VIEWER-AWARE INTELLIGENT MOBILE VIDEO SYSTEM FOR PROLONGED

BATTERY LIFE

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Peng Gao

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Electrical and Computer Engineering

October 2017

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Viewer-Aware Intelligent Mobile Video System for Prolonged Battery Life

By

Peng Gao

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota

State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Na Gong

Chair

Jinhui Wang

Zhibin Lin

Approved:

11/07/2017	Benjamin Braaten
Date	Department Chair

ABSTRACT

In the modern society, mobile is gradually going to become all about video streaming. The main reasons of video growth are mobile devices such as smartphones and tablets which enable people to have access to videos they would like to watch at anywhere and anytime. However, due to the large video data size and intensive computation, video processing leads to a huge power consumption. Mobile system designers typically focus on hardware-level power optimization techniques without considering how hardware performance interfaces with viewer experience.

In my research, I investigated how viewing context factors affect mobile viewing experience. Furthermore, a viewer-aware intelligent mobile video system was designed to optimize power efficiency automatically in real-time according to the viewing context and maintain the same viewing experience. Our research opened a door for developments of future viewer-aware mobile system design, accelerating low-cost mobile devices with longer battery life.

ACKNOWLEDGEMENTS

DEDICATION

To my grandfather and grandmother, Peimin Chu and Peiling Cong.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SRAM ....Static Random Access Memory

VCAS ....Viewing Context Aware SRAM

QoE ....Quality of Experiences

LSB ....Least Significant Bits

MSB ....Most Significant Bits

IC..Integrated Circuit

I^2C ....Inter-Integrated Circuit

PSNR..Peak Signal-to-Noise Ratio

ADC ....Analog-to-Digital Conversion

CIF ....Common Intermediate Format

FHD..Full High Definition

ITU...International Telecommunication Union

1. INTRODUCTION

1.1. Trends in Designing Power Efficient Mobile System

In the past decade, mobile devices such as smartphones and tablets are becoming day-to-day device of choice for users of all ages. According to research from Cisco, two-thirds of global mobile data traffic will be driven by video by 2017 [1]. One of the most annoying issue with mobile devices is their short battery life. Video storage consumes a huge amount of power. In video processing applications, the high frequency of video data read and write makes video memory become one of the most power consuming part in a mobile device. Research shows that embedded SRAM contributes to over 30% of the system power consumption of a mobile device [2, 3]. Researchers are working towards designing new low power devices due to the increasing amount of portable device usage [4], but viewer experience is rarely considered when developers are trying to improve hardware design to save power.

The unique contribution of this thesis is the proposed viewer-aware intelligent mobile video system. This system can be used to reduce the power consumption of video processing. By combining viewer experience, viewing context factors and mobile device hardware system, this system can significantly reduce power consumption while maintaining the same subjective video quality. A machine learning algorithm was also applied to analyze viewer's video quality tolerance threshold, which makes the proposed system be able to accurately determine the most power efficient video output quality.

The second chapter will introduce the viewer-aware intelligent mobile video system. The use of three different types of sensors including luminance sensor, proximity sensor and accelerometer can determine the different types of viewing contexts. By analyzing the viewing contexts, the system will enable power saving techniques accordingly using bit truncation and

1

voltage scaling to reduce the power consumption of video processing while maintaining the same subjective video quality.

The third chapter describes my contribution to Viewer-Aware Intelligent Efficient Mobile Video Embedded Memory. In this research project, I conducted video quality tests for the proposed Viewing Context Aware Storage (VCAS). By following the recommendations from ITU [19], I tested the video output acceptability from VCAS. In this testing, 6 CIF videos and 8 Full HD (1080p) videos were used to verify the proposed power saving method. The results clearly indicates that VCAS provides reliable video output quality while saving 37.5% and 50% power in overcast and sunlight viewing contexts. Compared to other existing power saving techniques, the VCAS has obvious advantages in area overhead and reduced power consumption.

1.2. Related Work

There is a rich body of literature in circuit and architectural techniques for power reduction for video processing systems. Psychophysical researchers have conducted researches on the impact of visual context on human visual system (HVS), hardware designers are mainly focusing on improving memory design from a hardware perspective to support high video power consumption.

1.2.1. Related Work on Quality of Experience

Psychophysical researchers have conducted researches on the impact of visual context. Many results showed that the viewing contexts that influence the mobile video watching experience fall into three major aspects: viewer movement, viewing distance and ambient luminance. Among these three aspects, ambient luminance appears to be the most effective factor [5-8]. The International Telecommunication Union has developed multiple mature and scientific methods for subjective video quality testing. These methods can be used for many

different purposes including, but not limited to, selection of algorithms, ranking of audiovisual system performance and evaluation of the quality level during an audiovisual connection [9].

1.2.2. Related Work on Low Power Memory Designs

In the past few years, many memory designs have been developed for mobile video applications. The property of SRAM with more transistors provides high reliability while in low voltage makes it the most popular video data memory. A memory's power-quality tradeoff is usually set during its design process, which leads to a limited power efficiency for different video applications. Many existing studies chose to adapt more than 6T bit cells to reduce power consumption. Recently, 6T/8T and 8T/10T hybrid structures are introduced to optimize the power efficiency of mobile video streaming [10, 11]. Also, asymmetric 7T [27], column-decoupled 8T [28], read-disturb-free 9T [29], 10T cells [30], and bit-interleaving 12T cells [31]. On the other hand, there are also many existing studies focused on adjustment of bit cell voltage. In 2015, a voltage-scaled SRAM design that can dynamically tradeoff between power and video quality was introduced, but the encoder and decoder circuits of this technique leads to a large computational complexity and silicon area [12]. There are also techniques including boosted word-line voltage [32], dual-rail supply [33], and read-modify-write or write-back [34]. However, these improvements in memory power efficiency often comes with increase in design complexity, large silicon area and power penalty for voltage regulator or boosting circuits.

1.3. Research Objective

The existing studies introduced many power efficient hardware designs. But these designs are missing on how viewer's QoE can affect the power efficiency of mobile devices. The goal in this research is to connect viewer experience, viewing context factors and the mobile device hardware system using machine learning algorithm to reduce the power consumption of

video processing. In this thesis, two popular low-power techniques are applied: bit truncation and voltage scaling. These two low-power techniques will lead to objective video quality degradation while reducing the power consumption, but the proposed intelligent mobile video system can analyze the viewing context and then find the most power efficient solution according to the viewing context using a decision tree model while maintaining the same perceptual video quality, therefore, the video degradation will not be noticed by viewer.
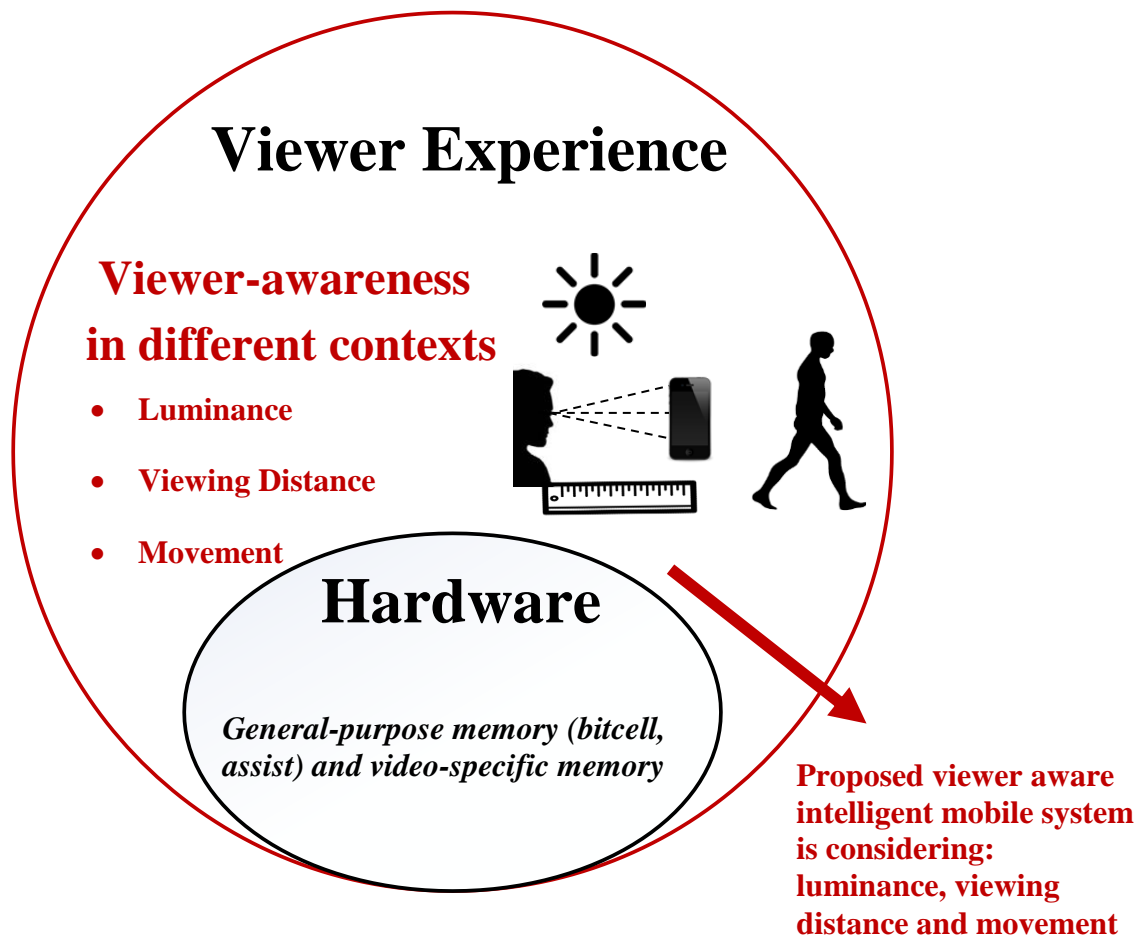

Figure 1. Thesis focus

1.4. Contribution in the Thesis

The existing power efficient techniques are mainly focusing on hardware optimization, but more power saving can be achieved by combining hardware modifications with the

surrounding viewing context. In this thesis, a viewer-aware intelligent mobile video system is proposed to maximize the power efficiency. Compared to other existing studies, this system took luminance, viewing distance and movement into consideration. Three different types of sensors are integrated to the system to collect viewing context data, including luminance sensor, accelerometer and proximity sensor. By using these three types of sensors we can collect viewers' subjective video quality acceptability data and train a high accuracy decision tree model. With the decision tree model, this system can achieve a real-time high accuracy video quality adaptation according to the surrounding viewing context. This adaptation can bring low quality videos to situations where the video quality degradation cannot be noticed by viewer in order to improve battery life.

