IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 13, 2019, accepted March 4, 2019, date of publication April 3, 2019, date of current version April 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2908997

Content-Adaptive Memory for Viewer-Aware Energy-Quality Scalable Mobile Video Systems

JONATHON EDSTROM[1], YIFU GONG[1], ALI AHMAD HAIDOUS[1], BRITTNEY HUMPHREY[1], MARK E. MCCOURT[2], YIWEN XU[3], JINHUI WANG[4], (Member, IEEE), AND NA GONG[4], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58108, USA
[2]Department of Psychology, North Dakota State University, Fargo, ND 58108, USA
[3]Department of Industrial and Manufacturing Engineering, North Dakota State University, Fargo, ND 58108, USA
[4]Department of Electrical and Computer Engineering, University of South Alabama, Mobile, AL 36608, USA

Corresponding author: Na Gong (nagong@southalabama.edu)

This work was supported in part by the National Science Foundation under Grant CCF-1855706, and in part by the Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under Grant P30 GM114748.

ABSTRACT Mobile devices are becoming ever more popular for streaming videos, which account for the majority of all the data traffic on the Internet. Memory is a critical component in mobile video processing systems, increasingly dominating the power consumption. Today, memory designers are still focusing on hardware-level power optimization techniques, which usually come with significant implementation cost (e.g., silicon area overhead or performance penalty). In this paper, we propose a video content-aware memory technique for power-quality tradeoff from viewer's perspectives. Based on the influence of video macroblock characteristics on the viewer's experience, we develop two simple and effective models–decision tree and logistic regression to enable hardware adaptation. We have also implemented a novel viewer-aware bit-truncation technique which minimizes the impact on the viewer's experience, while introducing energy-quality adaptation to the video storage.

INDEX TERMS Viewer's experience, video memory, video content, viewer-aware bit truncation, energy-quality adaptation.

I. INTRODUCTION

Video is *everywhere* today. According to the recent Cisco Visual Networking Index, Mobile video traffic accounted for 60% of total mobile data in 2016 [1]. It is expected to increase 9-fold between 2016 and 2021 and grow to approximately 78% in 2021, with the continuous evolution of mobile networks and the proliferation of mobile devices [1]. Consequently, video steaming has become one of the most energy-intensive applications on mobile devices. In particular, during the mobile video steaming process, the frequent memory access contributes to over 92% of the motion compensation energy [2] and 50% of the video decoding consumption [3]; the high energy consumption restraints are only expected to increase with the emerging of Ultra-High-Definition (UHD) (e.g., 4K and 8K) videos [4]. Accordingly, enhancing energy efficiency of video memories is of paramount importance to enable efficient mobile video systems, and is also one of the key design considerations to deliver 4K/8K UHD videos to mobile devices.

Designers have extensively exploited memory techniques for power reduction, but traditional memory designs are typically developed based on an objective video output metric such as the peak signal-to-noise ratio (*PSNR*), without dynamic energy-quality adaptation to viewer's true experience. Such hardware-viewer isolation is mainly due to the following design challenges. First, the existing models to represent viewer's experience, such as the recently developed human visual system (HVS) model [5], are too conceptual and too complex, to be useful in guiding hardware design. Second, hardware design, particularly memories, usually lack run-time adaptation and therefore new hardware design techniques that enable viewer-aware adaptation need

The associate editor coordinating the review of this manuscript and approving it for publication was Yun Zhang.

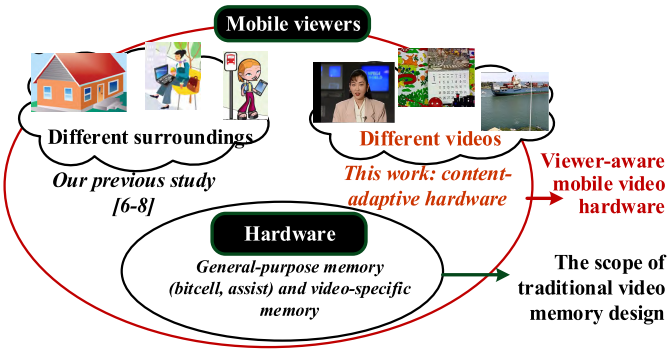


FIGURE 1. Proposed content-adaptive mobile video memory for viewer-aware mobile video systems.

to be explored. Last, but not least, it is challenging for mobile designers to directly connect hardware design to viewer's experience, which requires professional lab setup, human subject involvement, and psychophysical analysis.

We have recently explored viewer-aware video memory design by investigating the impact of illuminance levels in different viewing surroundings on the viewer's experience [6]–[8], as illustrated in Fig. 1. Specifically, we used a bit truncation technique to introduce memory failures in high noise-tolerance viewing contexts with high luminance levels by adaptively disabling the least significant bits (LSB) of the video data stored in memories. Our previous studies [6]–[8] illustrate a new dimension of power savings for hardware design through the introduction of viewer awareness, but the developed memory lacks adaptation across a wide variety of mobile videos. To enable an optimized trade-off between energy efficiency and video quality, in this paper, we propose a novel energy-quality scalable video memory design technique that takes into account video content to adjust the energy-quality trade-off according to viewer's experience. Specially, this paper makes the following contributions.

- We developed a Xilinx Zynq 7020 FPGA based H.264 decoder and display system and based on it, the contribution of different video memories to the output quality has been analyzed. We demonstrate that the frame buffer can tolerant significant memory failures, which enables power saving opportunities for hardware design (Section III.A and Appendix).
- The impact of video content on viewer's experience is studied from the psychological perspective. We conclude that the correlation characteristics between "banding distortion" to viewers caused by hardware noise and the areas in frames that exhibit low variance among pixel luminance values have the potential to enable content-adaptation opportunities for hardware design (Section III.B).
- Based on macroblock characteristics analysis and subjective video testing, two models including one decision tree model and one logistic regression model have been developed to enable effective connection of the video content to the hardware design process (Section IV).
- We further develop a novel viewer-aware bit truncation technique which enables better visual experience while maintaining similar power efficiency. Based on the developed models and viewer-aware bit truncation technique, a content-adaptive video memory design with dynamic energy-quality trade-off is implemented (Section V).
- Finally, a comprehensive suite of simulations on the proposed content-adaptive video memory is performed and the enriched results including performance, layout design, video output quality of various mobile videos, and power efficiency, are discussed (details are shown in Section VI).

To the best of the authors' knowledge, the proposed memory has made the first attempt to ***exploit viewer's experience and video content*** to enable energy-quality adaptive hardware design.

The organization of the paper is as follows. A review of related video memories is provided in Section II. In Section III, we study the contributions of video memories and the impact of video content on viewer's experience. In Section IV, we present our subjective testing procedure and model development process. Our hardware design, that implements the functionality for power savings through viewer-aware bit truncation, is presented in Section V. The evaluation results are presented in Section VI. Finally, we conclude the paper in Section VII.

II. RELATED WORK

There is a rich body of literature for power reduction for embedded memories and voltage scaling is particularly effective to reduce the memories' power consumption due to the strong dependency of dynamic and leakage power consumption on supply voltage. However, voltage-scaled SRAMs are susceptible to failures and many techniques have been developed, which mainly fall into the following three aspects: (i) *assist* schemes such as boosted wordline [9], negative bitline [10], and dual-rail supply [11]; (ii) *more-than-6T* bitcells to achieve low voltage operation, such as 8T [12], 9T [13], and 10T [14]; and (iii) *error-correction* techniques such as

