# Beyond H.264/AVC Coding Using Texture Analysis and Synthesis

P. Ndjiki-Nya, A. Smolic, and T. Wiegand

Fraunhofer Institut Nachrichtentechnik Heinrich-Hertz-Institut

# Outline

- Motivation
- Video Codec Architecture
  - Texture Analysis
  - Texture Synthesis
  - Video Quality Assessment
- Results
- Conclusions and Future Work

# Motivation

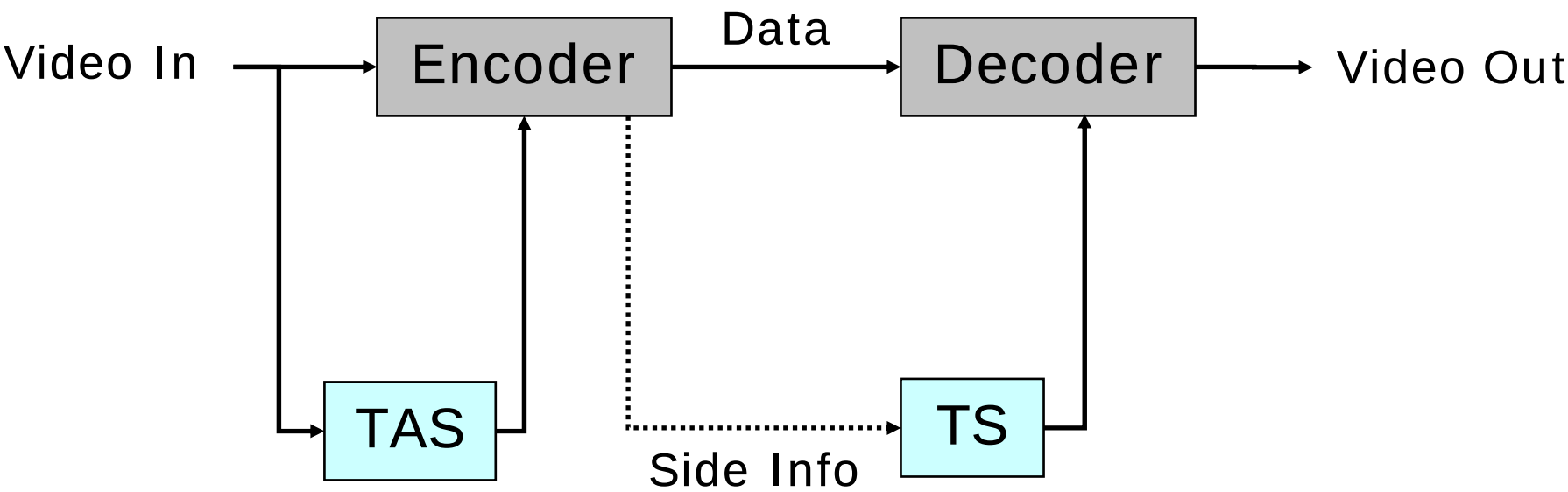
- Textures with large amount of visible details are difficult to code e.g., grass, trees, sand, clouds, water ...

Where is the original texture ?

