Optimizing Pixel Predictors for Steganalysis

Vojtěch Holub and Jessica Fridrich

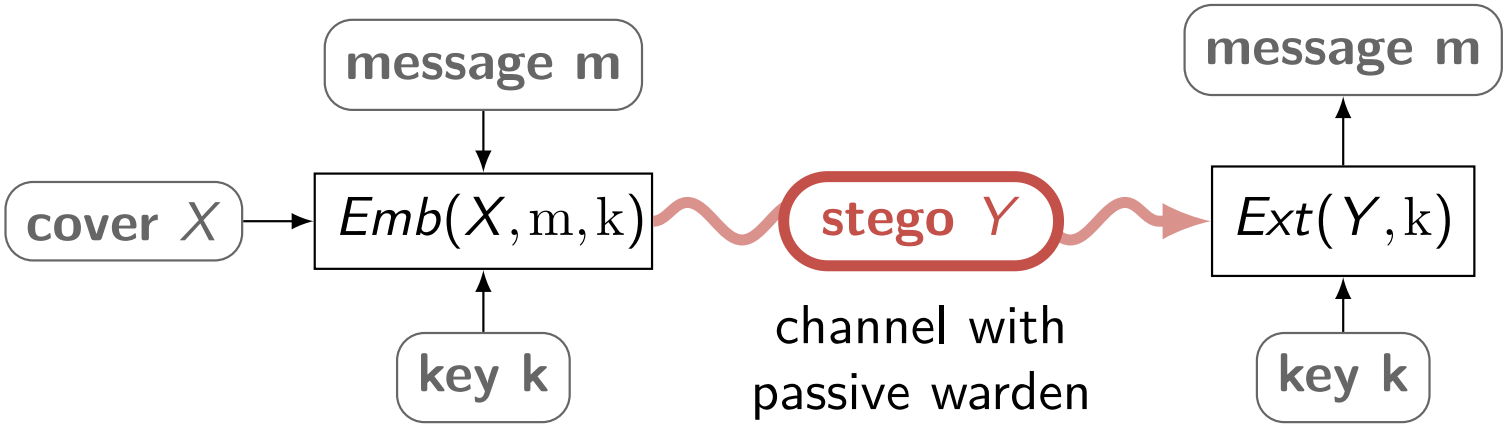
Dept. of Electrical and Computer Engineering
SUNY Binghamton, New York

IS&T / SPIE 2012, San Francisco, CA

BINGHAMTON
UNIVERSITY
STATE UNIVERSITY OF NEW YORK

Steganography

- **The art of secret communication**



- **Steganography by cover modification**
  X is slightly modified to Y to convey a secret message (by flipping LSBs, changing DCT coefficients, ...). Goal: make the embedding changes statistically undetectable.
- **Steganalysis**
  Warden's job: tell whether a cover or stego object is sent.

Pixel Predictor

Warden represents images by features computed from noise residuals and builds the detector as a classifier in the feature space.

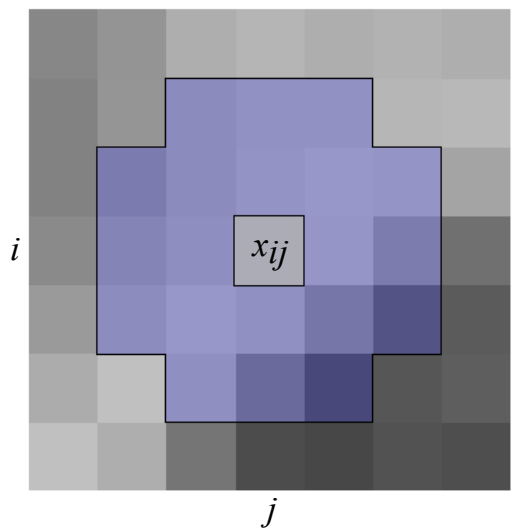
Noise residual

- **Narrower dynamic range than x_{ij}**
- **Increased SNR**

$$r_{ij} = x_{ij} - \mathrm{Pred}(x_{ij})$$

Predictor

- **Estimates the value of pixel x_{ij} from its neighborhood**
- **E.g., by fitting linear or quadratic polynomials, etc.**



