A Low-Energy Video Event Data Recorder Using Dual Image/Video Codec

Jongpil Jung†, Jinyeon Lim†, Seunghan Lee†, Jungeon Lee†, Jinyoung Yang‡, Chong-Min Kyung‡
KAIST†, CISS‡
Daejeon, South Korea
{jpjung08, jylim, shlee04, jelee}@vslab.kaist.ac.kr
{jyyang0308, kyung}@kaist.ac.kr

Abstract

Energy consumption is a crucial issue of mobile surveillance cameras owing to limited battery capacity. The lifetime of the system is significantly extended by the event-driven operation; the system mostly stays in sleep mode and wakes up only when an event is detected. In this paper, we propose a design of a low-energy surveillance camera that records events such as the abnormal movement of objects, or physical damage to the camera itself. Unlike conventional event-driven approaches, the proposed system records video from 10 seconds before the event detection because the most critical information is often before or at the moment of event detection, not after the detection. Two different encoders, a low-power encoder and a high-compression encoder, are employed together to implement the low-energy surveillance camera. Experimental results show that the energy consumption of the whole system is reduced by up to 74.9% (by 66.8% on average) compared with conventional always-on system.

1. Introduction

Demand for low-energy surveillance camera is steadily increasing due to growing needs in various places where power line is not necessarily available. However, conventional researches based on event-driven operation focuses on helping human operators rather than reducing energy consumption. In [11], video surveillance system is proposed which detects abandoned or removed objects. In [14], traffic surveillance framework is proposed which is for detection, classification, and tracking of vehicles. Pedestrian detection is yet another active research topic [12]. The researches are to help operators who constantly monitor all video streams. The system automatically detects abnormal events, then let the operators know [13].

Energy consumption of the surveillance camera is significantly reduced by the event-driven operation because system is mostly in sleep mode and wakes up only when an event is detected [9]. However, the system proposed in [9] is not practical because it takes some time for the system to wake up. Because the system starts recording after an event is detected, it cannot record the most important scenes which is the moment of event detection or scene prior to the detection. According to [2], it is required to save video data starting from 10 seconds prior to the time of event. The video during 10 seconds before the event is critical to figure out how the event occurred.

To reduce energy consumption of surveillance camera, conventional methods optimize video encoders [4][7] or adopt low-resolution image sensors [11][5]. Considering that the video encoder is not the only one which consumes power, system-level optimization beyond codec is more effective to reduce energy consumption. The resolution of cameras in commercial mobile device has already reached 4k UHD, i.e., 3840 × 2160. Demand for high-definition video is increasing in surveillance applications as well. Thus, it is important to develop low-power surveillance camera with high-resolution image sensor.

In this paper, we propose a design of low-energy event-driven surveillance camera which records video of abnormal events. While conventional event-driven methods are focusing on helping human operators by detecting abnormal events, the objective of the proposed design is to maximize the lifetime of battery-based mobile camera system by minimizing energy consumption. Two different video encoders are employed to record the abnormal events starting from 10 seconds before the event.

To verify the effectiveness of the proposed method, we applied the design to video event data recorder (VEDR) installed in a vehicle. VEDR is a system which records video of surroundings of vehicle so that the video can be used as evidence in case of an accident. It started from car black box system which records data from various sensors of vehicle [8]. Recently, people started to record high-definition video as well as the values of sensors on vehicle. Since the VEDR is operated by the energy saved in battery of vehicle, it is not free from the problem of energy scarcity. Especially in cold weather during winter, performance of car battery significantly degrades [10]. Many cars even fail to start engine in such situations, if the energy consumption of electrical devices during car parking is excessive.
2. Event-driven Operation

In conventional surveillance applications, camera is always recording even when there is no change in the captured scene. Considering that actual movement of object is very scarce in most cases, it is very inefficient if the camera is kept on at all time. The average power consumption of surveillance camera system can be significantly reduced by adopting event-driven operation method. In the event-driven system, most of the power-consuming parts are turned off to save energy consumption. The system wakes up only when abnormal event is detected.

