General-Purpose Image Forensics Using Patch Likelihood under Image Statistical Models

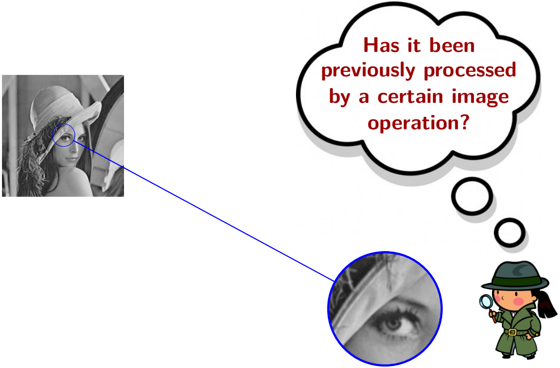
The 7th IEEE International Workshop on Information Forensics and Security

Wei Fan, Kai Wang, and François Cayre

GIPSA-lab, Grenoble, France

18-11-2015

Detecting Image Operations



1. Generality
   - Targeted
   - General-purpose
2. Size
   - whole image
   - small image block

Analysis of Current Image Forensics

- Targeted Forensics (*well studied*)
  - Exploit particular artifacts of *specific* image operation
  - Different features for different image operations
- General-Purpose Forensics (*little studied*)
  - Cope with *multiple* image operations
  - Possible to adopt powerful steganalytical features, *e.g.*, SPAM
- Forensic classification on small image blocks
  - Important for revealing *forgery semantics*
  - Image block size $\downarrow$ $\xrightarrow[\textit{leads to}]{\textit{usually}}$ forensic performance $\downarrow$

- X. Qiu *et al.*, "A universal image forensic strategy based on steganalytic model". In: *Proc. ACM IHMMSec*, 2014, pp. 165-170
- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Analysis of Current Image Forensics

- Targeted Forensics (*well studied*)
  - Exploit particular artifacts of *specific* image operation
  - Different features for different image operations

- Most current forensic methods are targeted, and few results are reported on small image blocks

1. **Generality**
2. **Classification on small blocks**

- Forensic classification on small image blocks
  - Important for revealing *forgery semantics*
  - Image block size ↓ $\xrightarrow[\textit{leads to}]{\textit{usually}}$ forensic performance ↓

▶ X. Qiu *et al.*, "A universal image forensic strategy based on steganalytic model". In: *Proc. ACM IHMMSec*, 2014, pp. 165-170
▶ T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Motivation

Question

Given an image block, is it more like a natural, original block or a processed one?

Proposed Solution

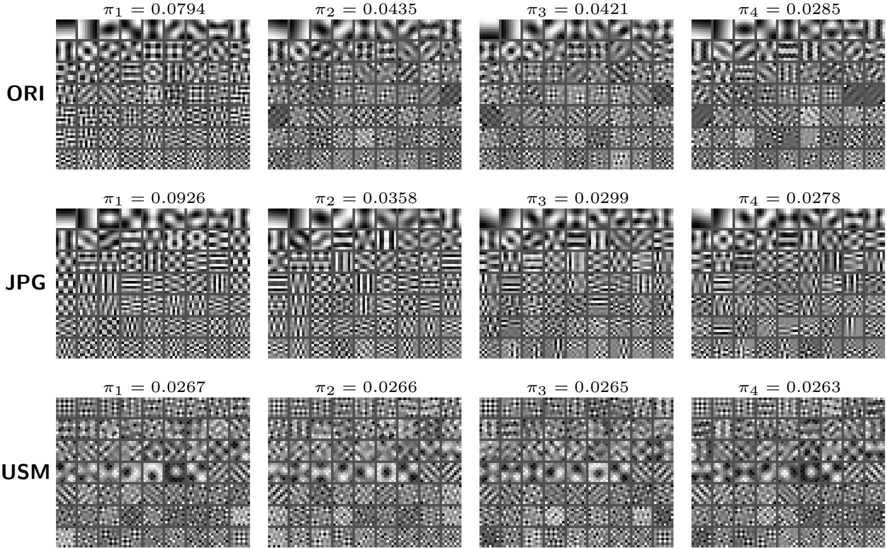
Compare the average patch likelihood values calculated under different natural image statistical models

Gaussian Mixture Model (GMM)

$$L(\theta|\mathbf{x}) = p(\mathbf{x}|\theta) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \mathbf{C}_k)$$

