**TRƯỜNG ĐAI HỌC TÀI NGUYÊN VÀ MÔI TRƯỜNG TP.HCM**

**KHOA: HỆ THỐNG THÔNG TIN VÀ VIỄN THÁM**



**BÁO CÁO ĐỒ ÁN MÔN HỌC**

**AN TOÀN BẢO MẬT MÁY TÍNH & HỆ THỐNG**

**ĐỀ TÀI: NGHIÊN CỨU KỸ THUẬT THỦY VÂN SỐ VÀ XÂY DỰNG ỨNG DỤNG BẢO VỆ BẢN QUYỀN ẢNH SỐ**

Giảng viênhướng dẫn: **ThS. Phạm Trọng Huynh**

Sinh viên thực hiện:  **Văn Trung Nghĩa - 0850080085**

**Nguyễn Phi Trường – 0850080110**

**Nguyễn Minh Quang - 0850080099**

Lớp: **08\_ĐH\_CNPM**

Khóa: **08**

**TP. Hồ Chí Minh, tháng 3 năm 2023**

**TRƯỜNG ĐAI HỌC TÀI NGUYÊN VÀ MÔI TRƯỜNG TP.HCM**

**KHOA: HỆ THỐNG THÔNG TIN VÀ VIỄN THÁM**



**BÁO CÁO ĐỒ ÁN MÔN HỌC**

**AN TOÀN BẢO MẬT MÁY TÍNH & HỆ THỐNG**

**ĐỀ TÀI: NGHIÊN CỨU KỸ THUẬT THỦY VÂN SỐ VÀ XÂY DỰNG ỨNG DỤNG BẢO VỆ BẢN QUYỀN ẢNH SỐ**

Giảng viênhướng dẫn: **ThS. Phạm Trọng Huynh**

Sinh viên thực hiện:  **Văn Trung Nghĩa - 0850080085**

**Nguyễn Phi Trường – 0850080110**

**Nguyễn Minh Quang - 0850080099**

Lớp: **08\_ĐH\_CNPM**

Khóa: **08**

**TP. Hồ Chí Minh, tháng 3 năm 2023**

**MỞ ĐẦU**

Với sự phát triển vũ bão của Internet, các phương tiện kĩ thuật số như phương tiện lưu trữ, phương tiện truyền thông, kỉ nguyên thông tin số đã được hình thành. Hầu hết thông tin ngày nay được lưu trữ dưới dạng số hoá trước khi trở thành một vật phẩm thông thường như sách báo, tạp chí. Việc trao đổi, phân bố, sao chép và xử lý các sản phẩm số này ngày càng nhanh chóng, đơn giản năm ngoài tầm kiểm soát của các tổ chức. Và do đó, tình trạng xâm phạm bản quyền sao chép lậu các sản phẩm kĩ thuật số đã xảy ra ở nhiều nơi. Tháng 3/1998 ở Hồng Kông, chính quyền đã tịch thu một số lượng đĩa CD âm nhạc lậu trị giá 85 triệu đôla, và vào 6/1998, ở Đức cũng đã tịch thu các sản phẩm phần mềm và kĩ thuật số lậu trị giá 1,9 triệu đô la.

Trong năm này, quốc hội Mỹ thông qua dự luật về “Hành động bảo vệ bản quyền kỷ nguyên kĩ thuật số” (Digital Millenium Copyright Act) để bảo vệ quyền sở hữu trí tuệ với các sản phẩm kĩ thuật số. Tất cả các việc đó nói lên sự cần thiết phải có một cách để bảo vệ các em phẩm kĩ thuật số khỏi các thao tác sao chép, xử lý trái phép và ngăn chặn các hoạt động này xảy ra. Một phương pháp chống sao chép trái phép và bảo vệ bản quyền cho các sản phẩm số – thuỷ vân số (digital watermarking) - đã ra đời trong hoàn cảnh này.

Thuỷ vân số là một phương pháp dùng để bảo vệ các sản phẩm số. Nó có thể được dùng trong các lĩnh vực bảo vệ bản quyền, chống sao chép, phân biệt già mạo, ...Các sản phẩm số này có thể là văn bản audio, video, phần mềm, ảnh...

Thuỷ vân số là một phương pháp mới dựa trên lý thuyết tổng hợp của nhiều lĩnh vực khác nhau như mật mã học, lý thuyết thông tin, lý thuyết truyền thông và xử lý tín hiệu số, xử lý ảnh. Mục đích của phương pháp này là dấu thêm một lượng thông tin có ích vào sản phẩm số và lượng thông tin này được gọi là thuỷ vẫn.

Chính vì tính hữu ích trong ứng dụng thực tiễn của thủy vân số nên em quyết định lựa chọn đề tài tiểu luận là: “Nghiên cứu kỹ thuật thủy vân số trong việc bảo vệ bản quyền ảnh số”.

**NHẬN XÉT**

**(Của giảng viên)**

………….., ngày….tháng….năm……

NGƯỜI NHẬN XÉT

*(ký tên)*

**MỤC LỤC**

[CHƯƠNG 1. TỔNG QUAN VỀ THỦY VÂN SỐ 1](#_Toc131330060)

[1.1 Định nghĩa vấn đề 1](#_Toc131330061)

[1.1.1 Khái niệm thủy vân số 1](#_Toc131330062)

[1.1.2 Phân loại thủy vân 2](#_Toc131330063)

[1.1.3 Phân loại thủy vân theo miền nhúng 2](#_Toc131330064)

[1.1.4 Phân loại theo đối tượng được nhúng thủy vân 2](#_Toc131330065)

[1.1.5 Phân loại thủy vân theo cảm nhận của con người 2](#_Toc131330066)

[1.2 Các hướng ứng dụng của thủy vân 4](#_Toc131330067)

[1.2.1 Bảo vệ bản quyền ảnh số 4](#_Toc131330068)

[1.2.2 Đặc tính của thủy vân 5](#_Toc131330069)

[1.3 Yêu cầu đối với phương pháp thủy vân 8](#_Toc131330070)

[1.4 Khả năng tấn công trên hệ thống thủy vân số 9](#_Toc131330071)

[1.5 Phạm vi đồ án 9](#_Toc131330072)

[1.6 Mục tiêu, sự cần thiết của đề tài, tại sao lại chọn đề tài này ? 10](#_Toc131330073)

[CHƯƠNG 2. KỸ THUẬT THỦY VÂN SỐ 12](#_Toc131330074)

[2.1 Giới thiệu về kỹ thuật giấu thông tin 12](#_Toc131330075)

[2.2 Phân loại về kỹ thuật giấu thông tin 12](#_Toc131330076)

[2.2.1 Giấu thông tin bí mật 12](#_Toc131330077)

[2.2.2 Giấu thông tin thủy vân 13](#_Toc131330078)

[2.3 Tìm hiểu phương pháp và mô hình thủy vân số 14](#_Toc131330079)

[2.3.1 Phương pháp thủy vân số 14](#_Toc131330080)

[2.3.2 Mô hình thủy vân số 14](#_Toc131330081)

[2.4 Tìm hiểu về các thuật toán thủy vân theo miền không gian ảnh (SW; WU- LEE; LBS; PCT) 17](#_Toc131330082)

[2.4.1 Thuật toán SW 17](#_Toc131330083)

[2.4.2 Thuật toán WU-LEE 19](#_Toc131330084)

[2.4.3 Thuật toán LBS 22](#_Toc131330085)

[2.4.4 Thuật toán PCT 25](#_Toc131330086)

[2.5 Tìm hiểu về các thuật toán thủy vân theo miền tần số (DCT; DWT) 29](#_Toc131330087)

[2.5.1 Hướng tiếp cận theo miền tần số 29](#_Toc131330088)

[2.5.2 Miền tần số DCT 30](#_Toc131330089)

[2.5.3 Miền tần số DWT 32](#_Toc131330090)

[CHƯƠNG 3. CÀI ĐẶT THỬ NGHIỆM 35](#_Toc131330091)

[3.1 Phát biểu bài toán 35](#_Toc131330092)

[3.2 Ứng dụng chương trình 35](#_Toc131330093)

[3.3 Hướng dẫn sử dụng 35](#_Toc131330094)

[CHƯƠNG 4. KẾT LUẬN 38](#_Toc131330095)

[TÀI LIỆU THAM KHẢO 39](#_Toc131330096)

**DANH MỤC HÌNH VẼ**

[Hình 1. 1 Phân loại thủy vân số 2](#_Toc135808172)

[Hình 2. 1 Quy trình nhúng thủy vân 15](#_Toc135808179)

[Hình 2. 2 Quy trình trích xuất và tìm kiếm thủy vân 16](#_Toc135808180)

[Hình 2. 3 Minh họa thuật toán SW: nhúng bit 1 vào khối ảnh B. 18](#_Toc135808181)

[Hình 2. 4 Minh họa chọn điểm ảnh giấu tin vào những khối ảnh màu 19](#_Toc135808182)

[Hình 2. 5 Minh họa thuật toán WU\_LEE nhúng đoạn bit 01 20](#_Toc135808183)

[Hình 2. 6 Ví dụ bảng các hệ số DCT 31](#_Toc135808184)

[Hình 2. 7 Phân chia 3 miền tần số thấp giữa, cao của phép biến đổi DCT. 31](#_Toc135808185)

[Hình 2. 8 Biến đổi Wavelet và cấu trúc dải thông 33](#_Toc135808186)

1. TỔNG QUAN VỀ THỦY VÂN SỐ

Định nghĩa vấn đề

* + 1. Khái niệm thủy vân số

[9] Kỹ thuật thủy vân trên giấy xuất hiện trong các tác phẩm nghệ thuật làm giấy thủ công cách đây khoảng 700 năm. Loại giấy có thủy vân cổ nhất được tìm thấy vào những năm 1929 và nguyên bản của nó bắt nguồn từ thị trấn Fabriano ở Ý đã đóng góp một vai trò rất lớn đối với sự tiến hóa của công nghiệp sản xuất giấy. Vào thời điểm này, kỹ thuật thủy vân được xem là phương pháp hữu hiệu để xác định nguồn gốc sản phẩm, giúp người dùng lựa chọn đúng hãng sản xuất giấy mà mình muốn mua.

Thuật ngữ watermark bắt nguồn từ một loại mực vô hình được viết trên giấy và chỉ hiển thị khi nhúng giấy đó vào nước. Thuật ngữ Thủy vân số được cộng đồng thế giới chấp nhận rộng rãi vào đầu thập niên 1990. Khoảng năm 1995, sự quan tâm đến thủy vân số bắt đầu phát triển nhanh

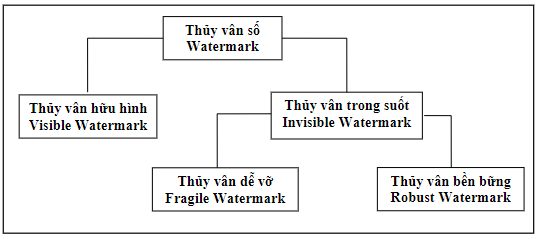
Thủy vân số là quá trình sử dụng các thông tin (ảnh, chuỗi bít, chuỗi số) nhúng một cách tinh vi vào dữ liệu số (ảnh số, audio, video hay text) nhằm xác định thông tin bản quyền của tác phẩm đó. Mục đích của thủy vân số là bảo vệ bản quyền cho phương tiện dữ liệu số mang thông tin thủy vân.

Thao tác đưa thủy vân vào một môi trường số được gọi là thủy vân số. Thủy vân số được xem như là một hình thức ẩn giấu tin. Theo sơ đồ phân loại kỹ thuật giấu tin của A.P. Pentitcolas 1999 theo hai hướng nghiên cứu chính là giấu tin mật và thủy vân số. Có thể xem watermarking là thao tác mà nhúng tin mà trong đó người dùng đầu cuối không cần quan tâm tới thông tin được giấu bên trong đối tượng chứa tin.

Như vậy, Thủy vân số là quá trình nhúng những dữ liệu vào một đối tượng đa phương tiện theo một phương pháp nào đó, để sau đó có thể phát 5 hiện hoặc trích xuất thủy vân cho mục đích xác thực nguồn gốc của sản phẩm. Thủy vân là một phần đặc trưng của thông tin nhúng vào dữ liệu cần bảo vệ. Một yêu cầu quan trọng đối với thủy vân là rất khó để trích xuất hoặc gỡ bỏ được nó từ đối tượng được nhúng thủy vân mà không biết được chìa khóa bí mật. [9]

* + 1. Phân loại thủy vân

[9] Thủy vân và kỹ thuật thủy vân tùy theo từng tiêu chí phân loại mà có thể được chia thành nhiều loại khác nhau [9] :



Hình 1. 1 Phân loại thủy vân số

* + 1. Phân loại thủy vân theo miền nhúng

[9] Một trong những tiêu chí để phân loại là “miền nhúng ” là nơi chứa thủy vân. Ví dụ, thủy vân có thể được thực hiện trong “miền không gian”. Một khả năng khác là thủy vân trong miền tần số [9]

* + 1. Phân loại theo đối tượng được nhúng thủy vân

[9] Kỹ thuật thủy vân có thể được phân loại theo đối tượng đa phương tiện cần nhúng thủy vân như sau:

+ Thủy vân trên ảnh

+ Thủy vân trên video

+ Thủy vân trên âm thanh

+ Thủy vân trên văn bản [9]

* + 1. Phân loại thủy vân theo cảm nhận của con người

[9] Tùy theo cảm nhận của con người, thủy vân có thể được chia ba loại khác nhau

+ Thủy vân hiện: hiển thị cho người xem thông tin về sản phẩm dưới dạng các hình mờ.

