

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

NGUYỄN VĂN TOÀN

**NGHIÊN CỨU PHƯƠNG PHÁP TRA CỨU ẢNH
SỬ DỤNG PHÂN CỤM GIA TĂNG
VỚI PHẢN HỒI LIÊN QUAN**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN, 2018

LỜI CẢM ƠN

Luận văn này được hoàn thành với sự hướng dẫn tận tình của PGS.TS Nguyễn Hữu Quỳnh – Khoa Công nghệ thông tin - Đại học Điện lực. Trước tiên tôi xin chân thành bày tỏ lòng biết ơn sâu sắc tới PGS.TS Nguyễn Hữu Quỳnh người đã tận tình hướng dẫn, động viên giúp đỡ tôi trong suốt thời gian thực hiện luận văn. Tôi cũng xin chân thành cảm ơn các thầy cô trong trường Đại học Công Nghệ thông tin và Truyền thông – Đại học Thái Nguyên, tạo điều kiện thuận lợi cho tôi hoàn thành tốt khóa học.

Xin chân thành cảm ơn các anh, các chị và các bạn học viên lớp Cao học CHK15A đã luôn động viên, giúp đỡ và nhiệt tình chia sẻ với tôi những kinh nghiệm học tập, công tác trong suốt khoá học.

Cuối cùng, tôi xin gửi lời cảm ơn sâu sắc đến gia đình, người thân, bạn bè đã động viên, khuyến khích và hỗ trợ cần thiết để tôi hoàn thành luận văn này.

Mặc dù rất cố gắng, song luận văn này không thể tránh khỏi những thiếu sót, kính mong được sự chỉ dẫn của các quý thầy cô và các bạn.

Thái Nguyên, ngày 2 tháng 5 năm 2018

Người viết

Nguyễn Văn Toàn

LỜI CAM ĐOAN

Tôi xin cam đoan rằng số liệu và kết quả nghiên cứu trong luận văn này là trung thực và không trùng lặp với các đề tài khác. Tôi cũng 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.

Thái Nguyên, ngày 2 tháng 5 năm 2018

Người cam đoan

Nguyễn Văn Toàn

MỤC LỤC

LỜI CẢM ƠN	i
LỜI CAM ĐOAN.....	iii
MỤC LỤC.....	iv
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT.....	vi
DANH MỤC CÁC BẢNG BIỂU	vii
DANH MỤC CÁC HÌNH.....	viii
MỞ ĐẦU.....	1
CHƯƠNG 1: TỔNG QUAN VỀ TRA CỨU ẢNH DỰA VÀO NỘI DUNG VỚI PHẢN HỒI LIÊN QUAN	4
1.1. Tổng quan về tra cứu ảnh dựa vào nội dung.....	4
1.1.1. Vấn đề tra cứu ảnh dựa vào nội dung	4
1.2. Tra cứu ảnh sử dụng phản hồi liên quan.....	14
1.3. Vấn đề phân cụm[1]	16
1.3.1. Thuật toán K-Means.....	20
1.3.2. Phân cụm phân cấp	21
1.3.3. Phân cụm dựa vào mật độ	23
1.3.4. Phân cụm dựa vào mô hình.....	23
1.3.5. Phân cụm dựa vào đồ thị.....	26
1.4. Tiểu kết chương 1.	26
CHƯƠNG 2: PHƯƠNG PHÁP TRA CỨU ẢNH VỚI PHẢN HỒI LIÊN QUAN SỬ DỤNG PHÂN CỤM GIA TĂNG.....	27
2.1. Tra cứu ảnh với ngữ nghĩa mức cao	27
2.1.1. Giới thiệu về tra cứu ảnh với ngữ nghĩa mức cao.....	27
2.1.2. Khoảng cách ngữ nghĩa.....	28
2.1.3. Phản hồi liên quan.....	29
2.2. Tra cứu ảnh với phản hồi liên quan	31
2.3. Kỹ thuật phân tích phân biệt tuyến tính (LDA-Linear Discriminant Analysis).....	32

