Sparse Coding based Frequency Adaptive Loop Filtering for Video Coding

Lehrstuhl und Institut für Nachrichtentechnik

RWTH AACHEN UNIVERSITY

Outline

1. Sparse Coding based Denoising

2. Frequency Adaptation Model

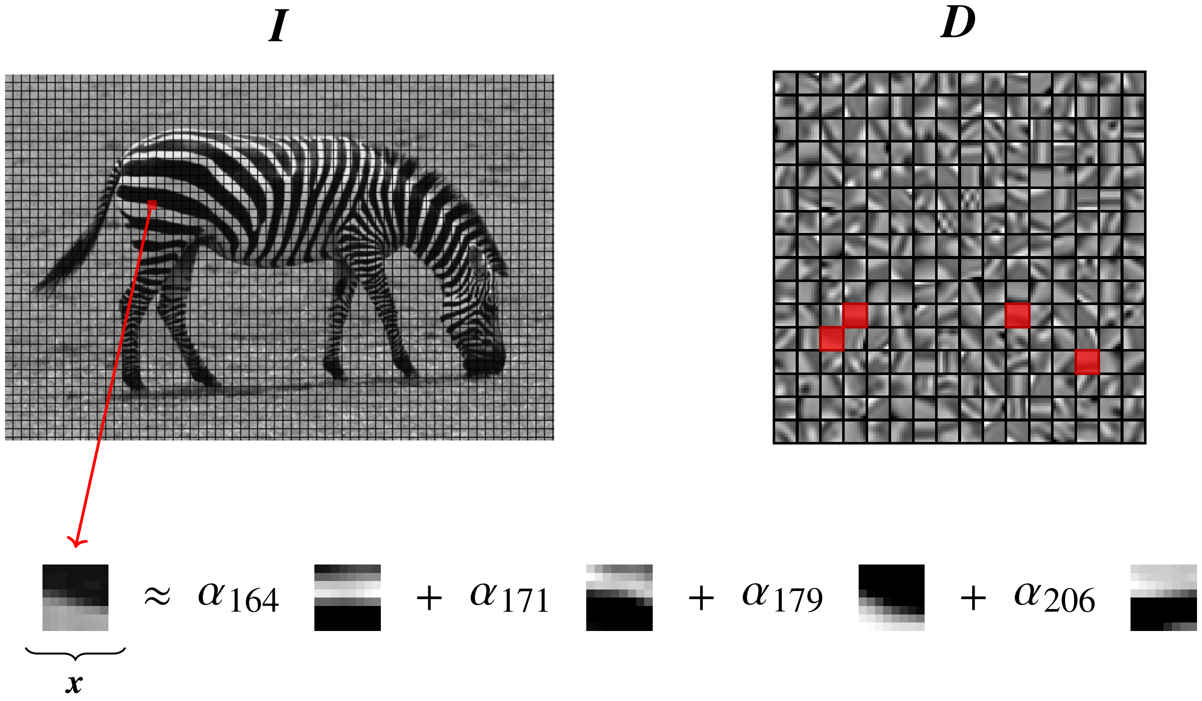
3. Simulation Setup and Results

4. Summary and Outlook

Sparse Image Representations

Image patches $\boldsymbol{x}$ can be represented by a sparse coefficient vector $\boldsymbol{\alpha}$ in certain dictionaries $\boldsymbol{D}$ up to an acceptable error

$$\boldsymbol{x} = \boldsymbol{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$$



Sparse Image Representations

- ➥ How to get a suitable dictionary $\boldsymbol{D}$ and the coefficient vector $\boldsymbol{\alpha}$?
  - **Choose** a dictionary such as the DCT basis functions or **learn** it such that it is optimized for a sparse representation
  - Calculate the coefficients via **correlation** in case of an orthogonal dictionary or via **sparse optimization**

