ĐẠI HỌC QUỐC GIA TP.HCM

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

******

**Môn học**

**ẨN THÔNG TIN TRÊN DỮ LIỆU SỐ VÀ ỨNG DỤNG**

**Đề tài:**

**ĐÁNH DẤU DỮ LIỆU TRÊN HỆ THỐNG THÔNG TIN ĐẠI LÝ:**

**ẨN THÔNG TIN TRONG BẢN ĐỒ VECTOR 2D**

**Giảng viên:** TS. DƯƠNG MINH ĐỨC

**Học viên thực hiện:**

TRẦN THÁI BÌNH - CH1702015

NGUYỄN QUANG ĐẠT - CH1702026

LÊ GIAI TỰ - CH1702048

***TP. Hồ Chí Minh, tháng 08 năm 2019***

ĐẠI HỌC QUỐC GIA TP.HCM

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

******

**Môn học**

**ẨN THÔNG TIN TRÊN DỮ LIỆU SỐ VÀ ỨNG DỤNG**

**Đề tài:**

**ĐÁNH DẤU DỮ LIỆU TRÊN HỆ THỐNG THÔNG TIN ĐẠI LÝ:**

**ẨN THÔNG TIN TRONG BẢN ĐỒ VECTOR 2D**

**Giảng viên:** TS. DƯƠNG MINH ĐỨC

**Học viên thực hiện:**

TRẦN THÁI BÌNH - CH1702015

NGUYỄN QUANG ĐẠT - CH1702026

LÊ GIAI TỰ - CH1702048

***TP. Hồ Chí Minh, tháng 08 năm 2019***

Mục lục

[I. Giới thiệu 4](#_Toc21263169)

[1) Sự cần thiết của ẩn thông tin trong dữ liệu hệ thống địa lý 4](#_Toc21263170)

[2) Đặc điểm của ẩn dữ liệu trong bản đồ vector 2D 5](#_Toc21263171)

[3) Cách tấn công 7](#_Toc21263172)

[II. Thực hiện watermarking loại mạnh mẽ 8](#_Toc21263173)

[1) Trong miền không gian: 8](#_Toc21263174)

[a) Dựa theo mối hệ cơ bản của các đỉnh. 8](#_Toc21263175)

[b) Dựa trên thống kê 10](#_Toc21263176)

[c) Dựa trên điểm tương quan 11](#_Toc21263177)

[2) Trong miền chuyển đổi: 12](#_Toc21263178)

[a) Lược đồ tên miền DFT 12](#_Toc21263179)

[b) Lược đồ tên miền DWT 14](#_Toc21263180)

[c) Lược đồ miền lưới phổ 15](#_Toc21263181)

[3) Phương pháp bảo quản hình dạng 15](#_Toc21263182)

[a) Động lực 15](#_Toc21263183)

[b) Trích xuất dữ liệu 17](#_Toc21263184)

[c) Nhúng dữ liệu 18](#_Toc21263185)

[d) Phát hiện hình mờ 19](#_Toc21263186)

[e) Xác định ngưỡng 20](#_Toc21263187)

[f) Hiệu suất 20](#_Toc21263188)

[4) Đánh giá 23](#_Toc21263189)

[III. Thực hiện ẩn dữ liệu có thể đảo ngược 23](#_Toc21263190)

[1) Lược đồ bổ sung biên: 24](#_Toc21263191)

[2) Lược đồ mở rộng những điểm khách biệt: 25](#_Toc21263192)

[a) Ý tưởng 25](#_Toc21263193)

[b) Hình thành ý tưởng 26](#_Toc21263194)

[c) Các điều kiện để thực hiện nhúng 26](#_Toc21263195)

[d) Nhúng và trích xuất 28](#_Toc21263196)

[e) Thực nghiệm 30](#_Toc21263197)

[3) Đánh giá 32](#_Toc21263198)

[IV. Thực hiện watermarking fragile 32](#_Toc21263199)

[1) Quy trình ẩn và nhúng dữ liệu: 32](#_Toc21263200)

[2) Xác thực bản đồ vector bằng các sử dụng watermark fragile trích xuất thông điệp nhúng: 34](#_Toc21263201)

[a) Quá trình nhúng 34](#_Toc21263202)

[b) Trích xuất và xác thực 35](#_Toc21263203)

[3) Kết quả thực nghiệm 36](#_Toc21263204)

[V. Tóm tắt và hướng phát triển 37](#_Toc21263205)

[1) Watermarking mạnh mẽ chống lại các cuộc tấn công 38](#_Toc21263206)

[2) Độ trung thực: Lỗi, Hình dạng, Cấu trúc liên kết 38](#_Toc21263207)

[3) Đánh giá độ trung thực 39](#_Toc21263208)

[4) Trích xuất và watermarking fragile. 39](#_Toc21263209)

[VI. Tài liệu tham khảo 39](#_Toc21263210)

# Giới thiệu

Sự phát triển nhanh chóng của truyền thông máy tính và kỹ thuật Internet làm cho nó rất dễ dàng để trao đổi dữ liệu qua mạng. Tương ứng, nó cũng trở nên khó khăn hơn để bảo vệ bản quyền kỹ thuật số của các phương tiện kỹ thuật số khác nhau. Kỹ thuật thủy ấn cung cấp một giải pháp tiềm năng cho vấn đề này. Watermarking cũng có thể đóng nhiều vai trò khác như giao tiếp ẩn, xác thực dữ liệu, theo dõi dữ liệu (dấu vân tay). Tập dữ liệu có thể được được sử dụng làm dữ liệu bao phủ của hình mờ, ví dụ hình ảnh kỹ thuật số, âm thanh, video, văn bản, mã vạch, mô hình 3D, dữ liệu CAD, dữ liệu đa giác 2D, phần mềm và VLSI. Trong số các loại dữ liệu này, hình mờ cho một số dữ liệu đa phương tiện chung, chẳng hạn như hình ảnh kỹ thuật số, âm thanh, video và mô hình 3D, đã được đưa ra nhiều chú ý hơn các loại dữ liệu khác. Đề tài này tập trung vào các kỹ thuật dữ liệu watermarking Hệ thống thông tin địa (GIS), cụ thể là 2D vector bản đồ watermarking.

## Sự cần thiết của ẩn thông tin trong dữ liệu hệ thống địa lý

Nói chung, dữ liệu GIS có giá trị lớn. Việc có bản đồ vector là tồn nhiều chi phí, bao gồm công cụ, vật chất và lao động. Số hóa và vector dữ liệu gốc cũng thể hiện công việc khó khăn để tạo ra các vectơ cần thiết bản đồ. Do đó, bản đồ số trong GIS thường không thể được sử dụng mà không tốn chi phú. Ứng dụng của GIS phải chia sẻ dữ liệu địa lý của nó trong nhóm người dùng theo nhiều cách. Ví dụ: qua web, hoặc CD. Nó có nghĩa là chủ sở hữu bản đồ phải cung cấp một số bản sao cho người dùng và khách hàng. Điều này làm cho việc bảo vệ các dữ liệu phân tán là một vấn đề quan trọng. Có một số môi trường đặc biệt, trong đó yêu cầu bảo mật bổ sung là cần thiết, đặc biệt là cho việc sử dụng nhạy cảm. Ví dụ, việc áp dụng bí mật bản đồ kỹ thuật số quân sự đòi hỏi khả năng xác thực nguồn dữ liệu bản đồ cũng như xác nhận tính toàn vẹn. Bản đồ vector kỹ thuật số vẫn không thể thay thế bản đồ giấy truyền thống cho đến khi những yêu cầu này có thể được đáp ứng. Watermarking có thể cung cấp giải pháp cho các vấn đề trên. Bản đồ chủ sở hữu có thể nhúng mã sản phẩm vào bản đồ vector trước khi phân phối để bảo vệ bản quyền. Các nguồn của bản đồ quân sự có thể được truy tìm bằng cách ẩn danh tính duy nhất trong chúng. Việc ẩn dữ liệu cũng có thể khả thi phương pháp xác minh tính toàn vẹn hoặc liên lạc bí mật bằng cách sử dụng hàm băm giá trị hoặc thông điệp bí mật cho dữ liệu ẩn. So với dữ liệu dữ liệu đa phương tiện, có tương đối ít công trình nghiên cứu được thực hiện trong khu vực của watermarking bản đồ vector. Theo hiểu biết của chúng tôi, một công ty ở Uruguay có tên The Digital Map Ltd sẽ cấp bằng sáng chế cho chương trình bảo vệ bản quyền cho các bản đồ vector trong đó sử dụng sơ đồ thủy ấn cho dữ liệu đa giác. Một phần mềm thương mại có tên MapSN cũng đã được phát triển cho nhà sản xuất bản đồ số.

## Đặc điểm của ẩn dữ liệu trong bản đồ vector 2D

Cấu trúc dữ liệu và ứng dụng của bản đồ vectơ 2D cũng được phân biệt được đánh dấu từ các hình ảnh pixel cũng như các loại dữ liệu đa phương tiện khác. Điều này dẫn đến sự khác biệt giữa watermarking bản đồ vector và đa phương tiện nói chung trong nhiều khía cạnh. Ví dụ: những điều liên quan đến nguyên tắc cơ bản để nhúng dữ liệu và các tiêu chí để đánh giá độ trung thực. Tất cả các tính năng trên nên được xem xét khi ẩn dữ liệu trong bản đồ vector.

* Bản đồ vector: Cấu trúc dữ liệu và nguyên tắc nhúng cơ bản

Nói chung, bản đồ vectơ được hình thành bởi dữ liệu không gian, dữ liệu thuộc tính và một số dữ liệu bổ sung được sử dụng như chỉ số hoặc mô tả thêm. Dữ liệu không gian bình thường có dạng một chuỗi các tọa độ 2D mô tả địa lý vị trí của các đối tượng bản đồ đại diện cho các đối tượng địa lý trong thế giới thực. Tất cả các đối tượng bản đồ có thể được phân loại thành ba hình học cơ bản các yếu tố. Đó là điểm, polylines và đa giác. Những yếu tố hình học được tạo bởi các đỉnh có tổ chức và dữ liệu không gian là tọa độ của các đỉnh này dựa trên một hệ tọa độ địa lý cụ thể. Tại dữ liệu mô tả các thuộc tính của các đối tượng bản đồ như tên của chúng, phân loại và một số thông tin khác. Rõ ràng là thông tin được ghi lại bởi dữ liệu phân bổ là rất quan trọng và thông thường không thể sửa đổi tùy ý, cũng như các dữ liệu bổ sung khác được đề cập ở trên. Cho đến bây giờ, tất cả các đề án watermarking được đề xuất đã nhúng hình mờ vào dữ liệu không gian của bản đồ bìa, cụ thể là tọa độ của các đỉnh. Mỗi bản đồ vector có dung sai chính xác được ký hiệu là τ. Điều này mang lại cho biên độ cực đại của các biến dạng cho phép đối với tọa độ. Bất kì biến dạng tọa độ chắc chắn dưới sẽ không làm giảm hiệu lực của bản đồ. Dung sai chính xác của bản đồ che phủ đóng vai trò tương tự như hình ảnh mô hình bìa trong watermarking hình ảnh kỹ thuật số. Nó cung cấp một ít dự phòng cho ẩn thông tin thêm. Nguyên tắc cơ bản của hình mờ bản đồ vector là rằng các biến dạng tọa độ gây ra bởi ẩn dữ liệu không được vượt quá dung sai τ.

* Độ trung thực của dữ liệu vector

Một nguyên tắc phổ biến theo sau bởi tất cả các sơ đồ watermarking là nhúng một tin nhắn ẩn sẽ không làm giảm tính hợp lệ của dữ liệu bao phủ. Thuật ngữ độ trung thực thường được sử dụng làm thước đo tính hợp lệ của dữ liệu trong thế giới watermarking. Theo các loại dữ liệu khác nhau và tương ứng của họ tập quán, thuật ngữ trung thực có thể có ý nghĩa khác nhau. Đối với hình ảnh kỹ thuật số, video, âm thanh và các bộ dữ liệu đa phương tiện chung khác, khi sử dụng trực tiếp dữ liệu là sử dụng cơ quan cảm giác của con người. Theo nghĩa đó, mắt người có thể được sử dụng để đo độ trung thực của hình ảnh. Nếu mắt người không thể phân biệt được, hai hình ảnh sau đó có thể được coi là có cùng giá trị sử dụng. Đó là, có cùng một sự trung thực. Một số thông số định lượng được định nghĩa để đo sự khác biệt giữa hai bộ dữ liệu, chẳng hạn như PSNR, MSE, ví dụ. Đối với việc đánh giá độ trung thực của vector dữ liệu bản đồ, cả nhận thức của con người và PSNR đều không thể cung cấp một dữ liệu phù hợp đo lường. Thứ nhất, người sử dụng trực tiếp bản đồ vector không còn là ý nghĩa của con người các cơ quan nhưng là máy tính. Trong một tỷ lệ điển hình, thậm chí hai bản đồ kỹ thuật số khá tương tự khi đánh giá bằng mắt. Vẫn có thể phối hợp khác biệt giữa hai bản đồ có thể vượt quá dung sai chính xác. Thứ hai, thuật ngữ PSNR chủ yếu phản ánh năng lượng của các lỗi. Nó thích hợp hơn nhiều để đánh giá hình ảnh nhưng không phải bản đồ vector. Đó là bởi vì ngay cả một PSNR cao của bản đồ vectơ không thể đảm bảo rằng tất cả các lỗi đỉnh nằm trong phần trước của bản đồ. Khi chúng tôi đánh giá độ trung thực của bản đồ vector, một số khác các yếu tố phải được tính đến. Đó là hình dạng và cấu trúc liên kết của các đối tượng bản đồ không thể được phản ánh trực tiếp bởi PSNR. Nói cách khác, một bản đồ độ trung thực thấp với hình dạng hoặc cấu trúc liên kết rõ ràng cũng có thể có PSNR cao. Do những yếu tố này, rất khó để đánh dấu hoàn hảo bản đồ vector. Hiện tại, không có số đo vector định lượng thích hợp bản đồ trung thực.

