Computational Harmonic Analysis Tools for Image Compression and Inpainting

Gerlind Plonka

Institute for Numerical and Applied Mathematics
University of Göttingen

Dagstuhl

November, 2016

Outline

- Interpolation versus decorrelating transforms for image compression
- Non-adaptive transforms
  - Tensor-product wavelet transforms
  - Directional frames
- Adaptive wavelet transforms
  - Generalized lifting schemes
  - Geometric approaches with adaptivity costs
  - Description of the EPWT algorithm
- Wavelet transforms for inpainting

Introduction: Interpolation versus decorrelating transforms for image compression

Idea 1 Interpolation/Inpainting:
Store only a subset of given image points and reconstruct the image via interpolation.

Idea 2 Decorrelation:
Apply a transform to achieve data decorrelation (sparse data representation) and store the significant coefficients.
Apply the inverse transform for reconstruction.

a) Apply a non-adaptive transform (DCT, wavelet transform)

b) Apply an adaptive transform (SVD, adaptive wavelet transform)

Decorrelating transforms

KLT The **Karhunen-Loéve transform** (KLT) is the SVD of the image as a matrix. The low-rank approximation using only the largest singular values gives an optimal compression result.

Drawbacks: basis of singular vectors is data dependent, expensive

Idea: Find non-adaptive transforms that work well for many images:

Discrete cosine transform (heart of lossy compression in JPEG)

$$\widehat{\mathbf{A}} = \mathbf{C}_8 \mathbf{A} \mathbf{C}_8^T \quad \mathbf{A} \in \mathbb{C}^{8\times 8}$$

Discrete tensor product wavelet transform

$$\widehat{\mathbf{A}} = \mathbf{W}_N \mathbf{A} \mathbf{W}_M^T \quad \mathbf{A} \in \mathbb{C}^{N\times M}$$

