Papadopoulos, M. A., Zhang, A., Agrafiotis, D., & Bull, D. (2017). An adaptive QP offset determination method for HEVC. In *2016 IEEE International Conference on Image Processing (ICIP 2016): Proceedings of a meeting held 25-28 September 2016, Phoenix, Arizona, USA* (pp. 4220-4224). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ICIP.2016.7533155

AN ADAPTIVE QP OFFSET DETERMINATION METHOD FOR HEVC

Miltiadis Alexios Papadopoulos, Fan Zhang, Dimitris Agrafiotis and David Bull

Department of Electrical and Electronic Engineering, University of Bristol, BS8 1UB, UK
{ma.papadopoulos, fan.zhang, d.agrafiotis, dave.bull}@bristol.ac.uk

ABSTRACT

This paper investigates the effect that the QP offset value has on the coding performance of HEVC. We relate QP offset to the type of texture content present in the sequence. These then are used to develop a low-complexity adaptive QP offset selection method. This enables in-loop configuration of the QP offset parameter in a way that is content dependent and utilizes available encoding statistics. The proposed adaptive method is found to offer average bitrate reductions ranging from 1.38% for dynamic texture sequences up to 1.59% for static texture sequences relative to the QP offset used in the JCT-VC common test conditions.

***Index Terms*—** HEVC, QP offset, adaptive method

1. INTRODUCTION

In HEVC as in previous lossy video coding standards, the level of compression applied and the bit rate generated when coding a sequence is (primarily) controlled through the quantisation parameter (QP). The QP determines how the coefficients of each transform block (TB) will be quantised. The value of the quantisation parameter can be fixed prior to encoding or adjusted continuously by the codec's rate controller (if used). Previous work has looked at ways of determining this QP value either prior to encoding [1–3] or in an adaptive manner during encoding [4–7] with the aim of improving the rate-distortion performance of the codec.

HEVC allows for a different QP offset (QP_{offset}) to be applied to the QP value of each picture (frame) in a group-of-pictures (GOP) relative to the QP value of the closest I frame (QP_I). The QP value of a picture is then given as QP_I + QP_{offset}. The value of the QP offset typically depends on the picture type or the picture temporal ID. Previous work on the H.264 standard introduced an adaptive chroma QP offset decision method based on an adaptive Lagrangian multiplier method [8].

The JCT-VC common test conditions (CTC) [9], that were established during the development of HEVC for evaluating the performance of different coding tools, specify different prediction structures including the random access (RA), one that we focus on in this paper. The random access prediction structure employs hierarchical B-frames following the coding order shown in Fig.1. The same figure also shows the QP offset applied to each frame in the GOP. It can be seen that the more times a frame in a GOP acts as a reference for other frames the lower its QP offset value is, with the latter ranging from 1 to 4. The QP offset parameter is permitted, within the scope of the standard, to acquire a zero or even a negative integer value, the latter resulting in the relevant frame having a QP value that is lower than that of the I frame.

In this work, we study if and when such QP offset values are beneficial to the rate-distortion performance of the codec and how they relate to the texture content of the coded sequence. Based on the findings of this study we propose a low-complexity adaptive method for in-loop configuration of the QP offset parameter that is content dependent and which utilizes available encoding statistics.

The rest of the paper is structured as follows. Section 2 describes the experiment that we conducted for estimating the QP offset that offers the best rate-distortion performance for the sequences tested. Section 3 presents the adaptive QP offset determination algorithm. Section 4 evaluates the performance of the proposed method and compares it with the CTC QP offset strategy. The paper ends with our conclusions and recommendations for future work.

2. OPTIMAL QP OFFSET EXPERIMENT

This experiment is motivated by the study presented in [10]. When coding a sequence, the encoder groups the frames of the sequence into groups of pictures (GOPs), with the pictures being of type I, P or B [11]. The frames of the sequence share the same configuration parameters (e.g. Lagrange multipliers, QP offsets, reference indices etc.) regardless of the GOP they belong to. Zhang and Bull [10] indicated that some of these parameters are far from optimal and could be adaptively configured based on certain attributes of the sequence. Doing this was shown to lead to a better overall rate-distortion performance. More specifically an optimal Lagrange multiplier value was estimated based on the ratio of the distortion exhibited in P frames versus the distortion exhibited in B frames as follows:

The work presented was supported by the "Marie Skłodowska-Curie actions - The EU Framework Programme for Research" project PROVISION and the Visual Information Laboratory of the University of Bristol, UK.

