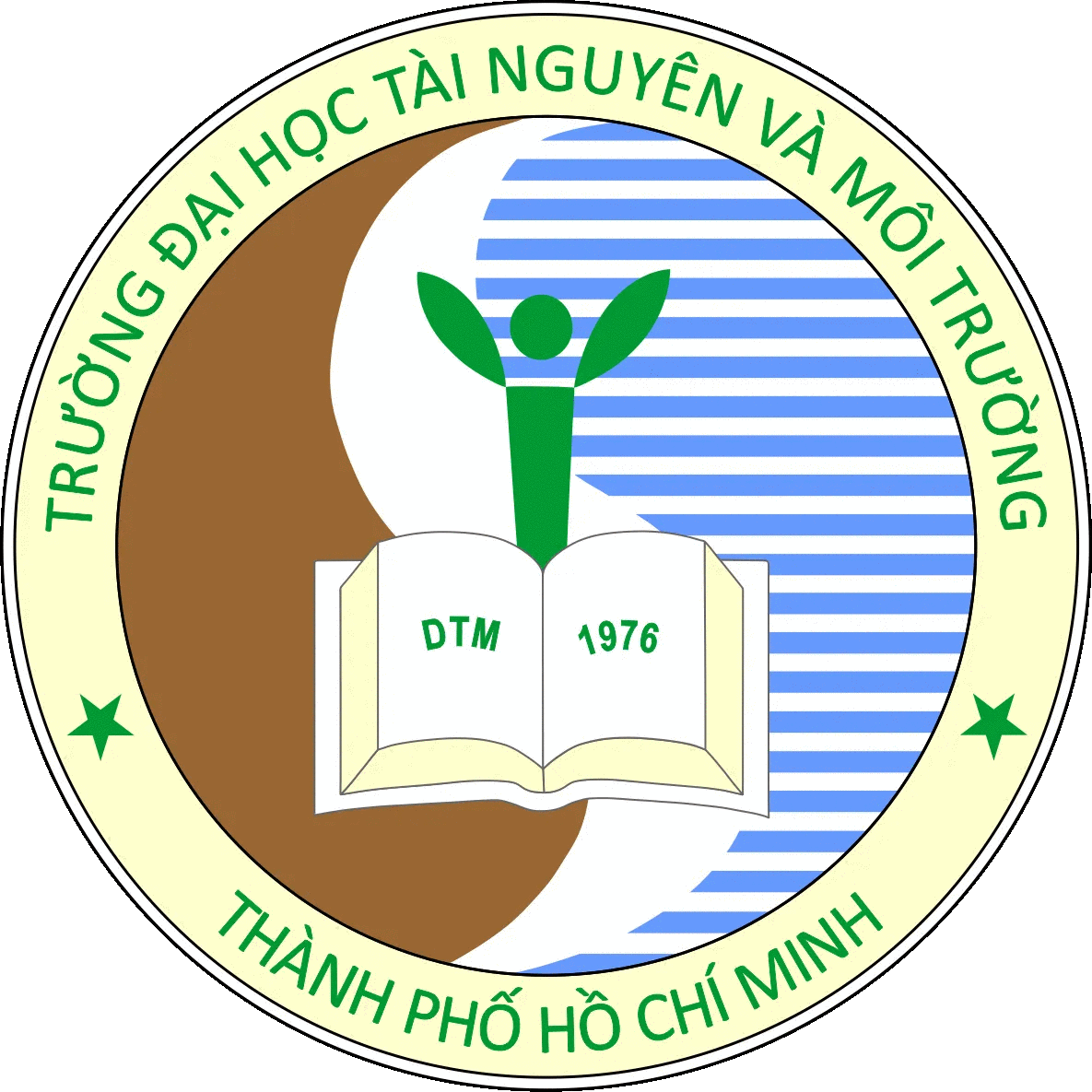
**BỘ TÀI NGUYÊN VÀ MÔI TRƯỜNG**

**TRƯỜNG ĐH TÀI NGUYÊN VÀ MÔI TRƯỜNG TP.HCM**

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

****

**BÁO CÁO MÔN HỌC**

**BẢO MẬT MẠNG MÁY TÍNH VÀ HỆ THỐNG**

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

Giáo viên hướng dẫn: ThS. Phạm Trọng Huynh

Sinh viên thực hiện: Lê Chí Bảo 0850080005

Phạm Thị Thảo Huyền 0850080024

Lê Thị Bảo Yến 0850080057

Lớp: 08\_ĐH\_CNPM

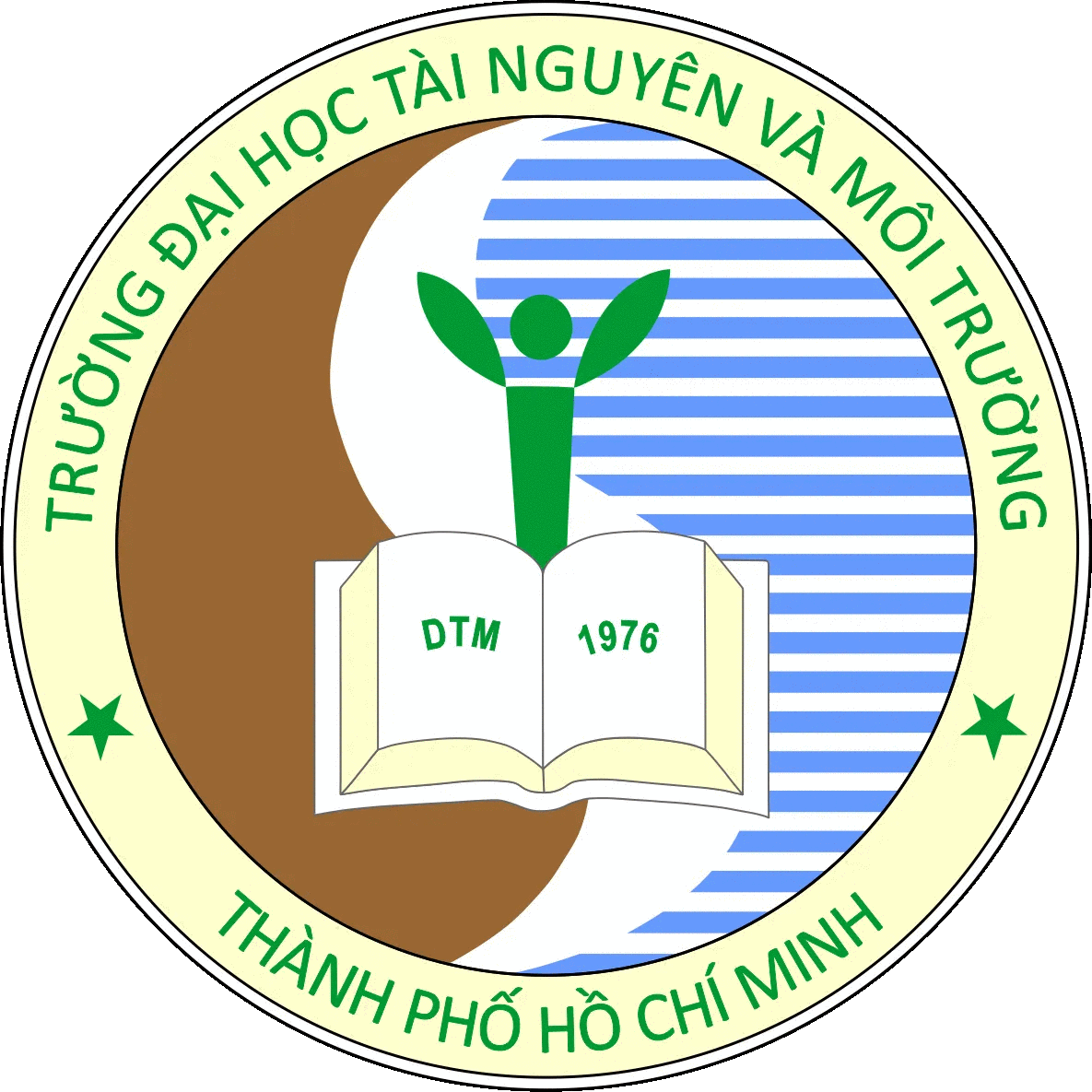
Khóa: 08

***TP. Hồ Chí Minh, Tháng 5 Năm 2023***

**BỘ TÀI NGUYÊN VÀ MÔI TRƯỜNG**

**TRƯỜNG ĐH TÀI NGUYÊN VÀ MÔI TRƯỜNG TP.HCM**

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

****

**BÁO CÁO MÔN HỌC**

**BẢO MẬT MẠNG MÁY TÍNH VÀ HỆ THỐNG**

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

Giáo viên hướng dẫn: ThS. Phạm Trọng Huynh

Sinh viên thực hiện: Lê Chí Bảo 0850080005

Phạm Thị Thảo Huyền 0850080024

Lê Thị Bảo Yến 0850080057

Lớp: 08\_ĐH\_CNPM

Khóa: 08

***TP. Hồ Chí Minh, Tháng 5 Năm 2023***

**LỜI CẢM ƠN**

Trong suốt quá trình học tập và hoàn thành đồ án, chúng em đã nhận được sự hướng dẫn, giúp đỡ, các kiến thức chuyên môn của thầy cô và các bạn đã giúp chúng em có thể hoàn thành đồ án một cách tốt nhất.

Với lòng kính trọng và biết ơn sâu sắc chúng em xin được bày tỏ lời cảm ơn chân thành tới:

Ban giám hiệu trường Đại Học Tài Nguyên Và Môi Trường Thành Phố Hồ Chí Minh, Khoa Hệ Thống Thông Tin Và Viễn Thám, cùng các quý thầy cô giáo đã truyền đạt vốn tri thức và nhiệt huyết rất quý báu cho chúng em trong suốt thời gian học tập tại trường.

ThS. Phạm Trọng Huynh đã tận tâm hướng dẫn chúng em qua từng buổi học trên lớp và những kiến thức từ thực tiễn đã góp phần giúp em hoàn thành đồ án này

ThS. Phạm Trọng Huynh đã nhiệt tình hướng dẫn em trong suốt quá trình thực hiện tiểu luận từ lúc sơ khai cho đến khi hoàn thành.

Trong quá trình thiết kế và thực hiện không tránh khỏi sai sót, mong thầy thông cảm và có những góp ý kiến quý báu nhằm hoàn thiện hơn cho sản phảm.

Em xin chân thành cảm ơn!

**NHẬN XÉT**

..………………………………………………………………………………………….……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

**MỤC LỤC**

[**CHƯƠNG I. GIỚI THIỆU ĐỒ ÁN** 1](#_Toc135807984)

[**1.1.** **Lý do chọn đề tài** 1](#_Toc135807985)

[**1.2.** **Mục đích đề tài** 1](#_Toc135807986)

[**1.3.** **Đối tượng và phạm vi nghiên cứu** 1](#_Toc135807987)

[**1.4.** **Phương pháp thực hiện** 2](#_Toc135807988)

[**1.5.** **Kết quả đạt được** 2](#_Toc135807989)

[**CHƯƠNG II. PHÂN TÍCH ĐỀ TÀI** 3](#_Toc135807990)

[**2.1.** **Kỹ thuật che giấu tập tin** 3](#_Toc135807991)

[**2.1.1.** **Khái niệm giấu tin** 3](#_Toc135807992)

[**2.1.2.** **Phân loại các kỹ thuật giấu tin và ứng dụng** 3](#_Toc135807993)

[**2.1.3.** **Mô hình kỹ thuật giấu ảnh** 4](#_Toc135807994)

[**2.2.** **Phương pháp và mô hình thủy vân số** 7](#_Toc135807995)

[**2.2.1.** **Một số vấn đề cơ bản về thủy phân** 7](#_Toc135807996)

[**2.2.1.1.** **Khái niệm về thủy vân** 7](#_Toc135807997)

[**2.2.1.2.** **Phân loại thủy vân** 7](#_Toc135807998)

[**2.2.1.3.** **Các ứng dụng của thuỷ vân với ảnh số** 8](#_Toc135807999)

[**2.2.1.4.** **Một số tính chất của sơ đồ thủy vân** 9](#_Toc135808000)

[**2.2.2.** **Mô hình thủy vân số** 10](#_Toc135808001)

[**2.2.2.1.** **Tạo thủy vân số** 10](#_Toc135808002)

[**2.2.2.2.** **Quy trình nhúng thủy vân** 11](#_Toc135808003)

[**2.2.2.3.** **Tách thủy vân** 12](#_Toc135808004)

[**2.2.2.4.** **Phát hiện thủy vân và giải mã** 12](#_Toc135808005)

[**2.3.** **Thuật toán thủy văn theo miền không gian ảnh** 13](#_Toc135808006)

[**2.3.1.** **Thuật toán SW** 14](#_Toc135808007)

[**2.3.2.** **Thuật toán WU-LEE** 16](#_Toc135808008)

[**2.3.3.** **Thuật toán LBS** 19](#_Toc135808009)

[**2.3.3.** **Thuật toán PCT** 23](#_Toc135808010)

[**2.4.** **Thuật toán thủy văn theo miền tần số** 29](#_Toc135808011)

