universität
innsbruck

Increasing Trust in Image Analysis by Detecting Trellis Quantization in JPEG Images

Nora Hofer

IEEE International Conference on Image Processing · Abu Dhabi, UAE, October 2024

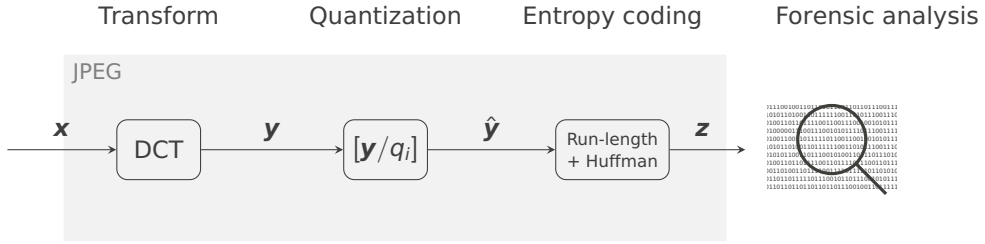
Digital Image Forensics

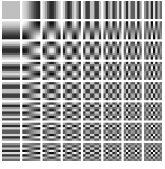
Methods for the verification of **image authenticity**, **source attribution**, and the detection of **traces of manipulation**

Image content

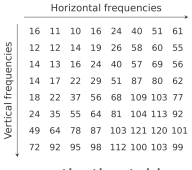
Statistical image properties

The Compression Pipeline

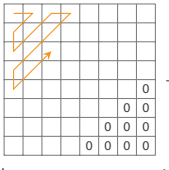




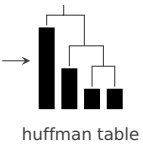
2D base vectors



quantization table

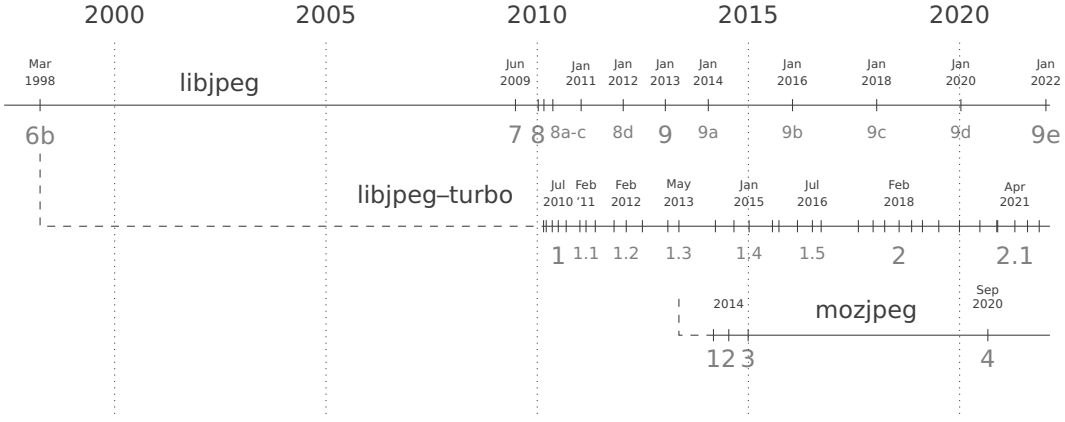


zig–zag arrangement



huffman table

JPEG Compressor Timeline



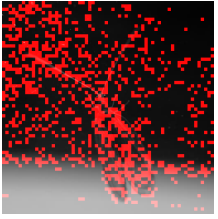
Beneš, M., Hofer, N., and Böhme, R. Know Your Library: How the libjpeg Version Influences Compression and Decompression Results. *IH&MMSEC*. ACM, 2022, pp. 19–25.

Do JPEG Implementations Differ? Yes!

libjpeg v6b

libjpeg-turbo

MozJPEG

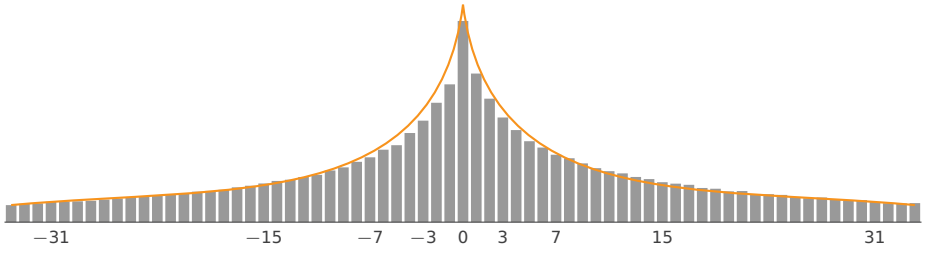


Red highlights: 8×8 pixel blocks containing numerical differences in at least one DCT coefficient.

