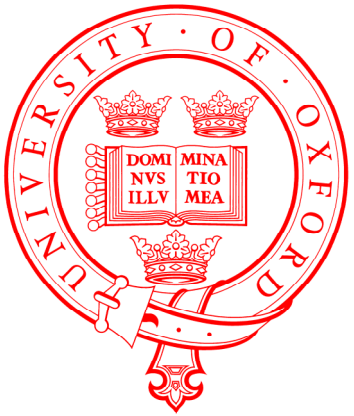
Resampling and the Detection of LSB Matching in Colour Bitmaps



Andrew Ker

adk@comlab.ox.ac.uk

Royal Society University Research Fellow
Oxford University Computing Laboratory

SPIE EI'05

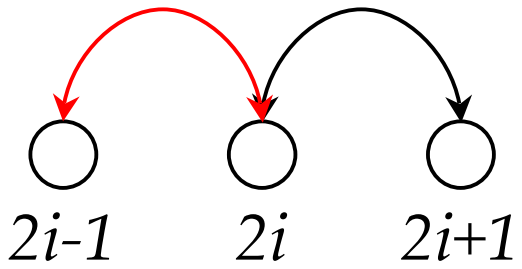
17 January 2005

LSB Matching

a.k.a. “plus/minus 1”

- *Consider cover in pseudorandom order*
- *Increment or decrement cover samples* ***at random*** *so that the LSBs match the hidden bit stream*

Differs from the standard LSB Replacement algorithm in that other bit planes may be changed.



LSB Matching

Why study a spatial-domain embedding method?

Because it can be performed without steganography software

LSB Matching

Why study a spatial-domain embedding method?

Because it can be performed without steganography software

```
perl -n0777 <cover-image.ppm >stego-image.ppm
     -e'split/(\s+)/,<STDIN>,5;@z=map ord,split"",pop@_;srand key;
        for(0..$#z){@p[$k,$_]=($_,$p[$k=int rand$_]);}
        map{$z[$q=shift@p]+=($z[$q]-ord()&1)*(rand 2<=>1)}
        split"",unpack"B*",$_;print@_,map chr,@z;' payload
```

LSB Matching

Why study a spatial-domain embedding method?

Because it can be performed without steganography software

```
perl -n0777 <cover-image.ppm >stego-image.ppm
     -e'split/(\s+)/,<STDIN>,5;@z=map ord,split"",pop@_;srand key;
        for(0..$#z){@p[$k,$_]=($_,$p[$k=int rand$_]);}
        map{$z[$q=shift@p]+=($z[$q]-ord()&1)*(rand 2<=>1)}
        split"",unpack"B*",$_;print@_,map chr,@z;' payload
```

- *impossible to prevent use of this mini-program*
- *if used carefully, probably undetectable*

Harmsen's HCF COM Detector

"Steganalysis of Additive Noise Modelable Information Hiding" [SPIE EI'03]

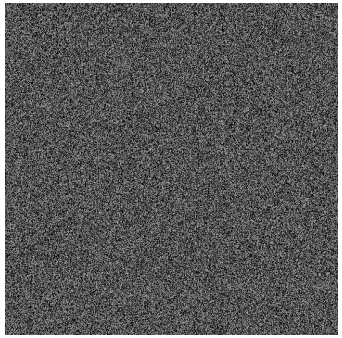
Model steganography as additive noise and examine the effects on the image histogram.

cover image

stego "noise"

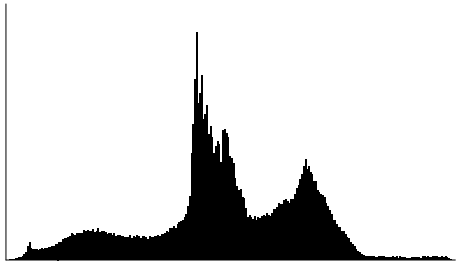
stego image

+

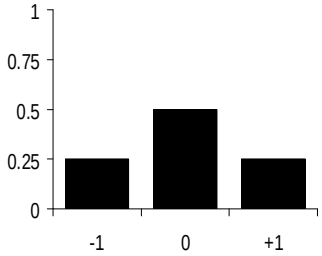


=

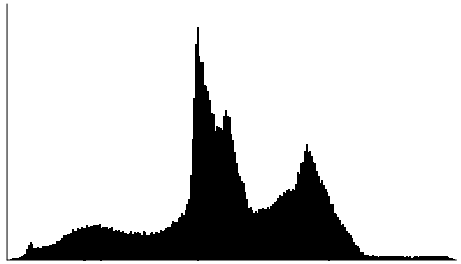
Histogram



*

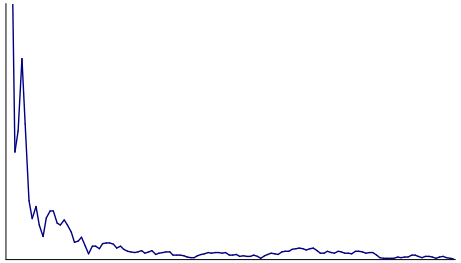


=

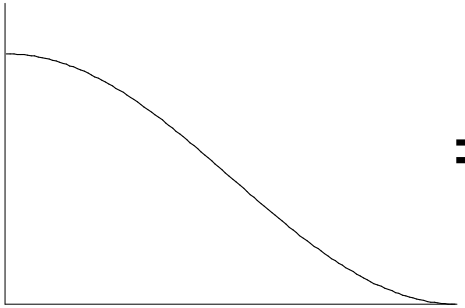


256-pt DFT
(first 128 points)

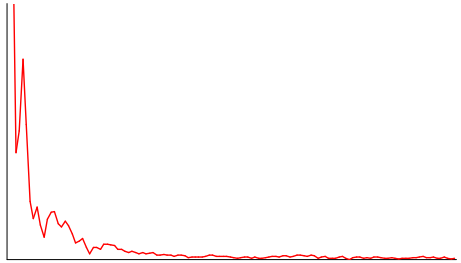
"HCF"



×



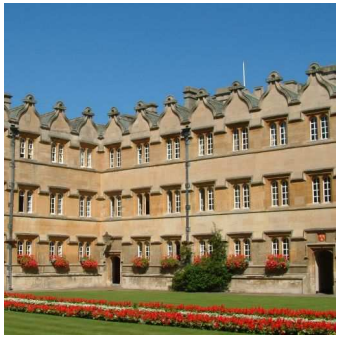
=



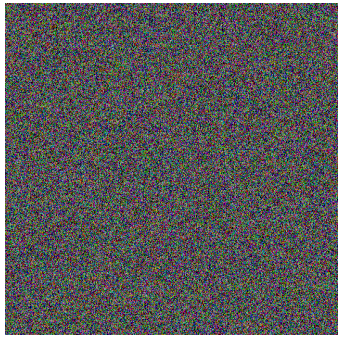
cover image

stego "noise"

