Printer Characterization and Signature Embedding for Security and Forensic Applications

Pei-Ju Chiang, Aravind K. Mikkilineni, Sungjoo Suh
Jan P. Allebach, George T.-C. Chiu, Edward J. Delp

http://shay.ecn.purdue.edu/~prints

Document Protection & Forensic Analysis

- Investigate methods that will allow one to determine if a given document was printed on a particular digital printer, model, manufacturer
- Use our experience in characterizing and improving print quality of ink-jet and electrophotography (laser printing) to develop techniques to authenticate a printer and a document

Protect & Prevent

- Copying – scan and print
- Forgery - Alterations – additions and deletions
- Fingerprint and Trace
- Authentication – is this a proper document?

Goals

- Develop methods to embed security features in printed documents by exploiting how the printer actually places marks on the paper to form the printed document
- Can be very robust and/or can be very fragile and tamperproof
- Use intrinsic signature of printer to identify as much information as possible from printed document about printer that produced it
- Embed auxiliary information in document at time of printing via extrinsic signature
- Intrinsic and extrinsic signatures are based on extraction and modulation of features generated by the physical characteristics of the printer mechanism

Intrinsic Printer Identification

- Graylevel co-occurrence texture features estimated from printed regions within individual text characters
- System works across various font types and sizes, paper types, and consumable age when trained with same font and paper type
- 90% classification accuracy when training on new data and testing on old data
- Similar technique is promising for forensic identification of inkjet printers

Extrinsic Signature Embedding

- Generate extrinsic signature by modulating laser intensity
- Ability to synchronize with individual text lines and embed different signals on a per line basis
- Embedded signature does not affect perceived image/text quality, but is still detectable from the scanned document

Character Level Decoding Error

- Test document generated using our Forensic Monkey Text Generator (FMTG) with 12pt Times Roman font (50 lines per page)
- 2 bits embedded every three lines (33 bits in 50 lines of text)
- All bits can be correctly detected with well chosen embedding parameters
- Tradeoff between embedding frequency/amplitude and detection error

References available at http://shay.ecn.purdue.edu/~prints