BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG

HỨA THANH LONG

KỸ THUẬT GIẤU TIN TRONG ẢNH 2D VÀ ỨNG DỤNG BẢO MẬT DỮ LIỆU VĂN BẢN

Chuyên ngành: Khoa học máy tính
Mã số : 60.48.01

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Đà Nẵng - Năm 2013

Công trình được hoàn thành tại
ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học: **TS. Nguyễn Tấn Khôi**

Phản biện 1: **PGS.TSKH. Trần Quốc Chiến**

Phản biện 2: **PGS.TS. Trần Cao Đệ**

Luận văn được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp Thạc sĩ Kỹ thuật họp tại Đại Học Đà Nẵng vào ngày 8 tháng 6 năm 2013.

Có thể tìm hiểu Luận văn tại:
- Trung tâm Thông tin - Học liệu, Đại học Đà Nẵng

MỞ ĐẦU

1. Lý do chọn đề tài

Ngày nay với sự tiến bộ vượt bậc của công nghệ thông tin được đánh giá là động lực thay đổi và là bước ngoặc trong lịch sử phát triển của xã hội, đưa thế giới chuyển từ kỷ nguyên công nghiệp sang kỷ nguyên thông tin và phát triển nền kinh tế tri thức. Cuộc cách mạng thông tin kỹ thuật số đã đem lại những thay đổi sâu sắc trong cuộc sống nhân loại. Hàng loạt máy móc và các thiết bị số hiện đại ra đời như: máy tính cá nhân, máy ảnh kỹ thuật số, máy quét ảnh, máy in, máy ghi âm kỹ thuật số đã đem lại nhiều tiện ích cho con người. bên cạnh những ích lợi to lớn, thiết thực mà mạng máy tính đem lại, chúng ta cũng đang đối đầu với những thử thách liên quan đến các vấn đề truyền thông bảo mật và đặc biệt là vấn đề phân phối các tài liệu đa phương tiện sao cho bảo đảm quyền sở hữu trí tuệ. Tình trạng sao chép bất hợp pháp, giả mạo các tác phẩm số hóa gây bức xúc không chỉ riêng các tác giả mà còn cho cả những người làm pháp luật. Những hành vi xâm phạm bản quyền như giả mạo, ăn cắp tác phẩm, sử dụng các tác phẩm không có bản quyền,… đang trở nên phổ biến và ngày càng tinh vi, khó kiểm soát. Tuy nhiên đã có một số phương pháp được đề xuất để khắc phục những vấn đề trên như là: mã hóa thông tin, chữ ký số, giấu tin trong các sản phẩm đa phương tiện.

Hiện nay phương pháp giấu tin được biết đến bởi hai lĩnh vực chủ yếu là Steganography (giấu thông tin mật) và Watermaking (thủy vân). Steganography là một phương pháp giấu thông tin mật vào các dữ liệu truyền thông (ảnh, văn bản, nhạc, phim,…) để chuyển tải thông tin đến người nhận mà người thứ ba không hề biết được sự tồn tại của thông tin mật trong quá trình truyền. Phương pháp Steganography cũng làm thay đổi tư duy trong lĩnh vực bảo mật

thông tin bởi tính khả thi của việc giấu một lượng thông tin mật vào trong một dữ liệu thông thường mà không bị phát hiện bằng giác quan của con người. Bên cạnh đó Watermaking được sử dụng chủ yếu trong lĩnh vực bảo vệ quyền sản phẩm số bằng cách đưa thông tin bản quyền như tên tác giả, logo, … vào sản phẩm. Với sự tồn tại của thông tin thủy vân nhà sản xuất có thể chứng minh được nguồn gốc của sản phẩm khi sản phẩm được phát tán. Cả hai phương pháp được sử dụng với các mục đích khác nhau song chúng đều có chung một đặc điểm chung là giấu thông tin vào sản phẩm số sao cho trong quá trình trao đổi thông tin trên mạng mà người thứ ba không phát hiện được.

Hiện nay phương pháp giấu thông tin mật đã được nhiều người quan tâm, nghiên cứu và được ứng dụng. Đặt biệt vấn đề bảo mật thông tin khi truyền trên mạng.

Xuất phát từ những nhu cầu trên cùng với sự đồng ý của người hướng dẫn, tôi chọn đề tài luận văn cao học:

"Kỹ thuật giấu tin trong ảnh 2D và ứng dụng bảo mật dữ liệu văn bản".

2. Mục tiêu nghiên cứu

2.1. Mục tiêu

Nhằm nghiên cứu, đánh giá các kỹ thuật giấu tin trong ảnh và ứng dụng giấu văn bản mật vào trong ảnh để đảm bảo an toàn trong quá trình trao đổi văn bản công trên mạng.

2.2. Mục tiêu cụ thể

- Tìm hiểu về an toàn và bảo mật thông tin.
- Tìm hiểu về các kỹ thuật giấu tin.
- Tìm hiểu mô hình kỹ thuật giấu tin.