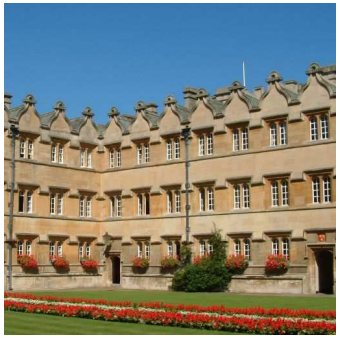
stego image



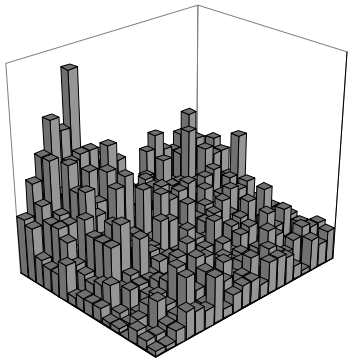
+



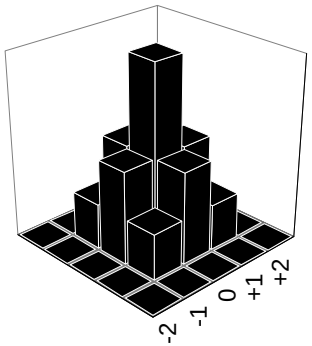
=



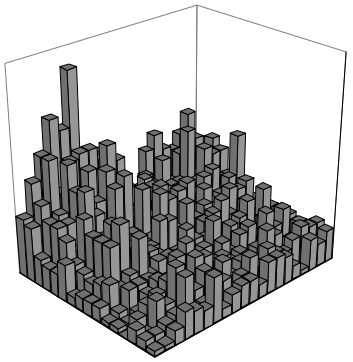
3D Histogram



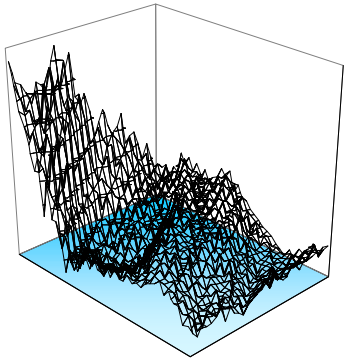
*



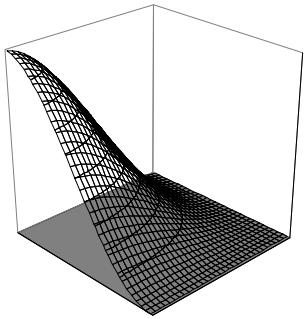
=



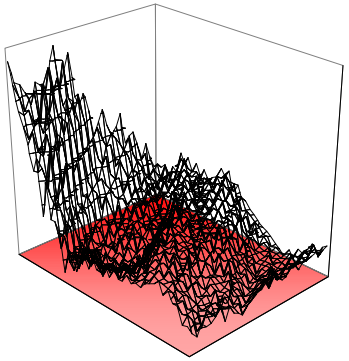
256^3-pt 3D DFT
(first 128^3 points)
"HCF"



×



=



3D "HCF COM"

(77, 77, 77)

(55, 54, 54)

HCF COM Detector Slogan

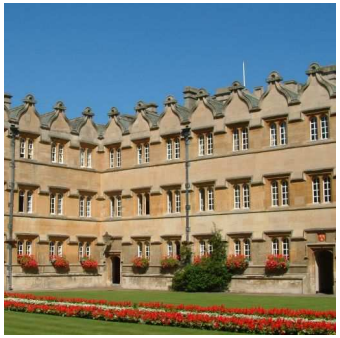
Steganography reduces the COM

(& longer messages reduce the COM by more than shorter messages)

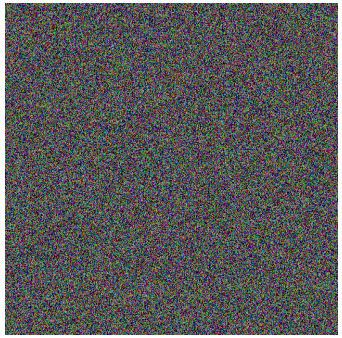
cover image

stego "noise"

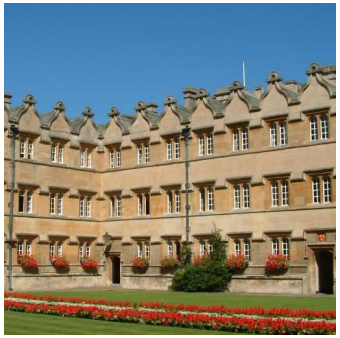
stego image



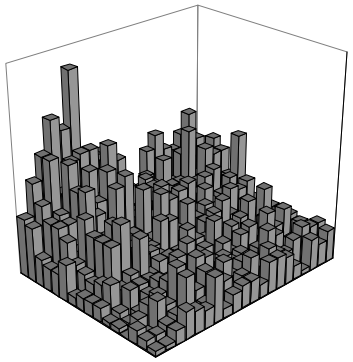
+



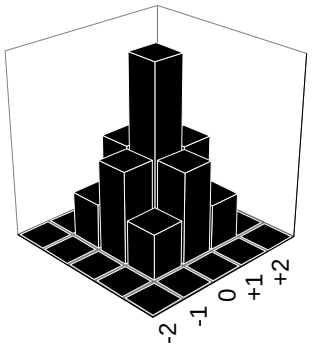
=



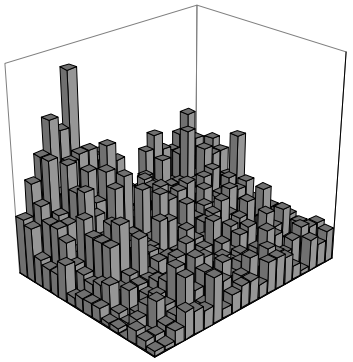
3D Histogram



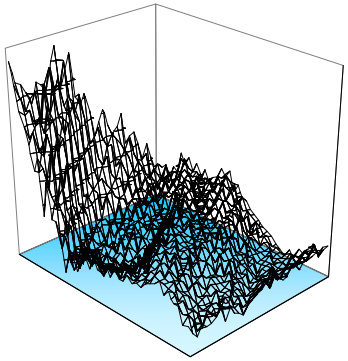
*



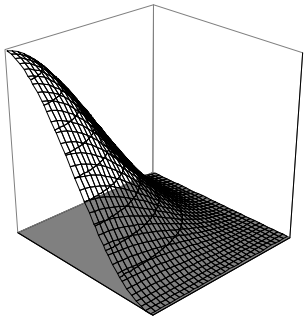
=



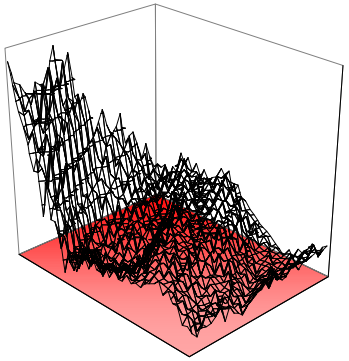
256^3-pt 3D DFT
(first 128^3 points)
"HCF"



×



=

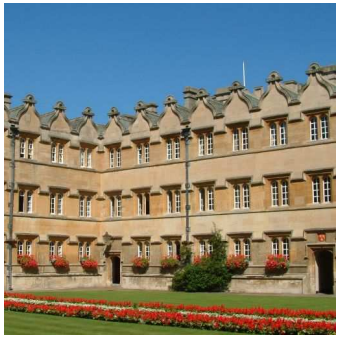


3D "HCF COM"

