LỜI CẢM ƠN

Em xin bày tỏ lòng biết ơn sâu sắc nhất tới giảng viên hướng dẫn là Thạc sỹ Hồ Thị Hương Thơm, cô đã tận tình hướng dẫn và giúp đỡ em rất nhiều trong quá trình tìm hiểu và nghiên cứu để em có thể hoàn thành tốt đề tài tốt nghiệp của mình.

Em xin gửi lời cảm ơn đến Ban giám hiệu và các Thầy Cô giáo của Trường Đại học Dân Lập Hải Phòng đã giảng dạy chúng em trong suốt 4 năm học, cung cấp cho chúng em những kiến thức chuyên môn cần thiết và quý báu giúp chúng em hiểu rõ hơn các lĩnh vực nghiên cứu để hoàn thành đề tài được giao.

Xin cảm ơn các bạn bè và gia đình đã động viên cổ vũ, đóng góp ý kiến, trao đổi trong suốt quá trình học tập cũng như làm tốt nghiệp, giúp em hoàn thành đề tài đúng thời hạn.

Hải Phòng, ngày 1 tháng 7 năm 2011

Sinh viên

Vũ Thùy Dung

MỤC LỤC

DANH MỤC HÌNH VẼ

DANH MỤC BẢNG BIỂU

DANH MỤC CHỮ VIẾT TẮT

LSB	Least Significant Bits	Các bit ít quan trọng nhất
RS	Regular / Singular	Kỹ thuật chính quy - đơn
SES	Steganography Evading Statistical analyses	Kỹ thuật giấu tin tránh phát hiện bằng thống kê
IMG	Image	Ảnh đen trắng img
PCX	Personal Computer Exchange	Ảnh xám PCX
GIF	Graphics Interchange Format	Định dạng ảnh đồ họa GIF
BMP	Bitmap	Ảnh không nén Bitmap
PNG	Portable Network Graphics	Ảnh PNG
JPEG	Joint Photographic Experts Group	Ảnh nén JPEG
PSNR	Peak signal to noise ratio	Tỉ số tín hiệu cực đại trên nhiễu
MSE	Mean Squared Error	Lỗi bình phương

MỞ ĐẦU

Công nghệ thông tin và đặc biệt là sự phát triển của hệ thống mạng máy tính đã tạo nên môi trường mở và là phương tiện trao đổi, phân phối tài liệu một cách tiện lợi, nhanh chóng. Tuy nhiên cũng đặt ra một vấn đề về bảo vệ tài liệu, ngăn chặn việc đánh cắp và sao chép tài liệu một cách bất hợp pháp. Vấn đề an toàn và bảo mật thông tin hiện nay luôn nhận được sự quan tâm đặc biệt của nhiều nhà nghiên cứu trong nhiều lĩnh vực.

Giấu tin trong ảnh là một bộ phận chiếm tỷ lệ lớn nhất trong các chương trình ứng dụng, các phần mềm, hệ thống giấu tin trong dữ liệu đa phương tiện bởi lượng thông tin được trao đổi bằng ảnh là rất lớn. Hơn nữa, giấu thông tin trong ảnh cũng đóng vai trò hết sức quan trọng trong hầu hết các ứng dụng bảo vệ an toàn thông tin như: nhận thức thông tin, xác định xuyên tạc thông tin, bảo vệ bản quyền tác giả, điều khiển truy nhập, giấu thông tin mật…

Đồ án trình bày về giấu và phát hiện ảnh có giấu thông tin. Đồng thời trình bày về kỹ thuật giấu tin trên miền dữ liệu ảnh bằng SES (Steganography Evading Statistical analyses). Từ đó đưa ra các thực nghiệm và đánh giá cho việc phát hiện thông tin ẩn giấu trên miền dữ liệu ảnh bằng SES.

Để nói rõ về nội dung này, đồ án của em được tổ chức gồm năm chương:

Chương 1: Tổng quan kỹ thuật giấu tin trong ảnh.

Chương 2: Cấu trúc chung của ảnh Bitmap.

Chương 3: Kỹ thuật phát hiện ảnh có giấu tin RS (Regular / Singular).

Chương 4: Kỹ thuật giấu tin SES (Steganography Evading Statistical analyses).

Chương 5: Cài đặt và thử nghiệm.

CHƯƠNG 1: TỔNG QUAN VỀ GIẤU TIN TRONG ẢNH

1.1 Tổng quan về giấu tin

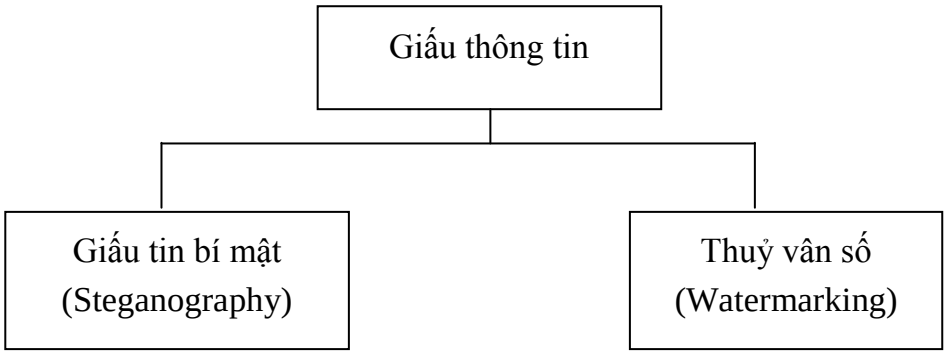
1.1.1 Định nghĩa về giấu tin

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

1.1.2 Mục đích của giấu tin

- Bảo mật cho những dữ liệu được đem giấu.
- Bảo đảm an toàn (bảo vệ bản quyền) cho chính các đối tượng chứa những dự liệu giấu trong đó.

Có thể thấy hai mục đích này hoàn toàn đối lập nhau và dần dần phát triển thành 2 lĩnh vực với những yêu cầu và tính chất khác nhau (Hình 1.1).



Hình 1.1. Hai lĩnh vực chính của kỹ thuật giấu thông tin

1.1.2.1 *Kỹ thuật giấu thông tin bí mật (Steganography)*

Đảm bảo an toàn và bảo mật thông tin tập trung vào các kỹ thuật giấu tin để có thể giấu được nhiều thông tin nhất. Thông tin mật được giấu trong đối tượng sao cho người khác không phát hiện được.

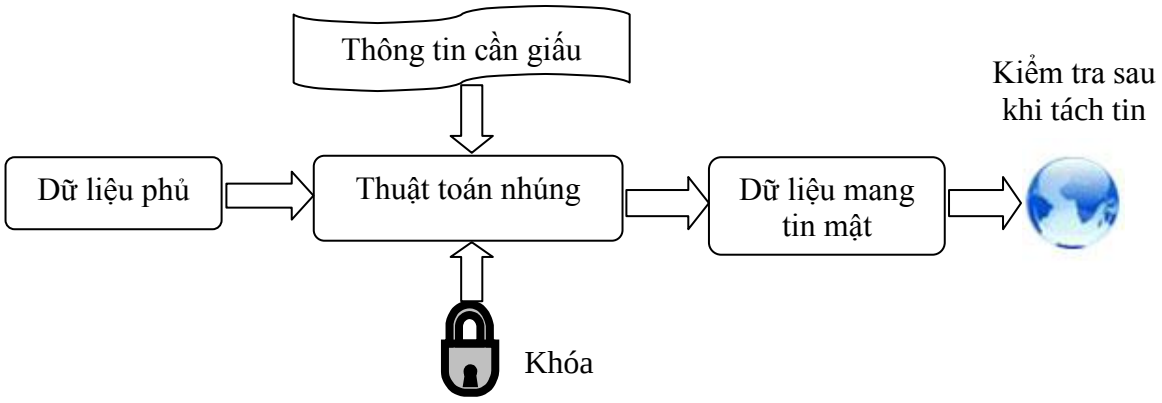
1.1.2.2 *Kỹ thuật giấu thông tin theo kiểu đánh dấu (Watermarking)*

Để bảo vệ bản quyền của đối tượng chứa thông tin, kỹ thuật giấu tin tập trung đảm bảo một số yêu cầu như đảm bảo tính bền vững… Đây chính là ứng dụng cơ bản nhất của kỹ thuật thuỷ vân số.

1.1.3 Mô hình kỹ thuật giấu tin

Giấu thông tin vào phương tiện chứa và tách lấy thông tin là hai quá trình trái ngược nhau và có thể mô tả qua sơ đồ khối của hệ thống như sau.

Mô hình kỹ thuật giấu tin cơ bản được trình bày trên hình 1.2.



Hình 1.2. Lược đồ chung cho quá trình giấu tin

Trên hình vẽ:

– Secret Message (M): Thông tin cần giấu tùy thuộc vào mục đích của người sử dụng, nó có thể là thông điệp hoặc các logo, các hình ảnh bản quyền.

– Cover Data (I): Dữ liệu phủ (môi trường sẽ giấu tin như: văn bản, audio, video, …).

– Embeding Algorithm (E): Thuật toán nhúng tin. Thông tin sẽ được giấu vào trong phương tiện chứa nhờ một bộ nhúng, bộ nhúng đó là chương trình, là những thuật toán để giấu tin và được thực hiện với khóa bí mật.

– Key (K): Khóa bí mật được sử dụng trong giấu tin.

– Stego Data (S): Dữ liệu mang tin mật (môi trường đã chứa tin mật).

– Control (C): Kiểm tra thông tin.

1.1.4 Mô hình kỹ thuật tách tin

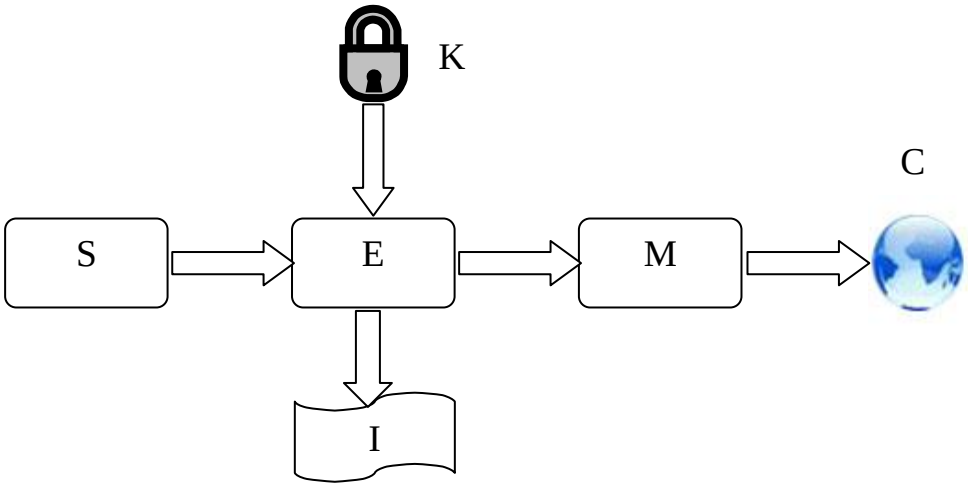
Tách thông tin từ các phương tiện chứa đã được giấu tin diễn ra theo quy trình ngược lại với đầu ra là thông tin đã được giấu vào phương tiện chứa.

Phương tiện chứa sau khi tách lấy thông tin có thể được sử dụng, quản lý theo những yêu cầu khác nhau.

Hình 1.3 dưới đây chỉ ra các công việc giải mã thông tin đã giấu.

Sau khi nhận được đối tượng phương tiện chứa có giấu thông tin, quá trình giải mã được thực hiện thông qua một bộ giải mã ứng với bộ giấu thông tin cùng với khoá của quá trình giấu.