Outline

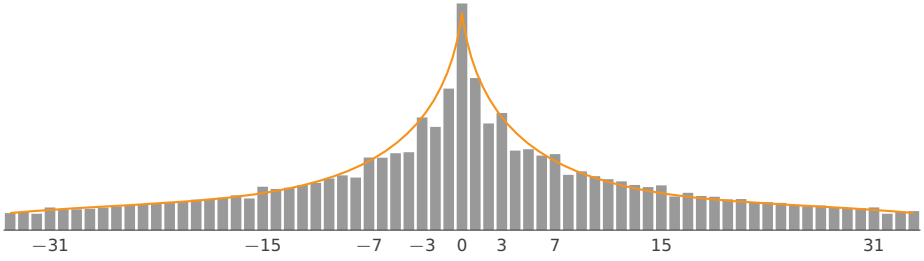
- **Forensic implications of JPEG optimizations**
- Detection of artifacts from JPEG optimizations

Histogram of DCT Coefficients



First AC subband of 100 **baseline** grayscale images (QF 99).

Histogram of DCT Coefficients



First AC subband of 100 **trellis quantized** grayscale images (QF 99).

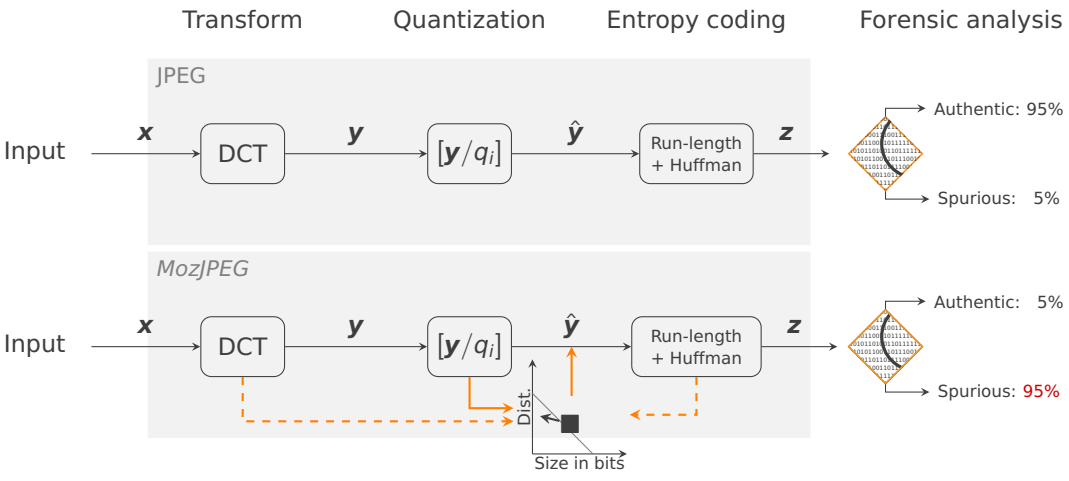
Detection of Steganography in *MozJPEG* Images

Performance of EfficientNet-B0 detectors			
Embedding	*libjpeg-turbo*		*MozJPEG*
	Baseline acc.	FPR	FPR
nsF5 (Fridrich et al., 2007.)	99%	1%	**99%**
UERD (Guo et al., 2015.)	93%	4%	**43%**
J-UNIWARD (Holub et al., 2014.)	91%	8%	**94%**

1 000 ALASKA2 test images of QF 75 and 0.4 bits per non-zero AC coefficients (bpnzAC) embedding.

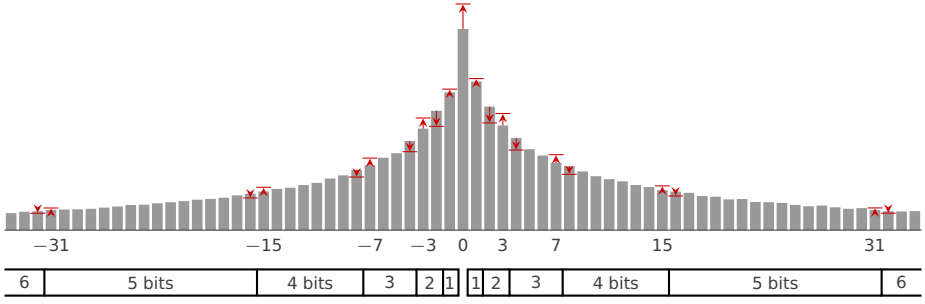
ImageNet-pretrained EfficientNet-B0 models: 32 batch size, 0.25 dropout rate, 0.0001 learning rate, Adam optimizer.