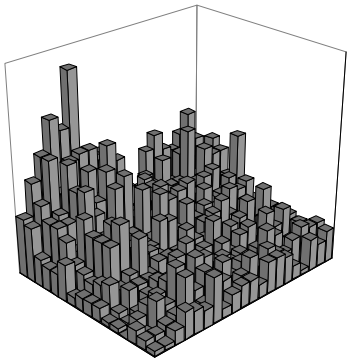
(77, 77, 77)

(55, 54, 54)

stego image

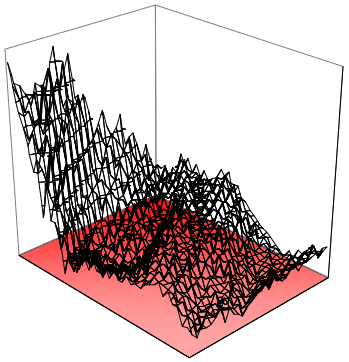


3D Histogram



unknown to detector

256^3-pt 3D DFT
(first 128^3 points)
"HCF"



3D "HCF COM"

(55, 54, 54)

Potential Problems

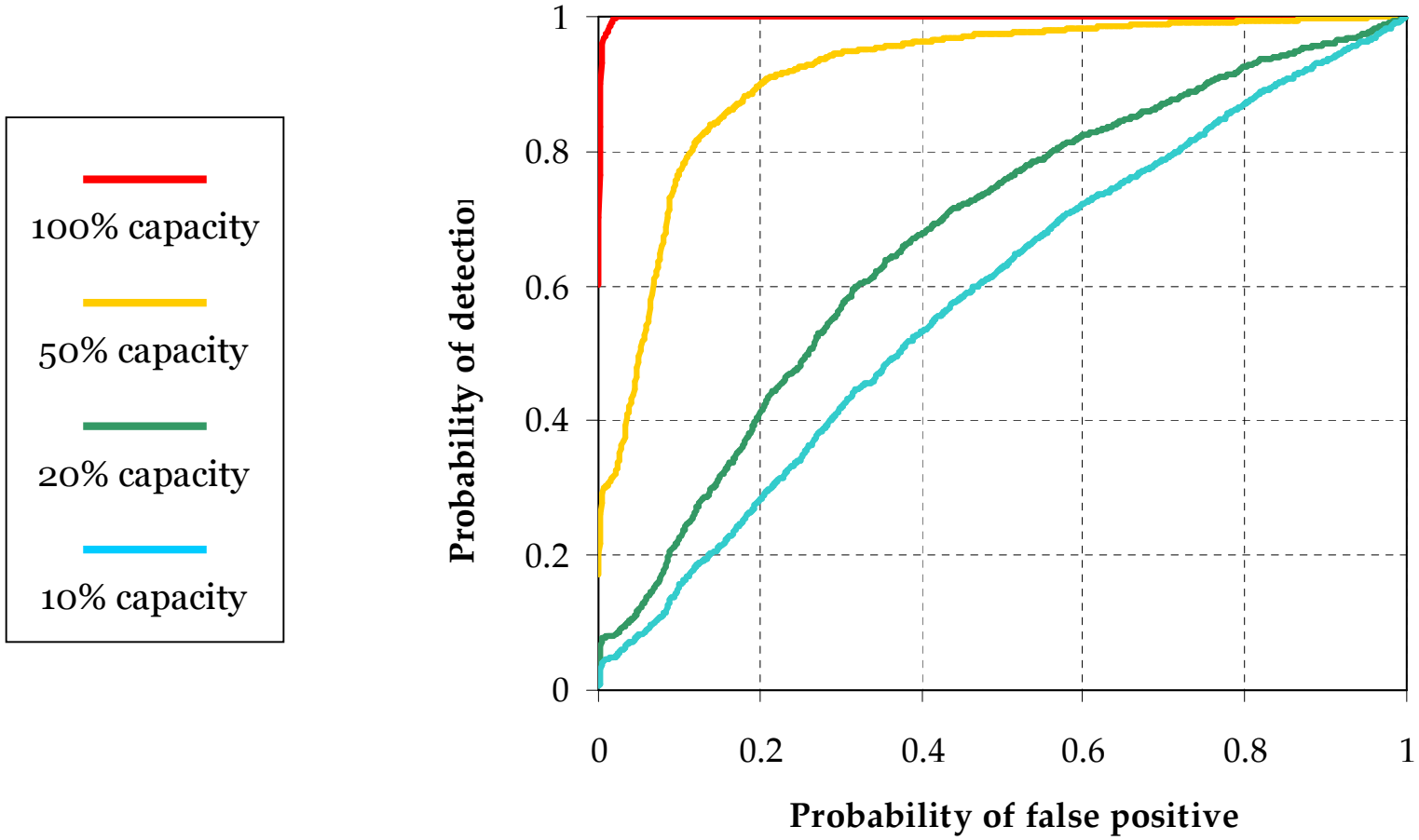
1. The detector cannot see the cover image – the COM cannot be compared with the cover COM.

2. This detector is detecting (any type of) noise, not just steganography.

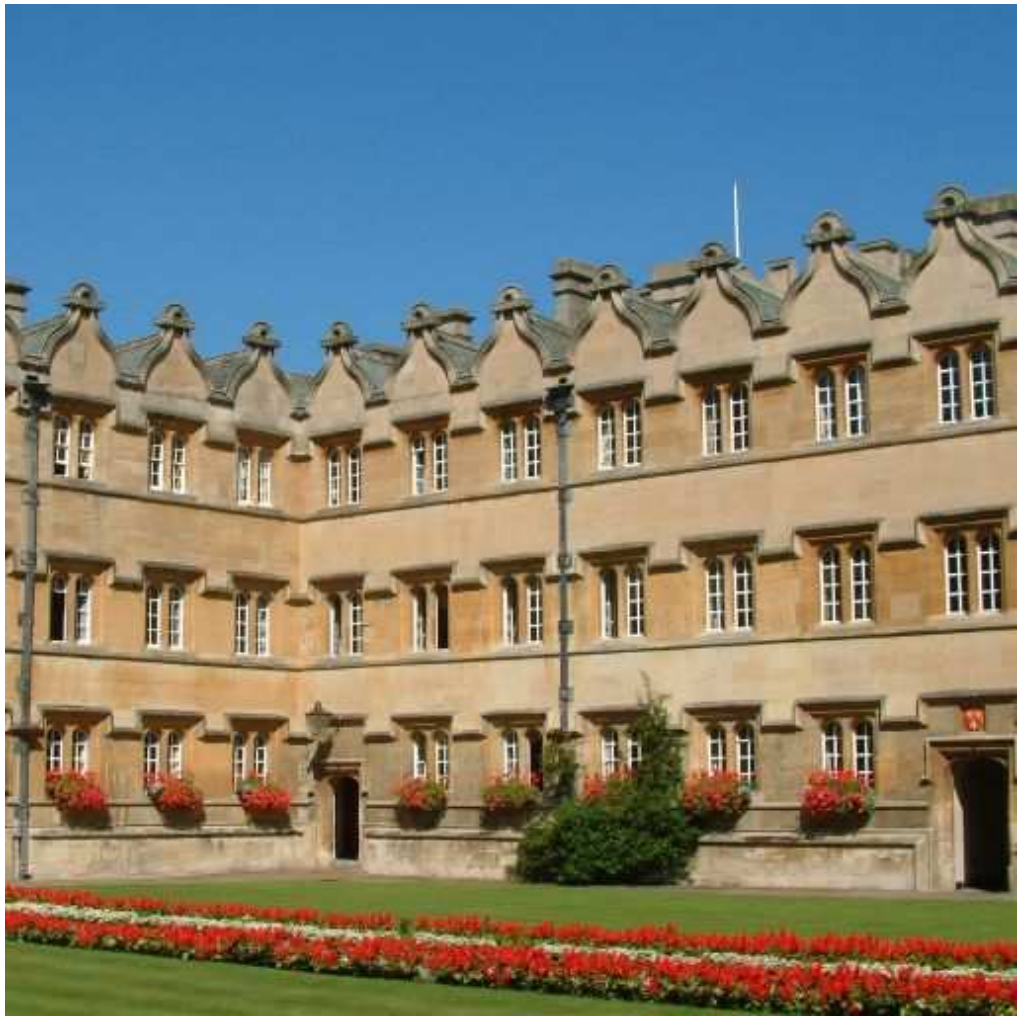
3. Methods which use only the histogram of the image are throwing away a lot of data.

