Fourth-order Structural Steganalysis and Analysis of Cover Assumptions

Andrew Ker

adk@comlab.ox.ac.uk

Royal Society University Research Fellow
Oxford University Computing Laboratory

SPIE EI'06
16 January 2006

LSB Replacement

- Extremely simple
- Quick
- High capacity
- Visually imperceptible

LSB Replacement

- Extremely simple
- Quick
- High capacity
- Visually imperceptible

- Extremely vulnerable to statistical analysis

RS	[Fridrich et al, ACM MMS’01]
Pairs	[Fridrich et al, SPIE EI’03]
Sample Pairs	[Dumitrescu et al, IHW’02]
Triples	[Ker, IHW’05]

…

LSB Replacement

- Extremely simple
- Quick
- High capacity
- Visually imperceptible

- Extremely vulnerable to statistical analysis

RS	[Fridrich et al, ACM MMS'01]
Pairs	[Fridrich et al, SPIE EI'03]
Sample Pairs	[Dumitrescu et al, IHW'02]
Triples	[Ker, IHW'05]
…	

The most sensitive detectors apply *structural steganalysis*.

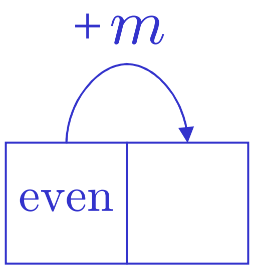
Structural property: even cover samples can only be incremented
odd cover samples can only be decremented

see [Ker, IHW'05]

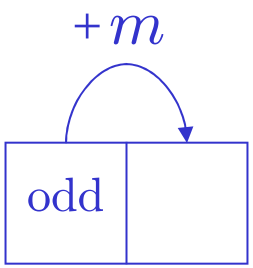
Trace Subsets

Classify all pairs of pixels

$\mathcal{E}_m$



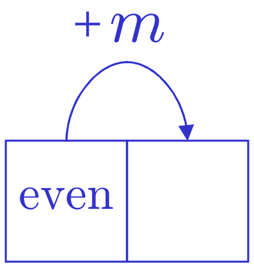
$\mathcal{O}_m$



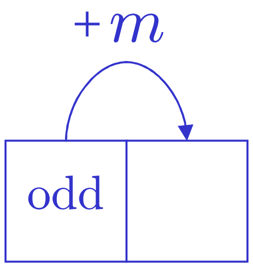
Trace Subsets

Classify all pairs of pixels

$\mathcal{E}_m$



$\mathcal{O}_m$

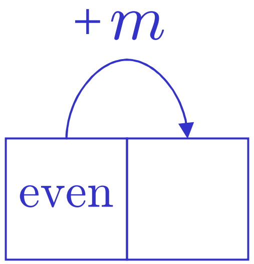


... and all triples of pixels

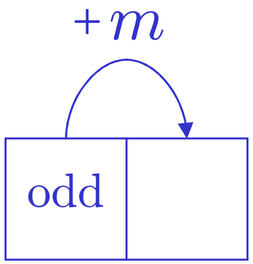
Trace Subsets

Classify all pairs of pixels

$\mathcal{E}_m$

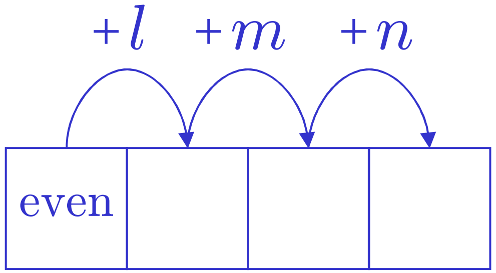


$\mathcal{O}_m$

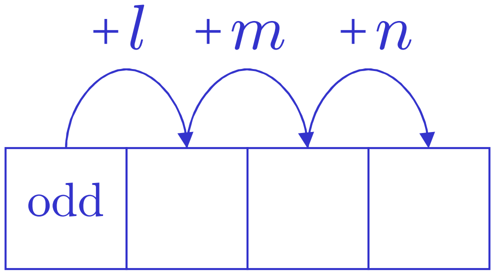


... and all triples of pixels ... and all quadruplets

$\mathcal{E}_{l\,,m,n}$



$\mathcal{O}_{l,m,n}$



Trace Subsets

Examples:

10	11	12	13

$\mathcal{E}_{1,1,1}$

21	24	24	21

$\mathcal{O}_{3,0,-3}$

Trace Subsets

Fix a cover. Embed a random message of proportionate length p**.**

The structural framework in [Ker, IHW'05] relates the sizes of the trace subsets, before and after embedding.

Sizes of $\mathcal{E}_m$ *and* $\mathcal{O}_m$
in stego image

Sizes of $\mathcal{E}_m$ *and* $\mathcal{O}_m$
in cover image

$$\begin{pmatrix} \downarrow \\ \vdots \\ E_{-1} \\ O_0 \\ E_0 \\ O_1 \\ E_1 \\ O_2 \\ \vdots \end{pmatrix} \approx \begin{pmatrix} \text{some matrix} \\ \text{parameterized by } p \end{pmatrix} \begin{pmatrix} \downarrow \\ \vdots \\ e_{-1} \\ o_0 \\ e_0 \\ o_1 \\ e_1 \\ o_2 \\ \vdots \end{pmatrix}$$

2nd Order Structural Detector

- Uses trace subsets $\mathcal{E}_m$ and $\mathcal{O}_m$ (pairs of pixels)

- Linear system $$\underset{\substack{\uparrow\\ \textbf{cover}}}{\begin{pmatrix} e_{2m} \\ o_{2m-1} \\ e_{2m+1} \\ o_{2m} \end{pmatrix}} \approx \frac{1}{(1-2p)^2} \begin{pmatrix} (1-p)^2 & -p(1-p) & -p(1-p) & p^2 \\ -p(1-p) & (1-p)^2 & p^2 & -p(1-p) \\ -p(1-p) & p^2 & (1-p)^2 & -p(1-p) \\ p^2 & -p(1-p) & -p(1-p) & (1-p)^2 \end{pmatrix} \underset{\substack{\uparrow\\ \textbf{stego}}}{\begin{pmatrix} E_{2m} \\ O_{2m-1} \\ E_{2m+1} \\ O_{2m} \end{pmatrix}}$$