## Cách tấn công

Một cuộc tấn công thành công có nghĩa là watermark có thể được loại bỏ trong khi tính hợp lệ của dữ liệu bao phủ có thể được bảo tồn. Không giống như hình ảnh pixel, dữ liệu không gian của bản đồ vector là một chuỗi các giá trị dấu phẩy động với một số nhất định độ chính xác. Do đó, các tấn công có thể trên watermarking bản đồ vector cũng khác với watermarking đa phương tiện nói chung.

* Tấn công hình học

Một số biến đổi hình học như dịch, xoay và chia tỷ lệ là hình thức chính của các cuộc tấn công hình học. Đối với watermarking hình ảnh kỹ thuật số, tấn công hình học rất khó để chống lại bởi vì những biến đổi này cần để cho phép nội suy các giá trị pixel. Đó là một quá trình không thể đảo ngược và luôn luôn gây mất thông tin. Tuy nhiên, đối với bản đồ vector, các cuộc tấn công như vậy là biến đổi hầu như có thể đảo ngược của tọa độ, trong đó hầu như ít hoặc không có thông tin sẽ bị mất. Tấn công hình học là tương đối dễ dàng hơn để bảo vệ chống lại trong sơ đồ watermarking bản đồ vector.

* Tấn công Vertex

Vertex có nghĩa là các cuộc tấn công ở cấp độ đỉnh. Đó là bằng cách thêm mới các đỉnh vào bản đồ (nội suy) hoặc xóa các đỉnh khỏi bản đồ (đơn giản cắt xén). Những cuộc tấn công như vậy có thể làm xáo trộn sự đồng bộ của máy dò bằng cách thay đổi số đỉnh. Do đó, tấn công rất nghiêm trọng watermarking bản đồ vector. Đơn giản hóa bản đồ là một hoạt động phổ biến trong công việc hàng ngày để tăng cường tốc độ xử lý dữ liệu bản đồ. Nó làm làm mất thông tin và để giữ tính hợp lệ của bản đồ được xử lý. Do đó, khả năng sống sót sau khi đơn giản hóa bản đồ là rất quan trọng cho một kế hoạch watermarking mạnh mẽ.

* Sắp xếp lại đối tượng

Sắp xếp lại đối tượng là một cuộc tấn công ở cấp đối tượng. Dữ liệu không gian của một vetor bản đồ chứa nhiều đối tượng được cấu tạo bởi các đỉnh được sắp xếp. Tất cả các đối tượng được lưu trữ trong tệp bản đồ theo một trình tự nhất định. Hoặc là thay đổi lưu chuỗi các đối tượng bản đồ hoặc sắp xếp lại các đỉnh trong một đối tượng có thể tạo một tệp bản đồ mới mà không làm giảm độ chính xác của dữ liệu. Đối với các sơ đồ watermarking phụ thuộc vào thứ tự đối tượng, thao tác này là một cuộc tấn công gây tử vong.

* Tấn công Noise

Chủ yếu có hai nguồn có thể đưa nhiễu vào bản đồ vector. Các đầu tiên là bởi một số loại công việc hàng ngày. Ví dụ, có một số phổ biến định dạng tệp trong thế giới GIS. Việc chuyển đổi giữa các định dạng có thể hơi bóp méo dữ liệu. Cái khác là một cuộc tấn công độc hại. Kẻ tấn công cố gắng để phá hủy hình mờ bằng cách thêm nhiễu vào tập dữ liệu. Biến dạng noise là một cuộc tấn công nghiêm trọng nhưng nhìn chung nó không phải là một lựa chọn tốt của kẻ tấn công bởi vì việc áp đặt noise có thể làm giảm hiệu lực của bản đồ.

# Thực hiện watermarking loại mạnh mẽ

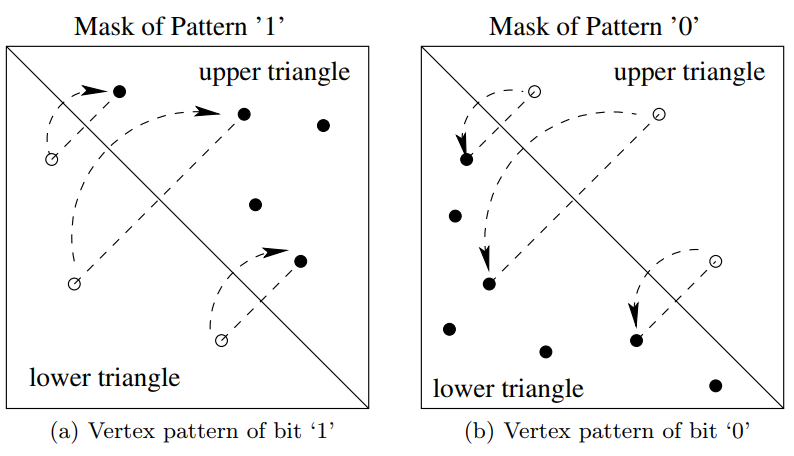
Dựa theo các tạo độ thực thi watermarking, các phương pháp mạnh mẽ trong watermarking vertor bản đồ có thể được chia thành 2 loại. Đó là thuật toán trong miền biến đổi không gian và trong miền chuyển đổi. Trong phần này, một sơ đồ đại diện của mỗi lớp được giới thiệu. Sau đó, một thuật toán bảo toàn hình dạng được trình bày để cải thiện các phương pháp trước đó.

## Trong miền không gian:

Watermarking vector bản đồ trong miền không gian là ẩn thông tin bằng các sửa đổi các giá trị tọa độ của các đỉnh. Các bit ẩn có thể được biểu diễn bằng một số tính năng không gian của các đỉnh. Đó là mối quan hệ vị trí giữa các đỉnh hoặc thuộc tính thống kê của tọa độ.

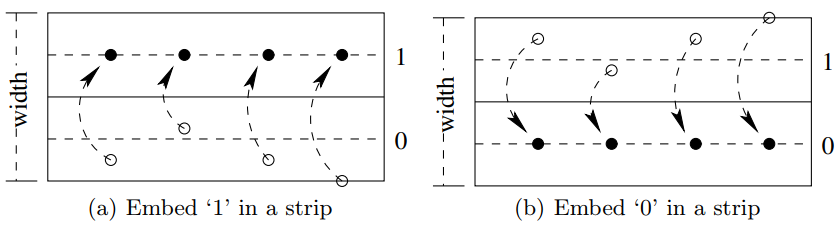
### Dựa theo mối hệ cơ bản của các đỉnh.

Thuật toán được đề xuất bởi Sakamoto [2] là một sơ đồ điển hình dựa trên khái niệm sửa đổi mối quan hệ vị trí giữa các đỉnh. Kang [3] tăng cường sự mạnh mẽ của nó đối với các cuộc tấn công noise. Ý tưởng cơ bản của hai sơ đồ đó là dịch chuyển các đỉnh trong mặt được xác định trước và làm cho mối quan hệ vị trí của chúng tuân theo các mẫu cụ thể đại diện cho 0 hoặc 1. Lấy sơ đồ nâng cao được đề xuất bởi Kang [3] làm ví dụ, hai mẫu đỉnh và quy trình nhúng được hiển thị trong Hình 6.1. Đầu tiên, bản đồ gốc được phân thành các khối và mặt được xác định cho mỗi khối. Tiếp theo, mặt được chia thành một tam giác trên và một tam giác dưới bởi đường chéo của nó được kết nối bởi các đỉnh đông nam và tây bắc. Để nhúng một bit 1, tất cả các đỉnh trong tam giác dưới được chuyển sang vị trí đối xứng của chúng trong tam giác trên đối thông qua đường chéo, như trong Hình 6.1 (a). Để nhúng một chút 0, chỉ là một quá trình nghịch đảo được hiển thị trong Hình 6.1 (b). Dữ liệu ẩn có thể được trích xuất bằng cách tìm ra tam giác nào có hầu hết các đỉnh trong mặt. Kỹ thuật này sẽ trích xuất mù và mạnh mẽ trước tấn công noise và đơn giản hóa.



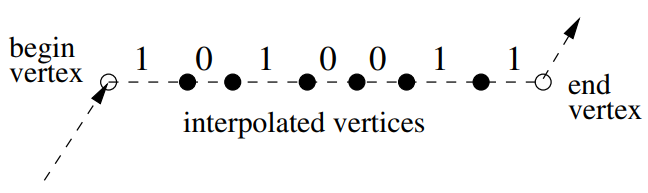
Hình 6.1. Hai mẫu đại diện cho 1 và và 0 (thuật toán trong [3])

Schulz và Voigt [4] đã đề xuất một sơ đồ dung lượng cao, mạnh mẽ cho các cuộc tấn công đơn giản hóa tiếng ồn và đơn giản hóa. Trong sơ đồ, bản đồ gốc được chia thành nhiều dải dọc và ngang. Độ rộng của các dải có thể là để đảm bảo độ dung sai. Các đỉnh trong một dải được chuyển sang đường tham chiếu được sử dụng để biểu diễn bit 1, hoặc 0. Hình 6.2 thể hiện quá trình dịch chuyển. Việc trích xuất bit ẩn sẽ phụ thuộc vào một nửa dải nơi có hầu hết các đỉnh. Thuật toán có thể tồn tại các tạp âm phụ gia với biên độ nhỏ hơn 1/4 dải chiều rộng.



Hình 6.2. Nhúng dữ liệu trong dải

Thuật toán được đề xuất bởi Huber [5] cũng sử dụng mối quan hệ vị trí giữa các đỉnh để ẩn bit. Lược đồ này khác với các đề cập trước đây, trong đó các đỉnh mới được nội suy để nhúng dữ liệu. Hình 6.3 là một ví dụ về sơ đồ này.



Hình 6.3. Một minh họa đơn giản về sơ đồ Huber

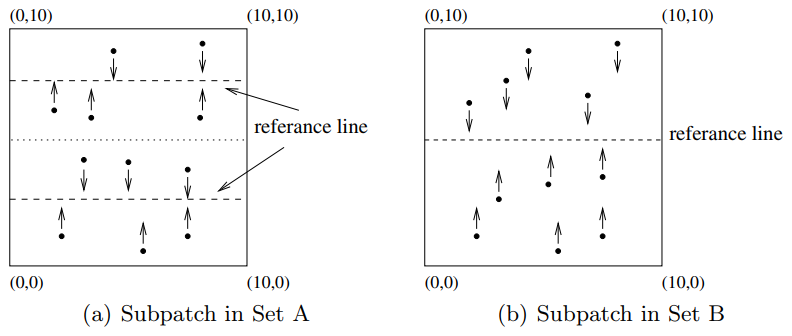
Trong Hình 6.3, đỉnh bắt đầu và đỉnh kết thúc được biểu thị bằng hai vòng tròn màu trắng là hai đỉnh liền kề trong bản đồ gốc. Tất cả các điểm được tô đen được nội suy thành các đỉnh mới thu được trong quá trình nhúng. Giả sử khoảng cách giữa đỉnh bắt đầu và đỉnh nội suy đầu tiên đại diện là bit 1 và một nửa khoảng cách đại diện là bit 0, dữ liệu được nhúng trong đoạn đường sau đó là ‘1010011. Sơ đồ không có biến dạng và nó mạnh mẽ cho các cuộc tấn công hình học. Watermarking rộng kích thước của dữ liệu bìa. Đề án cũng dễ vỡ với việc đơn giản hóa bản đồ.

