

**ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

-- o0o--

NGUYỄN ĐỨC HÙNG

**NGHIÊN CỨU PHƯƠNG PHÁP TRA CỨU ẢNH CON
SỬ DỤNG KỸ THUẬT MÁY HỌC**

LUẬN VĂN THẠC SĨ : KHOA HỌC MÁY TÍNH

Thái Nguyên, tháng 12 năm 2013

**ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

-- o0o--

NGUYỄN ĐỨC HÙNG

**NGHIÊN CỨU PHƯƠNG PHÁP TRA CỨU ẢNH CON
SỬ DỤNG KỸ THUẬT MÁY HỌC**

Chuyên ngành : Khoa Học Máy Tính

Mã số : 60480101

LUẬN VĂN THẠC SĨ : CÔNG NGHỆ THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC

1 : TS. NGUYỄN HỮU QUỲNH



LỜI MỞ ĐẦU

Những năm gần đây, chúng ta đã chứng kiến sự tăng nhanh lượng dữ liệu ảnh cùng với sự phát triển bùng nổ của các ứng dụng trên Internet. Hàng ngày, nhiều cơ quan, tổ chức và cá nhân đã tạo ra nhiều Exa-bytes dữ liệu ảnh và đưa chúng lên Internet. Khi số lượng ảnh trong tập ảnh còn ít, việc nhận diện một bức ảnh hay việc so sánh sự giống và khác nhau giữa các bức ảnh có thể thực hiện được bằng mắt thường, tuy nhiên khi có số lượng rất lớn ảnh thì việc so sánh bằng mắt thường là rất khó khăn, đòi hỏi phải có những phương pháp hiệu quả và chính xác hơn. Việc tìm kiếm được một bức ảnh mong muốn trong hàng triệu bức ảnh thuộc đủ loại chủ đề khác nhau là rất khó khăn. Vấn đề đặt ra là phải có những phương pháp tổ chức cơ sở dữ liệu ảnh tốt cùng với những kỹ thuật tìm kiếm, tra cứu ảnh hiệu quả, có độ chính xác cao và có hiệu năng tốt.

Ban đầu, tra cứu ảnh được thực hiện dựa trên các mô tả ngắn bằng một tập các thuộc tính độc lập nội dung (tên file, khuôn dạng, loại, cỡ, tên tác giả và vị trí đĩa) của các ảnh. Tuy nhiên, cách tiếp cận này giới hạn các truy vấn vào các thuộc tính đã có. Một sự thay thế khác là sử dụng các chú thích văn bản thủ công hoặc các từ khóa sao cho các kỹ thuật tra cứu thông tin cổ điển có thể được sử dụng để tìm kiếm các ảnh. Nhưng cách tiếp cận này vẫn có các vấn đề như sự nhập nhằng, không đầy đủ và chủ quan. Do dữ liệu ảnh rất phong phú về thông tin, để thu được nội dung của một ảnh sử dụng một số ít từ khóa là không khả thi, chưa đề cập đến công việc tẻ nhạt trong quá trình chú thích.

Một cách tiếp cận tự động và hiệu quả hơn được gọi là tra cứu ảnh dựa vào nội dung (CBIR), trong CBIR sử dụng các đặc trưng ảnh mức thấp để biểu diễn, so sánh và tra cứu các ảnh. Hầu hết các hệ thống CBIR theo cách tiếp cận hai bước để tìm kiếm các cơ sở dữ liệu ảnh. Đầu tiên (đánh chỉ số), một véc tơ đặc trưng biểu diễn các thuộc tính quan trọng nào đó của ảnh được trích rút và lưu trữ như siêu dữ liệu cho mỗi ảnh cơ sở dữ liệu. Thứ hai (tìm kiếm), một ảnh truy vấn được cho, các ảnh tương tự nhất với ảnh truy vấn được trả về cho người sử dụng bằng việc so sánh các véc tơ đặc trưng của các ảnh cơ sở dữ liệu với các đặc trưng của ảnh truy vấn. Tất cả các hệ thống CBIR này thuộc về sơ đồ truy vấn bởi mẫu (QBE).

