

Received 16 February 2023, accepted 5 May 2023, date of publication 15 May 2023, date of current version 22 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3276243

## RESEARCH ARTICLE

# Coding Unit-Based Region of Interest Encryption in HEVC/H.265 Video

JIN-YONG YU<sup>1</sup> AND YOUNG-GAB KIM<sup>1</sup>

Department of Computer and Information Security, and Convergence Engineering for Intelligent Drone, Sejong University, Seoul 05006, South Korea

Corresponding author: Young-Gab Kim (alwaysgabi@sejong.ac.kr)

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) Grant funded by the Korea Government (MSIT) (Development of Artificial Intelligence-Based Video Security Technology and Systems for Public Infrastructure Safety) under Grant 2019-0-00231.

**ABSTRACT** The region of interest (ROI) encryption in the video can reduce the complexity of calculation and improve encryption speed by encrypting only the area containing critical visual information. Above all, ROI encryption can expand the utilization domain (e.g., video surveillance), unlike general encryption methods that de-identify the entire frame. However, the traditional ROI encryption process in high-efficiency video coding (HEVC)/H.265 is more complex than in advanced video coding/H.264, and the encrypted area tends to be wider than the ROI. Thus, cryptographic algorithms are applied outside the ROI, which wastes computing resources and restricts the visual information that needs to be provided. Therefore, this paper proposes a coding unit (CU)-based ROI encryption for HEVC/H.265 video. The proposed method selectively encrypts HEVC/H.265 parameters, such as the intra prediction mode, motion vector (MV), MV sign, transform coefficient (TC), and TC sign, which have significant visual influence, and adopts the tile concept for parallel processing frames. The CU-based ROI encryption reduces complexity by identifying the encryption area based on the CU coordinates, applying encryption only to the CUs associated with the ROI. This approach can preserve the area around the ROI by restricting the reference area. Moreover, it provides up to about 30% faster encryption speed than the traditional method while maintaining performance (i.e., peak signal-to-noise ratio and structural similarity index measure).

**INDEX TERMS** HEVC/H.265, coding unit, ROI encryption, encryption propagation.

## I. INTRODUCTION

The high-efficiency video coding (HEVC)/H.265 attracts attention in various domains (e.g., video surveillance, medical, etc.) for its high-quality video compression and fast processing speed. The HEVC/H.265 is a current video coding standard issued by the Video Coding Experts Group and the International Organization for Standards/International Electrotechnical Commission Moving Picture Experts Group. In addition, HEVC/H.265 has undergone many improvements compared with the previous standard, advanced video coding (AVC)/H.264, and compression efficiency has approximately doubled and can also process up to 8K resolution [1]. Although HEVC/H.265 video provides object

identification and a high processing speed, these advantages can cause a significant threat to privacy from broadly deployed video devices (e.g., closed-circuit television, internet protocol cameras, dash cam, etc.). Piza et al. [2] stated that anybody could be used in crimes that infringe on privacy, especially because widespread video devices can collect sensitive information about individuals. Accordingly, a protection method (e.g., encryption, intervention, limited vision, secure processing, redaction, or data hiding) is needed for visual privacy [3]. Moreover, the requirements of each environment must also be met. For example, considering that public video cameras are used in surveillance for public safety [4], [5] surveillance video should be cognizant of the circumstances while protecting privacy. Moreover, the encrypted areas of the video should be decryptable if necessary, and real-time processing is needed for a time-sensitive

The associate editor coordinating the review of this manuscript and approving it for publication was Alessia Saggese<sup>1</sup>.

response. Considering these requirements, encryption that can be applied in diverse approaches would be suitable for protecting video in various domains.

Video encryption uses two types of compression and encryption relationships [6]. One is compression-independent encryption, which may not rely on the video codec and can be applied to most video systems [7]. However, independent execution may result in a format compliance issue, and it is not generally used for real-time video processing because it delays the overall process [8]. In contrast, joint compression and encryption meet these requirements while maintaining a lower rate of quality loss in a video [9]. Consequently, many studies on video encryption have adopted joint compression and encryption, and the proposed method also benefits from this encryption type. In addition, various video encryption methods exist, such as fully layered, selective, and perceptual encryption [10]. However, the region of interest (ROI) encryption process for HEVC/H.265 is quite different from that for AVC/H.264. In particular, ROI encryption reduces computational complexity and improves encryption speed by applying an encryption algorithm to only an administrator-specified area [11]. In AVC/H.264, the encryption is conducted for a relatively similar area as for the ROI encryption, whereas encrypting areas similar to the ROI is not easy in HEVC/H.265. The primary reason for these problems is the differently sized HEVC/H.265 video processing units. In other words, unlike AVC/H.264, which has a constant macroblock size, designating an ROI for the coding unit (CU) of HEVC/H.265 is difficult, owing to the various CU sizes and its size can change every frame. These HEVC/H.265 characteristics may allow for encrypting a larger area than is necessary. Moreover, if the ROI target is dynamic, the problem becomes exacerbated, and the surrounding area cannot be preserved. Although the existing studies proposed alleviating some problems by conducting encryption at the tile level, the issues such as applying encryption to an excessive area were still not solved (see details in Section III Problem Definition). As a result, the traditional ROI encryption method in HEVC/H.265 wastes computing resources by applying an encryption algorithm to an excessive area and can-not preserve areas outside the ROI.

Therefore, this study proposes a CU-based ROI encryption method that conducts encryption at the CU level. The proposed method conducts encryption using the inclusion relationship between the ROI and CU. It is intended for real-time processing and accurate ROI encryption. Accordingly, we adopt you only look once (YOLO) v4 as the object detection algorithm to detect the ROI (i.e., face) to offer real-time processing. We then provide a method to identify the CUs related to the ROI through the detected ROI boundary coordinates and simple calculation. Moreover, the efficiency of encoding is improved by parallel processing and selectively encrypting HEVC/H.265 parameters, such as motion vector (MV), MV sign, transform coefficient (TC), TC sign, and intra prediction mode (IPM). The parameters

represent effective encryption performance because they significantly affect the visual quality of HEVC/H.265 video [12] and are encrypted through the advanced encryption standard-cipher feedback (AES-CFB) algorithm, enabling real-time encryption. Consequently, CU-based ROI encryption can improve encryption speed while retaining similar encryption performance (i.e., in terms of the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM)). It also provides good visual information about the surrounding circumstances relative to traditional methods. The proposed method is novel in that it has not been dealt with in existing studies, but there are many things to consider compared to ROI encryption at the relatively simple tile level. The main contributions of this study are as follows:

- The proposed method is the first attempt to conduct ROI encryption at the CU level. Existing studies tend to rely on functions (i.e., tile) supported by HEVC/H.265, whereas this study presents a novel approach that considers the compression procedure and its complex factors.
- Because the proposed approach is based on the coordinates of the ROI and CU, there is no need to consider the scanning order. In other words, the proposed method does not require such tools as flexible macroblock ordering (FMO).
- In terms of performance, CU-based ROI encryption improves encryption speed while maintaining values of PSNR and SSIM similar to those of traditional methods. The improved speed can free up time to apply various computationally complex encryption algorithms for enhancing PSNR and SSIM.
- The proposed method provides an enhanced encryption ratio (ER) and intersection over union (IOU) compared to traditional methods because it applies to encryption algorithms only for identified CUs. Moreover, restricting the encryption propagation at the coding tree unit (CTU) level provides improved visual information about the surrounding area.