# Motivation

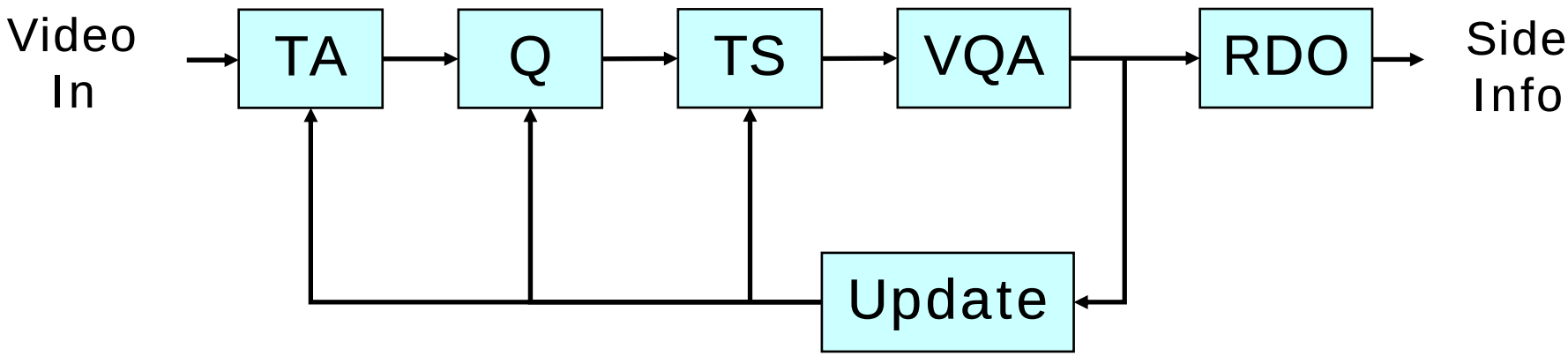
- Exact reproduction of details is irrelevant if
  - Textures are shown with limited spatial accuracy
  - Viewer does not know the original video
- Mean Squared Error (MSE) distortion criterion is not suitable for efficient coding of detail-irrelevant textures
- If information needed for approximate reproduction of detail-irrelevant textures less than bit-rate using MSE (residual error coding, ...) ➔ Bit-rate savings
- Could we use similarity criteria instead of MSE as coding distortion ?

# Video Codec Architecture



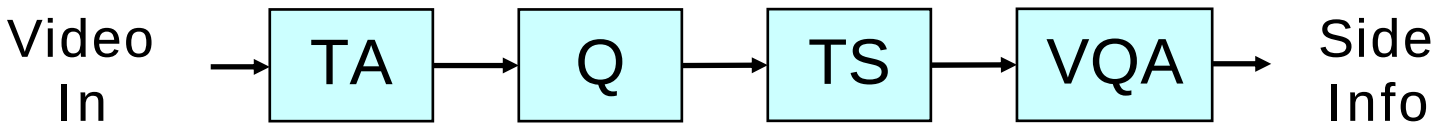

Texture Analysis-Synthesis (TAS) Video Coding

# Texture Analysis-Synthesis Loop




| | |
|---|---|
| Q | Quantizer |
| VQA | Video Quality Assessment |
| RDO | Rate-Distortion Optimization |

