Tìm hiểu phương pháp thủy vân số thuận nghịch và xây dựng ứng dụng

Phạm Công Hòa

Trường Đại học Công nghệ
Chuyên ngành: Hệ thống thông tin; Mã số: 60 48 05
Cán bộ hướng dẫn khoa học: PGS.TS. Ngô Quốc Tạo
Năm bảo vệ: 2012

Abstract. Đưa ra các kiến thức cơ bản về kỹ thuật giấu tin và thủy vân, từ đó chỉ ra thủy vân là một nhánh của giấu tin. So sánh các phương pháp thủy vân trên miền không gian, miền tần số và thủy vân thuận nghịch (TVTN). Tuy nhiên, mục đích của thủy vân khác hoàn toàn so với mục đích của giấu tin mật và mã hóa. Phân tích các hướng ứng dụng quan trọng của thủy vân trong đời thường. Có nhiều môi trường đa phương tiện để thực hiện giấu tin và cũng có chừng đó môi trường để thực hiện thủy vân. Trình bày một số thuật toán thủy vân trên các miền: miền không gian, miền tần số dựa vào biến đổi Cosine rời rạc DCT và miền tần số dựa vào biến đổi sóng nhỏ rời rạc DWT. Phân tích và thiết kế các modul cho hệ thống, cài đặt thuật toán TVTN và chạy thử nghiệm chương trình. Thuật toán được lựa chọn cài đặt là TVTN CPT trên miền không gian và TVTN dựa vào DCT trên miền tần số

Keywords: Hệ thống thông tin; An toàn dữ liệu; Thủy vân số; Kỹ thuật giấu tin

Content.

CHƯƠNG 1: TỔNG QUAN VỀ KỸ THUẬT GIẤU TIN VÀ THỦY VÂN SỐ

1.1. Kỹ thuật giấu tin trong phương tiện số

1.1.1. Khái niệm giấu tin và lý thuyết cơ sở

Steganography có nguồn gốc từ tiếng Hy Lạp và được sử dụng tới ngày nay, nó có nghĩa là "tài liệu được phủ" (covered writing). Các câu chuyện kể về kỹ thuật giấu thông tin đã có từ rất lâu. Những tài liệu tìm thấy ghi chép về kỹ thuật giấu thông tin sớm nhất thuộc về sử gia Hy Lạp Herodotus (khoảng năm 440 trước Công nguyên). Khi bạo chúa Hy Lạp Histiaeus bị vua Darius bắt giữ ở Susa vào thế kỷ thứ 5 trước Công nguyên, ông ta đã cố gửi thông báo bí mật cho con rể của mình là Aristagoras ở Miletus. Histiaeus đã cạo trọc đầu của một nô lệ tin cậy và xăm một thông báo trên da đầu của người nô lệ đó. Khi tóc của người nô lệ mọc đủ dài, anh ta được gửi tới Miletus.

Giấu thông tin là một kỹ thuật nhúng (giấu) một lượng thông tin số nào đó vào một đối tượng dữ liệu số khác. Kỹ thuật giấu thông tin nhằm bảo đảm an toàn và bảo mật thông tin với hai mục đích. Một là bảo mật cho dữ liệu được đem giấu, hai là bảo vệ cho chính đối tượng dùng để giấu dữ liệu vào.Yêu cầu cơ bản của kỹ thuật giấu tin là không làm ảnh hưởng đến dữ liệu gốc.

Kỹ thuật giấu thông tin nhằm mục đích đảm bảo an toàn và bảo mật mật thông tin ở cả hai khía cạnh. Một là bảo mật cho dữ liệu được đem giấu, hai là bảo mật cho chính đối tượng được dùng để giấu tin. Điều này dẫn đến hai khuynh hướng chủ yếu của giấu tin:

- Khuynh hướng thứ nhất là giấu tin mật (steganography). Khuynh hướng này tập trưng vào các kỹ thuật giấu tin sao cho thông tin giấu được nhiều và quan trọng là người khác khó phát hiện dược một đối tượng có bị giấu tin bên trong hay không.
- Khuynh hướng thứ hai là thủy vân số (watermarking). Khuynh hướng thủy vân số đánh giấu vào đối tượng nhằm khẳng định bản quyền quyền sở hữu hay phát hiện sự xuyên tạc thông tin.

1.1.2. Phân loại các kỹ thuật giấu tin

Hai mục đích khác nhau của kỹ thuật giấu tin dẫn đến hai hướng kỹ thuật chủ yếu là giấu tin mật và thủy vân. Giấu tin mật là kỹ thuật giấu một lượng thông tin lớn vào một dữ liệu chứa nào đó sao cho người khác khó phát hiện được một đối tượng có giấu tin bên trong hay không nhằm bảo vệ lượng thông tin đem nhúng. Đồng thời, các kỹ thuật giấu tin mật còn quan tâm lượng tin có thể được giấu, lượng thông tin giấu được càng nhiều càng tốt. Tuy nhiên, lượng thông tin giấu càng lớn thì tính ẩn của thông tin giấu càng thấp.

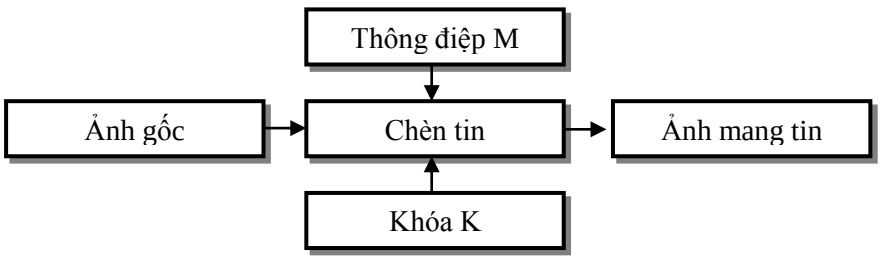
Các thành phần trong quá trình giấu tin:

- Dữ liệu nguồn (Host Signal): là dữ liệu gốc được dùng làm nơi giấu dữ liệu. Ví dụ nếu giấu tin trong bức ảnh thì bức ảnh là dữ liệu nguồn.
- Dữ liệu nhúng (Embed Data): là dữ liệu cần giấu, nó được nhúng vào dữ liệu nguồn, còn gọi là phương tiện giấu tin.
- Khoá và chìa (nếu cần): để mã hóa thông tin trước khi giấu vào ảnh.
- Dữ liệu mang thông tin ẩn: là sản phẩm của quá trình giấu tin.

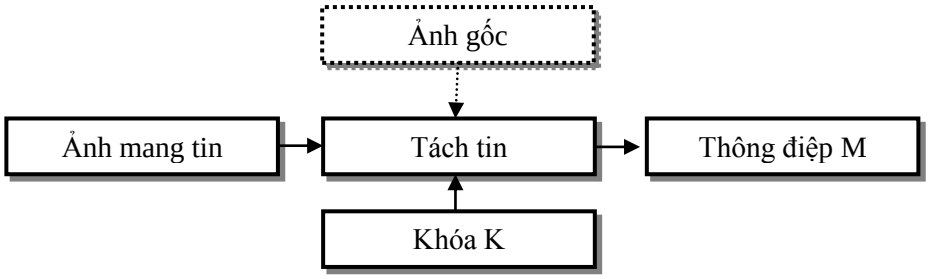
1.1.3. Mô hình kỹ thuật giấu tin

Hệ thống giấu tin nói chung bao gồm 2 phần chính: chèn tin và tách tin

Giai đoạn chèn tin, các thông tin khoá (công khai hoặc bí mật) và tin giấu được chèn vào ảnh gốc để được sản phẩm mang tin giấu. Giai đoạn tách tin, dữ liệu, khoá (bí mật) và hoặc ảnh gốc (ảnh không chèn tin) sẽ làm dữ liệu cơ sở để tách tin từ sảm phẩm mang tin giấu.



a) Chèn



b) Tách tin

Hình 1.1: Quá trình chèn (giấu) tin và tách tin

Trong đó:

- Mẩu tin mật: có thể là văn bản hoặc tệp ảnh hay bất kỳ một tệp nhị phân nào, vì quá trình xử lý chúng ta đều chuyển chúng thành chuỗi các bit.
- Ảnh phủ hay ảnh gốc: ảnh được dùng để làm môi trường nhúng tin mật.
- Khoá K: khoá viết mật tham gia vào quá trình giấu tin, tăng tính bảo mật.
- Ảnh mang tin: là ảnh sau khi đã nhúng tin mật vào đó.

1.1.4. Giấu tin trong dữ liệu đa phương tiện

Kỹ thuật giấu tin đã được nghiên cứu và áp dụng trong nhiều môi trường dữ liệu khác nhau như trong dữ liệu đa phương tiện (văn bản, hình ảnh, âm thanh, phim), trong sản phẩm phần mềm và gần đây là những nghiên cứu trên lĩnh vực cơ sở dữ liệu quan hệ. Trong các dữ liệu đó, dữ liệu đa phương tiện là môi trường chiếm tỷ lệ chủ yếu trong các kỹ thuật giấu tin.