2.3.1. Định nghĩa về LDA.....	32
2.3.2 Tính toán phương sai between-class (SB).....	32
2.3.3 Tính phương sai within-class (Sw).....	34
2.3.4 Xây dựng không gian thấp chiều	36
2.3.5. Sơ đồ phương pháp tra cứu ảnh sử dụng phân cụm gia tăng trong phản hồi liên quan.....	37
2.4. Tiểu kết chương 2.....	39
CHƯƠNG 3: CHƯƠNG TRÌNH THỬ NGHIỆM	40
3.1. Giới thiệu bài toán tra cứu ảnh dựa vào nội dung.....	40
3.2. Môi trường thực nghiệm.	41
3.2.1. Cơ sở dữ liệu ảnh.	42
3.2.2. Vec-tơ đặc trưng.....	43
3.2.3. Tập tin cây nền.....	44
3.2.4. Cấu hình đề xuất thiết bị chạy thực nghiệm	44
3.3. Đánh giá kết quả thực nghiệm.	44
3.3.1. Chiến lược mô phỏng phản hồi liên quan.	44
3.3.2. Kết quả đánh giá.....	45
3.4. Giao diện hệ thống	47
3.5. Tiểu kết chương 3.	51
KẾT LUẬN	52
TÀI LIỆU THAM KHẢO.....	53

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

Từ viết tắt	Diễn giải
CBIR	Tra cứu ảnh dựa vào nội dung.
RBIR	Tra cứu ảnh dựa vào vùng
CSDL	Cơ sở dữ liệu
CRT	Mẫu vùng cấu tạo
EMD	Earth Mover Distance
IRM	Lược đồ đối sánh vùng tích hợp
DPF	Hàm khoảng cách động một phần
MRMD	Khoảng cách đa tạp đa phân giải
RF	Phản hồi liên quan

DANH MỤC CÁC BẢNG BIỂU

Bảng 3.1. Bảng phân bố tập ảnh Corel.....	42
Bảng 3.2. Các loại đặc trưng.....	43
Bảng 3.3. Bảng cấu hình đề xuất thiết bị chạy thực nghiệm.	44
Bảng 3.4. Bảng kết quả của các phương pháp	46

DANH MỤC CÁC HÌNH

Hình 1.1. Kiến trúc tổng quan của hệ thống tra cứu ảnh dựa vào nội dung.	5
Hình 1.2. Không gian màu RGB.	8
Hình 1.3. Không gian màu HSV.	9
Hình 1.4. Lược đồ của một hệ thống CBIR với RF.	16
Hình 1.5. Các tập dữ liệu và các cụm	17
Hình 1.6. Các tập dữ liệu không thích hợp với K-Means.	21
Hình 1.7 Phân cụm phân cấp tập dữ liệu $D=\{a,b,c,d,e\}$	22
Hình 2.1. Dịch chuyển điểm truy vấn.	29
Hình 2.2. Hình dạng lõi (đa điểm).	30
Hình 2.3. Hình dạng lõm (đa điểm).	30
Hình 2.4. Tra cứu ảnh dựa vào nội dung với phản hồi liên quan.	31
Hình 2.5. Các bước được trực quan hóa để tính một không gian con chiều thấp hơn của kỹ thuật LDA.	33
Hình 2.6. Sơ đồ tra cứu ảnh sử dụng phân cụm gia tăng.	38
Hình 3.1. Mô hình tổng quát của hệ thống.	41
Hình 3.2. Biểu đồ so sánh kết quả thực nghiệm	46
Hình 3.3. Giao diện chính của hệ thống.	47
Hình 3.4. Chọn tập dữ liệu ảnh / đặc trưng.	48
Hình 3.5. Chọn ảnh truy vấn khởi tạo.	48
Hình 3.6. Tra cứu với truy vấn khởi tạo 84003, 84004, 84008 thuộc lớp 840. ..	48
Hình 3.7. Kết quả phân cụm tập huấn luyện.	49
Hình 3.8. Công cụ tra cứu và phân cụm LDA.	49
Hình 3.9. Kết quả tra cứu phản hồi liên quan.	50
Hình 3.10. Kết quả phân cụm gia tăng.	50
Hình 3.11. Kết quả tra cứu sau khi sử dụng phân cụm gia tăng.	51

MỞ ĐẦU

Tra cứu ảnh dựa vào nội dung (CBIR-Content Based Image Retrieval) đã nhận được nhiều sự quan tâm trong thập kỷ qua, do nhu cầu xử lý hiệu quả lượng dữ liệu đa phương tiện khổng lồ và tăng nhanh chóng. Nhiều hệ thống CBIR đã được phát triển, gồm QBIC, Photobook, MARS, VisualSEEK, SIMPLIcity và những hệ thống khác. Trong một hệ thống CBIR tiêu biểu, các đặc trưng ảnh trực quan mức thấp (tức là màu, kết cấu và hình dạng) được trích rút tự động cho mục tiêu đánh chỉ số và mô tả ảnh. Để tìm kiếm các ảnh mong muốn, người dùng đưa một ảnh làm mẫu và hệ thống trả lại một tập các ảnh tương tự dựa vào các đặc trưng được trích rút.