Also, a traditional SRAM design with control units to disable and enable LSBs is introduced in this thesis. By using the control unit, the memory can disable 0, 3, or 4 LSBs' Write Enable and Read Enable circuit to achieve bit truncation purpose. While under different luminance levels, this system will truncated different amounts of bit cells to reduce unnecessary power consumption. This system exhibits the lowest implementation cost (<0.01% area overhead) with dynamic power-quality tradeoff.

Compared to other existing solutions, these two proposed techniques combined viewer experience with hardware modifications which makes it orthogonal to the existing solutions. Therefore, these two techniques can be applied with existing techniques simultaneously to further improve the power efficiency of mobile video applications. In the future when HEVC/H.265 and other memory designs become more popular, videos with higher resolutions such as 4k and 8k videos will receive more attention. That will make the proposed techniques have an even better performance.

2. PROPOSED VIEWER-AWARE INTELLIGENT MOBILE VIDEO SYSTEM

This chapter introduces the viewer-aware intelligent mobile video system. This system is designed to maximize the memory power efficiency. By using the trained decision tree model from viewers' QoE experiment data collection result, this system will be able to fully take advantage from the tree most effective viewing context factors.

2.1. Context Influence on Mobile User Experience

2.1.1. Video Decoding and Processing

The proposed viewer-aware intelligent mobile video system is applied to the traditional H.264 video codec, which is one of the most popular video codec standards in mobile multimedia communications. But this technology is not only limited to H.264 video codec, any video codec that decodes a video to raw video data format can use this technology to improve the power efficiency of video processing. The new video codec H.265/High Efficiency Video Codec (HEVC) is becoming more and more popular as people's demand and chip fabrication technology increase, our technology can provide a much higher power savings due to the large video frame buffer memory space required by high resolution videos [13].

On mobile devices, most of the time people are using the video decoding function instead of the encoding function, because of the encoding process is usually done by video providers such as YouTube and Netflix. Fig. 2 shows the block diagram of H.264 decoder. In video decoding process, the reference frame memory is accessed frequently. After each frame is decoded, a new frame will be written into the memory and ready to be read out as the reference frame for next frame. Fig. 2 also shows the typical frame data stored in embedded memory. Now days, the most popular video formats are MP4, AVI, and WMV etc., these video formats are encoded video formats. But before a video frame is shown on the screen, the video data needs to

be decoded first. For a decoded YUV 4:2:0 format video data [14], each pixel contains 8-bit of

Luma data and 4-bit of Chroma data. To store a whole frame of this video sample, it requires 176

x144 (25344) x 8 bits of SRAM to store Luma (Y) data and 25344 x 4 bits of SRAM to store

Chroma (U, V) data. Hence, there are astronomical amounts of read and write operations in

memory during the decoding process, and this is the reason why video processing drains the

battery life of a mobile device so fast.

From the mathematics perspective, every bit of these video data does not represent same

amount of information. For an 8-bit number, the MSB can represent 128 in decimal, but the LSB

can only represent 1 in decimal. Therefore, bit truncation and voltage scaling can be applied on

LSBs to reduce power consumption without changing the video output quality too much.

For bit truncation technique, the amount of saved power can be calculated using the

power consumption model of mobile video memory:

$$P = P_w + P_r \tag{1}$$

$$P_w = \sum_{k=0}^{7} \sum_{\substack{i=0,1 \\ j=0,1}} [F_k(i,j) \cdot P_{wk}(i,j)] \tag{2}$$

$$P_r = \sum_{k=0}^{7} \sum_{i=0,1} [F_k(i) \cdot P_{rk}(i)] \tag{3}$$

Where P is the power consumption; k is the bit number; i and j are old and new values stored in

an SRAM. F indicates the bit change (switching) probability from i to j. According to this power

consumption model, by completely shutting down LSBs, the power consumption of video

processing can be significantly reduced while maintaining the more important information
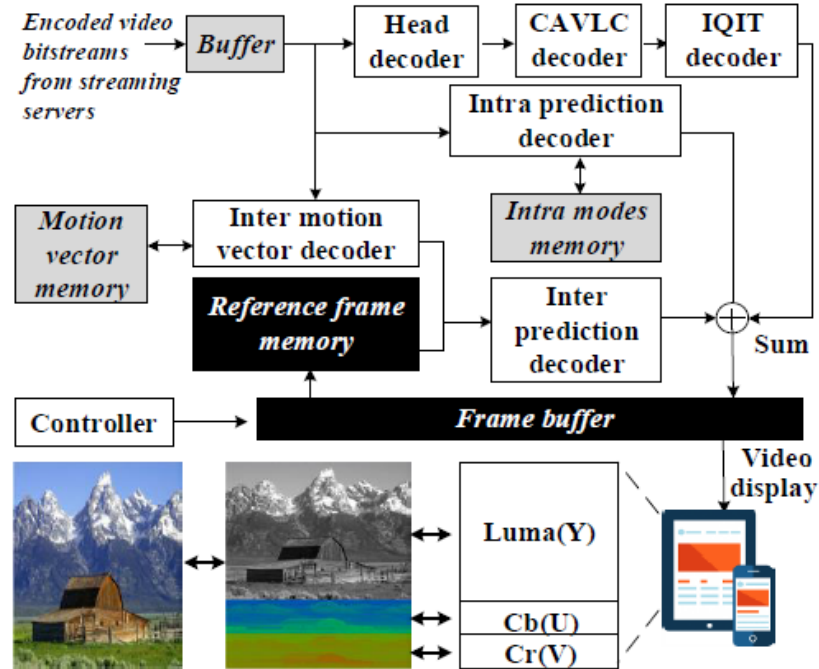
correct.

Figure 2. H.264 video decoder and embedded memory

2.1.2. Impact of Viewing Context Influence on Mobile Videos

Many existing studies have shown that perceptual video QoE is closely related to the viewing context. The three main factors that affects video quality the most are: body movements, viewing distance and ambient illumination. Among these three factors, ambient illumination has the biggest impact on video QoE. When the surrounding environment changes, viewer's tolerance of video quality will also increase. For example, when we look at our phone screen under direct sunlight, the screen becomes blurry because of the strong sunlight projected on the screen, and this effect makes viewer's tolerance on video quality increase (Fig. 3). Similarly, when viewer is watching mobile device screen from a long distance or watching while walking, the degradation of video quality will become harder to notice.
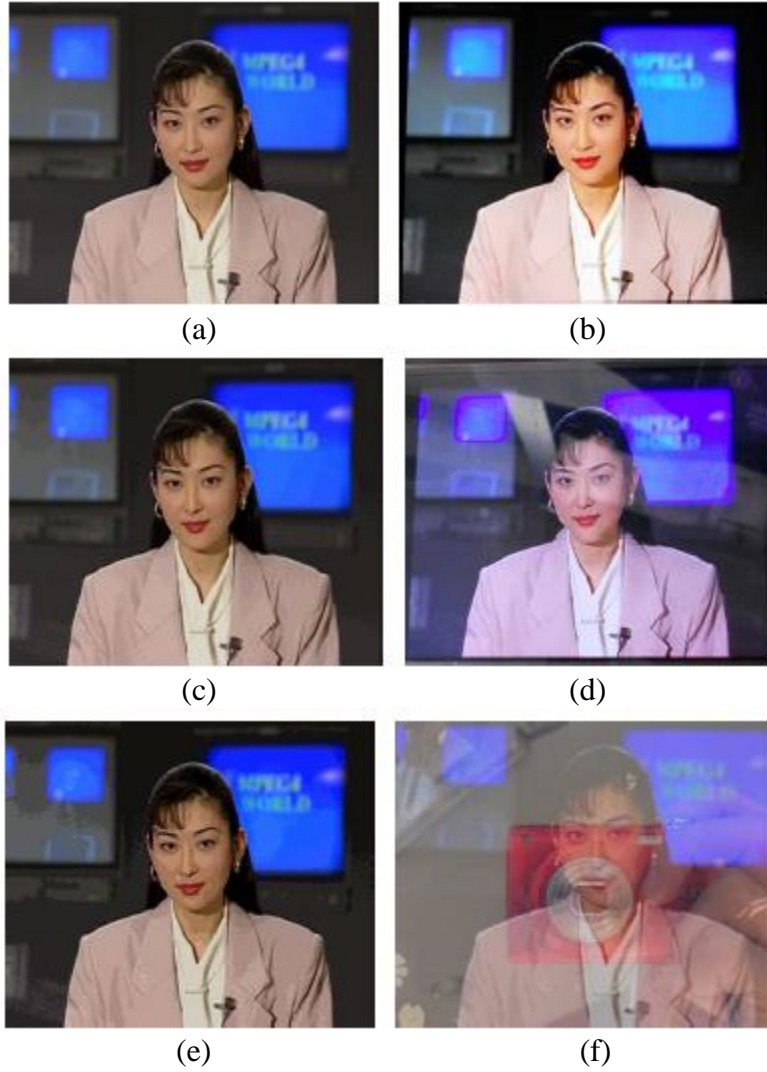
8

Figure 3. Video output images from screenshot (a, c, e) and camera (b, d, f)

Normally researchers use PSNR as a metric to measure the quality of a video which is defined as:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \tag{4}$$

Where MSE is the mean square error between the original videos and the degraded videos, expressed as:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [Org(i,j) - Deg(i,j)]^2 \tag{5}$$

9

But in subjective video QoE, PSNR does not provide any useful information anymore due to the fact that videos with different PSNR can provide a very different QoE under different viewing contexts [14]. Our approach is to use this fact to improve battery life. By analyzing the surrounding viewing context and reducing the voltage source on SRAM memory, the video quality degradation caused by reduced voltage will be unnoticeable to viewers. In our research, machine learning was also used to find the subjective QoE of different video qualities under different viewing contexts to obtain an accurate and reliable data set for model training.

2.2. Embedded System Design and Implementation

2.2.1. Embedded Hardware Setup

An embedded system was implemented using two development boards. As shown in Fig. 4, the board on the left is an Arduino Uno board [15] which is used to collect all the context information and send the processed data to the Odroid-XU3 board [16] using serial communication. The Odroix-XU3 board contains processed video samples for display.

