thu được nhiều thông tin không gian của vùng (đối tượng) trong ảnh, ảnh cần chia thành nhiều khối, kết quả của việc chia này là không gian lưu trữ các lược đồ màu biểu diễn ảnh tăng cao và độ phức tạp tính toán lớn. Hơn nữa, các đối

tượng của ảnh trong thực tế khó có thể ép vào các khối đều. Vì vậy, các kỹ thuật đề cập ở trên không cho các kết quả tốt [7, 40, 69].

1.3.2 Lượng hóa màu

Để sinh ra các lược đồ màu, lượng hoá màu phải được áp dụng. Lượng hoá màu là quá trình giảm số các màu được sử dụng để biểu diễn một ảnh. Một lược đồ lượng hoá được xác định bởi không gian màu và phân đoạn của không gian màu được sử dụng. Một không gian màu là biểu diễn của màu trong không gian ba chiều.

Áp dụng một lược đồ lượng hoá chuẩn trên một không gian màu, mỗi trục được chia ra thành một số phần. Khi các trục được chia ra thành k , l , và m phần, số các màu được sử dụng để biểu diễn một ảnh sẽ là $n = k.l.m$. Lượng hoá không gian màu thành n màu thường được xem như một lược đồ lượng hoá n dải. Hình 1.3 minh hoạ sự ảnh hưởng của lượng hoá các ảnh màu.



Hình 1.3. Từ trái sang: ảnh gốc sử dụng 256 màu, được lượng hoá trong 8 dải, và được lượng hoá trong 64 dải sử dụng không gian màu RGB.

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 dải được lượng hoá. Lược đồ màu được tính toán dễ dàng và hiệu quả trong mô tả phân bố màu toàn cục và cục bộ trong ảnh. Hơn nữa, lược đồ màu không nhạy cảm với quay và dịch chuyển về trục quan sát và thay đổi chậm với tỷ lệ và vị trí quan sát.

Do mọi điểm ảnh trong ảnh có thể được mô tả bởi ba thành phần màu trong một không gian màu nào đó (thí dụ, các thành phần đỏ, xanh lam và xanh lơ trong

không gian RGB, hoặc sắc màu, độ nét và giá trị trong không gian HSV), một lược đồ có thể được định nghĩa cho mỗi thành phần. Một lược đồ màu chứa nhiều dải hơn sẽ có khả năng phân biệt các ảnh tốt hơn. Tuy nhiên, điều này sẽ tăng độ phức tạp tính toán và khó khăn cho cơ chế đánh chỉ số CSDL ảnh.

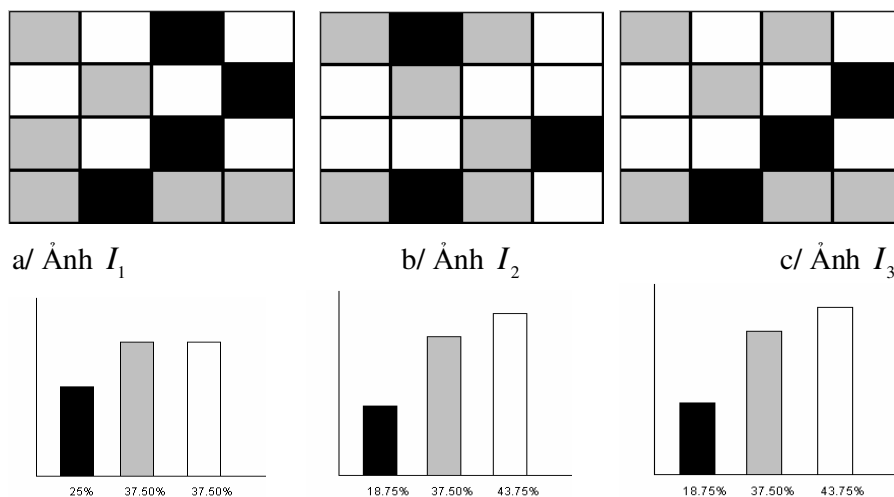
Hơn nữa, số lượng dải nhiều không cải tiến hiệu năng tra cứu trong nhiều ứng dụng. Một trong các cách để xác định số lượng các dải là sử dụng các phương pháp phân cụm để xác định K màu tốt nhất trong một không gian đã cho với một tập các ảnh đã cho và mỗi màu tốt nhất này sẽ được coi là một dải của lược đồ. Do quá trình phân cụm này lấy phân bố màu của các ảnh trên toàn bộ CSDL ảnh nên khả năng các dải lược đồ không có hoặc có rất ít điểm ảnh là cực tiểu. Một lựa chọn khác là sử dụng các dải có số điểm ảnh lớn nhất. Có lựa chọn này là do phần lớn các điểm ảnh của một ảnh thuộc về một số ít các dải của lược đồ [78]. Giảm số các dải của lược đồ theo cách này sẽ không làm giảm hiệu năng của so sánh theo lược đồ, mà còn có thể tăng cường hiệu năng, do các dải nhỏ của lược đồ coi như là nhiễu.

Khi một CSDL ảnh chứa một số lượng lớn các ảnh, so sánh theo lược đồ sẽ cho ra nhiều kết quả sai. Ngoài ra, lược đồ màu không quan tâm đến thông tin không gian của các điểm ảnh, vì thế các ảnh rất khác nhau có thể có các phân bố màu tương tự. Vấn đề này trở nên đặc biệt quan trọng với các CSDL ảnh lớn. Để giảm các kết quả sai, một số cải tiến đã được đề xuất để liên kết thông tin không gian vào lược đồ như kỹ thuật lược đồ liên kết [17]. Một số cách tiếp cận khác chia một ảnh thành các vùng con và tính lược đồ cho mỗi vùng con. Cách chia đơn giản là phân hoạch hình chữ nhật [51], phân hoạch hình quạt [48]. Cách chia phức tạp hơn là phân đoạn vùng [45, 46] hoặc thậm chí phân hoạch đối tượng [8, 25, 29]. Tăng số các vùng con sẽ tăng thông tin không gian, nhưng cũng tăng không gian lưu trữ các lược đồ màu biểu diễn ảnh và thời gian tính toán.

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 các

lược đồ màu này. Với GCH, chúng ta có thể sử dụng các độ đo khác nhau, sẽ được trình bày trong mục 1.4, để tính toán khoảng cách giữa các lược đồ màu. Ví dụ ở dưới (Hình 1.4) chỉ ra cách tính khoảng cách giữa hai ảnh sử dụng GCH.



Hình 1.4. Ba ảnh I_1 , I_2 và I_3 và các lược đồ màu tương ứng của chúng.

Trong các lược đồ màu này có ba dải: đen, xám, và trắng. Lược đồ màu của ảnh I_1 : {25%, 37.5%, 37.5%}; lược đồ màu của ảnh I_2 : {18.75%, 37.5%, 43.75%}; và ảnh I_3 có lược đồ màu giống như ảnh I_2 . Nếu chúng ta sử dụng khoảng cách Euclid để tính toán khoảng cách lược đồ, khoảng cách giữa các ảnh I_1 và I_2 theo GCH là:

$$d_{GCH}(I_1, I_2) = \sqrt{(0.25 - 0.1875)^2 + (0.375 - 0.375)^2 + (0.375 - 0.4375)^2} = 0.088$$

Khoảng cách giữa các ảnh I_1 và I_3 bằng khoảng cách giữa các ảnh I_1 và I_2 và khoảng cách giữa các ảnh I_2 và I_3 là 0.

GCH là phương pháp tra cứu ảnh truyền thống dựa vào màu. Tuy nhiên, GCH không gồm thông tin liên quan đến phân bố màu của các vùng, vì thế khoảng cách giữa các ảnh có thể không chỉ ra sự khác nhau thật giữa các ảnh. Chẳng hạn, khoảng cách giữa hai ảnh I_1 và I_3 phải nhỏ hơn khoảng cách giữa hai ảnh I_1 và I_2 , nhưng sử dụng GCH chúng ta thu được khoảng cách giống nhau. Đây là nhược điểm chính của GCH.

1.3.3.3 Lược đồ màu cục bộ LCH

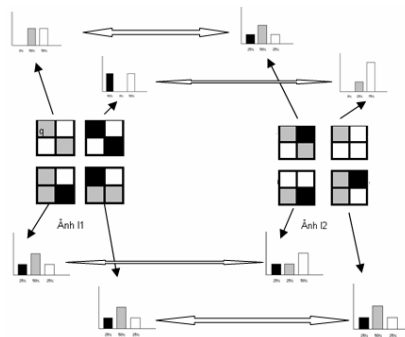
Cách tiếp cận LCH [51] gồm thông tin liên quan đến phân bố màu của các vùng. Bước đầu tiên là phân đoạn ảnh thành các khối và thu lược đồ màu cho mỗi khối. Sau đó một ảnh sẽ được biểu diễn bởi các lược đồ này. Khi so sánh hai ảnh, chúng ta tính khoảng cách lược đồ giữa một khối trong một ảnh và một khối ở cùng vị trí trong ảnh kia. Khoảng cách giữa hai ảnh sẽ được xác định bởi tổng tất cả các khoảng cách này. Nếu chúng ta sử dụng căn bậc hai của khoảng cách Euclid làm khoảng cách giữa các lược đồ, khoảng cách giữa hai ảnh Q và I theo LCH sẽ được xác định bằng:

$$d_{LCH}(Q,I) = \sum_{k=1}^M \sqrt{\sum_{i=1}^N (H_Q^k[i] - H_I^k[i])^2} \quad (1-1)$$

Ở đây M là số các khối trong ảnh, N là số các dải trong lược đồ màu, và $H_Q^k[i]$ ($H_I^k[i]$) là giá trị của dải i trong lược đồ màu H_Q^k (H_I^k), H_Q^k (H_I^k) biểu diễn khối k trong ảnh Q (I).

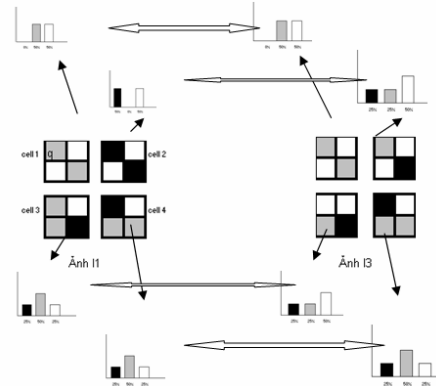
Các ví dụ ở dưới sử dụng các ảnh I_1 , I_2 và I_3 trong Hình 1.4 để chỉ ra cách tính khoảng cách giữa các ảnh sử dụng LCH.

Khoảng cách giữa I_1 và I_2 (xem Hình 1.5) được tính như sau:

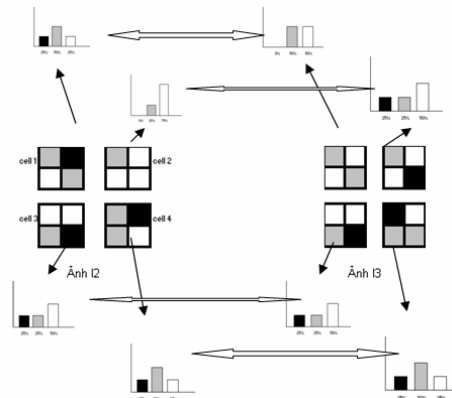


Hình 1.5. Tính khoảng cách giữa ảnh I_1 và I_2 sử dụng LCH, $d_{LCH}(I_1, I_2) = 1.319$,
 $d_{GCH}(I_1, I_2) = 0.088$.

Hình 1.6 chỉ ra việc sử dụng LCH để tính khoảng cách giữa các ảnh I_1 và I_3 (0.707), trong khi Hình 1.7 chỉ ra sự sử dụng của nó để tính khoảng cách giữa các ảnh I_2 và I_3 (1.768).



Hình 1.6. Tính khoảng cách giữa các ảnh I_1 và I_3 sử dụng LCH, $d_{LCH}(I_1, I_3) = 0.707$,
 $d_{GCH}(I_1, I_3) = 0.088$.



Hình 1.7. Tính khoảng cách giữa các ảnh I_2 và I_3 sử dụng LCH $d_{LCH}(I_2, I_3) = 0.707$,
 $d_{GCH}(I_2, I_3) = 0$.

Trong một số trường hợp sử dụng LCH có thể thu được độ chính xác tra cứu tốt hơn sử dụng GCH và các khoảng cách mới giữa các ảnh có thể hợp lý hơn các khoảng cách thu được sử dụng GCH.

Với phương pháp này, nếu muốn thu được nhiều thông tin không gian của các vùng trong ảnh, ảnh phải được chia thành nhiều khối. Điều đó làm tăng không gian lưu trữ các lược đồ biểu diễn ảnh và tăng thời gian tính toán.

1.3.3.4 Véc tơ gắn kết màu

Véc tơ gắn kết màu (CCV) [16] 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. Cho α_i biểu thị số các điểm ảnh gắn kết trong dải màu thứ i và β_i biểu thị số các điểm ảnh không gắn kết trong một ảnh. CCV của một ảnh được định nghĩa bằng véc tơ $\langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_N, \beta_N) \rangle$. Do thông tin không gian được liên kết vào lược đồ màu nên CCV cho các kết quả tra cứu tốt hơn lược đồ màu, đặc biệt cho các ảnh hoặc có phần lớn màu đồng nhất hoặc có phần lớn các vùng kết cấu.

Tuy nhiên, phương pháp này sử dụng lược đồ màu toàn cục làm cơ sở nên rất khó thu được thông tin không gian của các vùng của ảnh.

1.3.3.5 Tương quan màu

Tương quan màu [30] mô tả các phân bố màu của các điểm ảnh và chỉ ra tương quan không gian của các cặp màu. Chiều thứ nhất và thứ hai của lược đồ màu ba chiều là các màu của các cặp điểm ảnh và chiều thứ ba là khoảng cách không gian của chúng. Một tương quan màu là một bảng được đánh chỉ số bởi các cặp màu, ở đây mục thứ k cho (i, j) chỉ rõ xác suất tìm được một điểm ảnh có màu j tại một khoảng cách k từ một điểm ảnh có màu i trong ảnh. Cho I biểu diễn toàn bộ tập các điểm ảnh và $I_{c(i)}$ biểu diễn tập các điểm ảnh có màu $c(i)$. Tương quan màu xác định bằng:

$$\gamma_{i,j}^k = \text{Prob}_{p_1 \in I_{c(i)}, p_2 \in I} [p_2 \in I_{c(j)} | |p_1 - p_2| = k] \quad (1-2)$$

Ở đây $i, j \in \{1, 2, \dots, N\}$, $k \in \{1, 2, \dots, d\}$, và $|p_1 - p_2|$ là khoảng cách giữa các điểm ảnh p_1 và p_2 . Nếu xét tất cả các kết hợp có thể của các cặp màu, thì cỡ của tương quan màu sẽ rất lớn, do đó tự tương quan màu thường được sử dụng thay thế.

Tự tương quan màu chỉ sự tương quan không gian giữa các màu thuần nhất và vì thế giảm số chiều xuống.

So sánh với lược đồ màu và véc tơ gắn kết màu, tự tương quan màu cho các kết quả tra cứu tốt hơn. Tuy nhiên, tương quan màu có độ phức tạp tính toán cao, do véc tơ đặc trưng có số chiều cao.

1.3.3.6 Các màu trội

Các lược đồ màu thường rất thưa và thông thường chỉ cần số ít màu là đủ để miêu tả đặc trưng màu trong một ảnh màu, các màu trội [9, 38] 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 và phần trăm tương ứng của nó. Mỗi màu đại diện và phần trăm tương ứng này tạo ra một cặp các thuộc tính mô tả các đặc trưng màu trong một vùng ảnh.

Ký hiệu mô tả đặc trưng lược đồ màu trội F được xác định bởi một tập các cặp thuộc tính:

$$F = \{(c_i, p_i), i = 1..N\} \quad (1-3)$$

Ở đây N là tổng số các cụm màu trong ảnh, c_i là một véc tơ màu ba chiều, p_i là phần trăm của nó.

Tuy nhiên, phương pháp này cũng cho kết quả tra cứu không cao khi cơ sở dữ liệu ảnh có kích thước lớn, do nó chỉ biểu thị phân bố xác suất của các màu trội trong ảnh.

1.3.3.7 Mô men màu

Mô men màu là các mô men thống kê của các phân bố xác suất của các màu. Các mô men màu được sử dụng trong nhiều hệ thống tra cứu ảnh như QBIC [14, 74]. Các mô men màu bậc nhất (trung bình), bậc hai (phương sai) và bậc ba (độ lệch), đã được minh chứng là hiệu quả trong biểu diễn các phân bố màu của các ảnh [36].

Về mặt toán học, ba mô men đầu tiên được xác định bằng:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij} \quad (1-4)$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (1-5)$$

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right)^{\frac{1}{3}} \quad (1-6)$$

Ở đây f_{ij} là giá trị của thành phần màu thứ i của điểm ảnh j và N là số các điểm ảnh trong ảnh.

Do chỉ 9 số (ba mô men cho mỗi một trong ba thành phần màu) được sử dụng để biểu diễn đặc trưng màu của mỗi ảnh, các mô men màu là một biểu diễn rất nén so với các đặc trưng màu khác. Do biểu diễn rất nén này, các mô men màu có thể làm giảm khả năng phân biệt các ảnh. Thông thường, các mô men màu có thể được sử dụng như sơ duyệt lần đầu để giảm không gian tra cứu trước khi các đặc trưng màu phức tạp khác được sử dụng.

1.3.4 Thông tin không gian

Các vùng hoặc đối tượng với các đặc trưng màu và kết cấu tương tự có thể được phân biệt tốt hơn bằng việc kết hợp các thông tin không gian. Chẳng hạn, các vùng bầu trời màu xanh và biển xanh có thể có các lược đồ màu tương tự, nhưng thông tin không gian của chúng trong các ảnh là khác nhau. Do đó, thông tin không gian của các vùng (đối tượng) hoặc quan hệ không gian giữa nhiều vùng (đối tượng) trong một ảnh rất quan trọng cho tra cứu các ảnh.

Thu nhận thông tin không gian của các đối tượng trong một ảnh là một quá trình quan trọng trong phân biệt các ảnh. Quá trình này bao gồm việc biểu diễn vị trí không gian tuyệt đối và vị trí không gian tương đối của các đối tượng. Các thao tác như giao và chồng được sử dụng. Bố cục màu kết hợp thông tin không gian với đặc trưng màu trong ảnh tạo ra một đặc trưng rất quan trọng trong quá trình tra cứu.

Trong [47] đã đề xuất kỹ thuật sử dụng lược đồ hình quạt. Tác giả đã đề xuất một cách tiếp cận dựa vào lược đồ màu có đưa thông tin không gian vào bản miêu tả ảnh. Ban đầu ảnh được lượng hóa thành n màu và sau đó ảnh được chia thành các khối hình quạt và tính toán lược đồ của mỗi màu. Các điểm ảnh tuy có cùng màu, song chúng được phân vào các dải khác nhau tùy thuộc vào điểm ảnh thuộc khối hình quạt nào.

Biểu diễn quan hệ không gian được sử dụng rộng rãi nhất là xâu hai chiều được đề xuất bởi Chang và cộng sự [62]. Biểu diễn này được xây dựng bởi việc chiếu các ảnh dọc theo các hướng x và y . Hai tập ký hiệu V và A được xác định trên hình chiếu. Mỗi ký hiệu thuộc V biểu diễn một đối tượng trong ảnh. Mỗi ký hiệu thuộc A biểu diễn quan hệ không gian giữa các đối tượng. Một cải tiến của xâu hai chiều [61] là cắt tất cả các đối tượng dọc theo hộp bao tối thiểu và mở rộng các quan hệ không gian thành hai tập toán tử không gian. Một tập xác định quan hệ không gian cục bộ. Tập còn lại xác định quan hệ không gian toàn cục, quan hệ không gian toàn cục này chỉ ra rằng hình chiếu của hai đối tượng là rời nhau, liền kề hoặc chồng lên nhau.

Ngoài ra, cây tứ phân không gian [23] và biểu diễn các vùng liên quan với đối tượng ảnh [73] cũng được sử dụng cho biểu diễn thông tin không gian. Tuy nhiên, tra cứu ảnh dựa trên các quan hệ không gian của các vùng khó được thực hiện do phân đoạn các đối tượng hoặc các vùng thường cho độ chính xác không cao.

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. Tuy nhiên, trong hầu hết các trường hợp chúng ta chỉ nhận được các vùng. Đây là bước rất quan trọng đối với tra cứu ảnh. Cả đặc trưng hình và đặc trưng bố cục phụ thuộc vào phân vùng tốt. Trong phần này chúng tôi sẽ mô tả một số kỹ thuật phân vùng đã được sử dụng trong tra cứu ảnh.

Trong [35], Lybanon và cộng sự đã nghiên cứu một cách tiếp cận phép toán hình thái học trong phân vùng ảnh. Cách tiếp cận này hiệu quả trong xử lý các loại ảnh thiên văn và các ảnh hồng ngoại, nhưng hiệu năng của nó cần được tiếp tục đánh giá cho các ảnh tự nhiên phức tạp hơn. Trong [77], Li và cộng sự đã đề xuất cách tiếp cận phân vùng dựa vào entropy mờ. Cách tiếp cận này hiệu quả cho các ảnh có lược đồ không có các đỉnh và các rãnh rõ ràng. Các kỹ thuật phân vùng khác dựa trên phép đặc tam giác Delaunay, fractals và luồng biên [3, 28, 77].

Các thuật toán được đề cập ở trên là tự động. Ưu điểm chính của các thuật toán phân vùng loại này là nó trích rút các đường biên từ một số lượng lớn các ảnh mà không chiếm nhiều thời gian và công sức của con người. Tuy nhiên, với các ảnh tự nhiên nói chung phân vùng tự động thường không cho kết quả chính xác. Một thuật toán có thể phân vùng trong trường hợp này chỉ là các vùng, mà không là các đối tượng.

Trong [53], Samadani và Han đã đề xuất một cách tiếp cận trích rút đường biên được trợ giúp bởi máy tính. Kỹ thuật kết hợp các đầu vào từ người sử dụng với các biên ảnh được sinh ra bởi máy tính.

Trong các ứng dụng thị giác, các yêu cầu về phân vùng rất khác nhau cho các đặc trưng hình dạng và các đặc trưng bố cục. Với các đặc trưng hình dạng, phân vùng tốt là rất quan trọng, trong khi các đặc trưng bố cục chỉ cần một phân vùng thô.

1.4 Các độ đo tương tự

Các kết quả tra cứu thu được bởi so sánh sự tương tự giữa các đặc trưng của ảnh CSDL và ảnh truy vấn.

Người ta chọn độ đo dựa trên các tiêu chí sau đây:

Phù hợp với cảm nhận trực quan: Các ảnh thường hay được mô tả trong không gian đặc trưng và độ tương tự giữa các ảnh thường được đo bởi một độ đo

khoảng cách trong không gian đặc trưng. Tuy nhiên, nếu kết hợp thêm cảm nhận của con người vào bản miêu tả ảnh, chúng ta sẽ thu được kết quả tốt hơn.

Hiệu quả: Độ đo cần được tính toán nhanh để có phản hồi nhanh trong khi tra cứu. Các ứng dụng VFBIR đòi hỏi một phản hồi rất nhanh, không lâu hơn vài giây. Trong chu kỳ thời gian ngắn đó, quá trình tra cứu thường phải tính toán hàng ngàn khoảng cách phụ thuộc vào cỡ của CSDL ảnh. Do đó độ phức tạp của độ đo khoảng cách rất quan trọng.

Khả năng: Hiệu năng của hệ thống không được giảm quá nhiều cho các CSDL lớn do một hệ thống có thể tra cứu trong các CSDL chứa hàng triệu ảnh. Khi thực hiện, hệ thống VFBIR phải tính toán tất cả các khoảng cách giữa ảnh truy vấn và các ảnh trong CSDL. Sau đó các khoảng cách này được lưu trữ để tìm ra các ảnh tương tự nhất đối với ảnh truy vấn. Do đó độ phức tạp của quá trình tra cứu phải tương ứng với cỡ của CSDL ảnh.

Các tính chất đối với khoảng cách:

- Sự bất biến của độ đo tương tự: Khoảng cách giữa một ảnh với bản thân nó phải bằng với một hằng số độc lập với ảnh.

$$d(A, A) = d(B, B) \quad (1-7)$$

- Tối thiểu: Một ảnh phải tương tự với chính nó hơn là với các ảnh khác.

$$d(A, A) < d(A, B) \quad (1-8)$$

- Tính đối xứng: Nếu ảnh A tương tự với ảnh B thì ảnh B phải tương tự với ảnh A .

$$d(A, B) = d(B, A) \quad (1-9)$$

- Tính bắc cầu: nếu ảnh A tương tự với ảnh B , và B tương tự với C thì C tương tự với A .

Tuy nhiên, tính chất bắc cầu này có thể không giữ cho một chuỗi các ảnh. Ngay khi nếu ảnh I_i tương tự với ảnh I_{i+1} với $i = 1..N$. Điều này không có

nghĩa rằng ảnh I_1 tương tự với ảnh I_N . Thí dụ, trong một cảnh video mỗi frame tương tự với các frame lân cận của nó nhưng frame đầu tiên và frame cuối cùng của cảnh có thể rất khác nhau.

Tính chất bền vững: Hệ thống phải ổn định đối với các thay đổi trong các điều kiện ảnh của các ảnh CSDL.