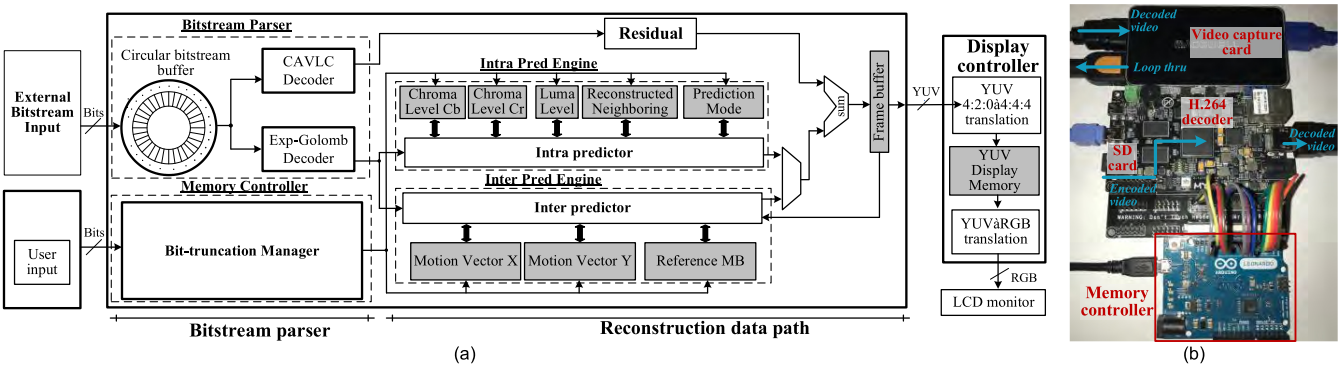


FIGURE 2. Mobile video memory architecture. (a) Block diagram of the mobile video decoding and display process and (b) Xilinx Zynq 7020 FPGA based system implementation. Different memories are shaded. MB: Macroblock.

error correction codes [15] and data remapping [16]. However, the improvements in embedded memory power efficiency are often achieved with significant design complexity, silicon area overhead, and performance penalty for voltage regulators and boosting circuits.

Several recent efforts have investigated application resilience of videos to approximations with ''good enough'' output and additional power savings. Chang *et al.* [17] present a hybrid 6T+8T SRAM to achieve quality-power optimization. In [18], a heterogeneous sizing scheme is presented to reduce the failure probability of conventional 6T bitcells. In [19], the correlation between the most-significant-bits (MSBs) of video data was utilized to design a hybrid 8T+10T memory for power savings.

At the same time, alternative metrics for the analyzing videos objectively, including Structural Similarity (*SSIM*) and *PSNR-B*, have recently been shown to outperform the traditional mean squared error (*MSE*) and *PSNR* [32], [33]. While *SSIM* and *PSNR-B* have more meaning in terms of the viewer's perception of a video, the complexity of their calculations makes them less useful to hardware designers when optimizing energy-quality tradeoff.

Very recently, we have investigated viewer-aware video memory design by studying the impact of illuminance levels in viewing contexts on the viewer's experience [6]–[8], where an increased amount of ambient luminance allows for a larger amount of bits to be truncated without noticeable degradation to the viewers. We developed a viewing context-aware SRAM (VCAS), which introduces memory failures in luminance contexts with high memory failure tolerance. Two low-power techniques - voltage scaling and bit truncation - are explored to implement. We conclude that those two techniques achieve similar *PSNR* values, but the video quality degradation caused by bit truncation is much less noticeable than that of the voltage scaling technique for the viewers. The hardware design from our previous study has been developed for general videos, although, we observed that the video characteristics significantly influence the viewer's experience [8]. In this paper, we study the impact of video content characteristics on viewer's experience to enable video content-adaptive memory with dynamic energy-quality tradeoff.

It is worthy to emphasize that, the proposed content-adaptive video memory as well as viewing luminance-aware video memories [6]–[8] are orthogonal to existing low-power hardware-level memories and they can be applied simultaneously to optimize power efficiency.

III. INFLUENCE OF VIDEO CONTENT ON VIEWER'S EXPERIENCE

A. MOBILE VIDEO MEMORY SYSTEM

Video streaming has become the most important energy-intensive application used in mobile devices [8]. Fig. 2 (a) shows the block diagram of a H.264 video decoding and display system [36]. After parsing compressed bitstream, the inter predictor uses the reconstructed frames stored in the reference frame buffer and the transmitted motion vectors to construct new frames. After the frames are decoded, the display controller sends them from the frame buffer to the display panel periodically. During this process, multiple memories are needed for storing the intermediate and final results of the frame data, as listed in Table 1.

To evaluate the contribution of different memories to the output video quality, we developed a video decoder and display system, as shown in Fig. 2 (b). For the memories listed in Table 1, we applied the bit truncation technique [7] to each memory during the video decoding process by disabling least-significant bits (LSBs) [8], [21] and then the output video is captured for quality evaluation. Specifically, LSB truncation starting with one bit with a maximum of five bits have been applied to each video memory. The encoded bitstream, which resided on an on-board SD card, is decoded using a Xilinx Zynq 7020 FPGA based H.264 decoder. An Arduino-based memory controller is implemented select the memory for truncation as well as the number of truncated LSB which are specified by the user input over a serial interface. A video capture card is utilized to capture the video output over the HDMI output for evaluation. It has been shown that, the frame buffer, the largest memory, can tolerant three truncated LSBs, which provides power saving opportunities for a hardware design. The detailed results are discussed in Appendix.

TABLE 1. Video memories and their functionality.

Video Memories	Size in Bits (Width x Depth)	Memory Functionality
Chroma Level Cb	32 x 8	Stores the blue-difference color space bottom line pixels for up macroblocks
Chroma Level Cr	32 x 8	Stores the red-difference color space bottom line pixels for up macroblocks
Luminosity Level	32 x 8	Stores the luminosity color space bottom line pixels for up macroblocks
Reconstructed Neighboring	32 x 7	Stores neighboring pixels of a luma block after the current macroblock is coded and reconstructed
Prediction Mode	16 x 7	Stores the current macroblock prediction mode for 4x4 blocks
Motion Vector X	64 x 7	Stores the horizontal motion vector prediction calculation of surrounding blocks' motion data
Motion Vector Y	64 x 7	Stores the vertical motion vector prediction calculation of surrounding blocks' motion data
Reference Macroblock	8 x 8	Stores the reference I, SI, P, or SP macroblock used for inter prediction
Frame buffer	**64 x 512**	**Stores the current and previous decoded frames for prediction and display, respectively**
Y Display	64 x 8	Stores the luma Y component of the display memory for HDMI output buffer
U Display	64 x 8	Stores the chrominance U component of the display memory for HDMI output buffer
V Display	64 x 8	Stores the chrominance V component of the display memory for HDMI output buffer

B. INFLUENCE OF VIDEO CONTENT ON VIEWER'S EXPERIENCE IN THE PRESENCE OF HARDWARE NOISE

Traditionally, hardware designers have used *PSNR*for evaluating video quality, which has been recently shown to be insufficient to demonstrate the viewer's experience [20], [31]. *PSNR* does not encompass the necessary information to hardware designers about viewer's experience, due to the fact that key influencing factors for viewer's experience, such as video content and environment conditions, are not included in *PSNR* [31]. In this paper, we aim to find a better method to analyze videos in a quantitative way that will also be useful to hardware researchers. We begin this process by using the *PSNR* metric to describe video quality. We continue by adding new insight to the traditional *PSNR* metric with the introduction of *content-aware* information. This new form of information allows us to gracefully scale the video quality with enhanced energy efficiency of hardware.

1) TRADITIONAL PSNR METRIC

The traditional *PSNR* metric is defined as [19]

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \quad (1)$$

where the *MSE* is the mean squared error between the original video (*Org*) and the degraded video (*Deg*), expressed as

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[Org\,(i,j) - Deg(i,j)]^2 \quad (2)$$

Although *PSNR* is simple for hardware designers to understand, it does not truly capture the effect that errors have on the user's perception of the video. To show the lack of complete information the *PSNR* provides in terms of user perception, we apply the bit truncation technique to two videos and calculate the *PSNR* values for 1 to 4 truncated LSBs of the luma data (i.e. the luminance channel, or Y component in raw YUV videos). We adopt the bit truncation technique to enable energy-quality adaption, which is due to the following two reasons: (i) bit truncation causes blurring in videos, which is similar to the "*banding distortion*" in the codec-algorithm field, and the video degradation is much less noticeable to viewers as compared to other low-power techniques such as voltage scaling [8] and (ii) the power/energy savings with bit truncation is much more significant than other low power techniques such as voltage scaling [21].

TABLE 2. Output quality of different videos with bit truncation.