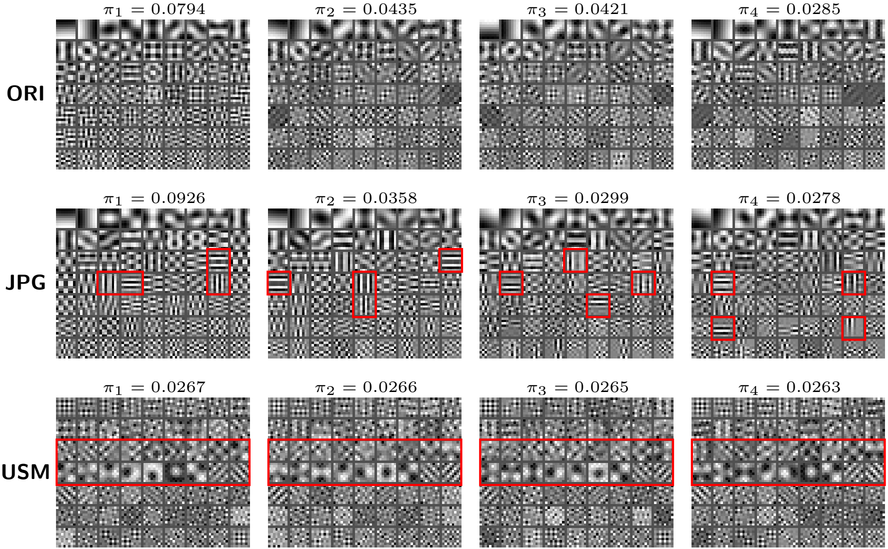
- D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration". In: *Proc. ICCV*. 2011, pp. 479-486

Eigenvectors of GMM Covariance Matrices



- D. Zoran and Y. Weiss, “Natural images, Gaussian mixtures and dead leaves”. In: *Proc. NIPS*. 2012, pp. 1736-1744

Eigenvectors of GMM Covariance Matrices



- D. Zoran and Y. Weiss, "Natural images, Gaussian mixtures and dead leaves". In: *Proc. NIPS*. 2012, pp. 1736-1744

Hypothesis Testing

Test

$$\Lambda(\mathbf{X}) = \frac{1}{N}\sum_{i=1}^{N} \log L(\theta_0|\mathbf{x}_i) - \frac{1}{N}\sum_{i=1}^{N} \log L(\theta_1|\mathbf{x}_i) \gtrless \eta$$

- $\mathbf{x}_i$: overlapping patches extracted from image (block) $\mathbf{X}$
- $\mathcal{H}_0$: $\mathbf{X}$ is original, unprocessed
  GMM parametrized by θ_0
- $\mathcal{H}_1$: $\mathbf{X}$ is processed by a certain image operation
  GMM parametrized by θ_1

Decision Rule

$$\begin{cases} \text{reject } \mathcal{H}_0 & \text{if } \Lambda(\mathbf{X}) \leq \eta \\ \text{do not reject } \mathcal{H}_0 & \text{if } \Lambda(\mathbf{X}) > \eta \end{cases}$$

Image Operations

ORI	no image processing
GF	*Gaussian filtering* with window size 3×3, and standard deviation 0.5 to generate the filter kernel
JPG	*JPEG compression* with quality factor 90
MF	*median filtering* with window size 3×3
RS	*resampling* with bicubic interpolation to scale the image to 80% of its original size
USM	*unsharp masking* with window size 3×3, and parameter 0.5 for the Laplacian filter to generate the sharpening filter kernel
WGN	*white Gaussian noise addition* with standard deviation 2

- 6 image operations, each of which is with one fixed parameter setting

Image Datasets

1. GFTR: 2457 images of size 512×512 for **training**
   - SPAM (686-dimensional), 2457 samples (whole image or block)
   - GMM (200 components), $\sim$1.2 million extracted 8×8 patches
2. GFTE: 2448 images of size 512×512 for **testing**
   - whole image ($512{\times}512$), 2448 samples for each image operation
   - image block (32×32, 16×16), 2448×10 samples for each image operation

- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224
- `ftp://firewall.teleco.uvigo.es:27244/DS_01_UTFI.zip`
- `ftp://lesc.dinfo.unifi.it/pub/Public/JPEGloc/dataset/`

Experimental Results

detection accuracy [%]

		GF	JPG	MF	RS	USM	WGN
512×512	SPAM-based	99.86	98.20	99.94	96.45	99.73	98.53
	Proposed-S	99.10	97.28	95.69	92.61	99.73	99.45
	Proposed-T	99.82	99.49	99.31	92.67	99.73	99.80
32×32	SPAM-based	99.35	94.18	99.43	89.23	98.76	95.04
	Proposed-S	97.69	95.83	93.81	90.96	99.22	95.50
	Proposed-T	97.73	96.04	93.99	90.96	99.21	97.55
16×16	SPAM-based	98.38	88.00	99.26	78.21	97.82	91.20
	Proposed-S	97.27	94.27	92.88	89.70	98.59	95.58
	Proposed-T	97.37	94.68	93.01	89.72	98.59	95.66

- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Experimental Results

Simple threshold: $\eta = 0$

detection accuracy [%]

		GF	JPG	MF	RS	USM	WGN
512×512	SPAM-based	99.86	98.20	99.94	96.45	99.73	98.53
	Proposed-S	99.10	97.28	95.69	92.61	99.73	99.45
	Proposed-T	99.82	99.49	99.31	92.67	99.73	99.80
32×32	SPAM-based	99.35	94.18	99.43	89.23	98.76	95.04
	Proposed-S	97.69	95.83	93.81	90.96	99.22	95.50
	Proposed-T	97.73	96.04	93.99	90.96	99.21	97.55
16×16	SPAM-based	98.38	88.00	99.26	78.21	97.82	91.20
	Proposed-S	97.27	94.27	92.88	89.70	98.59	95.58
	Proposed-T	97.37	94.68	93.01	89.72	98.59	95.66

Trained threshold η on GFTR dataset

- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Experimental Results

detection accuracy [%]

		GF	JPG	MF	RS	USM	WGN
512×512	SPAM-based	99.86	98.20	99.94	96.45	99.73	98.53
	Proposed-S	99.10	97.28	95.69	92.61	99.73	99.45
	Proposed-T	99.82	99.49	99.31	92.67	99.73	99.80
32×32	SPAM-based	99.35	94.18	99.43	89.23	98.76	95.04
	Proposed-S	97.69	95.83	93.81	90.96	99.22	95.50
	Proposed-T	97.73	96.04	93.99	90.96	99.21	97.55
16×16	SPAM-based	98.38	88.00	99.26	78.21	97.82	91.20
	Proposed-S	97.27	94.27	92.88	89.70	98.59	95.58
	Proposed-T	97.37	94.68	93.01	89.72	98.59	95.66

- At least comparable to the SPAM feature
- Especially advantageous on small blocks

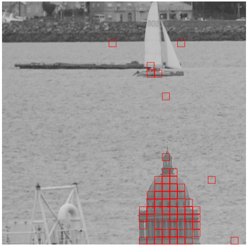
▶ T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Fine-Grained Image Tampering Localization

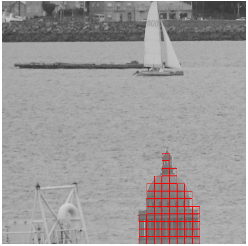
ORI

JPG

Forgery



SPAM-based



Proposed

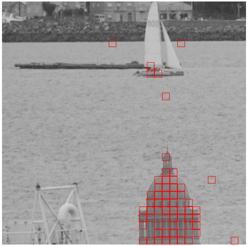
- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Fine-Grained Image Tampering Localization

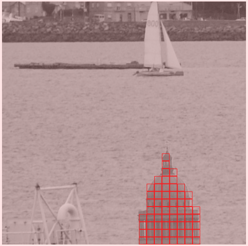
ORI

JPG

Forgery



SPAM-based



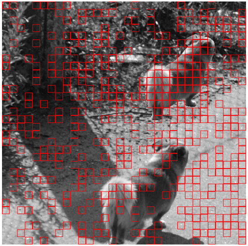
Proposed

- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

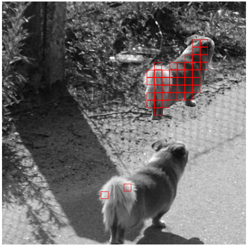
Fine-Grained Image Tampering Localization

ORI

Forgery (with **RS**)



SPAM-based



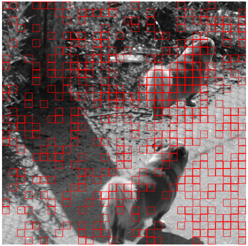
Proposed

- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

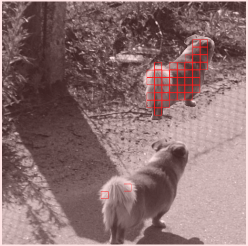
Fine-Grained Image Tampering Localization

ORI

Forgery (with **RS**)



SPAM-based



Proposed

- T. Pevný *et al.*, "Steganalysis by subtractive pixel adjacency matrix". *IEEE TIFS* 5, 2 (2010), pp. 215-224

Conclusions

1. A general-purpose framework for image forensics
   - Comparison of average patch likelihood values calculated under different image models
   - At least comparable performance compared with the SPAM feature
   - Conceptually simplicity, no hand-crafted feature extraction, and easiness to be extended

Perspectives

- Multi-class classification
- More image operations with more parameters
- Richer natural image statistical models

Thank you for your attention!

Q & A