Dưới đây là một số độ đo tương tự được sử dụng phổ biến.

- Lược đồ giao [69]:

Khoảng cách được xác định dựa trên kích thước của phần chung của hai lược đồ màu. Giả sử hai lược đồ màu được ký hiệu là h_1 và h_2 , khoảng cách giữa chúng có thể được xác định bằng:

$$dist_{HI} = 1 - \sum_{i=1}^N \min(h_{1i}, h_{2i}) \quad (1-10)$$

- Khoảng cách L_1 [36], khoảng cách dạng Minkowski L_p : khoảng cách dạng Minkowski L_p giữa hai lược đồ được xác định như sau:

$$dist_{Mp} = \left(\sum_i |h_{1i} - h_{2i}|^p \right)^{1/p} \quad (1-11)$$

- Khoảng cách dạng toàn phương [26]: khoảng cách giữa hai lược đồ màu N chiều h_1 và h_2 được xác định như sau:

$$dist_{QF} = (h_1 - h_2)' A (h_1 - h_2) \quad (1-12)$$

Ở đây $A = [a_{ij}]$ là một ma trận và các trọng số a_{ij} biểu thị sự tương tự giữa các dải i và j . Thông thường a_{ij} được cho bằng

$$a_{ij} = 1 - (d_{ij} / d_{max})^k \quad (1-13)$$

Ở đây d_{ij} là khoảng cách giữa màu i và màu j (thông thường d_{ij} là khoảng cách Euclid giữa hai màu trong một số không gian màu đồng nhất như La^*b^*

hoặc $Lu*v^*$) và $d_{max} = \max_{ij}(d_{ij})$. k là hằng số điều khiển trọng số giữa các màu lân cận.

- Khoảng cách EMD [60] dựa trên giá trị tối thiểu để biến đổi một phân bố thành một phân bố khác.

$$dist_{EMD} = \frac{\sum_{ij} g_{ij} d_{ij}}{\sum_{ij} g_{ij}} \quad (1-14)$$

Ở đây d_{ij} biểu thị sự không tương tự giữa các dải i và j , và $g_{ij} \geq 0$ là luồng tối ưu giữa hai phân bố sao cho tổng giá trị

$$dist_{EMD} = \sum_{ij} g_{ij} d_{ij} \quad (1-15)$$

là cực tiểu, tùy thuộc vào các ràng buộc sau:

$$\begin{aligned} \sum_i g_{ij} &\leq h_{1i} \\ \sum_j g_{ij} &\leq h_{2i} \\ \sum_{ij} g_{ij} &= \min(h_{1i}, h_{2i}) \end{aligned} \quad (1-16)$$

Với tất cả i và j . Mũ số trong công thức (1-14) là một hệ số chuẩn hoá cho phép đối sánh các phần của các phân bố với tổng khối lượng khác nhau.

Ngoài các độ đo khoảng cách trên, người ta còn quan tâm đến các độ đo khoảng cách sau:

- Khoảng cách Kolmogorov-Smirnov được đề xuất trong [18] được xác định bởi sự khác nhau cực đại giữa các phân bố tích lũy

$$dist_{Mp} = \max_i |h_{1i}^c - h_{2i}^c| \quad (1-17)$$

Ở đây h^c là lược đồ tích lũy của lược đồ h

- Thông kê kiểu Cramer/Von Mises dựa trên các phân bố tích lũy được xác định

$$dist_C = \sum_i (h_{1i}^c - h_{2i}^c)^2 \quad (1-18)$$

- Phân bố thống kê χ^2 được xác định bởi

$$dist_{\chi} = \sum_i \frac{\left(h_{1i} - \hat{h}_i \right)^2}{\hat{h}_i} \quad (1-19)$$

Ở đây $\hat{h}_i = \frac{h_{1i} + h_{2i}}{2}$ biểu thị ước lượng chung.

- Sự chênh lệch Kullback-Leibler được xác định bởi

$$dist_{KL} = \sum_i h_{1i} \log \frac{h_{1i}}{h_{2i}} \quad (1-20)$$

- Sự chênh lệch Jeffrey được xác định bởi

$$dist_{JD} = \sum \left(h_{1i} \log \frac{h_{1i}}{\hat{h}_i} + h_{2i} \log \frac{h_{2i}}{\hat{h}_i} \right) \quad (1-21)$$

- Bất biến trung bình trọng số [4] được xác định bởi:

$$dist_{WMV} = \frac{\mu_1 - \mu_2}{\sigma(\mu)} + \frac{\sigma_1 - \sigma_2}{\sigma(\sigma)} \quad (1-22)$$

Ở đây μ_1, μ_2 là các tham số thực nghiệm và σ_1, σ_2 là các độ lệch chuẩn của hai lược đồ h_1, h_2 . $\sigma(\cdot)$ biểu thị sự ước lượng của độ lệch chuẩn của thực thể tương ứng.

- Khoảng cách Bhattacharyya [15] được xác định bởi:

$$d_B^2(N(\mu_1, \Sigma_1), N(\mu_2, \Sigma_2)) = \frac{1}{8} (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2) + \frac{1}{2} \ln \frac{\det \Sigma}{\sqrt{\det \Sigma_1 \det \Sigma_2}} \quad (1-23)$$

Ở đây $\Sigma = 0.5 \times (\Sigma_1 + \Sigma_2)$

- Khoảng cách Mahalanobis [15] được xác định bởi

$$d_B^2(N(\mu_1, \Sigma), N(\mu_2, \Sigma)) = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2) \quad (1-24)$$

1.5 Đánh giá hiệu năng tra cứu

Nhiều kỹ thuật đã sử dụng các độ đo hiệu năng lấy từ tra cứu thông tin [71] hoặc các độ đo mới được phát triển cho tra cứu ảnh dựa vào đặc trưng thị giác [21, 39] để đánh giá hiệu năng tra cứu.

Để đánh giá một ứng dụng tra cứu ảnh dựa vào đặc trưng thị giác, CSDL ảnh và 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 một 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 CSDL.

Một trong các vấn đề để tạo CSDL ảnh là cỡ của CSDL phải đủ lớn và các ảnh phải đủ đa dạng thuộc các lĩnh vực khác nhau. Với tra cứu văn bản truyền thống, CSDL thường có hàng triệu tài liệu [71] trong khi trong VFBIR hầu hết các hệ thống làm việc chỉ với vài ngàn ảnh, thậm chí còn ít hơn.

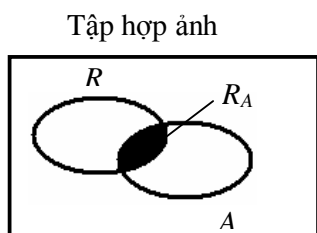
Ngay khi các ảnh được tập hợp, nhiệm vụ tiếp theo trong đánh giá hiệu năng của ứng dụng VFBIR là xác định một tập các truy vấn và chọn các ảnh làm cơ sở đánh giá dựa trên CSDL ảnh vào.

Khi CSDL ảnh được tập hợp và các truy vấn và các ảnh liên quan đến ảnh truy vấn trong CSDL được lựa chọn, các ảnh truy vấn được biểu diễn một-một đối với quá trình tra cứu của ứng dụng VFBIR, sau đó các kết quả tra cứu được so sánh với các ảnh liên quan đến ảnh truy vấn trong CSDL tương ứng. Đồ thị Recall - Precision [5] là một cách để đánh giá các kết quả tra cứu cho các hệ thống tra cứu ảnh dựa vào đặc trưng thị giác.

Ký hiệu R là tập các ảnh liên quan đến ảnh truy vấn trong CSDL ảnh, A là tập các ảnh được tra cứu từ CSDL ảnh, R_A là tập các ảnh liên quan với ảnh truy vấn trong tập A (xem Hình 1.8).

$$recall = \frac{|R_A|}{|R|} \quad (1-25)$$

$$precision = \frac{|R_A|}{|A|} \quad (1-26)$$



Hình 1.8. Recall và Precision cho các kết quả truy vấn.

Nếu có nhiều truy vấn, chúng ta có thể tính độ chính xác trung bình của tất cả các truy vấn:

$$precision_{avg}(l) = \sum_{i=1}^{|Q|} \frac{precision_i(l)}{|Q|} \quad (1-27)$$

Ở đây $precision_{avg}(l)$ là độ chính xác trung bình của mức l , $precision_i(l)$ là độ chính xác của truy vấn i tại mức l , $|Q|$ là số các truy vấn.

1.6 Các hệ thống VFBIR

Tra cứu ảnh dựa vào đặc trưng thị giác đã được nhiều nhà nghiên cứu quan tâm trong những năm gần đây. Nhiều hệ thống tra cứu ảnh đã được xây dựng. Dưới đây sẽ mô tả ngắn gọn một số hệ thống VFBIR đã được phát triển.

QBIC của IBM: QBIC là hệ thống tra cứu ảnh dựa vào đặc trưng thị giác thương mại đầu tiên. Các kỹ thuật được sử dụng trong hệ thống đã ảnh hưởng nhiều đến các hệ thống tra cứu ảnh về sau. QBIC hỗ trợ chính các truy vấn theo ảnh mẫu, các phác thảo được người sử dụng xây dựng và các mẫu kết cấu và màu được lựa chọn.

Đặc trưng màu sử dụng trong QBIC là trung bình (R,G,B), (Y,i,q), (L,a,b), các tọa độ MTM và một lược đồ k thành phần [13]. Đặc trưng kết cấu được sử dụng là một phiên bản cải tiến của biểu diễn kết cấu Tamura [24]. Đặc trưng hình dạng được sử dụng bao gồm diện tích hình, hình tròn, độ lệch tâm, hướng trục chính [11] và một tập các mô men bất biến [13, 65]. QBIC là một trong ít hệ thống sử dụng cơ chế đánh chỉ số đặc trưng nhiều chiều. Trong cách đánh chỉ số, KLT được sử dụng đầu tiên để giảm số chiều và sau đó cây R^* được sử dụng như cấu trúc đánh chỉ số nhiều chiều [33].

Blobworld: Blobworld [6] là hệ thống tra cứu ảnh dựa vào tìm kiếm các vùng ảnh gắn kết tương ứng với các đối tượng. Mỗi ảnh được tự động phân đoạn thành các vùng kết hợp với các mô tả màu và kết cấu. Truy vấn dựa vào các đặc tính của một hoặc hai vùng quan tâm. Để tra cứu trên các cơ sở dữ liệu lớn, hệ thống đánh chỉ số các mô tả vùng sử dụng cấu trúc cây.

RetrievalWare: RetrievalWare là một máy tra cứu ảnh dựa vào đặc trưng thị giác được phát triển bởi tập đoàn công nghệ Excalibur. Hệ thống đã ứng dụng mạng Noron để tra cứu ảnh [10]. Phiên bản gần đây sử dụng màu, hình, kết cấu, độ sáng, bố cục màu, và hướng tỷ lệ của ảnh làm các đặc trưng để truy vấn. Hệ thống cũng hỗ trợ các kết hợp của các đặc trưng này và cho phép người sử dụng điều chỉnh các trọng số kết hợp với mỗi đặc trưng.

VisualSeek và WebSeek: VisualSEEK [67] là máy tra cứu dựa vào đặc trưng thị giác và WebSEEK [68] là máy tìm kiếm văn bản/ảnh trên web, cả hai sản phẩm đã được phát triển tại đại học Columbia. Các đặc điểm nghiên cứu chính là truy vấn quan hệ không gian của các vùng ảnh và trích rút đặc trưng thị giác lĩnh vực. Các đặc trưng thị giác được sử dụng trong các hệ thống là các tập màu và các đặc trưng kết cấu dựa vào biến đổi sóng. Để tăng tốc quá trình tra cứu, hệ thống đã sử dụng cơ chế đánh chỉ số dựa vào cây nhị phân. VisualSEEK hỗ trợ các truy vấn dựa vào cả các đặc trưng thị giác và các quan hệ không gian. Điều này cho phép người sử dụng chuyển một truy vấn cảnh hoàng hôn bằng cảnh có vùng màu đỏ-cam trên đỉnh và

vùng xanh lơ hoặc xanh lá cây ở phía dưới bằng phác thảo. WebSEEk là một máy tìm kiếm hướng Web bao gồm ba module chính: module tập hợp ảnh/video, module phân loại chủ đề và đánh chỉ số, và module tìm kiếm, duyệt, và tra cứu. Hệ thống hỗ trợ các truy vấn dựa trên cả các từ khoá và đặc trưng thị giác.

CIRES: CIRES [49] là một hệ thống tra cứu ảnh dựa vào nội dung trong các thư viện ảnh số khá ổn định và được phát triển bởi đại học Texas. Hệ thống tra cứu sử dụng các đặc trưng gồm màu (các lược đồ màu), kết cấu (lọc Gabor) và sử dụng kỹ thuật nhóm các vùng đồng nhất cảm nhận. Hệ thống có thể thực hiện trên các ảnh chứa các đối tượng như cây, bầu trời, các tòa nhà, các cây cầu,...

Hệ thống trong [28] đã đề xuất cách tiếp cận để tăng hiệu năng tra cứu. Hệ thống sử dụng tiền phân lớp để cải tiến các kết quả tra cứu. Họ đề xuất phân loại lại các ảnh thành các loại như đồ thị/bức ảnh, kết cấu/không kết cấu. Các đặc trưng dựa vào vùng tương tự với cách tiếp cận trong BlopWorld được sử dụng, nhưng các mô tả vùng của các ảnh được đối sánh tự động.

Các nỗ lực trong lĩnh vực tra cứu ảnh cũng được thực hiện cho các ứng dụng y học. Dự án tra cứu ảnh trong các ứng dụng y học (IRMA) [80]. Mục tiêu của dự án là phát triển và thực thi các phương pháp mức cao cho tra cứu ảnh dựa vào nội dung đối với các công việc chuẩn đoán.

Tìm kiếm ảnh của Google: Tìm kiếm ảnh của tập đoàn Google [22] cho phép người sử dụng tìm kiếm trên Web theo nội dung ảnh. Các từ khóa cho tìm kiếm ảnh dựa vào tên tệp tin ảnh, văn bản mô tả ảnh và văn bản liên quan đến ảnh. Khi tìm kiếm ảnh, mỗi văn bản mô tả của ảnh truy vấn được so sánh với văn bản mô tả của các ảnh trong cơ sở dữ liệu. Các ảnh có độ tương tự cao nhất được hiển thị.

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 [79].

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 [16]. 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 [42, 43].

Thứ hai, chúng tôi sẽ đề xuất phương pháp kết hợp thông tin màu và không gian trong quá trình tra cứu [45, 46] nhằm nâng cao hiệu năng tra cứu.

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

Trong chương này, chúng tôi sẽ trình bày phương pháp tra cứu ảnh dựa vào đặc trưng thị giác sử dụng lược đồ màu khối. Trên cơ sở phương pháp lược đồ màu khối chúng tôi trình bày phương pháp tra cứu đề xuất, có tên là HG, sử dụng ít không gian lưu trữ các lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển. Bên cạnh đó, chúng tôi cũng trình bày một cải tiến hiệu năng của phương pháp HG, có tên là IHG, phương pháp này sẽ tăng tốc độ và độ chính xác của phương pháp HG.

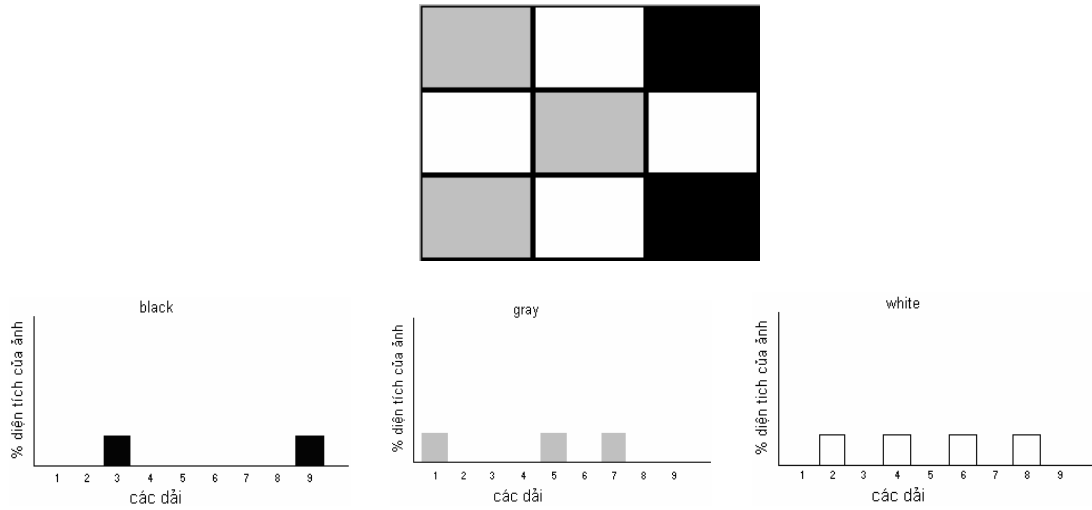
2.1 Lược đồ màu khối

Trong mục này, chúng tôi sẽ giới thiệu sơ lược cách tiếp cận lược đồ màu khối (CCH) sử dụng trong tra cứu ảnh [54].

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.

Trong cách biểu diễn này, một ảnh được cấu tạo bởi l màu sẽ được mô tả bởi l lược đồ màu khối, mỗi lược đồ màu khối mô tả phân bố không gian của một màu trong ảnh. Nếu một màu không xuất hiện trong một ảnh, không có lược đồ màu khối được lưu trữ. Hình 2.1 chỉ ra một ảnh ba màu với 9 khối (3×3), các khối được đánh thứ tự từ trái qua phải và từ trên xuống dưới và lược đồ màu khối của mỗi màu tương ứng.



Hình 2.1. Một ảnh được chia thành 9 khối ảnh và ba lược đồ màu khối của nó.

Xác định khoảng cách giữa hai ảnh

Khoảng cách giữa ảnh truy vấn h_{I_1} và ảnh CSDL h_{I_2} được xác định dựa trên độ đo khoảng cách L_1 (công thức 2-1).

$$D(h_{I_1}[i], h_{I_2}[i]) = \sum_{j=1}^{m \times m} |h_{I_1}[i][j] - h_{I_2}[i][j]| \quad (2-1)$$

Ở đây $h_{I_1}[i][j]$ và $h_{I_2}[i][j]$ biểu diễn dải thứ j của lược đồ màu khối thứ i được sử dụng để biểu diễn ảnh truy vấn h_{I_1} và ảnh CSDL h_{I_2} tương ứng.

Để chuẩn hoá kết quả thu được với khoảng cách L_1 , kết quả thu được sẽ đem chia cho tổng diện tích của các vùng được mô tả bởi mỗi lược đồ ($a_{I_1}[i] + a_{I_2}[i]$). Các diện tích này cũng được chuẩn hoá với các cỡ của ảnh. Khoảng cách được chuẩn hoá D_n được chỉ ra trong công thức 2-2.

$$D_n(h_{I_1}[i], h_{I_2}[i]) = \frac{D(h_{I_1}[i], h_{I_2}[i])}{a_{I_1}[i] + a_{I_2}[i]} \quad (2-2)$$

Sử dụng CCH, chúng ta sẽ thu được độ chính xác tra cứu tốt hơn GCH. Bên cạnh đó, sử dụng CCH, biểu diễn ảnh sẽ cô đọng hơn khi sử dụng cách tiếp cận LCH [54].

2.2 Phương pháp tra cứu dựa vào lược đồ màu khối

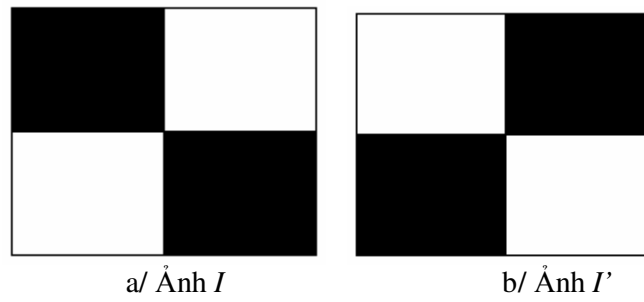
2.2.1 Giới thiệu

Lược đồ màu toàn cục GCH là một kỹ thuật áp dụng các đặc trưng màu đơn giản và hiệu quả. Kỹ thuật này có ưu điểm bất biến với quay, tỷ lệ và tính toán rất đơn giản. Tuy nhiên, kỹ thuật này chỉ đem sự phân bố của các màu vào bản miêu tả ảnh mà không quan tâm đến các quan hệ của các màu và vị trí không gian của các màu trong ảnh [41, 59, 72, 79].

Để khắc phục hạn chế của GCH, lược đồ màu cục bộ LCH đã được đề xuất. Với LCH, một ảnh được chia thành một số khối có kích cỡ bằng nhau và khoảng cách giữa hai ảnh là tổng các khoảng cách lược đồ của các khối tương ứng. Tuy nhiên, phương pháp này sử dụng nhiều không gian để lưu trữ các lược đồ màu biểu diễn ảnh và không có khả năng xử lý đối với các biến đổi hình học như quay và dịch chuyển [56] và biến đổi vị trí không gian.

Để giảm chi phí không gian lưu trữ các lược đồ màu biểu diễn ảnh của LCH, phương pháp CCH [54] được đề xuất. Tuy nhiên, phương pháp này cũng 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 CSDL. Vì vậy, khi ảnh bị quay hoặc dịch chuyển, nó không cho kết quả tốt. Ví dụ 2.1 ở dưới minh họa cho điều này.

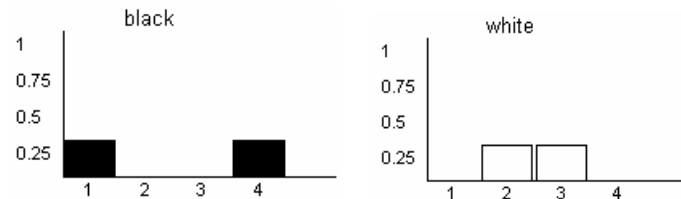
Ví dụ 2.1. Xét ảnh I và ảnh I' được cho trong Hình 2.2, ảnh I' là kết quả của việc điều chỉnh ảnh I quay một góc 90^0 .



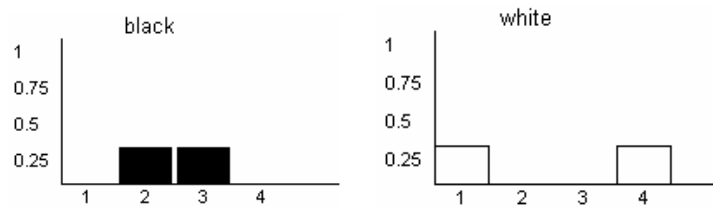
Hình 2.2. Ảnh I và ảnh I' .

Phương pháp CCH thực hiện tính khoảng cách giữa hai ảnh I và I' ở trên như sau:

Đầu tiên, lược đồ màu khối theo màu black và white của ảnh I và I' được tính như được chỉ ra trong Hình 2.3 và Hình 2.4 tương ứng.

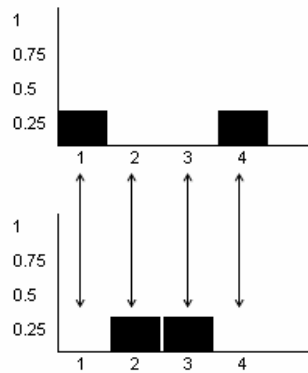


Hình 2.3. Lược đồ màu khối theo màu black và white biểu diễn ảnh I .



Hình 2.4. Lược đồ màu khối theo màu black và white biểu diễn ảnh I' .

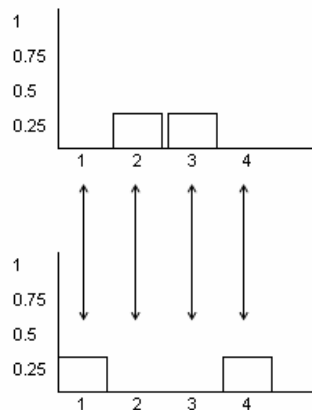
Tiếp theo, khoảng cách giữa ảnh I và I' theo màu black và white được thực hiện như chỉ ra trong Hình 2.5 và Hình 2.6.



Hình 2.5. Tính khoảng cách của ảnh I và I' theo màu black.

Tính khoảng cách của ảnh I và I' theo màu black. Khoảng cách này có giá trị:

$$D_n(I, I')_{black} = \frac{|0.25 - 0| + |0 - 0.25| + |0 - 0.25| + |0.25 - 0|}{2} = 0.5.$$



Hình 2.6. Tính khoảng cách của ảnh I và I' theo màu white.

Tính khoảng cách của ảnh I và I' theo màu white. Khoảng cách này có giá trị:

$$D_n(I, I')_{white} = \frac{|0 - 0.25| + |0.25 - 0| + |0.25 - 0| + |0 - 0.25|}{2} = 0.5$$

Khoảng cách giữa ảnh I và I' : $D_n(I, I') = D_n(I, I')_{black} + D_n(I, I')_{white} = 1$

Nhận xét: Hai ảnh I và I' xét trong Ví dụ 2.1 đáng lẽ phải giống nhau. Tuy nhiên, sử dụng cách tính của CCH, hai ảnh này lại rất khác nhau.

Để khắc phục hạn chế trên của CCH, chúng tôi đề xuất phương pháp HG [42].

2.2.2 Phương pháp tra cứu đề xuất HG

2.2.2.1 Khái niệm về đồ thị hai phía

Mục này sẽ giới thiệu một số khái niệm cơ bản về đồ thị hai phía [50]:

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

Trong phần này chúng tôi trình bày phương pháp HG. Phương pháp này đã được chúng tôi công bố trong [42].