2.1. Conventional event-driven system

Typical VEDR consists of a CMOS image sensor (CIS), a video encoder, a microprocessor, a memory system, a video data storage, and a communication module. If low-power event detector is added to the conventional VEDR as shown in the Fig. 1, average power consumption is significantly reduced. In standby state with no event detected, only the event detector and CIS are turned on. Since power consumption of event detector is much less than that of video encoder or microprocessor, the system power consumption in standby state is much less than that of conventional VEDR. When an event is detected, power consumption of event-driven VEDR is slightly more than that of conventional VEDR because of the event detector. Fig. 2 shows how power consumption of both systems vary with time. Average power consumption of event-driven VEDR is much less than power consumption of conventional VEDR unless events are detected too frequently.

For an event-driven VEDR, type of events must be defined. Motion in captured scene is widely used as type of event for surveillance camera. If there is no moving object in video, then the video does not provide any information. Another important event for vehicle is acceleration. If a vehicle is physically damaged by another vehicle or pedestrian, acceleration is detected. In this paper, we assume that it is an event when motion or acceleration is detected.

2.2. Proposed method

As it is mentioned in the introduction, it is required to start recording 10 seconds before detection of an event. In conventional event-driven systems, it was impossible to have video before an event detection because system wakes up after the event is detected. Considering that it takes some time for the system to wake up, conventional event-driven system starts recording few seconds after an event is detected. In such a case, the most important portion of video will not be recorded.

The problem comes from the difference between event detection time and record starting time. The problem is solved if we can delay video data for few seconds. If the few seconds of video is temporarily saved in volatile memory, the system is capable of having all required data even after the event is detected. In conventional works, this is done by using buffer memory between H.264 encoder and data storage. Video captured by CIS is encoded in H.264 format and saved in the buffer at all time. If something happens, then the data in the buffer is written to data storage. In such a system, however, energy consumption is wasted because all system components including high power consuming H.264 encoder must be kept on.

To save energy consumption, it is required to turn off as many components as possible. At least, the most power-consuming parts such as H.264 encoder and processor must be turned off for significant energy reduction. CIS must be turned on even when there is no event detected because the video data is required by event detector. For H.264 encoder to be turned off, uncompressed data from CIS must be temporarily saved. However, that is impractical because the data size of uncompressed video data is too large to be saved in a memory. The size of uncompressed 10-second 1080p video is 13.6 gigabits.

Energy consumption of the proposed system is reduced by using light-weight video codec addition to H.264 video encoder. The light-weight codec, which is JPEG, is added between CIS and H.264 encoder. The video data from CIS is passed on to JPEG encoder, and JPEG-compressed data is temporarily saved in DRAM, which is directly connected to the JPEG encoder. Other blocks such as H.264 encoder, processor, storage, and communication module are all turned off in this state. When an event is detected, all system components wake up. JPEG-compressed data in DRAM is decoded by JPEG decoder, then re-encoded by

![Figure 1. Event-driven video event data recorder](image)

![Figure 2. Event-driven operation of smart camera system](image)
H.264 encoder and saved in storage. More details on the system design are explained in Section 3.

In the proposed design, power consumption at standby state is significantly reduced compared with conventional always-on operation, because most of components such as H.264 encoder, processor, storage, communication module, and main memory are turned off. However, since event detector and JPEG codec are added, power consumption during event-detected state is slightly more than conventional system. Assuming situations with infrequent events, average power consumption of the proposed system is much less than conventional system. More details on energy reduction is explained in Sections 4 and 5.

3. System Design

Block diagram of the proposed system is shown in Fig. 3. Similar to conventional VEDR system, the proposed system consists of CIS, image signal processor, H.264 encoder, main processor, event detector, accelerometer, memory, storage, and communication module. Micro SD card is used as a storage of video data. The major difference of the proposed architecture from conventional system is addition of JPEG encoder and decoder. For typical surveillance camera, the flow of video data starts from CIS and ends at the storage. For the data flow in the proposed system, there are two different paths: 1) one going through JPEG codec, and 2) the other one which bypasses the JPEG codec. The path is decided by operating mode which depends on what state the car is in. The system takes ACC signal of a vehicle (on/off signal for accessories) as an input. When a driver starts engine, ACC signal goes high and our system starts operating in driving mode. When a driver turns off engine, ACC signal goes low and our system operates in parked mode.