Video output quality with 3 LSBs truncated	# LSBs truncated	*PSNR* (dB)
	**Video #1** (video tag: *EFv2FvnlLao*)	
	1	52.868
	2	44.433
	3	**37.490**
	4	30.985
	**Video #2** (video tag: *FNlpA4FME-8*)	
	1	52.741
	2	44.461
	3	**37.693**
	4	31.154

Table 2 shows two videos which were downloaded from Google's recently released Youtube-8M Dataset [22], which is the largest multi-label video dataset. In this paper, to maintain a short and consistent size label for all included YouTube video samples, we use the video tag to label each video, which is the last portion of the full URL address[1]. As observed in Table 2, using the bit truncation technique, the *PSNR*value is reduced by approximately 7dB, on average, for each additional truncated LSB. Both videos have very similar *PSNR* values with the same number of LSBs truncated, but the visual quality is significantly different. As compared to video #1 (video tag:*EFv2FvnlLao*), the "banding distortion" of video #2 (video tag: *FNlpA4FME-8*) is much more noticeable to the viewers. Accordingly, the traditional video quality metric *PSNR*cannot correlate well with the viewer's

***PSNR*: 48.1 dB**
Min. Plain MB %: 18.37%; Max. Plain MB %: 20.22%
Median Plain MB %: 19.35%; Mean Plain MB %: 19.34%

***PSNR*: 47.8 dB**
Min. Plain MB %: 53.26%; Max. Plain MB %: 54.56%
Median Plain MB %:53.69%; Mean Plain MB %: 53.71%

FIGURE 3. **Plain MBs visualization and video output comparison of two videos with varying plain MB % (with 2 LSBs truncated). White: plain MBs.**

experience and the video-content properties, such as the texture/motion characteristics, significantly affect the viewer's experience. In this paper, we introduce the video content information to study the viewer's experience. Specifically, we adapt the recently developed video macroblock (MB) characterization by analyzing the pixel-luminance values' variance [23], as described in the next subsection.

2) VIDEO MACROBLOCK VARIANCE ANALYSIS

The MB variance analysis is typically conducted during the video pre-processing stage when encoding videos [23], [24]. In our analysis, we adopt their defined calculation for determining whether a given MB is considered either *plain* or *textured*, which avoids introducing significant computational overhead. The calculation is based on the variance of pixel luminance values of a given MB and is defined as [23]

$$V_{MB} = \sum_{i=0}^{15} \sum_{i=0}^{15} (P(i,j) - \rho_{MB})^2 \gg 8$$
$$MB = \begin{cases} Plain & if\,(V_{MB} \leq Th_{Low}) \\ Textured & Else \end{cases} \quad (3)$$

where ρ_{MB} and V_{MB} are the average luminance and variance of luminance values in a given MB, respectively. The value we used for Th_{Low} was 1.25 as was determined in [24] through the use of regression analysis. For our purposes, this Th_{Low} value is an arbitrary number used to define our plain macroblock percentages in our model design process (Section IV). This MB characterization can be calculated during the encoding process and transmitted as metadata in the video bit stream. Currently we use an embedded system implementation for calculating this average plain MB calculation. To minimize computational overhead, we calculate a single, averaged plain MB percentage that represents an entire sample. However, it is possible to calculate a per frame MB percentage for videos that change scenes frequently for dynamic adaptation. Two benchmark videos, *Akiyo* and *News*, were initially retrieved from [25]; these videos contained static backgrounds with a low amount of motion from the reporter(s) in the videos. Both videos displayed low plain macroblock percentages when analyzed. We further obtain 32 video samples with similar broadcasting characteristics from the Youtube-8M Dataset [22] and calculate the percentages per frame for the *minimum*, *maximum*, *median*, and *average* percentage of each sample video. Fig. 2 displays two video samples with similar *PSNR* values but varying[1] plain MB percentages (with 2 LSBs truncated). The distribution of plain MBs and the resulting banding distortion effect are visualized in Fig. 3. An important observation is

[1] *These tags are used as a unique key that points to the corresponding YouTube video. For example, the video tag* EFv2FvnlLao *can be used to locate the original video sample on YouTube using the following URL: https://www.youtube.com/watch?v=* EFv2FvnlLao.

that ***a noticeable relationship exists between the banding distortion and plain MBs***;videos with large amounts of plain MBs, especially where the plain MBs are dense, tend to have decreased visual quality to the viewers. Accordingly, we utilize this relationship to develop a content-adaptive model to predict the number of truncated LSBs for different videos. Specifically, to minimize the computational overhead, we use the *average* plain MB percentage per video frame and focus on low-motion videos with a stationary camera or containing a reporter in our analysis.

IV. MODELING PROCESS

To determine the acceptable number of LSBs to truncate for different videos, we conduct subjective video testing and based on the collected data, develop two models using decision tree and logistic regression methods. For our initial study, described in this paper, we only consider the luma (Y) component when truncating LSBs.

A. SUBJECTIVE TESTING PROCEDURE FOR DATA COLLECTION

We conduct two sets of subjective video studies to collect viewers' feedback. Within each of the studies designed for subjective analysis of truncation techniques, participants were asked to view multiple versions of the same video. Our testing procedure follows guidelines from the ITU [26] and uses the Degradation Category Rating (DCR) method [20], which is also known as the Double Stimulus Impairment Scale (DSIS). The participants were asked to watch both original video and truncated video and then score from 1 to 5 based on the quality in their opinions (*imperceptible-5, perceptible but not annoying-4, slightly annoying-3, annoying-2, very annoying-1*). We used an average score of 4.0 or higher as the target for acceptable video quality [27]. The first (second) of two studies contained 10 (13) participants who were each asked to view 7 (9) individual videos from our 34 sample videos.

With these average scores for different amounts of LSBs truncated, we split the video samples into different regions. Based on this, we develop models that connect the average plain MB percentage to number of LSBs that can be truncated.

B. MODELING PROCESS

1) DECISION TREE MODEL

From our initial subjective studies, we aim to model the correlations between the calculated average plain MB percentage and the largest amount of LSBs that can be truncated for a given *PSNR* that will maintain an acceptable video quality. Fig. 4 displays the video samples average plain macroblock percentage and how many bits can be truncated based on the minimum acceptable impairment score of 4.0.

From these preliminary results, we discover an inverse relationship between plain MB percentage and acceptable number of LSBs to truncate. With the knowledge of this relationship and the subjective data gathered from participants, we develop a decision tree model using the Classification Learner tool in MATLAB, as shown in Fig. 5. By traversing the tree from the top to the bottom based on the plain MB percentages, the number of truncated LSBs can be obtained for different videos. It is worthy to mention that the majority of videos from the Youtube-8M dataset have plain MB percentages above 1.96405% (see Fig. 5) and therefore the number of videos with the decision for 3 LSBs truncation is much less than that of 1 LSB and 2 LSBs truncation.

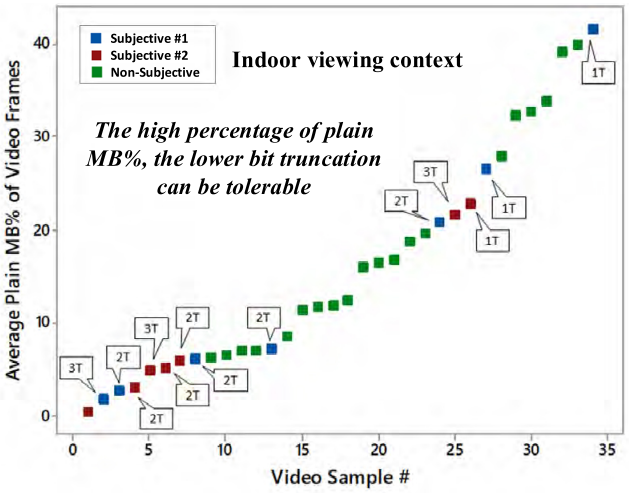


FIGURE 4. Acceptable truncated bits based on subjective feedback. 1T: 1 LSB truncated; 2T: 2 LSBs truncated; 3T: 3 LSBs truncated.

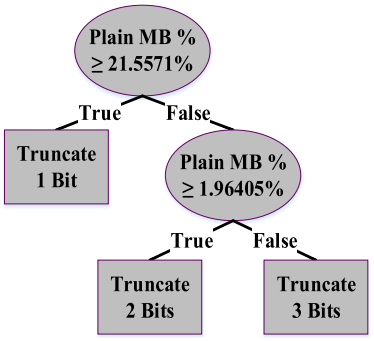


FIGURE 5. Developed decision tree model for bit truncation.