The remainder of this paper is organized as follows. Section II explains the background and reviews the related work. Among the functions provided by HEVC/H.265, the primary factors related to CU-based ROI encryption are analyzed, and existing ROI encryption studies on HEVC/H.265 are compared with the proposed method. Section III defines problems implementing the CU-based ROI encryption and the challenges of existing studies. Section IV presents an overview of CU-based ROI encryption and describes the detailed encryption process, specifying the scope of the CUs to be encrypted. Section V evaluates the performance of the proposed method using metrics, such as PSNR, SSIM, encryption time, ER, IOU, and others, by using the test videos of various situations. Section VI discusses problems with CU-based ROI encryption and the improvements it offers. Finally, Section VII presents the conclusion and future work.

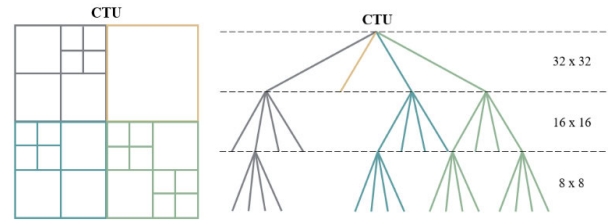
**II. RELATED WORK**

Video encryption, unlike general encryption, involves compression. Accordingly, this section classifies two subjects (i.e., HEVC/H.265 function, ROI encryption research in HEVC/H.265) and analyzes them. Subsection II-A presents an overview of the HEVC/H.265 process and describes its primary functions for CU-based ROI encryption. In Subsection II-B, existing studies on ROI encryption for HEVC/H.265 are analyzed to determine the problems with the traditional ROI encryption method. The proposed method is then compared with the traditional method and represents the result.

**A. A PRIMARY FUNCTION OF HEVC/H.265 FOR CU-BASED ROI ENCRYPTION**

It is necessary to analyze the changed HEVC/H.265 functions and processes to conduct encryption at the CU level. This Section shows the differences from AVC/H.264 and details the primary features for implementing CU-based ROI encryption.

HEVC/H.265 includes new tools and functions (e.g., coding blocks, quad-tree block partitioning process, more precise intra and inter predictions, optimized entropy coding, and the new in-loop sample adaptive offset (SAO) filter) compared with AVC/H.264 [1]. They improve the compression efficiency in the encoding process. As shown in Fig. 1, the frame is split into multiple units in the partitioning step, and the prediction is subtracted using inter or intra prediction in the prediction step. In detail, the frame is partitioned into a CTU, and each CTU contains one luma coding tree block (CTB) and two chromas CTBs. And then, subdividing CTU recursively yields CU and Coding Blocks (CBs). CUs are split into prediction units (PUs) for intra and inter prediction, and PUs are recursively divided into transform units (TUs) for residual coding. This is the main difference from the previous standard. Macroblock, the video processing unit of AVC/H.264, is generally fixed as  $16 \times 16$  and performs uniform work. On the other hand, the video processing unit (i.e., CU) of HEVC/H.265 is divided into  $8 \times 8$  to  $64 \times 64$  as shown in Fig. 2. The cause of the quadtree de-composition is a depth-first search strategy in diagnosis scanning order. That is

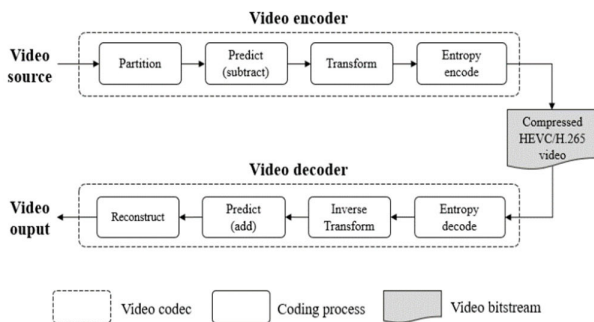


**FIGURE 2.** CTU split according to rate-distortion cost.

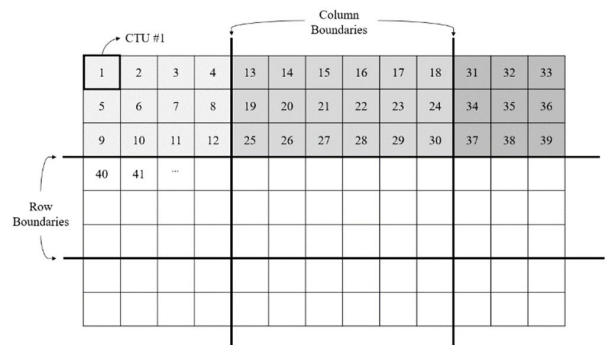
accurate rate-distortion (RD) costs are assigned to all evaluated CUs, and a split decision is made if it yields a lower RD cost than a non-split alternative during the search [13]. Afterward, transforming and quantizing the residual in the transform step. In the last process for encoding, entropy encodes the transformed output, prediction information, mode information, and headers.

Regarding the video coding efficiency, HEVC/H.265 is designed to process several parallel steps, considering the coding complexity. There are three parallel processing approaches, independent slice, tile, and wavefront, to simultaneously process multiple regions of a single frame [14]. Among them, the tile function is frequently used, and tiles are created when rows and columns intersect in the partitioning process as shown in Fig. 3. The CTU contained in the tile is scanned according to raster scanning order, and the CU split is conducted. The tiles are always aligned with CTU boundaries, and each tile contains a different number of CTUs, according to the analysis by Misra et al. [15]. The reason is that each tile's width varies to increase coding efficiency in parallel processing. In addition, a tile can be spatially more compact than a slice or other parallel functions while containing the same number of CTUs. And it does not contain headers to increase coding efficiency. Since the tile function has these advantages and can be processed independently, it is often used for existing ROI encryption.

Likewise, many studies adopt encryption within the entropy coding method. The context-adaptive binary arithmetic coding (CABAC) is a highly adaptive entropy coding engine. In AVC/H.264, context-adaptive variable length coding (CAVLC) was often used, but as video quality has



**FIGURE 1.** Overview of video encoding and decoding process in HEVC/H.265.



**FIGURE 3.** Example of  $3 \times 3$  tile division and CTU scanning order of the frame.

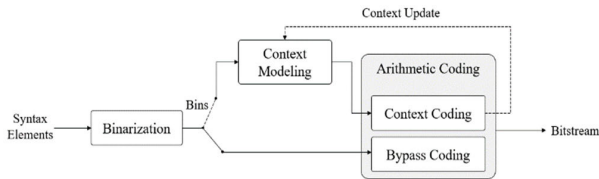


FIGURE 4. Main function of the CABAC.

improved, most entropy encoding processes adopt CABAC. As shown in Fig. 4, the main functions of the CABAC engine are binarization, context modeling, and arithmetic coding [16]. The binarization phase corresponds with syntax elements to binary symbols (i.e., bin). The probabilities of bins are updated in the context modeling phase. Finally, the bins are compressed into bits according to the estimated probabilities in the arithmetic coding phase. In the process, encryption is classified into whether to encrypt all syntax elements or selectively encrypt some syntax elements. It is generally considered safe to apply encryption to all syntax elements, but encrypting them cannot meet format compliance. Video encryption has several requirements (e.g., format compliance, encryption efficiency, compression efficiency, etc.), and any encryption method must stick to requirements. So, most studies adopting encryption within entropy coding encrypt only some syntax elements for satisfying format compliance.

## B. ROI ENCRYPTION IN HEVC/H.265

The traditional method is reviewed and compared with the proposed method by analyzing the existing studies on ROI encryption in HEVC/H.265.

