

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

**NGUYỄN HỒNG HẠNH**

**THUẬT TOÁN PHÁT HIỆN ẢNH GIẢ MẠO  
DẠNG CẮT - DÁN DỰA TRÊN CÁC PHÉP  
BIẾN ĐỔI MA TRẬN**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

**THÁI NGUYÊN - 2014**

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

NGUYỄN HỒNG HẠNH

**THUẬT TOÁN PHÁT HIỆN ẢNH GIẢ MẠO  
DẠNG CẮT - DẪN DỰA TRÊN CÁC PHÉP  
BIẾN ĐỔI MA TRẬN**

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã số: 60 48 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

HƯỚNG DẪN KHOA HỌC: PGS.TS PHẠM VĂN ẮT

THÁI NGUYÊN - 2014

## LỜI CAM ĐOAN

Tôi xin cam đoan luận văn thạc sĩ khoa học máy tính: ***“Thuật toán phát hiện ảnh giả mạo dạng cắt - dựa trên các phép biến đổi ma trận”*** là kết quả của quá trình học tập, nghiên cứu khoa học độc lập, nghiêm túc.

Các số liệu trong luận văn là trung thực, có nguồn gốc rõ ràng, được trích dẫn và có tính kế thừa, phát triển từ các tài liệu, tạp chí, các công trình nghiên cứu đã được công bố, các website, ...

Các phương pháp nêu trong luận văn được rút ra từ những cơ sở lý luận và quá trình nghiên cứu tìm hiểu của tác giả.

*Thái Nguyên, tháng    năm 2014*

Tác giả

**Nguyễn Hồng Hạnh**

## LỜI CẢM ƠN

*Lời đầu tiên tôi xin gửi lời cảm ơn thầy PGS. TS. Phạm Văn Ất đã tận tình giúp đỡ tôi trong suốt thời gian học tập vừa qua, đặc biệt là định hướng nghiên cứu và tận tình hướng dẫn cho tôi trong suốt quá trình làm luận văn.*

*Nếu không có sự giúp đỡ tận tình của thầy, tôi khó có thể hoàn thành luận văn này.*

*Bên cạnh đó tôi cũng xin chân thành cảm ơn các thầy cô Trường Đại học Công nghệ Thông tin & Truyền thông đã tận tình giảng dạy, chỉ bảo và cung cấp cho tôi những kiến thức cần thiết trong suốt thời gian học và cũng xin gửi lời cảm ơn chân thành đến những người thân, bạn bè và đồng nghiệp đã giúp đỡ và động viên tôi trong suốt thời gian học tập cũng như trong thời gian thực hiện luận văn.*

*Chân thành cảm ơn !*

*Thái Nguyên, tháng năm 2014*

Tác giả

**Nguyễn Hồng Hạnh**

## MỤC LỤC

<b>LỜI CAM ĐOAN</b> .....	i
<b>LỜI CẢM ƠN</b> .....	ii
<b>MỤC LỤC</b> .....	iii
<b>DANH MỤC HÌNH VẼ</b> .....	iv
<b>MỞ ĐẦU</b> .....	1
<b>CHƯƠNG 1: TỔNG QUAN VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN PHÁT HIỆN ẢNH GIẢ MẠO</b> .....	3
1.1. TỔNG QUAN VỀ XỬ LÝ ẢNH.....	3
1.1.1. Xử lý ảnh là gì?.....	3
1.1.2. Các khái niệm cơ bản trong xử lý ảnh.....	9
1.1.3. Các vấn đề cơ bản trong xử lý ảnh .....	9
1.1.4. Một số định dạng ảnh phổ biến.....	12
1.2. BÀI TOÁN PHÁT HIỆN ẢNH GIẢ MẠO .....	17
1.2.1. Ảnh giả mạo.....	17
1.2.2. Các loại ảnh giả mạo cơ bản .....	17
1.2.3. Các cách tiếp cận trong phát hiện ảnh giả mạo .....	21
<b>CHƯƠNG 2: THUẬT TOÁN PHÁT HIỆN ẢNH GIẢ MẠO DẠNG CẮT - DÁN</b> .....	24
2.1. THUẬT TOÁN ĐỐI SÁNH CHÍNH XÁC (EXACT MATCH) PHÁT HIỆN ẢNH GIẢ MẠO DẠNG CẮT - DÁN.....	24
2.1.1. Ý tưởng.....	24
2.1.2. Thuật toán.....	25
2.1.3. Phân tích thuật toán .....	26
2.1.4. Ví dụ minh họa .....	26
2.2. THUẬT TOÁN ĐỐI SÁNH BỀN VỮNG (ROBUST MATCH) PHÁT HIỆN ẢNH GIẢ MẠO DẠNG CẮT - DÁN DỰA TRÊN PHÉP BIẾN ĐỔI COSIN RỜI RẠC (DCT - Discrete cosine transform).....	35