+ Thủy vân ẩn bền vững: được nhúng bằng cách thay đổi trên điểm ảnh sao cho hệ thống cảm giác của con người không thể nhận thấy và phải chịu được các thông tác xử lý tín hiệu thông thường “tấn công” và nó chỉ có thể được phục hồi với cơ chế giải mã thích hợp mà thôi. Xét theo tính bí mật của thủy vân bền vững được phân loại nhỏ hơn như sau :

-Lược đồ “thủy vân” bí mật : Cần tới ảnh gốc để trích xuất thủy vân. Có 2 loại lược đồ thủy vân bí mật :

Loại 1: yêu cầu cả ảnh bị biến đổi và ảnh gốc khi trích xuất thủy vân. Ảnh gốc được sử dụng để tìm kiếm vị trí thủy vân trong bức ảnh bị biển đối.

Loại 2: trong đó yêu cầu một bản sao của thủy vân trong quá trình trích xuất và kiểm tra mới có thể biết được thủy vân tcos ở trong bức ảnh cần kiểm tra hay không

Trong hai loại trên khi trích xuất thủy vân cần đòi hỏi có chìa khóa bí mật. Đối với loại thứ nhất thì chìa khóa bí mật ở đây là ảnh gốc, còn đối với loại chìa khóa thứ 2 thì chìa khóa bí mật là dữ liệu bí mật được sử dụng để nhúng vào bức ảnh (hay nói cách khác là thủy vân )

- Lược đồ thủy vân nửa bí mật

Không sử dụng ảnh gốc trong quá trình xác định thủy vân. Tuy nhiên, lược đồ này chỉ đưa ra thông tin có sự hiện diện của thủy vân hay không .

- Lược đồ thủy vân mù

Trong lược đồ này, không yêu cầu ảnh gốc lẫn thủy vân được nhúng trong quá trình trích thủy vân

- Lược đồ thủy vân khóa công khai

Còn được gọi là thủy vân bất đối xứng. Trong lược đồ này, chìa khóa để tìm kiếm và trích xuất thủy vân được công khai với mọi người trái ngược với thủy vân bí mật chìa khóa để tìm kiếm và trích xuất thủy vân là chìa khóa bí mật. Biết được khóa công khai “khó ” mà tính được khóa bí mật và khóa bí mật được sử dụng để nhúng và loại bỏ thủy vân.

+ Thủy vân ẩn dễ vỡ : được nhúng theo cách mà bất kỳ biến đổi hay giả mạo đều làm thay đổi hay phá hủy “thủy vân”.

+ Thủy vân hiện và ẩn đồng thời : (dual watermark) là sự kết hợp giữa thủy vân ẩn và thủy vân hiện [9]

* 1. Các hướng ứng dụng của thủy vân
     1. Bảo vệ bản quyền ảnh số

[9] Mặc dù đã có nhiều quy định về bảo vệ bản quyền và đã có những chuyển biến tích cự trong việc thực thi quyền tác giả, nhưng vẫn chưa đủ. Nhưng hành động xâm phạm bản quyền tác giả diễn ra tràn lan, tinh vi và công khai trước sự bất lực của chủ sở hữu. Đặc biệt với dữ liệu số như ảnh số với nhiều định dạng thì vấn đề bảo vệ bản quyền trở nên khó khăn hơn.

Trong việc mua bán và trao đổi các tác phẩm số này nảy sinh các vấn đề cụ thể như sau:

+ Vấn đề thứ 1 là phải bảo đảm quyền tác giả.: Để bảo vệ được bản quyền của người sở hữu ảnh số thì ảnh số đó phải có những thông tin đặc biệt chứng minh nó là thuộc quyền sở hữu của minh.

+ Vấn đề thứ 2 là đảm bảo thông tin sẵn sang cho người dùng hợp pháp và chống phân phối bất hợp pháp nội dung tác phẩm: mua bán,…

+ Vấn đề thứ 3 lần vết thông tin phát hiện người phân phối sản phậm bất hợp pháp: khi vấn đề về vi phạm bản quyền xảy ra hoặc khi chủ sở hữu sản phẩm số nghi ngờ là có bản sao sản phẩm không hợp lệ.

Đây là ứng dụng cơ bản nhất của kỹ thuật thủy vân. Trong thực tế , nhiều tác phẩm đã có tác quyền nhưng vẫn bị sử dụng sai mục đích. Các thông báo tác quyền này thường được đặt ở một vị trí nào đó trên tác phẩm phân phối.

Do các dấu thủy vân có thể vừa không thể nhìn thấy vừa không thể tách rời tác phẩm chứa nó nên sẽ là giải pháp tốt nhất cho việc bảo vệ bản quyền tác giả. Dấu thủy vân (một thông tin nào đó mang ý nghĩa quyền sở hữu tác giả) sẽ được nhúng vào trong các sản phẩm, dấu thủy vân đó chỉ người chủ sở hữu hợp pháp các sản phẩm đó và được dùng làm minh chứng cho bản quyền sản phẩm.

**Xác thực thông tin và phát hiện xuyên tạc thông tin:** : dấu thủy vân không chỉ được dùng để chỉ ra thông tin bản quyền tác giả mà còn được dùng để xác thực thông tin và phát hiện ra xuyên tạc thông tin. Dấu thủy vân sẽ được nhúng trong một tác phẩm sau đó được lấy ra và so sánh với dấu thủy vân ban đầu. Nếu có sự sai lệch chứng tỏ tác phẩm gốc đã bị tấn công và xuyên tạc. Các thủy vân nên được ẩn để tránh sự tò mò của đối phương, hơn nữa việc làm giả các thủy vân hợp lệ hay xuyên tạc thông tin nguồn cũng cần xét đến. Trong các ứng dụng thực tế, người ta mong muốn tìm được vị trí bị xuyên tạc cũng như phân biệt được các thay đổi (ví dụ như phân biệt một đối tượng đa phương tiện chứa thông tin giấu bị thay đổi, xuyên tạc nội dung hay chỉ bị nén mất dữ liệu). Yêu cầu chung đối với ứng dụng này la khả năng giấu thông tin cao và thủy vân không bền vững.

**Dấu vân tay hay dán nhãn:** thủy vân trong những ứng dụng này được sử dụng để nhận diện người gửi hay người nhận một thông tin nào đó. Ví dụ các vân khác nhau sẽ được nhúng vào các bản copy khác nhau của thông tin gốc trước khi chuyển cho nhiều người.. Những ứng dụng này, yêu cầu là đảm bảo độ an toàn cao cho các thủy vân, tránh khả năng xóa dấu vết trong khi phân phối

**Điều khiển truy cập:** các thiết bị phát hiện thủy vân (ở đây sử dụng phương pháp phát hiện thủy vân đã giấu mà không cần thông tin gốc) được gắn sẵn vào trong các hệ thống đọc ghi, tùy thuộc vào việc có thủy vân hay không để điều khiển (cho phép/ cấm) truy cập. Ví dụ hệ thống quản lý sao chép DVD được ứng dụng ở nhật. [9]

* + 1. Đặc tính của thủy vân

[9] Trước đây, đã có một số bài báo thảo luận về đặc tính của thủy vân. Một số thuộc tính thường được thảo luận như: tính phức tạp, tính trung thực hình ảnh, độ tin cậy phát hiện, tính bền vững, dung lượng, bảo mật,…Trong thực tế, không thể để thiết kế một hệ thống thủy vân đảm bảo được tất cả các thuộc tính trên. Do đó, việc đảm bảo cân bằng giữa các thuộc tính là thực sự cần thiết và vấn đề đảm bảo cân bằng phải dựa trên sự phân tích ứng dụng một cách cẩn thận [9]

* **Độ trung thực**

[9] Độ trung thực nghĩa là người theo dõi không thể phát hiện ra dấu thủy vân hay nói cách khác dấu thủy vân không làm giảm chất lượng hình ảnh. Để tín hiệu thực sự là không thể cẩm thấy thì thông tin phải được nhúng vào những bít ít quan trọng. Tuy nhiên, tín hiệu lại dễ dàng bị loại bỏ trong quá trình nên có tổn thất thông tin.

Các nghiên cứu trước đây về thủy vân đều tập trung hầu hết vào việc thiết kế thủy vân không thể thấy được và thường nhúng thủy vân vào trong vùng tín hiệu ít quan trọng về mặt cảm nhận, ví dụ như tần số cao hoặc các bít ít quan trọng. Tuy nhiên, gần đây, các kỹ thuật khác (như kỹ thuật trải phổ)lại chèn giấu thủy ký không thấy được vào trong vùng tín hiệu quan trọng về mặt cảm nhận. Đặt dấu thủy ký trong vùng tín hiệu quan trọng về mặt cảm nhận còn có thể nâng cao tính bền vững chống lại các quá trình xử lý tín hiệu [9]

* **Tính bền vững**

[9] Hình ảnh được thủy vân có thể phải trải qua nhiều loại xử lý biến đổi khác nhau, ví dụ, tăng độ tương phản, lọc thông, làm mờ,…

Do vậy, dấu thủy ký phải có tính bền vững mới chịu được các phép biến đổi ảnh cũng như biến đổi tín hiệu số thành tín hiệu tương tự, tương tự thành số và nén.

Ngoài ra, ảnh chứa thủy vân phải chịu được các phép biến đổi hình học như di chuyển vị trí, co dãn kích thước và cắt xén

Thủy vân đạt được tính bền vững thực sự khi: dấu thủy vân ký vẫn còn trong dữ liệu sau khi biến đổi và bộ phát hiện/ trích xuất vẫn có thể phát hiện ra thủy vân. Ví dụ, dấu thủy vân vẫn còn tồn tại trong ảnh sau khi phép biến đổi hình học nhưng thuật toán trích xuất/ phát hiện chỉ phát hiện và đưa ra thủy vân sau khi loại bỏ phép biến đổi. Trong trường hợp, không xác định rõ phép biến đổi để thực hiện biến đổi ngược thì bộ phát hiện/ trích xuất không thể phát hiện và đưa ra thủy vân mặc dù thủy vân vẫn tồn tại trong ảnh số.

Thủy vân đạt được tính bền vững thực sự khi: dấu thủy vân ký vẫn còn trong dữ liệu sau khi biến đổi và bộ phát hiện/ trích xuất vẫn có thể phát hiện ra thủy vân. [9]

Ví dụ, dấu thủy vân vẫn còn tồn tại trong ảnh sau khi phép biến đổi hình học nhưng thuật toán trích xuất/ phát hiện chỉ phát hiện và đưa ra thủy vân sau khi loại bỏ phép biến đổi. Trong trường hợp, không xác định rõ phép biến đổi để thực hiện biến đổi ngược thì bộ phát hiện/ trích xuất không thể phát hiện và đưa ra thủy vân mặc dù thủy vân vẫn tồn tại trong ảnh số.

* **Tính dễ hỏng**

[9] Là thuộc tính đối ngược hoàn toàn với tính bền vững của thủy vân. Thuộc tính này thường được ứng dụng trong lược đồ thủy vân vỡ. Với lược đồ này yêu cầu đặt ra là dấu thủy ký hoặc bị phá hủy bởi bất cứ phương pháp sao chép nào ngoại trừ các phương pháp sao chép hợp pháp. Ví dụ, thủy vân đặt trong một văn bản hợp pháp tồn tại qua bất cứ lần sao chép nào mà không thay đổi nội dung nhưng sẽ bị phá hủy nếu có câu trong nội dung bị thay đổi. Yêu cầu này không giống với chữ ký số trong kỹ thuật mã hóa, trong đó, có thể xác thực tính nguyên vẹn của các bít một cách chính xác nhưng không thể phân biệt các mức biến đổi có thể chấp nhận được.[9]

* **Tỉ lệ lỗi sai đường**

[9] Tỉ lệ lỗi sai dương là xác suất hệ thống phát hiện nhầm: xác định một mẩu dữ liệu không mang dấu thủy ký là mang dấu thủy ký. Tùy theo ứng dụng mà ảnh hưởng của lỗi là khác nhau, trong một số ứng dụng có thể là rất nghiêm trọng. Do đó, trong ứng dụng, người ta phát tính toán trước sao cho tỷ lệ lỗi sai dương nhỏ hơn mức cho phép.

* **Tính dư thừa**

Tính dư thừa liên quan đến một thực tế là thủy vân được lặp lại ở những vùng tấn số khac nhau, do đó nếu có một lỗi trên một vùng tần số thì vẫn có thể được khôi phục thông điệp từ các dải tần khác. Tính dư thừa ánh xạ đến tính bền vững, có nghĩa là thủy vân có thể được khôi phục ngay cả khi nó bị biến đổi ở độ nhất dịnh do sự vô ý hay tấn công có chủ ý.

* **Đa thủy vân**

Một kẻ tấn công có thể thủy vân lại một đối tượng đã đóng dấu thủy vân và sau đó tuyên bố sản phẩm thuộc quyền sở hữu của mình. Một giải pháp đơn giản nhất trong trường hợp này là gán nhãn thời gian cho thông tin thủy vân với sự có mặt của cơ quan chứng thực hay có thể nhúng thủy vân khác nhau với những người sử dụng khác nhau. Với phương pháp nhúng nhiều thủy vân cho phép lần vết theo nội dung thủy vân nhưng lại tạo điều kiện cho phép tấn công loại bỏ bằng cách lấy trung bình xác suất (tấn công đồng thời).