JPEG

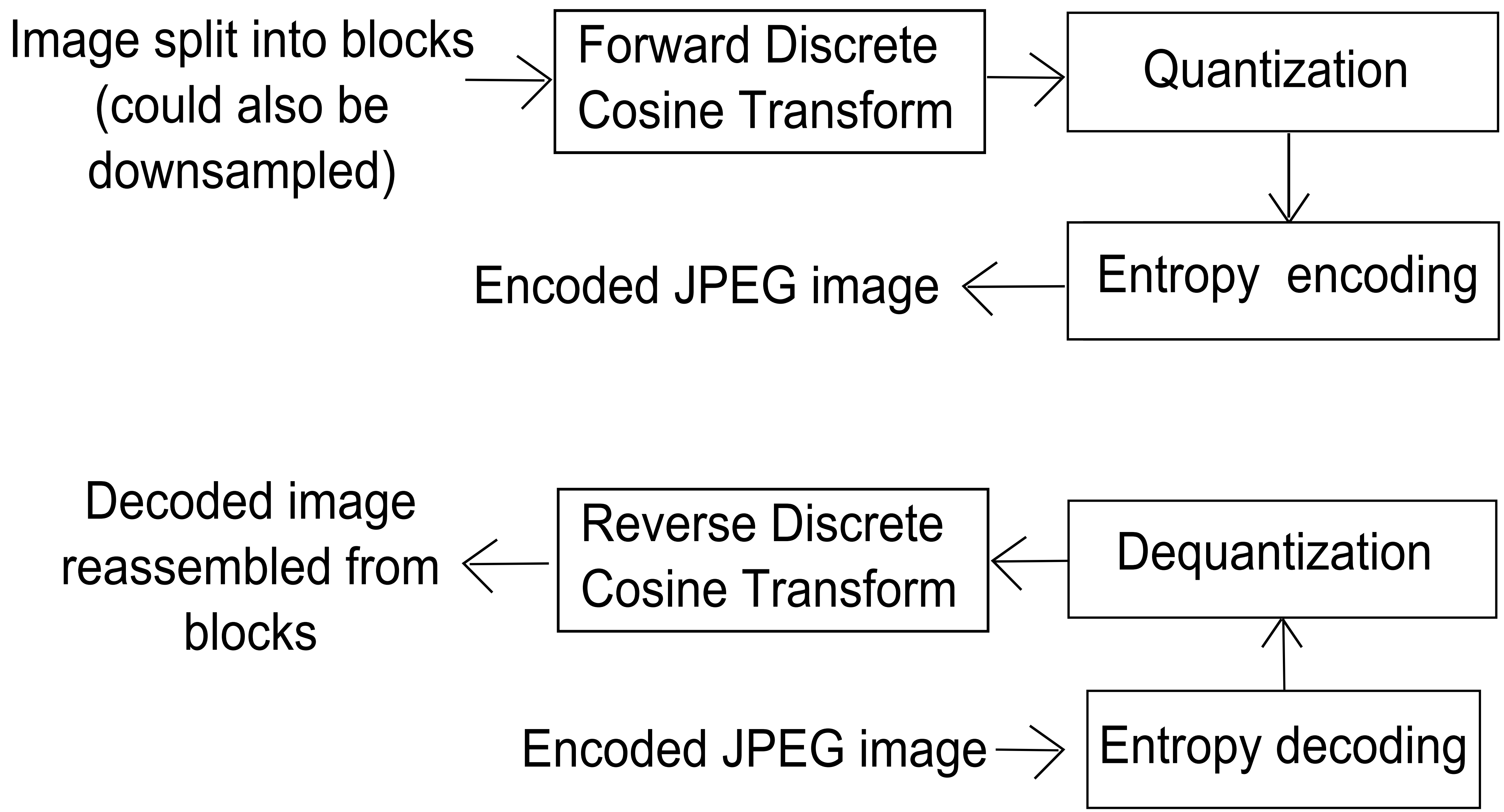


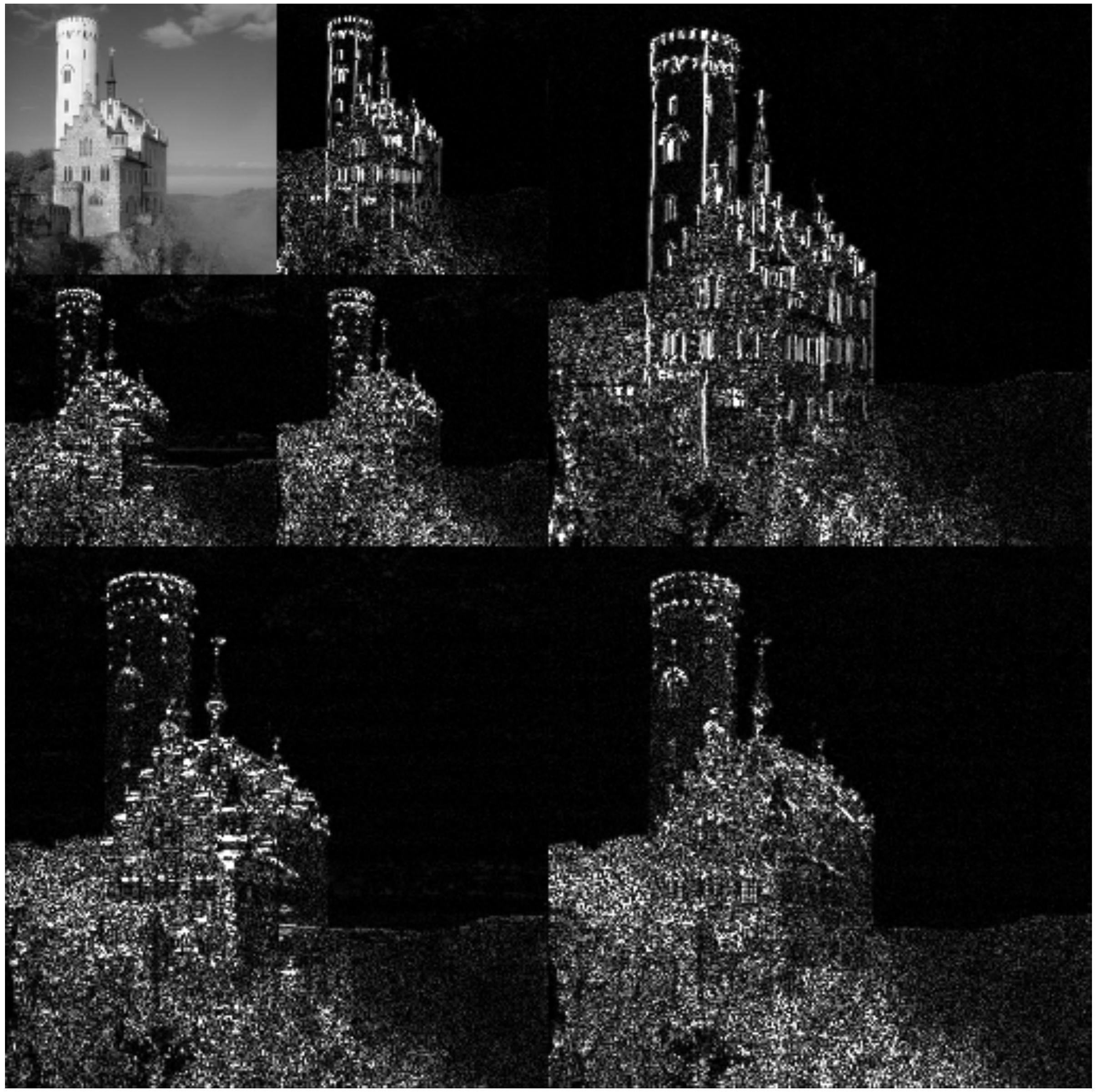
Image split into blocks (could also be downsampled)
Forward Discrete Cosine Transform
Quantization
Entropy encoding
Encoded JPEG image
Decoded image reassembled from blocks
Reverse Discrete Cosine Transform
Dequantization
Entropy decoding
Encoded JPEG image

Discrete wavelet transform

multiscale transform: $\widehat{\mathbf{A}} = \mathbf{W}_N^1 \mathbf{W}_N \mathbf{A} \mathbf{W}_N^T (\mathbf{W}_N^1)^T$

first level

$$W_N = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & & & & & & \\ & & 1 & 1 & & & & \\ & & & & \ddots & & & \\ & & & & & & 1 & 1 \\ 1 & -1 & & & & & & \\ & & 1 & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & & 1 & -1 \end{bmatrix}$$



second level

$$W_N^1 = \begin{bmatrix} \mathbf{W}_{N/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N/2} \end{bmatrix}$$

Many wavelet (frame) constructions for image analysis

1) steerable wavelets [Freeman and Adelson '91]
2) curvelets [Candes, Donoho '03]
3) shearlets [Labate, Lim, Kutyniok, Weiss '05]
4) contourlets [Do, Vetterli '05]
5) Gabor wavelets [Lee '08]
6) α-molecules [Grohs, Keiper, Kutyniok, Schäfer '14]
7) ...

Wanted properties of a new directional wavelet system

- Good space-frequency localization
- "Simple structure" of the wavelet system $\{\psi_\lambda\}_{\lambda\in\Lambda}$
- Orthonormal basis or Parseval frame of $L^2(\mathbb{R}^2)$, i.e.,

  $$f = \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda\rangle \psi_\lambda$$

  and

  $$\sum_{\lambda\in\Lambda} |\langle f, \psi_\lambda\rangle|^2 = \|f\|^2_{L^2(\mathbb{R}^2)} \quad \text{for all } f \in L^2(\mathbb{R}^2)$$

- Good approximation properties: If f is in a certain smoothness class, then f can be well approximated by a sparse wavelet frame expansion, such that e.g.

  $$\|f - f_N\|_2^2 \leq C\, N^{-\beta}$$

  for (piecewise) Hölder smooth functions of order β.