The light sensor, accelerometer and proximity sensor are connected to the Arduino Uno board. The light sensor can be used to detect luminance value, the accelerometer is used to measure the movement by collecting acceleration in X, Y and Z axis, and the proximity sensor can measure the distance between viewer and the screen.

These three sensors were chosen because they are commonly used in the modern smartphones. The modern smartphones are designed to be "smart". The light sensor on smartphone can be used to automatically adjust screen brightness, accelerometer can be used to record user's fitness activities, and proximity sensor will prevent user from accidentally touching the buttons on the phone while making a phone call.
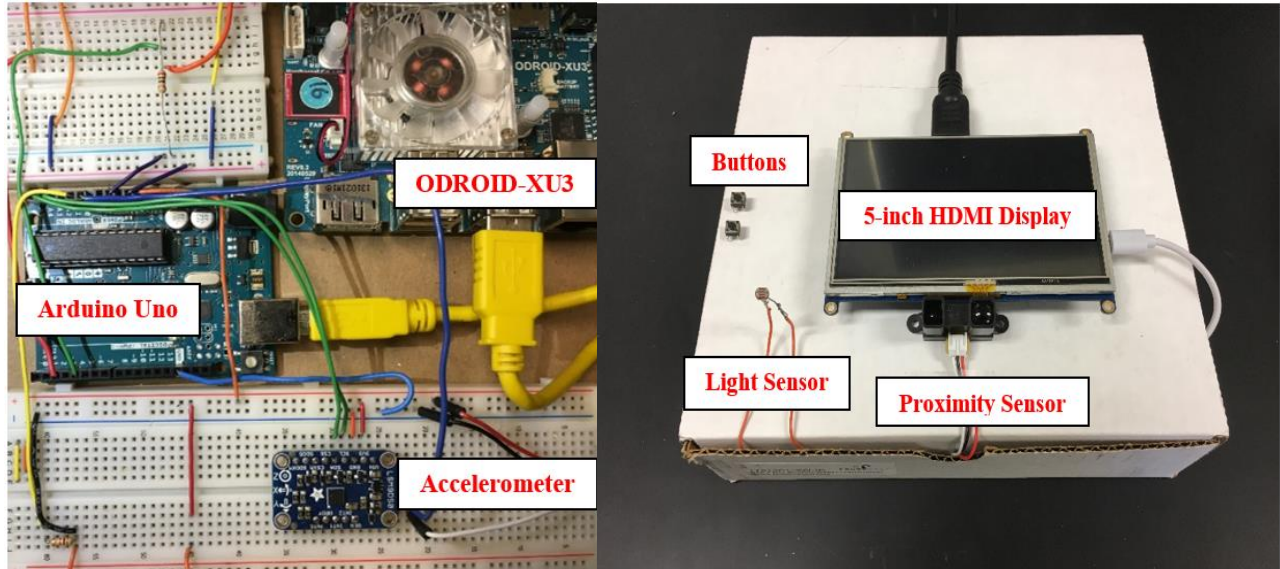
Figure 4. The embedded system device setup

On the outside of the device, a 5-inch HDMI screen is mounted as display. Two buttons on the left side of the device is serving the purpose of collecting feedback data from viewers. The button on top will be pressed when the video sample quality is acceptable for viewer and the button at the bottom is for unacceptable video samples. The light sensor and proximity sensor are located very close to the screen to monitor the surrounding luminance and viewing distance.

2.2.2. Embedded Software Utilization

The Arduino Uno board collects all the sensor data from the ADC pins on the board over I^2C. These ADC pins are updated frequently to collect the sensor information in real-time. In each cycle, the Arduino Uno board reads the analog voltage value from these three sensors and convert them into digital values, after each conversion, Arduino Uno board sends all the data to the Odroid-XU3 board.

To emulate the effect of bit truncation and voltage scaling, "errors" were injected into videos to emulate the result of bit truncation and voltage scaling. These errors were injected by modifying the LSBs in all pixel data. For bit truncation technique, The LSBs were overwrote to zeros. For voltage scaling technique, the data was passed through a digital filter which has 1%

chance to flip the LSBs. These two technologies provides different amounts of power saving and also different visual effects on the video samples.

These two types of techniques brings different types of video quality degradation. Bit truncation technique generates shades on the video frames which causes large PSNR degradation but also provides more power compared to voltage scaling. Voltage scaling generates noise dots on the video frames which causes low PSNR degradation and lower power saving compared to bit truncation technique (Fig. 5). Voltage scaling technique is always applied after bit truncation for additional power saving due to the fact that bit truncation is completely shutting down the memory columns and voltage scaling is only reducing the power supply voltage for memory columns.
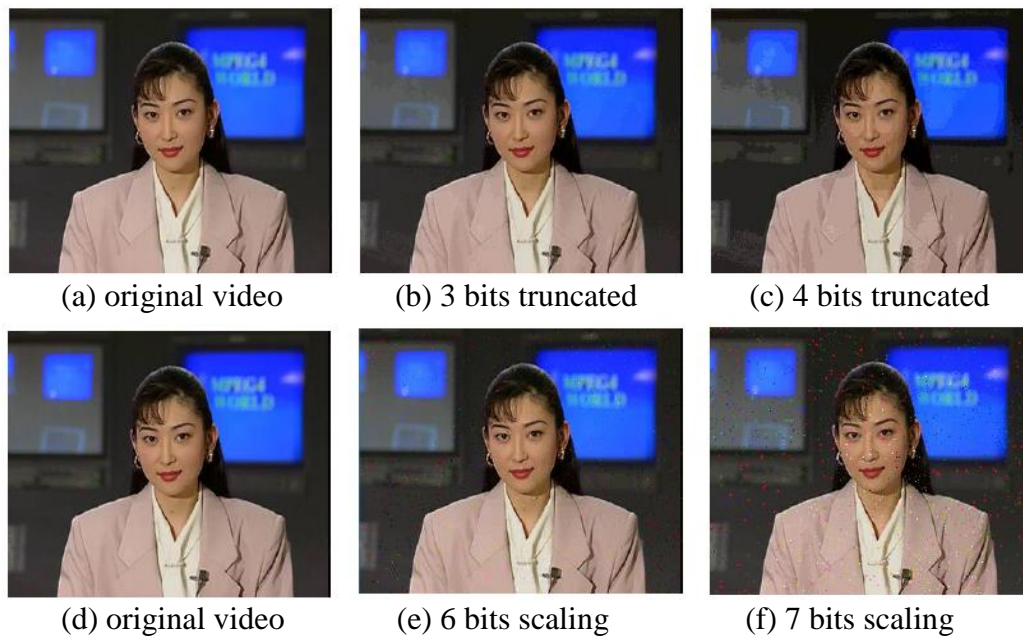

(a) original video	(b) 3 bits truncated	(c) 4 bits truncated
(d) original video	(e) 6 bits scaling	(f) 7 bits scaling

Figure 5. Different impacts of bit truncation and voltage scaling

2.3. Context Adaptation Decision Tree Model

2.3.1. Background of Decision Tree Classifier

Researcher have been using machine learning for data classification, regression and density estimation in a variety of application areas such as bioinformatics, speech recognition, spam detection and computer vision [25]. Among all the machine learning algorithms, decision tree is one of the most popular tool researchers choose to use now days. Decision tree is commonly used in operations research, specifically in decision analysis. The basic idea of decision tree is to break up a complex decision into a union of several simpler decisions, hoping the final solution obtained this way would resemble the intended desired solution [26].

2.3.2. Viewer-aware Mobile Video System Model Implementation

By following the flowchart shown in Fig. 6, 15 viewers were invited to participate in this data collection process. During the data collection process, the Odroid-XU3 board will display a random video sample with error injected. Viewer was asked to watch these video samples and select whether the video sample is acceptable or not under the current viewing context. The viewing context data will be collected simultaneously using the three onboard sensors. Our initial testing result shows that the limit for bit truncation is 4 bits and 5 bits for voltage scaling. The processed videos include: 2 bits truncation, 3 bits truncation and 4 bits truncation. Three videos were selected from [17] as our sample videos (Fig. 7). All these three samples are CIF format, with the same resolution 352 x 288. After injecting errors into these video samples, there are 72 different video qualities based on the bit truncation and voltage scaling methods. Each video sample is 6 seconds long and every video sample was played 5 times during each viewer's test. In this experiment, 15 people were invited from North Dakota State University, all of them have normal visions and participated the testing under my instructions.

13

After the data collection process, the Statistic and Machine Learning Toolbox was used to train the model using the collected. A decision tree model was developed for determining how many bits should be truncated or scaled. Our goal of training the model is to make each processed videos able to be matched up with at least one viewing context.
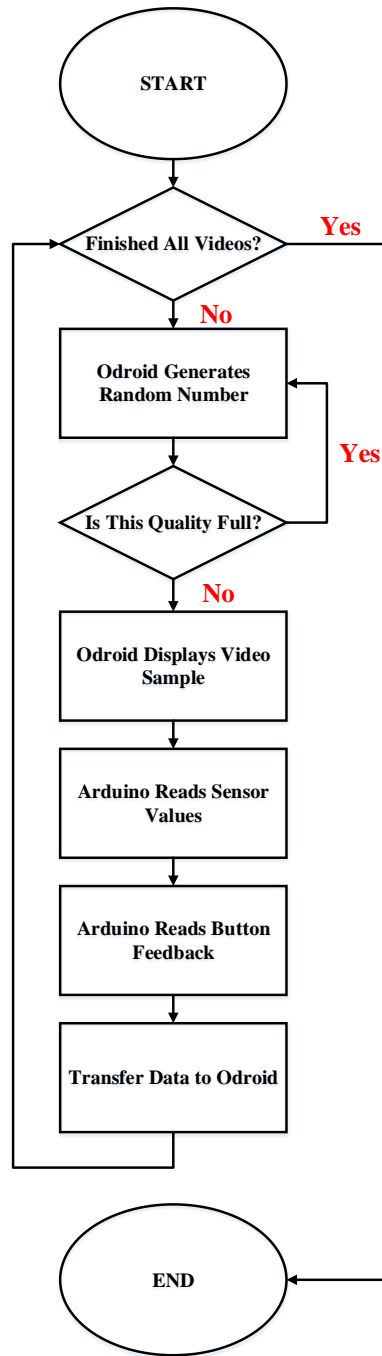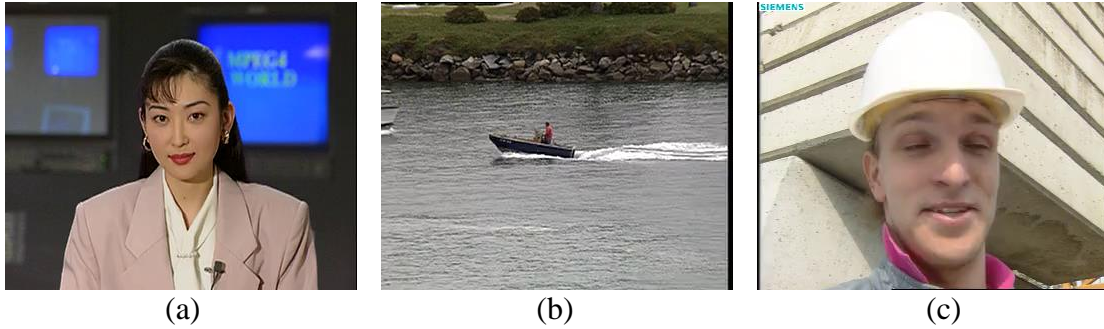

