ANALYSIS/SYNTHESIS CODING OF DYNAMIC TEXTURES BASED ON MOTION DISTRIBUTION STATISTICS

Olena Chubach, Patrick Garus, Mathias Wien, Jens-Rainer Ohm
Institut für Nachrichtentechnik, RWTH Aachen University

Motivation

Highly textured parts: challenging for conventional codec but perceptually irrelevant for humans.

Fig. 1: Examples of dynamic textures content

- Viewer rather perceives semantic equivalence of displayed content than specific details
- Exact positions of the texture patterns are irrelevant for humans
- Therefore textures may be displayed without a pixel fidelity, instead of conventional coding
- This allows to omit encoding prediction residuals and motion vector coding of dynamic textures, leading to substantial reduction of bits to be coded

Motion-based Characterization of Dynamic Textures

- Dynamic textures are represented by a set of first order motion features computed along the space and time dimensions
- Motion vectors from 3 spatial and 2 temporal neighboring positions are considered for motion co-occurrence matrix (MCM), providing efficient representation of motion distribution in dynamic textures

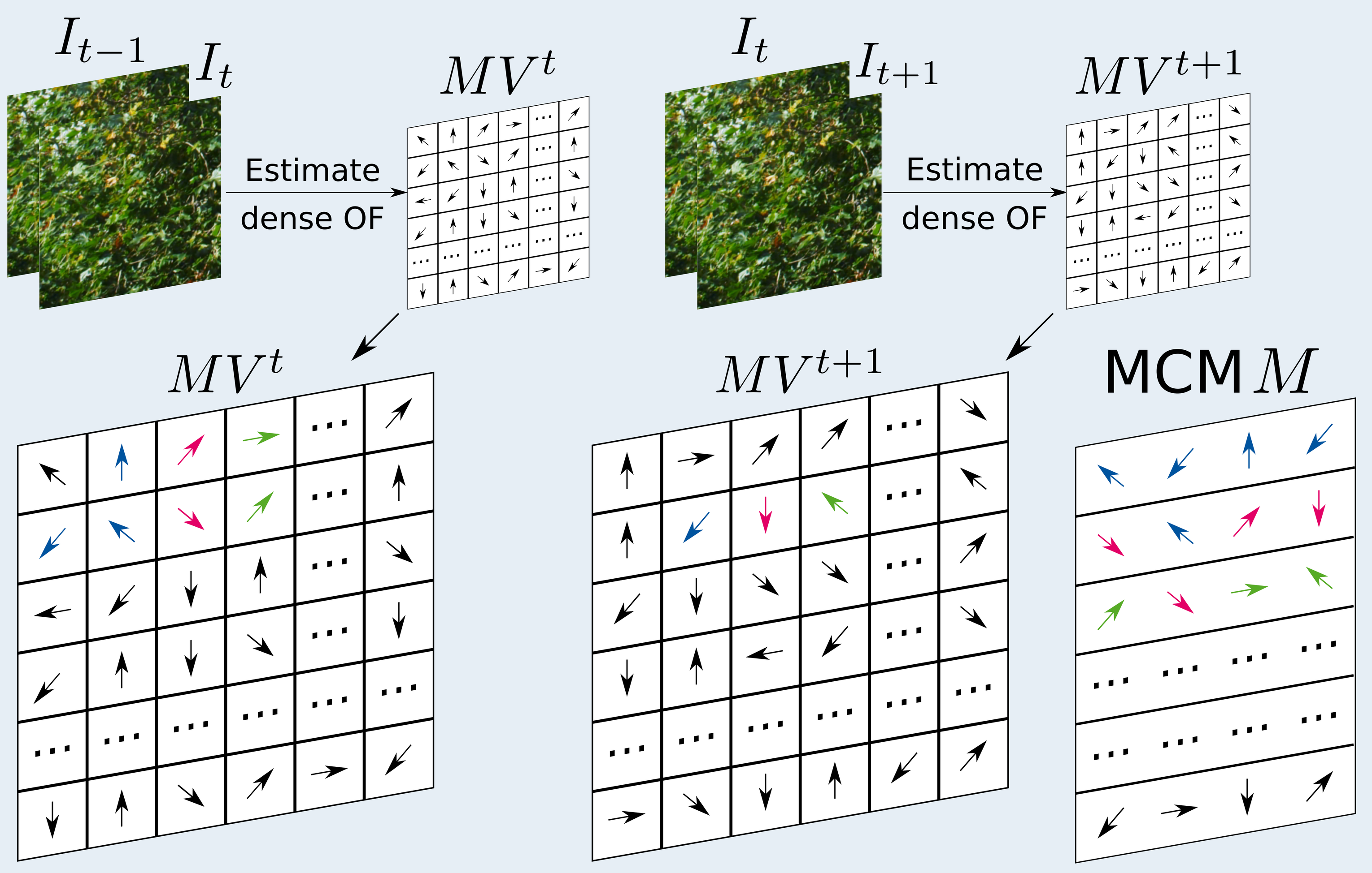

Fig. 2: Procedure of MCM computation

Modified Coding Structure

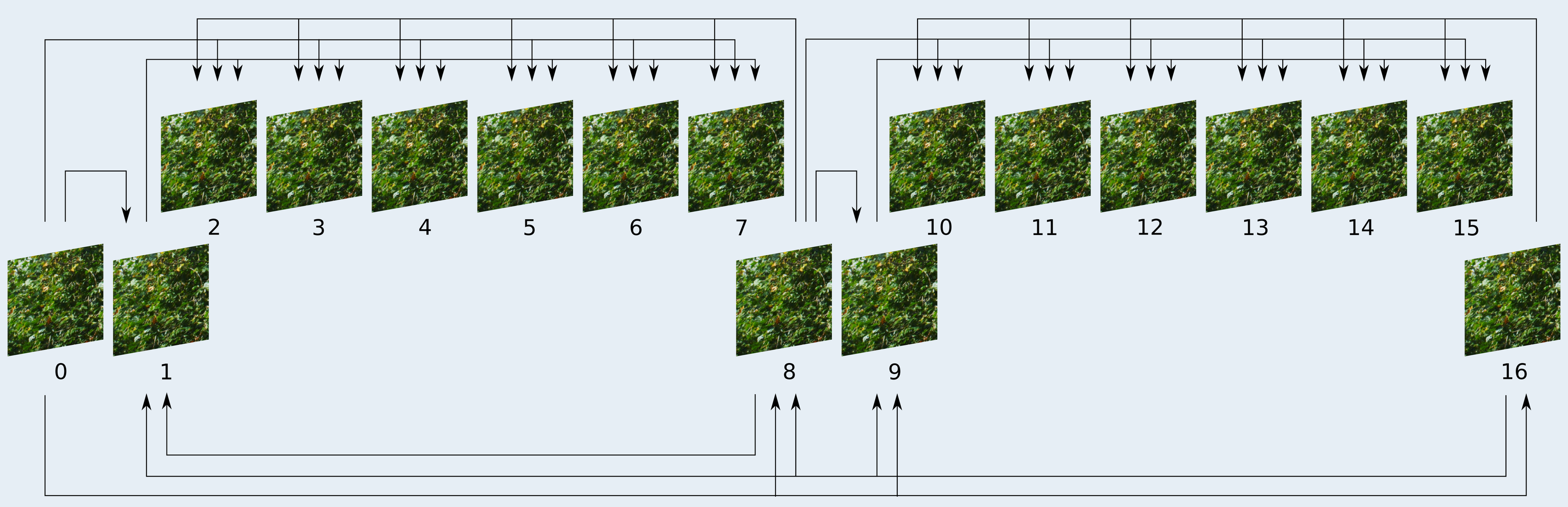

Fig. 3: Modified coding structure in case of sGOP size 16

- Frames 0, 1, 8, 9 and 16 are reference frames and reconstructed first
- Remaining 12 frames are skipped during encoding/decoding and will be synthesized