Reliability

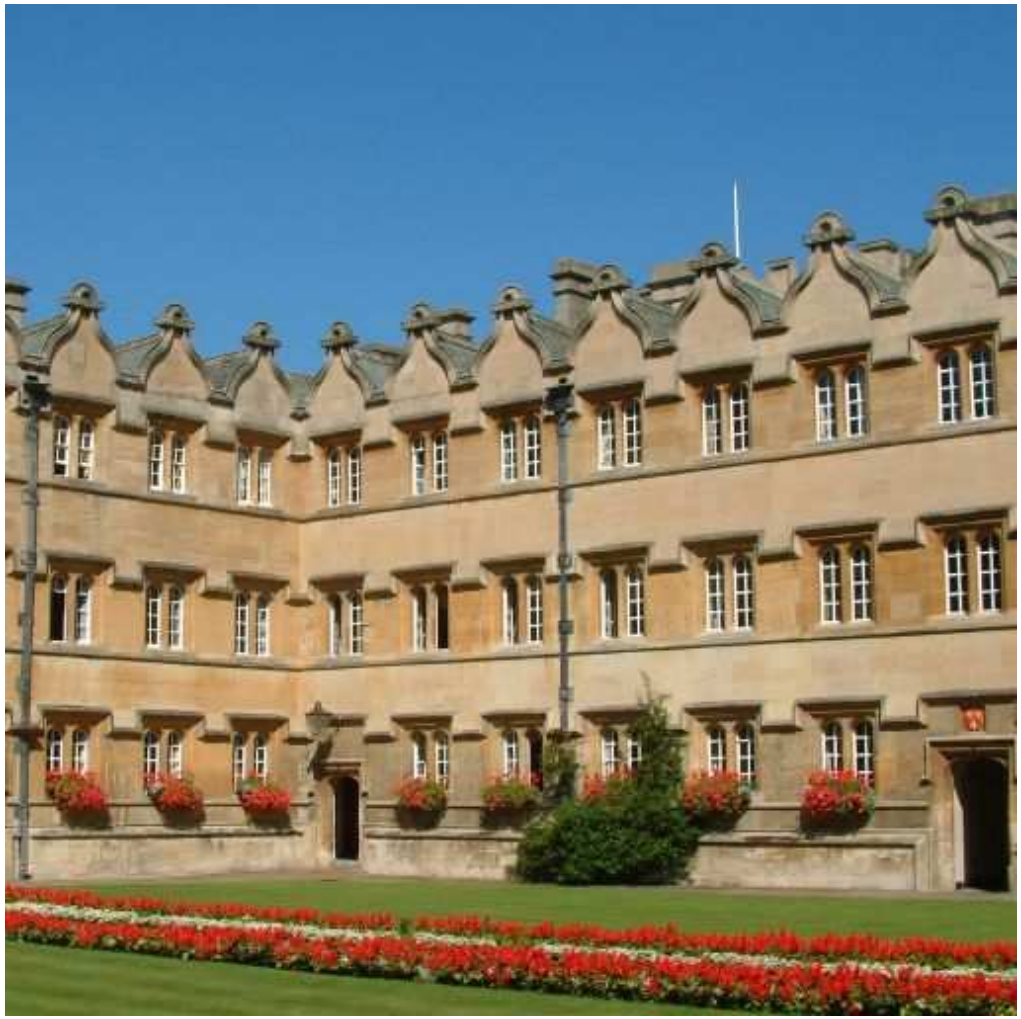
observed for 10000 colour bitmaps previously subject to moderate JPEG compression.