Taha et al. [17], [18] conduct two studies for ROI encryption in HEVC/H.265. In [17], they present end-to-end real-time encryption of ROI in HEVC/H.265 videos. The proposed ROI encryption method employs the tile concept of HEVC/H.265 to encrypt selected tiles. Selective encryption is conducted on the syntax element of HEVC/H.265 so that the bit-stream could create the encrypted video for only ROI. In inter coding, the independence of tiles is guaranteed by restricting the MV so that only the corresponding tiles can be utilized. Finally, the proposed method can derive low bitrate and complexity overhead through Kvazaar HEVC/H.265 encoder and openHEVC decoder. In [18], they present an extension of the previous study, which covers the ROI based on chaotic-based encryption. During entropy coding, they selectively encrypt HEVC parameters MV, MV sign, TC, and TC sign at the bin string (bins) level. IPM performs encryption in isolation but ensures format compliant encryption. A keystream is created using the chaos-based encryption system, and encryption is performed through syntax elements and XOR operation. Finally, they said that the proposed method could process in real-time.

Farajallah et al. [19] conducted a privacy protection study based on the tile concept and proposed two methods for ROI encryption. The first is the encryption of all syntax elements (i.e., naive encryption) within the ROI tile, and the second

is selective encryption. Naive encryption cannot meet format compliance. On the other hand, selective encryption is performed at the bins level in CABAC for MV, MV sign, TC, and TC sign and meets format compliance. The keystream is created through AES-CFB mode, and encryption is performed through syntax elements and XOR operation. They are concerned that tile-based ROI encryption will affect too many areas, so they generate and reconstruct a single tile containing the ROI. As a result, the proposed method makes the size and position of the ROI and the tiles similar.

Tew et al. [20] encrypt only three elements for the selected CTU in a slice group. They meet format compliance by selectively encrypting the sign bin, transform skip bin, and suffix bin on bin string level. The most contribution of this study is that ROI encryption is conducted at the CTU level. If the implementation is possible, the encrypted area is expected to appear most similar to ROI among existing studies.

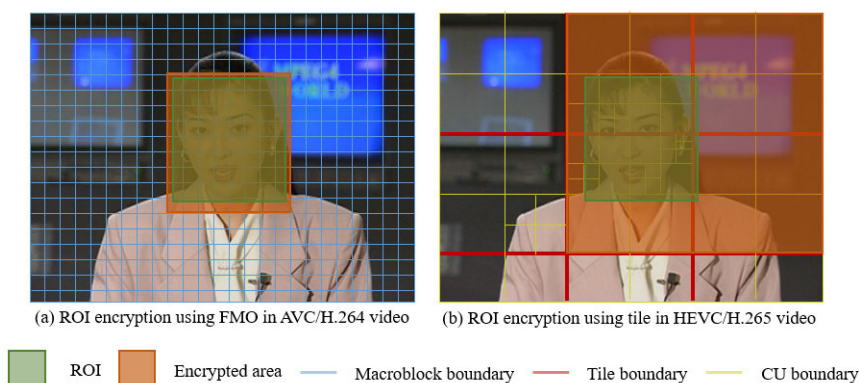
A comparison between the existing studies and the proposed method is presented in Table 1. Although there are a few related studies, such as Bergeron et al. [21], most associated studies are excluded owing to redundant content. Accordingly, we select and analyze papers detailing the process of conducting ROI encryption in HEVC/H.265. Some studies [17], [18], [19] adopted joint compression, encryption, and tiling and covered typical ROI encryption methods in HEVC/H.265. They also attempted to improve the efficiency of the encryption performance by applying different encryption methods (e.g., naive, selective, and chaotic encryption). However, they do not sufficiently explain how the tile containing the ROI is selected. The ROI encryption example defines ROI encryption ambiguously, as it encrypts a broader area than expected. Existing studies tend to treat the ROI and tiles the same. Although Farajallah et al. [19] adopted a method of re-constructing tiles that depended on the detected object, this method may reduce compression performance and processing speed because tiles must be reconstructed for every frame. Among other things, encryption at the tile level worsens encryption efficiency by increasing the ER when the ROI is located at the edge of the tile (see Section III for a detailed description). The approach used in another study [20] is the most similar in motive and purpose to this study in terms of conducting encryption in an area smaller than a tile. However, it is not easy to evaluate the use of the slice, which is less efficient than the tile. Tew et al. [20] also lack an explanation for the application of CTU encryption. First, the authors did not describe how the CTU is scanned. Second, the slice group and ROI mapping process are not proven. There is no specific explanation of how the ROI is detected or how the tile or slice group containing the ROI is specified. In AVC/H.264, the mapping description can be simplified because there is an FMO tool. However, in HEVC/H.265, this mapping method is not defined, and many similar tools are not supported. Thus, sufficient description is required for the proof. Furthermore, research studies on ROI encryption in HEVC/H.265 are rare, so it is challenging to find relevant evidence. Additionally,



**TABLE 1. Comparison of features of existing studies with the proposed method.**

	C1	C2	C3	C4	C5	C6	C7	Summarization
Taha et al. [18]	-	✓	Selective	Tile	Bins	Chaotic-based	✓	Based on previous studies (i.e., [17]), they have conducted in-depth research and proposed new chaotic-based selective ROI encryption at the tile level.
Farajallah et al. [19]	-	-	Naïve	Tile	Bins	AES-CFB	None	Two methods of encryption are applied to video to conduct encryption at the tile level. If the ROI spans multiple tiles, it is reconstructed into one tile for ROI encryption.
		✓	Selective					
Tew et al. [20]	-	✓	Selective	CTU	Bins	None	None	The proposed method conducts slice-based ROI encryption and partially encrypts the CTU included in the slice.
<b>Proposed</b>	-	✓	Selective	CU	Bins	AES-CFB	✓	Unlike the traditional method, which conducts encryption at the tile level, the proposed method conducts encryption at the CU level. In addition, it provides visual information about the video by restricting the encryption propagation at the CTU level.

C1: Compression independence, C2: Format compliance, C3: Encryption method, C4: Encryption unit, C5: Encryption domain, C6: Encryption algorithms, C7: Real-time processing



**FIGURE 5. Difference between ROI encryption between HEVC/H.265 and AVC/H.264.**

we found some studies [8], [22], [23], but these deal with compression-independent encryption types and cannot be evaluated because no corresponding comparative attribute is provided, such as the encryption method or unit. In conclusion, this study proposed a novel ROI encryption conducted at HEVC/H.265 video processing unit level, and it is the first attempt as far as we know. Moreover, the proposed method is expected to contribute to the expansion of research by suggesting new approaches to underdeveloped research areas.

### III. PROBLEM DEFINITIONS

This section defines the problems that arise in conducting ROI encryption at the CU level and the problems of existing studies.

Conducting ROI encryption at the CU level is complicated for several reasons. The main problem is that the different sizes of the video processing units in HEVC/H.265 cause difficulty in defining the area to be processed independently for the ROI. For example, as illustrated in Fig. 5, it is assumed that ROI encryption is conducted using a conventional method in AVC/H.264 and HEVC/H.265. In the case of ROI encryption in AVC/H.264, macroblocks have a constant

size of  $16 \times 16$  [24], and scanning tools, such as FMO, can be used to set an independently processable area (i.e., slice group) similar to an ROI [25], as depicted in Fig. 5(a). Unfortunately, HEVC/H.265 does not support FMO due to compression efficiency problems [15], and even if it did, it would be difficult to have a slice group size similar to that of the ROI owing to the different CU sizes. Assuming that the video is processed with  $3 \times 3$  tiles, the encryption scope can be represented as displayed in Fig. 5(b). The traditional method increases the area where the encryption algorithm is applied four-fold compared with the specified ROI. Moreover, the encryption area can significantly increase when the number of detected ROIs is large or when the ROIs border on tiles. Because this adversely affects the encryption speed and compression efficiency, a method that encrypts only the specified area is required.

