

PVQ Applied outside of Daala (IETF 97 Draft)

Yushin Cho
Mozilla Corporation
November, 2016

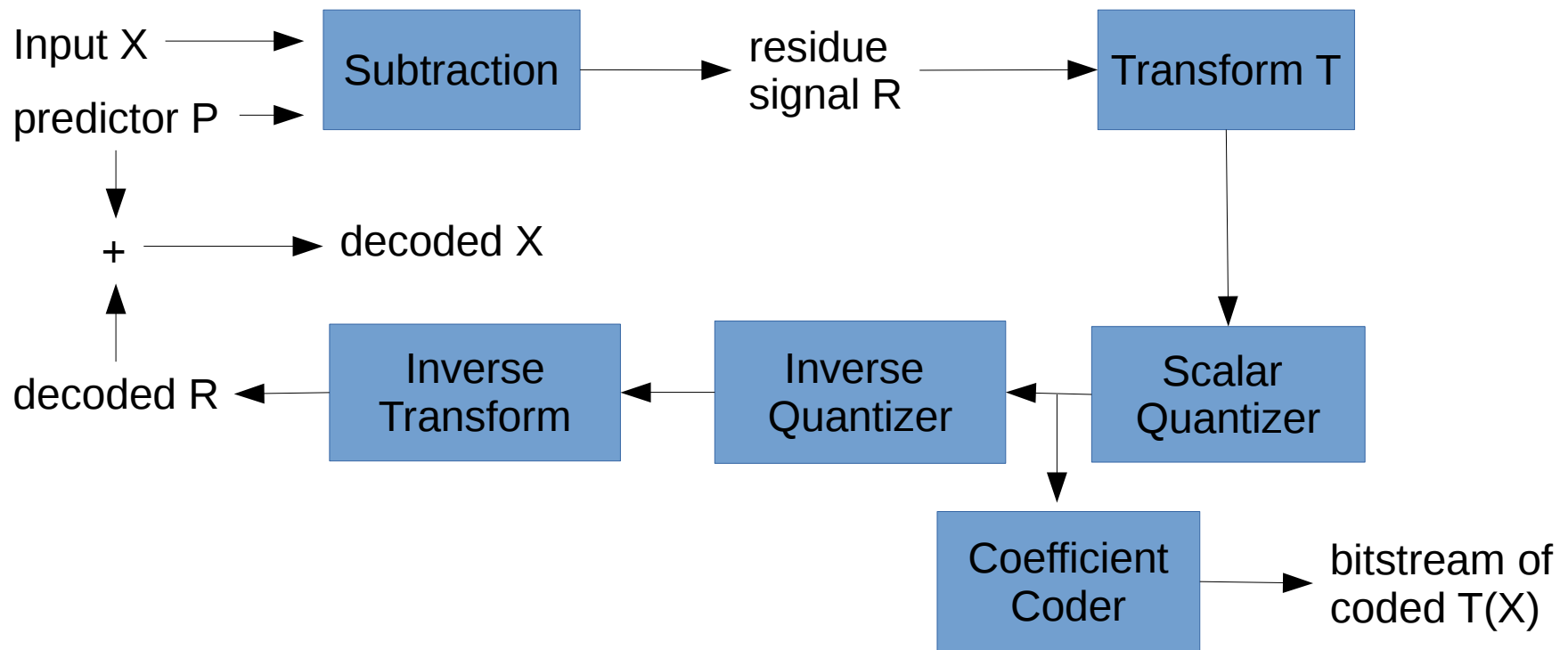
Introduction

- Perceptual Vector Quantization (PVQ)
 - Proposed as a quantization and coefficient coding tool for NETVC
 - Originally developed for the Daala video codec
 - Does a gain-shape coding of transform coefficients
- The most distinguishing idea of PVQ is the way it references a predictor.
 - PVQ does not subtract the predictor from the input to produce a residual