2.2.1. Ý tưởng.....	38
2.2.2. Thuật toán.....	39
2.2.3. Phân tích thuật toán.....	44
2.3. THUẬT TOÁN PHÁT HIỆN ẢNH GIẢ MẠO DẠNG CẮT - DÁN DỰA TRÊN PHÉP BIẾN ĐỔI DWT (Discrete Wavelet Transform) .....	45
2.3.1. Ý tưởng.....	48
2.3.2. Thuật toán.....	49
2.3.3. Phân tích thuật toán.....	54
2.4. SO SÁNH CÁC PHƯƠNG PHÁP PHÁT HIỆN ẢNH GIẢ MẠO DẠNG CẮT - DÁN.....	55
<b>CHƯƠNG 3: PHẦN MỀM THỬ NGHIỆM.....</b>	<b>56</b>
3.1. PHÂN TÍCH YÊU CẦU BÀI TOÁN .....	56
3.2. TẠO ẢNH GIẢ MẠO.....	56
3.3. SƠ ĐỒ CHUNG CỦA THUẬT TOÁN ĐỐI SÁNH CHÍNH XÁC - (exact match).....	57
3.4. SƠ ĐỒ CHUNG CỦA THUẬT TOÁN ĐỐI SÁNH BỀN VỮNG - (robust match).....	58
3.5. LỰA CHỌN CÔNG CỤ .....	59
3.6. MÔI TRƯỜNG CÀI ĐẶT VÀ XÂY DỰNG CÁC HÀM.....	60
3.7. KẾT QUẢ THỰC NGHIỆM.....	60
3.7.1. Phần mềm đối sánh bền vững dựa trên phép biến đổi DCT .....	61
3.7.2. Phần mềm đối sánh bền vững dựa trên phép biến đổi DWT.....	64
<b>KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....</b>	<b>67</b>
1. KẾT LUẬN.....	67
2. KIẾN NGHỊ.....	68
<b>TÀI LIỆU THAM KHẢO .....</b>	<b>69</b>