Kết quả thu được gồm phương tiện chứa gốc và thông tin đã giấu. Bước tiếp theo thông tin đã giấu sẽ được xử lý kiểm định so sánh với thông tin ban đầu.



Hình 1.3. Lược đồ chung cho quá trình tách tin

1.2 Giấu tin trong ảnh

1.2.1 Khái niệm giấu tin trong ảnh

Giấu tin trong ảnh hiện đang rất được quan tâm. Nó đóng vai trò hết sức quan trọng trong hầu hết các ứng dụng bảo vệ an toàn thông tin như: nhận thực thông tin, xác định xuyên tạc thông tin, bảo vệ bản quyền tác giả.

Lượng thông tin mang ý nghĩa tùy thuộc vào mục đích của người sử dụng sẽ được giấu vào dữ liệu ảnh nhưng chất lượng ảnh ít thay đổi và khó có thể biết được đằng sau ảnh đó mang những thông tin có ý nghĩa gì.

1.2.2 Các yêu cầu đối với giấu tin trong ảnh

Những yêu cầu cơ bản đối với giấu tin cho ảnh là:

– Tính ẩn của giấu tin được chèn vào ảnh: Sự hiện diện của giấu tin trong ảnh không làm ảnh hưởng tới chất lượng của ảnh đã chèn tin.

Tính ẩn của tin là một yêu cầu rất quan trọng của phương pháp giấu tin. Nếu tính ẩn của tin không được đảm bảo thì không những làm ảnh hưởng tới chất lượng của ảnh mà còn dễ dàng tạo điều kiện cho các hình thức tấn công nhằm loại bỏ tin ra khỏi ảnh. Với ảnh được đánh dấu một cách lý tưởng thì ảnh có bản quyền và ảnh gốc sẽ không thể phân biệt được bằng mắt thường. Vì vậy giá trị của bức ảnh sẽ không bị thay đổi và sẽ là rào cản lớn cho những kẻ phá hoại muốn xoá hoặc sửa đổi các thông tin về bản quyền ảnh.

– Tính bền của giấu tin: Cho phép các tin có thể tồn tại được qua các phép biến đổi ảnh, biến dạng hình học hay các hình thức tấn công cố ý khác.

Tính bền của giấu tin liên quan đến việc tách tin từ một ảnh có bản quyền sau khi được đánh dấu có thể được đem ra xử lý để phục vụ cho các mục đích khác nhau như: nén ảnh, biến đổi hình học, lọc ảnh cải thiện ảnh, các biến đổi cố tình để xoá dấu tin ra khỏi ảnh,…v.v. Vấn đề được đặt ra liệu sau khi ảnh bị xử lý còn có thể tách được lượng tin ra khỏi ảnh không và tách được thì chất lượng của tin có đảm bảo tin cậy không, do đó khi chèn một dấu ẩn vào ảnh thì trước hết phải đảm bảo tính ẩn của nó.

– Tính an toàn: Không thể xoá được tin ra khỏi ảnh trừ khi ảnh được biến đổi tới mức không còn mang thông tin.

1.2.3 Những đặc trưng và tính chất của giấu tin trong ảnh

Giấu tin trong ảnh chiếm vị trí chủ yếu trong các kỹ thuật giấu tin, vì vậy mà các kỹ thuật giấu tin phần lớn cũng tập trung vào các kỹ thuật giấu tin trong ảnh.

Các phương tiện chứa khác nhau thì cũng sẽ có các kỹ thuật giấu khác nhau nên các kỹ thuật giấu tin trong ảnh thường chú ý những đặc trưng và các tính chất cơ bản sau đây.

1.2.3.1 Phương tiện có chứa dữ liệu tri giác tĩnh

Dữ liệu gốc ở đây là dữ liệu tĩnh, dù đã giấu thông tin vào trong ảnh hay chưa thì khi xem ảnh bằng thị giác, dữ liệu ảnh không thay đổi theo thời gian, điều này khác với dữ liệu âm thanh và dữ liệu băng hình vì khi nghe hay xem thì dữ liệu gốc sẽ thay đổi liên tục với tri giác của con người theo các đoạn, các bài, các ảnh…

1.2.3.2 Giấu tin phụ thuộc ảnh

Kỹ thuật giấu tin phụ thuộc vào các loại ảnh khác nhau. Chẳng hạn đối với ảnh đen trắng, ảnh xám hay ảnh màu thì cũng có những kỹ thuật riêng cho từng loại ảnh có những đặc trưng khác nhau.

1.2.3.3 Kỹ thuật giấu tin lợi dụng tính chất hệ thống thị giác của con người

Giấu tin trong ảnh ít nhiều cũng gây ra những thay đổi trên dữ liệu ảnh gốc. Dữ liệu ảnh được quan sát bằng hệ thống thị giác của con người, nên các kỹ thuật giấu tin phải đảm bảo một yêu cầu cơ bản là những thay đổi trên ảnh phải rất nhỏ, sao cho bằng mắt thường khó nhận ra được sự thay đổi đó vì có như thế thì mới đảm bảo được độ an toàn cho thông tin giấu.

1.2.3.4 Giấu thông tin không làm thay đổi kích thước ảnh

Các phép toán giấu tin sẽ được thực hiện trên dữ liệu của ảnh. Dữ liệu ảnh bao gồm phần header (là nơi lưu các thông tin về tệp, kích thước, và địa chỉ offset về vùng dữ liệu), bảng màu (có thể có) và dữ liệu ảnh. Khi giấu tin, các phương pháp giấu đều biến đổi giá trị của các bit trong dữ liệu ảnh trước hay sau khi giấu tin là như nhau.

1.2.3.5 Đảm bảo chất lượng sau khi giấu tin

Đây là một yêu cầu quan trọng đối với giấu tin trong ảnh. Sau khi giấu tin bên trong, ảnh phải đảm bảo được yêu cầu không bị biến đổi, để có thể không bị phát hiện dễ dàng so với ảnh gốc.

Yêu cầu này dường như khá đơn giản đối với ảnh màu hoặc ảnh xám bởi mỗi điểm ảnh được biểu diễn bởi nhiều bit, nhiều giá trị và khi thay đổi một giá trị nhỏ nào đó thì chất lượng ảnh thay đổi không đáng kể, thông tin giấu khó bị phát hiện, nhưng đối với ảnh đen trắng mỗi điểm ảnh chỉ là đen hoặc trắng, và nếu biến đổi một bit từ trắng thành đen và ngược lại mà không khéo thì sẽ rất dễ bị phát hiện.

1.2.4 Các tiêu chí đánh giá kỹ thuật giấu tin trong ảnh

1.2.4.1 Tính vô hình

Kỹ thuật giấu thông tin trong ảnh phụ thuộc rất nhiều vào hệ thống thị giác của con người. Tính vô hình hay không cảm nhận được của mắt người thường giảm dần ở những vùng ảnh có màu xanh tím, thủy vân ẩn thường được chọn giấu trong vùng này.

1.2.4.2 Khả năng giấu thông tin

Khả năng giấu thông tin (Hiding Capacity) hay lượng thông tin giấu được (dung lượng) trong một ảnh được tính bằng tỉ lệ giữa lượng thông tin giấu và kích thước của ảnh. Các thuật toán giấu tin đều cố gắng đạt được mục tiêu giấu được nhiều tin và gây nhiễu không đáng kể.

Thực tế, người ta luôn phải cân nhắc giữa dung lượng thông tin cần giấu với các tiêu chí khác như chất lượng (Quality), tính bền vững (Robustness) của thông tin giấu.

1.2.4.3 Chất lượng của ảnh có giấu thông tin

Chất lượng của ảnh có giấu tin được đánh giá qua sự cảm nhận của mắt người. Nên chọn những ảnh có nhiễu, có những vùng góc cạnh hoặc có cấu trúc, làm ảnh môi trường vì mắt thường ít nhận biết được sự biến đổi, khi có tin giấu, trên những ảnh này.

1.2.4.4 Tính bền vững của thông tin được giấu

Tính bền vững thể hiện qua việc các thông tin giấu không bị thay đổi khi ảnh mang tin phải chịu tác động của các phép xử lý ảnh như nén, lọc, biến đổi, tỉ lệ,…

1.2.4.5 Thuật toán và độ phức tạp của thuật toán

Cần nắm được một số kiến thức cơ bản về cấu trúc của ảnh để chọn ra thuật toán tìm miền ảnh thích hợp cho việc giấu tin. Độ phức tạp của thuật toán mã hóa và giải mã là yếu tố quan trọng để đánh giá các phương pháp giấu tin trong ảnh. Yêu cầu về độ phức tạp tính toán phụ thuộc vào từng ứng dụng.

Những ứng dụng theo hướng Watermark thường có thuật toán phức tạp hơn hướng Steganography.

CHƯƠNG 2: TỔNG QUAN VỀ ẢNH BITMAP

2.1 Giới thiệu ảnh BITMAP (BMP)

Ảnh BITMAP (BMP) được phát triển bởi Microsoft Corporation, được lưu trữ dưới dạng độc lập thiết bị cho phép Windows hiển thị dữ liệu không phụ thuộc vào khung chỉ định màu trên bất kỳ phần cứng nào.

Tên tệp mở rộng mặc định của một tệp ảnh Bitmap là BMP, nét vẽ được thể hiện là các điểm ảnh. Qui ước màu đen, trắng tương ứng với các giá trị 0, 1.

Ảnh BMP được sử dụng trên Microsoft Windows và các ứng dụng chạy trên Windows từ version 3.0 trở lên. BMP thuộc loại ảnh mảnh.

Các thuộc tính tiêu biểu của một tập tin ảnh BMP là:

– Số bit trên mỗi điểm ảnh thường được ký hiệu bởi n. Một ảnh BMP n bit có 2^n màu. Giá trị n càng lớn thì ảnh càng có nhiều màu và càng rõ nét hơn.

Giá trị tiêu biểu của n là 1 (ảnh đen trắng), 4 (ảnh 16 màu), 8 (ảnh 256 màu), 16 (ảnh 65536 màu) và 24 (ảnh 16 triệu màu). Ảnh BMP 24 - bit có chất lượng hình ảnh trung thực nhất.

– Chiều cao của ảnh (height), cho bởi điểm ảnh.

– Chiều rộng của ảnh (width), cho bởi điểm ảnh.

Đặc điểm nổi bật nhất của định dạng BMP là tập tin ảnh thường không được nén bằng bất kỳ thuật toán nào. Khi lưu ảnh, các điểm ảnh được ghi trực tiếp vào tập tin một điểm ảnh sẽ được mô tả bởi một hay nhiều byte tùy thuộc vào giá trị n của ảnh.

Do đó, một hình ảnh lưu dưới dạng BMP thường có kích cỡ rất lớn, gấp nhiều lần so với các ảnh được nén (chẳng hạn GIF, JPEG hay PNG).

2.2 Cấu trúc ảnh BITMAP (BMP)

Cấu trúc một tệp ảnh BMP gồm có bốn phần:

– Bitmap File Header: Lưu trữ thông tin tổng hợp về tệp ảnh BMP.

– Bitmap Information: Lưu trữ thông tin chi tiết về ảnh bitmap.

– Color Palette: Lưu trữ định nghĩa của màu được sử dụng cho bitmap.

– Bitmap Data: Lưu trữ từng điểm ảnh của hình ảnh thực tế.

2.2.1 Bitmap File Header

Đây là khối bytes ở phần đầu tập tin, sử dụng để định danh tập tin.