2) LOGISTIC REGRESSION MODEL

In our model development process, we also considered another widely-applied statistical modeling method: *logistic regression*, in which we have

$$\begin{cases} \pi_i &= \pi_3 \exp(\beta_{i0} + \beta_{i1} x), i = 1, 2 \\ \pi_3 &= \frac{1}{1+\exp(\beta_{10}+\beta_{11}x)+\exp(\beta_{20}+\beta_{21}x)} \end{cases}, \quad (4)$$

where $\pi_i := P\{Y = i|x\}$ indicates the probability that the number of truncated LSBs is i for given average plain MB percentage which equals x. We use Matlab to fit the $\hat{\beta}$ coefficients and get $\hat{\beta}_{10} = -1.6636$, $\hat{\beta}_{11} = 12.7929$, $\hat{\beta}_{20} = 1.4408$, $\hat{\beta}_{21} = 1.0497$. However,

TABLE 3. Results of ordinal logistic regression.

x	P {1 LSB truncated}	P { 2 LSBs truncated }	P { 3 LSBs truncated }	*Decision for LSB truncation*
0.05	0.0876	0.7261	0.1863	2 LSBs
0.10	0.1354	0.7415	0.1231	2 LSBs
0.15	0.2034	0.7174	0.0793	2 LSBs
0.20	0.2939	0.6560	0.0502	2 LSBs
0.25	0.4043	0.5643	0.0314	2 LSBs
0.28504	***0.4888***	***0.4888***	***0.0224***	***2 LSBs***
0.30	0.5253	0.4552	0.0195	1 LSB
0.35	0.6434	0.3446	0.0120	1 LSB
0.40	0.7463	0.2463	0.0074	1 LSB
0.45	0.8275	0.1679	0.0046	1 LSB
0.50	0.8866	0.1105	0.0028	1 LSB
0.55	0.9273	0.0710	0.0017	1 LSB
0.6	0.9541	0.0448	0.0011	1 LSB

their corresponding p-values are 0.243, 0.103, 0.111, 0.881, respectively. This implies that all four coefficients are not significant in the regression under a 5% significance level. By observing the data, one can clearly see that this is due to noise.

In addition, we notice that, if a user chooses a video as satisfactory which is truncated by k LSBs, then he/she will be satisfied by the same video truncated by k' LSBs where $0 < k' < k$. The difference between k LSBs and k' LSBs truncation is the energy efficiency that can be enabled; the efficiency is higher for k LSB truncations. To this end, we further apply the *ordinal logistic regression*, which yields

$$\ln\left(\frac{\pi_1}{\pi_2+\pi_3}\right) = \beta_{10} + \beta_1 x \tag{5}$$

$$\ln\left(\frac{\pi_1+\pi_2}{\pi_3}\right) = \beta_{20} + \beta_1 x \tag{6}$$

Moreover, we have

$$\pi_1 + \pi_2 + \pi_3 = 1 \tag{7}$$

Solving (5), (6) and (7), one can get

$$\begin{aligned}
\pi_1 &= \frac{\exp(\beta_{10}+\beta_1 x)}{1+\exp(\beta_{10}+\beta_1 x)} \\
\pi_2 &= \frac{1}{1+\exp(\beta_{10}+\beta_1 x)} - \frac{1}{1+\exp(\beta_{20}+\beta_1 x)} \\
\pi_3 &= \frac{1}{1+\exp(\beta_{20}+\beta_1 x)}.
\end{aligned} \tag{8}$$

We use Matlab to fit the ordinal coefficients and get $\hat{\beta} = \left[\hat{\beta}_{10}, \hat{\beta}_{20}, \hat{\beta}_1\right] = [-2.8322, 0.9856, 9.7783]$, with p-values $\boldsymbol{p} = [0.0039, 0.1710, 0.0156]$, respectively. With this ordinal logistic regression, only β_{20} is not significant under a 5% significance level and the result is much better than the previous case using the standard logistic regression.

Table 3 lists the ordinal logistic regression results. One can see that there is no decision for 3 LSBs truncation based on the ordinal logistic regression model. This is mainly because very few videos with 3 truncated LSBs are considered acceptable by the participants; also, most of the video testing results with 3 LSBs truncation are considered to be noisy data. When the plain MB percentage (x) is 0.28504 (i.e., 28.504%), we have *P{1 LSB truncated}=P{2 LSBs truncated}*= 0.4888. Accordingly, if $x > 28.504\%$, 1 LSB is truncated; otherwise, 2 LSBs would be truncated.

The developed decision tree model and ordinal logistic regression model only involve very few parameters and the computation time is negligible. The comparison of results between the developed decision tree model and ordinal logistic regression model will be discussed in Section VI.

V. QUALITY OPTIMIZED BIT TRUNCATION DESIGN

In this paper, we also propose a new viewer-aware bit-truncation technique which has less visual quality degradation with the same number of LSBs truncated. Based on the developed bit-truncation technique and models, we implement an energy-quality scalable memory with content adaptation.

A. QUALITY OPTIMIZED BIT TRUNCATION

Bit truncation can adjust the video data's bit-depth by disabling LSBs to enable power savings and it has been applied widely in low-power hardware design [8], [21]. In this paper, we introduce viewer-awareness to the hardware-design process and develop a new hardware-implementation scheme for bit truncation with a minimized effect on the viewer's experience.

Suppose that we are truncating the lowest t LSBs of each luma (Y) byte. For a given video, we can calculate the true numerical value for these truncated bits. However, if we consider all videos in general, the true (decimal) value of these truncated t LSBs should be considered a random variable. These truncated t LSBs may express any decimal numbers among $0, 1, 2, \cdots, 2^t - 1$, because we do not have general prior knowledge that works for all videos. A crucial question is as follows: *what value should be set/given after the true value of these lowest t bits are truncated?* A natural and intuitive method is to make them all zeros. For example, if the true value of a byte is 10101110(B) and three bits are truncated, then the byte's value after truncation is 10101000(B). Setting the truncated bits as zeros has been widely adopted by designers [8], [21]. However, in the following proposition, we show that this value is not the best for minimizing the expected mean square error, E (MSE).

Proposition 1. Suppose that the lowest t LSBs of a byte are truncated. Without losing generality, it is assumed that the true value of these bits is evenly distributed. Then, the best value for these t truncated bits, in terms of minimizing E (MSE), is $10\cdots0(B)$ (with $t-1$ zeros).

Proof. Let random variable Y indicate the true numerical value which is expressed by the truncated t LSBs. Because Y is evenly distributed, we have the following probability mass function (pmf) for Y:

Y =	0	1	2	⋯	2^t-1
probability	$1/2^t$	$1/2^t$	$1/2^t$	⋯	$1/2^t$

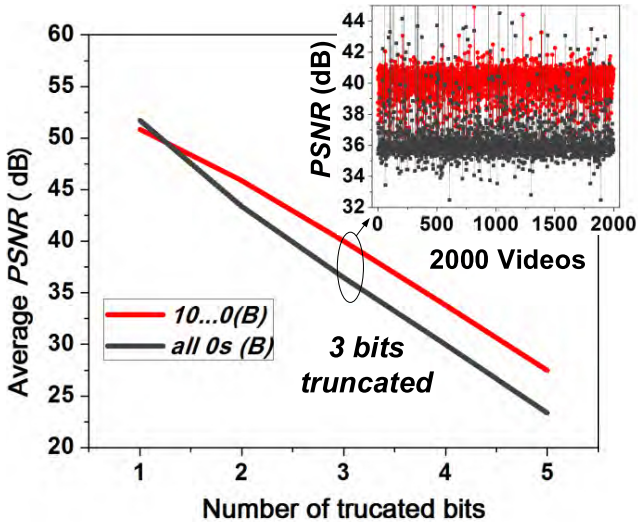


FIGURE 6. Average *PSNR* values of 2,000 YouTube-8M videos using two different truncation techniques.