1.1.4.1. Giấu tin trong ảnh số

Thông tin sẽ được giấu cùng với dữ liệu ảnh nhưng chất lượng ảnh ít thay đổi và không ai biết được đằng sau đó mang những thông tin có ý nghĩa. Ngày nay, khi ảnh số đã được sử dụng phổ biến, giấu thông tin trong ảnh đã đem lại nhiều những ứng dụng quan trọng trên nhiều lĩnh vực trong đời sống xã hội. Ví dụ đối với các nước phát triển, chữ ký tay đã được số hóa và lưu trữ sử dụng như hồ sơ các nhân của các dịch vụ ngân hàng và tài chính, nó được dùng để xác thực trong các thẻ tín dụng của người tiêu dùng. Các kỹ thuật giấu tin trong ảnh hiện nay đều thuộc vào một trong 3 nhóm:

- Giấu tin trong miền quan sát.
- Các phương pháp dựa vào kỹ thuật biến đổi ảnh.
- Các phương pháp sử dụng mặt nạ giác quan.

1.1.4.2. Giấu tin trong Audio

Giấu thông tin trong audio mang những đặc điểm riêng khác với giấu thông tin trong các đối tượng đa phương tiện khác. Một trong những yêu cầu cơ bản của giấu tin là đảm bảo tính chất ẩn của thông tin được giấu đồng thời không làm ảnh hưởng nhiều đến chất lượng của dữ liệu gốc. Để đảm

bảo yêu cầu này, kỹ thuật giấu thông tin trong ảnh phụ thuộc vào hệ thống thị giác của con người, còn kỹ thuật giấu thông tin trong audio lại phụ thuộc vào hệ thống thính giác.

Mã hóa bít thấp: Cũng như các file ảnh, phương pháp chèn vào các bit ít quan trọng cũng lưu trữ dữ liệu giấu vào trong các bit ít quan trọng của file audio. Phương pháp mã hóa LSB là cách đơn giản nhất để nhúng thông tin vào trong dữ liệu audio. Phương pháp này sẽ thay thế bít ít quan trọng nhất (thường là bít cuối) của mỗi mẫu dữ liệu bằng bit thông tin giấu.

Mã hóa pha là kỹ thuật thực hiện giấu tin trong audio thông qua việc thay thế pha của một segment audio ban đầu bằng một pha tham chiếu (referency phase) thể hiện dữ liệu. Pha của các segment tiếp theo sẽ được điều chỉnh sao cho duy trì mối quan hệ giữa các đoạn.

Kỹ thuật giấu dựa vào tiếng vang thực hiện giấu tin bằng cách thêm vào tiếng vang trong tín hiệu gốc. Dữ liệu nhúng được giấu bằng cách thay đổi 3 tham số của tiếng vang: biên độ ban đầu, tỉ lệ phân rã và độ trễ. Khi thời gian giữa tín hiệu gốc và tiếng vang giảm xuống, hai tín hiệu có thể trộn lẫn và người nghe khó có thể phân biệt giữa hai tín hiệu. Số lượng tin giấu có liên quan đến thời gian trễ của tiếng vang và biên độ của nó.

1.1.4.3. Giấu tin trong Video

Cũng giống như giấu tin trong ảnh hay trong audio, giấu tin trong video cũng được quan tâm và được phát triển mạnh mẽ cho nhiều ứng dụng như kiểm soát sao chép thông tin, nhận thực thông tin và bảo vệ bản quyền tác giả. Nhiều kỹ thuật giấu tin trong ảnh được áp dụng cho giấu tin trong video nhưng đảm bảo các ràng buộc:

- Do dữ liệu video là rất lớn nên thời gian để giấu tin cũng rất lớn. Vì vậy, việc giấu tin phải được thực hiện trong thời gian thực để truyền các tín hiệu video đi.
- Do giới hạn của băng thông nên việc giấu tin không được làm thay đổi kích thước của dữ liệu.
- Các dữ liệu video thường được lưu ở dạng nén, thông thường dựa vào quá trình xử lý sự thay đổi ảnh từ khung hình này đến khung hình khác. Vì vậy, khi giấu tin cần đảm bảo rằng thông tin không quá dễ phát hiện ra bằng mắt thường.

1.1.4.4. Giấu tin trong văn bản

Trong trao đổi thông tin qua hệ thống máy tính, văn bản chiếm một tỷ lệ rất lớn so với các loại phương tiện chứa khác. Tuy vậy, giấu tin trong văn bản lại chưa được quan tâm nghiên cứu nhiều. Các nghiên cứu về giấu tin trong văn bản được chia theo hai hướng:

Thứ nhất, văn bản được sử dụng để giấu tin là những văn bản được chụp lại và lưu trên máy như một bức ảnh nhị phân. Theo hướng này, các kỹ thuật giấu tin được thực hiện như kỹ thuật giấu tin trong ảnh.

Thứ hai, phương tiện chứa sử dụng cho quá trình giấu tin được lưu dưới dạng văn bản. Theo hướng này, giấu tin được thực hiện bằng cách điều chỉnh khoảng cách giữa các dòng hoặc thay đổi kích thước một số ký tự tại một số vị trí trên văn bản mà không làm ảnh hưởng nhiều đến nội dung văn bản gốc.

1.2. Kỹ thuật thủy vân số

1.2.1. Lịch sử phương pháp thủy vân số

Phương pháp thủy vân đầu tiền được thực hiện là thủy vân trên giấy. Đó là một thông tin nhỏ được nhúng chìm trong giấy để thể hiện bản gốc hoặc bản chính thức. Theo Hartung và Kutter, thủy vân trên giấy đã bắt đầu được sử dụng vào năm 1292 ở Fabriano, Italy – nơi được coi là nơi sinh của thủy vân. Sau đó, thủy vân đã nhanh chóng lan rộng trên toàn Italy và rồi trên các nước châu Âu và Mỹ. Ban đầu, thủy vân giấy được dùng với mục đích xác định nhãn hàng và nhà máy sản xuất. Sau này được sử dụng để xác định định dạng, chất lượng và độ dài, ngày tháng của sản phẩm.

Đến thế kỷ thứ 18, nó bắt đầu được dùng cho tiền tệ và cho đến nay thủy vân vẫn là một công cụ được dùng rộng rãi với mục đích bảo mật cho tiền tệ, chống làm tiền giả. Thuật ngữ "thủy vân" (watermarking) được đưa ra vào cuối thế 18, nó bắt nguồn từ một loại mực vô hình khi viết lên giấy và chỉ hiển thị khi nhúng giấy đó vào nước. Năm 1988, Komatsu và Tominaga đã đưa ra thuật ngữ "thủy vân số" (Digital watermarking).

Vậy thủy vân số là quá trình sử dụng các thông tin (ảnh, chuỗi bít, chuỗi số) nhúng một cách tinh vi vào dữ liệu số (ảnh số, audio, video hay text) nhằm xác định thông tin bản quyền của tác phẩm số. Mục đích của thủy vân số là bảo vệ bản quyền cho phương tiện dữ liệu số mang thông tin thủy vân. Tùy theo mục đích của hệ thủy vân mà người ta lại chia thành các hướng nhỏ như thủy vân dễ vỡ và thủy vân bền vững.

Thủy vân bền vững quan tâm nhiều đến việc nhúng những mẩu tin đòi hỏi độ bền vững cao của thông tin được giấu trước các biến đổi thông thường trên dữ liệu chứa. Hướng này được sử dụng để bảo vệ bản quyền tác giả.

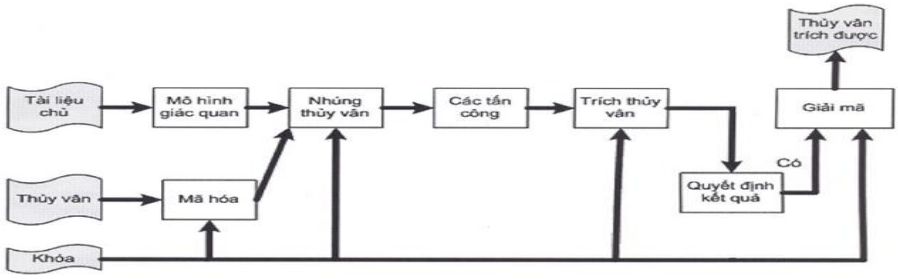
Thủy vân dễ vỡ yêu cầu thông tin giấu sẽ bị sai lệch nếu có bất kỳ sự thay đổi nào trên dữ liệu chứa. Hướng này được sử dụng để phát hiện xuyên tạc thông tin.

Ở mỗi loại thủy vân bền vững hoặc thủy vân dễ vỡ lại chia thành hai loại dựa theo đặc tính đó là thủy vân ẩn và thủy vân hiện. Thủy vân hiện cho phép nhìn thấy thông tin đem nhúng vào dữ liệu

chứa. Loại này được sử dụng cho mục đích công bố công khai về quyền sở hữu. Ngược lại, thủy vân ẩn không cho phép nhìn thấy nội dung thông tin nhúng và nó được sử dụng với mục đích gài bí mật các thông tin xác nhận quyền sở hữu.

Hình 1.3: Ví dụ về thủy vân ẩn và thủy vân hiện

1.2.2. Mô hình thủy vân số



Hình 1.4: Mô hình thủy vân do Sviatoslav Voloshynovskiy và các cộng sự đề xuất

Mô Mô hình thủy vân Sviatoslav Voloshynovskiy được chia làm 3 phần chính:

- Nhúng thủy vân
- Các tấn công trên thủy vân
- Trích thủy vân

Thủy vân có thể được mã hóa để tăng cường tính bền vững. Thông thường, tài liệu đã nhúng thủy vân sẽ trải qua một số bước tấn công trước khi được trích thủy vân. Sau quá trình trích được thực hiện, dựa vào kết quả trích rút để có quyết định tài liệu có được nhúng thủy vân hay không, và nếu có thì thực hiện giải mã trên dữ liệu trích để nhận được giá trị thủy vân.