Ứng dụng đọc khối bytes này để kiểm tra xem đó có đúng là tập tin BMP không và có bị hư hỏng không.

Bảng 2.1. Chi tiết khối bytes tiêu đề tập tin BMP

Offset	Size	Mục đích
0000h	2 bytes	Magic number sử dụng để định nghĩa tập tin BMP: 0x42 0x4D (mã hexa của kí tự B và M). Các mục dưới đây có thể được dùng: • BM - Windows 3.1x, 95, NT, ... etc • CI - OS/2 Color Icon • CP - OS/2 Color Pointer
0002h	4 bytes	Kích thước của tập tin BMP theo byte
0006h	2 bytes	Dành riêng; giá trị thực tế phụ thuộc vào ứng dụng tạo ra hình ảnh.
0008h	2 bytes	Dành riêng; giá trị thực tế phụ thuộc vào ứng dụng tạo ra hình ảnh.
000Ah	4 bytes	Offset, địa chỉ bắt đầu các byte dữ liệu ảnh bitmap.

2.2.2 Bitmap Information

Khối bytes này cho biết các thông tin chi tiết về hình ảnh sẽ được sử dụng để hiển thị hình ảnh trên màn hình.

Bảng 2.2 dưới đây miêu tả chi tiết khối bytes thông tin tập tin BMP.

Bảng 2.2. Chi tiết khối bytes thông tin tập tin BMP

Offset	Size	Mục đích
Eh	4	Kích thước của tiêu đề (40 bytes).
12h	4	Chiều rộng bitmap tính bằng pixel (Signed interger).
16h	4	Chiều cao bitmap tính bằng pixel (Signed interger).
1Ah	2	Số lượng các mặt phẳng màu sắc được sử dụng. Phải được thiết lập bằng 1.
1Ch	2	Số bit trên mỗi pixel, là độ sâu màu của hình ảnh. Giá trị điển hình là 1, 4, 8, 16, 24 và 32.
1Eh	4	Phương pháp nén được sử dụng.
22h	4	Kích thước hình ảnh. Đây là kích thước của dữ liệu bitmap và không nên nhầm lẫn với kích thước tập tin.
26h	4	Độ phân giải theo chiều ngang của hình ảnh (Signed interger).
2Ah	4	Độ phân giải theo chiều dọc của hình ảnh (Signed interger).
2Eh	4	Số lượng màu trong bảng màu.
32h	4	Số lượng các màu sắc quan trọng được sử dụng, hoặc 0 khi màu sắc nào cũng đều là quan trọng, thường bị bỏ qua.

2.2.3 Color Palette

Tiếp theo vùng Info là Color Palette của BMP, gồm nhiều bộ có kích thước bằng 4 byte xếp liền nhau theo cấu trúc Blue – Green - Red và một byte dành riêng cho Itensity.

Kích thước của vùng Palette màu = 4 * số màu của ảnh.

Bảng màu xuất hiện trong tập tin BMP sau tiêu đề BMP và tiêu đề DIB. Vì vậy, offset là kích cỡ của tiêu đề BMP cộng với kích thước của tiêu đề DIB.

Vì Palette màu của màn hình có cấu tạo theo thứ tự Red – Green - Blue nên khi đọc Palette màu của ảnh BMP vào thì phải chuyển đổi cho phù hợp.

Số màu của ảnh được biết dựa trên số bit cho 1 pixel cụ thể là:

- 8 bits / pixel: ảnh 256 màu.
- 4 bits / pixel: ảnh 16 màu.
- 24 bits / pixel ảnh 24 bit màu.

Có tất cả 2 ^ 24 màu RGB khác nhau, nhưng các loại Bitmap 1 bit (2 màu, hoặc chuẩn Windows là trắng - đen), 4 bits (16 màu), 8 bits (256 màu) không thể khai thác hết, nên chỉ liệt kê các màu được dùng trong tệp. Mỗi màu trong bảng màu được mô tả bằng 4 bytes (BlueByte, GreenByte, RedByte và ReservByte).

Ví dụ:

Bảng màu loại 1 bit chuẩn Windows có 8 bytes: 0, 0, 0, 0, 255, 255, 255, 0 (4 bytes đầu là màu thứ 0, 4 bytes sau là màu thứ 1. Do chỉ có 0 và 1 nên mô tả mỗi điểm ảnh chỉ cần dùng 1 bit).

Tương tự như vậy, bảng màu của tệp 4 bits có 64 bytes, lần lượt từ màu số 0 đến màu số 15, bảng màu của tệp 8 bits có 1024 bytes (từ 0 đến 255).

Chính vì các màu được liệt kê như vậy nên các màu trong tệp 1 bit, 4 bits, 8 bits được gọi là Indexed, còn các màu trong tệp 24 bits được gọi là True.

2.2.4 Bitmap Data

Phần Bitmap Data nằm ngay sau phần Color Palette của ảnh BMP. Đây là phần chứa các giá trị màu của các điểm ảnh trong BMP.

Dữ liệu ảnh được lưu từng điểm cho đến hết hàng ngang (từ trái sang phải), và từng hàng ngang cho đến hết ảnh (từ dưới lên trên).

Mỗi byte trong vùng Bitmap Data biểu diễn 1 hoặc nhiều điểm ảnh tùy theo số bits cho một pixel.

Đối với mỗi điểm ảnh loại màu Indexed, ta cần 1 bit, 4 bits hoặc 8 bits để đặc trưng cho điểm đang xét ứng với màu thứ mấy trong bảng màu.

Ví dụ:

Giá trị 0111 (=7) trong loại BMP 4 bits cho biết điểm đó có màu 7 (màu xám theo chuẩn Windows). Riêng loại 24 bits thì không mô tả màu bằng thứ tự trên bảng màu (nếu liệt kê hết bảng màu của nó thì đã tốn cả Gigabyte bộ nhớ và đĩa) mà được liệt kê luôn giá trị RGB của 3 màu thành phần.

Ví dụ:

Trắng ={255, 255, 255}, Đen = {0, 0, 0}.

Như vậy, mỗi điểm ảnh loại 1 bit tốn 1/8 bytes (nói cách khác, 1 byte lưu được 8 điểm 1 bit), loại 4 bits tốn 1/2 byte, loại 8 bits tốn 1 byte và loại 24 bits tốn 3 bytes.

Tuy nhiên, tính chung cả bức ảnh thì khối dữ liệu không hoàn toàn tỉ lệ thuận như vậy mà thường lớn hơn một chút.

Lý do chính ở chỗ ta ngầm quy ước số bytes cần dùng cho 1 hàng ngang phải là bội của 4. Nếu bạn có ảnh 1 x 1, 1 bit, thì cũng tốn 66 bytes như ảnh 32 x 1, 1 bit (54 cho header, 8 cho bảng màu, 4 cho 1 hàng tối thiểu).

Nếu thử xoay bức hình 32 x 1 (vừa đúng 4 bytes dữ liệu) thành 1 x 32, sự lãng phí sẽ xuất hiện, lúc đó mỗi hàng sẽ lãng phí 31 bits, tổng cộng 32 lần như thế (31*4 bytes = 124 bytes).

CHƯƠNG 3: KỸ THUẬT PHÁT HIỆN ẢNH CÓ GIẤU TIN RS

3.1 Các vấn đề phát hiện ảnh có giấu tin [2]

3.1.1 Phân tích tin ẩn giấu (Steganalynis)

Steganalysis là kỹ thuật phát hiện sự tồn tại của thông tin ẩn giấu trong multimedia. Giống như thám mã, mục đích của steganalysis là phát hiện ra thông tin ẩn và phá vỡ tính bí mật của vật mang tin ẩn.

Phân tích ảnh có giấu thông tin thường dựa vào các yếu tố sau:

- Phân tích dựa vào các đối tượng đã mang tin.
- Phân tích bằng so sánh đặc trưng: So sánh vật mang tin chưa giấu tin với vật mang tin đã được giấu tin, đưa ra sự khác biệt giữa chúng.
- Phân tích dựa vào thông điệp cần giấu để dò tìm.
- Phân tích dựa vào các thuật toán giấu tin và các đối tượng giấu đã biết: Kiểu phân tích này phải quyết định các đặc trưng của đối tượng giấu tin, chỉ ra công cụ giấu tin (thuật toán) đã sử dụng.
- Phân tích dựa vào thuật toán giấu tin, đối tượng gốc và đối tượng sau khi giấu tin.

3.1.2 Các phương pháp phân tích

3.1.2.1 Phân tích trực quan

Đây là phương pháp đơn giản nhất mặc dù kết quả thường không được đáng tin cậy.

Để phát hiện khả năng một ảnh có giấu tin hay không bằng việc phân tích ảnh một cách trực quan và tìm kiếm những điểm bất thường.

Thật vậy, việc thay đổi bảng màu (của một ảnh màu) dù nhỏ để giấu thông điệp bí mật có thể dẫn đến kết quả là sự thay đổi màu sắc lớn trên ảnh gốc, đặc biệt là nếu ảnh gốc có chứa các màu sắc khác nhau ở mức độ cao.

Thường dựa vào quan sát hoặc dùng biểu đồ histogram giữa ảnh gốc và ảnh chưa giấu tin để phát hiện ra sự khác biệt giữa hai ảnh căn cứ đưa ra vấn đề nghi vấn. Với phương pháp phân tích này thường khó phát hiện với ảnh có độ nhiễu cao và kích cỡ lớn.

3.1.2.2 Phân tích định dạng ảnh

Phương pháp này rất rộng và thường dựa vào các dạng ảnh Bitmap để đoán nhận kỹ thuật giấu hay sử dụng, như các ảnh Bitmap thường hay sử dụng giấu trên miền LSB.

Có nhiều định dạng tệp tin ảnh khác nhau như BMP, GIF, JPEG. Mỗi loại có đặc điểm và cấu trúc định dạng tệp tin khác nhau. Do đó, khi thực hiện giấu tin, chẳng hạn giấu tin theo LSB, sẽ cho sự thay đổi trên ảnh kết quả ở các điểm ảnh khác nhau. Và khi thực hiện phát hiện ảnh giấu tin cũng vậy.

3.1.2.3 Phân tích thống kê

Theo Plitzman và Westfeld, lý thuyết thống kê có thể áp dụng để phân tích thống kê các cặp giá trị (cặp giá trị điểm ảnh) để tìm sự khác biệt ở bit LSB.

Trước khi giấu tin, trên ảnh chứa thông điệp (cover image) thì mỗi cặp hai giá trị là phân phối không đều. Sau khi giấu tin, giá trị trong mỗi cặp có xu hướng trở nên bằng nhau. Hơn nữa, nếu các kỹ thuật giấu tin mật giấu các bit thông điệp một cách tuần tự vào các điểm ảnh liên tiếp nhau, bắt đầu từ góc trên trái thì ta sẽ quan sát được sự thay đổi đột ngột trong các thống kê.

Đây là phương pháp sử dụng các lý thuyết thống kê và thống kê toán sau khi đã xác định được nghi vấn đặc trưng. Phương pháp này thường đưa ra độ tin cậy cao hơn và đặc biệt là cho các ảnh dữ liệu lớn.

3.2 Kỹ thuật phát hiện RS (Regular Singular) [2] [4]

3.2.1 Giới thiệu về kỹ thuật RS

Kỹ thuật RS (Regular Singular) được đề xuất bởi nhóm tác giả J.Fridich, M.Goljan và R.Du. Đây là một kỹ thuật phát hiện và ước lượng thông tin cho ảnh có giấu tin trên miền LSB.