* **Độ phức tạp tính toán**

Cũng như bất cứ công nghệ nào sử dụng trong thương mại, độ phức tạp tính toán của lược đồ thủy vân đều rất quan trọng. Điều này, đặc biệt đúng khi sử lý với các dữ liệu thời gian thực.

Mặt khác, cần phải xem xét tính co giãn của độ phức tạp tính toán. Người thiết kế lược đồ thủy vân luôn mong muốn thiết kế được lược đồ mà quy trình nhúng và phát hiện thủy vân có tính co giãn theo các thế hệ của máy tính. Ví dụ, lược đồ thủy vân thế hệ đầu tiên có độ phức tạp tính toán không lớn nhưng độ tin cậy không cao so với lược đồ thủy vân thế hệ tiếp theo. Nhưng khi giải quyết một vấn đề tính toán lớn thì lược đồ thủy vân ở thế hệ sau lại làm việc tốt hơn. [9]

* 1. Yêu cầu đối với phương pháp thủy vân

Khi thực hiện thủy vân ảnh số, cần phải có một số tiêu chí để đánh giá chất lượng của giải thuật. Thông thường người ta dựa trên các tính chất sau:

* **Bảo đảm tính vô hình**

[9] Quá trình thủy vân sẽ làm biến đổi ảnh mang do thủy vân được nhúng vào. Tính “vô hình” thể hiện mức độ biến đổi ảnh mang.

Lược đồ thủy vân hiệu quả, sẽ làm cho thủy vân trở nên “vô hình” trên ảnh mang làm cho người khác khó có thể nhận ra, do vậy đảm bảo được tính bí mật của thủy vân. Tuy nhiên trong thực tế không phải khi nào người ta cũng cố gắng để đạt được tính vô hình cao nhất, ví dụ trong thủy vân hiện thủy vân được sử dụng để làm biểu tượng xác thực nguồn gốc sản phẩm, do vậy không nhất thiết phải là bí mật, nhiều khi cần lộ ra cho mọi người biết để mà dè chừng.

* **Khả năng chống giả mạo (tính toàn vẹn)**

Đối với thủy vân thì khả năng chống giả mạo là yêu cầu vô cùng quan trọng vì có như vậy mới bảo vệ được bản quyền, minh chứng cho tính pháp lý của sản phẩm. Để có thể chống lại giả mạo thì bất cứ sự thay đổi nào về nội dung của các ảnh số thì thủy vân này sẽ bị hủy đi. Do đó, rất khó làm giả các ảnh số có chứa thủy vân.

* **Tính bền vững**

Yêu cầu thứ 3 là thủy vân phải bền vững. Thủy vân phải có khả năng tồn tại cao với các hình thức tấn công có chủ đích và không có chủ đích. Các tấn công không có chủ đích đối với ảnh số bao gồm như nén ảnh, lấy mẫu, lọc, chuyển đối A/D và D/A

Tấn công có chủ đích có thể là việc xóa, thay đổi hoặc làm nhiễu thủy vân trong ảnh. Để thực hiện được điều này, thủy ấn phải được dấu trong các vùng quan trọng đối với trực giác. Phương pháp thủy vân phải đám bảo sao cho việc không thể lấy lại thủy vân tương đương với việc ảnh bị biến đổi quá nhiều , không còn giá trị về thương mại.

* **Dung lượng**

Với yêu cầu này, thủy vân nhúng vào ảnh phải đủ dùng trong ứng dụng mà không làm thay đổi quá nhiều chất lượng ảnh.

Việc giấu thủy vân trong ảnh thì ta bắt buộc phải thay đổi dữ liệu ảnh. ta có thể tăng tính bền vững cho thủy vân bằng cách tăng lượng thay đổi ảnh cho mỗi đơn vị cần giấu. nhưng, nếu thay đổi quá nhiều thì tính ẩn không còn được đảm bảo nữa. Còn nếu thay đổi ảnh quá ít thì các yếu tố dùng để xác định thủy vân trong ảnh sau các phép tấn công có thể không đủ để xác định thủy vân. nếu thông tin được giấu quá nhiều thì cũng dễ làm thay đổi chất lượng ảnh, và làm giảm tính bền vững. Vì vậy, lượng thay đổi ảnh lớn nhất có thể chấp nhận và tính bền vững là nhân tố quyết định cho khối lượng tin được giấu trong ảnh.

Trong thực tế, người ta luôn phải cân nhắc giữa chất lượng (tính bí mật, tính toàn vẹn, tính bền vững) và dung lượng thủy vân. [9]

* 1. Khả năng tấn công trên hệ thống thủy vân số

[9] Thủy vân bền vững phải vượt qua được các tấn công ngẫu nhiên và cố ý.

* **Tấn công đơn giản:** là dạng tấn công làm hỏng thủy vân đã được nhúng bằng cách thao tác lên toàn bộ dữ liệu được nhúng thủy vân mà không có ý định nhận dạng để lấy tách thủy vân.
* **Tấn công phát hiện:** Là sự tấn công với mục đích loại bỏ đi mối quan hệ và vô hiệu quá khả năng khôi phục thủy vân, làm cho bộ phát hiện không thể xác định được thủy vân. Điều này được thực hiện chủ yếu bằng cách thay đổi hình dạng hình học như phóng to, thu nhỏ, xoay, cắt xén, xóa hoặc chèn thêm các điểm ảnh và phép biến đổi hình học
* **Tấn công nhập nhằng:** là sự tấn công với mục đích gây nhầm lẫn bằng cách tạo ra dữ liệu gốc giả hoặc dữ liệu đã được nhúng thủy vân giả. Ví dụ: kẻ tấn công có thể làm giảm tính xác thực của thủy vân bằng cách nhúng một hoặc nhiều thủy vân bổ sung sao cho thủy vân mới không thể phân biệt được với thủy vân ban đầu – thủy vân dùng để xác thực.
* **Tấn công loại bỏ:** nhằm mục đích phân tích để xác định ra thủy vân hoạc dữ liệu gốc , tách dữ liệu đã được nhúng thủy vân thành dữ liệu gốc và thủy vân. [9]
  1. Phạm vi đồ án

Đồ án tập trung nghiên cứu các kỹ thuật thủy vân trên ảnh số. Ứng dụng mà luận văn xây dựng là hệ thống nhúng và tách thủy vân nhằm xác thực nội dung thông tin và bảo vệ bản quyền ảnh số.

* 1. Mục tiêu, sự cần thiết của đề tài, tại sao lại chọn đề tài này ?
* **Mục tiêu**

[9] Mục đích của luận văn là nghiên cứu hệ thống thủy vân số và các hướng ứng dụng của thủy vân số chủ yếu là ứng dụng trong bảo vệ bản quyền ảnh số.Tập trung vào phân tích các thuật toán thủy vân số. Từ đó, xây dựng chương trình thử nghiệm cài đặt một số thuật toán thủy vân nhằm ứng dụng xác thực thông tin và bảo vệ bản quyền cho dữ liệu ảnh số. [9]

* **Sự cần thiết của đề tài**

[3] Ngày nay, với sự phát triển mạnh mẽ của các mạng máy tính tốc độ cao, đặc biệt là Internet, các phương tiện kỹ thuật số như phương tiện lưu trữ, phương tiện truyền thông, đã mở ra một kỷ nguyên mới – kỷ nguyên thông tin số. Hầu hết các thông tin ngày nay đều được lưu trữ dưới dạng số hóa. Đồng thời, quá trình toàn cầu hóa mạng Internet đã biến xã hội ảo là nơi diễn ra trao đổi thông tin trong mọi lĩnh vực chính trị, quân sự, quốc phòng, kinh tế, thương mại. Tuy nhiên, công nghệ số cũng tạo ra khả năng sao chép hoàn hảo, không có bất kỳ khuyết điểm và phân phối lại những sản phẩm này trên toàn thế giới, có hoặc không sự cho phép của người sở hữu. Việc trao đổi, phân bố, sao chép và xử lý các sản phẩm số này ngày càng nhanh chóng, đơn giản, nằm ngoài tầm kiểm soát của các tổ chức. Vấn đề đặt ra cho tất cả các phương thức kinh doanh, phân phối tài nguyên số trên mạng là tuân thủ các nguyên tắc về quyền sở hữu trí tuệ, và không cản trở quá trình phân phối, trao đổi tài nguyên số. Nhu cầu được bảo vệ bản quyền và sở hữu trí tuệ các sản phẩm số đã trở thành một vấn đề quan trọng và đang được quan tâm [3]

Hiện nay, có hàng tỉ bức ảnh được phân phối trên các kênh truyền công cộng. Do chúng có đặc tính dễ sao chép, dễ chỉnh sửa nên nhiều đối tượng lợi dụng cố ý đánh cắp, làm sai lệch, giả mạo bức ảnh gốc. Từ đó, có thể gây thiệt hại đến uy tín, thiệt hại về kinh tế cho người sở hữu bức ảnh đặc biệt trong bối cảnh bùng nổ Internet.

* **Tại sao lại chọn đề tài này**

Để giải quyết cho các vấn đề an toàn truyền thông vào bảo vệ bản quyền tài liệu số đặc biệt là ảnh số thì việc xây dựng một hệ thống có sử dụng kỹ thuật nhúng thủy vân vẫn là một giải pháp tối ưu. Thuỷ vân số là một phương pháp mới dựa trên lý thuyết tổng hợp của nhiều lĩnh vực khác nhau như mật mã học, lý thuyết thông tin, lý thuyết truyền thông và xử lý tín hiệu số, xử lý ảnh. Bằng cách sử dụng thủy vân, dữ liệu số sẽ bảo vệ khỏi sự sao chép bất hợp pháp. Tạo thủy vân là một phương pháp nhúng một lượng thông tin nào đó vào trong dữ liệu đa phương tiện cần được bảo vệ sở hữu mà không để lại ảnh hưởng nào đến chất lượng của sản phẩm. Thủy vân luôn gắn kết với sản phẩm đó. Bằng trực giác khó có thể phát hiện được thủy vân trong dữ liệu chứa, nhưng có thể tách chúng bằng các chương trình có cài đặt thuật toán thủy vân. Thủy vân được tách từ dữ liệu số chính là bằng chứng kết luận dữ liệu số có bị xuyên tác thông tin hay vi phạm bản quyền hay không.

Chính vì tính hữu ích trong ứng dụng thực tiễn của thủy vân số nên em quyết định lựa chọn đề tài là: “Hệ thống thủy vân số và ứng dụng thủy vân số trong bảo vệ bản quyền ảnh số”.

1. KỸ THUẬT THỦY VÂN SỐ
   1. Giới thiệu về kỹ thuật giấu thông tin

[1] Giấu thông tin là kỹ thuật giấu (nhúng) một lượng thông tin nào đó vào trong một đối tượng dữ liệu số khác nhằm giữ bí mật và xác thực thông tin.Giấu thông tin trong ngôn ngữ, hành văn cũng được sử dụng từ lâu, với việc lựa chọn theo một quy luật xác định như ghép các từ đầu của mỗi câu hay nhặt các từ, các chữ cái theo một quy luật nào đó. Một hình thức cũng được người Trung Hoa sử dụng rất nhiều là sử dụng một tờ giấy mẫu có đục các lỗ làm một “mặt nạ” viết trước các nội dung cần giấu tại các vị trí đó, tiếp theo các nội dung “vô hại” khác lên sau, cách làm này tận dụng đặc điểm về chữ tượng hình và các từ đồng âm khác nghĩa của chữ Hán

Kỹ thuật thuỷ vân bắt đầu được sử dụng muộn hơn, vào cuối thế kỷ 13 tại Ý, khi các nhà sản xuất giấy làm các hình mờ chìm trong giấy in để bảo vệ bản quyền nhà sản xuất. Watermark xuất phát từ đó, sau đó đã nhanh chóng phát triển ra toàn Châu Âu, áp dụng trong các lĩnh vực xuất bản khác cho tới ngày nay. [1]

* 1. Phân loại về kỹ thuật giấu thông tin
     1. Giấu thông tin bí mật

[1] Đây là ứng dụng phổ biến nhất từ trước tới nay. Đối với giấu thông tin bí mật người ta quan tâm chủ yếu tới các mục tiêu sau:

* Độ an toàn của tin giấu (khả năng không bị phát hiện của tin giấu).
* Lượng thông tin tối đa có thể giấu trong một phương tiện chứa cụ thể mà vẫn có thể đảm bảo an toàn
* Độ bảo mật của thông tin trong trường hợp giấu tin bị phát hiện.