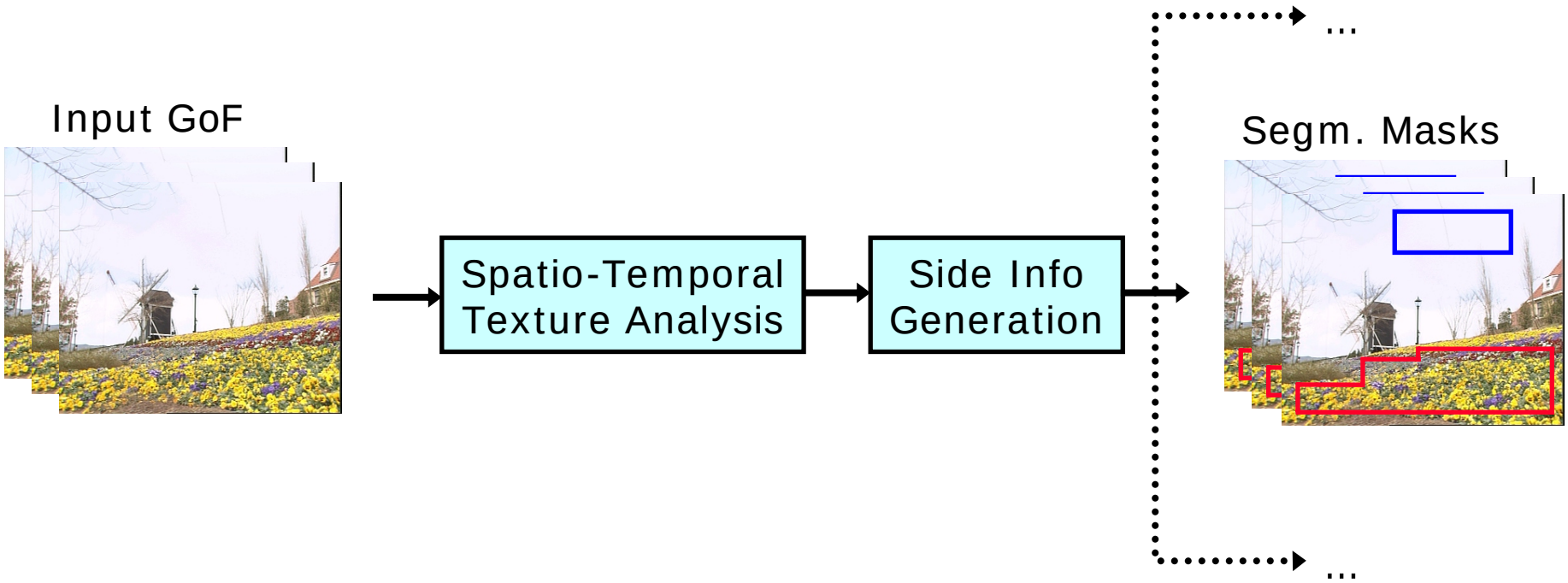
# Texture Analysis-Synthesis




Q Quantizer
VQA Video Quality Assessment

n1

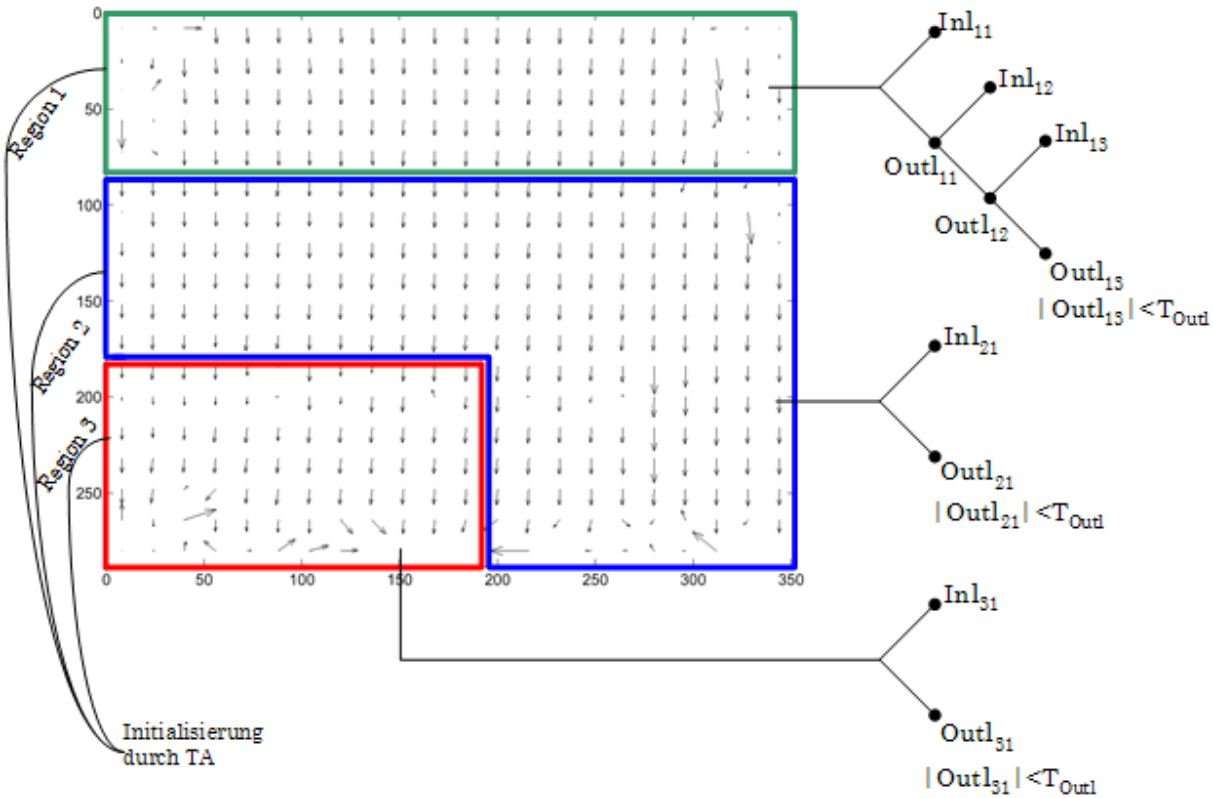