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 thuật toán HG:

Tiếp theo, chúng tôi mô tả chi tiết thuật toán HG [42] 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
 - 1.1 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 $D \leftarrow D + MCM(G(X, Y, E, c), n)$
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^2 đỉnh ở phía bên trái tương ứng với $n \times n$ dải của lược đồ màu khối $H(I_1, c, n)$ và n^2 đỉnh ở phía bên phải tương ứng với $n \times n$ dải của lược đồ màu khối $H(I_2, c, n)$. Hàm $MCM(,)$ trả lại khoảng cách giữa hai ảnh theo màu c đã cho.

Mô tả thực hiện của thuật toán HG:

Đầu tiên, thuật toán HG lượng hoá các màu của ảnh truy vấn và ảnh CSDL và tập màu này được ký hiệu là C . Sau đó HG chia ảnh thành các khối ảnh có kích cỡ bằng nhau, xây dựng lược đồ màu khối cho mỗi màu c ($0 < c \leq C$). Trong lược đồ màu khối này, mỗi dải của lược đồ màu khối theo màu c chỉ ra số điểm ảnh có màu c trong khối ảnh tương ứng. Khi xây dựng đồ thị hai phía có trọng số $G(X, Y, E, c)$ theo màu c , mỗi dải của lược đồ màu khối tương ứng với một đỉnh. Trong đồ thị $G(X, Y, E, c)$, trọng số của mỗi cạnh nối hai đỉnh là khoảng cách giữa hai đỉnh và khoảng cách giữa hai ảnh theo màu c là giá trị đối sánh cực tiểu trên đồ thị này. Do đó, khoảng cách giữa hai ảnh có thể được tính bằng tổng tất cả các khoảng cách giữa hai ảnh của tất cả các màu $c \in C$. Chẳng hạn, nếu ảnh truy vấn và ảnh CSDL được lượng hoá thành ba màu (black, gray và white) và mỗi ảnh được chia thành bốn khối ảnh thì chúng ta sẽ thu được ba đồ thị hai phía có trọng số $G(X, Y, E, black)$, $G(X, Y, E, gray)$ và $G(X, Y, E, white)$ tương ứng. Mỗi đồ thị này sẽ gồm tám đỉnh, bốn đỉnh ở phía bên trái của đồ thị tương ứng với bốn dải của lược đồ màu khối của ảnh truy vấn và bốn đỉnh ở phía bên phải tương ứng với bốn dải của lược đồ màu khối của ảnh CSDL. Tiếp theo, để tính khoảng cách giữa ảnh truy vấn và ảnh CSDL, chúng ta lần lượt tính khoảng cách của từng màu black, gray và white thông qua các đồ thị $G(X, Y, E, black)$, $G(X, Y, E, gray)$ và $G(X, Y, E, white)$ tương ứng và tính tổng khoảng cách theo các màu black, gray và white này.

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 $G(X, Y, E, c)$ thông qua thuật toán Hungarian [27]. 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 .

Trong hàm MCM, chúng tôi có sử dụng công thức:

$$w(i,j) \leftarrow |h_{I_1}[c][i] - h_{I_2}[c][j]| \quad (2-3)$$

Trong công thức này $h_{I_1}[c][i]$ biểu diễn dải thứ i của lược đồ theo màu c của ảnh truy vấn I_1 và $h_{I_2}[c][j]$ biểu diễn dải thứ j của lược đồ theo màu c của ảnh CSDL I_2 .

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,

Chứng minh: Để tính trọng số của tất cả các cạnh trong đồ thị $G(X, Y, E, c)$ ở bước 1 cần thời gian là $O(n^4)$. Để tính giá trị đối sánh cực tiểu của đồ thị $G(X, Y, E, c)$ ở bước 2 cần thời gian là $O(n^4)$ [27]. Hơn nữa, ở bước 3 hàm cần n^2 phép tính.

Do đó, độ phức tạp thời gian của hàm MCM là $O(n^4)$. Mệnh đề đã được chứng minh .

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,

Chứng minh: Độ phức tạp của hàm MCM là $O(n^4)$. Bên cạnh đó, quá trình tính toán các lược đồ của một ảnh đòi hỏi thời gian là $O(n^2)$. Hơn nữa, quá trình tính khoảng cách giữa hai ảnh cần $O(n^4)$, do nó cần thực hiện k lần (với $k=|C|$ và $k \ll n$) hàm $MCM(,)$ để tính khoảng cách giữa hai ảnh theo màu c .

Do đó, độ phức tạp thời gian của thuật toán HG là $O(n^4)$. Mệnh đề đã được chứng minh .

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.

Chứng minh: Giả sử lược đồ màu khối theo một màu của ảnh Q gồm $k \times k$ dải ($k=2,3,4,\dots,n$). Chúng tôi nhận thấy rằng, khi ảnh Q bị quay một góc $l\frac{\pi}{2}$ (với $l=1,2,3$), vị trí của các dải của lược đồ màu khối theo một màu của ảnh Q' (kết quả

của việc quay ảnh Q một góc $k\frac{\pi}{2}$) sẽ thay đổi. Điều này làm cho khoảng cách của một dải của lược đồ màu khối theo một màu của ảnh Q với một dải của lược đồ màu khối theo một màu ở vị trí tương ứng trong lược đồ màu khối theo một màu của ảnh Q' sẽ không còn thích hợp.

Hơn nữa, phương pháp CCH chỉ so sánh mỗi dải của lược đồ màu khối theo một màu của ảnh Q với một dải của lược đồ màu khối theo một màu ở vị trí tương ứng có cùng màu trong ảnh Q' và coi khoảng cách của hai ảnh là tổng các khoảng cách này. Điều này làm cho khoảng cách của ảnh Q và Q' không chính xác.

Ngoài ra, phương pháp HG so sánh mỗi dải của lược đồ màu khối theo một màu của ảnh Q với tất cả các dải của lược đồ màu khối theo một màu của ảnh Q' , do đó khoảng cách của ảnh Q và Q' thu được bởi HG sẽ không bị thay đổi.

Vậy, độ nhạy cảm với phép quay của phương pháp HG ít hơn phương pháp CCH. Đây là điều cần chứng minh .

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.

Chứng minh: Giả sử lược đồ màu khối theo một màu của ảnh Q gồm $k \times k$ dải ($k=2,3,4,\dots,n$). Khi ảnh Q bị dịch chuyển, chúng ta sẽ nhận được một ảnh Q' . Rõ ràng rằng, lược đồ màu khối theo một màu của ảnh Q' chính là lược đồ màu khối theo một màu của ảnh Q cộng thêm một lượng nhiễu. Tuy nhiên, chỉ một số dải của lược đồ màu khối theo một màu của ảnh bị cộng thêm nhiễu (có dải không ảnh hưởng nhiễu), các dải chứa nhiễu chủ yếu là các dải tương ứng với các khối nằm ở cạnh của ảnh. Điều này làm cho khoảng cách mỗi dải của lược đồ màu khối theo một màu của ảnh Q với mỗi dải của lược đồ màu khối theo một màu ở vị trí tương ứng của ảnh Q' sẽ không phù hợp.

Bên cạnh đó, phương pháp CCH chỉ so sánh mỗi dải của lược đồ màu khối theo một màu của ảnh Q với một dải của lược đồ màu khối theo một màu tương ứng của ảnh Q' . Điều này làm cho khoảng cách của ảnh Q và Q' tính được bởi CCH không chính xác.

Hơn nữa, phương pháp HG so sánh mỗi dải của lược đồ màu khối theo một màu của ảnh Q với tất cả các dải của lược đồ màu khối theo một màu của ảnh Q' , do đó khoảng cách của ảnh Q và Q' thu được bởi HG sẽ không bị ảnh hưởng nhiều.

Vậy, độ nhạy cảm với phép dịch chuyển của phương pháp HG ít hơn phương pháp CCH. Đây là điều cần chứng minh .

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 trong [43].

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

Mục này giới thiệu khái niệm về sự tương tự lý tưởng của hai dải của hai lược đồ màu khối tương ứng.

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 thoả 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 nhiều 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 [43].

2.3.3 Phương pháp IHG

Ý tưởng cơ bản của phương pháp IHG:

Chỉ số 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 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

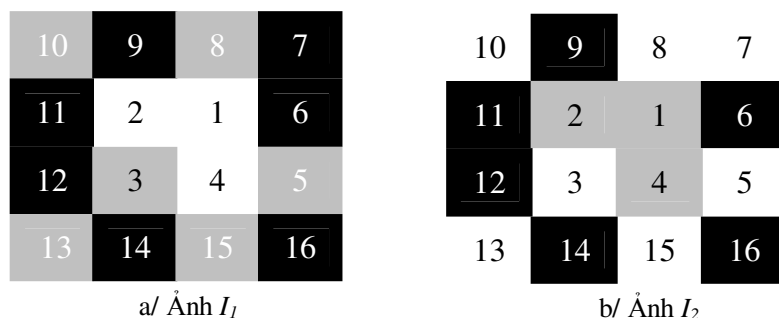
1. $dis \leftarrow 0$
2. $C \leftarrow \text{Quantization}(I_1, I_2)$
3. For mỗi c in C do
 - 3.1 $dis \leftarrow dis + \text{DistancebyColor}(I_1, I_2, n, c)$
4. Trả lại giá trị dis

Mô tả thuật toán IHG:

Đầu tiên, thuật toán tính lược đồ màu khối theo màu c và đánh số các dải của lược đồ màu khối tương ứng với các khối ảnh từ trong ra ngoài theo chiều ngược chiều kim đồng hồ. Tiếp theo, IHG chia bài toán tính khoảng cách của ảnh truy vấn và ảnh CSDL theo màu c thành các bài toán con phụ thuộc lược đồ màu khối tương ứng với các khối thuộc hình chữ nhật có kích thước $2*r$ hoặc $2*r+1$ và xây dựng đồ thị hai phía có trọng số cho mỗi bài toán con, ở đây mỗi đỉnh của đồ thị là một dải của lược đồ màu khối tương ứng với khối ảnh trên cạnh hình vuông kích thước $2*r$ hoặc $2*r+1$. Cuối cùng, thuật toán xác định giá trị đối sánh cực tiểu của mỗi bài toán con thông qua áp dụng thuật toán Hungarian [27]. Khoảng cách giữa hai ảnh theo màu c là tổng các giá trị khoảng cách của tất cả các bài toán con theo màu c và khoảng cách giữa hai ảnh là tổng tất cả các khoảng cách giữa hai ảnh theo tất cả các màu. Ví dụ 2.2 ở dưới minh họa sự thực hiện của thuật toán IHG.

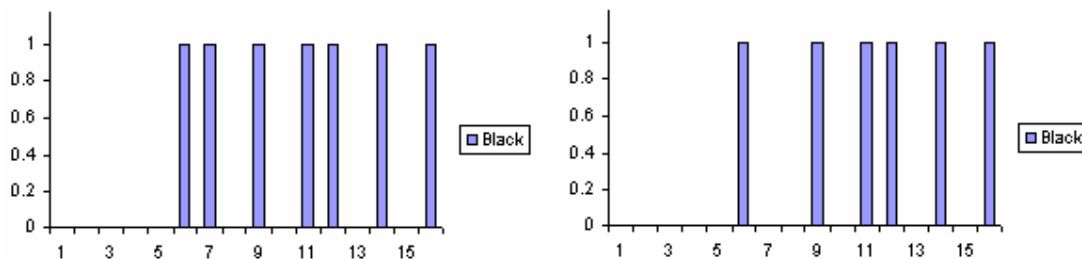
Ví dụ 2.2: So sánh khoảng cách giữa hai ảnh I_1 và I_2 .

Bước 1: Hai ảnh I_1 và I_2 được tách thành 16 khối ảnh (4×4) cho mỗi ảnh và đánh số các khối ảnh của mỗi ảnh theo như Hình 2.7.



Hình 2.7. Các khối ảnh của mỗi ảnh được đánh số từ trong ra và ngược chiều kim đồng hồ.

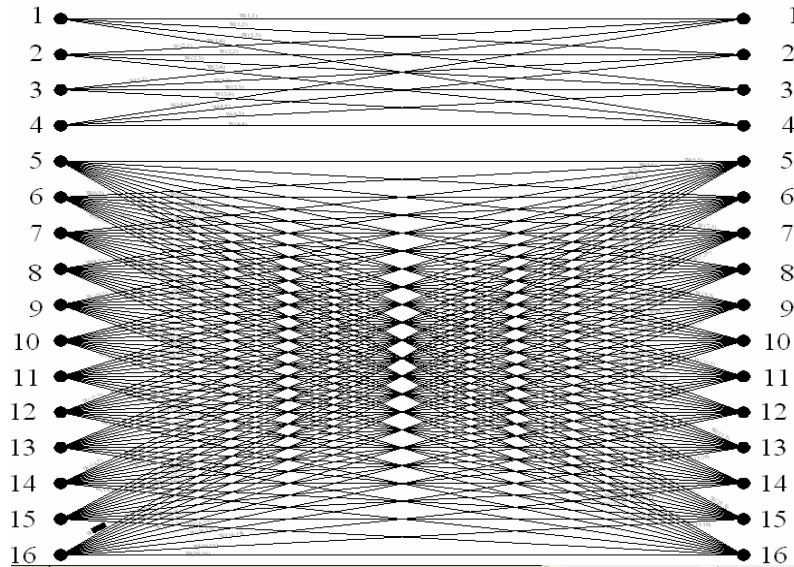
Xây dựng lược đồ màu khối theo màu c của mỗi ảnh. Trong ví dụ này màu black được chọn để minh họa như được chỉ ra trong Hình 2.8.



Hình 2.8. Lược đồ màu khối theo màu black của hai ảnh I_1 và I_2 .

Bước 2: Xây dựng đồ thị hai phía cho mỗi bài toán con như được chỉ ra trong Hình 2.9. Trong đồ thị này, chúng tôi nhận thấy: Thứ nhất, mỗi đỉnh trong các đỉnh có thứ tự từ 1 đến 4 của tập X chỉ được nối với bốn đỉnh có thứ tự từ 1 đến 4 của tập Y và ngược lại. Thứ hai, mỗi đỉnh trong các đỉnh có số thứ tự từ 5 đến 16 của tập X được nối với tất cả các đỉnh có số thứ tự từ 5 đến 16 của tập Y và ngược lại. Có điều thứ nhất là do dải của lược đồ màu khối tương ứng với mỗi khối ảnh của cạnh hình vuông kích thước 1 của ảnh I_1 chỉ được so sánh với các dải của lược đồ màu khối tương ứng với các khối ảnh của cạnh hình vuông kích thước 1 của ảnh I_2 và ngược lại. Tương tự đối với điều thứ hai, mỗi dải của lược đồ màu khối tương

ứng với mỗi khối ảnh của cạnh hình vuông kích thước 2 của ảnh I_1 chỉ được so sánh với các dải của lược đồ màu khối tương ứng với các khối ảnh của cạnh hình vuông kích thước 2 của ảnh I_2 và ngược lại.



Hình 2.9. Đồ thị hai phía biểu thị mối quan hệ của các dải của lược đồ màu khối của ảnh I_1 và I_2 theo màu black.

Bước 3: Từ đồ thị trong Hình 2.9 chúng tôi xây dựng ma trận kề cho bài toán con 1 là Mat_1 và bài toán con 2 là Mat_2 theo màu black. Ở đây $w_{i,j}$ là trọng số của cạnh nối đỉnh i của tập X với đỉnh j của tập Y . Giá trị trọng số này được nén vào khoảng $[0,1]$.

$$Mat_1 = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,4} \\ w_{2,1} & w_{2,2} & \dots & w_{2,4} \\ w_{3,1} & w_{3,2} & \dots & w_{3,4} \\ w_{4,1} & w_{4,2} & \dots & w_{4,4} \end{bmatrix} \quad Mat_2 = \begin{bmatrix} w_{5,5} & w_{5,6} & \dots & w_{5,16} \\ w_{6,5} & w_{6,6} & \dots & w_{6,16} \\ \dots & \dots & \dots & \dots \\ w_{16,5} & w_{16,6} & \dots & w_{16,16} \end{bmatrix}$$

Tìm giá trị đối sánh cực tiểu của đồ thị lược đồ màu khối theo màu black $Cost1_{\text{black}}$ của ma trận kề Mat_1 và giá trị đối sánh cực tiểu của đồ thị lược đồ màu khối theo màu black $Cost2_{\text{black}}$ của ma trận kề Mat_2 dựa trên thuật toán Hungarian

[27], giá trị $Cost_{black} = Cost1_{black} + Cost2_{black}$ biểu thị khoảng cách giữa hai ảnh I_1 và I_2 theo màu black.

Tương tự, chúng ta sẽ tính được giá trị đối sánh cực tiểu theo màu white $Cost_{white} = Cost1_{white} + Cost2_{white}$ và giá trị đối sánh cực tiểu của đồ thị lược đồ màu khối theo màu gray $Cost_{gray} = Cost1_{gray} + Cost2_{gray}$. Khoảng cách cuối cùng giữa hai ảnh được xác định bằng tổng giá trị khoảng cách của các màu black, white và gray, tức là $Cost = Cost_{black} + Cost_{white} + Cost_{gray}$.

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($G(X, Y, E, c), r$). Hàm EdgeDistance($G(X, Y, E, c), r$) 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
 $costr \leftarrow Costr + w(i,j)$
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,

Chứng minh: Để tính giá trị đối sánh cực tiểu của đồ thị $G(X, Y, E, c)$ gồm $(4r-5)^2$ đỉnh ở bước 1 cần $O(r^2)$ [27]. Hơn nữa, ở bước 2 cần $(4r-5)$ phép tính.

Do đó, độ phức tạp thời gian của hàm EdgeDistance là $O(r^2)$. Mệnh đề đã được chứng minh .

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,

Chứng minh:

Rõ ràng rằng, độ phức tạp của thân vòng lặp *while* được tính dựa trên thời gian tính giá trị các trọng số ở bước 2.3, cần n^2 phép tính và thời gian tính khoảng cách các dải của lược đồ màu khối tương ứng với các khối ảnh của cạnh hình vuông kích thước r của ảnh I_1 và các dải của lược đồ màu khối tương ứng với các khối ảnh của cạnh hình vuông kích thước r của ảnh I_2 theo màu c , cần $O(n^2)$. Vì vậy, độ phức tạp của thân vòng lặp *while* là $O(n^2)$.

Hơn nữa, vòng lặp *while* được thực hiện nhiều nhất là n lần. Vì thế, độ phức tạp của vòng lặp *while* là $O(n^3)$.

Do đó, độ phức tạp thời gian của hàm DistancebyColor (I_1, I_2, n, c) là $O(n^3)$. Mệnh đề đã được chứng minh .

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,

Chứng minh: Do độ phức tạp của thuật toán DistancebyColor(I_1, I_2, n, c) là $O(n^3)$ (Mệnh đề 2.6). Hơn nữa, thuật toán IHG(I_1, I_2, n) gọi hàm DistancebyColor(I_1, I_2, n, c) C lần (C là một hằng số), do đó độ phức tạp của thuật toán IHG(I_1, I_2, n) là $O(Cn^3) = O(n^3)$. Mệnh đề đã được chứng minh .

2.4 Các thực nghiệm

2.4.1 Môi trường thực nghiệm

CSDL gồm 7,812 ảnh jpeg. CSDL ảnh này là tập con của tập ảnh của GS Wang [76] và chúng tôi tập hợp từ Internet được sử dụng để đánh giá hiệu năng tra cứu. Các ảnh trong CSDL có kích cỡ là 128×85 điểm ảnh hoặc 85×128 điểm ảnh. Các ảnh gồm 256 màu (các ảnh được lượng hoá thành 15 màu). CSDL gồm các loại chính: Vườn hoa, cảnh cá dưới biển, thực vật, chim, ngựa, nhà cửa, thiết bị điện tử,

thời trang, trượt tuyết, lướt sóng, cảnh hoàng hôn, bãi biển, phong cảnh, người chơi gôn, bò tót, mây, trái cây, quốc kỳ, bệnh viện, ngôi sao, rừng, di tích cổ, thuyền buồm, ô tô.

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 kết quả 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 CSDL “Wang 1000” [76], 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



a/ Ngựa



b/ Voi



c/ Hoa



d/ Người Châu phi



e/ Xe buýt



f/ Cầu Thê Húc

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

Bốn truy vấn đầu thực hiện đối với ảnh truy vấn và tập ảnh liên quan không điều chỉnh quay và dịch chuyển. Hai truy vấn được thực hiện đối với ảnh truy vấn và tập ảnh liên quan quay và dịch chuyển. Mỗi truy vấn này được sử dụng cả ba phương pháp LCH, CCH và HG.

Trong thực nghiệm, chúng tôi sử dụng các tham số: số các màu trong tập C là 15, số các khối ảnh là 64.

Bảng 2.2. Các kết quả của truy vấn 1.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>CCH</i>	<i>LCH</i>
0.1	0.91	0.93	0.94
0.2	0.81	0.78	0.79
0.3	0.75	0.73	0.71
0.4	0.52	0.67	0.68
0.5	0.45	0.58	0.66
0.6	0.41	0.51	0.51
0.7	0.35	0.46	0.46
0.8	0.21	0.23	0.38
0.9	0.17	0.21	0.28
1	0.09	0.07	0.09

Bảng 2.3. Các kết quả của truy vấn 2.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>CCH</i>	<i>LCH</i>
0.1	0.89	0.94	0.92
0.2	0.83	0.81	0.76
0.3	0.67	0.72	0.71
0.4	0.5	0.58	0.61
0.5	0.39	0.56	0.59
0.6	0.35	0.51	0.48
0.7	0.34	0.47	0.46
0.8	0.21	0.27	0.35
0.9	0.18	0.19	0.18
1	0.09	0.13	0.13

Bảng 2.4. Các kết quả của truy vấn 3.

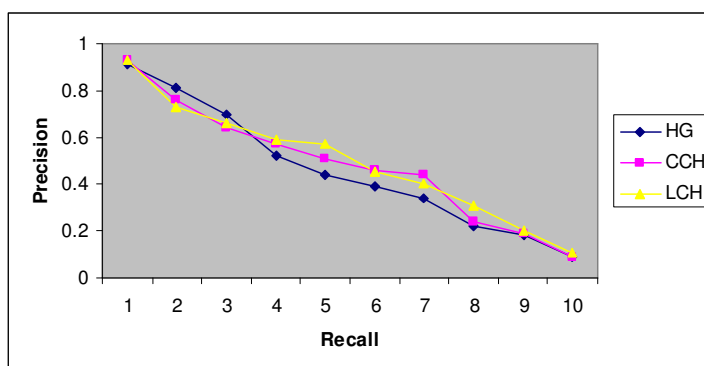
<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>CCH</i>	<i>LCH</i>
0.1	0.92	0.95	0.93
0.2	0.84	0.78	0.71
0.3	0.71	0.71	0.68
0.4	0.5	0.65	0.65
0.5	0.41	0.58	0.63
0.6	0.36	0.52	0.47
0.7	0.33	0.51	0.42
0.8	0.25	0.29	0.35
0.9	0.22	0.23	0.18
1	0.08	0.07	0.13

Bảng 2.5. Các kết quả của truy vấn 4.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>CCH</i>	<i>LCH</i>
0.1	0.9	0.91	0.91
0.2	0.74	0.65	0.64
0.3	0.65	0.39	0.54
0.4	0.56	0.37	0.43
0.5	0.51	0.33	0.41
0.6	0.45	0.29	0.32
0.7	0.32	0.31	0.27
0.8	0.2	0.17	0.17
0.9	0.16	0.13	0.15
1	0.08	0.07	0.09

Bảng 2.2, 2.3, 2.4 và 2.5 đưa ra tóm tắt các kết quả truy vấn 1, 2, 3 và 4 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Trong thực nghiệm đầu tiên, phương pháp HG được sử dụng cho quá trình tra cứu. Sau đó, CCH được sử dụng trong thực nghiệm thứ hai. Cuối cùng cách tiếp cận LCH được sử dụng trong thực nghiệm thứ 3.

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.

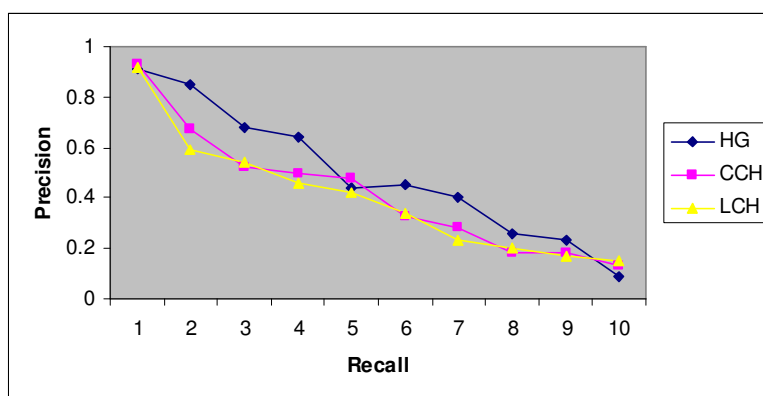
Bảng 2.6 và 2.7 đưa ra tóm tắt các kết quả truy vấn 5 và 6 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Trong thực nghiệm đầu tiên, phương pháp HG được sử dụng cho quá trình tra cứu. Sau đó, CCH được sử dụng trong thực nghiệm thứ hai. Cuối cùng cách tiếp cận LCH được sử dụng trong thực nghiệm thứ 3.

