IMAGE PREDICTION BASED ON NON-NEGATIVE MATRIX FACTORIZATION

Mehmet TÜRKAN

Christine GUILLEMOT

ICASSP 2011
IEEE Int. Conf. on Acoustics, Speech and Signal Processing
May 22-27, 2011
Prague, Czech Republic

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA

centre de recherche
RENNES - BRETAGNE ATLANTIQUE

AGENDA

Problem Addressed

Prior work on Image Prediction

... Template Matching

... Sparse Approximations

A new approach based on ...

Non-negative Matrix Factorization

Experimental Results

Conclusion

Closed-Loop Image Prediction

H.264/AVC Intra Prediction

- ✓ homogeneous regions
- ✓ contours (if any of modes support the orientation)
- ✘ more complex structures and textural regions

Closed-Loop Image Prediction

H.264/AVC Intra Prediction

- ✓ homogeneous regions
- ✓ contours (if any of modes support the orientation)
- ✗ more complex structures and textural regions

Template Matching

- ✓ to cope with the **H.264/AVC** intra prediction lacks

CLOSED-LOOP IMAGE PREDICTION

H.264/AVC Intra Prediction

- ✓ homogeneous regions
- ✓ contours (if any of modes support the orientation)
- ✗ more complex structures and textural regions

Template Matching

- ✓ to cope with the **H.264/AVC** intra prediction lacks

Sparse Approximations

- ✓ a generalization of template matching

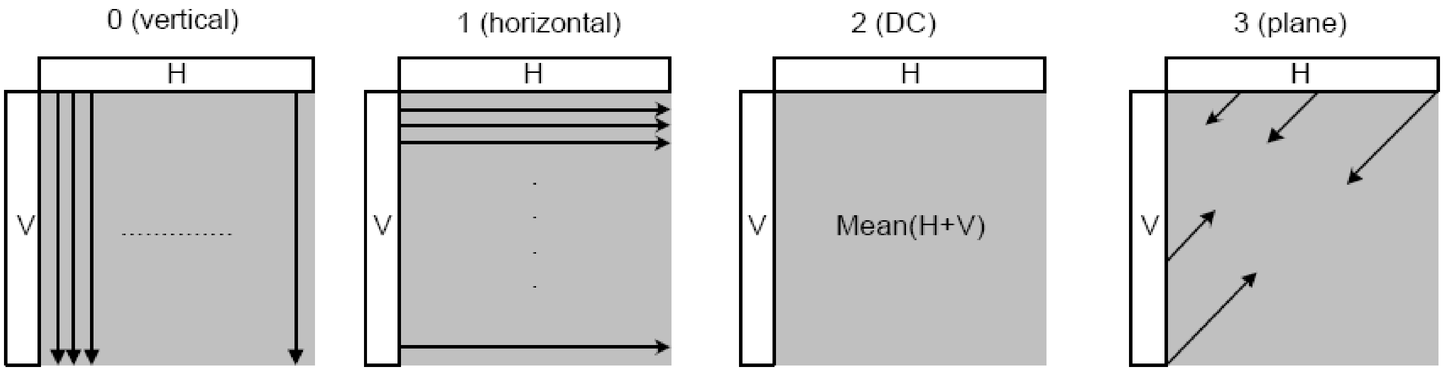
H.264/AVC Intra Prediction

Propagating pixel values along a specified direction using prior encoded samples from spatially neighbouring pixels

H.264/AVC Intra Prediction

Propagating pixel values along a specified direction using prior encoded samples from spatially neighbouring pixels

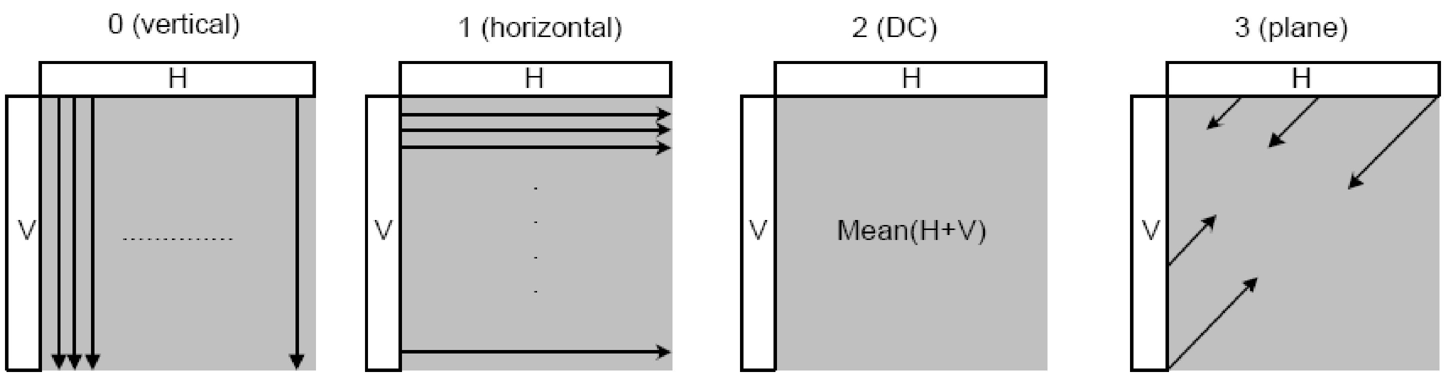
✓ **Intra-16x16** with 4 prediction modes (DC + 3 directional)



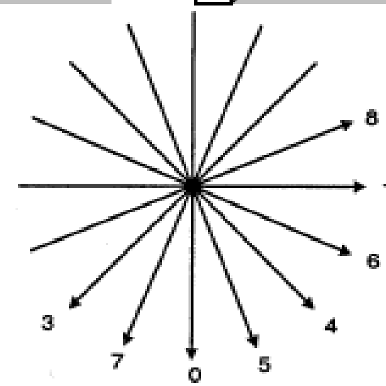
H.264/AVC INTRA PREDICTION

Propagating pixel values along a specified direction using prior encoded samples from spatially neighbouring pixels