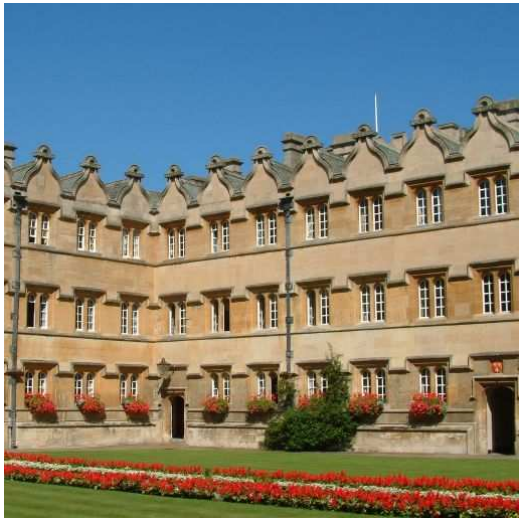


Calibration

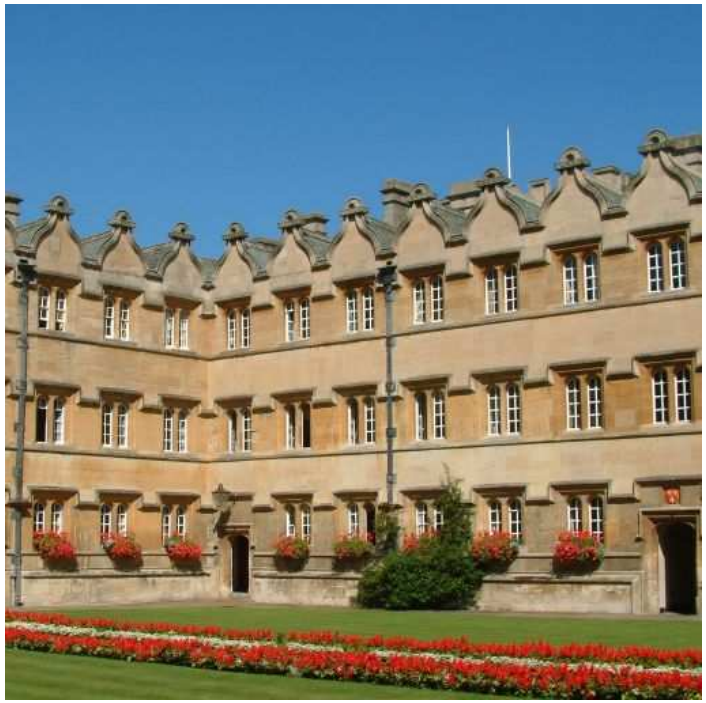


Calibration

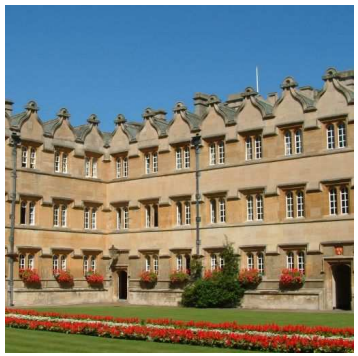




cover image

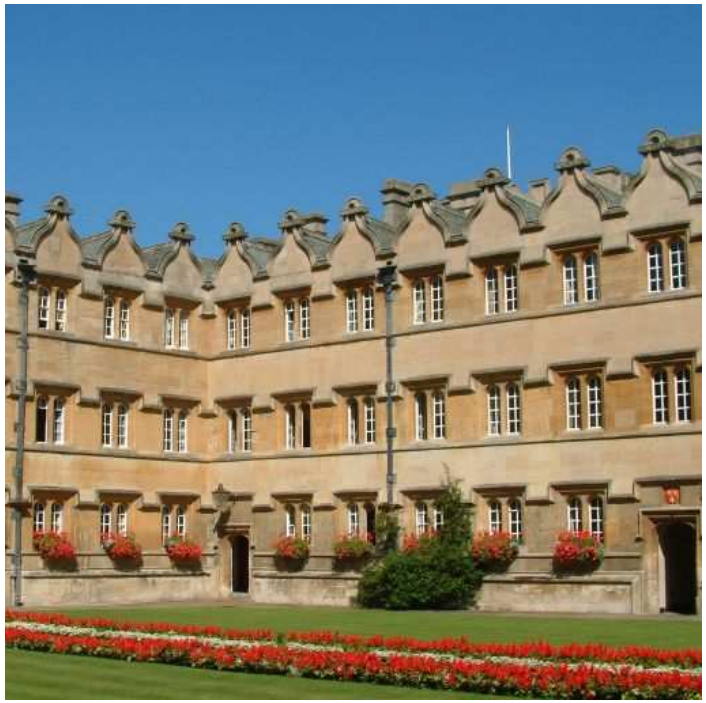


HCF COM = (77, 77, 77)

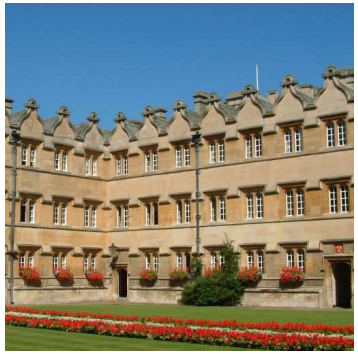


HCF COM = (76, 77, 77)

cover image

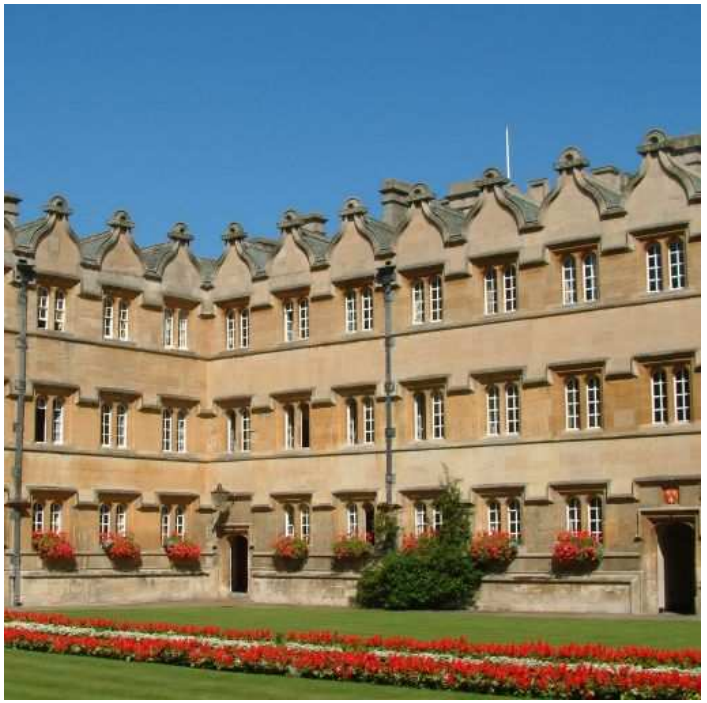


HCF COM = (77, 77, 77)



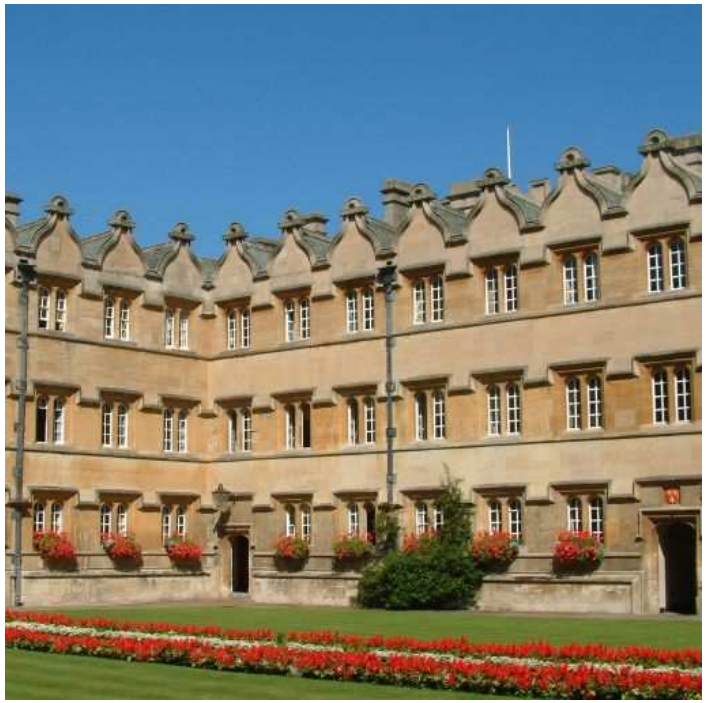
HCF COM = (76, 77, 77)

stego image

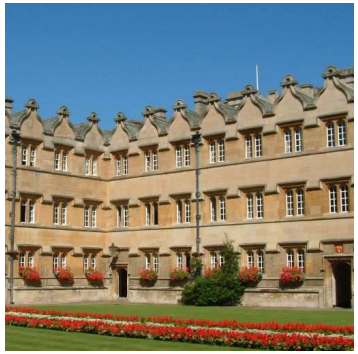


HCF COM = (55, 54, 54)

