INTRA-FRAME PREDICTION WITH LAPPED TRANSFORMS FOR IMAGE CODING

Rafael Galvão de Oliveira

Beatrice Pesquet

Overview

- Motivation
- Lapped transforms: background
- Intra-frame prediction
  - Difficulties for LT
  - Proposed Solution
  - Modification in the mode selection
- Experimental Results
- Conclusion and future work

Motivation: Lapped transforms in image coding

- **Reduction of blocking effects**
- **Exploitation of redundancies among neighboring blocks**
- **Superior performance for encoding**
  - Objectively and subjectively
  - Higher complexity
- **JPEG XR**

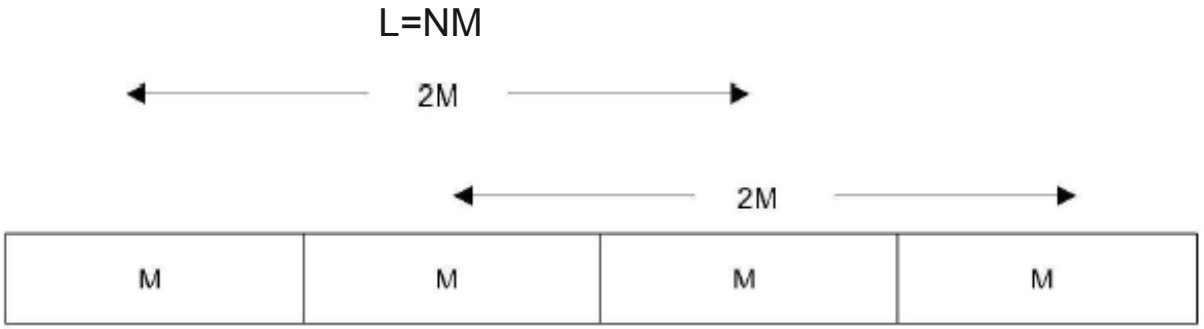
[1] de Oliveira, R.G.; de Queiroz, R.L.; , "Intra prediction versus wavelets and lapped transforms in an H.264/AVC coder," *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on* , vol., no., pp.137-140, 12-15 Oct. 2008

Overview

- Motivation
- Lapped transforms: background
- Intra-frame prediction
  - Difficulties for LT
  - Proposed Solution
  - Modification in the mode selection
- Experimental Results
- Conclusion and future work

Lapped Transforms: background

- **Differently from block transforms, lapped transforms have a bigger support than the traditional block**



$$\mathbf{v_m^T} = \left[\mathbf{x}\left(\mathbf{mM} - (\mathbf{N}-1)\frac{\mathbf{M}}{\mathbf{2}}\right) \cdots \mathbf{x}\left(\mathbf{mM} - (\mathbf{N}+1)\frac{\mathbf{M}}{\mathbf{2}} - \mathbf{1}\right)\right]$$

Lapped Transforms: background

- **Direct Transform**

$$\mathbf{Y_m} = \mathbf{P} \cdot \mathbf{v_m}$$

$$\mathbf{P} = [\mathbf{P_0 P_1} \ldots \mathbf{P_{N-1}}]$$

$$\mathbf{Q} = [\mathbf{Q_0 Q_1} \ldots \mathbf{Q_{N-1}}]$$

- **Inverse Transform**

$$\hat{\mathbf{v}}_\mathbf{m} = \mathbf{Q^T} \cdot \mathbf{Y_m}$$

$$\hat{\mathbf{v}}_\mathbf{m} \neq \mathbf{v_m}$$

$$\mathbf{Q^T} \cdot \mathbf{P} \neq \mathbf{I}$$

- **Perfect reconstruction**

$$\sum_{k=0}^{N-1-m} \mathbf{Q_k^T P_{k+m}} = \sum_{\mathbf{k=0}}^{\mathbf{N-1-m}} \mathbf{Q_{k+m}^T P_k} = \delta(\mathbf{m})\, \mathbf{I_M}$$

TELECOM ParisTech

Lapped Transforms: reconstruction

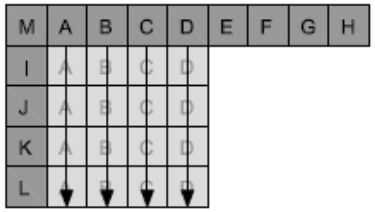
- The blocks cannot be reconstructed independently
- The signal is reconstructed accumulating the contribution of several neighboring blocks
- Border treatment has to be carried out

Overview

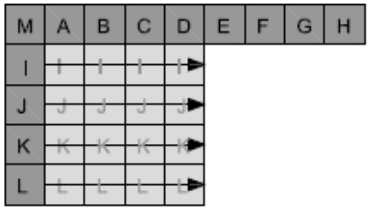
- Motivation
- Lapped transforms: background
- Intra-frame prediction
  - Difficulties for LT
  - Proposed Solution
  - Modification in the mode selection
- Experimental Results
- Conclusion and future work

Intra-frame prediction in H.264/AVC

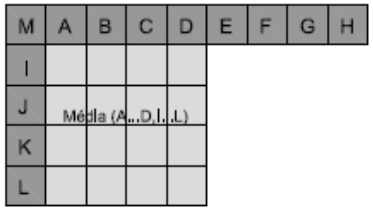
4x4 and
8x8 blocks



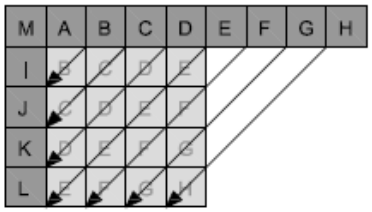
Mode 0



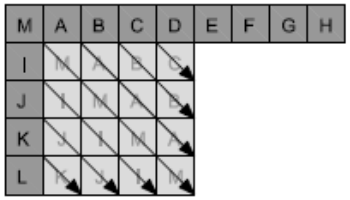
Mode 1



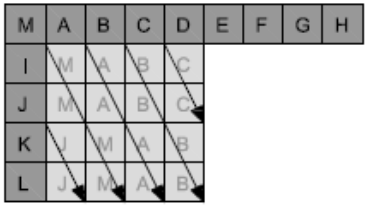
Mode 2



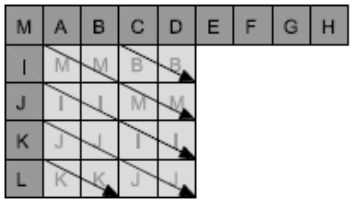
Mode 3



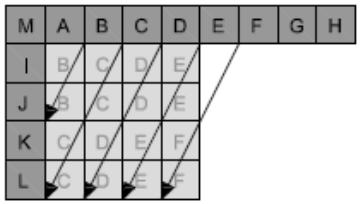
Mode 4



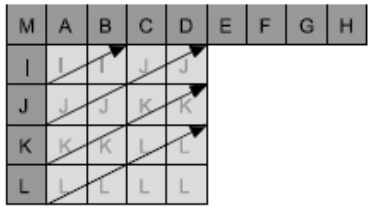
Mode 5



Mode 6



Mode 7

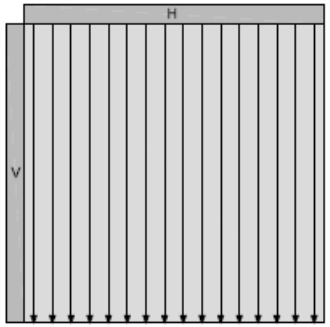


Mode 8

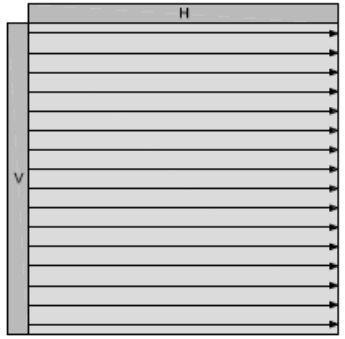
[2] de Queiroz, R.L.; Ortis, R.S.; Zaghetto, A.; Fonseca, T.A.; , "Fringe benefits of the H.264/AVC," *Telecommunications Symposium, 2006 International* , vol., no., pp.166-170, 3-6 Sept. 2006