### Dựa trên thống kê

Trong sơ đồ được đề xuất bởi Voigt và Busch [6], thuộc tính thống kê của tọa độ được sử dụng để ẩn dữ liệu. Một vùng hình chữ nhật trong bản đồ gốc được chọn làm dữ liệu bìa để nhúng bit watermark. Thứ nhất, khu vực được chia thành các bản vá. Sau đó, một chuỗi nhiễu giả được sử dụng để phân chia tất cả các bản vá thành hai tập hợp khác nhau A và B. Tiếp theo, mỗi bản vá trong hai bộ được chia thành các tập hợp con. Đối với mỗi subpatch, tọa độ của các đỉnh trong nó được chuyển thành các giá trị tương đối mới bằng cách lấy góc phía tây nam làm gốc. Nếu không có dữ liệu nhúng, tọa độ tương đối trong tập A và B đều được phân phối đồng đều với các thuộc tính tương tự như kỳ vọng và phương sai. Nếu một nhúng 1 bit, việc sửa đổi các đỉnh được thực hiện trong mỗi subpatch như trong Hình 6.4. Sau khi sửa đổi, phương sai mẫu của các giá trị tọa độ trong bộ A (ký hiệu là ) được tăng lên và giá trị của bộ B (ký hiệu là ) bị giảm. Một biến ngẫu nhiên F = / , phân phối F sau đó được sử dụng để phát hiện dựa trên một ngưỡng.

Bằng cách sử dụng thuộc tính thống kê của tọa độ, thuật toán có được chứng minh là mạnh mẽ đối với nhiều cuộc tấn công như dịch chuyển bản đồ, nội suy dữ liệu, đơn giản hóa, nhiễu phụ gia.

### Dựa trên điểm tương quan



Hình 6.4. Việc sửa đổi các đỉnh trong tập A và B (thuật toán trong [6]).

Ohbuchi [7] trình bày một sơ đồ dựa trên phát hiện tương quan. Pseudo Noise Sequence (PNS) được sử dụng để cải thiện tính bảo mật và độ tin cậy phát hiện trong sơ đồ. Bản đồ bìa được phân đoạn đầu tiên thành các khối hình chữ nhật có chứa số lượng đỉnh. Một khối ẩn một bit watermark. Dành cho một khối tùy ý, một bit ẩn có thể được nhúng bởi công thức. (6.1):



Trong đó và là tọa độ ban đầu và tọa độ watermarking của các đỉnh tương ứng trong khối thứ i. Watermark bit là i thứ tự và là bit thứ i của PNS. Hệ số biên độ là α. Trong quy trình trích xuất hình mờ, cần có bản đồ gốc và cùng một PNS được sử dụng để giải mã, có thể được xem như là một quy trình phát hiện dựa trên mối tương quan. Sơ đồ này mạnh mẽ đối với các loại tấn công tuy nhiên nó không phải là một sơ đồ trích xuất mù. Một kế hoạch watermark với phát hiện thông tin thường không thể được sử dụng thực tế.

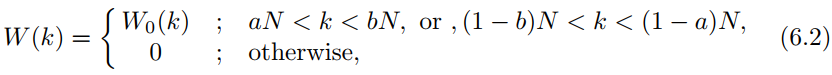
Voigt và Busch [8] đã đề xuất một thuật toán dựa trên phổ trải chuỗi trực tiếp. Nhúng và trích xuất hình mờ được giới hạn ở các vị trí kỹ thuật số thập phân cụ thể của tọa độ thô. Giả sử một bit được nhúng. Đầu tiên bit được thay thế bằng pseudo noise sequence (PNS) để hình thành dữ liệu hình mờ, về bản chất là một quy trình phổ trải chuỗi trực tiếp. PNS sau đó được nhúng vào tọa độ bằng cách sửa đổi 2 giá trị số thập phân thấp hơn của tọa độ. Biên độ của các sửa đổi bằng với dung sai của dữ liệu bản đồ. Trong giai đoạn trích xuất dữ liệu, một mối tương quan được sửa đổi giữa PNS và 2 giá trị số thập phân thấp hơn của tọa độ hình chìm được tính toán và sử dụng để phán đoán. Các thuật toán là mạnh mẽ để tiếng ồn phụ gia. Thêm hoặc loại bỏ các đỉnh có thể phá hủy sự đồng bộ của máy dò.

## Trong miền chuyển đổi:

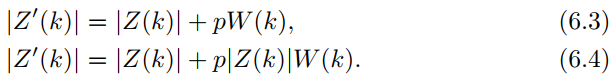
Một hình mờ được nhúng bằng cách sửa đổi các hệ số biến đổi của tọa độ. Các kế hoạch điển hình được thực hiện trong các lĩnh vực sau: DFT, DWT và Miền lưới phổ.

### Lược đồ tên miền DFT

Nikolaidis và cộng sự. [9, 10] đã đề xuất một sơ đồ watermark trích xuất mù bao gồm việc nhúng một bit đơn vào một đa tuyến bằng cách sửa đổi các hệ số Fourier rời rạc của chuỗi tọa độ đa tuyến. Lấy một đa tuyến có N đỉnh làm dữ liệu bìa, một bit ẩn có thể được nhúng như sau. Đầu tiên, cho tạo độ các đỉnh , một chuỗi phức z(k) thu được bằng cách kết hợp tọa độ x và y như sau . Sau khi DFT được thực hiện trên z (k), chúng ta sẽ nhận được chuỗi hệ số {Z (k) | k = 1, 2, · · ·, N}. Thứ hai, bắng cách thể hiện chuỗi Pseudo Noise Sequence . Việc tạo watermark W(k) là một quy trình trải phổ như công thức (6.2)



Trong đó a, b (0 < a < b < 1) được sửa dụng để chọn phảm nhúng trong miền phổ. Từ đó watermark W(k) được nhúng nhờ vào chỉnh biên độ Z(k) bằng cách sử dụng cộng toán tử (Công thức 6.3) hoặc nhân toán tử (Công thứ 6.4).



Để trích xuất dữ liệu ẩn, mối tương quan tuyến tính (ký hiệu là c) của | Z (k) | và W (k) được tính toán. Sau đó, mối tương quan được chuẩn hóa c ′ = c / µc được sử dụng cho một phán đoán dựa trên một ngưỡng, trong đó µc là giá trị trung bình lý thuyết của c. Lược đồ vốn đã mạnh mẽ đối với nhiều loại tấn công như dịch bản đồ, xoay, chia tỷ lệ và bắt đầu dịch chuyển đỉnh trong đa tuyến watermark. Thay đổi số đỉnh như đơn giản hóa hoặc nội suy có thể làm anh hưởng đồng bộ hóa trong quá trình phân tích.

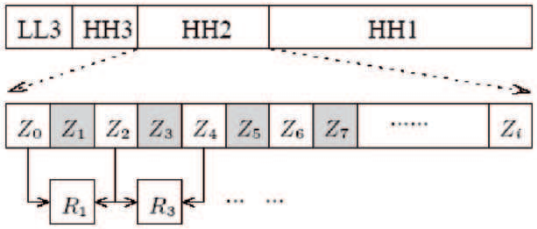
Để cải thiện độ tin cậy của quá trình phân tích, Nikolaidis [11] cũng đề xuất một thuật toán nâng cao trong đó nhiều polylines được sử dụng để tạo watermarking. Các bit ẩn được nhúng trong mỗi đa tuyến bằng một phương thức trong các tác phẩm trước đây của họ [9,10]. Khi phát hiện bit ẩn, mỗi polyline sẽ nhận được giá trị tương quan chuẩn hóa , trong đó M là số lượng polyline. Hàm fusion dữ liệu f (·) sau đó được chọn để tính toán một tham số kết hợp , sẽ được sử dụng cho phán quyết cuối cùng dựa trên một ngưỡng.

Dựa trên các tác phẩm trước đây [9, 10], Kitamura đề xuất một sơ đồ sửa đổi [12] nhúng các bit nhiều hình mờ trong miền DFT. Watermark là một chuỗi bit có ý nghĩa thay vì một bit được đại diện bởi PNS. Quy trình nhúng tương tự như các công trình trước trong khi quy trình trích xuất khác nhau vì bản đồ gốc sẽ là cần thiết. Do đó, nó không phải là một kế hoạch mù quáng và điều này giới hạn ứng dụng của nó.

### Lược đồ tên miền DWT

Li và Xu đã đề xuất một sơ đồ mù [13], bao gồm việc nhúng nhiều bit vào bản đồ vector trong miền DWT. Ý tưởng chính của sơ đồ được mô tả như sau.

1. Tương tự như công trình [9, 10], tọa độ của các đỉnh được kết hợp trước tiên thành một chuỗi phức tạp. Sau đó, một phân tách sóng con ba cấp được thực hiện trên chuỗi, kết quả là bốn băng con, tức là HH1, HH2, HH3, LL3. Xem xét dung sai độ chính xác của bản đồ, HH2 và HH3 được chọn để nhúng dữ liệu.
2. Lấy các hệ số trong HH2 làm ví dụ, phương pháp nhúng được hiển thị trong Hình 6.5.



Các giá trị là hệ số của HH2. Bắt đầu tại , các bit watermarking sẽ được nhúng vào mỗi hệ số. Đó là hệ số các chỉ số lẻ Ví dụ: giả sử nhúng bit watermark đầu tiên vào . Giá trị tham chiếu R được tính bằng . Hai hệ số lân cận và được sử dụng và α là hệ số công suất kiểm soát biên độ nhúng. Tính và sử dụng tính chẵn lẻ của làm đại diện cho một hình mờ. Nếu = 0, sửa đổi | | để làm cho là một số chẵn. Nếu = 1, sửa đổi | | để làm cho là một số lẻ. Các bit watermark khác , , · · · được nhúng vào theo cùng một cách.

1. Việc trích xuất dữ liệu có thể được thực hiện mà không cần sử dụng bản đồ gốc. Phân rã và được tính toán và các bit watermark có thể được trích xuất theo tính chẵn lẻ của từng .

Trong sơ đồ này, biên độ hệ số DWT tọa độ được sửa đổi thành các chế độ cụ thể để biểu diễn các bit watermark. Nó có thể chịu đưng các cuộc tấn công bằng tiếng ồn và một số cuộc tấn công hình học. Giống như thuật toán nội bộ các miền khác, nó rất mong manh đối với các cuộc tấn công đỉnh như đơn giản hóa bản đồ hoặc nội suy.

### Lược đồ miền lưới phổ

Mặc dù ứng dụng của bản đồ vector 2D và mô hình 3D rất khác nhau, nhưng vẫn có nhiều điểm tương đồng giữa hai loại dữ liệu. Chúng là dữ liệu vectơ và tập điểm đỉnh. Có thể áp dụng một số phương pháp watermarking mô hình 3D cho bản đồ vector 2D. Ohbuchi et al. đề xuất một sơ đồ thủy ấn bản đồ vector 2D [14] trong miền phổ lưới. Điều này ban đầu được sử dụng trên các mô hình 3D [15]. Trong thuật toán, một số kỹ thuật trong thế giới 3D có thể được điều chỉnh để xử lý dữ liệu vectơ 2D. Đầu tiên, tất cả các đỉnh trong bản đồ gốc được kết nối để thiết lập lưới 2D bằng cách sử dụng Delaunay Triangulation. Toàn bộ bản đồ được phân chia một cách thích ứng thành nhiều khối bằng cách sử dụng cây quy tắc k-d. Đối với mỗi khối, phân tích quang phổ lưới được thực hiện trên lưới 2D trong khối. Các hệ số quang phổ lưới thu được và có cùng kích thước với chuỗi tọa độ ban đầu trong khối. Trình tự hệ số quang phổ sau đó được coi là dữ liệu che cho hình mờ. Phương pháp nhúng dữ liệu tương tự như công việc [7] được mô tả trong Phần 6.2.1. Và sơ đồ cũng là một sơ đồ không mù. Nhờ sử dụng dữ liệu gốc, sơ đồ này mạnh mẽ đối với nhiều cuộc tấn công như xoay bản đồ, dịch, xoay, nội suy bản đồ, đơn giản hóa, nhiễu phụ gia và cắt xén.

## Phương pháp bảo quản hình dạng

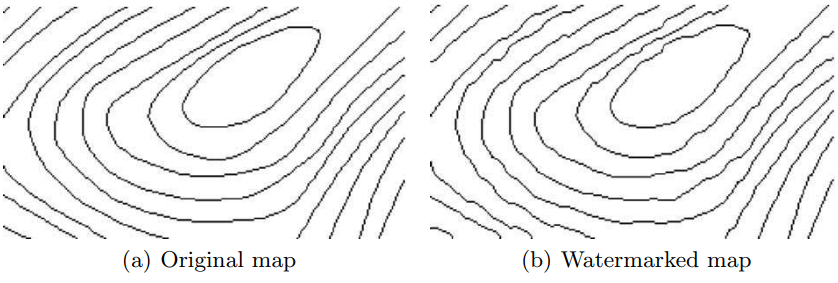
Các thuật toán được đề cập ở trên là các phương pháp điển hình cho các bản đồ vector 2D có watermarking mạnh mẽ trong miền không gian hoặc biến đổi. Bằng cách phân tích sự thiếu sót của các tác phẩm trước đây, một sơ đồ dựa trên thống kê sẽ được trình bày thêm trong phần còn lại của phần này.