Giấu thông tin bí mật không quan tâm nhiều tới các yêu cầu về khả năng bền vững của phương tiện chứa. Việc giải mã để nhận được thông tin cũng không cần phương tiện chứa gốc ban đầu. Các yêu cầu về khả năng chống tấn công không được quan tâm lắm, thay vào đó là thông tin giấu phải được bảo mật. Đối với các thuật toán giấu thông tin mật, người ta không chú trọng đến việc bảo vệ thông tin mật trước sự tấn công của các đối thủ mà thay vào đó quan tâm đến tính ẩn và tính an toàn đối với dữ liệu cần giấu [1]

* + 1. Giấu thông tin thủy vân

[1] Khác với kỹ thuật giấu thông tin để giữ bí mật thông tin, giấu thông tin thủy vân có mục tiêu là bảo vệ bản quyền và xác thực thông tin. Vì vậy, kỹ thuật này không chống lại việc khai thác thông tin, mà quan trọng nhất đối với nó là đảm bảo tuyệt đối tính bền vững, nghĩa là không thể hủy bỏ được thông tin giấu nó trừ khi hủy chính sản phẩm chứa. Ngoài ra các thông tin nhúng cần có ảnh hưởng tối thiểu đối với phương tiện chứa, vì vậy thông tin cần giấu càng nhỏ càng tốt. (Giới thiệu về kỹ thuật giấu thông tin) [1]

|  |  |  |
| --- | --- | --- |
| **Giấu thông tin mật thủy vân số (Watermarking)** | | |
|  | **Giấu thông tin mật** | **Thủy vân số** |
| **Mục tiêu** | * Tàng hình các phiên liên lạc để bảo mật thông tin * Dùng trong các liên lạc xác định | * Chủ yếu phục vụ cho mục đích bảo vệ bản quyền * Chủ yếu dùng trong các hoạt động xuất bản |
| **Cách thực hiện** | * không làm “thay đổi” phương tiện chứa thông tin | * Có thể tác động nhỏ về cảm nhận tới phương tiện chứa |
| **Yêu cầu** | * Giấu được nhiều thông tin nhất * Không cần quan tâm tới độ bền của phương tiện chứa * Không thể quan sát được việc giấu thông tin Không kiểm tra được nếu không có khoá thích hợp | * Chỉ cần nhúng ít dữ liệu * Dữ liệu nhúng cần phải bền vững * Đảm bảo trước các phương pháp nén dữ liệu * Dữ liệu nhúng có thể nhận thấy hay không nhận thấy * Không kiểm tra được nếu không có khoá thích hợp |

* 1. Tìm hiểu phương pháp và mô hình thủy vân số
     1. Phương pháp thủy vân số

[2] Phương pháp thủy vân số (digital watermarking) là một phương pháp nhúng thông tin vào một tài liệu (như hình ảnh, âm thanh, video) sao cho thông tin này không thể bị xóa bỏ mà không ảnh hưởng đến tính năng hoặc chất lượng của tài liệu ban đầu. Một số phương pháp thủy vân số phổ biến bao gồm:

* Kỹ thuật nhúng dấu số: Phương pháp này nhúng thông tin vào các bit của tài liệu, ví dụ như sử dụng các bit ít có giá trị để chứa thông tin.
* Kỹ thuật nhúng dấu hình ảnh: Phương pháp này thêm các đối tượng nhỏ và không nổi bật vào tấm hình để chứa thông tin thủy vân số.
* Kỹ thuật nhúng dấu số trong âm thanh: Phương pháp này nhúng thông tin vào các tần số không được sử dụng trong tín hiệu âm thanh.
* Kỹ thuật nhúng dấu số trong video: Phương pháp này nhúng thông tin vào các khung hình của video, ví dụ như thêm một số pixel ẩn vào khung hình để chứa thông tin thủy vân số.
* Kỹ thuật mã hóa dữ liệu thủy vân số: Phương pháp này mã hóa thông tin thủy vân số và sau đó nhúng nó vào tài liệu bằng cách sử dụng các thuật toán mã hóa. [2]
  + 1. Mô hình thủy vân số

1. **Tạo thủy vân số**

[9] Thủy vân có thể là một hình ảnh dạng logo hay văn bản với độ dài cho trước. Thủy vân dạng hình ảnh có khả năng chống chịu trước các phép xử lý ảnh tốt hơn nhiều só với dạng thủy vân dạng ký tự. Thủy vân có thể được biến đổi (bằng mã hóa, chuyển đổi định dạng), trước khi giấu vào ảnh. Các thuật toán nhúng thủy vân dạng logo được gọi là thuật toán thủy vân hợp nhất ảnh. Thủy vân dạng ảnh có lợi ích là dễ dàng nhận biết về mặt trực giác và đưa ra một chứng minh đúng đắn về quyền sở hữu ảnh. Bình thường sẽ có một khóa bí mật K dùng để tang tính bảo mật cho dữ liệu được nhúng. Do tính bền vững được đảm bảo hơn nên thủy vân dạng ảnh được sử dụng nhiều hơn.

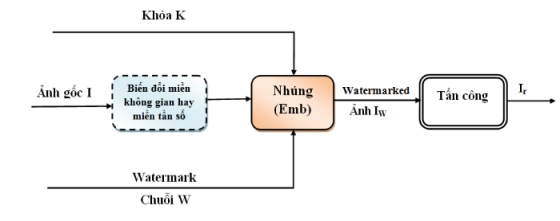
Để tăng thêm tính an toàn và dung lượng thì thủy vân trước khi nhúng vào ảnh mang có thể được mã hóa hay nén lại. Theo cơ chế này, đầu tiên thủy vân số sẽ được nén lại để lượng dữ liệu thủy vân có thể tăng lên, sau đó được mã hóa để tăng tính bảo mật cho thông tin trước khi được giấu vào ảnh mang. Tuy nhiên, giải pháp này làm tăng độ phức tạp của bài toán về phát hiện thủy vân. [9]

1. **Quy trình nhúng thủy vân**

[2] 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 được gọi là dữ liệu có bản quyền và 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



Hình 2. 1 Quy trình nhúng thủy vân

Hình trên trình bày và giải thích quá trình nhúng thủy vân cho ảnh tĩnh. Trong đó, Ảnh gốc được kí hiệu bằng I, “thủy vân” được kí hiệu bở W, hình ảnh chứa “thủy vân” là I w và K là khóa nhúng. Hàm nhúng EMB có đầu vào là ảnh gốc I, “thủy vân” W và khóa K và tạo ra một ảnh mới có chứa thủy vân mới thể hiện bằng Iw

Khóa nhúng K là thực sự cần thiết cho việc nâng cao khả năng bảo mật của hệ thống “thủy vân”. Trước quá trình nhúng, hình ảnh gốc có thể được chuyển đổi sang miền tần số hoặc nhúng có thể được thực hiện biến đổi sang miền không gian. Miền được chọn phụ thuộc vào việc lựa chọn kỹ thuật “thủy vân”. Nếu quá trình nhúng được thực hiện trong miền tần số, biến đổi nghịch đảo được áp dụng để thu được hình ảnh chứa “thủy vân”. Biểu thức toán học cho hàm nhúng có thể được thể hiện như sau [2]:

Đối với kỹ thuật biến đổi theo miền không gian :



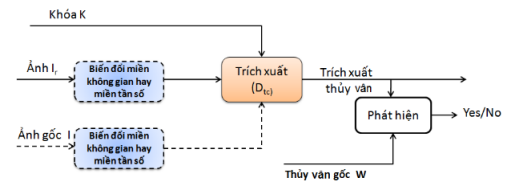
Đối với kỹ thuật biến đổi theo miền tần số :



Trong đó f là vecto hệ số cho phép biến đổi.

1. **Trích xuất và tìm kiếm thủy vân**

[2] 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 đi.



Hình 2. 2 Quy trình trích xuất và tìm kiếm thủy vân

Hình trên trình bày và giải thích quy trình trích xuất và tìm kiếm thủy vân ở trong ảnh tĩnh. Một hàm phát hiện D tc có đầu vào là hình ảnh I r có chức năng xác định quyền sở hữu sản phẩm. Các hình ảnh I r có thể chứa thủy vân hoặc không chứa thủy vân. Trong trường hợp tổng quát, hình ảnh có thể bị biến đổi. Hàm phát hiện có khả năng khôi phục thủy vân We từ bức ảnh hoặc kiểm tra sự hiện diện của thủy vân W trong bức ảnh đã cho I r hay không. Trong quá trình này hình ảnh gốc I cũng có thể yêu cầu, phụ thuộc vào lược đồ thủy vân được lựa chọn.

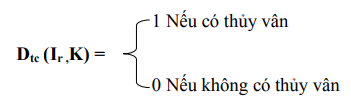
Biểu thức toán học cho thủ tục trích xuất mù (trích xuất không sử dụng ảnh gốc I) cụ thể như sau :



Biểu thức toán học cho thủ tục trích xuất không mù (trích xuất có sử dụng ảnh gốc I) cụ thể như sau :



Thuật toán phát hiện thủy vân mù có đầu ra là một giá trị nhi phân cho biết có sự hiện diện của thủy vân hay không. Bởi vậy, có thể giả sử:



Trong lược đồ tách thủy vân phải được trích xuất một cách chính xác, nguyên mẫu. Lược đồ trích xuất thủy vân có thể chúng mưng được quyền sở hữu, trong khi lược đồ phát hiện thủy vân có thể xác nhận có sự hiện diện của thủy vân hay không. [2]

* 1. Tìm hiểu về các thuật toán thủy vân theo miền không gian ảnh (SW; WU- LEE; LBS; PCT)
     1. Thuật toán SW

[9] Là một thuật toán được sử dụng trong xử lý tín hiệu số và xử lý ngôn ngữ tự nhiên. Thuật toán này hoạt động bằng cách di chuyển một cửa sổ trượt qua các tín hiệu đầu vào hoặc văn bản, thực hiện tính toán trên cửa sổ đó và trả về kết quả tương ứng.

**Quy trình nhúng thủy vân**

- Chia F thành các khối kích thước m x n.

- Với mỗi khối B trong F ta xét khả năng giấu một bit dữ liệu di của d theo các bước :

Bước 1: Tính tổng SUM[B] các điểm đen trong khối B, đặt t = SUM[B] mod 2

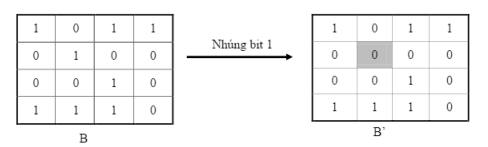
Bước 2: So sánh tính chẵn lẻ giữa t và di

Nếu t và di cùng tính chẵn lẻ thì khối B mặc nhiên đã giấu được bit di mà không cần làm gì.

Nếu t và di khác tính chẵn lẻ thì ta sẽ đảo 1 bit trong B. Chính sách đảo bit: nếu số điểm đen và điểm trắng xấp xỉ nhau thì chọn ngẩu nhiên 1 bit để đảo. Nếu có nhiều điểm đen và có điểm trắng thì sửa điểm trắng thành điểm đen. Ngược lại sẽ sửa điểm đen thành điểm trắng.

Ví dụ minh họa: Giả sử giấu một bit dữ liệu b = 1 vào khối B.

Ta có SUM (B) = 8. Do SUM (B) mod 2 = 0 nên khối B không thõa mãn yêu cầu để giấu bit 1. Muốn giấu bit 1 vào khối này ta cần phải thay đổi khối bằng cách chọn một bit bất kỳ và đổi từ 0 sang 1 và từ 1 sang 0. Giả sử ta đảo lại bit tại vị trí B [2,2] ta được khối B đã được nhúng bít 1.



Hình 2. 3 Minh họa thuật toán SW: nhúng bit 1 vào khối ảnh B.

Giả sử vẫn với khối B đã cho như trên nhưng ta phải giấu bit dữ liệu b= 0 và khối đó. Ta thấy do Sum (B) = 8 nên Sum (B) mod 2 = 0. Khối B được bảo toàn và bit dữ liệu b = 0 xem như được giấu. [9]

**Quá trình tách thủy vân**

[9] Trong thuật toán thủy vân này, khóa đơn giản là kích thước của khối, tức là bộ số (m, n). Nếu biết kích thước của khối thì dễ dàng trích lại dữ liệu d theo ccs bước:

Bước 1: Chia ảnh có nhúng thủy vân B‟ thành các khối có kích thước m x n với mỗi khối Bi‟ trong B‟ ta tính Sum [Bi‟]

Bước 2: tách thủy vân theo cách xét

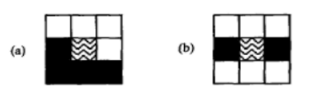
+ Nếu Sum [Bi‟] là chẵn thì bit di = 0

+ Ngược lại , nếu Sum [Bi‟] là lẻ thì bit di = 1

**Nhận xét**

Với thuật toán này việc chọn khối khá là đơn giản: ta có thể bắt đầu từ khối đầu tiên và các khối tiếp theo một cách tuần tự. Tuy nhiên, ta có thể chọn ngẫu nhiên một khối chưa giấu ở mỗi lần giấu, hoặc chọn các khối theo một thuật toán xác định kèm theo một khóa K. Khi đó, ta đã làm tang được độ an toàn của thuật toán vì khóa bây giờ còn thêm cả chỉ số khối đã giấu tin cho từng bit. Hoặc ta có thể thay đổi kích thước khối mỗi lần giấu, chẳng hạn như khối thứ nhất có kích thước là 8 x 8 thì khối thứ 2 có kích thước 8x12 trong trường hợp này khóa sẽ gồm cả kích thước khối của mỗi lần giấu.