Intra-frame prediction in H.264/AVC

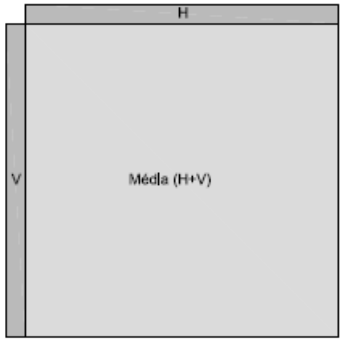
16x16 blocks



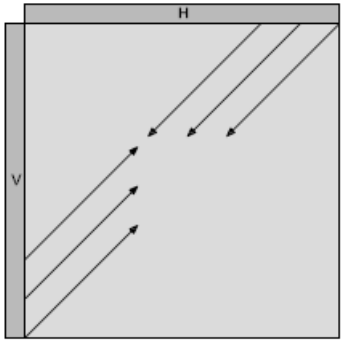
Mode 0



Mode 1

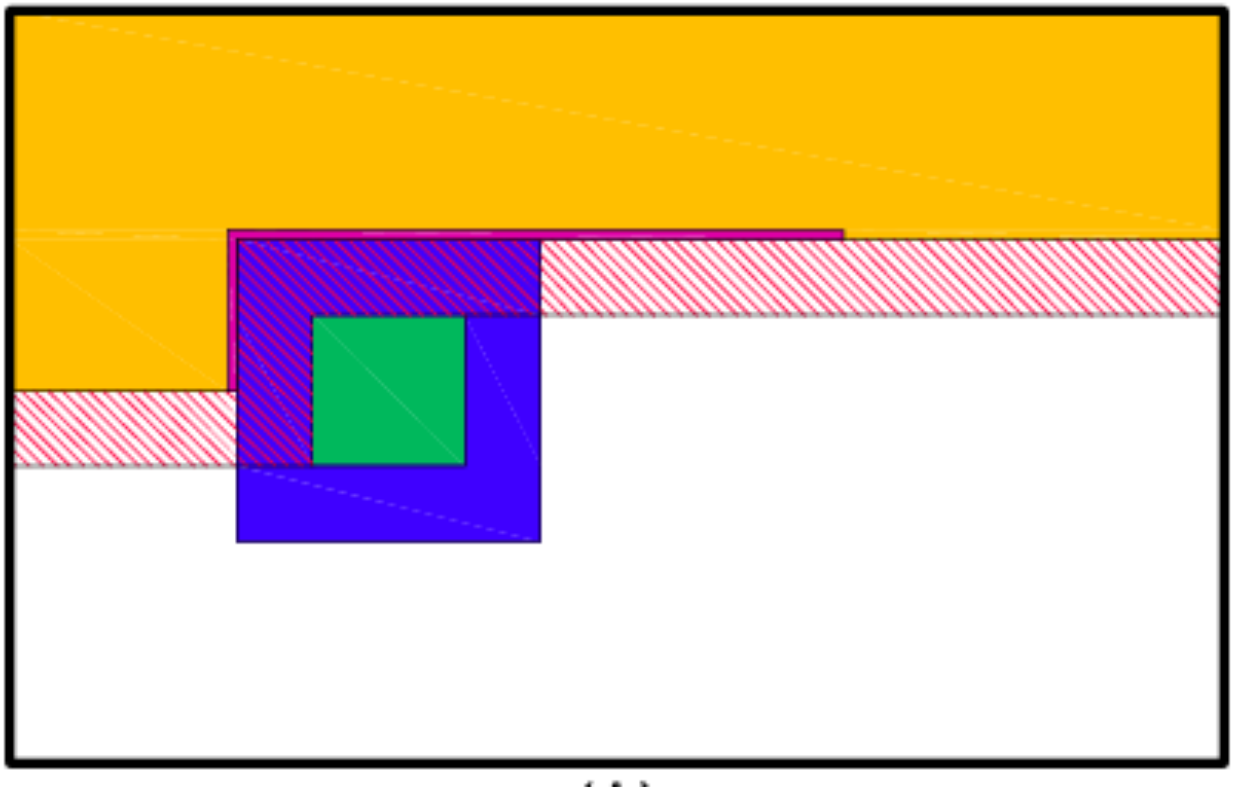


Mode 2