However: Not suitable for image compression because of redundancy!

Adaptive wavelet transforms

Idea Design adaptive approximation schemes respecting the local geometric regularity of two-dimensional functions

Basic adaptive wavelet approaches

a) Apply a generalized lifting scheme to the data using (nonlinear) data-dependent prediction and update operators

b) Adaptive approximation schemes using geometric image information, usually with extra adaptivity costs

Basic adaptive wavelet approaches

a) Apply a generalized lifting scheme to the data using (nonlinear) data-dependent prediction and update operators

Literature (incomplete)

- **discrete MRA and generalized wavelets** (Harten '93)
- **second generation wavelets** (Sweldens '97)
- **edge adapted multiscale transform** (Cohen & Matei '01)
- **Nonlinear wavelet transforms** (Claypoole et al. '03)
- **adaptive lifting schemes** (Heijmans et al. '06)
- **adaptive directional lifting based wavelet transf.** (Ding et al. '06)
- **edge-adapted nonlinear MRA (ENO-EA)** (Arandiga et al. '08)
- **meshless multiscale decompositions** (Baraniuk et al. '08)
- **nonlinear locally adaptive filter banks** (Plonka & Tenorth '09)

How does it work?

The general lifting scheme consists of three steps.

1. Split Split the given data $\mathbf{a} = (a(i,j))_{i,j=0}^{N-1}$ into two sets $\mathbf{a}^e$ and $\mathbf{a}^o$

2. Predict Find a good approximation $\tilde{\mathbf{a}}^o$ of $\mathbf{a}^o$ of the form

$$\tilde{\mathbf{a}}^o = P_1 \mathbf{a}^o + P_2 \mathbf{a}^e$$

Put

$$\mathbf{d}^o := \tilde{\mathbf{a}}^o - \mathbf{a}^o.$$

Assume that $(\mathbf{a}^e, \mathbf{a}^o) \mapsto (\mathbf{a}^e, \mathbf{d}^o)$ is invertible, i.e., $I - P_1$ is invertible.

3. Update Find a "smoothed" approximation of $\mathbf{a}^e$ (a low-pass filtered subsampled version of $\mathbf{a}$)

$$\tilde{\mathbf{a}}^e := U_1(\mathbf{d}^o) + U_2(\mathbf{a}^e)$$

Assume that $(\mathbf{a}^e, \mathbf{d}^o) \mapsto (\tilde{\mathbf{a}}^e, \mathbf{d}^o)$ is invertible, i.e., that U_2 is invertible.

How to choose the prediction and update operators?

Prediction operator local approximation of $\mathbf{a}^o$ by an adaptively weighted average of "neighboring" data

Example 1.

- Fix a stencil at a neighborhood of $a^o(i, j)$ (adaptively)
- Compute a polynomial p by interpolating/approximating the data on the stencil
- Choose $p(i, j)$ to approximate $a^o(i, j)$.

Example 2. Use nonlinear diffusion filters to determine the prediction operator

Update operator usually linear, non-adaptive

Basic adaptive wavelet approaches

wedgelets (Donoho ’99)

approximation of images using an adaptively chosen domain decomposition

bandelets (Le Pennec & Mallat ’05)

wavelet filter bank followed by adaptive geometric orthogonal filters

geometric wavelets (Dekel & Leviatan ’05)

binary space partition and polynomial approximations in subdomains

geometrical grouplets (Mallat ’09)

association fields that group points, generalized Haar wavelets

EPWT (Plonka et al. 09)

tetrolets (Krommweh ’10)

generalized Haar wavelets on adaptively chosen tetrolet partitions

Comparison of basic adaptive wavelet approaches

a) Generalized lifting scheme with nonlinear prediction

Advantages invertible transform, no side information necessary
usually a justifiable computational effort

Drawbacks bad stability of the reconstruction scheme
only slightly better approximation results compared with linear (nonadaptive) transforms

b) Adaptive wavelet approximation using geometric image information

Advantages very good approximation results

Drawbacks adaptivity costs for encoding
usually high computational effort

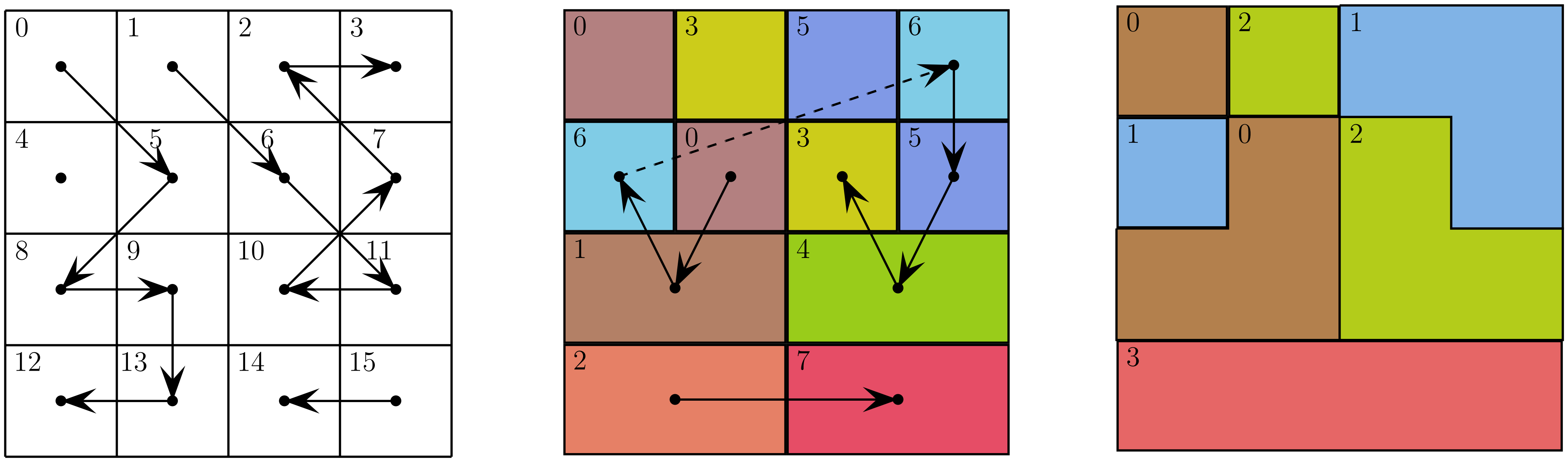
Problem Given a matrix of data points (image values), how to compress the data by a wavelet transform thereby exploiting the local correlations efficiently?