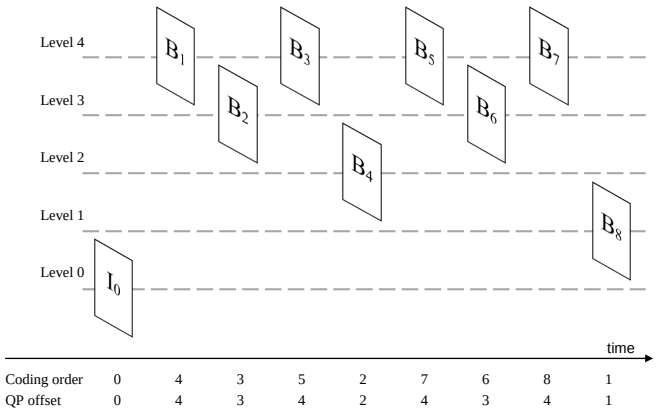

Fig. 1: The CTC random access prediction structure : Hierarchical B-frame structure, coding order and frame QP offset.

$$\mathrm{MSE}_{\mathrm{ratio}} = \frac{\mathrm{MSE}_{\mathrm{P}}}{\mathrm{MSE}_{\mathrm{B}}} \tag{1}$$

where $\mathrm{MSE}_{\mathrm{P}}$ denotes the mean square error (MSE) [12] of P frames and $\mathrm{MSE}_{\mathrm{B}}$ the MSE of B frames in a GOP.

For the purpose of the MSE ratio calculation in Eq.1, the last B frame in the GOP is considered to be a P frame, because it references a single past frame. Distortion was measured as the MSE between the original and compressed frame. The study of [10] additionally analysed the texture content of the sequences used and observed that sequences dominated by static texture content exhibited higher MSE ratio values than sequences with primarily dynamic texture content.

In the experiment described in this section we aim to investigate two things; a) if an optimal QP offset value can be predicted for the frames of the current GOP based on the MSE ratio of the frames in the previous GOP and b) if there is any relationship between the type of texture content in the examined frames (static or dynamic) and the RD behaviour of these frames under different QP offsets. Table 1 lists the sequences used for testing [13, 14] (CIF resolution, 60fps, 4:2:0 chroma) and the texture class that each sequence was assigned to (A - Static, B - Dynamic and C-Mixed). All sequneces were initially encoded using the random access coding structure shown in Table 2 to obtain anchor points. The sequences were subsequently encoded with different QP offsets and compared (bitrate difference) against the corresponding anchor point. A total of 121 frames from each sequence were encoded at QP values of 22, 27, 32, 37 and 42 with QP offsets ranging from -3 to 3 in increments of 1.

Compared to the CTC random access prediction structure shown in Fig.1 ours differs in that we set the initial $\mathrm{QP}_{\mathrm{offset}}$ to 0 and the $\mathrm{GOP}_{\mathrm{Size}}$ to 4 instead of 8. We additionally set the IntraPeriod to -1 thus allowing only the first frame to be intra-coded and the $\mathrm{QP}_{\mathrm{factor}}$ to 0.5 so that its value does not to influence our conclusions from the experiment. The incremental QP offset that is subsequently applied to each frame is multiplied by the level of the coding hierarchy in which the frame resides; Frame1 is in level 0, Frame2 is in level 1, Frame3 and Frame4 are in level 2. This ensures an even QP difference between levels of the GOP structure and simplifies analysis of the results. The encoder settings used are visible in Table 2.

Table 1: Test sequences used in the experiment.

Class	Sequence	Source
A	1. Carpet, 2. Miss-America, 3. Picture	BVI & Standard
B	4. Flag, 5. Spring, 6. Water	DynTex
C	7. Football, 8. Flower, 9. Mobile	Standard & DynTex

Table 2: Coding structure used in the experiment.

```
#======== Coding Structure =============
IntraPeriod :  -1 # Period of I-Frame ( -1 = only first)
DecodingRefreshType :  1 # Random Accesss 0:none ...
GOPSize :  4 # GOP Size (number of B slice = GOPSize-1)
# Type POC QPoffset QPfactor tcOffsetDiv2 ...
Frame1:  B 4 0 0.5 0 0 0 2 3 -4 -6 -8 0
Frame2:  B 2 0 0.5 0 0 1 2 3 -4 -2 2 2 2
Frame3:  B 1 0 0.5 0 0 2 2 4 -3 -1 1 3 2 1
Frame4:  B 3 0 0.5 0 0 2 2 3 -3 -1 1 2 -2
```