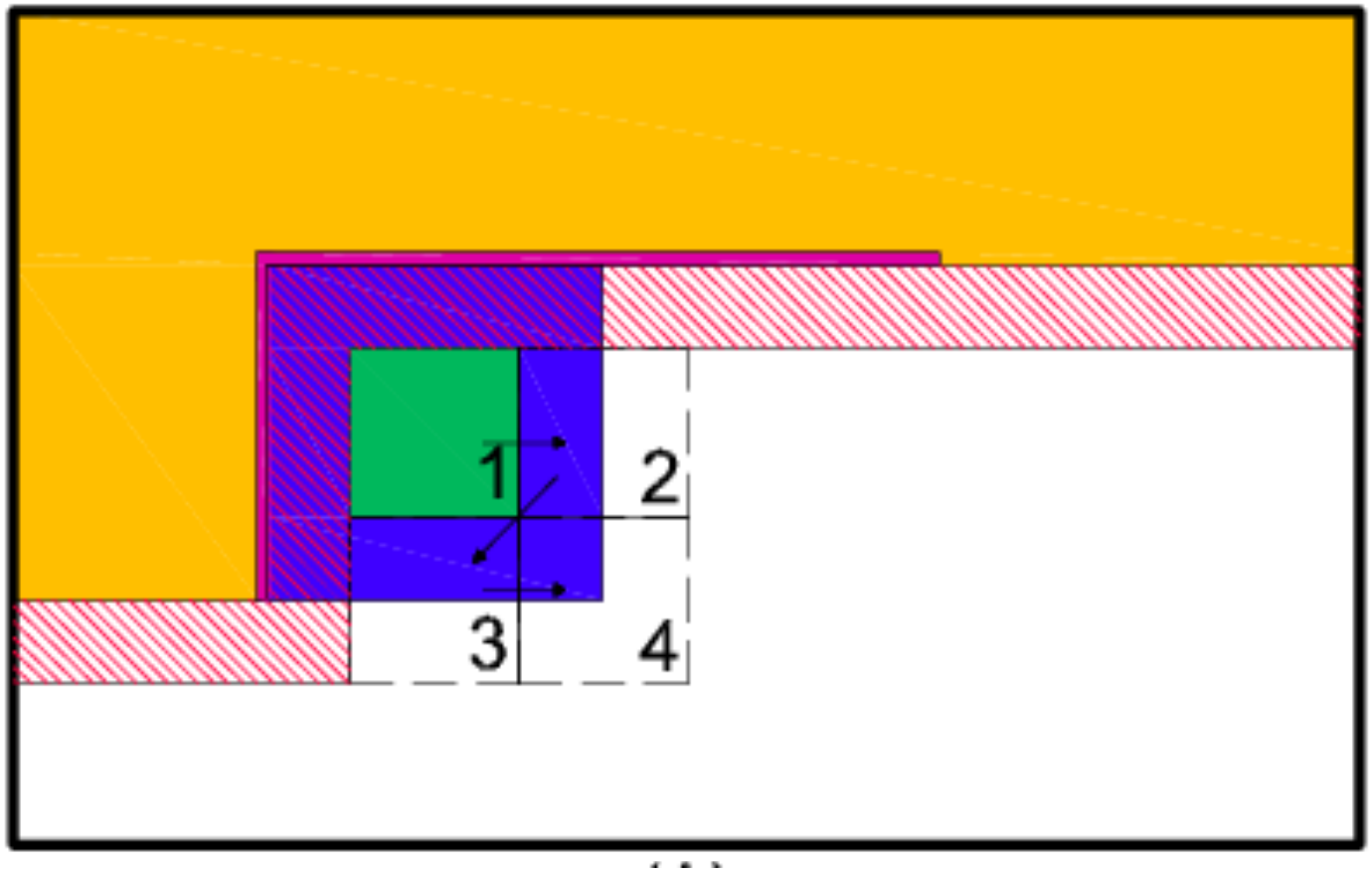


Mode 3

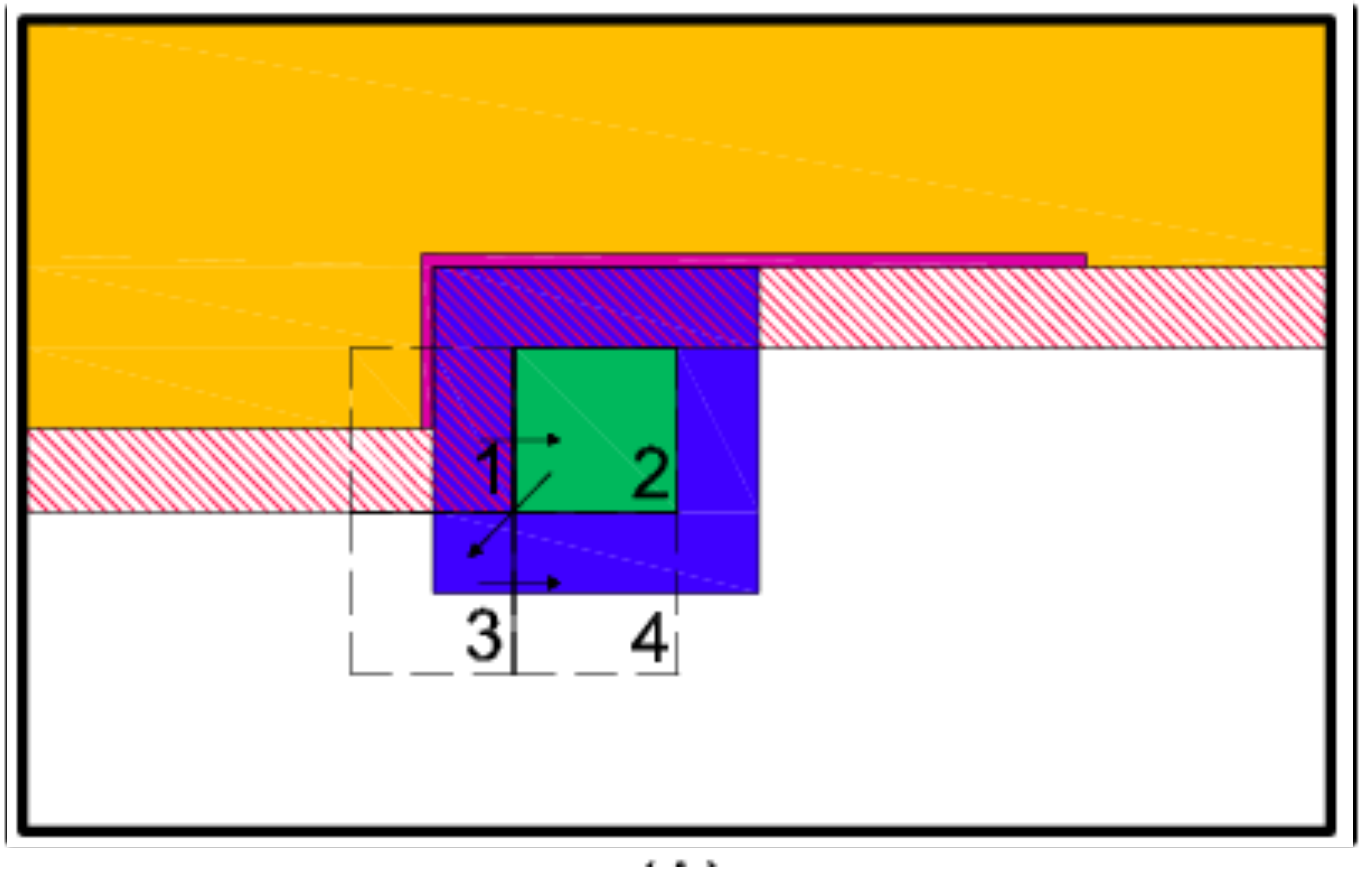
Intra frame prediction and Lapped Transforms



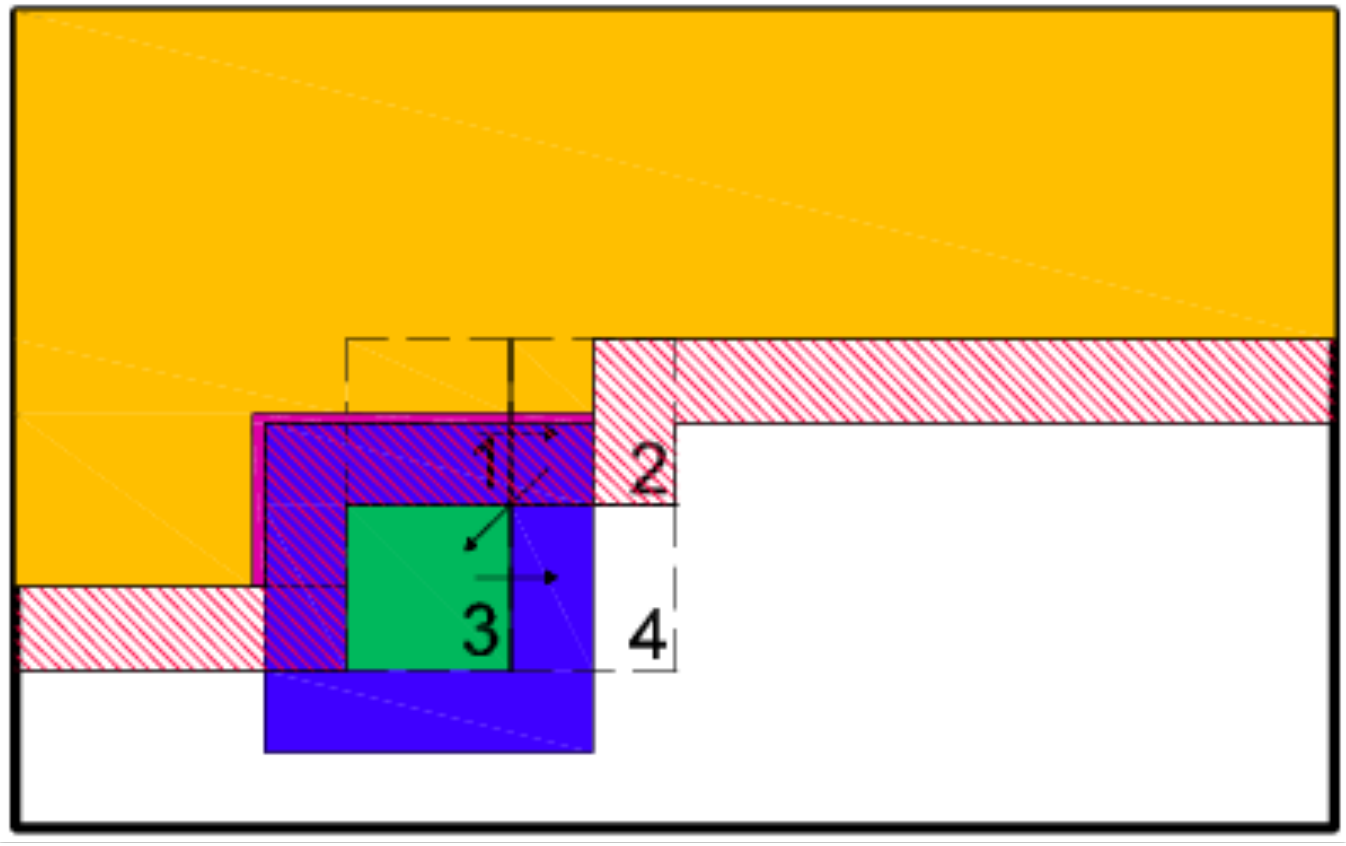
Intra frame prediction and Lapped Transforms