After identifying the CU to which the encryption algorithm is applied, the encryption propagation should be considered for the decoded video. If a frame is encrypted as one independent area without parallel processing, encryption propagation occurs for the whole frame because the functions affecting neighboring pixels, such as inter prediction [26] or intra

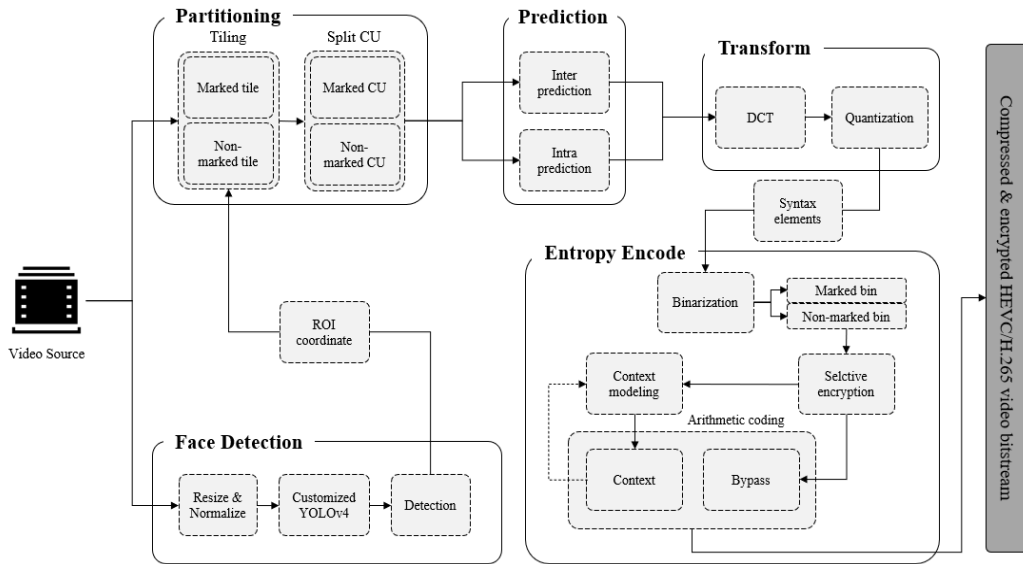


FIGURE 6. Overview of the process for CU-based ROI encryption.

prediction [27], operate over the entire image area rather than being restricted within a specific scope. Intra prediction is a function for predicting a prediction block (PB) by referring to data in the current frame and using a previously de-coded boundary sample from a neighboring block. The IPM has 33 angular modes, a planar mode, and a DC mode, and the blocks are copied to predict a PB. Inter prediction predicts the area of the current frame by referring to a partial area of the previous frame. Due to these video decoding characteristics, the encryption effect is propagated to unspecified areas; thus, appropriate measures are needed to restrict the reference scope. In this regard, the restriction includes MV, skip, and merge modes.

#### IV. CU-BASED ROI ENCRYPTION

Moreover, CU-based ROI encryption is a novel approach that can alleviate the problems found in previous studies (i.e., tile-based ROI encryption). This section describes such work as identifying the CUs to which encryption will be applied for ROI encryption at the CU level, restricting the propagation of encryption, and the encryption application process. As aforementioned, various factors were considered to conduct encryption at the CU level, and we simplified the complex process of identifying the CU corresponding to the ROI.

##### A. OVERVIEW OF THE CU-BASED ROI ENCRYPTION PROCESS

Encrypting human faces in video or images is a privacy-preserving method. The reason is that the human face usually contains considerable identifying information [28]. Thus, the ROI was set as a human face for this study. The YOLOv4 object detection algorithm was adopted to designate the ROI. This algorithm has been used in many studies because it performs well and supports real-time object detection [29]. However, faces were not included among the 80 previously

trained objects, so a new model trained on faces was required to detect the ROI. We used the dataset WIDER FACE [30] to train faces and generate a customized YOLOv4 for detecting only faces. Then, the face was detected, and the ROI boundary and coordinate were represented. The extracted ROI coordinate information was then passed to partitioning, as illustrated in Fig. 6. The proposed method conducts encryption based on the received coordinate information. First, the video source is framed and the frame is split into tiles for parallel processing. The received ROI coordinate is compared for classification to a tile (i.e., a marked tile) containing or spanning ROI coordinates and a tile (i.e., a non-marked tile) that does not. Likewise, the CU is classified into a CU (i.e., a marked CU) containing ROI coordinates and a CU (i.e., a non-marked CU) that does not. This work can classify the CUs where encryption algorithms will be applied and which CUs will not (see Section IV-B for a detailed description). Then, each CU is divided into one or more PUs, and the PU is obtained through intra or inter prediction in the prediction step. In the decoding prediction process, the prediction should be restricted so that pixels outside the specified ROI are not copied and that operations that refer to the encrypted area are not performed (see Section IV-D for a detailed description). This is because the prediction work affects the entire processed unit (e.g., tile, frame), so even if encryption is applied to a tiny area, encryption propagation may occur throughout the image. Next, discrete coefficient transformation and quantization is applied to conduct an entropy encoding step. In the entropy encoding step, the syntax elements transformed in the previous step are binarized and selectively encrypted. Syntax elements to which encryption is applied are MV, MV sign, TC, TC sign, and IPM, which significantly affect the visual quality of images among HEVC/H.265 parameters. The encryption is only applied to the marked bin corresponding to the identified

CU, and compressed and encrypted HEVC/H.265 video bitstream is derived through arithmetic coding (see Section IV-C for a detailed description).

**B. CU IDENTIFICATION FOR ENCRYPTION**

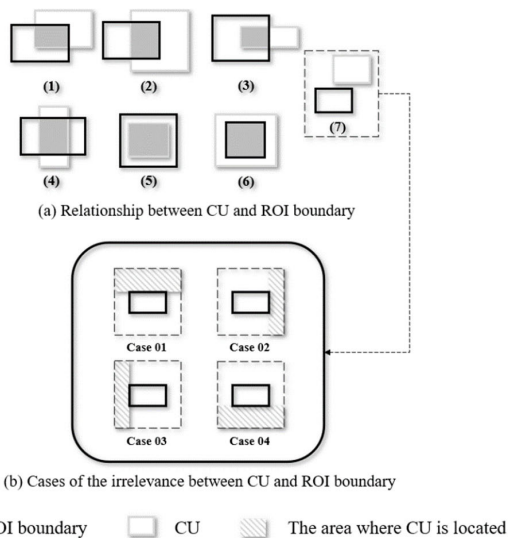
The key to the CU-based ROI encryption is identifying the CU included or spanned in the ROI at the partitioning step. The CUs identified first in the partitioning process were marked and can be distinguished from CUs that are unrelated to the ROI in subsequent processes. The CU identification method depends on whether there is a common region of two rectangles in two dimensions. The relationship between the two rectangles (i.e., the CU and ROI boundary) is indicated in Fig. 7(a), and the order of operation to reveal the relationship between CU and ROI in the partitioning process follows the CU scanning order [1]. For AVC/H.264, the scanning order is crucial to conducting ROI encryption, but the proposed method is not dependent on the scanning order method. The proposed method can identify CUs that correlate with the ROI based on the coordinates of the detected objects. In particular, we focused on instances where no relationship exists between the two rectangles to identify the CUs. More situations must be considered in cases where two rectangles span or contain six different correlations. In contrast, if no correlation exists, only a few cases must be considered. As a result, we derived the four cases where no common region exists between two rectangles, as depicted in Fig. 7(b). These correlations can also be applied to identify the relationship between the tile and the ROI. Above all, classifying the tile first, which is a relatively large unit according to correlation with ROI, can reduce the computational power in identifying CU because there is no need to confirm the relevance of the CU and the ROI for tiles that are not related to the ROI.