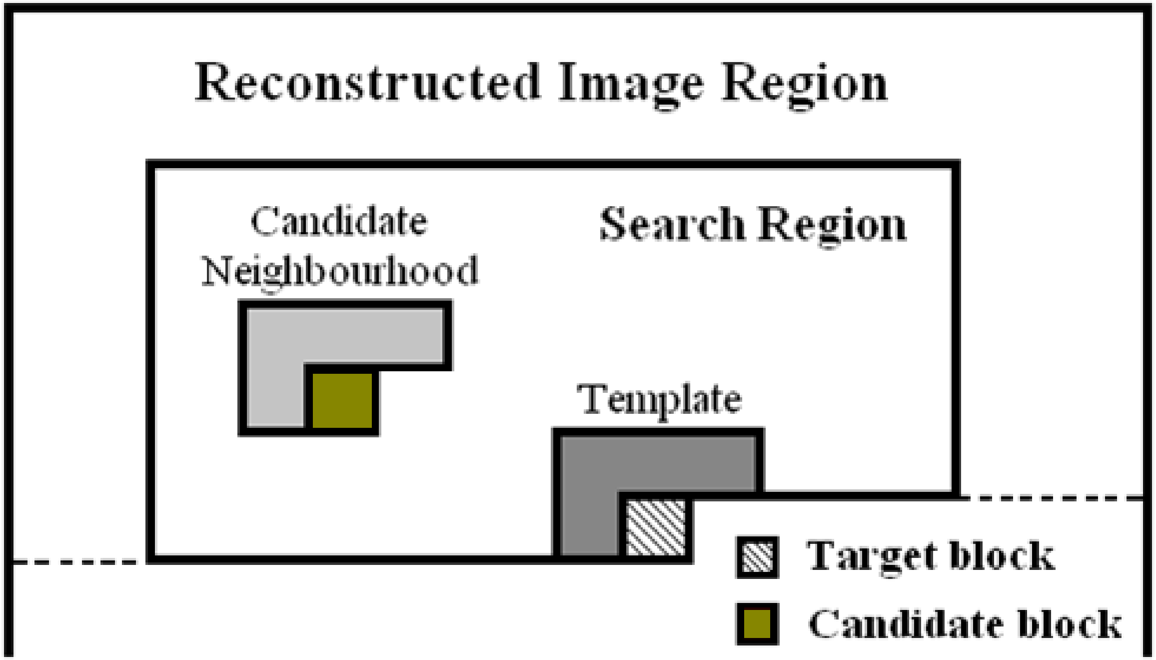
- ✓ **Intra-16x16** with 4 prediction modes (DC + 3 directional)



- ✓ **Intra-4x4** with 9 prediction modes (DC + 8 directional)



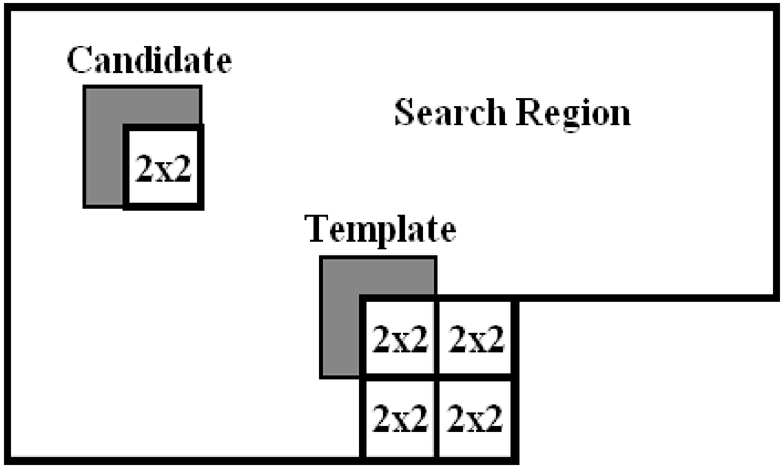
TEMPLATE MATCHING



The best candidate block is selected with

✓ the minimum distance between template and candidate block neighbourhood

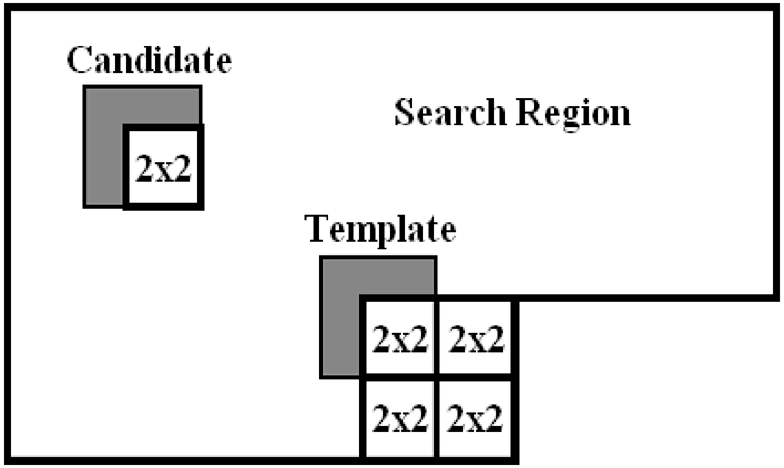
Template Matching – Example



An additional prediction mode in **H.264/AVC** (Intra-4x4)