- Tìm hiểu về một số ứng dụng giấu tin đang được triển khai.
- Tìm hiểu về độ an toàn của hệ thống giấu tin.
- Tìm hiểu về các tấn công trên giấu tin.
- Xây dựng chương trình giấu trong ảnh dựa trên kỹ thuật Wu - Lee

3. Đối tượng và phạm vi nghiên cứu

3.1. Đối tượng nghiên cứu

- Các kỹ thuật giấu tin.
- Giấu tin trong ảnh trắng đen, ảnh đa cấp xám, ảnh 2D.
- Văn bản trao đổi trên mạng.

3.2. Phạm vi nghiên cứu

- Nghiên cứu các kỹ thuật giấu tin trong ảnh màu 2D.
- An toàn bảo mật trong quá trình trao đổi văn bản.

4. Phương pháp nghiên cứu

4.1. Phương pháp lý thuyết

- Nghiên cứu để biểu diễn ảnh trên máy tính.
- Cách định dạng ảnh.
- Nét của ảnh.
- Một số tiêu chí đánh giá giấu thông tin trong ảnh số.
- Ứng dụng việc giấu tin và trao đổi thông tin an toàn trên mạng.

4.2. Phân tích hệ thống

- Phân tích các chức năng.
- Xây dựng chương trình.

4.3. Phương pháp thực nghiệm

- Triển khai thử nghiệm
- Đánh giá kết quả

5. Cấu trúc của luận văn

Nội dung luận văn bao gồm phần mở đầu, ba chương và phần kết luận.

Chương 1: TỔNG QUAN ĐỀ TÀI. Trình bày một số khái niêm cơ bản về kỹ thuật giấu thông tin, phân loại các kỹ thuật giấu tin, những ứng dụng cơ bản, mô hình tổng quát của kỹ thuật giấu tin và vài phần mềm hiện có.

Chương 2: KỸ THUẬT GIẤU TIN. Trình bày các nghiên cứu về kỹ thuật giấu tin trong môi trường ảnh, sự khác biệt của kỹ thuật giấu tin trong các loại ảnh khác nhau, các tính chất và yêu cầu của hệ giấu tin trong ảnh.

Chương 3: XÂY DỰNG HỆ THỐNG BẢO MẬT VĂN BẢN DỰA VÀO GIẤU TIN. Khảo sát, đánh giá, so sánh một số kỹ thuật giấu thông tin cơ bản. Phát triển một chương trình giấu tin thử nghiệm sử dụng kỹ thuật giấu tin "Wu – Lee". Đánh giá chất lượng của kỹ thuật này.

Phần kết luận nêu những kết quả đạt được, hướng nghiên cứu và những đề xuất từ hệ thống giấu tin và phát triển hoàn thiện hệ thống giấu tin đã xây dựng.

6. Tổng quan tài liệu

Căn cứ vào tên của luân văn là "giấu tin trong ảnh 2D và ứng dụng bảo mật dữ liệu văn bản" và từ đó có thể dựa trên các từ trong tên luận văn để tìm kiếm.

Các tài liệu đã đươc công bố (sách, bài báo, luận văn, luận án, văn bản), Các tài liệu chưa được công bố (báo cáo, bài trình bày hội thảo), Ý kiến chuyên gia.

- Tìm kiếm thủ công: Vào các thư viện của trường để tìm kiếm những quyển sách mà giống với nội dung minh cần tìm.

- Tìm kiếm từ Internet: Với sự phát triển mạnh của Internet đã có nhiều trang web hỗ trợ chúng ta tìm kiếm thông tin một cách dễ dàng, nhanh chóng, ít tốn kém. Trang web phổ biến nhất hiện nay để tìm kiếm như là: www.google.com.vn

Để tìm kiếm nhanh chúng ta cần xác định được từ khóa cần tìm.

Sử dụng OR, NOT, AND và cặp dấu " " nếu muốn đúng cụm từ.

Sau khi thực hiện hai phương pháp tìm kiếm thủ công và tìm kiếm từ Internet. Kiểm tra độ chính xác của tài liệu, trích chọn những trang, những phần nội dung đúng hoặc gần đúng với nội dung luận văn.

Tổng hợp lại những tài liệu sau khi đã được đánh giá để trích đưa vào luận văn.