However, preprocessing is required about the coordinates scheme of the tiled frame for identifying tile and CU related

to ROI. The reason is that YOLO represents the ROI in absolute coordinates, whereas during HEVC/H.265 encoding, the coordinates are initialized every tile. The preprocessing need not be considered if the parallel processing function is not used. Still, the proposed method should also consider the performance (e.g., encoding time) related to encryption time because the CU-based ROI encryption is one of the joint compression and encryption types. Therefore, we made it possible to indicate the absolute coordinates instead of the relative coordinates represented by each tile and assume the following. When the absolute coordinates of the top-left corner of the CU are  $(CU_x, CU_y)$  and  $CU_{width}$  is the width and height of each CU, the coordinates of the bottom-right corner are  $(CU_x + CU_{width}, CU_y + CU_{width})$ . Similarly, when the coordinates of the top-left corner of the ROI boundary are  $(ROI_x, ROI_y)$  and the width and height are  $w$  and  $h$ , respectively, the coordinates of the bottom-right corner are  $(ROI_x + w, ROI_y + h)$ . We confirmed that the following cases are met when the ROI boundary and CU do not have a common region under this assumption. There are four cases where ROI and CU do not correlate. The criterion for identifying these cases is when the CU is located in the colored area, as depicted in Fig. 7(b). For example, in Case 1, the four sides of the CU are located below the base side of the ROI. For the area in Case 1 that spans Cases 2 and 3, if any of the conditions for each case are attained, no area is in common with the ROI. Therefore, no additional work is required to compare the CU and ROI in the overlapping area. The conditions for each case are as follows:

- ⑩ **Case 1** :  $ROI_y > CU_y + CU_{width}$  (1)
- ⑩ **Case 2** :  $ROI_x + w < CU_x$  (2)
- ⑩ **Case 3** :  $ROI_x > CU_x + CU_{width}$  (3)
- ⑩ **Case 4** :  $ROI_y + h < CU_y$  (4)

If the CU and ROI coordinates satisfy any of the above cases, they have no common area. Conversely, if the above requirements are not met, it can be assumed that the CU requires encryption. This approach can effectively identify CUs that have a common area with the ROI through a simple comparison operation, and it does not burden the system. Moreover, it does not require reliance on scanning order tools, such as FMO. In addition, the scope of the identified CU completely covers the ROI and represents fewer ERs and IOUs than conventional methods. The CU split is determined by the RD cost, as mentioned in Section II-A. The CU split usually occurs when the difference in image gradient is significant. The CU division frequently occurs around contours, corners, angles, and surface boundaries [31]. As a result, ROIs are distinct from the background, and faces generally contain subdivided CUs or have areas in common with some CUs. In addition, CU-based ROI encryption can improve the ER and IOU compared with traditional methods (i.e., tile-based ROI encryption) owing to these characteristics.



**FIGURE 7. Correlation between CU and ROI.**

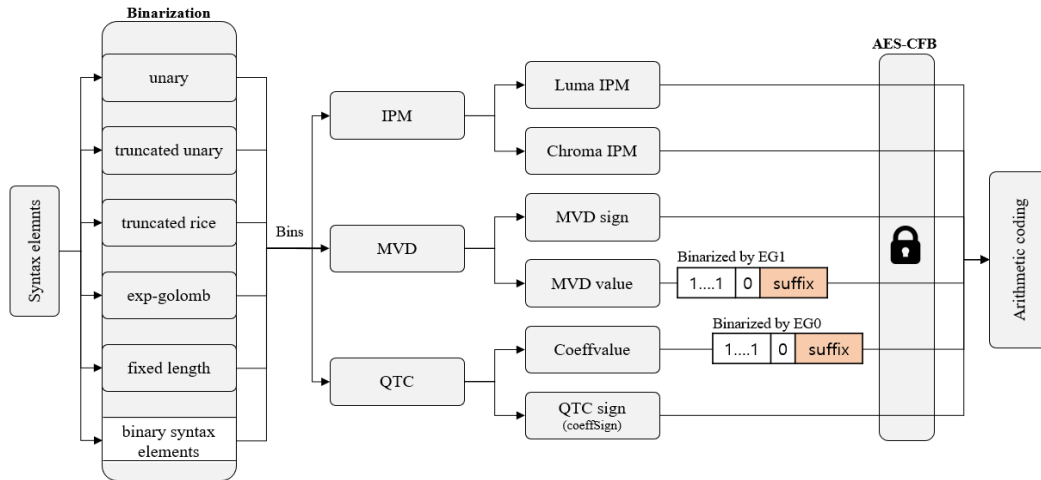


FIGURE 8. Selective syntax elements encryption in CABAC bin string.

C. SELECTIVE ENCRYPTION IN CABAC BIN STRING

The proposed method encrypts the HEVC/H.265 parameters MV, MV sign, TC, TC sign, and IPM for the identified CU during the entropy encoding process. As illustrated in Fig. 8, nonbinary syntax elements are converted to a bin string in the binarization methods, such as unary, truncated unary, fixed-length, truncated rice code, and  $k^{th}$ -order exp-Golomb (EG) codes. Then, encryption is conducted at the CABAC bin string level for the bin string corresponding to the identified CU among the binarized syntax elements. The AES-CFB is used for the real-time encryption of binarized syntax elements, and an initial vector (IV) is set through a pseudo-random number generator. The keystream  $S$  generated by the secret keys  $E_k$  and  $IV$  generates an encrypted syntax element  $C$  through an XOR operation with the syntax element  $P$ . The keystream  $S$  is as follows:

$$S_i = E_k(C_{i-1}), \quad i \geq 1 \tag{5}$$

The proposed encryption method guarantees format compliance and maintains the bit rate by encrypting the suffix part of some selected bin using the bypass mode in CABAC [32]. For example, the encryption process of the syntax element quantized TC (QTC) is that QTC is binarized by the truncated rice code and  $k^{th}$ order EG code and encrypted using a sign bit (i.e., TC sign) and a nonzero QTC value (i.e., TC) separately. The sign bit is encrypted before it is encoded in bypass mode. The value is binarized using the EG0 code and then encrypted before encoding with binary arithmetic coding. The QTC syntax element  $P(Q)$  encryption is as follows:

$$C(Q)_i = P(Q)_i \oplus S_i, \quad i \geq 1 \tag{6}$$

Encoding and encrypting the parameter MV means encoding and encrypting the syntax element MV difference (MVD). The MVD significantly influences visual quality because it closely relates to contour and motion. Therefore, the MVD exhibits high encryption performance. In addition, MVD

encryption conducts the encryption of sign and value separately, like QTC encryption. The MVD sign is encrypted through the XOR calculation and encoded in the bypass mode. The MVD value is binarized into the EG1 code, and encryption is performed by the XOR calculation of the suffix bit and keystream  $S$ . Then, it is encoded in the bypass mode. Luma IPM encryption is performed through an XOR calculation between the number of candidate mode lists and the keystream  $S$ . The coefficient scanning mode should be synchronized by mapping the position of the last coefficient because the decryption may be difficult due to IPM encryption [33]. In HEVC/H.265, there are five (i.e., planar, vertical, horizontal, DC, and corresponding luma IPM) chroma IPMs. In general, because chroma IPM encryption is affected by luma IPM encryption, the numbers in the chroma IPM list are encrypted by XOR calculation only in the unaffected cases.

