

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

HOÀNG HÀ HẢI NAM

**NGHIÊN CỨU KỸ THUẬT SIFT TRONG TRÍCH TRỌN
ĐẶC TRƯNG, XÂY DỰNG ỨNG DỤNG TÌM KIẾM ẢNH**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên, 2018

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

HOÀNG HÀ HẢI NAM

**NGHIÊN CỨU KỸ THUẬT SIFT TRONG TRÍCH TRỌN
ĐẶC TRƯNG, XÂY DỰNG ỨNG DỤNG TÌM KIẾM ẢNH**

Chuyên ngành: Khoa học máy tính

Mã số: 8 480 101

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Người hướng dẫn khoa học: TS. Vũ Vinh Quang

Thái Nguyên, 2018

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 kì 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.

Thái Nguyên, ngày 26 tháng 6 năm 2018

Tác giả luận văn

Hoàng Hà Hải Nam

LỜI CẢM ƠN

Trước hết, với lòng kính trọng và biết ơn sâu sắc, tôi xin gửi lời cảm ơn tới Tiến sĩ Vũ Vinh Quang – người đã tận tình hướng dẫn tôi trong suốt quá trình tìm hiểu, nghiên cứu và hoàn thiện luận văn.

Tôi xin gửi lời cảm ơn chân thành tới Ban giám hiệu nhà trường và các thầy cô đã trực tiếp giảng dạy, giúp đỡ tôi trong suốt thời gian học tập, nghiên cứu khoa học.

Cuối cùng, tôi xin gửi lời cảm ơn đến gia đình, người thân, bạn bè, đồng nghiệp đã giúp đỡ, tạo mọi điều kiện thuận lợi cho tôi trong suốt thời gian học tập và hoàn thiện luận văn này.

Thái Nguyên, ngày 26 tháng 6 năm 2018

Tác giả luận văn

Hoàng Hà Hải Nam

DANH MỤC CÁC HÌNH ẢNH

Hình 1.1: Kiến trúc chung của một MMDBMS.....	10
Hình 1.2: Mô hình hệ thống Tra cứu ảnh theo nội dung.....	16
Hình 2.1: Ví dụ hiển thị một ảnh.....	21
Hình 2.2: Truy vấn của Google “5D3”	23
Hình 2.3: Truy vấn của Google “Apple”	23
Hình 2.4: Một số loại kết cấu	24
Hình 2.5: Một kết quả trả về của Google Image.....	25
Hình 2.6: Một kết quả trả về của Bing.....	25
Hình 2.7: Một kết quả trả về của Flickr Images Search.....	26
Hình 2.8: Biểu đồ mô phỏng việc tính toán các DoG ảnh từ các ảnh kê mờ.....	35
Hình 2.9: Mỗi điểm ảnh được so sánh với 26 láng giềng của nó.....	36
Hình 2.10: Số lượng mẫu tỷ lệ trên mỗi Octave	36
Hình 2.11: Quá trình lựa chọn các điểm hấp dẫn.....	38
Hình 2.12: Biểu diễn các vector đặc trưng.....	41
Bảng 2.13: Một số phương pháp lựa chọn đặc trưng.....	42
Hình 2.14: Mô hình hệ thống IVFADCj; Hệ thống bên trái: chèn một vector vào danh sách chỉ mục ngược; hệ thống bên phải: tìm kiếm k láng giềng gần nhất.	50
Hình 2.15: Mô hình giải quyết bài toán	52
Hình 3.1: Tập dữ liệu ảnh thử nghiệm	55
Hình 3.2: Người dùng chọn ảnh truy vấn.....	56
Hình 3.3: Đặc trưng SIFT của ảnh được trích chọn.....	57
Hình 3.4: 20 ảnh tương đồng với ảnh truy vấn	57
Hình 3.5: Kết quả trả về với truy vấn iPhoneX.....	58
Hình 3.6: Kết quả trả về với truy vấn Apple.....	58
Hình 3.7: Một số kết quả truy vấn khác	59

DANH MỤC CÁC BẢNG BIỂU

Bảng 1.1: Mô tả kích thước kiểu dữ liệu.....	6
Bảng 3.1: Cấu hình phần cứng sử dụng trong thực nghiệm.....	54
Bảng 3.2: Công cụ phần mềm sử dụng trong thực nghiệm.....	54
Bảng 3.3: Một số thư viện sử dụng trong thực nghiệm	55
Bảng 3.4: Kết quả độ chính xác trung bình của 10 truy vấn.....	60

DANH MỤC CÁC TỪ VIẾT TẮT

Từ viết tắt	Tiếng Anh	Tiếng Việt
AI	Artificial intelligence	Trí tuệ nhân tạo
AR	Augmented reality	Tương tác thực tại ảo
CSDL	Database	Cơ sở dữ liệu
DBMS	Database Management System	Hệ quản trị cơ sở dữ liệu
IoT	Internet of things	Internet vạn vật
IR	Information Retrieval	Hệ thống tự động truy tìm thông tin
MIRS	Multimedia Indexing & Retrieval System	Hệ thống chỉ mục và truy tìm thông tin đa phương tiện
MMDBMS	Multimedia Database Management System	Hệ thống quản trị cơ sở dữ liệu đa phương tiện
MRI	Magnetic Resonance Imaging	
VR	Virtual reality	Thực tế ảo
WWW	World Wide Web	