Trong khi hầu hết các hệ thống CBIR tra cứu các ảnh dựa trên so sánh toàn bộ ảnh, tức là với một ảnh truy vấn được cho hệ thống trả về tất cả các ảnh tương tự. Tuy nhiên, những người sử dụng có thể rất quan tâm đến tìm kiếm đối tượng. Trong trường hợp này, người sử dụng cung cấp một ảnh con truy vấn (có thể là một đối tượng) mà hệ thống tra cứu sẽ tìm kiếm các ảnh mà chứa ảnh con truy vấn (theo nhận thức của con người) từ cơ sở dữ liệu ảnh. Truy vấn ảnh con cũng có thể là bản thân một ảnh. Nhiệm vụ này, gọi là tra cứu ảnh con dựa vào nội dung (CBsIR), là khó giải quyết bởi một loạt các hiệu ứng (như thay đổi cỡ và khác biệt về các vị trí quan sát,...) gây ra ảnh con mục tiêu khác nhau đáng kể trong các ảnh khác nhau. Một bài toán kết hợp với CBsIR là cách xác định vị trí ảnh con bên trong một ảnh cơ sở dữ liệu hiệu quả.

Đề tài nghiên cứu tổng quan về tra cứu ảnh dựa vào nội dung và đi sâu vào nghiên cứu bài toán tra cứu ảnh con dựa vào nội dung (CBsIR) để tìm kiếm các ảnh cơ sở dữ liệu có chứa các ảnh con truy vấn. Bên cạnh đó, đề tài cũng kết hợp kỹ thuật máy học vào quá trình tra cứu ảnh con để được các kết quả gần với nhận thức của con người hơn. Trên những cơ sở phương pháp tra cứu ảnh con truy vấn, tiến hành xây dựng chương trình thử nghiệm cho phép đọc vào một ảnh con truy vấn và tìm kiếm những ảnh tương tự với ảnh con truy vấn trong một tập hợp các ảnh cho trước và thú nhận đánh giá phản hồi từ người dùng.

Nội dung luận văn gồm có 3 chương:

CHƯƠNG 1: TỔNG QUAN VỀ TRA CỨU ẢNH CON DỰA VÀO NỘI DUNG VÀ HỌC TỪ THÔNG TIN NGƯỜI DÙNG.

CHƯƠNG 2: TRA CỨU ẢNH CON DỰA VÀO NỘI DUNG SỬ DỤNG KỸ THUẬT MÁY HỌC, giới thiệu phương pháp tra cứu ảnh con dựa vào cây phân cấp và sự kết hợp giữa tra cứu ảnh con và kỹ thuật máy học.

CHƯƠNG 3: CHƯƠNG TRÌNH THỬ NGHIỆM, xây dựng hệ thống và trình bày một số kết quả đạt được.

LỜI CẢM ƠN

Trong quá trình thực hiện luận văn này, tôi luôn nhận được sự hướng dẫn, chỉ bảo tận tình của TS Nguyễn Hữu Quỳnh, Khoa Công nghệ Thông tin thuộc trường Đại học Điện lực là cán bộ trực tiếp hướng dẫn khoa học cho tôi. Thầy đã giành nhiều thời gian trong việc hướng dẫn cách nghiên cứu, đọc tài liệu, cài đặt các thuật toán và giúp đỡ về xây dựng hệ thống thực nghiệm.

Tôi xin chân thành cảm ơn các Thầy, Cô giáo trong trường Đại học Công nghệ Thông tin và Truyền Thông, Đại học Thái Nguyên đã luôn nhiệt tình giúp đỡ và tạo điều kiện tốt nhất cho tôi trong suốt quá trình học tập tại trường.

Xin chân thành cảm ơn các anh, các chị và các bạn học viên lớp Cao học - trường Đại học Công nghệ Thông tin và Truyền Thông thuộc Đại học Thái Nguyên đã luôn đồng viên, giúp đỡ và nhiệt tình chia sẻ với tôi những kinh nghiệm học tập, công tác trong suốt khoá học.