Detection Framework

1. **Computing residual:** $r_{ij} = x_{ij} - \mathbf{Pred}(x_{ij})$
2. **Quantization and truncation:** $r_{ij} \leftarrow \mathrm{round}\left(\mathrm{trunc}_T\left(\frac{r_{ij}}{q}\right)\right)$, $q \in \mathbb{R}, T = 2$. **Thus,** $r_{ij} \in \{-2, 1, 0, 1, 2\}$
3. **Forming 4D co-occurrence matrix:** $\mathbf{C} = \mathbf{C}^{(\mathrm{h})} + \mathbf{C}^{(\mathrm{v})}$

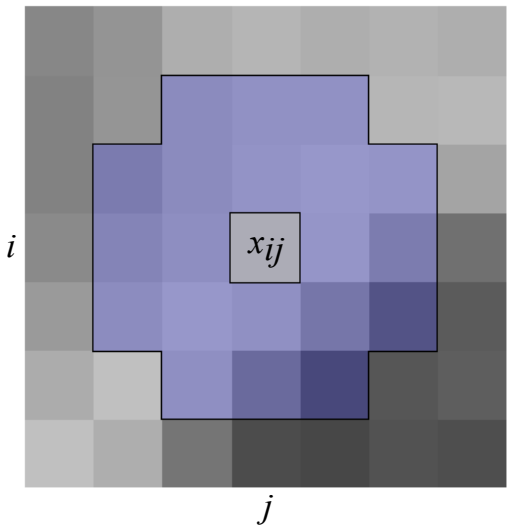
$$\mathbf{C}^{(h)}_{d_1 d_2 d_3 d_4} = \{\#(i,j) | r_{ij} = d_1, r_{ij+1} = d_2, r_{ij+2} = d_3, r_{ij+3} = d_4\}$$

$$\dim(\mathbf{C}) = 5^4 = 625$$

4. **Symmetrization of C – Dim. reduction** $625 \rightarrow 169$
   Sign-symmetry: $\mathbf{C}_{d_1 d_2 d_3 d_4} \leftarrow \mathbf{C}_{d_1 d_2 d_3 d_4} + \mathbf{C}_{-d_1 -d_2 -d_3 -d_4}$
   Directional symmetry: $\mathbf{C}_{d_1 d_2 d_3 d_4} \leftarrow \mathbf{C}_{d_1 d_2 d_3 d_4} + \mathbf{C}_{d_4 d_3 d_2 d_1}$
5. **Ensemble classifier [Kodovský-2011]**

Predictor Parametrization (structure)

Each predictor will be parametrized, for instance



$$\Longrightarrow \quad \mathbf{K} = \begin{pmatrix} 0 & d & c & d & 0 \\ d & b & a & b & d \\ c & a & 0 & a & c \\ d & b & a & b & d \\ 0 & d & c & d & 0 \end{pmatrix}$$

- **Parameters** a, b, c, d
- **Sum over all elements must equal to** 1
- **Free parameters** b, c, d **since** a **can be computed from the rest**