### Động lực

Xem xét các khiếm khuyết chính của các công trình trước đây, kế hoạch đề xuất sẽ tập trung vào hai khía cạnh sau đây.

1. Biến dạng của hình dạng và sự lựa chọn dữ liệu bìa.

Một nguyên tắc được sử dụng trong hầu hết các công trình trước đây là để tăng cường sự mạnh mẽ. Điều này được thực hiện bằng cách kiểm soát chặt chẽ sự biến dạng của mọi đỉnh dưới dung sai τ vì mục đích giữ độ trung thực của bản đồ thủy ấn. Tất cả các công trình trước đây đều lấy tọa độ 2D hoặc hệ số tần số của chúng làm dữ liệu che phủ để kiểm soát các biến dạng của các đỉnh. Tiêu chí này là không đủ để đáp ứng các yêu cầu của hình mờ bản đồ vector. Điều này là do hình dạng chi tiết của các đối tượng bản đồ có thể được sửa đổi ngay cả khi độ trung thực của dữ liệu bản đồ được bảo quản tốt. Một thí nghiệm có thể được thực hiện để minh họa vấn đề. Hình 6.6 (a) là một phần mở rộng của bản đồ gốc. Hình 6.6 (b) là phần hình chìm tương ứng của nó. Ở đây tọa độ thô được chọn làm dữ liệu bìa và quy trình nhúng đã không tính đến hình dạng bản đồ. Mặc dù độ méo cảm ứng của mọi đỉnh trong Hình 6.6 (b) nằm dưới dung sai τ, hình dạng cục bộ của các đối tượng trong bản đồ đã được sửa đổi. Hình dạng bị bóp méo làm giảm chất lượng của bản đồ watermarking và khả năng tàng hình của thông điệp. Điều này có thể thu hút sự chú ý của những kẻ tấn công. Nguyên nhân chính của vấn đề là các thuật toán chọn tọa độ hoặc hệ số biến đổi của chúng như dữ liệu bìa. Đây là một hình thức bất tiện để trình bày thông tin hình dạng của các đối tượng bản đồ. Do đó, dữ liệu bìa mới, có thể liên quan trực tiếp đến thông tin hình dạng bản đồ, phải được chọn thay vì tọa độ thô để điều khiển thuận tiện các biến dạng hình dạng do nhúng dữ liệu.



Hình 6.6. Một ví dụ về hình dạng bị bóp méo.

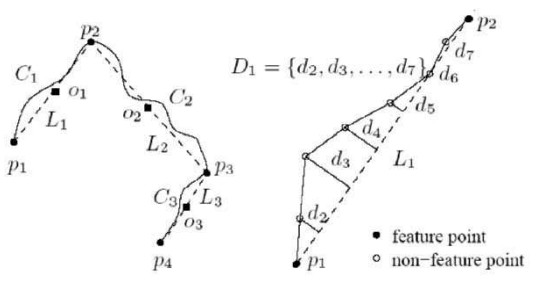
1. Watermarking mạnh mẽ trước cuộc tấn công

Một cuộc tấn công thành công thường phải loại bỏ thông điệp ẩn khỏi dữ liệu watermarking trong khi vẫn giữ được độ trung thực của nó. Do môi trường ứng dụng quan trọng của bản đồ vectơ 2D, ngay cả việc điều chỉnh nhẹ các tọa độ cũng có thể phá hủy độ trung thực của bản đồ. Nó không phải là một lựa chọn tốt cho kẻ tấn công để điều khiển bản đồ watermarking bằng cách xoay, dịch, thu nhỏ, thêm nhiễu ngẫu nhiên, v.v. Tuy nhiên, hầu hết các thuật toán hiện có đều sử dụng các cuộc tấn công này. Một số thao tác có thể thao túng dữ liệu bản đồ mà không mất tính trung thực của nó, bao gồm nội suy bản đồ, đơn giản hóa và sắp xếp lại dữ liệu, nên được coi là các cuộc tấn công nguy hiểm hơn. Điều này sẽ được nhấn mạnh trong sơ đồ sau

### Trích xuất dữ liệu

Để giảm các biến dạng hình dạng, dữ liệu che được sử dụng trong sơ đồ được dự kiến sẽ thể hiện rõ ràng thông tin hình dạng của các đối tượng bản đồ, không thể được thỏa mãn trực tiếp bởi các tọa độ thô. Một chuỗi khoảng cách được trích xuất từ bản đồ gốc và được sử dụng làm dữ liệu bìa mới. Với bản đồ gốc , các điểm đặc trưng của nó được phát hiện. Các này tách thành n tập con được ký hiệu là , · · ·, , · · ·, . Thuật toán Douglas Peucker [16] được sử dụng trong sơ đồ này để phát hiện điểm tính năng dựa trên ngưỡng . Lựa chọn sẽ phục vụ đánh sao sau này. Cho , giả sử tập hợp con chứa các đỉnh được ký hiệu . Rõ ràng là đỉnh đầu tiên và đỉnh cuối cùng là hai điểm đặc trưng liền kề trong tập F và các đỉnh khác là các điểm không có tính năng giữa chúng. Đầu tiên, 2 đỉnh và liên kết với nhau để trở thành phân khúc . Sau đó khoản cách đã được tính toán, cho là khoản cách từ mọi điểm không tính năng tới . Hình dạng cục bộ của có thể liên quan trực tiếp đến tập . Điểm giữa của đoạn (kí hiệu ) cũng được tính để sử dụng trực tiếp. Lập lại chương trình cho tất cả tập con , · · ·, , · · ·, và thu được dữ liệu cuối cùng bao gồm tập khoảng cách . Một tập hợp điểm giữa tương ứng sẽ được trích xuất từ . Tập sẽ được sử dụng dữ liệu bìa cho watermarking. Ưu điểm chính là có liên quan trực tiếp với các hình dạng cục bộ của bản đồ ban đầu và nó là bất biến đối với bản đồ dịch hoặc xoay.

Hình 6.7 trình bày quy trình trên bằng cách sử dụng một đa tuyến đơn giản làm ví dụ.



Hình 6.7. Khai thác dữ liệu dựa trên thuật toán Douglas Peucker.

Như được hiển thị trong phần bên trái của hình, các điểm đặc trưng được phát hiện chia polyline thành ba tập con . Tất cả các điểm tính năng được kết nối lần lượt để tạo thành các phân đoạn dòng . Điều này được hiển thị như một đường chấm chấm. Điểm giữa của các đoạn là . Phần bên phải của Hình 6.7 sau đó chỉ ra quy trình tính khoảng cách bằng cách lấy đoạn . Giả sử có sáu điểm không có tính năng giữa các điểm tính năng và , khoảng cách từ mỗi điểm không có tính năng đến đoạn đường được tính là = . Tính toán hoàn chỉnh của toàn bộ đa tuyến sẽ thu được một chuỗi khoảng cách là = .

### Nhúng dữ liệu

Hình mờ được nhúng bằng cách thay đổi phân phối dữ liệu bìa. Dữ liệu bìa được chia thành hai tập con và theo các bước sau:

1. Chia bản đồ gốc thành các miếng vá có kích thước đồng đều, sử dụng nắp lưới hình chữ nhật xếp chồng lên nhau.
2. Dựa trên khóa k, tạo chuỗi bit giả ngẫu nhiên có độ dài bằng số phân bản đồ. Mỗi một bit được gán cho một phần của bản đồ. Khóa k có thể được sử dụng làm khóa bí mật để phát hiện hình mờ.
3. Với mỗi i ∈ {1, 2, · · ·, n}, hãy kiểm tra cờ bit của các phần bản đồ trong đó có và sau đó gán cho tập hợp con và theo giá trị cờ 0 hoặc 1. Đối với một bản đồ nhất định, hai tập con và luôn có các bản phân phối tương tự nhau.

Thứ 2, để nhúng bit 0 thì chỉ cần giữ nguyên và . Để nhúng bit 1, thì hãy đễ và nhân với hệ số α (0 < α < 1) . Mục tiêu của chương trình là giới thiệu đủ sự khác biệt giữa các bản phân phối và , nó có thể coi là một tính năng để watermark một bit.

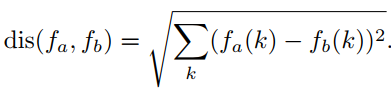
Cuối cùng, kết hợp các tập hợp con và thành một chuỗi khoảng cách . Nó được sử dụng để tính toán các vị trí mới của các đỉnh bản đồ watermark có .

Hãy biểu thị dung sai τ của . Các biến dạng vị trí gây ra bởi nhúng phải dưới để duy trì độ trung thực của bản đồ.

Để đáp ứng điều kiện, , trong là ngưỡng được sử dụng để phát hiện điểm tính năng. Do đó, được xác định như sau .

### Phát hiện hình mờ

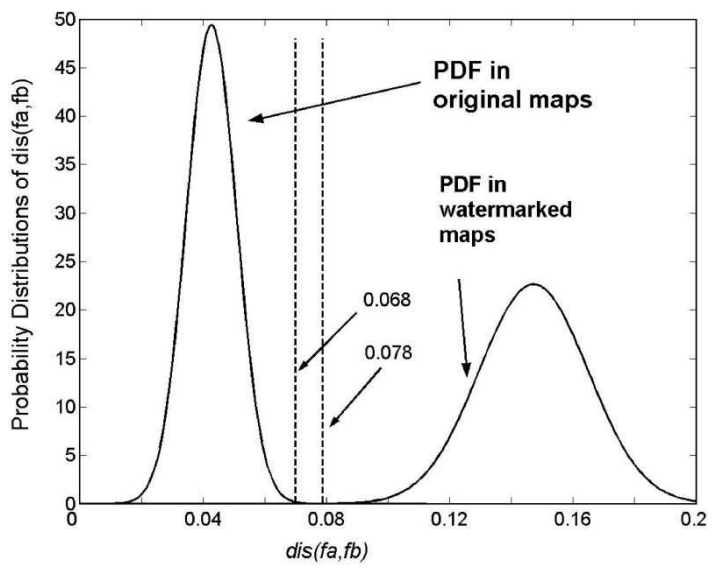
Đưa ra bản đồ dễ bị giả mạo , tập dữ liệu được trích xuất từ nó và chia thành và . Ở đây các tham số tương tự được sử dụng như và khóa bí mật k. Các hàm phân phối của và sau đó được tính toán và ký hiệu là và . Sự giống nhau giữa fa (k) và fb (k) có thể được sử dụng để phát hiện watermark. Khoảng cách Euclide được sử dụng làm thước đo để đánh giá sự giống nhau giữa hai bản phân phối:



Lưu ý rằng dis(, ) thấp hơn có nghĩa là mức độ tương tự cao hơn giữa và . Lấy dis(, ) làm thước đo tương tự, quy trình phát hiện là một phán đoán dựa trên ngưỡng T. Bản đồ không được đánh dấu nếu dis(, ) < T. Nếu dis(, ) ≥ T, bản đồ nên có hình ẩn vì quy trình nhúng có thể làm giảm sự tương đồng giữa và .

### Xác định ngưỡng

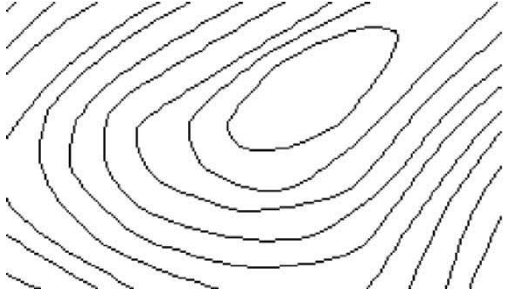
Phân phối dis(, ) được sử dụng để xác định ngưỡng T. Một 1000 khối khác nhau của bản đồ đường viền lớn được sử dụng làm bản đồ gốc và dis(, ) sau đó được tính cho từng khối trước và sau khi tạo watermarking. Hình 6.8 chỉ ra các phân phối kết quả của dis(, ) . Đường cong bên trái trong hình là phân bố dis(, ) trong bản đồ gốc và bên phải là phân bố trong bản đồ watermarking. Như trong hình, nếu T được chọn là 0,078, có thể đạt được lỗi dương phát hiện tối thiểu là 6.15 × 10-6 và lỗi âm tính giả là 4.33 ×10-5. Để tăng cường độ mạnh mẽ cho tấn công nhiễu, nên chọn T nhỏ hơn. Trong sơ đồ này, T được chọn là 0,068 với mục đích chống lại các tạp âm có biên độ thấp. Các kết quả thử nghiệm trong phần còn lại sẽ chỉ ra rằng lựa chọn là hợp lý.



Hình 6.8. Sự phân bố của dis(, )

### Hiệu suất

Bản đồ gốc được hiển thị trong Hình 6.6 được sử dụng cho các thí nghiệm. Nó là một phần của bản đồ đường viền với tỷ lệ 1: 10000. Dung sai (τ) của tọa độ là 1 mét (m). Kích thước của miếng vá bản đồ được chia là 20m × 20m. Hệ số α được chọn là 0,6. Ngưỡng phát hiện điểm tính năng có thể được tính là . Ngưỡng T được đặt thành 0,068. Buổi biểu diễn của chương trình được thảo luận trong các khía cạnh sau đây.