Ý tưởng:

Chia miền dữ liệu ảnh thành các miền nhỏ hơn cùng kích cỡ không trùng khớp, sau đó phân các miền nhỏ này vào 3 miền: miền đều đặn R (Regular), miền dị thường S (Singular) và miền không sử dụng được U (Unusable) bằng một mặt nạ phụ trợ M, hàm Hamming f và hàm F_1, F_{-1}.

Dựa vào thống kê số miền của R và S nhóm tác giả thấy một mối quan hệ khác nhau của ảnh trước khi giấu tin và sau khi giấu tin, ước lượng tỷ lệ thay đổi của miền R và S chính là tỷ lệ thông tin giấu trên các miền LSB của ảnh.

3.2.2 Các định nghĩa về kỹ thuật RS

Giả sử ta có một ảnh có M × N điểm ảnh. Tập P là tập tất cả các giá trị điểm ảnh có trên ảnh. Với ảnh đa cấp xám 8 bit thì P = {0, 1, …., 255}.

Định nghĩa 1

Hàm Hamming xác định khoảng cách giữa các điểm trong một tập:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n-1} |x_i - x_{i+1}| \qquad (3.1)$$

Trong đó $x_1, x_2, \ldots, x_n$ là giá trị các điểm ảnh trên nhóm G.

Việc giấu tin LSB làm tăng nhiễu trên ảnh do đó ta hy vọng rằng giá trị của hàm f sẽ tăng (hoặc giảm) sau khi giấu tin LSB.

Định nghĩa 2

Việc giấu tin LSB sử dụng các kiểu hàm trộn (Flip) bit $F_m(x)$ với m = -1, 0, 1 và x là giá trị điểm ảnh. Cụ thể như sau:

F_1: $0 \leftrightarrow 1, 2 \leftrightarrow 3, 4 \leftrightarrow 5 \ldots, 254 \leftrightarrow 255.$ (3.2)

F_{-1}: $-1 \leftrightarrow 0, 1 \leftrightarrow 2, 3 \leftrightarrow 4, \ldots, 253 \leftrightarrow 254, 255 \leftrightarrow 256.$ (3.3)

Hay $F_{-1}(x) = F(x+1) - 1$ với mọi x

$F_0(x) = x$, với $\forall x \in P$.

Định nghĩa 3

Hàm F_1 và F_{-1} được áp dụng lên nhóm G ($x_1, x_2, x_3, \ldots, x_n$) với một mặt nạ M (M là một n – bộ với các thành phần nhận giá trị -1, 0 hoặc 1) được định nghĩa như sau:

$$F_M(G) = (F_{M(1)}(x_1), F_{M(2)}(x_2), \ldots, F_{M(n)}(x_n)) \qquad (3.4)$$

trong đó $M(i) \in \{-1, 0, 1\}$

Ví dụ:

Nếu điểm ảnh trong nhóm G có các giá trị là (39, 38, 40, 41) và một mặt nạ M = (1, 0, 1, 0) thì $F_M(G) = (F_1(39), F_0(38), F_1(40), F_0(41)) = (38, 38, 41, 41)$.

Định nghĩa 4

Cho một mặt nạ M, hàm F và hàm khoảng cách f, một nhóm G các điểm ảnh được phân lớp vào một trong ba lớp như sau:

$$G \in R \Leftrightarrow f(F_M(G)) > f(G). \qquad (3.5)$$

$$G \in S \Leftrightarrow f(F_M(G)) < f(G). \qquad (3.6)$$

$$G \in U \Leftrightarrow f(F_M(G)) = f(G). \qquad (3.7)$$

Trong đó R gọi là các nhóm đều đặn (Regular), S là các nhóm dị thường (Singular) và U là các nhóm không sử dụng (Unusable).

Định nghĩa 5

Ta gọi:

R_M là số các nhóm R với mặt nạ M không âm, $M \in \{0, 1\}$.

S_M là số các nhóm S với mặt nạ M không âm, $M \in \{0, 1\}$.

R_{-M} là số các nhóm R với mặt nạ M không dương, $M \in \{-1, 0\}$.

S_{-M} là số các nhóm S với mặt nạ M không dương, $M \in \{-1, 0\}$.

Giả thuyết thống kê của phương pháp này là trong một ảnh điển hình (chưa giấu thông tin) thì giá trị của R_M gần bằng giá trị của R_{-M} và tương tự giá trị của S_M gần bằng giá trị của S_{-M}.

$$R_M \cong R_{-M} \text{ và } S_M \cong S_{-M} \qquad (3.8)$$

3.2.3 Phương pháp phát hiện RS

Cho phép kí hiệu số miền đều đặn cho mặt nạ M là R_M và số miền dị thường cho mặt nạ M là S_M. Ta có $R_M + S_M \leq 1$ và $R_{-M} + S_{-M} \leq 1$ cho mặt nạ là âm. Các giả thiết thống kê về phương pháp phát hiện là một trong những hình ảnh tiêu biểu.

Giá trị của R_M xấp xỉ bằng giá trị của R_{-M}, và cũng tương tự cho S_M và S_{-M}

$$R_M \approx R_{-M} \text{ và } S_M \approx S_{-M}$$

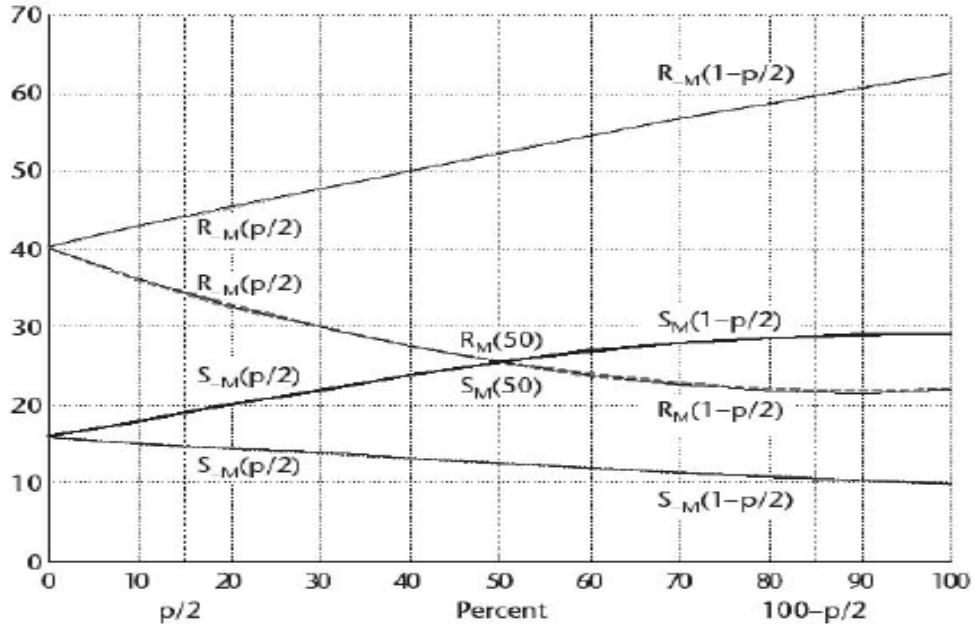
Áp dụng các thao tác trộn hàm F_{-1} cũng giống như hàm F_1 tới một ảnh màu, có thể thay đổi vị trí bằng 1.

Cho một hình ảnh tiêu biểu, số nhóm R và S thay đổi đáng kể khi thay đổi vị trí màu sắc bằng 1. Sự ngẫu nhiên của miền LSB tập trung vào sự khác biệt giữa R_M và S_M tới 0 như độ dài thông điệp m được nhúng tăng lên. Sau khi trộn LSB của 50% pixels thì ta được $R_M \approx S_M$.

Điều này tương đương với việc nói rằng khả năng nhúng miền LSB không mất thông tin là 0. Điều bất ngờ là ảnh hưởng của miền LSB thay đổi ngẫu nhiên cũng làm ảnh hưởng trực tiếp trên R_{-M} và S_{-M}.

Chúng làm tăng sự khác biệt cùng với độ dài thông điệp m được nhúng.

Đồ thị RS trong hình 3.1 dưới đây cho ta thấy R_M, S_M, R_{-M}, S_{-M} có chức năng như là số điểm ảnh được trộn với LSB. Giải thích riêng cho việc làm tăng sự khác biệt giữa R_{-M} và S_{-M} là sự bỏ qua cho tính không bền vững.



Hình 3.1. Đồ thị RS cho một hình ảnh tiêu biểu.
Trục x có liên quan số điểm ảnh được trộn với LSB, trục y có liên quan số miền đều đặn và dị thường với mặt nạ M và –M, M = [0 1 1 0].

Nguồn gốc của phương pháp RS là ước tính bốn đường cong của biểu đồ RS và tính toán khác, sự giao nhau của chúng bằng cách sử dụng ngoại suy. Thông thường hình dạng của bốn đường cong trong biểu đồ gần như là các đường thẳng hoàn toàn đối với ảnh gốc.

Những thực nghiệm chứng tỏ đường cong R_{-M} và S_{-M} là được biểu diễn với những đường thẳng hợp lý, trong khi bên trong đường cong R_M và S_M có thể được biểu diễn gần đúng hơn với phương trình bậc 2.

Các tham số của đường cong xác định từ các điểm đánh dấu trong hình 3.1.

Theo ước lượng ban đầu số nhóm R và S tương ứng với các điểm $R_M(p/2)$, $S_M(p/2)$, $R_{-M}(p/2)$, và $S_{-M}(p/2)$.

Giả định thông điệp là một chuỗi bit ngẫu nhiên trung bình chỉ có một nửa các pixels sẽ được trộn. Nếu trộn LSB của tất cả các pixels trong ảnh và tính toán được số nhóm R và S sẽ thu được bốn điểm sau: $R_M(1 - p/2)$, $S_M(1 - p/2)$, $R_{-M}(1 - p/2)$ và $S_{-M}(1 - p/2)$.

Bằng cách thay đổi ngẫu nhiên miền LSB của ảnh giấu thông tin thì sẽ thu được hai trung điểm R_M (1/2) và S_M (1/2). Có thể gắn các đường thẳng qua các đỉnh R_{-M} (p/2), R_{-M} (1 - p/2) và S_{-M} (p/2), S_{-M} (1 - p/2).

Các điểm R_M (p/2), R_M (1/2), R_M (1 - p/2) và S_M (p/2), S_M (1/2), S_M (1 - p/2) xác định được hai parabol.

Nó có thể tránh được thời gian sử dụng thống kê dự toán của trung điểm R_M (1/2) và S_M (1/2) tại cùng thời điểm, muốn làm cho độ dài thông điệp tăng lên hợp lý bằng cách chấp nhận thêm hai điều kiện (tự nhiên):

– Điểm giao nhau của đường cong R_M và R_{-M} như điểm giao nhau của đường cong S_M và S_{-M} có cùng trục tọa độ x.

– Giao điểm đường cong R_M và S_M tại m = 50%, hay $R_M(1/2) = S_M(1/2)$.

Giả thiết này tương đương với việc nói rằng khả năng thu nhận của miền LSB ngẫu nhiên được nhúng mà không mất thông tin là 0 [1].

Sự biến đổi nó có thể xuất phát từ một công thức đơn giản cho độ dài thông điệp p bí mật.

