Data Hiding in Curves with Fingerprinting Applications for Maps

Hongmei Gou and Min Wu
Communication and Signal Processing Lab

1. Introduction and Motivation

- Why maps and curves?
  - Maps contain important information for military and intelligence agencies
  - Curve is one of the major components in maps
- Representation of digital map
  - Raster Map
    - An image represented by a 2D array of pixels
  - Vector Map
    - Geometrical primitives are used to represent objects
- Challenges
  - Resilience to geometric attacks, such as rotation, scaling, and translation (RST)
  - Resilience to D/A-A/D conversion, such as printing-and-scanning
- Current embedding techniques for binary documents
  - Fragile and used for authentication and annotation e.g., flip pixels or perturb vertices
  - Can not survive D/A-A/D conversion
  - Difficult to deal with the RST issue
- Our approach
  - New feature domain
    - Coordinates of control points in B-spline representation of curves
  - Spread spectrum embedding and correlation-based detection
  - Robust to collusion and printing-and-scanning attacks

2. Data Hiding Algorithms

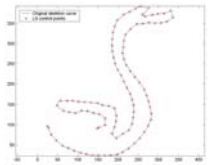
- Feature extraction
  - B-spline representation of curves

$$\mathbf{p}(t) = \sum_{i=0}^{n} \mathbf{c}_i B_{i,k}(t)$$

Curve Points; B-spline Basis Function; Control Points

  - Control points
    - Given a set of properly chosen samples on the curve, its B-spline control points can be obtained using the least square technique

Sample Curve



Control Points

  - Advantages
    - Compact set of salient features for curves
    - Invariant to affine transform, e.g., RST
    - Decoupling of x and y coordinates
- Embedding fingerprints
  - Orthogonal noise-like sequences are taken as digital fingerprints
  - Spread spectrum additive embedding: a scaled version of the fingerprint sequence is added to the coordinates of the set of control points

$$\mathbf{c}_i' = \mathbf{c}_i + \alpha \mathbf{w}_i$$

    - $\mathbf{c}_i$: original control points
    - $\mathbf{w}_i$: fingerprinting sequence
    - α : scaling factor
    - $\mathbf{c}_i'$: marked control points
  - A fingerprinted curve can be constructed from the marked control points
- Detecting fingerprints
  - Registration with the original unmarked curve, which is available to a detector in fingerprinting applications
  - Extract control points of the test curve and compute the difference to arrive at an estimated fingerprint sequence

$$\mathbf{w}_i^T = (\mathbf{c}_i^T - \mathbf{c}_i)/\alpha$$

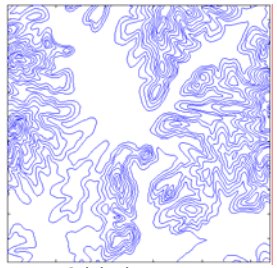
    - $\mathbf{c}_i^T$: control points extracted from the test curve
    - $\mathbf{w}_i^T$: estimated fingerprinting sequence
  - Correlation-based detection
    - Evaluate the similarity between the estimated fingerprinting sequence and each fingerprint sequence in the database by correlation coefficient ρ and Z-statistics

$$Z = \log\left(\frac{1+\rho}{1-\rho}\right)\frac{\sqrt{2(n+1)-3}}{2}$$

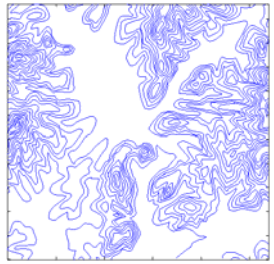
    - If the similarity is higher than a threshold (3-6 for Z-statistics), then with high probability the corresponding fingerprint sequence is present in the test curve

3. Experimental Results

- Fingerprinting topological map

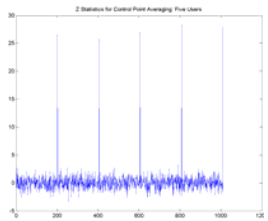


Original Map

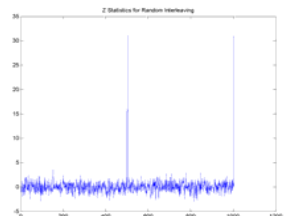


Fingerprinted Map

- Resilient to collusion attacks

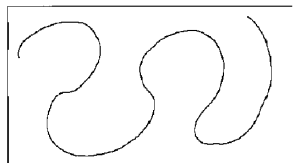


5-user Average Collusion

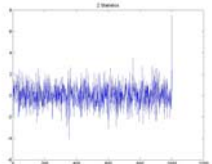


2-user Interleaving Collusion

- Resilient to the printing-and-scanning attack



Curve after Printing-and-Scanning



Detection Statistics

4. Conclusions and Future Work

- New data hiding algorithm for curves
  - Parameterizing curves using the B-spline model
  - Resistant to collusion attacks
  - Resistant to the printing-and-scanning attack
- Applications for protecting maps and drawings
- Future work
  - Automatic registration for RST resilient fingerprinting
  - Printing-and-scanning tests for large scale maps
  - Fingerprinting other documents, such as drawings, signatures, and handwritten notes.