cover image

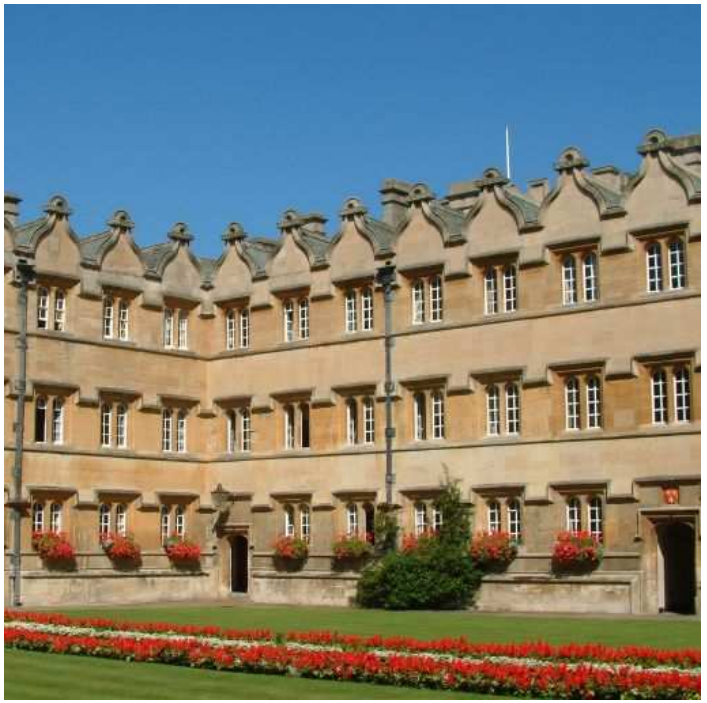


HCF COM = (77, 77, 77)

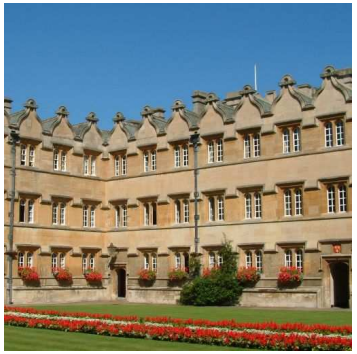


HCF COM = (76, 77, 77)

stego image



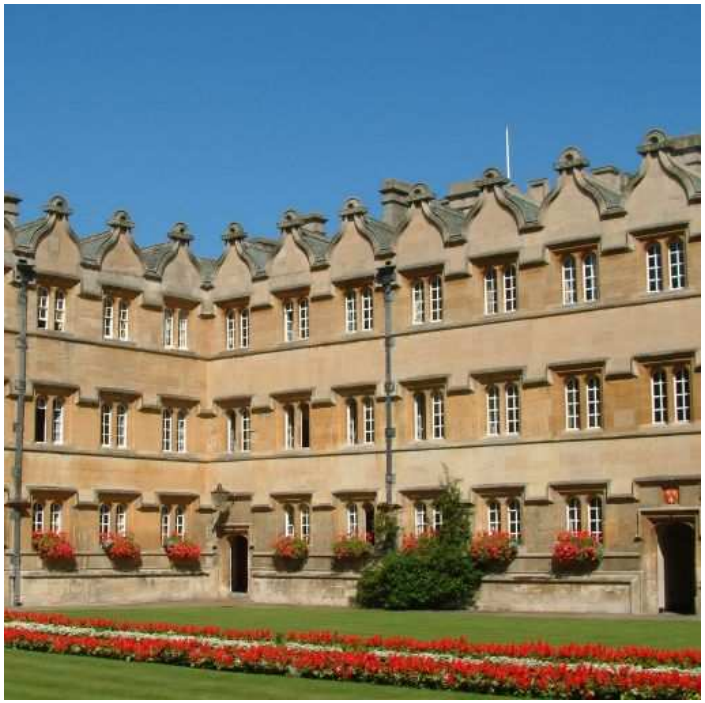
HCF COM = (55, 54, 54)



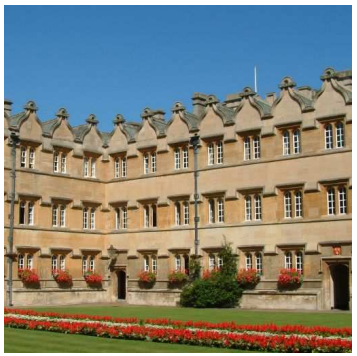
HCF COM = (64, 64, 64)

unknown to detector

stego image



HCF COM = (55, 54, 54)



HCF COM = (64, 64, 64)

Improved Detector

i) When a cover image is halved in size the HCF COM is largely unchanged.

ii) Steganography reduces the full-size image HCF COM by more than the half-size image. (“Downsampling tends to reduce the effect of noise”).

Given a suspect image, downsample it:
If the HCF COM increases, suspect steganography.

(use multidimensional classifier on 3D vector: COM divided by downsampled image COM)

Improved Detector

i) When a cover image is halved in size the HCF COM is largely unchanged.

ii) Steganography reduces the full-size image HCF COM by more than the half-size image. (“Downsampling tends to reduce the effect of noise”).

Given a suspect image, downsample it:
If the HCF COM increases, suspect steganography.

(use multidimensional classifier on 3D vector: COM divided by downsampled image COM)

Experimental Results

Generally an improvement over the standard HCF COM detector, but occasional major failures

cover image

HCF COM=(69, 69, 69)

HCF COM=(69, 69, 69)

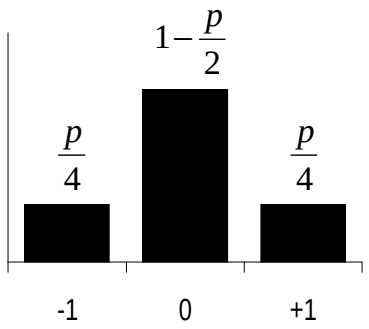
stego image (50% embedding)

HCF COM=(58, 57, 57)