1.2.3. Các tính chất quan trọng của kỹ thuật thủy vân số

Tính bền vững: Chất lượng của thuật toán phụ thuộc vào tính bền vững của thủy vân. Đặc biệt đối với thủy vân bền vững, yêu cầu quan trọng là thủy vân không bị thay đổi sau một số phép xử lý trên đối tượng được nhúng. Đối với ảnh số, các phép xử lý này có thể là phép nén thông tin, lọc, tính tiến, quay, làm sắc ảnh, xén ảnh,…

Tính vô hình: Đối với thủy vân ẩn thì mọi thuật toán đều cố gắng nhúng thủy vân sao cho chúng không bị phát hiện bởi người sử dụng. Thông thường đối với một thuật toán nếu tính bền vững cao thì tính vô hình kém và ngược lại, do đó cần có sự cân nhắc giữa tính bền vững và tính vô hình để đảm bảo thủy vân đạt được cả tính bền vững cũng như tính vô hình.

Tính bảo mật: Bảo mật đối với khóa, thủy vân sao cho nếu một ai đó không có quyền thì không thể dò được thủy vân.

1.2.4. Hệ thống thủy vân số

Hệ thống thủy vân số là quá trình sử dụng một thủy vân nhúng vào trong một dữ liệu số để được một dữ liệu số có chứa thủy vân hay gọi là dữ liệu có bản quyền. Dữ liệu có bản quyền này sẽ được phân phối trên kênh truyền tin. Trong quá trình phân phối, dữ liệu bản quyền có thể bị tấn công trái phép hoặc yếu tố gây nhiễu. Nếu dữ liệu số bản quyền bị nghi ngờ sao chép trái phép hoặc chỉnh sửa thông tin thì có thể xác minh nhờ quá trình tách thủy vân đã nhúng. Như vậy, hệ thống thủy vân số nói chung bao gồm 2 quá trình là quá trình nhúng thủy vân và quá trình tách thủy vân.

Thủy vân mang thông tin bảo mật hoặc bản quyền về dữ liệu chứa.

Khóa thủy vân được dùng cho cả phiên nhúng và phát hiện thủy vân. Khóa thủy vân là duy nhất với mỗi thủy vân. Khóa đó là khóa bí mật, chỉ tác giả mới biết. Điều đó nói lên rằng chỉ tác giả mới phát hiện ra được thủy vân. Tùy từng bộ nhúng thủy vân mà có các yêu cầu với khóa thủy vân.

1.2.4.1. Quá trình nhúng thủy vân

Giai đoạn này gồm thông tin khóa thủy vân, thủy vân, dữ liệu chứa và bộ nhúng thủy vân.

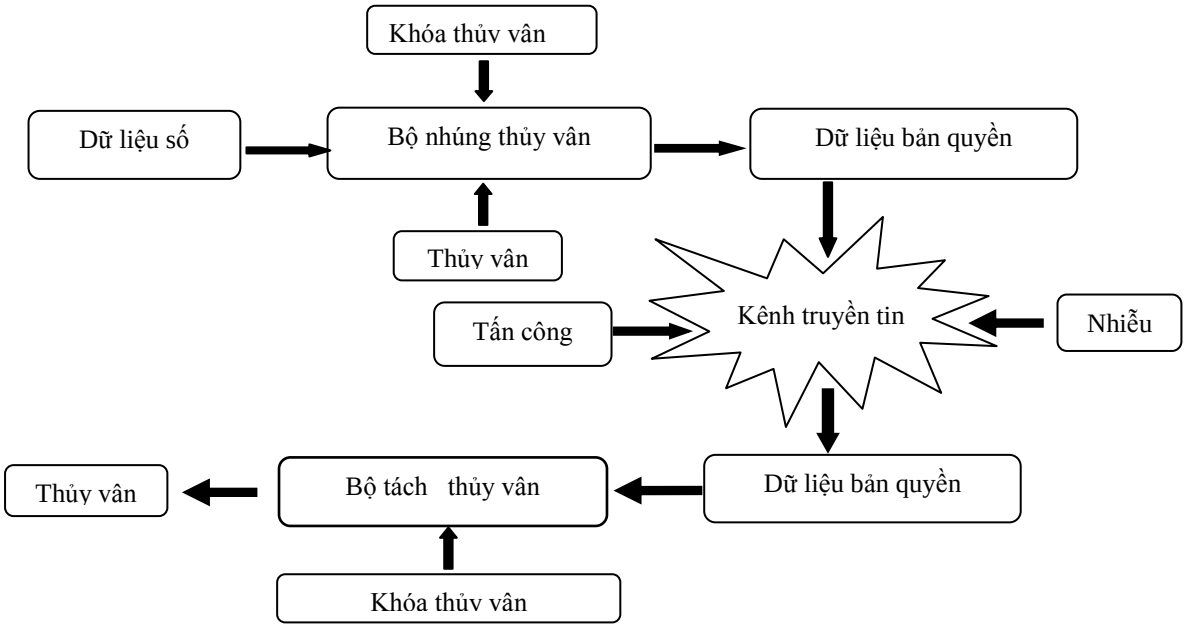
Dữ liệu chứa bao gồm các đối tượng như văn bản, audio, video, ảnh… dạng số, được dùng làm môi trường để giấu tin.

Bộ nhúng thủy vân là chương trình được cài đặt những thuật toán thủy vân và được thực hiện với một khóa bí mật.

Thủy vân sẽ được nhúng vào trong dữ liệu chứa nhờ một bộ nhúng thủy vân. Kết quả quá trình này là được dữ liệu chứa đã nhúng thủy vân gọi là dữ liệu có bản quyền và được phân phối trên các môi trường khác nhau. Trên đường phân phối có nhiễu và sự tấn công từ bên ngoài. Do đó yêu cầu các kỹ thuật thủy vân số phải bền vững với cả nhiễu và sự tấn công trên.

1.2.4.2. Quá trình trích thủy vân

Quá trình tách thủy vân được thực hiện thông qua một bộ tách thủy vân tương ứng với bộ nhúng thủy vân cùng với khóa của quá trình nhúng. Kết quả thu được là một thủy vân. Thủy vân thu được có thể giống với thủy vân ban đầu hoặc sai khác do nhiễu và sự tấn công trên đường truyền.



Hình 1.5: Sơ đồ hệ thống thủy vân số

1.2.5. Các hướng ứng dụng của thủy vân

Bảo vệ bản quyền tác giả CP (copyright protection).

Xác thực thông tin và phát hiện xuyên tạc thông tin (authentication and tamper detection)

Dấu vân tay hay dán nhãn (fingerprinting and labeling)

Điều khiển truy nhập (copy control)

1.3. Phân biệt giữa giấu tin và thủy vân

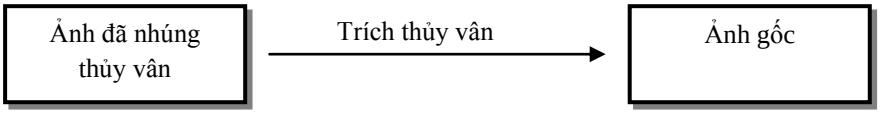
Xét về tính chất, thủy vân giống giấu tin ở chỗ cả hai hướng này đều tìm cách nhúng thông tin mật vào một môi trường. Nhưng về bản chất thì thủy vân và giấu tin có những nét khác ở một số điểm sau:

- Mục tiêu của thủy vân là nhúng thông tin không lớn, thường là biểu tượng, chữ ký hay các đánh dấu khác vào môi trường phủ nhằm phục vụ việc xác nhận bản quyền. Ngược lại, giấu tin mật yêu cầu lượng thông tin giấu là lớn.
- Thủy vân khác với giấu tin mật ở chỗ giấu tin sau đó cần tách lại tin còn thủy vân tìm cách biến tin giấu thành một thuộc tính của vật mang.
- Chỉ tiêu quan trọng nhất của một thủy vân là tính bền vững, của giấu tin là dung lượng.
- Thủy vân có thể vô hình hoặc hữu hình trên vật mang còn giấu tin chỉ được vô hình.

CHƯƠNG 2: THỦY VÂN THUẬN NGHỊCH TRÊN ẢNH SỐ

2.1. Thủy vân số thuận nghịch.

Thủy vân là phương pháp nhúng thông tin vào một tín hiệu kỹ thuật số như âm thanh, hình ảnh, video. Thủy vân thuận nghịch có thể khôi phục lại hình ảnh ban đầu mà không có bất kỳ sự biến dạng sau khi dữ liệu được trích ra từ sản phẩm đã nhúng thủy vân.



Hình 2.1: Mô hình thủy vân thuận nghịch

Trong thời gian gần đây giấu thuận nghịch được quan tâm một cách đặc biệt. Các nghiên cứu gần đây về thủy vân thuận nghịch: quá trình trích thủy vân để nhận thông tin nhúng có thể cần hoặc không cần đến ảnh gốc nhưng kết quả thu về ngoài thông tin nhúng ta có thêm một ảnh có các thuộc tính giống như ảnh gốc.