D. ENCRYPTION PROPAGATION RESTRICTION

Although the CU related to the ROI was identified, and encryption was applied only to that CU, the encryption propagation appears over a wide area, as displayed in the decoded video in Fig. 9(a) and Fig. 9(b). In the prediction decoding process, references to the ROI should be restricted because intra and inter prediction propagate the encryption to areas outside the ROI by referencing the encrypted ROI. Intra prediction is used to remove spatial redundancy. Intra prediction is a tool in HEVC/H.265 that uses some data prediction spatially from region to region within a specific frame, but it has no dependence on other pictures in the video frames. In other words, intra prediction uses the previously decoded boundary samples from a spatially neighboring block to predict a new PB. The PBs might have been created by referencing encrypted blocks. Accordingly, intra prediction is restricted in the prediction step by manipulating the IPM or restricting the scope of prediction to mitigate encryption propagation. Inter prediction is used to remove temporal or spatial redundancy. The coding system searches the previously encoded video images for the image region most similar to the encoded block





FIGURE 9. The decoded video according to reference area restriction.

Sequence	ROI detection	Tile-based ROI encryption	CU-based ROI encryption
vidyo 1 (1280x720) 600 frames			
vidyo 3 (1280x720) 600 frames			
vidyo 4 (1280x720) 600 frames			

FIGURE 10. Comparison of encryption scope between traditional ROI encryption and CU-based ROI encryption.

to eliminate temporal redundancy. Once found, this block’s samples (i.e., pixels) are used as an estimate or prediction of the pixel values of the current block. The predicted samples are subtracted from the current block samples, resulting in a difference signal (i.e., residual). Likewise, referenced samples may be included in an encrypted area. Therefore, we restricted references to encrypted blocks. In particular, the motion estimation and the skip and merge modes are constrained in the encryption region boundary [26]. In the traditional method, these restrictions are conducted at the tile level, and the HEVC main profile restricts the width of tiles to a minimum of 256 pixels [1]. In contrast, the proposed method provides improved visual information relative to the traditional method by restricting the reference scope at the CTU level, as presented in Fig. 9(c).

### V. EXPERIMENTS

The Kvazaar is an academic software video encoder about HEVC/H.265 [34]. The proposed CU-based ROI encryption was implemented using Kvazaar. The processor used in this experiment was a 64-bit 16-core AMD Threadripper PRO 3955WX running at 3.90 GHz with 64 GB of main memory. The operating system was Ubuntu 20.04. The dataset for the experiment was provided by <https://media.xiph.org> [35], and each video contains one or more people. Also made

some fixes for YOLOv4. Although ROI encryption, such as the proposed method, requires perfect object detection, most object detection algorithms cannot achieve complete accuracy. Similarly, face detection using YOLOv4 failed in some frames. Still, the detection work was revised so that the coordinate information of the previous frame can be retrieved to rectify the detection failure of objects in the current frame.

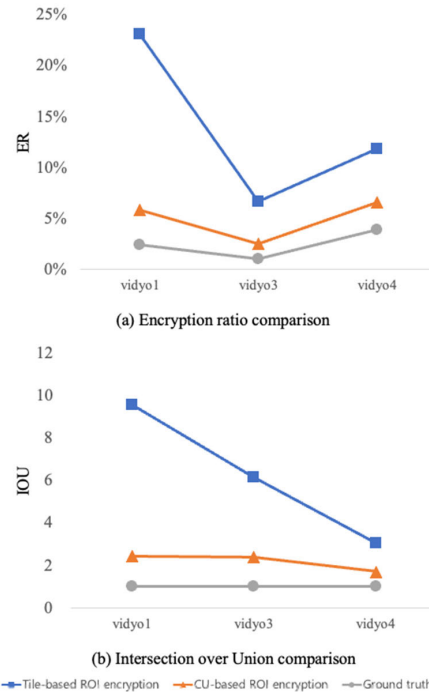
In this experiment, we compared the traditional ROI encryption with CU-based ROI encryption, as depicted in Fig. 10 using video dataset “vidyo” series. The original video displays the ROI boundary, revealing that the proposed method can represent a more similar encryption scope for the ROI than the traditional method. Whole-frame encryption cannot serve as a comparator because it encrypts all visual information within the frame, including the ROI. This method is suitable for video storage rather than providing users with real-time video for monitoring. Generally, all the test videos adopted the tiling concept for parallel processing. Most studies on ROI encryption in HEVC/H.265 have adopted the same method as the tile-based ROI encryption. After the object to be encrypted is detected, the tile containing the detected object is encrypted. The results reveal a difference between the detected ROI and the encryption scope when tile-based ROI encryption is used. The difference becomes even more pronounced when the video includes a dynamic

object or many objects. Moreover, if the size of the tile is large, in some frames where tile-based ROI encryption is applied, the encryption scope is the same as in whole-frame encryption. This result indicates that ROI encryption is not working properly and wastes resources. However, the proposed method encrypts the area similar to the detected ROI and can preserve the surrounding area more than the traditional method. These improvements can provide video that can monitor the overall situation while protecting privacy, and this video can be used as a video surveillance domain in public infrastructure.

This study focuses on how accurately the user-specified ROI should be encrypted rather than on improving encryption performance measured by the PSNR or SSIM. The encryption time was shortened while maintaining the ROI encryption performance without degradation. Table 2 presents the performance of each encryption method and includes the PSNR and SSIM. The PSNR and SSIM measured at the ROI level have similar values for each encryption method, as the same encryption algorithm was applied. The slight differences are due to the different encryption scopes. In addition, the ER and IOU were evaluated to compare the encryption methods. Whole-frame encryption is excluded because ER and IOU are not suitable for whole frame measurement. The ER is calculated as the area with encryption divided by the total frame area. For example, in a  $1280 \times 720$  resolution video, assuming that point A (500,500) to point B (600,600) is encrypted, the ER is  $(100 \times 100) / (1280 \times 720)$ . Likewise, the IOU is calculated as the encrypted area divided by the ROI. The ER and IOU of tile-based and CU-based ROI encryption are presented in Fig. 11, and the ground truth is the boundary area (i.e., the ROI) of the face detected using YOLOv4. The graph indicates that the proposed method has more compact encryption for the ROI than before and provides improved visual information about the surrounding area. The ground truth in Fig. 11(a) represents the ratio of the area occupied by the ROI in the entire frame and reveals that the proposed method has a lower encryption rate than the existing method. In addition, Fig. 11(b) refers to the degree of encryption versus the ROI. The closer the value is to the ground truth, the more compact encryption. If the value is higher than the ground truth, a wider area that includes the ROI is encrypted, and if the value is lower than the ground truth, the ROI is not properly encrypted. Accordingly, the CU-based ROI

**TABLE 2. PSNR and SSIM at ROI for whole-frame encryption, tile-based region of interest encryption, and CU-based ROI encryption.**

Sequence	Whole-frame encryption		Tile-based ROI encryption		CU-based ROI encryption	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
vidyo 1	9.36	0.21	9.49	0.23	9.29	0.19
vidyo 3	6.42	0.21	6.64	0.22	6.73	0.20
vidyo 4	8.19	0.23	8.26	0.22	8.22	0.24



**FIGURE 11. Comparison of ER and IOU for tile-based ROI encryption, CU-based ROI encryption, and ground truth.**

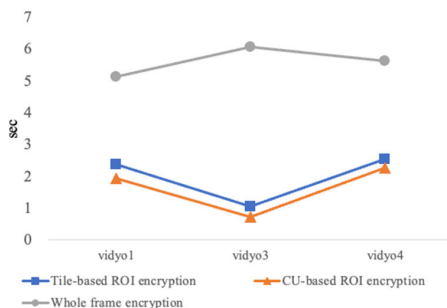
encryption significantly reduces the time required for encryption, as shown in Fig. 12. The experimental results demonstrated that the proposed method improves the encryption speed by up to 30% and provides improved visual information about areas other than the ROI while maintaining an encryption performance (i.e., PSNR and SSIM) similar to that of the traditional method.

