

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM**



**LÊ CÔNG KHANH**

**TÌM KIẾM ẢNH TRONG TẬP DỮ LIỆU ẢNH  
LỚN DỰA TRÊN ĐẶC TRƯNG**

**LUẬN VĂN THẠC SĨ**

**Chuyên ngành : CÔNG NGHỆ THÔNG TIN**

**Mã số ngành: 60480201**

**TP. HỒ CHÍ MINH, tháng 11 năm 2016**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM**



**LÊ CÔNG KHANH**

**TÌM KIẾM ẢNH TRONG TẬP DỮ LIỆU ẢNH  
LỚN DỰA TRÊN ĐẶC TRƯNG**

**LUẬN VĂN THẠC SĨ**

**Chuyên ngành : CÔNG NGHỆ THÔNG TIN**

**Mã số ngành: 60480201**

**CÁN BỘ HƯỚNG DẪN KHOA HỌC: TS. NGUYỄN THANH BÌNH**

**TP. HỒ CHÍ MINH, tháng 11 năm 2016**

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH TẠI**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM**

Cán bộ hướng dẫn khoa học: **TS. NGUYỄN THANH BÌNH**

*(Ghi rõ họ, tên, học hàm, học vị và chữ ký)*

Luận văn Thạc sĩ được bảo vệ tại Trường Đại học Công nghệ TP. HCM vào ngày 17 tháng 12 năm 2016.

Thành phần Hội đồng đánh giá Luận văn Thạc sĩ gồm:

*(Ghi rõ họ, tên, học hàm, học vị của Hội đồng chấm bảo vệ Luận văn Thạc sĩ)*

TT	Họ và tên	Chức danh Hội đồng
1		Chủ Tịch
2		Phản biện 1
3		Phản biện 2
4		Ủy viên
5		Ủy viên, Thư ký

Xác nhận của Chủ tịch Hội đồng đánh giá Luận văn sau khi Luận văn đã được sửa chữa (nếu có).

Chủ tịch Hội đồng đánh giá Luận văn

TRƯỜNG ĐH CÔNG NGHỆ TP. HCM

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

PHÒNG QLKH – ĐTSĐH

Độc lập – Tự do – Hạnh phúc

*TP. HCM, ngày..... tháng..... năm 2016*

## **NHIỆM VỤ LUẬN VĂN THẠC SĨ**

Họ tên học viên: Lê Công Khanh.....Giới tính: Nam.....

Ngày, tháng, năm sinh: 17-6-1977.....Nơi sinh: Tiền Giang.....

Chuyên ngành: Công Nghệ Thông Tin .....MSHV: 1441860050 .....

I- Tên đề tài:

Tìm kiếm ảnh trong tập dữ liệu ảnh lớn dựa trên đặc trưng.

II- Nhiệm vụ và nội dung:

- Nghiên cứu kỹ thuật tìm kiếm ảnh theo nội dung, từ đó, đề xuất mô hình tìm kiếm ảnh trong tập dữ liệu ảnh lớn dựa trên đặc trưng ảnh.

- Xây dựng hệ thống tìm kiếm ảnh trong tập dữ liệu ảnh lớn dựa trên đặc trưng ảnh.

III- Ngày giao nhiệm vụ: 23/01/2016

IV- Ngày hoàn thành nhiệm vụ: .....

V- Cán bộ hướng dẫn: Tiến Sỹ Nguyễn Thanh Bình

**CÁN BỘ HƯỚNG DẪN**

**KHOA QUẢN LÝ CHUYÊN NGÀNH**

(Họ tên và chữ ký)

(Họ tên và chữ ký)

## **LỜI CAM ĐOAN**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu, kết quả nêu trong Luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tôi xin cam đoan rằng mọi sự giúp đỡ cho việc thực hiện Luận văn này đã được cảm ơn và các thông tin trích dẫn trong Luận văn đã được chỉ rõ nguồn gốc.

Học viên thực hiện Luận văn

**Lê Công Khanh**

## LỜI CẢM ƠN

Trước hết em xin bày tỏ lòng biết ơn sâu sắc nhất tới thầy giáo hướng dẫn TS. Nguyễn Thanh Bình đã tận tình giúp đỡ em rất nhiều trong suốt quá trình tìm hiểu nghiên cứu và hoàn thành báo cáo luận văn.

Em xin chân thành cảm ơn các thầy cô đã trang bị cho em những kiến thức cơ bản cần thiết để em có thể hoàn thành luận văn này.

Xin gửi lời cảm ơn đến bạn bè những người luôn bên em đã động viên và tạo điều kiện thuận lợi cho em, tận tình giúp đỡ chỉ bảo em những gì em còn thiếu sót trong quá trình làm báo cáo luận văn.

Cuối cùng em xin bày tỏ lòng biết ơn sâu sắc tới những người thân trong gia đình đã giành cho em sự quan tâm đặc biệt và luôn động viên em.

Vì thời gian có hạn, trình độ hiểu biết của bản thân còn nhiều hạn chế. Vì vậy, trong đề án không tránh khỏi những thiếu sót, em rất mong nhận được sự đóng góp ý kiến của tất cả các thầy cô giáo cũng như các bạn bè để luận văn của em được hoàn thiện hơn.

Em xin chân thành cảm ơn!

Tác giả luận văn

**Lê Công Khanh**

## TÓM TẮT

Sự tăng không ngừng về lượng ảnh trên Web tạo nguồn ảnh phong phú đáp ứng được nguồn cung ảnh cho nhu cầu của con người. Mặc dù một số mô hình tìm kiếm ảnh đã ra đời đáp ứng phần nào nhu cầu tìm kiếm ảnh, song nâng cao chất lượng tìm kiếm luôn là vấn đề được đặt ra. Bài toán tìm kiếm ảnh, và nâng cao chất lượng xếp hạng ảnh đã và đang nhận được sự quan tâm đặc biệt.

Đề tài của luận văn: “Tìm kiếm ảnh trong tập dữ liệu ảnh lớn dựa trên đặc trưng” nhằm để giải quyết bài toán trên. Nhiệm vụ của luận văn là nghiên cứu các kỹ thuật tìm kiếm ảnh theo nội dung và đề xuất mô hình tìm kiếm ảnh trong tập dữ liệu ảnh lớn dựa trên đặc trưng ảnh.

Đầu tiên, luận văn nêu một số đặc trưng cơ bản của ảnh cũng như một số nghiên cứu liên quan về các đặc trưng ảnh. Tiếp theo, luận văn đề xuất phương pháp tìm kiếm ảnh theo nội dung, đưa ra mô hình phối hợp các đặc trưng kết hợp với toán tử LTP và xây dựng giải thuật truy vấn.

Luận văn đã được hiện thực quá trình truy vấn ảnh bằng giải thuật đã đề xuất. Quá trình thực nghiệm được thực hiện trên nhiều tập dataset khác nhau, chứa các đối tượng khác nhau như người, xe, hoa, cảnh... trong tập dữ liệu của Wang tại địa chỉ <http://wang.ist.psu.edu/docs/related/>. Tập dữ liệu này bao gồm 10.000 ảnh và thử nghiệm trên 2 phương pháp: Truy xuất bằng query và truy xuất dựa trên phương pháp support vector machine (SVM).

Qua kết quả thực nghiệm, chúng ta thấy rằng phương pháp đề xuất cho kết quả tốt hơn phương pháp SVM. Phương pháp đề xuất đã kết hợp các đặc trưng cục bộ và toàn cục lại với nhau. Đây là lý do giải thích tại sao phương pháp đề xuất tốt hơn phương pháp SVM.

## ABSTRACT

The continuous increase in volume images on the web to create images rich source supply to meet the demand picture for the people. Although some models were launched image search partly meet the need to look for images, but improve the quality of search is always a problem arises. Image search problem and improve image quality ratings have been received special attention.

The theme of the thesis: "Content-based image retrieval in the large image database based on the characteristic" in order to solve the problem on. The task of the thesis is to study the technical content image search and search suggestions model images in large image data sets based on image features.

Firstly, the thesis outlined some basic characteristics of the image as well as a number of studies related to the image features. Secondly, the thesis proposed the method for image searching by contents, making coordination model characteristics associated with LTP and building operators query algorithms.

Thesis has been realized by the image retrieval algorithm was proposed. The processing of this work do on many different training dataset, containing various objects such as people, cars, flowers, landscape ... in the data set of Wang at <http://wang.ist.psu.edu/docs/related/>. This dataset includes 10,000 images. We test on two methods: Retrieve query and retrieval using methods based on support vector machine (SVM).

Through experimental results, we found that the proposed method gives better results than SVM method. The method propose combining local characteristics and wikis together. This is the reason why the proposed method better methods based on SVM method.



## MỤC LỤC

LỜI CAM ĐOAN .....	i
LỜI CẢM ƠN .....	ii
TÓM TẮT.....	iii
ABSTRACT .....	iv
MỤC LỤC .....	v
DANH MỤC TỪ VIẾT TẮT .....	vii
DANH MỤC CÁC BẢNG.....	viii
DANH MỤC CÁC BIỂU ĐỒ, ĐỒ THỊ, SƠ ĐỒ, HÌNH ẢNH.....	ix
CHƯƠNG 1: GIỚI THIỆU .....	1
1.1 Giới thiệu đề tài.....	1
1.2 Mục tiêu của đề tài.....	2
1.3 Nội dung đề tài.....	2
1.4 Giới hạn của đề tài .....	2
1.5 Phương pháp nghiên cứu .....	2
1.6 Cấu trúc luận văn .....	3
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VÀ CÁC NGHIÊN CỨU LIÊN QUAN .....	4
2.1 Cơ sở lý thuyết.....	4
2.1.1 Đặc trưng màu sắc .....	4
2.1.2 Đặc trưng kết cấu.....	10
2.1.3 Đặc trưng hình dạng .....	16
2.1.4 Mô tả Boundary .....	16
2.1.5 Mô tả theo Vùng (Region).....	17
2.2 Độ đo tương đồng .....	19
2.2.1 Độ đo về màu sắc .....	19
2.2.2 Độ đo về kết cấu .....	20
2.2.3 Độ đo về hình dạng.....	21
2.3 Các nghiên cứu liên quan.....	24
2.3.1 Trong nước .....	24
2.3.2 Ngoài nước .....	24

CHƯƠNG 3: PHƯƠNG PHÁP ĐỀ XUẤT TÌM KIẾM ẢNH THEO NỘI DUNG	26
3.1 Yêu cầu bài toán	26
3.2. Mô hình nghiên cứu	28
3.2.1. Mô hình phối hợp các đặc trưng ảnh	28
3.2.2 Kết hợp toán tử LTP (Local Ternary Pattern)	29
3.3 Phương pháp rút trích đặc trưng ảnh và truy vấn ảnh	32
3.3.1 Phối hợp các đặc trưng để truy vấn ảnh	32
3.3.1.1 Xử lý cơ sở dữ liệu	32
3.3.1.2 Xử lý ảnh truy vấn	32
3.3.1.3 Đo sự tương tự giữa các véc tơ ảnh	32
3.3.1.4 Hiện thị kết quả trả về	33
3.3.2 Truy vấn ảnh dùng moments của LTP	33
3.3.2.1 Moments	33
3.3.2.2 Mẫu tam phân (LTP) và Moment	34
3.3.2.3 Phương pháp đề xuất	35
CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM	37
4.1 Tập dữ liệu thử nghiệm	38
4.2 Kết quả truy vấn	39
4.3 Code đặc trưng	45
CHƯƠNG 5: KẾT LUẬN	48
5.1 Kết quả đạt được	48
5.2 Ưu điểm và nhược điểm của giải thuật đề xuất	48
5.2.1 Ưu điểm của giải thuật đề xuất	48
5.2.2 Nhược điểm của giải thuật đề xuất	48
5.3 Đóng góp của luận văn	48
5.3.1 Đóng góp khoa học	49
5.3.2 Đóng góp thực tiễn	49
5.4 Hướng mở rộng	49
TÀI LIỆU THAM KHẢO	50

## DANH MỤC TỪ VIẾT TẮT

CBIR	: Content Based Image Retrieval
LBP	: Local Binary Pattern
LTP	: Local Ternary Pattern
RGB	: Red - Green – Blue
CMY	: Cyan-Magenta-Yellow
HSI	: Hue-Saturation-Intensity
HOG	: Histogram of oriented gradients
SVM	: Support Vector machine
MSVM	: Multiple Support Vector Machines

## **DANH MỤC CÁC BẢNG**

Bảng 4.1 Kết quả truy xuất ảnh của phương pháp đề xuất so với phương pháp khác.. 42

## **DANH MỤC CÁC BIỂU ĐỒ, ĐỒ THỊ, SƠ ĐỒ, HÌNH ẢNH**

Hình 2.1: Mô hình phân lớp của CBIR .....	4
Hình 3.1: Cấu trúc của hệ thống CBIR .....	26
Hình 3.2 Sơ đồ truy xuất hình ảnh từ cơ sở dữ liệu ảnh .....	35
Hình 4.1 Một số hình trong tập dataset Wang .....	39
Hình 4.2 Giao diện chính chương trình.....	39
Hình 4.3 Giao diện chọn dataset .....	40
Hình 4.4 Giao diện chọn thư mục chứa ảnh.....	40
Hình 4.5 Giao diện rút trích đặc trưng của ảnh.....	40
Hình 4.6 Giao diện chọn ảnh ban đầu .....	41
Hình 4.7 Kết quả truy xuất.....	42

## CHƯƠNG 1: GIỚI THIỆU

### 1.1 Giới thiệu đề tài

Cùng với sự phát triển của công nghệ thông tin và Internet đã làm gia tăng nhanh chóng dữ liệu đa phương tiện, khiến cho nhu cầu chia sẻ và tìm kiếm những loại tài nguyên này cũng tăng theo, trong đó có tài nguyên là dữ liệu kiểu hình ảnh. Tìm kiếm dữ liệu hình ảnh được sử dụng với nhiều mục đích khác nhau và được ứng dụng nhiều trong các lĩnh vực quan trọng của cuộc sống như trong hệ thống bảo mật, an ninh, y tế hay các hệ thống phát hiện chuyển động... Vì thế việc nghiên cứu và phát triển các hệ thống tìm kiếm ảnh ngày càng trở nên cấp thiết.

Có hai kiểu tìm kiếm đó là tìm kiếm theo từ khóa và tìm kiếm theo nội dung ảnh (CBIR-Content Based Image Retrieval), tìm kiếm theo từ khóa dễ thỏa mãn được nhu cầu người dùng với các nhu cầu tìm kiếm ảnh mới theo như mong muốn xuất hiện trong suy nghĩ của họ, tìm kiếm theo từ khóa thì nhanh hơn so với tìm kiếm theo nội dung bởi vì nó hoạt động dựa trên việc phân tích và so sánh các từ hoặc cụm từ tương ứng với nhau để đưa ra kết quả và không đòi hỏi người dùng phải có ảnh mẫu. Nhược điểm của phương pháp này là kết quả không chính xác. Một phương pháp khác để tìm kiếm hình ảnh là tìm kiếm ảnh theo nội dung của hình ảnh. Phương pháp này cần một ảnh mẫu cho đầu vào để tìm ra những bức ảnh tương ứng. Phương pháp này cho kết quả tốt hơn về tính đúng đắn bởi vì thông qua nội dung của bức ảnh sẽ được biểu diễn các đặc trưng trong bức ảnh và đưa ra những kết quả tương ứng với nội dung bức ảnh đầu vào. Nó đáp ứng tốt hơn cho người dùng, tuy nhiên người dùng sẽ cần phải có một ảnh mẫu để trích chọn và biểu diễn các đặc trưng trong bức ảnh đó trước khi tìm kiếm. Tìm kiếm ảnh theo nội dung là một phương pháp mới, hiệu quả để tìm ra những kết quả hình ảnh tương tự với thông tin được lấy trực tiếp từ nội dung của ảnh.

Nói chung, đối với hệ thống này người dùng sẽ cung cấp ảnh truy vấn và hệ thống sẽ trả về kết quả là tập các ảnh tương tự. Do đó, làm thế nào để mô tả, mô hình một hình ảnh để so sánh các ảnh khác nhau và đánh chỉ số cho các ảnh trong

cơ sở dữ liệu và để tìm kiếm ảnh một cách hiệu quả là một vấn đề hết sức quan trọng. Đây là bài toán đặt ra và đang được xã hội quan tâm.

Vì vậy, nghiên cứu phương pháp tìm kiếm ảnh dựa trên một số đặc trưng cho trước trong tập dữ liệu lớn được mạnh dạn đề xuất nghiên cứu trong đề tài này.

## **1.2 Mục tiêu của đề tài**

Nghiên cứu các kỹ thuật tìm kiếm ảnh theo nội dung và đề xuất mô hình tìm kiếm ảnh trong tập dữ liệu ảnh lớn dựa trên đặc trưng ảnh.

## **1.3 Nội dung đề tài**

Tìm kiếm ảnh theo một vài đặc điểm của ảnh trong tập dữ liệu lớn là một công đoạn quan trọng trong rất nhiều ứng dụng thị giác máy tính. Những hệ thống thuộc loại này có thể kể ra như là: tìm kiếm ảnh trên internet, tìm kiếm ảnh y khoa trong bệnh viện theo đặc điểm của bệnh nhân, hệ thống giao diện người dùng dựa vào cảm nhận, những căn phòng thông minh và những hệ thống thông minh hỗ trợ tài xế lái xe tự động, .... Nội dung chính của đề tài là nghiên cứu phương pháp nhằm tìm kiếm ảnh dựa trên một số đặc điểm được mô tả trước trong tập dữ liệu ảnh lớn.

