IEEE *Access*

Multidisciplinary : Rapid Review : Open Access Journal

Towards a Low Complexity Scheme for Medical Images in Scalable Video Coding

MUHAMMAD SHOAIB[1], **MUHAMMAD IMRAN**[1], **FAZLI SUBHAN**[2],
AND IFTIKHAR AHMAD[3]

[1]College of Computer and Information Sciences, King Saud University, Riyadh 11451, Saudi Arabia
[2]Department of Computer Science, National University of Modern Languages, Islamabad 44000, Pakistan
[3]Faculty of Computer and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Muhammad Imran (cimran@ksu.edu.sa)

ABSTRACT Medical imaging has become of vital importance for diagnosing diseases and conducting noninvasive procedures. Advances in eHealth applications are challenged by the fact that Digital Imaging and Communications in Medicine (DICOM) requires high-resolution images, thereby increasing their size and the associated computational complexity, particularly when these images are communicated over IP and wireless networks. Therefore, medical research requires an efficient coding technique to achieve high-quality and low-complexity images with error-resilient features. In this study, we propose an improved coding scheme that exploits the content features of encoded videos with low complexity combined with flexible macroblock ordering for error resilience. We identify the homogeneous region in which the search for optimal macroblock modes is early terminated. For non-homogeneous regions, the integration of smaller blocks is employed only if the vector difference is less than the threshold. Results confirm that the proposed technique achieves a considerable performance improvement compared with existing schemes in terms of reducing the computational complexity without compromising the bit-rate and peak signal-to-noise ratio.

INDEX TERMS eHealth, medical imaging, scalable video coding, flexible macroblock ordering, slice coding, low complexity.

I. INTRODUCTION

Recent advancements in video coding, supporting information technologies, and infrastructure are evident in telemedical services, such as electrocardiographic (ECG) communications in teleradiology [1]. High-speed data transmission is required for future multimedia service priorities. For example, in medical scenarios, a single whole-slide image file occupies >15 GB. Moreover, the accumulated file size enlarges to terabytes when multiple focal planes (Z-stack images) are included [2].

Teleradiological systems transmit medical images to medical centers for telesurgery, remote patient monitoring, and remote health diagnoses [1], [2]. This requires high-resolution multi-frame medical images to be transmitted remotely for diagnostic purposes. Therefore, efficient low complexity compression techniques are in high demand for time-sensitive healthcare services. Prominent among such services are real-time teleradiological systems, frozen section diagnoses, and dynamic telepathology.

With scalable video coding (SVC), state-of-the-art video coding standards like MPEG-4 AVC/H.264 can extract appropriate partial bit streams. This partial bit stream is vital for target bit rates, which have low temporal scalability, spatial resolution, or overall low quality. The bit stream extraction process retains a rebuilding quality appropriate to the rate of the partial bit streams [3], [4]. Nonetheless, the encoded SVC hierarchical prediction architecture with probable error propagation imposes penalties on video stream transmission. Minimal packet loss rates in macroblocks can translate into significantly high frame loss rates [5]. Hence, a compression scheme that supports healthcare communication services is critically required. In the medical field, these requirements have not received adequate attention from researchers. SVC standards also support several error-resilient tools, such as slide coding and flexible macroblock ordering (FMO) to improve robustness against errors in the bit stream. With FMO, the picture is divided into a number of slice groups;

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Hugo Albuquerque.

each is mapped by several macroblocks. Subsequently, different slice groups are assigned to prioritized areas and regions-of-interest. In this study, we accomplished using FMO for error resilience for encoding medical videos with low complexity, and identified how FMO benefits coding and error concealment efficiency.

The organization of the remainder of this paper is as follows: Section 2 presents the related work, and Section 3 briefly highlights the importance of FMO and slice coding. In Section 4, we detail supportive medical imaging in eHealth. Section 5 describes the implementation. The performance analysis is provided in Section 6. Finally, Section 7 presents our conclusion.

II. RELATED WORK

The research into video coding has attempted to augment eHealth services, particularly medical imaging. Recent advances in the Internet of Healthcare Things (IoHT), smartphones with internet and wearable healthcare devices such as body-mounted sensors e.g. the Holter heart monitor provide feasible communications between machine and patient. Systematic analysis based on healthcare surveys reveals a trend towards online systems for heart monitoring [6]. In healthcare, electroencephalograms (EEGs) are diagnostic tools to study physiological regions of the brain. The authors in [7] applied nonlinear analysis tools and models that revealed the EEG autism spectrum disorder (ASD) in children. Inspired by the ASD outcomes, various other EEG-based diagnoses were presented. Cerebral palsy is an infantile disease caused by low levels of brain oxygenation during pregnancy and at birth. In [8], the authors devised a supporting technology called REHAB FUN that developed virtual reality scenarios to help in the treatment of children with cerebral palsy.