Figure 6. Flowchart of data collection process

14

| (a) | (b) | (c) |

Figure 7. Three video samples used for data collection

2.4. Experimental Results

2.4.1. Results

Fig. 8 shows the overall video testing results. First I verified the existing studies' conclusion which shows the three factors that affects perceptual video QoE the most. The results perfectly matched the existing studies, luminance has the most significant impact on video QoE. The graph of movement and viewing distance does not show a strong relationship between these two factors and video quality acceptability, but we can still see the acceptability of video quality becomes much high when the values of these two factors approached to their maximum. Note that, to accelerate the processing speed of our system, all the data values we collected in this experiment are raw data values.
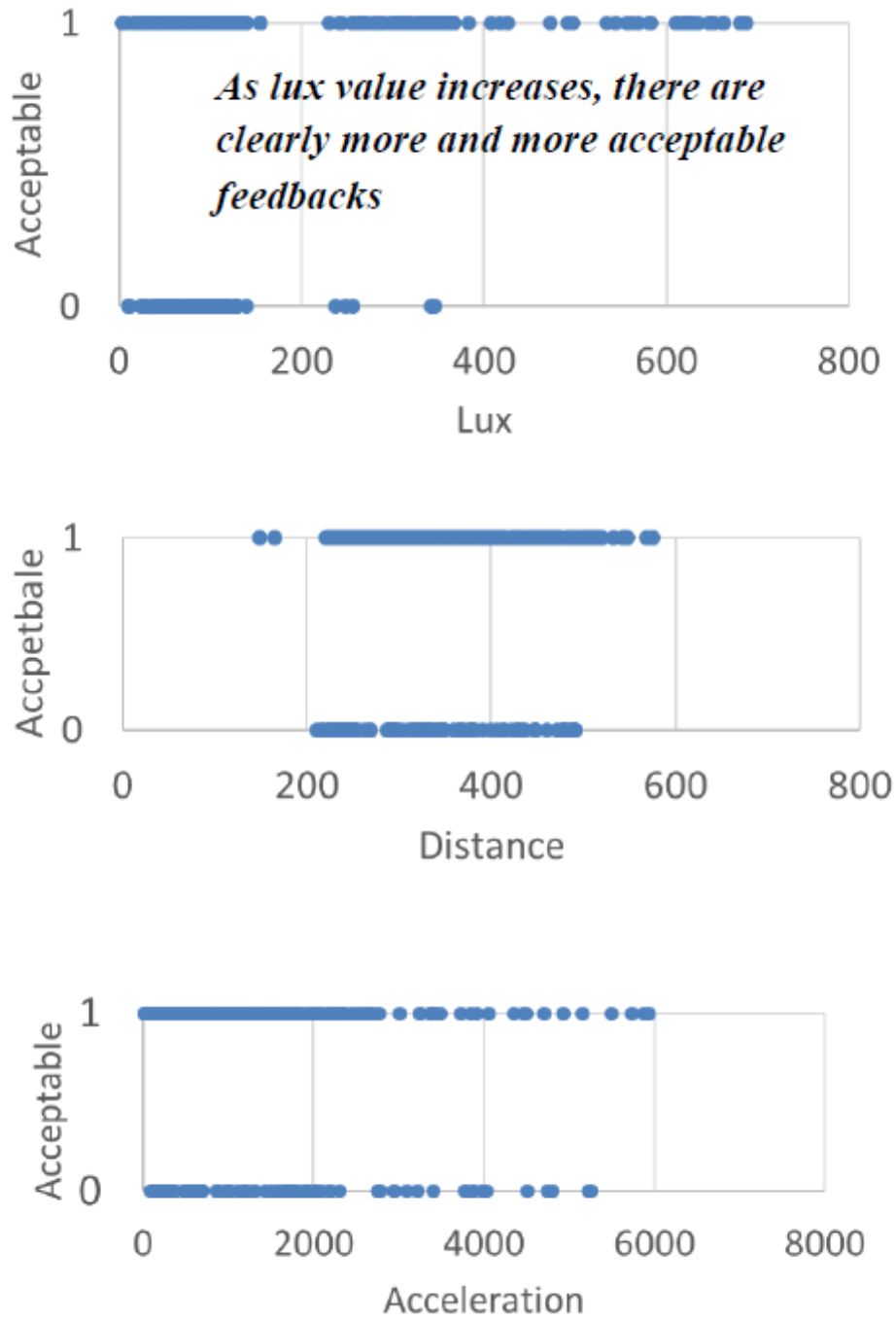
As lux value increases, there are clearly more and more acceptable feedbacks

Figure 8. Video data collection results. Each point on the graph represents one feedback

Table 1. Data sample used for model training

Acceleration	Distance	Luminance	Acceptability
41	400	492	1
54	545	355	1
100	295	57	1
111	345	28	0
115	245	47	0
138	293	33	0
164	239	44	0
166	289	683	1
169	366	78	0
192	217	102	0
213	493	360	1
224	251	90	0
241	414	95	0
244	271	544	1

Because of the impact of luminance is too strong, luminance is in a total dominance in the collected data set. The characteristics of movement and viewing distance are not obvious enough for model training. Viewers' opinion was mainly influenced by luminance, where movement and viewing distance does not contribute enough weight in this model training. I decided to train a model for bit truncation and luminance only. Therefore, the relationship between these two most significant predict factors can be easily observed. In the future, this model may be further improved by adding voltage scaling according to movement and viewing distance due to the fact that these factors all have a lesser weight in model training.

Another data collection experiment was conducted with 15 participants. In this experiment, the three same video samples were used to apply bit truncation effects. All of these three samples were processed with the effect of two, three and four bits truncated. Each video lasts 6 seconds long and will be played five times each.
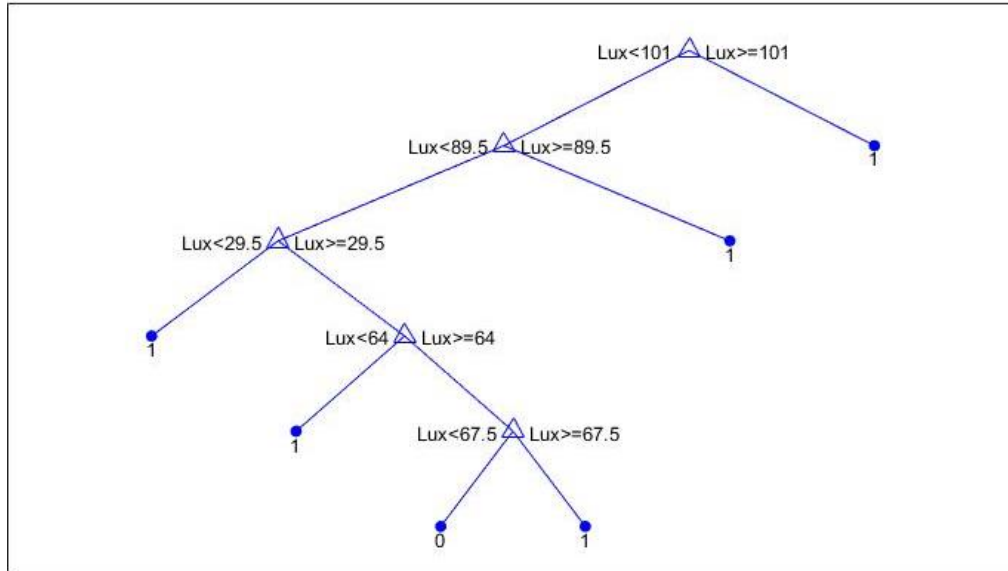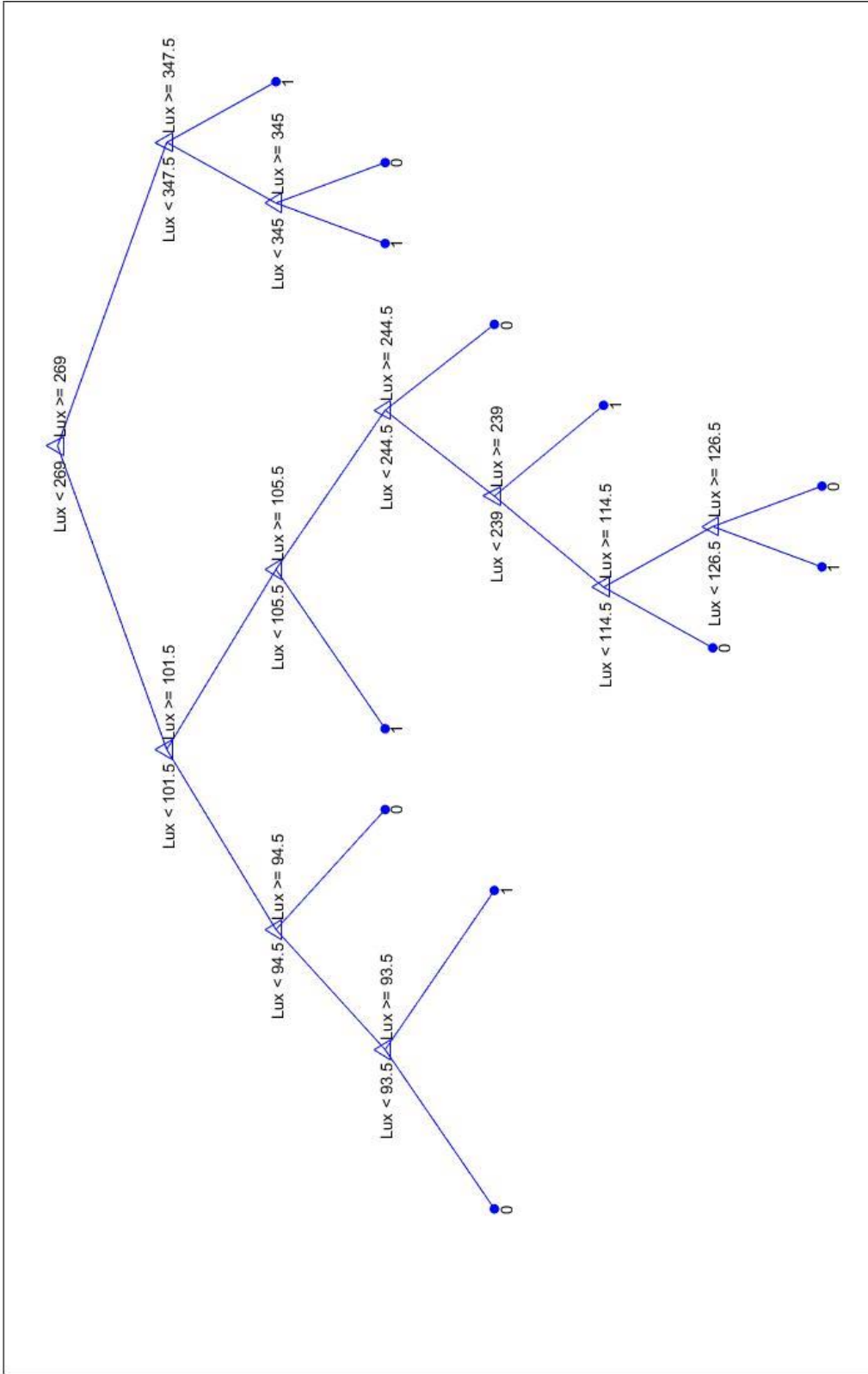
(a) two bits truncation
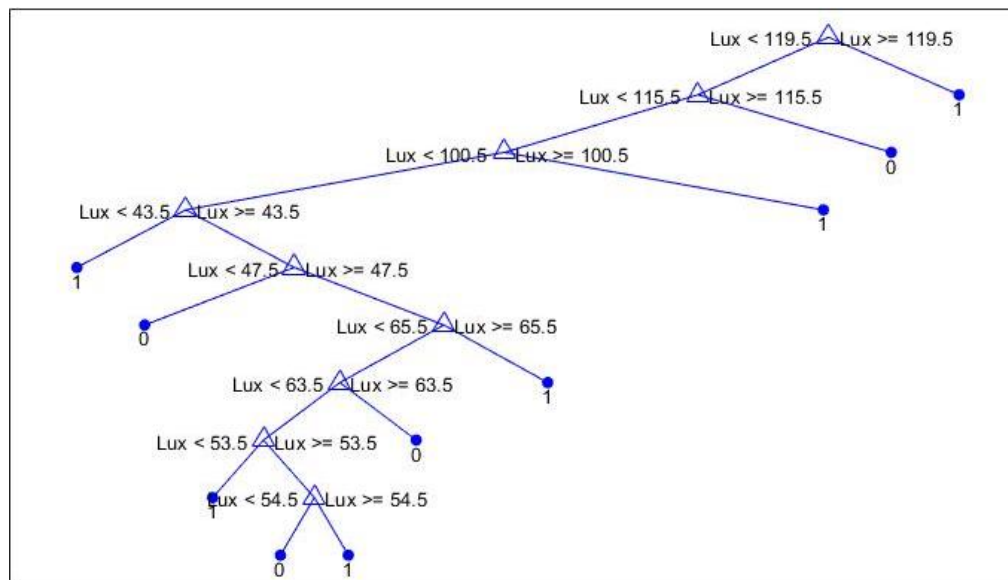
(b) three bits truncation