CHƯƠNG 1
TỔNG QUAN ĐỀ TÀI

1.1. GIỚI THIỆU CHUNG

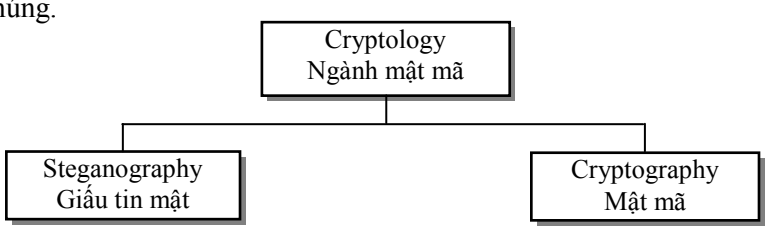
Sự phát triển của công nghệ thông tin đã đem lại những thay đổi sâu sắc trong xã hội và trong cuộc sống của chúng ta. Những thuận lợi mà thông tin kỹ thuật số mang lại cũng sinh ra những thách thức và cơ hội cho quá trình đổi mới. Sự ra đơi những phần mềm có tính năng rất mạnh, các thiết bị mới như máy ảnh kỹ thuật số, máy quét ảnh chất lượng cao, máy in, máy ghi âm kỹ thuật số, v.v…, đã được sáng tạo trên cơ sở thỏa mãn thế giới tiêu dùng rộng lớn, để xử lý và thưởng thức các dữ liệu đa phương tiện (multimedia data). Mạng Internet toàn cầu đã hình thành một xã hội ảo nơi diễn ra quá trình trao đổi thông tin trong mọi lĩnh vực chính trị, quân sự, quốc phòng, kinh tế thương mại,… Chính trong môi trường mở và tiện nghi như thế xuất hiện những vấn nạn, tiêu cực đang rất cần đến các giải pháp hữu hiệu cho vấn đề an toàn thông tin như nạn ăn cắp bản quyền, nạn xuyên tạc thông tin, truy cập thông tin trái phép, … Tìm giải pháp cho những vấn đề nêu trên không chỉ tạo điều kiện đi sâu vào lĩnh vực công nghệ phức tạp đang phát triển rất nhanh này mà còn dẫn đến những cơ hội phát triển kinh tế.

Quá trình phát triển lâu dài, có nhiêu phương pháp bảo vệ thông tin đã được đưa ra, trong đó giải pháp dùng mật mã học là giải pháp được ứng dụng rộng rãi nhất. Các hệ mã đã được phát triển nhanh chóng và được ứng dụng rất phổ biến cho đến tận ngày nay. Thông tin ban đầu được mã hóa thành các ký hiệu vô nghĩa, sau đó sẽ được lấy lại thông qua việc giải mã nhờ vào khóa của hệ mã. Đã có rất nhiều những hệ mã phức tạp được sử dụng như DES, RSA,…

các phương pháp này trong thực tế tỏ ra rất hiệu quả và được ứng dụng phổ biến.

Tuy nhiên trong luận văn này không đi sâu vào nghiên cứu về các hệ mật mã mà chỉ tiếp cận với một phương pháp đã và đang được nghiên cứu, pháp triển ở nhiều nước trên thới giới, đó là phương pháp che giấu thông tin.

Để đảm bảo an toàn cho nội dung của thông tin, người ta thường sử dùng phương pháp mã hóa thông tin, nhằm giấu đi ý nghĩa của nó. Để giữ bí mật cho thông tin người ta tìm ra cách che giấu đi sự hiện diện của nó. Xu hướng hiện nay là kết hợp hai kỹ thuật: mã hóa thông tin sau đó che giấu thông tin. Mã hóa và che giấu thông tin có quan hệ chặt chẽ với nhau. Nhiều ý tưởng của kỹ thuật mật mã (Steganography) rất hữu ích trong việc che giấu sự hiện hữu của thông tin. Nghiên cứu việc kết hợp của hai kỹ thuật mật mã và che giấu thông tin, nhằm khắc phục những nhược điểm hoặc những hạn chế của từng loại. Cho phép xây dựng những hệ thống bảo mật, an toàn cho việc chuyển tải dữ liệu trên phương tiện thông tin đại chúng.



Hình 1.1 Phân cấp các lĩnh vực nghiên cứu của mật mã học

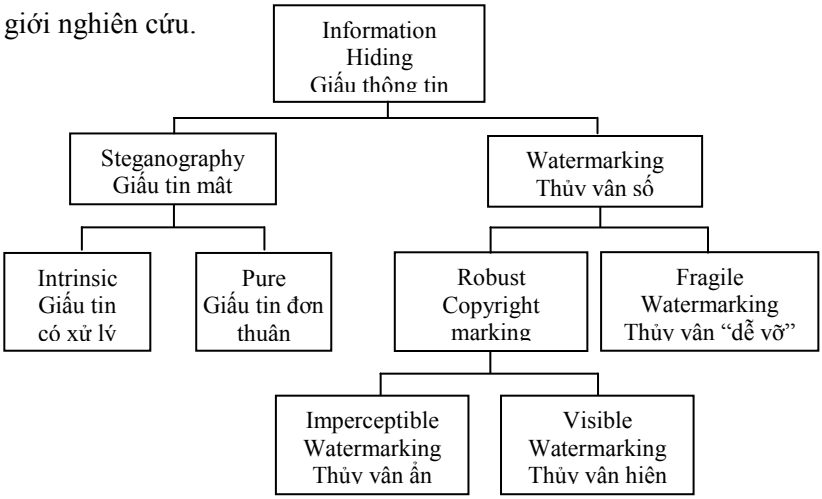
1.2. GIẤU THÔNG TIN VÀ VÀI NÉT VỀ LỊCH SỬ CỦA NÓ

1.2.1. Định nghĩa giấu thông tin

Giấu thông tin là một kỹ thuật nhúng (giấu) một lượng thông tin số nào đó vào trong một đối tượng dữ liệu khác.

1.2.2. Phân loai kỹ thuật giấu tin

Năm 1999, Fabien A.P. Petitcolas đưa ra, sau hội nghị quốc tế lần thứ hai về giấu tin năm 1998 và đã được chấp nhận rộng rãi trong giới nghiên cứu.