## VI. DISCUSSION

This section discusses several issues in CU-based ROI encryption that should be considered in the future. Although improved results were obtained compared to the traditional methods for encryption speed and decoded video, the following issue analysis and solutions are sought to improve the proposed method further.

### A. OBJECT DETECTION FAILURE

Object detection specifies the encryption target and boundary in ROI encryption. Various detection algorithms have been developed, such as YOLO, that perform decently for object detection. However, perfect object detection is still challenging in all situations. Due to ROI encryption, which conducts encryption based on object location information, if object detection fails, encryption is not conducted, and visual information can be leaked as it is. Although this risk cannot be completely ruled out, countermeasures exist to alleviate it. Object detection algorithms generally learn the characteristics of various objects through a training network. In other words, the performance of an object detection algorithm depends on the variety, quantity, and quality of the data it learns. For example, in this study, we generated an algorithm to detect faces using the YOLOv4 and WIDER



**FIGURE 12.** Comparison of encryption time for whole frame encryption, tile-based ROI encryption, and CU-based ROI encryption.

FACE datasets. Although faces were detected in most frames, routine detection was difficult in some frames, such as for a facial side view, because the WIDER FACE dataset mainly consists of frontal faces. Detection failure occurred in very few cases with the test video, so it was alleviated with some improvements, but a fundamental countermeasure is required. Therefore, it is necessary to construct and train on the dataset while considering various situations.

### B. ENCRYPTION PERFORMANCE IMPROVEMENT

There is a need to investigate and analyze various algorithms and methods in terms of encryption performance (i.e., in terms of PSNR and SSIM). Unlike whole-frame encryption, which protects the entire region, ROI encryption deals with a specific area, including privacy, so the performance related to deidentification, defined by such metrics as PSNR and SSIM is important. However, although there are various algorithms for video encryption, such as RC6 [36], RSA [37], blowfish [38], and puzzle [39], existing studies have only dealt with AES-CFB and chaos-based stream cipher, as mentioned. Moreover, the proposed method has time to apply more complex encryption algorithms by improving the encryption speed relative to traditional methods. This work is expected to discover which encryption algorithm works effectively for encrypting video.

### C. COMPRESSION EFFICIENCY

The proposed method must take measures against the loss of compression efficiency. When we identified the CUs included in the ROI and performed encryption, we observed that the encryption influence propagated to the surrounding area. Although this guarantees safe interpolation and deblocking at the CTU boundary because it splits the frame by the CTU size and processes it independently, it burdens the system and reduces the compression efficiency. Conversely, encryption propagation occurs if an independent area is increased to improve encoding efficiency. The encryption propagation is still superior to the traditional method in terms of performance, but it may be difficult to use from a practical viewpoint. However, if these restrictions are conducted on each CU, the compression efficiency decreases as the resolution increases, and interpolation and deblocking problems occur. Even if it is assumed that all these problems can be solved,

to obtain a decoded video with restricted encryption propagation at the CU level, an encrypted and unencrypted bit-stream must be generated for the original video during encoding. The decoder uses the generated bitstream to predict the area inside and outside of the ROI. However, having the bitstream of the original video in the decoder is a significant security threat; thus, this method should not be considered. Therefore, it is necessary to devise an optimized CU split method to improve the compression efficiency of the proposed method.

## VII. CONCLUSION

As high-quality video processing is needed, HEVC/H.265 has added various functions to improve coding efficiency. However, the video coding improvement has made it difficult to conduct ROI encryption in HEVC/H.265 without the tile or slice function. The objective of ROI encryption in the video is to protect the confidentiality of a designated area while maintaining the surrounding scene. However, with the existing method (i.e., tile-based ROI encryption), it is difficult to encrypt only a designated area without affecting the surrounding pixels. Thus, the domains where encrypted video can be used are extremely limited. Therefore, we proposed a novel ROI encryption method to conduct encryption at the CU level. To the best of our knowledge, encrypting the ROI at the CU level is the first attempt, and it restricts the encryption scope by providing a method to identify the CUs relevant to the ROI. Through experiments, it was confirmed that the proposed method could preserve the surrounding area of ROI by only conducting encryption close to the specified ROI. It also represents an encryption area comparable to ROI encryption in AVS/H.264. The encryption time is shortened, retaining encryption performance similar to that of the existing method. Thus, the proposed method can significantly increase the scope of the domains that can be used. The traditional method is not easy to use other than for storage, but the proposed method is expected to be used in various domains by conducting partial de-identification. In terms of performance, CU-based ROI encryption performs better than the traditional method as the number or movement of objects increases or decreases. However, there are still some points that require improvement. When we tried to implement CU-based ROI encryption, the compression efficiency was reduced when some functions in the existing encoding method were restricted. In addition, this study focused on implementing ROI encryption that preserves the surrounding content, so CU-based ROI encryption must consider encryption performance in terms of metrics, such as PSNR and SSIM. Therefore, we plan to research encryption performance improvement while maintaining encoding efficiency in future work. Eventually, this study will serve as a basis for future work and is expected to be employed in various domains, such as video surveillance in public infrastructure.

## REFERENCES

- [1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Over-view of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.