[**2.4.1.** **Biến đổi cosine rời rạc (DCT)** 29](#_Toc135808012)

[**2.4.2.** **Thuật toán thủy văn dựa trên miền DWT** 38](#_Toc135808013)

[**CHƯƠNG III. THIẾT KẾ VÀ HIỆN THỰC** 44](#_Toc135808014)

[**3.1.** **Mô tả chức năng hệ thống** 44](#_Toc135808015)

[**3.2.** **Cài đặt và thực nghiệm** 44](#_Toc135808016)

[**3.2.1.** **Giao diện chính phần mềm thử nghiệm** 44](#_Toc135808017)

[**3.2.2.** **Giao diện nhúng ảnh thủy vân bằng phương pháp LSB** 45](#_Toc135808018)

[**3.2.3.** **Giao diện trích xuất thủy vân bằng phương pháp LSB** 45](#_Toc135808019)

[**3.2.4.** **Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DCT** 46](#_Toc135808020)

[**3.2.5.** **Giao diện trích xuất thủy vân bằng phương pháp biến đổi DCT** 46](#_Toc135808021)

[**3.2.6.** **Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DWT** 46](#_Toc135808022)

[**3.2.7.** **Giao diện trích xuất thủy vân bằng phương pháp biến đổi DWT** 47](#_Toc135808023)

[**CHƯƠNG IV. KẾT LUẬN** 48](#_Toc135808024)

[**TÀI LIỆU THAM KHẢO** 49](#_Toc135808025)

**DANH MỤC HÌNH ẢNH**

[Hình 1. Hình ảnh phân loại các kỹ thuật giấu tin 3](#_Toc135808026)

[Hình 2. Mô hình thuật toán nhúng tin 4](#_Toc135808027)

[Hình 3. Mô hình trích tin 5](#_Toc135808028)

[Hình 4. Phân loại thủy vân theo mục đích ứng dụng 7](#_Toc135808029)

[Hình 5. Mô hình hệ thống thủy vân số 10](#_Toc135808030)

[Hình 6. Quy trình nhúng ảnh thủy vân 11](#_Toc135808031)

[Hình 7. Minh họa thuật toán WS: nhúng bit 1 bào khối ảnh B 15](#_Toc135808032)

[Hình 8. Minh họa chọn điểm ảnh giấu tin vào những khối ảnh màu 16](#_Toc135808033)

[Hình 9. Minh họa thuật toán WU\_LEE nhúng đoạn bit 01 18](#_Toc135808034)

[Hình 10. Hình ảnh bảng hệ số DCT 30](#_Toc135808035)

[Hình 11. Bảng phân chia 3 miền tần số thấp giữa, cao của phép biến đổi DCT 31](#_Toc135808036)

[Hình 12. Quy trình tách và nhúng thủy vân theo kỹ thuật vân trên miền DCT 32](#_Toc135808037)

[Hình 13. Biến đổi Wavelet và cấu trúc giải thông 38](#_Toc135808038)

[Hình 14. (a)Thủy vân gốc (b)Thủy vân được trích xuất từ các khối (c)Thủy vân kết hợp. 40](#_Toc135808039)

[Hình 15. Dải thông LL2 được chia các khối nhỏ hơn 40](#_Toc135808040)

[Hình 16. Hình (a)- Ảnh gốc (b)-Ảnh đã thủy vân với Q=35 41](#_Toc135808041)

[Hình 17. Giao diện chính phần mềm thử nghiệm 44](#_Toc135808042)

[Hình 18. Giao diện nhúng ảnh thủy vân bằng phương pháp LSB 45](#_Toc135808043)

[Hình 19. Giao diện trích xuất ảnh thủy vân bằng phương pháp LSB 45](#_Toc135808044)

[Hình 20. Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DCT 46](#_Toc135808045)

[Hình 21. Giao diện trích xuất ảnh thủy vân bằng phương pháp biến đổi DCT 46](#_Toc135808046)

[Hình 22. Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DWT 47](#_Toc135808047)

[Hình 23. Giao diện trích xuất ảnh thủy vân bằng phương pháp biến đổi DWT 47](#_Toc135808048)

**DANH MỤC VIẾT TẮT**

|  |  |
| --- | --- |
| **Chữ viết tắt** | **Ý nghĩa** |
| **SW** | Thuật toán thủy vân đơn giản (Simple Watermarking) |
| **WU-LEE** | Thuật toán thủy vân được đặt theo tên của 2 tác giả: M.Y.Wu và J.H.LEE |
| **PCT** | Thuật toán thủy vân đặt theo tên của 3 tác giả: Hsiang – Kuang Pan, Yu-Yuan Chen và Yu-chee Treng |
| **LSB** | Least Significant Bit |
| **DCT** | Biến đổi Cosine rời rạc (Discrete Cosine Tranform) |
| **DWT** | Biến đổi sóng rời rạc (Discrete Wavelet Tranform) |

# **CHƯƠNG I. GIỚI THIỆU ĐỒ ÁN**

* 1. **Lý do chọn đề tài**

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. 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 nhóm chúng em quyết định lựa chọn đề tài là: **“Nghiên Cứu Kỹ Thuật Thủy Vân Số Và Xây Dựng Ứng Dụng Bải Vệ Bản Quyền Ảnh Số”**.

* 1. **Mục đích đề tài**

Mục đích của đồ án 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.

* 1. **Đối tượng và phạm vi nghiên cứu**

Đồ án tập trung nghiên cứu các kỹ thuật thủy vân trên ảnh số. Ứng dụng mà đồ án 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. **Phương pháp thực hiện**

Phương pháp thực hiện đề tài là nghiên cứu các vấn đề liên quan đến giấu tin, tập trung nghiên cứu tiến hành xây dựng chương trình và cài đặt chương trình thử nghiệm.

* 1. **Kết quả đạt được**

Đồ án đã hệ thống lại các kiến thức cơ bản về thủy vân số, nghiên cứu một số thuật toán trên miền không gian và miền tần số. Đồng thời cài đặt thành công thuật toán thủy vân trên miền tần số và miền không gian nhằm ứng dụng xác thực bản quyền ảnh số của tác giả.

# **CHƯƠNG II. PHÂN TÍCH ĐỀ TÀI**

* 1. **Kỹ thuật che giấu tập tin**
     1. **Khái niệm giấu tin**

Giấu tin là một kỹ thuật giấu một lượng thông tin nào đó vào một đối tượng dữ liệu khác nhằm hai mục đích đó là: Một là bảo vệ cho chính đối tượng được giấu tin bên trong. Đây chính là thủy vân số, đây là lĩnh vực rất đa dạng, có nhiều mục đích và đang được quan tâm, nghiên cứu rất nhiều; tính ứng dụng của nó trong hiện tại rất lớn và đã có nhiều kỹ thuật được đề xuất.

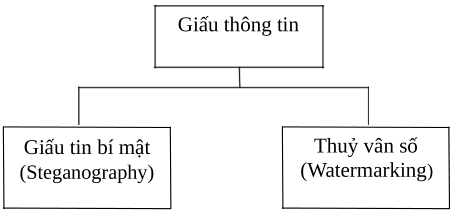
Hai là bảo mật cho dữ liệu được đem giấu. Chính là giấu tin mật, tập trung vào các kỹ thuật giấu tin sao cho người khác rất vất vả, khó khăn mới phát hiện được đối tượng có chứa thông tin mật bên trong.

* + 1. **Phân loại các kỹ thuật giấu tin và ứng dụng**

Dựa theo mục đích sử dụng, ta có chia các lược đồ giấu tin thành hai nhóm chính giấu tin và thủy vân:

* Thủy vân số: Là nhúng một lượng thông tin có ích vào các sản phẩm số. Có thể chia thành hai loại thủy vân như sau: Thủy vân bền vững và thủy vân dễ vỡ. Thủy vân bền vững nghiên cứu đến việc nhúng những mẩu tin với yêu cầu độ bền vững cao của thông tin được giấu trước những biến đổi của môi trường dữ liệu. Còn thủy vân dễ vỡ quan tâm đến thông tin giấu sẽ bị sai lệch nếu có bất kỳ sự thay đổi nào trên dữ liệu chứa thông tin giấu.
* Giấu tin mật: Với kỹ thuật giấu tin mật này luôn quan tâm đến tính che giấu thông tin, với mục đích nhằm làm sao cho việc phát hiện được đối tượng có chứa thông tin mật ở bên trong hay không; nếu phát hiện có giấu tin thì việc giải mã thông tin mật cũng sẽ mất rất nhiều thời gian và rất khó thực hiện được. Các kỹ thuật giấu tin cũng rất quan tâm đến lượng tin có thể giấu được.

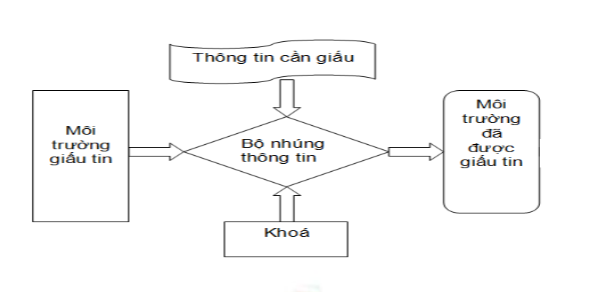
Như vậy, kỹ thuật giấu tin có thể được phân loại như hình sau:



Hình 1. Hình ảnh phân loại các kỹ thuật giấu tin

* + 1. **Mô hình kỹ thuật giấu ảnh**

Giống như các lược đồ mã hóa thông tin, một kỹ thuật giấu tin gồm thuật toán nhúng và thuật toán trích tin. Sơ đồ của thuật toán nhúng tin tổng quát có mô hình như sau:

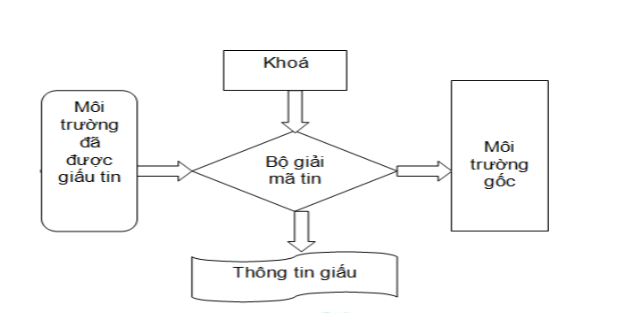


Hình 2. Mô hình thuật toán nhúng tin

Trong Hình trên, môi trường giấu tin là những dữ liệu dùng để chứa thông tin mật, dữ liệu môi trường thường là những dữ liệu được dùng phổ biến trên Internet như: tệp ảnh, tệp âm thanh, tệp video, tệp text...Thông tin cần giấu là một lượng thông tin mang ý nghĩa và mục đích nào đó tùy thuộc vào mục đích và yêu cầu của người sử dụng (tin mật). Tin mật được nhúng vào trong môi trường chứa thông tin nhờ một bộ nhúng thông tin. Trong quá trình nhúng tin, thuật toán có thể sử dụng hệ thống khóa làm tăng tính an toàn cho hệ thống. Sau khi nhúng tin mật vào dữ liệu môi trường ta nhận được dữ liệu môi trường có chưa tin. Dữ liệu này được truyền tải trên các môi trường truyền thông khác nhau.

Khi nhận được dữ liệu có chưa tin, người nhận thực hiện thuật toán trích tin tương ứng để trích rút tin mật từ dữ liệu môi trường. Trong một số trường hợp, người dùng có thể khôi phục lại dữ liệu môi trường từ dữ liệu chứa tin.

Thuật toán trích tin có mô hình thực hiện như:



Hình 3. Mô hình trích tin

Trong hình trên đã chỉ ra các nhiệm vụ của quá trình giải mã thông tin đã được giấu. Quá trình giải mã này phải được thực hiện thông qua một bộ giải mã tương ứng với bộ nhúng thông tin và kết hợp với khóa để giải mã tin. Khóa để giải mã tin này có thể giống hay khác với khóa đã nhúng tin. Kết quả thu được bao gồm môi trường gốc và thông tin đã được che giấu. Tùy theo các trường hợp cụ thể, thông tin tách được ra có thể phải cần xử lý, kiểm định và so sánh với thông tin đem giấu ban đầu. Thông qua dữ liệu được tách ra từ môi trường chứa thông tin

giấu, người ta có thể biết được trong quá trình truyền tải, phân phát dữ liệu có bị xâm phạm, tấn công hay không.

Đối với các hệ thống giấu thông tin mật này rất quan tâm đến tính an toàn và bảo mật thông tin của dữ liệu cần giấu. Hệ thống giấu tin mật có độ bảo mật cao nếu có độ phức tạp của các thuật toán thám mã khó có thể thực hiện được trên máy tính. Tuy nhiên, cũng có các hệ thống chỉ quan tâm đến số lượng thông tin có thể được che giấu, hay quan tâm đến sự ảnh hưởng của thông tin mật đến các môi trường chứa dữ liệu.