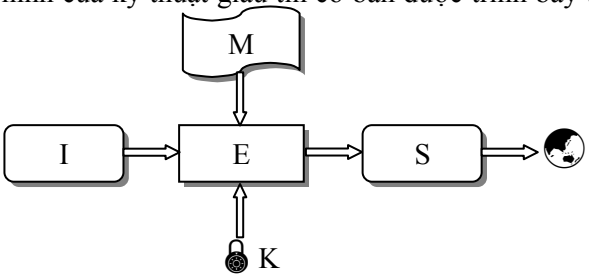
Hình 1.2. Phân loại các kỹ thuật giấu thông tin[8]

1.2.3. Lịch sử giấu tin

Từ Steganography bắt nguôn từ thời Hy Lạp cổ và được sử dụng cho tới ngày nay, nó có nghĩa là tài liệu được phủ (covered writing).

1.3. MÔ HÌNH KỸ THUẬT GIẤU TIN

Mô hình của kỹ thuật giấu tin cơ bản được trình bày trên hình vẽ sau:



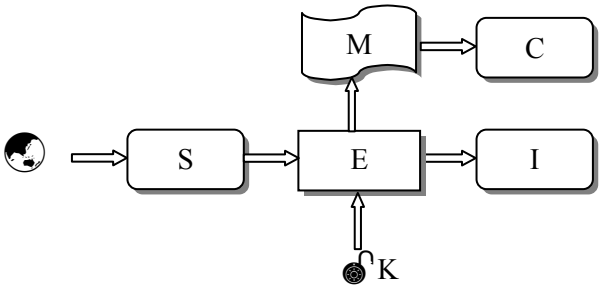
Hình 1.3. Lược đồ chung cho quá trình giấu thông tin

Hình vẽ trên biểu diễn quá trình giấu thông tin cơ bản. Phương tiện chứa bao gồm các đối tượng được dùng làm môi trường để giấu thông tin như văn bản, ảnh, audio, video,… dữ liệu giấu là một lương thông tin mang ý nghĩa nào đó, tùy thuộc vào mục đích của người sử dụng.

Ký hiệu:

a. Secret Message (M): thông tin cần giấu.
b. Cover Data (I): dữ liệu phủ, môi trường giấu thông tin.
c. Embedding Algorithm (E): bộ mã hóa / giải mã là những chương trình, những thuật toán nhúng tin.
d. Key (K): khóa bí mật, sử dụng trong kỹ thuật giấu tin.
e. Stego Data (S): dữ liệu mang tin mật.
f. Control (C): kiểm tra thông tin sau khi giải mã

Thông tin được giấu vào phương tiện chứa, theo một thuật toán, sử dụng khóa bí mật dùng chung giữa người gởi và người nhận thông tin.



Hình 1.4. Lược đồ của quá trình giải mã thông tin

1.4. ỨNG DỤNG CỦA GIẤU TIN

a. Bảo vệ quyền tác giả (copyright protection)

b. Nhận thực thông tin hay phát hiện xuyên tạc thông tin (authentication and tamperdection)

c. Giấu vân tay hay dán nhãn (fingerprinting and labeling):

d. Điều khiển truy cập (copy control).

e. Giấu tin mật (Steganography).

1.5. GIẤU TIN TRONG DỮ LIỆU ĐA PHƯƠNG TIỆN

1.5.1. Giấu thông tin trong ảnh

1.5.2. Giấu thông tin trong audio

1.5.3. Giấu thông tin trong video

1.6. ĐỘ AN TOÀN CỦA MỘT HỆ THỐNG GIẤU TIN

Việc phá vỡ một hệ thống giấu tin thông thường gồm ba phần: phát hiện, giải tin và hủy thông tin đã giấu. Một hệ thống giấu tin mật được gọi là thực sự an toàn khi kẻ tấn công không phát hiện được sự tồn tại của thông tin giấu trong một đối tượng chứa.

1.7. CÁC TẤN CÔNG TRÊN GIẤU TIN

Tấn công trên đối tượng đã giấu tin là những phép biến đổi sao cho có thể làm mất thông tin giấu. Các kỹ thuật tấn công phân làm hai nhóm chính:

- Một là biến đổi tạo nhiễu đối với dữ liệu đã được giấu tin
- Hai là làm mất tính đồng bộ giữa đối tượng vỏ và thông tin giấu để không thể khôi phục lại tin đã giấu.

Vấn đề đặt ra là liệu có thể tồn tại một hệ giấu tin bền vững trước các tân công trên? Đến nay, vẫn chưa tim được một hệ giấu tin nào bền vững trước tất cả mọi kiểu tấn công.

