A Novel Approach to Video Compression using Region of Interest (ROI) Method on Video Surveillance Systems

DewiAnggraini Puspa Hapsari[1], Sarifuddin Madenda[2], Muhammad Subali[3], Aini Suri Talita[4]
Faculty of Computer Science and Information Technology, Gunadarma University, Jakarta, Indonesia[1, 2, 4]
Faculty Information Technology, Universitas Cendekia Abditama, Tangerang, Indonesia[3]

***Abstract*—With the increasing of criminal actions, people use various surveillance techniques to create a sense of security. One of the most widely used surveillance technique is installing CCTV cameras at various locations. On the surveillance systems, there are other supporting devices apart from CCTV cameras. One of such supporting devices is a hard disk to save the recorded data. The recording on CCTV has two modes: motion detection mode and continuous mode. The continuous mode will record continuously, which affects the amount of hard disk space used. Motion detection mode records one event only, not all recordings, saving hard disk space, however, it may miss some events. Based on these two modes, compression technology is required. The current compression technology applies the ROI method. A ROI (Region of Interest) is the part of the image that wants to filter to form some operations against it. ROI allows coding differently in certain areas of the digital image to have a higher quality than the surrounding area (background). This paper offers a novel approach to saving the foreground frame generated from the ROI method and compressing it. The novel approach will be applied to the AVI, MJPEG 2000, and MPEG-4 video formats. The decompression process is used to restore the original video data to measure the method's performance. To measure the proposed method's performance, it will compare the compression ratio and Peak Signal-to-Noise Ratio (PSNR) with the traditional method without implementing the ROI-based method. The PSNR value in this paper, that measures the quality of the compression result,s are above 40 dB. It indicates that the resulting video is similar to the original video. The ROI-based compression method can increase the compression ratio 5-7 times higher than the existing method for lossy AVI format video. While on MJPEG-2000 and MPEG-4 format video, it increases the compression ratio 7-15 times and 1-3 times, respectively. The PSNR value for the proposed method is above 40 dB, which indicates that the reconstructed video is similar to the original video, even though the pixel values have changed slightly.**

***Keywords*—Compression; decompression; foreground; region of interest; video surveillance systems**

I. Introduction

The increasing criminal actions result in people using various surveillance techniques to ensure security and manage communities to create a sense of security [1, 2]. One commonly used surveillance technique is to use video surveillance cameras installed in several places [3]. The development of the internet also supports the increase in surveillance cameras to analyze the recordings [4]. Video surveillance is currently an active field of research that aims to analyze video captured by cameras. This functionality makes it suitable for security applications, object identification and tracking, gender classification, et al. [5].

Nowadays, CCTV is facilitated by remote location monitoring with mobile phones. CCTV does not stand alone but has other supporting devices [6]. These supporting devices are used to monitor and record image objects seen by CCTV cameras [7]. CCTV cameras that produce a low-quality video which is generally recorded at a resolution of 352x240 and have a frame rate of 30 fps, occupy 10 GB of storage in one day [8].

The footage on CCTV is mainly stored on secondary storage devices. CCTV recording equipment has two modes, continuous and motion detection [9]. In continuous mode, CCTV will record continuously so that the hard disk capacity runs out quickly because it stores all the information captured by the camera without choosing whether the data is needed or not. In motion detection mode, also called event recording mode, CCTV only records when specific events occur. Because it only saves when an event occurs, the hard disk capacity does not run out quickly. Its limitation is that not all events are held because CCTV only stores when an object is moving, so some information is lost [10]. A video compression technique is needed that can combine the two modes' advantages.

Several types of video compression used in CCTV are Motion JPEG (MJPEG), Motion JPEG 2000 (MJPEG 2000), MPEG-4, and H.264 [11, 12]. MPEG-4 can be three times more efficient in compression than MJPEG. When the number of frame rates is less than 5-6 frames per second, then MPEG-4 will be less favorable. MPEG-4 uses H-264 video compression and will be efficient when bandwidth conditions are limited and stable. MJPEG uses the Discrete Continuous Transform (DCT) algorithm, while MJPEG 2000 uses the Discrete Wavelet Transform (DWT) algorithm to produce better image quality at higher compression levels. On MJPEG 2000 it is possible to perform a decompression process with a lower resolution representation than the original image, so it is suitable for motion detection algorithms [12].