Bảng 2.6. Các kết quả của truy vấn 5.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>CCH</i>	<i>LCH</i>
0.1	0.92	0.89	0.91
0.2	0.87	0.71	0.53
0.3	0.71	0.53	0.51
0.4	0.66	0.51	0.45
0.5	0.43	0.48	0.41
0.6	0.47	0.35	0.33
0.7	0.41	0.31	0.23
0.8	0.24	0.21	0.19
0.9	0.21	0.16	0.16
1	0.08	0.15	0.15

Bảng 2.7. Các kết quả của truy vấn 6.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>CCH</i>	<i>LCH</i>
0.1	0.89	0.96	0.93
0.2	0.83	0.63	0.64
0.3	0.64	0.5	0.56
0.4	0.61	0.49	0.46
0.5	0.45	0.47	0.43
0.6	0.43	0.31	0.34
0.7	0.39	0.25	0.23
0.8	0.28	0.14	0.2
0.9	0.25	0.2	0.17
1	0.1	0.11	0.15



Hình 2.12. So sánh LCH, CCH với HG theo các truy vấn 5 và 6 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.

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 và tốc độ của kỹ thuật tra cứu IHG, sáu truy vấn được chọn cho mục tiêu đánh giá. Tập truy vấn của chúng tôi gồm 6 truy vấn. Các truy vấn 1, 3, 4 và 5 cùng với tập ảnh liên quan được tạo ra từ CSDL “Wang 1000” [76], 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



a/ Xe buýt



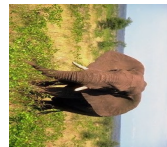
b/ Cầu Thê Húc



c/ Khủng long



d/ Ngựa



e/ Voi



f/ Hoa

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

Hai truy vấn đầu thực hiện đối với ảnh truy vấn và tập ảnh liên quan được điều chỉnh dịch chuyển. Bốn truy vấn sau được thực hiện đối với ảnh truy vấn và tập ảnh

liên quan được điều chỉnh quay. Các truy vấn 1 và 2 được sử dụng hai phương pháp HG và IHG, các truy vấn từ 3 đến 6 sử dụng thêm phương pháp SR [59, 75, 76].

Trong thực nghiệm, chúng tôi sử dụng các tham số: số các màu trong tập C là 15, số các khối là 64.

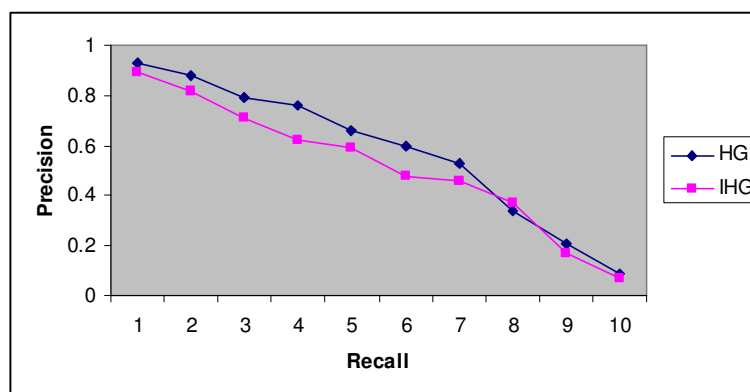
Bảng 2.9 và 2.10 đưa ra tóm tắt các kết quả truy vấn 1 và 2 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Trong thực nghiệm đầu tiên, phương pháp HG được sử dụng cho quá trình tra cứu. Sau đó, IHG được sử dụng trong thực nghiệm thứ hai.

Bảng 2.9. Các kết quả của truy vấn 1. Bảng 2.10. Các kết quả của truy vấn 2.

<i>Recall</i>	<i>Precision</i>	
	<i>HG</i>	<i>IHG</i>
0.1	0.93	0.88
0.2	0.89	0.83
0.3	0.79	0.69
0.4	0.77	0.61
0.5	0.7	0.57
0.6	0.64	0.48
0.7	0.56	0.44
0.8	0.34	0.37
0.9	0.21	0.15
1	0.09	0.05

<i>Recall</i>	<i>Precision</i>	
	<i>HG</i>	<i>IHG</i>
0.1	0.92	0.9
0.2	0.86	0.81
0.3	0.79	0.73
0.4	0.75	0.63
0.5	0.61	0.61
0.6	0.56	0.47
0.7	0.49	0.48
0.8	0.33	0.37
0.9	0.2	0.19
1	0.08	0.09

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.

Bảng 2.11, 2.12, 2.13 và 2.14 đưa ra tóm tắt các kết quả truy vấn 3, 4, 5 và 6 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Trong thực nghiệm đầu tiên, phương pháp HG được sử dụng cho quá trình tra cứu. Sau đó, IHG được sử dụng trong thực nghiệm thứ hai. Cuối cùng, SR [59, 75, 76] được sử dụng trong thực nghiệm thứ ba.

Bảng 2.11. Các kết quả của truy vấn 3.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>IHG</i>	<i>SR</i>
0.1	0.93	0.92	0.9
0.2	0.72	0.82	0.7
0.3	0.6	0.69	0.57
0.4	0.59	0.67	0.51
0.5	0.49	0.6	0.49
0.6	0.47	0.46	0.41
0.7	0.38	0.47	0.3
0.8	0.22	0.3	0.29
0.9	0.11	0.18	0.1
1	0.09	0.08	0.07

Bảng 2.12. Các kết quả của truy vấn 4.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>IHG</i>	<i>SR</i>
0.1	0.95	0.91	0.9
0.2	0.74	0.84	0.79
0.3	0.62	0.67	0.61
0.4	0.61	0.66	0.52
0.5	0.52	0.58	0.51
0.6	0.49	0.44	0.42
0.7	0.4	0.45	0.37
0.8	0.24	0.28	0.32
0.9	0.13	0.16	0.12
1	0.11	0.06	0.04

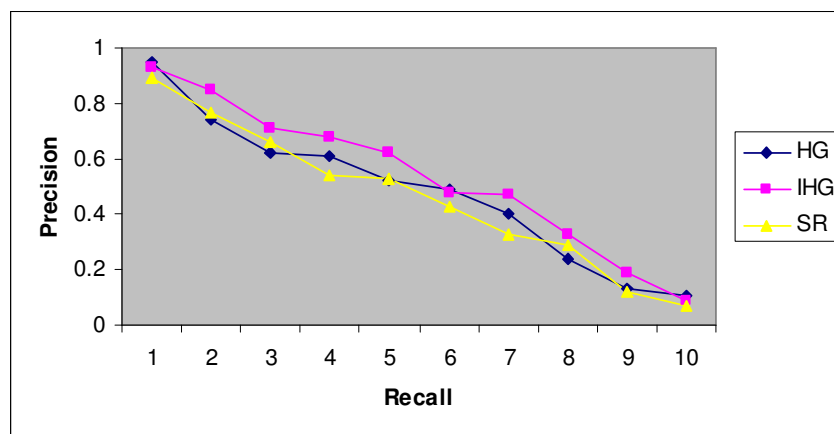
Bảng 2.13. Các kết quả của truy vấn 5.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>IHG</i>	<i>SR</i>
0.1	0.96	0.94	0.89
0.2	0.75	0.85	0.78
0.3	0.63	0.72	0.7
0.4	0.62	0.69	0.54
0.5	0.53	0.63	0.55
0.6	0.5	0.49	0.45
0.7	0.41	0.45	0.33
0.8	0.25	0.37	0.29
0.9	0.14	0.22	0.12
1	0.12	0.09	0.07

Bảng 2.14. Các kết quả của truy vấn 6.

<i>Recall</i>	<i>Precision</i>		
	<i>HG</i>	<i>IHG</i>	<i>SR</i>
0.1	0.95	0.95	0.88
0.2	0.74	0.89	0.8
0.3	0.62	0.77	0.77
0.4	0.61	0.71	0.58
0.5	0.52	0.65	0.57
0.6	0.49	0.54	0.43
0.7	0.4	0.52	0.32
0.8	0.24	0.35	0.27
0.9	0.13	0.21	0.13
1	0.11	0.11	0.09

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 phương pháp HG và SR.

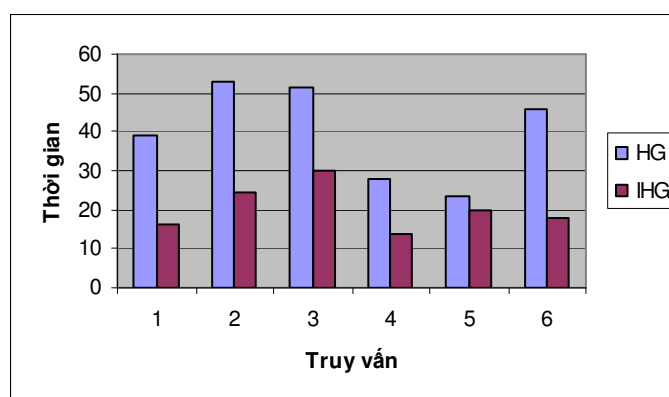


Hình 2.15. So sánh HG với IHG và SR theo các truy vấn 3, 4, 5 và 6 dưới dạng Recall-Precision.

Bảng 2.15 đưa ra thời gian thực hiện của phương pháp HG và IHG theo các truy vấn 1, 2, 3, 4, 5 và 6 tương ứng. Hình 2.16 biểu thị tốc độ thực hiện của phương pháp HG và IHG theo các truy vấn 1, 2, 3, 4, 5 và 6 tương ứng.

Bảng 2.15. Thời gian thực hiện của phương pháp HG và IHG.

<i>Truy vấn</i>	<i>Thời gian thực hiện (giây)</i>	
	<i>HG</i>	<i>IHG</i>
1	39.1	16.4
2	53.1	24.4
3	51.6	29.8
4	28	13.8
5	23.4	19.7
6	45.7	17.6



Hình 2.16. Biểu đồ so sánh tốc độ của phương pháp HG và IHG.

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. Tốc độ của phương pháp IHG là nhanh hơn HG.

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 lược đồ biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển, được gọi là HG. Kỹ thuật này gồm các bước: lượng hoá các màu của ảnh,

tính toán các lược đồ màu khối cho mỗi màu, xây dựng đồ thị hai phía có trọng số cho mỗi màu và tính tổng tất cả các khoảng cách của hai ảnh theo mỗi màu. 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

Trong chương này, chúng tôi sẽ trình bày phương pháp biểu diễn ảnh sử dụng cây tứ phân. Sau đó, chúng tôi sẽ trình bày phương pháp tra cứu ảnh dựa vào vùng ảnh, có tên là CSI, sử dụng đặc trưng của vùng ảnh vào quá trình tra cứu làm tăng hiệu năng tra cứu. Bên cạnh đó, chúng tôi cũng đề xuất kỹ thuật tra cứu ảnh sử dụng màu và các cụm màu thuần nhất, có tên là CCS.

3.1 Biểu diễn ảnh sử dụng phương pháp cây tứ phân

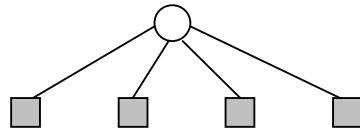
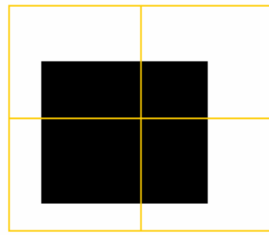
Với phương pháp cây tứ phân [58], một vùng trong ảnh được tách thành bốn vùng có kích cỡ bằng nhau. Với mỗi vùng, nếu vùng là thuần nhất, thì vùng này không được xem xét là ứng cử viên cho lần tách tiếp theo. Nếu vùng không thuần nhất, nó sẽ được tách thành bốn vùng con cho đến khi tất cả các vùng đều 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.

Ví dụ 3.1: Xét ảnh được cho như trong Hình 3.1.

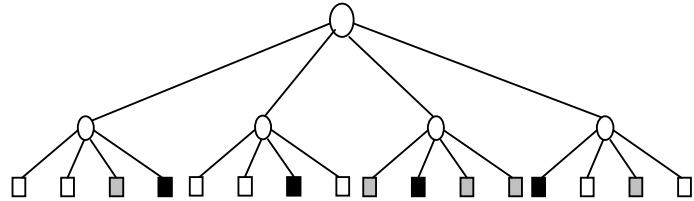
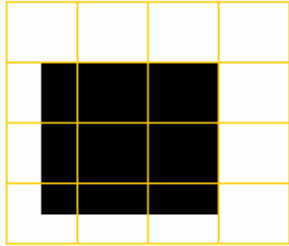


Hình 3.1 Ảnh gốc.

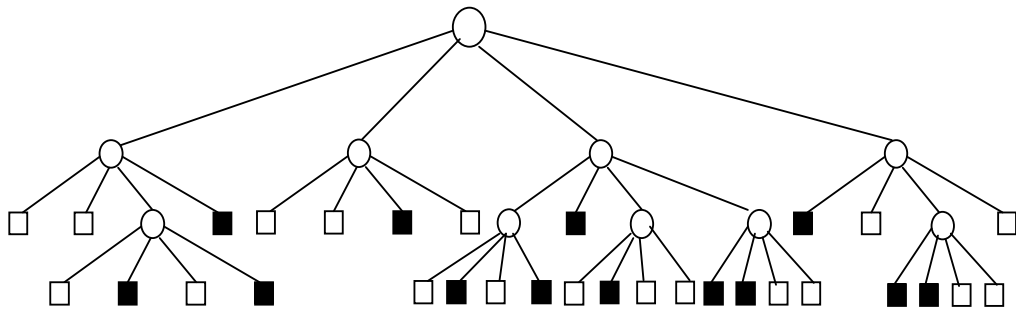
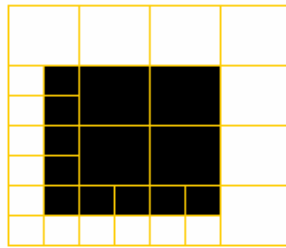
Nếu sử dụng phương pháp cây tứ phân, chúng ta sẽ được cây tứ phân như được chỉ ra trong Hình 3.2.



a/ Tách lần thứ nhất



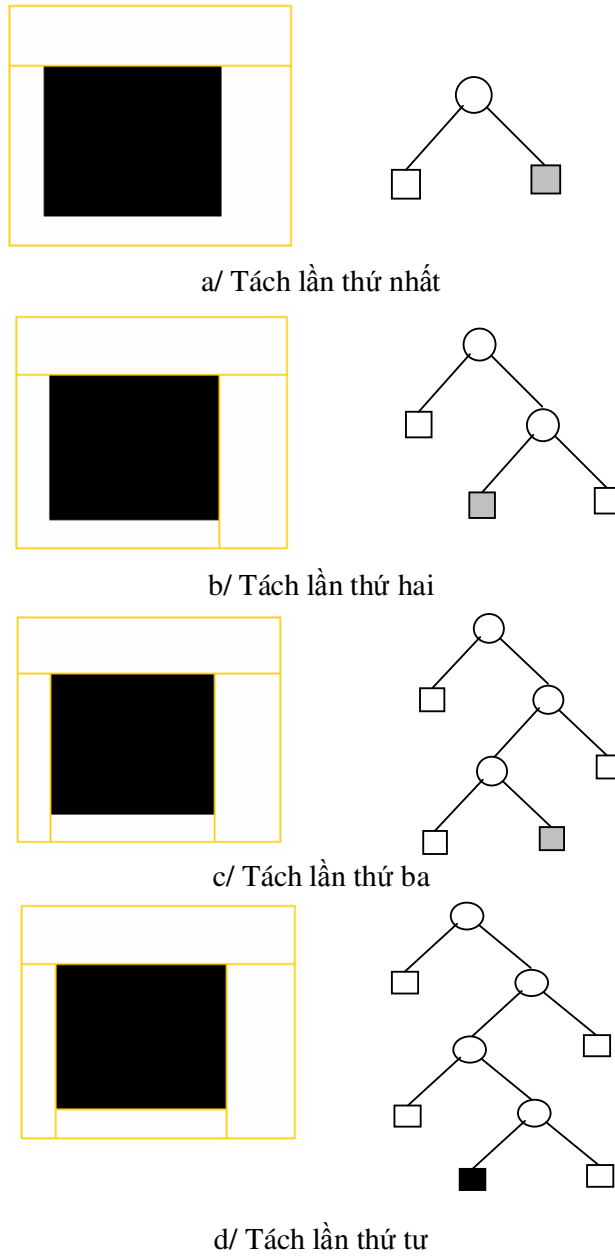
b/ Tách lần thứ hai



c/ Tách lần thứ ba

Hình 3.2. Cây tứ phân biểu diễn ảnh cho trong Hình 3.1.

Chúng ta nhận thấy, 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. Chẳng hạn, như ảnh được chỉ ra trong Hình 3.1, phương pháp cây tứ phân phải tách thành 31 khối. Trong khi ảnh này có thể được chia thành các vùng tốt hơn như được chỉ ra trong Hình 3.3 (chỉ cần 5 khối).



Hình 3.3 Cây biểu diễn ảnh cho trong Hình 3.1.

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 [51]. Một biến thể

của cách tiếp cận này là cách tiếp cận dựa vào cây tứ phân [58], 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 trên, chúng tôi trình bày kỹ thuật CSI và CCS. Hai kỹ thuật này hướng tới việc chia ảnh thành các vùng có kích cỡ khác nhau (tùy thuộc vào cấu tạo của từng ảnh).

3.2.2 Trích rút đặc trưng

Dựa vào một màu chúng tôi thực hiện trích rút các thông tin không gian. Các thông tin này được sử dụng để đánh giá khoảng cách của các ảnh.

3.2.2.1 Trích rút màu và thông tin không gian

Trong phần này, chúng tôi trình bày thuật toán CSI trích rút đặc trưng màu và thông tin không gian trong ảnh. Thuật toán này của chúng tôi đã công bố trong [45].

Chúng tôi áp dụng thuật toán cân bằng lược đồ để phân cụm các màu được chọn. Với mỗi màu được chọn, thuật toán được áp dụng với một không gian ảnh theo cả trục x và y . Kết quả của thuật toán là một tập các vùng với mỗi màu. Vùng BR_i được biểu diễn bởi hình chữ nhật $(x_{ul}^i, y_{ul}^i, x_{br}^i, y_{br}^i)$.

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 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.

Kỳ vọng có thể được tính toán theo kinh nghiệm về phân bố mẫu. Nếu các mẫu quan sát quá xa 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 kỳ vọng được mô tả như dưới.

Độ lệch giữa một quan sát i và 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 thoả 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:

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

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

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

Ký hiệu DX_{toprow} và $DX_{bottomrow}$ là 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 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 $\text{Cost}(BR_i)$ bằng $DX_{selectedrow}$, thì BR_i được phân hoạch theo chiều dọc. Trong trường hợp ngược lại, nếu $\text{Cost}(BR_i)$ bằng $DX_{selectedcol}$, BR_i được phân hoạch theo chiều ngang.

Thuật toán trích rút màu và thông tin không gian CSI được mô tả như dưới đây.

Thuật toán CSI:

Vào: - Img - ảnh
 - $minarea$ – ngưỡng diện tích của vùng
 - T – ngưỡng nhiễu
 - C - Tập màu

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

```

For (i=1; i<=|C|; i++) {
1. Stack ←  $Img$ 
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  $\text{Cost}(BR)=\text{Max}(DX_{selectedrow}, DX_{selectedcol})$ .
    2.2.2 for mỗi  $BR_j$ 
      ▪ tính  $DX \leftarrow \frac{obs(j) - E(j)}{\sqrt{E(j)}}$ 
      ▪ If ( $DX < T$ ) Stack ←  $BR_j$ 
      ▪ else xuất ( $C_i, BR_j$ )
3. while (Stack #  $\emptyset$ )
  
```

Trong thuật toán CSI, có ba tham số $minarea$, $Cost(BR_i)$ và T , trong đó:

- ❖ Tham số $minarea$ 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 $Cost(BR_i)$ bằng $DX_{selectedrow}$, vùng này được phân hoạch theo chiều dọc. Trong trường hợp ngược lại, nếu $Cost(BR_i)$ bằng $DX_{selectedcol}$ vùng được phân hoạch theo chiều ngang.
- ❖ Tham số T là ngưỡng nhiễu được chấp nhận của mỗi vùng.

Kết quả của thuật toán này là thông tin không gian của mỗi màu của ảnh được biểu diễn bởi một danh sách $\langle (c_1; br_1), (c_2; br_2), \dots, (c_n; br_n) \rangle$, trong đó:

- ❖ c_i là màu được chọn,
- ❖ br_i là danh sách các vùng có màu c_i .

Mỗi br_i là một danh sách $\langle (x_{tb}^1, y_{tb}^1, x_{br}^1, y_{br}^1); \dots; (x_{tb}^n, y_{tb}^n, x_{br}^n, y_{br}^n) \rangle$. Ký hiệu $(x_{tb}^i, y_{tb}^i, x_{br}^i, y_{br}^i)$ là một hình chữ nhật với (x_{tb}^i, y_{tb}^i) và (x_{br}^i, y_{br}^i) là các tọa độ góc trên bên trái và góc dưới bên phải của hình chữ nhật tương ứng.

Ví dụ 3.2: Hình 3.4 chỉ ra một ảnh I có cỡ 10×10 điểm ảnh.

1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	0	0	1	1	1
1	1	1	0	0	1	1	1	1	1
1	1	1	0	0	1	0	0	1	1
1	1	1	0	0	1	0	0	1	1
1	1	1	0	0	1	0	0	1	1
1	1	1	1	0	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1

Hình 3.4. Ảnh I cỡ 10×10 điểm ảnh.

Trong ví dụ này, ảnh I được tách ra thành hai vùng con theo chiều ngang và hướng từ trái qua phải như được chỉ ra ở Hình 3.5.

1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1

1	1	1	1	1	1	1
1	1	1	1	1	1	1
1	1	0	0	1	1	1
0	0	1	1	1	1	1
0	0	1	0	0	1	1
0	0	1	0	0	1	1
0	0	1	0	0	1	1
1	0	1	1	1	1	1
1	0	1	1	1	1	1
1	0	1	1	1	1	1

a/ Vùng BR₁
b/Vùng BR₂

Hình 3.5. Ảnh I sau khi được tách ra thành hai vùng BR₁ và BR₂.

Bảng 3.1. Tính độ lệch $DX_{selectedrow}$ cho phân hoạch theo dòng của ảnh I .

Dòng	Tần suất	Tổng tần suất (row)	Tổng tần suất từ (row+1)	Trung bình khối trên (row)	Phần còn lại	DX_{toprow}	$DX_{bottomrow}$	$\text{Max}(DX_{toprow}, DX_{bottomrow})$	$DX_{selectedrow}$
1	10	10	71	9	81	0.33	-1.11	0.33	0.47
2	10	20	61	18	72	0.47	-1.3	0.47	
3	8	28	53	27	63	0.19	-1.26	0.19	
4	8	36	45	36	54	0	-1.22	0	
5	6	42	39	45	45	-0.45	-0.89	-0.45	
6	6	48	33	54	36	-0.82	-0.5	-0.5	
7	6	54	27	63	27	-1.13	0	0	
8	9	63	18	72	18	-1.06	0	0	
9	9	72	9	81	9	-1	0	0	
10	9	81	0	90	0	-0.95			

Ảnh I được tách ra thành hai vùng BR_1 và BR_2 như được chỉ ra trong Hình 3.5 ở trên là do giá trị độ lệch theo cột $DX_{selectedcol}$ lớn hơn giá trị độ lệch theo dòng $DX_{selectedrow}$. Hơn nữa, $DX_{selectedcol}$ đạt giá trị lớn nhất tại cột thứ 3 tính từ bên trái. Do đó, vùng được tạo ra từ cột thứ 1 đến cột thứ 3 của ảnh I tính từ bên trái có độ thuần nhất lớn nhất (xem Bảng 3.1 và Bảng 3.2).

Bảng 3.2. Tính độ lệch $DX_{selectedcol}$ cho phân hoạch theo cột của ảnh I .

Cột	Tần suất	Tổng tần suất (column)	Tổng tần suất từ (column+1)	Trung bình khối trái(column)	Phần còn lại	$DX_{leftcol}$	$DX_{rightcol}$	$\text{Max}(DX_{leftcol}, DX_{rightcol})$	$DX_{selectedcol}$
1	10	10	71	9	81	0.33	-1.11	0.33	0.58
2	10	20	61	18	72	0.47	-1.3	0.47	
3	10	30	51	27	63	0.58	-1.51	0.58	
4	6	36	45	36	54	0	-1.22	0	
5	3	39	42	45	45	-0.89	-0.45	-0.45	
6	9	48	33	54	36	-0.82	-0.5	-0.5	
7	6	54	27	63	27	-1.13	0	0	
8	7	61	20	72	18	-1.3	0.47	0.47	
9	10	71	10	81	9	-1.11	0.33	0.33	
10	10	81	0	90	0	-0.95			

Tiếp theo, chúng ta xét vùng BR_1 và BR_2 . Đối với vùng BR_1 do vùng này là thuần nhất nên không cần đưa vào lần phân hoạch tiếp theo. Tuy nhiên, vùng BR_2 không thuần nhất nên vùng này cần được đưa vào lần phân hoạch tiếp theo.

Ở đây vùng BR_2 được tách ra thành hai vùng con theo chiều ngang và hướng từ phải qua trái như được chỉ ra ở Hình 3.6.