Một vài lĩnh vực như y học, quân đội hoặc nghiên cứu thực nghiệm vật lý phân tử hạt nhân,... nó đòi hỏi không những tách đúng thông điệp mà còn khôi phục xấp xỉ đúng ảnh gốc ban đầu. Vào năm 2001, phương pháp giấu thuật nghịch đầu tiên được đề xuất bởi Honsinger cùng các đồng nghiệp [6], từ đó đến nay nhiều kỹ thuật giấu thuật nghịch được công bố với hai hình thức giấu chính là trong miền dữ liệu và trong miền dữ liệu biến đổi.

2.2. Thủy vân số trên miền không gian

Các thuật toán thủy vân trong miền không gian tập trung vào việc thay đổi trực tiếp trong miền điểm ảnh. Thế mạnh của phương thức thủy vân trong miền điểm ảnh là đơn giản và có độ phức tạp tính toán thấp. Tuy nhiên, kỹ thuật này chỉ đảm bảo thuộc tính ẩn mà không có tính bền vững. Vì vậy, các thuật toán này được cài đặt cho ứng dụng xác thực thông tin của ảnh số.

2.2.1. Cơ sở lý thuyết

Ý tưởng cơ bản của thuật toán trong kỹ thuật này là chia một ảnh gốc thành các khối nhỏ, số lượng bit giấu trong mỗi khối tùy thuộc vào từng thuật toán. Thuật toán này dùng cho cả ảnh màu, ảnh đa mức xám và ảnh đen trắng nhưng để dễ trình bày thuật toán chúng ta sẽ sử dụng ảnh đen trắng.

Một số phép toán thường dùng khi thủy vân trên miền không gian:

Phép đảo bit: là một phép biến đổi trên các bit nhị phân. Đảo bit b được hiểu là phép biến đổi thay b bởi 1-b, tức là nếu ban đầu b nhận giá trị 0 thì sau khi đảo bit nó sẽ nhận giá trị 1 và ngược lại, nếu ban đầu b có giá trị là 1 thì sau khi đảo b mang giá trị 0.

Phép XOR (kí hiệu ⊕): là phép cộng loại trừ các phần tử tương ứng trên hai ma trân:

$$C = A \oplus B, \text{ với } \{\ C_{ij}=1 \text{ nếu } A_{ij} \neq B_{ij};\ C_{ij}=0 \text{ nếu } A_{ij} = B_{ij}\ \}$$

Phép Sum ma trận A (ký hiệu là Sum[A]) được định nghĩa là tổng tất cả các phần tử của ma trận A.

Phép nhân bit hai ma trận A, B (ký hiệu là A^ B) được định nghĩa:

C=A ^ B, với C_{ij} =1 nếu A_{ij} =B_{ij}=1,

C_{ij}=0 trong các trường hợp còn lại.

Phép nhân hai ma trận số nguyên A, B (ký hiệu A⊗B) được định nghĩa:

C= A ⊗ B, với C_{ij} = A_{ij} * B_{ij}

2.2.2. Một số thuật toán

2.2.2.1. Thuật toán 1(SW)

Đây là một thuật toán đơn giản (Simple Watermarking). Cho một file ảnh Bitmap đen trắng F, dữ liệu thủy vân d được biểu diễn dưới dạng nhị phân (dãy bit 0/1). Các bit 1 gọi là điểm đen, bit 0 gọi là điểm trắng.

Ý tưởng cơ bản của thuật toán này là chia một ảnh gốc thành các khối nhỏ, trong mỗi khối nhỏ sẽ giấu không quá một bit thông tin.

2.2.2.2. Thuật toán 2 (Wu-Lee)

Thuật toán này của 2 tác giả M.Y. Wu và J.H.Lee đưa ra cải tiến hơn thuật toán 1 bằng việc đưa thêm khóa K sử dụng trong quá trình nhúng và tách thủy vân đồng thời đưa thêm các điều kiện đảo bit trong mỗi khối. Với thuật toán này, có thể nhúng một bít vào mỗi khối bằng cách hiệu chỉnh nhiều nhất 1 bít của khối. Kỹ thuật này có khả năng làm tăng dữ liệu có thể nhúng.

Xét ảnh gốc F, khóa bí mật K và một số dữ liệu được nhúng vào F. Khóa bí mật K là một ma trận ảnh có kích thước mxn. Để đơn giản ta giả sử kích thước của ảnh gốc F là bội số của mxn. Quá trình nhúng thu được ảnh F có một số bit đã bị hiệu chỉnh. Thuật toán thực hiện như sau:

2.2.2.3. Thuật toán 3 (PCT)

Thuật toán này được đưa ra bởi 3 tác giả Hsiang-Kuang Pan, Yu-Yuan Chen, and Yu-Chee Tseng. Thuật toán cho phép nhúng nhiều bit vào 1 khối bằng cách có thể đảo 2 bit trong 1 khối. Trong thuật toán có sử dụng khóa K và ma trận trọng số W nhằm bảo đảm an toàn cho thủy vân được nhúng

- Khóa bí mật K: là một ma trận nhị phân có cùng kích thước mxn với kích thước của khối ảnh. Khóa được dùng một cách bí mật giữa người gửi và người nhận.

- Ma trận trọng số W cấp r: ma trận này có kích thước bằng kích thước của một khối ảnh (mxn) và thỏa mãn các điều kiện sau:

+ W là một ma trận số nguyên có các phần tử nằm trong khoảng (0..2^r-1) với r cho trước thỏa mãn điều kiện 2^r<(mxn)

+ Mỗi phần tử có giá trị từ (1..2^r-1) phải xuất hiện ít nhất 1 lần trong W.

Với mỗi n,m,r thỏa mãn 2^r-1<=mxn sẽ có:

$$C_{mn}^{2^r-1} \times (2^r-1)! \times (2^r-1)^{mn-(2^r-1)}$$

Khả năng chọn W là rất lớn. Ví dụ với m=n=4, r=2 → có 5.356.925.280 khả năng lựa chọn W. Con số này đủ lớn để làm giảm nguy cơ thủy vân bị phát hiện.

2.2.2.4. Thuật toán 4 (LSB)

Đây là thuật toán thủy vân dựa vào các bit ít quan trọng LSB. Các loại ảnh màu và đa mức xám có giá trị của mỗi điểm ảnh được biểu diễn bằng dãy nhiều bit. Trong dãy các bit này có một bit được gọi là bit it quan trọng nhất (LSB – Least Significant Bit). Bit ít quan trọng nhất là bit mà khi ta đảo giá trị của nó thì điểm màu bị thay đổi ít nhất.

2.3. Thủy vân số trên miền tần số

Các thuật toán này sử dụng phương pháp biến đổi cosine rời rạc DCT (Discrete Cosine Transform) để chuyển từng khối ảnh từ miền không gian ảnh sang miền tần số. Thủy vân sẽ được nhúng trong miền không gian tần số của ảnh theo kỹ thuật trải phổ trong truyền thông. Đây là kỹ thuật phổ biến nhất với nhiều thuật toán và là phương pháp có thể đảm bảo được tính mạnh mẽ và chính xác của thủy vân sau khi nhúng.

2.3.1. Biến đổi cosin rời rạc (DCT)

Biến đổi cosin rời rạc DCT (Discrete Cosine Transform) được đưa ra bởi Ahmed và các đồng nghiệp vào năm 1974. Từ đó cho đến nay, nó được sử dụng rất phổ biến trong nhiều kỹ thuật xử lý ảnh số nói riêng và xử lý tín hiệu số nói chung. Trong các kỹ thuật thủy vân ảnh dựa trên phép biến đổi dữ liệu ảnh sang miền tần số thì phép biến đổi DTC cũng được sử dụng rất nhiều. Nó được sử dụng trong chuẩn nén JPEG để mã hóa ảnh tĩnh và chuẩn MPEG để mã hóa ảnh động.

Công thức biến đổi DCT thuận từ $I(k,l) \rightarrow I(u,v)$

$$I(u,v) = \frac{C(u)C(v)}{4} \sum_{k=0}^{7} \sum_{l=0}^{7} I(k,l) \cos(\frac{(2k+1)u\pi}{16}) \cos(\frac{(2l+1)v\pi}{16})$$

I(u,v) được gọi là hệ số DCT và là số thực.

Công thức biến đổi ngược IDCT từ $I(u,v) \rightarrow I(k,l)$

$$I(k,l) = \sum_{k=0}^{7} \sum_{l=0}^{7} \frac{C(u)C(v)}{4} I(u,v) \cos(\frac{(2k+1)u\pi}{16}) \cos(\frac{(2l+1)v\pi}{16})$$

Ở đây 0<=k, l, u, v<=7.

Phép biến đổi DCT ảnh hai chiều thể hiện đặc tính nội dung về tần số của thông tin ảnh . Hầu hết các thuật toán, ảnh gốc được chia thành các khối ma trận ảnh 8x8. Áp dụng biến đổi DCT cho mỗi khối ta sẽ thu được khối 8x8 chứa các hệ số DCT. Gọi $C_b(j,k)$ là giá trị các hệ số trong đó: b là số thứ tự của khối, (j,k) là vị trí của hệ số. Hệ số đầu tiên $C_b(0,0)$ được gọi là DC và chứa thông tin độ sáng của khối đó. Các hệ số còn lại biểu diễn cho các thành phần tần số cao theo hướng ngang và theo hướng thẳng đứng gọi là hệ số AC.