# Texture Analyzer Principle



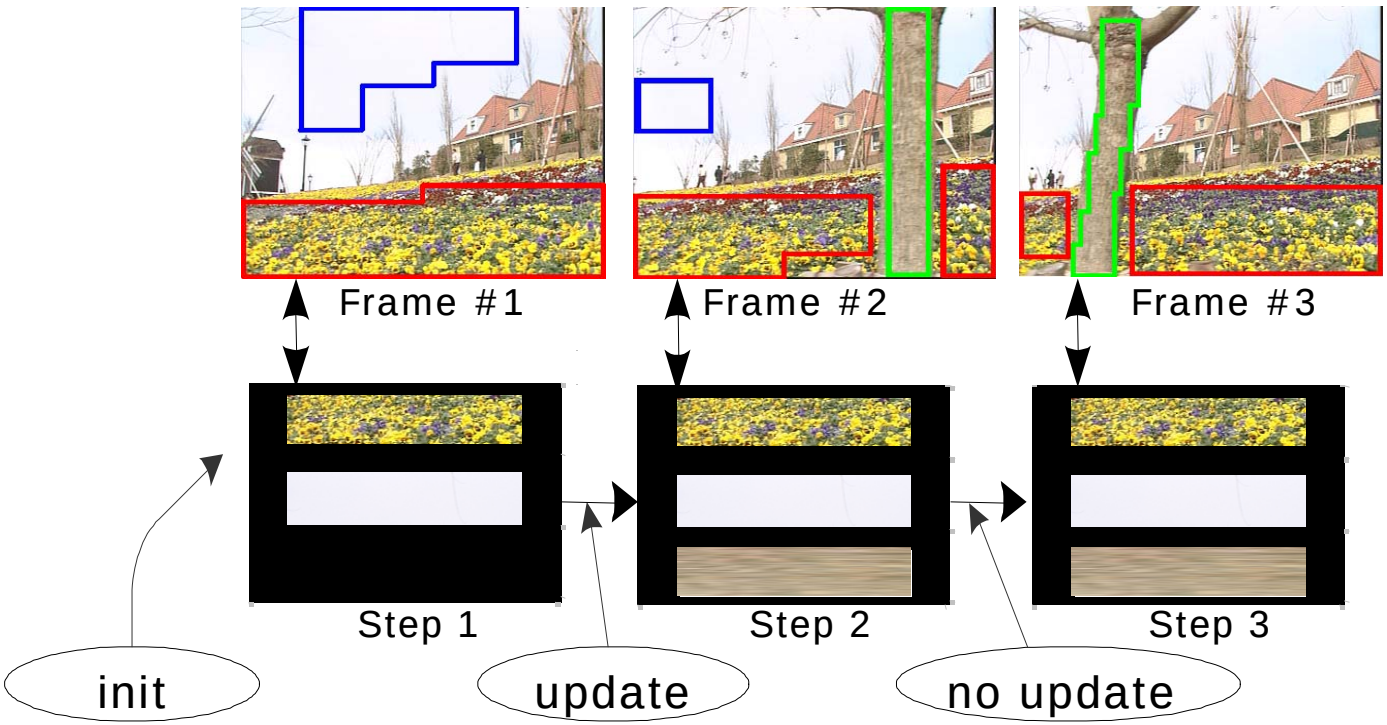
**n1** "...": Additional side info. to generic mask