Video contains data redundancy by which the differences can be recorded within frames or between frames. There are two types of video compression, based on the structure's color component (spatial redundancy) and based on changes between frames (temporal redundancy). The first type is based on the fact that the human eye cannot distinguish colors well enough because it is affected by the object's brightness. While the other type only encodes the changing frames and directly stored the rest data [13].

The frame changes because there is movement in the object (foreground) captured in the video. Each frame in video data generally contains objects (foreground) and background. The foreground and background detection processes are based on Region of Interest (ROI) that was chosen. It allows different encoding in some regions of the digital image to have a better quality than the surrounding area (background) [14]. ROI is a part of the image that you want to filter to perform some operations.

The application of ROI uses the Max-shift method. In [15], the Max-shift method is used to choose the ROI on an image. The Max-shift method works by shifting the bit-plane from the selected ROI so that it occupies a higher position than the surrounding bit-plane (background), where the shift is conducted to the maximum extent resulting in the maximum quality on foreground compared to the surrounding area. ROI can also reduce the size of the data volume.

Based on these descriptions, this research offers a compression and decompression technique for video surveillance systems by applying a modified ROI method. Modifications are made by using a cropping process to the foreground by implementing the ROI method. And then compress the foreground frame so that the background does not change. The proposed method is implemented on three types of video compression format, lossy video compression in AVI, MJPEG 2000, and MPEG-4.

The rest of this article is as follows; Section II describes the related works of the proposed method. Section III describes the dataset and explains the detail of the proposed method. Section IV describes the results and discussion. The paper is concluded in Section V.

II. Related Work

In a technical review paper conducted by Seagate in 2012 showed that the compression techniques, used to store CCTV footage, can result in a decrease in quality and have little effect on video size [8]. Seagate conducted the review using a very large variation in hard drive recording capacity based on 24x7 video streaming resolutions. With many security applications requiring dozens of cameras and streaming 24x7 video, this can result in storage volumes reaching hundreds of gigabytes.

Cisco estimates that bandwidth usage for data globally will grow 71% in 2017 and forecasts an annual growth rate of 46% from 2017 to 2022 [16]. A CCTV camera connected to the internet with a stable video stream to the cloud will share a usage share of 2% and will grow seven times to increase to 3% by 2022 [17]. Based on the growth of data usage from 2019 to 2020, it is predicted in [18] that video data usage will continue to increase until 2025. It is due to the size of the recorded storage capacity depending on the hard disk installed (typically 160 GB, but some are up to 1 TB), whereas many security applications require dozens of cameras and 24x7 video streams and result in storage volumes running up to hundreds of Gigabytes.

Recordings are mostly stored on secondary storage devices such as hard disks. So, to reduce storage space, a compression technique is applied. In some research, it is recommended to use cloud storage for CCTV footage using IP Camera and storage using NVR (Network Video Recorder) [16, 19]. The limitations are that it requires a fast internet connection on each CCTV camera and requires a large bandwidth when transmitting. So, the storage method using DVR is still being used.

Research on compression to reduce data volume has been conducted by developing motion detection recording modes [20]. In [21], other providing motion detection information, the CCTV video processing also provides two other output that is face detection information and identification information. ROI based method to classify and detecting vehicles was given in [22].

In [23], the author designs a lossy and lossless algorithm by applying ROI to video compression. This algorithm design applies lossy and lossless video compression to raw video. Raw video in the form of cif or qcif is an uncompressed image file. The proposed approach is in which the selected ROI area is based on the human face (face detection). The limitation of this study is that it does not show any effect on data volume.

III. Material and Proposed Method

Fig. 1 is describing the research framework. There are two stages in this research, the compression stage, and the decompression stage. This stage is generating a modified algorithm design from the proposed research topic. The result of this algorithm design will be implemented in MATLAB software.

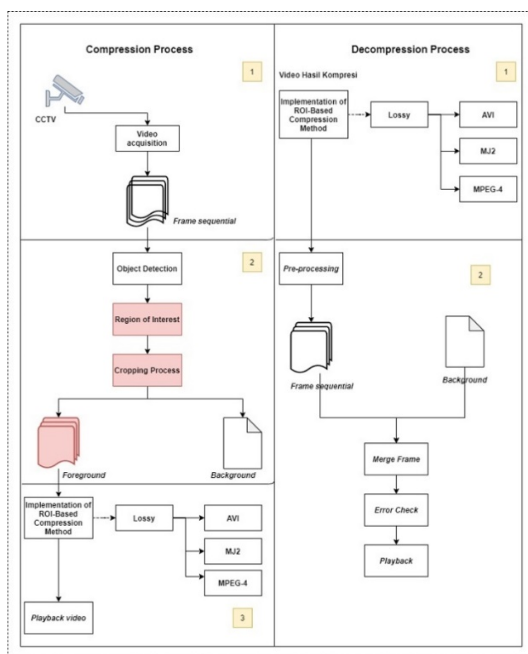

Fig. 1. The Research Framework.