Let x be the targeted (decimal) value that is set for these truncated LSBs. We aim to minimize E (MSE), namely to minimize

$$f(x)=\frac{1}{2^t}\left[(x-0)^2+(x-1)^2+\cdots+\left(x-(2^t-1)\right)^2\right] \quad (9)$$

Let

$$0=\mathrm{f}'(\mathrm{x})=\frac{1}{2^{t-1}}\left[x+(x-1)+\cdots+\left(x-\left(2^t-1\right)\right)\right]\Rightarrow$$
$$x=2^{t-1}-\frac{1}{2} \quad (10)$$

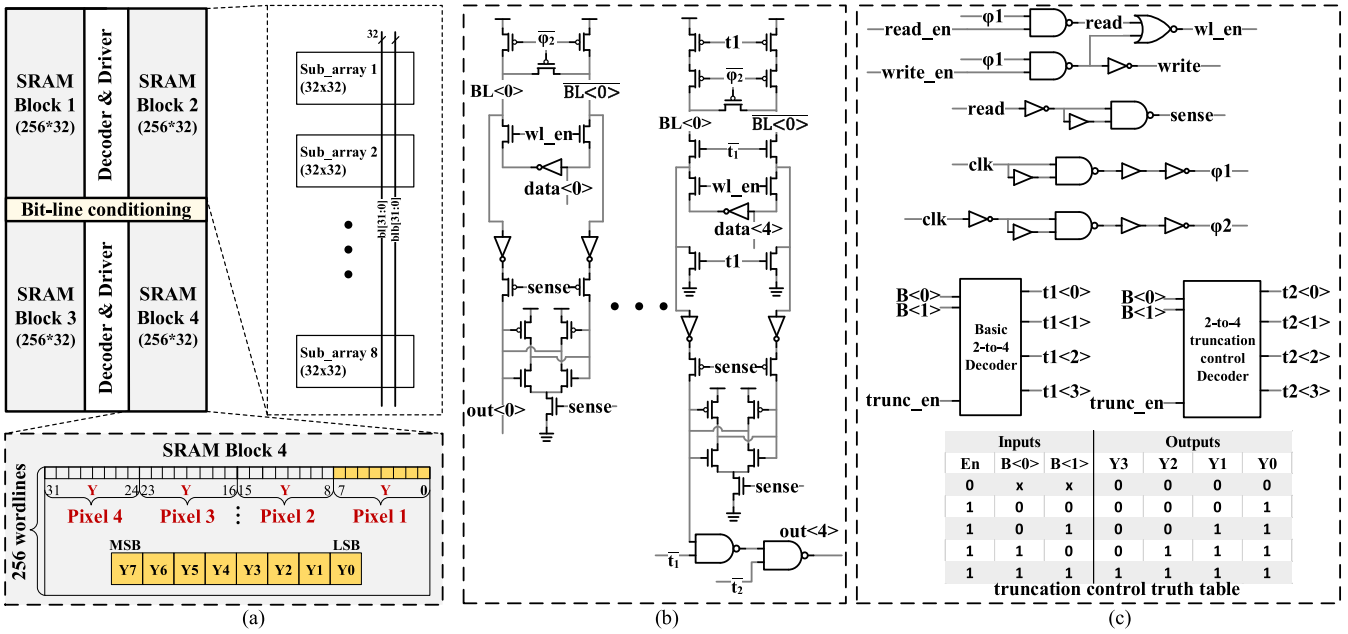
Because x is an integer, we take $x = 2^{t-1} = 10\cdots0(B)$ (with $t-1$ zeros).

The significance of Proposition 1 is that it shows the dependence between the value set for the truncated bits and the expected MSE and that it gives the best value, in general. We randomly selected 2,000 unique videos, representing 100,000 individual frames, from YouTube-8M [22]. As illustrated in Fig. 6, setting the truncated bits to be $10\cdots0(B)$ (with $t-1$ zeros) can enable much higher *PSNR* values, thereby providing a better viewing experience for the same videos in the same surroundings.

B. CONTENT-ADAPTATION VIDEO MEMORY DESIGN

Fig. 7 (a) shows the architecture of the proposed viewer-aware dynamic bit-truncation memory with 512 words × 64 bits, which contains 32kb 6T SRAM bit-cells. To enable viewer-aware bit truncation for LSBs, two different bit-line conditioning circuitries are applied to the memory. The normal bit-line conditioning circuitries have a pre-charge unit, write driver, and sense amplifier, and they are connected to the 4 most significant bits (MSBs) in a byte; the remaining bit-lines contain extra circuitry to enable bit truncation, and they are applied for the 4 LSBs in a byte as shown in Fig. 7 (b).

The truncation controller is shown in Fig. 7 (c). φ *1* and φ *2* are signals generated from peripheral circuitry based on the clock signal. φ *1* controls read and write operations depending on which period it is in; φ *2* controls the pre-charging circuity of the memory. The *sense* signal only turns on for a very short time at the end of the reading operation in order to reduce the power consumption during the read operation. The truncation process is controlled by three external signals.



Inputs			Outputs			
En	B<0>	B<1>	Y3	Y2	Y1	Y0
0	x	x	0	0	0	0
1	0	0	0	0	0	1
1	0	1	0	0	1	1
1	1	0	0	1	1	1
1	1	1	1	1	1	1

FIGURE 7. Content-adaptive video memory. (a) Memory structure. (b) Bit-line conditioning circuitry. (c) Bit truncation controller.

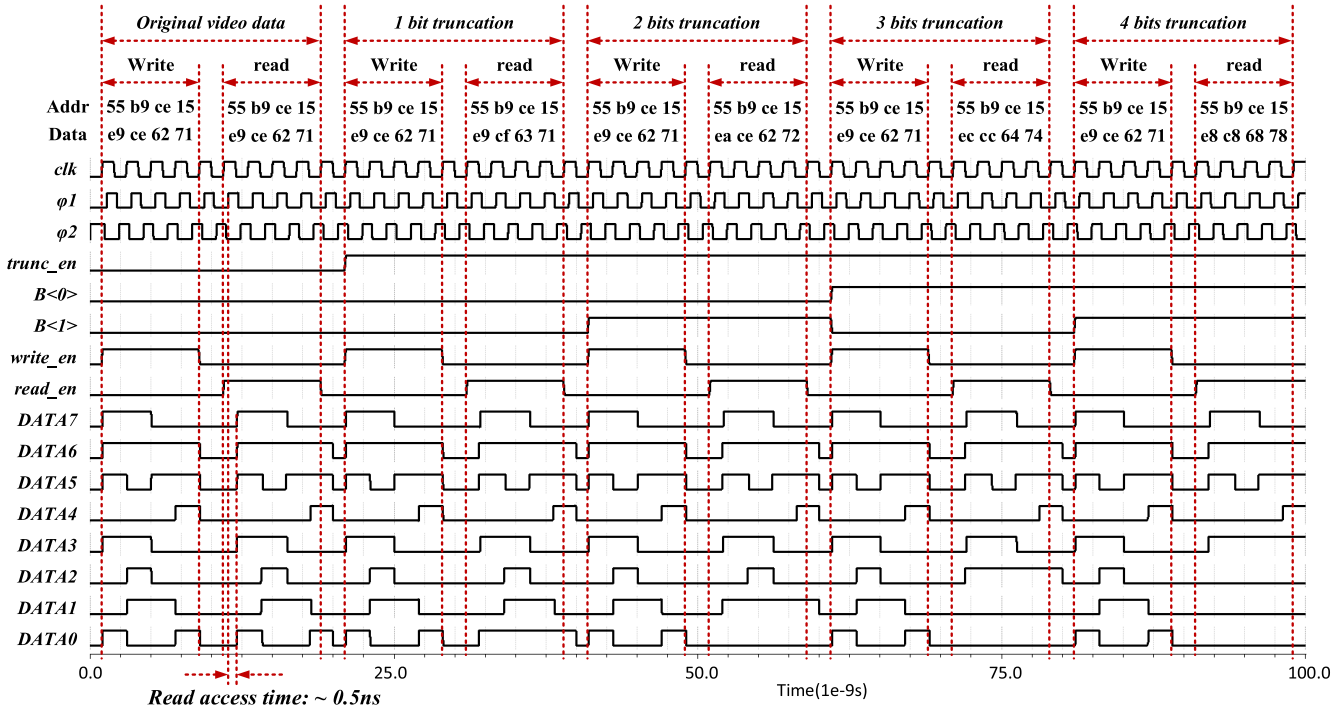


FIGURE 8. Timing diagram. DATA7: MSB; DATA0: LSB.