✓ up to 11.3% bit-rate saving *(Tan et al. ICIP'06)*

TEMPLATE MATCHING – EXAMPLE



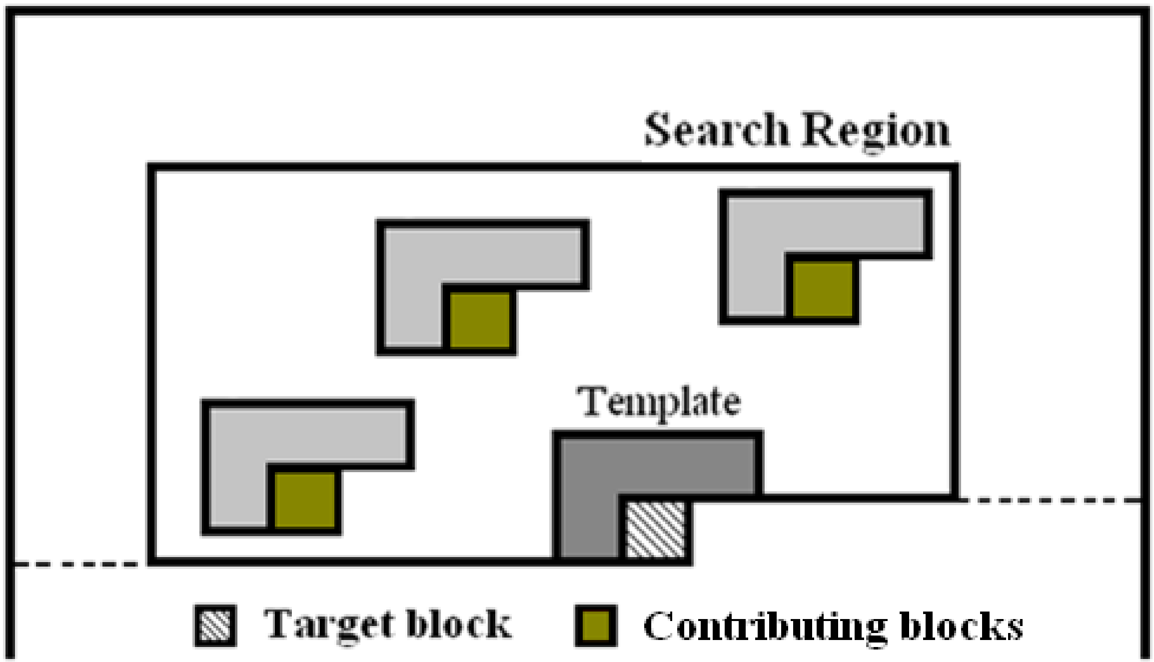
An additional prediction mode in **H.264/AVC** (Intra-4x4)

- ✓ up to 11.3% bit-rate saving *(Tan et al. ICIP'06)*

Averaging multiple predictors (larger and directional template)

- ✓ more than 15% bit-rate saving *(Tan et al. CCNC'07)*

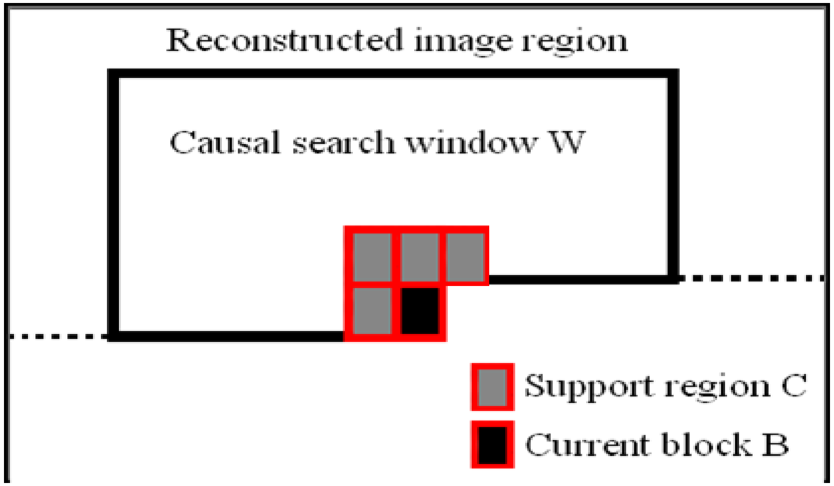
SPARSE PREDICTION



A linear combination approximation of the template

- ✓ weighting coefficients are calculated with a greedy sparse approximation algorithm such as **OMP**

NOTATION

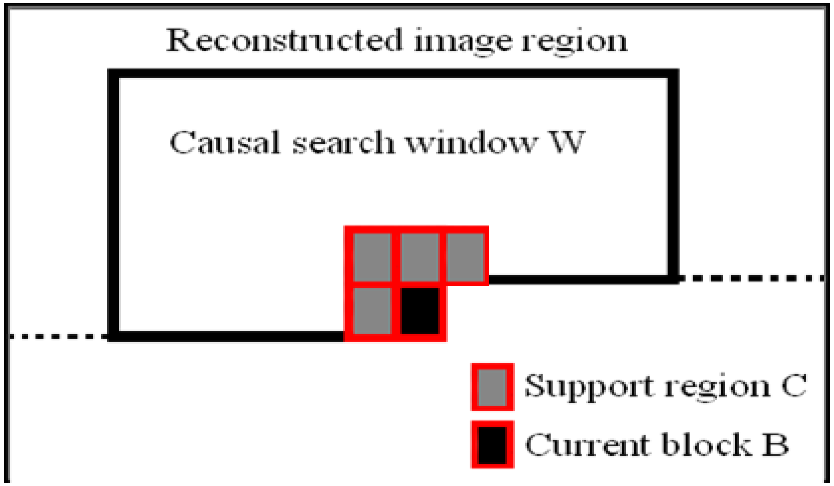


$\vec{b} \in \mathbf{R}^N$: stacked sample values of region S = B U C

$\vec{b}_c$: compacted data vector (support region C)

$\vec{b}_t$: actual values of the current block B

NOTATION



$A \in \mathbf{R}^{N \times M}$: stacked luminance values of all patches in W

A_c : compacted dictionary (corresponds to region C)

A_t : compacted dictionary (corresponds to region B)