The research phase begins with video acquisition. Video acquisition here is intended to collect video data used in this research. The next stage is to separate the video into some consecutive frames (sequential frames). Frames (digital images) are collection of pixels such as RGB or YCbCr. It is continued by video compression process. The difference between this proposed method and the common traditional compression method by AVI, MJPEG 2000 and MPEG-4 is that on this research the compression was implemented on the cropped frame. Based on the ROI method, the cropping process was held on the foreground frame that separates it from the background. At the compression stage, there are three stages: object detection by applying the background subtraction algorithm, followed by determining the ROI area to be separated between background and foreground, and third one is cropping process. The DCT algorithm will compress the foreground cropping results for AVI and MPEG-4 video formats and the DWT algorithm for MJPEG 2000 video format.

The decompression stage consists of three stages, namely, reading the compressed video, merging the foreground and background of each sequential frame, and generating the reconstructed video. The merging of the foreground and background of each sequential frame begins with the pre-processing stage. The pre-processing process is performed by reading the length of the video, determining the minimum and maximum of the width and height of each sequential frame. Next, separate the video data into sequential frames. After getting the sequential frames, read the background frame and combine it with the sequential frames based on its maximum width and height. Perform an error check on the mixed results to see whether each frame can be merged correctly.

A. Video Acquisition

Video acquisition is the process of capturing or scanning video so that digital video will be obtained [24]. In the video acquisition process, several factors need to be the main concern in the process. These factors are the types of acquisition tool, camera resolution, lighting techniques, zooming techniques (enlarge and reduce the camera) and the angle of data collection [25].

Each type of camera has its characteristics. Therefore, this research uses some different cameras. It consists of 1 CCTV camera and two non-CCTV cameras with different camera resolutions, i.e., 1.3 MP for CCTV cameras and 5 MP and 13 MP for Non-CCTV cameras. The video capture process is conducted simultaneously. The three cameras are placed at the height of one meter from the ground and distances two meters from the observed object. The dataset at the research location was taken at night, requiring additional artificial light. The object of this research focuses on the coordinated movement of objects that move at a certain speed that can be captured by the camera and it consists of an object only, a remote-controlled toy car.

B. Sequential Frames

Video is a collection of images that are recorded sequentially, and their movements are according to a function of time. After the video acquisition process, frames from the video will be separated [26]. The frame splitting stage begins by reading the length of the image on the video to find out how many frames the video has. Next will be determined the size of the image that is the width and height of the image.

C. Object Detection using Background Subtraction Method

After sequential frame stage, next is detecting the object. Video acquisition in this research uses a static camera so that it does not move and the point of view of taking the objects is fixed. The resulting background remains same from the beginning of the video to the end. It makes the research direction focuses on the movement of objects, so the algorithm used for object detection is the background subtraction algorithm [26].

The initial step to be conducted is to read the first frame of the video data. This initial frame will be used as the background frame. The next step is to apply edge detection using the background subtraction method to produce a foreground mask. This foreground mask will be used as a reference for the next stage. The background image is the frame that will be compared with the next frame (second frame, third frame, etc.), and then the result is reduced to get the difference. The difference will be used to detect the movement of the object. The result of image reduction is converted into a greyscale image. If there is a change in the value of a particular pixel, assign value 1 (marked in white). If there is no change in the pixel value, give the value 0 (marked in black). The difference from this greyscale image will produce a foreground mask.

D. Region of Interest

After getting the foreground mask by applying background subtraction algorithm, it will use the foreground mask as the area of interest for ROI. ROI area is marked by a bounding box, which is a box with x and y coordinates representing the width and length of the detected object [27]. Building the bounding box requires a background image, foreground image, and foreground mask. In making a bounding box, start by measuring the size in the region of the image. Then read the foreground mask generated by the background subtraction algorithm. Continue with displaying the box from the bounding box so that it will visually illustrate the separation.