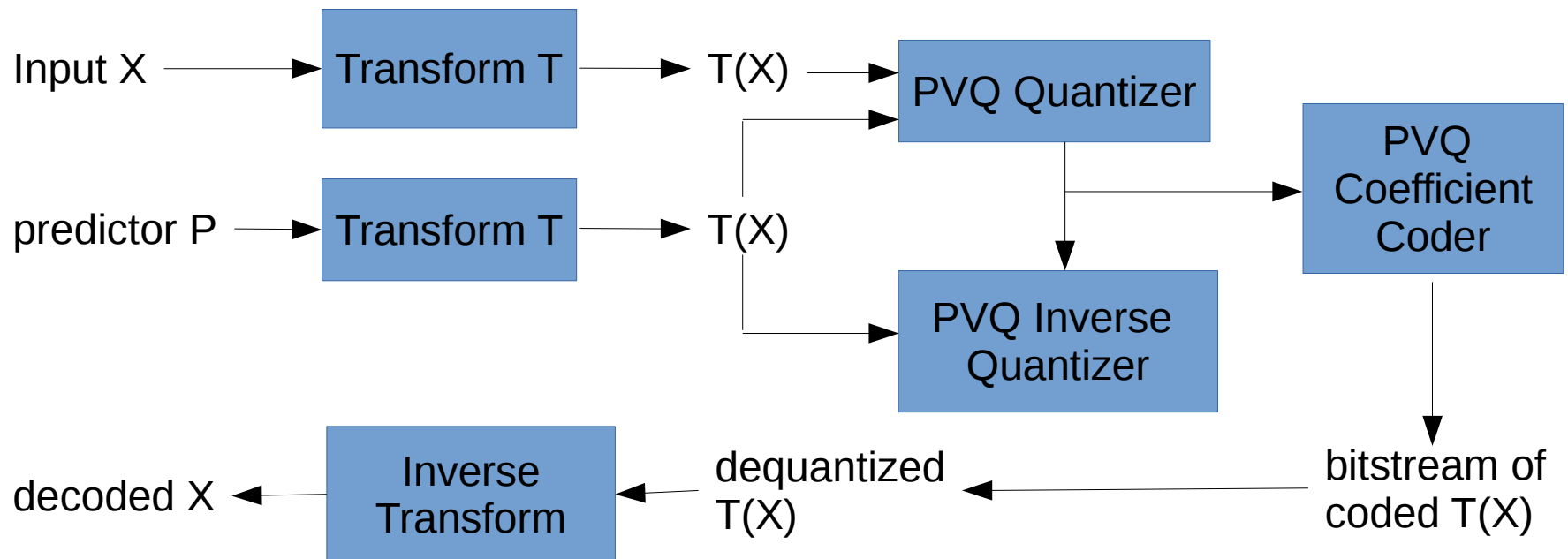
Integrating PVQ into AV1

- Introduction of a transformed predictors both in encoder and decoder
 - Because PVQ references the predictor in the transform domain, instead of using a pixel-domain residual as in traditional scalar quantization
- Activity masking, the major benefit of PVQ, is not enabled yet
 - Encoding RDO is solely based on PSNR

Traditional Architecture



AV1 with PVQ



Coding Gain Change

Metric	AV1 --> AV1 with PVQ
PSNR Y	-0.17
PSNR-HVS	0.27
SSIM	0.93
MS-SSIM	0.14
CIEDE2000	-0.28

- For the IETF test sequence set, "objective-1-fast".
- IETF and AOM for high latency encoding options are used.

Speed

- Increase in total encoding time due to PVQ's search for best codepoint
 - PVQ's time complexity is close to $O(n*n)$ for n coefficients, while scalar quantization has $O(n)$
- Compared to Daala, the search space for a RDO decision in AV1 is far larger
 - For the 1st frame of grandma_qcif (176x144) in intra frame mode, Daala calls PVQ 3843 times, while AV1 calls 632,520 times, that is $\sim 165x$.

Speed - Contd.

Speed Level	# of calls to AV1 quantizer	# of calls to PVQ quantize	# of calls to PVQ search inside a band
5	28,028	26,786	365,913
4	57,445	56,980	472,222
3	505,039	564,724	3,680,366
2	505,039	564,724	3,680,366
1	535,100	580,566	3,990,327
0	589,931	632,520	4,109,113

Note: Above results are from the encoding of 1st frame as a intra frame mode.

Issues arising in Integration

- PVQ does RDO during quantization
 - While most traditional scalar quantization as used in AV1 does not
- Scan orders for AV1's hybrid transforms
 - AV1 uses both DCT and ADST, which required PVQ to add corresponding bandwise scan orders.

Future Work

- Enable activity masking
- Enable high bit depth (10/12 bits)
- RDO with more model-driven decision making

References

- [I-D.valin-netvc-pvq] Valin, J., "Pyramid Vector Quantization for Video Coding", draft-valin-netvc-pvq-00, June 2015.
 - <https://tools.ietf.org/html/draft-valin-netvc-pvq-00>