- Cover assumptions: $e_m \approx o_m$ for odd m

⇨ Well-known estimator for p which we call *Couples* steganalysis

4th Order Structural Detector

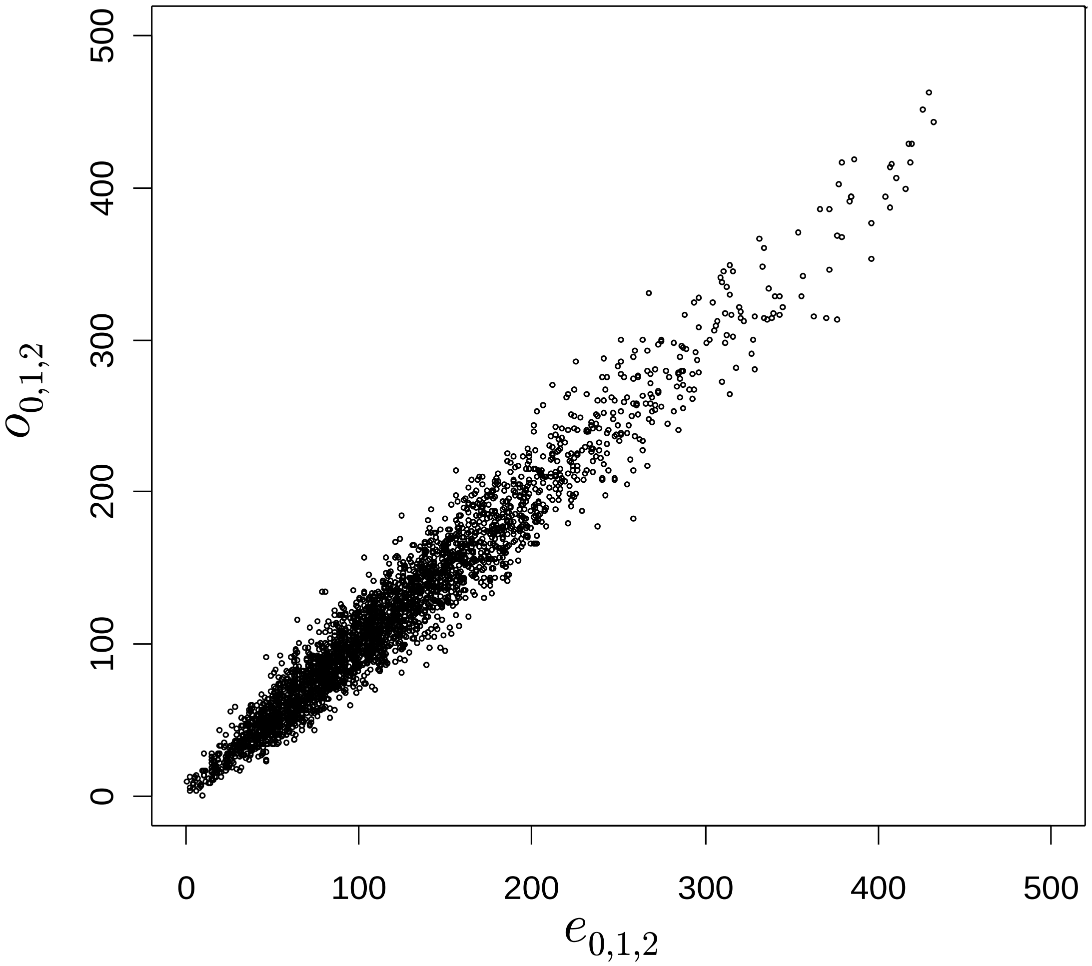
- Uses trace subsets $\mathcal{E}_{l,m,n}$ and $\mathcal{O}_{l,m,n}$ (quadruples of pixels)

- Linear system

$$\underset{\substack{\uparrow \\ \textbf{cover}}}{\begin{pmatrix} e_{2l,2m,2n} \\ o_{2l-1,2m,2n} \\ e_{2l+1,2m-1,2n} \\ o_{2l,2m-1,2n} \\ e_{2l,2m+1,2n-1} \\ o_{2l-1,2m+1,2n-1} \\ e_{2l+1,2m,2n-1} \\ o_{2l,2m,2n-1} \\ e_{2l,2m,2n+1} \\ o_{2l-1,2m,2n+1} \\ e_{2l+1,2m-1,2n+1} \\ o_{2l,2m-1,2n+1} \\ e_{2l,2m+1,2n} \\ o_{2l-1,2m+1,2n} \\ e_{2l+1,2m,2n} \\ o_{2l,2m,2n} \end{pmatrix}} \approx \begin{pmatrix} \text{16×16 matrix} \\ \text{parameterized by } p \end{pmatrix} \underset{\substack{\uparrow \\ \textbf{stego}}}{\begin{pmatrix} E_{2l,2m,2n} \\ O_{2l-1,2m,2n} \\ E_{2l+1,2m-1,2n} \\ O_{2l,2m-1,2n} \\ E_{2l,2m+1,2n-1} \\ O_{2l-1,2m+1,2n-1} \\ E_{2l+1,2m,2n-1} \\ O_{2l,2m,2n-1} \\ E_{2l,2m,2n+1} \\ O_{2l-1,2m,2n+1} \\ E_{2l+1,2m-1,2n+1} \\ O_{2l,2m-1,2n+1} \\ E_{2l,2m+1,2n} \\ O_{2l-1,2m+1,2n} \\ E_{2l+1,2m,2n} \\ O_{2l,2m,2n} \end{pmatrix}}$$