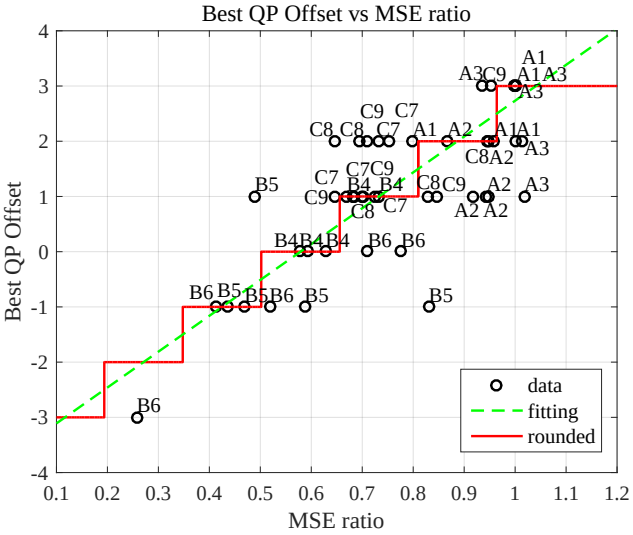

Fig. 2: Optimal QP offset data (black circles), data fitting (green) and quantized curve (red). The labels next to the dots correspond to those in Table 1.

In order to establish the optimal QP offset we examine the change in bitrate resulting from the new QP offsets relative to the achor (anchor : zero QP offset for all frames in the GOP). For each coding point (QP value) the optimal QP offset is the one that offers the highest bitrate reduction for the same PSNR. In Fig.2 the optimal QP offset is plotted against the MSE ratio of Eq.1 as calculated for the anchor. It can be clearly seen that sequences with low MSE ratio benefit more from zero or negative QP offsets (points in the lower left quadrant) as opposed to sequences that have MSE ratios closer to 1 that benefit from positive QP offset values (upper

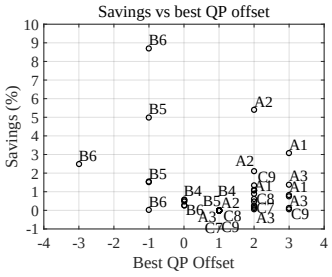

Fig. 3: Bitrate savings for the chosen best QP offset over the CTC QP offset.

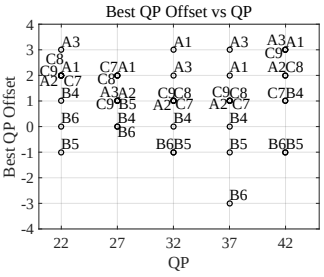

Fig. 4: Best QP offset for each QP.

right quadrant). Given that sequences that include mostly dynamic textures (denoted with a B followed by the number of the sequence in Fig.2) generate lower MSE ratios than those containing static textures (denoted with a A followed by the number of the sequence in Fig.2), we can conclude that coding of dynamic texture content would benefit from the use of zero and negative QP offsets (lower left quadrant of Fig.2). The case is more evident in Fig.4 where the optimal QP offsets are plotted against the QP of the sequence. The dynamic (B) sequences benefit from zero or negative QP offsets regardless of the QP whereas static (A) sequences prefer positive QP offsets. Finally, Fig.3 suggests that bitrate savings of up to 8.7% (compared to the CTC recommended QP offset) can be achieved by using different QP offset values for the sequences in Table 1.

3. ADAPTIVE QP OFFSET DETERMINATION METHOD

Based on the data collected from the experiment above, we formed the linear regression model described in Eq.2 [15, 16]:

$$\mathrm{QP}_{\mathrm{offset\_best}} = a * \mathrm{MSE}_{\mathrm{ratio}} - b \tag{2}$$

where a, b are the fitting coefficients and $\mathrm{MSE}_{\mathrm{ratio}}$ the ratio of the total MSE of P frames to that of B frames as defined in Eq.1. For a GOP size of 4 the fitting coefficients are set to $a = 6.493$ and $b = 3.759$. In this case the mean absolute error of the model equals to 1.