1	1	1	1	1
1	1	1	1	1
1	1	0	0	1
0	0	1	1	1
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
1	0	1	1	1
1	0	1	1	1
1	0	1	1	1

1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	1

a/ Vùng $BR_{2,1}$ b/ Vùng $BR_{2,2}$

Hình 3.6. Vùng BR_2 sau khi được tách ra thành hai vùng $BR_{2,1}$ và $BR_{2,2}$.

Bảng 3.3. Tính độ lệch $DX_{selectedrow}$ cho phân hoạch theo dòng của vùng BR_2 .

Dòng	Tần suất	Tổng tần suất (row)	Tổng tần suất từ (row+1)	Trung bình khối trên (row)	Phần còn lại	DX_{toprow}	$DX_{bottomrow}$	$\text{Max}(DX_{toprow}, DX_{bottomrow})$	$DX_{selectedrow}$
1	7	7	44	6.3	57	0.28	-1.68	0.27	0.39
2	7	14	37	12.6	50	0.39	-1.88	0.39	
3	5	19	32	18.9	44	0.02	-1.82	0.02	
4	5	24	27	25.2	38	-0.24	-1.75	-0.23	
5	3	27	24	31.5	32	-0.8	-1.33	-0.80	
6	3	30	21	37.8	25	-1.27	-0.83	-0.83	
7	3	33	18	44.1	19	-1.67	-0.20	-0.20	
8	6	39	12	50.4	13	-1.61	-0.16	-0.16	
9	6	45	6	56.7	6	-1.55	-0.11	-0.12	
10	6	51	0	63	0	-1.51			

Vùng BR_2 được tách ra thành hai vùng $BR_{2,1}$ và $BR_{2,2}$ như được chỉ ra trong Hình 3.6 ở trên là do giá trị độ lệch theo cột $DX_{selectedcol}$ lớn hơn giá trị độ lệch theo dòng $DX_{selectedrow}$. Hơn nữa, $DX_{selectedcol}$ đạt giá trị lớn nhất tại cột thứ 6 tính từ bên trái. Do đó, vùng được tạo ra từ cột thứ 6 đến cột thứ 7 của vùng BR_2 tính từ bên trái có độ thuần nhất lớn nhất (xem Bảng 3.3 và Bảng 3.4).

Bảng 3.4. Tính độ lệch $DX_{selectedcol}$ cho phân hoạch theo cột của vùng BR_2 .

Cột	Tần suất	Tổng tần suất (column)	Tổng tần suất từ (column+1)	Trung bình khối trái(column)	Phần còn lại	$DX_{leftcol}$	$DX_{rightcol}$	$\text{Max}(DX_{leftcol}, DX_{rightcol})$	$DX_{selectedcol}$
1	6	6	45	9	54	-1	-1.22	-1	0.47
2	3	9	42	18	45	-2.12	-0.44	-0.44	
3	9	18	33	27	36	-1.73	-0.5	-0.5	
4	6	24	27	36	27	-2	0	0	
5	7	31	20	45	18	-2.09	0.47	0.47	
6	10	41	10	54	9	-1.77	0.33	0.33	
7	10	51	0	63	0	-1.51	-1.22		

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,

Chứng minh:

Hiển nhiên rằng thời gian thực hiện của thuật toán CSI là thời gian thực hiện lệnh *do...while* trong bước 2. Do đó, chúng ta cần xác định số lần lặp của lệnh *do...while* và thời gian thực hiện thân của vòng lặp.

Chúng ta cũng thấy rằng, số lần lặp tối đa là $\frac{n}{min\ area}$ (với $min\ area$ là số điểm ảnh của cửa sổ và là hằng số).

Rõ ràng rằng, để xác định thời gian thực hiện thân vòng lặp, chúng ta phải xác định thời gian thực hiện của bước 2.2. Trong bước này chúng ta nhận thấy, thời gian thực hiện lệnh điều kiện $area(BR) > min\ area$ là $O(n)$. Hơn nữa, thời gian tính $Cost(BR) = Max(D_{selected\ row}, D_{selected\ col})$ trong bước 2.2.1 là $O(n)$. Ngoài ra, thời gian thực hiện hai lần lệnh tính độ lệch $DX = \frac{obs(j) - E(j)}{\sqrt{E(j)}}$ trong bước 2.2.2 là $O(n)$. Vì thế, thời gian thực hiện bước 2.2 là $O(n)$.

Từ đây chúng ta suy ra thời gian thực hiện vòng lặp **do...while** là $O(kn^2)$ (với $k = \frac{1}{min\ area}$ là hằng số).

Do đó, độ phức tạp thời gian của thuật toán CSI là $O(n^2)$. Mệnh đề đã được chứng minh .

3.2.2.2 Trích rút các cụm màu thuần nhất.

Trong phần này, chúng tôi trình bày thuật toán trích rút màu và các cụm màu thuần nhất của các màu được lựa chọn, gọi là CCS. Thuật toán này của chúng tôi đã được công bố trong [46].

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.

Thuật toán CCS có thể được viết như sau:

Thuật toán CCS:

Vào: I – ảnh, C - tập màu

$minsize$ – ngưỡng diện tích của một vùng, $tolerance$ - ngưỡng nhiễu

Ra: Các cụm màu thuần nhất trong ảnh.

For mỗi $c \in C$ do {

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) Stack $\leftarrow Rec_1$

2.2.1.3 If (size(Rec_2) > 0) Stack $\leftarrow Rec_2$

2.2.2 else xuất (c, REC)

3. while (Stack $\neq \emptyset$) }

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.

Kết quả của thuật toán này là màu và các cụm màu của nó trong một ảnh được biểu diễn bởi danh sách $\langle (c_1; rec_1), (c_2; rec_2), \dots, (c_n; rec_n) \rangle$. Ở đây c_i là màu được lựa chọn, và rec_i là vùng có màu c_i . Ký hiệu $(x_{tl}^i, y_{tl}^i, x_{br}^i, y_{br}^i)$ là một hình chữ nhật với (x_{tl}^i, y_{tl}^i) và (x_{br}^i, y_{br}^i) là các tọa độ góc trên bên trái và góc dưới bên phải của hình chữ nhật tương ứng.

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

Thủ tục Split() phân hoạch cụm REC thành hai cụm Rec_1 và Rec_2 .

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

$$\{ row+ \leftarrow p_{i,j}; afterrow+ \leftarrow p_{i+1,j} \}$$
 - 1.2 $v_i \leftarrow \lfloor \frac{\min(i, n-i)}{n} * (row - afterrow) \rfloor$
 - 1.3 $k \leftarrow \text{Arg}(\max(\langle v_1, v_2, \dots, v_n \rangle)); kv \leftarrow \max(\langle v_1, v_2, \dots, v_n \rangle)$
2. for $j \leftarrow 0$ to $n-1$ do
 - 2.1 for $i \leftarrow 0$ to $n-1$ do

$$\{ col+ \leftarrow p_{i,j}; aftercol+ \leftarrow p_{i,j+1} \}$$
 - 2.2 $h_j \leftarrow \lfloor \frac{\min(j, n-j)}{n} * (col - aftercol) \rfloor$
 - 2.3 $l \leftarrow \text{Arg}(\max(\langle h_1, h_2, \dots, h_n \rangle)); lv \leftarrow \max(\langle h_1, h_2, \dots, h_n \rangle)$
3. if $(kv > lv)$ then
 - 3.1 Tách REC theo chiều đứng tại dòng k
 - 3.2 $Rec_1 \leftarrow \text{size}((0,0);(k,n-1)); c_1 \leftarrow \text{color}(Rec_1)$
 - 3.3 $Rec_2 \leftarrow \text{size}((k,0);(n-1,n-1)); c_2 \leftarrow \text{color}(Rec_2)$
4. else if $(kv < lv)$ then
 - 4.1 Tách REC theo chiều ngang tại cột l
 - 4.2 $Rec_1 \leftarrow \text{size}((0,0);(n-1,l)); c_1 \leftarrow \text{color}(Rec_1)$
 - 4.3 $Rec_2 \leftarrow \text{size}((0,l);(n-1,n-1)); c_2 \leftarrow \text{color}(Rec_2)$
5. else if $(v_1 = v_2 = \dots = v_n = h_1 = h_2 = \dots = h_n)$
 - 5.1 $Rec_1 \leftarrow 0; c_1 \leftarrow \text{Null}$
 - 5.2 $Rec_2 \leftarrow 0; c_2 \leftarrow \text{Null}$
6. Trả lại $\langle (c_1, Rec_1), (c_2, Rec_2) \rangle$

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. kv nhận giá trị lớn nhất của danh sách $\langle v_1, v_2, \dots, v_n \rangle$ và lv nhận 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. Thủ tục Split() sẽ xuất ra các vùng Rec_1, Rec_2 và các màu c_1, c_2 tương ứng của nó.

Ví dụ 3.3: Hình 3.7 chỉ ra một ảnh gồm 6×10 điểm ảnh.

1	1	1	0	0	1	1	1	1	1	1
1	1	1	1	0	0	1	1	1	1	1
1	1	1	0	1	0	0	1	1	1	1
1	1	1	0	0	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10	

Hình 3.7. Ảnh gồm 6×10 điểm ảnh.

Bảng 3.5. Tính toán giá trị của v_i .

Thứ tự	$\sum_j p_{i,j}$	$\sum_j p_{i+1,j}$	v_i
1	8	8	0
2	8	7	0.3
3	7	7	0
4	7	10	1
5	10	10	0

Bảng 3.6. Tính toán giá trị của h_j .

6	6	6	3	3	3	5	6	6	$\sum_i p_{i,j}$
6	6	3	3	3	5	6	6	6	$\sum_i p_{i,j+1}$
0	0	0.9	0	0	0.8	0.3	0	0	h_j
1	2	3	4	5	6	7	8	9	Thứ tự

Trong ví dụ này, các vùng được tách ra thành hai vùng con theo chiều đứng tại dòng thứ tư ($k=4$) (xem trong Bảng 3.5 và Bảng 3.6).

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,

Chứng minh:

Rõ ràng rằng, độ phức tạp thời gian của thủ tục Split() là độ phức tạp thời gian lớn nhất của một trong các bước từ 1 đến 6. Chúng ta cũng dễ dàng nhận thấy, bước 1 có độ phức tạp thời gian là $O(n)$, bước 2 có độ phức tạp thời gian là $O(n)$ và các bước 3, 4, 5 và 6 đều có độ phức tạp thời gian là $O(1)$.

Do đó, độ phức tạp thời gian của thủ tục Split() là $O(n)$. Mệnh đề đã được chứng minh .

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,

Chứng minh:

Rõ ràng rằng, độ phức tạp của thuật toán CCS là thời gian thực hiện lệnh *do...while*. Đối với lệnh này, chúng ta dễ dàng nhận thấy số lần lặp tối đa là $\frac{n}{min\ size}$ (với *minsize* là số điểm ảnh của cửa sổ và là hằng số) và thời gian thực hiện của thân vòng lặp chính là độ phức tạp của thủ tục Split(). Hơn nữa, thủ tục Split() có độ phức tạp thời gian là $O(n)$ (xem Mệnh đề 3.3). Vì vậy, thời gian để thực hiện lệnh *do...while* là $O(\frac{n^2}{min\ size})$.

Do đó, độ phức tạp thời gian của thuật toán CCS là $O(n^2)$. Mệnh đề đã được chứng minh .

3.2.3 Độ tương tự giữa hai ảnh

Trong phần này, chúng tôi sử dụng thông tin màu và không gian để tính khoảng cách giữa hai ảnh Img_1 và Img_2 . 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 hàm DRC (Distance by Region Comparing) để tính khoảng cách giữa ảnh Img_1 và Img_2 .

Hàm DRC tính khoảng cách giữa hai ảnh Img_1 và Img_2 được mô tả như sau:

Hàm DRC:

Vào: T_c - tổng số các màu của tập màu
 R_{Img_1} - các vùng của ảnh Img_1
 R_{Img_2} - các vùng của ảnh Img_2

Ra : sim - độ tương tự giữa ảnh Img_1 và Img_2

1. sim ← 0;
2. for k ← 1 to T_c do
 - 2.1 for i ← 1 to T_k^{s1} do
 - 2.1.1 for j ← 1 to T_k^{s2} do

if $(R_{Img_1}(i, k) \cap R_{Img_2}(j, k)) \neq \emptyset$ then

sim ← | $R_{Img_1}(i, k) \cap R_{Img_2}(j, k)$ |
3. Trả lại giá trị sim

Trong hàm DRC, $R_{Img}(i,k)$ là vùng thứ i của màu thứ k trong ảnh Img . T_c là tổng số các màu của tập màu, $T_k^{s_1}$ là số các vùng có màu k của ảnh Img_1 , $T_k^{s_2}$ là số các vùng có màu k của ảnh Img_2 .

Với mỗi màu k trong tập màu T_c , hàm kiểm tra mỗi vùng thứ i ($i=1, \dots, T_k^{s_1}$) của ảnh Img_1 có chồng lên vùng j ($j=1, \dots, T_k^{s_2}$) của ảnh Img_2 không, nếu chồng thì số điểm ảnh của phần giao giữa vùng i và vùng j được cộng vào khoảng cách giữa hai ảnh Img_1 và Img_2 .

3.2.4 Các thực nghiệm

3.2.4.1 Môi trường thực nghiệm

Hiệu năng tra cứu được đánh giá sử dụng một CSDL gồm 7,812 ảnh jpeg. CSDL này là tập con của tập ảnh của GS WANG [76] và chúng tôi tập hợp qua Internet. Các CSDL ảnh này sẽ được sử dụng để phản ánh hiệu quả của phương pháp tra cứu đối với phương pháp CSI và CCS. Các ảnh trong CSDL có kích cỡ là 128 x 85 điểm ảnh hoặc 85 x 128 điểm ảnh. Các ảnh gồm 256 màu (các ảnh được lượng hoá thành 12 màu). CSDL gồm các loại chính: Vườn hoa, cá biển, thực vật, chim, ngựa, nhà, thiết bị điện tử, thời trang, trượt tuyết, lướt sóng, cảnh hoàng hôn, bãi biển, phong cảnh, chơi gôn, bò tót, mây, trái cây, quốc kỳ, bệnh viện, ngôi sao, rừng, di tích cổ, thuyền buồm, ô tô.

3.2.4.2 Kết quả thực nghiệm

Chúng tôi sử dụng đồ thị Recall-Precision như được trình bày ở phần đánh giá hiệu năng (mục 1.5 ở chương 1) để đánh giá hiệu quả tra cứu.

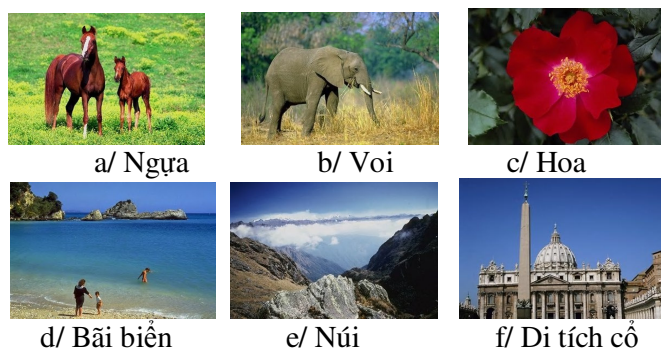
Phương pháp CSI:

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 (Quad Tree) [58] và CBC (Color Based Cluster) [55], các truy vấn từ 3 đến 6 sử dụng thêm phương pháp SR [59, 75, 76]. Các truy vấn từ 1 đến 6 cùng với tập ảnh liên quan

được tạo ra từ CSDL “WANG 1000” [76]. Thục nghiệm của chúng tôi đã sử dụng các tham số $minarea$ và T , ở đây $minarea=36$ và $T= 0.42$. 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.

Hai truy vấn đầu được thực hiện đối với ảnh truy vấn và tập ảnh liên quan tương đối hỗn tạp. Bốn truy vấn sau thực hiện đối với ảnh truy vấn và tập ảnh liên quan tương đối thuần nhất.

Các Bảng 3.8 và 3.9 đưa ra tóm tắt các kết quả của các truy vấn 1 và 2 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Trong thực nghiệm đầu tiên, chúng tôi sử dụng cách tiếp cận CSI cho quá trình tra cứu. Phương pháp QT được sử dụng cho thực nghiệm thứ hai. Cuối cùng, cách tiếp cận CBC [55] được sử dụng trong thực nghiệm thứ ba.

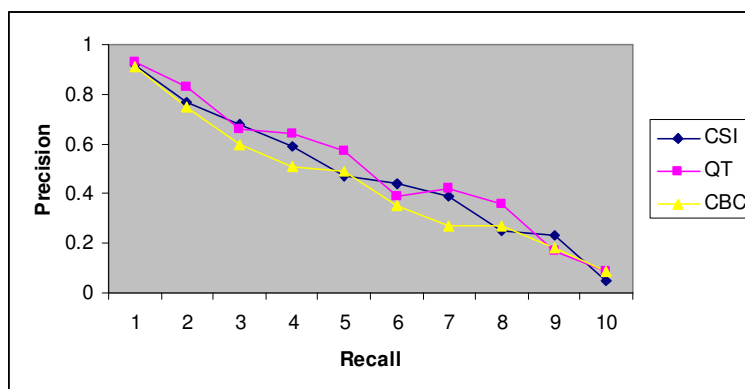
Bảng 3.8. Các kết quả của truy vấn 1.

<i>Recall</i>	<i>Precision</i>		
	<i>CSI</i>	<i>QT</i>	<i>CBC</i>
0.1	0.91	0.92	0.9
0.2	0.75	0.81	0.71
0.3	0.69	0.66	0.52
0.4	0.59	0.64	0.48
0.5	0.4	0.63	0.47
0.6	0.42	0.39	0.31
0.7	0.38	0.41	0.24
0.8	0.24	0.4	0.27
0.9	0.2	0.16	0.17
1	0.04	0.11	0.13

Bảng 3.9. Các kết quả của truy vấn 2.

<i>Recall</i>	<i>Precision</i>		
	<i>CSI</i>	<i>QT</i>	<i>CBC</i>
0.1	0.93	0.93	0.92
0.2	0.79	0.84	0.79
0.3	0.66	0.65	0.68
0.4	0.59	0.63	0.54
0.5	0.54	0.51	0.51
0.6	0.45	0.38	0.39
0.7	0.39	0.43	0.29
0.8	0.26	0.31	0.27
0.9	0.25	0.17	0.18
1	0.05	0.07	0.04

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.

Các Bảng 3.10, 3.11, 3.12 và 3.13 đưa ra tóm tắt các kết quả của các truy vấn 3, 4, 5 và 6 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Trong thực nghiệm đầu tiên, cách tiếp cận CSI được sử dụng cho quá trình tra cứu. Phương pháp QT được sử dụng cho thực nghiệm thứ hai. Thực

nghiệm thứ ba sử dụng phương pháp CBC. Cuối cùng, học nghiệm thứ tư sử dụng cách tiếp cận SR.

Bảng 3.10. Các kết quả của truy vấn 3.

<i>Recall</i>	<i>Precision</i>			
	<i>CSI</i>	<i>QT</i>	<i>CBC</i>	<i>SR</i>
0.1	0.92	0.91	0.91	0.89
0.2	0.91	0.67	0.83	0.69
0.3	0.74	0.58	0.72	0.59
0.4	0.7	0.49	0.68	0.52
0.5	0.45	0.47	0.49	0.51
0.6	0.53	0.31	0.48	0.38
0.7	0.44	0.29	0.34	0.34
0.8	0.42	0.15	0.36	0.27
0.9	0.39	0.14	0.26	0.19
1	0.08	0.03	0.07	0.08

Bảng 3.11. Các kết quả của truy vấn 4.

<i>Recall</i>	<i>Precision</i>			
	<i>CSI</i>	<i>QT</i>	<i>CBC</i>	<i>SR</i>
0.1	0.87	0.83	0.86	0.85
0.2	0.84	0.67	0.71	0.72
0.3	0.69	0.57	0.52	0.52
0.4	0.65	0.54	0.5	0.51
0.5	0.42	0.52	0.42	0.48
0.6	0.48	0.36	0.34	0.41
0.7	0.39	0.35	0.29	0.36
0.8	0.37	0.25	0.27	0.35
0.9	0.34	0.24	0.21	0.27
1	0.03	0.13	0.15	0.16

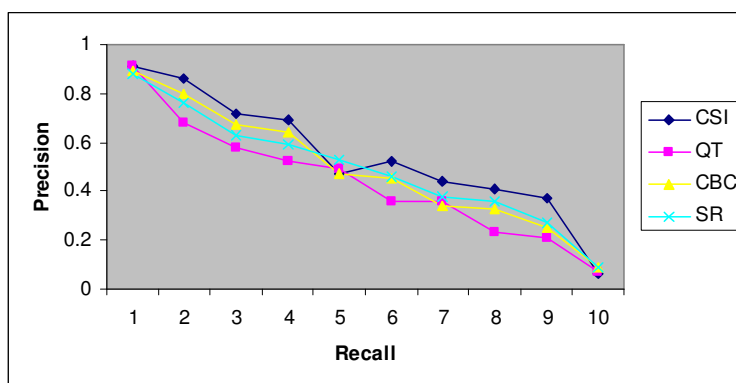
Bảng 3.12. Các kết quả của truy vấn 5.

<i>Recall</i>	<i>Precision</i>			
	<i>CSI</i>	<i>QT</i>	<i>CBC</i>	<i>SR</i>
0.1	0.88	0.97	0.87	0.89
0.2	0.81	0.69	0.81	0.78
0.3	0.7	0.58	0.7	0.71
0.4	0.66	0.52	0.65	0.62
0.5	0.47	0.49	0.49	0.56
0.6	0.49	0.43	0.47	0.5
0.7	0.4	0.45	0.34	0.41
0.8	0.38	0.31	0.31	0.39
0.9	0.35	0.29	0.27	0.3
1	0.05	0.06	0.05	0.07

Bảng 3.13. Các kết quả của truy vấn 6.

<i>Recall</i>	<i>Precision</i>			
	<i>CSI</i>	<i>QT</i>	<i>CBC</i>	<i>SR</i>
0.1	0.95	0.92	0.92	0.88
0.2	0.89	0.68	0.84	0.83
0.3	0.75	0.59	0.75	0.71
0.4	0.73	0.51	0.69	0.7
0.5	0.54	0.46	0.5	0.57
0.6	0.59	0.32	0.49	0.54
0.7	0.51	0.33	0.36	0.42
0.8	0.47	0.19	0.37	0.41
0.9	0.39	0.15	0.28	0.31
1	0.08	0.05	0.07	0.06

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 [16] và CSI, các truy vấn 4, 5 và 6 được sử dụng thêm phương pháp SR [59, 75, 76]. Các truy vấn từ 1 đến 6 cùng với tập ảnh liên quan được tạo ra từ CSDL “WANG 1000” [76]. 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



a/ Thức ăn



b/ Ngựa



c/ Voi



d/ Bãi biển



e/ Núi



f/ Di tích cổ

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

Ba truy vấn đầu được thực hiện đối với ảnh truy vấn và tập ảnh liên quan tương đối hỗn tạp. Ba truy vấn sau thực hiện đối với ảnh truy vấn và tập ảnh liên quan tương đối thuần nhất.

Các Bảng 3.15, 3.16 và 3.17 đưa ra tóm tắt các kết quả dưới dạng Recall – Precision của các truy vấn 1, 2 và 3 tương ứng.

Bảng 3.15. Các kết quả của truy vấn 1.

<i>Recall</i>	<i>Precision</i>		
	<i>CCS</i>	<i>CCV</i>	<i>CSI</i>
0.1	0.9	0.94	0.93
0.2	0.82	0.91	0.89
0.3	0.77	0.76	0.84
0.4	0.72	0.73	0.82
0.5	0.63	0.62	0.73
0.6	0.59	0.61	0.69
0.7	0.49	0.51	0.54
0.8	0.47	0.47	0.47
0.9	0.39	0.23	0.32
1	0.08	0.07	0.07

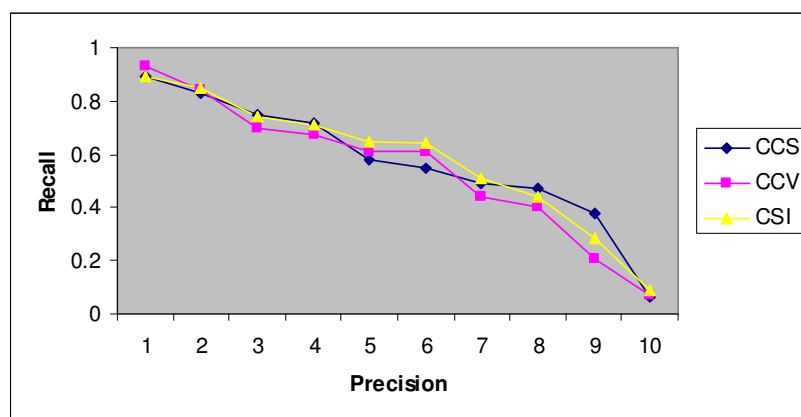
Bảng 3.16. Các kết quả của truy vấn 2.

<i>Recall</i>	<i>Precision</i>		
	<i>CCS</i>	<i>CCV</i>	<i>CSI</i>
0.1	0.9	0.92	0.89
0.2	0.81	0.86	0.86
0.3	0.72	0.7	0.71
0.4	0.69	0.67	0.68
0.5	0.52	0.64	0.62
0.6	0.51	0.65	0.61
0.7	0.46	0.43	0.48
0.8	0.45	0.38	0.43
0.9	0.35	0.22	0.24
1	0.05	0.08	0.15

Bảng 3.17. Các kết quả của truy vấn 3.