Intra frame prediction and Lapped Transforms

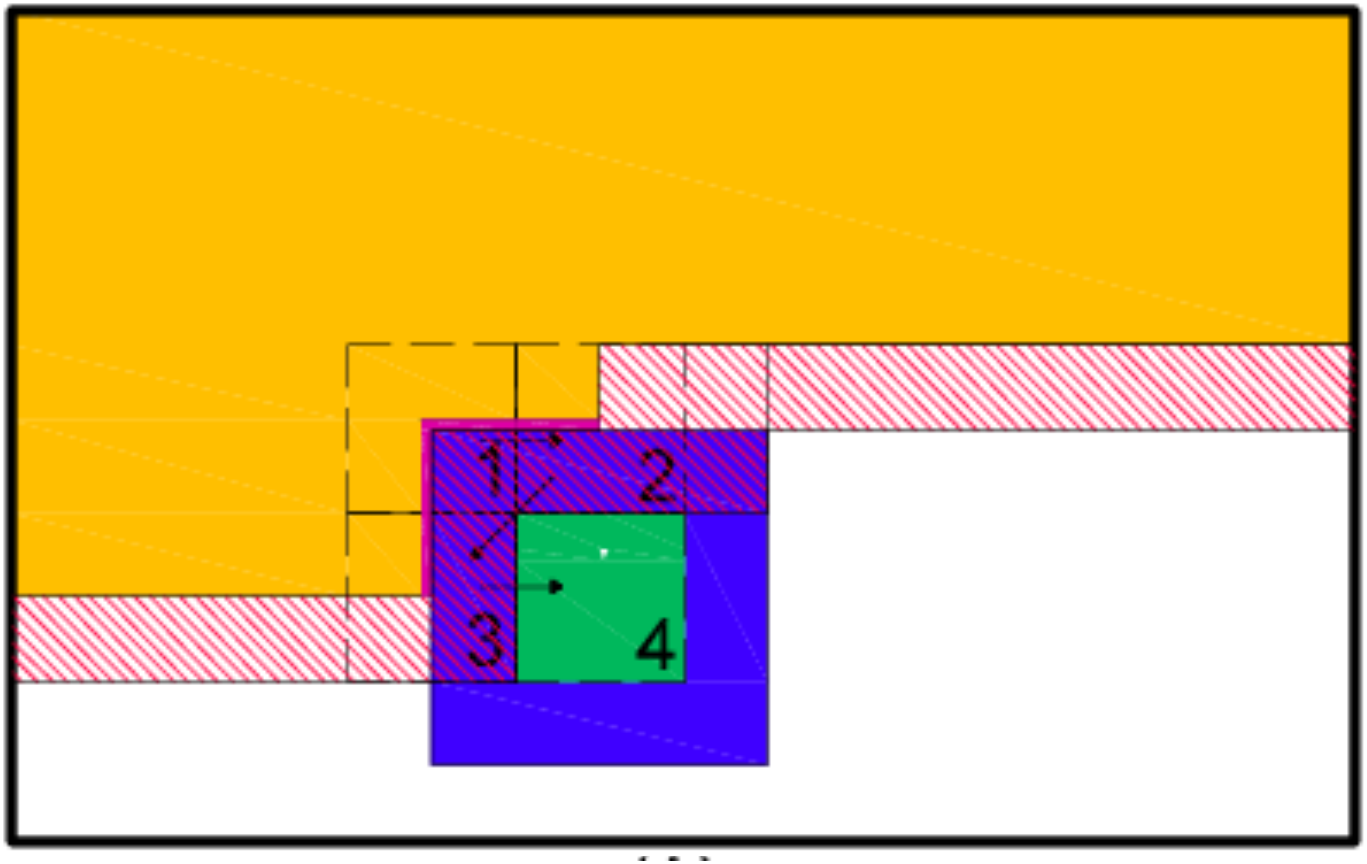


Intra frame prediction and Lapped Transforms



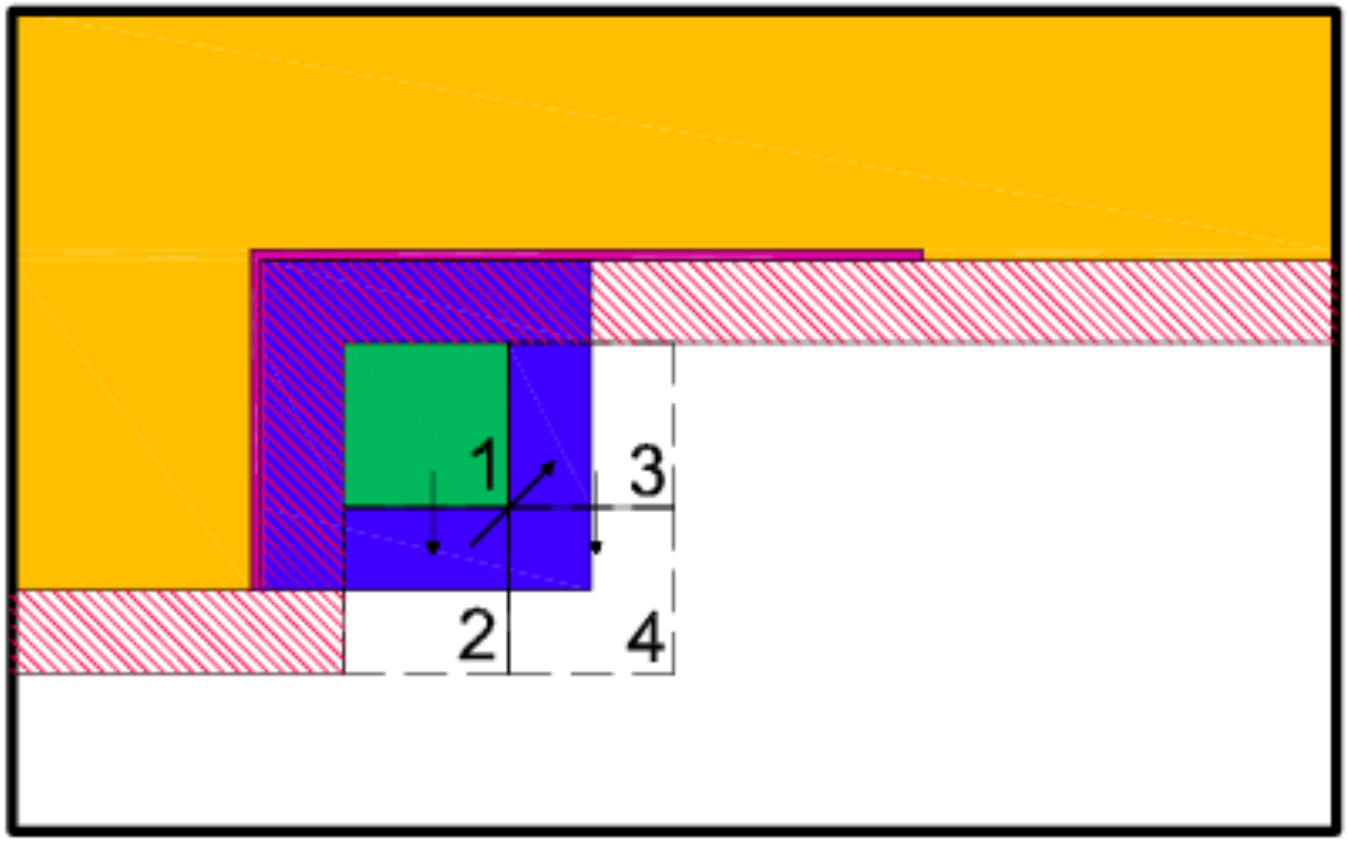
TELECOM ParisTech

Intra frame prediction and Lapped Transforms



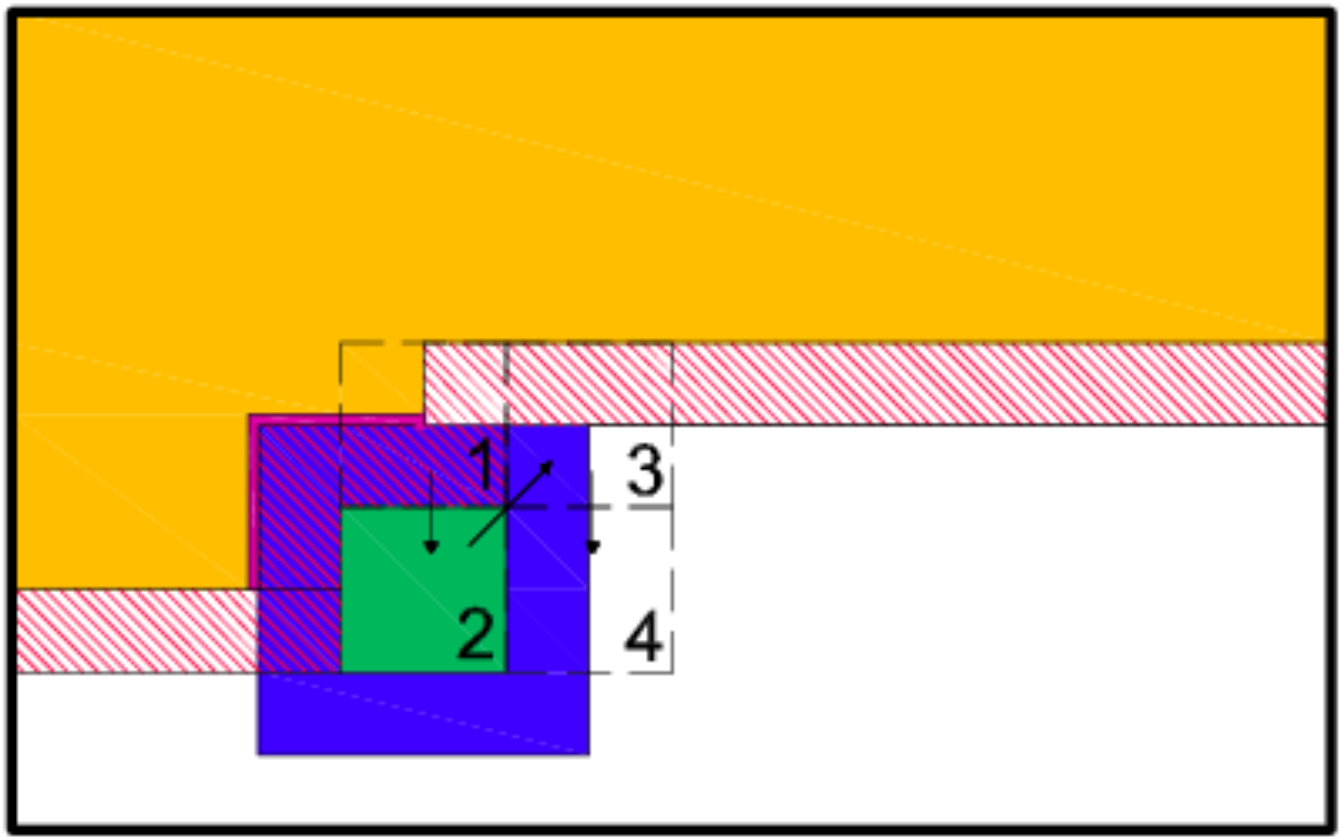
TELECOM ParisTech

Intra frame prediction and Lapped Transforms

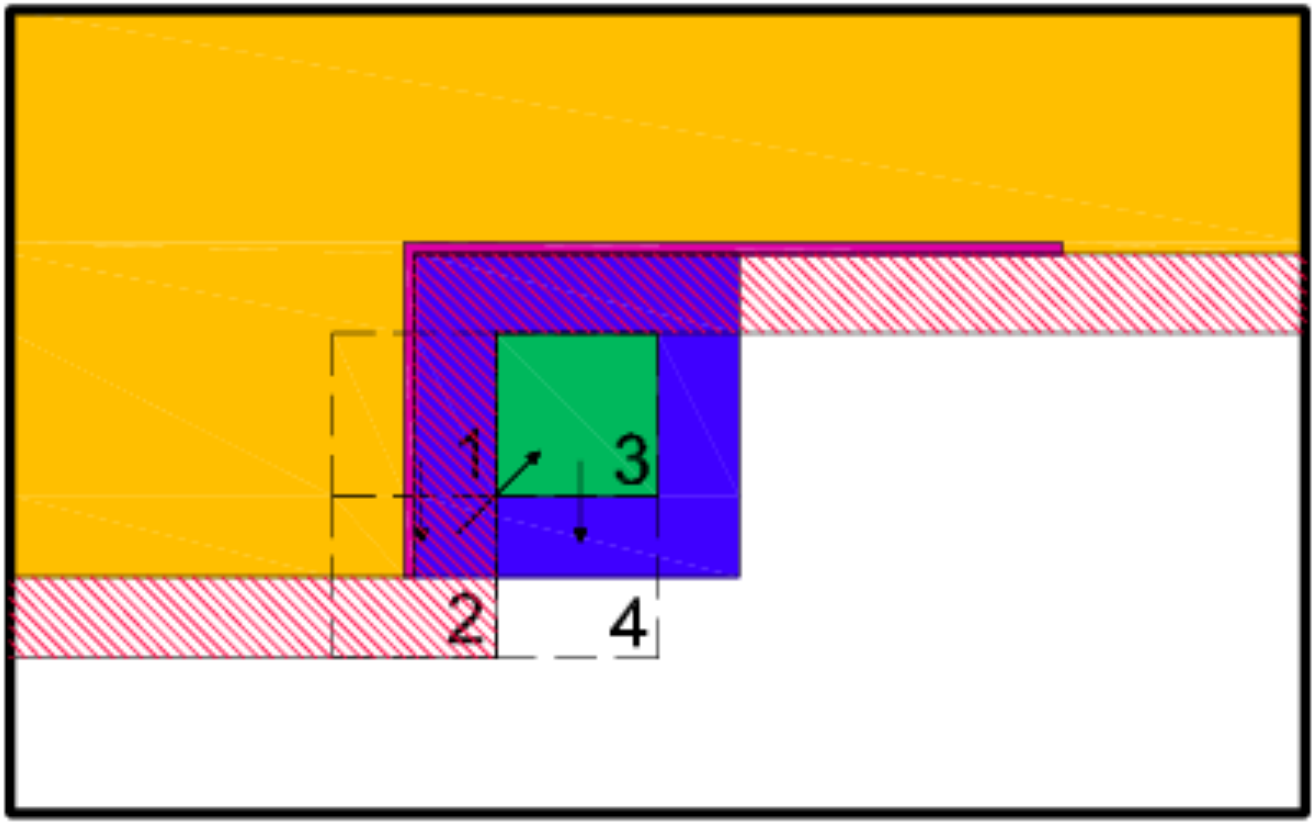


TELECOM ParisTech

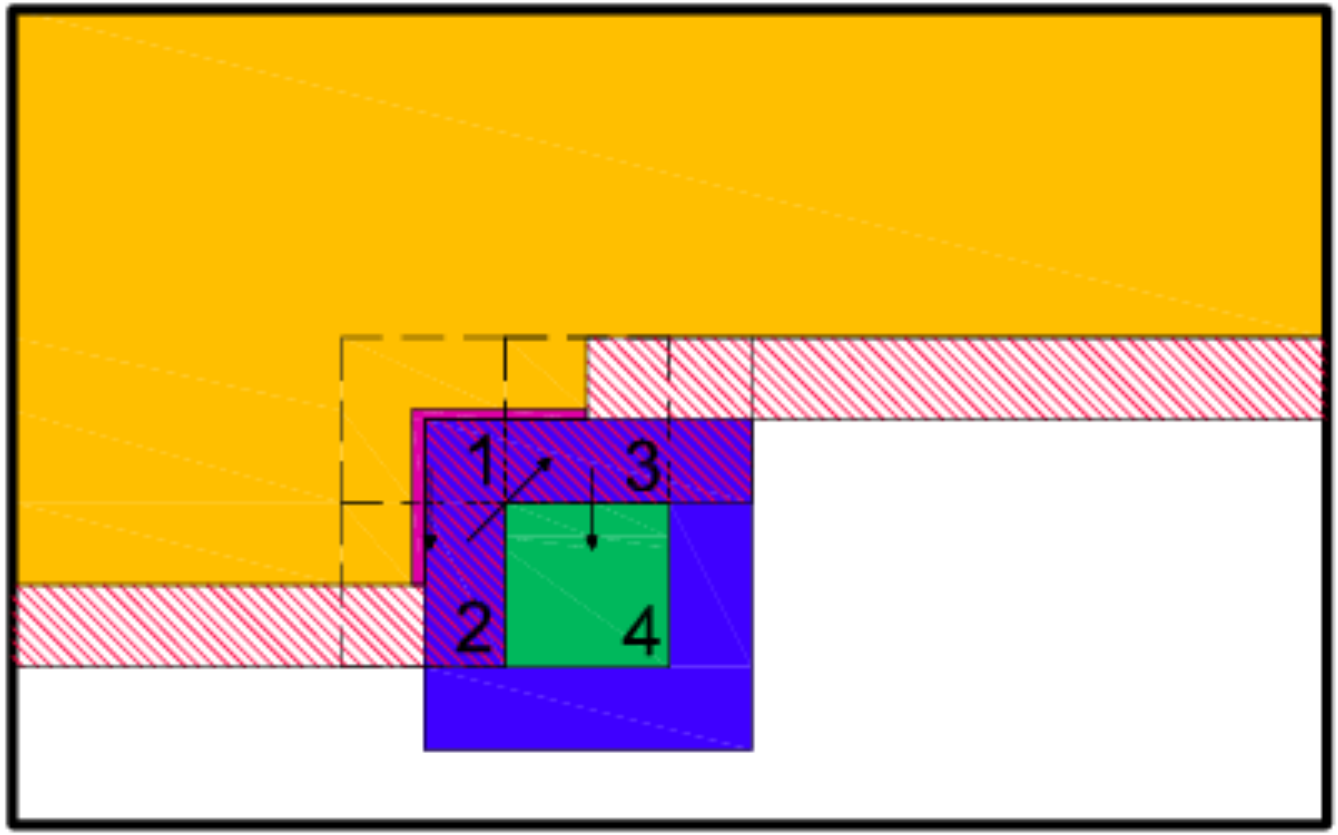
Intra frame prediction and Lapped Transforms