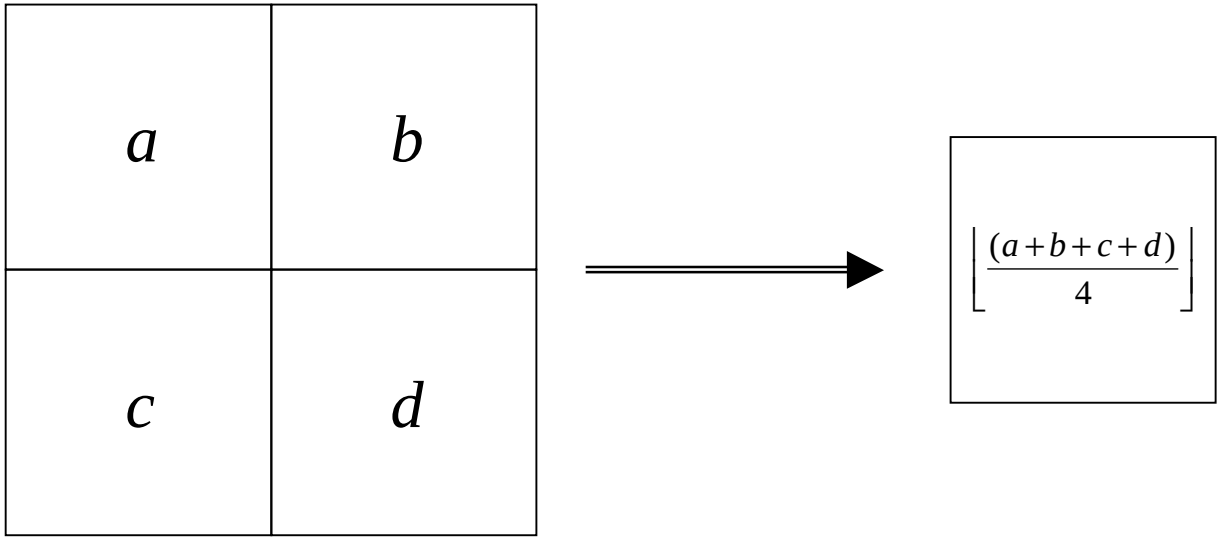
HCF COM=(54, 54, 53)

Why Did This Happen?

If proportion p of the maximal message is embedded, the stego noise is



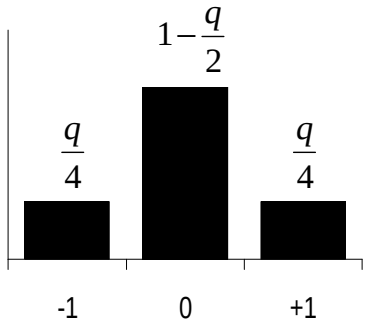
The downsampling procedure is



Lemma

Assuming that the sums of groups of 4 original pixels are uniformly distributed mod 4,

the effect on the downsampled image is to add noise with histogram



where $q < p$

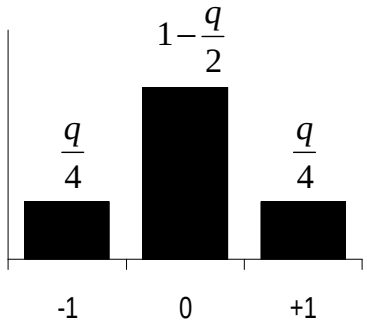
i.e. downsampling reduces stego noise

(so increases the HCF COM when steganography is present)

Lemma

Assuming that the sums of groups of 4 original pixels are uniformly distributed mod 4,

the effect on the downsampled image is to add noise with histogram



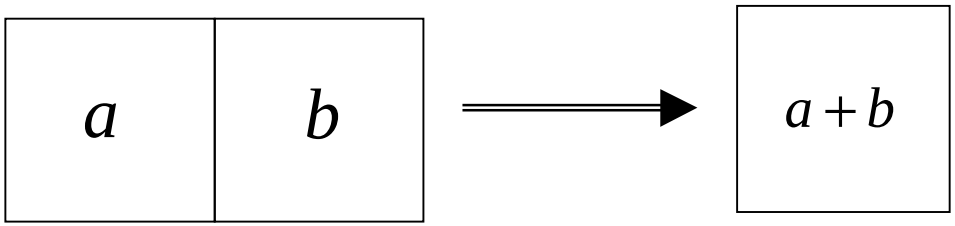
where $q < p$

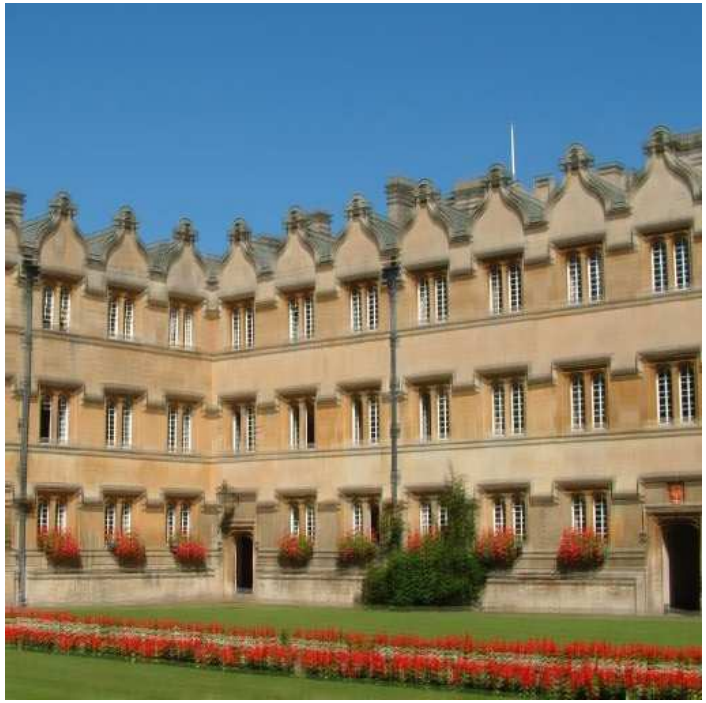
i.e. downsampling reduces stego noise

(so increases the HCF COM when steganography is present)

Better Calibration

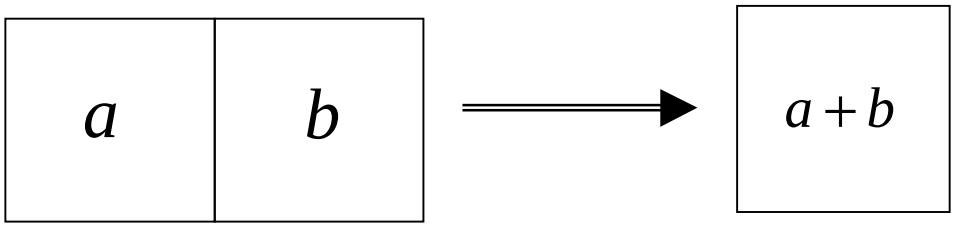
Don’t round down.





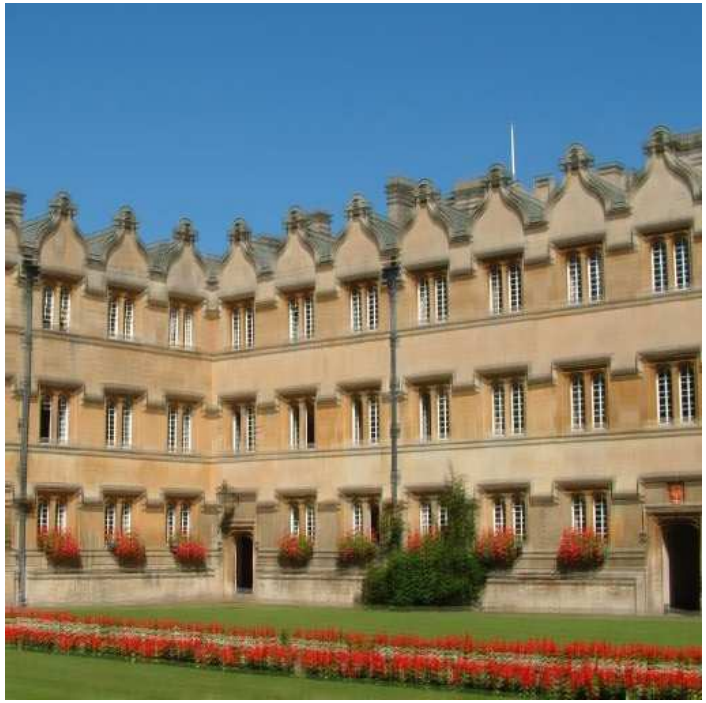
Better Calibration

Don't round down.