## **1.4 Giới hạn của đề tài**

Nghiên cứu một số kỹ thuật tìm kiếm ảnh, đặc biệt là các kỹ thuật tìm kiếm ảnh dựa vào màu sắc, cấu trúc và nội dung của ảnh.

Tìm kiếm ảnh dựa trên các đặc trưng cơ bản như màu sắc, hình dạng cũng như cấu trúc.

## **1.5 Phương pháp nghiên cứu**

Phương pháp nghiên cứu được sử dụng: định tính, định lượng và so sánh kết quả với các giải thuật khác.

Tìm kiếm tài liệu liên quan trên Internet, các bài báo, kỷ yếu, các báo cáo hội thảo, hội nghị trong nước và trên thế giới có liên quan đề tài.

Nghiên cứu một số kỹ thuật tra cứu ảnh cần thiết và cài đặt chúng, đánh giá và so sánh với các kỹ thuật đang tồn tại.

Khảo sát yêu cầu bài toán, từ đó phân tích và thiết kế bài toán.

Trao đổi với các chuyên gia về lĩnh vực tìm kiếm ảnh theo nội dung ảnh.

Sử dụng công cụ Matlab để tiến hành hiện thực giải thuật đề thu thập kết quả.

### **1.6 Cấu trúc luận văn**

Luận văn được tổ chức thành 5 chương như sau:

Chương 1: Giới thiệu khái quát vấn đề được đề cập trong luận văn; trình bày mục tiêu, nội dung thực hiện, giới hạn của đề tài và phương pháp nghiên cứu.

Chương 2: Trình bày các cơ sở lý thuyết và các nghiên cứu liên quan trong và ngoài nước.

Chương 3: Trình bày chi tiết về phương pháp đề xuất trong luận văn.

Chương 4: Trình bày kết quả thực nghiệm, đánh giá hiệu quả của giải thuật đề xuất.

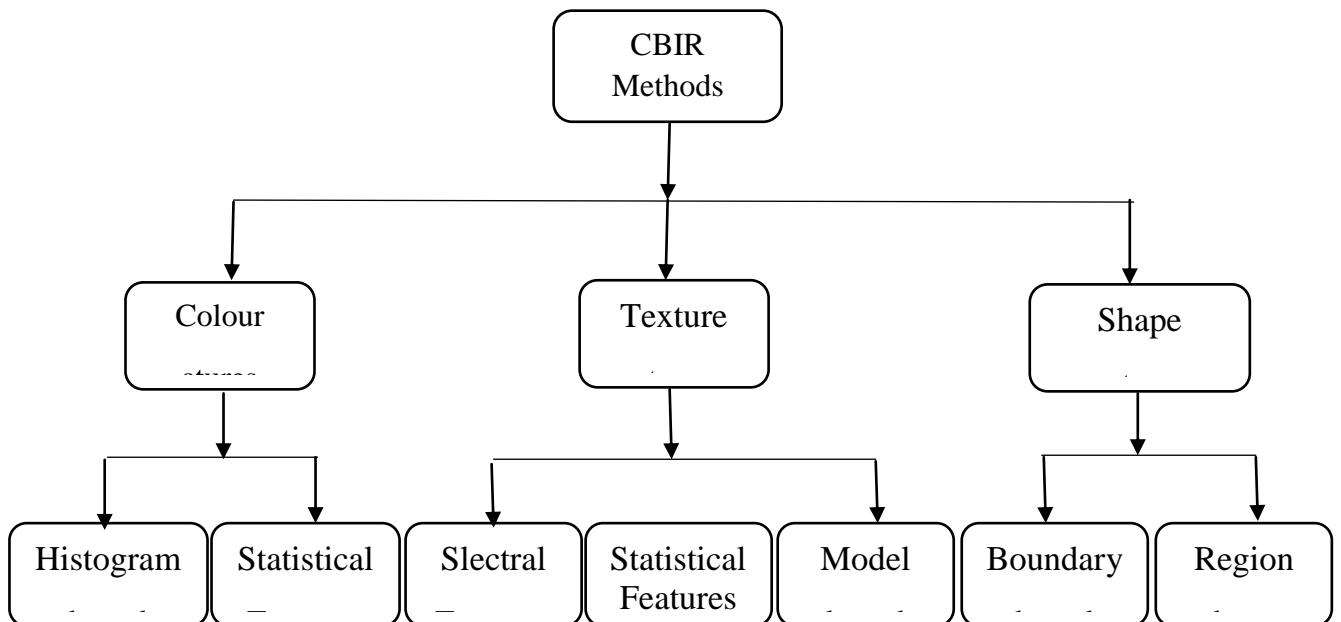
Chương 5: Dựa trên các kết quả đạt được để kết luận, trình bày ưu và nhược điểm của giải thuật đề xuất, hướng mở rộng cho đề tài.



## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VÀ CÁC NGHIÊN CỨU LIÊN QUAN

### 2.1 Cơ sở lý thuyết

Phương pháp tìm kiếm ảnh dựa trên nội dung có thể được phân thành các lớp phụ thuộc vào các đặc trưng như màu sắc, kết cấu và hình dạng như hình 2.1. Mỗi lớp đặc trưng được chia thành các lớp con tùy theo sử dụng thuật toán để xây dựng các véc tơ đặc trưng. Đặc trưng hình dạng được chia làm 02 đặc trưng: đặc trưng biên và đặc trưng vùng. Trong nghiên cứu, các nhà nghiên cứu phân loại các đặc trưng không gian của ảnh trong một lớp riêng biệt.



Hình 2.1: Mô hình phân lớp của CBIR

#### 2.1.1 Đặc trưng màu sắc

Đặc trưng màu sắc là một trong những đặc trưng quan trọng nhất của việc tìm kiếm tập ảnh màu của đối tượng nội dung bất kỳ. Màu sắc đóng vai trò rất quan trọng trong cơ chế nhận thức thị giác của con người.

Một ảnh màu điển hình được lấy từ camera số hoặc download từ Internet thường có ba kênh màu (ảnh xám chỉ có một kênh), những giá trị của dữ liệu ba

chiều này từ ảnh màu có thể cho ta biết vị trí của những điểm ảnh này trong không gian màu.

Những điểm ảnh có giá trị (1, 1, 1) cho những màu khác nhau trong những không gian màu khác nhau. Như vậy, mô tả đầy đủ của một ảnh màu điển hình gồm thông tin không gian hai chiều với điểm ảnh trong vùng không gian này và dữ liệu màu ba chiều với điểm ảnh màu trong không gian mà chúng ta đang đề cập. Ở đây giả thuyết không gian màu là cố định, bỏ qua thông tin không gian, thông tin màu trong ảnh có thể coi như là tín hiệu ba chiều đơn giản.

Nếu chúng ta coi thông tin màu của ảnh là tín hiệu một, hai, hoặc ba chiều đơn giản thì việc phân tích các tín hiệu sử dụng ước lượng mật độ xác suất là một cách dễ nhất để mô tả thông tin màu của ảnh.

Tất cả các phương pháp biểu diễn đặc trưng màu sắc của một ảnh có thể được phân thành hai nhóm: biểu đồ màu và phương pháp thống kê biểu đồ màu. Các không gian màu thường được sử dụng nhiều nhất như:

\* Không gian màu chuẩn RGB

Mắt người có thể phân biệt hàng ngàn màu sắc khác nhau, những con số chính xác hơn vẫn còn đang được bàn cãi nhiều. Ba màu RGB (Red – Green- Blue) mã hóa hệ thống đồ họa sử dụng ba byte ( $2^8$ ) hay khoảng chừng 16 triệu màu phân biệt. Máy tính có thể phân biệt bất kỳ màu gì sau khi được mã hóa, nhưng việc mã hóa có thể không trình bày được những sự khác biệt trong thế giới thực. Mỗi điểm ảnh RGB bao gồm một byte cho màu R, một byte cho màu G và một byte cho màu B. Việc mã hóa một màu tùy ý trong dãy hiển thị được làm bằng cách tổ hợp ba màu chính.

Ví dụ: Red(255,0,0), Green(0,255,0), Blue(0,0,255), Black(0,0,0). Hệ thống màu RGB là một hệ thống màu cộng vào bởi vì mỗi màu được tạo nên bằng cách cộng thêm các phần tử vào màu đen(0,0,0). Khuôn dạng của không gian màu RGB là định dạng phổ biến nhất của ảnh số, lý do chính là tính tương thích với màn hình hiển thị chính là màn hình vi tính. Tuy nhiên không gian màu RGB có hạn chế lớn

nhất là không phù hợp với cách con người cảm nhận về màu sắc. Do đó, không phù hợp cho việc ứng dụng vào tìm kiếm ảnh.

#### \* Hệ thống màu CMY

Hệ thống màu CMY theo mô hình in trên giấy trắng và theo khuôn mẫu trừ từ màu trắng thay vì thêm vào từ màu đen như hệ thống màu RGB. CMY là viết tắt của Cyan-Magenta-Yellow (màu lục lam, màu đỏ tươi, màu vàng), đó là ba màu chính tương ứng với ba màu mực in. Cyan hấp thụ sự chiếu sáng của màu đỏ, Magenta hấp thụ màu xanh lục, Yellow hấp thụ màu xanh dương. Do đó, tạo ra sự phản ánh tương ứng như khi in ảnh được chiếu sáng với ánh sáng trắng. Hệ thống dưới dạng âm tính vì mã hóa theo dạng hấp thụ màu. Có một số mã hóa như sau: trắng (0,0,0) vì không có ánh sáng trắng được hấp thụ, đen (255,255,255) vì tất cả các thành phần của màu trắng đều được hấp thụ.

Hệ thống màu CMY dường như là một sự đảo ngược của hệ thống màu RGB. Đặc tính của nó là sự đơn giản, ứng dụng nhiều trong thực tế. Tuy nhiên khuyết điểm của nó cũng tương tự như không gian màu RGB, tức là cách mã hóa khác với cách mà con người cảm nhận về màu sắc. Không thích hợp cho bài toán tìm kiếm ảnh dựa vào nội dung.

#### \* Hệ thống màu $L^*a^*b^*$ :

Mô hình  $L^*a^*b^*$  được đề cử bởi CIE cho việc lượng hóa sự khác biệt của màu sắc trong vật chiếu sáng của ánh sáng ban ngày. Tuy nhiên nó có một sự chuyển đổi được ghi vào để mà tính toán cho việc thích nghi với những nguồn sáng. Đây là hệ thống màu có sự tách riêng ánh sáng và màu sắc ra riêng. Do đó, cũng có khả năng lớn cho việc tìm kiếm dựa vào nội dung.

#### \* Hệ thống màu HSI: Hue-Saturation-Intensity

Hệ thống màu HSI mã hóa thông tin màu sắc bằng cách chia giá trị intensity I từ hai giá trị được mã hóa thuộc về độ hội tụ của màu- hue H và saturation S.

Thành phần không gian màu HSI gồm có ba phần: Hue được định nghĩa có giá trị  $0-2\pi$ , mang thông tin về màu sắc. Saturation có giá trị  $0-1$ , mang giá trị về độ thuần khiết của thành phần Hue. Intensity (Value) mang thông tin về độ sáng của

điểm ảnh. Ta có thể hình dung không gian màu HSI như là vật hình nón. Với trục chính biểu thị cường độ sáng Intensity. Khoảng cách đến trục biểu thị độ tập trung Saturation. Góc xung quanh trục biểu thị cho sắc màu Hue.

Đôi khi, hệ thống màu HSI được coi như là hệ thống màu HSV dùng Value thay vì Intensity. Hệ thống màu HSI thì thích hợp hơn với một số thiết kế đồ họa bởi vì nó cung cấp sự điều khiển trực tiếp đến ánh sáng và hue. Hệ thống màu HSI cũng hỗ trợ tốt hơn cho những thuật toán xử lý ảnh vì sự tiêu chuẩn hóa về ánh sáng và tập chung vào hai tham số về độ hội tụ màu và cường độ màu.

Hệ thống màu HSI có sự phân chia rõ rệt giữa ánh sáng và màu sắc. Do đó có khả năng rất lớn được áp dụng cho việc tính đặc trưng và so sánh sự giống nhau về màu sắc của hai ảnh. Do đó nó rất thích hợp cho việc tìm kiếm ảnh dựa vào màu.

Sự giống và khác nhau giữa hai ảnh về mặt màu sắc đối với mắt người chỉ mang ý nghĩa tương đối. Do đó khi áp dụng vào bài toán này trên máy tính thì ta cũng giả lập sự tương đối này.

Phương pháp chính của việc tìm kiếm theo màu sắc là dùng lược đồ màu để làm đặc trưng cho từng ảnh. Do những đặc điểm riêng của mô hình màu HSI và đặc trưng của việc tìm kiếm nên tính lược đồ màu cũng được dùng một mô hình rất đặc biệt để phù hợp cho những đặc điểm riêng này.

Cách đơn giản và sử dụng thường xuyên nhất để biểu diễn cho màu sắc là biểu đồ màu. Biểu đồ màu là một bảng tóm tắt thông tin về màu sắc cho một ảnh màu bất kỳ. Đối với mỗi điểm của không gian màu sắc, số lượng điểm ảnh của một màu được tính toán nhất định. Các thông tin về màu sắc thường đơn giản và tự nhiên. Tuy nhiên, nó có nhược điểm là khoảng cách độ tương đồng giữa hai hình ảnh có màu sắc giống nhau là lớn. Ngoài ra, biểu đồ như vậy là rất thưa thớt và do đó dễ bị nhiễu.

Stricker và Orengo đề xuất sử dụng biểu đồ màu tích lũy [1]. Một biểu diễn màu sắc như vậy là ít nhiễu và cũng làm giảm số lượng các lỗi Type II nếu các phần tử liền kề của biểu đồ tương ứng với màu sắc tương tự. Một phương pháp tiếp cận khác để tính độ tương đồng của màu sắc được trình bày trong [2]. Trong tác phẩm này, số

liệu khác dựa trên không gian véc tơ của màu sắc (biểu đồ) được đề xuất. Các biểu đồ màu tự nó không lưu trữ thông tin về cách bố trí không gian của màu sắc vào hình ảnh. Một giải pháp cho vấn đề này đã được đề xuất trong [3]. Sau khi xây dựng một biểu đồ màu nơi mà chỉ có màu chính của một hình ảnh được xem xét cho mọi phần tử khác không của biểu đồ, tọa độ của khối tâm của vùng màu tương ứng được tính toán. Thông tin này được sử dụng để đo độ tương đồng giữa các hình ảnh với số lượng điểm ảnh thuộc vùng màu này. Giải pháp này có thể thực hiện, trong một nghĩa nào đó, phải được xem xét bố trí không gian của màu sắc, nhưng nó có một bất lợi đáng kể. Nếu hình ảnh có chứa một số thành phần hợp nhất của cùng một màu sắc, thực tế này sẽ không được phản ánh trong các véc tơ đặc trưng của ảnh. Thay vào đó, một khối tâm chung cho tất cả các thành phần sẽ được tính toán.

Một sự thay đổi của mô hình này đã được đề xuất bởi Stricker và các cộng sự trong [4]: kênh phân phối các màu riêng biệt được coi là một phần của một phân phối ba chiều chứ không phải là phân phối độc lập. Cho véc tơ đặc trưng, giá trị trung bình cho mỗi kênh màu và ma trận đồng nhất của kênh phân phối được sử dụng. Để lấy đồ họa và hình ảnh cùng lúc, công việc này được áp dụng một phương pháp truy hồi thích nghi [5]. Các phương pháp được đề xuất sử dụng Histogram of oriented gradients (HOG) như các đặc trưng dựa trên điểm ảnh. Tuy nhiên, đặc điểm của đồ họa và hình ảnh là khác nhau, điều này ảnh hưởng đến trích xuất đặc trưng và độ chính xác truy xuất. Như vậy, một phương pháp thích hợp được đề nghị chọn đặc trưng khác nhau dựa trên HOG cho truy xuất đồ họa và hình ảnh. Trong [6], một phương pháp trích chọn đặc trưng màu sắc và kết cấu của hình ảnh một cách nhanh chóng cho truy xuất ảnh dựa trên nội dung (CBIR) được đề xuất. Đầu tiên, không gian màu HSV được định lượng một cách hợp lý. Biểu đồ màu và đặc trưng kết cấu dựa trên ma trận đồng nhất được trích xuất từ véc tơ đặc trưng. Sau đó, các đặc điểm của biểu đồ màu sắc toàn cục, đặc trưng kết cấu biểu đồ màu cục bộ được so sánh và phân tích cho CBIR. Dựa trên những sản phẩm này, một hệ thống CBIR được thiết kế sử dụng đặc trưng màu sắc và kết cấu hợp nhất bằng cách xây dựng trọng số của véc tơ đặc trưng.

Trong [7], các đặc trưng như hình dạng và kết cấu được so sánh bằng khoảng cách Euclide giữa ảnh truy vấn và ảnh tham chiếu. Các thao tác hình thái với cấu trúc phân tử ảnh bất biến không gian được sử dụng cho trích xuất đặc trưng. Sau khi quá trình trích xuất đặc trưng, các véc tơ đặc trưng được tính toán bằng cách áp dụng phương pháp mã hóa xử lý khối (BTC) trong trích xuất đặc trưng hình ảnh. Nó cải thiện hiệu suất của truy xuất hình ảnh và giảm độ phức tạp tính toán để thực hiện truy vấn. Dựa trên mô hình màu HSV, một phương pháp dựa trên đối tượng đặc trưng không gian màu (OSCF) cho truy xuất màu của ảnh được đề xuất trong [8]. Thứ nhất, đối tượng này được trích xuất từ màu sắc, sau đó là các đặc trưng hình ảnh biểu diễn bởi các đối tượng trong nó. Màu sắc và đặc trưng không gian màu được áp dụng để mô tả các đối tượng. Phương pháp mới chỉ quan tâm đến các đối tượng chính trung tâm. Trong [9], tác giả đề xuất một cách tiếp cận mờ mới để phân loại màu sắc hình ảnh dựa trên nội dung của nó, để đặt ra một truy vấn về ngôn ngữ tự nhiên và kết hợp các truy vấn dựa trên mạng lưới thần kinh để truy xuất nhanh chóng và hiệu quả.