Idea

1. Find a (one-dimensional) path through all data points such that there is a strong correlation between neighboring data points.
2. Apply a one-dimensional wavelet transform along the path.
3. Apply the idea repeatedly to the low-pass filtered array of data.

Toy Example

$$\mathbf{f} = \begin{bmatrix} 115 & 108 & 109 & 112 \\ 106 & 116 & 107 & 109 \\ 112 & 110 & 108 & 108 \\ 108 & 109 & 103 & 106 \end{bmatrix} \qquad \text{array of data.}$$



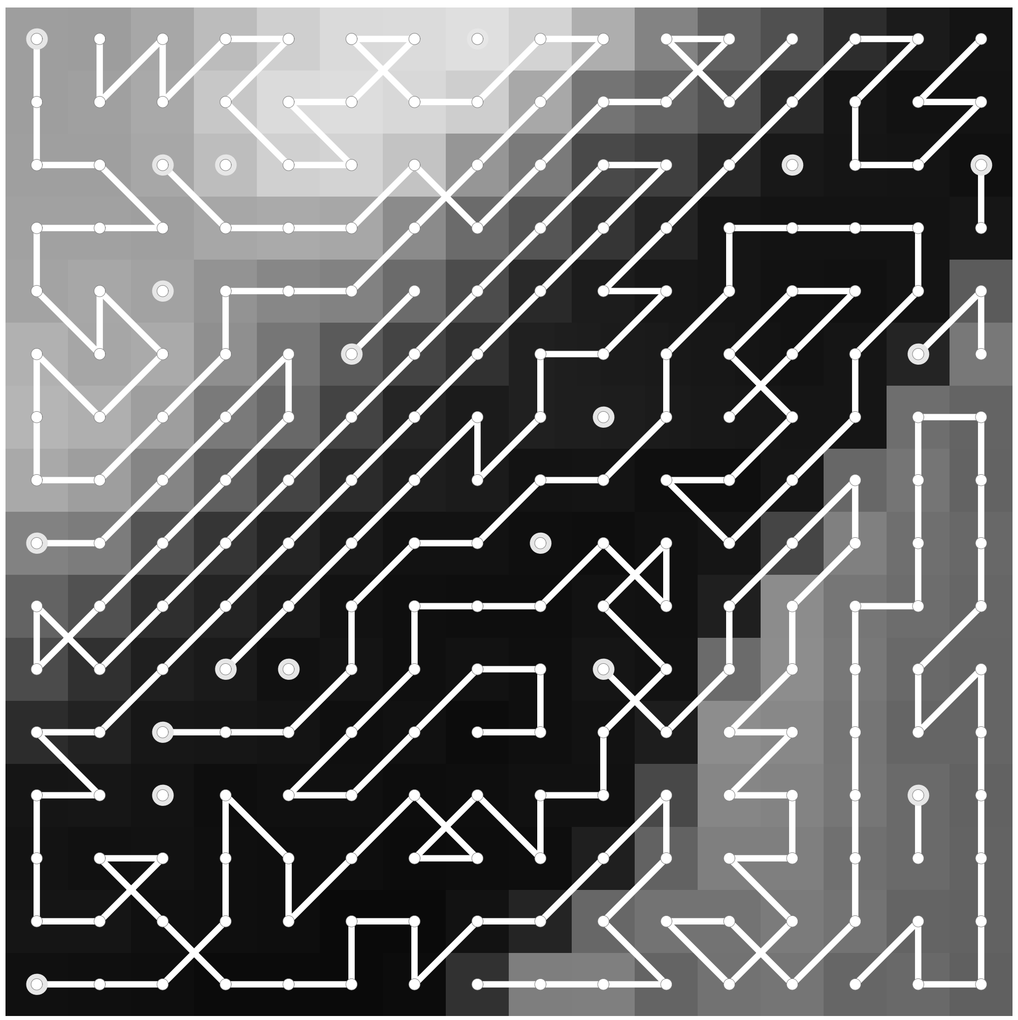
$$\mathbf{p}^4 = ((0,5,8,9,13,12),(1,6,11,10,7,2,3),(4),(15,14)),$$