Optimization Methodology

Optimized parameters

- **Free parameters of the predictor structure**
- **Quantization step q**

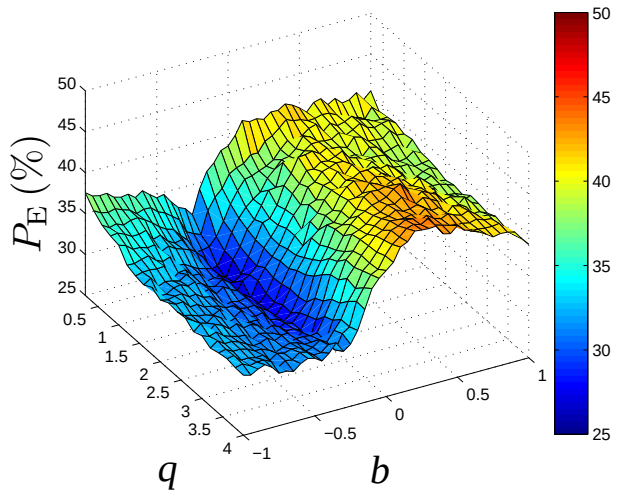
Objective function

- `L2R_L2LOSS` **(margin width of linear SVM) proposed by [Filler-2011] – Problematic**
- $P_{\mathrm{E}} = \min_{P_{FA}} \frac{1}{2}\left(P_{FA} + P_{MD}\left(P_{FA}\right)\right)$ **calculated using ensemble classifier on a subset of 2000 images.**

Optimization method

Nelder-Mead – Derivative-free simplex-reflection algorithm

$$\mathbf{K} = \begin{pmatrix} b & a & b \\ a & 0 & a \\ b & a & b \end{pmatrix}$$



Cover Sources

Three image databases

- **BOSSbase ver. 0.92 [BOSS-2010] – 9074 images, grayscale, 7 cameras, resized to 512×512**
- **NRCS512 – 6644 images, grayscale, NRCS scans, two 512×512 cropped from the center of every image**
- **LEICA512 – 8626 images, grayscale, Leica M9, 18 Mpixels, two 512×512 cropped from the center of every image**

Steganographic Algorithms

Three stego algorithms

- **HUGO (Highly Undetectable steGO) [Pevný *et al.*-2010]**
- **EA (Edge-Adaptive) [Luo *et al.*-2010]**
- **± 1 embedding with optimal ternary coder**

Two payloads

- **0.1 bits per pixel (bpp)**
- **0.4 bits per pixel (bpp)**

Optimizing the 3×3 Predictor

We optimized symmetric 3×3 predictors with structure

$$\begin{pmatrix} b & a & b \\ a & 0 & a \\ b & a & b \end{pmatrix}$$

Predictor parameters: (a, b), q

(b = free parameter, q = quantization step)

Initial predictor parameters for optimization:

- **Optimal cover predictor in the LSE sense**
- $q = 1.5$

Reference predictors

- **Predictor derived by [Böhme&Ker-2008]:**

$$\mathbf{KB} = \begin{pmatrix} -0.25 & 0.5 & -0.25 \\ 0.5 & 0 & 0.5 \\ -0.25 & 0.5 & -0.25 \end{pmatrix}$$

- **Optimal 3×3 cover predictor in the LSE sense (LSE)**
- **Quantization q selected as best $q \in \{1, 1.25, 1.5, 1.75, 2\}$**