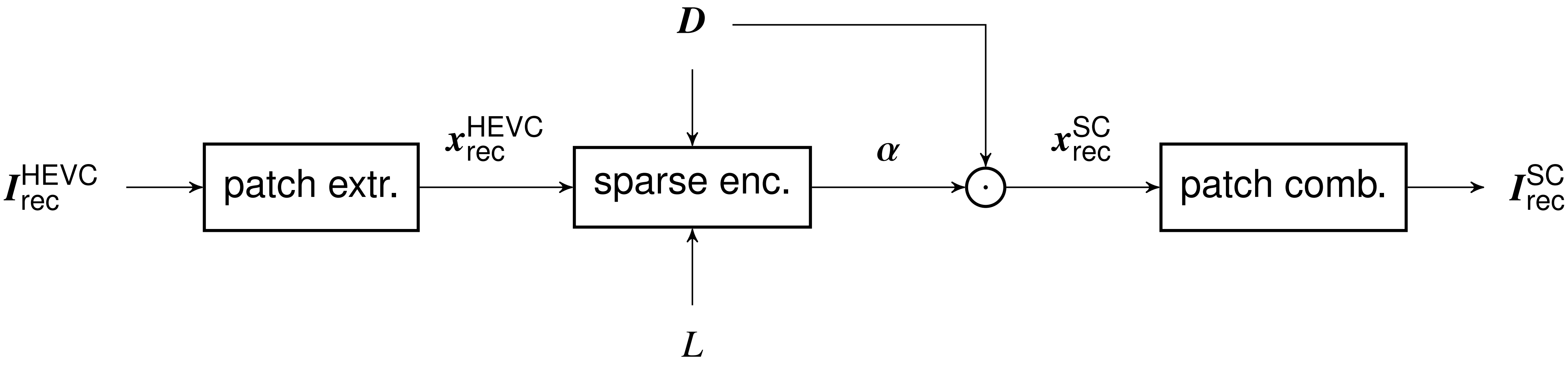
$$\boldsymbol{D} = \arg\min_{\boldsymbol{D}} \sum_{i=1}^{n} \frac{1}{2} \|\boldsymbol{x}_i - \boldsymbol{D}\boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\alpha}_i\|_1$$

$$\boldsymbol{\alpha} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{\alpha}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{\alpha}\|_0 = L$$

Important parameters: The sparsity parameters λ and L

Sparse Coding based In loop Filtering

The main idea comes from sparse coding based denoising [Ela10]



1. **Extract** overlapping patches $\boldsymbol{x}_{\text{rec}}^{\text{HEVC}}$ from the image to be filtered $\boldsymbol{I}_{\text{rec}}^{\text{HEVC}}$ and center these patches
2. Choose a suitable parameter L and **calculate** sparse codes $\boldsymbol{\alpha}$ in the dictionary $\boldsymbol{D}$ for all extracted patches $\boldsymbol{x}_{\text{rec}}^{\text{HEVC}}$
3. **Reconstruct** the patches calculating $\boldsymbol{x}_{\text{rec}}^{\text{SC}} \approx \boldsymbol{D}\boldsymbol{\alpha}$
4. **Combine** the reconstructed patches $\boldsymbol{x}_{\text{rec}}^{\text{SC}}$ back to an image $\boldsymbol{I}_{\text{rec}}^{\text{SC}}$ via averaging in overlapping areas

Sparse Coding based In loop Filtering

Interim conclusion:

✓ The learned dictionary introduces some prior knowledge on the image characteristics to denoising/loop filtering problem.

✓ It is known that sparse coding based denoising shows good results in case of additive Gaussian noise

Open questions:

➥ How to chose the parameter L?
  - The higher the noise level the lower L should be chosen, as choosing too many dictionary atoms will result in the representation of the noise itself.

➥ Is there any guarantee that it will also perform well in the case of coding noise?
  - No. Only in the case of no correlation between the noise and all dictionary atoms there is a guarantee to find the same sparse coefficients for a noisy patch as for the corresponding raw patch.
  - Therefore, we designed a frequency adaption model and a corresponding signaling of parameters.