Besides path of video data, there are paths for other information as well. The information for event detection is delivered from accelerometer and JPEG encoder. For motion detection, the event detector uses DCT coefficients. To detect physical contact with other vehicles, the event detector utilizes values from the accelerometer. When an event is detected, the event detector passes the information to the processor. Then the processor controls all blocks, including H.264 encoder, according to the information.

3.1. Parked Mode

The event-driven feature explained in Section 2.2 is employed during parked mode. Within parked mode, the system switches between two different sub-modes, standby mode and event mode, depending on the detection of events. Since the light-weight compression unit is utilized during the parked mode, mux and demux are switched accordingly.

Normally, when there is no event detected, the system operates in standby mode, which is shown in Fig. 4. In the standby mode, the system is at the state of waiting for an event. We turn off as many components as possible because the system is in standby mode most of the time. The power consumption of standby mode significantly affects total energy consumption of the system. Only the minimal required parts for event detection are turned on: CIS, event detector, accelerometer, JPEG encoder, and low-power memory.

The important feature of the proposed system is recording video prior to the detection of event. The system saves the 10 seconds of data in the memory, which is connected to JPEG encoder. Bold arrow in Fig. 4 indicates video data flow in standby mode. CIS always captures scene even when there is no event. Then the data is temporarily saved in memory. Only the recent 10-second video is kept in memory, and old data is overwritten by new data. The size of uncompressed data from CIS becomes problem in terms of energy consumption and cost. According to [1], read/write power of LPDDR2 memory is quite a lot whereas the power consumption of deep power down (DPD) mode is very small. It is important to reduce the size of data to be written and read, so that DPD mode is utilized as much as possible. Considering the cost of memory device with large capacity is expensive compared with low-capacity memory, it is required to reduce the data size. Due to above reasons, JPEG encoder is employed in the proposed design. Uncompressed data from CIS is passed on to JPEG encoder. After the JPEG encoding, data size reduces by up to 95% with quality of PSNR 36dB. Thus, the size of 10-second video reduces from 13.6 gigabits to 0.7 gigabits. Considering margin, 1Gb LPDDR2 memory is sufficient for implementation.

It is also very important to design the event detector
is ready for H.264 encoding, JPEG decoder starts decoding and storage, get ready for H.264 encoding. After the system sor wakes up, then other blocks, including H.264 encoder. enable signal is generated by the event detector, the processing the bitstream in the storage. All blocks that were turned off during the sleeping state, the system switches to event mode. In this mode, the event detector is always turned on during standby mode. We adopt basic motion detection method based on inter-frame difference [3][6]. Motion is detected when difference between two consecutive frames exceeds threshold. Data of previous frame is required to get inter-frame difference. If all uncompressed data from CIS is delivered to the even detector, power consumption of event detector is enormous because of huge data size. In the proposed design, movement within captured scene is detected by using DCT coefficients from JPEG encoder, not uncompressed pixel data. DC coefficient from 8 × 8 block DCT represents average intensity of the block. It is the effect of downsizing image from 1920 × 1080 to 240 × 135. Since only luminance information is enough for motion detection (i.e., color information can be discarded), data size required for a frame is reduced down to 32.4KB.

Physical movement of a vehicle is another important event, which is detected using accelerometer. The accelerometer is initialized by processor at the very beginning so that it operates stand alone. It operates even when processor is in sleep mode. The accelerometer triggers an interrupt signal when acceleration at any direction exceeds threshold.

When any kinds of event is detected by the event detector, the system switches to event mode. In this mode, the system starts encoding in H.264 format and saves encoded bitstream in the storage. All blocks that were turned off during the parked mode are turned back on. Data flow during the event mode is shown in Fig. 5 as a bold arrow. When enable signal is generated by the event detector, the processor wakes up. Then other blocks, including H.264 encoder and storage, get ready for H.264 encoding. After the system is ready for H.264 encoding, JPEG decoder starts decoding data at low-power memory that are the video prior to the event detection. The 10-second delayed video is passed on to the H.264 encoder and finally it is saved in the storage. Meanwhile, video from CIS is still being saved in memory, which will be saved in storage through H.264 encoder at the end. After the event is finished, the system switches back to the standby mode. Timing of this process is expressed in the Fig. 6.