Sparse prediction

Support region approximation with a constraint:

$$\vec{x}_{opt} = \min_{\vec{x}} \left\| \vec{b}_c - A_c \vec{x} \right\|_2^2 \quad \text{subject to}$$

$$\min_{\vec{x}} \left\| \vec{b}_t - A_t \vec{x} \right\|_2^2 \quad \text{and} \quad \left\| \vec{x} \right\|_0 \leq K$$

Sparse prediction

Support region approximation with a constraint:

$$\vec{x}_{opt} = \min_{\vec{x}} \left\| \vec{b}_c - A_c \vec{x} \right\|_2^2 \quad \text{subject to}$$

$$\min_{\vec{x}} \left\| \vec{b}_t - A_t \vec{x} \right\|_2^2 \quad \text{and} \quad \left\| \vec{x} \right\|_0 \leq K$$

*The **selected sparsity level** needs to be transmitted so that the decoder can exactly do the same prediction*

Sparse Prediction

Support region approximation with a constraint:

$$\vec{x}_{opt} = \min_{\vec{x}} \left\| \vec{b}_c - A_c \vec{x} \right\|_2^2 \quad \text{subject to}$$

$$\min_{\vec{x}} \left\| \vec{b}_t - A_t \vec{x} \right\|_2^2 \ \text{ and } \ \left\| \vec{x} \right\|_0 \leq K$$

*The **selected sparsity level** needs to be transmitted so that the decoder can exactly do the same prediction*

The predicted signal: $\hat{b}_t = A_t \vec{x}_{opt}$

Non-Negative Matrix Factorization

NMF: Low-rank representation of high-dimensional data

- Dimensionality reduction
- Data mining
- Noise removal

Non-negative matrix factorization

Given a non-negative matrix $B \in \mathbf{R}^{N \times L}$ and $M < \min(N, L)$

NMF tries to find matrix factors $A \in \mathbf{R}^{N \times M}$ and $X \in \mathbf{R}^{M \times L}$

Non-negative matrix factorization

Given a non-negative matrix $B \in \mathbf{R}^{N\times L}$ and $M < \min(N, L)$

NMF tries to find matrix factors $A \in \mathbf{R}^{N\times M}$ and $X \in \mathbf{R}^{M\times L}$

$$\min_{A,X}\left[\frac{1}{2}\left\|B - AX\right\|_F^2\right] \quad \text{subject to} \quad A \geq 0,\ X \geq 0$$

NON-NEGATIVE MATRIX FACTORIZATION

Given a non-negative matrix $B \in \mathbf{R}^{N \times L}$ and $M < \min(N, L)$

NMF tries to find matrix factors $A \in \mathbf{R}^{N \times M}$ and $X \in \mathbf{R}^{M \times L}$

$$\min_{A,X} \left[\frac{1}{2} \left\| B - AX \right\|_F^2 \right] \quad \text{subject to} \quad A \geq 0, \;\; X \geq 0$$

Multiplicative update equations *[Lee and Seung, 2000]*

$$X_{a,\mu} \leftarrow X_{a,\mu} \frac{\left(A^T B\right)_{a,\mu}}{\left(A^T A X\right)_{a,\mu} + 10^{-9}} \quad \text{and} \quad A_{i,a} \leftarrow A_{i,a} \frac{\left(B X^T\right)_{i,a}}{\left(A X X^T\right)_{i,a} + 10^{-9}}$$

Image Prediction based-on NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}} \left[\frac{1}{2} \left\| \vec{b} - A\vec{x} \right\|_2^2 \right] \quad \text{subject to} \quad A \geq 0, \;\; \vec{x} \geq 0$$

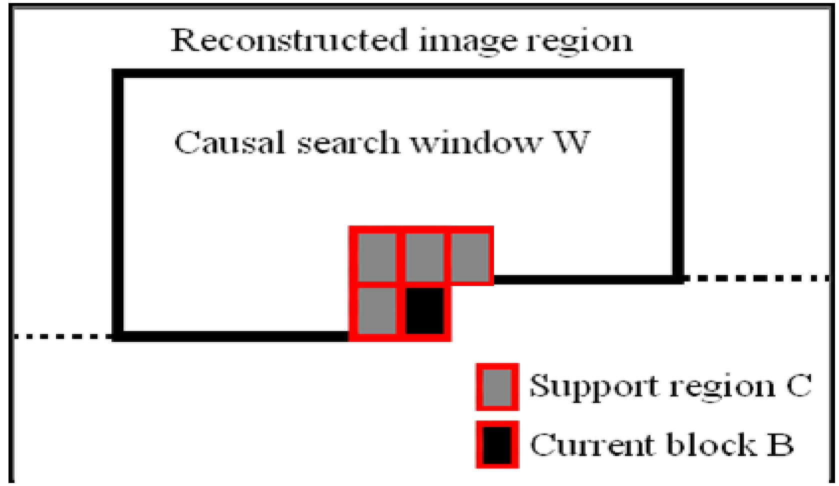
Image Prediction Based-on NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}} \left[\frac{1}{2} \left\| \vec{b} - A\vec{x} \right\|_2^2 \right] \quad \text{subject to} \quad A \geq 0, \;\; \vec{x} \geq 0$$

Idea: Fix A and b,

DICTIONARY CONSTRUCTION



$A \in \mathbf{R}^{N \times M}$: stacked luminance values of all patches in W

A_c : compacted dictionary (corresponds to region C)

A_t : compacted dictionary (corresponds to region B)

IMAGE PREDICTION BASED-ON NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}} \left[\frac{1}{2} \left\| \vec{b} - A\vec{x} \right\|_2^2 \right] \quad \text{subject to} \quad A \geq 0, \ \ \vec{x} \geq 0$$

Idea: Fix A and b,

IMAGE PREDICTION BASED-ON NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}} \left[\frac{1}{2} \left\| \vec{b} - A\vec{x} \right\|_2^2 \right] \quad \text{subject to} \quad A \geq 0, \ \ \vec{x} \geq 0$$

Idea: Fix A and b,

Find an NMF representation of the support region, and approximate the unkown block with the same parameters