Hà Nội, ngày 25 tháng 12 năm 2013

Ý KIẾN NHẬN XÉT

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

MỤC LỤC

Lời mở đầu

Trang

Chương 1: TỔNG QUAN VỀ TRA CỨU ẢNH CON DỰA VÀO NỘI DUNG VÀ HỌC TỪ THÔNG TIN NGƯỜI DÙNG.....	10
1.1 Giới thiệu.....	10
1.1.1 Tra cứu ảnh dựa vào nội dung	10
1.1.2 Tra cứu ảnh con dựa vào nội dung	12
1.2 Tra cứu ảnh dựa vào màu.....	12
1.2.1 Các không gian màu	13
1.2.2 Biểu diễn và mô tả ảnh dựa vào màu.....	14
1.2.3 Hàm khoảng cách.....	15
1.3 Tra cứu thông tin.....	16
1.4 Học từ người dùng trong tra cứu thông tin	17
1.5 Một số nghiên cứu liên quan về tra cứu ảnh con dựa vào nội dung	21
1.5.1 Tra cứu ảnh con dựa vào phân hoạch	21
1.5.2 Tra cứu ảnh con dựa vào điểm.....	23
1.6 Kết luận chương 1	24
Chương 2 : TRA CỨU ẢNH CON DỰA VÀO NỘI DUNG SỬ DỤNG KỸ THUẬT MÁY HỌC.....	25
2.1 Giới thiệu bài toán.....	25
2.2 Kỹ thuật tra cứu ảnh con dựa vào nội dung	25
2.2.1 Cấu trúc cây và phân hoạch phân cấp của phương pháp HTM	26
2.2.2 Trích rút đặc trưng sử dụng màu trung bình.....	28
2.2.3 Các hàm khoảng cách véc tơ	28
2.2.4 Chiến lược tìm kiếm	31
2.3 Học trong tra cứu ảnh con dựa vào nội dung	32
2.3.1 Lược đồ đánh lại trọng số	33
2.3.2 Cập nhật trọng số	36
2.3.3 Tương tự ảnh.....	37
2.4 Kết luận chương 2	39
Chương 3: CHƯƠNG TRÌNH THỬ NGHIỆM	40
3.1. Biểu đồ hệ thống	40
3.1.1. Sơ đồ hệ thống tra cứu ảnh sử dụng phản hồi liên quan.....	40
3.1.2. Biểu đồ Use Case.....	41
3.1.3. Biểu đồ trình tự và biểu đồ hoạt động	41
3.1.4 Thiết kế cơ sở dữ liệu	45
3.2 Mô tả chương trình.....	46
3.3 Kết quả đạt được	47
3.4 Thực nghiệm thu hồi ảnh bằng thông tin phản hồi	49
3.5 Kết luận chương 3	51
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....	52
TÀI LIỆU THAM KHẢO	53

DANH MỤC CÁC BẢNG, SƠ ĐỒ, HÌNH

Hình 1.1: Kiến trúc hệ thống tra cứu ảnh dựa vào nội dung	9
Hình 1.2: Các không gian màu.....	11
Hình 1.2.1: Phản hồi liên quan tìm kiếm trên các ảnh.	16
Hình 1.2.2: Truy vấn tối ưu Rocchio để tách các tài liệu liên quan.	17
Hình 1.2.3: Ứng dụng của thuật toán Rocchio's.	19
Hình 2.1: Phân hoạch phân cấp của một ảnh với cấu trúc cây kết quả và chuỗi chỉ số tương ứng cho lưu trữ.....	25
Hình 2.2: Ví dụ về các cấu trúc cây cho ảnh cơ sở dữ liệu và ảnh con truy vấn.	28
Hình 2.3: Thuật toán HTM.....	30
Hình 2.4: Hình ảnh thu hồi với thông tin phản hồi của người dùng.	33
Hình 2.5: So sánh độ bất lợi của ảnh.....	34
Hình 3.1: Sơ đồ hệ thống tra cứu ảnh sử dụng phản hồi.....	38
Hình 3.2: Biểu đồ Use Case.	39
Hình 3.3: Biểu đồ trình tự tác nhân tra cứu ảnh.	39
Hình 3.4: Biểu đồ hoạt động của tác nhân tra cứu ảnh.....	40
Hình 3.5: Biểu đồ trình tự tác nhân phản hồi ảnh.	41
Hình 3.6: Biểu đồ hoạt động của tác nhân phản hồi ảnh.....	41
Hình 3.7: Biểu đồ trình tự của tác nhân Upload ảnh.	42
Hình 3.8: Biểu đồ hoạt động của tác nhân Upload ảnh.....	43
Hình 3.9 : Bảng cơ sở dữ liệu Image.....	43
Hình 3.10 : Bảng cơ sở dữ liệu Cluster	44
Hình 3.11: Mối quan hệ giữa các bảng.....	44
Hình 3.12: Số liệu các loại ảnh trong cơ sở dữ liệu.	45
Hình 3.13: Giao diện chính hệ thống tra cứu ảnh.....	46
Hình 3.14: Giao diện hiển thị kết quả truy vấn.	47
Hình 3.15: Hình ảnh truy vấn.	47
Hình 3.16: Hình ảnh thu hồi lần lặp thứ nhất.....	48
Hình 3.17: Hình ảnh thu hồi lần lặp thứ hai.....	48

