**1. Chương 1: Tổng quan về thủy vân số**

**1.1 Vấn đề về bản quyền số hiện nay**

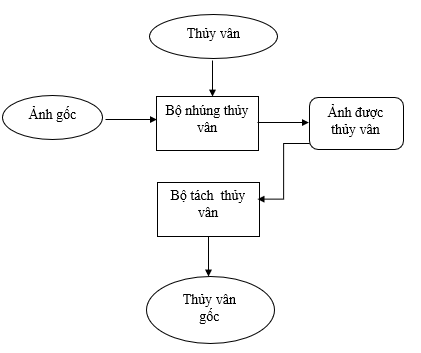
Ngày nay, sự phát triển mạnh mẽ của công nghệ thông tin đã mở ra một kỷ nguyên số hóa mới. Chính vì thế, việc mua bán, lưu trữ, chia sẻ các sản phẩm truyền thông đa phương tiện như video và hình ảnh dưới dạng số hóa cũng đã và đang trở nên vô cùng tiện lợi và nhanh chóng. Đây là một bước tiến lớn của khoa học kỹ thuật mà con người tạo ra. Nhưng cũng vì lý do trên mà dữ liệu số có thể dễ dàng được sao chép, lưu trữ và chỉnh sửa tùy ý. Việc trao đổi các sản phẩm kỹ thuật số này cũng ngày càng trở nên nhanh chóng, đơn giản. Điều này dấy lên những thách thức về bảo mật thông tin, chống xâm phạm quyền sở hữu trí tuệ một cách cấp bách hơn bao giờ hết. Nhu cầu được bảo vệ bản quyền và sở hữu trí tuệ các sản phẩm kỹ thuật số cũng đã và đang trở thành vấn đề quan trọng và được quan tâm. Hiện nay, có hàng tỉ bức ảnh số được tải lên trên mạng Internet, trong đó có cả những bức ảnh là sản phẩm bản quyền với mục đích thương mại. Vì chúng có đặc tính là dễ dàng bị sao chép, chỉnh sửa nên nhiều đối tượng lợi dụng việc này để cố ý đánh cắp, sao chép giả mạo bức ảnh gốc.

Thủy vân số ra đời như một giải pháp để giải quyết các vấn đề về bản quyền số. Đây là một phương pháp nhúng thông tin mật vào trong dữ liệu số của ảnh với mục đích là để khẳng định quyền sở hữu của tác giả của tác phẩm đó mà không gây ảnh hưởng đến chất lượng của sản phẩm. Thủy vân khó có thể bị phát hiện bởi trực giác đơn thuần mà cần có thuật toán để có thể tách được hoàn toàn chúng ra khỏi sản phẩm. Đây có thể được coi là bằng chứng để chứng minh được quyền sở hữu bản quyền của người sở hữu tác phẩm số cũng như phát hiện ra dữ liệu số có bị xâm hại bản quyền hay không.

* 1. **Tổng quan về kỹ thuật giấu tin và thủy vân số**
     1. ***Giới thiệu về thủy vân số***

Thủy vân thường là những ký hiệu hoặc thông tin mật được thêm vào chất liệu giấy trong quá trình sản xuất, các ký hiệu này được sử dụng để đại diện cho hãng sản xuất. Thủy vân bắt nguồn từ việc con người dùng một chất hóa học làm mực vô hình để viết lên trên giấy, loại mực này chỉ hiện lên khi hơ vào lửa hoặc nhúng xuống nước. Trong lịch sử thì thủy vân đầu tiên được ra đời ở Italia vào thế kỷ thứ 13. Đến thế kỷ 18, thủy vân xuất hiện ở Mỹ và Châu Âu và được sử dụng trong lĩnh vực tiền tệ và giao dịch. Cho đến nay, thủy vân vẫn được sử dụng trong lĩnh vực in tiền giấy và bảo mật cho tiền tệ, chống làm tiền giả.