Image prediction based-on NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}}\left[\frac{1}{2}\left\|\vec{b}-A\vec{x}\right\|_2^2\right] \quad \text{subject to} \quad A \geq 0,\ \ \vec{x} \geq 0$$

Idea: Fix A and b,

Find an NMF representation of the support region, and approximate the unkown block with the same parameters

$$\min_{\vec{x}:\vec{x}\geq 0}\left[\frac{1}{2}\left\|\vec{b}_c - A_c\vec{x}\right\|_2^2\right]$$

IMAGE PREDICTION BASED-ON NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}} \left[\frac{1}{2} \left\| \vec{b} - A\vec{x} \right\|_2^2 \right] \quad \text{subject to} \quad A \geq 0, \;\; \vec{x} \geq 0$$

Idea: Fix A and b,

Find an NMF representation of the support region, and approximate the unkown block with the same parameters

$$\min_{\vec{x}:\vec{x}\geq 0} \left[\frac{1}{2} \left\| \vec{b}_c - A_c\vec{x} \right\|_2^2 \right] \qquad x_a \leftarrow x_a \frac{\left(A_c^T \vec{b}_c\right)_a}{\left(A_c^T A_c \vec{x}\right)_a + 10^{-9}}, \;\; a = 1 \ldots M$$

IMAGE PREDICTION BASED-ON NMF

We can write NMF cost function in the **vector** form

$$\min_{A,\vec{x}} \left[\frac{1}{2} \left\| \vec{b} - A\vec{x} \right\|_2^2 \right] \quad \text{subject to} \quad A \geq 0, \;\; \vec{x} \geq 0$$

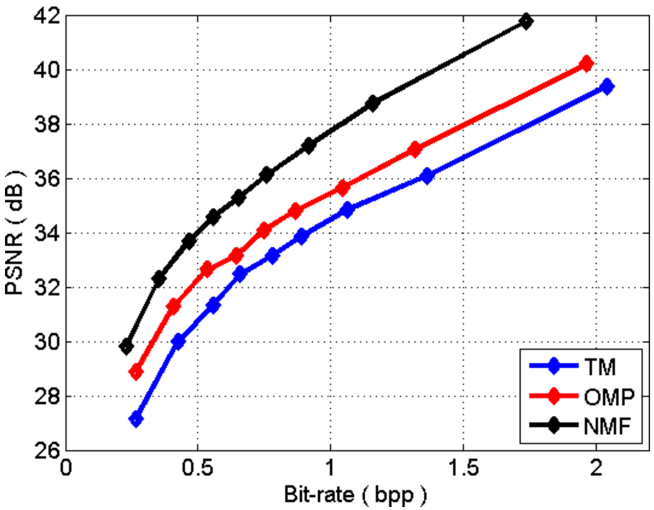
Idea: Fix A and b,

Find an NMF representation of the support region, and approximate the unkown block with the same parameters

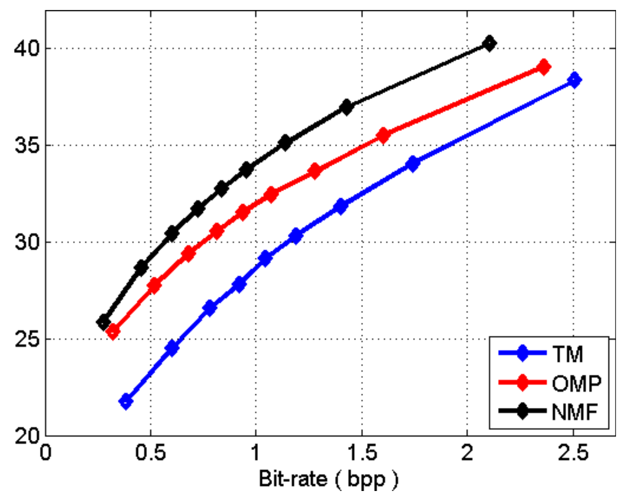
$$\min_{\vec{x}:\vec{x}\geq 0} \left[\frac{1}{2} \left\| \vec{b}_c - A_c\vec{x} \right\|_2^2 \right] \qquad x_a \leftarrow x_a \frac{\left(A_c^T \vec{b}_c\right)_a}{\left(A_c^T A_c \vec{x}\right)_a + 10^{-9}}, \quad a = 1 \ldots M$$

$$\hat{b}_t = A_t \vec{x}_{opt}$$

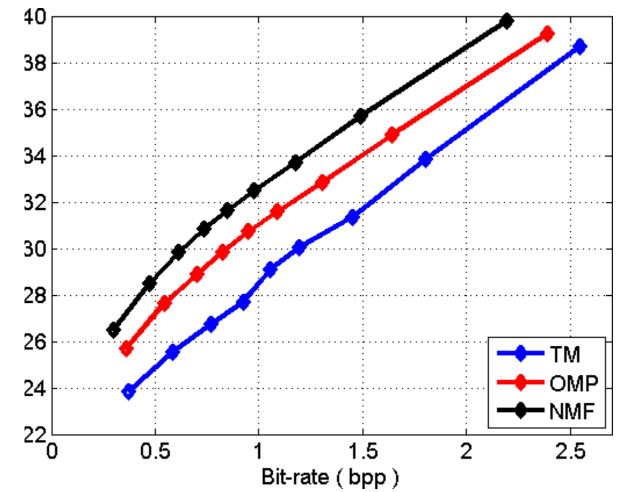
Experimental Results
Compression Efficiency



Foreman (CIF)



Barbara (512x512)



Cameraman (256x256)

OMP is iterated 8 iterations. Iteration number is Huffman encoded.
Prediction residue is transform encoded as in JPEG. (8x8 block size.)
The quantization is weighted by a factor (QP) varying between 10…90.

EXPERIMENTAL RESULTS RECONSTRUCTION QUALITY

Template Matching

(31.29dB @0.56bpp)

Sparse Approx.