$$\mathbf{f}^3 = (115.5, 111, 108.5, 107.5, 108, 109, 109, 104.5),$$

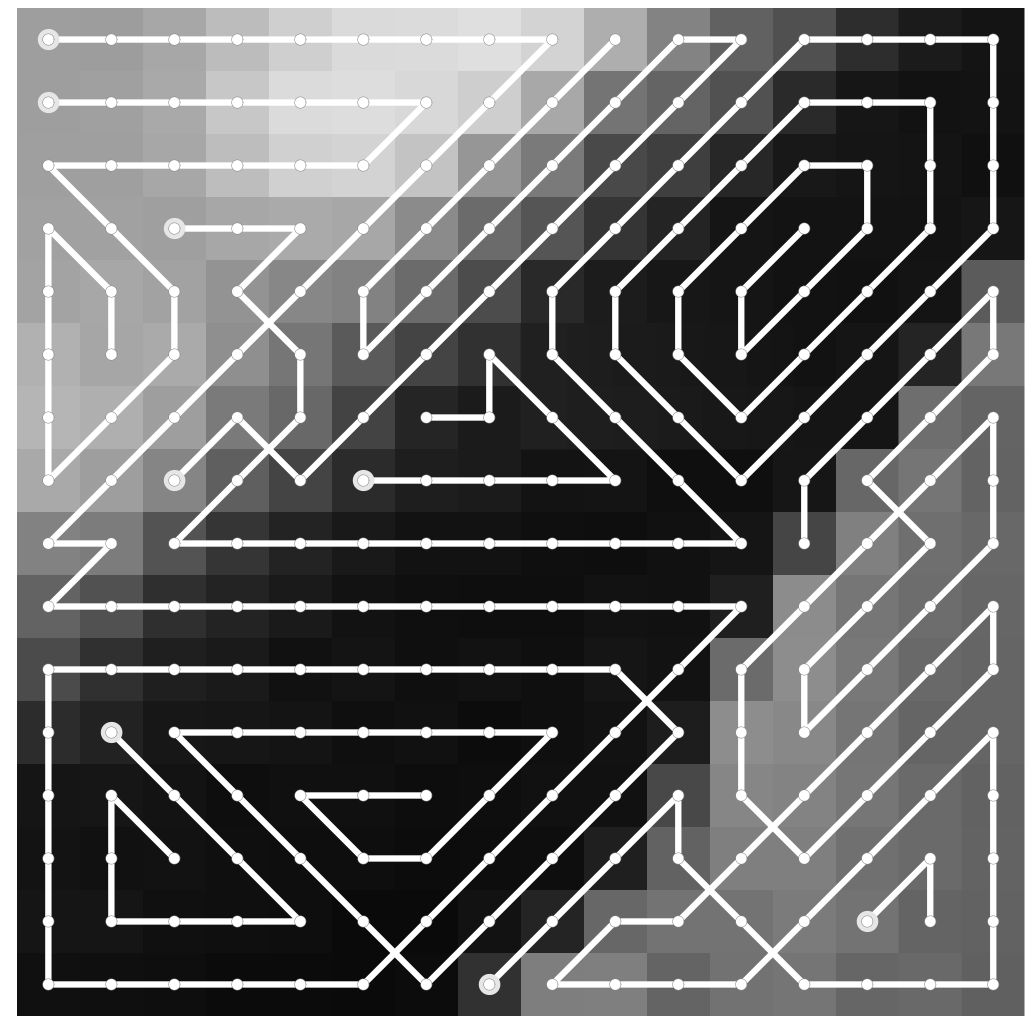
$$\mathbf{p}^3 = ((0,1,6,5,4,3),(2,7)), \qquad \mathbf{p}^2 = (0,1,2,3).$$

The relaxed EPWT

Idea: Change the direction of the path only if the difference of data values is greater than a predetermined value θ.



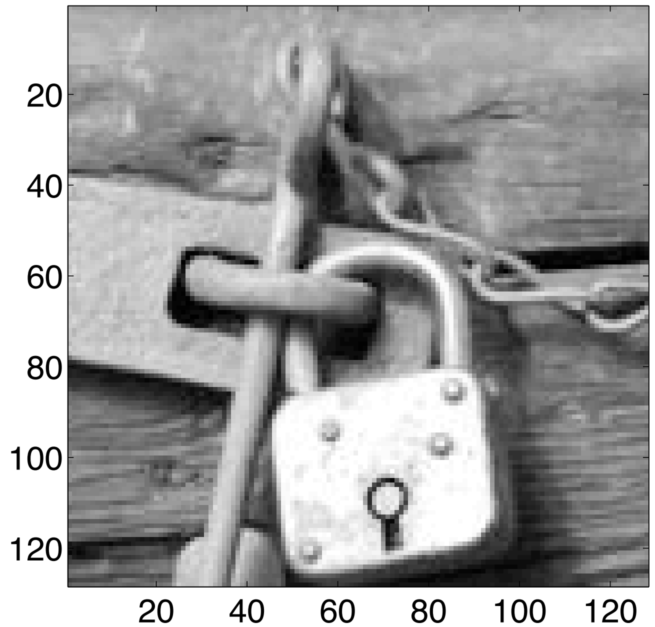
rigorous EPWT ($\theta = 0$)
Entropy 2.08 bit per pixel



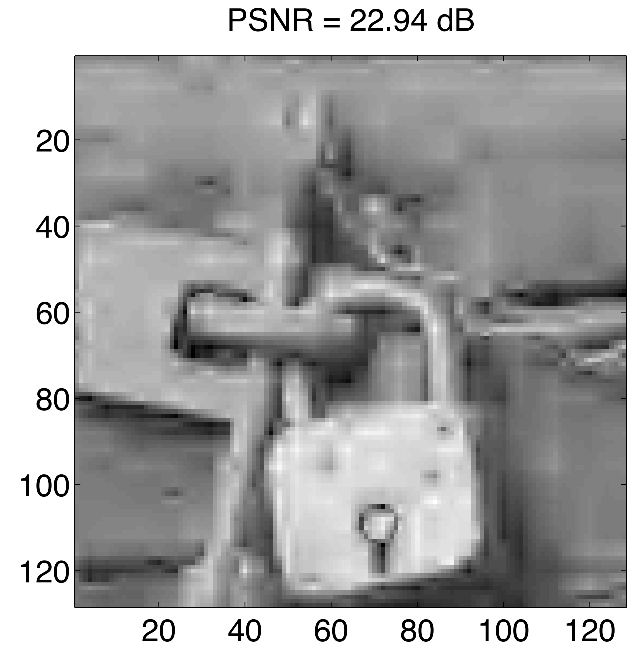
relaxed EPWT ($\theta = 0.14$)
Entropy 0.39 bit per pixel