KÍ HIỆU CÁC CỤM TỪ VIẾT TẮT

CBIR	Content Base Image Retrieval	Tra cứu ảnh dựa vào nội dung
QBIC	Query By Image Content	Truy vấn theo nội dung ảnh
CIE	Commission internationale de l'éclairage	Ủy ban quốc tế về màu sắc
HAC	Hierarchical Agglomerative Clustering	Phân cụm tích lũy phân cấp
CBC	Color Base Clustering	Phân cụm dựa vào màu
MST	Minimum Spanning Tree	Cây mở rộng tối thiểu
IRM	Integrated Region Matching	Đối sánh vùng tích hợp
CSDL	DataBase	Cơ Sở Dữ Liệu

Chương 1: TỔNG QUAN VỀ TRA CỨU ẢNH CON DỰA VÀO NỘI DUNG VÀ HỌC TỪ THÔNG TIN NGƯỜI DÙNG

1.1 Giới thiệu

1.1.1 Tra cứu ảnh dựa vào nội dung

Khi ta có nhu cầu tìm kiếm một vài bức ảnh trong một kho dữ liệu ảnh có thể lên tới vài trăm nghìn bức ảnh để minh họa cho một đề tài nào đó thì tuyệt đối không phải là chuyện đơn giản nếu chúng ta tìm kiếm một cách thủ công tức là xem lần lượt từng bức ảnh cho tới khi ta tìm thấy được bức ảnh có nội dung cần tìm. Song song với sự phát triển của các phương tiện kỹ thuật số trong tương lai, số lượng ảnh sẽ còn tăng nhiều hơn nữa. Do vậy nhu cầu thật sự đòi hỏi phải có một công cụ hỗ trợ tìm kiếm ảnh chính xác và hiệu quả. Vì vậy tra cứu ảnh dựa vào nội dung ra đời để góp phần đáp ứng nhu cầu này.

“Tra cứu ảnh dựa vào nội dung” - Đây là một chủ đề nghiên cứu mới trong lĩnh vực công nghệ thông tin. Mục đích chính của nó là lấy những ảnh từ cơ sở dữ liệu phù hợp với tiêu chí truy vấn. Các yếu tố mô tả nội dung của một bức ảnh có liên quan đến cảm nhận như màu sắc, kết cấu, hình dạng, cấu trúc, mối liên hệ về không gian và chuyển động. Do vậy, phân tích ảnh, nhận dạng ảnh và thị giác máy tính đóng vai trò cơ bản trong các hệ thống tra cứu ảnh. Nó cho phép tự động trích chọn hầu hết các thông tin cảm nhận, thông qua việc phân tích phân bố điểm ảnh và rút ra các độ đo nội dung trực quan. Các hệ thống truy vấn ảnh dựa vào nội dung hiện nay rất đa dạng nhưng nhìn chung được phân biệt bởi: các đặc trưng mà hệ thống rút trích từ ảnh để làm cơ sở truy vấn, phương pháp trích rút đặc trưng ảnh được sử dụng trong hệ thống truy vấn, độ đo sự tương tự giữa hai ảnh, phương pháp đánh chỉ số nhiều chiều để tối ưu việc tìm kiếm.

Một số chức năng chính của một hệ thống tra cứu ảnh dựa vào nội dung:

- Phân tích nội dung của nguồn thông tin và biểu diễn nội dung của các nguồn thông tin được phân tích phù hợp với sự đối sánh truy vấn của người sử dụng (không gian của thông tin nguồn được chuyển đổi thành không gian đặc điểm với mục đích đối sánh nhanh trong bước tiếp theo). Bước này thường mất nhiều thời gian cho việc xử lý tuần tự các thông tin nguồn (ảnh) trong cơ sở dữ liệu. Nó chỉ phải làm một lần và có thể làm độc lập.