E. Cropping Process

At this stage, a cropping process is conducted from the ROI results. The input image from the previous stage will be read and then cropped to get a new image. Image cropping in general, can be done based on the coordinates, the number of pixels or the results of zooming a certain area.

This research uses the foreground's width and height information to calculate the intersection point's coordinates after reading the bounding box image. The image cropping process is conducted based on the intersection point coordinates. The result of this stage is a sequential frame containing only objects and a background frame. Sequential frames have objects (foreground) that will be used as input for the compression stage, while the background frame will be saved for the decompression stage. It will merge the foreground and background later in the decompression stage.

F. Implementation of ROI-based Compression Method

This stage implements a lossy compression method in AVI, MJPEG 2000 and MPEG-4 video compression formats on the cropped foreground frame. The lossy compression method is a compression method that removes some information from the original data during the compression process, such as reducing the pixel value but does not significantly cross the limits of the visual perception of the human eye.

The AVI and MPEG-4 video formats use the DCT transformation algorithm, while the MJPEG 2000 video format uses the DWT transformation algorithm [28]. A quantization process will be applied to each video format to reduce the amount of information needed to represent the frequency so that unnecessary information can be eliminated. The entropy coding process is conducted after the quantization process is complete. The result of this process is a compressed image where for the AVI video format, it will be saved using a bitmap (.bmp) file, while for the MJPEG 2000 it will be saved using a JPEG 2000 file (.jp2) and for the MPEG-4, it will be saved using a JPEG (.jpg) file. The last step is to combine all compressed foreground frames into a video file for each video formats.

G. Implementation of ROI-based Decompression Method

The decompression process aims to restore the compressed image into an image representation like the original image [29]. It begins by reading the compressed video for three formats and then saving the information about the length of the video, minimum and maximum value of the width and height of each sequential frame. After separating, the video into a sequence of frames, combine the background and foreground for each frame. It turns the frame sequence into a video representation.

H. Cropping Process

The proposed method performance evaluation measures include compression ratio and Peak Signal-to-Noise Ratio (PSNR). The performance of a data compression algorithm can be measured by calculating the compression ratio. It measures the physical substance, how much capacity the original compressed file has. The compression ratio (CR) is defined as the ratio of the size of the original data (uncompressed) and the size of the compressed data, which is expressed by (1) [13].

$$CR = \frac{X}{x'} \quad (1)$$

where x = the size of the original data (bytes), and x' = the size of the compressed data (bytes). The compression ratio is closely related to how much memory space can be saved.

Calculating the quality of the decompressed image can be done by calculating the PSNR value. To calculate PSNR, it takes the Mean-Square Error (MSE) value given by (2) [13]:

$$MSE = \frac{1}{CM.N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [|I(i,j) - R(i,j)|]^2 \quad (2)$$

where, I (i, j) represent the original image before compression, while R (i, j) represents the compressed image, C represents the color component of the image where C=1 for binary images or grey level, C=3 for color images and the size of the image is N x M. A significant MSE value indicates that the deviation or difference between the reconstructed and original images is quite significant. Meanwhile, for calculating the PSNR value we can use (3) [13]:

$$PSNR = 10log_{10}(\frac{max_1^2}{MSE}) = 20log_{10}(\frac{max_1}{\sqrt{MSE}}) \quad (3)$$

where, max1 is the maximum value of a pixel in the original image I. The smaller the MSE noise value, the higher the PSNR value, the higher the image quality. On the other hand, the higher the MSE noise value, the smaller the PSNR value, the lower the image quality. Typical values for PSNR in the lossy image and video compression are between 30 dB and 50 dB, where higher is better [30, 31]. Values above 40 dB are usually considered very well, and values below 20 dB are usually unacceptable [31, 32]. When the two images are identical, the MSE value will be zero, which will result in infinite PSNR [31].

IV. Results and Discussion

A. Result of Video Acquisition

Table I shows the specifications of the video used. Data was taken with three cameras, one CCTV camera and two non-CCTV cameras, with different specifications to produce three video data. The video acquisition is held at night, so it requires additional lighting, which is placed in front of the object.