In the "smeared" image, pixel values have twice the range, 0..511

NB: must still use only the lowest 128 frequencies in the COM calculation



When an image is smeared and the HCF COM observed to increase, suspect steganography.

Further Improvements

HCF COM calibrated by smearing requires a DFT on 512^3 points

Don't treat RGB values as a 3D vector – add up the components $r+g+b$. The sum has "three times as much noise" due to steganography.

⇨ *DFT on 768 points*

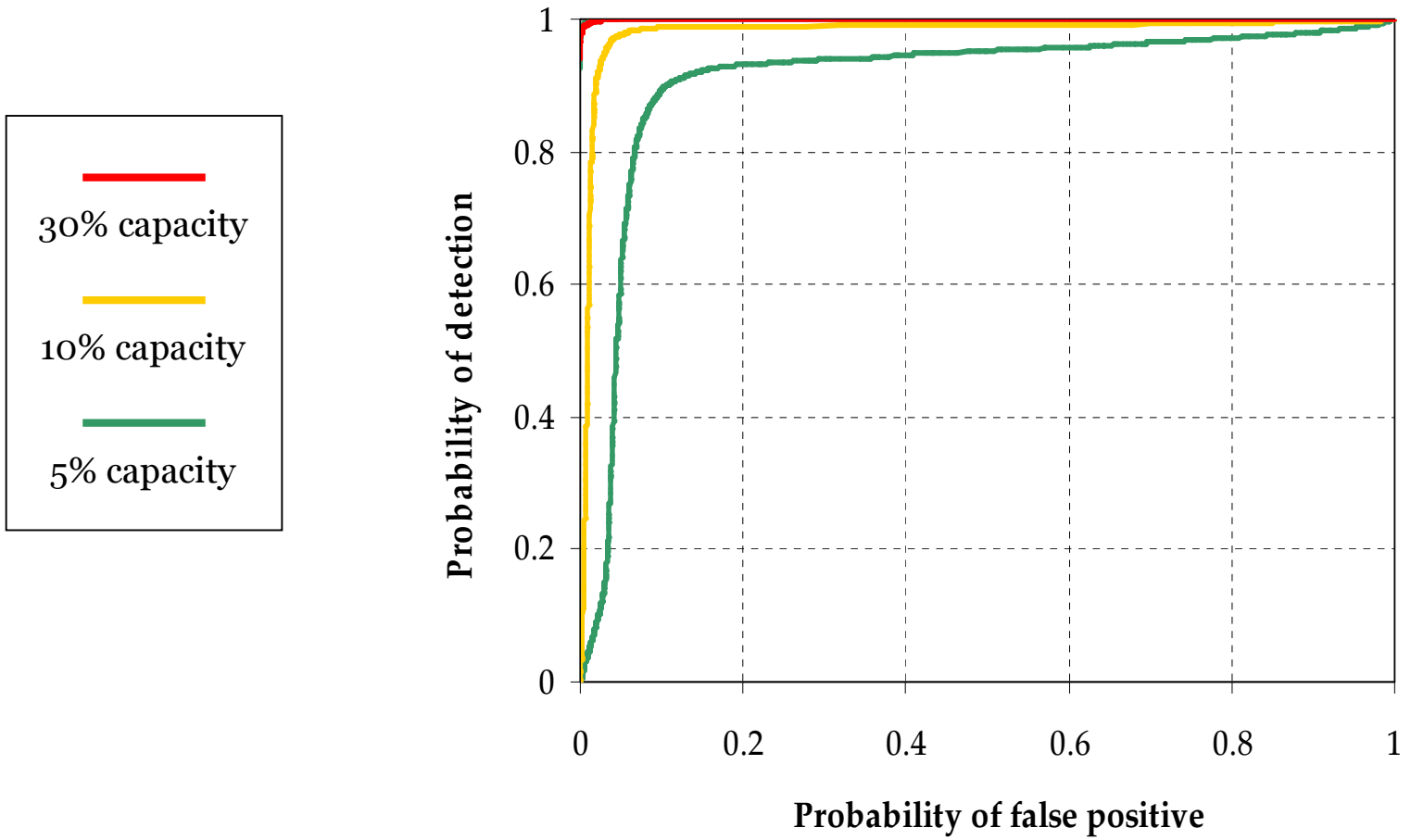
Form a 2D "adjacency histogram" (co-occurrence matrix) and calibrate using the "smeared" image

⇨ *DFT on 1536^2 points*

Faster and more reliable

Reliability

observed for 10000 colour bitmaps previously subject to moderate JPEG compression.



Detectors for LSB Matching

JPEG Compatability

[Fridrich, SPIE ITCom'01]

Close Colour Pairs

[Westfeld, IHW'02]

"HCF COM"

[Harmsen, SPIE EI'03]

Detectors for LSB Matching

	JPEG Covers	Resampled JPEG Covers	Uncompressed Covers
JPEG Compatability *[Fridrich, SPIE ITCom'01]*			
Close Colour Pairs *[Westfeld, IHW'02]*		***Different types of cover image can give very different results***	
"HCF COM" *[Harmsen, SPIE EI'03]*			

Detectors for LSB Matching

	JPEG Covers	Resampled JPEG Covers	Uncompressed Covers
JPEG Compatability *[Fridrich, SPIE ITCom'01]*	✓ ?	✘	✘
Close Colour Pairs *[Westfeld, IHW'02]*	>1% capacity	✘	✘
"HCF COM" *[Harmsen, SPIE EI'03]*	>50% capacity	>50-75% capacity	100% capacity

Detectors for LSB Matching

	JPEG Covers	Resampled JPEG Covers	Uncompressed Covers
JPEG Compatability *[Fridrich, SPIE ITCom'01]*	✓ ?	✘	✘
Close Colour Pairs *[Westfeld, IHW'02]*	>1% capacity	✘	✘
"HCF COM" *[Harmsen, SPIE EI'03]*	>50% capacity	>50-75% capacity	100% capacity
Calibrated Detectors	>5% capacity	>5-10% capacity	>50% capacity

Conclusions

- LSB Matching is almost as simple as LSB Replacement, but much harder to detect.

- Harmsen's standard "HCF COM" detector is usable for colour bitmaps of all types, but not very sensitive.

- We have suggested ways to improve the sensitivity by comparing the HCF COM of an image with that of a downsampled/smeared image.

- More performance is gained by totalling up the RGB components of a colour image.

LSB Matching is still very difficult to detect in cover images which have never been JPEG compressed (or in grayscale images) unless the hidden payload is very large.

End

adk@comlab.ox.ac.uk