Intra frame prediction and Lapped Transforms



Intra frame prediction and Lapped Transforms



TELECOM ParisTech

Intra frame prediction and Lapped Transforms: analysis/reconstruction

- **Prediction, transform and reconstruction**

Prediction $\mathbf{R_{mn}} = \mathbf{V_{mn}} - \mathbf{Pred_{mn}}$

Direct transform $\mathbf{Y_{mn}} = \mathbf{P} \cdot \mathbf{R_{mn}} \cdot \mathbf{P^T}$

Inverse LT $\widehat{\mathbf{R}}_{\mathbf{mn}} = \mathbf{Q^T} \cdot \mathbf{Y_{mn}} \cdot \mathbf{Q}$

$$\widehat{\mathbf{V}}_{\mathbf{mn}} = \widehat{\mathbf{R}}_{\mathbf{mn}} + \mathbf{Q^T} \cdot \mathbf{P} \cdot \mathbf{Pred_{mn}} \cdot \mathbf{P^T} \cdot \mathbf{Q}$$

TELECOM ParisTech

Mode Selection

- **SAD**

$$SAD = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \left| \mathbf{x}_{nm}(i,j) - \mathbf{Pred}_{nm}(i,j) \right|$$

- **WSAD**

$$WSAD = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \left| W_{ij} \left(\mathbf{v}_{nm}(i,j) - \mathbf{Pred}_{nm}(i,j) \right) \right|$$