Test: door lock image (128×128)

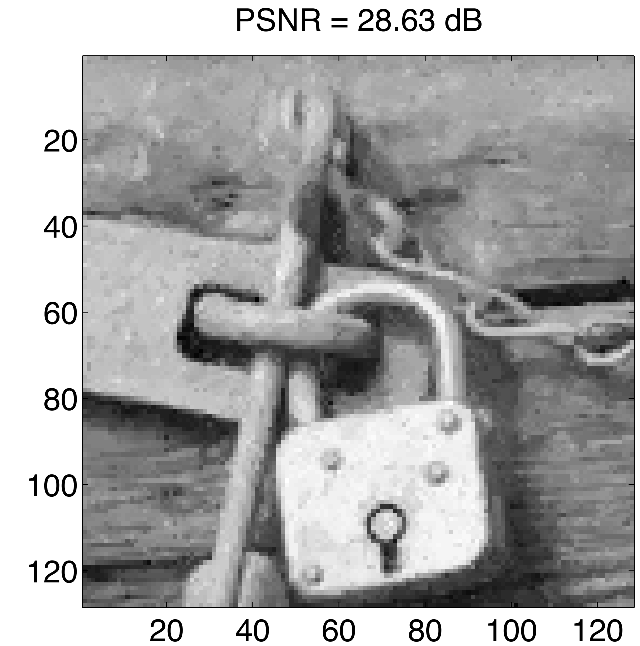
WT	θ_1	levels	nonzero coeff	PSNR	entropy of $\tilde{\mathbf{p}}^{14}$
tensor prod. Haar	-	7	512	22.16	-
tensor prod Daub.	-	6	512	22.94	-
tensor prod 7-9	-	4	512	22.49	-
EPWT Haar	0.00	14	512	28.04	2.22
EPWT Haar	0.05	14	512	28.37	1.11
EPWT Haar	0.10	14	512	27.74	0.55
EPWT Daub.	0.00	12	512	28.63	2.22
EPWT Daub.	0.05	12	512	29.23	1.11
EPWT Daub.	0.10	12	512	28.67	0.55
EPWT Daub.	0.15	12	512	27.65	0.32
EPWT 7-9	0.00	10	512	28.35	2.22
EPWT 7-9	0.05	10	512	28.99	1.11
EPWT 7-9	0.10	10	512	28.38	0.55