(32.63dB @0.53bpp)

NMF

(33.68dB @0.46bpp)

EXPERIMENTAL RESULTS PREDICTION QUALITY

Template Matching
23.30db @QP=30

Sparse Approx.
26.14db @QP=30

NMF
24.37db @QP=30

Experimental Results
Prediction Quality

Template Matching
23.30db @QP=30

Sparse Approx.
26.14db @QP=30

NMF
24.37db @QP=30

Impose a sparsity constraint on NMF

Image prediction based-on NMF

Constraint: Use only k-NN patches, and keep track of the sparse vectors to optimize the prediction (k = 1...K)

IMAGE PREDICTION BASED-ON NMF

Constraint: Use only k-NN patches, and keep track of the sparse vectors to optimize the prediction (k = 1...K)

$$\vec{x}_{opt} = \min_{\vec{x}:\vec{x}\geq 0}\left[\frac{1}{2}\left\|\vec{b}_c - A_c\vec{x}\right\|_2^2\right] \quad \text{subject to}$$

$$\min_{\vec{x}}\left\|\vec{b}_t - A_t\vec{x}\right\|_2^2 \quad \text{and} \quad \left\|\vec{x}\right\|_0 \leq K$$

Image Prediction Based-on NMF

Constraint: Use only k-NN patches, and keep track of the sparse vectors to optimize the prediction (k = 1...K)

$$\vec{x}_{opt} = \min_{\vec{x}:\vec{x}\geq 0}\left[\frac{1}{2}\left\|\vec{b}_c - A_c\vec{x}\right\|_2^2\right] \quad \text{subject to}$$

$$\min_{\vec{x}}\left\|\vec{b}_t - A_t\vec{x}\right\|_2^2 \quad \text{and} \quad \left\|\vec{x}\right\|_0 \leq K$$

*The **selected k value** needs to be transmitted so that the decoder can run with the same number of patches*

IMAGE PREDICTION BASED-ON NMF

Constraint: Use only k-NN patches, and keep track of the sparse vectors to optimize the prediction (k = 1...K)

$$\vec{x}_{opt} = \min_{\vec{x}:\vec{x}\geq 0}\left[\frac{1}{2}\left\|\vec{b}_c - A_c\vec{x}\right\|_2^2\right] \quad \text{subject to}$$

$$\min_{\vec{x}}\left\|\vec{b}_t - A_t\vec{x}\right\|_2^2 \quad \text{and} \quad \left\|\vec{x}\right\|_0 \leq K$$

The ***selected k value*** *needs to be transmitted so that the decoder can run with the same number of patches*

The predicted signal: $\hat{b}_t = A_t\vec{x}_{opt}$

Image prediction based-on NMF

Computational load is reduced with sparsity constraint

IMAGE PREDICTION BASED-ON NMF

Computational load is reduced with sparsity constraint

→ more prediction modes can be introduced

IMAGE PREDICTION BASED-ON NMF

Computational load is reduced with sparsity constraint

→ more prediction modes can be introduced

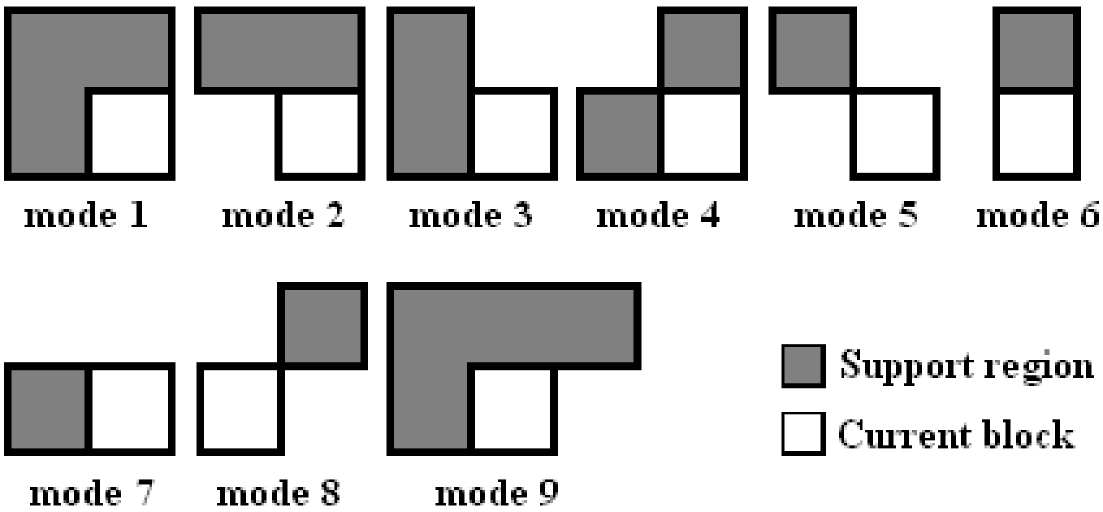
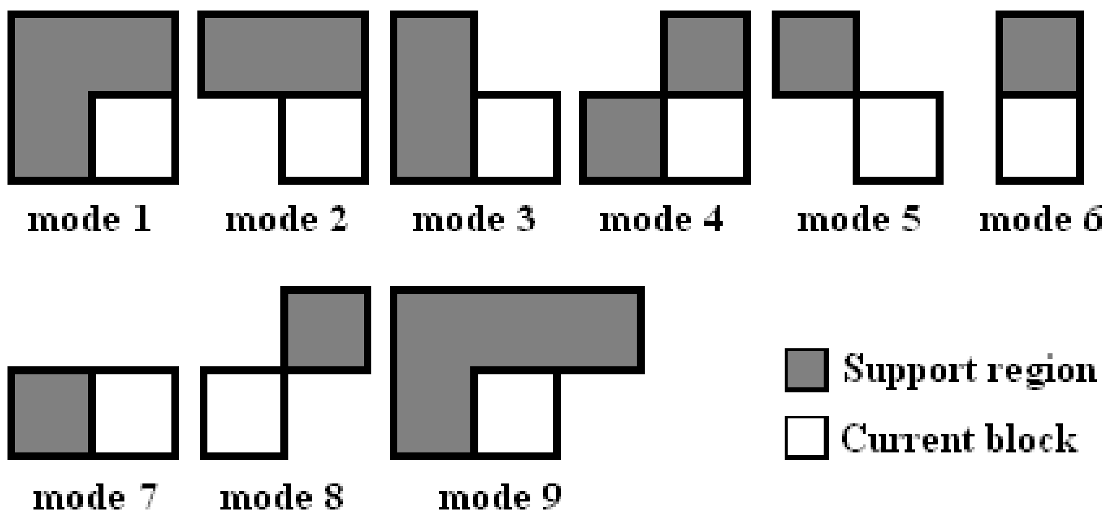