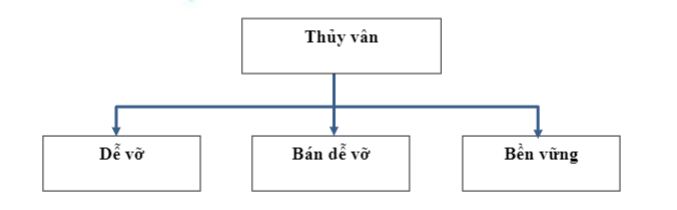
Có nhiều kỹ thuật che giấu tập tin khác nhau, sau đây là một số trong số đó:

1. Đổi tên tập tin: Kỹ thuật đơn giản nhất để che giấu tập tin là đổi tên của nó. Bằng cách đổi tên tập tin thành một cái tên khác, bạn có thể làm cho nó khó tìm thấy hơn đối với người dùng khác.
2. Ẩn tập tin: Trên hệ điều hành Windows, bạn có thể ẩn tập tin bằng cách chọn thuộc tính "Hidden" (ẩn) cho tập tin đó. Tập tin được ẩn sẽ không hiển thị trong thư mục mặc định, nhưng nó vẫn có thể được tìm thấy bằng cách sử dụng tìm kiếm hoặc hiển thị các tập tin ẩn.
3. Nén tập tin: Bằng cách sử dụng phần mềm nén tập tin như WinZip hoặc WinRAR, bạn có thể nén tập tin vào một file nén. File nén này có thể được đặt tên một cách khác biệt và được lưu trữ trong một thư mục bình thường như bất kỳ tập tin nào khác.
4. Sử dụng mã hóa: Bạn có thể mã hóa tập tin của mình bằng cách sử dụng phần mềm mã hóa như VeraCrypt. Phần mềm này sẽ tạo ra một đĩa ảo có thể được bảo vệ bằng mật khẩu. Bạn có thể lưu trữ các tập tin của mình trên đĩa ảo này và chúng sẽ được mã hóa và bảo vệ bằng mật khẩu.
5. Sử dụng các dịch vụ lưu trữ trực tuyến: Bạn có thể lưu trữ các tập tin của mình trên các dịch vụ lưu trữ trực tuyến như Dropbox, Google Drive hoặc OneDrive. Điều này có thể giúp che giấu các tập tin của bạn bằng cách lưu trữ chúng trên đám mây, trong khi đồng thời cung cấp tính di động và truy cập từ xa. Tuy nhiên, bạn cần phải đảm bảo rằng tài khoản của mình được bảo mật và chỉ có bạn mới có thể truy cập vào các tập tin của mình.
   1. **Phương pháp và mô hình thủy vân số**
      1. **Một số vấn đề cơ bản về thủy phân**
         1. **Khái niệm về thủy vân**

Thuỷ vân (Watermarking) là kỹ thuật nhúng thông tin vào dữ liệu môi trường (đa phương tiện) như: tệp ảnh, tệp video, tệp âm thanh,... nhằm bảo vệ dữ liệu đa phương tiện trong quá trình trao đổi không an toàn. Thông tin nhúng (dấu thủy vân) được dùng để xác định tính toàn vẹn, phát hiện vị trí bị thay đổi hoặc chứng minh quyền tác giả đối với dữ liệu chứa dấu thủy vân.

* + - 1. **Phân loại thủy vân**

Dựa vào mục đích sử dụng, thủy vân số có thể được phân loại như hình sau



Hình 4. Phân loại thủy vân theo mục đích ứng dụng

Đối với các lược đồ thủy vân dễ vỡ (Fragile), dấu thủy vân sẽ bị thay đổi trước sự biến đổi của dữ liệu đã thủy vân dù chỉ vài bít. Vì vậy, thủy vân dễ vỡ thường được ứng dụng trong bài toán xác thực tính toàn vẹn của dữ liệu thủy vân. Trong một số trường hợp, thủy vân dễ vỡ có thể khoanh vùng được dữ liệu thay đổi.

Trái với thủy vân dễ vỡ, thủy vân bền vững (Robust) yêu cầu dấu thủy vân phải tồn tại trước sự tấn công vô tình hay có chủ định nhằm loại bỏ dấu thủy vân.

Đối với thủy vân bền vững, dấu thủy vân phải luôn tồn tại trước mọi sự tấn công trừ khi dữ liệu thủy vân không còn giá trị sử dụng. Một số phép tấn công thường được sử dụng như: nén, cắt, xoay, lọc, làm mờ, cân bằng sáng...Do dấu thủy vân bền vững (tồn tại) trước các sự tấn công nên nhóm các lược đồ này thường được sử dụng trong bài toán bảo vệ bản quyền hay xác định chủ sở hữu.

Ngoài thủy vân dễ vỡ và thủy vân bền vững, các lược đồ thủy vân có khả năng bền vững trước các phép tấn công nén, nhiễu, print, cân bằng sáng nhưng lại dễ vỡ với các phép tấn công copy/paste, co dãn thì được gọi là thủy vân bán dễ vỡ. Một cách tổng quát, dấu thủy vân của lược đồ thủy vân bán dễ vỡ cần phải dễ thay đổi trước các phép tấn công nhằm thay đổi nội dung hay ý nghĩa của dữ liệu nhưng lại bền vững trước sự thay đổi định dạng dữ liệu hoặc các lỗi đường truyền. Các lược đồ thủy vân bán dễ vỡ thường sử dụng trong bài toán phát hiện sự giả mạo nội dung, xuyên tạc thông tin.

* + - 1. **Các ứng dụng của thuỷ vân với ảnh số**

Các ứng dụng của thuỷ vân đối với ảnh số bao gồm các lĩnh vực như bảo vệ bản quyền, xác thực ảnh và bảo toàn dữ liệu, giấu dữ liệu và gán nhãn ảnh, ta sẽ lần lượt đề cập từng ứng dụng.

a) Bảo vệ bản quyền

Mục đích của thuỷ vân với bảo vệ bản quyền là gắn một “dấu hiệu” vào dữ liệu ảnh mà có thể xác định được người nắm giữ bản quyền. Và ta cũng có thể gắn thêm một dấu hiệu khác gọi là vân tay để xác định người dùng của sản phẩm. Dấu hiệu có thể là một dãy số như mã hàng hoá quốc tế, một message hoặc một logo... Thuật ngữ thuỷ vân xuất phát từ phương thức đánh dấu giấy tờ với một logo từ thời xa xưa với mục đích tương tự.

b) Xác thực ảnh và bảo toàn dữ liệu

Một ứng dụng khác của thuỷ vân là xác thực ảnh và phát hiện giả mạo. Ảnh số ngày càng được sử dụng như các bằng chứng trước pháp luật ngày nay. Vấn đề là cần xác thực được tính hợp pháp của ảnh này. Thuỷ vân được sử dụng ở đây để xác định xem ảnh là nguyên bản hay đã chịu tác động của con người, bằng các ứng dụng xử lý ảnh. Thuỷ vân được dấu lúc đầu phải mang tính chất không bền vững, để bất kỳ sự thay đổi nhỏ nào tới ảnh cũng có thể làm hỏng thuỷ vân hoặc phát hiện được thay đổi đối với thuỷ vân này. Tuy vậy, thuỷ vân vẫn phải tồn tại với các phép biến đổi ảnh thông thường như chuyển đổi định dạng, lấy mẫu, nén...

c) Giấu dữ liệu và gán nhãn ảnh

Giấu dữ liệu là nhằm trao đổi dữ liệu bí mật thông qua một bức ảnh. Điều này cho phép trao đổi thông tin mà không gây chú ý đối với người ngoài. Khối lượng dữ liệu dấu được là quan trọng nhất đối với mục đích này.

Còn gán nhãn ảnh là ứng dụng dùng để cung cấp thêm thông tin cho người dùng hoặc để phục hồi ảnh từ cơ sở dữ liệu.

* + - 1. **Một số tính chất của sơ đồ thủy vân**

Tuỳ thuộc vào từng loại thuỷ vân số và ứng dụng của nó mà ta có các yêu cầu khác nhau đối với các phương pháp tạo thuỷ vân. Ở đây chỉ đề cập đến yêu cầu đối với thuỷ vân số ẩn.

Đối với loại thuỷ vân này, ba yêu cầu sau đây là cơ bản và cần thiết đối với các ứng dụng bảo vệ bản quyền:

**Tính ẩn:** Thứ nhất, thuỷ vân phải ẩn đối với trực giác của con người (imperceptibility hay perceptual tranperancy), tức là con người phải không nhận biết được sự có mặt của thuỷ vân trong ảnh. Điều này cũng có nghĩa là việc dấu thuỷ vân chỉ gây ra sự thay đổi rất nhỏ đối với ảnh, không ảnh hưởng đến chất lượng ảnh.

**Tính bền vững:** Yêu cầu thứ hai là thuỷ vân phải bền vững (robustness), thuỷ vân phải có khả năng tồn tại cao với các 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 đổi A/D và D/A ... còn các tấn công có chủ đích có thể là việc xoá, thay đổi hoặc làm nhiễu thuỷ vân trong ảnh. Để thực hiện được điều này, thuỷ vân phải được dấu trong các vùng quan trọng đối với trực giác(perceptualsignificant). Phương pháp thuỷ vân số phải đảm bảo sao cho việc không thể lấy lại thuỷ vân tương đương với việc ảnh đã bị biến đổi quá nhiều, không còn giá trị về thương mại.

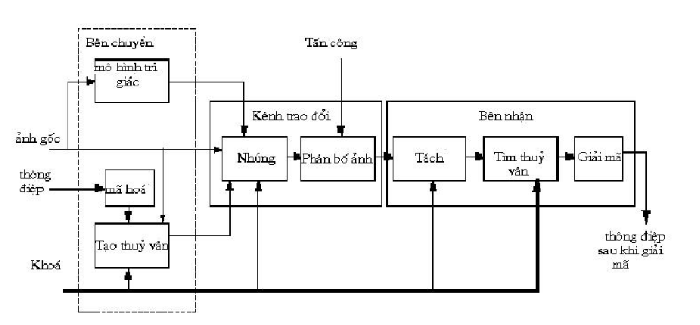
**Khả năng mang tin cao**: Với yêu cầu này, lượng tin cần thêm vào ảnh phải đủ dùng trong ứng dụng mà không làm thay đổi quá nhiều chất lượng ảnh.

Tuy vậy, việc làm tốt cả ba yêu cầu trên là một điều rất khó. Để dấu thuỷ vân trong ảnh thì ta bắt buộc phải thay đổi dữ liệu ảnh. Ta có thể tăng tính bền vững cho thuỷ vân bằng cách tăng lượng thay đổi ảnh cho mỗi đơn vị tin cần dấu. 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 thuỷ vân trong ảnh sau các phép tấn công có thể không đủ để xác định thuỷ vân. Nếu thông tin được dấu 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 hai nhân tố quyết định cho khối lượng tin được dấu trong ảnh.

Còn đối với các ứng dụng để phát hiện giả mạo ảnh gốc thì thuỷ vân nhúng vào phải có tính giòn (fragile) nghĩa là sẽ bị vỡ nếu chịu sự biến đổi mất thông tin. Một cách thực hiện điều này mà vẫn đảm bảo tính ẩn của thuỷ vân là dấu nó vào các phần dữ liệu ít đáng chú ý về mặt trực giác (perceptual insignificant). Đối với ảnh, đó có thể là các bít thấp của các điểm ảnh (LSB).

* + 1. **Mô hình thủy vân số**

****

Hình 5. Mô hình hệ thống thủy vân số

Trên đây là mô hình hệ thống thủy vân đối với ảnh số. Quá trình liên quan đến xử lý thủy vân gồm 4 bước chính mà ta có thể mô tả như sau:

* + - 1. **Tạo thủy vân số**

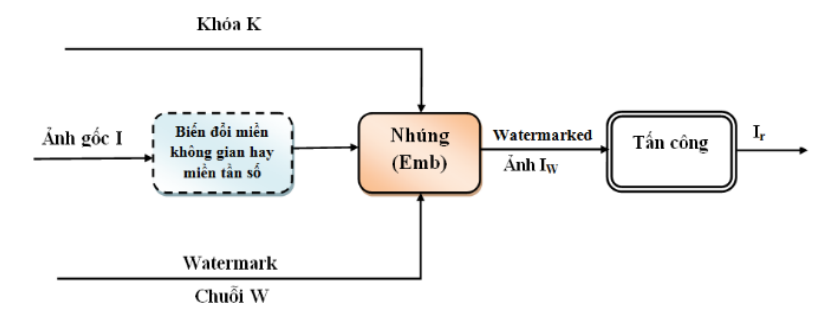
Thủy vân có thể là một hình ảnh dạng logo hay văn bản với độ dài cho trước. Thủy vân dạng hình ảnh có khả năng chống chịu trước các phép xử lý ảnh tốt hơn nhiều só với dạng thủy vân dạng ký tự. Thủy vân có thể được biến đổi (bằng mã hóa, chuyển đổi định dạng), trước khi giấu vào ảnh. Các thuật toán nhúng 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.

* + - 1. **Quy 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 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 6. Quy trình nhúng ảnh thủy vân

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:

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. **Tách thủy vân**

Để tách thủy vân ra khỏi ảnh, ta sẽ dùng khóa K trong quá trình nhúng, và ảnh cần tách thủy vân. Thuật toán tách thủy vân có các bước ngoặc với thuật toán nhúng ở trên.