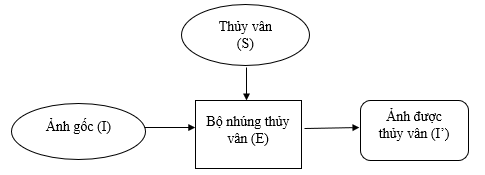
Thủy vân số là các thông tin mật được nhúng vào các phương tiện truyền thông như ảnh số, video, văn bản,… Thủy vân số dần trở nên phổ biến và được nghiên cứu rộng rãi ở những năm 90 của thế kỷ trước. Sự phát triển của công nghệ số như một cuộc cách mạng về khoa học kỹ thuật đã dấy lên nhu cầu về việc bảo vệ bản quyền nội dung số như đã nói ở trên. Thủy vân số nói cách khác là bảo vệ quyền sở hữu nội dung số

**

*Hình 1.1: Sơ đồ tổng quan về một hệ thống thủy vân số*

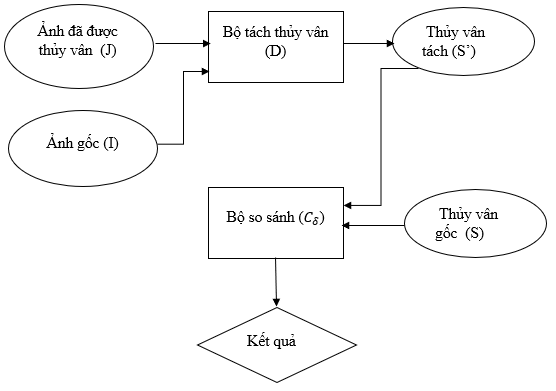
Hệ thống thủy vân số là quá trình nhúng thủy vân (còn được gọi là chữ ký số) vào trong dữ liệu số để dữ liệu số đó trở thành dữ liệu bản quyền. Thủy vân sau đó còn có thể được tách ra với mục đích để đánh giá và xác thực đối tượng. Đối tượng ở đây có thể là ảnh số, video, audio, văn bản,… Nói cách khác, trong hệ thống này, ta có một watermark được gắn vào một ảnh để ra được một sản phẩm được đánh dấu bản quyền theo phương trình toán học tổng quan như sau:

|  |  |  |
| --- | --- | --- |
|  |  |  |

**

*Hình 1.2: Sơ đồ quá trình nhúng thủy vân*

Trong đó E là thuật toán sử dụng để nhúng thủy vân. Với quá trình tách watermark, ta chỉ cần đảo ngược lại thuật toán E để có thể lấy được thủy vân ra khỏi ảnh gốc. Đối với quá trình tách thủy vân, ta có sơ đồ mô hình tổng quan như sau:

**

*Hình 1.3: Sơ đồ tổng quan quá trình tách thủy vân*

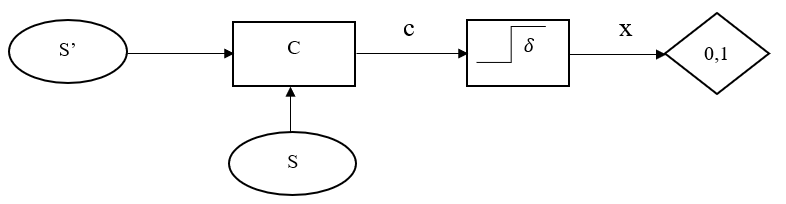
Trong đó, đầu vào là ảnh đã được thủy vân và ảnh gốc được cho vào bộ tách sử dụng thuật toán . Trong quá trình này, ảnh gốc được thêm vào như là phiên bản gốc của khi chưa được thủy vân. Mục đích của việc làm này là để tăng độ tin cậy của thủy vân bởi vì có một số thuật toán nhúng có thể tận dụng ảnh gốc trong quá trình nhúng thủy vân để tránh sự mất mát giá trị của các điểm ảnh. Ta có phương trình toán học tổng quan như sau:

|  |  |  |
| --- | --- | --- |
|  |  |  |

Thủy vân được lấy ra khỏi bộ tách sau đó sẽ được so sánh với thủy vân gốc bởi một bộ so sánh . Ở đây delta chính là ngưỡng (threshold). Kết quả của quá trình so sánh sẽ là một số nhị phân. Nếu nó là 1 thì tức là và là giống nhau và là 0 sẽ là ngược lại:

|  |  |  |
| --- | --- | --- |
|  |  |  |

Hình 1.4 dưới đây là minh họa chi tiết hơn về bộ so sánh giữa và . Ta có là sự tương ứng giữa và và là một ngưỡng nào đó để kết luận rằng và có là một hay không. Ở đây, là thuật toán dùng để so sánh trong bộ so sánh. Một hệ thống thủy vân thường được tóm tắt bằng một bộ có 3 phần tử

**

*Hình 1.4: Minh họa bộ so sánh giữa S và S’*

Dựa vào các thuật toán khác nhau mà thủy vân có thể được nhúng vào ảnh theo nhiều cách khác nhau và phương pháp tiếp cận để nhúng thủy vân cũng thường là không giống nhau. Chính vì vậy mà quá trình lấy thủy vân ra khỏi ảnh cũng là khác nhau. Trong một số trường hợp, chúng ta chỉ có thể biết được liệu có thủy vân được nhúng trong ảnh hay không thông qua một thuật toán dò thủy vân. Trong trường hợp này, bộ dò thủy vân chỉ có thể xác nhận được ảnh là bản quyền hay không còn bộ tách thủy vân sẽ chứng thực được bản quyền này thuộc về bên nào.

Thủy vân số có một số đặc tính thường được thảo luận như: tính phức tạp, tính trung thực, hình ảnh, độ tin cậy phát hiện, tính bền vững, … Trong thực tế, để thiết kế được một hệ thống thủy vân đảm bảo được tất cả các thuộc tính trên là điều gần như không thể. Do đó, việc đảm bảo 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. Dưới đây, ta sẽ đi vào phân tích một số đặc tính quan trọng của thủy vân số.

1. *Tính bền vững*

Hình ảnh được thủy vân có thể phải trải qua nhiều quá trình xử lý ví dụ như quá trình chuyển đổi ảnh số sang tương tự và ngược lại; nén tổn hao, … Chính vì thế, thủy vân phải có tính bền vững thì mới nguyên vẹn sau khi thực hiện các phép biến đổi đó. Ngoài ra, ảnh chứa thủy vân cũng phải chịu được các phép biến đổi hình học như di chuyển vị trí, co dãn kích thước và cắt ghép.

Một thủy vân được coi là bền vững nếu nó vẫn toàn vẹn sau khi ảnh gốc bị chỉnh sửa và biến đổi mà không làm chất lượng của ảnh gốc giảm cho tới một mức nhất định nào đó.

Thủy vân có thể được nhúng trong hình ảnh bằng cách điều chỉnh các giá trị điểm ảnh. Đây được gọi là biến đổi miền không gian. Tuy nhiên, đây được coi là phép nhúng không bền vững bởi vì thủy vân dễ bị xâm hại nếu bị chỉnh sửa các điểm ảnh thông qua nén tổn hao. Vào năm 1997, tác giả Cox et.al đưa ra một bài báo nghiên cứu về “Thủy vân dựa trên trải phổ bảo vệ dữ liệu đa phương tiện” trong đó chỉ ra thuật toán nhúng thủy vân vào miền tần số của hình ảnh. Đây được coi là phương pháp bảo mật cao hơn phương pháp trước đó là miền không gian. Chi tiết về các phương pháp này sẽ được đề cập ở các phần tiếp theo của chương 1 và 2.

1. *Tính ẩn*

Tính ẩn của thủy vân là khả năng khó bị phát hiện ra bởi mắt người và khả năng không làm giảm chất lượng của ảnh sau khi thực hiện nhúng. Một hệ thống thủy vân hiệu quả sẽ làm cho thủy vân trở nên vô hình trên ảnh để đảm bảo 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 thủy vân hiện có thể được sử dụng làm biểu tượng xác thực nguồn gốc của sản phẩm mà tác giả cố tình muốn để lộ ra cho người xem biết.