3.2. Driving Mode

In driving mode, the system operates as shown in Fig. 7. The system operates similar to conventional VEDR, which does not have light-weight compression unit. When a driver starts engine, the proposed system reads ACC signal from vehicle (which switches from low to high at the moment) and the mux is set to skip the JPEG codec parts. The components for the parked mode such as JPEG encoder, JPEG decoder, and low-power memory, are turned off during the driving mode. Other components such as CIS, H.264 encoder, processor, event detector, and storage are always kept on. The role of this mode is very simple because it is just required to capture and save. Event-driven feature explained in Section 2.2 is not employed in this mode. The video data from CIS module is directly passed on to the H.264 encoder, then they are saved at the storage in H.264 format.

Since the vehicle moves during the driving mode, it has no meaning to detect movement of object as an event. Thus, only the accelerometer is used as an input of event detector during the driving mode. The role of event detection in this mode is quite different from the parked mode. It is not for saving energy consumption, but for saving important data in more secure place. If an event is detected, the video of the event is duplicated and saved in separate space of storage. It helps to be prepared for memory failure.

The power consumption of the driving mode is quite a lot because the most power-consuming part, application processor with H.264 encoder and processor, are always kept on. The only event-driven feature in this mode is the storage. Considering energy consumption of micro SD card is very small compared with total energy consumption, it can be said that the system constantly consumes a lot of energy.
4. Energy Model

Power-gating is applied to reduce the power consumption of the proposed system. Thus, total energy consumption depends on which blocks are turned on for how long. Total power consumption is sum of power consumption of components, which are turned on at the time. Power consumption of each component is assumed to be constant.

There are two different operating modes in the proposed design of VEDR: driving mode and parked mode. As shown in Fig. 7, most of the blocks are turned on during driving mode, including CIS, H.264 encoder, processor, main memory, micro SD card, and other peripherals. Since power consumption during driving mode ($P_{drv}$) is constant, energy consumption during driving mode ($E_{drv}$) is simply equal to power consumption multiplied by duration of the driving mode ($T_{drv}$) as expressed in (1).

$$ E_{drv} = T_{drv} \cdot P_{drv} $$  (1)

Power consumption of whole system during driving mode is sum of power consumption of blocks that are turned on during the mode as shown in (2).

$$ P_{drv} = P_{cis} + P_{h264} + P_{proc} + P_{dram} + P_{sd} + P_{peri} $$  (2)

where $P_{cis}$, $P_{h264}$, $P_{proc}$, $P_{dram}$, $P_{sd}$, and $P_{peri}$ are power consumption of CIS, H.264 encoder, processor, main memory, micro SD card, and other peripherals, respectively.

Energy consumption during parked mode is different from driving mode because some blocks are occasionally turned on and off when an event is detected. Total energy consumption during parked mode ($E_{park}$) can be divided into energy consumption during standby mode ($E_{stby}$) and energy consumption during event mode ($E_{evt}$). Total energy consumption also depends on the number of detected events ($N_{evt}$). Total energy consumption during parked mode is expressed in (3).

$$ E_{park} = E_{stby} + N_{evt} \cdot E_{evt} $$  (3)

Energy consumption of standby mode ($E_{stby}$) is equal to power consumption of standby mode ($P_{stby}$) multiplied by duration of standby mode ($T_{stby}$) as shown in (4).

$$ E_{stby} = T_{stby} \cdot P_{stby} $$  (4)

(5) shows power consumption during standby mode which is sum of power consumption of CIS ($P_{cis}$), JPEG encoder ($P_{jpegenc}$), low-power DRAM connected to JPEG encoder ($P_{lpdram}$), and event detector ($P_{ed}$).

$$ P_{stby} = P_{cis} + P_{jpegenc} + P_{lpdram} + P_{ed} $$  (5)

Other blocks are turned off during standby mode.

Energy consumption for recording one event video ($E_{evt}$) is equal to power consumption of event mode ($P_{evt}$) multiplied by average duration of an event ($T_{evt}$) as shown in (7).

$$ E_{evt} = T_{evt} \cdot P_{evt} $$  (6)

Power consumption of event mode is sum of power consumption of all blocks including CIS, event detector, JPEG encoder, JPEG decoder ($P_{jpegdec}$), H.264 encoder, processor, micro SD card, and other peripherals.

$$ P_{evt} = P_{cis} + P_{ed} + P_{jpegenc} + P_{jpegdec} + P_{h264} + P_{proc} + P_{sd} + P_{peri} $$  (7)

The event mode is the most power-consuming mode because all components are turned on including JPEG encoder and decoder.