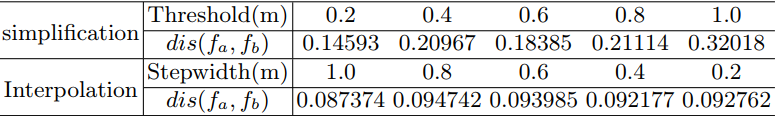


Hình 6.9. Bản đồ Watermarked

1. Tàng hình

Bản đồ kết quả được đánh dấu bằng sơ đồ được hiển thị trong Hình 6.9. Điều này cho thấy lược đồ tàng hình tốt hơn. So sánh với hình 6 (b), các biến dạng hình dạng gây ra bởi thuật toán rõ ràng là thấp hơn so với các công trình trước đây. Bằng cách chọn làm dữ liệu bìa, quy trình nhúng dữ liệu có thể giữ nguyên hình dạng cục bộ của bản đồ gốc.

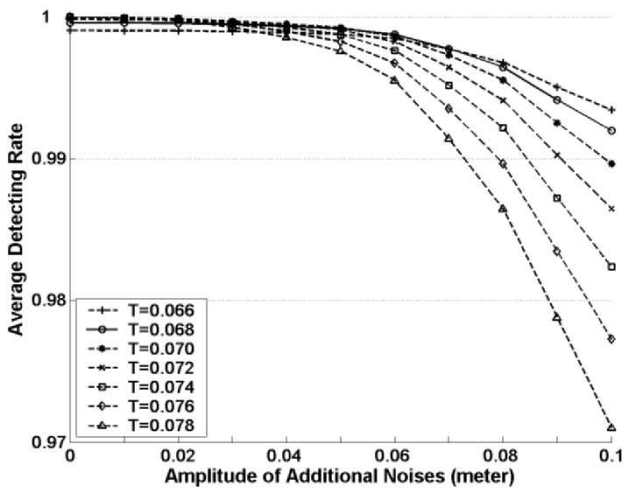
Bảng 6.1. Kết quả kiểm tra độ bền để đơn giản hóa và nội suy bản đồ



1. Mạnh mẽ

Dữ liệu bìa được chọn là bất biến đối với bản đồ dịch và xoay. Việc phân chia là độc lập theo thứ tự lưu trữ của các đối tượng bản đồ. Kết quả là, lược đồ được đề xuất vốn đã mạnh mẽ để sắp xếp lại, xoay và sắp xếp dữ liệu. Các bài kiểm tra độ bền sau đó tập trung vào các hình thức tấn công nghiêm trọng nhất đối với các sơ đồ watermarking bản đồ vector như đơn giản hóa bản đồ, nội suy và nhiễu biên độ thấp. Để đơn giản hóa bản đồ, thuật toán Douglas-Peucker được sử dụng với phạm vi ngưỡng [0, τ]. Thuật toán Spline được áp dụng cho phép nội suy bản đồ với một loạt các bước băng thông. Tất cả các kết quả phát hiện được thể hiện trong Bảng 6.1. Lấy ngưỡng T = 0,068, hình mờ có thể được bảo vệ thành công dưới mỗi lần tấn công. Điều này thể hiện sự mạnh mẽ của sơ đồ đối với các cuộc tấn công trên.

Tấn công tiếng ồn chủ yếu đến từ một số công việc hàng ngày (ví dụ: chuyển đổi định dạng bản đồ) và thường có biên độ nhỏ. Đưa ra một bản đồ để phát hiện, một loạt các tạp âm phụ gia được áp đặt với biên độ từ 0,01m đến 0,1m để mô phỏng các quy trình tấn công tiếng ồn. Đối với mỗi biên độ nhiễu, các thủ tục phát hiện và tấn công tương ứng được lặp lại 1000 lần. Giả sử rằng xác suất của bản đồ thủy ấn là 0,5, chúng ta có thể tính tỷ lệ phát hiện trung bình của sơ đồ với ngưỡng T nhất định.



Hình 6.10. Tỷ lệ phát hiện trung bình dưới các cuộc tấn công tiếng ồn.

Các đường cong trong Hình 6.10 là tốc độ phát hiện trung bình trong các cuộc tấn công tiếng ồn với một loạt ngưỡng T từ 0,066 đến 0,078. Nó được chỉ ra rằng thuật toán được đề xuất là mạnh mẽ đối với các tạp âm cộng gộp trong phạm vi biên độ đã cho. Nếu ngưỡng T giảm từ 0,078 xuống 0,068, hiệu suất chống nhiễu có thể được tăng cường đồng thời. T không nên nhỏ hơn vì tốc độ phát hiện dưới biên độ nhiễu thấp hơn có thể giảm đáng kể, như thể hiện bởi đường cong T = 0,066 trong Hình 6.10. Ngưỡng T được sử dụng trong sơ đồ là 0,068 và tốc độ phát hiện được hiển thị bằng đường cong rắn trong hình.

## Đánh giá

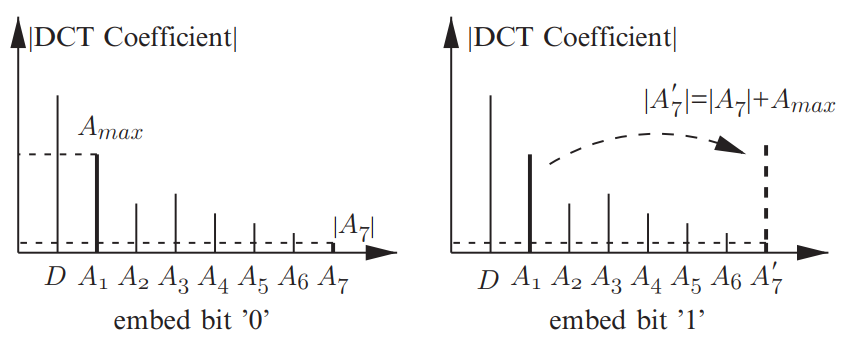
Trong phần này, kỹ thuật watermarking mạnh mẽ với bản đồ vector được giới thiệu. Một đánh giá ngắn gọn về các phương pháp điển hình được thực hiện trong cả miền không gian hoặc miền biến đổi được trình bày. Cùng với các điểm yếu chung của các công trình trước đây, một kế hoạch tạo hình được đề xuất. Bằng cách tính đến các hình dạng cục bộ của các đối tượng bản đồ, sơ đồ có thể tạo hình mờ mạnh mẽ cho bản đồ vectơ 2D trong khi đưa ra các biến dạng hình dạng thấp. Kết quả thí nghiệm cho thấy hiệu suất của sơ đồ trong khả năng ẩn mạnh mẽ. Bản đồ gốc được sử dụng trong các thí nghiệm là bản đồ đường viền với tính năng hình dạng mịn. Chúng tôi cũng đã thử nghiệm thuật toán trên một số loại bản đồ khác như bản đồ sông và bản đồ ranh giới chẳng hạn. Trong hầu hết các trường hợp, thuật toán hoạt động tốt. Tuy nhiên, thuật toán không thể hoạt động trên các bản đồ bao gồm các đường thẳng, hình chữ nhật. Cũng cần lưu ý rằng năng lực của công việc được đề xuất là nhỏ. Trong thử nghiệm, hơn hai nghìn đỉnh được sử dụng để nhúng một bit watermark để đạt được tính năng thống kê ổn định.

# Thực hiện ẩn dữ liệu có thể đảo ngược

Ẩn dữ liệu có thể đảo ngược có nghĩa là quy trình nhúng có thể đảo ngược. Nó có thể được phục hồi mà không mất sau khi dữ liệu ẩn được trích xuất hoàn toàn. Do môi trường ứng dụng quan trọng của bản đồ vector, việc sửa đổi dữ liệu bản đồ thường không được mong đợi. Do đó, các sơ đồ đảo ngược thích hợp hơn để ẩn dữ liệu trong bản đồ vectơ vì các biến dạng có thể được loại bỏ sau khi trích xuất dữ liệu ẩn. Các kỹ thuật chính của ẩn dữ liệu đảo ngược đã được nghiên cứu trên một số loại dữ liệu đa phương tiện như hình ảnh và âm thanh. Khả năng đảo ngược của sơ đồ có thể đạt được bằng cách bổ sung modulo khả nghịch [17], nén không tổn hao [18], dịch chuyển biểu đồ [19, 20], mở rộng chênh lệch [21] và nén [22]. Tuy nhiên, rất ít tác phẩm liên quan đến chủ đề ẩn dữ liệu đảo ngược trong bản đồ vector 2D. Trong phần này, hai ẩn dữ liệu có thể đảo ngược cho các bản đồ vectơ 2D dựa trên các kỹ thuật bổ sung biên độ và mở rộng chênh lệch tương ứng sẽ được thảo luận.

## Lược đồ bổ sung biên:

Voigt đã thiết kế sơ đồ hình chìm có thể đảo ngược [23] nhúng các bit hình mờ trong miền biến đổi Cosine rời rạc (DCT) nguyên. Ý tưởng chính của sơ đồ là sử dụng một tính năng quan trọng của dữ liệu bản đồ. Đó là sự tương quan cao của tọa độ đỉnh. Nói chung, do hình dạng liên tục và trơn tru của các đối tượng bản đồ, tọa độ của các đỉnh liên tiếp trong một đối tượng luôn có mối tương quan cao. Người ta biết rằng Biến đổi Cosine rời rạc có đặc tính nén năng lượng cho dữ liệu tương quan cao. Sau DCT, năng lượng của dữ liệu được chuyển đổi sẽ tập trung vào các hệ số AC tần số thấp và DC. Tận dụng các đặc điểm trên, thuật toán được đề xuất kết hợp tám đỉnh thành một đơn vị và với mỗi đơn vị, một bit hình mờ sẽ được nhúng vào nó bằng cách thay đổi các hệ số DCT trong tám điểm của nó. Phương pháp nhúng cơ bản được thể hiện trong hình 6.11.



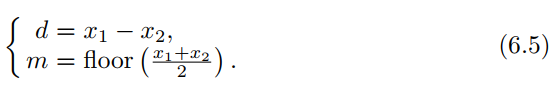
Hình 6.11. Phương pháp nhúng của [23]

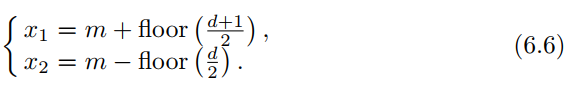
Trục dọc trong hình biểu thị các giá trị tuyệt đối của tám hệ số DCT của một đơn vị. D là hệ số DC và đến là hệ số AC. Do tương quan cao của tọa độ, giá trị tuyệt đối của hệ số tần số cao nhất, , là nhỏ nhất trong hầu hết các trường hợp. Nó thường nhỏ hơn giá trị tuyệt đối tối đa từ đến (ký hiệu là ). Nếu bit nhúng 0, đơn vị vẫn không thay đổi. Nếu bit nhúng 1, thì hệ số được sửa đổi thành . Phương pháp nhúng ở trên có thể đảo ngược vì gốc có thể thu được bằng cách trừ khỏi sau khi trích xuất bit ẩn. Sơ đồ này là thuật toán đầu tiên trong đó ý tưởng ẩn dữ liệu đảo ngược được đưa vào các bản đồ vector kỹ thuật watermarking. Một nhược điểm của sơ đồ là các biến dạng gây ra bởi hình mờ tương đối quá lớn. Các tác giả đã thực hiện một số bổ sung.

## Lược đồ mở rộng những điểm khách biệt:

### Ý tưởng

Mở rộng những điểm khác biệt là một phương pháp ẩn dữ liệu đảo ngược trong hình ảnh pixel và lần đầu tiên được đề xuất bởi Tian [21]. Ý tưởng cơ bản của việc mở rộng khác biệt là sử dụng sự tương quan cao của dữ liệu bao phủ. Đối với một cặp phần tử liền kề được ký hiệu là số nguyên và , là dữ liệu che phủ có tương quan cao, một biến đổi số nguyên được xác định để tính chênh lệch của chúng (d) và trung bình (m). Nó được hiển thị trong biểu thức sau. (6.5)





Để ẩn các bit dữ liệu, và được biến đổi thành d và m bởi công thức. (6.5). Sự tương quan cao của dữ liệu bao phủ có nghĩa là hai yếu tố và thường rất gần nhau. Đó là, sự khác biệt của d có thể rất nhỏ trong hầu hết các trường hợp. Có thể cung cấp i bit để đặt dữ liệu ẩn bằng cách dịch chuyển d bởi i bit. Điều này được thực hiện bằng cách mở rộng d lên lần trong khi vẫn giữ các biến dạng cảm ứng ở mức chấp nhận được. Giả sử sự khác biệt mở rộng mang i bit ẩn được ký hiệu là , các phần tử watermarking và có thể được tính bằng và m thông qua công thức. (6.6). Mở rộng điểm khác biệt có các thuộc tính sau khi được sử dụng để ẩn dữ liệu đảo ngược:

1. Dữ liệu bao phủ phải là số nguyên để sử dụng mở rộng những điểm khác biệt.
2. Tổng các lỗi gây ra bởi sự mở rộng sẽ được chia sẻ bởi hai phần tử và .
3. Dữ liệu bao phủ có tương quan cao được khuyến nghị để ẩn dữ liệu vì tương quan cao hơn thường có nghĩa là biến dạng thấp hơn và dung lượng cao hơn.