- Cover assumptions: **?**

Cover Symmetries



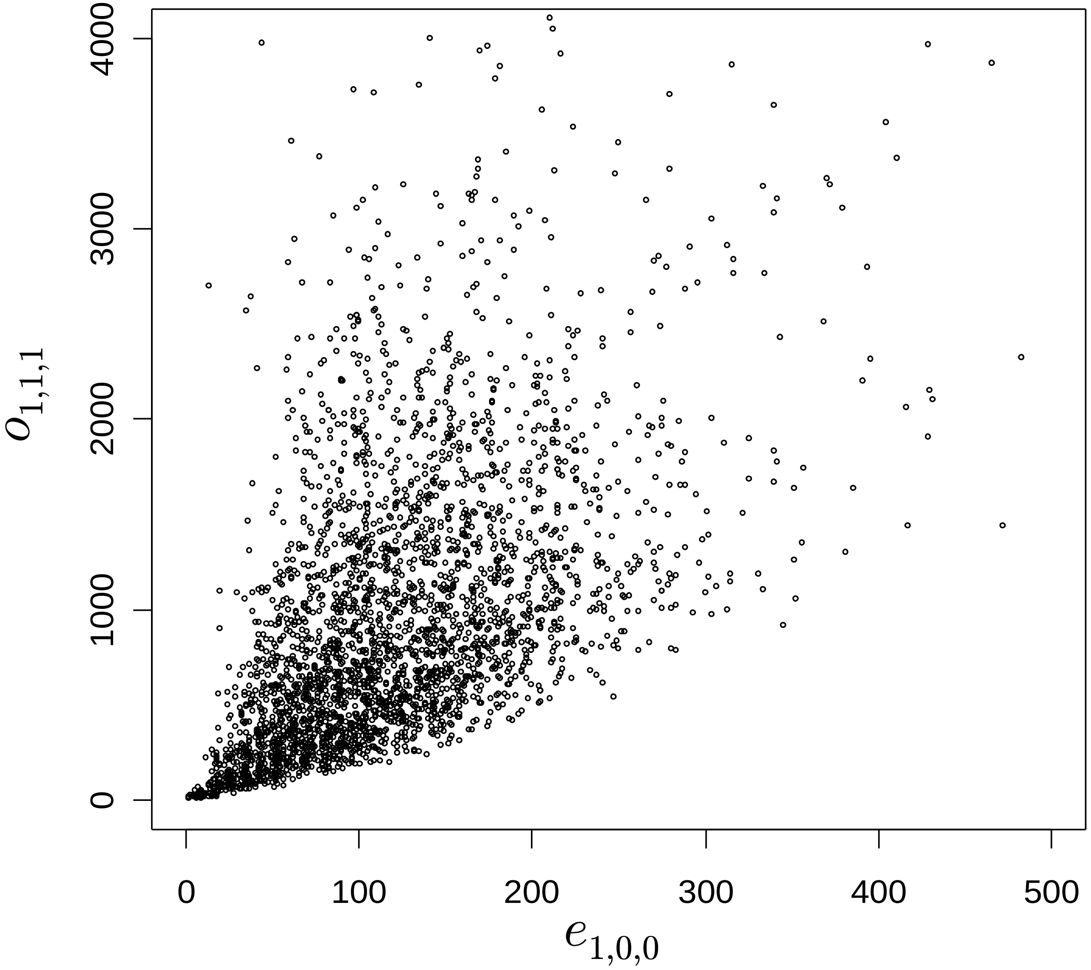
We say that

$e_{0,1,2} \approx o_{0,1,2}$

is a *symmetry*

Source: 3000 grayscale never-compressed images

Cover Symmetries



Not a symmetry

Source: 3000 grayscale never-compressed images

A Search for all Symmetries

Using a set of 3000 natural images,

- computed each $e_{l,m,n}$ and $o_{l,m,n}$

  (for l, m, n in the range -4 to 4: about 1500 trace subsets)

- for every pair of trace subsets computed the “closeness”*****.

  (about 1 million pairs)

- tried to find a small set of rules explaining all close trace subsets.

**sensible definition of closeness requires some care*

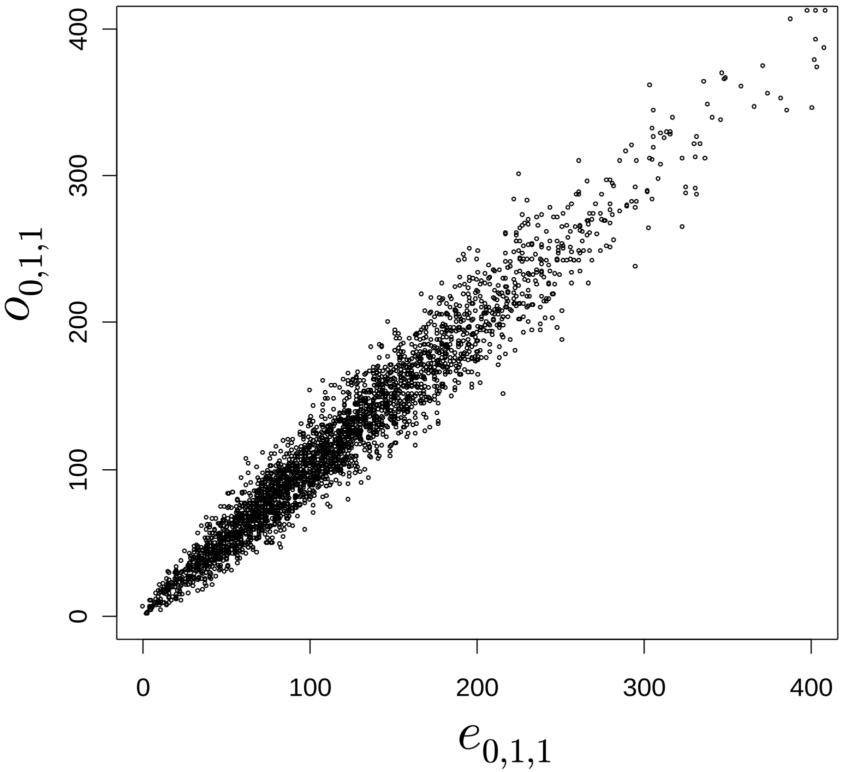
A Search for all Symmetries

All the symmetries we found were generated by:

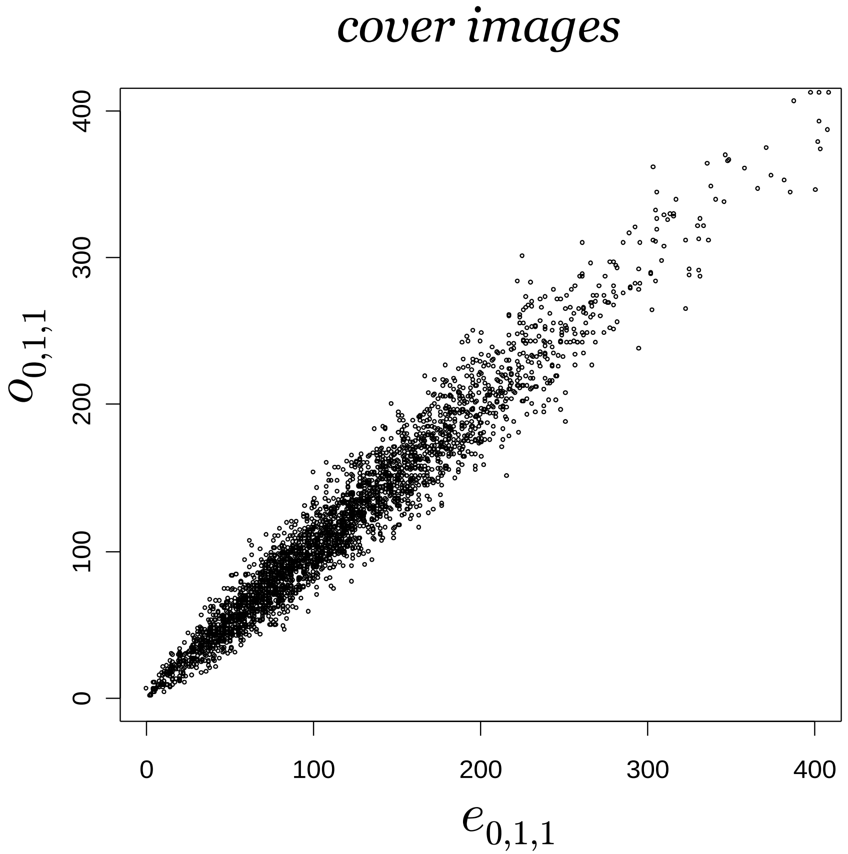
- Parity Symmetry: $e_{l,m,n} \approx o_{l,m,n}$
- Inversion Symmetry: $e_{l,m,n} \approx o_{-l,-m,-n}$
- Permutative Symmetry: $e_{l,m,n} \approx e_{\pi(l,m,n)}$
  for all cyclic permutations π

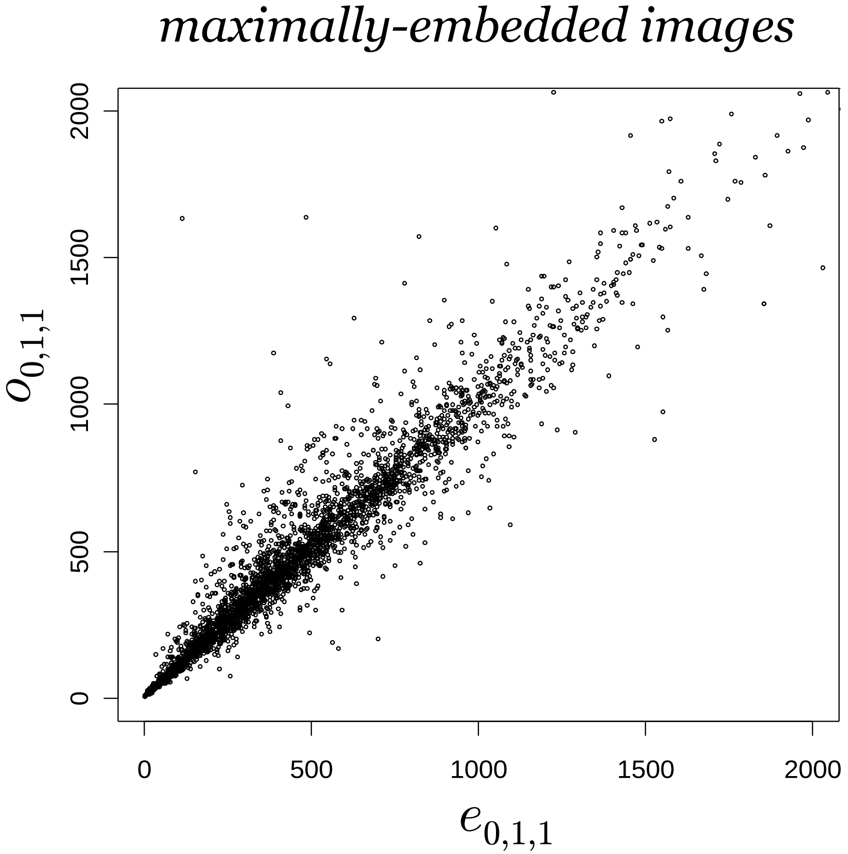
(not quite the whole story, see paper for details)