5. Experiment

In this section, experimental results are explained that show how much energy consumption is reduced by the proposed method. The energy consumption is calculated by the energy model in Section 4. With the model, we compared the energy consumption of the proposed system with the energy consumption of conventional VEDR on various event occurrence scenarios. The value of each component’s power consumption is based on the measurement from the prototype. A commercially available application processor is embeded in the prototype. The chip includes an ARM Cortex-A8 processor which operates at 800MHz, and a H.264 encoder which is capable of encoding 1080p video at 30fps.

The energy consumption of the proposed system depends on the frequency of event occurrence, whereas energy consumption of conventional VEDR is almost constant no matter how frequent events are detected. For fair comparison, scenarios are selected carefully considering following parameters. The frequency at night is much lower than the frequency during a day because people are more active during a day. The type of location is another important factor. At crowded areas such as parking lots at mall, stadiums, or the main street, the frequency of event occurrences is extremely high. The frequency at residential area or office area is relatively low compared with a crowded area. It also depends on whether the parking lot is indoor or outdoor. Considering all, 11 scenarios are selected for the experimental results.

Fig. 8 shows examples of event occurrence. Each peak indicates one detected event and the position of the peak in horizontal axis indicates the time of the event. Event occurrence at underground parking lot of university research facility is shown in Fig. 8 (a). Three events are detected during three hours. However, as shown in the Fig. 8 (b), event occurrence rate at an apartment is 27.1 events per hour, which is much more than the rate at parking lot of research facility. Fig. 9 shows example images of detected events: vehicle and pedestrian.

\[ \text{As explained earlier, the energy consumption is different} \]

\[ \text{between day and night, and indoor and outdoor.} \]

\[ \text{Considering all, 11 scenarios are selected for the} \]

\[ \text{experimental results.} \]

\[ \text{Fig. 8 shows examples of event occurrence. Each peak} \]

\[ \text{indicates one detected event and the position of the peak in} \]

\[ \text{horizontal axis indicates the time of the event. Event} \]

\[ \text{occurrence at underground parking lot of university research} \]

\[ \text{facility is shown in Fig. 8 (a). Three events are detected} \]

\[ \text{during three hours. However, as shown in the Fig. 8 (b),} \]

\[ \text{event occurrence rate at an apartment is 27.1 events per} \]

\[ \text{hour, which is much more than the rate at parking lot} \]

\[ \text{of research facility. Fig. 9 shows example images of detected} \]

\[ \text{events: vehicle and pedestrian.} \]
Results are shown in Fig. 10. To compare energy consumption of the proposed system with that of conventional VEDR, each value for the proposed system are normalized to the value for conventional VEDR. As shown in the figure, energy consumption is reduced in every scenario. The energy consumption is reduced by up to 74.9%, and by 66.8% on average. Reduction of the energy consumption is most significant in the case of residential area during night because event occurrence rate is as low as 0.6 events per hour. The energy consumption is reduced even in the crowded area. However, the degree of reduction is 30.6%, which is not as significant as residential area at night.

Among detected events in the crowded area, not every event was meaningful. Important video in this application is about moving objects that can damage my vehicle. Most of events detected in the crowded area were pedestrians at a distance, which are not a threat to vehicles. If we design better event detector that eliminates false alarm, total energy consumption will be reduced even more.

6. Conclusions

In this paper, the design of a low-energy surveillance camera is proposed. The proposed system records video only when abnormal events are detected. Unlike conventional video event data recorders, the proposed system does not miss any important moment because it records the video of abnormal event starting from 10 seconds before the event. Energy consumption of the system is reduced by up to 74.9% and by 66.8% on average, thanks to the low power consumption during standby mode. The energy consumption is expected to be reduced even more when an event detector with less false positive is designed.

Acknowledgement

This work was supported by the Center for Integrated Smart Sensors funded by the Ministry of Science, ICT & Future Planning as Global Frontier Project (CISS-2013066998).

References