Figure 9. Decision tree model for bit truncation and luminance

Figure 9. Decision tree model for bit truncation and luminance (continued)

As shown in Fig. 9, these tree decision tree models are developed for different bit truncation levels. As shown, for good video quality (two bits truncated), by looking at the highest level of decision branch, our model shows that when lux is greater than 101, the output quality is acceptable. For medium video quality (three bits truncated), the model shows that when lux is greater than 119.5, the video quality is acceptable. And for bad quality video (four bits truncated), it requires lux value to be greater than 269 to make the video degradation invisible.

After pruning the decision tree model, the first level of the tree was taken into consideration due to its dominance in separation (Fig. 10). By analyzing the luminance value in real-time, the system can select how many bits to truncate using the programmed model.
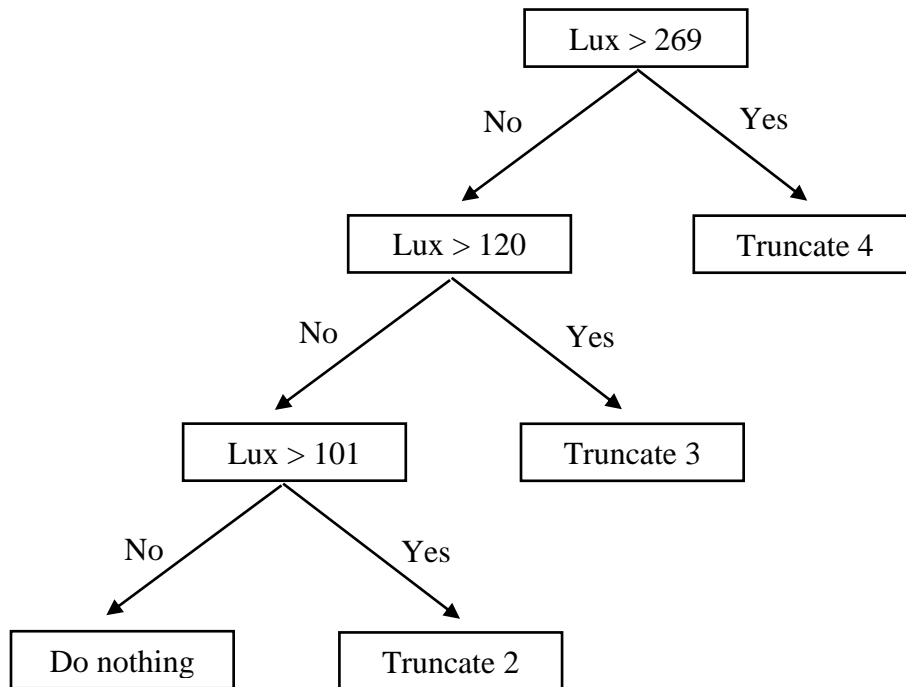
Figure 10. Simplified decision tree model

2.4.2. Conclusion of Viewer-Aware Intelligent Mobile Video System

The existing studies found that luminance, viewing distance and movement are the three main factors effecting perceptual video quality, but no existing study tried to quantify the

relationship between video quality and the value of these three factors. Also, in the existing studies, there are researchers introduced voltage scaling and bit truncation techniques for power saving purposes. In our research, I combined these two technologies and applied them on mobile devices. The machine learning algorithm can accurately find the best solution at different viewing conditions. Compared with existing studies, the system I designed increased the elements that are taken into consideration when calculating power saving, including light, movement and viewing distance. Our research provides an accurate decision tree model that can be directly applied on modern mobile devices since all the sensors that are used in our experiments are used by almost all smartphones and tablets.

3. PROPOSED VIDEO CONTEXT AWARE SYSTEM[1]

In this chapter, a Video Context Aware System (VCAS) is introduced to reduce the power consumption. This system analyzes user's viewing context and enables hardware power saving techniques to bring power saving to a new level. The VCAS also uses bit truncation technique, compared to viewer-aware intelligent mobile video system, the VCAS focus more on the effectiveness of luminance and hardware chip level improvements. These two research projects were conducted at the same time.

3.1. VCAS Technique

3.1.1. VCAS Hardware Implementation

We designed a SRAM architecture of the proposed VACS. The SRAM size is 32 kbits structured with 256 word lines times 32 bits. The VCAS control unit consists of a Write Enable (WE) control circuit and Read Enable (RE) control circuit (Fig. 11).

In this research project, we considered dark, overcast and sunlight viewing contexts. The two-bit control circuit will enable bit truncation in the following ways:

$$\{Bit2, Bit1\} \begin{cases} 00: \text{in dark, use original video data} \\ 01: \text{in overcast, disable 3 LSB WE (or RE) and PRE} \\ 10: \text{in sunlight, disable 4 LSB WE (or RE) and PRE} \end{cases}$$