Razzak et al. [9] discuss challenges in the development of data analytic techniques in healthcare for disease avoidance. They outline relevant algorithms commonly used for classification, clustering, anomalies detection, and association among diseases to select appropriate models. Ullah et al. [10] propose an integrated infrastructure of unmanned aerial vehicles (UAV) and body area networks (BANs). The proposed framework would collect and process health data in real-time by linking a UAV with a BAN. In [11], the authors contributed to the classification of myocardial infarction and atrial fibrillation to recognize the signal pattern for different cardiac diseases by applying deep learning techniques. In [12], the authors propose a zero-watermarking algorithm to secure the patients' identity during the transmission of their medical information via the Internet of Things. For the proposed implementation, an encrypted key is embedded in the patient's identity image.

The authors in [2] describe the provision of best healthcare services and the necessities for next-generation telehealth and telemedicine systems. This research provides super-fast broadband, ultralow latency across various telemedicine modalities to support remote healthcare information systems

and medical diagnostics. However, this architecture is not practical for patient home care services.

The authors in [13] discuss the demands of multiuser video streaming with the quality of service (QoS) over wireless networks. These researchers provide a solution for video streaming over multichannel, multi-radio, multi-hop wireless networks. They propose a distributed scheduling scheme to minimize video distortion to achieve the desired fairness level. However, this study ignores video quality degradation due to inherent transmission errors. The authors in [14] examine several QoS approaches and estimate different combinations to increase the quality of a video stream communicated over an unstable network. An error control technique is employed with a channel coder, a multiplexer and unequal error-protection schemes to overcome the losses. However, transcoding or re-encoding is then necessary to decode for diverse devices.

Cui et al. [3] proposed a fast-mode decision algorithm to expedite the SVC encoding and exploit the relationship between rate-distortion (RD) and the statistical mode decision of the enhancement layer. A limitation of this algorithm is that it cannot provide any evidence of compatibility with slice coding. Paluri et al. [15] propose a low complexity and low delay generalized model for predicting the loss of coded video slices for prioritization. They applied the proposed model for loss prediction to the cumulative mean squared error caused by the loss of different coded slices. The approach uses unequal protection for the slice. The main drawback of this model is that the approach cannot exploit local video content characteristics. Buhari et al. [16] propose a human visual system-based watermarking algorithm with low complexity to take advantage of texture features in the frame. The main disadvantage of this algorithm is the increase in computational complexity. Koziri et al. [17] explore the challenges to decrease load inequalities among various threads in slice-based parallelization for appropriate slice sizing. Their method that manipulates group of picture structures is important for low delay scenarios but its main disadvantage is that it increases computational complexity. Santos et al. [1] added bidirectional prediction support to improve the efficiency of the rate predictor's lossless encoder. The main drawback of this work is that it cannot provide solutions evidence for error-prone scenarios. Ali et al. [18] introduced a subpartition in addition to three extant partitions in the H.264/AVC codec for the coding of prioritized information. The main drawback of this work is the increase in complexity. Grois and Hadar [19] proposed coding for regions of interest, which were flexibly selected from the pre-encoded scalable video bit stream and that with adaptive settings enabled the extraction of desirable regions of interest by location, size, and resolution. Their methodology provides an efficient mechanism to meet the requirement of heterogeneous end-user devices. The main deficiency of this approach is the increase in complexity under conditions inherent to an erroneous environment.

All these methods are efficient with acceptable quality deterioration. However, the impact of communication errors along with single-encoded streams has not yet been explored, specifically from the perspective of error-resilient slice coding. The present work focuses on the transmission of high-resolution medical images in error-prone environments with low complexity by using FMO for resilience. The proposed research has the potential to deal with the challenges of low-delay teleradiology, pathological procedures for rapid microscopic sample analysis. Target medical procedures include videos for magnetic resonance imaging (MRI), endoscopy, orthopedics, bronchoscopy, laparoscopy, gynecology, ophthalmology, and gastroenterology. In the following section, we briefly discuss FMO, slice coding, SVC, and Digital Imaging and Communications in Medicine (DICOM) standards relevant to the proposed scheme.

III. FMO AND SLICE CODING

The H.264 video coding standard augments various tools used to enhance error robustness. In this paper, we focus on slice coding and error concealment efficiency. FMO is an extremely effective error-resilient tool for H.264 video coding. The key advantage of using the FMO tool is that it allows slices consisting of neighboring macroblocks. By configuring the FMO, any macroblock can be assigned freely to any slice group. The advantage of using a dispersed macroblock order is the ease with which missing blocks can be reconstructed because information from the surrounding macroblocks can be used. Consequently, the errors are dispersed evenly across the entire frame instead of being bound to a particular area. FMO has considerable prospects, particularly in error resilience. Performance is significantly improved, and FMO is recommended especially in environments that experience significant packet losses [15], [17], [18].