Một phương pháp truy xuất dựa trên hình ảnh màu mới được đề xuất trong [10]. Độ chính xác của thuật toán này là cao hơn so với phương pháp học có giám sát và hiệu quả của nó cũng hơn cách học có giám sát. Thứ nhất, thông qua các ma trận khoảng cách của màu sắc, ảnh mẫu được phân cụm lại theo cách tự tổ chức, do đó bảng màu của nó được xây dựng. Dựa trên bảng màu này, hình ảnh khác trong cơ sở dữ liệu được ánh xạ về khoảng cách nhỏ nhất. Bằng cách này, một biểu đồ chuẩn theo bảng màu giống nhau có thể được thu được cho mỗi hình ảnh trong cơ sở dữ liệu. Bên cạnh đó, thuật toán này cũng kết hợp các vùng màu chính để đại diện cho phân bố không gian của màu sắc.

Trong [11], tác giả đã thảo luận về các phương pháp so sánh được sử dụng trong biểu đồ màu sắc dựa trên hai phương pháp chính được sử dụng thường xuyên trong CBIR: biểu đồ màu sắc chuẩn bằng cách sử dụng GLCM và biểu đồ màu sắc sử dụng KMeans. Sử dụng khoảng cách Euclide để tính toán sự giống nhau giữa hình ảnh truy vấn và các hình ảnh ứng viên. Kết quả thí nghiệm cho

thấy rằng biểu đồ màu sắc với phương pháp K-Means có độ chính xác cao và chính xác so với GLCM.

Trong [12], một phương pháp được đề xuất để truy xuất ảnh nhị phân, đối với phương pháp này thì ảnh đen và trắng được biểu diễn bởi một tên đặc trưng mới cho biểu đồ phân cấp mật độ thích nghi, trong đó khai thác sự phân bố của các điểm ảnh trên một miền hai chiều. Một loại biểu đồ mới trong đó kết hợp chỉ có các thông tin thị giác ngoại vi cạnh của các hình ảnh được giới thiệu trong [13]. Các thao tác trích xuất cạnh được thực hiện với việc sử dụng điều khiển bao quanh trung tâm của hệ thống thị giác con người. Các biểu đồ màu quanh trung tâm (Center-Surround Histogram - CSH) được đề nghị có hai ưu điểm so với biểu đồ kinh điển. Đầu tiên, nó làm giảm số lượng thông tin trực quan cần phải được xử lý lần hai, nó kết hợp một mức độ của thông tin không gian khi được sử dụng trong các ứng dụng truy xuất hình ảnh dựa trên nội dung.

Một truy xuất hình ảnh mạnh mẽ dựa trên biểu đồ màu sắc của đặc trưng vùng cục bộ (LFR) được trình bày trong [14]. Thứ nhất, điểm đặc trưng ảnh ổn định được trích xuất bằng cách sử dụng bộ điều chỉnh đa tỷ lệ Harris-Laplace. Sau đó, các vùng đặc trưng chữ ký cục bộ được xác định thích nghi theo tỉ lệ đặc trưng lý thuyết. Cuối cùng, các biểu đồ màu sắc của vùng đặc trưng cục bộ được cấu thành, sự giống nhau giữa các hình ảnh màu được tính bằng cách sử dụng các biểu đồ màu của LFRs. Một hệ thống CBIR mới được đề xuất trong [15] tên là iSearch và đối sánh toàn cục/cục bộ của các đặc trưng cục bộ được kết hợp để truy xuất hình chính xác. Nhiều đặc trưng cục bộ được trích xuất bao gồm phép biến đổi đặc trưng bất biến (SIFT), moment vùng màu và đối tượng phân mảnh đường viền đủ để biểu diễn xuất hiện thị giác của các phần tử; trong khi đối sánh toàn cục và cục bộ của hình ảnh trong bộ dữ liệu là lớn. Để làm điều này, một mã hóa phân mảnh đường viền hiệu quả và phương pháp lập chỉ mục được phát triển.

### **2.1.2 Đặc trưng kết cấu**

Kết cấu là một khái niệm trực quan, là một thành phần chủ yếu về nhận thức thuộc tri giác của con người. Giống như màu sắc, kết cấu trở thành đặc điểm cần

thiết để xem xét khi truy vấn cơ sở dữ liệu ảnh. Mọi người đều có thể nhận thấy kết cấu tuy nhiên nó rất khó để xác định, điều này được thể hiện bằng một số khái niệm khác nhau về kết cấu. Mặc dù không có một khái niệm chung cho kết cấu nhưng tất cả các nhà nghiên cứu đều tập trung thống nhất trên hai điểm chính:

- Trong phạm vi một kết cấu có sự biến đổi đáng kể về mức độ cường độ giữa các điểm ảnh liên kề, đó là giới hạn của độ phân giải, không có sự đồng nhất.

- Kết cấu là thuộc tính đồng nhất ở một vài không gian lớn hơn độ phân giải của ảnh, hàm ý trong những thuộc tính này của cấu trúc là ảnh có độ phân giải nhất định.

Khác với màu sắc, kết cấu diễn ra trên cả một vùng hơn là tại một điểm, nó thường được định nghĩa bằng những mức xám được hiểu như là màu sắc. Một số nhà nghiên cứu giải quyết bài toán về xác định kết cấu bằng cách mô tả nó trong những thuật ngữ của hệ thống thị giác của con người như hướng, độ thô, độ tương phản.... Một số nhà nghiên cứu khác lại định nghĩa kết cấu bởi những ứng dụng. Điều này tạo cho kết cấu đa dạng và cho nhiều cách để trích chọn kết cấu.

Định nghĩa của kết cấu dựa trên nhận thức của con người là phù hợp cho nghiên cứu và cho bàn luận về nét tự nhiên của kết cấu. Mặc dù vậy một định nghĩa đưa ra những vấn đề khi được sử dụng như là học thuyết cơ bản cho thuật toán phân tích kết cấu.

Kết cấu cho chúng ta thông tin về sự sắp xếp cấu trúc của bề mặt và các đối tượng trên hình ảnh. Kết cấu không được định nghĩa cho một điểm ảnh riêng biệt; nó phụ thuộc vào sự phân bố cường độ lên ảnh. Kết cấu sở hữu tính chu kỳ và tính khả năng mở rộng; nó có thể được mô tả bởi hướng cơ bản, độ tương phản, và độ sắc nét. Phân tích kết cấu đóng vai trò quan trọng trong việc so sánh các hình ảnh bổ sung các đặc trưng màu sắc. Thống kê các đặc trưng thường được sử dụng nhất bao gồm:

- Các thông số tổng hợp thống kê tính toán từ các giá trị cường độ điểm ảnh.
- Thông số tính toán dựa trên các ma trận đồng xuất hiện.
- Biểu đồ kết cấu xây dựng dựa trên các đặc trưng Tamura.



Một trong những phương pháp đầu tiên để biểu diễn cho các đặc trưng kết cấu của hình ảnh là các ma trận đồng nhất xám (GLCM) đề xuất bởi Haralick và các cộng sự. [16]. Các tác giả đề xuất 14 mô tả, bao gồm hai moment gốc, tương phản (bất biến, moment khác), tương quan và những vấn đề khác. Mỗi mô tả biểu diễn cho một thuộc tính kết cấu. Vì thế, nhiều công trình ví dụ như mô tả trong [17], được dành cho việc lựa chọn thống kê mô tả có nguồn gốc từ các ma trận đồng nhất, mô tả kết cấu một cách tốt nhất. Trong [18], trước hết, chuyển không gian màu từ mô hình RGB đến mô hình HSI, sau đó trích xuất biểu đồ màu tạo thành véc tơ đặc trưng màu. Thứ hai, trích xuất đặc trưng kết cấu bằng cách sử dụng ma trận đồng nhất xám. Thứ ba, áp dụng moment Zernike để trích xuất các đặc trưng hình dạng. Cuối cùng, kết hợp màu sắc, kết cấu và đặc trưng hình dạng để tạo thành các véc tơ đặc trưng hợp nhất của toàn bộ hình ảnh. Các thực nghiệm trên được sử dụng cho bộ dữ liệu hình ảnh cho thấy kế hoạch đề xuất đạt được một hiệu suất rất tốt về độ chính xác truy xuất so với các phương pháp khác.

Một phương pháp được đề xuất [19] cho truy xuất hình ảnh hiệu quả áp dụng một sự kết hợp có trọng số của màu sắc và kết cấu cho các biến đổi wavelet, dựa trên không gian màu và thứ tự thống kê tương ứng. Bộ mô tả đề xuất là đặc biệt hữu dụng cho việc tìm kiếm hình ảnh đa độ phân giải.

#### *Mô tả kết cấu dựa trên Wavelet*

Trong mô tả kết cấu dựa trên wavelet, một đặc trưng cụ thể của phương pháp này là biểu diễn và phân tích các tín hiệu với tỷ lệ khác nhau, tức là, theo độ phân giải khác nhau. Các hình ảnh được mô tả bởi một cấu trúc phân cấp, mỗi cấp biểu diễn cho các tín hiệu gốc với một mức độ chi tiết.

Smith và Chang sử dụng đặc điểm thống kê (trung bình và bất biến) tính toán cho mỗi dải con như là đặc trưng kết cấu [20]. Họ đã so sánh hiệu quả của phân loại kết cấu cho các đặc trưng được xây dựng bằng cách tiếp cận phương pháp wavelet, phân tích đồng nhất vào dải con (mà không mở rộng quy mô, mỗi dải con có chứa một phần của một tín hiệu của tần số nhất định), biến đổi rời rạc cosin, phân tích không gian. Trong [21], hệ số phân phối độ lệch trung bình và tiêu chuẩn của biến

đổi wavelet được sử dụng để xây dựng các véc tơ đặc trưng. Trong trường hợp chuyển đổi với các bộ lọc  $N$ , hình ảnh được biểu diễn bởi một véc tơ đặc trưng của  $2N$  chiều.

Trong [22], tác giả tính toán một đặc trưng kết cấu mới bằng cách áp dụng hàm mật độ Gauss chung đến sự phân bố của hệ số Curvelet được gọi là đặc trưng kết cấu Curvelet GGD. Mục đích là để nghiên cứu đặc trưng kết cấu curvelet GGD và so sánh hiệu suất truy xuất của Curvelet, Wavelet và đặc trưng kết cấu Wavelet GGD. Kết quả thí nghiệm cho thấy rằng cả hai curvelet và các đặc trưng curvelet GGD thực hiện tốt hơn đáng kể so với đặc trưng kết cấu wavelet và Wavelet GGD. Trong số hai loại đặc trưng dựa trên curvelet, trong CBIR đặc trưng curvelet có hiệu suất làm tốt hơn hơn đặc trưng kết cấu curvelet GGD. Công việc bao gồm việc giảm thiểu các đặc trưng cấp thấp mô tả một hình ảnh bằng cách sử dụng một mô tả giảm kết hợp thông tin màu sắc và kết cấu đó là chuyển đổi wavelet được khám phá trong [23]. Một phương pháp được đề xuất để mô tả ảnh bởi các dải con tần số cao là biến đổi Wavelet rời rạc (DWT) liên quan đến các khu vực trọng số nổi bật sau một bước phân đoạn mờ.

Trong [24], một chữ ký hình ảnh đơn giản dựa trên những moment chuẩn của phân bố hệ số Wavelet được đề xuất. Chữ ký này có thể được tính cho mỗi thể wavelet lọc nhanh. Do đó một bản đồ chữ ký hình ảnh thu được được sử dụng như là một đặc tính hình ảnh cho tìm kiếm hình ảnh dựa trên nội dung (CBIR). Việc biểu diễn phép biến đổi Curvelet (MCT) và sự kết hợp với cây từ vựng (VT) cho bộ sưu tập đặc trưng và truy xuất các hình ảnh từ cơ sở dữ liệu [25]. MCT đã được thực hiện bằng cách sử dụng dải con Gabor wavelet. Các thuật toán đề xuất bắt thông tin cạnh trong một hình ảnh chính xác hơn Gabor transform (GT) và chuyển đổi Curvelet trong đó sử dụng thuật giải a trous cho biến đổi Wavelet (ACT) cho phân tích của một ảnh.

Trong đề xuất [26], một kỹ thuật tìm kiếm thông minh dựa trên meta-heuristic, hybrid, được gọi là hỗn hợp Thuật toán tìm kiếm hấp dẫn (MGSA), được sử dụng. Một số thông số trích xuất đặc trưng được tối ưu hóa để đạt được độ chính xác tối

đa của hệ thống CBIR. Một hệ thống CBIR cực kỳ nhanh chóng trong đó sử dụng tập hợp Multiple Support Vector Machines (MSVM) được đề xuất trong [27]. Tác giả sử dụng phép biến đổi wavelet Daubechies để trích xuất véc tơ đặc trưng của hình ảnh. Trong [28], một cơ sở wavelet khác được sử dụng để mô tả từng ảnh truy vấn. Một hàm hồi quy được sử dụng để điều chỉnh hiệu suất tối đa truy xuất trong các tập dữ liệu huấn luyện, nó sử dụng để ước tính tốt nhất bộ lọc wavelet về mặt hiệu suất truy xuất dự kiến cho mỗi hình ảnh truy vấn.

Tamura và các cộng sự. [29] trình bày một cách tiếp cận để mô tả kết cấu trên cơ sở nhận thức về thị giác con người. Họ đề nghị sáu tham số: *thô*, *độ tương phản*, *hướng từ*, *dòng giống*, *đều đặn*, và *độ nhám* tương ứng với sáu thuộc tính kết cấu đã được công nhận là ý nghĩa thị giác trong quá trình thí nghiệm tâm lý. Howarth và Ruger [30] - [31] nhận thấy rằng các thông số mô tả ba đặc tính đầu tiên: tính thô, độ tương phản và tính định hướng có hiệu quả hơn trong phân loại và tìm kiếm hình ảnh bởi kết cấu. Tập hợp của tất cả các điểm đó của một hình ảnh được đưa ra như ảnh *Tamura*.

Phân tích kết cấu bằng phương pháp của các bộ lọc Gabor là một trường hợp đặc biệt của phương pháp wavelet và được sử dụng thường xuyên trong truy xuất ảnh bởi kết cấu. Trong hầu hết các hệ thống CBIR dựa trên Gabor wavelet [32] - [34], hệ số phân phối giá trị trung bình và độ lệch chuẩn của biến đổi wavelet được sử dụng để tạo các véc tơ đặc trưng.

Trong trường hợp chuyển đổi với các bộ lọc  $N$ , hình ảnh được biểu diễn bởi một véc tơ đặc trưng  $2N$  chiều. Ý tưởng khác của hệ thống CBIR được dựa trên việc mở rộng các hình ảnh trong điều kiện cơ sở thu được bằng cách phân tích một tập huấn luyện của ảnh. Ví dụ là bộ lọc thông dụng ICA bằng cách áp dụng các thành phần phân tích độc lập của tập huấn luyện. Cách các bộ lọc ICA được xây dựng tương tự như quá trình huấn luyện của hệ thống thị giác con người. Các bộ lọc ICA thu được bằng phân tích thành phần cạnh biên cục bộ độc lập và tương tự như các bộ lọc Gabor. Không giống như cách thứ hai, các bộ lọc ICA được tạo tự nhiên

và phản ánh tính chất kết cấu chính của ảnh đã được sử dụng để có được chúng. Việc xây dựng các bộ lọc ICA có thể được tìm thấy trong [35-38].

Ba đặc trưng ảnh được đề xuất để truy xuất ảnh trong [39]. Các đặc trưng ảnh đầu tiên và thứ hai là dựa trên đặc trưng màu và kết cấu, lần lượt gọi là ma trận đồng nhất màu (CCM) và sự khác biệt giữa các điểm ảnh của mẫu ảnh (DBPSP). Các đặc trưng ảnh thứ ba là dựa trên phân phối màu sắc, được gọi là biểu đồ màu K-mean (CHKM). Trong [40], đầu tiên không gian màu HSV được định lượng hợp lý, Biểu đồ màu và đặc trưng kết cấu dựa trên ma trận đồng nhất được trích xuất từ véc tơ đặc trưng. Sau đó, các đặc điểm của biểu đồ màu cục bộ, biểu đồ màu cục bộ và đặc trưng kết cấu được so sánh và phân tích cho CBIR. Dựa trên kết quả công việc này, một hệ thống CBIR được thiết kế sử dụng màu sắc và kết cấu hợp nhất các đặc trưng bằng cách tạo trọng số của các véc tơ đặc trưng.

Trong [41], một phương pháp truy vấn ảnh dựa trên nội dung được đề xuất dựa trên hiệu quả tích hợp của đặc trưng màu sắc và kết cấu.

Một sơ đồ đặc trưng mới sử dụng biểu đồ *Gabor wavelet* nâng cao (EGWC) được đưa ra lập chỉ mục và truy xuất ảnh trong [42]. EGWC sử dụng Gaborwavelets để phân tích các hình ảnh theo tỉ lệ định hướng khác nhau. Các hệ số Gabor wavelet được lượng tử hóa bằng cách tối ưu ngưỡng lượng tử. Trong bước tiếp theo, biểu đồ màu của hệ số lượng tử hóa wavelet được tính trong từng tỷ lệ và định hướng wavelet. Một phương pháp mới được đề xuất trong đó sử dụng một thuật toán phân cụm k-means nổi tiếng và một cơ sở dữ liệu chỉ mục cấu trúc B ± cây để tạo điều kiện lấy hình ảnh có liên quan một cách hiệu quả [43].