IMAGE PREDICTION BASED-ON NMF

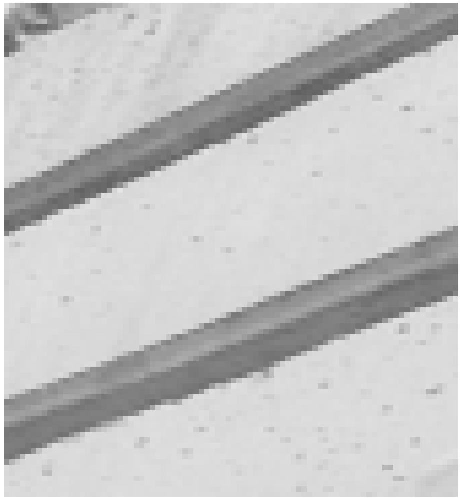
Computational load is reduced with sparsity constraint

→ more prediction modes can be introduced

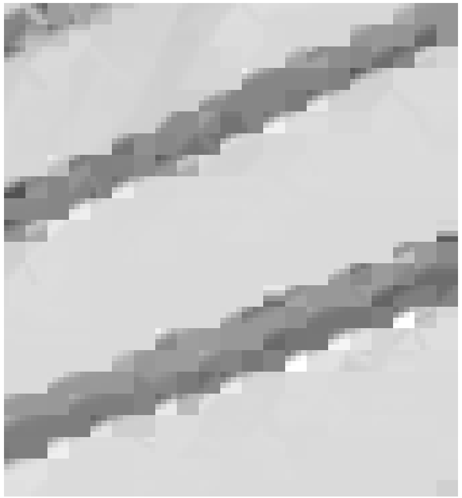


*The **selected mode** needs to be signalled so that the decoder can do the same prediction*

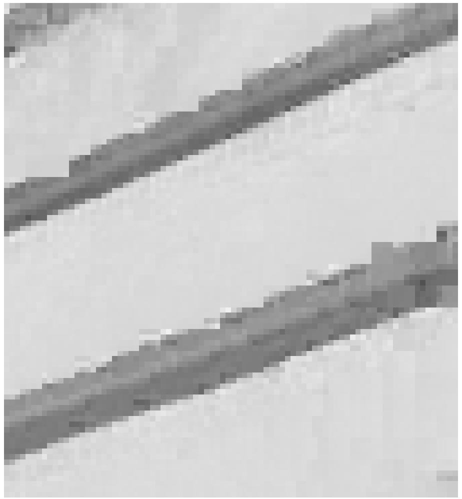
Experimental results
Prediction quality with sparsity



