ADAPTIVE POST-FILTERING BASED ON LOCAL BINARY PATTERNS

Ying Liu[1], Dake He[2], Paul Fieguth[1]

[1]Department of Systems Design Engineering, University of Waterloo, Waterloo, Canada

[2]Compression Technology, Research In Motion, Waterloo, Canada

Motivation

This work seeks to enhance the performance of post-filtering (Fig. 1) for the next-generation coding standard HEVC. We intend to develop an effective Wiener-based filtering method to deal with the quantization noise for reconstructed video frames which possess complicated textures (such as in Fig. 2). In addressing this goal, how to achieve an efficient segmentation for designing the adaptive filters is the key problem.

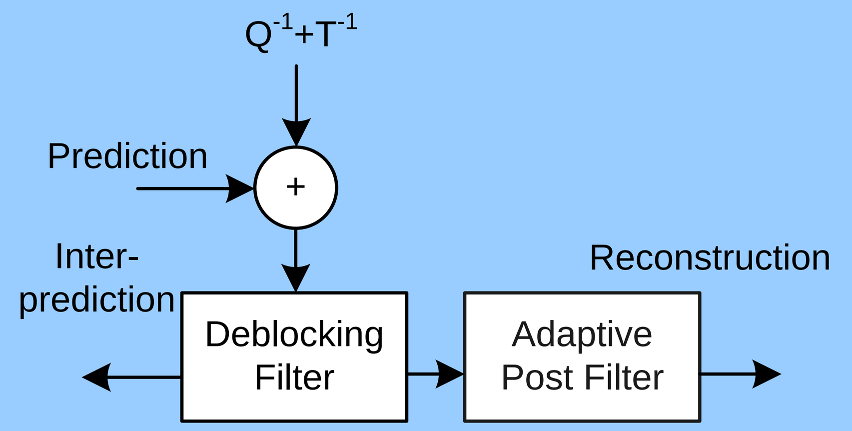

Fig. 1 A coding scheme of adaptive post-filtering

Fig. 2 An intercepted frame with complicated texture.

Related Methods

Most proposed adaptive post-filters are based on Wiener filtering which provides an optimal linear filter to minimize the Mean Square Error (MSE)

$$E=\sum_x\sum_y\left[S(x,y)-\sum_i\sum_j h(i,j)R(x+i,y+j)\right]^2$$

In approaches such as [1], instead of using a single filter h, they adopt multiple Wiener filters { $h^m, m\in\{0,1,\cdots,\mathrm{M}\}$ } for different segments of a reconstructed frame R, with m chosen as

$$m=\min\left(\lfloor v(x,y)+0.5\rfloor,15\right)$$

where $v(x,y)$ corresponds to a local variation measurement, based on $(2K+1)\times(2K+1)$ region centered on (x, y).

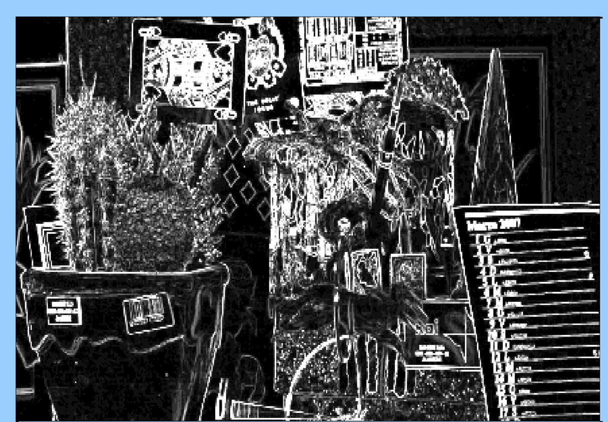

Fig. 3 A segmentation of Fig. 1 with local variation measurement [1].

Limitation:

- A segmentation is determined by variation information. Other information, such as texture features, is ignored.

For example, the adaptive filters of Fig.4 are generated based on [1] and are quite similar in shape.