Optimization Results – RAW

Alg.	Pld.	Ker	BOSSbase (a, b), q	BOSSbase P_E	NRCS512 (a, b), q	NRCS512 P_E	LEICA512 (a, b), q	LEICA512 P_E
HUGO	0.1	KB	(0.50, -0.25), 1.00	43.90	(0.50, -0.25), 2.00	48.62	(0.50, -0.25), 1.75	38.13
		LSE	(0.45, -0.20), 2.00	44.31	(0.51, -0.26), 1.75	48.90	(0.48, -0.23), 1.50	38.43
		Opt	(0.49, -0.24), 2.00	43.78	(0.60, -0.35), 1.69	48.86	(0.57, -0.32), 1.52	36.54
	0.4	KB	(0.50, -0.25), 1.00	26.37	(0.50, -0.25), 1.00	43.95	(0.50, -0.25), 1.75	13.58
		LSE	(0.45, -0.20), 1.50	27.65	(0.51, -0.26), 2.00	43.91	(0.48, -0.23), 1.50	13.35
		Opt	(0.51, -0.26), 1.58	26.49	(0.37, -0.12), 2.37	43.50	(0.38, -0.13), 1.98	12.07
EA	0.1	KB	(0.50, -0.25), 2.00	37.85	(0.50, -0.25), 2.00	47.66	(0.50, -0.25), 2.00	24.77
		LSE	(0.45, -0.20), 2.00	35.64	(0.51, -0.26), 1.75	47.66	(0.48, -0.23), 2.00	23.94
		Opt	(0.46, -0.21), 1.91	35.42	(0.67, -0.42), 1.84	47.36	(0.37, -0.12), 2.34	17.96
	0.4	KB	(0.50, -0.25), 1.75	17.93	(0.50, -0.25), 1.00	39.56	(0.50, -0.25), 1.75	4.62
		LSE	(0.45, -0.20), 1.75	16.00	(0.51, -0.26), 1.50	39.48	(0.48, -0.23), 2.00	4.30
		Opt	(0.26, -0.01), 1.92	13.74	(0.39, -0.14), 1.58	37.06	(0.40, -0.15), 2.09	3.52
± 1	0.1	KB	(0.50, -0.25), 1.00	31.05	(0.50, -0.25), 1.00	47.82	(0.50, -0.25), 1.00	36.89
		LSE	(0.45, -0.20), 1.00	32.56	(0.51, -0.26), 1.50	48.54	(0.48, -0.23), 1.50	38.19
		Opt	(0.55, -0.30), 0.58	31.42	(0.67, -0.42), 0.72	47.41	(0.56, -0.31), 0.93	37.11
	0.4	KB	(0.50, -0.25), 1.00	12.50	(0.50, -0.25), 1.00	40.52	(0.50, -0.25), 1.00	10.49
		LSE	(0.45, -0.20), 1.00	13.66	(0.51, -0.26), 1.00	41.99	(0.48, -0.23), 1.50	11.09
		Opt	(0.52, -0.27), 1.03	12.48	(0.73, -0.48), 0.55	39.70	(0.32, -0.07), 1.27	8.28

Interpretation of EA Results (1/2)

EA, BOSSbase, payload 0.4 bpp

- **KB** $= \begin{pmatrix} -0.25 & 0.5 & -0.25 \\ 0.5 & 0 & 0.5 \\ -0.25 & 0.5 & -0.25 \end{pmatrix} \longrightarrow P_{\mathrm{E}} = 17.93\%$
- **Opt** $= \begin{pmatrix} -0.01 & 0.26 & -0.01 \\ 0.26 & 0 & 0.26 \\ -0.01 & 0.26 & -0.01 \end{pmatrix} \longrightarrow P_{\mathrm{E}} = 13.74\%$

Why?

Message is embedded only to horizontal/vertical pixel pairs depending only their value difference.

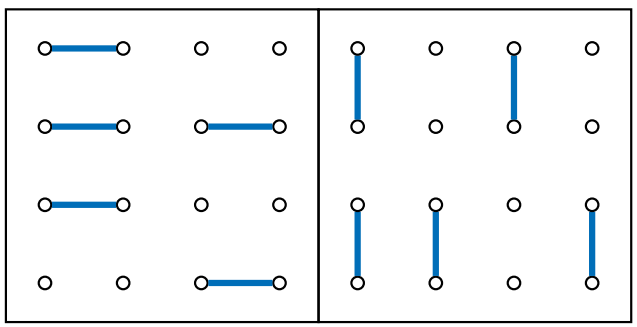
$\Longrightarrow$ **Adding diagonal neighbors does not improve steganalysis.**

Interpretation of EA Results (2/2)

EA algorithm

- **Image is divided into square blocks of a randomly selected size** $B \times B$, $B \in \{1, 4, 8, 12\}$
- **Every block is randomly rotated by** d **degrees,** $d \in \{0, 90, 180, 270\}$
- **Embedding into two horizontally neighboring pixels** $(x_{i,j}, x_{i,j+1})$**,**i **odd, where** $x_{i,j} - x_{i,j+1} > T$**. At most one value from the pair is modified.**
- **Blocks are rotated back to their original direction.**