Sau khi định lại kích thước trục x từ mức p/2 → 0 và 100 - p/2 → 1, giao điểm của trục tọa độ x là một nghiệm của phương trình bậc hai:

$$2(d_1 + d_0)x^2 + (d_{-0} – d_{-1} – d_1 – 3d_0)x + d_0 – d_{-0} = 0 \quad (3.9)$$

$$d_0 = R_M (p/2) – S_M (p/2), \quad (3.10)$$

$$d_1 = R_M (1- p/2) – S_M (1 - p/2), \quad (3.10)$$

$$d_{-0} = R_{-M} (p/2) – S_{-M} (p/2), \quad (3.11)$$

$$d_{-1} = R_{-M} (1 - p/2) – S_{-M} (1 - p/2). \quad (3.12)$$

Độ dài thông điệp p là được tính toán từ nghiệm của x, có giá trị tuyệt đối nhỏ nhất của phương trình. Được tính bằng:

$$p = x / (x - 1/2).$$

Nó thỏa mãn để nói rằng các đường thẳng được xác định bởi số nhóm R và S tại p/2 và 1 - p/2.

3.2.4 Thuật toán RS

Đầu vào:

- I là một ảnh cấp xám.
- n: số phần tử của một nhóm.
- M_n: mặt nạ gồm các phần tử nhận giá trị trong tập {-1, 0, 1}.

Đầu ra :

Đánh giá S là một ảnh đã giấu tin hay chưa giấu tin.

Cách bước thực hiện

Bước 1: Đọc ảnh I.

Bước 2: Đọc giá trị điểm ảnh vào một ma trận $A_{M \times N}$.

Bước 3: $P = P \cup \{x_i\}$ với $x_i \in [0, 255]$.

Bước 4:

Chia ảnh I thành M × N/n nhóm khác nhau. Mỗi nhóm n điểm ảnh.

Với mỗi nhóm $G = (x_1, x_2, \ldots, x_n)$ ta thực hiện các bước sau.

Bước 5:

Để thực hiện việc phân dữ liệu ảnh thành miền có giá trị đều đặn và miền giá trị dị thường ta sử dụng hàm phụ trợ.

Tính hàm f (G):

$$f(G) = \sum_{i=1}^{n-1} |x_i - x_{i+1}|$$

Bước 6:

Cho mặt nạ $M = \{M(i)\}$ với i=1,…., n và $M(i) \in \{-1, 0, 1\}$. Tính:

$$F_M(G) = (F_{M(1)}(x_1), F_{M(2)}(x_2), \ldots\ldots, F_{M(n)}(x_n))$$

Bước 7:

Nhóm G được quyết định là thuộc nhóm nào trong 3 nhóm (R (Regular), S (Singular và U (Unusable)) khi và chỉ khi:

- $G \in R \Leftrightarrow f(F_M(G)) > f(G)$
- $G \in S \Leftrightarrow f(F_M(G)) < f(G)$
- $G \in U \Leftrightarrow f(F_M(G)) = f(G)$

Bước 8: Tính

- R_M = số các nhóm R với mặt nạ không âm, $M \in \{0, 1\}$.
- S_M = số các nhóm S với mặt nạ không âm, $M \in \{0, 1\}$.
- R_{-M} = số các nhóm R với mặt nạ không dương, $M \in \{-1, 0\}$.
- S_{-M} = số các nhóm S với mặt nạ không dương, $M \in \{-1, 0\}$.

Bước 9:

Nếu $| R_M | = | S_M |$ thì p =1

Ngược lại thực hiện các bước 9 đến bước 12.

Bước 10:

Tính các hệ số:

- $d_0 = R_M (p/2) – S_M (p/2)$;
- $d_0 = R_M (1 - p/2) – S_M (1 - p/2)$;
- $d_{-0} = R_{-M} (p/2) – S_{-M} (p/2)$;
- $d_{-1} = R_{-M} (1 - p/2) – S_{-M} (1 - p/2)$;

Bước 11:

Tính x_p là nghiệm của phương trình:

$$2(d_1 + d_0) x^2_p + (d_{-0} – d_{-1} - d_1 - 3d_0) x_p + d_0 – d_{-0} = 0$$

Bước 12:

Tính ước lượng độ dài thông điệp p:

$$P = x_p / (x_p - 1/2)$$

CHƯƠNG 4: KỸ THUẬT GIẤU TIN SES TRÁNH PHÁT HIỆN BẰNG RS

4.1 Giới thiệu kỹ thuật giấu tin SES

- Kỹ thuật SES được đề xuất bởi tác giả Jeong Jae Yu.
- Kỹ thuật giấu tin trong ảnh SES (Steganography Evading analyses) là kỹ thuật giấu tin tránh sự phát hiện bằng phương pháp thống kê RS [4].

4.2 Phương pháp giấu tin SES

4.2.1 Quá trình giấu tin [4]

Đầu vào:

- Ảnh gốc.
- Thông điệp cần giấu.

Đầu ra:

- Ảnh đã được giấu thông điệp.

Các bước thực hiện:

Bước 1:

- Đọc ảnh gốc P.
- Tính toán thống kê RS của ảnh gốc.

Bước 2:

- Đọc thông điệp.
- Chuyển thông điệp sang dạng số trong bảng mã Ascii.
- Ghép độ dài thông điệp vào phía trước tin đã được mã hóa.

Bước 3:

- Chuyển mã Ascii của thông điệp sang dạng nhị phân.

Bước 4:

Thực hiện giấu tin:

$$x' = \begin{cases} x & \text{if} \quad s_i = \text{LSB}(x_i) \\ F_r(x_i) & \text{otherwise} \quad r \in \{-1, 1\} \end{cases} \tag{4.1}$$

- Tính LSB của điểm ảnh.
- So sánh LSB của điểm ảnh với lần lượt các phần tử thông điệp đã được chuyển đổi sang nhị phân ở bước trên.

Nếu phần tử thông điệp bằng với LSB của điểm ảnh thì điểm ảnh sẽ được giữ nguyên. Ngược lại, nếu không bằng nhau thì xét tiếp:

- Nếu phần tử thông điệp = 1, LSB = 0 => Ta sử dụng hàm F_1 .
- Nếu phần tử thông điệp = 0, LSB = 1 => Ta sử dụng hàm F_{-1} .

Bước 5: Điều chỉnh các thống kê RS với các bộ phận không sử dụng nhúng trong ảnh gốc.

– Kiểm tra RS của ảnh đã được giấu tin.

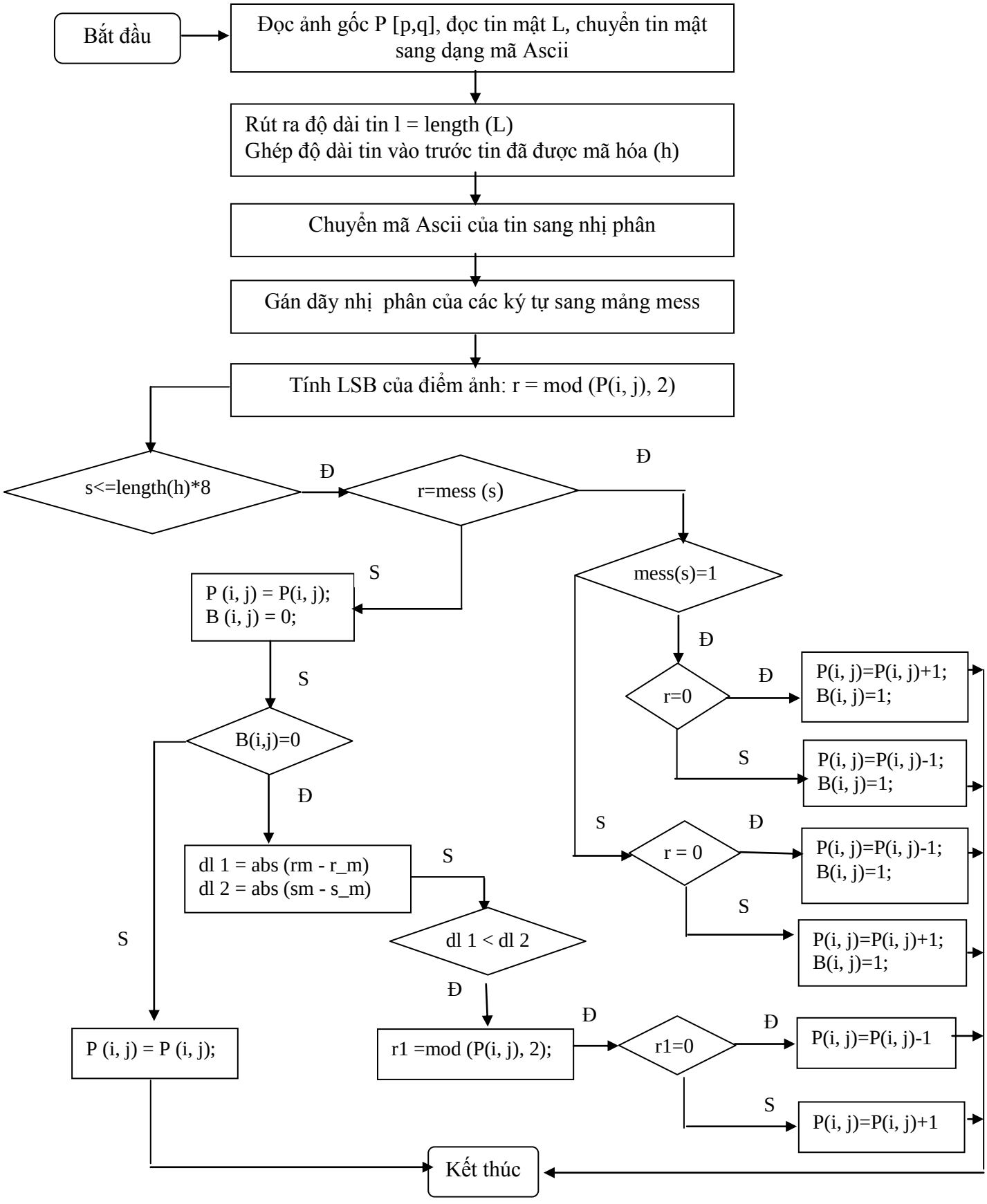
– Sử dụng một mảng để đánh dấu: những điểm ảnh nào được giấu tin thì gán bằng 1, những điểm ảnh nào chưa được giấu tin thì gán bằng 0.

– Trên phần mà những điểm ảnh gán bằng 0 thì ta điều chỉnh RS sao cho ảnh sau khi được giấu tin tránh phát hiện bằng thống kê RS.

Áp dụng:

$$F_M(R_M) \in S, F_M(R_M) \in R \qquad (4.2)$$

$$\Rightarrow R_M \cong R_{-M}, S_M \cong S_{-M} \qquad (4.3)$$



Hình 4.1. Sơ đồ giấu tin

4.2.2 Quá trình tách tin

Input:

– Ảnh đã được giấu tin.

Output:

– Ảnh gốc.

– Thông điệp được giấu.