Kỹ thuật trên sẽ gặp phải hiện tượng gây bất thường đối với ảnh sau khi giấu thông tin đặc biệt khi chọn vào những khối ảnh một màu, chẳng hạn như một khối màu đen hoặc toàn trắng. Khi đó, nếu cần đảo giá trị một bit thì vị trí bit đảo sẽ khác biệt hoàn toàn với các bit trong khối và dễ nhận biết có sự thay đổi. Vì vậy, để xác định nên thay đổi bit nào khối bit ta phải tính hệ số ảnh hưởng của bit đó khi nó bị thay đổi. Hệ số này tính bằng cách xét sự thay đổi về tính trơn và tính liên kết với các điểm láng giềng. Tính trơn được đo theo sự chuyển đổi mức xám theo chiều ngang và chiều dọc, đường chéo trong cửa sổ 3x 3. Tính liên kết được tính bằng số nhóm điểm đen và số điểm trắng. Ví dụ: Nếu đảo một điểm ảnh trong hình (a) sẽ ít bị chú ý hơn điểm ảnh trong hình. [9]



Hình 2. 4 Minh họa chọn điểm ảnh giấu tin vào những khối ảnh màu

* + 1. Thuật toán WU-LEE

[9] Thuật toán này của hai 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 bit và mỗi khối bằng cách hiệu chỉnh nhiều nhất 1 bit 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 m x n . Để đơn giản ta giả sử kích thước của ảnh gốc F là bội số của m x n . Quá trình nhúng thu được ảnh F có một số bit đã bị hiệu chỉnh. Thuật toán thực hiện như sau [9]:

**Quá trình nhúng thủy vân**

+[9] Bước 1 : Chia F thành các khối, mỗi khối có kích thước m x n .

+ Bước 2: Với mỗi khối Fi thu được ở bước 1. Kiểm tra điều kiện:



Nếu điều kiện trên đúng thì tiếp tục thực hiện bước 3 để nhúng một bit vào Fi . Ngược lại, dữ liệu sẽ không được nhúng vào Fi và Fi sẽ được giữ nguyên.

+ Bước 3: Giả sử bit được nhúng vào Fi là b. Được hiệu chỉnh Fi ta làm như sau:

Trường hợp 1: Nếu SUM (Fi ^ K) mod 2 = b thì không thay đổi Fi và bit b hiện nhiên được nhúng vào khối Fi .

Trường hợp 2: Nếu SUM (Fi ^ K) mod 2 # b và SUM (Fi ^ K) = 1 thì chọn ngẫu nhiên một bit của Fi tại vị trí (i,j) mà Fi (j,k)=0 và K (j,k)=1 và đảo Fi (j,k) thành 1.

Trường hợp 3: Nếu SUM (Fi ^ K) mod 2 # b và SUM (Fi ^ K) = SUM (K) – 1 thì chọn ngẫu nhiên một bit của Fi tại vị trí (j,k) mà K (j,k)=1 và đảo ngược Fi (j,k) thành 0.

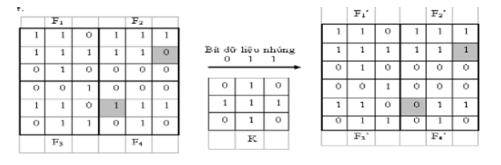
Trường hợp 4: Nếu SUM (Fi ^ K) mod 2 # b và 1< SUM (Fi ^ K) < SUM (K) – 1 thì chọn ngẫu nhiên một bit của Fi tại vị trí (j,k) mà K (j,k)=1 và đảo ngược Fi (j,k).

Trong bước 3 chỉ thực hiện tối đa một phép đảo một bit của Fi để thu được khối Fi‟ nhằm đảm bảo tính bất biến

Ví dụ minh họa:

Giả sử ta cần nhúng dãy bit d = 011 và một ảnh F có kích thước 6 x 6 với một ma trận khóa K có kích thước 3 x 3 như trong hình 2.3. Ta có Sum (K) = 5

Chia ảnh F thành bốn khối nhỏ mỗi khối sẽ có kích thước là 3 x 3 ta thu được F1, F2, F3, F4 .



Hình 2. 5 Minh họa thuật toán WU\_LEE nhúng đoạn bit 01

Áp dụng thuật toán, lần lượt nhúng các bit vào các khối như sau:

- Với F1, Vì SUM (F1 ^ K) = 5 = SUM (K) không thỏa mãn điều kiện nhúng nên không nhúng dữ liệu vào trong F1.

- Với F2, SUM (F2 ^ K) = 3 thỏa mãn điều kiện nhúng và bít cần nhúng là 0.

Vì SUM (F2 ^ K) mod 2 = 3 mod 2 # 0 và 1 < SUM (F2 ^ K) < SUM (K) – 1 nên ta chọn ngẫu nhiên một vị trí để đảo bit trong khối F2, chẳng hạn vị trí (2,3) thỏa mãn K [2,3] =1 (theo trường hợp 4). Sau khi đảo bit F2 [2,3] ta thu được khối F2 „ như trên hình ảnh (bit bị đảo được tô xám)

- Với F3, SUM (F3 ^ K)= 3 thỏa mãn điều kiện nhúng và bit cần nhúng là 1.

Ta có SUM (F3 ^ K) mod 2 = 3 mod 2 = 1 = b. Khối F3‟ thu được giữ nguyên khối F3 nhung với ý nghĩa là khối đã được giấu bit = 1 (theo trường hợp 1)

- Với F4, SUM (F4 ^ K) = 4 thỏa mãn điều kiện nhúng và bit cần nhúng là 1. Ta có SUM (F4 ^ K) mod 2 = 4 mod 2 = 0 # b và SUM (F4 ^ K) = SUM (K) – 1. Theo trường hợp 3 trong thuật toán, ta chọn vị trí (2,1) để đảo bit trong khối F4 vì với phần tử này ta có F4[2,1] = 1 và K [2,1] =1 . Sau khi đảo bit F 4 [2,1] ta thu được khối F4‟ như trên hình vẽ (bit bị đảo được tô xám). [9]

**Quá trình trích thủy vân**

[9] Phương pháp giấu tin Wu- Lee cho phép giấu nhiều nhất 1 bit dữ liệu trong 1 khối, Giả sử có được bất biến 0 < SUM (Fi ^ K) < SUM (K)

Trong thuật toán nhúng tin, tiến hành đảo 1 bit trong mỗi khối F I, sao cho tổng số bit 1 của Fi ^ K bằng tổng số bit 1 của K tức là :

SUM (Fi ^ K) = [b mod 2]

Do đó, khi xác định được 0 < SUM (Fi ^ K) < SUM ( K) thì có nghĩa là khối đó có giấu tin, Bit tin được giấu xác định bởi công thức:

b= [SUM (Fi‟ ^ K)] mod 2

**Nhận xét**

Việc chọn khóa K nhằm làm tăng độ bảo mật của thuật toán. Nếu thuật toán 1 chỉ biết kích thước khối là mx n thì đối phương rất dễ khai thác thủy vân.

Phép toán F i ^ K quy định thuật toán chỉ được phép sửa các bit trong khối Fi ứng với bit 1 trong khóa K. Như vậy, khóa K được xem như một mặt nạ, tạo ra khung nhìn cho thuật toán. Ta có thể thay phép toán ^ bằng một phép toán khác chẳn hạn phép + .

Điều kiện 0< SUM (Fi ^ K) < SUM (K) quy định nếu khối Fi ^ K toàn 0 hoặc giống như khóa K thì không được giấu tin để tránh bị lộ.

Do việc giấu tin vào khối chỉ cần thay đổi tối đa một bit nên việc chọn bit nào trong F để đảo cần tuân thủ theo nguyên tắc: Nếu F i ^ K có nhiều bit 1 thì chọn bit 1 , ngược lại nếu Fi ^ K có quá it bit 1 thì chọn bit 0. Ngược tắc này làm giảm khả năng bit đảo bị phát hiện.

Vì khóa K là bí mật nên thông tin đã nhúng là bí mật. Thuật toán này đã thay đổi nhiều nhất của một bit của khối Fi khi giấu một bit thông tin vào bên trong khối nên với một khối có kích thước m x n đủ lớn thì sự thay đổi của Fi là nhỏ. [9]

* + 1. Thuật toán LBS

[9]Về cơ bản, kỹ thuật thủy vân LBS dựa trên tần suất xuất hiện của các bit 0 và 1 trong file ảnh gốc và trong thông điệp cần mã hóa, từ đó đưa ra sự thay thế các bit này để thực hiện việc giấu tin .

Cụ thể hơn, trong kỹ thuật thủy vân LSB, bit cuối cùng của mỗi byte được đặt giá trị 0, sau đó tùy thuộc vào giá trị 0 hoặc 1 của dữ liệu mà thay đổi. Nếu bit của dữ liệu là 0 thì giữ nguyên, còn nếu bit của dữ liệu là 1 thì sẽ đổi giá trị này trên ảnh thành 1.

Để thực hiện kỹ thuật thủy vân này, cần một ảnh gốc, hay còn gọi là cover image. Do phương pháp này sử dụng những bits của từng pixcel trong ảnh, nó đòi hỏi một định dạng nén không mất thông tin. Khi ta sử dụng ảnh màu 24 bit, từng bit của mỗi màu thành phần R, G, B đều có thể được sử dụng, như vậy có thể giấu được 3 bit trong mỗi điểm ảnh, đồng nghĩa với việc nhúng được nhiều thông tin hơn.

**Dữ liệu vào**

+ Ảnh gốc

+ Dữ liệu thủy vân

+ Khóa bí mật

**Dữ liệu ra**

+ Ảnh mang: có chứa thông tin thủy vân. Ảnh mang có sự thay đổi không đáng kể so với ảnh gốc.

+ Để có thể lưu trữ lượng thông tin lớn và sự thay đổi màu sắc của ảnh là không đáng kể, chúng ta sử dụng file bitmap 24 bit . Cụ thể hơn, một pixel của ảnh được biểu diễn 3 màu đỏ, xanh lá cây và xanh da trời (R, G, B), mỗi màu sử dụng 8 bit. Tuy nhiên trong lược đồ trình bày dưới đây, chúng ta chỉ sử dụng bit cuối cùng màu xanh da trời để giấu thông tin.

**Thuật toán nhúng thủy vân**

+ Bước 1 : Tính tổng số bytes cần dùng để nhúng thủy vân. Giá trị này được lưu trong biến int S.

+ Bước 2: Đọc các kí tự từ file text sau đó chuyển giá trị ASCII của chúng sang dạng nhị phân 8 bit, lưu giữ trong một mạng từ A [7] về A [0] (A[0] là LSB)

+ Bước 3: Tính toán xem có tất cả bao nhiêu bit 0 và 1 xuất hiện trong mỗi byte, lưu tổng các giá trị này lần lượt là i 0 và i1.

+ Bước 4: Lặp lại từ bước 1 -> 4 cho đến khi kết thức toàn bộ văn bản EOF.

+ Bước 5 : Từ ảnh gốc , đọc giá trị RGB của mỗi pixel.

+ Bước 6: Đọc giá trị bit cuối cùng của mỗi pixels. Với ảnh RGB 24 bit thì ta sẽ đọc bit cuối cùng trong số 8 bit của màu xanh da trời.

+ Bước 7: Kiểm tra xem bit này có giá trị 0 hay l, sau đó tính tổng số lần xuất hiện các bit này trong S pixcels, lưu lần lượt vào hai biến c 0 và c 1.

+ Bước 8: Lặp lại từ bước 5 cho đến bước 7 [8\*S] lần. Đây là số pixel cần đọc để có thể giấu toàn bộ các byte thông điệp

+ Bước 9: Nếu [ (c0 > c1) và (i0 > i1) ] và [ (c1> c0) và (i1>i0)], đặt gí trị cho flag = 0, ngược lại đạt giá trị flag = 1.

+ Bước 10 : Ghi giá trị của flag vào phía bên trái của bit cuối cùng của pixel đầu tiên trong ảnh giấu.

+ Bước 11: Mở ảnh gốc ở chế độ đọc. Tạo một ảnh mang giống ảnh gốc ở chế độ ghi.

+ Bước 12: đọc Header của file gốc. Ghi thông tin header này lên ảnh mang. Từ ảnh gốc, đọc giá trị RGB của mỗi pixel.

+ Bước 13: đọc bit stream của dữ liệu. Nếu giá trị của cờ là 0 thì giữ nguyên giá trị bit của dữ liệu, sau đó ghi đè lên bit cuối cùng của màu xanh của pixel, ngược lại, nếu giá trị cờ là 1 thì đảo lại bit dữ liệu rời mới ghi lên pixcel (0 thành 1 hoặc 1 thành 0). Ghi pixel này vào ảnh stego.

+ Bước 14: Nếu toàn bộ các LSB đã được sửa đổi thành công, thì ghi nốt các bit còn lại của các pixel vào ảnh mang. Ngược lại, quay lại bước 13. [9]

**Thuật toán tách thủy vân**

+ [9] Bước 1: Mở ảnh Stego dưới chế độ đọc

+ Bước 2: Đọc bit liền kề bit cuối của pixel đầu tiên trong ảnh. Dựa trên giá trị của nó, đặt giá trị flag là 0 hoặc 1.

+ Bước 3: Đọc từng pixel của ảnh Stego.

+ Bước 4: Nếu flag là 0 thì đọc bit cuối cùng của mỗi pixel và đưa vào một mảng. Ngược lại nếu flag = 1 thì đảo bit rồi mới chuyển vào mảng.

+ Bước 5: Đọc mỗi 8 pixel theo cách trên, sau đó chuyển nội dung của mỗi 8 phần tử của mạng sang hệ thập phân, dây chính là giá trị ASCII của kí tự.

+ Bước 6: Nếu chưa gặp giá trị EOF thì in kí tự và quay lại bước 3.