$B = 4$



rot. 180° **rot.** 90°

Interpretation of LEICA512 Results

± 1, LEICA512, payload 0.4 bpp

$$S = \begin{pmatrix} b & a & b \\ a & 0 & a \\ b & a & b \end{pmatrix}$$

- **KB** $= \begin{pmatrix} -0.25 & 0.5 & -0.25 \\ 0.5 & 0 & 0.5 \\ -0.25 & 0.5 & -0.25 \end{pmatrix} \longrightarrow P_{\mathrm{E}} = 10.49\%$
- **Opt** $= \begin{pmatrix} -0.07 & 0.32 & -0.07 \\ 0.32 & 0 & 0.32 \\ -0.07 & 0.32 & -0.07 \end{pmatrix} \longrightarrow P_{\mathrm{E}} = 8.28\%$

LEICA512 images are 512×512 crops of 18 Mpix originals

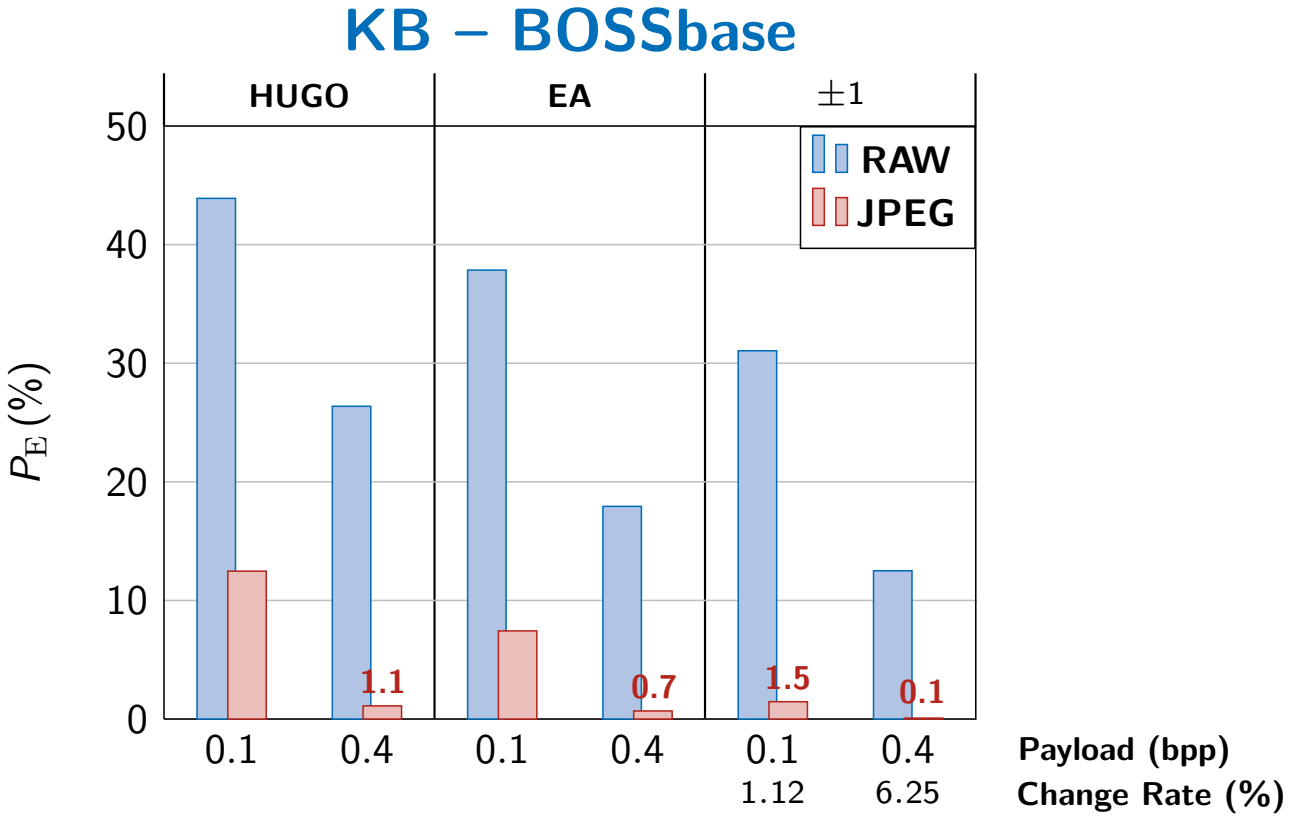
$\Longrightarrow$ **Strong dependencies among neighboring pixels**