	Low	Horizontal						High
Low	1	2	6	7	15	16	28	29
	3	5	8	14	17	27	30	43
	4	9	13	18	26	31	42	44

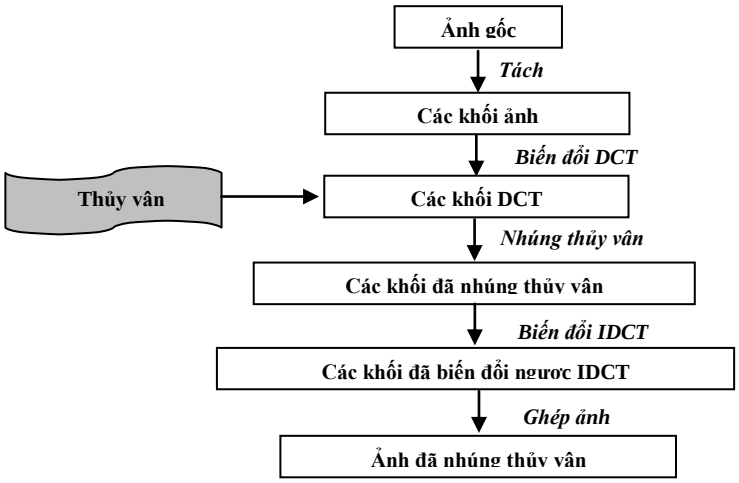
	10	12	19	25	32	41	45	54
	11	20	24	33	40	46	53	55
	21	23	34	39	47	52	56	61
	22	35	38	48	51	57	60	62
High	36	37	49	50	58	59	63	64

Bảng 1.1. Ví dụ bảng các hệ số DCT

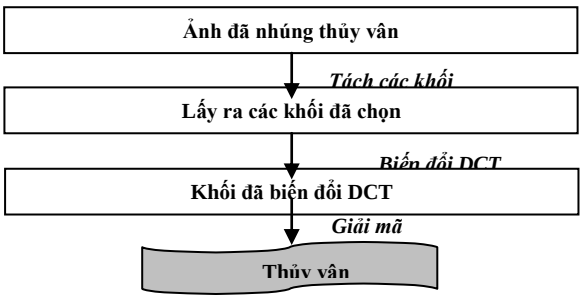
Theo nguyên lý chung, khi biến đổi chi tiết giữa các điểm ảnh càng lớn theo một hướng nào đó trong khối các điểm ảnh (hướng ngang, hướng thẳng đứng hay theo hướng đường chéo) thì các hệ số biến đổi DCT tương ứng cũng lớn.

2.3.2. Kỹ thuật thủy vân trên miền DCT

Thủy vân trên miền DCT là một kỹ thuật được sử dụng phổ biến với nhiều thuật toán. Nhìn chung, các thuật toán đều thực hiện các bước giống nhau trong quy trình nhúng và tách thủy vân như hình 2.6. Tuy nhiên, các thuật toán khác nhau thì khác nhau về cách lựa chọn vị trí nhúng thủy vân và phương thức nhúng.



(a)Quá trình nhúng thủy vân



(b)Quá trình tách thủy vân

Hình 2.8: Quy trình nhúng và tách thủy vân theo kỹ thủy vân trên miền DCT

2.3.2.1. Thuật toán DCT 1

Thuật toán được nhóm tác giả Nguyễn Xuân Huy và Trần Quốc Dũng đưa ra trên bài báo: "Một thuật toán thủy vân ảnh trên miền DCT – An Image Watermarking Algorithm Using DCT Domain". Nội dung bài viết này đề xuất một thuật toán nhúng thuỷ vân vào trong ảnh sao cho thoả mãn các tính chất và yêu cầu của một hệ thuỷ vân trên ảnh số. Thuật toán trong bài viết tập trung vào kỹ thuật chọn miền tần số để giấu tin nhằm nâng cao tính bền vững của thủy vân.

- ***Mô tả thuật toán***

− Input:

+ Một chuỗi các bít thể hiện bản quyền

+ Một ảnh

− Output:

+ Một ảnh sau khi thuỷ vân. Khoá để giải mã.

2.3.2.2. Thuật toán DCT 2

- ***Mô tả thuật toán***

Cùng ý tưởng nhúng tín hiệu thủy vân vào miền tần số giữa của khối biến đổi cosin rời rạc, tác giả Chris Shoemarker đã sử dụng phép biến đổi DCT để phân tích khối được chọn từ ảnh gốc thành các miền tần số, rồi chọn một cặp hệ số trong miền tần số giữa để thực hiện quá trình nhúng một bit thủy vân. Quá trình nhúng luôn đảm bảo sau khi nhúng bit thủy vân thì khoảng cách về giá trị giữa hai hệ số được chọn có giá trị lớn hơn hoặc bằng k cho trước.

2.3.2.3. Thuật toán DCT 3

- ***Mô tả thuật toán:***

Trong thuật toán DCT 3 này tác giả Benham lựa chọn vị trí nhúng tin có sự loại bỏ các khối không phù hợp. Các khối bị loại bỏ là các *khối nhẵn* hoặc *khối sắc* không cao. Các khối được chọn nhúng thủy vân là các *khối sắc* lớn.

Khối nhẵn: chúng ta có thể phát hiện ra các khối này bằng cách đếm số lượng hệ số cao tần có giá trị là "0". Nếu tất cả các hệ số này hay chỉ cần tồn tại ít nhất 1 hệ số ở nửa trên của đường zig-zắc bằng "0" thì khối đó được xem là khối nhẵn.

Khối sắc: được phát hiện bằng cách tìm giá trị tuyệt đối lớn nhất của hệ số AC tần số thấp. Ngưỡng được sử dụng là 100.

Thuật toán sử dụng 3 hệ số để nhúng 1 bit.

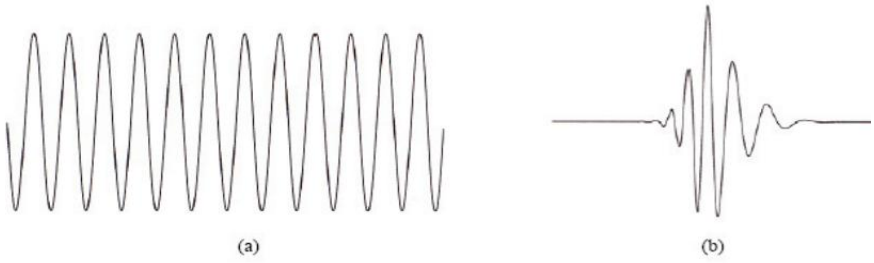
2.4. Thủy vân số trên miền Wavelet

Các thuật toán này sử dụng phương pháp biến đổi sóng nhỏ rời rạc DWT (Descrete Wavelet Transform) để chuyển miền không gian ảnh sang miền đa phân giải. Tính đa phân giải của sóng nhỏ hữu ích trong việc phân phối thông tin vào đối tượng bao phủ trong khi vẫn đảm bảo tính chất bền

vững và chất lượng hiển thị. DWT được dùng như một tiêu chuẩn trong nén JPEG200. Trong nhiều ứng dụng, DWT tỏ ra ưu thế hơn so với DCT do đặc tính đa phân giải.

2.4.1. Phép biến đổi DWT

Biến đổi tín hiệu chỉ là dạng biểu diễn khác của tín hiệu. Nó không làm thay đổi nội dung thông tin của tín hiệu. Thông thường, mỗi sóng là một hàm dao động của thời gian hoặc không gian và tuần hoàn. Sóng nhỏ là sóng cục bộ, chúng có năng lượng tập trung ở thời gian hoặc không gian và thích hợp để phân tích tín hiệu tạm thời. DWT sử dụng sóng nhỏ của năng lượng hữu hạn.



Hình 2.9: Sự khác nhau giữa sóng(a) và sóng nhỏ(b)

Trong phân tích sóng nhỏ, tín hiệu được phân tích nhiều lần với hàm sóng nhỏ và sự biến đổi được tính toán với mỗi phần được tạo ra. Với tần số cao, biến đổi sóng nhỏ cho độ phân dải rất tốt về thời gian và độ phân dải không tốt về tần số. Với tần số thấp, biến đổi sóng nhỏ cho độ phân dải tốt về tần số và độ phân dải không tốt về thời gian.

Ý tưởng của DWT cho tín hiệu một chiều: tín hiệu được chia thành 2 phần, phần tần số cao và phần số thấp. Thành phần tần số thấp lại được chia tiếp thành hai phần có tần số cao và thấp. Với các bài toán nén và thủy vân thường áp dụng không quá năm lần bước phân chia trên. Ngoài ra, từ các hệ số DWT, ta có thể tạo lại ảnh ban đầu bằng IDWT.

Có thể mô tả bằng toán học DWT và IDWT như sau:

Gọi $H(\omega)=\sum_k h_k e^{-jkw}$ là lọc thông thấp và $G(\omega)=\sum_k g_k e^{-jkw}$ là lọc thông cao thỏa mãn một vài điều kiện cho việc tái xây dựng ảnh ban đầu.

Một tín hiệu F(n) có thể được phân tích đệ quy như sau:

$$f_{j-1}^{low}(k)=\sum_n h_{n-2k} f_j(n) \text{ và } f_{j-1}^{high}(k)=\sum_n g_{n-2k} f_j(n)$$

Trong đó, j=J+1, J,…, J_0

J+1 là chỉ số mức phân giải cao còn J_0 là chỉ số mức phân giải thấp.