ndjiki; 29.10.2004

# Spatio-Temporal Analysis (STA)



# STA: Temporal Texture Mapping

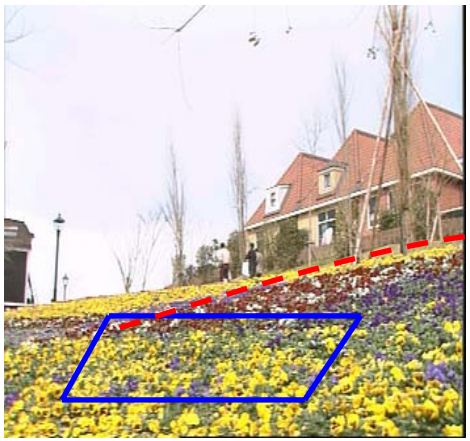


# Side Information Generation

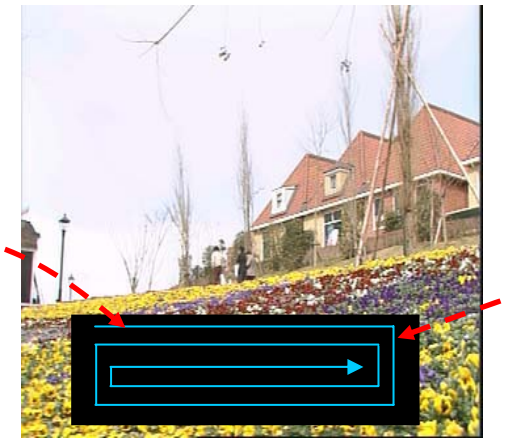
- Side info. per synthesizable texture region :
  - 8 motion parameters (Planar perspective motion model)
  - 1-8 quantization parameter(s)
  - 1 control flag
  - Coarse mask (macroblock-accurate)

# Texture Synthesizer for Rigid Textures

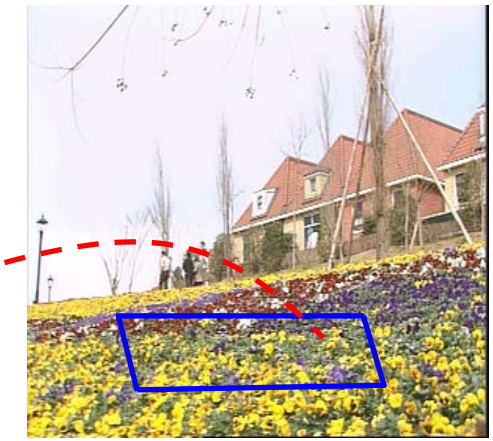
Left reference frame



Current frame



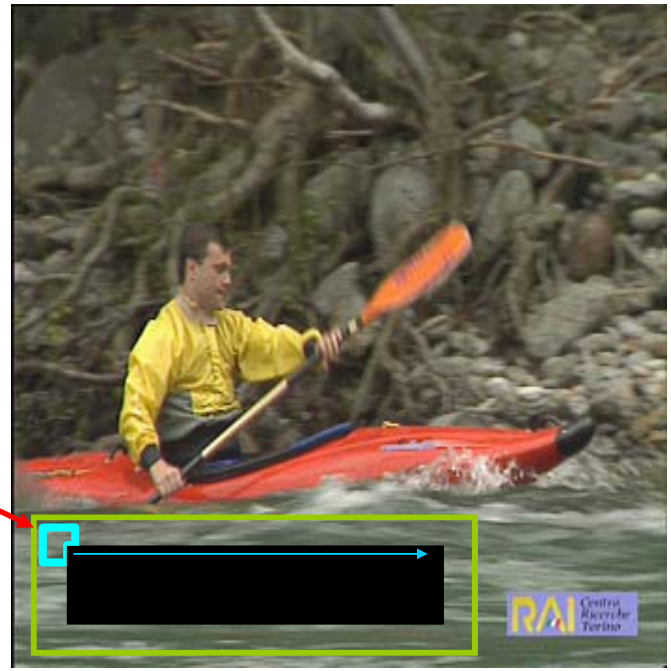
Right reference frame



# Texture Synthesizer for Textures with Local Motion

Left reference frame

Current frame



Match surrounding samples between current picture and warped pictures to find synthesis sample for the current sample