cover images

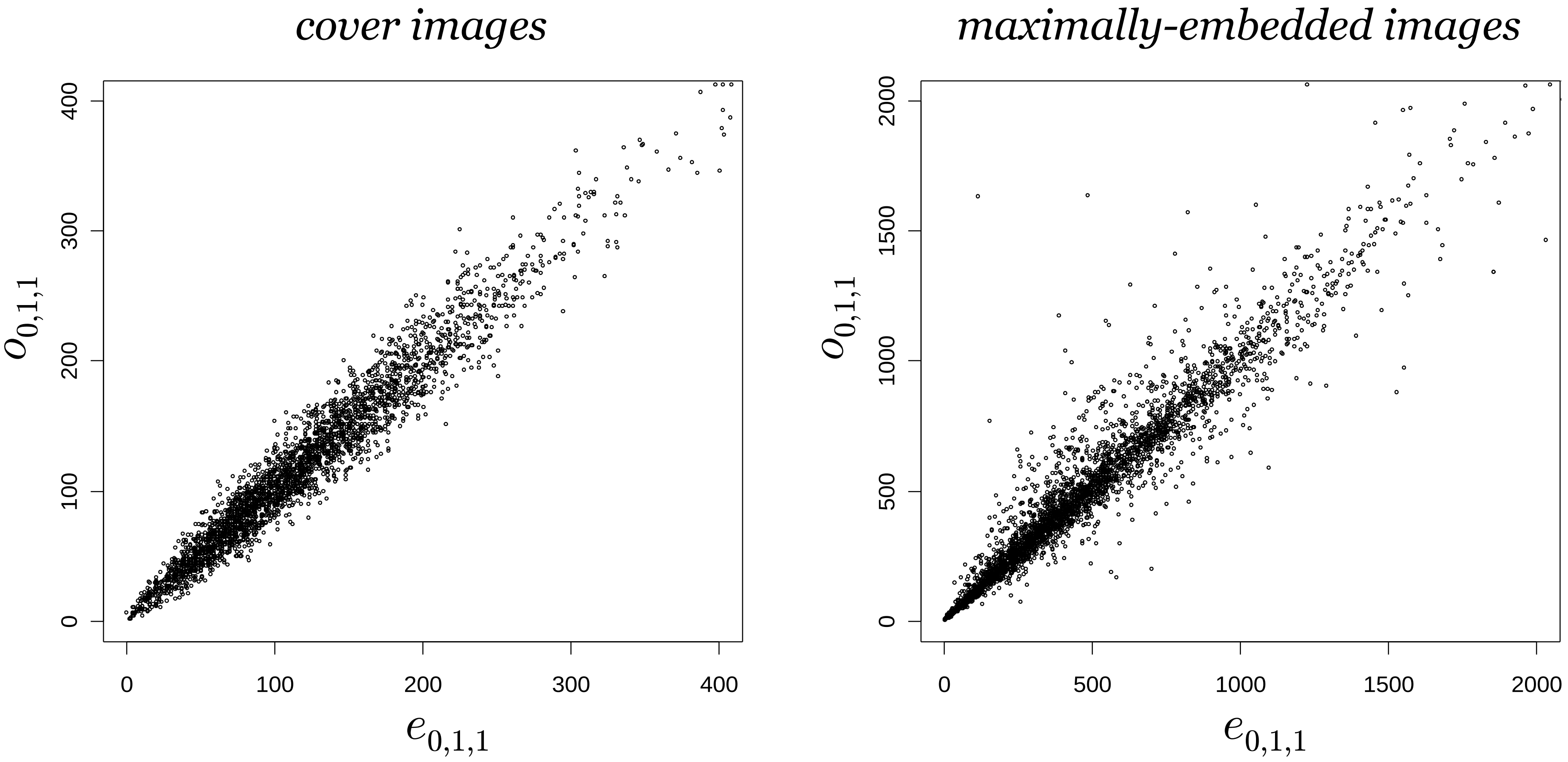


Source: 3000 grayscale never-compressed images


cover images
$o_{0,1,1}$
0
100
200
300
400
0
100
200
300
400
$e_{0,1,1}$

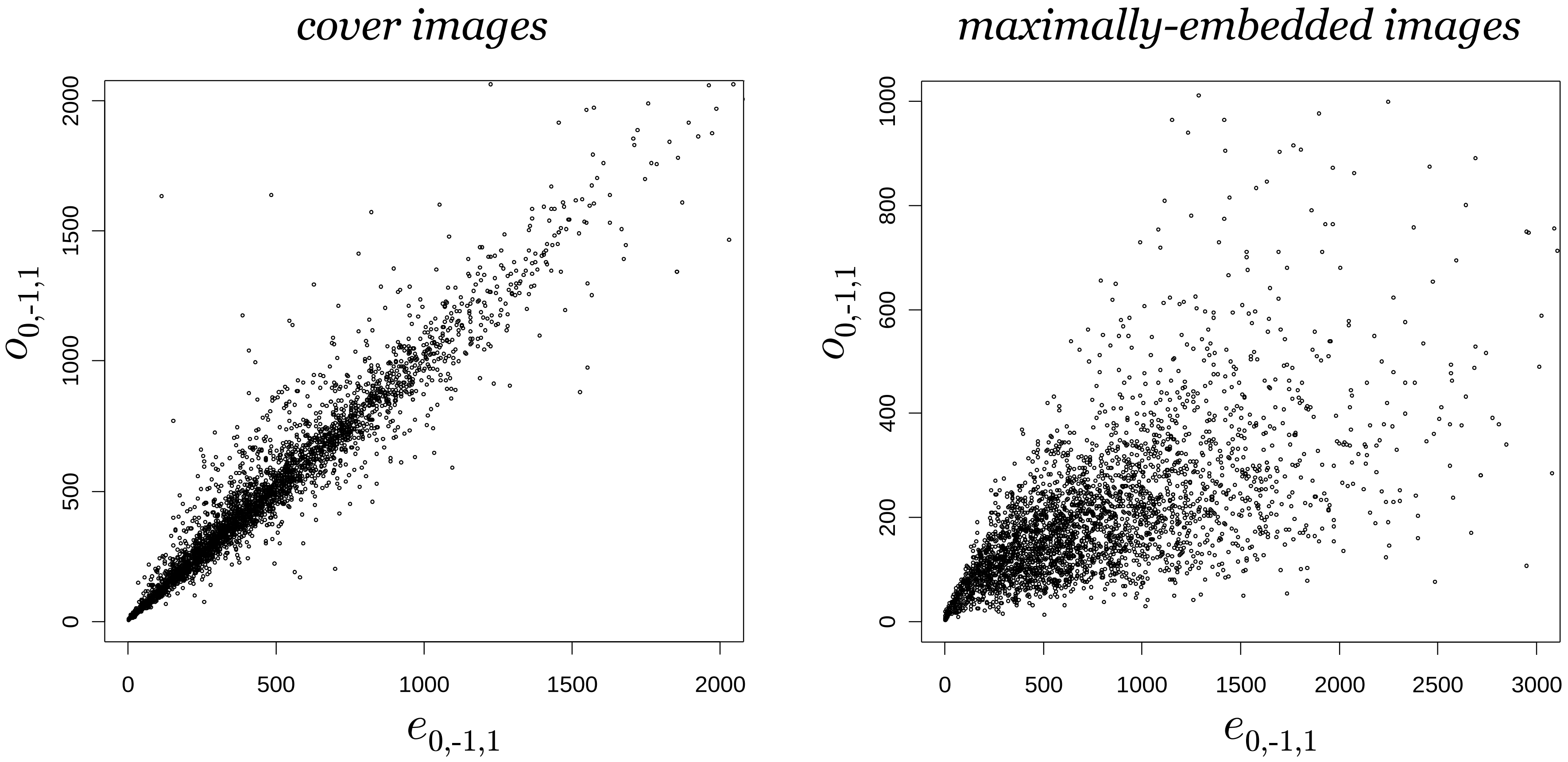

maximally-embedded images
$o_{0,1,1}$
0
500
1000
1500
2000
0
500
1000
1500
2000
$e_{0,1,1}$

$e_{0,1,1} \approx o_{0,1,1}$ **does not** discriminate covers from stego images



$e_{0,1,1} \approx o_{0,1,1}$ **does not** discriminate covers from stego images

$e_{0,-1,1} \approx o_{0,-1,1}$ **does** discriminate covers from stego images



Why Some Symmetries Fail To Discriminate

$\mathcal{E}_{0,1,1}$

$\mathcal{O}_{0,1,1}$

$e_{0,1,1} \approx o_{0,1,1}$

Why Some Symmetries Fail To Discriminate

$$\begin{tikzcd} \mathcal{E}_{0,1,1} \arrow[r, "\textit{Flip all LSBs}"] & \mathcal{O}_{0,-1,3} \end{tikzcd}$$

$$\begin{tikzcd} \mathcal{O}_{0,1,1} \arrow[r, "\textit{Flip all LSBs}"] & \mathcal{E}_{0,3,-1} \end{tikzcd}$$

$e_{0,1,1} \approx o_{0,1,1}$

$o_{0,-1,3} \approx e_{0,-1,3}$ *(parity symmetry)*
$\approx e_{0,3,-1}$ *(permutative symmetry)*

A Search for All Symmetries

All the symmetries we found were generated by:

- Parity Symmetry: $e_{l,m,n} \approx o_{l,m,n}$
- Inversion Symmetry: $e_{l,m,n} \approx o_{-l,-m,-n}$
- Permutative Symmetry: $e_{l,m,n} \approx e_{\pi(l,m,n)}$
  for all cyclic permutations π

A Search for All Symmetries

All the discriminating symmetries we found were generated by:

- Parity Symmetry: $e_{l,m,n} \approx o_{l,m,n}$
  if one or three of l, m and n are odd, or two of them are odd and not equal
- ~~Inversion Symmetry: $e_{l,m,n} \approx o_{-l,-m,-n}$~~
- Permutative Symmetry: $e_{l,m,n} \approx e_{\pi(l,m,n)}$
  for all cyclic permutations π such that (a fairly complex condition holds)

A Search for All Symmetries

All the discriminating symmetries we found were generated by:

- Parity Symmetry: $e_{l,m,n} \approx o_{l,m,n}$
  if one or three of l, m and n are odd, or two of them are odd and not equal

- ~~Inversion Symmetry: $e_{l,m,n} \approx o_{-l,-m,-n}$~~

- Permutative Symmetry: $e_{l,m,n} \approx e_{\pi(l,m,n)}$
  for all cyclic permutations π such that (a fairly complex condition holds)

Also in the paper: variance stabilization, and independence, of deviation from cover symmetries

4th Order Structural Detector

- Uses trace subsets $\mathcal{E}_{l,m,n}$ and $\mathcal{O}_{l,m,n}$ (quadruples of pixels)
- Linear system

$$\underset{\substack{\uparrow\\ \textbf{cover}}}{\begin{pmatrix} e_{2l,2m,2n} \\ o_{2l-1,2m,2n} \\ e_{2l+1,2m-1,2n} \\ o_{2l,2m-1,2n} \\ e_{2l,2m+1,2n-1} \\ o_{2l-1,2m+1,2n-1} \\ e_{2l+1,2m,2n-1} \\ o_{2l,2m,2n-1} \\ e_{2l,2m,2n+1} \\ o_{2l-1,2m,2n+1} \\ e_{2l+1,2m-1,2n+1} \\ o_{2l,2m-1,2n+1} \\ e_{2l,2m+1,2n} \\ o_{2l-1,2m+1,2n} \\ e_{2l+1,2m,2n} \\ o_{2l,2m,2n} \end{pmatrix}} \approx \left(\begin{array}{c} 16\times16 \text{ matrix} \\ \text{parameterized by } p \end{array} \right) \underset{\substack{\uparrow\\ \textbf{stego}}}{\begin{pmatrix} E_{2l,2m,2n} \\ O_{2l-1,2m,2n} \\ E_{2l+1,2m-1,2n} \\ O_{2l,2m-1,2n} \\ E_{2l,2m+1,2n-1} \\ O_{2l-1,2m+1,2n-1} \\ E_{2l+1,2m,2n-1} \\ O_{2l,2m,2n-1} \\ E_{2l,2m,2n+1} \\ O_{2l-1,2m,2n+1} \\ E_{2l+1,2m-1,2n+1} \\ O_{2l,2m-1,2n+1} \\ E_{2l,2m+1,2n} \\ O_{2l-1,2m+1,2n} \\ E_{2l+1,2m,2n} \\ O_{2l,2m,2n} \end{pmatrix}}$$

- Cover assumptions:

$e_{l,m,n} \approx o_{l,m,n}$ for l, m, n such that …

$e_{l,m,n} \approx e_{\pi(l,m,n)}$ for l, m, n, π such that …

4th Order Structural Detector

- Cover assumptions:

$$e_{l,m,n} \approx o_{l,m,n} \quad \text{for } l, m, n \text{ such that …}$$

$$e_{l,m,n} \approx e_{\pi(l,m,n)} \quad \text{for } l, m, n, \pi \text{ such that …}$$

(also restricting to low values of l, m, n)

gives rise to 400 discriminating cover symmetries.