The model is plotted against the experimental data in Fig.2 (green line) alongside a quantised (rounded to the closest integer) version of it that will be used in QP offset determination algorithm (since QP offset can only receive integral values). The Pearson and Spearman correlation coefficients for the data in Fig.2 are computed to be 0.771 and 0.712 respectively, which validates the choice of a linear model. The algorithm was implemented in HEVC HM 16.2 following the pseudocode shown in Algorithm 1.

Line 5-9 calculates the average MSE for the P and B frames separately. Line 10-13 in algorithm 1 updates a buffer K that is used for storing MSE statistics from previously

Algorithm 1 Estimating best $\mathrm{QP}_{\mathrm{offset}}$ for frame m.

1: **for** each frame m **do**
2: **if** frame m is not $\mathrm{I}_{\mathrm{frame}}$ **then**
3: apply previously calculated $\mathrm{QP}_{\mathrm{offset}}$
4: encode frame m
5: **if** frame m is $\mathrm{P}_{\mathrm{frame}}$ **then**
6: update $\mathrm{MSE}_{\mathrm{Pframes}}$ with $\mathrm{MSE}_{\mathrm{m}}$ in $\mathrm{buffer}_{\mathrm{entry}}$ k
7: **else**
8: update $\mathrm{MSE}_{\mathrm{Bframes}}$ with $\mathrm{MSE}_{\mathrm{m}}$ in $\mathrm{buffer}_{\mathrm{entry}}$ k
9: **end if**
10: **if** $\mathrm{buffer}_{\mathrm{entry}}$ k is complete **then**
11: add $\mathrm{buffer}_{\mathrm{entry}}$ k to buffer K
12: keep buffer K size == 3
13: **end if**
14: **else**
15: encode frame m
16: **end if**
17: **for** each GOP i in buffer K **do**
18: calculate $\mathrm{totalMSE}_{\mathrm{Xframes}}$ where $X = P, B$
19: **end for**
20: $\mathrm{MSE}_{\mathrm{ratio}} = \frac{\mathrm{totalMSE}_{\mathrm{Pframes}}}{\mathrm{totalMSE}_{\mathrm{Bframes}}}$
21: $\mathrm{QP}_{\mathrm{offset}} = \mathrm{QP}_{\mathrm{offset}} + round(6.493 * \mathrm{MSE}_{\mathrm{ratio}} - 3.759) * \mathrm{mframe\_level}$
22: $\mathrm{CLIP}(-4 * \mathrm{mframe\_level} < \mathrm{QP}_{\mathrm{offset}} < 4 * \mathrm{mframe\_level})$
23: **end for**

encoded GOPs. Line 20 uses these values to calculate the MSE ratio. Line 21 and 22 calculates the appropriate QP offset for the next GOP making sure that the QP offset values of the different hierarchy levels in the GOP don't differ by more than 3. Note that scene cut detection is not included in our implementation. We can overcome this limitation by applying a scene cut detection method similar to [10].

In our implementation, we tested various $\mathrm{QP}_{\mathrm{offset}}$ update procedures, in which the $\mathrm{MSE}_{\mathrm{ratio}}$ was calculated from up to three previous GOPs. The encoding results were compared against the QP offset behaviour recommended in the HEVC common test conditions and shown in Fig.1 (albeit with a GOP size of 4) where there is a $\mathrm{QP}_{\mathrm{offset}}$ of 1 between the levels of the GOP hierarchy of every GOP. Initially we examined the case of calculating the $\mathrm{MSE}_{\mathrm{ratio}}$ of the first GOP only or the previous available GOP in every encoding iteration. In another case, statistics from the three previous GOPs (size of buffer K was equal to 3) were used to calculate the simple moving average (SMA) or the weighted moving average (WMA) [17] of the $\mathrm{MSE}_{\mathrm{ratio}}$ for the P and B frames. The $\mathrm{QP}_{\mathrm{offset}}$ computed was then replacing the $\mathrm{QP}_{\mathrm{offset}}$ value of the currently encoding GOP. The update procedure that was eventually selected for implementation increments the QP offset by a value computed from the WMA of the latest 3 GOPs using the weights in [17]. By employing a WMA, we allow the most recent GOP statistics to affect the calculated $\mathrm{QP}_{\mathrm{offset}}$ more than those of older GOPs.