Các bước thực hiện:

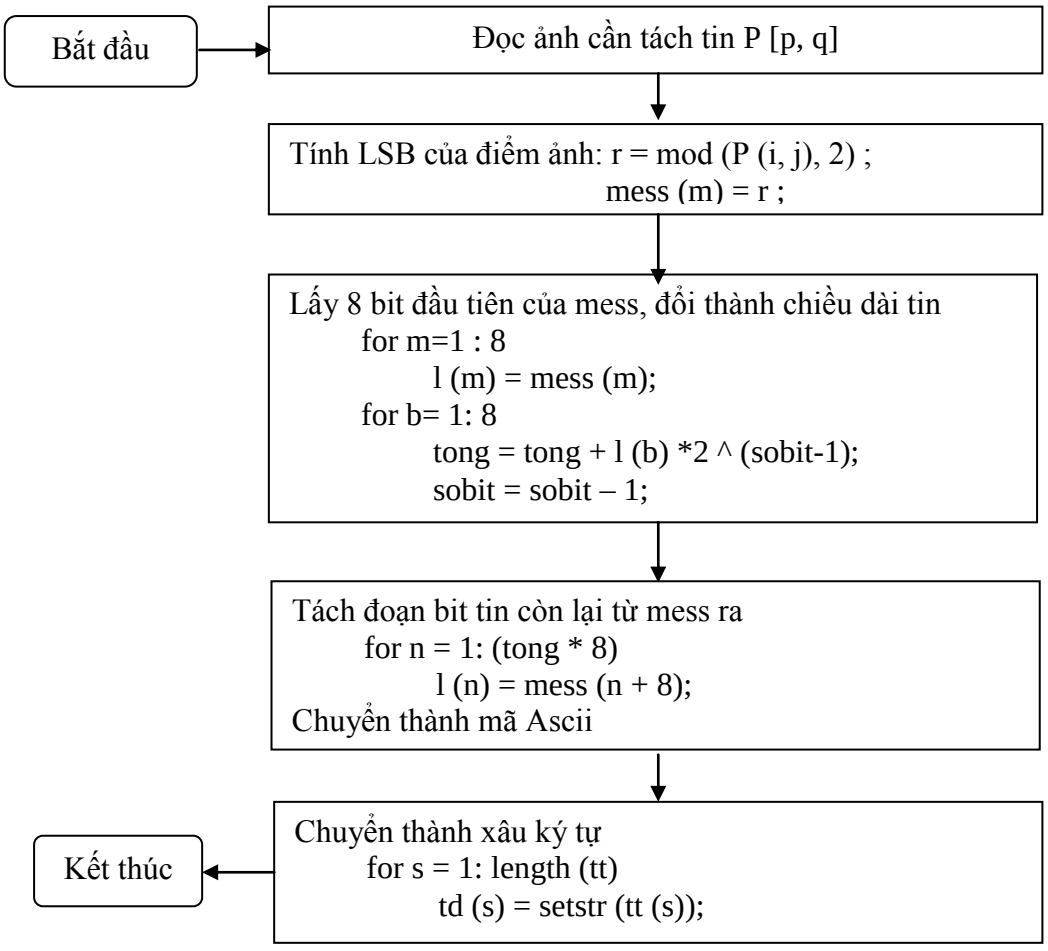
– Bước 1: Đọc ảnh giấu thông điệp.

– Bước 2: Lấy LSB của điểm ảnh cho vào mảng mess.

– Bước 3: Lấy 8 bit đầu tiên của mảng mess, tiến hành chuyển đổi từ số nhị phân sang dạng thập phân, được chiều dài của thông điệp.

– Bước 4: Những bit còn lại trong mảng mess chuyển thành số trong bảng mã Ascii.

– Bước 5: Chuyển từ số sang xâu kí tự , được thông điệp đã giấu.



Hình 4.2. Sơ đồ tách tin

CHƯƠNG 5: CÀI ĐẶT VÀ THỬ NGHIỆM

5.1 Môi trường thử nghiệm

Các thử nghiệm dùng để đánh giá thuật toán và kỹ thuật giấu và phát hiện được thực hiện trên môi trường MATLAP phiên bản 2008b.

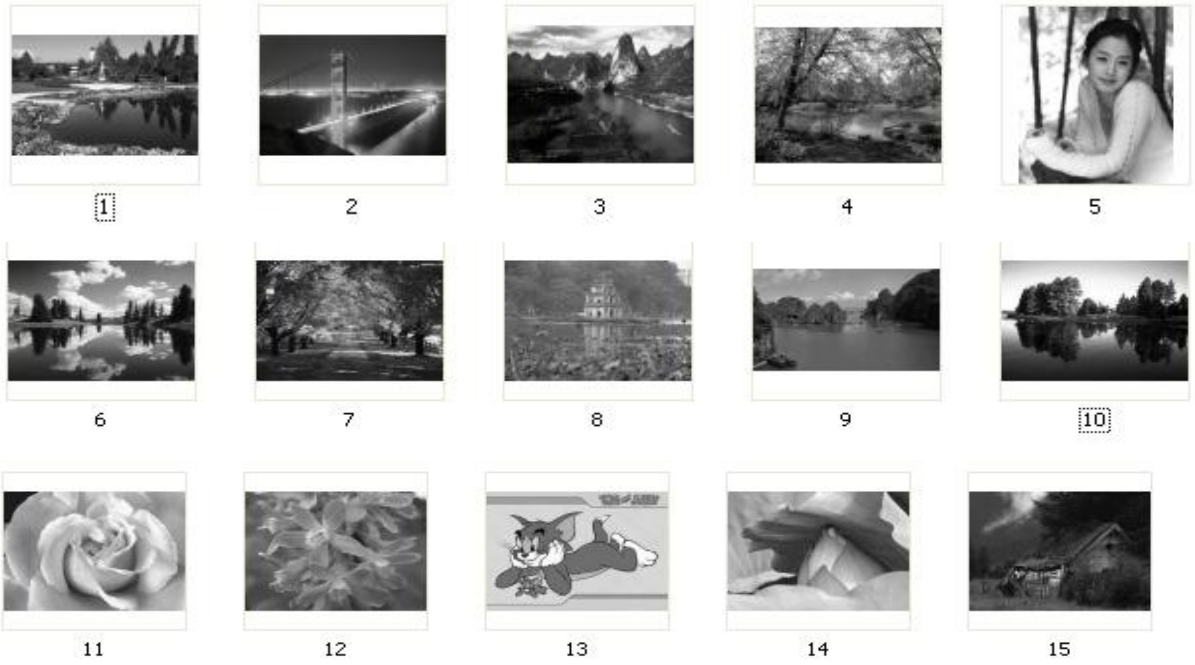
Chạy trên máy tính cấu hình Pentium (R) Dual-Core CPU T4200 2.00GHz, bộ nhớ trong 1 Gb, bộ nhớ ngoài có dung lượng trống khả dụng 15GB.

Quá trình thực nghiệm cần sự hỗ trợ của phần mềm xử lý ảnh Photoshop phiên bản CS2 8.0 để chuyển đổi dữ liệu ảnh từ màu sang ảnh cấp xám thuận tiện cho các thuận toán.

5.1.1 Tập ảnh thử nghiệm

Tập dữ liệu thử nghiệm gồm 5 ảnh chuẩn kích thước 512 x 512 trong Hình 5.1 và 15 ảnh được chụp từ máy ảnh kỹ thuật số, được chuyển đổi thành ảnh xám 8 bit bởi phần mềm Adobe Photoshop CS2 với nhiều kích cỡ khác nhau trong Hình 5.2 dưới đây.

Hình 5.1. 5 ảnh chuẩn



Hình 5.2. 15 ảnh chụp bằng máy ảnh kỹ thuật số với nhiều kích cỡ.

5.1.2 Đo độ đánh giá PSNR

Tỉ số tín hiệu cực đại trên nhiễu PSNR (peak signal-to-noise ratio), thường được viết tắt là PSNR, là một thuật ngữ dùng để tính tỉ lệ giữa giá trị năng lượng tối đa của một tín hiệu và năng lượng nhiễu ảnh hướng đến độ chính xác của thông tin.

PSNR được sử dụng để đo chất lượng tín hiệu khôi phục của các thuật toán nén có mất mát dữ liệu (ví dụ: dùng trong nén ảnh). Tín hiệu trong trường hợp này là dữ liệu gốc, và nhiễu là các lỗi xuất hiện khi nén.

Khi so sánh các thuật toán nén thường dựa vào sự cảm nhận gần chính xác của con người đối với dữ liệu được khôi phục, chính vì thế trong một số trường hợp dữ liệu được khôi phục của thuật toán này dường như có chất lượng tốt hơn những cái khác, mặc dù nó có giá trị PSNR thấp hơn (thông thường PSNR càng cao thì chất lượng dữ liệu được khôi phục càng tốt).

Nó được định nghĩa thông qua bình phương trung bình lỗi MSE (mean squared error) được dùng cho ảnh 2 chiều có kích thước $m \times n$ trong đó I và K là ảnh gốc và ảnh được khôi phục tương ứng:

$$MSE = \frac{1}{m\,n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \qquad (5.1)$$

PSNR được định nghĩa:

$$\begin{aligned} PSNR &= 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \\ &= 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \end{aligned} \qquad (5.2)$$

MAX_I là giá trị tối đa của pixel trên ảnh. Khi các pixel được biểu diễn bởi 8 bits, thì giá trị của nó là 255. Khi hai hình ảnh giống hệt nhau, MSE sẽ bằng 0 và PSNR đi đến vô hạn.

5.1.3 Áp dụng giấu tin trên ảnh

Hình 5.3. anh1.bmp

Bảng 5.1. Thống kê RS và PSNR của anh1.bmp

	anh1.bmp	kqanh1-10kytu.bmp	kqanh1-100kytu.bmp	kqanh1-600kytu.bmp
R_M	3829	3584	3600	3672
R_{-M}	3778	3677	3680	3721
S_M	2488	1244	1267	1375
S_{-M}	2534	1194	1218	1339
PSNR (dB)		50.7441	50.7393	50.7326

Nhận xét : Kỹ thuật giấu được lượng thông tin lớn, quá trình xử lý nhanh, chất lượng hình ảnh sau khi giấu tin là tốt (PSNR >35 dB).

5.1.4 Một số giao diện chương trình

Giao diện chương trình chính bao gồm:

- Hệ thống
    + Trợ giúp
    + Thoát
- Giấu thông điệp
    + Thực hiện giấu tin
    + Thống kê RS
    + Đánh giá PSNR
- Tách thông điệp
    + Thực hiện tách tin