TABLE I. SPECIFICATIONS OF THE VIDEO USED

Video Name	Number of frames	Video file sizes
Car1	157 frames	138.87 MB
Car2	294 frames	290.70 MB
Car3	175 frames	461.43 MB

B. Result of Sequential Frames

On this stage, the results of sequence frames are saved in BMP format. For video Car1, the data size of each frame is 676 KB with the time required for sequential frame processing is 6.765 seconds. While for video Car2, the data size of each frame is 1 MB with the time required for sequential frame processing is 12.607 seconds. For the third video, the data size of each frame is 2.7 MB with the time required for sequential frame processing is 9.095 seconds. Fig. 2 gives examples of frame sequential results from three video data.

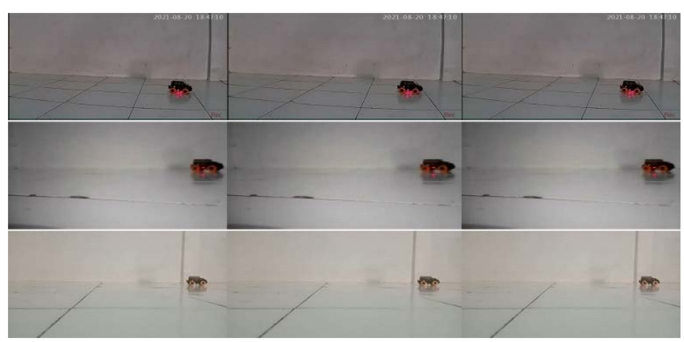

Fig. 2. Example of Frame Sequential Results.

C. Result of Object Detection using Background Subtraction Method

The background subtraction method takes the difference from the frame value by calculating it based on subtracting the current frame from the predefined background frame. Video input in sequential frames will be separated from the first frame by the next frame. The resulting image will be converted into a grey-level image. The grey-level image will be filtered and binarized to determine the foreground and background edges threshold.

Fig. 3 gives examples of object detection resulting from three video data, which shows the difference between foreground and background. The foreground is marked in white, while the background is marked in black. The time required for the object detection process in the first video is 6.763 seconds. While the time required for the object detection process in the second and third video is 13.140 seconds and 9.348 seconds, respectively.

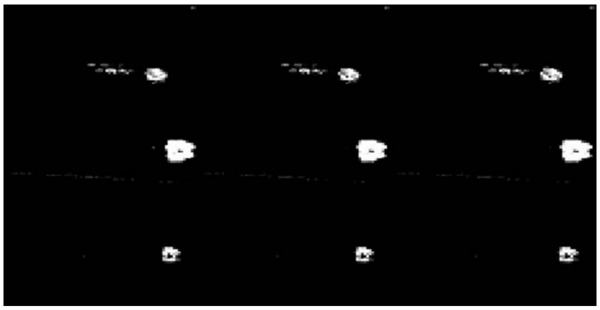

Fig. 3. Example of Object Detection Results.

D. Result of Region of Interest

Fig. 4 gives examples of the result of applying the ROI method on three video data. The yellow box can be seen visually as a separation sign of the foreground and inside the yellow box, while the background area is outside the yellow box.

E. Result of Cropping Process

The result of the bounding box process will be the input image for the cropping process. This cropped image contains information about the object of interest from each sequential frame. The cropping process will only keep the foreground from the sequential frame to produce a new image. The new image that contains information about the object will later become the input image of the compression process.

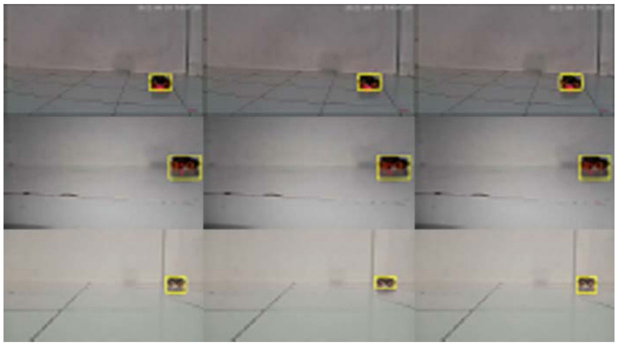

Fig. 4. Example of ROI Results.

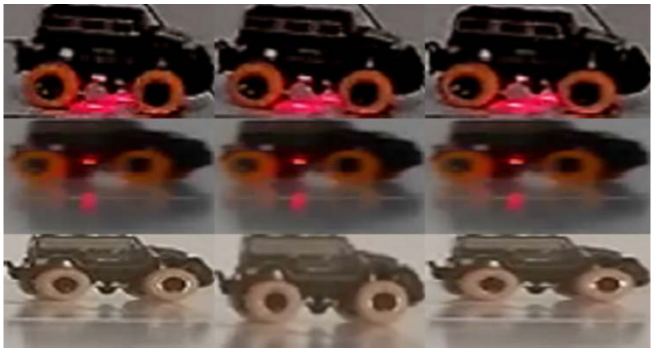

Fig. 5. Example of Cropping Process.