CHƯƠNG 2
KỸ THUẬT GIẤU TIN

2.1. ĐẶC TRƯNG VÀ TÍNH CHẤT GIẤU TIN TRONG ẢNH

Các kỹ thuật giấu tin phần lớn tập trung vào giấu thông tin trong ảnh. Mỗi phương tiện chứa khác nhau sẽ có những kỹ thuật giấu tin khác nhau. Có nhiều định dạng cũng như tính chất của các ảnh khác nhau nên các kỹ thuật giấu tin trong ảnh có những đặc trưng và các tính chất cơ bản sau đây:

1) *Phương tiện chứa dữ liệu tri giác tĩnh:*
2) *Kỹ thuật giấu tin phụ thuộc ảnh:*
3) *Kỹ thuật giấu tin lợi dụng tính chất hệ thống thị giác của con người*
4) *Giấu thông tin trong ảnh tác động lên dữ liệu ảnh nhưng không làm thay đổi kích thước ảnh*
5) *Đảm bảo yêu cầu chất lượng ảnh sau khi giấu thông tin*

2.2. GIẤU THÔNG TIN TRONG ẢNH TRẮNG ĐEN, ẢNH 2D VÀ ẢNH ĐA CẤP XÁM

Khởi nguồn của giấu tin thông tin trong ảnh là thông tin được giấu trong ảnh 2D hoặc ảnh xám, trong đó mỗi pixel ảnh mang nhiều giá trị, được biểu diễn bằng nhiều bit. Với loại ảnh này, một thay đổi giá trị nhỏ ở mỗi pixel hầu như không làm thay đổi chất lượng ảnh và khả năng bi phát hiện dưới sự quan sát của mắt thường là rất thấp.

2.3. HỆ THỐNG THỊ GIÁC CỦA CON NGƯỜI VÀ CÁC MÔ HÌNH MÀU CỦA ẢNH

Hiểu biết về hệ thống thị giác của con người sẽ góp một phần không nhỏ trong việc cải tiến và nâng cấp các thuật toán về giấu tin trong ảnh.

Để nhận biết một ảnh, nào người xử lý các thông tin thu nhận như vị trí không gian, đường nét, màu sắc (độ chói, động tương phản, tần số, …) của ảnh.

Ánh sáng là dải sóng mà mắt người có thể cảm nhận được, khi đi qua lăng kính được phân tích thành những ánh sáng đơn sắc có phổ màu từ tím đến đỏ, với dải bước sóng tương ứng từ 400 – 700 nm. Ánh sáng màu là tổ hợp của ánh sáng đơn sắc. Mắt người có thể cảm nhận được vài chục màu, song lại có thể phân biệt được tới hàng ngàn màu khác nhau trong không gian màu. Có ba thuộc tính chủ yếu trong cảm nhận màu:

– Sắc màu (Hue): dùng để phân biệt sự khác nhau giữa các màu.

– Mức bão hòa (Saturation): chỉ ra mức độ thuần của một màu hay khoảng cách của màu tới điểm có cường độ cần bằng.

– Độ sáng (Lightness): mô tả cường độ (intensity) sáng, là ánh sáng phản xạ nhận được từ đối tượng. tỉnh thoảng thuật ngữ Brightness (độ phát sáng) được dùng thay cho độ sáng, nó mô tả cường độ ánh sáng do đối tượng tự phát ra.

2.4. BIỄU DIỄN ẢNH TRÊN MÁY TÍNH

2.4.1. Ảnh Vector

Ảnh vectơ được tạo mới từ rất nhiều đối tượng khác nhau. Đối tượng Vectơ được xác định bởi các hàm toán học mà không chỉ là các điểm ảnh, ảnh vectơ cho chất lượng ảnh cao hơn ảnh Bitmap. Các đối tượng cơ bản gồm đường thẳng, đường cong và một số hình toán học gốc với các thuộc tính màu sắc, độ dài đường, …

2.4.2. Ảnh mành

Ảnh Bitmap được xây dựng từ các điểm ảnh màu (pixels) là một khối nhỏ màu hình chữ nhật. Tất cả các điểm màu được sắp xếp

với nhau theo một trật tự tạo thành ảnh.

2.5. CÁC ĐỊNH DẠNG ẢNH THÔNG DỤNG

Ảnh thu được sau quá trình số hóa có nhiều loại khác nhau, phụ thuộc vào kỹ thuật số hóa ảnh. Được chia thành hai loại ảnh: ảnh trắng đen và ảnh màu sau đây là một số định dạng ảnh thông dụng.

2.5.1. Định dạng ảnh IMG

2.5.2. Định dạng ảnh PCX

2.5.3. Định dạng ảnh TIFF (Targed Image File Format)

2.5.4. Định dạng ảnh GIF (Graphics Interchanger Format)

2.5.5. Định dạng anh JPEG (Joint Photographic Experts Group)

2.6. NÉN ẢNH

Nén dữ liệu là quá trình làm giảm dung lượng thông tin "dư thừa" trong dữ liệu gốc, do vậy lượng thông tin thu được sau nén thường nhỏ hơn dữ liệu gốc rất nhiều (khoảng 10%), những kỹ thuật nén mới như fratal cho tỷ lệ nén đến 30%. Tùy thuộc vào ứng dụng để chọn thuật toán nén theo một số các chỉ tiêu như:

- Tỷ lệ nén cao.
- Đảm bảo chất lượng, không làm mất thông tin ảnh.
- Tốc độ nén và tải nén cao.
- Phù hợp với các quy định hiện có.

2.6.1. Tỷ lệ nén (Compression rate)

Tỷ lệ nén = kích thước dữ liệu gốc / kích thước dữ liệu thu được.