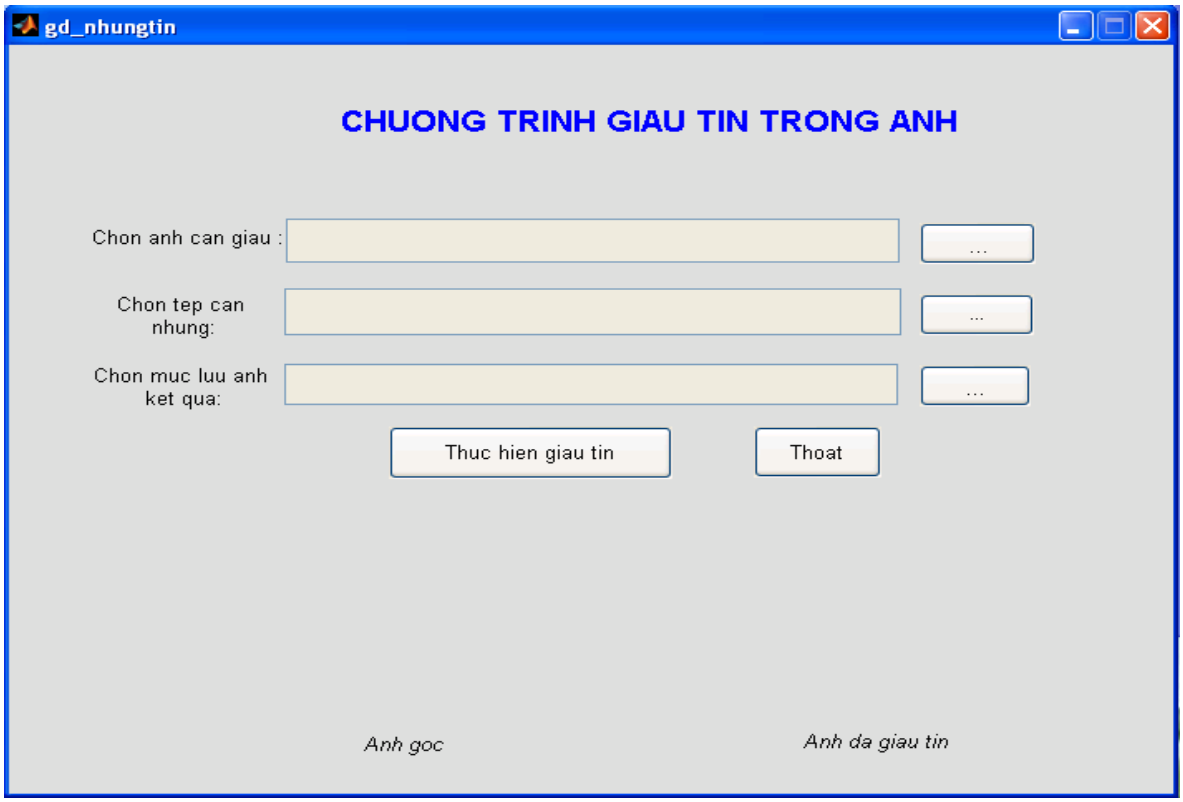
Hình 5.4. Trang chủ

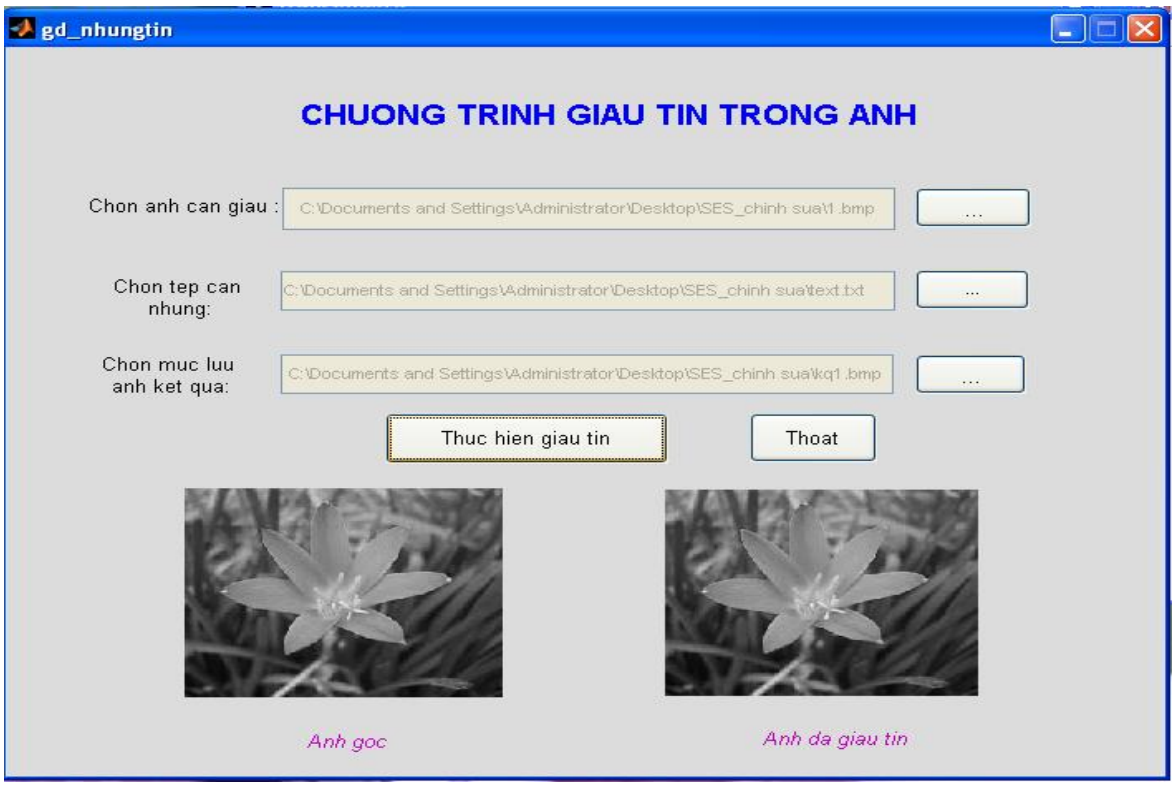
Hình 5.5. Chức năng Hệ thống

Hình 5.6. Chức năng Giấu thông điệp

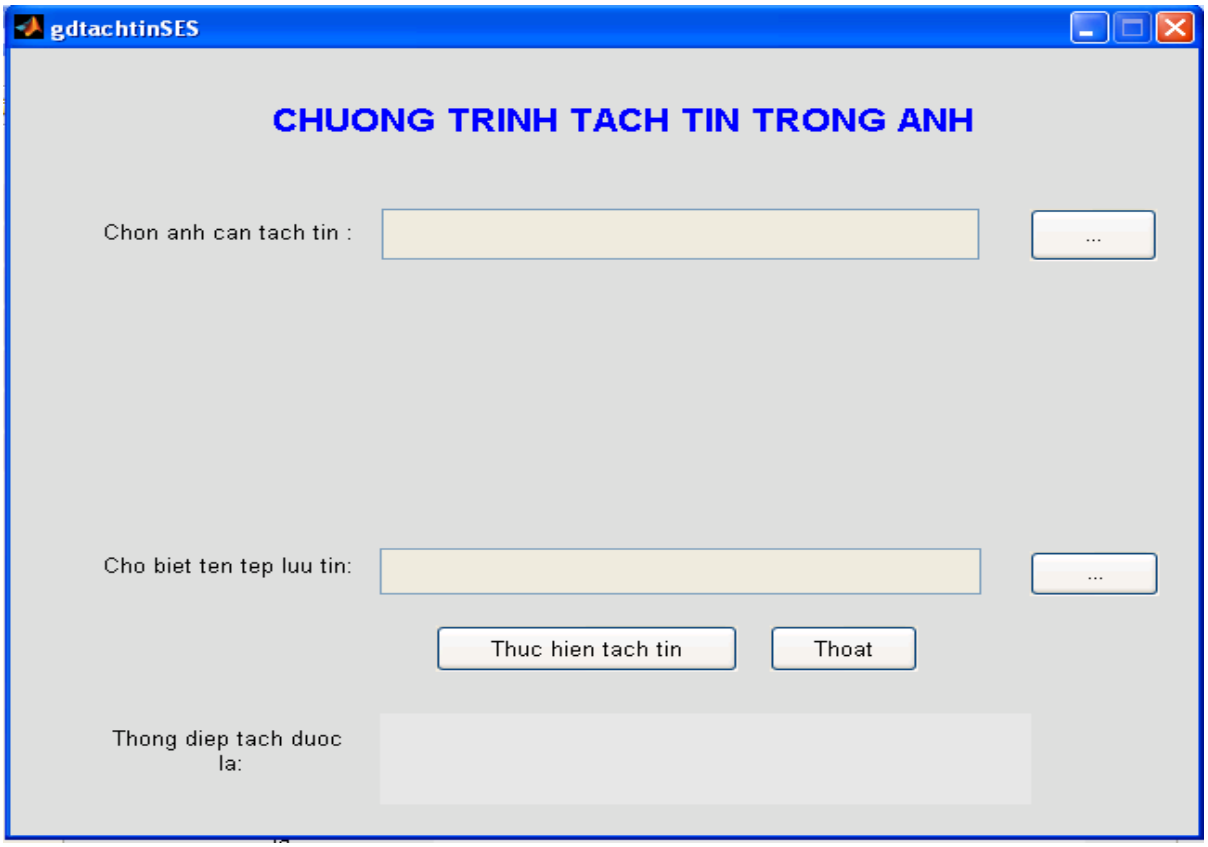
Hình 5.7. Chức năng Tách thông điệp

Hình 5.8. Trợ giúp

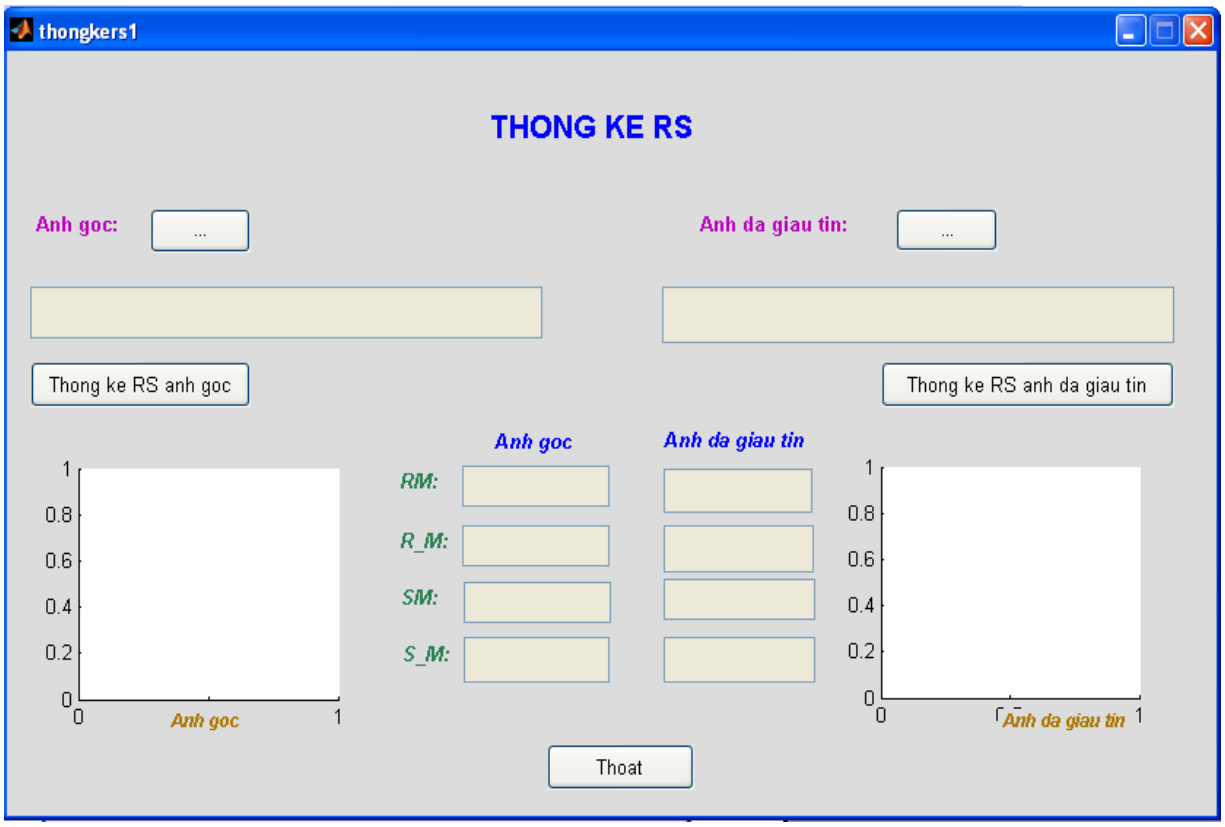




Hình 5.9. Chương trình giấu tin trong ảnh



Hình 5.10. Chương trình tách tin trong ảnh



Hình 5.11. Chương trình thống kê RS

Hình 5.12. Chương trình đánh giá PSNR

5.2 Các modul cài đặt

5.2.1 Chức năng: Thực hiện giấu tin trong ảnh

Ô nhập dữ liệu:

- Chọn ảnh cần giấu: click vào nút  bên cạnh để chọn ảnh giấu tin.
- Chọn tệp cần nhúng: click vào nút  bên cạnh để chọn tệp cần giấu tin.
- Chọn mục lưu ảnh kết quả: Lưu trữ ảnh kết quả sau khi giấu tin.