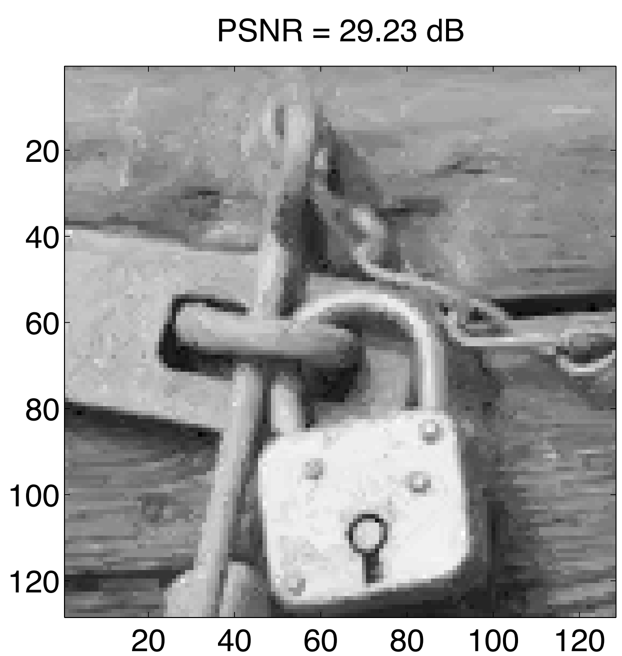
original image



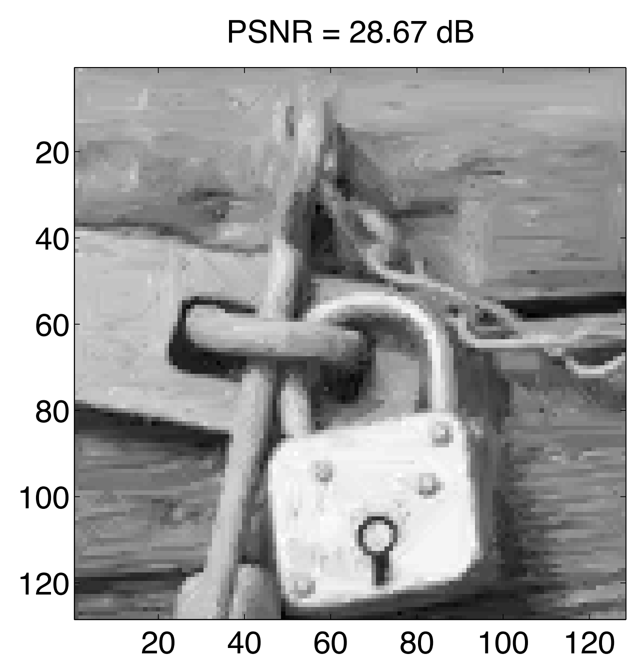
D4, 512 coeff.
PSNR= 22.94



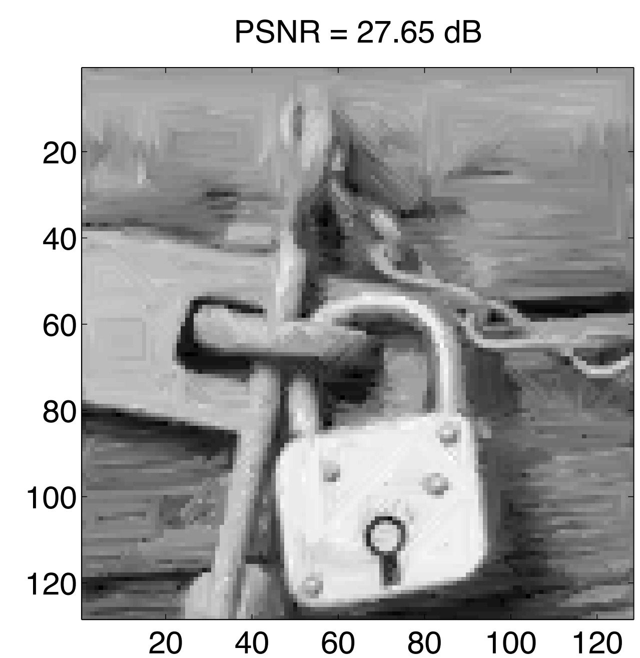
EPWT $\theta_1 = 0$
PSNR=28.63



EPWT, $\theta_1 = 0.05$
PSNR=29.23



EPWT, $\theta_1 = 0.1$
PSNR=28.67



EPWT, $\theta_1 = 0.15$
PSNR = 27.65

Results for N-term approximation

Theorem 1 (Plonka, Tenorth, Iske (2011))

The EPWT (with the Haar wavelet transform) leads for suitable path vectors to an N-term approximation of the form

$$\|f - f_N\|_2^2 \leq C\, N^{-\alpha}$$

for piecewise Hölder continuous functions of order α (with $0 < \alpha \leq 1$) possessing discontinuities along curves of finite length.

Theorem 2 (Plonka, Iske, Tenorth (2013))

The application of the EPWT leads for suitably chosen path vectors to an N-term approximation of the form

$$\|f - f_N\|_2^2 \leq C\, N^{-\alpha}$$

for piecewise Hölder smooth functions of order $\alpha > 0$ possessing discontinuities along curves of finite length.

Wavelet transforms for image inpainting

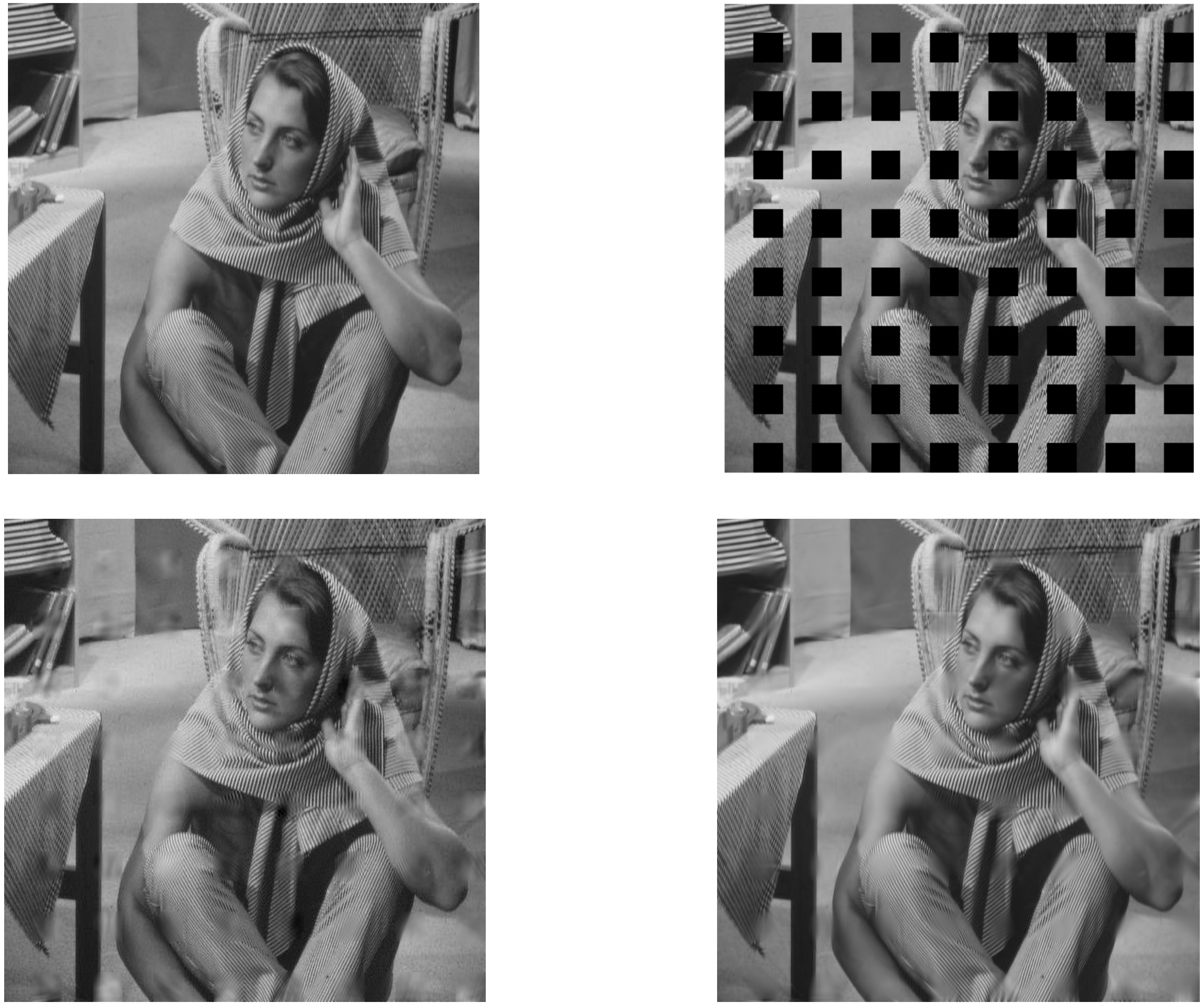
Known approaches

1. Apply a direction wavelet transform for regularization:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\Phi\mathbf{x}\|_1 \qquad \text{subject to} \qquad P_K\mathbf{x} = P_K\mathbf{x}^0.$$

see e.g. King et al. (2013)