4. PERFORMANCE VALIDATION

To validate the performance of the proposed adaptive QP offset method and the derived model, we employed a different set of sequences from those used in the previous sections of this paper. The sequences used for the validation part are listed in Table 3. The roman numerals in Table 3 denote video resolution (I - 416x240, II - 832x480, III - 640x360, IV - 1024x576, V - 1920x1080). As before the letters A,B, and C refer to the texture content of the sequence (static, dynamic and mixed respectively). More information about these sequences including how to get access to them can be found in [18].

Table 3: Sequences used for results validation.

Group & Class	Sequence	Source
I-II.A	Clouds, Fungus, Squirrel	BVI
I-II.B	Drops, Plasma, Sparkler	BVI
I-II.C	Cactus, ParkScene	HEVC
III-V.A	PaintingTilting, PaperStatic	BVI
III-V.B	Sparkler, CalmingWater	BVI
III-V.C	BricksBushesStatic, TreeWills	BVI

Table 4: Adaptive QP_{offset} vs CTC recommended QP offset.

Anchor	Default QP_{offset}=1		Adaptive		BD-Rate
Group & Class	BD-PSNR	BD-Rate	BD-PSNR	BD-Rate	Default-Adaptive
I.A	0.09dB	-1.84%	0.13dB	-2.72%	-0.90%
I.B	-0.08dB	1.46%	0.00dB	0.18%	-1.25%
I.C	0.12dB	-2.94%	0.13dB	-3.09%	-0.25%
II.A	0.07dB	-1.63%	0.11dB	-2.44%	-0.84%
II.B	-0.07dB	1.35%	-0.01dB	0.22%	-1.08%
II.C	0.11dB	-2.85%	0.12dB	-3.01%	-0.26%
III.A	0.19dB	-5.13%	0.26dB	-7.85%	-2.97%
III.B	-0.08dB	1.52%	0.02dB	-0.29%	-1.74%
III.C	0.03dB	-0.51%	0.03dB	-0.53%	-0.06%
IV.A	0.11dB	-3.41%	0.17dB	-5.83%	-2.56%
IV.B	-0.08dB	1.40%	0.02dB	-0.35%	-1.63%
IV.C	0.05dB	-0.93%	0.05dB	-0.89%	0.01%
V.A	0.08dB	-2.94%	0.11dB	-4.26%	-1.40%
V.B	-0.07dB	1.37%	0.01dB	-0.13%	-1.38%
V.C	0.06dB	-1.39%	0.04dB	-1.11%	0.20%
I	0.04dB	-0.88%	0.08dB	-1.72%	-0.87%
II	0.03dB	-0.82%	0.07dB	-1.59%	-0.79%
III	0.04dB	-1.37%	0.10dB	-2.89%	-1.59%
IV	0.03dB	-0.98%	0.08dB	-2.36%	-1.39%
V	0.02dB	-0.99%	0.05dB	-1.83%	-0.86%
A	0.10dB	-2.78%	0.15dB	-4.28%	**-1.59%**
B	-0.08dB	1.42%	0.01dB	-0.03%	**-1.38%**
C	0.07dB	-1.72%	0.07dB	-1.73%	**-0.07%**
Overall	0.03dB	-0.99%	0.08dB	-2.03%	**-1.07%**

Note: *Negative values indicate a reduction in PSNR or bitrate.*

The validation sequences were used to compare the changes in bitrate resulting from i) our method and ii) the QP_{offset}=0 anchor (using the CTC recommended QP_{offset}=1 between the GOP hierarchy levels). Table 4 lists the relative gains in PSNR and bitrate against the anchor reference configuration. The Bjøntegaard metric [19] was used to facilitate RD comparisons. The final column of the table displays the BD-Rate changes relative to the CTC QP offset approach.

As it can be observed regarding the different resolutions, the bitrate savings are relatively consistent across the groups I-V, indicating that our method is independent of the resolution of the sequence.

Regarding the content type, Static (A) and Dynamic (B) textures exhibit higher encoding performance improvement compared to Mixed (C) sequences. An average reduction in bitrate of 1.59% was observed for Static (A) sequences and a 1.38% for Dynamic (C) sequences, reaching 2.97% for Group III.A.

5. CONCLUSIONS

This paper investigated the effect that the QP offset value has on the coding performance of HEVC. The findings of the investigation were related to the type of texture content appearing in the sequence and were used to develop an adaptive QP offset determination method. Our method uses available encoding statistics from previously coded GOPs in the sequence to determine the best QP_{offset} for the current GOP. The performance of the proposed adaptive method was validated and was found to offer on average bitrate reductions of 1.07% for all sequences relative to the QP offset used in the JCT-VC common test conditions QP offset strategy.