2.6.2. Một số phương pháp nén ảnh

Kích thước ảnh gốc = Số phần tử x Độ dài bit tối đa

a. Phương pháp Huffman:

b. Phương pháp nén loạt dải RLC (Run Length Coding)

c. Phương pháp LZW (Lempel Ziv Welch)

2.7. TIÊU CHÍ ĐÁNH GIÁ KỸ THUẬT GIẤU TIN TRONG ẢNH SỐ

2.7.1. Tính vô hình

2.7.2. Khả năng giấu thông tin

2.7.3. Chất lượng của ảnh có giấu thông tin

2.7.4. Tính bền vững của thông tin được giấu

2.7.5. Thuật toán và độ phức tạp tính toán

2.8. KHẢO SÁT MỘT SỐ CÔNG CỤ GIẤU TIN TRONG ẢNH

2.8.1. Hide And Seek V4.1

2.8.2. Stego Dos

2.8.3. While Noise Storm

2.8.4. S – Tools for Windows

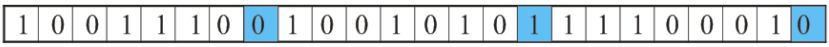
2.9. CÁC KỸ THUẬT XỬ LÝ ĐIỂM ẢNH

Xử lý các điểm ảnh là một kỹ thuật được sử dụng thường xuyên trong các kỹ thuật giấu tin trong ảnh. Các giá trị điểm ảnh được lấy ra rồi biến đổi theo thuật toán giấu tin. Tuy nhiên, miền giá trị của các điểm ảnh lại khác nhau phụ thuộc vào các loại ảnh, chính vì thế ta cần dùng đến kỹ thuật tách bit thông tin từ giá trị điểm ảnh.

Kỹ thuật này được sử dụng nhiều trong kỹ thuật giấu tin sử dụng các bit ít quan trọng nhất của điểm ảnh (LSB – Least Significant Bit) sẽ được trình bày cụ thể trong chương III. Kỹ thuật LSB là kỹ thuật sử dụng các bit ít quan trọng về trị giác nhất trong các bit mang giá trị điểm ảnh để giấu tin. Ví dụ như ảnh với ảnh 256 màu thi bit cuối cùng trong 8 bit biểu diễn một điểm ảnh được coi là bit quan trọng nhất theo nghĩa là nếu thay đổi bit này thì ảnh hưởng ít nhất đến cảm nhận của mắt người về điểm ảnh. Hay đối với ảnh 16 bit thì 15 bit là biểu diễn 3 màu RGB của điểm ảnh còn bit cuối cùng

không dùng đến thì ta sẽ tách bit này ra ở mỗi điểm ảnh để giấu tin… Như vậy, kỹ thuật tách bit trong xử lý điểm ảnh được sử dụng rất nhiều trong kỹ thuật giấu tin. Sau đây ta sẽ khảo sát một số kỹ thuật tách bit ít quan trọng trên một số loại ảnh phổ biến:

Tách bit cuối cùng trong 8 bit biểu diễn mỗi điểm ảnh của ảnh 256 màu



Hình 2.1. Mỗi điểm ảnh biểu diễn bởi 8 bit, bit cuối cùng (đổi màu) được coi là bít quan trọng nhất

Trong phép tách này bit cuối cùng được coi là bit quan trọng nhất, thay đổi giá trị của bit này sẽ tăng hoặc giảm giá trị của điểm ảnh đúng một đơn vị. ví dụ một điểm ảnh có giá trị là 234, nếu thay đổi bit cuối cùng từ 0 thành 1 giá trị mới. Của điểm ảnh là 235. Sự thay đổi nhỏ đó sẽ không làm màu của điểm ảnh thay đổi nhiều.

Với các ảnh 16, 24 bit ta cũng thực hiện tương tự.

- *Tách phần Blue trong RGB.*
- *Biến đổi không gian màu cho ảnh 24 bit màu.*

CHƯƠNG 3
XÂY DỰNG HỆ THỐNG BẢO MẬT VĂN BẢN DỰA VÀO GIẤU TIN

3.1. GIỚI THIỆU

Để thực hiện việc giấu thông tin trong môi trường ảnh, trước hết cần số hóa các bức ảnh theo những chuẩn phổ biến như: JPEG, PCX, GIF,…

Sau bước số hóa, tùy thuộc cấp độ màu khác nhau, có thể phân chia các loại ảnh trắng đen, ảnh xám hay ảnh màu. Ảnh trắng đen là ảnh nhị phân có 1 bit màu. Ứng với điểm đen, bit mang giá trị 0 và ứng với điểm trắng, bit mang giá trị 1. Giấu tin trong ảnh trắng đen thường gây nhiễu, dễ nhận biết được bằng mắt thường, số lượng thông tin giấu được cũng hạn chế.

Ảnh màu trong máy tính là mảng các số thể hiện cường độ sáng tại mỗi điểm ảnh. Các điểm ảnh cấu trúc theo dạng ảnh mảnh, số điểm ảnh thay đổi tuy thuộc độ phân giải của màn hình máy tính.