1. *Tính bảo mật*

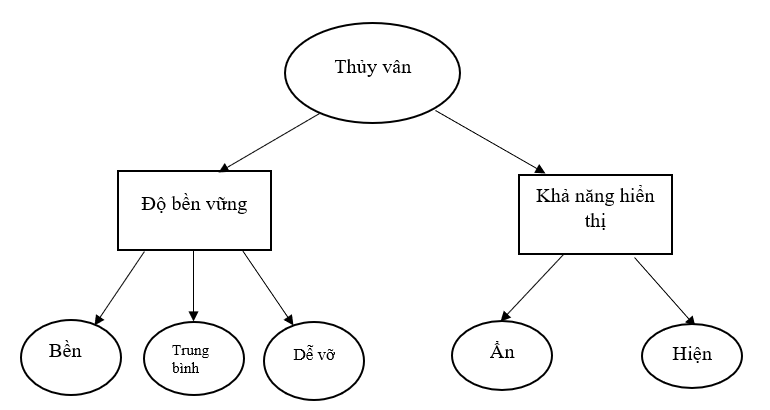
Tính bảo mật của thủy vân được thể hiện khi chỉ có tác giả của thủy vân mới biết được thuật toán để chỉnh sửa và tách thủy vân.

1. *Dung lượng*

Thuật toán nhúng thủy vân luôn mong muốn cho phép giấu càng nhiều thông tin càng tốt mà không làm giảm chất lượng của ảnh gốc. Việc giấu thủy vân trong ảnh khiến cho chúng ta bát buộc phải thay đổi dữ liệu của ảnh. Ta có thể tăng tính bền vững cho thủy vân bằng cách tăng dung lượng đơn vị cần giấu nhưng phải cân bằng sao cho hợp lý để ảnh không bị thay đổi quá nhiều khiến cho tính ẩn của thủy vân không còn được đảm bảo. Điều này sẽ góp phần dẫn đến giảm tính bền vững của thủy vân và chất lượng của ảnh gốc.

* + 1. ***Phân loại thủy vân số***

Dựa trên từng tiêu chí phân loại mà người ta chia thủy vân thành nhiều loại khác nhau như sau

**

*Hình 1.5: Phân loại thủy vân số*

1. *Phân loại theo độ bền vững*

Thủy vân dễ vỡ (fragile watermark): Đây là loại thủy vân được sử dụng bởi những thuật toán khá đơn giản và dễ dàng bị tấn công bởi sự chỉnh sửa, cắt ghép ảnh khiến cho thủy vân bị phá hủy trước khi đến tay người dùng

Thủy vân trung bình (semi-fragile watermark): Thủy vân có thể chống lại được một số kỹ thuật tấn công đơn giản.

Thủy vân bền vững (robust watermark): Đây là loại thủy vân có thể kháng lại nhiều kỹ thuật tấn công, chỉnh sửa để giữ cho thủy vân được nguyên vẹn nhất có thể trước khi đến tay người dùng.

1. *Phân loại theo khả năng hiển thị*

Thủy vân ẩn (invisible watermark): Đây là loại thủy vân được nhúng vào ảnh gốc nhưng lại ẩn đi. Đối với loại thủy vân này thì mắt người khó có thể nhận biết được mà phải dùng những công cụ hoặc phần mềm trong đó có những thuật toán để phát hiện cũng như tách thủy vân ra khỏi ảnh gốc.

Thủy vân hiện (visible watermark): Đây là loại thủy vân hiện hữu trên sản phẩm mà bằng trực giác người xem dễ dàng thấy được vị trí của thủy vân ở ảnh gốc.

Dưới đây là một số ví dụ về thủy vân ẩn và hiện:

**

*Hình 1.6: Ví dụ về thủy vân hiện*



*Hình 1.7: Ví dụ về thủy vân ẩn*

1. *Phân loại theo kỹ thuật thực hiện*

Thủy vân trên miền không gian (spatial domain): Đây là kỹ thuật nhúng thủy vân trực tiếp vào các điểm ảnh.