### Hình thành ý tưởng

Mở rộng khác biệt là thích hợp để ẩn dữ liệu trong dữ liệu bìa có tương quan cao. Một hình ảnh tự nhiên luôn có mối tương quan cao bởi vì hầu hết các pixel liền kề có giá trị tương tự nhau. Đây là màu xám hoặc màu chẳng hạn. Một bản đồ vector bao gồm một chuỗi các tọa độ của các đỉnh. Do mật độ của các đỉnh, vị trí của hai đỉnh liền kề thường rất gần nhau và sự khác biệt giữa các tọa độ của chúng là rất nhỏ. Do đó, chuỗi phối hợp cũng có thể được coi là có mối tương quan cao. Trong sơ đồ sau, tọa độ thô của bản đồ che được lấy làm dữ liệu che để ẩn dữ liệu bằng cách mở rộng chênh lệch. Thuật toán bắt đầu bằng việc phân chia bản đồ gốc trong đó cứ hai đỉnh liền kề được nhóm lại thành một cặp đỉnh. Sau đó, một điều kiện nhúng được sử dụng để đánh giá liệu một cặp nhất định có phù hợp để ẩn dữ liệu bằng cách mở rộng khác biệt hay không. Đối với tất cả các cặp phù hợp, các bit ẩn được nhúng bằng cách mở rộng sự khác biệt của tọa độ. Mặt khác, các bit ẩn được nhúng bởi thay thế Least Significant Bit (LSB) của sự khác biệt.

### Các điều kiện để thực hiện nhúng

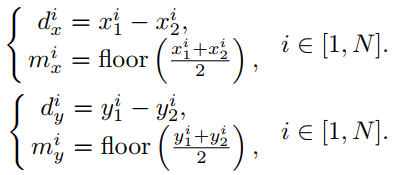
**Phân chia các điểm trong bản đồ:**

Nhìn chung, tọa độ của bản đồ vectơ là các số có dấu phẩy động với độ chính xác cố định. Để thực hiện mở rộng chênh lệch, tất cả các tọa độ trước tiên được chuyển đổi thành số nguyên bằng cách nhân với p là số chữ số sau dấu thập phân. Sau đó, bản đồ ban đầu được chia thành các nhóm. Mỗi nhóm chứa hai đỉnh liền kề. Ví dụ: giả sử một đối tượng bản đồ (đa tuyến hoặc đa giác) được tạo bởi các đỉnh , thì đối tượng được chia phải là

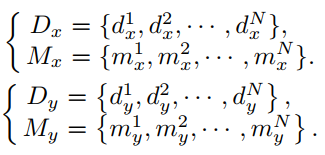
Lặp lại quy trình cho tất cả các đối tượng bản đồ (loại trừ điểm), bản đồ gốc có thể được chia thành N cặp đỉnh có thể được ký hiệu là



Trong đó và là tọa độ 2D của 2 đỉnh và . Mỗi cặp đĩnh sẽ được sẻ dụng là là đơn vị nhúng. Trong một cặp, chênh lệch d và m được tính tương ứng cho x và y.



Cho N các cặp đỉnh, từ số lượng các đỉnh ta có được công thứ sau, , và , . Đó là chuỗi khác biệt và chỗi trung bình tương ướng cho tọa độ x và y.



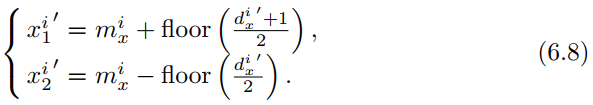
Dữ liệu ẩn sau đó có thể được nhúng vào tọa độ x (hoặc y) bằng cách áp dụng mở rộng chênh lệch cho chuỗi Dx (hoặc Dy). Cụ thể, bằng cách dịch chuyển trái (mở rộng) các phần tử phù hợp trong Dx (hoặc Dy), một số không gian bổ sung có thể được cung cấp để đặt các bit ẩn. Việc lựa chọn các yếu tố phù hợp trong Dx (hoặc Dy) dựa trên điều kiện nhúng có liên quan đến dung sai chính xác của bản đồ gốc.

**Các điều kiện nhúng**

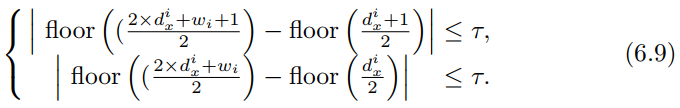
Độ sai tối đa được dùng trong bản đồ vectơ được gọi là dung sai của bản đồ và được ký hiệu là τ. Để đảm bảo tính hợp lệ của dữ liệu bản đồ, độ sai tối đa do ẩn dữ liệu không được vượt quá dung sai τ. Lấy tình huống nhúng dữ liệu theo tọa độ x làm ví dụ, một cặp đỉnh có thể phù hợp để ẩn dữ liệu bằng cách dịch chuyển bit giá trị chênh lệch nhưng chỉ khi điều kiện nhúng được thỏa mãn. Xem xét trường hợp dịch chuyển trái 1 bit là mở rộng lên gấp đôi. Giả sử một bit ẩn ∈ {0, 1} được nhúng vào tọa độ x của cặp đỉnh bằng cách mở rộng chênh lệch, điểm khác biệt chứa



Tạo độ nhúng và được tính bằng và



Để đảm bảo tính hợp lệ của dữ liệu bản đồ, nó phải đồng thời thỏa mãn và **.** Tổng hợp công thức (6.7) và (6.8) ta có điều kiện tương đương



Theo sự tương đương của và giá trị (0 hoặc 1), hai điều kiện phụ trong (6.9) có thể đơn gian hóa. Giao chúng lại với nhau có thể coi là điều kiện cuối cùng (6.10) xác định xem cặp cố phủ hợp mở rộng những điểm khác biệt không.



Phương trình (6.10) sau đó được sử dụng để kiểm tra sự phù hợp của tất cả các cặp đỉnh N. Một tập F với độ dài N được tạo ra để ghi kết quả. . Khi , cặp đỉnh thứ i đáp ứng điều kiện nhúng. Một bit ẩn sau đó sẽ được nhúng bằng cách mở rộng sự khác biệt . Theo đó, thì không thỏa điều kiện. Trong sơ đồ này, một bit ẩn cũng sẽ được nhúng như như việc thay băng LSB của . Không nên loại bỏ của ảnh gốc LSB vì mục đích đảm bảo tính thuận nghịch của sơ đồ. Một chuỗi bit L được tạo để ghi lại tất cả các LSB được thay thế để tránh mất thông tin. Cả F và L là thông tin cần thiết sẽ cần sau này để phục hồi dữ liệu. Chúng được nhúng trong bản đồ như một phần của dữ liệu ẩn.

### Nhúng và trích xuất

**Cấu trúc của dữ liệu ẩn:**

Để đảm bảo khả năng đảo ngược và cải thiện công suất, dữ liệu ẩn (ký hiệu như W) nên được trình bày như sau:



Trong đó comp (·) là thuật toán nén lossless. Ý nghĩa của mỗi thành phần được liệt kê dưới đây

*H*: Thông tin tiêu đề ghi lại độ dài của ba thành phần khác. Nó rất hữu ích để phân tách đáng tin cậy từng thành phần dữ liệu khỏi W trong quá trình trích xuất dữ liệu.

*comp (F)*: Cờ nén không tổn thất F. Cờ là cần thiết để khôi phục dữ liệu gốc vì chúng ta phải biết phương pháp nhúng của từng cặp đỉnh. Đó là, bằng cách mở rộng khác biệt hoặc bằng cách thay thế LSB. F được nén để tiết kiệm không gian cho tải trọng.

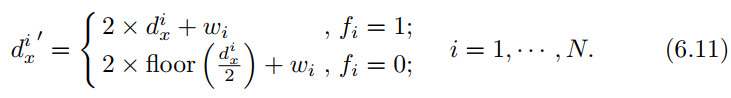
*comp (L)*: Lossless Compressed L. Để khôi phục dữ liệu gốc khi bit ẩn được nhúng bởi LSB thay thế, phải biết LSB gốc của di x. Điều này được ghi lại trong L.

*P:* Tải trọng của sơ đồ. Nó có thể là hàm băm, ví dụ MD5 trong bản đồ bìa để xác thực dữ liệu, dữ liệu meta của bản đồ bìa hoặc nơi thông điệp bí mật được truyền bởi người dùng (Truyền thông ẩn)

Comp (F) và comp (L) đều được sử dụng để phục hồi dữ liệu không mất. P là tải trọng thực được gọi là công suất của sơ đồ. W cuối cùng là một chuỗi nhị phân có độ dài N có thể được ký hiệu là .

**Ẩn dữ liệu**

Chúng tôi sẽ sử dụng phương pháp ẩn dữ liệu trong tọa độ x làm ví dụ. Phương thức này hoàn toàn giống nhau cho tọa độ y. Cho đến nay, ba chuỗi độ dài N có thể thu được bằng các phương pháp cũ. Đó là chuỗi khác biệt , cờ F và dữ liệu ẩn W. Nhiệm vụ nhúng là ẩn W trong tọa độ x bằng cách sửa đổi chuỗi khác biệt . Cho mỗi phần tử trong được chỉnh giông công thức (6.11) theo cờ :

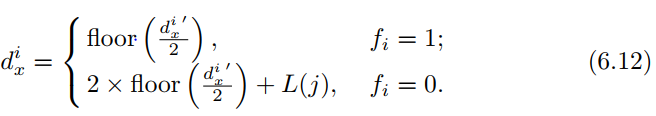


Cho khi , giá trị được nhúng bời phần mở rộng những điểm khác biệt. Là dịch chuyển trái bởi 1 bit và thay tại chổ bit thấp nhất của . Nếu không thì cho khi , giá trị được nhúng bằng cách thay thế trực tiếp LSB của . Biểu thị trình tự khác biệt được sửa đổi là tọa độ x hình mờ có thể được tính bằng và bởi phép biến đổi nghịch đảo như trong phương trình (6.6).

**Khai thác dữ liệu và phục hồi bản đồ gốc**

Với bản đồ watermarking chứa dữ liệu ẩn trong tọa độ x của nó, các phương thức tương tự được sử dụng trong qua trình nhúng là sử dụng để nhóm dữ liệu bản đồ và để tính cả chuỗi khác biệt và chuỗi tích phân . Dữ liệu ẩn W có thể thu được bằng cách trích xuất các bit thấp nhất của tất cả các phần tử của . Tải trọng P và hai chuỗi nén comp (F) và comp (L) có thể được tách ra khỏi W với sự trợ giúp của tiêu đề thông tin H.

Để khôi phục tọa độ x gốc, cần phải lấy chuỗi khác biệt ban đầu . Hai chuỗi comp (F) và comp (L) không nén, thu được cờ F và LSB L gốc. Sau đó thu được bằng cách sử dụng , F và L. Cho i = 1, …, N, mỗi phần tử của được tính như công thức sau (6.12):



và tương đương các phần tử của và F. Cho , thu được bằng các loại bỏ các bit thấp của và dich sang phải 1 bit. Cho , sẽ được phục hồi bằng việc trực tiếp thay thế LSB của bởi LSB gốc L(j), j là thứ tự trong L.

Cuối cùng, tọa độ x ban đầu có thể được phục hồi hoàn toàn bằng cách sử dụng

của phương trình (6.6) và chuỗi khác biệt được khôi phục và chuỗi trung bình .

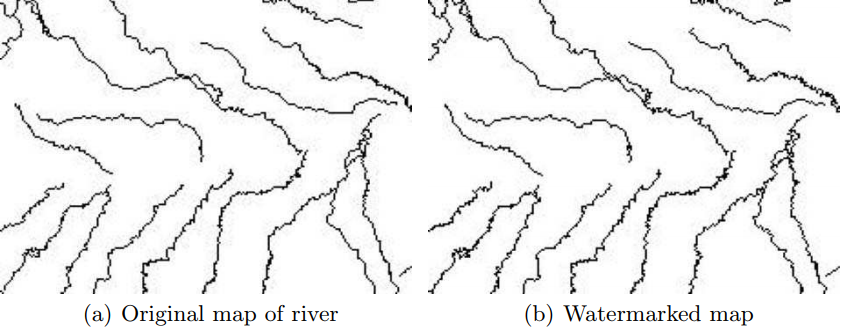
### Thực nghiệm

Hiệu suất của sơ đồ đã được kiểm tra bằng các thí nghiệm sử dụng bản đồ sông (Hình 13 (a)) với tỷ lệ 1: 4000000 làm dữ liệu bìa. Tổng số đỉnh trong bản đồ là 252000. Có 6 chữ số sau dấu thập phân của mọi giá trị tọa độ. Tất cả các tọa độ trước tiên được chuyển đổi thành số nguyên bằng cách nhân 106. Dung sai chính xác của bản đồ là τ = 0,5 km.