## DANH MỤC HÌNH VẼ

Hình 1.1. Các giai đoạn trong xử lý ảnh.....	4
Hình 1.2. Các thành phần chính trong hệ thống xử lý ảnh .....	8
Hình 1.3. Quá trình nén JPEG cơ .....	16
Hình 1.4. Minh họa về việc giả mạo ảnh .....	17
Hình 1.5. Minh họa cho loại ảnh giả mạo tăng cường ảnh .....	18
Hình 1.6. Ảnh che phủ và bỏ đi đối tượng .....	20
Hình 1.7 thể hiện một dạng khác thường thấy của giả mạo ảnh, đó là việc bổ sung thêm đối tượng. ....	21
Hình 1.8. Sơ đồ phát hiện ảnh giả mạo dựa vào cơ sở dữ liệu .....	22
Hình 1.9. Sự khác biệt của các hướng nguồn sáng khác nhau. ....	23
Hình 2.1. Đánh số thứ tự các phần tử của ma trận hệ số DCT theo đường zigzag .....	41
Hình 2.2. Vectơ dịch chuyển của vùng cắt - dán .....	44
Hình 2.3. Áp dụng phép biến đổi DWT theo hàng và cột .....	48
Hình 2.4. Áp dụng phép biến đổi DWT theo hai mức .....	48
Hình 2.5. Vectơ dịch chuyển của vùng cắt - dán .....	54
Hình 3.1 Kết quả thực nghiệm 1 phát hiện ảnh giả mạo dạng cắt - dán.....	61
Hình 3.2. Kết quả thực nghiệm 2 phát hiện ảnh giả mạo dạng cắt - dán.....	62
Hình 3.3. Kết quả thực nghiệm 3 phát hiện ảnh giả mạo dạng cắt - dán.....	63
Hình 3.4. Kết quả thực nghiệm 4 phát hiện ảnh giả mạo dạng cắt - dán.....	64
Hình 3.5. Kết quả thực nghiệm 5 phát hiện ảnh giả mạo dạng cắt – dán, ảnh giả nén JPEG với hệ số $Q=60$ .....	65
Hình 3.6. Kết quả thực nghiệm 6 phát hiện ảnh giả mạo dạng cắt - dán nhiễu 1% ....	66
Hình 3.7. Kết quả thực nghiệm 7 phát hiện ảnh giả mạo dạng cắt – dán nhiễu 1% làm mờ 1px nén JPEG $Q=60$ .....	66

## MỞ ĐẦU

Ngày nay, ảnh số là phương tiện truyền thông được sử dụng rộng rãi, đóng vai trò quan trọng trong đời sống con người, có tác động đến xã hội, tham gia vào các quá trình pháp lý và kinh tế như: làm bằng chứng trong điều tra, xử án, bảo hiểm, gian lận khoa học,... Hơn nữa, với sự phổ biến của máy ảnh kỹ thuật số và các phần mềm chỉnh sửa (Photoshop, GIMP,...) dẫn đến ảnh số có thể dễ dàng được chỉnh sửa mà không cần đến các kiến thức chuyên gia và việc chỉnh sửa hầu như không để lại dấu vết mà mắt thường có thể nhận biết được. Kết quả là khi những hình ảnh được chỉnh sửa sử dụng cho mục đích xấu nó có thể dẫn đến những hậu quả nghiêm trọng. Do vậy, việc phát hiện ảnh giả mạo là vấn đề đặt ra ngày càng cấp bách và càng trở lên khó khăn.

Mặc dù nhu cầu về việc phát hiện các giả mạo ảnh số đã được công nhận bởi cộng đồng các nhà nghiên cứu, nhưng hiện nay số tài liệu về lĩnh vực này vẫn chưa nhiều. Trong việc chống giả mạo ảnh, người ta đã nghiên cứu các kỹ thuật về tạo bản quyền ảnh trên cơ sở giấu các thông tin cần thiết vào các bức ảnh trước khi phát hành để tránh tình trạng sao chép bất hợp pháp hoặc để tiện cho việc phát hiện các sửa đổi hoặc cắt ghép sau này. Theo cách tiếp cận này, các thiết bị máy ảnh số và camera ngày nay thường bổ sung các tính năng bản quyền ngay trong quá trình thu nhận ảnh. Dựa vào đó sau này ta có thể biết được nguồn gốc của bức ảnh, biết được đó có phải là ảnh gốc hay không? Cách tiếp cận này dựa vào giả thiết rằng việc giả mạo ảnh sẽ làm thay đổi thông tin bản quyền. Tuy nhiên, hạn chế của cách tiếp cận này là thông tin bản quyền phải được chèn vào tại thời gian thu nhận ảnh nên chỉ giới hạn với các camera số được trang bị đặc biệt. Ngoài ra, một giải pháp khác không cần dùng đến các thông tin nhúng vào từ trước là các phương pháp thụ động phát hiện ảnh giả mạo.

Trong các loại giả mạo thì ảnh giả mạo dạng cắt - dán tương đối phổ biến và dễ dàng thực hiện, nên chúng đã và đang được nhiều nhà nghiên cứu quan tâm.