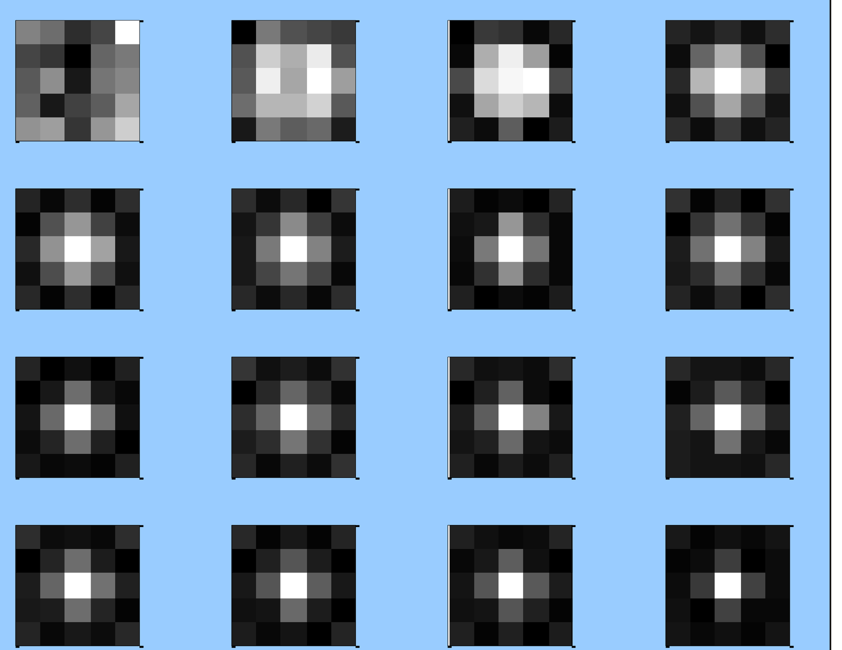

Fig. 4 Adaptive post-filters (5x5-tap) learned from Fig. 1 with method [1].

Proposed Method

The proposed adaptive post-filtering:

- Partition a reconstructed frame into non-overlapping segments based on LBP with variation information;
- For each segment, a post-filter is a 2D Wiener filter to reduce quantization noise.

First, as in [1], given a reconstructed frame *R*, calculate its local variation measurement.

Second, classify *R* into a M-label field L^M by thresholding the range of variation:

$$L^{\mathrm{M}}(x,y)=\begin{cases}0, & \text{if } \ell(x,y)<t_\beta\\ 1, & \text{if } t_\beta<\ell(x,y)<t_\alpha\\ ?, & \text{if } \ell(x,y)>t_\alpha\end{cases}$$

Then, for the reconstruction regions with undetermined label states, continue classifying the region into N-label states according to its LBP index

$$L^{\mathrm{N}}(x,y)=LBP_{\mathrm{P}}^{\mathrm{N}}(x,y)$$

LBPs can be set from one of two neighbour sizes: P=4 and P=8. For example, for P=8:

$$LBP_{\mathrm{P}}^{\mathrm{N}}=\begin{cases}LBP_{\mathrm{P=8}}^{riu}, & \text{if } \dfrac{size(R\mid L^{\mathrm{M}}=?)}{size(R\mid L^{\mathrm{M}}=0\cup L^{\mathrm{M}}=1)}\geq thr\\ merge(LBP_{\mathrm{P=8}}^{riu}), & otherwise\end{cases}$$

where, thr is a ratio threshold, $merge(\cdot)$ is a merging operator, and both of them decided by experiments.

Finally, a total of (M+N) filters can be learned by minimizing two set of MSE energy functions sequentially:

$$E^m=\sum_{(x,y)\in\Omega^m}\left[S_{x,y}^m-\sum_{i=-\mathrm{K}}^{\mathrm{K}}\sum_{j=-\mathrm{K}}^{\mathrm{K}}h_{i,j}^m R_{x+i,y+j}^m\right]^2$$

$$E^n=\sum_{(x,y)\in\Omega^n}\left[S_{x,y}^n-\sum_{i=-\mathrm{K}}^{\mathrm{K}}\sum_{j=-\mathrm{K}}^{\mathrm{K}}h_{i,j}^n R_{x+i,y+j}^n\right]^2$$

where $\Omega^n=\{(x,y)\mid L^{\mathrm{N}}(x,y)=n\}$ and

$\Omega^m=\{(x,y)\mid L^{\mathrm{M}}(x,y)=m\}$

In the encoder side, filter information needs to be generated for a reconstructed frame based on the pattern information (Fig. 5). On the decoder side, the filter coefficients can be obtained from the decoded side information, and then applied on the segments of a reconstructed frame (Fig.6).