Khi chuyển một ảnh tương tự (analog) sang ảnh số, người ta có thể chọn những cách thể hiện màu khác nhau:

- Ảnh 8 bit màu: mỗi điểm ảnh có thể nhận 1 trong 2^8 (256) màu, chọn từ bảng màu.

- Ảnh 8 bit màu dải xám: mỗi điểm ảnh có thể nhận 1 trong 2^8 (256) sắc thái xám.

Mức xám là kết quả sự mã hóa tương ứng một cường độ sáng của mỗi điểm ảnh với một giá trị số sau quá trình lượng hóa. Cách mã hóa kinh điển thường dùng là 16, 32, 64 mức. Mã hóa $2^8 = 256$ (0,1,…255) mức là phổ dụng nhất, mỗi điểm ảnh sẽ được mã hóa bởi 8 bit.

Ảnh 24 bit màu: mỗi điểm ảnh có thể nhận 1 trong 2^{24} (trên 16 triệu) màu, mỗi màu là sự pha trộn của 3 màu cơ bản RGB, nhận một giá trị từ 0 đến 255.

3.2. THUẬT TOÁN GIẤU THÔNG TIN TRONG KHỐI BIT

Thuật toán giấu tin trong khối bit thường được ứng dụng cho các tập tin dữ liệu môi trường kiểu hai màu (trắng đen) như các bản photocopy, Fax, mã vạch hay ảnh trắng đen. Ảnh trắng đen khó giấu tin do đặc điểm mỗi điểm ảnh chỉ được biểu diễn bởi một giá trị bit (0 hoặc 1). Nếu ta thay đổi giá trị bit từ 0 sang 1 hay ngược lại từ 1 sang 0 thì đều làm cho trên ảnh xuất hiện những điểm đen, điểm trắng lạ, dễ bị phát hiện.

3.2.1. Kỹ thuật giấu tin ngây thơ

Kỹ thuật giấu tin này khá đơn giản, bí mật của phương pháp chỉ là kích thước của khối ảnh. Bản chất của giấu tin được thực hiện trong kỹ thuật này là cách thức giấu thông tin theo quy ước chẵn lẻ.

3.2.2. Kỹ thuật giấu tin Chen-Pan-Tseng

a. Ý tưởng

Thuật toán giấu tin trong ảnh trắng đen được Yu Yuan Chen, Hsiang Kuang Pan và Yu Chee Tseng, khoa Công nghệ thông tin và Khoa học máy tính của trường Đại học quốc gia Đài Loan đề nghị. Trong phương pháp này, ngoài ma trận khóa (K) còn sử dụng thêm một ma trận trọng số (W) khi giấu thông tin. Thuật toán này đảm bảo tốt, an toàn và giấu được nhiều thông tin trong ảnh, bằng cách thay đổi nhiều nhất 2 bit mỗi khối. Nhược điểm của phương pháp này là chất lượng ảnh chưa cao, dễ bị phát hiện, chỉ áp dụng cho ảnh màu. Thuật toán này cải tiến sẽ cải thiện rất nhiều chất lượng ảnh bằng kỹ thuật chọn hệ số phân bố bit trắng đen và số bit giấu tương đương.

b. Một số khái niệm

- Khóa bí mật (Secret Key): K
- Ma trận trọng số (Weight Matrix) cấp r.
- Phép đảo bit:
- Phép toán trên ma trận dùng trong thuật toán.

c. Thuật toán

d. Phân tích đánh giá thuật toán

3.2.3. Kỹ thuật giấu tin Wu – Lee

a. Ý tưởng

b. Phân tích thuật toán

c. Nhận xét

3.3. THUẬT TOÁN GIẤU THÔNG TIN THAY THẾ BIT CÓ TRONG SỐ THẤP

Những năm gần đây, kỹ thuật giấu thông tin vào miền bit có trọng số thấp của ảnh được quan tâm nghiên cứu khá nhiều. LSB là bit có ảnh hưởng ít nhất tới việc quyết định màu sắc của mỗi điểm ảnh, vì vậy khi ta thay đổi bit này thì màu sắc của điểm ảnh mới sẽ gần như không khác biệt so với điểm ảnh cũ. LSB của một điểm ảnh có vị trí tương tự như chữ số hàng đơn vị của một số tự nhiên, khi bị thay đổi, giá trị chênh lệch giữa số cũ và số mới sẽ ít nhất, so với khi ta thay đổi giá trị của chữ số hàng chục hoặc hàng trăm. Việc xác định LSB của mỗi điểm ảnh trong một bức ảnh phụ thuộc vào định dạng của ảnh và số bit màu dành cho mỗi điểm ảnh của ảnh đó. Nội dung của phương pháp là chọn ra các bit ít quan trọng (ít làm thay đổi chất lượng của ảnh nền) và thay thế chúng bằng các bit thông tin cần giấu. Để khó bị phát hiện, thông tin giấu thường được nhúng vào những vùng mắt người kém nhạy cảm với màu sắc. Với ảnh 24 bit, mỗi màu được chứa trong 3 byte, theo thứ tự từ phải sang, byte đầu

tiên chứa giá trị biểu thị cường độ màu lam (B), byte thứ hai chứa giá trị biểu thị cường độ màu lục (G), byte thứ ba chứa giá trị biểu thị cường độ màu đỏ (R). Như vậy, mỗi màu được xác định bởi một số nguyên có giá trị trong khoảng từ 0 – 255.

