

JVET-J0038

Signal adaptive diffusion filters for video coding

Jonathan Pfaff, Jennifer Rasch, Michael Schäfer, Martin Winken, Anastasia Henkel, Mischa Siekmann, Heiko Schwarz, Detlev Marpe, Thomas Wiegand

Basic Approach: Smoothing of the intra- resp. inter-prediction signal *pred* :

- Scale Space of filters: Family of filtered prediction signals
$$F_t(pred), t \geq 0,$$
with $F_0(pred) = pred$. Here, $F_t(pred)$ smoother, the larger t .
- $F_t(pred)$ solves the discretization of
$$\frac{\partial}{\partial t} F_t(pred) = \text{div}(c \nabla F_t(pred))$$
with an appropriate boundary condition.
- Two variants for c :
 - Linear diffusion: $c=1$. Uniform smoothing.
 - Non-linear diffusion: c is a function that depends on the sample position and *pred* such that smoothing is performed region-wise and edges are kept.

Implementation of the diffusion filter:

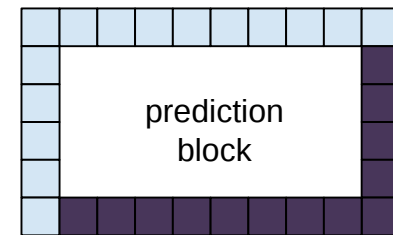
- Replace the filter parameter t by an integer n . The larger n , the stronger the filter.
- Perform n correlations of $pred$ with a mask h and boundary condition:

$$F_n(pred) = \underbrace{h * \dots * h}_{n \text{ times}} * pred.$$

- Linear diffusion:

$$h = \frac{1}{4} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

rec. border samples



reflected samples

- Non-linear diffusion: Mask depends on $pred$ and sample position:

$$h(x, y) = \begin{bmatrix} a_{-1,-1}(x, y; pred) & a_{-1,0}(x, y; pred) & a_{-1,1}(x, y; pred) \\ a_{0,-1}(x, y; pred) & a_{0,0}(x, y; pred) & a_{0,1}(x, y; pred) \\ a_{1,-1}(x, y; pred) & a_{1,0}(x, y; pred) & a_{1,1}(x, y; pred) \end{bmatrix}$$

Diffusion configuration:

- Up to two values for n , depending on linear/nonlinear and on inter/intra.
- Up to five diffusion configurations depending on blocksize and temporal layer:
 1. No diffusion
 2. Up to two configurations for linear diffusion
 3. Up to two configurations for non-linear diffusion.
- Only no diffusion for skip, small blocks or intra with DC- or Planar mode.

Decoder algorithmic description

- Determine diffusion filter configuration:
 - For blocks in merge mode that are not skip, inherit diffusion filter configuration from merging candidate.
 - For blocks which are not in merge mode and on which diffusion is allowed, parse a flag indicating if diffusion is to be applied.
 - If diffusion is to be applied:
 1. Parse if linear- or non-linear diffusion is to be applied.
 2. Parse the parameter n .
- If diffusion is to be applied:
 1. Compute the filtered prediction signal $F_n(\text{pred})$.
 2. Replace prediction signal pred by $F_n(\text{pred})$.

Encoder algorithmic description

- For merge-mode: For each merging-candidate in search, apply the same diffusion configuration as merging candidate to prediction signal.
- For non-merge inter prediction:
 1. Perform conventional motion estimation to get initial prediction signal.
 2. Perform RD-search over all applicable diffusion configurations using initial prediction signal.
- For intra prediction: Perform fast search and full RD-search only over a subset of all possible combinations of diffusion and intra mode.

Experimental Results:

	Random Access Main10 (High Tier, QP values 22, 27, 32, 37)			Random Access Main10 (Main Tier, QP values 27, 32, 37, 42)		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
JVET A1	-1.50%	-2.92%	-1.79%	-0.97%	-2.77%	-1.61%
JVET A2	-1.50%	-1.86%	-1.35%	-0.78%	-0.93%	-0.53%
JVET B	-0.97%	-0.91%	-1.72%	-0.52%	-1.06%	-1.38%
CfP SDR-A	-1.13%	-1.49%	-1.09%	-0.63%	-0.76%	-0.45%
CfP SDR-B	-1.15%	-1.22%	-1.77%	-0.57%	-1.05%	-1.02%
CfP HDR-A	-0.83%	-0.96%	-1.30%	-0.76%	-0.88%	-1.07%
CfP HDR-B	-0.52%	-0.81%	-0.36%	-0.30%	-0.53%	-0.34%
Overall	-1.09%	-1.45%	-1.34%	-0.65%	-1.14%	-0.91%
Enc Time [%]	111%			111%		
Dec Time [%]	102%			101%		

Anchor: HM+QT+BTS , JVET 0035

Summary

- We propose the technique of diffusion filters to smooth the conventional prediction signal in a video codec.
- Several types of diffusion filtering are possible. The type of diffusion is either merged or signalled.
- For a Random Access scenario, coding gains of 1.09% resp. 0.65% Y BD bit rate reduction for high resp. low bit rates corresponding to encoder run time of 111% are observed.
- The decoder runtime overhead of the proposed diffusion filters is negligible.