Hình 6.12. Hình ảnh xám (Lena, 130x130)

Cả hai tọa độ x và y được sử dụng để ẩn dữ liệu nhằm cải thiện dung lượng. Mặc dù các biến dạng cảm ứng của một số đỉnh có thể vượt quá dung sai, tính hợp lệ của dữ liệu bản đồ có thể được đảm bảo vì bản đồ gốc có thể được phục hồi chính xác sau khi trích xuất dữ liệu ẩn. Dữ liệu ẩn có độ dài tối đa 146550 bit có thể được nhúng. Ở đây công suất tối đa của sơ đồ là khoảng 0,58 bit / đỉnh. Băm MD5 của bản đồ bìa (128 bit) và hình ảnh thang độ xám có kích thước 130 × 130 (135200 bit, Hình 6.12) được lấy làm dữ liệu ẩn và được nhúng vào bản đồ bìa. Không gian còn lại của tải trọng cũng có thể được sử dụng để ẩn dữ liệu người dùng khác. Ví dụ: dữ liệu meta của bản đồ bìa. Hình 6.13 (b) là bản đồ thủy ấn chỉ ra rằng các biến dạng gây ra bởi việc ẩn dữ liệu có thể được kiểm soát tốt bằng cách sử dụng sơ đồ này. Sau khi dữ liệu ẩn được trích xuất thành công, bản đồ bìa được phục hồi. Kết quả thử nghiệm chỉ ra rằng bản đồ được khôi phục hoàn toàn giống với bản đồ gốc khi băm MD5 của chúng khớp chính xác.



Hình 6.13. Bản đồ gốc và bản đồ watermarking

## Đánh giá

Phần này thảo luận về vấn đề Ẩn dữ liệu có thể đảo ngược trong bản đồ vector. Hai thuật toán được giới thiệu. Việc đầu tiên được thực hiện trong miền DCT nguyên. Mối tương quan cao của tọa độ bản đồ và thuộc tính nén năng lượng của DCT được sử dụng bởi sơ đồ. Dữ liệu ẩn được nhúng bằng cách sửa đổi biên độ của các hệ số DCT nguyên. Các biến dạng trong sơ đồ này không dễ kiểm soát. Lược đồ thứ hai được triển khai trong miền không gian và dựa trên Mở rộng khác biệt. Một công suất tương đối cao có thể đạt được và các biến dạng cảm ứng có thể được kiểm soát tốt.

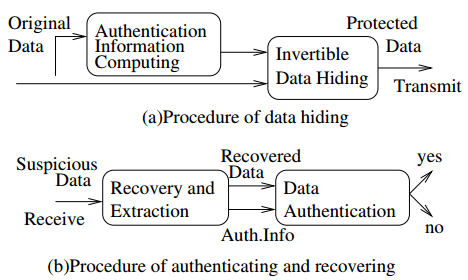
Một số khiếm khuyết là rõ ràng trong khả năng phân tích của thuật toán. Thứ nhất, tính vô hình của sơ đồ không thể tốt. Vì dữ liệu ẩn được nhúng vào tọa độ x (hoặc y), nên các hướng di chuyển của các đỉnh được giới hạn theo hướng x hoặc y. Do đó, hình dạng của bản đồ watermarking đã xuất hiện nhiễu loạn. Điều này làm suy giảm tính vô hình của sơ đồ. Thứ hai, tải sơ đồ phụ thuộc vào một mức độ nào đó bản đồ. Loại bản đồ che phủ xác định mối tương quan của tọa độ thô của nó. Một số bản đồ không có tọa độ tương quan cao. Ví dụ các bản đồ đường viền với tỷ lệ lớn. Sự tương quan thấp của tọa độ có thể dẫn đến giảm công suất.

# Thực hiện watermarking fragile

Như được miêu tả, ở phần đầu giới thiệu thì một số ứng dựng được yêu cầu độ bảo mật cao. Ví dụ như bản đồ bí mật quân sự phải được xác thực nghiêm ngặt trước khi sử dụng. Nếu một bản đồ được biết là sửa đổi, thì cần phải biết chổ nào đã sửa. Khả năng xác minh tính toàn vẹn của dữ liệu bản đồ và xác định vị trí các thay đổi là quan trọng. Có rất ít trong hoạt động lĩnh vực này. watermarking fragile rất hữu ích cho việc xác thực dữ liệu và nội địa hóa của sự can thiệp. Bằng cách giới thiệu ý tưởng về xác thực hình ảnh, phần này giới thiệu một sơ đồ watermarking fragile và trích xuất watermarking fragile để xác thực bản đồ vector 2D.

## Quy trình ẩn và nhúng dữ liệu:

Ý tưởng cơ bản là đơn giản. Đối với dữ liệu gốc được bảo vệ, dữ liệu là thông tin xác thực được tính toán trước và được sử dụng làm dữ liệu ẩn. Thông tin xác thực có thể là hàm băm, mã xác thực hoặc chữ ký số được tính trên dữ liệu gốc. Thông tin thu được sẽ chèn vào ảnh gốc bằng watermarking nghịch đảo (Hình 6.14(a)). Các dữ liệu gốc giờ có thể được truyền đi vì đã được bảo vệ bởi watermarking. Theo đó việc xác minh tính toàn vẹn dữ liệu khi có nghi ngờ sửa đổi (Hình 6.14(b)), thông tinh nhúng được xác thực trích xuất. Dữ liệu ban đầu sau đó được xây dựng lại vì watermarking được thông qua phương pháp là không thể đảo ngược. Sau đó, thông tin xác thực của dữ liệu được tính toán. Cuối cùng, tính toàn vẹn của dữ liệu nhận được có thể được xác minh bằng cách so sánh thông điệp nhúng với thông điệp được trích xuất. Nếu hai thông điệp được khớp chính xác, dữ liệu nhận được là được coi là xác thực.



Hình 6.14. Mô hình cơ bản của quá trính xác thực nghịch đảo

Một đặc điểm mong muốn khác của sở đồ là khả năng phát hiện giả mạo. Trong xác thực hình ảnh, đặc tính này có thể thu được bằng cách chia hình ảnh gốc thành các khối và xác thực từng khối một cách độc lập. Độ chính xác của nội địa hóa được xác định bởi kích thước của các khối. Dựa trên ý tưởng cơ bản, khả năng định vị giả mạo của bản đồ vector 2D có thể đạt được bằng cách chia bản đồ bao phủ thành các phân đoạn và thực hiện xác thực trên mỗi phân khúc.

Sơ đồ xác thực cho các bản đồ vector nên dễ vỡ đối với mọi nỗ lực giả mạo. Sự thay đổi của một bit đơn là một môi trường ứng dụng quan trọng trong dữ liệu bản đồ. Phần còn lại của phần trình bày nhúng và trích xuất dữ thông điệp trong watermarking fargile với bản đồ vector 2D

## Xác thực bản đồ vector bằng các sử dụng watermark fragile trích xuất thông điệp nhúng:

Lược đồ xác thực bản đồ Vector được mô tả ở đây dựa trên phương pháp đề xuất của Fridrich et al. [18], được sử dụng để xác thực hình ảnh. Trong sơ đồ này, Least Significant Bit (LSB) của ảnh gốc được thay thế bằng một bit chứa thông tin xác thực và dạng nén của LSB gốc. Dung lượng bổ sung được cung cấp bởi nén không mất dữ liệu của LSB gốc, cho phép tái tạo lại hình ảnh gốc.

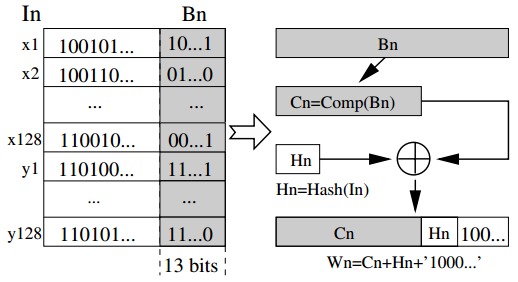
Đối với bản đồ vectơ 2D, chúng tôi tạo không gian để lưu trữ thông báo băm một cách dễ dàng nén mặt phẳng bit của tọa độ. Chúng được chọn theo phương sai (τ) của dữ liệu bản đồ. Mặt phẳng bit được chọn sau đó thay thế bằng thông điệp. Điều này được sáng tác bởi bản tóm tắt băm của bản gốc dữ liệu và dạng nén của mặt phẳng bit đã chọn. Trong hình mờ thủ tục trích xuất, cùng mặt phẳng bit được trích xuất từ ​​dữ liệu bản đồ. Các mặt phẳng bit gốc được nén, được giải nén và được sử dụng để tái tạo lại dữ liệu bản đồ gốc. Bằng cách so sánh các hàm băm được trích xuất và tính toán bản tóm tắt băm của bản đồ được khôi phục, tính toàn vẹn của dữ liệu có thể được xác thực. Để có được thông tin và vị trí của người giả mạo, bản đồ gốc nên chia thành các phân khúc. Các thủ tục trên được thực hiện cho mỗi bộ phận. Đề án watermarking fragile, sẽ làm biến dạng của dữ liệu được bảo vệ và chắc chắn sẽ chỉ ra khu vực xác thực thất bại.

**Chi tiết quá trình nhúng và chiết xuất**

Xác thực bản đồ với tỷ lệ 1: 10000 được sử dụng làm ví dụ. Giả sử dung sai của dữ liệu bản đồ là τ = 1 (mét), chi tiết như sau:

### Quá trình nhúng

*Bước 1*: Chia dữ liệu vectơ gốc thành các phân đoạn. Mỗi phân đoạn chứa 128 đỉnh, . Sau đó dịch tọa độ từ thập phân sang nhị phân.



Hình 6.15. Quy trình tạo watermark

*Bước 2*: Để đảm bảo các biến dạng cảm ứng không vượt quá dung sai, chọn mặt phẳng 13 bit cuối cùng của tọa độ x và y làm vị trí được sử dụng để nhúng hình mờ. Điều này có thể được ký hiệu là . Nén mặt phẳng bit đã chọn mà không mất. Sau đó biểu thị mặt phẳng bit được nén là . Ở đây Comp (·) là hàm nén không cần thiết.

*Bước 3*: Tính toán giá trị băm của bản đồ gốc là . Sau đó, dữ liệu hình mờ có thể được xây dựng là . Độ dài của được mở rộng bằng với độ dài của Bn bằng cách nối thêm dấu kết thúc theo sau là đủ số 0 để hoàn thành hàng đến cuối Wn. Quy trình tạo hình mờ được thể hiện trong hình 6.15.

*Bước 4*: Thay thế mặt phẳng bit đã chọn của bằng watermark để có được một đoạn watermarked . Lặp lại quy trình cho tất cả các phân khúc. Sau đó, bản đồ ban đầu được bảo vệ bằng hình mờ.

### Trích xuất và xác thực

*Bước 1:* Chia bản đồ nghi phạm thành cùng một số phân khúc.

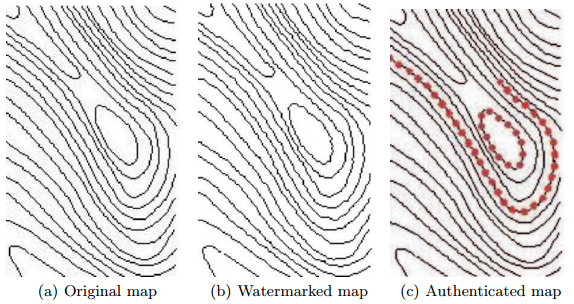
*Bước 2:* Đối với mỗi , trích xuất hình mờ . Sau đó tách thành Hash Digest và các mặt phẳng bit được nén . Giải nén thành .

*Bước 3:* Thay mặt phẳng bit đã chọn bằng và tính toán giá trị băm của bản đồ được phục hồi.

*Bước 4:* Nếu hai giá trị băm và khớp chính xác, dữ liệu bản đồ trong phân khúc hiện tại được coi là xác thực và dữ liệu gốc đã được phục hồi như một phần của quy trình. Nếu hai giá trị băm khác nhau, thì phân đoạn đang xem xét đã bị giả mạo.