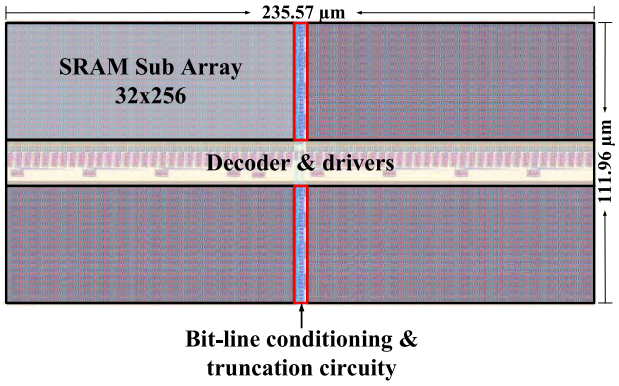


FIGURE 9. Physical layout design.

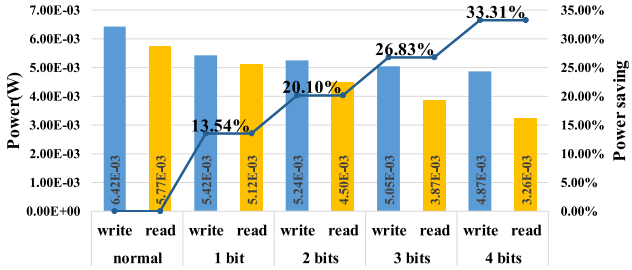


FIGURE 10. Power savings.

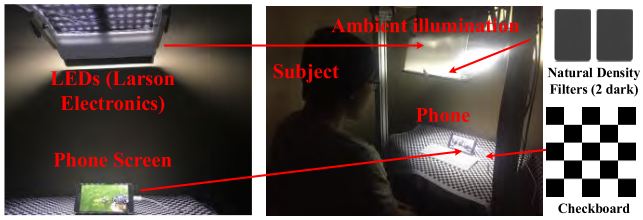


FIGURE 11. Psychological experiment set-up at North Dakota State University Center for Visual and Cognitive Neuroscience.

*trunc_en*controls whether the truncation function is on, and the other two signals, $B<0>$ and $B<1>$,determine how many bits to truncate. t_1 and t_2 are generated from $B<0>$ and $B<1>$ through two decoders. The decoder for t_1 is a normal 2-to-4 decoder. A special 2-to-4 truncation control decoder is applied for generating t_2, and the truth table is also shown in Fig. 7 (c). When t_1 and t_2 are both**0s**, the normal operations are applied; whenever t_1 is **1**, the pre-charging, writing, and reading operations are suspended; on the basis of t_1 being **1**, if t_2 is **1** then the output will be **0**, otherwise the output will be **1**; the data pattern **01** for t_1 and $t_2$will never appear.

The detailed evaluation results including performance, power efficiency, layout, and video quality will be presented in Section VI.

VI. EXPERIMENTAL RESULTS

The proposed memory is implemented based on a 45 nm CMOS technology [28]. In addition to hardware-level implementation and verification, psychological experiments are conducted to test the video output quality from the viewers' perspective.

FIGURE 12. Video quality testing results using the decision tree model.

(a) (b)

FIGURE 13. Output quality of video (***tag wF6IvdXXwc4***): (a) with 3 LSBs truncated using decision tree model and (b) with 2 LSBs truncated using the developed ordinal logistic regression model.

A. SPEED

Fig. 8 shows the timing diagram for the proposed memory. To test the functionality of the memory, the data: *0xe9*, *0xce*, *0x62*, and *0x71*, are written to the addresses: *0x55*, *0xb9*, *0xce*, and *0x15*,respectively, and then read out from the same addresses. For example, during a 3 bit truncation operation, the values read out are: *0xec*, *0xcc*, *0x64*, and *0x74*, which the last 3 LSBs for these values are **100**(*B*). The access delay of the reading operation is about 0.5 ns, which is fast enough to deliver the typical mobile video sequences (11MHz for CIF/QCIF and 72MHz for HD720 [29]).

B. LAYOUT

The layout design for 512 words × 64 bits SRAM with viewer-aware bit truncation is shown in Fig. 9. Only a few gates are added to the bit-line conditioning circuit to enable the truncation function. Also, after careful design, the decoders for truncation controlling can be fit into the free space of the original layout, without introducing additional overhead. The proposed memory consumes only 0.32% more silicon area as compared to the traditional SRAM, which is negligible.

C. POWER SAVINGS

Input patterns that cover all data switching possibilities have been tested for the memory. Normal operation, and 1 to 4 LSB truncations, are simulated based on these input patterns, and the power consumption for each scenario is shown in Fig. 10. As compared to normal operation, the average power consumption of reading and writing operations for 1 to 4 LSB truncations can enable 13.54%, 20.10%, 26.83%, and 33.31% power savings, respectively.

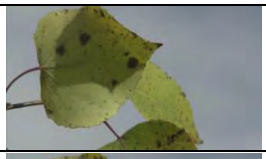
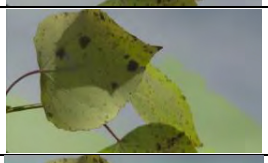
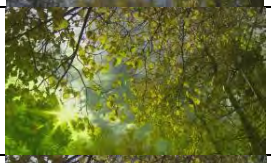
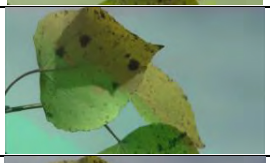
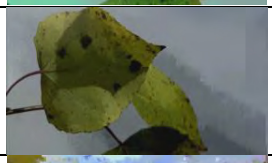
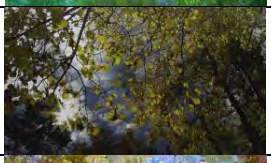
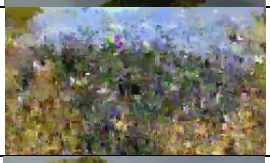
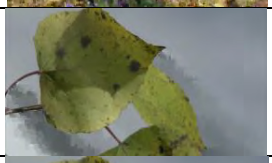
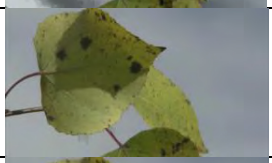
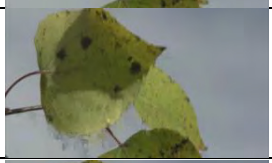
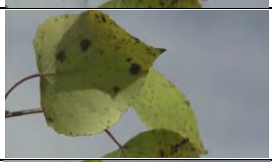
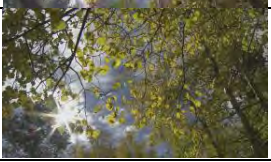
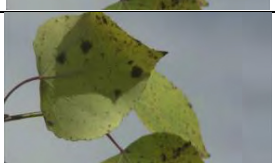
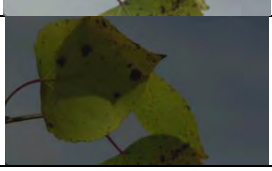
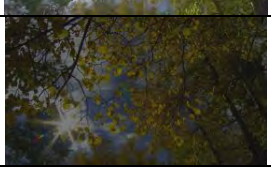
D. VIDEO QUALITY

Finally, in order to verify the effectiveness of our technique on the viewer's experience, we conduct psychological experiments at the North Dakota State University Center for Visual and Cognitive Neuroscience. The psychophysical experiment setup is shown in Fig. 11. The ambient illumination was provided using a rectangular array of 60 high-intensity LEDs capable of emitting a maximum of 64,000 Lux (Larson Electronics, model LEDP5W-60-D-1227-F5.15). An illumination meter (Extech model 401027) was used to accurately measure the ambient illumination of the phone used for testing, a Samsung Galaxy Note 4. In our experiments, we adjust the output of the high-intensity light source using neutral-density filters. The luminance level measured by the illumination meter was approximately 811 Lux, which is a typical indoor light level.