4th Order Structural Detector

- Cover assumptions: $e_{l,m,n} \approx o_{l,m,n}$ for l, m, n such that …

  $e_{l,m,n} \approx e_{\pi(l,m,n)}$ for l, m, n, π such that …

(also restricting to low values of l, m, n)

gives rise to 400 discriminating cover symmetries.

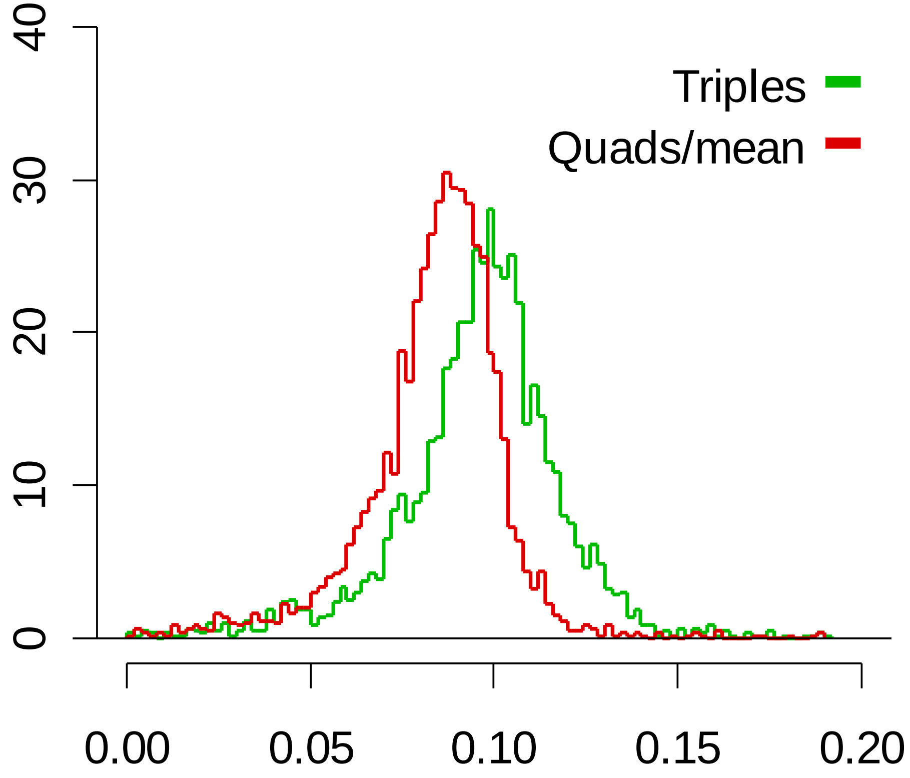
Each symmetry gives an equation for p; take all 400 and throw out:

- equations with no real root
- equations giving obviously-wrong answers (p<<0 or p>>1)

and take the mean or median of all remaining individual estimators for p.