## Kết quả thực nghiệm

Một số thí nghiệm được thực hiện để kiểm tra hiệu suất của thuật toán đề xuất. Hàm băm được thông qua là MD5. Đây là một chức năm một chiều có thể tùy ý chiều dài input và trả về signature với độ dài cố định là 128 bit. Hình 6.16 là hình ảnh kết quả thực nghiệm. Sự khác nhau của bản đồ gốc (Hình 6.16(a)) và bản đồ watermarking (Hình 6.16(b)) đã kiểm xoát tốt. Nếu không có sự giả mạo nào, bản đồ vector ban đầu có thể được phục hồi chính xác trong quá trình khai thác. Để kiểm tra sơ đồ khả năng bản địa hóa, chúng tôi đã can thiệp vào một số tọa độ của một phân đoạn tùy ý trong bản đồ watermarking. Hình 6.16(c) phục hồi lại bản đồ giảo mạo watermarking. Phân đoạn giả mạo có thể được phát hiện một cách đáng tin cậy và có được đánh dấu trong hình. Các phân khúc khác được khôi phục mà không mất.



Hình 6.16. Kết quả thực nghiệm.

Lược đồ hoàn toàn dễ vỡ, trong đó xác thực chắc chắn sẽ thất bại ngay cả khi chỉ có một chút dữ liệu được bảo vệ đã bị thay đổi. Lưu ý rằng độ chính xác định vị của sơ đồ được giới hạn trong chiều dài của bản đồ phân khúc. Việc lựa chọn độ dài phân khúc phải đảm bảo rằng mọi phân khúc có thể cung cấp đủ không gian để chèn thông tin xác thực. Điều này liên quan nhiều đến loại bản đồ che phủ và phương pháp nén không tổn hao. Đối với bản đồ có tọa độ tương quan cao, phân đoạn chiều dài có thể được chọn ngắn hơn so với trong bản đồ có tương quan thấp hơn. Loại bản đồ che phủ có thể có ảnh hưởng quan trọng đến độ chính xác về khả năng định vị giả mạo của chương trình. Thuật toán nén không mất dữ liệu được áp dụng trong thử nghiệm là thuật toán mã hóa Huffman được sử dụng phổ biến để nén dữ liệu. Do tọa độ của bản đồ đường viền ban đầu có mối tương quan cao, nên có thể cung cấp đủ không gian bằng cách nén LSB của đoạn chứa 128 đỉnh. Tuy nhiên, không thể được áp dụng thực tế vì sách mã Huffman không được xem xét.

# Tóm tắt và hướng phát triển

Bản đồ vectơ kỹ thuật số là một phần quan trọng của hệ thống thông tin địa lý (GIS) và đã được sử dụng rộng rãi ở nhiều khu vực. Trong chương này, chủ đề về hình mờ của GIS đã được xem xét. Đó là, các kỹ thuật ẩn dữ liệu trong bản đồ vector 2D. Điều này có nhiều ứng dụng tiềm năng trong các ứng dụng của dữ liệu bản đồ vector. Theo các thuộc tính và mục đích của các đề án, các công việc liên quan trong lĩnh vực này đã được phân loại thành ba lớp. Đó là, watermarking mạnh mẽ, thực hiện nhúng và trích xuất trong watermarking fragile. Ranh giới giữa các lớp này không hoàn toàn khác biệt. Ví dụ, sơ đồ dễ vỡ được trình bày trong Mục 6.4 cũng là một sơ đồ không thể đảo ngược. Trạng thái của nghệ thuật ẩn dữ liệu bản đồ vector 2D đã được đưa ra trong chương này bằng cách giới thiệu các đề án chính được đề xuất trong mỗi lớp. Ngoài ra một số vấn đề mở trong nghiên cứu hiện tại đã được trình bày một số kế hoạch. Một số nhược điểm của các công trình trước đây đã được cải thiện trong các đề án được trình bày. Bằng cách điều tra hiệu suất của các chương trình hiện tại, nó cho thấy vẫn còn nhiều vấn đề chưa được giải quyết. Sức mạnh mẽ của bản đồ cần được tăng cường trước cuộc tấn công đơn giản hóa và tấn công tiếng ồn. Thứ hai, các biến dạng được giới thiệu bởi watermarking nên được kiểm soát ở nhiều khía cạnh. Thứ ba, việc đánh giá chất lượng bản đồ cần một thước đo hiệu suất phù hợp. Hơn nữa, các sơ đồ có thể đảo ngược và dễ vỡ cần được xem xét thêm vì mức độ bảo mật cao cần thiết trong một số ứng dụng đặc biệt

## Watermarking mạnh mẽ chống lại các cuộc tấn công

Hầu hết các thuật toán được đề xuất có hiệu suất tốt để chống lại các cuộc tấn công hình học, nhưng kém mạnh mẽ hơn trước các cuộc tấn công noise. Hầu hết các sơ đồ là dễ vỡ để khi đơn giản hóa bản đồ hoặc cắt xén. Điều này đặc biệt cho các cuộc tấn công được thực hiện trong miền biến đổi. Có thể hiểu rằng càng mất nhiều thông tin trong cuộc tấn công thì càng nguy hiểm. Tiếng ồn áp đặt độc hại lên dữ liệu bản đồ không phải là lựa chọn tốt cho kẻ tấn công, những biến dạng tiếng ồn nhỏ xảy ra trong một số công việc hàng ngày như chuyển đổi định dạng bản đồ. Do đó, một sơ đồ mạnh mẽ ít nhất phải mạnh mẽ đối với các tạp âm có biên độ thấp. Đơn giản hóa bản đồ là một hoạt động phổ biến trong GIS và có thể gây mất thông tin mà không làm giảm tính hợp lệ của dữ liệu bản đồ. Điều rất quan trọng đối với watermark để chống lại cuộc tấn công này. Đây có thể được coi là một vấn đề mở và cần phải được tăng cường.

## Độ trung thực: Lỗi, Hình dạng, Cấu trúc liên kết

Nguyên tắc quan trọng nhất, phải được tuân thủ nghiêm ngặt bởi các sơ đồ watermarking bản đồ vector, là các biến dạng gây ra bởi việc ẩn dữ liệu phải được kiểm soát chặt chẽ để đảm bảo chất lượng của bản đồ được đánh dấu. Đối với bản đồ vector, thuật ngữ biến dạng có nhiều ý nghĩa. Một là biến dạng có thể có nghĩa là sai số tọa độ của các đỉnh. Hầu như tất cả các thuật toán hiện có đã tính đến yếu tố này. Một quy tắc nổi tiếng là các sai số của các đỉnh gây ra bởi hình mờ không được vượt quá độ chính xác của bản đồ che phủ map. Đối với các lược đồ trong miền không gian, việc nhúng dữ liệu trực tiếp sửa đổi tọa độ và thuận tiện để kiểm soát các lỗi. Trong trường hợp biến đổi tên miền, tình hình là khác nhau. Các lược đồ dựa trên miền biến đổi đó trực tiếp sửa đổi các hệ số biến đổi thay vì tọa độ. Nó không thuận tiện để kiểm soát các lỗi được giới thiệu. Ngoài các lỗi đỉnh, biến dạng cũng có nghĩa là lỗi của nhiều tính năng khác như hình dạng hoặc cấu trúc liên kết của các đối tượng bản đồ. Loại biến dạng này có tầm quan trọng lớn đối với chất lượng bản đồ. Các tác phẩm trước đây, họ đã bị bỏ qua bởi hầu hết. Ngay cả sự biến dạng nhẹ của các đỉnh cũng có thể gây ra những thay đổi không thể chấp nhận được về thông tin hình dạng hoặc cấu trúc liên kết. Ví dụ: bằng cách phá hủy độ mịn của các đối tượng bản đồ hoặc làm cho các đối tượng được phân tách ban đầu bị chồng chéo. Đây là một trong những công việc tương lai cần thiết để kiểm soát các biến dạng trong các sơ đồ watermarking bản đồ vector.

## Đánh giá độ trung thực

Mặc dù các loại thuật toán cho bản đồ vector watermarking đã được đề xuất, nhưng vẫn không có biện pháp thích hợp để đánh giá độ trung thực của bản đồ thủy ấn. Một biện pháp như vậy có thể rất quan trọng đối với sơ đồ watermarking để đánh giá mức độ biến dạng được đưa ra bởi quy trình nhúng. Một thước đo lý tưởng về độ trung thực của bản đồ cần tính đến không chỉ các lỗi về tọa độ, mà cả các biến dạng của các hình dạng bản đồ. Chúng tôi đã mô tả rằng cả nhận thức của con người và PSNR đều không phải là ứng cử viên phù hợp. Một phương pháp để đánh giá một cách hợp lý bản đồ vector độ trung thực vẫn là một chủ đề cho một công việc trong tương lai.

## Trích xuất và watermarking fragile.

Cho đến bây giờ, rất ít tác phẩm trong văn học thế giới hiện nay đã xem xét chủ đề này thiết kế sơ đồ watermarking và trích xuất cho bản đồ vector. Hạn chế quan trọng của hình mờ bản đồ vector là độ chính xác của bìa dữ liệu, watermarking và trích xuất cung cấp giải pháp tốt nhất cho yêu cầu này. Trong quảng cáo, hầu hết các công trình trước đây tập trung vào việc thiết kế các chương trình mạnh mẽ, nhưng ít có những fragile. Xác thực dữ liệu bản đồ vẫn là một vấn đề quan trọng cho nhiều ứng dụng đặc biệt, đặc biệt là cho các ứng dụng quân sự. Thiết kế các kế hoạch dễ vỡ cho mục đích xác thực dữ liệu bản đồ là một cách khác công việc quan trọng trong tương lai.

# Tài liệu tham khảo

1. <http://www.thedigitalmap.com>

2. Sakamoto, M., Matsuura, Y., and Takashima, Y. (2000): A scheme of digital watermarking for geographical map data. Symposium on cryptography and Information security

3. Kang, H. (2001): A vector watermarking using the generalized square mask. Proc. International Symposium on Information Technology: Coding and Computing, 234–236

4. Schulz, G. and Voigt, M. (2004): A high capacity watermarking system for digital maps. Proceedings of the 2004 multimedia and security workshop on Multimedia and security, 180–186

5. Huber, B.: Gis & steganography – part 3: Vector steganography. Available: http://www.directionsmag.com/

6. Voigt, M. and Busch, C. (2003): Feature-based watermarking of 2d-vector data. Proceedings of SPIE, Security and watermarking of Multimedia Content, 5020, 359–366

7. Ohbuchi, R., Ueda, H., and Endoh, S. (2002): Robust watermarking of vector digital maps. Proc. IEEE International Conference on Multimedia and Expo, (ICME ’02), 1, 577–580

8. Voigt, M. and Busch, C. (2002): Watermarking 2d-vector data for geographical information systems. Proceedings of SPIE, Security and watermarking of Multimedia Content, 4675, 621–628

9. Nikolaidis, N., Pitas, I., and Solachidis, V. (2000): Fourier descriptors watermarking of vector graphics images. Proc. IEEE International Conference on Image Processing, 3, 9–12

10. Solachidis, V., Nikolaidis, N. and Pitas I. (2000): Watermarking polygonal lines using Fourier descriptors. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP’2000), 1955–1958

11. Nikolaidis, N., Pitas, I., and Giannoula, A. (2002): Watermarking of sets of polygonal lines using fusion techniques. Proc. IEEE International Conference on Multimedia and Expo (ICME ’02), 2, 549–552

12. Kitamura, I., Kanai, S. and Kishinami, T. (2001): Copyright protection of vector map using digital watermarking method based on discrete Fourier transform. Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS

’01), 3, 1191–1193

13. Li, Y. and Xu, L. (2003): A blind watermarking of vector graphics images. Proc. Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2003), 424–429

14. Ohbuchi, R., Ueda, H., and Endoh, S. (2003): Watermarking 2d vector maps in the mesh-spectral domain. Proc. International Conference on Shape Modeling and Applications, 216–225

15. Ohbuchi, R., Takahashi, S., Miyazawa, T., and Mukaiyama, A. (2001): Watermarking 3d polygonal meshes in the mesh spectral domain. Proc. Graphics Interface, 9–17

16. Douglas, D.H. and Peucker, T.K. (1973): Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Canadian Cartographer, 10, 112–122

17. Honsinger, C., Jones, P., Rabbani, M., and Stoffel, J. (1999): Lossless recovery of an original image containing embedded data. U.S. Patent 6 278 791

18. Fridrich, J., Goljan, M., and Du, R. (2001): Invertible authentication. Proc. SPIE, Security and Watermarking of Multimedia Contents, 3971, pp. 197–208

19. Ni, Z., Shi, Y., Ansari, N., and Su, W. (2003): Reversible data hiding. Proceedings IEEE of ISCAS’03, 2, II-912–II-915

20. Leest, A., Veen, M. and Bruekers, F. (2003): Reversible image watermarking. Proc. IEEE International Conference on Image Processing, 2, 731–734

21. Tian, J. (2002): Reversible watermarking by difference expansion. Proc. of Workshop on Multimedia and Security, 19–22

22. Veen, M., Bruekers, F., Leest, A., and Cavin, S. (2003): High capacity reversible watermarking for audio. Proceedings of the SPIE, 5020, 1–11

23. Voigt, M., Yang, B., and Busch, C. (2004): Reversible watermarking of 2dvector data. Proceedings of the 2004 Multimedia and Security Workshop on Multimedia and Security, 160–165