To assess the degree to which observers can accept the truncated videos as compared to the reference videos using the developed models, we collected a total of 20 videos: 10 videos that we classify as having a stationary camera and 10 videos containing a reporter. Each video sample was evaluated at a single quality point, encoded using a constant rate factor of 0 (i.e. lossless compression), had a 640 × 360 resolution, was 10 seconds in length, and was downloaded from [22]. Based on these videos we calculated the average plain MB percentages and used the developed models to predict what the expected amount of acceptable LSBs to truncate would be for different videos. We then created another two versions of each video from the reference, one with the predicted amount of acceptable bits to truncate and another with one bit beyond the predicted acceptable amount. We created sequences of numbers to represent each video and randomized the order they would be presented. During testing, each participant would compare a total of 40 truncated videos to the original, non-truncated version and give their opinion of whether they would consider the video acceptable for viewing on the mobile device.

The testing results for the developed decision tree model are shown in Fig. 12. In our analysis, the plain macroblock percentages, the number of bits truncated, and the video quality metric (VQM) [34] calculation are included for comparison among samples. VQM is one widely used objective video quality metric that has been shown to have a strong correlation to the subjective viewer ratings. When calculating the VQM for each sample, we used the NTIA General Model

TABLE 4. Results of videos with different LSBs truncated in different memories.

Memories	Number of LSBs Truncated	*PSNR* (Max MB %)	*PSNR* (Min MB %)	Max Plain MB % Frame	Min Plain MB % Frame
Original video frames without any truncation	-	-	-		
Chroma Level Cb	5 LSBs truncated	43.0 dB	31.1 dB		
Chroma Level Cr	5 LSBs truncated	36.3 dB	27.0 dB		
Luminosity Level	5 LSBs truncated	18.0 dB	16.7 dB		
Restructured Neighboring	1 LSB truncated	13.7 dB	11.0 dB		
Prediction Mode	1 LSB truncated	23.0 dB	19.5 dB		
Motion Vector X	1 LSB truncated	29.2 dB	13.5 dB		
Motion Vector Y	1 LSB truncated	29.1 dB	13.3 dB		
Reference Macroblock (MB)	**5 LSBs truncated**	**42.8 dB**	**32.8 dB**		
Frame Buffer	**3 LSBs truncated**	**44 dB**	**25.5 dB**		
YUV Display	2 LSBs truncated in each vector	11.8 dB	13.2 dB		

with Full Reference Calibration, which have been standardized by both the ITU and ANSI [35]. The developed decision tree model works well for nearly all of videos. There was only one video, with *tag wF6lvdXXwc4*, out of 20 videos that was considered to not be acceptable by the vast majority of participants. As shown in Fig. 13 (a), this video displayed banding distortion, caused by bit truncation, appearing on the reporter's face; which is likely the viewer's focus point. Due to this particularly noticeable distortion, viewers were less likely to accept the displayed degradation. All other samples were considered acceptable by the majority of the 15 total participants, with the lowest acceptance rate being 73% for the video with *tag 2AQ6rhVhwRc*; another video with banding appearing very close to the viewer's focus point, the kitten playing with a string in the video.

We further compared the results using the ordinal logistic regression to the decision tree model. Those two models achieve the same prediction results for the majority of videos; only 4 out of 20 videos are different. For those 4 videos, decision tree model predicts 3 LSBs truncated, but the ordinal logistic regression model predicts 2 LSBs truncated. One of those 4 videos is the video with tag *wF6lvdXXwc4*; it was the only one that was considered to not be acceptable using the decision tree model. With 2 LSBs truncated predicted by the ordinal logistic regression model, the visual quality is significantly improved, as illustrated in Fig. 13 (b). For the other 3 videos (with tags *Lp3H1XOcKCE*, *dgAu_Wsd7Fo*, and *lcVPxLFlq1c*), the visual output with 3 LSBs truncated are acceptable by the majority of participants. Particularly, for the video with tag dgAu_Wsd7Fo, all of the participants said it was acceptable. From the above analysis, we can conclude that as compared to the decision tree model, the ordinal logistic regression model is a more conservative model which can avoid the worst video quality degradation case, but it may lose energy optimization opportunities for some videos.

Another interesting observation that was made during the video testing process is that if we can detect where the viewer's focus is in different videos (e.g., mobile gaze tracker [30]), we can further remove noticeable degradation in these sensitive areas of videos in the future.

VII. CONCLUSION

During the hardware implementation process, we use a single percentage for the entire video in order to minimize the overhead of the design. In order to better suit the applicability and energy-quality scalability, future research will investigate the capability of calculating the macroblock percentage for each frame. This per frame calculation could allow for real-time adjustment of truncated bits at the cost of additional area overhead. We also intend to expand our number of participants and video samples in order to create a more comprehensive model. Finally, we plan to further study the relationship between the content information described in this paper and the psychophysical human visual system models to better understand what other metrics we can use to support hardware design.

APPENDIX

Table 4 lists the results with LSB truncation in different video memories using the video system shown in Fig. 2. The standard video sequence *aspen_1080p.y4m* [25], which has a wide range of plain MB percentages across different frames, is used for evaluation. The average plain MB percentage was 20.90%; the maximum and minimum were 50.89% at frame #367 and 3.03% at frame #113, respectively. The video was encoded with the following ffmpeg [37] command:

ffmpeg -i aspen_1080p.y4m -profile:v baseline -pixel_format yuv420p -level 3.1 -framerate 30 -preset 1 -cavlc 1 -pix_fmt yuv420p aspen_1080p.264

As the encoded video is processed using the Xilinx Zynq 7020 FPGA based H.264 decoding and display system (Fig. 2), the number of truncated LSBs in each memory (Table I) changes from one to a maximum of five. The truncated bits were set to zeros [8], [21].

As shown in Table IV, truncation in different memories significantly influences the output quality. For example, truncation in chroma memories (Chroma Level Cb and Cr) causes color distortion, while truncation in Luminosity Level memory leads to the banding distortion effect. Additionally, truncation in the restructured neighboring memory results in significant output quality degradation. Among those memories, the reference MB memory and frame buffer can tolerant considerable memory failures. Particularly, for the frame buffer, considering its large size (Table I), the tolerance to three LSBs truncation provides power saving opportunities.

ACKNOWLEDGMENT

(Jonathon Edstrom and Yifu Gong contributed equally to this work.) The authors would like to thank Mr. Enrique Alvarez Vazquez and Mr. Ganesh Padmanabhan from NDSU Center for Visual and Cognitive Neuroscience for their support in the psychological experiments.

REFERENCES

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper*. Accessed: Dec. 1, 2017. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] F. Sampaio, M. Shafique, B. Zatt, S. Bampi, and J. Henkel, "Energy-efficient architecture for advanced video memory," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2014, pp. 132–139.

[3] T. Liu *et al.*, "A 125 μW, fully scalable MPEG-2 and H.264/AVC video decoder for mobile applications," *IEEE J. Solid-State Circuits*, vol. 42, no. 1, pp. 161–169, Jan. 2007.

[4] D. Zhou *et al.*, "A 4 Gpixel/s 8/10b H. 265/HEVC video decoder chip for 8K Ultra HD applications," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2016, pp. 266–267.

[5] M. Zhao, X. Gong, J. Liang, W. Wang, X. Que, and S. Cheng, "QoE-driven cross-layer optimization for wireless dynamic adaptive streaming of scalable videos over HTTP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 451–466, Mar. 2015.

[6] D. Chen, X. Wang, J. Wang, and N. Gong, "VCAS: Viewing context aware power-efficient mobile video embedded memory," in *Proc. 28th IEEE Int. Syst.-Chip Conf. (SOCC)*, Sep. 2015, pp. 333–338.

[7] J. Edstrom *et al.*, "Luminance-adaptive smart video storage system," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 734–737.

[8] D. Chen *et al.*, "Viewer-aware intelligent efficient mobile video embedded memory," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 4, pp. 684–696, Apr. 2018.