* + - 1. **Phát hiện thủy vân và giải mã**

Đối với thủy vân là một logo thì sau khi giải mã, việc xác định thủy vân có tồn tại hay không là đơn giản. Còn nếu thủy vân là một dãy số có phân bố Gauss thì có thể dựa vào kiểu tương quan, kiểu phân bố của dãy số thu được để đánh giá

Ta có thể dùng một số tiêu chuẩn sau để làm phép đo cho sự thành công của việc tách thủy vân

Tỉ số của tín hiệu nhọn đối với nhiễu – PSNR ( Peak Signal to Noise Ratio )

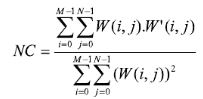


Với MSE là sai số bình phương trug bình (mean-square error). Với ảnh kích thước MxN thì MSE được định nghĩa như sau:

****

Với X(i,j) và XW(i,j) là các giá trị của điểm ảnh (i,j) trong ảnh ban đầu và ảnh đã thủy phân tươmg ứng.

Một giá trị khác hay được sử dụng đó là hệ số tương quan chuẩn – NC (Normarlized Correlation), được định nghĩa như sau:



Các ứng dụng gán nhãn ảnh hoặc dấu dữ liệu thường cố gắng trả về nguyên vẹn thông tin (thủy vân) được nhúng ban đàu. Vì thông tin có thể sai lệch nên người ta thường sử dụng các loại mã sửa sai với các ứng dụng này.

Cơ chế thủy vân cho các ứng dụng xác thực ảnh hoặc bảo toàn dữ liệu thì đưa ra câu trả lời có/không về việc ảnh có giải mạo, đưa ra xác định về nguồn dữ liệu hợp pháp hoặc chỉ ra các vùng dữ liệu đã thay đổi.

* 1. **Thuật toán thủy văn theo miền không gian ảnh**

Các thuật toán 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, thuật toán này được cài đặt cho ứng dụng xác thực thông tin của ảnh số.

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 là 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 sử dụng ảnh đen trắng. Ảnh đen trắng hay còn gọi là ảnh nhị phân là ảnh chỉ có hai mức giá trị mức xám là mức 0(đen) và 1 (trắng). Để tạo thủy vân cho ảnh đen trắng ta đem nhúng thủy vân vào ảnh nhị phân. Thông thường việc nhúng thủy vân vào ảnh đen trắng khó thực hiện hơn ảnh đa cấp xám hay ảnh màu. Lý do là ảnh nhị phân chỉ có hai mức xám duy nhất, vì thế nếu thay đổi một bít của điểm ảnh thì đồng nghĩa với thay đổi toàn bộ điểm ảnh.

Có hai cách để nhúng dữ liệu vào ảnh nhị phân là thay đổi giá trị của từng bit riêng lẻ hoặc thay đổi giá trị của một nhóm bít. Cách thứ nhất sẽ đảo ngược một điểm đen thành trắng hoặc một điểm trắng thành đen. Cách tiếp cận thứ 2 sẽ làm thay đổi một số đặc trưng của ảnh như độ dày của cạnh, vị trí tương quan giữa các bit…Cách tiếp cận này tùy thuộc nhiều vào kiểu ảnh (kiểu văn bản, kiểu bản đồ). Vì số tham số có thể thay đổi là hữu hạn, đặc biệt là yêu cầu thủy vân ẩn, tổng số dữ liệu có thể dấu được là hữu hạn.

* + 1. **Thuật toán SW**

Đây là một thuật toán đơn giản. 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, các 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.

* Quá 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

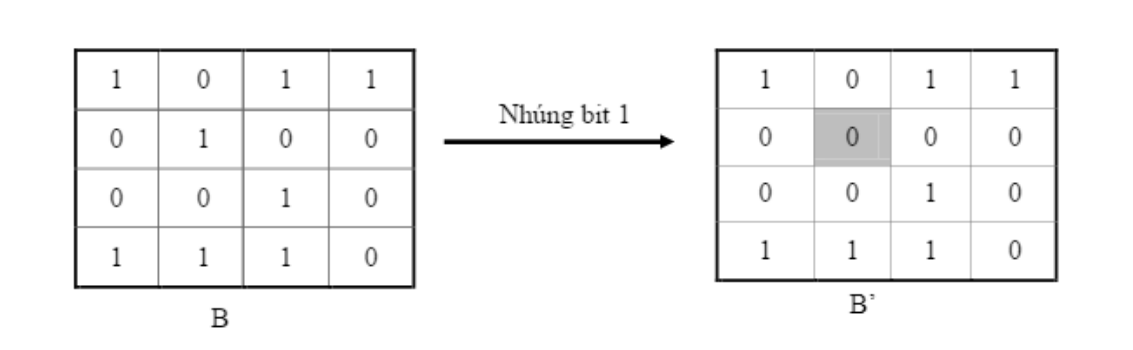
Nếu t và di cùng tính chẵn lẻ thì khối B mặc nhiên đã giấu được bit di mà không cần làm gì.

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 7. Minh họa thuật toán WS: nhúng bit 1 bào 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.

* Quá trình tách thủy vân

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:

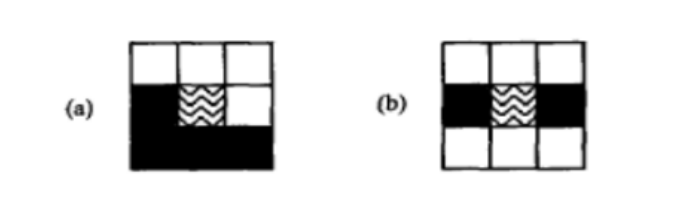
Bước 1: Chia ảnh có nhúng thủy vân B’ thành các khối có kích thước m x n với mỗi khối Bi’ trong B’ ta tính Sum [Bi’]

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.



Hình 8. Minh họa chọn điểm ảnh giấu tin vào những khối ảnh màu

* + 1. **Thuật toán WU-LEE**

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:

* Quá trình nhúng thủy vân
* 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.

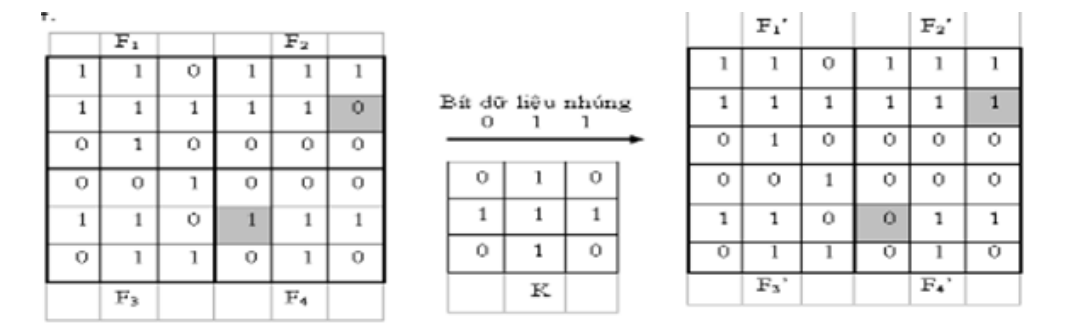
Trường hợp 4: Nếu SUM (Fi ^ K) mod 2 # b và 1< SUM (Fi ^ K) < SUM (K) – 1 thì chọn ngẫu nhiên một bit của Fi tại vị trí (j,k) mà K (j,k)=1 và đảo ngược Fi (j,k).

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 9. 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).

* Quá trình trích thủy vân

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:

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.

Ảnh F được lựa chọn để nhúng tin có quá nhiều điểm trắng hay quá nhiều điểm đen đều làm giảm tỷ lệ bít giấu được.

Thuật toán Wu – Lee đơn giản, lượng tin giấu được không thấp nhưng tính bảo mật không cao, không thích hợp với ảnh có mảng đen và trắng rộng.

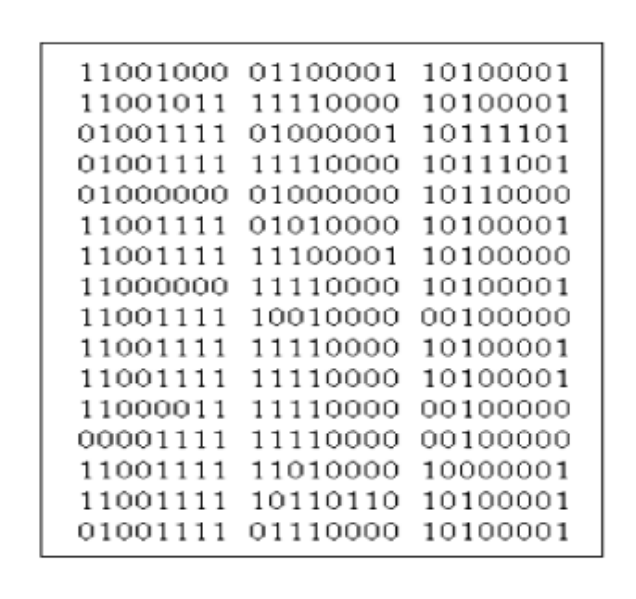
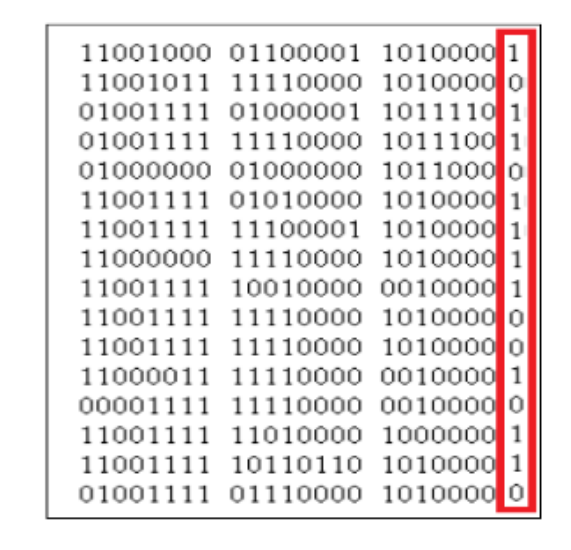
* + 1. **Thuật toán LBS**

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. (Hương N. D., 2012)

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 giá 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.
* Thuật toán tách thủy vân
* 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.
  + 1. **Thuật toán PCT**

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

{ [W]i,j|Ii=1…m = 1....n } = {1,2,……,2r-1}

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.

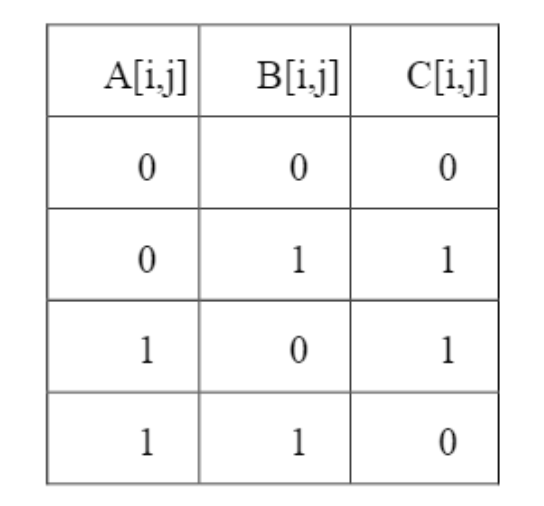
Ma trận W kích thước m\*n được gọi là ma trận trọng số cấp r nếu mỗi phần tử của tập hợp {1,2,….,2r-1} xuất hiện trong W ít nhất một lần và các phần tử của W chỉ nhận giá trị trong tập hợp {1,2,….,2r-1} với m, n, r là các số tự nhiên thỏa mãn 2 r – 1 ≤ m \* n.

Từ định nghĩa, ta nhận thấy với mỗi m, n, r thỏa mãn 2 r – 1 ≤ m \* n sẽ có:

* 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.
* Các phép toán trên ma trận:

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

* Phép cộng C = A + BSD

Trong đó C[i,j] = A[i,j] + B[i,j] ;

* Phép nhân C= A x B

Trong đó C[i,j] = A [i,j] \* B[i,j].