Trong [44], một đặc trưng bất biến xoay curvelet đại diện cho kết cấu được đề xuất làm tốt hơn và được sử dụng rộng rãi so với đặc trưng kết cấu Gabor. Một phương pháp mới cũng được đề nghị áp dụng phép biến đổi curvelet để phân vùng ảnh truy xuất. Một phương pháp được đề xuất gần đây nhất của hệ thống CBIR là sử dụng nhiều máy véc tơ hỗ trợ (Multiple Support Vector Machines) [45]. Phép biến đổi wavelet Daubechies wavelet trích xuất các véc tơ đặc trưng của ảnh. Truy vấn ảnh dựa trên nội dung (CBIR) để hỗ trợ chẩn đoán trong lĩnh vực y học được

trình bày trong [46]. Trong hệ thống đề xuất, hình ảnh được lập chỉ mục trong lĩnh vực thời trang mà không cần trích xuất đặc trưng miền: chữ ký được xây dựng cho mỗi ảnh từ phép biến đổi wavelet của nó. Những đặc điểm chữ ký của ảnh được phân phối hiệu quả qua phép biến đổi wavelet. Phép đo khoảng cách được định nghĩa để so sánh hai chữ ký hình ảnh và lấy ảnh truy xuất tương tự trong cơ sở dữ liệu khi một hình ảnh truy vấn được gửi bởi một bác sĩ.

### 2.1.3 Đặc trưng hình dạng

Cùng với đặc điểm màu sắc và kết cấu, hình dạng của các đối tượng (hình ảnh tương phản) cũng thường được sử dụng để so sánh hình ảnh. Phương pháp biểu diễn và mô tả hình dạng có thể được chia thành hai nhóm: các phương pháp bên ngoài, nó biểu diễn cho vùng trong giới hạn của đặc tính ngoài (ranh giới của nó) và những đặc tính nội bộ, trong đó biểu diễn cho vùng trong những điều khoản của nó đặc điểm nội bộ (các điểm ảnh bao gồm vùng). Các đặc trưng hình dạng được chia thành hai loại: mô tả đường biên và mô tả vùng. Hơn nữa chúng được phân loại như cấu trúc (a) và toàn cục (b). Mô tả đường biên toàn cục bao gồm chữ ký khác, mô tả Fourier và mô tả wavelet.

### 2.1.4 Mô tả Boundary

- *Các mã chuỗi*: Nó mô tả một đối tượng đường biên như là một chuỗi các phân đoạn với một hướng nhất định. Để xây dựng một mã chuỗi, các hình ảnh được xếp chồng với một lưới, và các điểm biên xấp xỉ bởi các nút lưới gần nhất. Các đường phân đoạn kết nối các nút lân cận.

- *Chữ ký*: Chữ ký là biểu diễn của một đối tượng đường biên hai chiều bằng hàm của một biến, được giả định là dễ dàng hơn để mô tả sự so với các đường biên hai chiều nguyên bản.

- *Mô tả Fourier*: Các mô tả Fourier là một trong những phương pháp phổ biến nhất của đường viền tham số hóa. Các Ý tưởng cơ bản của phương pháp này bao gồm trong ứng dụng của Fourier rời rạc vào chữ ký và sử dụng Hệ số Fourier thu được như tham số mô tả các đường viền.

Trong [15], một phương pháp mới được đề xuất có tên là iSearch và đối sánh toàn cục và cục bộ của các đặc trưng cục bộ được kết hợp để truy xuất chính xác hình ảnh kiểu tương tác. Đầu tiên tác giả trích xuất nhiều đặc trưng cục bộ bao gồm đặc trưng chuyên đổi bất biến (SIFT), những moment vùng màu và các phân mảnh đường viền đối tượng để biểu diễn hiệu quả cho sự xuất hiện trực quan các mục; khi kết hợp đối sánh toàn cục và cục bộ của hình ảnh bộ dữ liệu quy mô lớn được phép. Để nâng cao thuật toán SIFT, một cách tiếp cận mạnh mẽ được đề xuất để truy xuất hình ảnh dựa trên sự kết hợp của điểm hấp dẫn và thông tin cạnh trong [47]. Cách tiếp cận này là mạnh mẽ để dịch, luân chuyển và giữ một phần của đối tượng.

Một mô tả đặc trưng biểu đồ màu A trous wavelet cho biểu diễn hình ảnh được sử dụng trong [48]. Bằng cách mở rộng hơn nữa trong mô tả này, mô tả cấu trúc á trous Gradient (AGSD) được đề xuất cho truy xuất ảnh dựa trên nội dung. Đặc trưng đơn giản hóa AGSD được tính toán với a trous wavelet một cách cục bộ. Các thông tin cục bộ của hình ảnh được trích xuất thông qua mô tả cấu trúc vi mô (MSD); nó tìm thấy mối quan hệ giữa pixel lân cận. Trong [49], một sơ đồ đặc trưng mới gọi biểu đồ màu Gabor wavelet nâng cao (EGWC) được đề xuất cho chỉ mục và truy xuất hình ảnh. EGWC sử dụng Gabor wavelet để phân tích các hình ảnh trong tỷ lệ và định hướng khác nhau. Các hệ số wavelet Gabor sau đó được lượng tử hóa bằng cách sử dụng các tối ưu hóa ngưỡng lượng tử. Trong tiếp theo bước, biểu đồ tương quan của các hệ số wavelet được lượng tử hóa trong từng tỷ lệ và định hướng wavelet.

### **2.1.5 Mô tả theo Vùng (Region)**

Vùng có thể được mô tả bằng các thông số hình học đơn giản, chẳng hạn như một khu vực.

*- Phương pháp dựa trên Lưới:*

Sajjanhar và Lu [50] đề xuất một phương pháp trực quan rõ ràng để mô tả hình dạng của đối tượng, gọi là phương pháp dựa trên lưới. Ý tưởng cơ bản của phương pháp được đề xuất có thể được thể hiện qua hai bước: một lưới với các tế

bào của kích thước nhất định được đặt lên trên các đối tượng, và các tế bào của lưới điện nằm ngang từ cánh phải tới bên trái và từ trên xuống dưới.

- *Moments và bất biến của nó:*

Moment bất biến đang được sử dụng phổ biến nhất và mô tả vùng được sử dụng rộng rãi nhất. Ý tưởng sử dụng những Moment để mô tả hình dạng lần đầu tiên được đưa ra bởi Hu vào năm 1962 [51]. Tác giả coi những moment hình học của hàm có hai biến.

Trong [52], Luren và Fritz đưa ra một phương pháp nhanh nhất để tính moment cho ảnh nhị phân dựa trên việc sử dụng bất biến mật mã của định lý Green. Trong [53], bất biến khác được suy ra từ những *Moment* hình học. Ngoài Moment hình học (chúng được đưa ra đôi khi là số ít hoặc số nhiều), những Moment khác cũng được sử dụng. Các Mô tả Fourier (GFD) được đề xuất bởi Zhang và Lu [54], như mô tả dựa trên moment khác, dựa vào ý tưởng mở rộng theo quan điểm cơ sở nhất định.

Một phương pháp mới biểu diễn đặc trưng hình dạng dựa trên cạnh và xử lý hình thái học để truy xuất ảnh hiệu quả được trình bày trong [55]. Moment Pseudo Zernike dựa trên đặc trưng hình dạng toàn cục, nó là bất biến để biến đổi hình học cơ bản, được trích xuất và được sử dụng cho truy xuất hình ảnh tương tự với khoảng cách Canberra metric. Hệ thống CBIR được biểu diễn bằng cách sử dụng mô tả đặc trưng hình dạng và hiệu chỉnh moment Zernike dựa trên các moment Zernike với lỗi hình học tối thiểu và lỗi tích hợp số [56]. Trong [57], phân tích thí nghiệm của mô tả dày đặc điểm ảnh dựa trên như mô hình nhị phân cục bộ (LBP), mô hình tam phân cục bộ (LTP) và các biến thể của nó được thực hiện. Những mô tả này được sử dụng như là các đặc trưng cục bộ cùng với ZMS đặc trưng toàn cục đạt được tỷ lệ truy xuất cao hơn và chính xác trong hệ thống SBIR. Một phương pháp mới được đề xuất cho truy xuất ảnh theo nội dung dựa trên các điểm quan tâm [58]. Điểm quan tâm được phát hiện từ tỷ lệ hướng quay bình thường của ảnh.

Sau đó, các hình ảnh bình thường được chia thành một loạt các vùng khu vực con với các vùng khác theo sự phân bố điểm quan tâm.

## 2.2 Độ đo tương đồng

### 2.2.1 Độ đo về màu sắc

Bước quan trọng của quá trình tìm kiếm dữ liệu ảnh dựa vào nội dung là xác định độ trùng khớp của hai lược đồ màu. Do đó, phát sinh ra một giá trị để biểu thị cho sự trùng khớp này, có nhiều cách để tính giá trị này. Ta gọi những giá trị được tính từ những cách khác nhau này là các loại độ đo màu. Một cách đơn giản, độ đo màu là được coi một giá trị để biểu thị cho độ so khớp sự trùng khớp của hai lược đồ màu. Tùy theo từng trường hợp, từng loại độ đo màu giá trị này có thể âm hoặc dương lớn hoặc nhỏ tương ứng với mức độ giống nhau như thế nào của các loại lược đồ màu.

Mỗi loại độ đo màu có những ưu và khuyết điểm riêng, trong từng trường hợp cụ thể. Gọi  $h(I)$  và  $h(M)$  tương ứng là 2 lược đồ màu của hai ảnh  $I$  và ảnh  $M$ . Khi đó các loại độ đo màu được định nghĩa là một số nguyên (hoặc số thực) theo các loại độ đo tương ứng như sau:

#### ➤ Độ đo khoảng cách min - max

Được thực hiện dựa trên ý tưởng lấy phần giao của của hai lược đồ cần so sánh, ta sẽ được một lược đồ, tính tổng các giá trị có được từ lược đồ này cho ta được độ đo min-max.

Đối với độ đo min: ta tính dựa vào giá trị min tại mỗi  $K$  bin.

$$\text{Intersection}(h(I),h(M)) = \sum_{j=1}^K \min\{h(I)[j],h(M)[j]\}$$

Đối với độ đo max: ta tính dựa vào giá trị max tại mỗi  $K$  bin.

$$\text{Intersection}(h(I),h(M)) = \sum_{j=1}^K \max\{h(I)[j],h(M)[j]\}$$

$$\text{Matching}(h(I),h(M)) = \frac{\text{Intersection}(h(I),h(M))}{\max(\sum_i h(I)[i], \sum_i h(M)[i])}$$

#### ➤ Độ đo khoảng cách euclide

Đây là cách tính khoảng cách euclide thông thường giữa các  $K$  bin:



$$\text{Intersection}(h(I),h(M))=\sum_{j=1}^K \sqrt{(h(I)-h(M))^2}$$

Hoặc có thể là:

$$\text{Intersection}(h(I),h(M))=\sum_{j=1}^K \sqrt{(h(I)-h(M))^2}$$

c. Độ đo khoảng cách toàn phương

$$\text{Intersection}(h(I),h(M))=\sum_{i=1}^K \sum_{j=1}^K [(h(i)-h(i))a_{ij}[h(i)-h(j)]$$

➤ **Độ đo có trọng số**

$$d_{\text{hist}}(I,)= (h(I)-h(Q))^T A (h(I)-H(Q))$$

Trong đó,  $h(I)$  và  $h(Q)$  là những lược đồ tương ứng của ảnh I và Q, và A là ma trận đồng dạng  $K \times K$ . Trong ma trận này, những màu mà rất giống nhau thì gần với giá trị một, còn những màu rất khác nhau thì sẽ có giá trị gần với không.

### 2.2.2 Độ đo về kết cấu

Dùng phương pháp tương quan tự động và quang phổ năng lượng để đo độ tương đồng về kết cấu. Hàm tương quan tự động của một ảnh có thể được dùng để dò tìm những phần tử hoa văn của vân và cũng đồng thời mô tả tính mịn/tính thô của vân. Hàm tương quan tự động  $\rho(dr,dc)$  của một ảnh với kích thước  $(N+1) \times (N+1)$  với một độ dời  $d = (dr, dc)$  là:

$$\begin{aligned} \rho(dr, dc) &= \frac{\sum_{r=0}^N \sum_{c=0}^N I[r, c] I[r+dr, c+dc]}{\sum_{r=0}^N \sum_{c=0}^N I^2[r, c]} \\ &= \frac{I[r, c] o I_d[r, c]}{I[r, c] o I[r, c]} \end{aligned}$$

Nếu vân ở dạng thô, khi đó hàm tương quan tự động sụt giảm chậm, ngược lại nó sẽ sụt giảm rất nhanh. Đối với những vân theo quy tắc, hàm tự động tương quan sẽ yếu. Khi  $I[r+dr, c+dc]$  không được xác định rõ là đường bao của ảnh, một phương pháp để tính toán những giá trị ảnh ảo phải được định nghĩa.

Hàm tự động tương tác liên quan với quang phổ năng lượng của phép biến đổi Fourier. Nếu  $I[r,c]$  là hàm ảnh và  $F(u,v)$  là biến đổi Fourier của nó,  $F(u,v)$  2 được coi như là năng lượng quang phổ.

### 2.2.3 Độ đo về hình dạng

#### ➤ Độ so khớp đường biên của hình dạng

Thuật toán so khớp đường biên đòi hỏi sự trích rút và trình bày đường biên của cả ảnh cần truy vấn và ảnh mang ra so khớp. Đường biên có thể được trình bày bởi một dãy những điểm ảnh hay có thể được xấp xỉ bởi một đa giác. Đối với một dãy những điểm ảnh, một loại so khớp cổ điển là dùng mô tả Fourier để so sánh hai hình dạng với nhau. Trong toán học hàm liên tục, mô tả Fourier là những hệ số của dãy triển khai Fourier của hàm mà định nghĩa đường biên của hình dạng ảnh. Trong trường hợp đặc biệt, hình dạng được trình bày bởi dãy của  $m$  điểm  $\langle V_1, V_2, \dots, V_{m-1} \rangle$ . Từ những dãy điểm này, một dãy của véc tơ đơn vị:

$$v_k = \frac{V_{k+1} - V_k}{|V_{k+1} - V_k|}$$

và một dãy của những sai phân:

$$l_k = \sum_{i=1}^k |V_i - V_{i-1}|, \quad k > 0$$

$$l_0 = 0$$

có thể được tính.

Mô tả Fourier  $\{a_{-M}, \dots, a_0, \dots, a_M\}$  sau đó được xấp xỉ bởi:

$$a_n = \frac{1}{L \left( \frac{n2\pi}{L} \right) 2} \sum_{k=l}^m (v_{k-l} - v_k) e^{-jn(2\pi/L)l_k}$$

Những mô tả này có thể được dùng để định nghĩa độ đo khoảng cách hình dạng. Giả sử,  $Q$  là ảnh truy vấn và  $I$  là ảnh mang hình dạng được so sánh với  $Q$ . Gọi  $\{a_n^Q\}$  là dãy của những mô tả Fourier cho ảnh truy vấn, và  $\{a_n^I\}$  là mô tả Fourier cho ảnh. Khi đó, độ đo khoảng cách Fourier như sau:

$$d_{\text{Fourier}}(I, Q) = \left[ \sum_{N=m}^m |a_n^I - a_n^Q|^2 \right]^{1/2}$$

Như đã mô tả, khoảng cách này chỉ bất biến đối với phép tịnh tiến. Nếu mà những bất biến khác đòi hỏi, có thể dùng sự kết hợp với nhiều hàm số học để giải quyết vấn đề tỷ lệ, xoay và điểm bắt đầu để cực tiểu hoá  $d_{\text{Fourier}}(I, Q)$ .

Nếu đường biên được trình bày dưới dạng một đa giác, chiều dài của các cạnh và góc giữa chúng có thể được tính và dùng để trình bày hình dạng. Một hình dạng có thể được trình bày bởi một dãy những điểm nối liền nhau  $(X_i, Y_i, \alpha_i)$ , với cặp đường thẳng giao nhau tại điểm  $(X_i, Y_i)$  với góc lớn  $\alpha_i$ . Cho một dãy  $Q = Q_1, Q_2, \dots, Q_n$  của những điểm nối liền nhau trình bày đường biên của đối tượng truy vấn  $Q$  và một dãy tương tự  $I = I_1, I_2, \dots, I_m$  trình bày đường biên của đối tượng  $I$ , mục tiêu là tìm một ánh xạ từ  $Q$  tới  $I$  mà ánh xạ từ dãy phân đoạn của ảnh truy vấn tới dãy có chiều dài tương tự phân đoạn của ảnh và sao cho cặp của dãy phân đoạn truy vấn gần kề mà gặp tại một góc đặc biệt  $\alpha$  nên ánh xạ tới một cặp những dãy phân đoạn gần kề mà gặp nhau tại một góc  $\alpha$  tương tự.

Một kỹ thuật so khớp đường biên khác là so khớp mềm dẻo (elastic matching) trong đó hình dạng truy vấn được làm biến dạng để trở nên càng giống với hình dạng ảnh mẫu càng tốt. Sự cách biệt giữa hình dạng ảnh truy vấn và hình dạng của ảnh mẫu dựa vào hai thành phần chính: năng lượng đòi hỏi trong quá trình biến dạng từ hình dạng ảnh truy vấn tới mức độ phù hợp nhất với hình dạng trong ảnh mẫu. Và sự đo lường về độ giống nhau giữa hình dạng ảnh truy vấn sau khi bị biến dạng khớp với hình dạng trong ảnh mẫu.