**Ví dụ minh họa**

Giả sử thông tin cần nhúng là Hi. Trong bảng mã ASCII, H có mã là 72 và i có mã là 105:

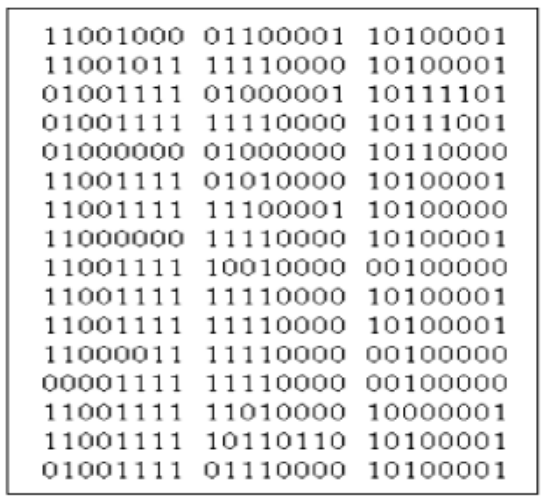
+ Chuyển sang hệ nhị phân ta có H = 01001000 và i = 01101001

+ Thông điệp Hi được mã hóa có dạng : 0100100001101001

+ Trong chuỗi trên có 6 bit 1 và 10 bit 0

+ Cần 16 pixel để lưu giữ 16 bit dữ liệu trên

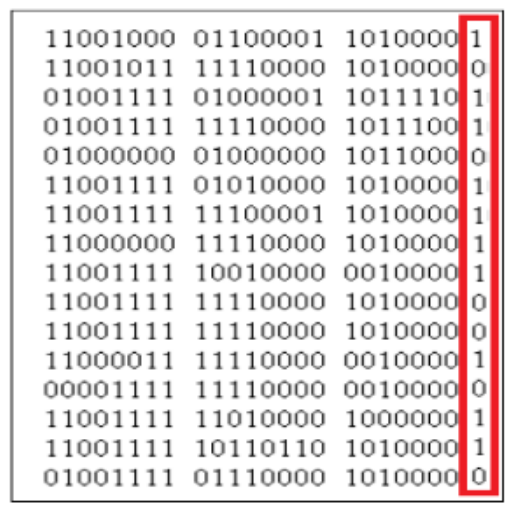
+ Giả sử có một bảng 16 pixel RGB có giá trị như sau:

****

+ Dựa theo thuật toán ta có : i0 = 10 i1 = 6 c0 = 5 c1 =11

+ Ở đây ta có i0> i1 nhưng c0 < c1 , do đó cần thực hiện phép đổi bit trên thông điệp thành 101101111001010 và đặt flag có giá trị bằng 1.

+ Sau quá trình nhúng thủy vân các điểm ảnh của ảnh Stego sẽ có dạng

****

Quá trình tách thủy vân tiến hành ngược lại :

+ Flag được đọc và nhận giá trị 1

+ Đọc 16 bits cuối của ảnh Stego ta nhận được chuỗi 1011011110010110

+ Do flag = 1 , ta cần thực hiện phép đảo bit để nhận được chuổi nguyên bản 0100100001101001. [9]

* + 1. Thuật toán PCT

[9] Việc nhúng thông tin vào ảnh nhị phân là một thách thức không nhỏ. Thuật toán giấu bít thông tin vào khối ảnh nhị phân (WL) được WU và LEE đề xuất. Tuy nhiên, mỗi khối giấu được không nhiều thông tin và khả năng bảo mật cũng không được tốt. Thuật toán CPT của Y. Chen , H. Pan, Y. Tseng cũng có tư tưởng giấu tin theo khối bít.

Theo thuật toán, ảnh được phân hoạch thành nhiều khối có cùng kích thước m x n . Với mỗi khối dữ liệu ảnh, có thể giấu được tối đa r bit thông tin, với r <= [log 2 (m\*n + 1)] bằng cách thay đổi không quá 2 bit trong khối dữ liệu ảnh.

So với thuật toán WL, thuật toán CPT có tỷ lệ giấu tin cao hơn nhiều, trong khi số bít cần thay đổi cũng rất ít. Ví dụ với khối 25 \* 25 thuật toán WL, ta chỉ giấu được 1 bit, nhung với thuật toán CPT có thể giấu tối đa là 8 bit.

Ngoài cách sử dụng một khóa K, thuật toán CPT còn sử dụng một ma trận trọng số nhằm giấu được một dãy nhiều bít vào mỗi khối, và ma trận trọng số này cũng là thành phần bí mật cùng với ma trận khóa K. Do vậy, độ an toàn, tính bảo mật của thuật toán CPT sẽ cao hơn.

**Dữ liệu vào:**

+ Ảnh nhị phân A dùng làm môi trường giấu tin. A được coi như một ma trận nhị phân, và được chia thành các ma trận con F cấp m \* n. Mỗi ma trận F là một khối bit được dùng làm môi trường giấu tin.

+ (b1,b2….br) là dãy r bít cần giấu vào trong mỗi khối ảnh kích thước m \* n và r phải thỏa mãn bất đẳng thức 2 ^r – 1 < = m\* n.

+ B là k \* r bit dữ liệu cần giấu, được tách thành k chuỗi r bit.

+ K là ma trận nhị phân cấp m \* n (KHóa)

+ W là ma trận trọng số cấp m \* n. Các phần tử của W cần thỏa mãn điều kiện

****

Số khả năng có thể lựa chọn K và W là khả năng (trong đó là tổ hợp m\*n phần tử). Vì vậy, với m và n, đủ lớn thì khả năng kẻ gian dò tìm ra được W là vô cùng khó nên thuật toán CPT có độ an toàn giấu tin rất cao.

Các ma trận K và W được sử dụng như khóa bí mật: người gửi sử dụng khóa K và ma trận trọng số W trong quá trình giấu tin và người nhận cần phải có khóa K, W để khôi phục lại thông tin đã giấu.

**Dữ liệu ra:**

Ảnh nhị phân A‟ chứa thông tin cần bảo mật. A‟ cùng gồm các ma trận con F‟ cấp m x n , trong đó mỗi F‟ giấu được r bít, và F‟ khác F tối đa hai phần tử.

- Các khái niệm cơ bản :

+ Ảnh nhị phân và ma trận nhị phân :

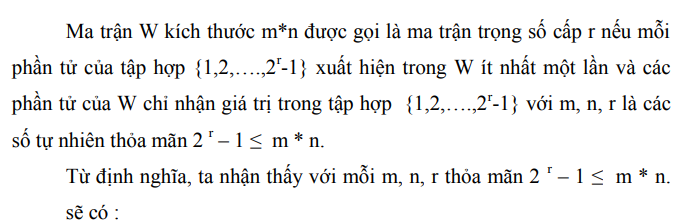
Trước hết ta quan tâm tới đối tượng chính là các ảnh nhị phân hay ảnh 1 bit màu. Đó là những bức ảnh mà mỗi điểm ảnh chỉ là điểm đen hoặc trắng, được quy định bởi một bit. Nếu bit mang giấ trị 0 thì điểm ảnh là đen, nếu là 1 thì điểm ảnh là trắng. Do đó để biểu diễn ảnh đen trắng ta có thể dùng ma trận nhị phân, là ma trận mà mỗi phần tử chỉ nhận một trong hai giá trị là 0 hoặc 1.

+ Khóa bí mật :

Là ma trận nhị phân có cùng kích thước với kích thước khối ảnh được dùng chung bởi người giấu tin và người tách thông tin.

+ Ma trận trọng số

Cũng là ma trận số cùng kích thước với khóa và được sử dụng bởi người giấu tin và người tách thông tin.



****

+ Ví dụ như với m = n = 4, r =2 thì ta có 5.356.925.280 khả năng lựa chọn. Con số này đủ lớn để làm giảm nguy cơ thông tin bị giải mã bởi những kẻ phá hoại.

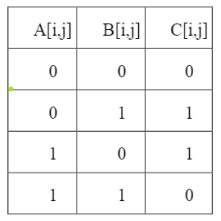
+ Phép đảo bít là một phép biến đổi trên các bit nhị phân. Đảo bit b tương đương với phép biển đổi thay b bởi 1- b, tức là nếu ban đầu b nhận giá trị 0 thì sau khi đảo nó sẽ nhận giá trị 1 và ngược lại, nếu ban đầu b có giá trị là 1 thì sau khi đảo nó sẽ có giá trị 0. [9]

• **Các phép toán trên ma trận :**

[9] Giả sử cho hai ma trận nhị phân A và B có cùng kích thước

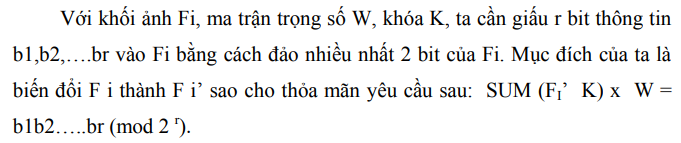
+ Phép cộng C = A + B

Trong đó C[i,j] = A[i,j] + B[i,j] ;

****

Trong đó C[i,j] = A [i,j] \* B[i,j] .

**• Thuật toán**

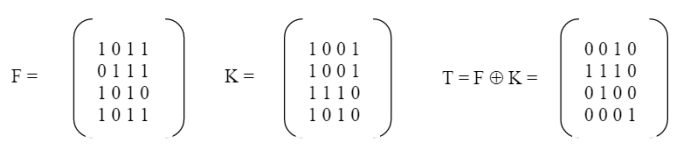


Thuật toán được thực hiện tuần tự cho từng khối F theo các bước sau:

+ Tính T = F + W

Trong đó: Phép + là phép toán XOR theo các vị trị tương ứng của hai ma trận nhị phân cùng bậc

Ví dụ :

****

+ Bước 2 : Tính S = SUM (T x W)

Trong đó X là phép nhân hai phần tử tương ứng của hai ma trận cùng bậc.

Phép SUM là dùng để tính tổng các phần tử của một ma trận .

Ví dụ:

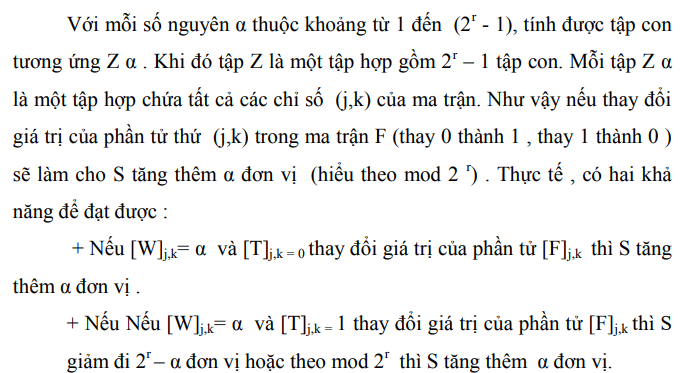
Giả sử :

****

Thì SUM (T x W) = 3+1+2+5+1+2 =14

+ Bước 3 : Xây dựng tập :

****



+ Bước 4:

Gọi F‟ là khối ảnh sau khi đã giấu r bit thông tin vào F ( F‟ khác F tối đa hai phần tử) và S‟ = (SUM (F‟ + K) \* W ). Khi đó sẽ thực hiện giấu tin bằng cách thay đổi các bit trong F để biến F thành F‟ sao cho đạt được bất biến :  (\*)

Trong đó b = (b1b2b3………br) . Ví dụ nếu r = 8 và (b1b2b3…..b8) =11111111 thì b = 255 .



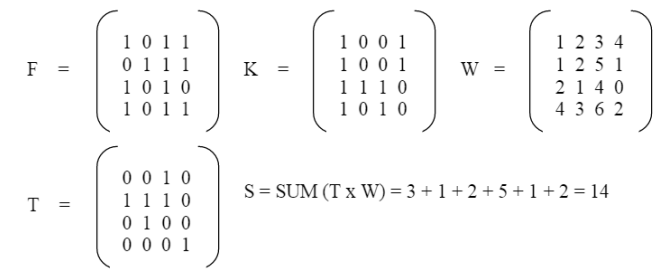
+ Trường hợp 1: . Do đó trong trường hợp này giấu được b vào F mà không cần biến đổi F tức là F‟ = F và S‟ = S

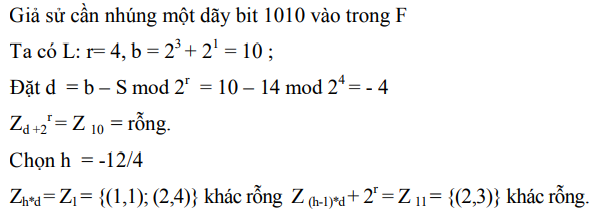
+ Trường hợp 2: Nếu d> 0 thì cần phải biến đổi F sao cho đạt được bất biến. Trong trường hợp này có 2 khả năng xảy ra.

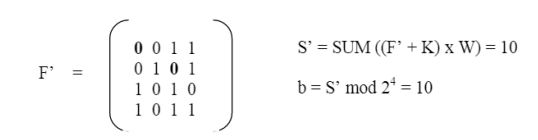
+ Trường hợp 3: Nếu d< 0 thì cần phải biến đổi F sao cho đạt được bất biến. Trong trường hợp này có hai khả năng xảy ra:

+ Bước 5

Khôi phục lại thông tin đã giấu. Khi người nhận được F‟ từ người gửi và biết ma trận mặt nạ K, ma trận trọng số W người nhạn chỉ cần tính S‟ =SUM (F‟ + K) x W => , từ đó xác định được dãy bít (b1b2…br) đã giấu trong F [9]

****



****

* 1. Tìm hiểu về các thuật toán thủy vân theo miền tần số (DCT; DWT)
     1. Hướng tiếp cận theo miền tần số

[2] Các thuật toán này sử dụng phương pháp biến đổi cosine rời rạc DCT 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 sau khi nhúng.