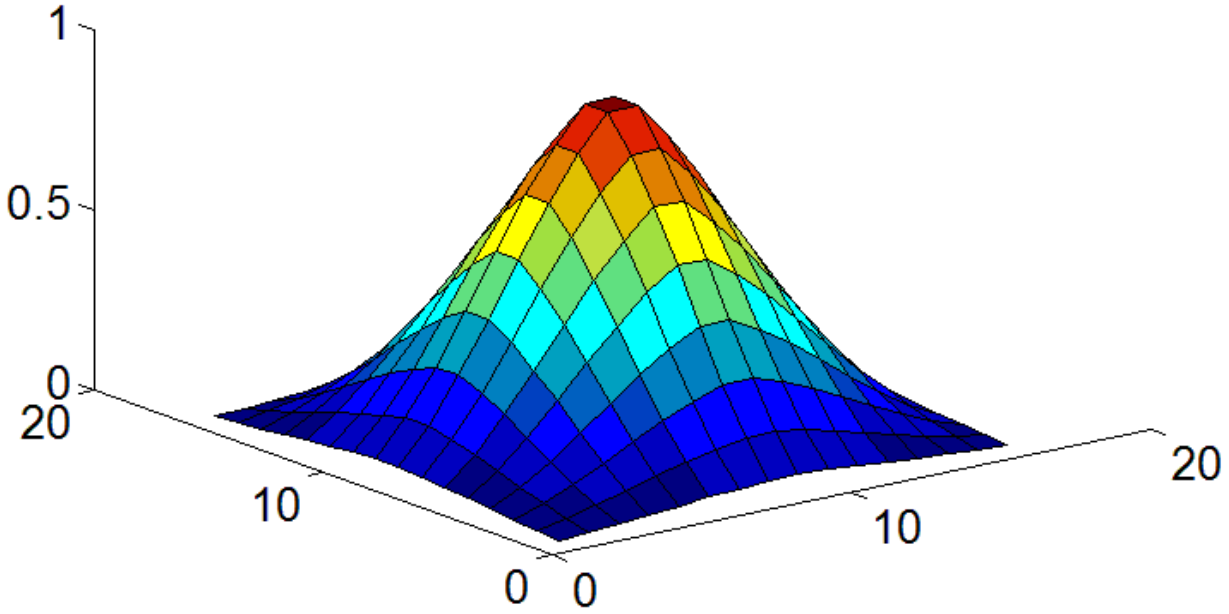
Weighting matrix $\mathbf{W} = \mathbf{Q}^{\mathbf{T}} \cdot \mathbf{P} \cdot \mathbf{Ones} \cdot \mathbf{P}^{\mathbf{T}} \cdot \mathbf{Q}$

- **RDO**
  - spatial reconstruction not fully available
  - distortion measure in the transform domain
  - e_{ij} is the energy of the synthesis filter

$$D = \sum_{ij} e_{ij} \left(Y_{ij} - Y'_{ij} \right)^2$$

TELECOM ParisTech

Weighting Window: example



- GLBT 8x16
- P,Q optimized for maximizing the coding gain of an AR(1) model with ρ=0.95

Overview

- Motivation
- Lapped transforms: background
- Intra-frame prediction
  - Dificulties for LT
  - Proposed Solution
  - Modification in the mode selection
- Experimental Results
- Conclusion and future work

TELECOM ParisTech

Experimental results

- **Implementation has been carried in a JPEG coder**
- **The coding block fixed to 8x8 pixels**
- **Lapped transform used in the tests: GLBT 8x16**
  - Overlap of 2: more blocks benefit from the prediction
  - GLBT: best performance for this overlapping factor
  - Overlapping of 2: prediction block 16x16
  - Only four modes avaible for the prediction using LT

TELECOM ParisTech

Implementation Details

- **For comparison reasons, the traditional intra frame prediction was also implemented in JPEG and the modes of the traditional scheme were encoded in the same way as proposed in H.264/AVC**

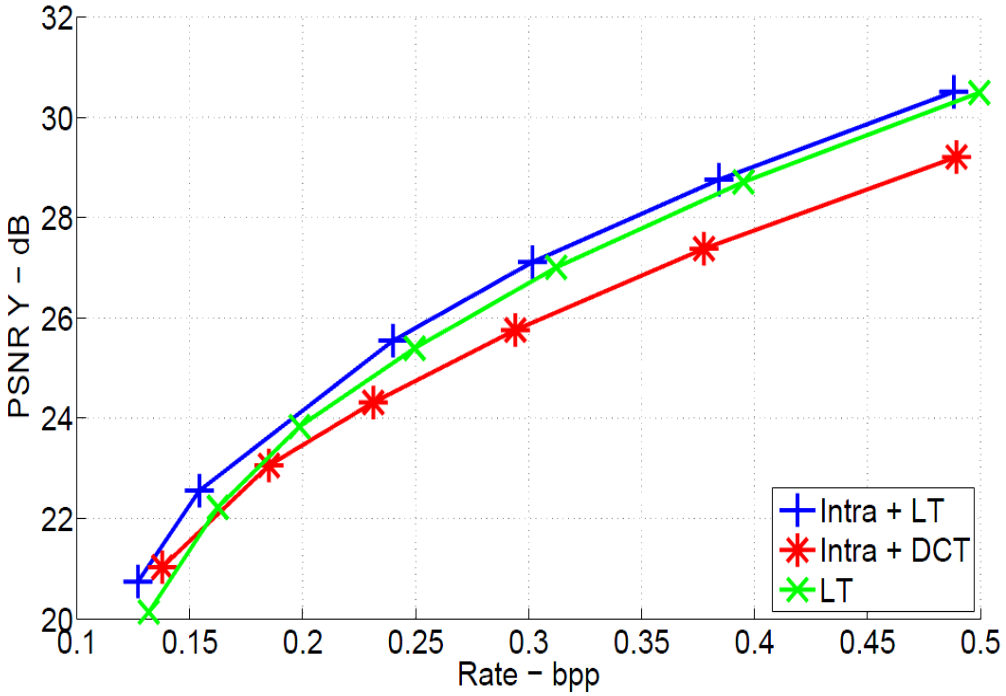
- **Modes of the proposed scheme encoded using 2 bits**
  - Using the mode prediction represented a rate increase (not both neighbors available for the prediction)

- **The direct application (no prediction) of lapped transforms also tested**

TELECOM ParisTech

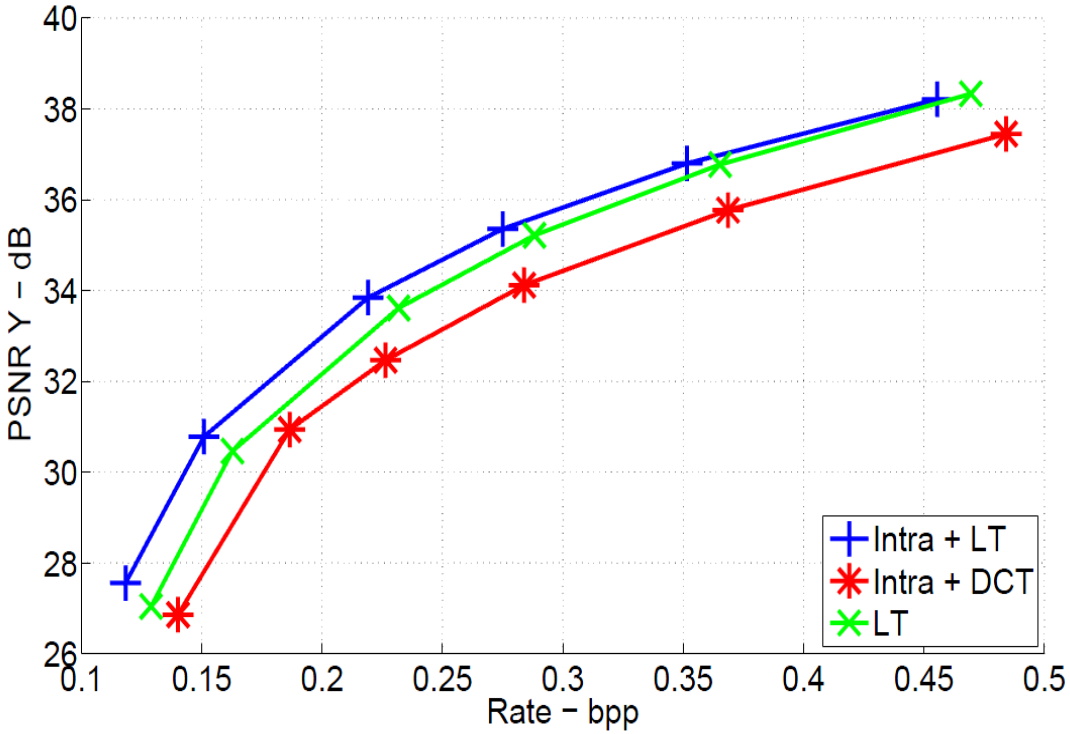
RD Results (1)