### ➤ So khớp ảnh phát họa

Hệ thống so khớp phát họa cho phép người dùng nhập vào một bản phát họa của những đường biên chính trong một ảnh và sau đó hệ thống sẽ tìm kiếm những ảnh màu hay ảnh xám mà có đường biên khớp nhất. Những ảnh màu được tiền xử lý như sau để đạt được một dạng trung gian gọi là ảnh được trích rút (abstract image).

- Áp dụng phép biến đổi Affine để giảm kích thước ảnh về kích thước chỉ định trước. Dùng một mặt nạ trung vị để lọc nhiễu. Kết quả của bước này cho ra một ảnh được tiêu chuẩn hóa.

- Sử dụng thuật toán dò biên để dò tìm đường biên dựa trên gradient. Dò biên được tiến hành qua hai bước: những biên toàn cục sẽ được tìm thấy trước tiên với một ngưỡng toàn cục dựa trên giá trị trung bình và biến đổi của gradient; sau đó, những đường biên cục bộ sẽ được chọn từ toàn cục theo những ngưỡng cục bộ. Kết quả của bước này cho ra ảnh gọi là ảnh đã được lọc biên (refined edge image).

- Tiến hành làm mảnh và rút ngắn trên ảnh đã được lọc biên. Kết quả cuối cùng được gọi là ảnh được trích rút.

Khi người sử dụng nhập vào một bức ảnh phát họa ở dạng thô như là một ảnh truy vấn, nó cũng lần lượt được đưa qua các giai đoạn chuẩn hóa kích thước, nhị phân hóa, làm mảnh hoá và rút gọn. Kết quả của quá trình xử lý này cho ra một ảnh gọi là bản phát họa đều nét. Giờ đây, bản phát họa đều nét phải được so khớp với những ảnh được trích rút ở trên. Thuật giải so khớp là thuật giải dựa vào mối tương quan. Hai ảnh sẽ được chia ra làm những hệ thống lưới ô vuông. Đối với mỗi hệ thống lưới ô vuông của ảnh truy vấn, tương quan cục bộ với hệ thống lưới ô vuông tương ứng của ảnh cơ sở dữ liệu sẽ được tính. Để thiết thực hơn, tương quan cục bộ được tính cho một vài những dịch chuyển khác biệt trong vị trí trong hệ thống ô lưới trên những ảnh cơ sở dữ liệu và giá trị tương quan cực đại qua tất cả những dịch chuyển là kết quả của cho hệ thống ô lưới đó. Độ đo sự giống nhau cuối cùng là tổng của mỗi tương quan cục bộ. Độ đo khoảng cách là nghịch đảo của độ đo sự giống nhau này.

Từ những chú giải trên, nó có thể được biểu diễn lại dưới dạng sau:

1

$$d_{\text{sketch}}(I, Q) = \frac{1}{\sum_g \max_n [d_{\text{correlation}}(\text{shift}(A^1(g)), L^Q(g))]}$$

trong đó  $A^I(g)$  quy cho hệ thống ô lưới  $g$  của ảnh được trích rút được tính từ ảnh cơ sở dữ liệu  $I$ ,  $\text{shift}(A^I(g))$  quy cho phiên bản được dịch chuyển của hệ thống lưới  $g$  của cùng ảnh được trích rút, và  $L(g)$  quy cho hệ thống ô lưới  $g$  của bản họa đều nét có kết quả từ ảnh truy vấn  $Q$ .

### **2.3 Các nghiên cứu liên quan**

#### **2.3.1 Trong nước**

Hiện nay ở Việt Nam, nhất là ở thành phố Hồ Chí Minh có ít nhóm nghiên cứu liên quan đến vấn đề truy vấn ảnh. Kết quả thu được còn hạn chế. Các bài báo khoa học trong nước khá khiêm tốn.

#### **2.3.2 Ngoài nước**

##### **➤ Hệ thống QBIC (Query By Image Content)**

Hệ thống QBIC của hãng IBM là một hệ thống tra cứu ảnh thương mại đầu tiên và nổi tiếng nhất trong số các hệ thống tra cứu ảnh dựa trên nội dung. Nó cho phép người sử dụng tra cứu ảnh dựa vào màu sắc, hình dạng và kết cấu. QBIC cung cấp một số phương pháp: Simple, Multi-feature và Multi-pass. Trong phương pháp truy vấn Simple chỉ sử dụng một đặc điểm. Truy vấn multi-feature bao gồm nhiều đặc điểm và mọi đặc điểm đều có trọng số như nhau trong suốt quá trình tìm kiếm. Truy vấn Multi-pass sử dụng đầu ra của các truy vấn trước làm cơ sở cho bước tiếp theo. Người sử dụng có thể vẽ ra và chỉ định màu, kết cấu mẫu của hình ảnh yêu cầu. Trong hệ thống QBIC màu tương tự được tính toán bằng độ đo bình phương sử dụng biểu đồ  $k$  phần tử và màu trung bình được sử dụng như là bộ lọc để cải tiến hiệu quả của truy vấn.

##### **➤ Hệ thống Photobook**

Hệ thống này được phát triển ở viện kỹ thuật Massachusetts. Nó cho phép người sử dụng tra cứu ảnh dựa trên màu sắc, hình dạng và kết cấu. Hệ thống này cung cấp một tập các thuật toán đối sánh gồm: Euclidean, mahalanobis, vector space angle, histogram, Fourier peak, và wavelet tree distance như là những độ đo khoảng cách. Hệ thống như là một công cụ bán tự động và có thể sinh ra một mẫu truy vấn dựa vào những ảnh mẫu được cung cấp bởi người sử dụng. Điều này cho

phép người sử dụng trực tiếp đưa những yêu cầu truy vấn của họ với những lĩnh vực khác nhau, và mỗi lĩnh vực họ có thể thu được những mẫu truy vấn tối ưu.

➤ **Hệ thống RetrievalWare**

Hệ thống này được phát triển bởi tập đoàn công nghệ Excalibur cho phép người sử dụng tra cứu ảnh bởi nội dung màu, hình dạng, kết cấu, độ sáng, kết cấu màu và hệ số co. Người sử dụng có thể điều chỉnh tỷ trọng của những đặc điểm này trong suốt quá trình tìm kiếm.

➤ **Hệ thống Imatch**

Hệ thống này cho phép người sử dụng tra cứu ảnh bởi nội dung màu, hình dạng và kết cấu. Nó cung cấp một số phương pháp để tra cứu ảnh tương tự: Màu tương tự, màu và hình dạng, màu và hình dạng mờ, phân bố màu. Màu tương tự để tra cứu những ảnh tương tự với ảnh mẫu dựa trên sự phân bố màu toàn cục. Màu và hình dạng thực hiện tra cứu bởi việc kết hợp cả hình dạng, kết cấu và màu. Màu và hình dạng mờ thực hiện thêm những bước xác định đối tượng trong ảnh mẫu. Phân bố màu cho phép người sử dụng vẽ ra sự phân bố màu hoặc xác định tỷ lệ phần trăm của một màu trong hình ảnh mong muốn. Imatch cũng cung cấp những đặc điểm khác nội dung để xác định ảnh: ảnh nhị phân, ảnh định dạng khác và những ảnh có tên tương tự.

➤ **Hệ thống VisualSEEK và WebSEEK**

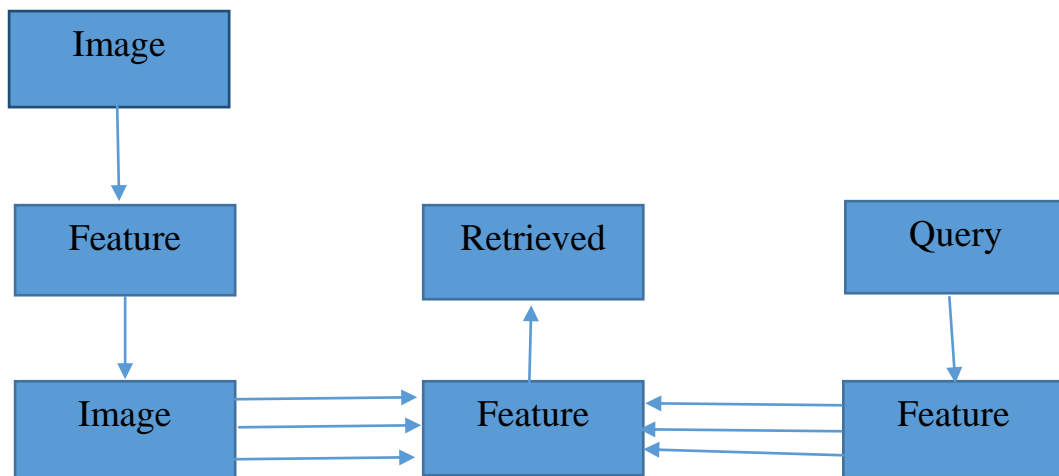
Cả hai hệ thống này đều được phát triển tại Trường Đại học Colombia. Visual SEEK là hệ thống cơ sở dữ liệu ảnh. Nó cho phép người sử dụng tra cứu ảnh dựa trên màu sắc, không gian miền và đặc điểm kết cấu. Thêm vào đó Visual SEEK còn cho phép người sử dụng tạo truy vấn bằng việc chỉ định vùng màu và những không gian vị trí của chúng. WebSEEK là một catalog ảnh và là công cụ tìm kiếm trên trang web (website).

## CHƯƠNG 3: PHƯƠNG PHÁP ĐỀ XUẤT TÌM KIẾM ẢNH THEO NỘI DUNG

### 3.1 Yêu cầu bài toán

Tìm kiếm ảnh theo nội dung là sử dụng các nội dung của ảnh để miêu tả và truy xuất từ cơ sở dữ liệu lớn. Hệ thống tìm kiếm ảnh theo nội dung (Content Based Image Retrieval - CBIR) điển hình được chia làm hai loại: trích chọn đặc trưng ngoại tuyến và truy xuất ảnh trực tuyến. Hình 3.1, biểu diễn kiến trúc của tìm kiếm ảnh theo nội dung, hệ thống tự động trích xuất thuộc tính thị giác (màu, hình dạng và kết cấu) của mỗi bức ảnh trong cơ sở dữ liệu, được thể hiện bằng các giá trị điểm ảnh và lưu trữ chúng trong một cơ sở dữ khác trong hệ thống gọi là cơ sở dữ liệu véc tơ đặc trưng. Dữ liệu đặc trưng như là chữ ký ảnh hoặc đặc trưng ảnh cho mỗi thuộc tính thị giác của mỗi ảnh là nhỏ hơn nhiều so với kích thước của dữ liệu ảnh.

Trong truy xuất ảnh trực tuyến là người sử dụng gửi một hình ảnh truy vấn đến hệ thống tìm kiếm ảnh theo nội dung để tìm ảnh mong muốn. Hệ thống sẽ biểu diễn ảnh truy vấn này với một véc tơ đặc trưng. Độ đo tương đồng giữa các véc tơ đặc trưng trong ảnh mẫu của các ảnh trong cơ sở dữ liệu đặc trưng được tính và xếp hạng. Sự truy xuất được tính bằng sơ đồ lập chỉ mục để cung cấp một phương pháp tìm kiếm hiệu quả trong cơ sở dữ liệu ảnh. Cuối cùng, hệ thống xếp hạng sẽ trả về kết quả các bức ảnh gần giống nhất so với ảnh truy vấn.



*Hình 3.1: Cấu trúc của hệ thống CBIR*

Cấu trúc của hệ thống CBIR đề xuất được thể hiện trong Hình 3.1. Đối với các cơ sở dữ liệu ảnh đã cho, các đặc trưng trích xuất đầu tiên từ các hình ảnh riêng biệt. Các đặc trưng có tính trực quan như màu sắc, kết cấu, hình dạng, vùng miền, đặc trưng không gian hoặc một số đặc trưng vùng nén ảnh. Đặc trưng sau khi trích xuất được mô tả bởi các véc tơ đặc trưng. Những véc tơ đặc trưng này sau đó được lưu trữ để tạo thành cơ sở dữ liệu đặc trưng hình ảnh. Đối với ảnh truy vấn nhất định, trích xuất đặc trưng của nó thể hiện bằng véc tơ đặc trưng. Véc tơ đặc trưng này được đối sánh với các véc tơ đã được lưu trữ trong cơ sở dữ liệu đặc trưng ảnh. Đôi khi kỹ thuật thu giảm số chiều được sử dụng để giảm các tính toán. Khoảng cách giữa các véc tơ đặc trưng của ảnh truy vấn với những ảnh trong cơ sở dữ liệu ảnh được tính toán. Khoảng cách của ảnh truy vấn với chính nó là số không nếu cùng trong cơ sở dữ liệu. Khoảng cách này sau đó được lưu trữ theo thứ tự tăng dần và được truy xuất với sự giúp đỡ của sơ đồ lập chỉ mục.

Kỹ thuật trích xuất đặc trưng ảnh hưởng đến tốc độ truy xuất của hệ thống CBIR. Một véc tơ đặc trưng là một tập hợp các thông số mô tả một ảnh.

Đa số các véc tơ được biểu diễn bởi một đặc trưng ảnh, chẳng hạn như màu sắc, kết cấu, hoặc hình dạng của đối tượng. Các véc tơ đặc trưng được tạo ra bởi các giải thuật giống nhau từ không gian véc tơ đặc trưng. Văn bản chú thích cho ảnh mô tả như lớp đặc trưng cấp cao. Các đặc trưng như màu sắc và kết cấu, hình dạng được gọi là đặc trưng cấp thấp.

Những vấn đề quan trọng của hệ thống đề xuất tìm kiếm hình ảnh dựa trên nội dung:

- Lựa chọn các cơ sở dữ liệu hình ảnh.
- Độ đo tương đồng.
- Đánh giá hiệu suất quá trình truy xuất.
- Trích xuất đặc trưng hình ảnh cấp thấp.

Đánh giá về hiệu suất truy xuất là một vấn đề rất quan trọng trong hệ thống tìm kiếm ảnh theo nội dung (CBIR). Nhiều phương pháp khác nhau để đo lường hiệu quả hoạt động của một hệ thống đã được đưa ra và được sử dụng bởi các nhà



ngiên cứu. Các biện pháp đánh giá phổ biến nhất được sử dụng trong CBIR là độ chính xác (Precision rate) và độ phủ (Recall rate) được định nghĩa như sau:

$$\text{Độ chính xác} = \frac{\text{Số ảnh tìm được đúng}}{\text{Số ảnh tìm được}}$$

$$\text{Độ phủ} = \frac{\text{Số ảnh tìm được đúng}}{\text{Tổng số ảnh đúng thực có trong CSDL}}$$

### 3.2. Mô hình nghiên cứu

#### 3.2.1. Mô hình phối hợp các đặc trưng ảnh

Trong mô hình này, có thể sử dụng một hoặc nhiều đặc trưng ảnh được trích dựa trên biến đổi wavelets kết hợp với các đặc trưng ảnh khác. Để đơn giản giả sử dùng 3 véc tơ đặc trưng ảnh là  $F^{W1}$ ,  $F^{W2}$  (các đặc trưng trích dựa trên wavelets) và  $F^{\text{other}}$  (là đặc trưng ảnh khác, ví dụ: đặc trưng màu) trích từ ảnh truy vấn; và tương ứng  $F_j^{W1}$ ,  $F_j^{W2}$ ,  $F_j^{\text{other}}$  được trích từ ảnh thứ j trong cơ sở dữ liệu. Mô hình này có thể được mở rộng cho nhiều (hoặc ít) hơn các đặc trưng trích dựa trên wavelets và các đặc trưng khác.

Bước đầu tiên của mô hình là ánh xạ các véc tơ đặc trưng đến kết quả đo đặc sự tương tự giữa hai véc tơ của hai ảnh. Cụ thể:

-  $D^{w1}: \mathbb{R}^{N_{w1}} \rightarrow \mathbb{R}$ , với  $N_{w1}$  là số chiều của véc tơ đặc trưng  $F^{w1}$ , theo phép đo:  $D^{w1}(F^{w1}, F_j^{w1})$ .

-  $D^{w2}: \mathbb{R}^{N_{w2}} \rightarrow \mathbb{R}$ , với  $N_{w2}$  là số chiều của véc tơ đặc trưng  $F^{w2}$ , theo phép đo:  $D^{w2}(F^{w2}, F_j^{w2})$ .

-  $D^{\text{other}}: \mathbb{R}^{N_{\text{other}}} \rightarrow \mathbb{R}$ , với  $N_{\text{other}}$  là số chiều của véc tơ đặc trưng  $F^{\text{other}}$ , theo phép đo:  $D^{\text{other}}(F^{\text{other}}, F_j^{\text{other}})$ .

Các phép đo  $D^{w1}$ ,  $D^{w2}$ ,  $D^{\text{other}}$  có thể là một trong các phép đo sự tương tự trình bày trong Mục 2.3. Bước tiếp theo của mô hình thực hiện ánh xạ  $D^{\text{matching}}:$

$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ , bởi:

$$D^{\text{matching}}(F^{w1}, F^{w2}, F^{\text{otherw1}}, F_j^{w1}, F_j^{w2}, F_j^{\text{otherw1}}) = \text{Matching}\{\alpha \cdot D^{w1}(F^{w1}, F_j^{w1}) \beta \cdot D^{w2}(F^{w2}, F_j^{w2}) \gamma \cdot D^{\text{other}}(F^{\text{other}}, F_j^{\text{otherw1}})\}$$

với  $\alpha, \beta, \gamma$  là các giá trị trọng số. *Matching* là hàm xác định từ thực nghiệm. Từ kết quả tính toán  $D^{\text{matching}}$  của ảnh truy vấn và các ảnh  $j$  trong cơ sở dữ liệu, căn cứ trên sắp xếp theo thứ tự từ nhỏ đến lớn của các giá trị để đưa ra các ảnh kết quả tương tự nhất với ảnh truy vấn.