---

[1] The material in this chapter was co-authored by Peng Gao, Dongliang Chen, Jonathon Edstrom, and Yifu Gong. Peng Gao was responsible for video processing, subjective video quality testing, testing result analysis and also part of hardware design.
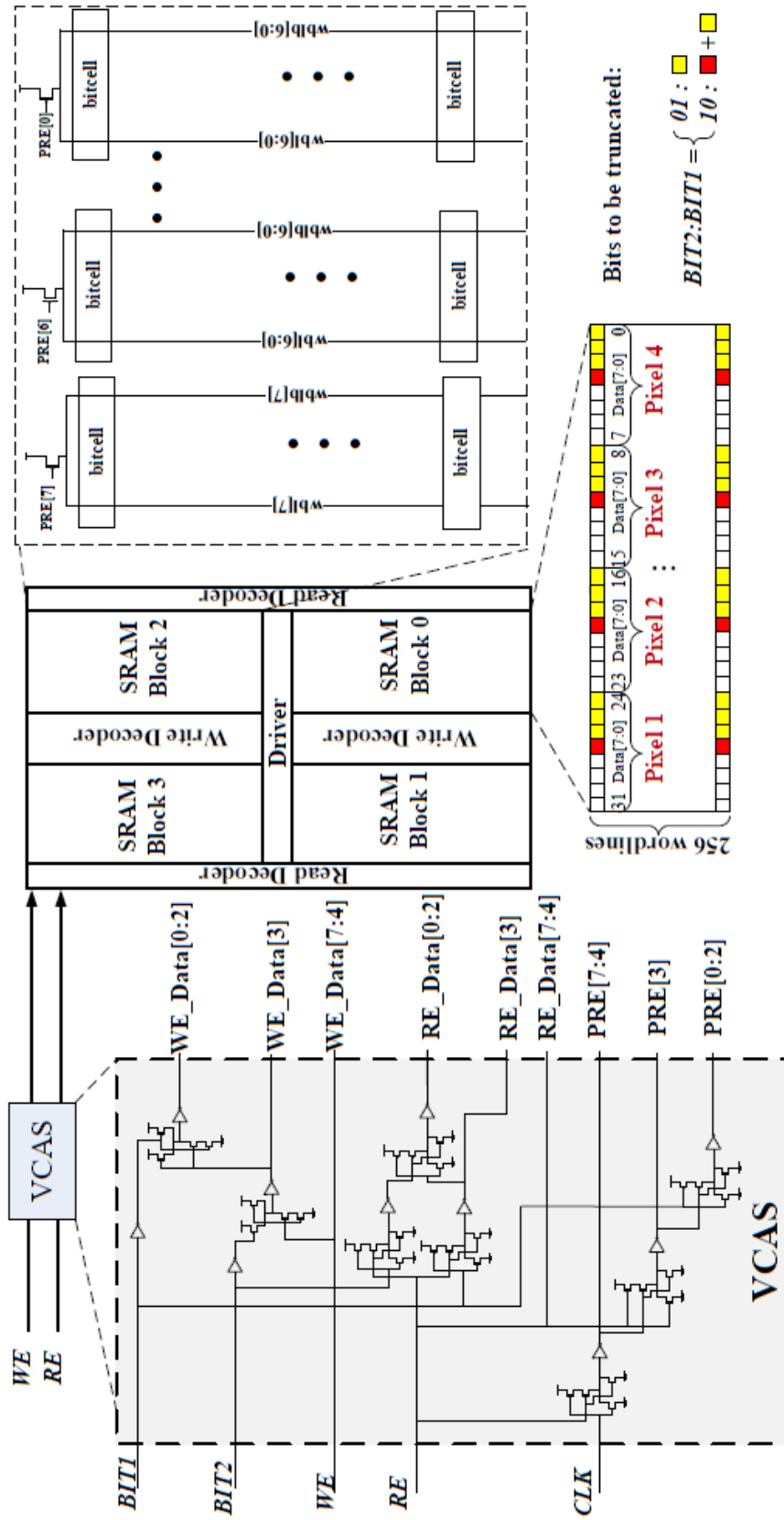
Figure 11. SRAM circuit with VCAS control circuit

23

The hardware simulation results are shown in Fig. 12. This timing diagram of VCAS in three different viewing contexts shows that this memory design provides enough speed to support the typical mobile video data transmission. The data (0x97, 0xf3, 0xc6, 0x0e) is written to the address (0x0a, 0x1a, 0x25, 0x3b) and then read out from the same address. These data sets were chosen for testing because they can clearly show the bit truncation results. If the mobile luminance sensor detects dark environment (0-1000 lux), then BIT2 and BIT1 are both 0s, and VCAS stores original video data for display; if the mobile user is watching the videos in overcast environment (1000-10000 lux) and then BIT1 becomes 1, VCAS will truncate 3 LSBs and the data will become (0x90, 0xf0, 0xc0, 0x08); In sunlight environment (10000+ lux) and BIT2 also becomes 1, VCAS will truncate 4 LSBs and data becomes (0x90, 0xf0, 0xc0, 0x00).
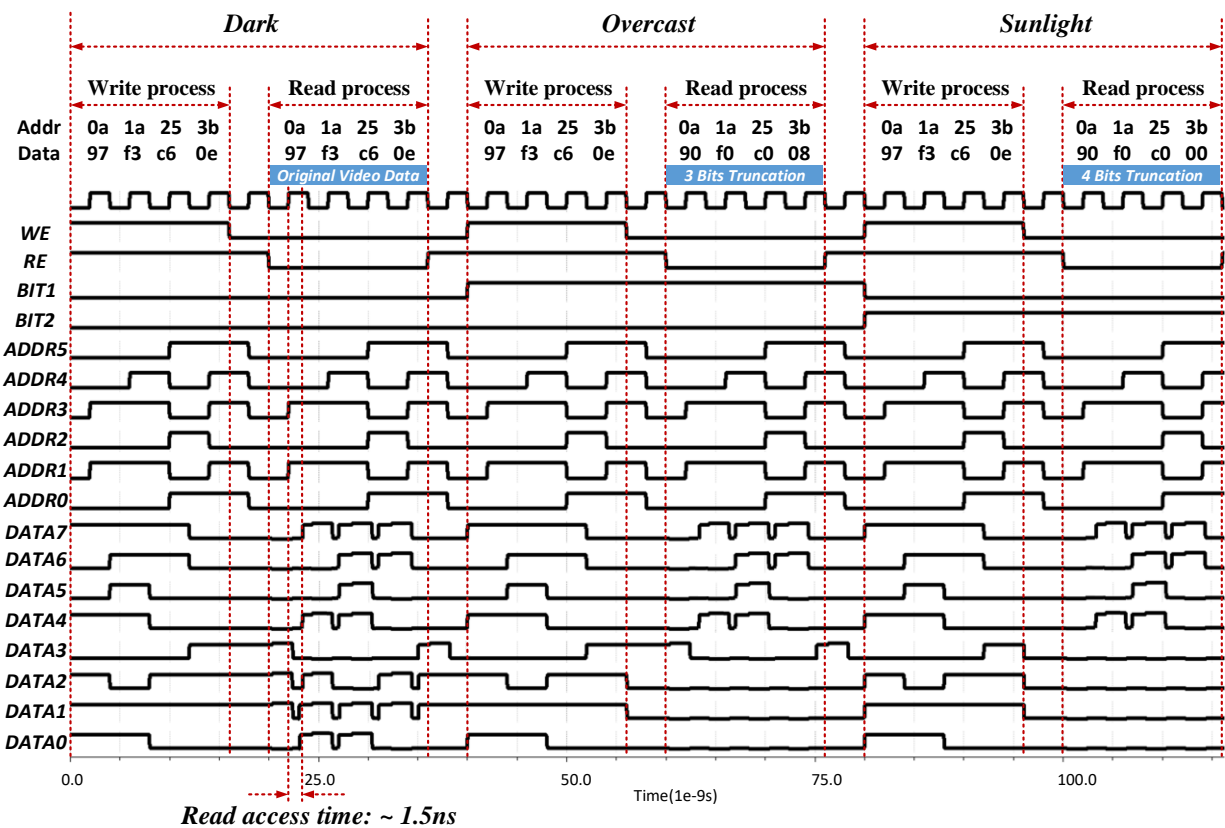

Figure 12. Timing diagram of VCAS in different viewing contexts

3.2. VCAS Experiment Methodology

3.2.1. Degradation Category Rating Method

The testing procedure closely followed ITU's subjective video quality testing recommendation. In the video quality evaluation test, Degradation Category Rating (DCR) method [18, 19] was used to collect viewer feedback, this method is also known as the Double Stimulus Impairment Scale (DSIS). During the experiment, the participants were asked to watch both original video and truncated video. After they have finished watching both video samples, I ask viewers to give their opinion score from 1 to 5 based on the quality difference between the original video and the truncated video. The five-grade impairment scale should be used: 5 imperceptible, 4 perceptible, but not annoying, 3 slightly annoying, 2 annoying, 1 very annoying. According to [20], 4.0 or higher score is set to be the target for acceptable video quality.

3.2.2. Subjective Video Quality Test

The perceptual video quality test was conducted using CIF (352x288 px) first. Six video samples were selected for this test, including Akiyo, Coastguard, Container, Foreman, Hall and Waterfall. These video samples are recommended by the Joint Collaborative Team on Video Coding (JCT-VC) [21, 22]. I clipped the videos to 6 seconds long and invited 15 participants to watch these videos on a Samsung Galaxy Note 4 smartphone.