[9] O. Hirabayashi *et al.*, "A process-variation-tolerant dual-power-supply SRAM with 0.179 μm^2 cell in 40 nm CMOS using level-programmable wordline driver," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 458–459.
[10] P. Wang *et al.*, "A 45 nm dual-port SRAM with write and read capability enhancement at low voltage," in *Proc. IEEE Int. SOC Conf.*, Sep. 2007, pp. 211–214.
[11] F. Tachibana *et al.*, "A 27% active and 85% standby power reduction in dual-power-supply SRAM using BL power calculator and digitally controllable retention circuit," *IEEE J. Solid-State Circuits*, vol. 49, no. 1, pp. 118–126, Jan. 2014.
[12] T.-H. Kim, J. Liu, and C. H. Kim, "A voltage scalable 0.26 V, 64 kb 8 T SRAM with V_{min} lowering techniques and deep sleep mode," *IEEE J. Solid-State Circuits*, vol. 44, no. 6, pp. 1785–1795, Jun. 2009.
[13] M.-F. Chang, S.-W. Chang, P.-W. Chou, and W.-C. Wu, "A 130 mV SRAM with expanded write and read margins for subthreshold applications," *IEEE J. Solid-State Circuits*, vol. 46, no. 2, pp. 520–529, Feb. 2011.
[14] H. Noguchi *et al.*, "A 10T non-precharge two-port SRAM for 74% power reduction in video processing," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI Circuits*, Mar. 2007, pp. 107–112.
[15] M. K. Qureshi and Z. Chishti, "Operating SECDED-based caches at ultra-low voltage with FLAIR," in *Proc. 43rd Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, Jun. 2013, pp. 1–11.
[16] A. Ansari, S. Feng, S. Gupta, and S. A. Mahlke, "Archipelago: A polymorphic cache design for enabling robust near-threshold operation," in *Proc. IEEE Symp. High Perform. Comput. Archit. (HPCA)*, Feb. 2011, pp. 539–550.
[17] I. J. Chang, D. Mohapatra, and K. Roy, "A priority-based 6T/8T hybrid SRAM architecture for aggressive voltage scaling in video applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 2, pp. 101–112, Feb. 2011.
[18] J. Kwon, I. Lee, and J. Park, "Heterogeneous SRAM cell sizing for low power H.264 applications," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 99, no. 2, pp. 1–10, Feb. 2012.
[19] N. Gong, S. Jiang, A. Challapalli, S. Fernandes, and R. Sridhar, "Ultra-low voltage split-data-aware embedded SRAM for mobile video applications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 12, pp. 883–887, Dec. 2012.
[20] L. Kerofsky, R. Vanam, and Y. Reznik, "Adapting objective video quality metrics to ambient lighting," in *Proc. 7th Int. Workshop Qual. Multimedia Exper. (QoMEX)*, 2015, pp. 1–6.
[21] F. Frustaci, D. Blaauw, D. Sylvester, and M. Alioto, "Approximate SRAMs with dynamic energy-quality management," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 6, pp. 2128–2141, Jun. 2016.
[22] (2017). *YouTube-8M Dataset*. [Online]. Available: https://research.google.com/youtube8m/
[23] M. Shafique *et al.*, "Application-guided power-efficient fault tolerance for H.264 context adaptive variable length coding," *IEEE Trans. Comput.*, vol. 66, no. 4, pp. 560–574, Apr. 2017.
[24] M. Shafique, B. Molkenthin, and J. Henkel, "An HVS-based adaptive computational complexity reduction scheme for H.264/AVC video encoder using prognostic early mode exclusion," in *Proc. Design, Autom. Test Europe Conf. Exhibit. (DATE)*, 2010, pp. 1713–1718.
[25] (2017). *Xiph.org Video Test Media [Derf's Collection]*. [Online]. Available: https://media.xiph.org/video/derf/
[26] (2012). *Methodology for the Sujective Assessment of the Quality of Television Pictures*. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf
[27] *Methodology for the Subjective Assessment of the Qualityof Television Pictures*, document Rec. ITU-R BT.500-13, 2012. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf
[28] *FreePDK45*. Accessed: Dec. 4, 2017. [Online]. Available: http://www.eda.ncsu.edu/wiki/FreePDK45:Contents
[29] J. S. Wang, P. Y. Chang, T. S. Tang, J. W. Chen, and J. I. Guo, "Design of subthreshold SRAMs for energy-efficient quality-scalable video applications," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 1, no. 2, pp. 183–192, Jun. 2011.
[30] Y. Feng, G. Cheung, W.-T. Tan, P. Le Callet, and Y. Ji, "Low-cost eye gaze prediction system for interactive networked video streaming," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 1865–1879, Dec. 2013.
[31] T. Zhao, Q. Liu, and C. W. Chen, "QoE in video transmission: A user experience-driven strategy," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 285–302, 1st Quart., 2017.
[32] M. Ruggiero, A. Bartolini, and L. Benini, "DBS4video: Dynamic luminance backlight scaling based on multi-histogram frame characterization for video streaming application," in *Proc. ACM EMSOFT*, Oct. 2008, pp. 109–118.
[33] C. Yim and A. C. Bovik, "Quality assessment of deblocked images," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 88–98, Jan. 2011.
[34] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312–322, Sep. 2004.
[35] *NTIA General Model (Aka VQM) and Full Reference Calibration Standards*. Accessed: Dec. 9, 2017. [Online]. Available: https://www.its.bldrdoc.gov/resources/video-quality-research/standards/hidden-general-model.aspx
[36] Q. Bin. (2018). *Osen Logic OSD10 H.264 Decoder*. [Online]. Available: http://bbs.eetop.cn/viewthread.php?tid=628991
[37] (2018). *FFmpeg*. [Online]. Available: https://www.ffmpeg.org/

JONATHON EDSTROM received the B.S. degree in computer engineering and the M.S. degree in electrical and computer engineering from North Dakota State University (NDSU), Fargo, ND, USA, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. His researches focus on embedded systems' implementation of low-power mobile video and vision technology.

YIFU GONG received the B.S. degree in electrical engineering from North Dakota State University (NDSU), in 2015, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. His researches focus on low-power very large-scale integration (VLSI) design and embedded vision systems.

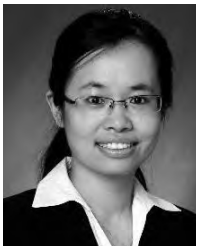
ALI AHMAD HAIDOUS received the B.S. degree in electrical and computer engineering (ECE) from Lipscomb University, in 2015, and the M.Eng. degree in ECE from North Dakota State University (NDSU), where he is currently pursuing the Ph.D. degree. He is also a Software Design Engineer with John Deere. His research primarily focuses on embedded hardware/software system integration for low-power applications.

BRITTNEY HUMPHREY is currently pursuing the bachelor's degree in electrical engineering with North Dakota State University. She is also an Undergraduate Research Assistant. Her research focuses on low-power mobile video storage systems.

MARK E. MCCOURT received the Ph.D. degree from the University of California, Santa Barbara, in 1982, under the supervision of Dr. G. Jacobs and Dr. J. Foley. He received postdoctoral training from Cambridge University with Dr. F. Campbell and from The Australian National University with Dr. G. Henry. He joined the Department of Psychology, North Dakota State University (NDSU), in 1991. He was named as a James A. Meier Professor, in 2001, and as a Dale Hogoboom Professor, in 2009. He is currently the Director of the NIH/NIGMS/IDeA Center of Biomedical Research Excellence (COBRE)-funded NDSU Center for Visual and Cognitive Neuroscience. He was a recipient of the NDSU Waldron Award for Excellence in Research, in 2005.

YIWEN XU received the Ph.D. degree in systems and industrial engineering from The University of Arizona. He is currently an Assistant Professor with the Department of Industrial and Manufacturing Engineering, North Dakota State University, Fargo, ND, USA. His research interests include applied operations research (especially probabilistic network optimization and applied integer programming) and reliability engineering.

JINHUI WANG (M'13) received the B.E. degree in electrical engineering from Hebei University, China, and the Ph.D. degree in electrical engineering from the Beijing University of Technology, Beijing, China. He is currently an Associate Professor with the Department of Electrical and Computer, University of South Alabama, Mobile, AL, USA. He has authored or coauthored more than 120 publications and holds 20 patents in the emerging semiconductor technologies. His research interests include neuromorphic computing, low-power, high-performance, and reliable integrated circuit design, 3-D IC, and thermal issue solution in very large-scale integration (VLSI).



NA GONG (M'13) received the B.E. degree in electrical engineering, in 2004, the M.E. degree in microelectronics from Hebei University, China, in 2007, and the Ph.D. degree in computer science and engineering from The State University of New York, Buffalo, in 2013. She is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of South Alabama, Mobile, AL, USA. Her research interests include power-efficient computing circuits and systems, video memory optimization, and neuromorphic hardware.