### 3.2.2 Kết hợp toán tử LTP (Local Ternary Pattern)

Mẫu nhị phân cục bộ (LBP – Local Binary Pattern) là một bộ mã có 2 giá trị (nhị phân) được sử dụng thành công trong nhiều ứng dụng. Ý tưởng toán tử LBP được dựa trên hai giá trị là bit 1 hoặc bit 0. Điều này không cho phép toán tử LBP phân biệt giữa nhiều mẫu. Toán tử LBP có hai hạn chế là:

- Toán tử LBP không thể phân biệt giữa hai giá trị pixel nếu một trong những điểm ảnh đầu tiên gần các điểm ảnh trung tâm, nhưng một chút sau đó điểm ảnh và một trong những không thể phân biệt thứ hai là thấp hơn nhiều so với giá trị trung tâm điểm ảnh.

- Trong hình ảnh khu vực bằng phẳng, chẳng hạn như trong hình ảnh khuôn mặt, nơi mà tất cả các điểm ảnh gần như có cùng giá trị màu xám, nếu một số lượng nhỏ của tiếng ồn được thêm vào những khu vực khai thác LBP sẽ cung cấp cho một số bit giá trị 0 và khác giá trị 1. Vì vậy, tính năng LBP sẽ không ổn định và do đó toán tử LBP sẽ không thích hợp cho việc phân tích các khu vực này.

Để giải quyết những vấn đề này một toán tử kết cấu 3 giá trị mới (Mẫu tam phân cục bộ - LTP), có thể được coi như là một phần mở rộng cho LBP được giới thiệu gần đây. Thay vì một ngưỡng mà chỉ dựa trên các giá trị điểm ảnh trung tâm của miền lân cận, người sử dụng sẽ xác định một ngưỡng  $t$  và giá trị điểm ảnh bất kỳ trong khoảng thời gian  $-t$  và  $t+$ , do đó gán giá trị 0 cho điểm ảnh đó, trong khi người dùng gán giá trị 1 cho điểm ảnh đó nếu nó là ở trên ngưỡng này và một giá trị -1 nếu nó là dưới nó khi so sánh với các giá trị điểm ảnh trung tâm. Phương trình sau đây cho thấy làm thế nào để tính toán toán tử LTP.

$$LPT(i) = \begin{cases} 1 & \text{if } -p_c \geq t \\ 0 & \text{if } |p_i - p_c| \\ 1 & \text{if } -p_i - p_c \geq -t \end{cases} \quad (1)$$

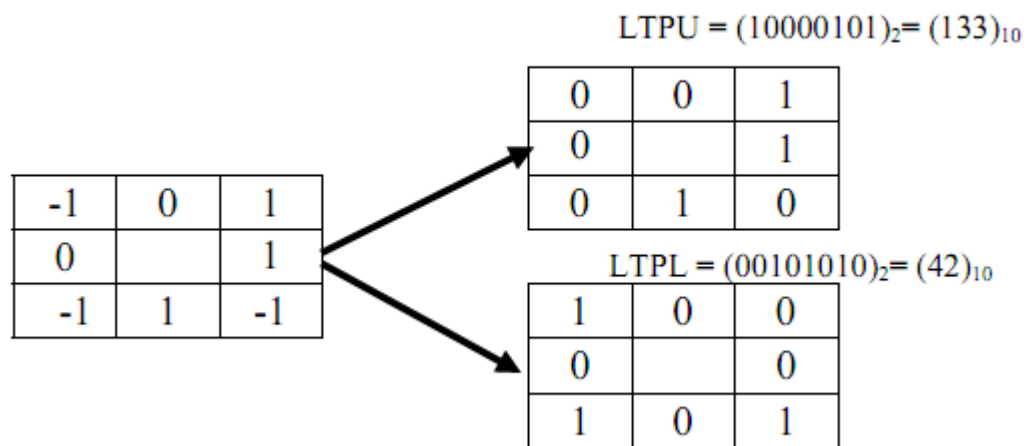
Trong đó  $t$  là ngưỡng do người sử dụng quy định,  $p_i$  là một giá trị điểm ảnh trong miền lân cận và  $p_c$  là giá trị điểm ảnh trung tâm.

Định nghĩa này dẫn đến việc có được một toán tử kết cấu ít nhạy cảm với tiếng ồn (vì nó không còn chủ yếu dựa trên giá trị của điểm ảnh trung ương) nhưng không còn chặt chẽ bất biến thành màu xám cấp phép biến đổi. Hình sau cho thấy một ví dụ về toán tử LTP hoạt động bằng cách sử dụng một giá trị ngưỡng  $t = 5$ :

12	34	45
38	35	55
11	65	23

-1	0	1
0		1
-1	1	-1

Để thoát khỏi những giá trị âm trong hình trên, các giá trị LTP được chia thành hai kênh LBP, phía trên LTP (LTP) và thấp hơn LTP (LTPL) như trong hình bên dưới. LTPU thu được bằng cách thay thế các giá trị tiêu cực trong LTP gốc bằng số không. Các LTPL thu được trong hai bước: đầu tiên, thay thế tất cả các giá trị của 1 trong việc LTP ban đầu là số không sau đó thay đổi các giá trị âm là 1.



### 3.2.3 Sử dụng support vector machine (SVM) cho bài toán truy vấn.

Xây dựng chương trình cho việc huấn luyện SVM là khó khăn và mất nhiều công sức, vì vậy đa số thuật chương trình SVM đã được xây dựng sẵn trong một số công cụ lập trình và cho bởi nguyên hàm và chỉ cần điền tham số, ta sẽ có kết quả đầu vào và đầu ra. Tuy nhiên, về cơ bản chương trình đó được xây dựng trên một số thuật toán như sau:

Những thuật toán thông dụng được thiết kế để huấn luyện SVM, có hai thuật toán kinh điển đã được cung cấp trong hầu hết ứng dụng SVM: thuật toán chặt khúc, thuật toán phân rã. Ý tưởng chính của những thuật toán này có thể trình bày tóm tắt như sau:

- *Thuật toán chặt khúc*

Thuật toán này bắt đầu với một tập con bất kỳ (*chunk*) của tập dữ liệu huấn luyện, sau đó huấn luyện SVM theo một phương án tối ưu trên tập con dữ liệu vừa chọn. Tiếp đến, thuật toán giữ lại các véc tơ hỗ trợ (các mẫu có  $\alpha_i > 0$ ) từ tập con sau khi đã loại bỏ các phần tử khác (tương ứng với  $\alpha_i = 0$ ) và dùng các véc tơ này để kiểm tra các phần tử trong phần còn lại của tập dữ liệu. Công việc này được lặp đi lặp lại, việc khởi tạo lại  $\alpha$  cho mỗi bài toán con mới phụ thuộc vào giá trị đầu ra của trạng thái trước đó và tiếp tục tối ưu bài toán con mới với các tham số tối ưu đã được lựa chọn. Thuật toán sẽ dừng lại khi thỏa mãn điều kiện tối ưu. Kích thước của tập con luôn thay đổi, nhưng cuối cùng nó bằng số lượng  $\alpha_i > 0$  (bằng số lượng véc tơ hỗ trợ).

Phương pháp này sử dụng ma trận Gram dùng để lưu tích vô hướng của từng cặp các véc tơ hỗ trợ phù hợp với kích thước bộ nhớ. Trong thực nghiệm, có thể xảy ra trường hợp số lượng véc tơ quá lớn, làm cho ma trận Gram vượt quá khả năng lưu trữ của máy tính.

- *Thuật toán phân rã*

Thuật toán này khắc phục nhược điểm của thuật toán chặt khúc bằng cách cố định kích thước của bài toán con (kích thước của ma trận Gram). Vì vậy tại mọi thời điểm, một phần tử mới được bổ sung vào tập làm việc thì một phần tử khác bị loại

ra. Điều này cho phép SVM có khả năng huấn luyện với tập dữ liệu lớn. Tuy nhiên, thực nghiệm cho thấy phương pháp này hội tụ chậm.

### **3.3 Phương pháp rút trích đặc trưng ảnh và truy vấn ảnh**

#### **3.3.1 Phối hợp các đặc trưng để truy vấn ảnh**

Phối hợp các đặc trưng để truy vấn ảnh bao gồm đặc trưng histogram màu, đặc trưng kết cấu dựa trên Gabor wavelets và đặc trưng dựa trên các hệ số của biến đổi contourlets.

##### **3.3.1.1 Xử lý cơ sở dữ liệu**

- Thực hiện rút trích đặc trưng histogram màu cho mỗi ảnh trong cơ sở dữ liệu ảnh, lưu tất cả các đặc trưng trích được của các ảnh vào tập tin.
- Thực hiện rút trích đặc trưng Gabor wavelets cho mỗi ảnh  $I_i$  trong cơ sở dữ liệu ảnh, lưu tất cả các đặc trưng trích được của các ảnh vào tập tin.
- Thực hiện rút trích đặc trưng contourlet cho mỗi ảnh  $i$  trong cơ sở dữ liệu ảnh, lưu tất cả các đặc trưng trích được của ảnh vào tập tin.

##### **3.3.1.2 Xử lý ảnh truy vấn**

- Chọn ảnh cần truy vấn.
- Rút trích đặc trưng Histogram màu và rút trích đặc trưng Gabor wavelets, contourlet.

##### **3.3.1.3 Đo sự tương tự giữa các véc tơ ảnh**

- Chọn cơ sở dữ liệu ảnh cần truy vấn ảnh.
- Tính véc tơ  $f_{\text{histogram}}$  là kết quả đo đặc sự tương tự giữa véc tơ đặc trưng histogram màu của ảnh truy vấn và mỗi véc tơ đặc trưng histogram màu lưu trong file `color_histogram_feature.txt`. Sử dụng phép đo Bhattacharyya cho tính toán này.
- Tính véc tơ  $f_{\text{gabor}}$  là kết quả đo đặc sự tương tự giữa véc tơ đặc trưng Gabor wavelet của ảnh truy vấn và mỗi véc tơ đặc trưng Gabor wavelet lưu trong file `gabor_wavelet_feature.txt`. Sử dụng phép đo Euclide.
- Tính véc tơ  $f_{\text{contourlet}}$  là kết quả đo đặc sự tương tự giữa véc tơ đặc trưng contourlet của ảnh truy vấn và mỗi véc tơ đặc trưng contourlet lưu

trong tập tin `contourlet_feature_SD.txt`. Sử dụng phép đo Manhattan cho phép đo này.

- Tính véc tơ đo đặc sự tương tự cho phương pháp phối hợp các đặc trưng như sau:

$$F = \alpha \times [1-f\_histogram] + \beta \times [f\_gabor] + \gamma \times [f\_contourlet]$$

Với  $\alpha, \beta$  là các hệ số.

#### 3.3.1.4 Hiển thị kết quả trả về

- Chọn  $R_n$  ngưỡng xác định số ảnh trả về.
- Sắp xếp các kết quả trong tập tin theo thứ giá trị tăng dần, hiển thị  $R_n$  ảnh ứng với  $R_n$  giá trị nhỏ nhất.

### 3.3.2 Truy vấn ảnh dùng moments của LTP

#### 3.3.2.1 Moments

Moment là một thước đo đối tượng hình dạng. Moment hình ảnh rất hữu ích để mô tả đối tượng sau khi phân khúc. Moment hình ảnh và nhiều loại hình bất biến dựa trên thời điểm đóng một vai trò quan trọng trong việc nhận dạng đối tượng và phân tích hình dạng.  $(P + q)$  thứ tự moment hình học  $M_{pq}$  của một cấp màu xám  $f(x, y)$  được định nghĩa như sau:

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (2)$$

Trong trường hợp rời rạc [59], các tích trong phương trình trên giảm xuống còn phép cộng và phương trình (2) trở thành như sau:

$$M_{pq} = \sum_{x=1}^n \sum_{y=1}^m x^p y^q f(x, y) \quad (3)$$

trong đó  $n, x, m$  là kích thước của ảnh cấp xám  $f(x, y)$ .

Thuộc tính đơn giản của hình ảnh đó được tìm thấy qua những moment hình ảnh bao gồm khu vực, trọng tâm của nó và thông tin về các định hướng. Đặc trưng moment là bất biến đối hình học. Đặc trưng này rất hữu ích để xác định các đối tượng

với hình dạng độc đáo bất kể kích thước và định hướng của chúng. Là bất biến theo tuyến tính phối hợp biến đổi, các bất biến thời điểm là các tính năng hữu ích trong vấn đề nhận dạng mẫu. Moments đã được sử dụng để phân biệt giữa hình dạng của khí cầu khác nhau, nhận dạng ký tự, và các ứng dụng phù hợp với bối cảnh. Thuộc tính sau của những moment hình ảnh là rất hữu ích trong việc truy xuất hình ảnh.

- Moment đặc trưng là bất biến đối hình học.
- Moment đặc trưng cung cấp đủ khả năng phân biệt giữa các đối tượng hình dạng khác nhau.
- Moment đặc trưng cung cấp mô tả cục bộ hiệu quả để xác định hình dạng của đối tượng.
- Dãy vô hạn các moment duy nhất nhận dạng đối tượng.

### **3.3.2.2 Mẫu tam phân (LTP) và Moment**

- Đặc trưng duy nhất không nắm bắt đầy đủ thông tin của một hình ảnh. Sự kết hợp của các đặc trưng được yêu cầu phải kết hợp các chi tiết tốt của một hình ảnh trong khi xây dựng véc tơ đặc trưng. Sự kết hợp của các tính năng cục bộ và toàn cục là một trong những cách tiếp cận như vậy theo hướng này. Các đặc trưng cục bộ giúp việc nắm bắt các biến cục bộ. Mặt khác, các đặc trưng toàn cục nắm bắt ý tưởng tổng thể của một hình ảnh. Ngoài ra, phương pháp này kết hợp những ưu điểm của cả các đặc trưng. Sự kết hợp của LTP và các moment giúp trong việc thực hiện các tiêu chí này. LTP, một đặc trưng cục bộ ghi lại chi tiết kết cấu và hoạt động như một bộ phân lớp mạnh mẽ. Moment, một đặc trưng toàn cục xác định hình dạng của một đối tượng trong hình ảnh và là bất biến để chuyển đổi hình học. Những lợi thế của sự kết hợp này được tóm tắt như sau:

- LTP so với LBP, ít nhạy cảm với tiếng ồn và do đó sự kết hợp của LTP với những moment ít bị ảnh hưởng bởi sự hiện diện của tiếng ồn.
- Việc sử dụng moment hình học như là một tính năng duy nhất tạo ra bất ổn số vì nó có giá trị cao hơn những moment thứ tự [60]. Nhưng sự kết hợp của LTP và moment để khắc phục nhược điểm này là moment giá trị của LTP không phải là rất cao.

- Moment hình học là bất biến đối hình học. Do đó sự kết hợp với LTP có lợi thế hơn trong các véc tơ đặc trưng LTP-Moment.

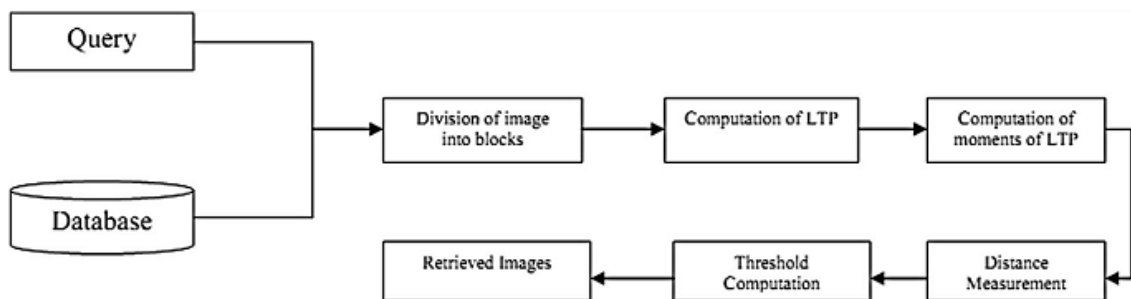
### 3.3.2.3 Phương pháp đề xuất

Phương pháp đề xuất bao gồm ba bước:

- Bước đầu tiên là có liên quan với các bộ phận của hình ảnh vào khối và cách tính mã LTP của mỗi khối.

- Trong bước thứ hai, chúng tôi tính toán các moment hình học của LTP mã số của ảnh truy vấn và hình ảnh cơ sở dữ liệu.

- Threshold được tính toán để thực hiện thu hồi ở bước 3.



Hình 3.2 Sơ đồ truy xuất hình ảnh từ cơ sở dữ liệu ảnh

#### ➤ Tính toán các mã LTP

Các thuật toán để tính toán mã LTP là như sau:

- Chuyển đổi hình ảnh thành màu xám.
- Rescale hình  $252 \times 252$ .
- Chia ảnh thành các khối  $84 \times 84$  và tính LTP mã số của từng khối.
- Tính toán các LTP mang hai giá trị: trên LTP (ULTP) và thấp hơn LTP (LLTP).

#### ➤ Tính toán các moment

Mã moment hình học của OLTP và OLTP được tính bằng cách sử dụng công thức (3). Trình tự của những moment được chọn từ 0-15. Các giá trị moment của OLTP và OLTP được tính riêng biệt.