Các hệ số: $f_{J_0}^{low}(k), f_{J_0}^{high}(k), f_{J+1}^{low}(k),..., f_J^{high}(k)$ được gọi là các hệ số của tín hiệu F(n), với $f_{J_0}^{low}(k)$ là phần tử nhỏ nhất (xấp xỉ) của F(n) và $f_J^{high}(k)$ là phần chi tiết của F(n) tại

các dải tần khác nhau. Tín hiệu ban đầu F(n) được xây dựng lại từ các hệ số DWT bằng cách đệ quy như sau:

$$f_J^{low}(n) = \sum_k h_{n-2k} \cdot f_{j-1}^k + \sum_k g_{n-2k} \cdot f_{j-1}^{high}(k)$$

Để đảm bảo quan hệ giữa DWT và IDWT, $H(\omega)$ và $G(\omega)$ phải thỏa mãn điều kiện trực giao sau : $|H(\omega)|2 + |G(\omega)|2=1$.

Biến đổi DWT và IDWT cho mảng hai chiều mxn: định nghĩa tương tự bằng cách thực hiện các biến đổi một chiều DWT và IDWT cho mỗi chiều tương ứng.

2.4.2. Các thuật toán thủy vân trên miền DWT

2.4.2.1. Thuật toán DWT-1

- ***Ý tưởng***

Dựa trên kỹ thuật biến đổi sóng nhỏ hai tác giả Raval Mehul và Rege Priti đã đề xuất kỹ thuật thủy vân sử dụng phép biến đổi sóng nhỏ hai chiều để phân tích ảnh gốc thành bốn băng LL, HL, LH và HH rồi nhúng tín hiệu thủy vân thứ nhất vào băng LL, nhúng một thủy vân khác vào băng HH. Kết quả thử nghiệm cho thấy thủy vân bền vững trước một số phép xử lý ảnh thông thường.

2.4.2.2. Thuật toán DWT-2

Ở thuật toán DWT-1 sử dụng phép biến đổi sóng nhỏ hai chiều để phân tích ảnh gốc thành các băng tần khác nhau, rồi nhúng tín hiệu thủy vân vào một hoặc một số các băng tần. Theo cách đó, thủy vân có thể bền vững trước một số tấn công nhưng lại kém bền vững với một nhóm các tấn công khác. Khắc phục yếu điểm trên, trong thuật toán này, các tác giả Peining Tao và Ahmet M. Eskicioglu đã nhúng tín hiệu thủy vân vào cả bốn băng tần trong phép phân tích sóng nhỏ, mỗi băng tần có thể sử dụng các hệ số khác nhau.

CHƯƠNG 3: XÂY DỰNG CHƯƠNG TRÌNH THỬ NGHIỆM

3.1. Phát biểu bài toán

Nhằm mục đích xác thực chủ sở hữu các hình ảnh số trong hệ thống phân phối và sản xuất các học liệu điện tử, tránh các sao chép, sửa chữa không có sự đồng ý của người chủ sở hữu, tác giả luận văn xây dựng ứng dụng thủy vân số trên các file hình ảnh số với phương pháp thủy vân thuận nghịch. Bên cạnh yêu cầu chứng minh quyền sở hữu, ứng dụng cũng thực hiện cách trích rút thông tin thủy vân ra khỏi các file hình ảnh để thu được lại hình ảnh gốc ban đầu.

3.2. Phân tích và thiết kế hệ thống

Bài toán trên được giải quyết bằng phương pháp thủy vân số thuận nghịch. Quá trình bảo vệ bản quyền và tách thủy vân để thu được hình ảnh gốc ban đầu được tiến hành qua các bước:

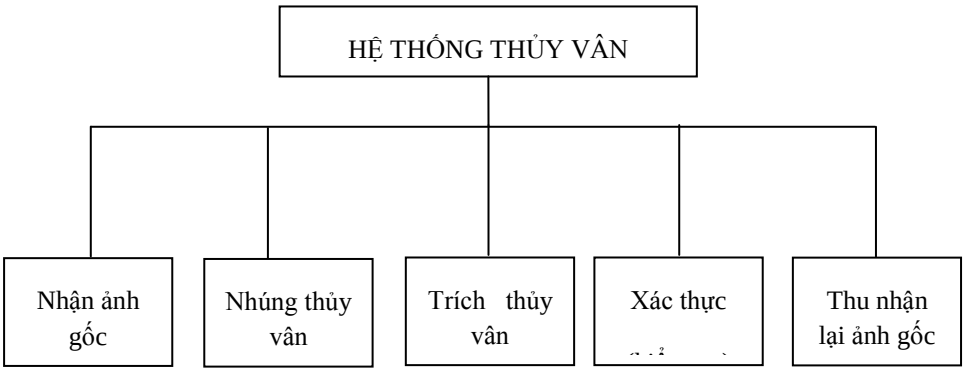
+ Chủ sở hữu bức ảnh số sẽ gửi các thông tin bí mật của mình cho hệ thống để tạo thủy vân và yêu cầu hệ thống nhúng thủy vân vào ảnh gốc của mình.

+ Hệ thống thủy vân tiến hành tạo và nhúng thủy vân cho chủ sở hữu. Sau khi nhúng thông tin, hệ thống sẽ gửi lại chủ sở hữu cả ảnh gốc và ảnh đã nhúng.

+ Khi có tranh chấp về vấn đề bản quyền, chủ sở hữu ảnh phải cung cấp ảnh có nghi ngờ sao chép trái phép cho hệ thống thủy vân để xác định nguồn gốc ảnh. Hệ thống thủy vân sẽ tách thủy vân từ ảnh nghi ngờ này. Sau đó, so sánh thủy vân nhận được với thủy vân mà chủ sở hữu cung cấp. Nếu hai thủy vân do hệ thống trích ra và thủy vân do chủ sở hữu cung cấp trùng nhau thì người này đúng là chủ sở hữu của tác phẩm, ngược lại chủ sở hữu trên là giả mạo và đã vi phạm sao chép sản phẩm không hợp pháp hoặc đã sửa đổi sản phẩm gốc thành tác phẩm của mình (nhái sản phẩm). Quá trình trích rút thông tin sẽ không yêu cầu chủ sở hữu cung cấp ảnh gốc, thêm nữa, sau khi trích thông tin ta thu được ảnh có các đặc tính giống như ảnh gốc.

3.2.1. Mô tả chức năng hệ thống

Hệ thống được xây dựng theo mô hình phân rã chức năng sau:



Hình 3.1: Sơ đồ chức năng hệ thống

◈ Chức năng nhận ảnh gốc

Các điểm ảnh trong tệp ảnh gốc gồm 3 thành phần màu: G,R và B. Ứng dụng duyệt qua toàn bộ ảnh gốc để nhận các điểm ảnh.

◈ Chức năng nhúng thủy vân

Thủy vân là chuỗi ký tự/file văn bản: chứa các thông tin bản quyền như tên tác giả, số chứng minh thư, mã số bản quyền… Khi nhúng, mỗi ký tự này sẽ chuyển thành mã ASCII tương ứng, sau đó đổi mã này thành chuỗi bít để đưa vào ảnh cần nhúng. Nếu so với ảnh đa cấp xám thì mỗi ký tự tương

đương với một điểm ảnh. Nếu so với ảnh 24 bít mầu thì mỗi điểm ảnh tương đương với 3 ký tự. Vì thế thủy vân là ký tự thì lượng thông tin nhúng được sẽ rất nhiều.

Thủy vân là một ảnh: ảnh này có thể là một logo đặc trưng cho công ty hoặc là dấu vân tay đặc trưng cho một cá nhân. Ảnh thủy vân phải có kích thước nhỏ hơn nhiều so với ảnh gốc.

Nếu ảnh thủy vân là ảnh đen trắng thì việc tạo thủy vân chỉ đơn thuần nhặt ra từng điểm ảnh để nhúng vào các khối ảnh. Nếu ảnh có kích thước MxN thì chuỗi nhị phân biểu diễn cho ảnh nhị phân cần nhúng có độ dài là MxN bít.

Nếu ảnh thủy vân là ảnh đa cấp xám: lấy giá trị của từng điểm ảnh theo cách duyệt ảnh từ trên xuống dưới, từ trái qua phải xếp thành chuỗi số biểu diễn cho thủy vân cần nhúng. Mỗi số trong dãy số trên lại được chuyển thành một dãy 8 bit nhị phân. Vậy nếu ảnh có kích thước MxN thì dãy thủy vân biểu diễn cho ảnh thủy vân có kích thước là MxNx8 bits.

Nếu ảnh thủy vân là ảnh 24 bits màu, mỗi thành phần mầu R, G, B chiếm 1 byte nhớ. Khi đó có 2 cách tạo thủy vân:

◈ Chức năng trích thủy vân

Từ các khối ảnh (16x16) có thủy vân ta sẽ lấy ra được một số bit thủy vân. Ghép các bít này lại với nhau để được dãy bit. Thực hiện cắt từng đoạn 8 bit một của dãy bit này để thu được mã ASCII của ký tự hoặc giá trị mức xám của một điểm ảnh.

Với mục đích xác thực thông tin thì yêu cầu của hệ thống phải là thủy vân dễ vỡ. Khi đó chỉ việc so sánh thủy vân tách được từ ảnh nghi ngờ với thủy vấn gốc mà chủ sở hữu đang có, nếu không giống nhau thì có nghĩa là tác phẩm đã bị sửa chữa thông tin trái phép, không phải là sản phẩm nguyên bản của tác giả.