Fig. 5 gives examples of the cropped images from three video data. The sizes of the resulting frames have a different size according to the size of each bounding box.

F. Result of Implementation of ROI-based Compression Method

Table II shows video compression results and the time required to perform the compression process based on this proposed method. It can see in Table II that the sizes of compressed data with the ROI-based compression method are smaller than the size of the original data. For example, on the first video, Car1 with AVI format, the compressed data size is 1.50 MB, much smaller than the original video size of 138.87 MB (Table I).

G. Processing Time for Implementation of ROI-based Decompression Method

Table III shows the time that it takes to process the decompression stage. All sequential frames result from combining background and foreground frames, will then be saved in AVI, MJEPG 2000 and MPEG-4 video formats. For each video playback process, the processing time will be calculated. The reconstructed video from decompression method will be used later to count PSNR.

H. Compression Ratio and PSNR Values

Compression ratio that is calculated by (1) is defined as the ratio of the size of the original data (uncompressed) and the size of the compressed data. Table IV gives the comparison of compression ratio between the result by proposed method (with ROI) and existing method (without ROI).

TABLE II. THE RESULTS OF VIDEO COMPRESSION USING THE ROI-BASED COMPRESSION METHOD FROM THE DATA CIDEO

	Video compression format	**Size of compressed data with ROI's**	**Processing time (seconds)**
Car1	AVI	1.50 MB	0.904519
	MJPEG 2000	0.61 MB	0.978303
	MPEG-4	0.11 MB	0.981535
Car2	AVI	2.12 MB	1.837451
	MJPEG 2000	1.12 MB	1.790168
	MPEG-4	0.15 MB	1.772530
Car3	AVI	1.45 MB	3.526940
	MJPEG 2000	0.61 MB	3.324557
	MPEG-4	0.19 MB	3.473233

TABLE III. PROCESSING TIME FOR VIDEO DECOMPRESSION USING THE ROI-BASED COMPRESSION METHOD

	Processing time for AVI format videos (second)	Processing time for MJPEG 2000 format videos (second)	Processing time for MPEG-4 format videos (second)
Car1	0.321361	0.314064	0.215553
Car2	0.208899	0.193403	0.153938
Car3	0.180601	0.192206	0.112877

In Table IV, the compression ratio of the proposed method (with ROI) in all cases is higher than the compression ratio of the existing method (without ROI). So, it can claim that this proposed method works better in compressing the video data than the current method if it uses a compression ratio to measure the performance.

Table V gives the result of PSNR calculation using ROI-based compression method from three video file formats. The higher PSNR value means the higher reconstructed video quality (result of decompression process). On the other hand, the smaller the PSNR value, the lower the video quality.

TABLE IV. PERFORMANCE COMPARISON OF THE COMPRESSION RATIO FOR THE PROPOSED METHOD (WITH ROI) AND WITHOUT ROI ON VARIOUS VIDEO

	Video compression format	Original size data (a)	Size compressed data without ROI (b)	Size Compressed data with ROI (c)	Compression ratio without ROI (a/b)	Compression ratio with ROI (a/c)	Improved compression ratio (b/c)
Car1	AVI	138.87 MB	10.3 MB	1.50 MB	13.4825	92.58	6.8667
	MJPEG 2000		7.05 MB	0.61 MB	19.6979	227.6557	11.5574
	MPEG-4		0.21 MB	0.11 MB	670.7683	1262.4545	1.9091
Car2	AVI	290.70 MB	15.7 MB	2.12 MB	18.516	137.1226	7.4057
	MJPEG 2000		8.27 MB	1.12 MB	35.1511	259.5536	7.3839
	MPEG-4		0.61 MB	0.19 MB	476.5574	1530	3.2105
Car3	AVI	461.43 MB	7.71 MB	1.45 MB	59.8482	318.2276	5.3172
	MJPEG 2000		9.61 MB	0.61 MB	48.0156	756.4423	15.7541
	MPEG-4		0.53 MB	0.19 MB	870.6226	2428.5789s	2.789

TABLE V. PERFORMANCE OF THE PROPOSED METHOD BY PSNR

	PSNR for AVI format videos with ROI (dB)	PSNR for MJPEG 2000 format videos with ROI (dB)	PSNR for MPEG-4 format videos with ROI (dB)
Car1	56.9328	52.3866	64.2107
Car2	62.7869	57.8663	63.1144
Car3	51.0671	46.9997	46.8808

In Table V, all PSNR values are above 40 dB, indicating that the reconstructed video resulting from decompression is similar to the original video.