<i>Recall</i>	<i>Precision</i>		
	<i>CCS</i>	<i>CCV</i>	<i>CSI</i>
0.1	0.88	0.92	0.86
0.2	0.86	0.76	0.79
0.3	0.77	0.63	0.68
0.4	0.74	0.61	0.64
0.5	0.59	0.58	0.59
0.6	0.56	0.57	0.62
0.7	0.51	0.37	0.51
0.8	0.5	0.36	0.41
0.9	0.4	0.18	0.27
1	0.06	0.05	0.04

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.

Các Bảng 3.18, 3.19 và 3.20 đưa ra tóm tắt các kết quả của các truy vấn 4, 5 và 6 tương ứng. Các kết quả tra cứu được tóm tắt dưới dạng Recall - Precision. Với mỗi truy vấn, bốn thực nghiệm được thực hiện. Trong thực nghiệm đầu tiên, cách tiếp cận CCS được sử dụng cho quá trình tra cứu. Phương pháp CCV được sử dụng

cho thực nghiệm thứ hai. Cách tiếp cận CSI được sử dụng cho thực nghiệm thứ ba. Cuối cùng, cách tiếp cận SR được sử dụng trong thực nghiệm thứ tư.

Bảng 3.18. Các kết quả của truy vấn 4.

<i>Recall</i>	<i>Precision</i>			
	<i>CCS</i>	<i>CCV</i>	<i>CSI</i>	<i>SR</i>
0.1	0.89	0.9	0.92	0.92
0.2	0.81	0.81	0.89	0.83
0.3	0.71	0.68	0.82	0.76
0.4	0.7	0.63	0.78	0.63
0.5	0.56	0.53	0.61	0.58
0.6	0.52	0.52	0.59	0.54
0.7	0.47	0.34	0.47	0.36
0.8	0.46	0.33	0.42	0.35
0.9	0.39	0.23	0.32	0.25
1	0.08	0.07	0.07	0.08

Bảng 3.19. Các kết quả của truy vấn 5.

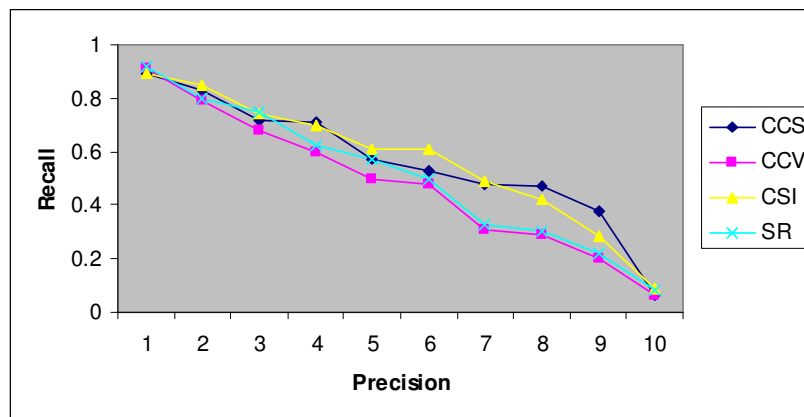
<i>Recall</i>	<i>Precision</i>			
	<i>CCS</i>	<i>CCV</i>	<i>CSI</i>	<i>SR</i>
0.1	0.9	0.91	0.89	0.92
0.2	0.81	0.8	0.86	0.81
0.3	0.7	0.69	0.72	0.77
0.4	0.69	0.57	0.68	0.61
0.5	0.55	0.48	0.62	0.56
0.6	0.51	0.46	0.61	0.48
0.7	0.46	0.29	0.48	0.31
0.8	0.45	0.25	0.43	0.27
0.9	0.35	0.19	0.24	0.21
1	0.05	0.07	0.15	0.09

Bảng 3.20. Các kết quả của truy vấn 6.

<i>Recall</i>	<i>Precision</i>			
	<i>CCS</i>	<i>CCV</i>	<i>CSI</i>	<i>SR</i>
0.1	0.88	0.92	0.86	0.92
0.2	0.86	0.76	0.79	0.77
0.3	0.75	0.66	0.68	0.72
0.4	0.74	0.61	0.64	0.63
0.5	0.6	0.48	0.59	0.57
0.6	0.56	0.46	0.62	0.47
0.7	0.51	0.29	0.51	0.31
0.8	0.5	0.28	0.41	0.29
0.9	0.4	0.18	0.27	0.21
1	0.06	0.05	0.04	0.06

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 CSDL.

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

Trong chương này, chúng tôi mô tả thiết kế và thực hiện của hệ thống tra cứu ảnh dựa vào đặc trưng màu cục bộ LVFIR, một hệ thống tra cứu ảnh dựa trên bốn kỹ thuật sau: Thứ nhất, phương pháp sử dụng ít không gian lưu trữ số lược đồ màu biểu diễn ảnh và ít nhạy cảm với quay và dịch chuyển, có tên là HG. Thứ hai, phương pháp tra cứu ảnh cải tiến dựa trên cơ sở phương pháp HG có tên là IHG. Thứ ba, kỹ thuật tra cứu ảnh dựa vào đặc trưng màu và thông tin không gian, gọi là CSI. Cuối cùng là kỹ thuật tra cứu khác sử dụng đặc trưng màu và các cụm màu thuần nhất của để phục vụ quá trình tra cứu, gọi là CCS.

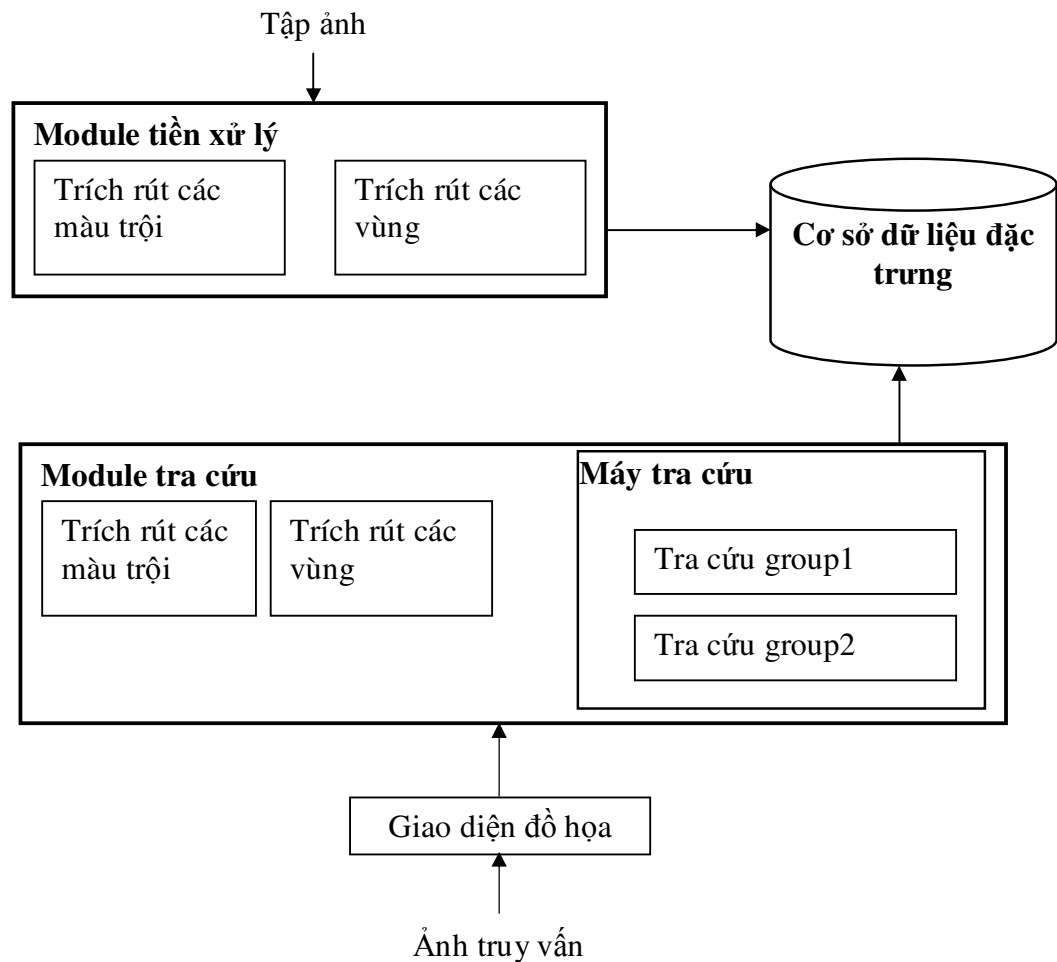
4.1 Thiết kế hệ thống tổng quát LVFIR

Hệ thống LVFIR được thiết kế và thực hiện trên hệ điều hành Windows XP sử dụng ngôn ngữ lập trình C# trong môi trường Visual Studio 2005. Kiến trúc toàn bộ hệ thống được chỉ ra trong hình 4.1. Kiến trúc này gồm hai module chính: module tiền xử lý và module tra cứu.

Ban đầu, CSDL ảnh được tiền xử lý (bởi module tiền xử lý) để trích rút các véc tơ đặc trưng. Module tra cứu nhận ảnh truy vấn từ người sử dụng thông qua giao diện đồ họa, trích rút các véc tơ đặc trưng từ ảnh truy vấn, và so sánh với các ảnh trong cơ sở dữ liệu ảnh.

Trong module tra cứu lại được chia ra làm hai module con, sử dụng chung chức năng trích rút đặc trưng ảnh: module tra cứu group1 và module tra cứu group2:

- Module tra cứu group1 sử dụng kỹ thuật tra cứu HG và kỹ thuật cải tiến IHG.
- Module tra cứu group2 thực hiện quá trình tra cứu áp dụng kỹ thuật CSI và CCS.



Hình 4.1. Kiến trúc của hệ thống LVFIR.

Trong hệ thống tra cứu LVFIR, các màu trội và vùng của các ảnh trong tập ảnh được trích rút và được mô tả bởi các véc tơ đặc trưng. Các véc tơ đặc trưng của các ảnh trong tập ảnh tạo ra CSDL đặc trưng.

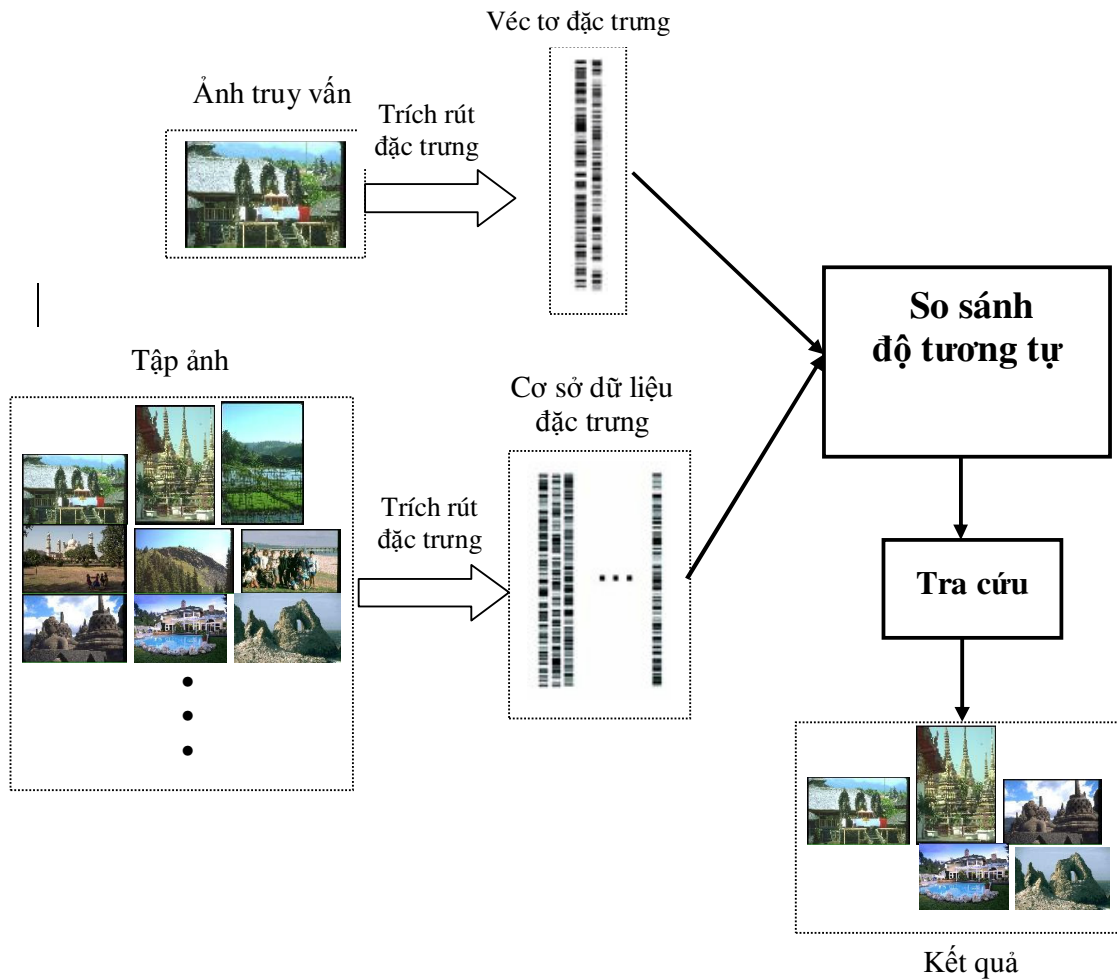
Để tra cứu các ảnh, người sử dụng cung cấp cho hệ thống ảnh truy vấn (thông qua giao diện đồ họa). Sau đó hệ thống trích rút các màu trội và các vùng của ảnh truy vấn và biểu diễn bởi véc tơ đặc trưng.

Độ tương tự giữa véc tơ đặc trưng của ảnh truy vấn và các véc tơ đặc trưng trong CSDL đặc trưng được tính toán và phân hạng thông qua máy tra cứu.

CSDL ảnh gồm 7,812 ảnh được sử dụng cho thực nghiệm. Tập con ảnh này được lấy từ tập ảnh của GS Wang [76] và một phần tác giả thu qua Internet và camera số. Tập ảnh đã được chuẩn hoá có cỡ 128 x 85 điểm ảnh hoặc 85 x 128 điểm ảnh. Các ảnh được tiền xử lý để trích rút véc tơ đặc trưng thích hợp và được lưu trữ trong CSDL đặc trưng.

Giao diện đồ hoạ được thiết kế để hiển thị 50 ảnh trên cùng mà được phân hạng theo thứ tự giảm dần của độ tương tự. Người sử dụng có thể chỉ rõ các truy vấn dựa vào đặc trưng thị giác bằng việc lựa chọn ảnh truy vấn.

4.2 Module tra cứu group1



Hình 4.2. Kiến trúc của Module tra cứu group1.

Trong module tra cứu group1 (Hình 4.2), các màu trội và các khối (các khối có kích cỡ bằng nhau) của các ảnh trong tập ảnh được trích rút và được mô tả bởi các véc tơ đặc trưng.

Để tra cứu các ảnh, người sử dụng cung cấp cho hệ thống ảnh truy vấn (thông qua giao diện đồ họa). Sau đó hệ thống trích rút các màu trội và các khối của ảnh truy vấn và biểu diễn bởi véc tơ đặc trưng.

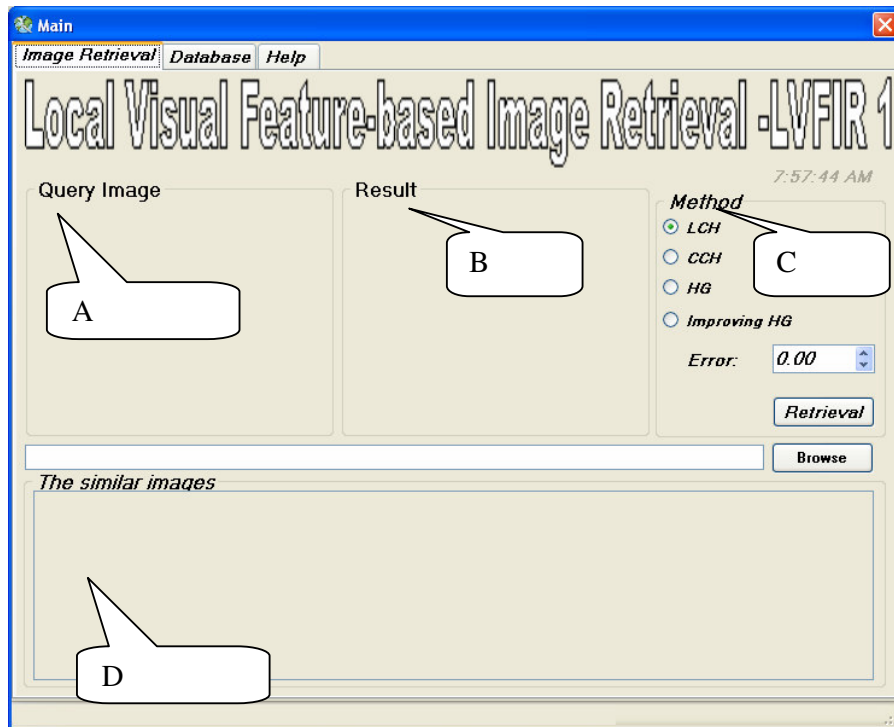
Độ tương tự giữa véc tơ đặc trưng của ảnh truy vấn và các véc tơ đặc trưng trong CSDL đặc trưng được tính toán và phân hạng thông qua module tra cứu group1, module này sử dụng kỹ thuật tra cứu HG và IHG đã được trình bày trong chương 2.

Hình 4.3 chỉ ra giao diện người sử dụng được thực hiện, và một truy vấn mẫu với một số kết quả của nó. Giao diện cũng cho phép người sử dụng “click” lên một ảnh để xem thông tin chi tiết hơn về ảnh. Từ ảnh truy vấn mẫu này, hệ thống có thể tra cứu các ảnh liên quan (thông qua các véc tơ đặc trưng). Giao diện chính của chương trình gồm có 4 vùng chính là A, B, C và D trên cửa sổ ứng dụng. Trong đó:

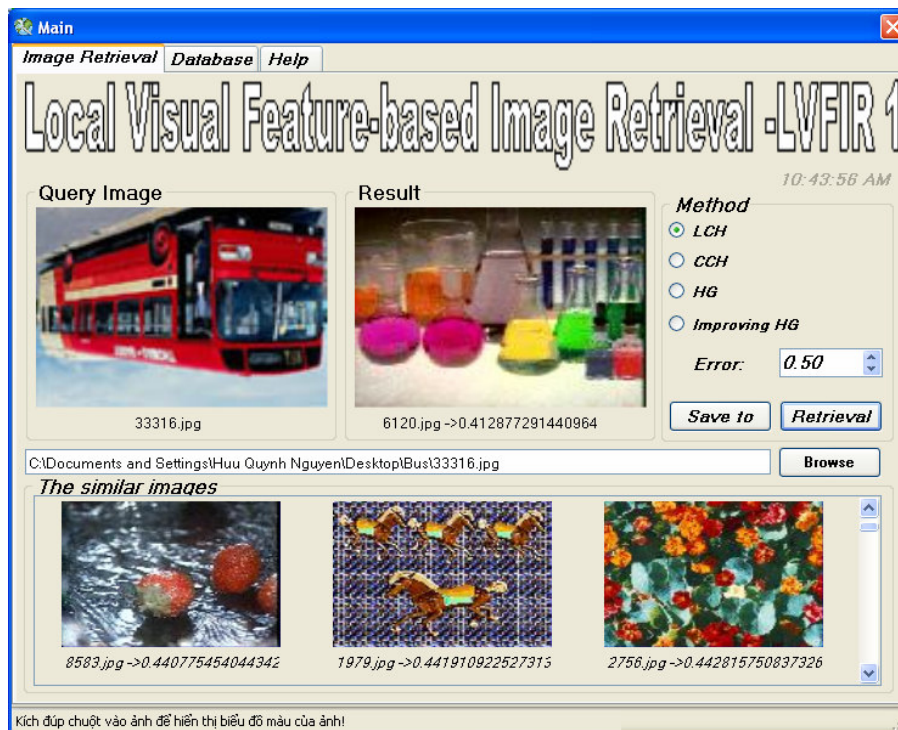
A. Ảnh truy vấn (Query image) và B. Ảnh tương tự với ảnh truy vấn nhất (Result).

C. Các phương pháp tra cứu: LCH (phương pháp lược đồ màu cục bộ), CCH (phương pháp lược đồ màu khối), HG (phương pháp HG), Improving HG (phương pháp HG cải tiến).

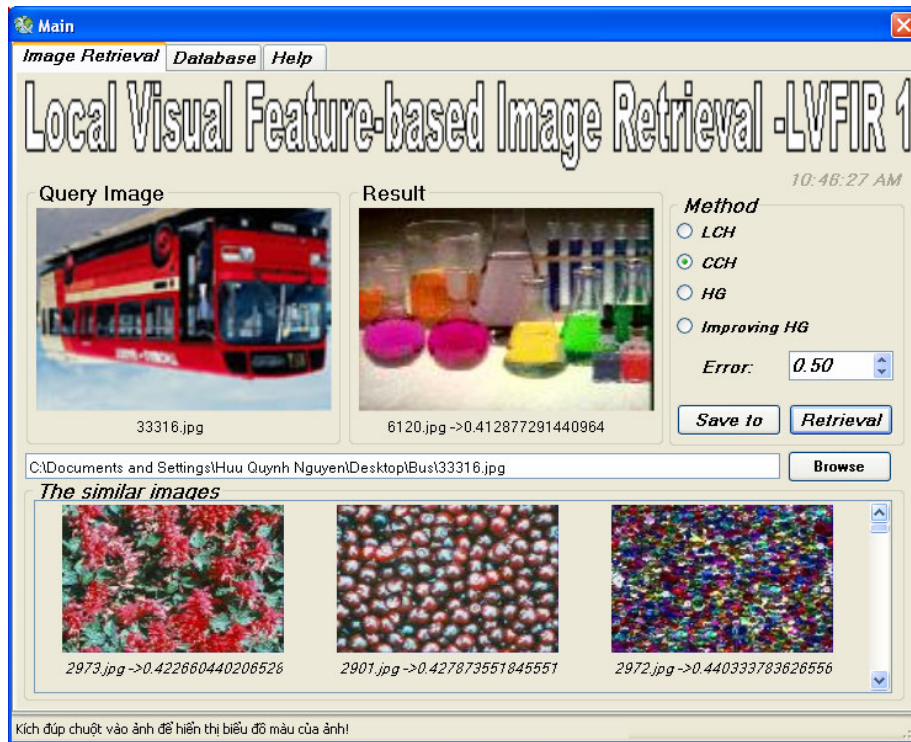
D. 50 ảnh tương tự nhất với ảnh truy vấn được hiển thị. Độ tương tự được hiển thị cùng với mỗi ảnh trong đó giá trị đầu tiên là LCH, thứ hai là CCH, sau đó đến HG và cuối cùng là HG cải tiến.



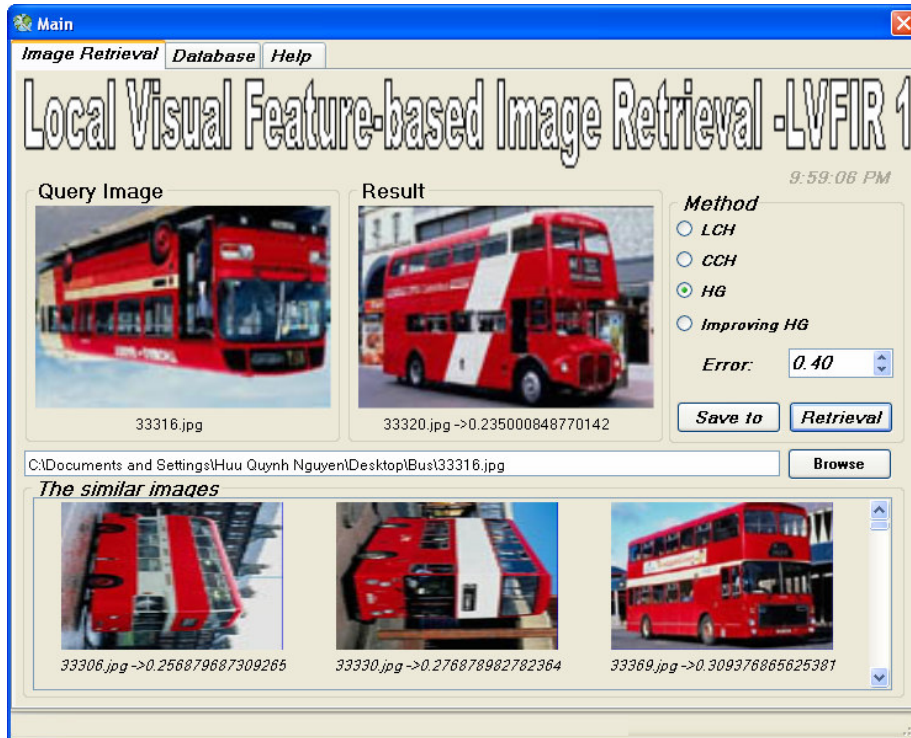
Hình 4.3. Màn hình chính của module tra cứu group1.