- [2] E. L. Piza, B. C. Welsh, D. P. Farrington, and A. L. Thomas, "CCTV surveillance for crime prevention: A 40-year systematic review with meta-analysis," *Criminol. Public Policy*, vol. 18, no. 1, pp. 135–159, Feb. 2019.
- [3] J. R. Padilla-López, A. A. Chararoui, and F. Flórez-Revueita, "Visual privacy protection methods: A survey," *Exp. Syst. Appl.*, vol. 42, no. 9, pp. 4177–4195, Jun. 2015.
- [4] *What's Wrong With Public Video Surveillance?* Accessed: Feb. 4, 2023. [Online]. Available: <https://www.aclu.org/other/whats-wrong-public-video-surveillance>
- [5] Q. M. Rajpoot and C. D. Jensen, "Video surveillance: Privacy issues and legal compliance," in *Promoting Social Change and Democracy Through Information Technology*. Hershey, PA, USA: IGI Global, 2015, pp. 69–92.
- [6] F. Liu and H. Koenig, "A survey of video encryption algorithms," *Comput. Secur.*, vol. 29, no. 1, pp. 3–15, Feb. 2010.
- [7] F. Liu and H. Koenig, "Puzzle—An efficient, compression independent video encryption algorithm," *Multimedia Tools Appl.*, vol. 73, no. 2, pp. 715–735, Nov. 2014.
- [8] L. Duan, D. Zhang, F. Xu, and G. Cui, "A novel video encryption method based on faster R-CNN," in *Proc. Int. Conf. Comput. Sci., Electron. Commun. Eng. (CSECE)*, 2018, pp. 100–104.
- [9] N. Sethi and S. Vijay, "Comparative image encryption method analysis using new transformed-mapped technique," in *Proc. Conf. Adv. Commun. Control Syst. (CACCS)*, 2013, pp. 46–50.
- [10] J. Yu, Y. Kim, and Y. Kim, "Intelligent video data security: A survey and open challenges," *IEEE Access*, vol. 9, pp. 26948–26967, 2021.
- [11] X. Ma, W. K. Zeng, L. T. Yang, D. Zou, and H. Jin, "Lossless ROI privacy protection of H.264/AVC compressed surveillance videos," *IEEE Trans. Emerg. Topics Comput.*, vol. 4, no. 3, pp. 349–362, Jul. 2016.
- [12] W. Hamidouche, M. Farajallah, N. Sidaty, S. E. Assad, and O. Deforges, "Real-time selective video encryption based on the chaos system in scalable HEVC extension," *Signal Process., Image Commun.*, vol. 58, pp. 73–86, Oct. 2017.
- [13] A. Lemmetti, M. Viitanen, A. Mercat, and J. Vanne, "Kvazaar 2.0: Fast and efficient open-source HEVC inter encoder," in *Proc. 11th ACM Multimedia Syst. Conf.*, May 2020, pp. 237–242.
- [14] C. C. Chi, M. Alvarez-Mesa, B. Juurlink, G. Clare, F. Henry, S. Pateux, and T. Schierl, "Parallel scalability and efficiency of HEVC parallelization approaches," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1827–1838, Dec. 2012.
- [15] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth, and M. Zhou, "An overview of tiles in HEVC," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 969–977, Dec. 2013.
- [16] V. Sze and M. Budagavi, "High throughput CABAC entropy coding in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1778–1791, Dec. 2012.
- [17] M. A. Taha, N. Sidaty, W. Hamidouche, O. Dforges, J. Vanne, and M. Viitanen, "End-to-end real-time ROI-based encryption in HEVC videos," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 171–175.
- [18] M. A. Taha, W. Hamidouche, N. Sidaty, M. Viitanen, J. Vanne, S. El Assad, and O. Deforges, "Privacy protection in real time HEVC standard using chaotic system," *Cryptography*, vol. 4, no. 2, pp. 18–39, 2020.
- [19] M. Farajallah, W. Hamidouche, O. Deforges, and S. E. Assad, "ROI encryption for the HEVC coded video contents," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 3096–3100.
- [20] Y. Tew, K. Wong, and R. C.-W. Phan, "Region-of-interest encryption in HEVC compressed video," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-TW)*, May 2016, pp. 1–2.
- [21] C. Bergeron, N. Sidaty, W. Hamidouche, B. Boyadjis, J. Le Feuvre, and Y. Lim, "Real-time selective encryption solution based on ROI for MPEG—A visual identity management AF," in *Proc. 22nd Int. Conf. Digit. Signal Process. (DSP)*, Aug. 2017, pp. 1–5.
- [22] A. Shifa, M. B. Intiaz, M. N. Asghar, and M. Fleury, "Skin detection and lightweight encryption for privacy protection in real-time surveillance applications," *Image Vis. Comput.*, vol. 94, Feb. 2020, Art. no. 103859.
- [23] Y. Hu, W. Zhou, S. Zhao, Z. Chen, and W. Li, "SDM: Semantic distortion measurement for video encryption," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 764–768.
- [24] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [25] F. Peng, X.-W. Zhu, and M. Long, "An ROI privacy protection scheme for H.264 video based on FMO and chaos," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 10, pp. 1688–1699, Oct. 2013.
- [26] X. Cheng and H. Li, "Encryption system integrated with ROI and tracking for high efficiency video coding," in *Proc. Comput. Sci. Inf. Technol.*, 2014, pp. 177–184.
- [27] D. Patel, T. Lad, and D. Shah, "Review on intra-prediction in high efficiency video coding (HEVC) standard," *Int. J. Comput. Appl.*, vol. 132, no. 13, pp. 27–30, 2015.
- [28] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. La-Gendijk, and T. Toft, "Privacy-preserving face recognition," in *Proc. Int. Symp. Privacy Enhancing Technol. Symp.*, 2009, pp. 235–253.
- [29] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.
- [30] *WIDER FACE: A Face Detection Benchmark*. Accessed: Feb. 4, 2023. [Online]. Available: <http://shuoyang1213.me/WIDERFACE/>
- [31] G. Sanchez, M. Saldanha, G. Balota, B. Zatt, M. Porto, and L. Agostini, "A complexity reduction algorithm for depth maps intra prediction on the 3D-HEVC," in *Proc. IEEE Vis. Commun. Image Process. Conf.*, Dec. 2014, pp. 137–140.
- [32] Z. Shahid, M. Chaumont, and W. Puech, "Fast protection of H.264/AVC by selective encryption of CAVLC and CABAC for I and P frames," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 565–576, May 2011.
- [33] B. Boyadjis, C. Bergeron, B. Pesquet-Popescu, and F. Dufaux, "Extended selective encryption of H.264/AVC (CABAC)- and HEVC-encoded video streams," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 4, pp. 892–906, Apr. 2017.
- [34] M. Viitanen, A. Koivula, A. Lemmetti, A. Ylä-Outinen, J. Vanne, and T. D. Hämmäläinen, "Kvazaar: Open-source HEVC/H.265 encoder," in *Proc. 24th ACM Int. Conf. Multimedia*, Oct. 2016, pp. 1179–1182.
- [35] *Xiph.org Video Test Media*. Accessed: Feb. 4, 2023. [Online]. Available: <https://media.xiph.org/>
- [36] A. I. Sallam, E.-S.-M. El-Rabaie, and O. S. Faragallah, "CABAC-based selective encryption for HEVC using RC6 in different operation modes," *Multimedia Tools Appl.*, vol. 77, no. 21, pp. 28395–28416, Nov. 2018.
- [37] A. Chadha, S. Mallik, A. Chadha, R. Johar, and M. M. Roja, "Dual-layer video encryption using RSA algorithm," *Int. J. Comput. Appl.*, vol. 116, no. 1, pp. 33–40, Apr. 2015.
- [38] R. Huang and C. Lu, "Research of H.264 video transmission encryption technology based on blowfish algorithm," in *Proc. 4th Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT)*, Dec. 2015, pp. 931–935.
- [39] M. K. Lee and E. S. Jang, "Cryptanalysis of start code-based encryption method for HEVC," *IEEE Access*, vol. 9, pp. 92568–92577, 2021.



**JIN-YONG YU** received the B.E. degree in computer science from the Academic Credit Bank System, Chung-Ang University, Seoul, South Korea, in 2017. He is currently pursuing the Ph.D. degree with the Department of Computer and Information Security, and Convergence Engineering for Intelligent Drone, Sejong University. His current research interests include the Internet of Things security and video data security based on artificial intelligence.



**YOUNG-GAB KIM** received the B.S. degree in biotechnology and genetic engineering and minored in computer science and engineering and the M.S. and Ph.D. degrees in computer science and engineering from Korea University, Seoul, South Korea, in 2001, 2003, and 2006, respectively. He was an Assistant Professor with the School of Information Technology, Catholic University of Daegu. He is currently a Professor with the Department of Computer and Information

Security, and Convergence Engineering for Intelligent Drone, Sejong University. He has published over 200 research articles in the field of computer science and information security. His current research interests include the Internet of Things (IoT) security, big data security, network security, home networks, security risk analysis, and security engineering. As a Korean ISO/IEC JTC 1 Member, he is contributing to developing data exchange standards.

...