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
DANH MỤC CÁC HÌNH ẢNH	iii
DANH MỤC CÁC BẢNG BIỂU	iv
DANH MỤC CÁC TỪ VIẾT TẮT	v
MỤC LỤC	vi
MỞ ĐẦU	1
CHƯƠNG 1. TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU ĐA PHƯƠNG TIỆN	2
1.1. Tổng quan về hệ quản trị cơ sở dữ liệu đa phương tiện.....	2
1.1.1. Giới thiệu chung.....	2
1.1.2. Các kiểu truyền thông và multimedia	5
1.1.3. Hệ quản trị cơ sở dữ liệu đa phương tiện (MMDBMS)	7
1.1.4. Truy tìm thông tin tài liệu văn bản.....	8
1.1.5. Truy xuất và chỉ số hoá multimedia.....	9
1.2. Kiến trúc tổ chức nội dung của một hệ thống cơ sở dữ liệu Multimedia.....	10
1.2.1. Kiến trúc hệ thống quản trị cơ sở dữ liệu đa phương tiện	10
1.2.2. Tổ chức dữ liệu đa phương tiện trên cơ sở nguyên lý thống nhất.....	12
1.2.3. Cấu trúc tóm tắt media.....	14
1.3. Tổng quan về tra cứu ảnh.....	15
1.3.1. Mô hình hệ thống tra cứu ảnh theo nội dung	16
1.3.2. Một số hệ thống tra cứu ảnh theo nội dung tiêu biểu.....	17
CHƯƠNG 2. TRÍCH CHỌN ĐẶC TRƯNG ẢNH	20
2.1. Tổng quan về trích chọn đặc trưng ảnh.....	20
2.1.1. Đặc trưng văn bản đi kèm ảnh và tìm kiếm ảnh theo văn bản đi kèm ảnh	20
2.1.2. Đặc trưng nội dung ảnh và tìm kiếm theo đặc trưng nội dung.....	23
2.2. Các phương pháp trích chọn đặc trưng ảnh và độ đo tương đồng giữa các ảnh	26
2.2.1. Đặc trưng màu sắc.....	27
2.2.2. Đặc trưng kết cấu	29
2.2.3. Đặc trưng hình dạng.....	30

2.2.4. Đặc trưng cục bộ bất biến SIFT	31
2.2.5. Lựa chọn đặc trưng	42
2.3. Mô hình k láng giềng gần nhất sử dụng bộ lượng tử hóa	44
2.3.1. Cơ sở lý thuyết	45
2.3.2. Thuật toán K láng giềng	50
2.3.3. Mô hình bài toán	51
CHƯƠNG 3. THỰC NGHIỆM VÀ ĐÁNH GIÁ	54
3.1. Bài toán truy vấn ảnh	54
3.1.1. Phát biểu bài toán	54
3.1.2. Cách giải quyết bài toán.....	54
3.2. Môi trường và các công cụ sử dụng cho thực nghiệm	54
3.2.1. Cấu hình phần cứng	54
3.2.2. Công cụ phần mềm sử dụng	54
3.2.3. Thư viện sử dụng.....	55
3.3. Xây dựng tập dữ liệu ảnh	55
3.4. Quy trình, phương pháp thực nghiệm	56
3.5. Kết quả thực nghiệm	59
KẾT LUẬN	62
TÀI LIỆU THAM KHẢO	63

MỞ ĐẦU

Cuộc cách mạng công nghiệp lần thứ tư diễn ra từ những năm 2000 được gọi là cuộc cách mạng số, thông qua các công nghệ như Internet vạn vật (IoT), trí tuệ nhân tạo (AI), thực tế ảo (VR), tương tác thực tại ảo (AR), mạng xã hội, điện toán đám mây, di động, phân tích dữ liệu lớn... để chuyển hóa toàn bộ thế giới thực thành thế giới số.

Sự phát triển mạnh mẽ của công nghệ ảnh số đã làm lượng ảnh lưu trữ tăng lên một cách nhanh chóng đòi hỏi phải có các công cụ hỗ trợ tìm kiếm ảnh hiệu quả và tiện lợi. Mặc dù các công cụ tìm kiếm ảnh theo văn bản đi kèm ảnh ra đời cho phép người dùng tìm kiếm ảnh với thời gian đáp ứng khá nhanh, tuy nhiên, các công cụ này vẫn còn hạn chế trong việc giải quyết sự không thống nhất giữa nội dung câu truy vấn và nội dung hiển thị của ảnh trả về. Sự ra đời của các công cụ tìm kiếm ảnh theo nội dung ảnh đã giải quyết được những hạn chế nêu trên. Từ những lí do trên, tôi chọn đề tài *Nghiên cứu kĩ thuật SIFT trong trích chọn đặc trưng, xây dựng ứng dụng tìm kiếm ảnh*.

Mục tiêu của luận văn là nghiên cứu đặc trưng SIFT để tối ưu hóa chất lượng tìm kiếm ảnh. Nội dung chính của luận văn tập trung khảo sát phương pháp trích chọn đặc trưng ảnh trong tìm kiếm và xếp hạng ảnh. Trên cơ sở ứng dụng phương pháp lượng tử hóa tích của Hervé Jégou và cộng sự [3], luận văn nghiên cứu mô hình tìm kiếm K láng giềng gần nhất kết hợp độ đo tương đồng về khoảng cách giữa các vector đặc trưng, từ đó tiến hành xây dựng mô hình thực nghiệm tìm kiếm ảnh sản phẩm trên phần mềm Matlab.

Luận văn bao gồm các nội dung sau:

Chương 1: Tổng quan về cơ sở dữ liệu đa phương tiện

Chương 2: Trích chọn đặc trưng ảnh

Chương 3: Thực nghiệm và đánh giá