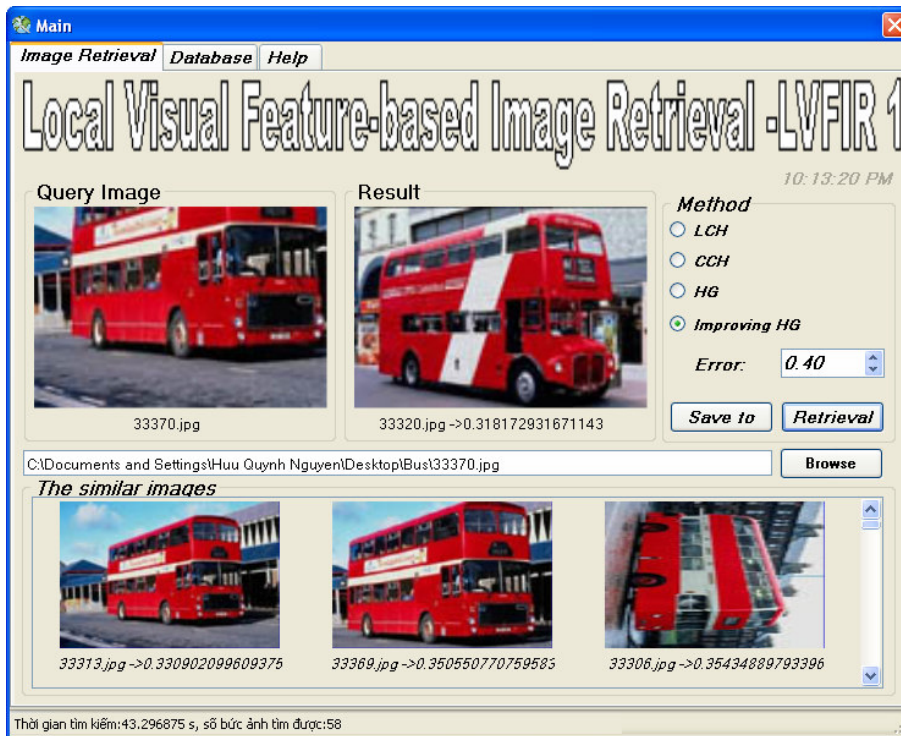
Hình 4.4. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng LCH.



Hình 4.5. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng CCH.

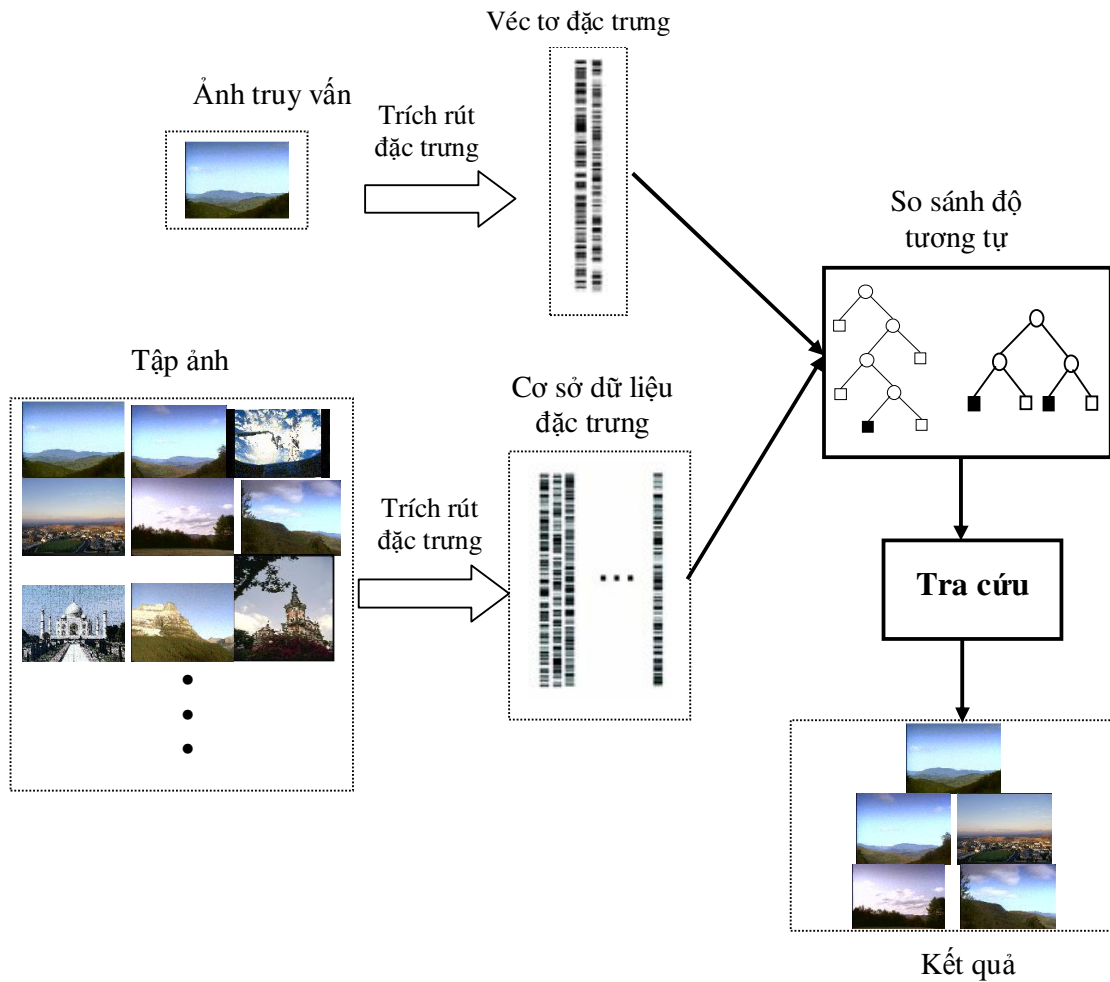


Hình 4.6. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng HG.



Hình 4.7. Giao diện tra cứu khi lựa chọn đặc điểm màu sử dụng IHG.

4.3 Module tra cứu group2



Hình 4.8. Kiến trúc của Module tra cứu group2.

Trong module tra cứu group2 (Hình 4.8), các màu trội và vùng của các ảnh trong tập ảnh được trích rút và được mô tả bởi các véc tơ đặc trưng.

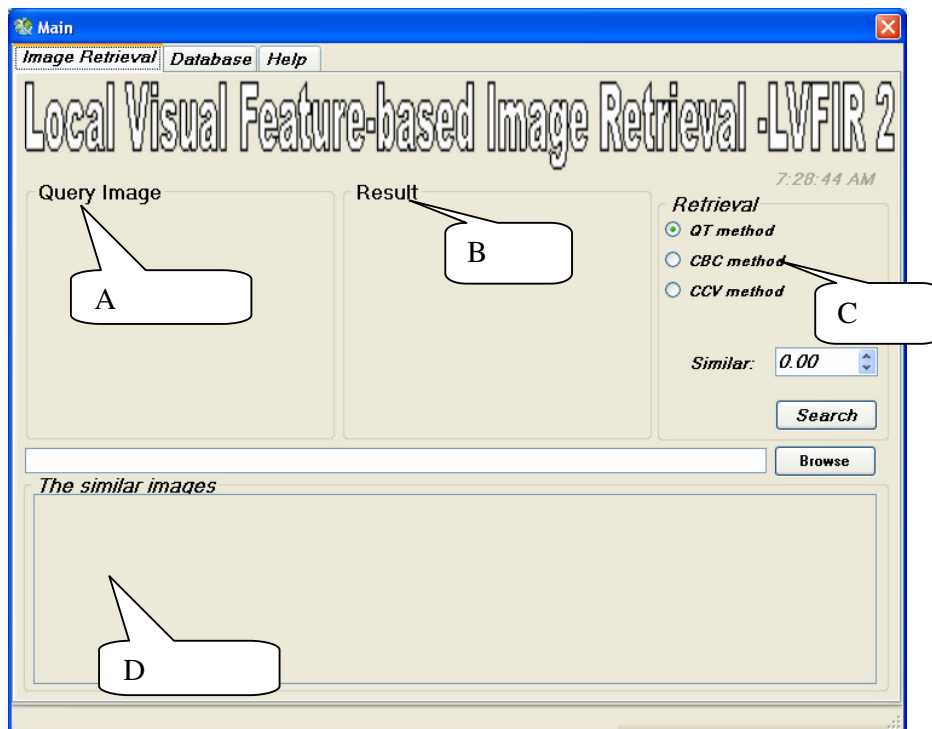
Để tra cứu các ảnh, người sử dụng cung cấp cho hệ thống ảnh truy vấn (thông qua giao diện đồ họa). Sau đó hệ thống trích rút các màu trội và các vùng của ảnh truy vấn (các vùng này thường có kích cỡ khác nhau) và biểu diễn bởi véc tơ đặc trưng.

Độ tương tự giữa véc tơ đặc trưng của ảnh truy vấn và các véc tơ đặc trưng trong CSDL đặc trưng được tính toán và phân hạng thông qua module tra cứu

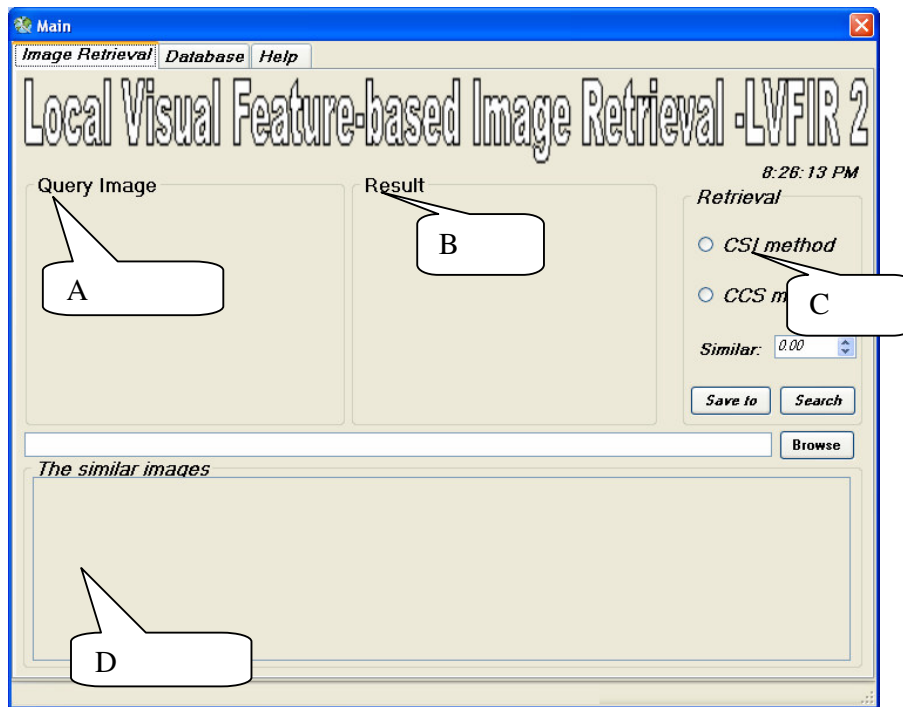
group2, module này sử dụng kỹ thuật tra cứu CSI và CCS đã được trình bày trong chương 3.

Hình 4.9 và 4.10 chỉ ra giao diện người sử dụng. Giao diện cũng cho phép người sử dụng “click” lên một ảnh để xem thông tin chi tiết hơn về ảnh. Từ ảnh truy vấn mẫu này, hệ thống có thể tra cứu các ảnh liên quan (thông qua các véc tơ đặc trưng). Giao diện chính của chương trình gồm có 4 vùng chính là A, B, C và D trên cửa sổ ứng dụng. Trong đó:

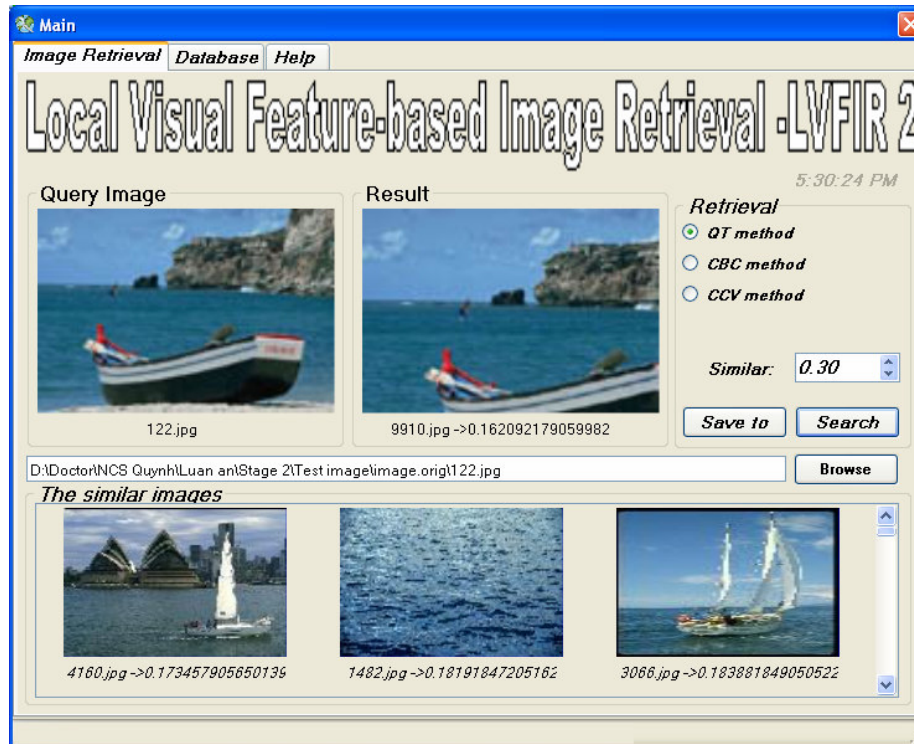
- A. Ảnh truy vấn (Query image) và B. Ảnh tương tự với ảnh truy vấn nhất (Result).
- C. Phương pháp tra cứu và độ tương tự cho phép.
- D. 50 ảnh tương tự nhất với ảnh truy vấn được hiển thị.



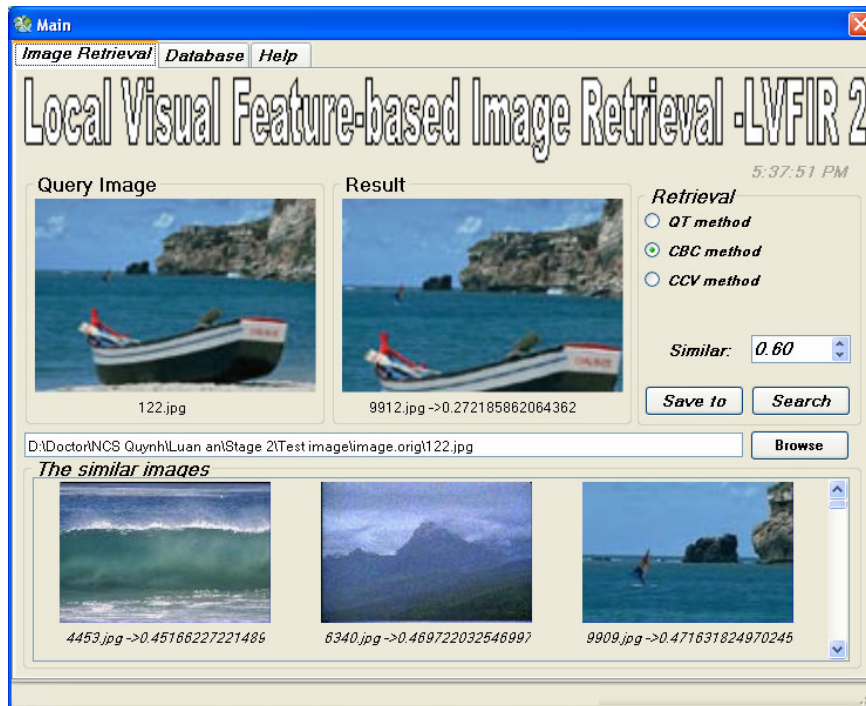
Hình 4.9. Giao diện sử dụng kỹ thuật QT, CBC và CCV của module tra cứu group2.



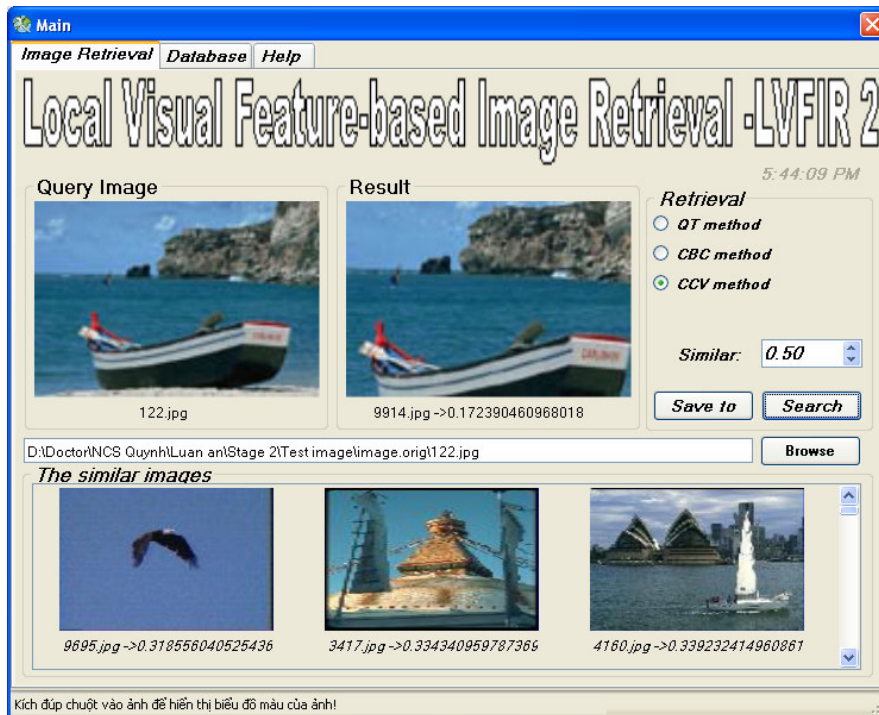
Hình 4.10. Giao diện sử dụng kỹ thuật CSI và CCS của module tra cứu group2.



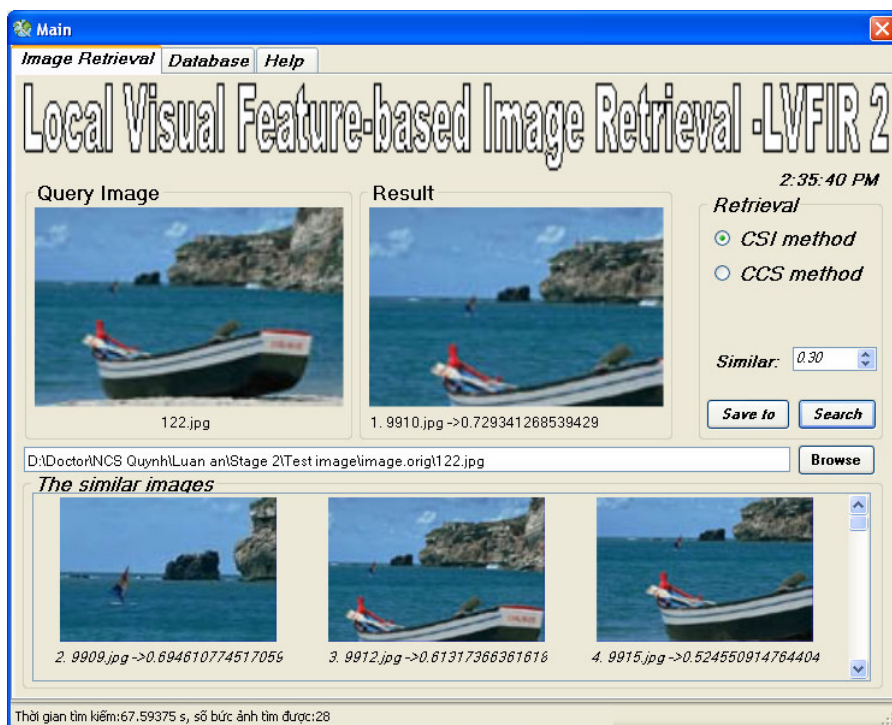
Hình 4.11. Giao diện tra cứu khi sử dụng phương pháp QT với ảnh truy vấn.



Hình 4.12. Giao diện tra cứu khi sử dụng phương pháp CBC với ảnh truy vấn.



Hình 4.13. Giao diện tra cứu khi sử dụng phương pháp CCV với ảnh truy vấn.



Hình 4.14. Giao diện tra cứu khi sử dụng phương pháp CSI với ảnh truy vấn.

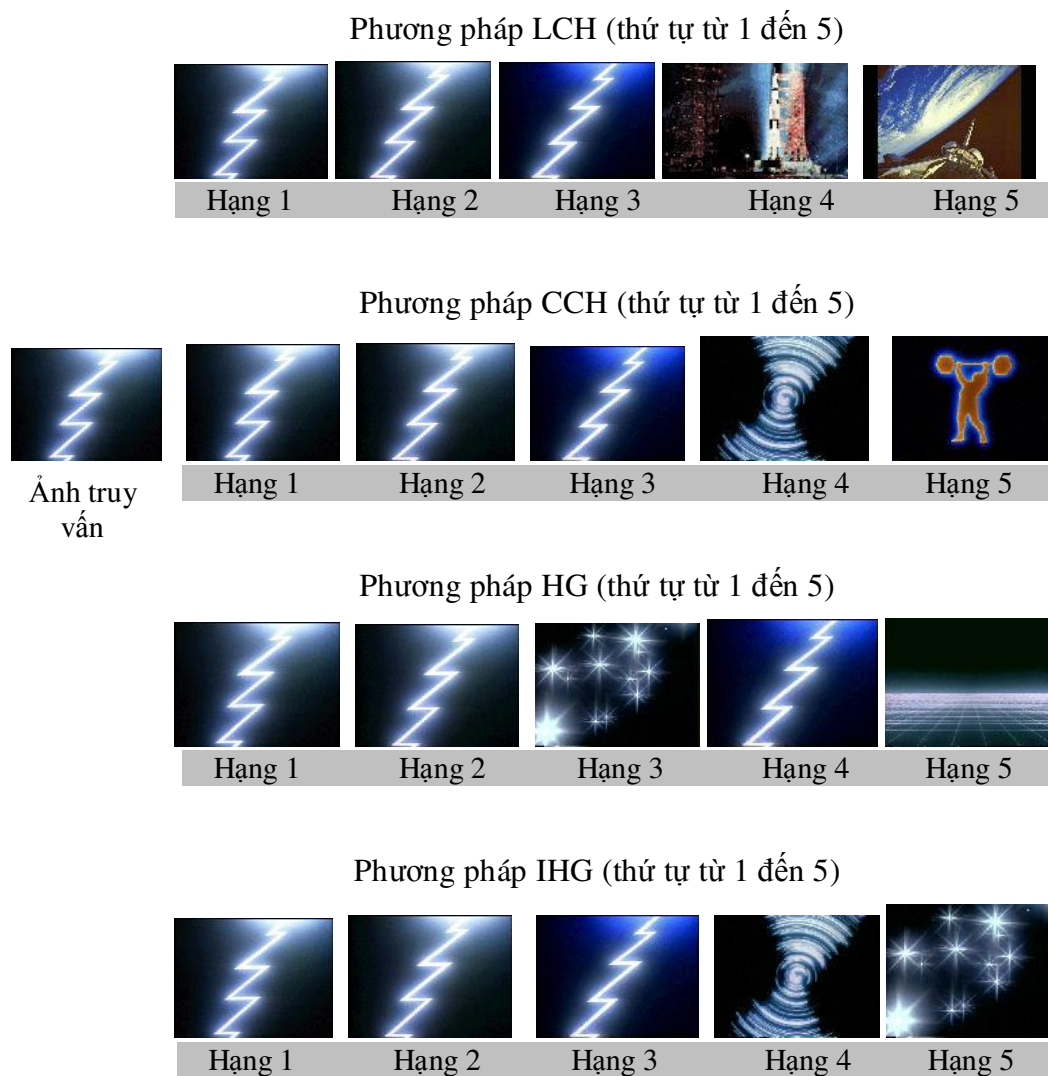


Hình 4.15. Giao diện tra cứu khi sử dụng phương pháp CCS với ảnh truy vấn.

4.4 Một số kết quả

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

Truy vấn 1:



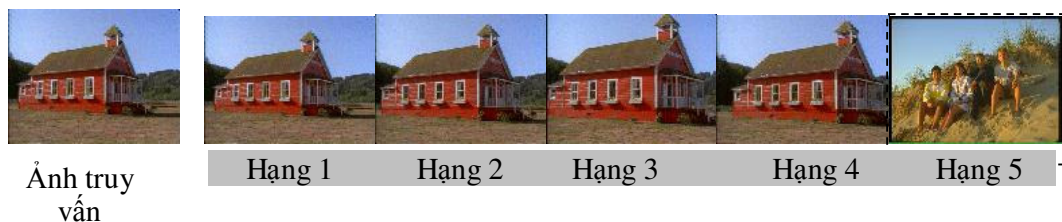
Hình 4.16. Kết quả thực hiện truy vấn 1.