Thủy vân trên miền tần số (frequency domain): Đây là kỹ thuật nhúng thủy vân vào vùng biến đổi tần số của ảnh.

* 1. **Khả năng phát triển ứng dụng của thủy vân số**
     1. ***Theo dõi phát sóng***

Hiện nay có rất nhiều chương trình tốn kém chi phí để sản xuất cũng như phát sóng trên truyền hình cũng như Internet: thể thao, phim ảnh, giải trí,quảng cáo,… Xuất phát từ yêu cầu bản vệ quyền sở hữu trí tuệ và theo dõi các sản phẩm giải trí được phát sóng trên truyền hình, nhà sản xuất và các bên liên quan cần một hệ thống giám sát thời lượng phát sóng và nội dung được phát sóng trên các phương tiện thông tin đại chúng của các nhà đài được liên kết. Hệ thống theo dõi chủ động được phát triển có áp dụng kỹ thuật thủy vân thì thủy vân sẽ tồn tại bên trong nội dung tín hiệu phát sóng. Từ đó, các công ty truyền thông và các nhà đài có thể quản lý và ngăn chặn được các hoạt động phát sóng lại một cách bất hợp pháp.

* + 1. ***Xác nhận nội dung***

Các tác phẩm kỹ thuật số ngày nay đứng trước nguy cơ bị làm giả về nội dung. Một giải pháp sử dụng chữ ký trực tiếp vào tài liệu dùng kỹ thuật thủy vân. Dấu thủy vân sẽ được nhúng vào một tác phẩm sau đó được lấy ra và so sánh với dấu thủy vân ban đầu. Chữ ký được thiết kế sao cho dù chỉ là sai lệch nhỏ nhất thì chứng tỏ rằng sản phẩm gốc đã bị xâm nhập và có thể bị làm giả. Trong thực tế, người ta mong muốn tìm được vị trí bị giả mạo trên sản phẩm gốc. Chính vì vậy, thủy vân cần phải có khả năng giấu tin cao và bền vững.

* + 1. ***Dấu vân tay và dán nhãn***

Thủy vân trong ứng dụng này được sử dụng để nhận diện người gửi hay người nhận một thông tin nào đó. Các thủy vân khác nhau sẽ được nhúng vào các bản copy khác nhau của thông tin gốc trước khi gửi cho nhiều người. Việc này rất hữu dụng trong việc giám sát và theo dõi các bản sao được sản xuất bất hợp pháp. Nó tương tự như số seri trong sản phẩm phần mềm, mỗi một sản phẩm sẽ mang một số seri riêng. Yêu cầu của thủy vân trong những ứng dụng này là phải bền vững, tránh khả năng bị xâm nhập và xóa dấu vết khi phân phối

* + 1. ***Giấu tin***

Do thủy vân là một dạng đặc biệt của giấu tin nên thủy vân có thể được sử dụng để truyền thông tin bí mật. Ở đây để đảm bảo tính bí mật, người gửi sẽ nhúng một mẩu tin mật vào sản phẩm kỹ thuật số và gửi đi, người nhận sẽ sử dụng thuật toán ngược (còn được gọi là key) để tách mẩu tin đó ra khỏi sản phẩm.

Như vậy, ở chương 1 này, em đã trình bày về các vấn đề liên quan đến bản quyền số hiện nay. Đồng thời, em cũng giới thiệu tổng quan về kỹ thuật thủy vân số nói chung với mục tiêu giải quyết được các vấn đề về bảo vệ bản quyền tác giả trong thời đại kỹ thuật số và đề cập đến việc phát triển ứng dụng mở rộng của thủy vân có thể được áp dụng vào thực tế. Chi tiết về các kỹ thuật được sử dụng trong quá trình nhúng thủy vân số vào sản phẩm kỹ thuật số sẽ được trình bày cụ thể hơn ở chương 2