Contents

1. Sparse Coding based Denoising

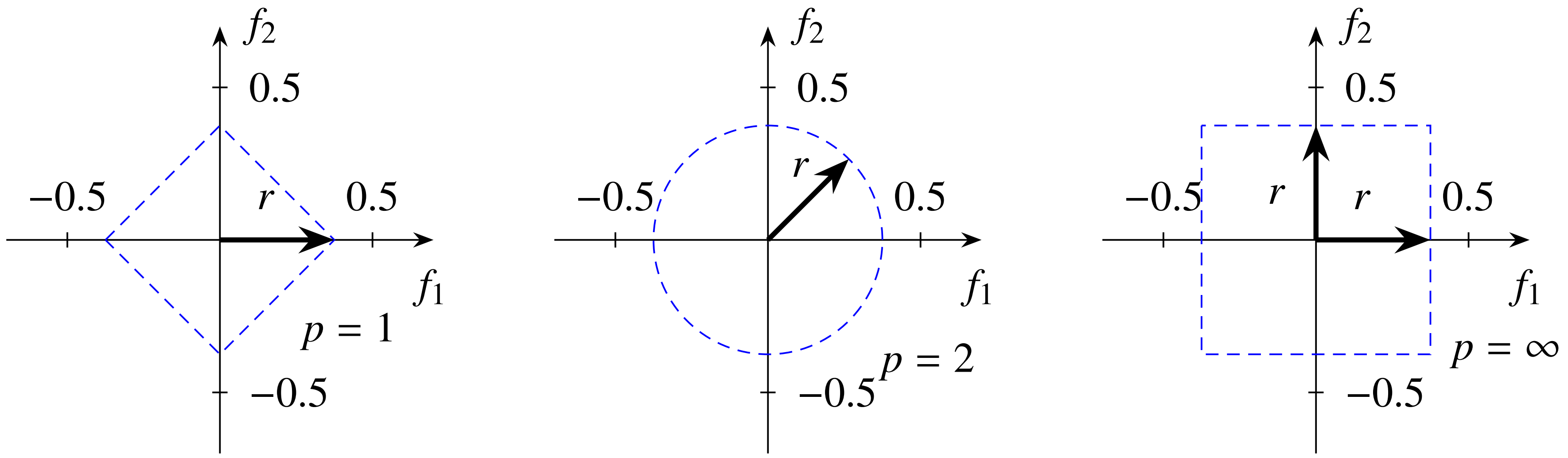
2. Frequency Adaptation Model

3. Simulation Setup and Results

4. Summary and Outlook

l_p-ball Energy Analysis

Energy measure of a 2 dimensional signal in the Fourier domain in dependence of the "radius" r and the shape p of an l_p-ball



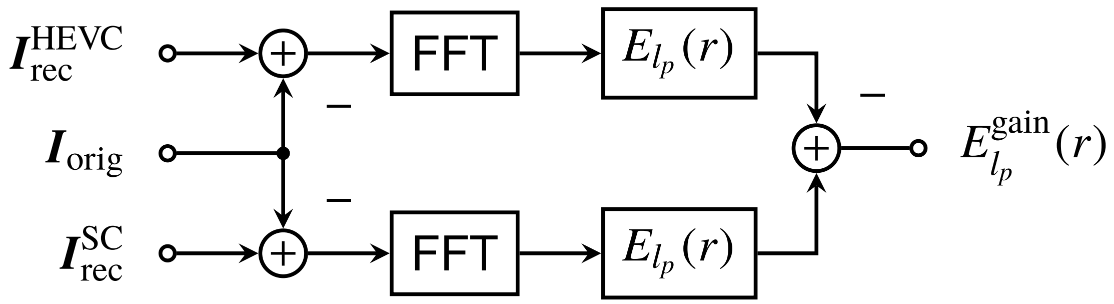
$$r = \left(|f_1|^p + |f_2|^p\right)^{\frac{1}{p}}$$