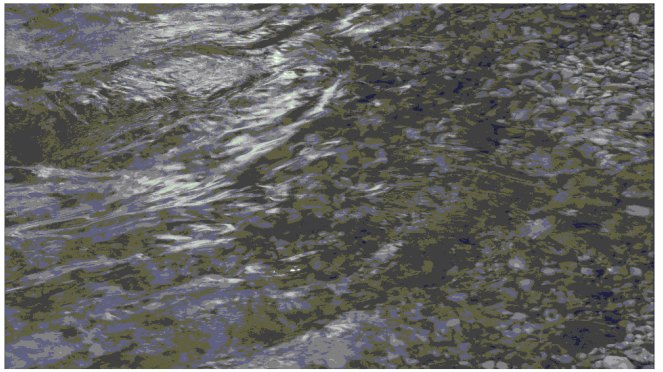
- Barbara (512x512)



RD Results (2)

- Riverbed (1920x1080) – first frame





Results: Bjontegaard metric (low bitrate)

Image (resolution)	G1(dB)	% R1	G2(dB)	% R2
Barbara (512×512)	0.79	-11.21	0.31	-3.98
Goldhill (512×512)	0.51	-8.16	0.32	-4.57
Lena (512×512)	0.40	-4.07	0.35	-3.90
Riverbed (1080p)	1.12	-14.64	0.90	-4.72
Sunflower (1080p)	1.95	-7.65	1.48	-7.02

1. Comparison between the proposed scheme and Intra-prediction with DCT
2. Comparison between the proposed scheme and the direct application of lapped transforms

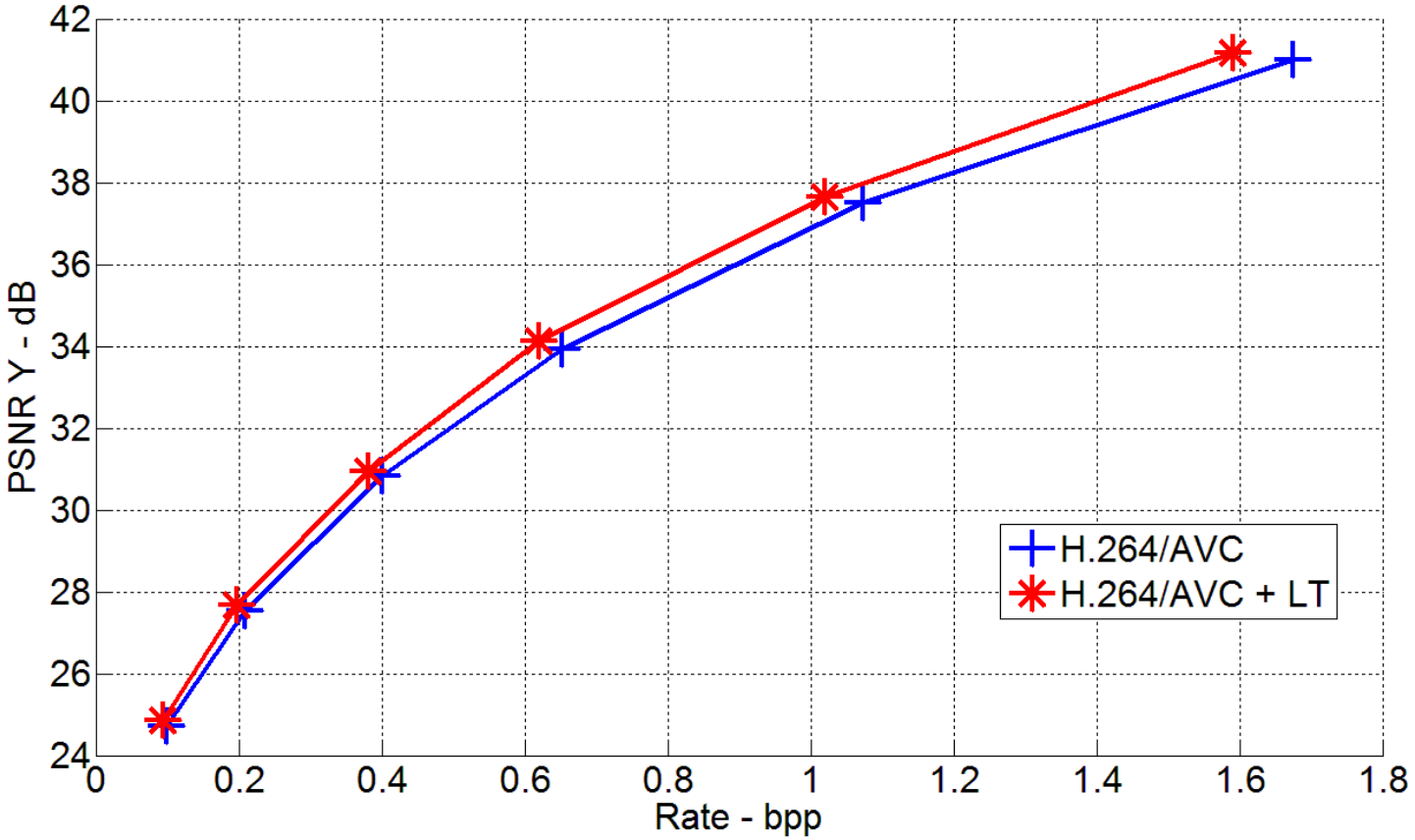
Overview

- Motivation
- Lapped transforms: background
- Intra-frame prediction
  - Difficulties for LT
  - Proposed Solution
  - Modification in the mode selection
- Experimental Results
- Conclusion and future work

Conclusion

- The results show that intra-frame prediction can be adapted to be compatible with lapped transforms
- The proposed scheme outperforms the direct application of lapped transforms and the traditional intra prediction with DCT in all tested images
- For the moment, fewer prediction modes were implemented for 16x16 pixel blocks

Work in progress: RD Resuts(Barbara)



Work in progress

- **Implementation in H.264/AVC available for the same conditions**
  - Competition with all the prediction modes
  - CABAC adapted to the integer DCT

- **Future work**
  - Variable size prediction blocks
  - Test other transforms (with other overlapping factors)
  - A extension to video of this work has been accepted in ICIP2011

Thank you!