The lower performance gains observed for mixed texture sequences (0.07%) suggest that, in order for the method to provide the best improvement possible, a frame partitioning should be applied that divides the frame into regions of homogeneous static and dynamic textures. A combination of this work with the method proposed in [10] is another possibility for future work. Subjective quality evaluation could also prove useful in locating the QP offset limits that prevent the introduction of perceivable quality fluctuations.

6. REFERENCES

[1] Zhenyu Wu, Hongyang Yu, Bin Tang, and Chang Wen Chen, "Adaptive initial quantization parameter determination for H.264/AVC video transcoding," *IEEE Transactions on Broadcasting*, vol. 58, no. 2, pp. 277–284, 2012.

[2] Thibaud Biatek, Michael Raulet, J-F Travers, and Olivier Deforges, "Efficient quantization parameter estimation in hevc based on ρ-domain," in *2014 Proceed-*

ings of the 22nd European Signal Processing Conference (EUSIPCO). IEEE, 2014, pp. 296–300.

[3] László Czúni, Gergely Császár, and Attila Licsár, "Estimating the optimal quantization parameter in H.264," in *2006. ICPR 2006. 18th International Conference on Pattern Recognition*. IEEE, 2006, vol. 4, pp. 330–333.

[4] Manoranjan Paul, Abhilash Antony, and G Sreelekha, "Performance improvement of hevc using adaptive quantization," in *Advances in Computing, Communications and Informatics (2014 International Conference on ICACCI*. IEEE, 2014, pp. 1428–1433.

[5] Jianying Zhu, Zhelei Xia, Haibing Yin, and Qiang Hua, "A novel quantization parameter estimation model based on neural network," in *2012 International Conference on Systems and Informatics (ICSAI)*. IEEE, 2012, pp. 2020–2023.

[6] Sudeng Hu, Hanli Wang, and Sam Kwong, "Adaptive quantization-parameter clip scheme for smooth quality in H.264/AVC," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1911–1919, 2012.

[7] Chuohao Yeo, Hui Li Tan, and Yih Han Tan, "Ssim-based adaptive quantization in hevc," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1690–1694.

[8] Cheolhong An and Truong Q Nguyen, "Adaptive lagrange multiplier selection using classification-maximization and its application to chroma qp offset decision," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 21, no. 6, pp. 783–791, 2011.

[9] Frank Bossen, "Common test conditions and software reference configurations," *JCTVC-H1100, Document*, 2012.

[10] Fan Zhang and David R Bull, "An adaptive lagrange multiplier determination method for rate-distortion optimisation in hybrid video codecs," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 671–675.

[11] Thomas Wiegand, Gary J Sullivan, Gisle Bjøntegaard, and Ajay Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[12] Zhou Wang and Alan C Bovik, "Mean squared error: love it or leave it? a new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[13] Renaud Péteri, Sándor Fazekas, and Mark J Huiskes, "Dyntex: A comprehensive database of dynamic textures," *Pattern Recognition Letters*, vol. 31, no. 12, pp. 1627–1632, 2010.

[14] Miltiadis Alexios Papadopoulos, Fan Zhang, Dimitris Agrafiotis, and David Bull, "Bvi video texture database," `http://data.bris.ac.uk/data/dataset/1if54ya4xpph81fbo1gkpk5kk4`, 2015, Accessed Jan. 28, 2016.

[15] Michael Schmidt and Hod Lipson, "Distilling free-form natural laws from experimental data," *science*, vol. 324, no. 5923, pp. 81–85, 2009.

[16] Michael Schmidt and Hod Lipson, "Eureqa (version 0.98 beta)[software]," 2014.

[17] John Devcic, "Weighted moving averages: The basics," `http://www.investopedia.com/articles/technical/060401.asp`, 2010, Accessed Jan. 28, 2016.

[18] Miltiadis Alexios Papadopoulos, Fan Zhang, Dimitris Agrafiotis, and David Bull, "A video texture database for perceptual compression and quality assessment," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 2781–2785.

[19] Gisle Bjontegaard, "Calcuation of average PSNR differences between RD-curves," *Doc. VCEG-M33 ITU-T Q6/16, Austin, TX, USA, 2-4 April 2001*, 2001.