$$E_{l_p}(r) = \oint |\boldsymbol{S}(f_1, f_2)|^2 \, \mathrm{d}s$$

$$E_{l_p}(r) = \sum |\boldsymbol{S}(f_1, f_2)|^2$$

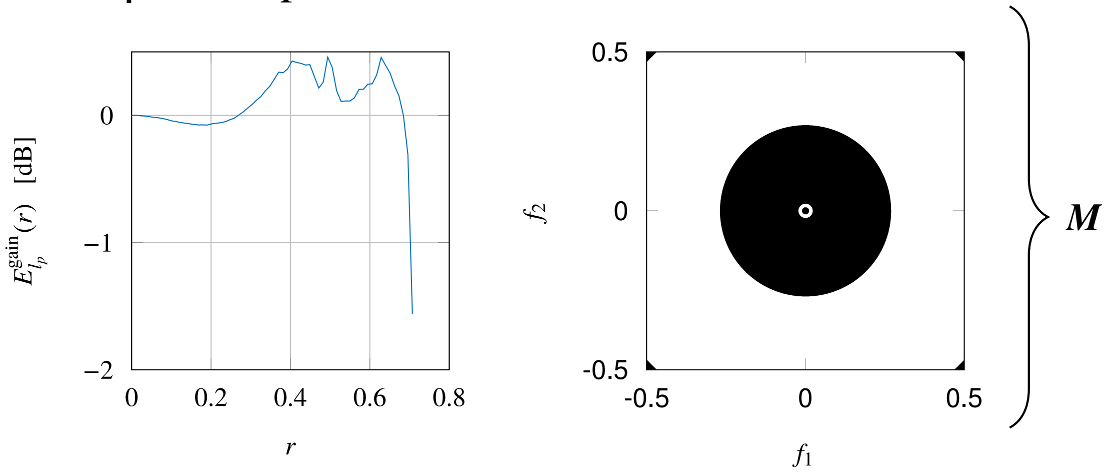
Frequency Adaption Model

Encoder side



- Calculate $E_{l_p}(r)$ for different parameters p and evaluate the performance of the sparse coding based in loop filter
- $E_{l_p}^{\text{gain}}(r)$ indicates frequency ranges in which the sparse coding based in loop filter outperforms the HEVC reconstruction

Example for $p = 2$



- $\boldsymbol{S}_{\text{rec}}^{\text{SCALF}} = \boldsymbol{M} \circ \boldsymbol{S}_{\text{rec}}^{\text{SC}} + \hat{\boldsymbol{M}} \circ \boldsymbol{S}_{\text{rec}}^{\text{HEVC}}$
- inverse Fourier transform of $\boldsymbol{S}_{\text{rec}}^{\text{SCALF}}$ results in the filtered Image $\boldsymbol{I}_{\text{rec}}^{\text{SCALF}}$
- The binary mask $\boldsymbol{M}$ needs to be transmitted to the decoder side

Frequency Adaption Model

Decoder side
Parsing of signaled parameters in the slice segment header

- `SCALFenabledFlag` indicating whether SCALF should be applied for the picture
- `ShapeIdx` indicating the parameter $p \in \{1, 2, \infty\}$ for the shape of the l_p-ball

The binary mask $\boldsymbol{M}$ is runlength coded with the following syntax elements:

- `NumRadiusIdc` indicating the total number of indices to the radius vector
- `RadiusIdc` containing the indices to the radius vector
- `MaskStartVal` containing the value of the filter mask at $r = 0$

Lehrstuhl und Institut für Nachrichtentechnik

Contents

1. Sparse Coding based Denoising

2. Frequency Adaptation Model

3. Simulation Setup and Results

4. Summary and Outlook

Simulation Setup

Training of the dictionary:

- # atoms in the dictionary $K = 512$
- patch size $s_{\mathrm{p}} = 8 \times 8$
- patches are fully overlapping, i.e. by 7 pixels
- the dictionary was trained with l_1-norm regularization with $\lambda = 0.15$

➥ No coded data was used for training

Sparse coding in encoder and decoder:

- $L = -\mathrm{QP} + 42$
  - for QP $\in \{22, 27, 32, 37\}$ this results in $L \in \{20, 15, 10, 5\}$

Results

BD-rate measurements for an All-Intra and a Randomaccess coding configuration

	AI		RA	
seq.	SCLF	SCALF	SCLF	SCALF
BQTerrace	−0.15 %	−0.24 %	−0.56 %	−0.85 %
BasketballDrive	−0.45 %	−0.43 %	−0.79 %	−0.88 %
Cactus	−1.16 %	−1.2 %	−0.96 %	−1.23 %
Kimono	−0.85 %	−0.86 %	−0.81 %	−0.97 %
ParkScene	−1.43 %	−1.5 %	−0.25 %	−0.44 %
PeopleOnStreet	−2.78 %	−2.86 %	−4.23 %	−4.6 %
Traffic	−1.84 %	−1.91 %	−2.37 %	−3.2 %
AVG	−1.24 %	−1.29 %	−1.42 %	−1.74 %

Table: BD-rate savings against HM-16.9 for different coding configurations and different sparse coding based in-loop filters.