Là một vấn đề quan trọng trong CBIR, độ đo tương tự lượng hóa sự giống nhau về nội dung giữa từng cặp ảnh. Phụ thuộc vào kiểu đặc trưng được trích rút mà chúng ta lựa chọn độ đo tương tự thích hợp. Tất cả các kỹ thuật tra cứu ảnh dựa vào nội dung hiện nay đều thừa nhận thông tin tương hỗ giữa độ đo tương tự ảnh và ngữ nghĩa của ảnh. Bằng nhiều cách khác nhau, độ đo tương tự cố gắng nắm được một khía cạnh nào đó của nội dung ảnh, đó là ngữ nghĩa kế thừa từ độ tương tự hay đặc trưng mức thấp. Tuy nhiên, ngữ nghĩa kế thừa từ độ tương tự nhiều khi không giống với khái niệm mức cao được truyền tải bởi một ảnh (ngữ nghĩa của ảnh). Đó chính là khoảng cách ngữ nghĩa, nó phản ánh sự khác biệt giữa năng lực mô tả hạn chế của đặc trưng trực quan mức thấp và khái niệm mức cao. Các kỹ thuật trong việc rút ngắn “khoảng cách ngữ nghĩa” gồm có 5 loại chính: (1) sử dụng bản thể đối tượng để xác định các khái niệm mức cao, (2) sử dụng các công cụ học máy để kết hợp các đặc trưng mức thấp với các khái niệm truy vấn, (3) đưa phản hồi liên quan vào lặp tra cứu cho học ý định của người dùng, (4) sinh ra mẫu ngữ nghĩa để hỗ trợ tra cứu ảnh mức cao, (5) Cách sử dụng cả nội dung trực quan của các ảnh và thông tin văn bản thu được từ Web cho tra cứu ảnh trên Web.

Từ những nhận định trên và được sự gợi ý của giáo viên hướng dẫn, tôi quyết định chọn đề tài: **“Nghiên cứu phương pháp tra cứu ảnh sử dụng phân cụm gia tăng với phản hồi liên quan”**. Đề tài sẽ kết hợp hai hướng tiếp cận (2) và (3), đưa phản hồi liên quan của người dùng vào quá trình tra cứu và sử dụng phương pháp phân cụm gia tăng để phân cụm tập ảnh phản hồi nhằm biểu diễn nhu cầu thông tin người dùng hiệu quả.

Phản hồi liên quan là một quá trình trực tuyến mà cố gắng học mục đích của người dùng trong quá trình tra cứu, là một công cụ mạnh được sử dụng truyền thống trong các hệ thống tra cứu thông tin. Nó được giới thiệu đối với CBIR khoảng đầu những năm 1990, với mục đích mang người dùng vào lặp tra cứu để giảm khoảng cách ngữ nghĩa giữa những gì mà truy vấn biểu diễn và những gì người dùng nghĩ. Bằng việc tiếp tục học thông qua tương tác với các người dùng cuối, phản hồi liên quan đã được chỉ ra là cung cấp cải tiến hiệu năng đáng kể trong các hệ thống CBIR.

Phân cụm là một phương pháp học không giám sát để tạo thành các nhóm hay các cụm dữ liệu. Lý thuyết phân cụm giả thuyết rằng “các đối tượng gần nhau có xu hướng liên quan tới cùng một yêu cầu”. Đã có nhiều thuật toán thực hiện việc phân cụm như: K-mean, K-medoid, EM...Tuy nhiên, các thuật toán này thường được gọi là phân cụm ngoại tuyến (off-line), tức là, các thuật toán này thực hiện phân cụm trên toàn bộ cơ sở dữ liệu ảnh đã có sẵn (gồm rất nhiều ảnh), mỗi khi có ảnh mới bổ sung vào, quá trình lại phải phân cụm lại từ đầu. Các thuật toán ngoại tuyến không phù hợp trong các trường hợp đòi hỏi trực tuyến (on-line), chẳng hạn, trường hợp mà áp dụng trên một tập ảnh nhỏ (là kết quả của một lần thực hiện tra cứu) nhưng đòi hỏi phân cụm ngay lập tức trong khi vẫn còn nhiều ảnh cần được bổ sung và phân cụm tiếp theo không cần phải tiến hành với dữ liệu đã được phân cụm trước đó. Thuật toán mà đáp ứng trường hợp trực tuyến này phải có tính chất “gia tăng” hay gọi là *phân cụm gia tăng*.