#### ➤ Đo khoảng cách



Các moment của mã LTP cho các khối khác nhau của truy vấn hình ảnh được biểu diễn như sau  $m_Q = (m_{Q1}, m_{Q2}, \dots, m_{Qn})$ .

Các moment của mã LTP cho các khối khác nhau của hình ảnh cơ sở dữ liệu được biểu diễn như là  $m_{DB} = (m_{DB1}, m_{DB2}, \dots, m_{DBn})$ . Sau đó, khoảng cách Euclide giữa những khoảnh khắc khối LTP của ảnh truy vấn và hình ảnh cơ sở dữ liệu được tính như sau:

$$D(m_Q, m_{DB}) = \sqrt{(m_{Qi} - m_{DBi})^2} \quad (4)$$

### ➤ Tính ngưỡng

Ngưỡng được sử dụng để thực hiện truy xuất. Sử dụng ngưỡng để cải thiện kết quả truy xuất so với các kết quả truy xuất thu được mà không sử dụng ngưỡng.

Ý tưởng cơ bản đằng sau ngưỡng tính toán là để tìm ra khoảng các giá trị khoảng cách mà trả lại hình ảnh giống với ảnh truy vấn.

Chỉ số tương tự Euclide tăng dần để các hình ảnh được sắp xếp theo độ tương tự để truy vấn hình ảnh. Đó là, các hình ảnh tương tự như hầu hết các đầu tiên và những người khác sau đó.

Các chỉ số tương tự của hình ảnh được lưu trữ cùng với các giá trị khoảng cách của chúng để xác định phạm vi giá trị tối thiểu và tối đa.

Điều này xác định khoản tương tự với một hình ảnh truy vấn. Thủ tục này được lặp lại cho mọi hình ảnh của cơ sở dữ liệu để tìm ra khoảng tương tự. Cuối cùng, các phạm vi tối thiểu và tối đa của các giá trị được ngăn chặn, khai thác. Những giá trị này xác định ngưỡng của toàn bộ danh mục các hình ảnh tương tự.

Điều này được thực hiện cho tất cả các loại hình ảnh trong cơ sở dữ liệu. Các giá trị ngưỡng cho LTP trên và LTP dưới được tính cách riêng biệt. Tính toán ngưỡng như sau:

- N là tổng số các hình ảnh có liên quan trong cơ sở dữ liệu và NDB được tổng số hình ảnh trong cơ sở dữ liệu.

- Sortmat là ma trận được sắp xếp (thứ tự tăng dần) của khoảng cách giá trị và Minix được N chỉ số đầu tiên của hình ảnh trong ma trận sortmat.

- *Start\_range* và *end\_range* được phạm vi của các hình ảnh có liên quan trong cơ sở dữ liệu.

- *Maxthresh* và *minthresh* lần lượt là các giá trị khoảng cách tối đa và tối thiểu của mỗi ảnh truy vấn.

- *Mthreshmat* là tối đa của tất cả các giá trị của *maxthresh*.

Thuật toán tính ngưỡng:

1. For  $u \leftarrow 1$  to  $N$ 
  - 1.1 if ( $minix(u) \geq start\_range$  and  $minix(u) \leq end\_range$ ) then
    - 1.1.1  $mthresh \leftarrow sortmat(u)$ ;
  - 1.2 endif
2. endfor
3.  $maxthreshmat \leftarrow mthresh$
4.  $mthreshmat \leftarrow \max(maxthreshmat)$ ;
5. For  $u \leftarrow 1$  to  $NDB$ 
  - 5.1 if ( $sortmat(u) \geq minthresh$  and  $sortmat(u) \leq maxthresh$ ) then
    - 5.1.1 if ( $ini(u) \geq start\_range$  and  $ini(u) \leq end\_range$ ) then
      - 5.1.1.1  $freq \leftarrow freq + 1$ ;
    - 5.1.2 endif
  - 5.2 endif
6. endfor
7. end

## CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM

### 4.1 Tập dữ liệu thử nghiệm

Ở chương này sẽ hiện thực quá trình truy vấn ảnh bằng giải thuật đã đề xuất. Các kết quả thí nghiệm được thực hiện trên máy tính có cấu hình: CPU Intel Core I5-4200U 1.6 GHz (4 CPUs), RAM 8GB trên hệ điều hành Windows 7. Toàn bộ phương pháp này được trình bày cũng như so sánh đều được hiện thực trên ngôn ngữ Matlab, phiên bản 2013b.

Để đánh giá kết quả của giải thuật, quá trình thực nghiệm được thực hiện trên nhiều tập dataset khác nhau, chứa các đối tượng khác nhau như người, xe, hoa, cảnh... trong tập dữ liệu của Wang (<http://wang.ist.psu.edu/docs/related/>). Tập dữ liệu này bao gồm 10.000 ảnh.

Một số hình ảnh trong tập dữ liệu test:

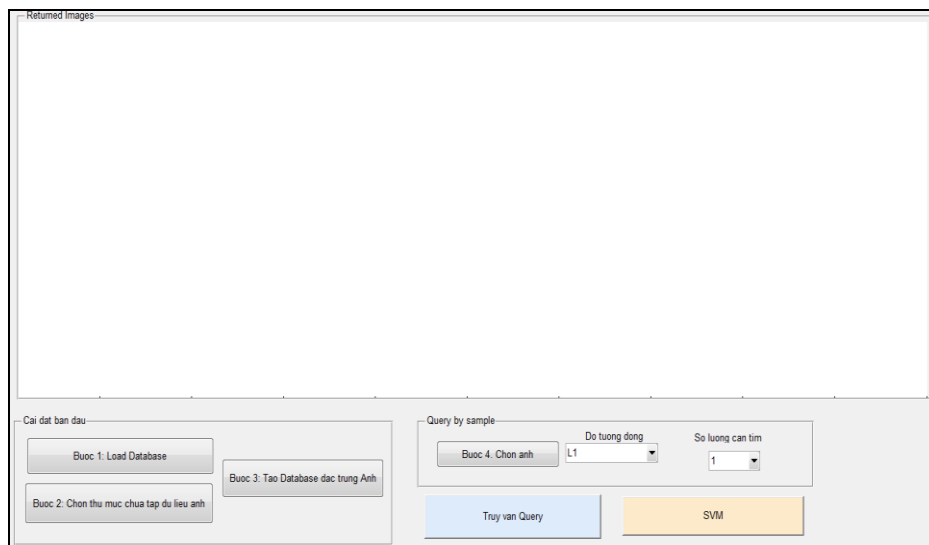




Hình 4.1 Một số hình trong tập dataset Wang

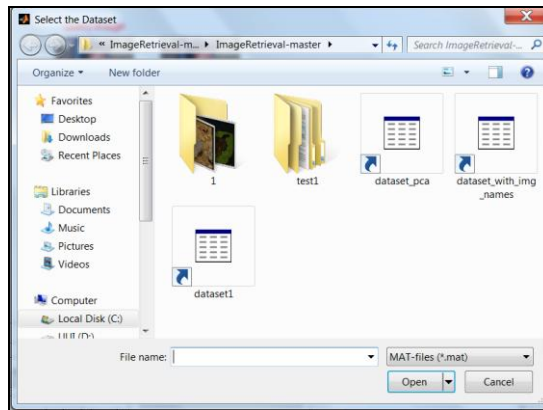
## 4.2 Kết quả truy vấn

Giao diện chính chương trình như hình 4.2.



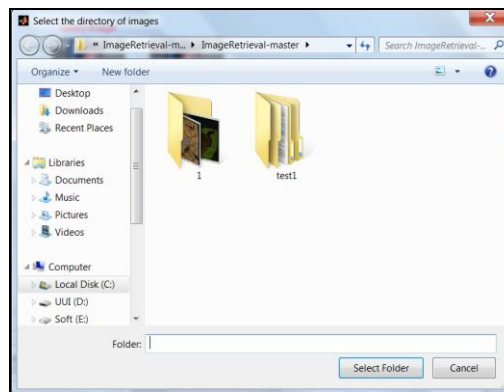
Hình 4.2 Giao diện chính chương trình

+ Bước 1: Load database. Trong bước này, chúng ta chọn thư mục chứa database như hình 4.3



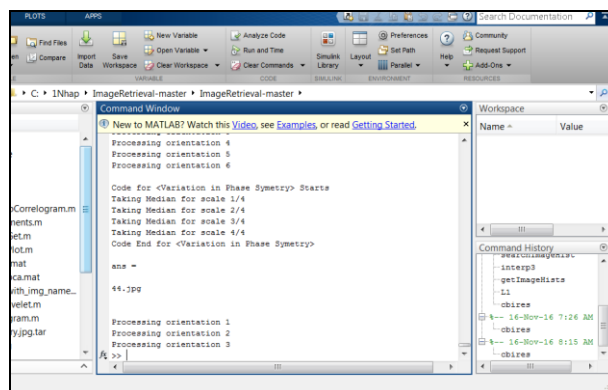
Hình 4.3 Giao diện chọn dataset

+ Bước 2: Chọn thư mục chứa dữ liệu ảnh như hình 4.4



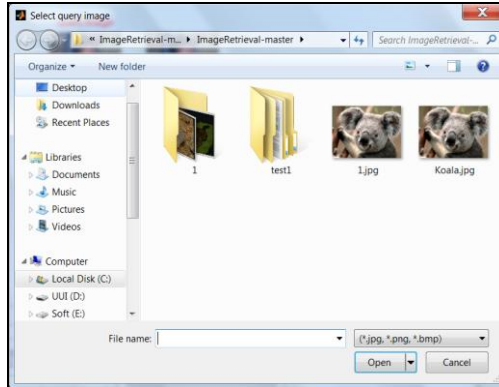
Hình 4.4 Giao diện chọn thư mục chứa ảnh

+ Bước 3: Rút trích đặc trưng và tạo database chứa các đặc trưng của ảnh như hình 4.5



Hình 4.5 Giao diện rút trích đặc trưng của ảnh

Bước 4. Chọn ảnh cần truy vấn (ảnh input vào)



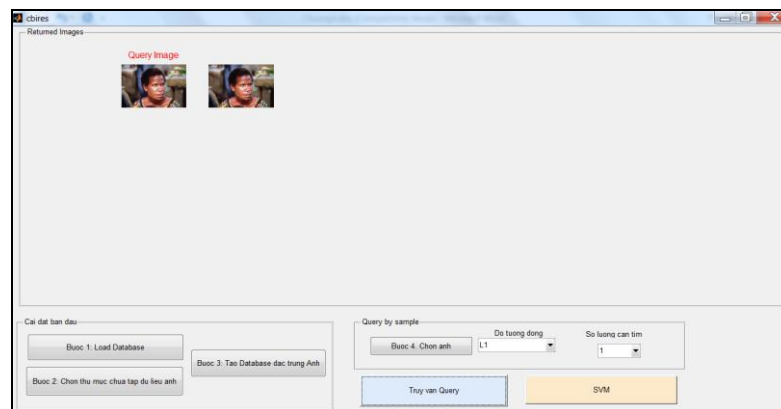
Hình 4.6 Giao diện chọn ảnh ban đầu

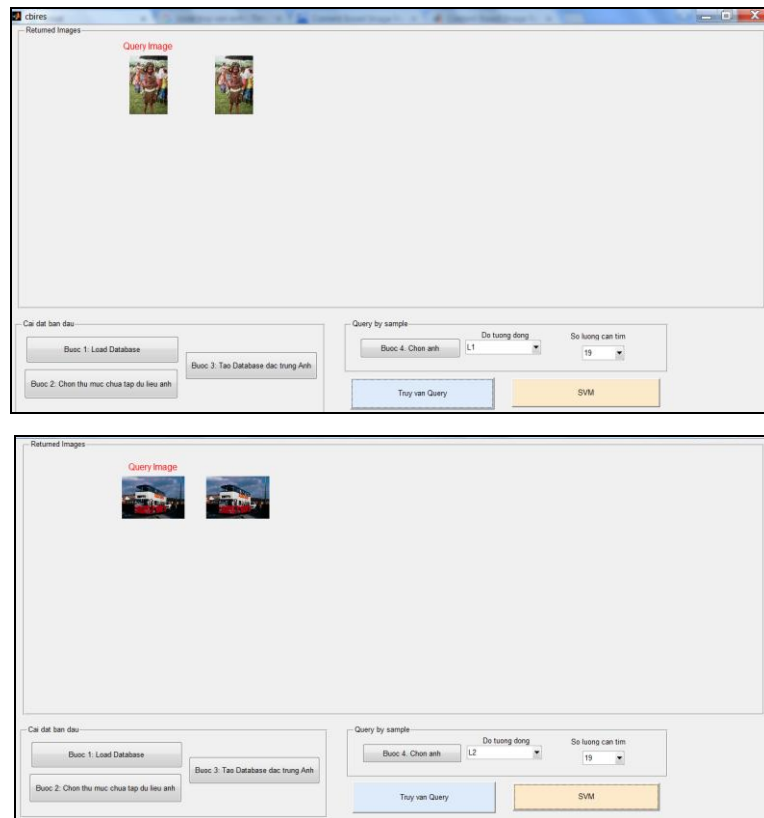
Bước 5: Chọn độ tương đồng giữa hình đầu vào và hình cần tìm

Bước 6: Chọn số lượng ảnh cần tìm

Bước 7: Chọn phương pháp truy xuất.

Ở đây chúng ta thử nghiệm trên 2 phương pháp: Truy xuất bằng query và truy xuất dựa trên phương pháp Support Vector Machine (SVM).





Hình 4.7 Kết quả truy xuất

Bảng 4.1 Kết quả truy xuất ảnh của phương pháp đề xuất so với phương pháp khác

Số thứ tự ảnh test	Số ảnh trùng thực tế	Phương pháp truy xuất dựa trên SVM	Phương pháp đề xuất
1	12	12	12
2	10	10	10
3	21	21	21
4	14	14	14
5	15	13	15
6	5	5	5
7	14	12	14
8	3	2	3

9	6	6	6
10	2	2	2
11	12	12	12
12	10	10	10
13	21	20	21
14	14	14	14
15	15	13	15
16	5	5	5
17	14	12	14
18	3	2	3
19	6	6	6
20	2	2	2
21	2	2	2
22	1	1	1
23	12	11	12
24	15	15	15
25	11	10	11
26	15	15	15
27	11	11	11
28	9	7	9
28	3	3	3
30	1	1	1
31	12	12	12
32	7	7	7



33	5	4	5
34	12	12	12
35	11	11	11
36	4	4	4
37	14	12	14
38	3	2	3
39	0	0	0
40	1	1	1
41	4	4	4
42	3	3	3
43	1	1	1
44	17	17	17
45	11	12	11
46	7	7	7
47	10	10	10
48	1	0	1
49	9	9	9
50	10	11	10

Chúng tôi thử trên nhiều thí nghiệm khác nhau. Trong khuôn khổ luận văn này, chúng tôi chỉ nêu 50 trường hợp trong cơ sở dữ liệu 300 ảnh. Từ kết quả bảng 4.1 và nhiều kết quả khác, chúng tôi thấy rằng phương pháp đề xuất cho kết quả tốt hơn phương pháp SVM. Phương pháp đề xuất đã kết hợp các đặc trưng cục bộ và toàn cục lại với nhau. Đây là lý do giải thích tại sao phương pháp đề xuất tốt hơn phương pháp SVM.

### 4.3 Code đặc trưng

#### Hàm truy xuất ảnh

```
function btn_BrowseImage_Callback(hObject, eventdata, handles)

[query_fname, query_pathname] = uigetfile('*.jpg; *.png; *.bmp', 'Select query
image');

if (query_fname ~= 0)
    query_fullpath = strcat(query_pathname, query_fname);
    imgInfo = imfinfo(query_fullpath);
    [pathstr, name, ext] = fileparts(query_fullpath); % fileparts returns char type

    if ( strcmp(lower(ext), '.jpg') == 1 || strcmp(lower(ext), '.png') == 1 ...
        || strcmp(lower(ext), '.bmp') == 1 )

        queryImage = imread( fullfile( pathstr, strcat(name, ext) ) );
        % Truy xuất image features

    queryImage = imresize(queryImage, [384 256]);
    if (strcmp(imgInfo.ColorType, 'truecolor') == 1)
        hsvHist = hsvHistogram(queryImage);
        autoCorrelogram = colorAutoCorrelogram(queryImage);
        color_moments = colorMoments(queryImage);

        % Scale image

        img = double(rgb2gray(queryImage))/255;
        [meanAmplitude, msEnergy] = gaborWavelet(img, 4, 6); % 4 = number of
scales, 6 = number of orientations
```

```

    wavelet_moments = waveletTransform(queryImage, imgInfo.ColorType);
    % construct the queryImage feature vector
    queryImageFeature = [hsvHist autoCorrelogram color_moments
meanAmplitude msEnergy wavelet_moments str2num(name)];
    elseif (strcmp(imgInfo.ColorType, 'grayscale') == 1)
        grayHist = imhist(queryImage);
        grayHist = grayHist/sum(grayHist);
        grayHist = grayHist(:)';
        color_moments = [mean(mean(queryImage)) std(std(double(queryImage)))];
        [meanAmplitude, msEnergy] = gaborWavelet(queryImage, 4, 6); % 4 =
number of scales, 6 = number of orientations
        wavelet_moments = waveletTransform(queryImage, imgInfo.ColorType);
        % construct the queryImage feature vector
        queryImageFeature = [grayHist color_moments meanAmplitude msEnergy
wavelet_moments str2num(name)];
    end
    % update handles
    handles.queryImageFeature = queryImageFeature;
    handles.img_ext = ext;
    handles.folder_name = pathstr;
    guidata(hObject, handles);
    helpdlg('Proceed with the query by executing the green button!');

    % Clear workspace
    clear('query_fname', 'query_pathname', 'query_fullpath', 'pathstr', ...
        'name', 'ext', 'queryImage', 'hsvHist', 'autoCorrelogram', ...
        'color_moments', 'img', 'meanAmplitude', 'msEnergy', ...
        'wavelet_moments', 'queryImageFeature', 'imgInfo');
else

```

```
        errorDlg('You have not selected the correct file type');  
    end  
else  
    return;  
end
```

## CHƯƠNG 5: KẾT LUẬN

### 5.1 Kết quả đạt được

Qua quá trình thực hiện luận văn, tôi đã hiểu được tính chất của các ảnh trong cơ sở dữ liệu và sau đó tìm ra phương pháp phù hợp với từng loại ảnh.