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

Phương pháp LCH (thứ tự từ 1 đến 5)



Phương pháp CCH (thứ tự từ 1 đến 5)



Ảnh truy
vấn

Phương pháp HG (thứ tự từ 1 đến 5)



Phương pháp IHG (thứ tự từ 1 đến 5)



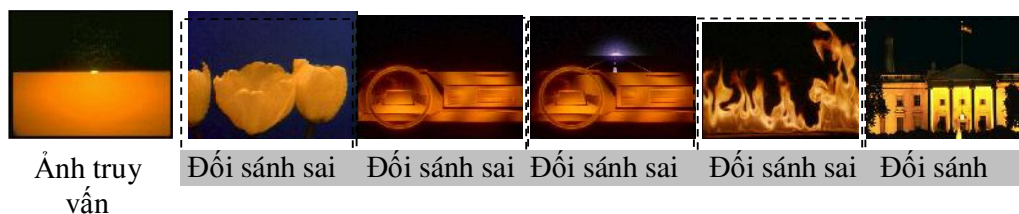
Hình 4.17. Kết quả thực hiện truy vấn 2.

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

Phương pháp LCH (thứ tự từ 1 đến 5)



Phương pháp CCH (thứ tự từ 1 đến 5)



Phương pháp HG (thứ tự từ 1 đến 5)



Phương pháp IHG (thứ tự từ 1 đến 5)



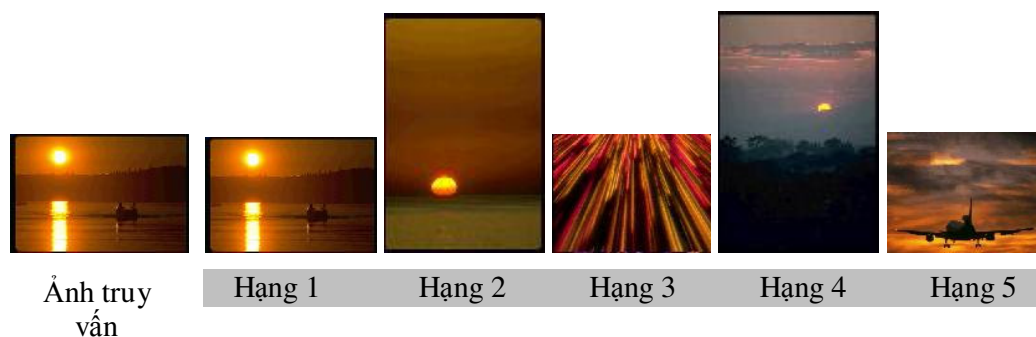
Hình 4.18. Kết quả thực hiện truy vấn 3.

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.

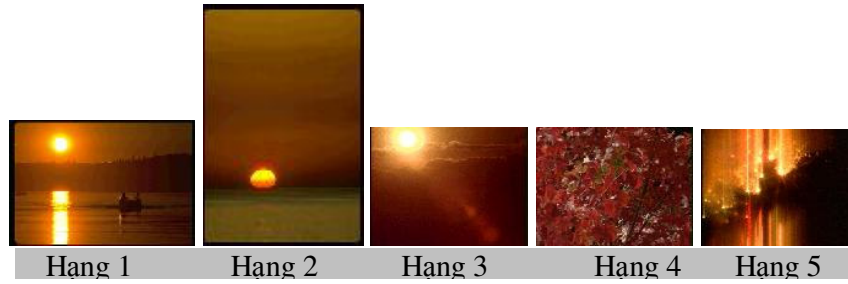
4.4.2 So sánh kỹ thuật QT, CBC và CCV với CSI và CCS

Truy vấn 1:

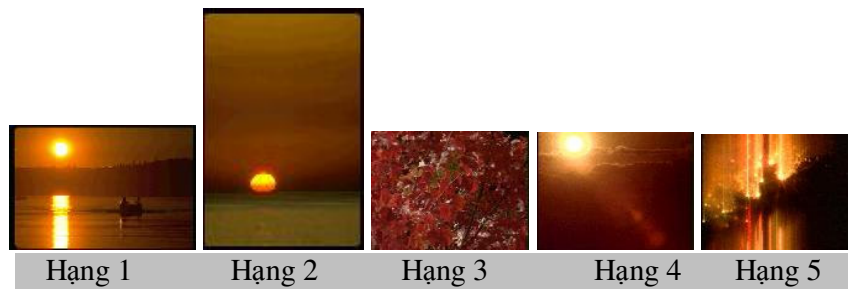
Phương pháp QT



Phương pháp CCV



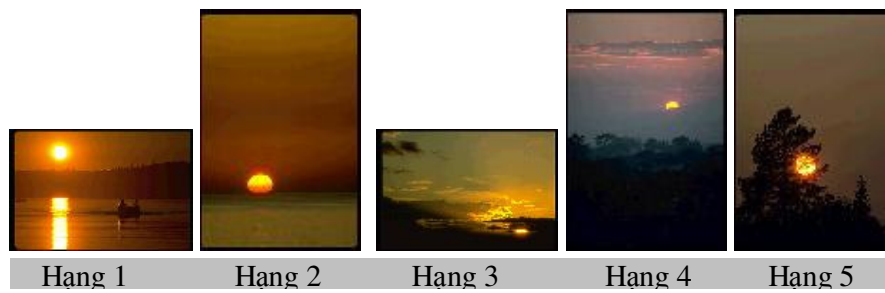
Phương pháp CBC



Phương pháp CSI

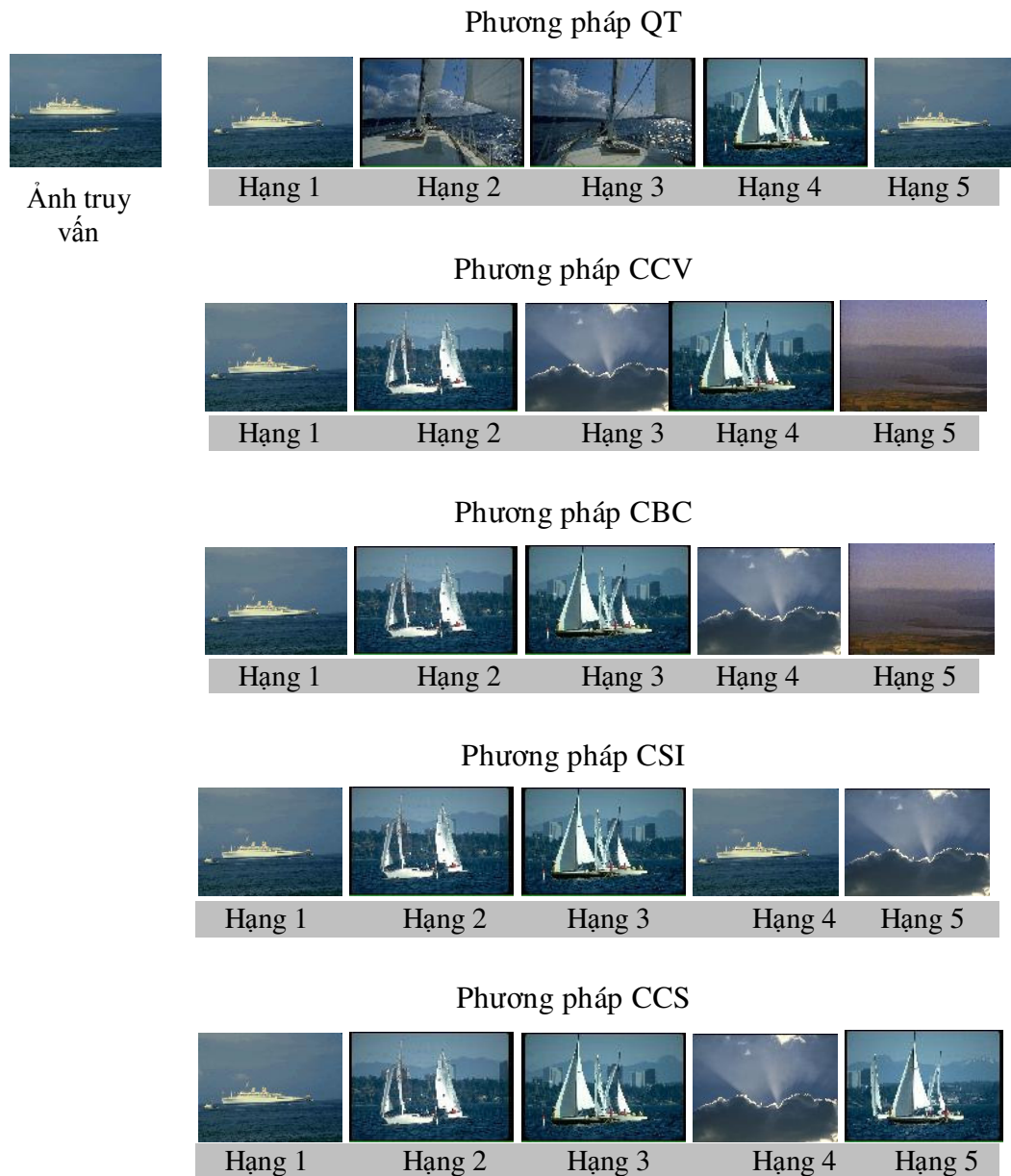


Phương pháp CCS



Hình 4.19. Kết quả thực hiện truy vấn 1.

Truy vấn 2:



Hình 4.20. Kết quả thực hiện truy vấn 2.

Truy vấn 3:

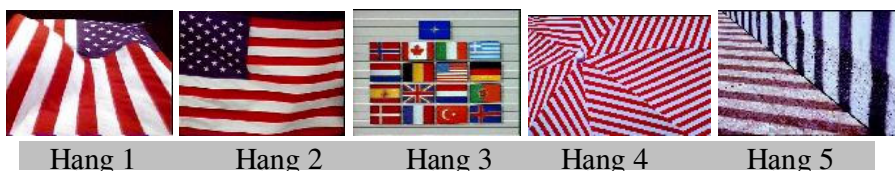
Phương pháp QT



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 công cụ C# và hệ quản trị cơ sở dữ liệu SQL Server 2005 trên hệ điều hành Windows XP, bộ xử lý Pentium 1.73 GHz, 512 MB bộ nhớ với CSDL ảnh gồm 7,812 ảnh. Đối với CSDL 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 CSDL ảnh này, phương pháp CSI và CCS cho độ chính xác cao hơn phương pháp QT, CCV và CBC.

KẾT LUẬN

Nghiên cứu về đặc trưng thị giác và trích rút các đặc trưng thị giác đã được thực hiện trong một thời gian dài. 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 các lược đồ màu biểu diễn ảnh.

Để giải quyết vấn đề giảm không gian lưu trữ 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 lược đồ màu toàn cục GCH, lược đồ màu cục bộ 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 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 biểu diễn ảnh. Phương pháp này đã được công bố trên tạp chí quốc tế IJCSES.

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 IJCSES.

Thứ ba, luận án đã đề xuất được kỹ thuật tra cứu ảnh CSI dựa vào đặc trưng màu và thông tin không gian. 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 CSDL ảnh để tăng tốc độ quá trình tra cứu ảnh.
- Thực nghiệm trên CSDL 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 (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.
- [2] 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. 43-47.
- [3] Quynh, N. H. and Tao, N. Q., Giang, N. T. (2008), “An efficient method for content based image retrieval using histogram graph”, *Proc. of IEEE on Control, Automation, Robotics and Vision*, pp. 874-878.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1]. 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.
- [2]. Ngô Quốc Tạo, Ngô Trường Giang, Nguyễn Hữu Quỳnh (2005), “Tra cứu ảnh dựa trên nội dung sử dụng biểu đồ màu cục bộ cải tiến”, *Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông*, tr. 543-550, Hải Phòng.

Tiếng Anh

- [3]. A. C. She and T. S. Huang (1994), “Segmentation of road scenes using color and fractal-based texture classification”, In *Proc. ICIP, Austin*, pp. 1026-1030.
- [4]. B. S. Manjunath, and W. Y. Ma (1996), "Texture features for browsing and retrieval of image data", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, pp. 837-842.
- [5]. B. Yates and R. Neto (1999), *Modern Information Retrieval*, Addison Wesley.
- [6]. Carson C, Belongie S, Greenspan H, Malik J (2002), Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8): pp.1026–1038.
- [7] Chua T. S., Lim S. K., Pung H. K. (1994), “Content-based retrieval of segmented images”, *ACM Multimedia*, San Francisco, Ca., USA, pp. 211-218.
- [8] D. Tegolo (1994), "Shape analysis for image retrieval", *Proc. of SPIE, Storage and Retrieval for Image and Video Databases -II*, no. 2185, San Jose, CA, pp. 59-69.
- [9] Deng, Y., Manjunath, B. S., Kenney, C., Moore, M. S., and Shin, H. (2001). “An efficient color representation for image retrieval”, *IEEE Trans. on Image Processing*, 10(1), pp.140–147.
- [10] Dow, J. (1993), “Content-based retrieval in multimedia imaging”, In *Proc. of SPIE Storage and Retrieval for Image and Video Databases*, pp. 164-167.
- [11] Equitz, W. and Niblack, W. (1994), *Retrieving images from a database using texture algorithms from the QBIC system*, Technical Report RJ 9805, Computer Science, IBM Research.

- [12] Forsyth D A, Ponce J (2002), *Computer Vision: A Modern Approach*, Prentice Hall, pp. 599–619.
- [13] Faloutsos, C., Flickner, M., Niblack, W., Petkovic, D., Equitz, W., and R.Barber (1993), *Efficient and effective querying by image content*, *Journal of Intelligent Information Systems*, pp. 231-262.
- [14] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., and Yanker, P. (1995), “Query by image and video content: The QBIC project”, *IEEE Computer*, 28(9), pp. 23 - 32.
- [15] Fukunaga, K. (1990), *Introduction to Statistical Pattern Recognition*. Academic Press.
- [16] G. Pass, and R. Zabith (1996), "Histogram refinement for content-based image retrieval", *IEEE Workshop on Applications of Computer Vision*, pp. 96-102.
- [17] G.Pass, and R. Zabith (1999), "Comparing images using joint histograms", *Multimedia Systems*, Vol.7, pp. 234-240.
- [18] German, D. (1990), “Boundary detection by constrained optimization”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 609- 628.
- [19] Geusebroek, J. M., van den Boomgaard, R., Smeulders, A. W. M., and Geerts, H. (2001), “Color invariance”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(12), pp. 1338–1350.
- [20] Gevers, T. and Smeulders, A. W. M. (1999), “Color based object recognition”, *Pattern Recognition*, 32, pp. 453–464.
- [21] Gunther, N. and Beretta, G. (2001), “A benchmark for image retrieval using distributed systems over the internet: BIRDS-I”, *SPIE Vol. 4311*, pp. 252-267.
- [22] Google Corporation (2009), <http://images.google.com>
- [23] H. Samet (1984), "The quadtree and related hierarchical data structures", *ACM Computing Surveys*, Vol.16, No.2, pp. 187-260.
- [24] H. Tamura, S. Mori, and T. Yamawaki (1978), “Texture features corresponding to visual perception”, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-8, no. 6, pp. 460 - 473.
- [25] H. V. Jagadish (1991), "A retrieval technique for similar shapes", *Proc. of Int. Conf. on Management of Data, SIGMOID'91*, Denver, CO, pp. 208-217.
- [26] Hafner, J., Sawhney, H. S., Equitz, W., Flickner, M., and Niblack, W. (1995), “Efficient color histogram indexing for quadratic form”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(7), pp. 729–736.
- [27] Hungarian algorithm http://en.wikipedia.org/wiki/Hungarian_algorithm.

- [28] James Z. Wang, Jia Li, Gio Wiederhold (2001), "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol 23, no.9, pp. 947-963.
- [29] J. E. Gary, and R. Mehrotra (1992), "Shape similarity-based retrieval in image database systems", *Proc. of SPIE, Image Storage and Retrieval Systems*, Vol. 1662, pp. 2-8.
- [30] J. Huang, *et al.*(1997), "Image indexing using color correlogram", *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pp. 762-768.
- [31] J. Kender and B. Yeo (1998), "Video scene segmentation via continuous video coherence", In *Proc. of IEEE Computer Vision and Pattern Recognition*, Santa Barbara, CA, IEEE Computer Society, pp. 367-373.
- [32] K. Ravishankar, B. Prasad, S. Gupta, and K. Biswas (1999), "Dominant color region based indexing for CBIR", *Proc. of the International Conference on Image Analysis and Processing*, pp. 887-892.
- [33] Lee, D., Barber, R., Niblack, W., Flickner, M., Hafner, J., and Petkovic, D. (1994), "Indexing for complex queries on a query-by-content image database", In *Proc. of IEEE Int'l Conf. on Image Processing*, vol.1, pp. 142-146.
- [34] M. A. Stricker and M. J. Swain (1994) "The capacity of color histogram indexing", In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Madison, Wisconsin, pp. 704-708.
- [35] . M. Lybanon, S. Lea, and S. Himes (1994), "Segmentation of diverse image types using opening and closing", In *Proc. IEEE Int. Conf. on Image Proc*, vol.1, pp. 347-351.
- [36] M. Stricker, and M. Orengo (1995), "Similarity of color images", *SPIE Storage and Retrieval for Image and Video Databases III*, vol. 2185, pp. 381-392.
- [37] M. Worring and Th. Gevers (2001), "Interactive retrieval of color images", *International Journal of Image and Graphics*, 1(3), pp. 387-414.
- [38] Ma, W.-Y. and Manjunath, B. S. (1997), "Netra: A toolbox for navigating large image databases", In *Proc. of IEEE Int. Conf. on Image Processing*, vol.1, pp. 568-571.
- [39] Manjunath, B. S., Ohm, J. R., Vasudevan, V. V., and Yamada, A. (2001), "Color and texture descriptors", *IEEE Tran. on Circuits and Systems for Video Technology*, 11(6), pp. 703-715.
- [40] Nagasaka A., Tanaka Y.(1992), "Automatic video indexing and full-video search for object appearances", *Journal of Information Processing*, vol.15, no.2, Information Processing Society of Japan, Tokyo, pp. 113-127.

- [41] Pi, M., Mandal, M. K., and Basu, A. (2005), “Image retrieval based on histogram of fractal parameters”, *IEEE Trans. Multimedia* 7, 4, pp. 597–605.
- [42] 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, 5 pp. 1-6.
- [43] 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. 43-47.
- [44] Quynh, N. H. and Tao, N. Q (2008), “Improving Harbin method for retrieving landscape images”, In Proc. of IEEE on Intelligent Information Hiding and Multimedia Signal Processing, pp. 771-774.
- [45] 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, pp. 480-484.
- [46] 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.
- [47] Quynh, N. H. and Tao, N. Q., Giang, N. T. (2008), “A efficient method for content based image retrieval using histogram graph”, In Proc. of IEEE on Control, Automation, Robotics and Vision, pp. 874-878.
- [48] Quynh, N. H. and Tao, N. Q., Giang, N. T. (2008), “Efficient content based image retrieval through sector histogram”, In Proc. of IEEE on Circuits and Systems, pp. 1814-1817.
- [49] Q. Iqbal and J. K. Aggarwal (2002), “CIRES: A System for Content-based Retrieval in Digital Image Libraries”, *International Conference on Control, Automation, Robotics and Vision*, pp. 205-210.
- [50] R. Diestel (1997), *Graph theory: Graduate texts in mathematics*, 173, New York : Springer.
- [51] R. Haralick and L. Shapiro (1993), *Computer and Robot Vision*, Addison-Wesley.
- [52] R. Datta, J. Li, and J. Z. Wang (2008), “Algorithmic Inferencing of Aesthetics and Emotion in Natural Images: An Exposition”, *Proc. IEEE ICIP*, pp. 105-108.
- [53] R. Samadani and C. Han (1993), “Computer-assisted extraction of boundaries from images”, In *Proc. SPIE Storage and Retrieval for Image and Video Databases*, pp. 219-225.

- [54] R.O Stehling, M.A. Nascimento, A.X. Falcão (2003), "Cell histograms versus color histograms for image representation and retrieval", *Knowledge and Information Systems (KAIS) Journal*, pp. 151-179.
- [55] R.O. Stehling, M.A. Nascimento, and A.X Falcão (2001), An adaptive and efficient clustering-based approach for content based image retrieval in image databases, In Proc. of the Intl. Data Engineering and Application Symposium, pp. 356–365.
- [56] R.O. Stehling, M.A. Nascimento, A.X. Falcão (2002), *Techniques for color-based image retrieval*, in: C. Djeraba (Ed.), *Multimedia Mining—A Highway to Intelligent Multimedia Documents*, Kluwer Academic, Dordrecht (Chapter 4).
- [57] Rafael C. Gonzalez, Richard E. Woods (2000), *Digital Image Processing*, Addison-Wesley, New York.
- [58] Ramesh Jain, Rangachar Kastun, Brian G. Schunck (1995), *Machine Vision* (Chapter 3), McGRAW-HILL, pp. 89-91.
- [59] Ritendra Datta, Dhiraj Joshi, Jia Li and James Z. Wang (2008), "Image Retrieval: Ideas, Influences, and Trends of the New Age," *ACM Computing Surveys*, vol. 40, no. 2, pp. 1-60.
- [60] Rubner, Y., Tomasi, C., and Guibas, L. J. (1998), "A metric for distributions with applications to image databases", In *Proc. of IEEE Computer Vision*, 1998. Sixth International Conference on, pp. 59-66.
- [61] S. K. Chang, E. Jungert, and Y. Li (1989), "Representation and retrieval of symbolic pictures using generalized 2D string", In: *SPIE Proceedings on Visual Communications and Image Processing*, Philadelphia, pp. 1360-1372.
- [62] S. K. Chang, Q. Y. Shi, and C. Y. Yan (1987), "Iconic indexing by 2-D strings", *IEEE Trans. on Pattern Anal. Machine Intell.*, vol.9, no.3, pp. 413-428.
- [63] S. Wang (2001), "A Robust CBIR Approach Using Local Color Histogram", *Technique Report*, Edmonton, Alberta, Canada.
- [64] S.-F. Chang, W. Chen, H. J. Meng, H. Sundaram, and D. Zhong (1997), "Videoq: An automated content based video search system using visual cues", In *Proceeding of The Fifth ACM International Multimedia Conference*, Seattle WA, ACM Press, pp. 313-324.
- [65] Scassellati, B., Alexopoulos, S., and Flickner, M. (1994), "Retrieving images by 2D shape: a comparison of computation methods with human perceptual judgments", In *Proc. of SPIE Storage and Retrieval for Image and Video Databases*, pp. 2-14.
- [66] Schettini, R., Ciocca, G., and Zuffi, S. (2001), "Color Imaging Science: Exploiting Digital Media, Ed. R. Luo and L. MacDonald", *chapter A Survey on Methods for Colour Image Indexing and Retrieval in Image Database*, John Wiley.

- [67] Smith, J. R. and Chang, S.-F. (1996), *Intelligent Multimedia Information Retrieval*, Ed. M. T. Maybury, chapter Querying by color regions using the VisualSeek content-based visual query system, MIT Press.
- [68] Smith, J. R. and Chang, S.-F. (1997), *Visually searching the web for content*, IEEE Multimedia, volume 4, issue 3, pp. 12 - 20.
- [69] Swain, M. J. and Ballard, D. H. (1991), "Color indexing", *International Journal of Computer Vision*, 7(1), pp. 11–32.
- [70] Smeulders A W M, Worring M, Santini S, Gupta A, Jain R (2000), "Content-Based Image Retrieval at the End of the Early Years", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1349–1380.
- [71] TREC (2002), *Text retrieval conference*, <http://trec.nist.gov>.
- [72] V. Castelli and L. D. Bergman (2002), *Image Database Search and Retrieval of Digital Imagery*, John Wiley & Sons, Inc., New York.
- [73] V. N. Gudivada, and V. V. Raghavan (1995), "Design and evaluation of algorithms for image retrieval by spatial similarity", *ACM Trans. on Information Systems*, Vol. 13, No. 2, pp. 115-144.
- [74] W. Niblack et al.(1993), "Querying images by content, using color, texture, and shape", *SPIE Conference on Storage and Retrieval for Image and Video Database*, Vol. 1908, pp.173-187.
- [75] Wang, Y. H. (2003), "Image indexing and similarity retrieval based on spatial relationship model". *Inf. Sci. Inf. Comput. Sci.* 154, 1-2, pp. 39–58.
- [76] Wang's research group (2004), <http://wang.ist.psu.edu/docs/related.shtml>
- [77] X. Q. Li, Z. W. Zhao, H. D. Cheng, C. M. Huang, and R. W. Harris (1994), "A Fuzzy logic approach to image segmentation", In *Proc. IEEE Int. Conf. on Image Proc.*, pp. 337-341.
- [78] Y. Gong, H. J. Zhang, and T. C. Chua (1994), "An image database system with content capturing and fast image indexing abilities", *Proc. IEEE International Conference on Multimedia Computing and Systems*, Boston, pp.121-130.
- [79] Y. Rui, T. Huang, and S. Chang (1999), "Image retrieval: current techniques, promising directions and open issues", *Journal of Visual Communication and Image Representation*, 10(4), pp. 39–62.
- [80] T. Lehmann, M. G'uld, C. Thies, B. Fischer, K. Spitzer, D. Keysers, H. Ney, M. Kohnen, H. Schubert, B. Wein (2003), The IRMA Project – A State of the Art Report on Content-Based Image Retrieval in Medical Applications. Proc. Korea-Germany Joint Workshop on Advanced Medical Image Processing, pp. 161–171.