Proposed Scheme

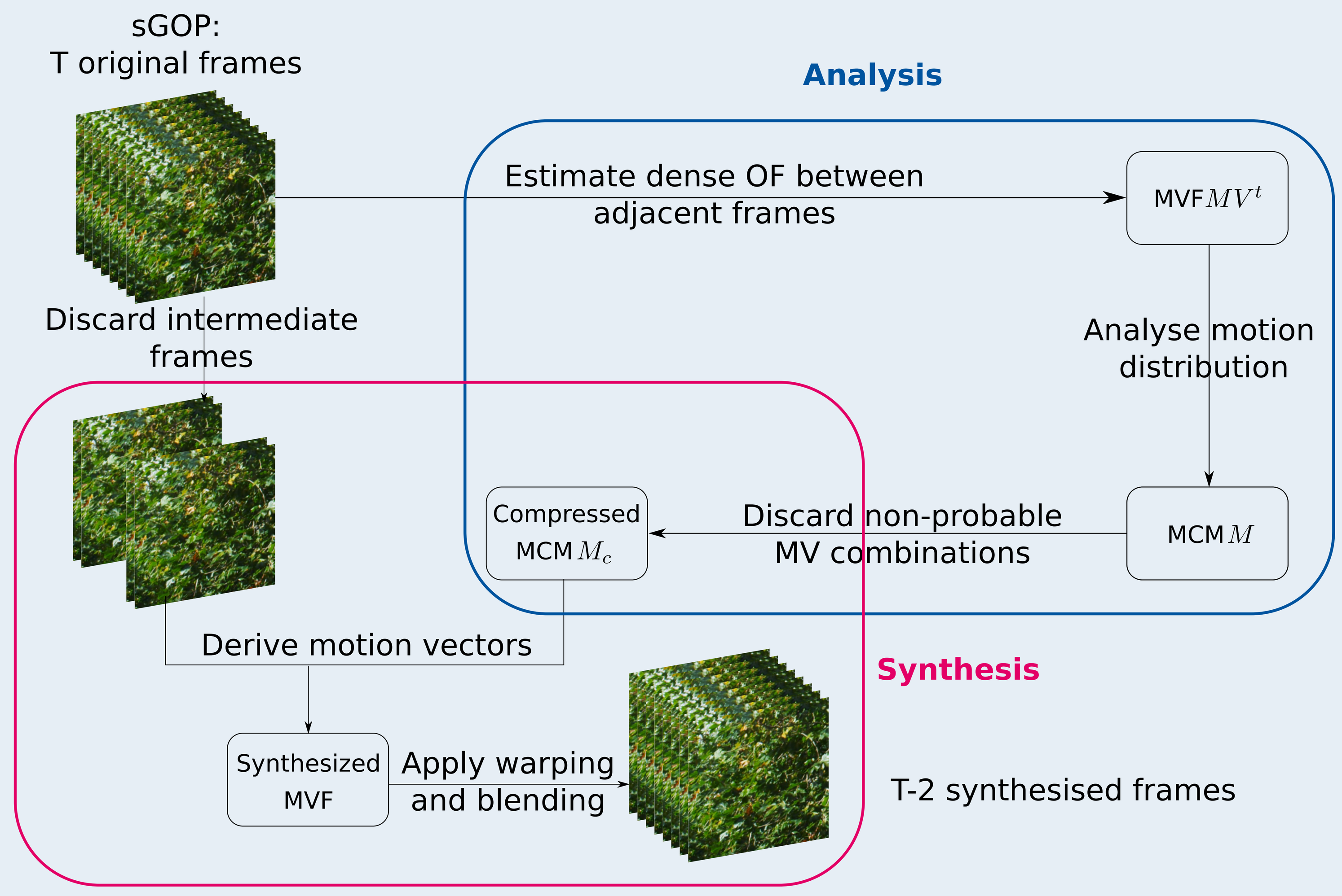

Fig. 4: Proposed scheme of DT analysis/synthesis based on motion distribution statistics

- Initial MVF is estimated between adjacent reference frames
- Compressed MCM is signaled to the decoder side for synthesis
- Synthetic MVFs are predicted based on initial MVF and compressed MCM and utilized for generating intermediate frames
- Synthesis procedure is performed twice: in forward direction using frames from the past and in backward direction using frames from the future
- Corresponding synthesized frames $\hat{I}_t^f$ and $\hat{I}_t^b$ are then blended:

$$\hat{I}_t = (1-\lambda(t))\hat{I}_t^f + \lambda(t)\hat{I}_t^b$$

Results

- Proposed method tested on sequences from HomTex database, containing water, leaves and smoke
- Test sequences: 256x256 pixels in width and height; 250 frames; 25 or 60 fps
- Encoded with HEVC Test Model (HM-16.6) with modified RA config., QP=22
- sGOP size considered - 16 frames
- 50% of the most probable MV combinations are kept for every sGOP
- MCMs and and MV interval ranges are compressed by arithmetic coding

Sequence	HEVC rate, CTC RA, [kb]	Synth, rate, modif. RA, [kb]	Rate reduction, %
BricksBushes Static-Bushes1	1744.5	$766.1 + \mathbf{5.7}$	-55.7
BricksBushes Static-Bushes2	1579.1	$717.3 + \mathbf{2.9}$	-55.4
LampLeaves Bushes1	1578.9	$728.3 + \mathbf{11.9}$	-53.1
LampLeaves Bushes2	1146.4	$507.4 + \mathbf{54.1}$	-51.0
LampLeaves Bushes3	1294.6	$545.8 + \mathbf{135.5}$	-47.4
LampLeaves BushesBackground	552.4	$281.2 + \mathbf{7.5}$	-47.8
PetibatoCropped	735.1	$391.5 + \mathbf{153.3}$	-25.9
TreeWills-Cropped	970.6	$584.3 + \mathbf{0.44}$	-39.8

Table 1: Rate comparison: 2nd column - HEVC CTC RA configuration; 3rd column - reference frames encoded with modified RA configuration (1st term); parameters required for synthesis, compressed by arithmetic coding (2nd term); 4th column - rate reduction, under the assumption of acceptable quality drop

chubach@ient.rwth-aachen.de
www.ient.rwth-aachen.de
Institut für Nachrichtentechnik, Melatener Str. 23, 52074 Aachen

ICIP 2017