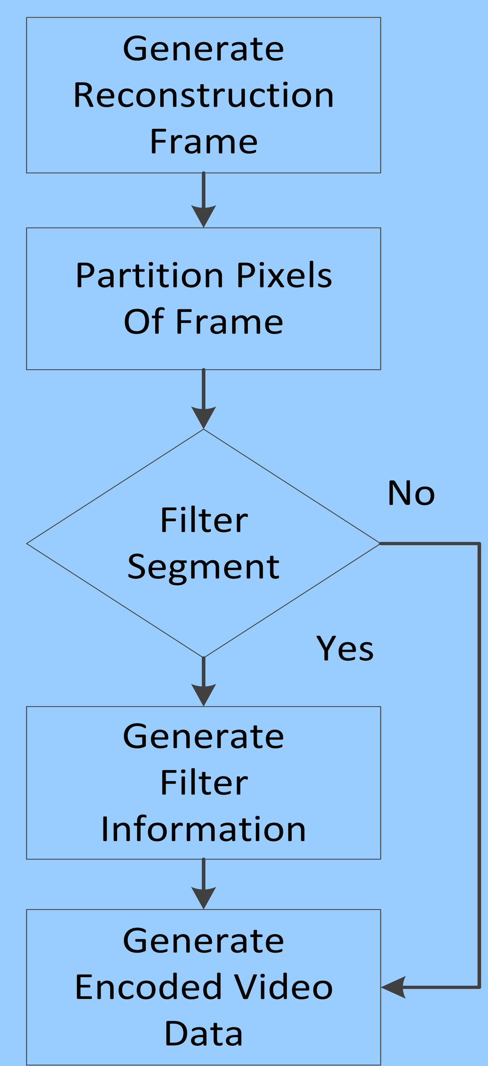

Fig. 5 A flowchart of an encoding procedure

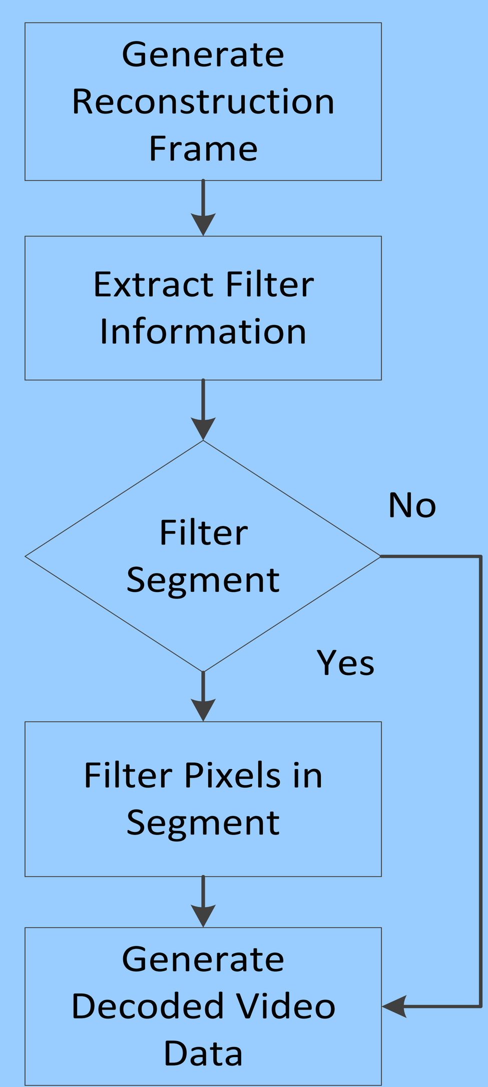

Fig. 6 A flowchart of an decoding procedure

Results

The reconstruction performance of the proposed method, based on pattern information, was compared the Qualcomm method [1], with the tests applied to five 1080p test video sequences having 24 prediction frames; the results are shown in Table1.