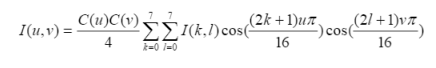
Kỹ thuật thủy vân trên miền tần số sử dụng các phương pháp biến đổi như Cosine rời rạc, biến đổi Fourier rời rạc, … để chuyển miền không gian ảnh sang miền tần số. Thủy vân sẽ được nhúng trong miền tần số của ảnh theo kỹ thuật trải phổ trong truyền thông, kỹ thuật này được đề xuất lần đầu tiên bởi Cox cùng cộng sự trong bài báo về “Thủy vân dựa trên trải phổ bảo vệ cho sự đa phương tiện” và đã được trích dẫn trong nhiều tài liệu. Đây là kỹ thuật phổ biến nhất với nhiều thuật toán được đề xuất và là phương pháp tốt giải quyết vấn đề về tính bền vững của thủy vân. [2]

* + 1. Miền tần số DCT

[9] Biến đổi cosin rời rạc DCT được đưa ra bởi Ahmed và các đồng nghiệp vào năm 1974. Từ đó đến nay, nó được sử dụng 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 DCT là được sử dụng nhiều. Nó được sử dụng chuẩn nén JPEG để mã hóa ảnh tĩnh và chuyển MPEG để mã hóa ảnh động.

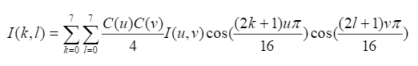
Biến đổi DCT hai chiều tổng quát là biến đổi trong khối hai chiều bất kỳ M x N. Sau đây trình bày công thức biến đổi DCT2 chiều trên khối kích thước 8 x 8 được sử dụng nhiều nhất hoặc 16 x16

Công thức biến đổi DCT thuận từ I (k,l) -> I (u,v)



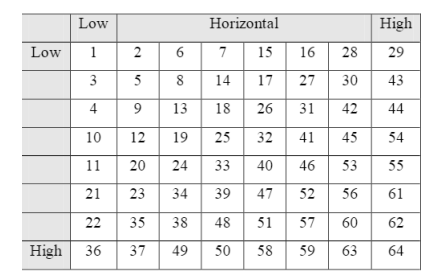
I (u,v) được gọi là hệ số DCT và là số thực.

Công thức biến đổi ngược IDCT từ I (u,v) -> I (k,l)



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 8 x 8. Áp dụng biến đổi DCT cho mỗi khối ta sẽ thu được khối 8 x 8 chứa các hệ số DCT. Gọi Cb (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 Cb(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

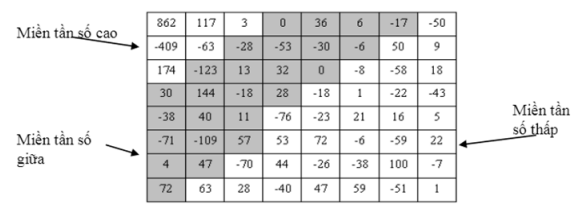


Hình 2. 6 Ví dụ bảng các hệ số DCT

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 đườngchéo) thì các hệ số biến đổi DCT tương ứng cũng lớn

Tóm lại, DCT làm giảm độ tương quan không gian của thông tin trong khối ảnh. Điều đó, cho phép biểu diễn thích hợp ở miền DCT do các hệ số DCT có xu hướng có phần dư thừa ít hơn. Hơn nữa, các hệ số DCT chứa thông tin về nội dung tần số không gian của thông tin trong khối. Nhờ các đặc tính tần số không gian của hệ thống nhìn của mắt người, các hệ số DCT có thể được mã hóa phù hợp, chỉ các hệ số DCT quan trọng nhất mới được mã hóa để truyền đi.

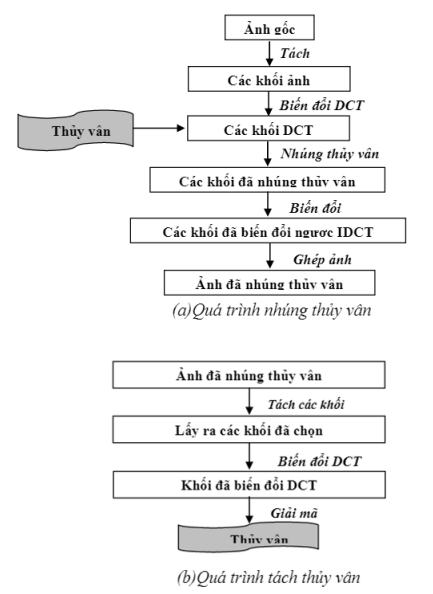
Khối hệ số DCT có thể chia thành ba miền ; miền tần số thấp, miền tần số giữa và miền tần số cao. Miền tần số thấp chứa các thông tin quan trọng ảnh hưởng đến tri giác. Các thông tin trong miền tần số cao thường không mang tính tri giác cao, khi nén JPEG thì thường loại bỏ thông tin trong miền này



Hình 2. 7 Phân chia 3 miền tần số thấp giữa, cao của phép biến đổi DCT.

Trong các thuật toán thủy vân, miền hệ số DCT tần số cao thường không được sử dụng do nó thường không bền vững với các phép xử lý ảnh hoặc nén ảnh JPEG. Miền tần số thấp cũng khó được sử dụng do một sự thay đổi dù nhỏ trong miền này cũng ảnh hưởng đến chất lượng tri giác của ảnh. Vì vậy, miền tần số ở giữa thường hay được sử dụng nhất và cũng cho kết quả tốt nhất

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. [9]

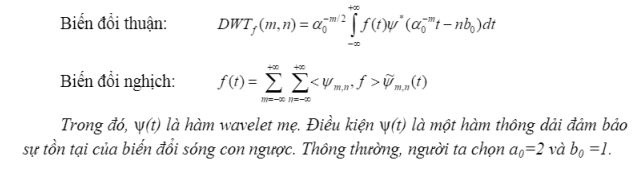


* + 1. Miền tần số DWT

**Phép biến đổi sóng rời rạc**

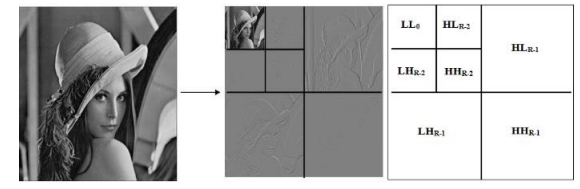
[9] Trong phép biến đổi này, Wavelets là các hàm được định nghĩa trong khoảng hữu hạn và có giá trị trung bình bằng 0. Ý tưởng cơ bản của phép biến đổi con sóng con là khai triển hàm f(t) bất kỳ như một xếp chồng của các con sóng con hay các hàm cơ sở. Các hàm cơ sở này có được từ một con sóng con nguyên mẫu được gọi là con sóng mẹ bằng cách lấy tỷ lệ và dịch.

Trong thực tế tính toán, biến đổi con sóng con rời rạc thuận và nghịch (DWT và IDWT) thường được thực hiện bởi phương trình sau :



Trong các thuật toán nghiên cứu có các thông số đáng chú ý sau đây: Tại một miền phân giải cấp 1 thì các hệ số của băng tần xấp xỉ (LL1 được mô tả tại hình 2.13) sẽ được gọi là v1 (x,y) . Các hệ số của băng tần HH1 sẽ được gọi là f 1,1 (x,y)của LH1 sẽ là f 2,1 (x,y ) và của HL1 là f 3,1 (x,y)

Với vài thuật toán các hệ số này sẽ được thăm theo đường zig zag. Khi đó, ta sẽ gọi các hệ số này lần lượt theo thứ tự như sau : v1 (i), f 1,1, (i) , f 2,1 (i) , f 3,1 (i)



Hình 2. 8 Biến đổi Wavelet và cấu trúc dải thông

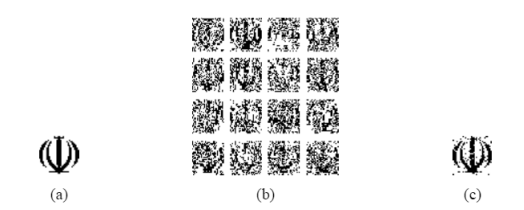
Trong một số trường hợp, sơ đồ dùng biến đổi sóng con đã tỏ ra ưu thế so với biến đổi Fourier rời rạc DFT hay biến đổi cosin rời rạc DCT. Do đặc tính đa phân giải, sơ đồ mã hóa Wavelets đặc biệt thích hợp cho các ứng dụng mà tính vô hướng và suy biến đóng vai trò quan trọng. Minh chứng cho điều này là biến đổi sóng con đã được dùng như một tiêu chuẩn trong nén JPEG2000. Ngoài ra, tính đa phân giải của Wavelets còn hữu ích trong việc phân phối thông điệp vào đối tượng bao phủ trong khi vẫn đảm bảo tính bền vững và chất lượng hiện thị. Do đó, lược đồ thủy vân sử dụng DWT vẫn đảm bảo được tính bền vững của thủy vân sau khi nén có mất mát thông tin theo chuẩn nén JPEG2000.

Tổng quát, biến đổi sóng con thực hiện triển khai tần số không gian đa tỷ lệ của một ảnh. Khai triển này tạo ra các hệ số xấp xỉ và các hệ số chi tiết ngang, dọc và chéo. Quá trình khai triển lại tiếp tục với các hệ số xấp xỉ ở mức phân tích cao hơn. Các hệ số xấp xỉ sau cùng chứa thông tin về băng tần thấp nhất trong khi các hệ số chi tiết chứa thông tin về băng tần cao hơn. [9]

**Lược đồ thủy vân sử dụng biến đổi DWT**

[9] Ngày nay, có nhiều thuật toán thủy vân sử dụng biến đổi sóng con và các kỹ thuật lượng tử hóa, thủy vân sử dụng miền biến đổi wavelet có lợi thế làm cho các thủy vân mạnh mẽ hơn chống lại được nhiều dạng tấn công như thay đổi thành phần tần số cao của hình ảnh, nén , lọc thông thấp qua, tuy nhiên nó không thể chống lại cuộc tấn công như cắt ảnh hay phá hủy một thành phần hình ảnh chứa thủy vân.

Hầu hết các phương pháp thủy vân dựa trên biến đổi wavelet chia dải thông con thành các khối nhỏ và sau đó nhúng từng bit logo thủy vân nên chúng hoàn toàn trong mỗi khối con, tức là mỗi bit của thủy vân được lưu trữ trong một hệ số của một khối con và kích thước của khối con phải lớn hơn kích thước của hình ảnh thủy vân. Khi một vùng của hình ảnh phủ bị phá hủy. Thủy vân nguyên vẹn có thể được trích xuất thủy vân hoàn chỉnh. Ví dụ như hình 2.14 cho thấy kết quả của thủy vân kết quả được trích xuất từ một ảnh đã bị nén (với thuật toán JPEG2000) và bị cắt đi một phần. [9]



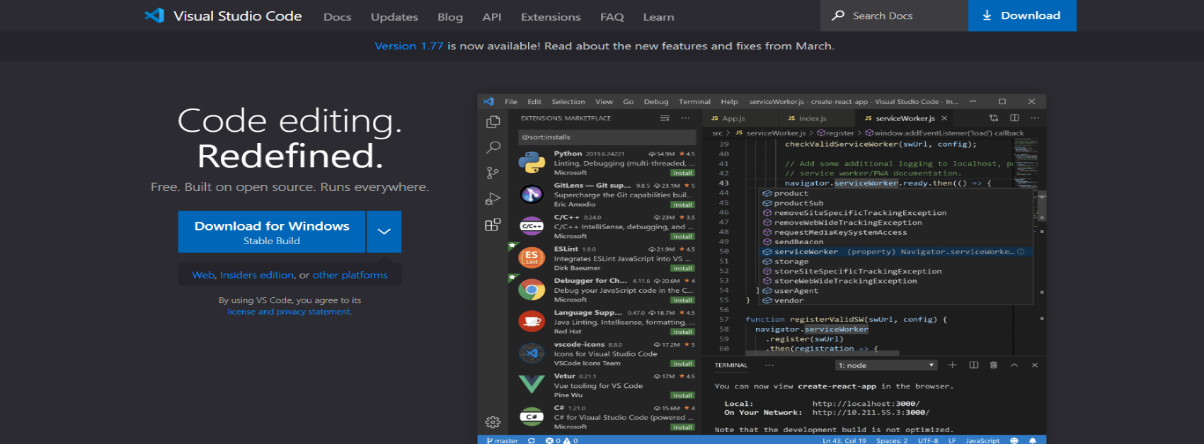
Hình a Thủy vân gốc, b thủy vân tách được từ các khối, c thủy vân kết hợp

1. CÀI ĐẶT THỬ NGHIỆM
   1. Phát biểu bài toán

Digital watermark là một loại dấu ấn số được nhúng vào tệp số như hình ảnh, video hoặc tài liệu văn bản để định danh và bảo vệ quyền sở hữu trí tuệ. Nó được sử dụng để xác định nguồn gốc của tài liệu, ngăn chặn việc sao chép trái phép hoặc sửa đổi nội dung của tài liệu. Digital watermark có thể được hiển thị hoặc ẩn, phụ thuộc vào mục đích sử dụng. Nó cũng có thể được mã hóa để ngăn chặn việc loại bỏ hoặc thay đổi nó bởi người khác.