* Thuật toán

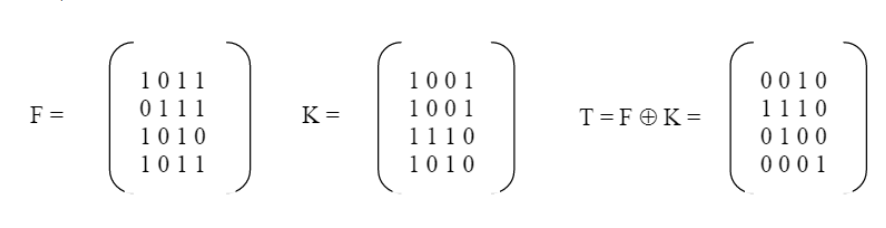
Với khối ảnh Fi, ma trận trọng số W, khóa K, ta cần giấu r bit thông tin b1,b2,….br vào Fi bằng cách đảo nhiều nhất 2 bit của Fi. Mục đích của ta là biến đổi F i thành F i’ sao cho thỏa mãn yêu cầu sau: SUM (FI’ K) x W = b1b2…..br (mod 2r). (Hoan, năm 2006) (Huy & Dũng)

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

* Bước 1: 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:

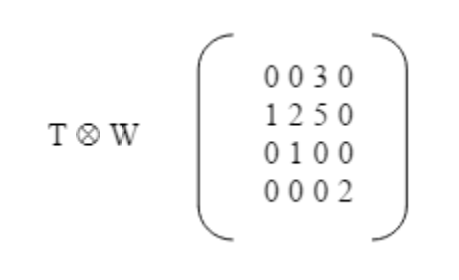


* Bước 2: Tính S = SUM (T x W)

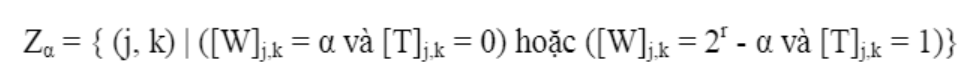
Trong đó:

* 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

Với mỗi số nguyên α thuộc khoảng từ 1 đến (2r - 1), tính được tập con tương ứng Z α. Khi đó tập Z là một tập hợp gồm 2r – 1 tập con. Mỗi tập Z α là một tập hợp chứa tất cả các chỉ số (j,k) của ma trận. Như vậy nếu thay đổi giá trị của phần tử thứ (j,k) trong ma trận F (thay 0 thành 1 , thay 1 thành 0 ) sẽ làm cho S tăng thêm α đơn vị (hiểu theo mod 2r) . Thực tế , có hai khả năng để đạt được :

* Nếu [W]j,k= α và [T]j,k = 0 thay đổi giá trị của phần tử [F]j,k thì S tăng thêm α đơn vị.
* Nếu Nếu [W]j,k= α và [T]j,k = 1 thay đổi giá trị của phần tử [F]j,k thì S giảm đi 2r – α đơn vị hoặc theo mod 2r thì S tăng thêm α đơn vị.
* 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: S’ = b (mod 2r) (\*)

Trong đó b = (b1b2b3………br).

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

Đặt d = b- S (mod 2r)

* Trường hợp 1: Nếu d = 0 thì S = b mod 2r. 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:

Nếu Zd # Rỗng thì cần chọn một cặp (j,k) bất kỳ thuộc Zd rồi thay đổi giá trị phần tử [F]j,k khi đó S sẽ tăng thêm d đơn vị (theo mod 2r), do đó đã đạt được bất biến. Trong trường hợp này giấu được b vào trong F chỉ cần thay đổi 1 phần tử trong F.

Nếu Zd = Rỗng thì thực hiện các bước sau:

Chọn h là số tự nhiên đầu tiên thỏa mãn Zhd # Rỗng và Z (1-h)d # Rỗng. Chọn cặp (j,k) bất kỳ thuộc Zhd và thay đổi giá trị của phần tử [F]j,k , khi đó S tăng thêm h \* d .

Chọn cặp (u,v) bất kỳ thuộc Z- (h-1)d và thay đổi giá trị của phần tử [F]u,v khi đó S tăng thêm (1-h)\*d = d- h\*d. Vậy khi thay đổi giá trị [F]j,k và [F]u,v thì S tăng một lượng là h \*d + d – h\* d= d. Trong trường hợp này giấu b vào trong F, cần thay đổi tới hai phần tử trong F.

* 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:

Nếu Zd+ 2r khác rỗng thì cần chọn cặp (j,k) bất kỳ thuộc Zd+ 2r rồi thay đổi giá trị phần tử [F]j,k khi đó S sẽ tăng d đơn vị (theo mod 2r) do đó đã đạt được bất biến. Trong trường hợp này giấu được b vào trong F chỉ cần thay đổi một phần tử trong F.

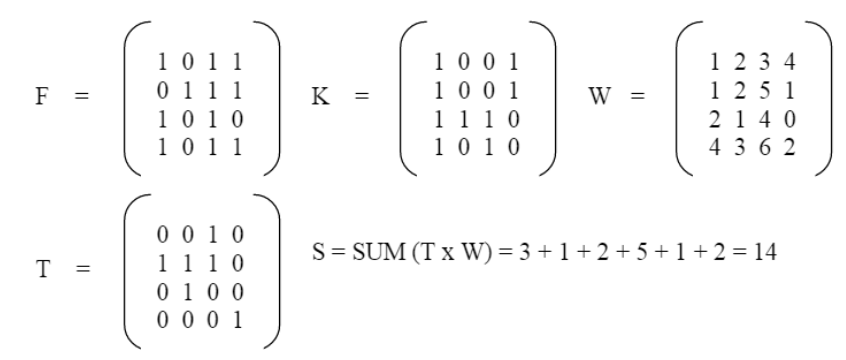
Nếu Zd+ 2r = rỗng thì thực hiện các bước sau:

Chọn h là số đầu tiên thỏa mãn Zhd khác rỗng và Z(1-h)d + 2r khác rỗng. Chọn cặp (j,k) bất kỳ thuộc Zhd và thay đổi giá trị của phần tử [F]j,k , khi đó S tăng thêm h \* d

Chọn cặp (u,v) bất kỳ thuộc Z(1-h)d + 2r và thay đổi giá trị của phần tử [F]u,v khi đó S cũng tăng thêm (1-h)\*d + 2r = d- h\*d + 2r

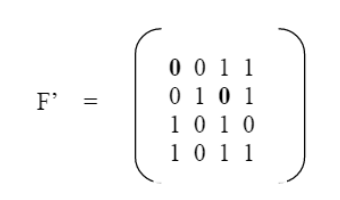
Vậy khi thay đổi giá trị của hai phần tử [F]j,k và [F]u,v , thì S tăng một lượng là h \* d + d- d\*h + 2r = d + 2r (theo mod 2r) thì S tăng thêm d. Trong trường hợp này giấu b vào F, cần thay đổi đến hai phần tử trong F.

* 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’ =35 SUM (F’ + K) x W => b = S’ mod 2r, từ đó xác định được dãy bít (b1b2…br) đã giấu trong F

Ví dụ:

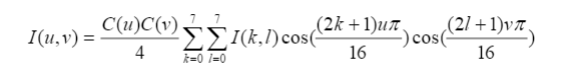
* Giả sử cần nhúng một dãy bit 1010 vào trong F
* Ta có L: r= 4, b = 23 + 21 = 10;
* Đặt d = b – S mod 2r = 10 – 14 mod 24 = - 4
* Zd +2r = Z10 = rỗng.
* Chọn h = -12/4
* Zh\*d= Z1= {(1,1); (2,4)} khác rỗng Z (h-1)\*d+ 2r = Z 11= {(2,3)} khác rỗng.



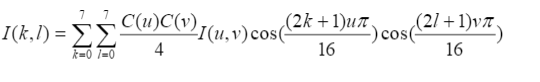
* 1. **Thuật toán thủy văn theo miền tần số**
     1. **Biến đổi cosine rời rạc (DCT)**

Biến đổi cosine 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. (Nguyễn Xuân Huy, 2002)

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.

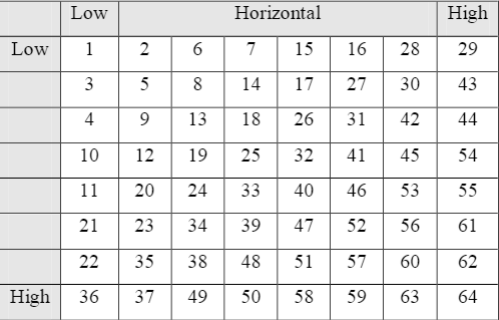
Công thức biến đổi DCT thuận từ I (k,l) -> I (u,v)

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. (Cường, 2012)

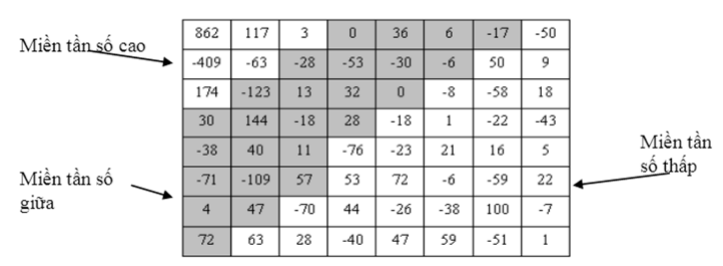


Hình 10. Hình ảnh bảng hệ số DCT

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.

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 38 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.

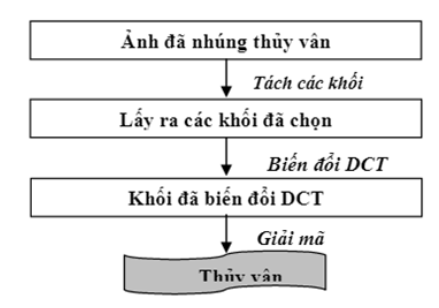
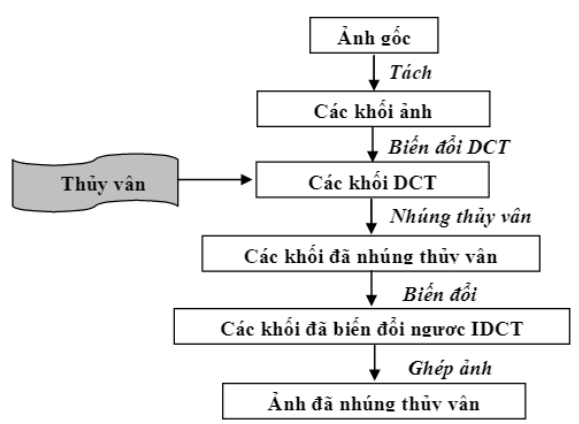
Khối hệ số DCT có thể chia thành ba miền; miền tần số thấp, miền tần số giữa và miền tần số cao. Miền tần số thấp chứa các thông tin quan trọng ảnh hưởng đến tri giác. Các thông tin trong miền tần số cao thường không mang tính tri giác cao, khi nén JPEG thì thường loại bỏ thông tin trong miền này. (Cường, 2012)



Hình 11. Bảng 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. Tuy nhiên, các thuật toán khác nhau thì khác nhau về cách lựa chọn vị trí nhúng thủy vân và phương thức nhúng. (Huy & Dũng)

Hình 12. Quy trình tách và nhúng thủy vân theo kỹ thuật vân trên miền DCT

1. Thuật toán DCT1

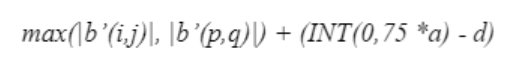
Thuật toán được nhóm tác giả Nguyễn Xuân Huy và Trần Quốc Dũng đưa ra trên bài báo: “Một thuật toán thủy vân trên miền DCT - An Image Watermarking Algorithm Using DCT domain”. Nội dung bài viết đề xuất một thuật toán nhúng thủy vân vào trong ảnh sao cho thỏa mãn các tính chất và yêu cầu của một hệ thủy vân trên ảnh số. Thuật toán chọn miền tần số để giấu tin nhằm nâng cao tính bền vững của thủy vân. (Huy & Dũng)

* Mô tả thuật toán
* Đầu vào:
* Một chuỗi các bit thể hiện bản quyền
* Một ảnh.
* Đầu ra:
* Một ảnh sau khi thủy vân. Khóa để giải mã
* Quá trình thủy vân
* Chia ảnh có kích thước m x n thành (mxn )/64 khối 8x 8 , mỗi bit sẽ được giấu trong một khối.
* Chọn một khối bất kì B và biến đổi DCT khối đó thu được B’
* Chọn hai hệ số ở vị trí bất kì trong miền tần số ở giữa của khối DCT, giả sử đó là b ‘(i,j) và b’ (p,q). Ta tính:

d= || b ‘(i,j) - b ‘(p,q) || mod a

Trong đó a là một tham số thỏa mãn điều kiện a = 2 (2t +1), t là một số nguyên dương.

Bít si sẽ được nhúng sao cho thỏa mãn điều kiện sau:

* Nếu d < 2t + 1 và si = 1 thì một trong hai hệ số DCT b’(i,j) hoặc b’(p,q) có trị tuyệt đối lớn hơn sẽ bị thay đổi để d > = 2t + 1 theo công thức sau:

Với hàm max (|b’(i,j) | , |b’(p,q)| ) là hàm chọn ra hệ số có trị tuyệt đối lớn hơn, hệ số được chọn sẽ được cộng thêm một lượng là INT (0,75 \*a) – d .

Hoặc cũng có thể biến đổi một trong hai hệ số theo công thức:

Với hàm min (|b’(i,j) | , |b’(p,q)| ) là hàm chọn ra hệ số có trị tuyệt đối nhỏ hơn, hệ số được chọn sẽ bị trừ đi một lượng là INT (0,25 \*a) + d .

INT () là hàm làm lấy phần nguyên của một số thực.

* Tương tự, nếu d > = 2t + 1 và si = 0 thì một trong hai hệ số DCT b’(i,j) hoặc b’(p,q) có giá trị tuyệt đối lớn hơn sẽ được thay đổi để thõa mãn d < 2t + 1 như sau:

Với hàm max (|b’(i,j) | , |b’(p,q)| ) là hàm chọn ra hệ số có trị tuyệt đối lớn hơn, hệ số được chọn sẽ bị trừ đi một lượng là INT d – INT (0,25 \*a).