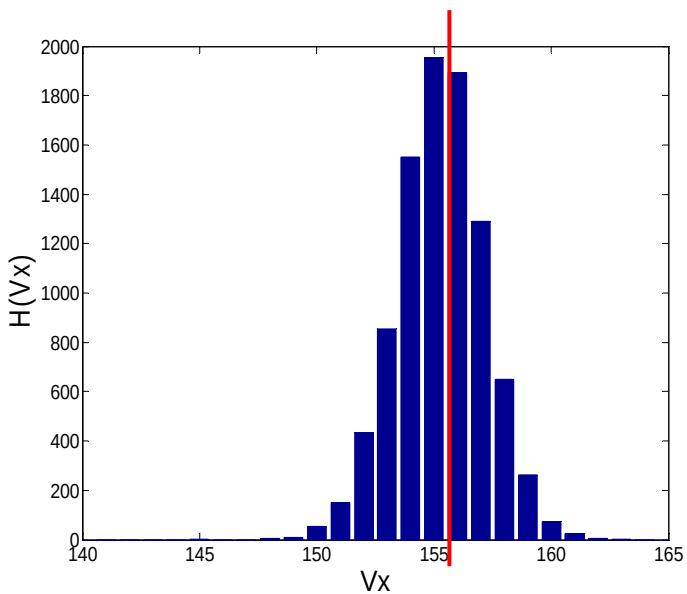
# Video Quality Assessment

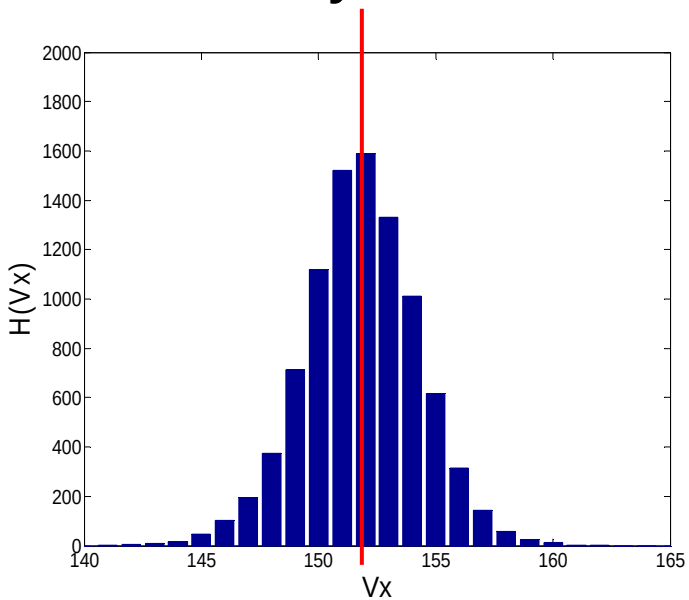
Temporal Artefacts

Spatial Artefacts

# Temporal VQA

## Original - Motion Distribution - Synthetic




## Artefact Existence Criterion

$$w_x \cdot \left\| \vec{H}_{x,orig} - \vec{H}_{x,synth} \right\| + w_y \cdot \left\| \vec{H}_{y,orig} - \vec{H}_{y,synth} \right\| \geq T_{TVQA}$$

# Spatial VQA

## Kirsch‘s Edge Detector

$$K_x = \frac{1}{15} \cdot \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \qquad K_y = \frac{1}{15} \cdot \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$$