Fig. 13 and Fig.14 show all the test results. According to the data I collected during the test, in overcast context, CIF videos received 55.6% imperceptible rating and 38.9% perceptible but not annoying rating. In direct sunlight context, CIF videos received 72.2% imperceptible rating and 22.2% perceptible but not annoying rating. The overall DCR score is higher than 4.3. The average score under overcast context is 4.5 and for direct sunlight context, the average score is 4.7, which indicates that VCAS provides acceptable video quality.
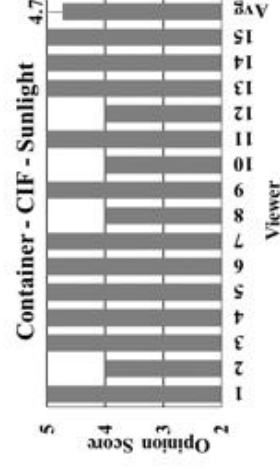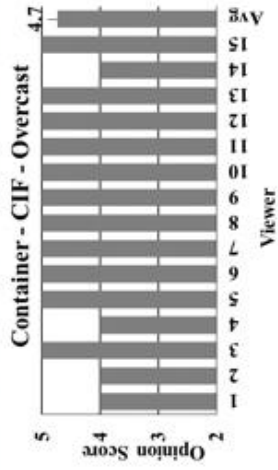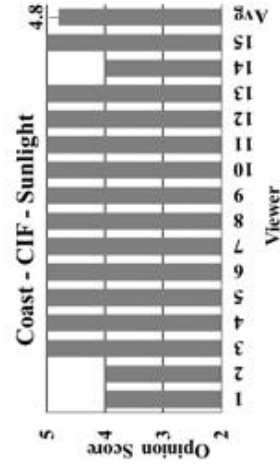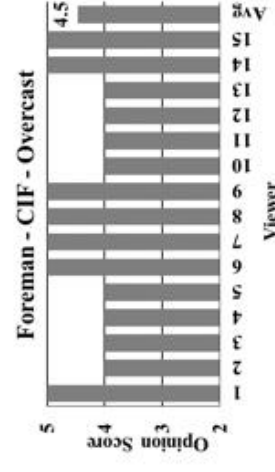
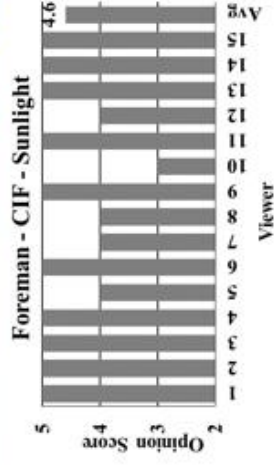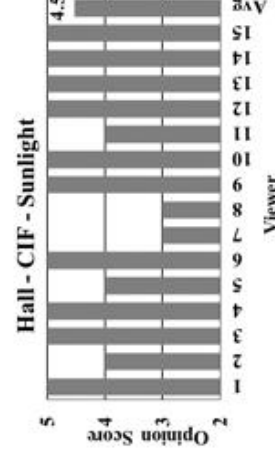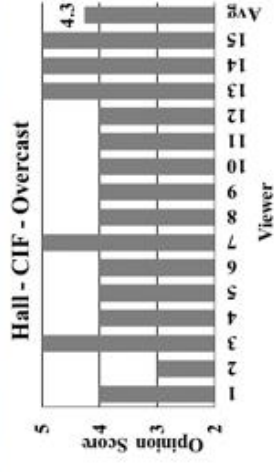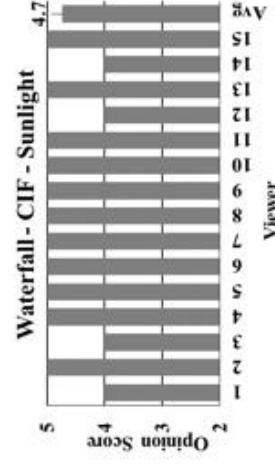Figure 13. VCR scores of CIF videos (Akiyo, Coast, and Container)
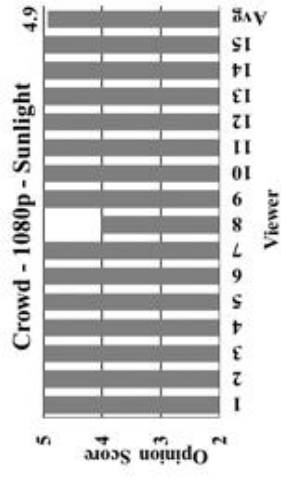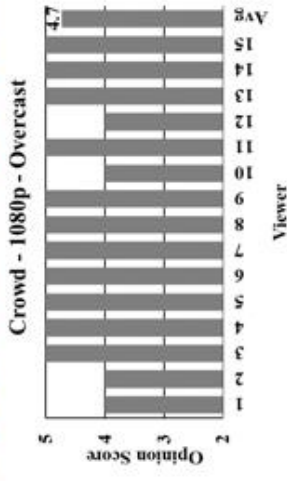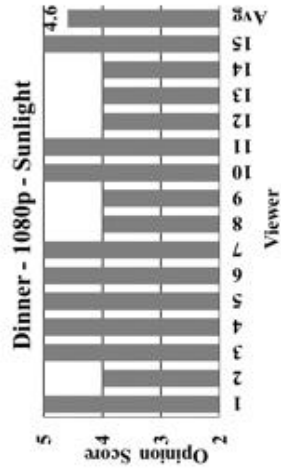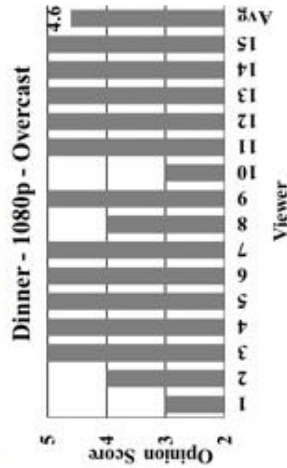
26

Figure 14. VCR scores of CIF videos (Foreman, Hall, and Waterfall)
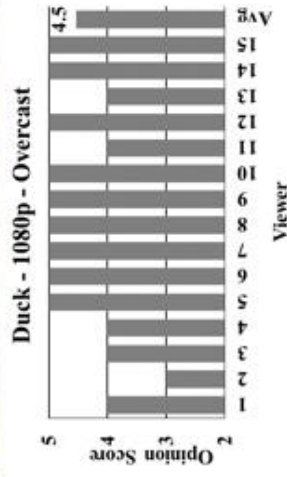
Figure 15. VCR scores of FHD videos (Crowd, Dinner, and Duck)

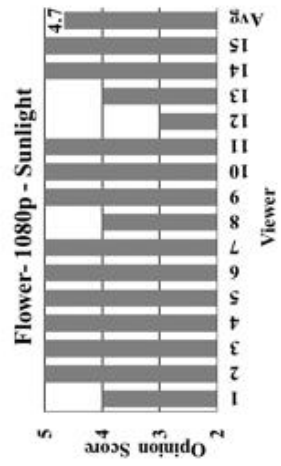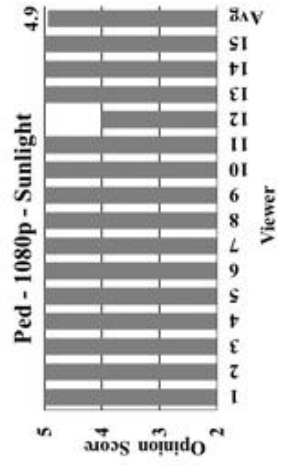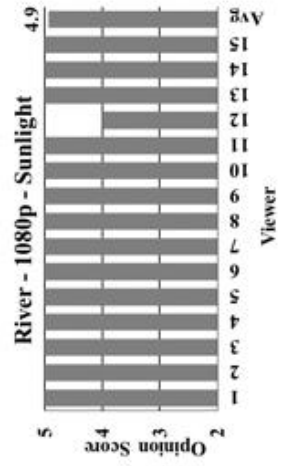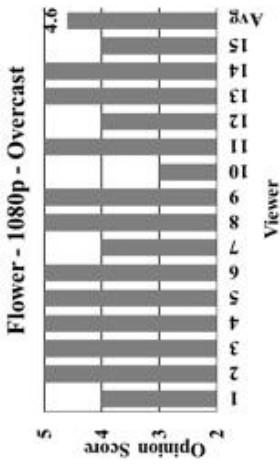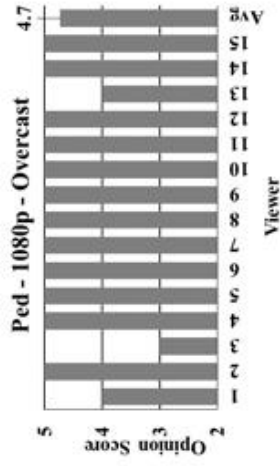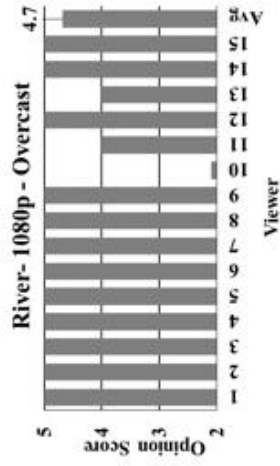Figure 16. VCR scores of FHD videos (Flower, Pedestrian, and River)
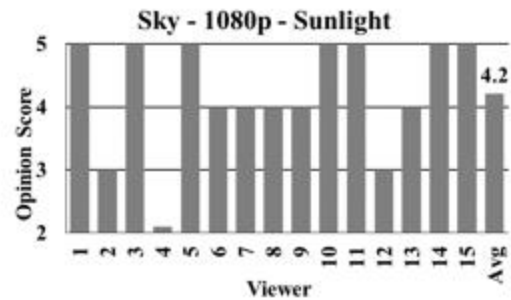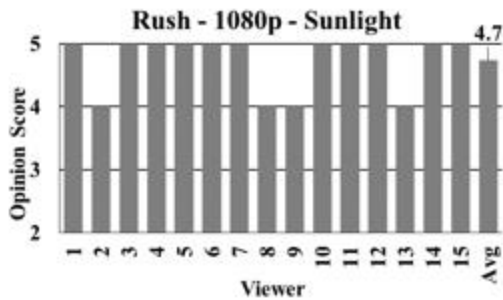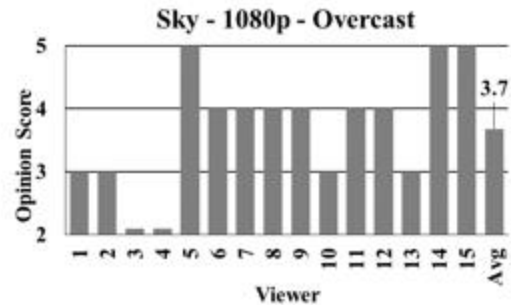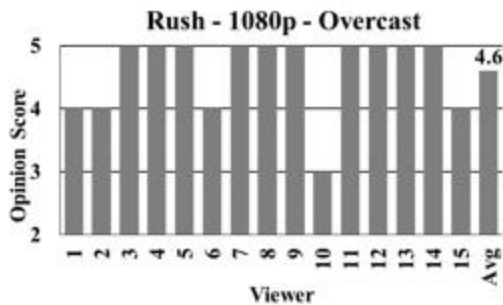
Figure 17. VCR scores of FHD videos (Rush and Sky)

Since the modern video transmission technology has been developing so fast and it is capable of making FHD video streaming popular, 8 FHD videos were also tested in this experiment. These video samples are: Crowd, Dinner, Duck, Sunflower, Pedestrian, River, Rush and Sky [23]. Fig.15, Fig.16 and Fig.17 show that the feedback scores for FHD video tests. In overcast context, FHD videos received 65% imperceptible rating and 19.2% perceptible but not annoying rating. In direct sunlight context, FHD videos received 77.5% imperceptible rating, and 19.2% perceptible but not annoying rating. The reason is that bit truncation technique brings shade into video frames, and the shades are much more obvious in frames that have a large area

30

with bright and unique color. Therefore, the bits need to be truncated will also need to take the video frame characteristics into consideration.

3.2.3. Video Quality Test Results Accuracy

The standard deviation for the video quality testing results was calculated to quantify the amount of variation of viewers' opinion scores. As shown in Table 1 and Table 2, all the testing scores are within 1 score away from the mean score. Viewers have very close ratings for all the video samples provided, which proofs that VCAS can bring a good power and quality tradeoff.

Table 2. Standard deviation of test results (CIF videos)

	STDEV(Overcast)	STDEV(Sunlight)
Akiyo	0.96	0.74
Coast	0.52	0.41
Container	0.46	0.46
Foreman	0.52	0.63
Hall	0.59	0.74
Waterfall	0.63	0.46

Table 3. Standard deviation of test results (FHD videos)

	STDEV(Overcast)	STDEV(Sunlight)
Crowd	0.46	0.26
Dinner	0.74	0.51
Duck	0.64	0.35
Flower	0.63	0.62
Pedestrian	0.59	0.26
River	0.82	0.26
Rush	0.63	0.46
Sky	0.98	0.94

3.2.4. Power Savings

Disabling RE and WE signals of truncated bits can provide 18% and 23% power savings in overcast and sunlight contexts. To further improve the power efficiency, we found out that the pre-charge circuits can be disabled to bring additional power saving without reducing

performance and area overhead. By disabling the pre-charge circuits, we can bring the power savings for overcast and sunlight contexts up to 36% and 48% respectively.