$\Longrightarrow$ **[Böhme-2008] recommends optimal LSE predictors for steganalysis satisfying $\left|\frac{a}{b}\right| = \frac{1}{2\rho}$, where ρ is the correlation among neighboring pixels.**

$\Longrightarrow$ **In contrast, our study suggests that $\left|\frac{a}{b}\right|$ should increase**

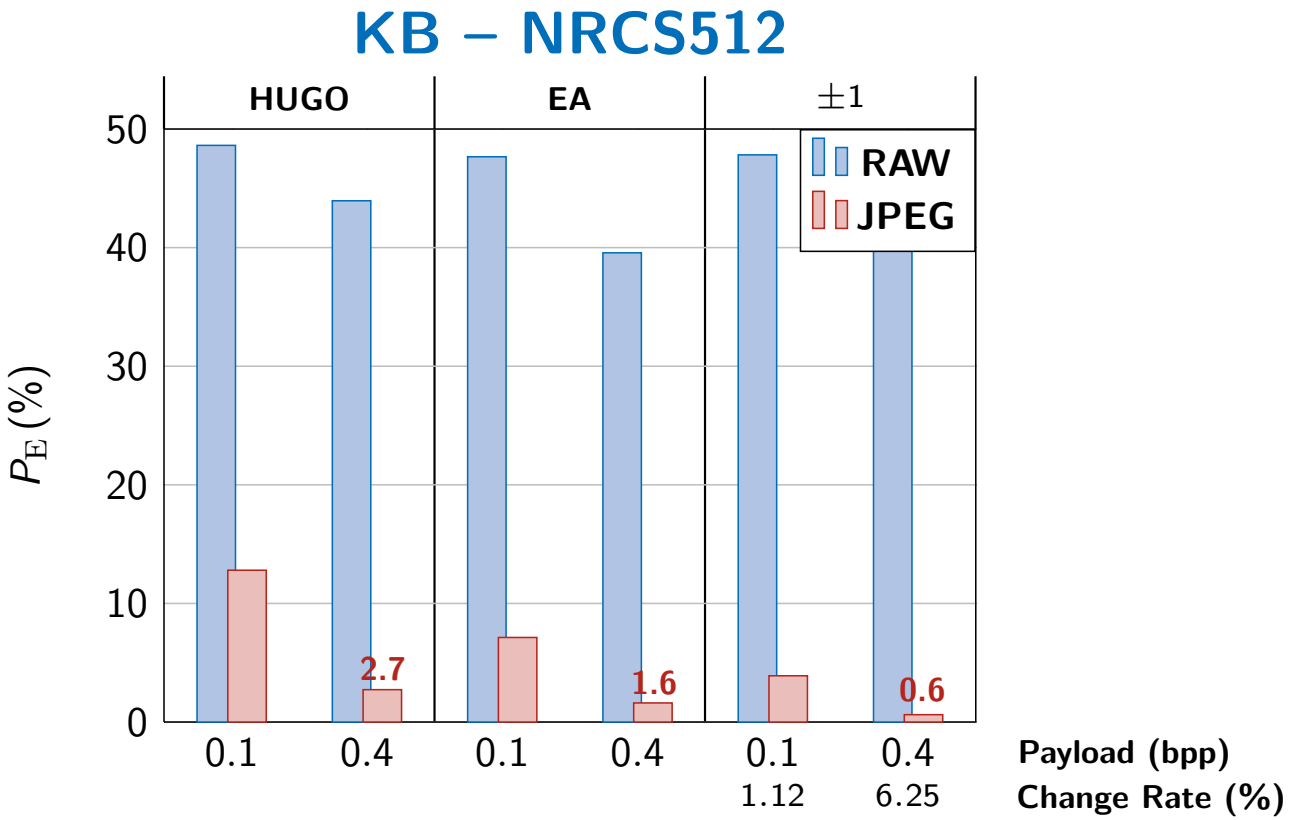