Nhận thức được ý nghĩa khoa học và tính thực tiễn của vấn đề này nên em đã chọn đề tài “*Thuật toán phát hiện ảnh giả mạo dạng cắt - dán dựa trên các phép biến đổi ma trận*” cho luận văn tốt nghiệp của mình.

\* Nội dung chính của luận văn gồm:

**- Phần mở đầu.**

***Chương 1: Tổng quan về xử lý ảnh và bài toán phát hiện ảnh giả mạo.***

Chương này trình bày tổng quan về một hệ thống xử lý ảnh, các vấn đề cơ bản trong xử lý ảnh, một số định dạng ảnh phổ biến như định dạng Bitmap, định dạng JPEG, đồng thời trình bày bài toán phát hiện ảnh giả mạo.

***Chương 2: Thuật toán phát hiện ảnh giả mạo dạng cắt - dán***

Chương này trình bày một số phép biến đổi ma trận, giới thiệu tổng quan về ảnh giả mạo dạng cắt - dán và phát hiện ảnh giả mạo. Bên cạnh đó trình bày một số thuật toán phát hiện ảnh giả mạo như: Thuật toán đối sánh chính xác, thuật toán đối sánh bền vững và thuật toán phát hiện ảnh giả mạo dạng cắt - dán dựa trên các phép biến đổi DWT.

***Chương 3: Phần mềm thử nghiệm***

Tiến hành xây dựng chương trình thực nghiệm để phát hiện ảnh kỹ thuật số giả mạo dạng cắt - dán.

***Kết luận***

Do thời gian và trình độ còn hạn chế nên luận văn khó tránh khỏi những thiếu sót, kính mong nhận được sự đóng góp, chỉ bảo của các thầy giáo, cô giáo và các bạn.

## CHƯƠNG 1

# TỔNG QUAN VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN PHÁT HIỆN ẢNH GIẢ MẠO

*Chương này trình bày tổng quan về một hệ thống xử lý ảnh, các vấn đề cơ bản trong xử lý ảnh, một số định dạng ảnh phổ biến như định dạng Bitmap, định dạng JPEG đồng thời trình bày bài toán phát hiện ảnh giả mạo.*

### 1.1. TỔNG QUAN VỀ XỬ LÝ ẢNH

#### 1.1.1. Xử lý ảnh là gì?

Xử lý ảnh là một khoa học tương đối mới mẻ so với nhiều ngành khoa học khác, nhất là trong quy mô công nghiệp, song trong xử lý ảnh đã bắt đầu xuất hiện những máy tính chuyên dụng. Để có thể hình dung cấu hình một hệ thống xử lý ảnh chuyên dụng hay một hệ thống xử lý ảnh dùng trong nghiên cứu, đào tạo, trước hết chúng ta sẽ xem xét các bước cần thiết trong xử lý ảnh.

Trước hết là quá trình thu nhận ảnh. Ảnh có thể thu nhận qua camera. Thường ảnh thu nhận qua camera là tín hiệu tương tự (loại camera ống kính CCIR), nhưng cũng có thể là tín hiệu số hóa (loại-Change Coupled Device).

Ảnh cũng có thể thu nhận từ vệ tinh qua các bộ cảm ứng (sensor), hay ảnh, tranh được quét trên scanner. Tiếp theo là quá trình số hóa để biến đổi tín hiệu liên tục sang tín hiệu rời rạc (lấy mẫu) và số hóa bằng lượng hóa, trước khi chuyển sang giai đoạn xử lý, phân tích hay lưu trữ lại.

Quá trình phân tích ảnh bao gồm nhiều công đoạn nhỏ. Trước hết là công việc tăng cường ảnh để nâng cao chất lượng ảnh. Do những nguyên nhân khác nhau: Có thể do chất lượng thiết bị thu nhận ảnh, do nguồn sáng hay do nhiễu ảnh có thể bị suy biến. Do vậy cần phải tăng cường và khôi phục lại ảnh để làm