3.2.5. Comparison with Prior Work

The existing solutions for video memory power optimization are mainly focusing on hardware-level low-power techniques. The proposed VCAS, however investigated the power saving opportunities provided by QoE under environmental visual interference. Our research is orthogonal to the existing techniques. Furthermore, the proposed VCAS is not only limited to one specific type of memory design, any memory design that uses bit cell array is capable of using this technique to achieve high efficiency power saving.

Table 4. Comparison with prior low-power mobile video SRAM designs

| | *TVLSI'08 [13]* | *TCASVT'11 [9]* | *TCASII'12 [3]* | *ISVLSI'07 [26]* | **This work** | | |
					In dark	**In overcast**	**In sunlight**
video specific characteristics	correlation of MSB	contribution of MSB and LSB	different contribution of MSB and LSB	reconstructed image memory	viewer experience awareness		
no. data bits for each pixel	8 bits/pixel	8 bits/pixel	8 bits/pixel (SDA-I) /6 bits/pixel (SDA-II)	8 bits/pixel	8 bits/pixel	5 bits/pixel	4 bits/pixel
dynamic adaption	No	No	No	No	**Yes**		
low-power technique	data filpping	6T+8T bitcells	8T+10T bitcells	10T non-precharge	bit truncation for run-time adaption		
bitcell array modification	Yes	Yes	additional word line	Yes	No		
additional hardware needed	majority logic and data flipping block	single-ended 6T, peripheral circuitries	No	read and write circuits	VCAS control circuit		
power penalty for extra bits	Yes	No	No	No	No	No	No
power consumption	-14%	-32%	-91.8% (SDA-I)/ -95% (SDA-II)	-74%	**0%**	**-36%**	**-48%**
video quality	good (without any quality loss)	acceptable (without significant quality loss)	acceptable (without significant quality loss)	good (without any quality loss)	good (without any quality loss)	acceptable (without significant quality loss)	acceptable (without significant quality loss)
area overhead	+14%	+11.64%	+52% (SDA-I)/ +18.5% (SDA-II)	+14.4%	**<0.01%**		
speed overhead	4%	-	-	65% faster	negligible		
technology	90nm	65nm	45nm	90nm	45nm		

33

4. SUMMARY

In this thesis, I introduced two power saving techniques. The main goal of these two techniques is to combine hardware power saving technique with viewers' QoE to maximize the benefit we can receive from the tradeoff between quality and power. Compared to the existing power saving techniques, my proposed methods that can combine hardware and QoE together to achieve higher power saving and result in lower perceptual video quality degradation. For the viewer-aware intelligent mobile video system, I used machine learning to differentiate the different viewing contexts, which provides much more space for power saving using voltage scaling technique. Compared with existing studies, I increased the elements that are considered when introducing power saving techniques to improve power efficiency. Our research uses the most common sensors on smartphones, so it can be directly applied to modern mobiles without additional cost or space.

4.1. Future Work

4.1.1. Bit Truncation and Voltage Scaling for Different Video Types

During our research, I found out that some videos have lower average opinion score compared to the others. The reason is that these types of videos have very slow color transaction and usually takes most of the space of a frame. It makes the visual effect of bit truncation much more obvious to see. A part of future work can focus on investigating this type of video quality degradation.

4.1.2. Improve Video Testing Device

The current embedded system device I built for machine learning training is based on two development boards, the space, weight and processing speed can be further improved by developing everything on a mobile device.

REFERENCES

[1] Cisco Systems, Inc., "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2012- 2017," Feb. 2013. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html. [Accessed 20 October 2017].

[2] M. A. Hoque, M. Siekkinen, and J. K. Nurminen, "Energy Efficient Multimedia Streaming to Mobile Devices – A Survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 579-597, First Quarter, 2014.

[3] Y. Benmoussa, J. Boukhobza, E. Senn, and D. Benazzouz, "Energy Consumption Modeling of H.264/AVC Video Decoding for GPP and DSP," in *Proc. 16th Euromicro Conference on Digital System Design*, pp. 890-896, 2013.

[4] A. Pathak, D. Sachan, H. Peta and M. Goswami, "A Modified SRAM Based Low Power Memory Design," in *29th International Conference on VLSI Design and 15th International Conference on Embedded Systems (VLSID)*, Kolkata, Jan. 2016.

[5] J. Xue and C.-W. Chen, "A Study on Perception of Mobile Video with Surrounding Contextual Influences," in *Proceedings of the 4th International Workshop on Quality of Multimedia Experience (QoMEX)*, Melbourne, Austrailia, Jul. 2012.

[6] J. Xue and C.-W. Chen, "Mobile JND: Environmental adapted perceptual model and mobile video quality enhancement," in *MMSys '12 Proceedings of the 3rd Multimedia Systems Conference*, Chapel Hill, Feb. 2012.

[7] J. Xue and C.-W. Chen, "Mobile Video Perception: New Insights and Adaptation Strategies," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 3, pp. 390-401, Jun. 2014.

[8] J. Xue and C.-W. Chen, "Towards Viewing Quality Optimized Video Adaptation," in *International Conference on Multimedia and Expo (ICME)*, Barcelona, Jul. 2011.

[9] "Recommendation ITU- R BT.500 13." [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf. [Accessed 20 October 2017].

[10] J. Kwon, I. Lee, and J. Park, "Heterogeneous SRAM Cell Sizing for Low Power H.264 Applications," *IEEE Transactions on Circuits and Systems I*, vol. 99, no. 2, pp. 1-10, Feb. 2012.

[11] J. Wang, N. Gong, L. Hou, X. Peng, S. Geng, and W. Wu, "Low Power and High Performance Dynamic CMOS XOR/XNOR Gate Design," *Microelectronic Engineering*, vol. 88, no. 8, pp. 2781-2784, 2011.

[12] F. Frustaci, M. Khayatzadeh, D. Blaauw, D. Sylvester and M. Alioto, "SRAM for Error-Tolerant Applications With Dynamic Energy-Quality Management in 28 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 5, pp. 1310-1323, May 2015.

[13] F. Sampaio, M. Shafique, B. Zatt, S. Bampi, and J. Henkel, "Energy-efficient Architecture for Advanced Video Memory," in *Proc. 2014 IEEE/ACM International Conference on Computer-Aided Design*, pp. 132-139, Nov. 2014.

[14] J. Edstrom, D, Chen, J. Wang, H. Gu, E. A. Vazquez, M. E. McCourt, N. Gong, "Luminance-adaptive smart video storage system," *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, QC, 2016, pp. 734-737.

[15] "Arduino Uno", [Online]. Avaliable: https://store.arduino.cc/usa/arduino-uno-rev3. [Accessed 20 October 2017].

[16] "Odroid-XU3", [Online]. Avaliable: https://ameridroid.com/products/odroid-xu3-lite. [Accessed 20 October 2017].

[17] "YUV Video Sequences," [Online]. Available: http://trace.eas.asu.edu/yuv/. [Accessed 20 March 2017].

[18] A. Ichigaya and Y. Nishida, "Required Bit Rates Analysis for a New Broadcasting Service Using HEVC/H.265," *IEEE Trans. on Broadcasting*, vol. 62, no. 2, Jun. 2016.

[19] "Recommendation ITU- R BT.500-13.," [Online]. Available: https://www.itu.int/dms_pubre c/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf

[20] [Online]. Available: https://www.telchemy.com/appnotes/Understanding%20IP%20Video% 20Quality%20Metrics.pdf. [Accessed 20 August 2017]

[21] "Joint collaborative team on video coding (JCT- VC)," ITU, [Online]. Available: http://www.itu.int/en/ITU-T/studygroups/2013–2016/16/Pages/video/jctvc.aspx. [Accessed 20 March 2017]

[22] J. Wang, N. Gong, L. Hou, X. Peng, R. Sridhar, and W. Wu, "Leakage Current, active power, and Delay Analysis of Dynamic Dual Vt CMOS Circuits under P-V-T Fluctuations," *Microelectronics Reliability*, 2011, vol. 51, no. 9, pp. 1498-1502, Sep. 2011.

[23] [Online]. Available: https://media.xiph.org/video/derf/. [Accessed 20 August 2017]

[24] N. Gong, S. Jiang, A. Challapalli, S. Fernandes and R. Sridhar, "Ultra-Low Voltage Split-Data-Aware Embedded SRAM for Mobile Video Applications," *IEEE Transactions on Circuits and Systems II*, vol. 59, no. 12, pp. 883-887, Dec. 2012.

[25] M. A. Alsheikh, S. Lin, D. Niyato and H. P. Tan, "Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications," in *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1996-2018, Fourthquarter 2014.

[26] Kamiński, B.; Jakubczyk, M.; Szufel, P. "A framework for sensitivity analysis of decision trees". *Central European Journal of Operations Research*: 1-25

[27] K. Takeda *et al.*, "A read-static-noise-margin-free SRAM cell for low-VDD and high-speed applications," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 113-121, Jan. 2006.

[28] R. V. Joshi, R. Kanj, and V. Ramadurai, "A novel column-decoupled 8T cell for low-power differential and domino-based SRAM design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 5, pp. 869–882, May 2011.

[29] S. A. Verkila, S. K. Bondada, and B. S. Amrutur, "A 100 MHz to 1 GHz, 0.35V to 1.5V supply 256 64 SRAM block using symmetrized 9T SRAM cell with controlled read," in *Proc. Conf. VLSI Design*, Jan. 2008, pp. 560-565.

[30] F. Abouzeid, A. Bienfait, K. C. Akyel, A. Feki, S. Clerc, L. Ciampolini, F. Giner, R. Wilson, and P. Roche, "Scalable 0.35 V to 1.2 V SRAM Bitcell Design From 65 nm CMOS to 28 nm FDSOI," *IEEE J. Solid-State Circuits*, vol. 49, no. 7, pp. 1499-1505, Jul. 2014.

[31] Y.-W. Chiu, Y.-H. Hu, M.-H. Tu, J.-K. Zhao, Y.-H. Chu, S.-J. Jou, and C.-T. Chuang, "40 nm Bit-Interleaving 12T Subthreshold SRAM with Data-Aware Write-Assist," *IEEE Trans. Circuits Syst. I*, vol. 61, no. 9, pp. 2578-2585, Sep. 2014.