JPEG Results

- **RAW images compressed to 80% quality JPEG, then decompressed.**
- **Predictor optimization did not improve performance, why?**



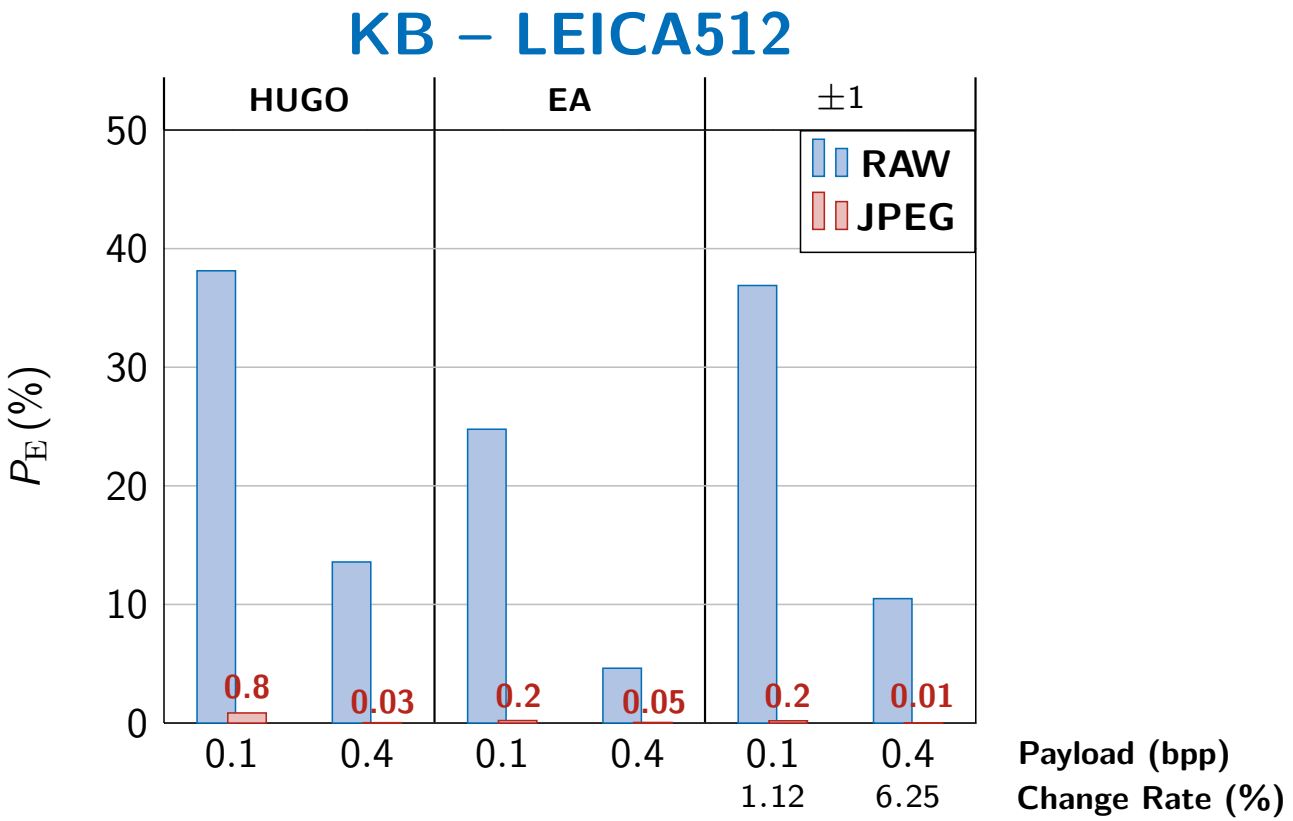
JPEG Results

- **RAW images compressed to 80% quality JPEG, then decompressed.**
- **Predictor optimization did not improve performance, why?**