V. CONCLUSION AND FUTURE WORK

This paper proposes an ROI-based compression-decompression method that separates foreground and background on each sequential frame by the background subtraction method. Based on experiment results, for all videos that were acquired, this proposed method gives better results in compression ratio and PSNR values compared with existing methods (without ROI). That is, it increases the compression ratio and PSNR values.

The results of increasing the compression ratio for the video compression method with ROI on the AVI compression video format are 6.8667, 7.4057, and 5.3172. The results of increasing the compression ratio for the video compression method with ROI on the MJPEG 2000 compression video format are 11.5574, 7.3839, and 15.7541. At the same time, the results of increasing the compression ratio for the video compression method with ROI on the MPEG 4 compression video format are 1.9091, 3.2105, and 2.789. The ROI-based compression method can increase the compression ratio 5-7 times higher than the existing method for lossy AVI format video. While on lossy MJPEG-2000 and MPEG-4, it increases the compression ratio 7-15 times and 1-3 times, respectively.

The results of the PSNR value for the video compression method in the AVI compression video format are 56.9328, 62.7869, and 51.0671. The PSNR values for the video compression method in the MJPEG 2000 compression video format are 52.3866, 57.8663, and 46.9997. And the results of the PSNR value for the video compression method in the MPEG-4 video compression format are 64.2107, 63.1144, and 46.8808. The PSNR value for the proposed method is above 40 dB [31, 32], which indicates that the reconstructed video is similar to the original video, even though the pixel values have changed slightly.

This study only uses one object, while there can be more than one object in the field (crowd people). The camera used in this study is still, while the camera used in CCTV may be a camera that moves which results in changes in the background. So further research can be conducted to detect more than one object, moving cameras, or changes in the camera's background.

ACKNOWLEDGMENT

This work was supported by Research Program supported by the Kemendikbudristek Indonesia in Hibah Penelitian Disertasi Contract Number: 064/E4.1/AK.04.PT/2021 dated July 12, 2021, and Yayasan Pendidikan Gunadarma Jakarta Indonesia, Contract Number: 09.25/LP/UG/VII/2021 dated July 13, 2021.

REFERENCES

[1] Cumming D, Johan S. Cameras tracking shoppers: the economics of retail video surveillance. Eurasian Business 16 Review 2015; 1 (2): 235–257.

[2] Tang Y, Ma B, Yan H. Intelligent video surveillance system for elderly people living alone based on ODVS. Scientific Research Publishing 2013.

[3] Porikli F, Bremond F, Dockstader SL, Ferryman J, Hoogs A, Lovell BC, Pankanti S, Rinner B, Tu P, VenetianerPL. Video surveillance: past, present, and now the future [DSP Forum]. IEEE Signal Processing Magazine 2013; 30 (3): 190–198.

[4] Such JM, Espinosa A, Garc´ıa-Fornes A. A survey of privacy in multi-agent systems. The Knowledge Engineering Review 2014; 29 (3): 314–344.

[5] Ye Y, Ci S, Katsaggelos A, Liu Y, Qian Y. Wireless video surveillance: a survey. IEEE Access 2013; 1: 646–660.

[6] Elharrouss A, Al-Maadeed. A combined multiple action recognition and summarization for surveillance video sequences. Applied Intelligence 2021; 51 (2): 690–712.

[7] Kruegle H. CCTV Surveillance: Video Practices and Technology. Great Britain, England: Elsevier Butterworth Heinemann, 2011.

[8] Suganya D, Shyla R, Jayasudha V, Marirajan S. Storage optimization of video surveillance from CCTV camera. International Research Journal of Engineering and Technology (IRJET) 2018; 5 (3): 1356–1359.

[9] Liu H, Chen S, Kubota N. Intelligent video systems and analytics: a survey. IEEE Transactions on Industrial Informatics 2013; 9 (3): 1222–1233.