The Compression Pipeline – Trellis Quantization



Histogram of DCT Coefficients

Trellis quantization amplifies **zeros** and **maximum values** of shorter encoding.



Cost of encoding a DCT block when replacing non-zero coefficients with the *highest* value of every lower size group.

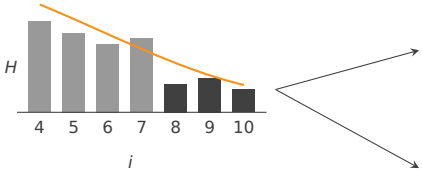
First AC subband of 100 grayscale images (QF 99).

Outline

- Forensic implications of JPEG optimizations
- **Detection of artifacts from JPEG optimizations**

Detection of Trellis Artifacts

Methods



Analytical modelling

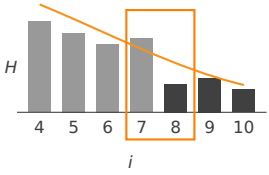
- Calibration
- “Vampire” neighborhoods

Statistical learning

- Cartesian calibration
- Vampire neighborhoods
- JRM features

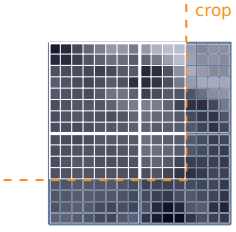
Analytical Detection of Trellis Artifacts

Calibration



$$\alpha_i = \frac{H_i - \hat{H}_i}{\hat{H}_{i+1}}$$

$$\alpha_{\mathcal{C}} = \sum_{i \in \mathcal{C}} (\alpha_i - \alpha_{i+1})$$



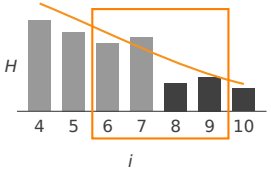
Calibration

- Estimate the histogram of DCT coefficients before trellis quantization.
- α: relative frequency of coefficients being changed to inner neighbors.

Fridrich, J., Goljan, M., and Hogea, D. Steganalysis of JPEG Images: Breaking the F5 Algorithm. *IH*. Springer, 2003, pp. 310–323.

Analytical Detection of Trellis Artifacts

Vampire neighborhoods



$$\beta_i = H_i - \frac{H_{i-1}+H_{i+2}}{2} + H_{i+1} - \frac{H_{i-1}+H_{i+2}}{2}$$

$$\beta_{\mathcal{C}} = \sum_{i \in \mathcal{C}} \left(H_i - H_{i-1} + H_{i+1} - H_{i+2} \right)$$

Vampire neighborhoods

- Measure deviation of monotonous histogram at candidate pairs . . .
- . . . with regard to candidate neighborhoods.

Statistical Detection of Trellis Artifacts

Ensemble of Fisher linear discriminant base learners

Calibration

- candidate neighborhood features of calibrated and input images
- feature dimensions: $4 \times 8 \times 10 \times 2 = 640$

Vampire neighborhoods

- candidate neighborhood features of input images
- feature dimensions: $4 \times 8 \times 10 = 320$

JRM features

- JRMs model dependencies between adjacent subbands
- features from an ensemble of JPEG Rich Models of 11 255 dimensions

Fridrich, J., Kodovský, J. Rich Models for Steganalysis of Digital Images. *TIFS*. IEEE, 2012, pp. 868–882.

Detection of Trellis Artifacts

Detection accuracies

	Analytic detectors		Learning detectors		
QF	calibration	vampires	calibration	vampires	JRM
100	95%	99%	100%	100%	100%
95	84%	92%	99%	99%	100%
90	80%	88%	99%	98%	100%
85	75%	82%	99%	97%	100%
80	72%	79%	98%	96%	100%
75	71%	76%	98%	96%	99%
50	72%	75%	97%	93%	98%

Test set of 5 000 ALASKA2 images compressed with *MozJPEG v*4.0.3, default compression settings.

Hofer, N. Increasing Trust in Image Analysis by Detecting Trellis Quantization in JPEG Images. *ICIP*. IEEE, 2024.

Conclusion

- Trellis quantization can **compromise the reliability** of forensic methods.
- To **prevent misclassifications** we can **detect trellis artifacts**.

When methods are applied in forensic investigations, we need to be aware of differences of compression implementations in practice.

Thank You

Contact:

Nora Hofer

nora.hofer@uibk.ac.at

https://informationsecurity.uibk.ac.at/people/nora-hofer/

Funding Notice:

UNCOVER

This work is funded by the EU's Horizon 2020 programme under grant agreement No. 101021687 (UNCOVER).

universität
innsbruck

Thank You

Increasing Trust in Image Analysis
by Detecting Trellis Quantization in JPEG Images

Nora Hofer

IEEE International Conference on Image Processing · Abu Dhabi, UAE, October 2024