Hoặc

* Quy trình trích để lấy lại thông tin:
* Đầu vào
* Một ảnh đã nhúng thủy vân
* Khóa để giải mã
* Đầu ra

Thủy vân là một dãy bit đã nhúng

Thực hiện: Đọc khối DCT từ ảnh chứa thủy vân và vị trí hai hệ số đã biến đổi, sau đó tính:

1. Thuật toán DCT2

* Mô tả thuật toán

Cùng ý tưởng nhúng thủy vân vào miền tần số giữa của khối biến đổi cosine 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 bảo đảm 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.

* Quá trình thủy vân

Thủy vân là một chuỗi các bit hoặc một ảnh nhị phân được nhúng vào ảnh gốc. Ảnh gốc có kích thước m x n sẽ được chia thành m x n / 64 khối 8 x 8, mỗi bit của thủy vân sẽ được nhúng trong một khối.

Chọn một khối ảnh gốc Fi, thực hiện phép biến đổi DCT với Fi để được Fi’

Chọn hai hệ số thuộc miền tần số giữa của Fi’, giả sử đó là Fi’(u,v) và Fj ‘(p,q), đọc thủy vân cần nhúng giả sử đó là si

Nếu bit cần nhúng si = 0 và nếu Fi’ (u,v) < Fi’ (p,q) thì đổi chỗ hai hệ số này.

Nếu bit cần nhúng si = 1 và nếu Fi’ (u,v) >= Fi’ (p,q) thì đổi chỗ hai hệ số này. Nếu Fi’ (u,v) > Fi’ (p,q) và nếu Fi’ (u,v) – Fi’ (p,q) < k thì tăng Fi’ (u,v) đồng thời giảm Fi’ (p,q) k/2 lần.

Nếu Fi’ (u,v) <= Fi’ (p,q) và nếu Fi’ (p,q) – Fi’ (u,v) < k thì tăng Fi’ (p,q) đồng thời giảm Fi’ (u,v) k/2 lần Dùng phép biến đổi ngược IDCT với mỗi khối đã nhúng thủy vân Fi’ . Ghép các khối ảnh để được ảnh đã nhúng thủy vân.

* Quá trình trích để lấy lại thông tin:

Đọc vào khối DCT đã nhúng thủy vân Fi’ và vị trí hai hệ số đã biến đổi (u,v) và (p.q), sau đó tính k = Fi’ (u,v) – Fi’ (p,q) .

Nếu k > 0 thì gán si = 0

Nếu k < 0 thì gán si = 1

Ghép dãy bit si để được thủy vân đã nhúng.

* Nhận xét:

Sau khi thử nghiệm cho thấy, hệ thống thủy vân trên đáp ứng tốt tính chất bảo đảm tính bền vững của thủy vân trước đa số các phép biến đổi ảnh thông thường. Hệ số k được gọi là hệ số tương quan giữa tính ẩn của thủy vân với tính bền vững của thủy vân. Hệ số k càng lớn, tính bền vững của thủy vân càng cao, đồng thời chất lượng ảnh sau khi nhúng thủy vân ẩn bền vững đó là mâu thuẫn giữa chất lượng thương mại của ảnh sau khi nhúng thủy vân với tính bền vững của thủy vân trước các tấn công. Trong thực tế, có thể xây dựng một hệ thống thủy vân với đề xuất về thông số giữa việc chọn hệ số k, chất lượng ảnh sau khi nhúng thủy vân và đồ bền vững của thủy vân trước các tấn công để người sử dụng tùy theo mục đích mà lựa chọn các thông số phù hợp.

Khóa để giải mã trong việc phát hiện thủy vân gồm kích thước khối và vị trí cặp hệ số được chọn trong khối. Do đó, độ phức tạp của việc dò tìm thủy vân khi không biết khóa phụ thuộc rất nhiều vào kỹ thuật chọn cặp hệ số trong quá trình nhúng thủy vân. Có thể chọn cố định một cặp số cho tất cả các khối, cũng có thể chọn vị trí thay đổi cho mỗi khối, khi đó vị trí tương ứng của cặp hệ số trong mỗi khối sẽ là một phần trong khóa để phát hiện thủy vân.

Quá trình tách thủy vân không cần sử dụng ảnh gốc.

1. Thuật toán DCT3

* Mô tả thuật toán :

Trong thuật toán DCT3 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 khổi nhẵn hoặc khối sắc không cao.

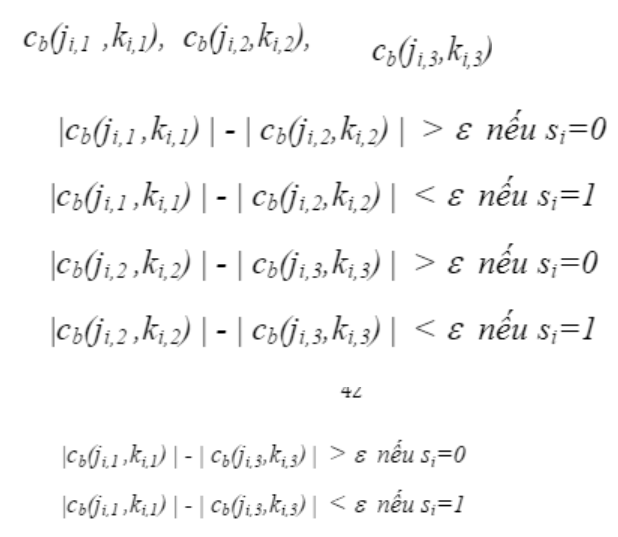
Các khối được chọn nhúng thủy vân là các khối sắc lớn.

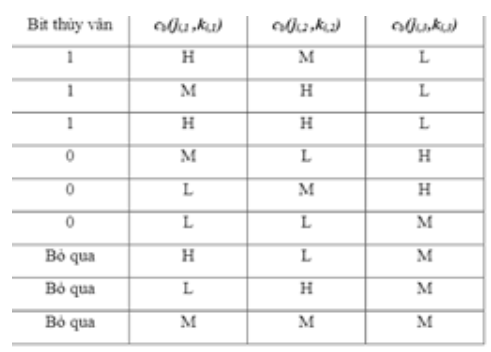
Khối nhẵn : chúng ta có thể phát hiện ra các khối này bằng cách đếm số lượng hệ số cao tần có giá trị là “0”. Nếu tất cả các hệ số này hay chỉ cần tồn tại ít nhất 1 hệ số ở nửa trên đường zig zắc bằng “0’ thì khối đó được xem là khối nhẵn.

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.

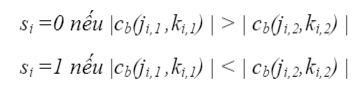
Thuật toán sử dụng 3 hệ số để nhúng 1 bit.

* Quá trình thủy vân:

Để nhúng bit si vào trong khối, 3 hệ số chọn ngẫu nhiên

Nếu thay đổi một trong 3 hệ số là quá lớn thì đơn giản là bỏ qua khối đó và bit đó sẽ được nhúng vào khối tiếp theo. H là hệ số có giá trị cao nhất trong 3 hệ số, M là hệ số ở giữa, L là hệ số thấp nhất.

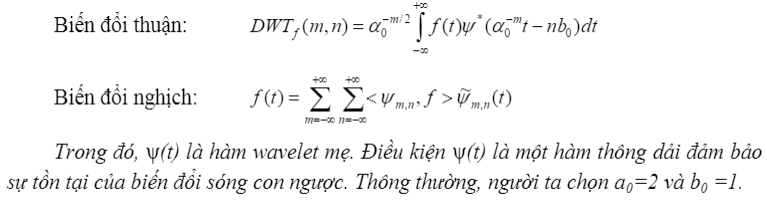
* Quá trình trích để lấy lại thông tin:

Lặp lại các vị trí nhúng tương ứng và các hệ số đã chọn. Lấy thủy vân theo điều kiện

* + 1. **Thuật toán thủy văn dựa trên miền DWT**
* Phép biến đổi sóng rời rạc.

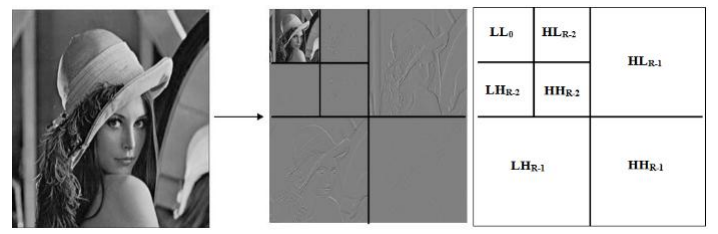
Trong phép biến đổi này, Wavelets là các hàm được định nghĩa trong khoảng hữu hạn và có giá trị trung bình bằng 0. Ý tưởng cơ bản của phép biến đổi con sóng con là khai triển hàm f(t) bất kỳ như một xếp chồng của các con sóng con hay các hàm cơ sở. Các hàm cơ sở này có được từ một con sóng con nguyên 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 LH1sẽ 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 13. Biến đổi Wavelet và cấu trúc giải thông

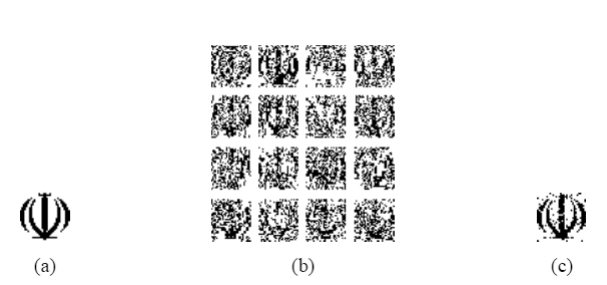
Trong một số trường hợp, sơ đồ dùng biến đổi sóng con đã tỏ ra ưu thế so với biến đổi Fourier rời rạc DFT hay biến đổi cosine rời rạc DCT. Do đặc tính đa phân giải, sơ đồ mã hóa Wavelets đặc biệt thích hợp cho các ứng dụng mà tính vô hướng và suy biến đóng vai trò quan trọng. Minh chứng cho điều này 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.

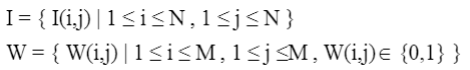
* Lược đồ thủy vân sử dụng biến đổi DWT

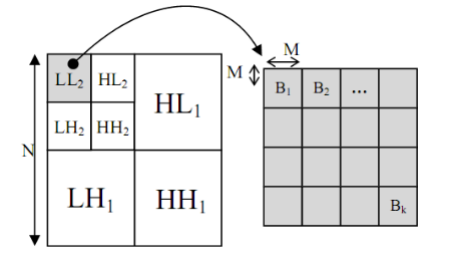
Ngày nay, có nhiều thuật toán thủy vân sử dụng biến đổi sóng con và các kỹ thuật lượng tử hóa, thủy vân sử dụng miền biến đổi wavelet có lợi thế làm cho các thủy vân mạnh mẽ hơn chống lại được nhiều dạng tấn công như thay đổi thành phần tần số cao của hình ảnh, nén, lọc thông thấp qua, tuy nhiên nó không 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 ảnh gốc và ảnh đã thủy văn Q=35 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.

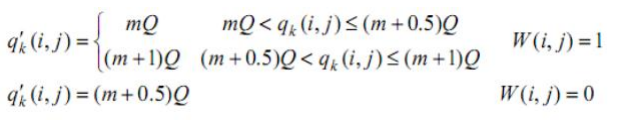
****Hình 14. (a)Thủy vân gốc (b)Thủy vân được trích xuất từ các khối (c)Thủy vân kết hợp.

a) Thuật toán nhúng

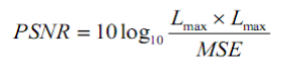
Đầu vào là ảnh mang I có kích thước N x N, ảnh nhị phân logo thủy vân W kích thước M x M

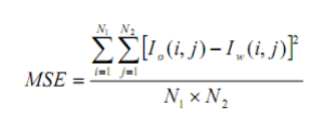
* Quy trình nhúng thủy vân :
* Bước 1: Hình ảnh mang được phân ra thành n mức sử dụng biến đổi wavelet rời rạc. Sau đó, chúng ta chọn dải thông con LLn để nhúng thủy vân.
* Bước 2: Chia dải thông con được lựa chọn thành các khối nhỏ hơn Bk với kích thước M x M. Như hình dưới đây:

Hình 15. Dải thông LL2 được chia các khối nhỏ hơn

* Bước 3: Logo thủy vân được chèn vào tất cả các khối con bằng cách lượng tử hóa các hệ số của khối theo công thức sau:

Trong đó qk (i,j) được sử dụng để biểu diễn các hệ số wavelet của khối con Bk và q’k (i,j) được sử dụng để biểu diễn các hệ số sau khi được lượng tử hóa. W (i,j) là logo thủy vân, m là một số nguyên và Q là kích thước bước lượng tử hóa.

Lựa chọn một giá trị Q tốt cho lược đồ thủy vân là rất quan trọng bởi việc tăng Q có thể làm tăng tính bền vững của thủy vân trước các cuộc tấn công nhưng lại suy giảm chất lượng ảnh mang hay nói cách khác ảnh hưởng đến tính vô hình của thủy vân. Vì vậy Q và PSNR có mối quan hệ với nhau và Q tỉ lệ nghịch với PSNR. Trong đó PSNR được tính theo công thức:

Với L max là giá trị cực đại của các điểm ảnh và MSE là tỷ lệ lỗi trung bình được định nghĩa :

Trong đó Io và Iw biểu diễn giá trị điểm ảnh tương ứng trên ảnh gốc và ảnh chứa thủy vân và N1, N 2 là kích thước của ảnh.