3.3.1. Thuật toán

3.3.2. Phân tích, đánh giá thuật toán

3.4. MỘT SỐ KỸ THUẬT GIẤU TIN KHÁC

3.4.1. Kỹ thuật giấu tin dựa trên bảng màu

3.4.2. Kỹ thuật trãi phổ (Spread Spectrum Communication)

3.4.3. Kỹ thuật dùng hệ số DCT (Discrete Cosine Transform)

3.5. KẾT QUẢ THỬ NGHIỆM

Ảnh gốc được xử lý là ảnh có định dạng bitmap "coquan.bmp" có kích thước là 797 x 565

Hình 3.1. Ảnh gốc kích thước 797 x 565

Văn bản giấu vào trong ảnh là văn bản dạng text

Hình 3.2. Thực hiện giấu tin trong ảnh kích thước 797 x 565

Sau khi giấu một lượng văn bản vào trong ảnh. Ta được một tấm ảnh có giấu tin.

Hình 3.3. Ảnh được giấu tin kích thước 797 x 565

Kết quả là ảnh không thay đổi về dung lượng cũng như kích thước

Kết quả trích xuất văn bản được tách ra từ trong ảnh là không bị sai lệch so với văn bản đưa vào lúc ban đầu.

Hình 3.4. Thực hiện trích xuất thông tin trong ảnh kích thước 797 x 565

Thuật toán này áp dụng tốt cho ảnh màu thì có thể nói thuật toán đã đạt được yêu cầu cơ bản của một ứng dụng giấu tin mật, như đảm bảo tính ẩn của thông tin được giấu và an toàn. Tuy nhiên số lượng thông tin giấu chưa cao.

3.6. ĐÁNH GIÁ KẾT QUẢ

3.6.1. Ưu điểm

- Ảnh gốc sau khi được giấu thông tin mật có sự sai khác với ảnh gốc trong phạm vi chấp nhận được.

- Nội dung thông tin giấu sau khi được giải tin không sai lệch so với nội dung gốc ban đầu. Sự thay đổi của ảnh sau khi giấu tin gần như không cảm nhận được bởi hệ thống thị giác của con người, đảm bảo an toàn cho thông tin giấu.

3.6.2. Khuyết điểm

- Số lượng văn bản được giấu vào trong ảnh còn hạn chế.

- Khi kích thước ảnh càng lớn thì lượng văn bản giấu được càng nhiều.

- Khóa bảo vệ khi nhập vào chưa được mã hóa khi hiển thị

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Luận văn này đã trình bày và đánh giá những kỹ thuật giấu tin cơ bản và cài đặt thử nghiệm một chương trình giấu tin mật kỹ thuật giấu tin mật Wu – Lee để trao đổi văn bản mật.

Chương trình này thử nghiệm giấu văn bản mật trong môi trường ảnh. Thuật toán trên có thể sử dụng với ảnh đa mức xám và các định dạng ảnh sử dụng bảng màu nói chung. Cài đặt, tiến hành thử nghiệm thuật toán trên môi trường định dạng ảnh Bitmap phù hợp với xu hướng thực tế trong trao đổi thông tin. Sự thay đổi của ảnh sau khi giấu tin gần như không cảm nhận được bởi hệ thống thị giác của con người, cho thấy thuật toán có thể áp dụng trong thực tế. Chương trình cũng đã sử dụng hệ thống khóa bí mật cho ảnh, đảm bảo tính mật, an toàn cao đối với lượng văn bản giấu, đông thời ảnh chứa tin giấu không có những thay đổi so với ảnh gốc.

Trên đây mới chỉ là những ý kiến chủ quan, em luôn rất mong muốn nhận được sự quan tâm đóng góp của quý thầy, cô và những người quan tâm đến lĩnh vực giấu văn bản trong ảnh.

Hướng phát triển tiếp theo là cần nghiên cưu sâu hơn nữa việc đưa nội dung văn bản vào trong ảnh bằng một file văn bản mà không cần phải nhập trực tiếp vào chương trình.

Mã hóa thông tin trước khi giấu là biện pháp hữu hiệu nhất để đảm bảo an toàn cho thông tin. Mã hóa hiện nay được sử dụng kỹ thuật mật mã khóa công khai (khóa lập mã là công khai, giải mã là bí mật). Lý thuyết đã chứng minh, có thể công bố phương pháp mã hóa, công khai khóa lập mã, chỉ giữ bí mật khóa giải mã, mà vẫn giữ được bí mật thông tin, việc tìm khóa giải mã là bài toán nan giải, theo nghĩa có độ phức tạp về thời gian rất lớn

Cần tiếp tục nghiên cứu hoàn thiện bài toán giấu tin trên các loại dữ liệu khác như: Audio, Video,.. để tăng độ bảo mật của nội dung thông tin và tránh được sự nghi ngờ của những kẻ thám tin.