Với mục đích bảo vệ bản quyền thì yêu cầu của hệ thống phải là thủy vân bền vững. Nghĩa là dù sản phẩm có bị sửa chữa theo một hình thức nào đó (do các đối tượng nhái lại từ bản gốc) thì thủy vân vẫn được bảo vệ. Do đó, tác giả có thể trích thủy vân từ bản nghi ngờ ăn cắp bản quyền để chứng minh rằng đây là tác phẩm của mình đã bị chỉnh sửa (bản nhái lại).

◈ Chức năng kiểm tra

Kiểm tra tính bền vững của các thuật toán thủy vân. Với chức năng này người sử dụng có thể kiểm tra xem thuật toán mình chọn có thể chống lại những biến đổi tấn công như: nén, nhiễu, tăng giảm độ sáng...từ đó có thể lựa chọn giải pháp hợp lý cho thuật toán nhúng thủy vân. Kẻ vi phạm bản quyền có thể dùng các tấn công trái phép để làm biến đổi dấu thủy vân. Nếu sau khi tấn công chất lượng ảnh thấp, không còn giá trị thương mại thì thuật toán thành công về khía cạnh bền vững.

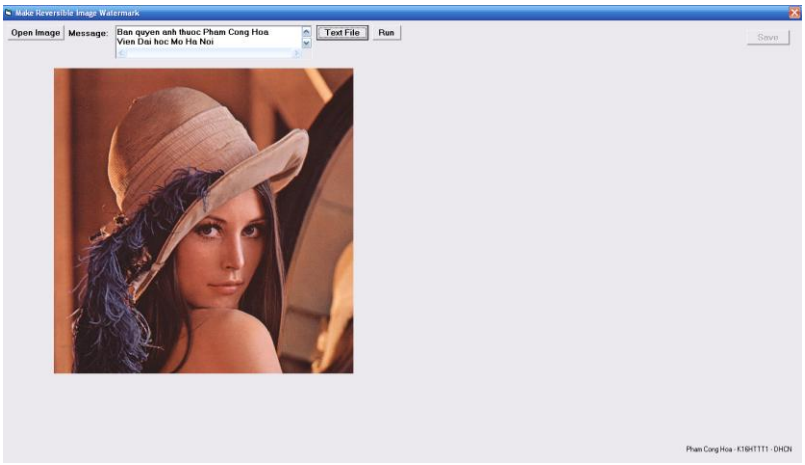
◈ Chức năng thu nhận lại ảnh gốc

Trong quá trình trích rút thông tin thủy vân, hệ thống cũng sửa đổi lại các giá trị hệ số DCT tương ứng, tiếp đến sử dụng DCT nghịch để chuyển đổi hệ số từ ảnh thủy vân về ảnh gốc. Do phép biến đổi DCT là thuận nghịch nên ta sẽ thu được giá trị thật của ảnh gốc.

3.2.2. Ứng dụng thử nghiệm

a). Chức năng nhúng thủy vân

◈ **Mở ảnh gốc**: chọn mục "Open File" để mở một file ảnh cần nhúng thủy vân. Ảnh gốc sẽ hiện ra bên trái của cửa sổ ứng dụng.



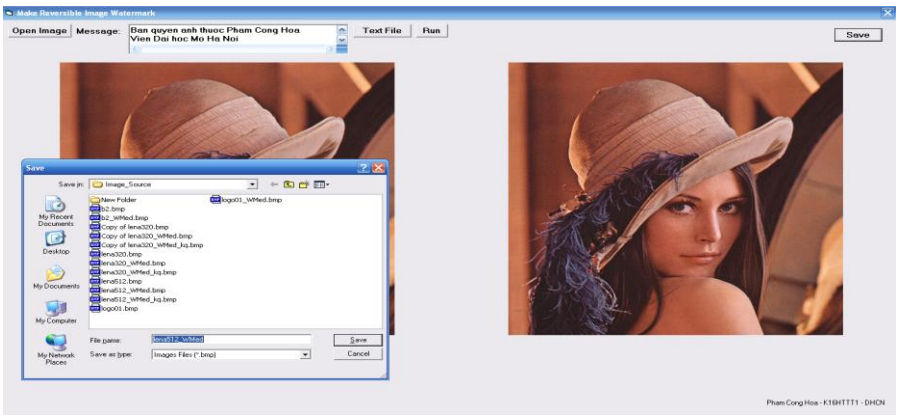
Hình 3.2: Giao diện khi mở ảnh gốc

◈ **Tạo thủy vân**: thủy vân là một chuỗi kỹ tự nhập trực tiếp vào một textbox (ô nhập liệu trên cửa sổ ứng dụng.

◈ **Nhúng thủy vân**: Chọn mục "Run" để bắt đầu quá trình nhúng. Các bước chính trong quá trình này bao gồm:

- Nhận các khối (16x16) điểm ảnh của ảnh gốc và lưu vào 3 mảng tương ứng với 3 thành phần màu G,R và B.
- Chuyển đổi các giá trị của 3 mảng trên sang giá trị đa cấp xám ứng với 3 mảng Y,U và V.
- Biến đổi DCT thuận trên các giá trị đa cấp xám của mảng Y. Sau bước này ta có mảng các hệ số DCT (mảng Coefficients)
- Chuyển đổi chuỗi ký tự thủy vân sang chuỗi các bit, mỗi ký tự tương ứng 8 bit.
- Sử dụng thuật toán DCT-1 (chương 2) để thay đổi giá trị các phần tử của mảng Coefficients.
- Chuyển đổi DCT nghịch để nhận các giá trị đa cấp xám Y mới.

◈ **Tạo ảnh thủy vân**: Sau khi biến đổi DCT nghịch ở bước trên, ta đưa từng điểm ảnh (đã biến đổi) vào một khung ảnh (pictureBox) trên cửa sổ ứng dụng. Ảnh đã nhúng thủy vân nằm bên phải cửa sổ ứng dụng. Bấm "Save" để lưu lại ảnh đã nhúng thủy vân.



Hình 3.3: Giao diện ghi lại ảnh đã thủy vân

◈ **Đánh giá:**

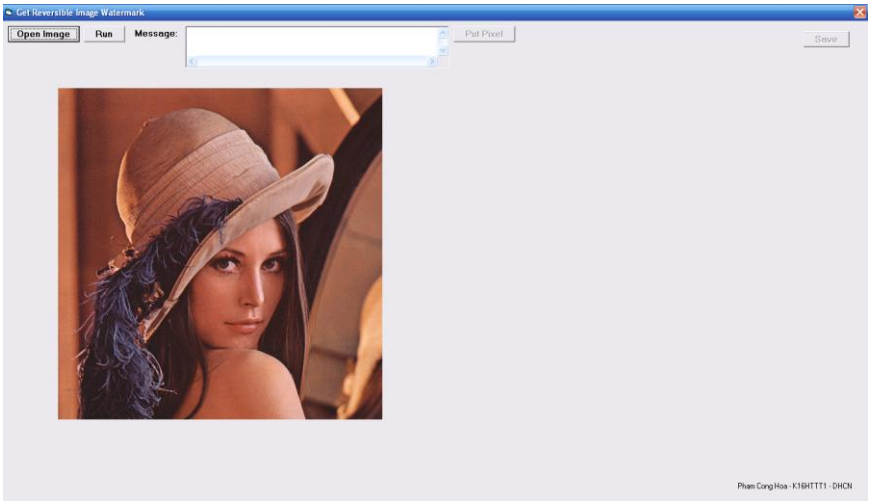
Vì chương trình xử lý với ảnh 24 bit mầu, nghĩa là chỉ có một thành phần R hoặc G hoặc B tham gia nhúng thủy vân nên khả năng "lộ" của ảnh rất thấp. Nếu quan sát kỹ trên các kết quả tại các vùng sáng của ảnh thì có khả năng nhận ra được một vài điểm ảnh có sự thay đổi trên gam màu tương ứng.

Để giảm khả năng bị lộ, có thể tăng kích thước khối.

Ở đây dung lượng thủy vân cần nhúng và khả năng bị lộ của ảnh là tỷ lệ nghịch. Khi dung lượng thông tin nhúng lớn thì khả năng bị lộ càng tăng và ngược lại.

b). Chức năng tách thủy vân

◈ **Mở ảnh đã thủy vân**: chọn mục "Open File" để mở một file đã nhúng thủy vân. Ảnh đã thủy vân sẽ hiện ra bên trái của cửa sổ ứng dụng.