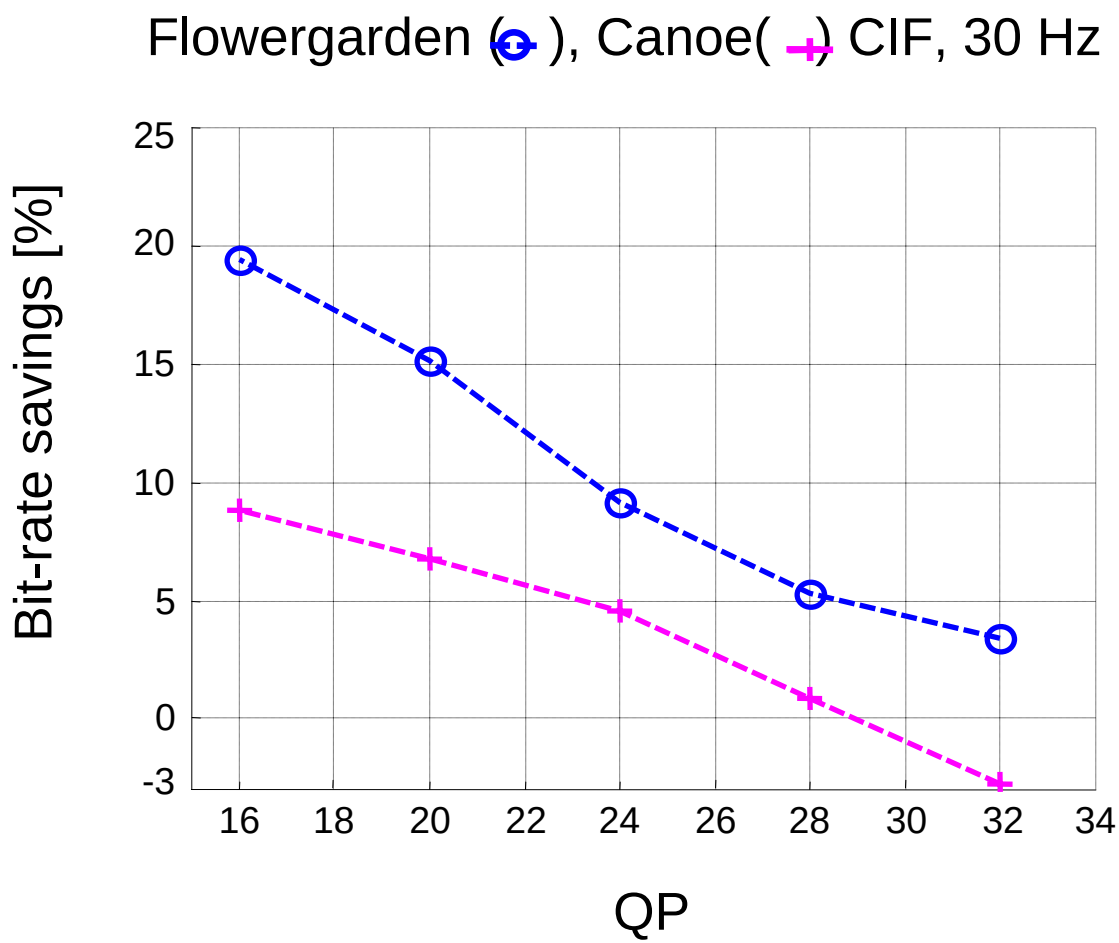
## Artefact Existence Criterion

$$R_{SVQA} = \frac{TP}{FP} \geq T_{SVQA}$$

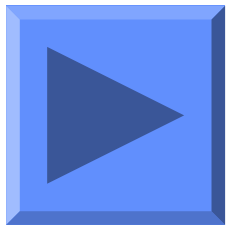
# Results: Test Conditions

- Video Codec: H.264/AVC
- “Flowergarden”, “Canoe”, “Raven”, and “Waterfall” test sequences
- Up to 101 frames each
- 3 synthesis frames between 2 key frames

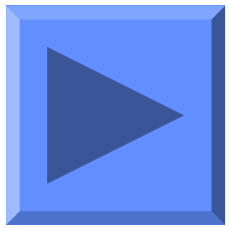
# Bit-rate Reduction vs. Quantizer



# Raven



# Waterfall



# Conclusions and Future Work

- Modular framework for improved H.264/AVC coding
- Analysis : Increase robustness
- Synthesis: Tackle universality
- Loop
  - Rate-distortion optimization
  - Iterative analysis-synthesis process
  - Additional criteria for similarity evaluation (also include MSE)