Đối với ảnh nhiễu do ánh sáng ngoài trời ta có thể dùng hàm lọc để xóa đi những pixel bị nhiễu. Từ đó đem lại những kết quả rất tốt.

Dựa trên các công trình nghiên cứu trong và ngoài nước trong khoảng thời gian gần đây, luận văn đã đề xuất một giải thuật truy vấn ảnh dựa trên các đặc trưng của ảnh. Sau quá trình thực nghiệm và nghiên cứu lý thuyết, tôi đã có một vài kết quả khả quan như: hoàn thiện những ý tưởng chính cũng như hiện thực để kiểm chứng trong thực tế.

Đề tài đã tiến hành đánh giá kết quả dựa trên các phương pháp định tính và định lượng, cũng như tiến hành so sánh kết quả từng giai đoạn với các phương pháp xử lý gần đây đã được đề xuất.

### 5.2 Ưu điểm và nhược điểm của giải thuật đề xuất

#### 5.2.1 Ưu điểm của giải thuật đề xuất

Giải quyết được bài toán truy vấn ảnh dựa trên đặc trưng của ảnh, giúp ích cho quá trình tìm kiếm ảnh. Kết quả thu được đã cải thiện được phần nào chất lượng tìm kiếm ảnh.

#### 5.2.2 Nhược điểm của giải thuật đề xuất

Mặc dù giải quyết được việc tìm kiếm ảnh, nhưng dựa trên quá trình xây dựng thuật giải cũng như thực nghiệm vẫn còn tồn tại một số vấn đề sau đây mà có thể xem như là nhược điểm của đề tài:

- Giải thuật chỉ mới xử lý với ảnh số, ảnh 2D.
- Tập dữ liệu ảnh thử nghiệm còn hạn chế do một số nguyên nhân hạn chế khi thu thập nguồn dữ liệu ảnh mẫu.
- Chưa tính đến tốc độ xử lý cũng như độ phức tạp của giải thuật.

### 5.3 Đóng góp của luận văn

### 5.3.1 Đóng góp khoa học

Đề tài đã đóng góp:

- Lý thuyết về các loại ảnh.
- Lý thuyết về các phương pháp truy vấn ảnh
- Đề tài đề xuất được một giải thuật kết hợp các đặc trưng cục bộ và toàn cục với nhau để truy vấn ảnh.
- Tạo được một tài liệu tham khảo tốt cho những ai quan tâm đến lĩnh vực xử lý ảnh.

### 5.3.2 Đóng góp thực tiễn

Đề tài đã tiến hành tìm hiểu được các phương pháp truy vấn ảnh phổ biến. Các đặc trưng cơ bản của một ảnh đã được nêu rõ cũng như các yêu cầu cơ bản khi truy vấn ảnh, không làm mất mát thông tin của ảnh.

#### 5.4 Hướng mở rộng

Do quá trình thu thập dữ liệu mẫu là các ảnh tự nhiên gặp nhiều khó khăn nên số lượng ảnh còn khá hạn chế, để có thể khẳng định kết quả giải thuật một cách triệt để hơn thì việc bổ sung thêm tập dữ liệu mẫu. Không những thế, việc bổ sung số lượng ảnh vẫn cần bổ sung thêm việc đa dạng về thể loại ảnh và kích cỡ.

Giải thuật vẫn tồn tại hạn chế là chỉ giải quyết trên ảnh ảnh 2D, nên vấn đề đặt ra còn là khắc phục nhằm giúp giải thuật xử lý được với ảnh 3D và nhiều kích cỡ hơn. Đây là những xu hướng phát triển trong thực tế hiện nay.

Do những hạn chế về mặt phần cứng nên đề tài chưa thể tiến hành đánh giá được thời gian thực hiện.

## TÀI LIỆU THAM KHẢO

- [1] Stricker, M. and Orengo, M. (1995), 'Similarity of Color Images', Proc. of the SPIE Conf, vol. 2420, pp. 381–392, 1995.
- [2] Ioka, M. (1989), 'A Method of Defining the Similarity of Images on the Basis of Color Information', Tech. Report RT-0030, IBM Tokyo Research Lab.
- [3] Vassilieva, N. and Novikov, B. (2005) 'Construction of Correspondences between Low-level Characteristics and Semantics of Static Images', Proc. of the 7th All-Russian Scientific Conf. 'Electronic Libraries: Perspective Methods and Technologies, Electronic Collections' RCDL'2005, Yaroslavl', Russia.
- [4] Stricker, M. and Dimai, A. (1997), 'Spectral Covariance and Fuzzy Regions for Image Indexing', Machine Vision Applications, vol. 10, pp. 66–73.
- [5] Hong-Bo Zhang & Shang-An Li & Shu-Yuan Chen & Song Zhi Su & Der-Jyh Duh & Shao Zi Li.(2012) 'Adaptive photograph retrieval method Multimedia Tools & Applications', DOI 10.1007/s11042-012-1233-7.
- [6] Jun Yue, Zhenbo Li, Lu Liu and Zetian Fub (2011), 'Content-based image retrieval using color and texture fused features', Mathematical and Computer Modelling, vol.54, pp. 1121–1127.
- [7] Daisy, M.M.H., TamilSelvi, S. and Prinza, L. (2012) 'Gray Scale Morphological Operations for Image Retrieval', 2012 International Conference on Computing, Electronics and Electrical Technologies [ICCEET], pp. 571-575.
- [8] Chaobing Huang, Yarong Han, Yu Zhang (2012), 'A Method for Object-based Color Image Retrieval', Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on , pp:1659-1663.
- [9] Fernando, R. and Kulkarni, S. (2012), 'Hybrid Technique for Colour Image Classification and Efficient Retrieval based on Fuzzy Logic and Neural

- Networks’, Neural Networks (IJCNN), The 2012 International Joint Conference on, pp:1-6.
- [10] Zhu Qiaoqiao, Huang Yuanyuan. (2012), ‘A New Image Retrieval Method Based on Color Feature’, Intelligent System Design and Engineering Application (ISDEA), 2012 Second International Conference on, pp:56-59, 2012.
- [11] Rasli, R.M.Muda, T.Z.T., Yusof, Y., Bakar, J.A. (2012) ‘Comparative Analysis of Content Based Image Retrieval Technique using Color Histogram. A Case Study of GLCM and K-Means Clustering’, Intelligent Systems, Modelling and Simulation (ISMS), Third International Conference on, pp: 283 – 286.
- [12] Content-based binary image retrieval using the adaptive hierarchical density histogram Panagiotis Sidiropoulos, Stefanos Vrochidis, Ioannis Kompatsiaris, Pattern Recognition vol. 44, pp. 739–750.
- [13] Konstantinos Konstantinidis, Vasileios Vonikakis, Georgios Panitsidis and Ioannis Andreadis (2011), ‘A Center-Surround Histogram for content-based image retrieval’, Pattern Anal Applic, vol. 14, pp. 251–260.
- [14] Xiang-Yang Wang & Jun-Feng Wu & Hong-Ying Yang(2010), ‘Robust image retrieval based on colour histogram of local feature regions’, Multimed Tools Appl vol. 49, pp. 323–345.
- [15] Haojie Li, Xiaohui Wang, Jinhui Tang and Chunxia Zhao(2013), ‘Combining global and local matching of multiple features for precise item image retrieval’, Multimedia Systems vol. 19, pp. 37–49.
- [16] Haralick, R.M., Shanmugam, K., and Dinstein, I.(1973), ‘Textural Features for Image Classification’, IEEE Trans.Systems, Man Cybernetics, vol. 3, no. 6, pp. 610– 621.
- [17] Howarth, P. and Rüger, S. (2004), ‘Evaluation of Texture Features for Content-based Image Retrieval’, Proc. Of CIVR'04, pp. 326–334.



- [18] Jiayin Kang and Wenjuan Zhang, 'A Framework for Image Retrieval with Hybrid Features', Control and Decision Conference (CCDC), 2012 24th Chinese, pp: 1326 – 1330, 2012.
- [19] Yong-Hwan Lee, Sang-Burm Rhee, Bonam Kim, 'Content-based Image Retrieval Using Wavelet Spatial-Color and Gabor Normalized Texture in Multi-resolution Database', Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on, pp: 371 – 377, 2012.
- [20] Smith, J.R. and Chang, S.-F. (1994), 'Transform Features For Texture Classification and Discrimination in Large Image Databases', Proc. of IEEE Int. Conf. on Image Processing (ICIP-94), Austin.
- [21] Do, M.N. and Vetterli, M. (2000), 'Texture Similarity Measurement Using Kullback–Leibler Distance on Wavelet Subbands', Proc. of Int. Conf. on Image Processing, 2000, vol. 3, pp. 730–733.
- [22] Sumana, I.J., Guojun Lu, Dengsheng Zhang (2012), 'Comparison of Curvelet and Wavelet Texture Features for Content Based Image Retrieval', Multimedia and Expo (ICME), IEEE International Conference on, pp.290 – 295.
- [23] Gallas, A., Barhoumi, W., Zagrouba, E. (2012), 'Image Retrieval Based on Wavelet Sub-bands and Fuzzy Weighted Regions', Communications and Information Technology (ICCIT), 2012 International Conference on , pp: 33 – 37.
- [24] Quellec, G., Lamard, M., Cochener, B., Roux, C., Cazuguel, G.(2012), 'Comprehensive Wavelet-Based Image Characterization for Content Based Image Retrieval', Content-Based Multimedia Indexing (CBMI), 2012 10th International Workshop on , pp:1-6.
- [25] Anil Balaji Gonde, R.P. Maheshwari and Balasubramanian (2013), 'Modified curvelet transform with vocabulary tree for content based image retrieval', Digital Signal Processing vol. 23, pp: 142–150.

- [26] Esmat Rashedi, Hossein Nezamabadi-pour and Saeid Saryazdi (2013), 'A simultaneous feature adaptation and feature selection method for content-based image retrieval systems', Knowledge-Based Systems Volume 39, Pages 85–94.
- [27] Ela Yildizer, Ali Metin Balci, Mohammad Hassan and Reda Alhadj (2012), 'Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble', Expert Systems with Applications Volume 39, Issue 3, Pages 2385–2396.
- [28] Quellec, G., Lamard, M., Cazuguel, G., Cochener, B. (2012), 'Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval', Ieee Transactions On Image Processing, Vol. 21, No. 4.
- [29] Tamura, H., Mori, S., and Yamawaki, T. (1978), 'Textural Features Corresponding to Visual Perception', IEEE Trans. Systems, Man Cybernetics, vol. 8, pp. 460–472.
- [30] Howarth, P. and Rüger, S. (2005), 'Robust Texture Features for Still Image Retrieval', IEEE Proc. Vision, Image Signal Processing, vol 152, no. 6, pp. 868–874.
- [31] Howarth, P. and Rüger, S. (2004), 'Evaluation of Texture Features for Content-based Image Retrieval', Proc. Of CIVR'04, pp. 326–334.
- [32] Sebe, N. and Lew, M.S.(2000), 'Wavelet Based Texture Classification, Proc. of Int. Conf. on Pattern Recognition', vol. 3, pp. 959–962.
- [33] Manjunath, B.S. and Ma, W.Y. (1996), 'Texture Features for Browsing and Retrieval of Image Data', IEEE Trans. Pattern Analysis Machine Intelligence, vol. 18, no. 8, pp. 837–842.
- [34] Manjunath, B.S., Wu, P., Newsam, S., and Shin, H.D. (2000), 'A Texture Descriptor for Browsing and Similarity Retrieval, Proc. Signal Processing Image Commun.', nos. 1–2, pp. 33–43.
- [35] Bell, A.J. and Sejnowsky, T.J. (1997), 'The 'Independent Components' of Natural Scenes are Edge Filters, Vision Research', no. 37, pp 3327–3338.

- [36] Borgne, H., Guerin-Dugue, A., and Antoniadis, A.(2004), ‘Representation of Images for Classification with Independent Features, Pattern Recognition Letters’, vol. 25, pp. 141– 154.
- [37] Snitkowska, E. and Kasprzak, W. (2006), ‘Independent Component Analysis of Textures in Angiography Images, Computational Imaging Vision’, vol. 32, pp. 367– 372.
- [38] Field, D.J. (1987), Relations Between the Statistics of Natural Images and the Response Properties of Cortical Cells, *J. Optical Soc. America*, vol. 12, no. 4, pp. 2370– 2393.
- [39] Chuen-Horng Lin, Rong-Tai Chen, Yung-Kuan Chan (2009), ‘A smart content-based image retrieval system based on colour and texture feature’, *Image and Vision Computing*, vol. 27, 658–665.
- [40] Jun Yue, Zhenbo Li, Lu Liu and Zetian Fu(2011), ‘Content-based image retrieval using colour and texture fused features’, *Mathematical and Computer Modelling*, vol. 54, pp. 1121– 1127.
- [41] Xiang-Yang Wang & Bei-Bei Zhang & Hong-Ying Yang, ‘Content-based image retrieval by integrating colour and texture features’, *Multimed Tools Appl*, DOI 10.1007/s11042-012-1055-7.
- [42] H. Abrishami Moghaddam and M. Nikzad Dehaji, ‘ Enhanced Gabor wavelet correlogram feature for image indexing and retrieval’, *Pattern Anal Applic*.
- [43] Ela Yildizer, Ali Metin Balci, Tamer N. Jarada, Reda Alhadjj, ‘Integrating wavelets with clustering and indexing for effective content-based image retrieval’, *Knowledge-Based Systems* 31 (2012) 55–66.
- [44] Dengsheng Zhang - M. Monirul Islam - Guojun Lu and Ishrat Jahan Sumana, ‘ Rotation Invariant Curvelet Features for Region Based Image Retrieval’, *Int J Comput Vis* (2012) 98:187–201.
- [45] Ela Yildizer, Ali Metin Balci, Mohammad Hassan, Reda Alhadjj, ‘Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble’, *Expert Systems with Applications* 39 (2012) 2385–239

- [46] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux, 'Wavelet optimization for content-based image retrieval in medical databases', *Medical Image Analysis* 14 (2010) 227–241.
- [47] Liang-Hua Chen, Yao-Ling Hung, and Li-Yun Wang (2012), 'An Integrated Approach to Image Retrieval', *Telecommunications and Signal Processing (TSP)*, 2012 35th International Conference on, pp: 695 – 699, 2012.
- [48] Meng Fanjie, Guo Baolong and Wu Xianxiang (2012), 'Localized Image Retrieval Based on Interest Points', 2012 International Workshop on Information and Electronics Engineering (IWIEE), pp. 3371 – 3375.
- [49] H. Abrishami Moghaddam and M. Nikzad Dehaji (2013), 'Enhanced Gabor wavelet correlogram feature for image indexing and retrieval', *Pattern Analysis and Applications*, Vol. 16, issue 2, pp:163-177.
- [50] Teague, M. (1980), 'Image Analysis via the General Theory of Moments', *J. Optical Society America*, vol. 70, no. 8, pp. 920–930.
- [51] Hu, M. K. (1962), 'Visual Pattern Recognition by Moment Invariants', *IEEE Trans. Information Theory*, vol. 8, issue 2, pp. 179–187.
- [52] Luren, Y. and Fritz, A. (1994), 'Fast Computation of Invariant Geometric Moments: A New Method Giving Correct Results', *Proc. of IEEE Int. Conf. on Image Processing*.
- [53] Hew, P., *Geometric and Zernike Moments* (1996), 'Diary', Department of Mathematics, The University of Western Australia, 1996. [http://citeseer.ist.psu.edu/hew96\\_geometric.html](http://citeseer.ist.psu.edu/hew96_geometric.html).
- [54] Zhang, D.S. and Lu, G., 'Generic Fourier Descriptor for Shape-based Image Retrieval (2002)', *Proc. of IEEE Int. Conf. on Multimedia and Expo (ICME2002)*, Lausanne, Switzerland, vol. 1, pp. 425–428.
- [55] R. Krishnamoorthy, S. Sathiyadevi (2013), 'Image retrieval using edge based shape similarity with multiresolution enhanced orthogonal polynomials model', *Digital Signal Processing* vol. 23, 555–568.

- [56] Z. M. Ma, Gang Zhang and Li Yan(2011), ‘Shape feature descriptor using modified Zernike moments’, *Pattern Anal Applic* vol. 14, pp. 9–22
- [57] Anjali Goyal, Ekta Walia, ‘Variants of dense descriptors and Zernike moments as features for accurate shape-based image retrieval’, *SIViP* DOI 10.1007/s11760-012-0353-x.
- [58] Meng Fanjie, Guo Baolong, Wu Xianxiang (2012), ‘Localized Image Retrieval Based on Interest Points’, *Procedia Engineering*, vol. 29 pp. 3371–3375.
- [59] Flusser J (2005) Moment invariants in image analysis. *Enformatika* 11.
- [60] Kotoulas L, Andreadis I (2005) Image analysis using moments. *5th International Conference on Technology and Automation, Thessaloniki, Greece* 360–364.