[10] Han J, Jeong D, Lee S. Analysis of the HIKVISION DVR file system. In: International Conference on Digital 18 Forensics and Cyber Crime; Seoul, South Korea; 2015. pp. 189–199.

[11] Damjanovski V. CCTV: Networking and Digital Technology. Great Britain, England: Elsevier Butterworth Heine-mann, 2005.

[12] Nihal KNAK, Shanavas ARM. A novel approach for compressing surveillance system videos. International Journal of Advanced Engineering, Management and Science 2016; 2 (2): 239374.

[13] Sarifuddin M. Pengolahan Citra dan Vidio Digital. Jakarta, Indonesia: Erlangga, 2015 (in Bahasa).

[14] Acharya T, Tsai P. JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures. Hokoben, NJ, USA: John Wiley & Sons, 2005.

[15] Tahoces P, Varela R, Lado M, Sauto M. Image compression: maxshift ROI encoding options in JPEG2000. Computer Vision and Image Understanding 2008; 109 (2): 139–145.

[16] Shin H, Jung J, Koo Y. Forecasting the video data traffic of 5G services in South Korea. Technological Forecasting and Social Change 2020; 153: 119948.

[17] Index, Cisco Visual Networking. Global Mobile Data Traffic Forecast Update, 2017–2022. USA: Cisco White Paper, 2019.

[18] Khan PW, Byun Y, Park N. A data verification system for CCTV surveillance cameras using blockchain technology in smart cities. Electronics, Multidisciplinary Digital Publishing Institute 2020; 9 (3): 484.

[19] Shidik GF, Noersasongko E, Nugraha A, Andono PN, Jumanto, Kusuma EJ. A systematic review of intelligence video surveillance: trends, techniques, frameworks, and datasets. IEEE Access 2019; 7: 170457–170473.

[20] Sagala JP, Candradewi I, Harjoko A. Penggunaan deteksi gerak untuk pengurangan ukuran data rekaman video kamera CCTV. IJEIS (Indonesian Journal of Electronics and Instrumentation Systems) 2020; 10 (1): 99–108 (in Bahasa).

[21] Nurhopipah A, Harjoko A. Motion detection and face recognition for CCTV surveillance system. IJCCS (Indonesian 40 Journal of Computing and Cybernetics Systems) 2018; 12 (2): 107–118.

[22] Pratomo AH, Kaswidjanti W, Mu'arifah S. Implementasi algoritma region of interest (ROI) untuk meningkatkan performa algoritma deteksi dan klasifikasi kendaraan. Jurnal Teknolologi Informatika dan Ilmu Komputer 2020; 7(1): 155–162 (In Bahasa with an abstract in English).

[23] Ramadoss, M. (2019). Design of Lossy and Lossless Algorithms for ROI-based Video Compression.

[24] Camastra F, Vinciarelli A. Machine Learning for Audio, Image, and Video Analysis. London: Springer, 2015.

[25] Tjin E. Kamera DSLR Itu Mudah. Jakarta, Indonesia: Bukune, 2011 (in Bahasa).

[26] Das K, Bhowmik MK, De BK, Bhattacharjee D. Background subtraction algorithm for moving object detection using SAMEER-TU dataset. In: Proceedings of Fourth International Conference on Soft Computing for Problem Solving; India; 2015. pp. 279–291.

[27] Fauzi A, Madenda S, Ernastuti, Wibowo EP, Masruriyah AFN. The importance of bounding box in motion detection. In: 2020 Fifth International Conference on Informatics and Computing (ICIC); Indonesia; 2020. pp. 1–5.

[28] Furht B, Greenberg J, Westwater R. Motion Estimation Algorithms for Video Compression. Berlin/Heidelberg, Germany: Springer Science & Business Media, 2012.

[29] Putra D. Pengolahan Citra Digital. Yogyakarta, Indonesia: Penerbit Andi, 2010 (in Bahasa).

[30] Murpratiwi SI, Widyantara, IMO. Pemilihan algoritma kompresi optimal untuk citra digital bitmap. Majalah Ilmiah Teknologi Elektro 2018; 17 (1): pp. 94-101 (in Bahasa).

[31] Bul DR. Digital picture format and representations. Communicating pictures 2014. pp. 99-132.

[32] Susanto A, Mulyono IUW, Fajar MR, Febrian GAR. A combination of hill cipher and LSB for image security. Scientific Journal of Informatics 2019; 6 (1): pp. 241-248.