Nút bấm:

- Thực hiện giấu tin.
- Thoát: Thoát khỏi giao diện chương trình giấu thông điệp.

5.2.2 Chức năng: Thực hiện tách tin

Ô nhập dữ liệu:

- Chọn ảnh cần tách: click vào nút  bên cạnh để chọn ảnh tách tin.
- Cho biết tên tệp lưu tin: click vào nút  bên cạnh để chọn tên tệp cần lưu tin.

Nút bấm:

- Thực hiện tách tin.
- Thoát: Thoát khỏi giao diện chương trình tách thông điệp.

5.2.3 Chức năng: Đánh giá PSNR

Ô nhập dữ liệu:

- Chọn ảnh gốc: click vào nút  bên cạnh để chọn ảnh chưa giấu tin.
- Chọn ảnh nhúng: click vào nút  bên cạnh để chọn ảnh giấu tin.

Nút bấm:

- Thực hiện đánh giá PSNR.
- Thoát: Thoát khỏi giao diện chương trình đánh giá PSNR.

5.2.4 Chức năng: Thống kê RS

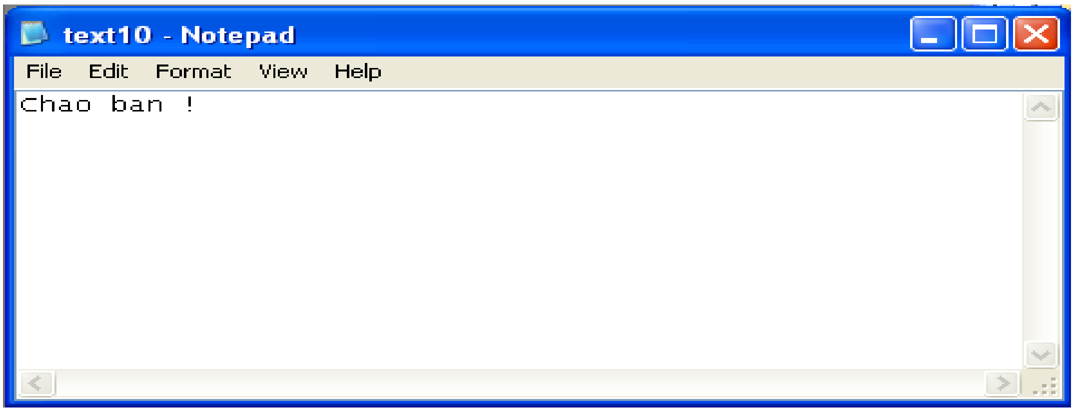
Ô nhập dữ liệu:

- Ảnh gốc: click vào nút  bên cạnh để chọn ảnh chưa giấu tin.
- Ảnh nhúng: click vào nút  bên cạnh để chọn ảnh giấu tin.

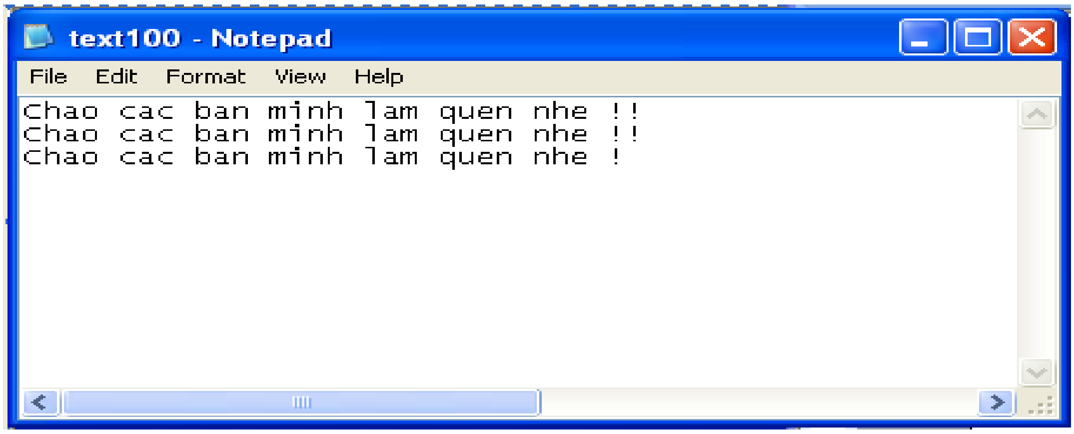
Nút bấm:

- Thống kê RS của ảnh gốc.
- Thống kê RS của ảnh đã giấu tin.
- Thoát: Thoát khỏi giao diện chương trình kiểm tra RS.

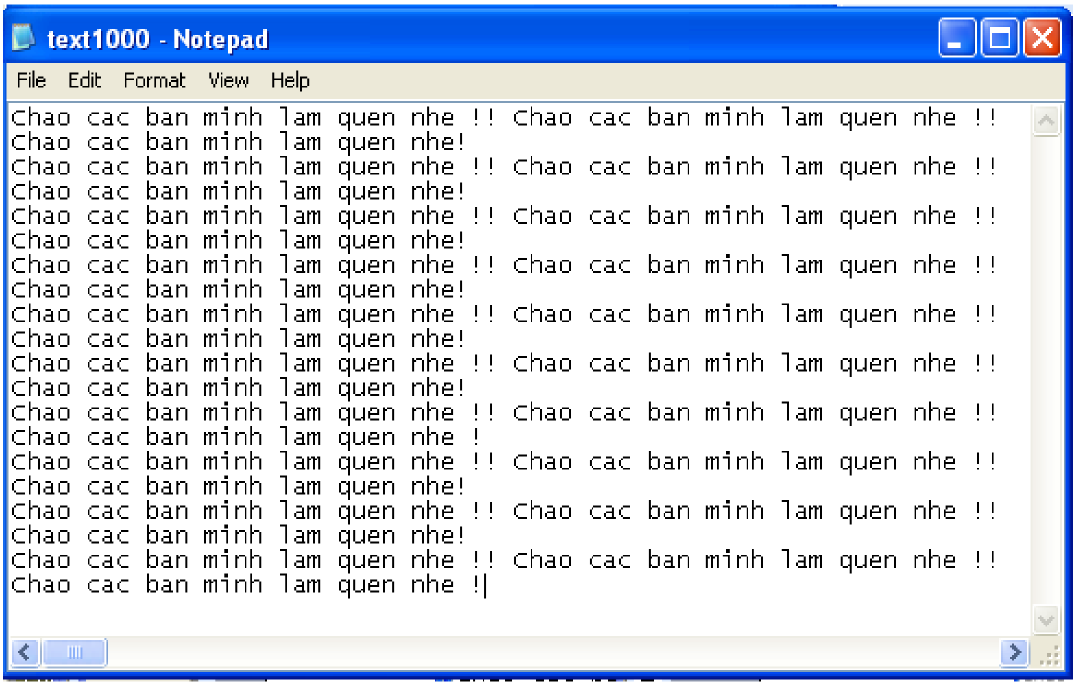
5.3 Thực nghiệm và đánh giá



Hình 5.13. Tệp thông điệp (10 ký tự)



Hình 5.14. Tệp thông điệp (100 ký tự)



Hình 5.15. Tệp thông điệp (1000 ký tự)

Hình 5.16. 5 ảnh chuẩn trước khi giấu và sau khi giấu

Bảng 5.2. Kết quả thực nghiệm trên 5 ảnh chuẩn

Lượng giấu **Ảnh giấu**	**PSNR**		
	10 kí tự	**100 kí tự**	**1000 kí tự**
Airplane.bmp	51.081	51.0694	50.9529
Barbara.bmp	51.132	51.1208	51.0018
Elanie.bmp	51.1388	51.1271	51.0122
Lake.bmp	51.1260	51.1138	50.9957
Pepper.bmp	50.6205	50.6086	50.4914

Hình 5.17. 15 ảnh bất kì trước khi giấu tin và sau khi giấu tin.

Bảng 5.3. Kết quả thực nghiệm trên 15 ảnh bất kỳ

Lượng giấu Ảnh giấu	PSNR		
	10 kí tự	**100 kí tự**	**1000 kí tự**
1	51.1315	51.0596	50.4984
2	51.1036	51.0379	50.4574
3	51.1721	51.1103	50.5271
4	51.1035	51.0416	50.4652
5	50.9447	50.8795	50.3334
6	51.1163	51.0592	50.4878
7	51.0658	50.9999	50.4233
8	51.1517	51.0858	50.5101
9	50.5113	50.4570	49.8941
10	51.1134	51.0530	50.4728
11	51.1666	51.1017	50.5096
12	51.0390	50.9729	50.3996
13	51.0443	50.9944	50.4150
14	50.9601	50.9000	50.3110
15	51.0622	51.0004	50.4302

Nhận xét :

Kỹ thuật giấu được lượng thông tin lớn, quá trình xử lý nhanh, chất lượng hình ảnh sau khi giấu tin là tốt (PSNR >35 dB). Giấu ảnh càng nhiều ký tự thì PSNR càng giảm, điều này chứng tỏ rằng chất lượng ảnh giấu thông điệp càng cao.

KẾT LUẬN

Đồ án của em đã thực hiện những nhiệm vụ sau:

1. Trình bày tổng quan kỹ thuật giấu tin.
2. Tổng quan về ảnh Bitmap.
3. Kỹ thuật phát hiện ảnh có giấu tin RS.
4. Kỹ thuật giấu tin SES tránh phát hiện bằng thống kê RS.

Đây là một kiến thức rất hữu ích và cần thiết để có thể khai thác ngày một hiệu quả các thành tựu của tin học. Đó cũng là một lý do để em chọn đề tài này làm đồ án tốt nghiệp. Em mong muốn giới thiệu và phổ biến những kiến thức rất cơ bản đến người đọc.

Việc kết hợp giấu thông tin và công nghệ thông tin là một vấn đề mới đang được nghiên cứu và phát triển để phục vụ nhiều lĩnh vực khác nhau. Trên thế giới người ta đã nghiên cứu nhiều về vấn đề này. Kỹ thuật giấu thông tin trong ảnh nói chung và giấu thông tin trong ảnh xám nói riêng là một hướng nghiên cứu chính của kỹ thuật giấu thông tin hiện nay và đã đạt nhiều kết quả khả quan.

Trong đề tài này em đã trình bày một số khái niệm liên quan đến kỹ thuật giấu tin nói chung và cụ thể là thuật toán giấu tin SES trong ảnh nói riêng.

Do còn nhiều hạn chế về thời gian nghiên cứu nên đề tài này không tránh khỏi những thiếu sót, vì vậy em rất mong nhận được sự đóng góp ý kiến của các thầy cô để đồ án được hoàn thiện.

Em xin chân thành cảm ơn!

TÀI LIỆU THAM KHẢO

Tài liệu Tiếng Việt

[1]. Nguyễn Xuân Huy, Trần Quốc Dũng, (2003), *Giáo trình giấu tin và thủy vân ảnh*, Trung tâm thông tin tư liệu, TTKHTN – CN.

[2]. Đỗ Thị Nguyệt (2009), *Đồ án tốt nghiệp*, Trường ĐHDL Hải Phòng.

Tài liệu Tiếng Anh

[3]. Jessica Fridrich, Miroslav Golian, Rui Du, *Reliable Detection of LSB Steganography in Color and Grayscale Images*, in preparation for the special is-sue on security in Magazine IEEE Multimedia.

[4]. Jeong-Jae Yu, *SES (Steganography Evading Statistical analyses),* CIST.