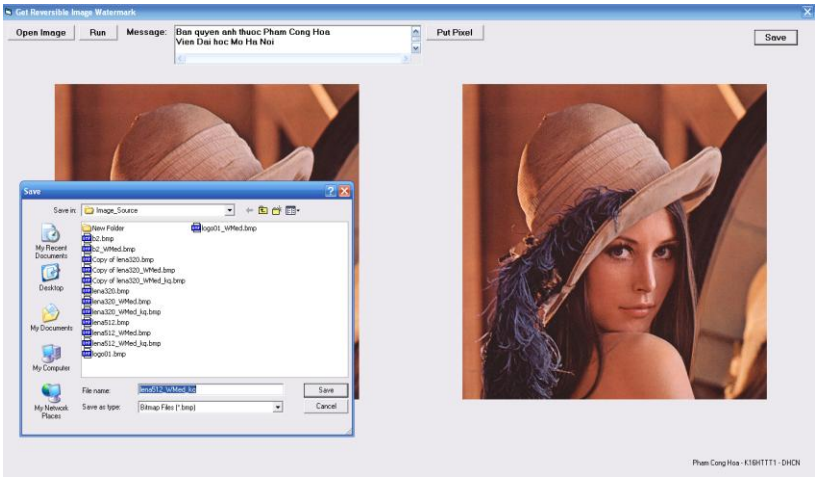


Hình 3.4: Giao diện mở ảnh đã thủy vân

Chọn “Run” để bắt đầu quá trình trích thủy vân. Các bước trong quá trình này bao gồm:

- Nhận các khối (16x16) điểm ảnh của ảnh đã thủy vân và lưu vào 3 mảng tương ứng với 3 thành phần màu G,R và B.
- Chuyển đổi các giá trị của 3 mảng trên sang giá trị đa cấp xám ứng với 3 mảng Y,U và V.
- Biến đổi DCT thuận trên các giá trị đa cấp xám của mảng Y. Sau bước này ta có mảng các hệ số DCT (mảng Coefficients)
- Sử dụng thuật toán DCT-1 (chương 2) để xác định giá trị các bit đã nhúng tương ứng trong mỗi khối. Nối các bit này thành các chuỗi 8 bit và chuyển đổi sang mã ASCII để nhận được chuỗi ký tự thủy vân. Trong bước này ta cũng thay đổi giá trị các phần tử của mảng Coefficients từ file văn bản đã lưu ở quá trình nhúng thủy vân. Mục đích của thao tác này là để phục vụ việc chuyển đổi lại ảnh đã thủy vân trở lại ảnh có tính chất nguyên dạng ảnh gốc.
- Chuyển đổi DCT nghịch để nhận các giá trị đa cấp xám Y mới.

◈ **Biến đổi ảnh đã thủy vân về nguyên dạng ảnh gốc**: Sau khi biến đổi DCT nghịch ở bước trên, chọn “Put Pixel” để đưa từng điểm ảnh (đã biến đổi) vào một khung ảnh (pictureBox) trên cửa sổ ứng dụng. Ảnh nguyên dạng ảnh gốc nằm bên phải cửa sổ ứng dụng. Bấm “Save” để lưu lại ảnh đã trích thủy vân.



Hình 3.5: Giao diện ghi ảnh đã trích thủy vân

◈ **Đánh giá:**

Vì các phép biến đổi DCT là thuận nghịch nên ta sẽ thu được ảnh cuối có nguyên dạng (thuộc tính từng điểm ảnh) như ảnh gốc.

Thời gian để trích thủy vân phụ thuộc vào thời gian nhận từng điểm ảnh và thời gian biến đổi DCT thuận. Như vậy ảnh thủy vân càng lớn (về dung lượng) thì thời gian trích thủy vân càng cao.

◈ **Một số hạn chế cần giải quyết:**

- Các tấn công khác như: xoay ảnh, tạo nhiễu,… chương trình chưa đảm bảo được tính bền vững.

- Tốc độ thực thi còn chậm.

- Chưa ứng dụng với các định dạng ảnh khác như JPG, GIF, PNG, …

KẾT LUẬN

Vấn đề bảo vệ trí tuệ và quyền tác giả xu hướng ngày càng được quan tâm trên thế giới. Kỹ thuật thủy vân (watermarking) được nghiên cứu và được ứng dụng trong nhiều các lĩnh vực khác nhau. Để watermarking đạt được tính mạnh hơn nữa thì chúng nên được công bố và thảo luận rộng rãi.

Thủy vân số là một công nghệ mới rất phức tạp, để thực sự có những ứng dụng trong thực tế phải cần có nhiều thời gian nghiên cứu và thẩm định. Tuy nhiên, đây cũng là một công nghệ được các nhà khoa học khẳng định là đầy hứa hẹn cho vấn đề bảo mật và an toàn thông tin. Thủy vân số có thể thực hiện ở nhiều môi trường khác nhau. Có nhiều thuật toán thủy vân, tùy từng mục đích cụ thể mà ta chọn thuật toán thủy vân phù hợp. Mỗi thuật toán khác nhau đều có những ưu điểm và nhược điểm riêng và thông thường chỉ chịu được một số tấn công, không có thuật toán nào có thể bền vững với tất cả các tấn công. Tính bền vững của thủy vân tỷ lệ nghịch với chất lượng ảnh sau khi nhúng.

◈ Kết quả đạt được:

Luận văn đã trình bày một cách có hệ thống các kiến thức liên quan đến giấu tin và thủy vân số, tập trung nghiên cứu các thuật toán thủy vân trên các miền khác nhau của ảnh số : miền không gian, miền tần số dựa vào phép biến đổi Cosine rời rạc DCT và miền tần số dựa vào phép biến đổi sóng nhỏ DWT.

Xây dựng chương trình thử nghiệm có cài đặt các thuật toán trên miền không gian và trên miền tần số dựa vào phép biến đổi Cosine rời rạc DCT , chương trình có khả năng thu nhận lại ảnh có tính chất tương đương ảnh gốc sau khi trích thủy vân, đánh giá tính bền vững của thủy vân qua một số phép tấn công đơn giản.

◈ Hướng phát triển của luận văn:

Với thủy vân bền vững, chương trình mới chỉ cài đặt thử nghiệm được một thuật toán trên miền DCT. Do đó cần tiếp tục nghiên cứu cài đặt thêm nhiều thuật toán khác.

Luận văn mới chỉ thực hiện nhúng thủy vân ẩn trên dữ liệu ảnh số, kỹ thuật nhúng, trích còn nhiều hạn chế về độ phức tạp. Trong tương lai cần tiếp tục xây dựng chương trình có thể nhúng thủy vân trên nhiều phương tiện khác nhau như audio, video.

References.

Tiếng Việt

[1]. Lương Mạnh Bá - TS. Nguyễn Thanh Thuỷ - "Nhập môn xử lý ảnh số"(1999).

[2]. Nguyễn Xuân Huy, Trần Quốc Dũng (2002), "Một thuật toán thủy vân ảnh trên miền DCT - An Image Watermarking Algorithm Using DCT Domain", *Kỷ yếu Hội thảo Quốc gia: Một số vấn đề chọn lọc của Công nghệ Thông tin, Thái Nguyên*, 29-31/08/2003, NXB Khoa học Kỹ thuật, Hà Nội, 2005, tr. 146-151.

[3]. Lê Tiến Thường, Nguyễn Thanh Tuấn (2004), "Giải pháp hiệu quả dùng kỹ thuật watermarking cho ứng dụng bảo vệ bản quyền ảnh số", *Tạp chí khoa học ĐH Bách Khoa TPHCM*, tr. 5-8.

[4]. Nguyễn Xuân Huy, Bùi Thị Thúy Hằng (2001), "Một số cải tiến của kỹ thuật giấu dữ liệu trong ảnh", *Kỷ yếu Hội nghị kỷ niệm 25 năm thành lập Viện Công nghệ thông tin ,* Hà Nội 24-25/12/2001, tr. 553 – 559.

[5]. Nguyễn Văn Tảo, Bùi Thế Hồng (2007), "Về một lược đồ thủy vân dựa trên phép biến đổi sóng nhỏ rời rạc và các ma trận số giả ngẫu nhiên", *Tạp chí Khoa học và Công nghệ, tập 45, số 3 năm 2007*, tr. 27-34.

Tiếng Anh

[6] C. W. Honsinger, P. Jones, M. Rabbani, J. C. Stoffel, "Lossless recover y of an original image containing embedded data", US Patent application, Docket no: 77102/E-D, 2001.

[7] Zhicheng Ni, Yun-Qing Shi, Nirwan Ansari, and Wei Su, "Reversible Data Hiding", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No.3 (2006) 354.

[8] J.H. Hwang, J. W. Kim, J. U. Choi, "A Reversible Watermarking Based on Histogram Shifting", IWDW 2006, LNCS 4283 (2006) pp. 384-361.

[9]. I.J. COX, J. KILIAN, T. LEIGHTON, AND T. SHAMOON, *A secure, robust watermark for multimedia,* in Proc First Int. Workshop on Information Hiding, R. Anderson, ed., no. 1174 in Lecture Notes in Computer Science, pp. 185–206, May/June 1996.

[10]. Frank Hartung, Martin Kutter, "Multimedia Watermarking Techniques", Proceedings of The IEEE, Vol. 87, No. 7, pp. 1085 – 1103, July 1999.

[11]. Hsiang-Kuang Pan, Yu-Yuan Chen, and Yu-Chee Tseng, " A Secure Data Hiding Scheme for Two-Color Images", *IEEE Symposium on Computers and Communications*, pp.750–755, July 2000.

[12]. I. J. Cox, F. T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia”, in *Proceedings of the IEEE ICIP '97*, vol. 6, pp.1673-1687, Santa Barbara, California, USA, 1997.

[13]. Martin Vetterli and Jelena Kovacevic (1995), *Wavelets and Subband Coding*, Prentice Hall.

[14]. M. Wu, E. Tang, and B. Liu, “Data hiding in digital binary image,” in Proc. Of IEEE Int. Conf. on Multimedia and Expo, New York City, pp. 393-396, July 31 to August 2