JPEG Results

- **RAW images compressed to 80% quality JPEG, then decompressed.**
- **Predictor optimization did not improve performance, why?**



Interpretation of JPEG Results

1. **JPEG compression nearly empties some co-occurrence bins.**
2. **Embedding repopulates them from neighboring bins.**

Example: ± 1 embedding, BOSSbase 80, payload 0.4 bpp

k–best	bin	avg. RAW bin		avg. JPEG bin		JPEG P_{E}	
		Cover	Stego	Cover	Stego	indiv.	merged
1.	$(1,-1,2,-1)$,...	5889	7100	1407	3847	11.06	11.06
2.	$(1,1,0,0)$,...	3492	3481	5774	5220	**45.21**	**0.36**
3.	$(2,0,0,0)$,...	2644	2786	5874	4858	38.84	0.27

Detection exploits a cover-source singularity rather than effects of embedding.

Conditional optimization

- **Predictor optimization with respect to already existing predictors – cascading**

Example 1: HUGO, BOSSbase, 0.4 bpp

Structure	Optimized predictor, q	$P_{\mathrm{E}}^{\mathrm{indiv}}$	$P_{\mathrm{E}}^{\mathrm{merged}}$	Dim
$\begin{pmatrix} a & 0 & a \end{pmatrix}$	$\begin{pmatrix} 0.5 & 0 & 0.5 \end{pmatrix}$, 1.95	28.76	28.76	169
$\begin{pmatrix} a & 0 & b \end{pmatrix}$	$\begin{pmatrix} 0.048 & \mathbf{0} & \mathbf{-0.952} \end{pmatrix}$, 0.93	30.04	25.09	338

The second-order difference is optimally supplemented by the first-order difference

Conditional optimization

- **Predictor optimization with respect to already existing predictors – cascading**

Example 2: HUGO, BOSSbase, 0.4 bpp

Structure	Optimized predictor, q	P_E^{indiv}	P_E^{merged}	Dim
$\begin{pmatrix} b & a & b \\ a & 0 & a \\ b & a & b \end{pmatrix}$	$\begin{pmatrix} -0.259 & 0.509 & -0.259 \\ 0.509 & 0 & 0.509 \\ -0.259 & 0.509 & -0.259 \end{pmatrix}$, 1.58	26.49	26.49	169
$\begin{pmatrix} c & b & c \\ a & 0 & a \\ c & b & c \end{pmatrix}$	$\begin{pmatrix} -0.034 & \mathbf{0.503} & -0.034 \\ 0.064 & \mathbf{0} & 0.064 \\ -0.034 & \mathbf{0.503} & -0.034 \end{pmatrix}$, 2.23	27.22	21.77	338
$\begin{pmatrix} c & b & c \\ a & 0 & a \\ c & b & c \end{pmatrix}$	$\begin{pmatrix} -0.044 & -0.092 & -0.044 \\ \mathbf{0.682} & \mathbf{0} & \mathbf{0.682} \\ -0.044 & -0.09 & -0.044 \end{pmatrix}$, 2.03	32.21	20.25	507

Result comparable with HUGO BOSS winners only with 507 features

Summary

Predictor optimization for covers is a different problem than for a binary detection (cover/stego) within a framework.

Advantages

- **Noticeable improvement for some cover sources (LEICA512) and steganographic algorithms (EA).**
- **Conditional optimization to improve the performance–dimensionality ratio or to build a rich model.**

Limitations

- **Optimization only over a small parameter vector (e.g., up to dimension of five) due to noisy objective function.**

Other

- **Astonishingly accurate detection in decompressed JPEGs (future direction).**