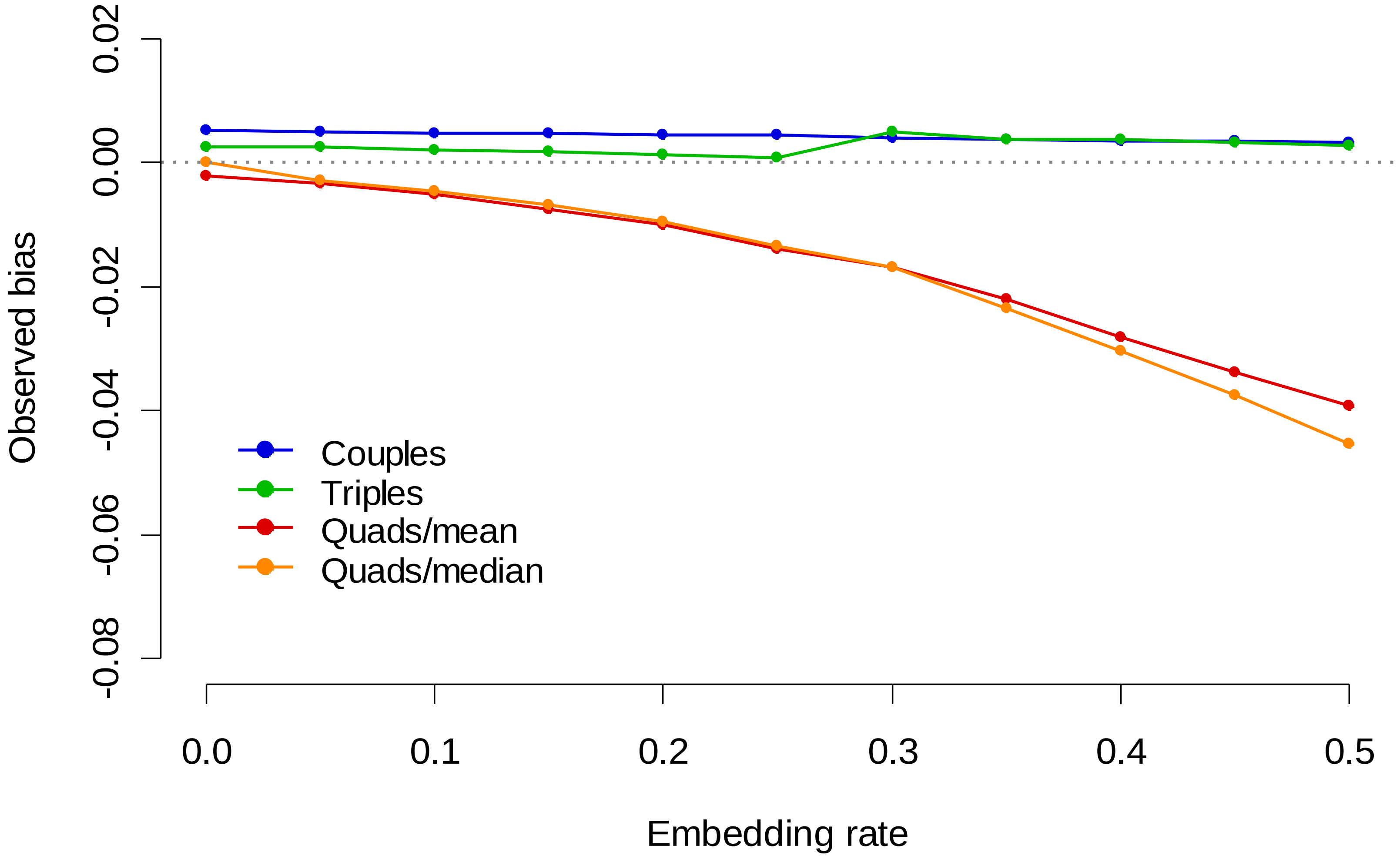
Detector Response

True embedding rate $p=0.1$



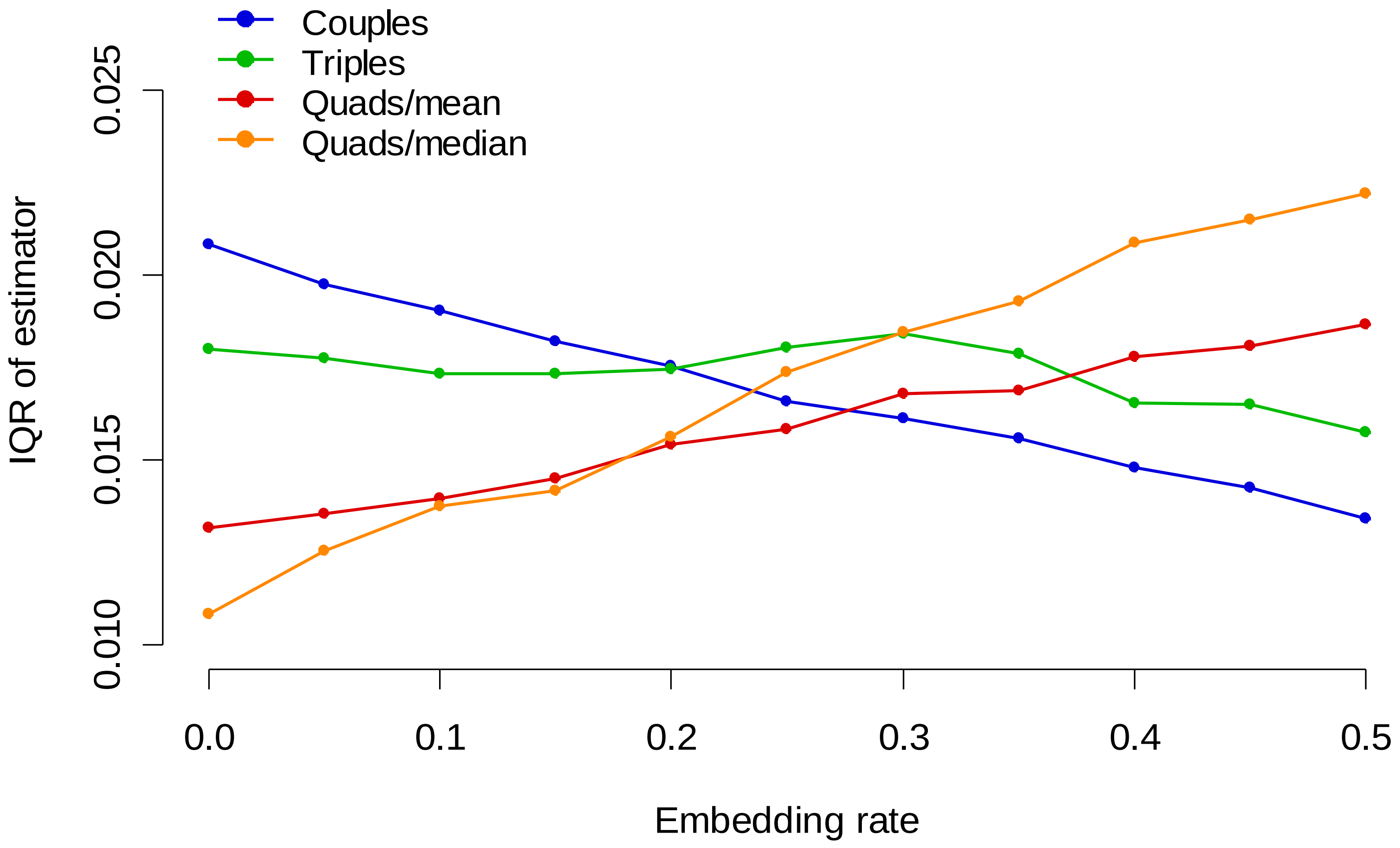
Source: 3000 colour never-compressed images

Observed Bias



Source: 3000 colour never-compressed images

Estimator Dispersion



Source: 3000 colour never-compressed images

Conclusions

- Have successfully made a structural detector based on quadruplets of pixels.

  The difficulty was in deciding the cover assumptions, which we determined by searching for symmetries in a set of natural images.

- The detector is not fully mature.

  Can we explain/correct the negative bias?
  Is there a better way to treat the hundreds of different equations estimating p?

- There is experimental evidence of (somewhat) improved performance.

  Further extension ("Quintuples Steganalysis") might not be valuable.
  Perhaps combination of trace subsets will provide progress.

Final Comparison

Table shows standard deviations of various estimators for no embedding (NB: Quads performance decreases as embedding rate increases)

	Never-compressed images		*JPEG compressed images*	
	Grayscale	*Colour*	*Grayscale*	*Colour*
RS	3.25	2.67	2.02	10.94
Couples (Sample Pairs)	3.29	2.56	1.73	8.56
Triples	3.45	2.36	1.90	2.08
Quadruples/mean	2.73	1.56	1.69	**2.03**
Quadruples/median	**2.56**	**1.45**	**1.62**	**2.03**

The End

adk@comlab.ox.ac.uk