Ngoài ra ngưỡng giá trị Q còn khác nhau phụ thuộc vào ảnh mang và miền tần số. Ví dụ như trong hình dưới:



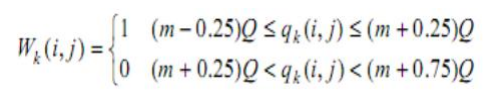
Hình 16. Hình (a)- Ảnh gốc (b)-Ảnh đã thủy vân với Q=35

* Bước 4: Cuối cùng với những hệ số giá trị mới, sử dụng biến đổi sóng con ngược để thu được hình ảnh chứa thủy vân. Việc lựa chon n mức phân giải sóng con, phải đảm bảo cân bằng giữa tính bền vững và tính vô hình. Với một lựa chọn tốt, thủy vân có thể có được tính bền vững cao hơn chống lại được sự suy giảm chất lượng hình ảnh. Chọn giá trị n nhỏ có thể làm tăng tốc độ thực hiện của thuật toán nhưng lại làm giảm tính bền vững và làm suy giảm chất lượng hình ảnh. Ngược lại nếu lựa chọn giá trị n lớn có thể tăng tính bền vững nhưng sẽ làm giảm kích thước của miền LLn do đó có thể gây ra giảm số lượng các khối con K. Như vậy rõ ràng có một mối quan hệ giữa N, M, n, K. Theo kết quả thực nghiệm tác giả chỉ ra rằng trường hợp tối ưu khi N = M x K x 2n

b) Thuật toán trích xuất

Trong khi hầu hết các phương pháp trích xuất thủy vân đều đòi hỏi hình ảnh gốc, thuật toán được trình bày là một trong những thuật toán trích xuất mù không yêu cầu ảnh gốc trong quá trình trích xuất thủy vân. Trong thuật toán, để trích xuất thủy vân cần khóa bí mật chính là kích thước bước lượng tử Q, mức phân giải n số lượng các khối con K .

c) Quy trình trích xuất thủy vân

* Bước 1: Những hình ảnh mang được phân giải theo n mức sử dụng biến đổi wavelet n rời rạc. Dải thông con LLn của hình ảnh phân giải được chia thành các khối con Bk với kích thước M x M.
* Bước 2: Những điểm ảnh của logo thủy vân tương ứng với mỗi khối con Bk được trích xuất theo công thức sau:

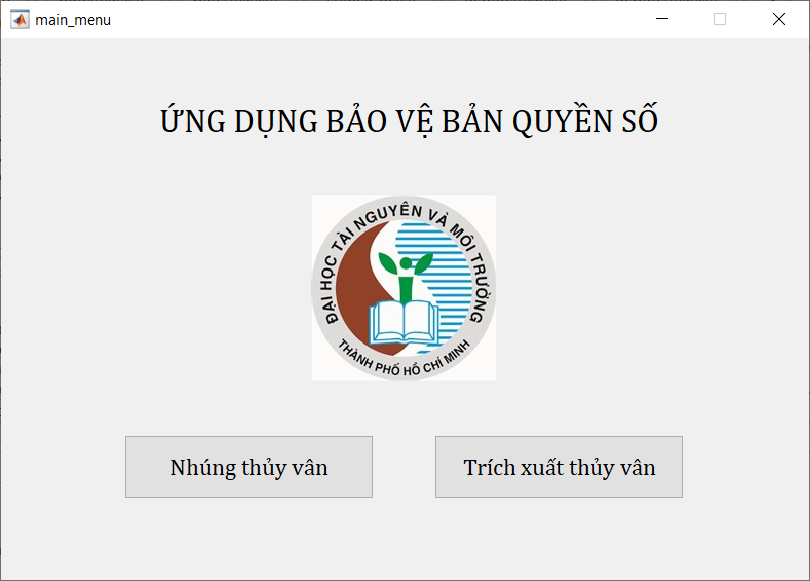
Trong đó, qk (i,j) được sử dụng biểu diễn các hệ số con sóng con của khối con Bk, m là một số nguyên và Q là kích thước bước lượng tử.

* Bước 3: Nếu không xảy ra các biến đổi với hình ảnh chứa thủy vân, thì tất cả các logo thủy vân được trích xuất giống như logo được nhúng ban đầu. Nhưng nếu đã xảy ra bất kỳ biến đổi nào hình ảnh mang, thủy vân được trích xuất nên được kết hợp lại một cách phù hợp để có kết quả cuối cùng. Kết hợp thủy vân được thực hiện theo công thức sau đây :

Trong đó Wk (i,j) được sử dụng để biểu diễn thủy vân được trích xuất từ khối con Bk và W (i,j) được sử dụng để biểu diễn thủy vân được kết hợp. K là số khối con.

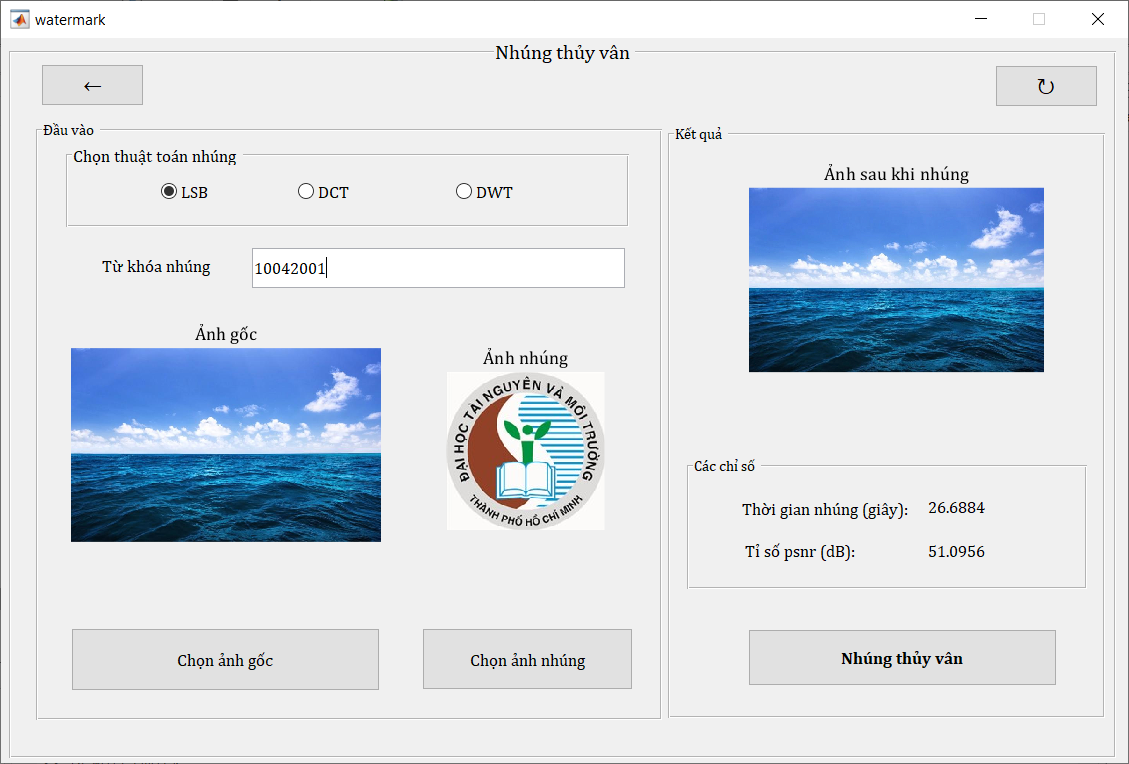
# **CHƯƠNG III. THIẾT KẾ VÀ HIỆN THỰC**

* 1. **Mô tả chức năng hệ thống**
* Chức năng nhận ảnh gốc
* Chức năng nhận ảnh cần thủy vân
* Chức năng nhúng thủy vân
* Chức năng trích xuất thủy vân
* Chức năng thủy vân bằng phương pháp LSB
* Chức năng thủy vân bằng phương pháp biến đối DWT
* Chức năng thủy vân bằng phương pháp DCT
  1. **Cài đặt và thực nghiệm**
     1. **Giao diện chính phần mềm thử nghiệm**



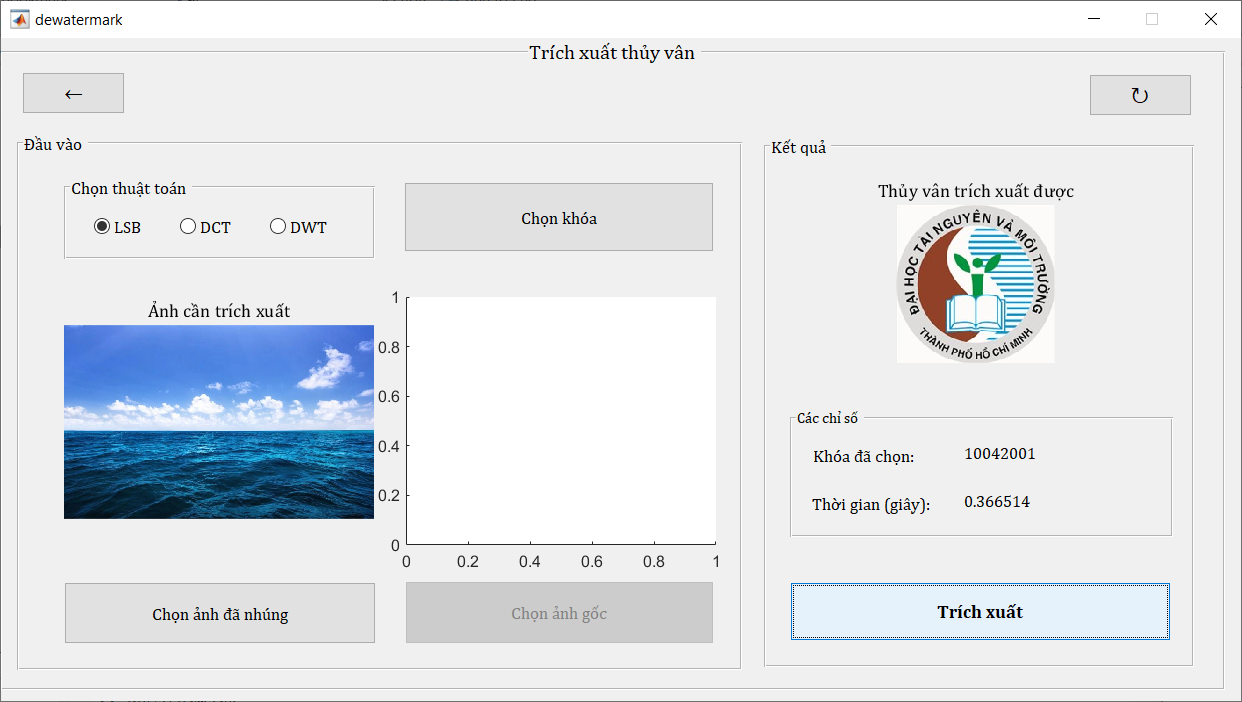
Hình 17. Giao diện chính phần mềm thử nghiệm

* + 1. **Giao diện nhúng ảnh thủy vân bằng phương pháp LSB**



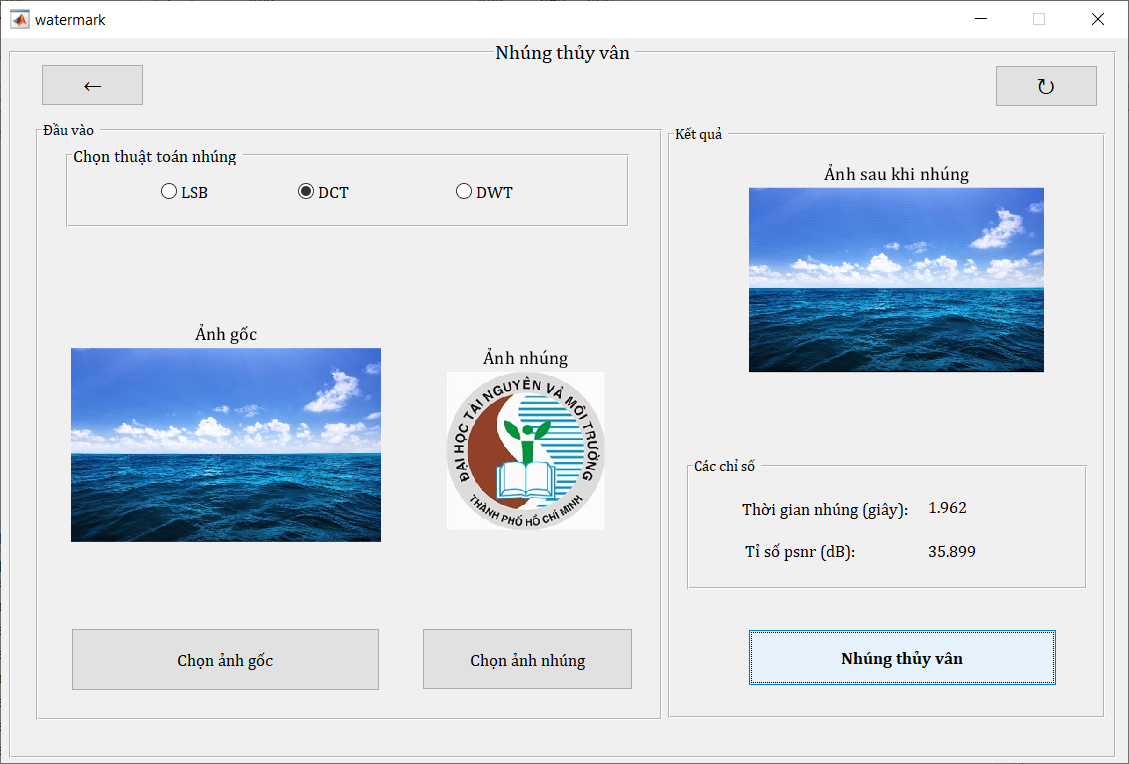
Hình 18. Giao diện nhúng ảnh thủy vân bằng phương pháp LSB