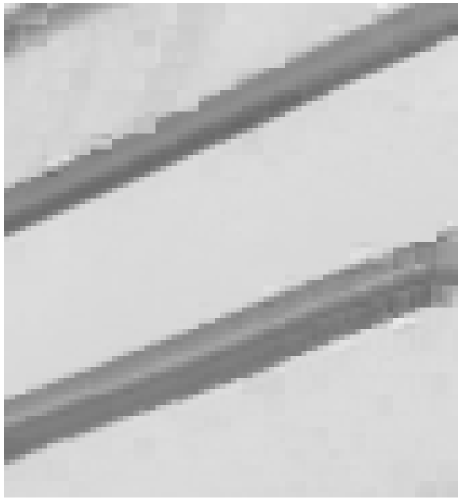
Original



H.264 intra
@QP=10

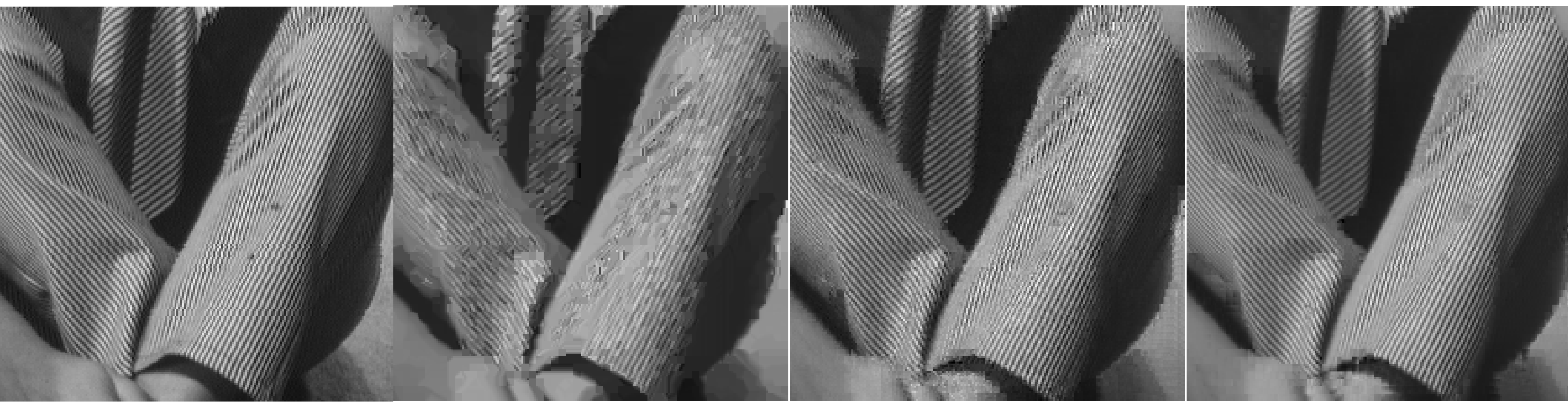


Sparse Approx.
@QP=10



Sparse NMF
@QP=10

Experimental results
Prediction quality with sparsity



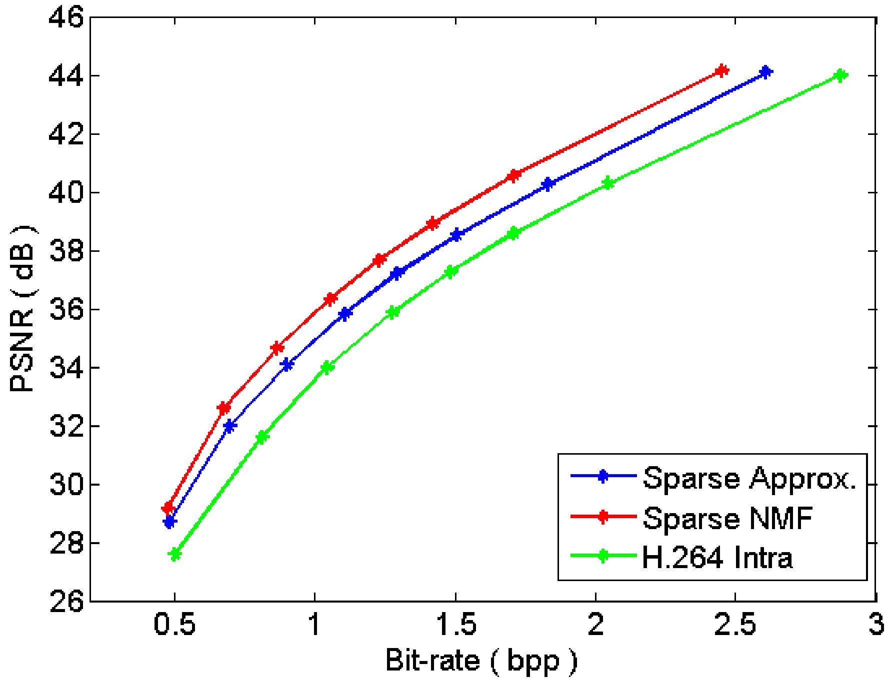
Original

H.264 intra
@QP=10

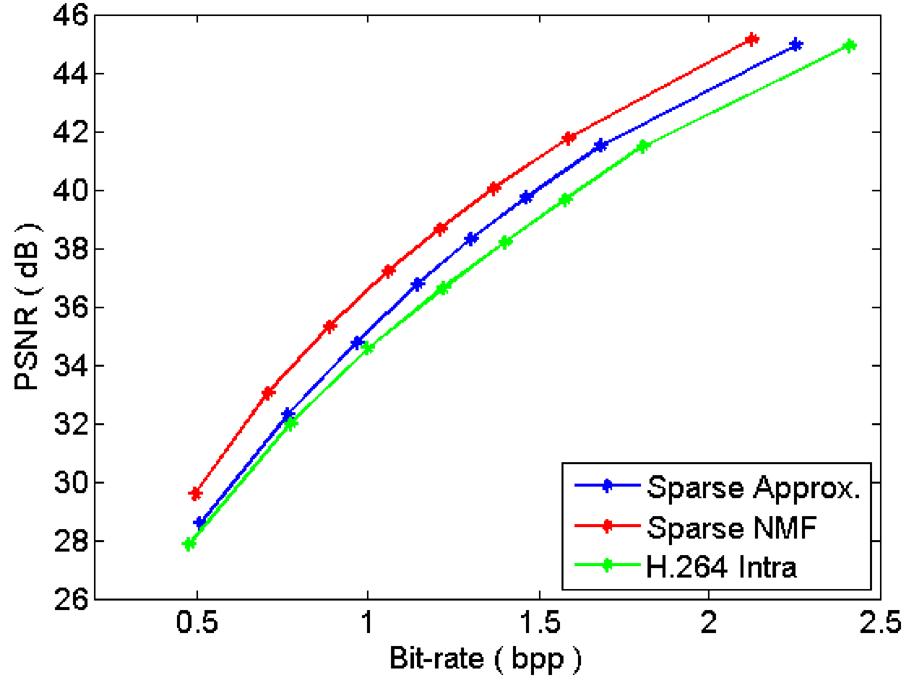
Sparse Approx.
@QP=10

Sparse NMF
@QP=10

Experimental Results
Compression Efficiency



Barbara (512x512)



Roof (512x512)

OMP is iterated 8 iterations, also K= 8. (Huffman encoded)
Prediction residue is transform encoded. 4x4 block size.
Best prediction mode and k value is selected by an RD cost function

CONCLUSION

A spatial texture prediction method is introduced

- ✓ in the context of a data dimensionality reduction method

CONCLUSION

A spatial texture prediction method is introduced

- ✓ in the context of a data dimensionality reduction method
- ✓ with sparsity constraints, it works even better

Conclusion

A spatial texture prediction method is introduced

- ✓ in the context of a data dimensionality reduction method
- ✓ with sparsity constraints, it works even better
- ✓ can be applied also to image inpainting and loss concealment applications

CONCLUSION

A spatial texture prediction method is introduced

- ✓ in the context of a data dimensionality reduction method
- ✓ with sparsity constraints, it works even better
- ✓ can be applied also to image inpainting and loss concealment applications

❑ An effective alternative when compared to other methods

CONCLUSION

A spatial texture prediction method is introduced

- ✓ in the context of a data dimensionality reduction method
- ✓ with sparsity constraints, it works even better
- ✓ can be applied also to image inpainting and loss concealment applications

❑ An effective alternative when compared to other methods

Questions ?