Contents

1. Sparse Coding based Denoising

2. Frequency Adaptation Model

3. Simulation Setup and Results

4. Summary and Outlook

Summary and Outlook

Our contribution:

- sparse coding based in-loop filtering without any use of coded data in the training process
- general frequency adaption model, which can also be used in different applications or other in loop filters

Outlook:

- introduce an angular component to the frequency adaption model
  – specific directional structures can be supported

Thank you for your attention!

Any questions?

Jens Schneider
schneider@ient.rwth-aachen.de

Institut für Nachrichtentechnik, RWTH Aachen University
www.ient.rwth-aachen.de

References I

- Gisle Bjontegaard. *Calculation of average PSNR differences between RD-curves*. Technical report Doc. VCEG-M33. Austin, USA: ITU-T SG16/Q6 VCEG, 2001.
- Frank Bossen. *Common test conditions and software reference configurations*. Technical report. 12th Meeting, Geneva: JCT-VC, Jan. 2013.
- B. Bross, W. Han, J. Ohm, G. Sullivan, Y. Wang, and T. Wiegand. *High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS and Last Call)*. JCTVC-L1003. 12th Meeting, Geneva, Jan. 2013.
- Jianle Chen, Elena Alshina, Gary J. Sullivan, Jens-Rainer Ohm, and Jill Boyce. *Algorithm Description of Joint Exploration Test Model 3*. Technical report. 3rd Meeting, Geneva: JVET, May 2016.
- Yuanying Dai, Dong Liu, and Feng Wu. “A Convolutional Neural Network Approach for Post-Processing in HEVC Intra Coding”. In: *CoRR* abs/1608.06690 (2016). arXiv: 1608.06690.
- Michael Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. 1st. Springer Publishing Company, Incorporated, 2010.
- C. M. Fu, E. Alshina, A. Alshin, Y. W. Huang, C. Y. Chen, C. Y. Tsai, C. W. Hsu, S. M. Lei, J. H. Park, and W. J. Han. “Sample Adaptive Offset in the HEVC Standard”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 22.12 (Dec. 2012), pages 1755–1764.
- Cheolkon Jung, Licheng Jiao, Hongtao Qi, and Tian Sun. “Image deblocking via sparse representation”. In: *Signal Processing: Image Communication* 27.6 (2012), pages 663–677.
- X. Liu, X. Wu, J. Zhou, and D. Zhao. “Data-Driven Soft Decoding of Compressed Images in Dual Transform-Pixel Domain”. In: *IEEE Transactions on Image Processing* 25.4 (Apr. 2016), pages 1649–1659.
- Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. “Online dictionary learning for sparse coding”. In: *Proceedings of the 26th annual international conference on machine learning*. ACM. 2009, pages 689–696.
- A. Norkin, G. Bjontegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. Zhou, and G. Van der Auwera. “HEVC Deblocking Filter”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 22.12 (Dec. 2012), pages 1746–1754.
- W. S. Park and M. Kim. “CNN-based in-loop filtering for coding efficiency improvement”. In: *2016 IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. July 2016, pages 1–5.
- Jens Schneider, Johannes Sauer, and Mathias Wien. “Dictionary Learning based High Frequency Inter-Layer prediction for Scalable HEVC”. In: *Proc. of IEEE Visual Communications and Image Processing VCIP ’17*. St. Petersburg, USA: IEEE, Piscataway, Dec. 2017.

References II

- M. Wien. *High Efficiency Video Coding*. 1st edition. Springer-Verlag Berlin Heidelberg, 2015.
- J. Wang and B. Shim. “On the Recovery Limit of Sparse Signals Using Orthogonal Matching Pursuit”. In: *IEEE Transactions on Signal Processing* 60.9 (Sept. 2012), pages 4973–4976.
- Roman Zeyde, Michael Elad, and Matan Protter. “On single image scale-up using sparse-representations”. In: *International conference on curves and surfaces*. Springer. 2010, pages 711–730.