* + 1. **Giao diện trích xuất thủy vân bằng phương pháp LSB**



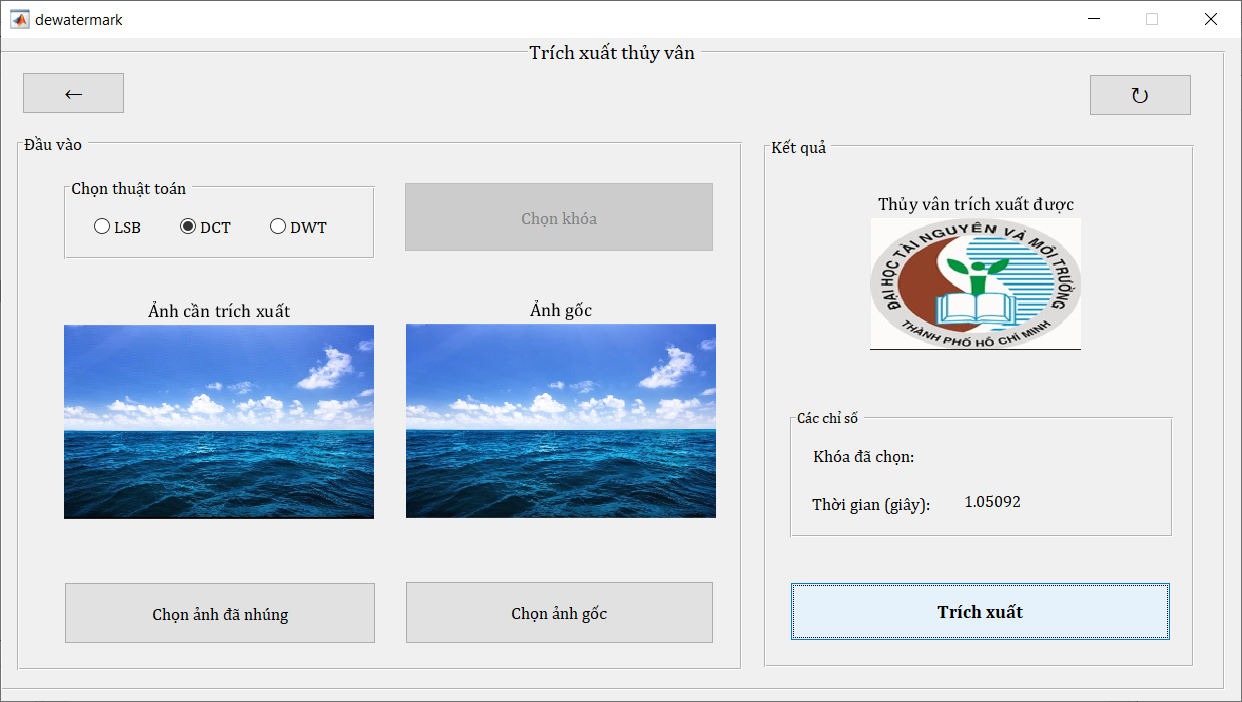
Hình 19. Giao diện trích xuất ảnh thủy vân bằng phương pháp LSB

* + 1. **Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DCT**



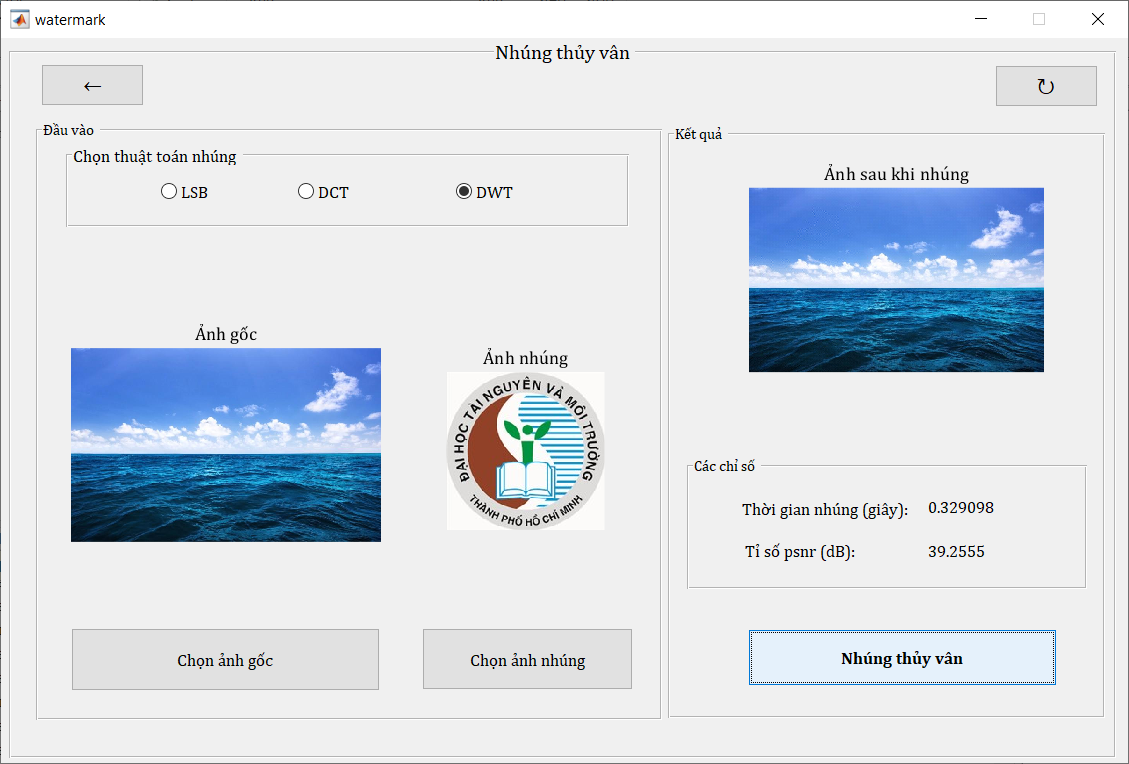
Hình 20. Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DCT

* + 1. **Giao diện trích xuất thủy vân bằng phương pháp biến đổi DCT**



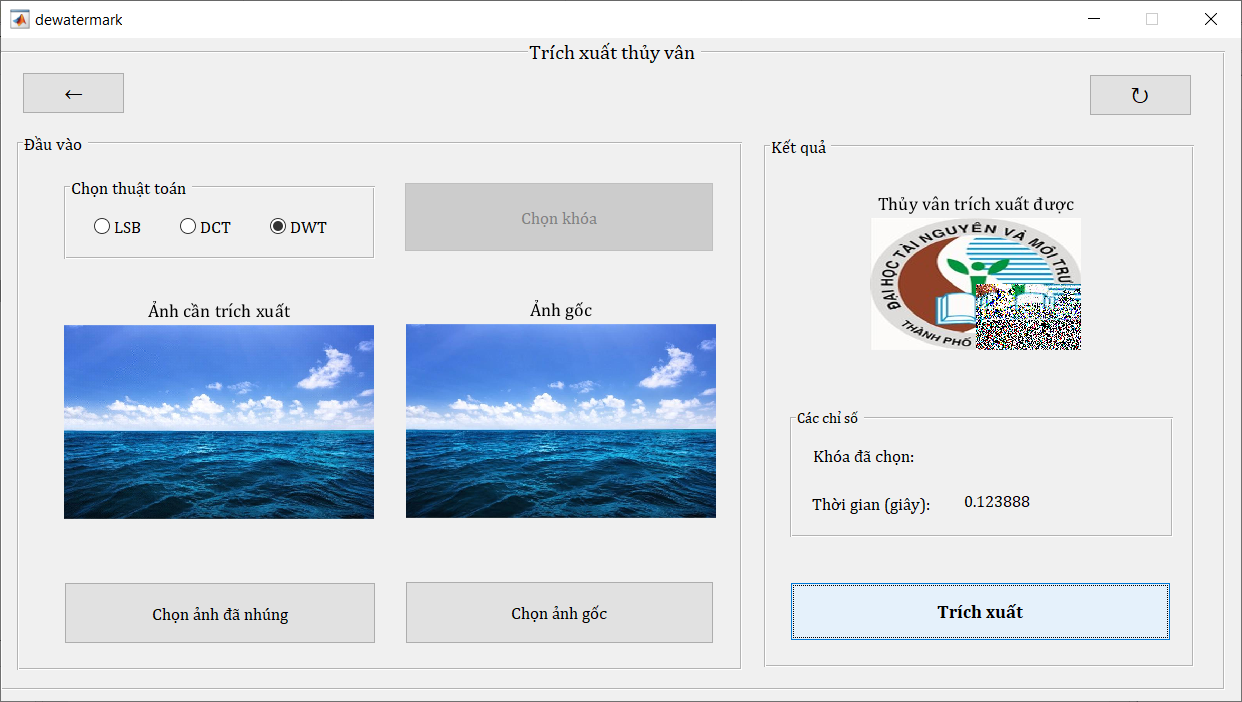
Hình 21. Giao diện trích xuất ảnh thủy vân bằng phương pháp biến đổi DCT

* + 1. **Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DWT**



Hình 22. Giao diện nhúng ảnh thủy vân bằng phương pháp biến đổi DWT

* + 1. **Giao diện trích xuất thủy vân bằng phương pháp biến đổi DWT**

****

Hình 23. Giao diện trích xuất ảnh thủy vân bằng phương pháp biến đổi DWT

# **CHƯƠNG IV. KẾT LUẬN**

Sau một thời gian nghiên cứu và tìm hiểu cùng với sự nỗ lực của cả nhóm, nhóm chúng em đã hoàn thành đồ án của mình. Nội dung chủ yếu của đồ án 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.

Qua quá trình tìm hiểu nghiên cứu đồ án đã đạt được một số kết quả như sau:

* Tổng hợp nghiên cứu về hệ thống thủy vân khái niệm, phân loại, ứng dụng, mô hình, các khả năng tấn công, yêu cầu đối với phương pháp thủy vân.
* Sử dụng 3 kỹ thuật LSB, DCT, DWT với đầu vào là một ảnh đen trắng và ảnh nhúng là một ảnh. Kết hợp phương pháp tấn công gây nhiễu nhằm so sánh tính bền vững của ảnh thủy vân trích xuất đối với các thuật toán.
* Các kết quả nhận được cho thấy được mô hình thủy vân xây dựng đã thành công khi nhúng một ảnh mang vào một ảnh đầu vào mà không làm thay đổi chất lượng hình ảnh.

Do vậy, đồ án hoàn toàn có tính khả thi, có ý nghĩa quan trọng trong việc bảo vệ bản quyền, sở hữu trí tuệ , không những áp dụng được trong thương mại mà còn mở ra một tiếp cận mới cho vấn đề bảo vệ bản quyền ảnh số.

# **TÀI LIỆU THAM KHẢO**

[1] Cường, N. L. (2012). *Nghiên cứu giải pháp nâng cao chất lượng thủy vân sử dụng biến đổi cosine rời rạc trên ảnh*.

[2] Hoan, P. N. (năm 2006). *Xử lý ảnh.* Học viện công nghệ Bưu Chính Viễn Thông.

[3] Hương, L. T. (2014). *Nghiên cứu và phát triển kỹ thuật thủy vân cơ sở dữ liệu quan hệ.* Được truy lục từ tailieudientu.

[4] Hương, N. D. (2012). *Kỹ thuật che giấu tập tin trên K bit LSB của ảnh.* Trường đại học Dân Lập Hải Phòng.

[5] Huy, N. X., & Dũng, T. *Một thuật toán thủy vân ảnh trên miền DCT.*

[6] Nguyễn Xuân Huy, T. Q. (2002). *Một thuật toán thủy vân ảnh trên miền DCT JPEG*.

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

[8] Shen Tao, X. D. (2009). *WATERMARKING GIS DATA FOR DIGITAL MAP COPYRIGHT.*

[9] Uyên, T. T. (2017). *Luận văn thạc sĩ hệ thống thủy văn số và ứng dụng thủy vân số trong bảo vệ ảnh số.*