* 1. Ứng dụng chương trình
* Chương trình được dùng để nhúng ảnh vào một ảnh gốc nhằm bảo vệ bản quyền ảnh số của tác giả.
  1. Hướng dẫn sử dụng
* **Môi trường chạy**
* Chương trình được cài đặt trên công cụ Visual Studio Code và ngôn ngữ Python
* Có thể download chương trình tại đỉa chỉ <https://www.python.org/downloads/>

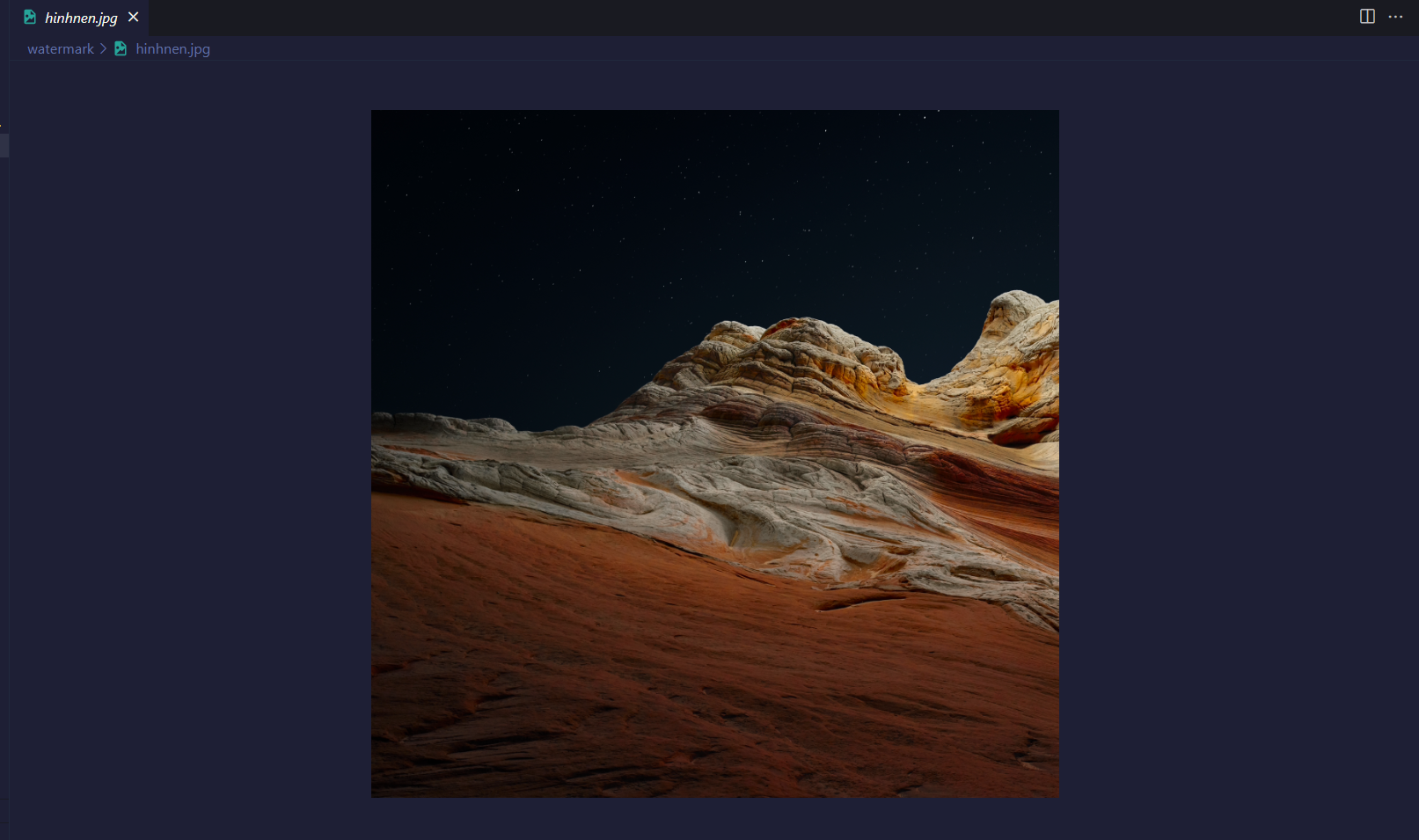




* **Chạy chương trình**
* Cài đặt extension Python trên Visual studio code



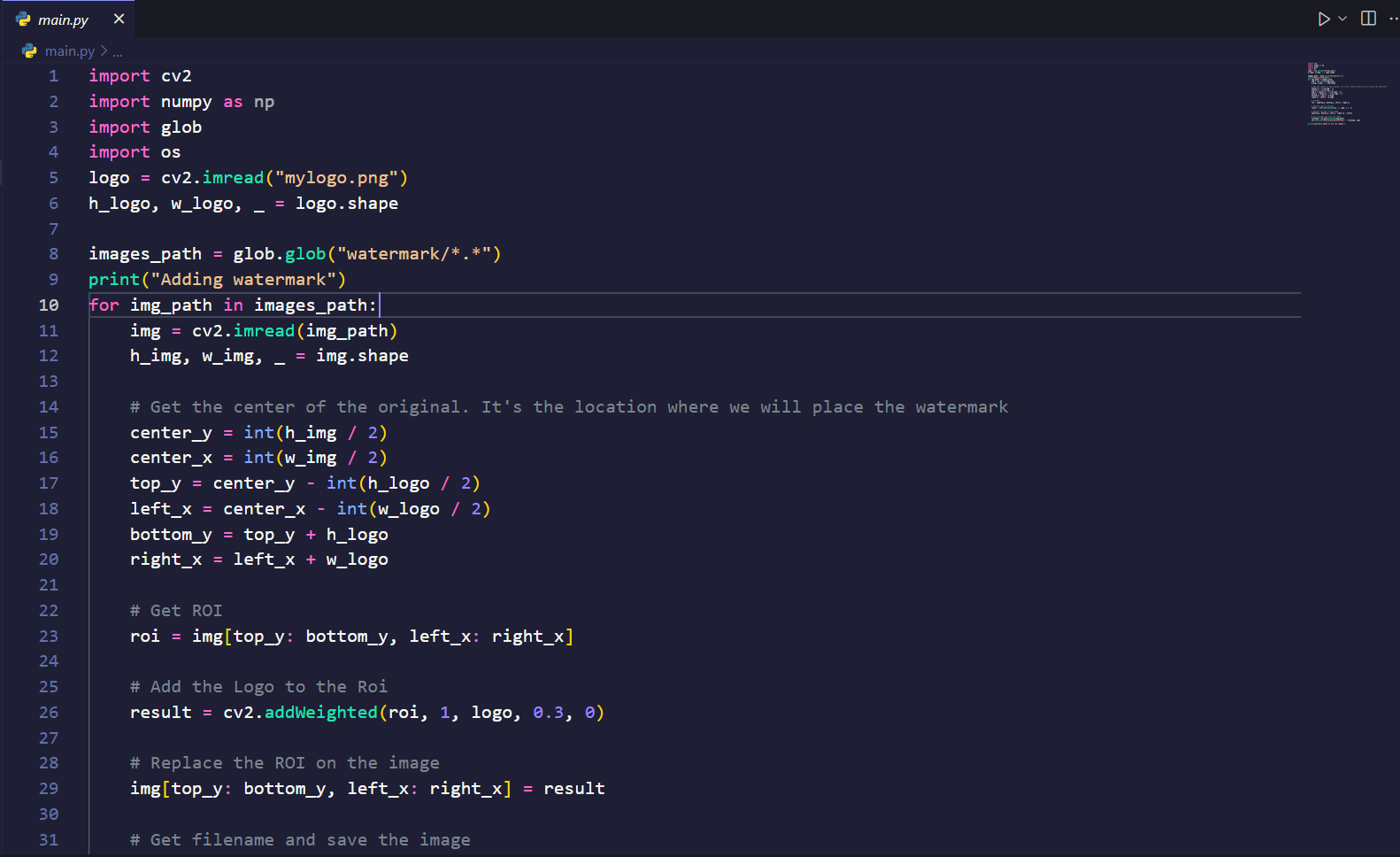
* Thêm hình gốc vào thư mục watermark



* Thêm chữ ký mylogo.png



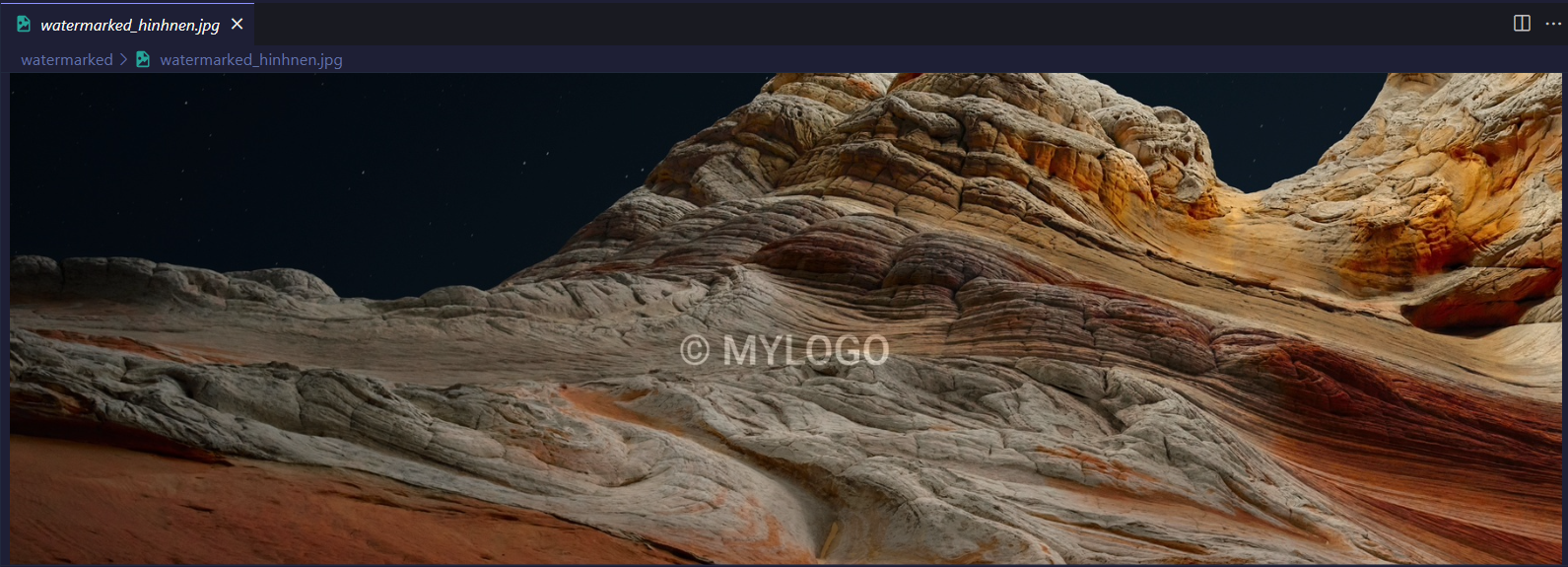
* Sau đó tạo File main.py để thực hiện chức năng thêm chữ ký vào bức hình



* Mở New Termianl gõ dòng lệnh: python main.py để tiến hành nhúng chữ ký vào ảnh gốc



* Hình sau khi được nhúng thêm vào thư mục waternmarked



1. KẾT LUẬN

Sau một thời gian nghiên cứu và tìm hiểu cùng với sự nỗ lực của bản thân và sự hướng dẫn tận tình của thầy giáo hướng dẫn, nhóm em đã hoàn thành đề tài của mình. Nội dung chủ yếu của đề tài là nghiên cứu về hệ thủy vân số, các hướng ứng dụng của thủy vân số chủ yếu là ứng dụng trong bảo vệ bản quyền ảnh số. Từ đó, xây dựng chương trình thử nghiệm cài đặt một số thuật toán thủy vân nhằm ứng dụng xác thực thông tin và bảo vệ bản quyền cho dữ liệu ảnh số

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

TÀI LIỆU THAM KHẢO

[1] Giới thiệu về kỹ thuật giấu thông tin. (n.d.). pp. https://antoanthongtin.gov.vn/giai-phap-khac/gioi-thieu-ve-ky-thuat-giau-thong-tin-100319.

[2] Hòa, P. C. (n.d.). *Tìm hiểu phương pháp thủy vân số thuận nghịch và xây dựng ứng dụng, Trường Đại học Công nghệ.* 2012.

[3] Hoan, N. Q. (n.d.). *Giáo trình xử lý ảnh, Học viện bưu chính viễn thông.* 2006.

[4] Hương, L. T. (n.d.). *Luận án tiến sĩ, Nghiên cứu và phát triển kỹ thuật thủy vân cơ sở dữ liệu quan hệ, Viện Công nghệ thông tin.* 2014.

[5] Lê Tiến Thường, N. T. (n.d.). *“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.* 2004.

[6] Nguyễn Xuân Huy, T. Q. (n.d.). *Một thuật toán thủy vân ảnh trên miền DCT, Hội thảo quốc gia Các vấn đề chọn lọc Công nghệ thông tin và Truyền thông.* 2002.

[7] Nhật, N. M. (n.d.). *Luận văn Kỹ thuật thủy vân số, Đại học Duy Tân – Đà Nẵng.*

[8] Tiến, T. N. (n.d.). *Bài giảng An toàn dữ liệu.* 2008.

[9] UYÊN, T. T. (n.d.). *Hệ thống thủy vân số và ứng dụng thủy vân số, Luận văn thạc sĩ công nghệ thông tin.* Hà Nội: 2017.