Wavelet transforms for image inpainting



upper left: original image; upper right: missing blocks
lower left: reconstruction with curvelets and local cosine,
lower right: reconstruction with shearlets (King et al. (2013))

2. **Reconstruction of missing or damaged wavelet coefficients**

model 1 (noiseless case)

Minimize $TV(u_\beta)$ where $u_\beta(x)$ has the wavelet transform

$$u_\beta(x) = \sum_{j,k} \beta_{j,k}\psi_{j,k}(x), \qquad \beta = (\beta_{jk}),\ j \in \mathbb{Z}, k \in \mathbb{Z}^2$$

subject to $\qquad \beta_{j,k} = \alpha_{j,k} \qquad (j,k) \in I$

model 2 (noisy case)

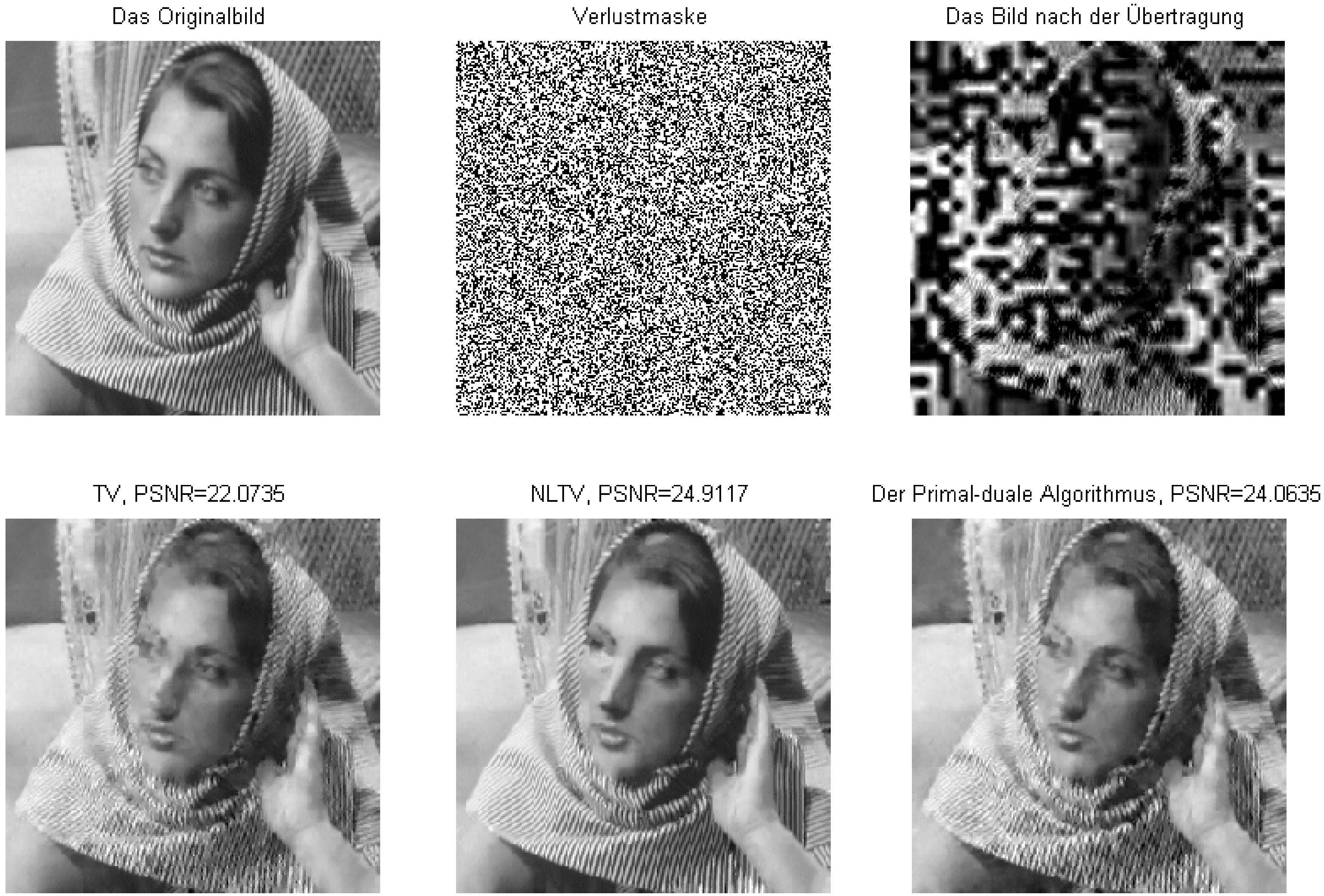
Minimize

$$TV(u_\beta(x)) + \sum_{(j,k)} \lambda_{j,k}(\beta_{j,k} - \alpha_{j,k})^2$$

where $\lambda_{j,k} = 0$ if $(j,k) \notin I$

see e.g. Chan et al. (2006); Zhang & Chan (2010)

Wavelet transforms for image inpainting



upper line: original image; mask of lost coefficients, reconstruction
lower line: reconstruction using TV, NLTV, Primal-dual algorithm
(Pototskaia, Master thesis (2013)

Summary

- Decorrelating non-adaptive transforms (cosine transform, tensor-product wavelet transform) are still the industrial standard for image compression.
- Adaptive transforms can perform better, but need higher computational effort, are less stable, or produce extra adaptivity costs.
- There is no inpainting method for image compression available that is comparable to the current compression standard performance.
- New approaches connection computational harmonic analysis tools may be:
  - the application of a multiscale approach for discrete Green's functions of diffusion operators
  - the application of “interpolatory” wavelets that can be interpreted (locally) in spatial domain

\thankyou