By partitioning reconstructed frame pixels with undetermined states (white label in Fig. 7) into multiple segments based on pattern information, respective filter information can be generated for each of the segments (Fig. 8).

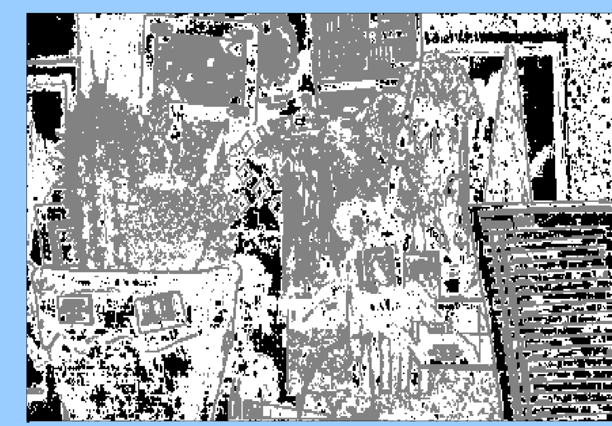

Fig. 7 A segmentation of Fig. 1 with two determined labels (black and gray) and one undetermined label (white).

Comparing the resulting filters from the proposed method (Fig.8) to the benchmark approach (Fig.4), we can see the proposed method is more adaptive to different characteristics of the input frame, such as filter shape and orientation.

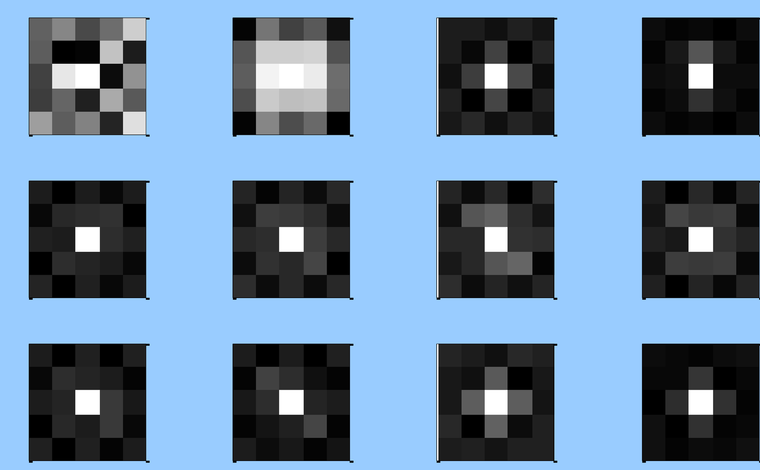

Fig. 8 Adaptive post-filters (5x5-tap) for Fig. 1 generated by the proposed method with an 8-neighbor rotation invariant LBP

Sequence (1080P, 24 P)	Qualcomm method[1] #Filters	PSNR(dB)	Proposed 8-LBP #Filters	PSNR(dB)	Proposed 4-LBP #Filters	PSNR(dB)
ParkScene	16	0.205	12	0.222	13	0.217
Kimono	16	0.220	6	0.208	13	0.216
Basketball	16	0.144	6	0.144	13	0.152
Cactus	16	0.112	12	0.159	13	0.138
BQSquare	16	0.183	12	0.220	13	0.239
Average	--	0.173	--	0.191	--	0.192

Table 1. A comparison of the proposed and Qualcomm [1] methods

Conclusions

The proposed Adaptive Post-Filter has two significant strengths:

- Be able to adapt to different texture and orientation characteristics;
- Require fewer filters allowing simpler filter implementations.

References

1. M. Karczewicz, G. Motta, Y. Ye, and P. Chen, "Poster-filter with reduced delay requirements," *VCEG, COM 16-C128-E*, Jan., 2009.
2. T. Ojala, M. Pietikinen, and T. Menp, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. PAMI*, vol. 24, no. 7, pp. 971–987, 2002.