In slice coding, macroblocks can be generously allocated to diverse slice groups during coding. So, the decoder must be informed about the allocation of macroblocks to the slice group. This information is communicated by the macroblock allocation map which is transmitted together with the coded information of the macroblocks embedded within the picture parameter set. Given that there can be up to eight slice groups in a picture three bits are required for each macroblock to identify its associated slice group. This gives slice coding a computational aspect. In most cases, however, certain patterns appear in the macroblock allocation map. One of the benefits of a pattern is that its regular structure can often be described using a simple function with few variables. Transmitting a pattern can be reduced to transmitting the predefined pattern type (i.e., the function). This decreases the variables needed to construct the macroblock allocation map. The macroblock allocation map can often be stored in two to eight bytes.

The H.264/AVC standard offers seven options to maintain the macroblock allocation map information within the picture parameter set. The first six alternatives are patterns, whereas the seventh is used when the allocation map cannot
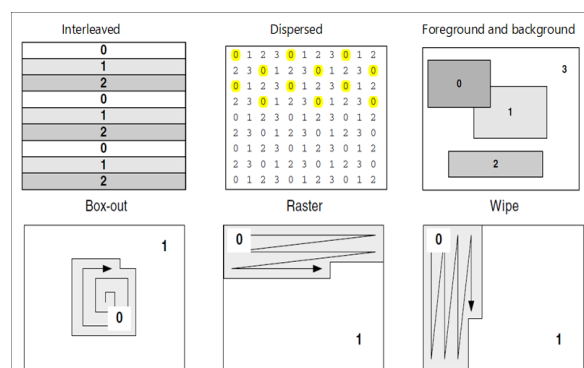

FIGURE 1. Different FMO types; 0, 1, 2, and 3 each represent a slice group.

be characterized by any of the six predefined patterns. In the subsequent discussions, we provide a brief overview of types 0 to 5. For FMO type 0, each slice group in a picture contains a maximum number of macroblocks that follow the sequential raster scan order. FMO type 1 uses a predefined dispersed pattern. The number of slice groups determines the macroblock arrangement of the allocation map. FMO type 2 is used with the region of interest where high quantization and resilience will be applied compared with the background which will have less.. FMO types 3–5 share the macroblocks over two different slice groups and are commonly known as evolving types. Fig. 1 shows the different FMO types.

IV. SVC IN SUPPORTING MEDICAL IMAGING

The SVC standard is an extension of MPEG-4 H.264/AVC video coding. The Joint Video Team from the VCEG and MPEG jointly standardized SVC [4]. Encoded SVC bit streams can transmit and decode a partial bit stream by using feature extraction. The subsequent (decoded) video has low-temporal, low-special-resolution, or low-quality scalabilities. The video reconstructed of partial bit streams is comparable to the single-layer H.264 design. Hence, SVC can provide a solution for medical imaging for heterogeneous networks and receiving devices that have an adaptation capability.

SVC encoding is comprised of intra/interlayer coding and *P and B* hierarchical prediction structures. SVC supports a flexible format for the real-time transport protocol interface. The bit stream scalability allows an adaptive bitrate for media streaming without requiring transcoding or re-encoding. The frame rate, special resolution, and picture quality variation in the IP network can also be managed easily for varying network conditions. The inherent characteristics of SVC, which include transmission robustness, compression efficiency, and flexibility for the inter-operability of heterogeneous modalities and networks, make it a good choice for eHealth imaging. However, these medical images follow certain health standards regarding format, structure, and transmission as discussed in the following paragraphs.

DICOM is a standard for distributing, exchanging, and viewing medical images required for communication between medical imaging equipment [20], [21]. The objectives of DICOM are to achieve compatibility and improve

workflow between various modalities and imaging systems in healthcare. Working Group 13 of the DICOM Committee has developed a new enhancement for MPEG-4 H.264/AVC encoded video sequences [22]. However, some healthcare applications in DICOM, such as diagnostic videos, still require functional development to integrate with the professional profile of MPEG-4 H.264/AVC.

eHealth research requires an efficient transmission scheme to transfer patient diagnostic images and videos. To address the evolving needs of medical imaging, an improved coding scheme and framework are presented using the scalable extension of MPEG-4 H.264/AVC. The proposed scheme can support the emerging requirements of the DICOM standard, such as high resolution, color sampling, and bit depth.

V. PROPOSED IMPLEMENTATION

The feature contents of video frames in H.264/AVC can generally be detected in macroblock mode by motion vector mapping. In this research, we expand our earlier research to evaluate the motion activity in the macroblock to identify the background [23]. We also extend our scheme using the FMO in slice coding mode and measure its performance.

The partition types in the background regions always contain 8×16, 16×8, and 16×16 partitions. Conversely, the partition types for macroblocks (MBs) in the foreground regions involve 4×4, 4×8, and 8×8 partitions with active motion or textures. Identical mode scattering is also found in the base layer frames of SVC [24], [25].

Figure 2 shows MB size and the partition and subpartitions in a frame from a stomach ulcer endoscopy. The image in Fig. 2(a) is developed from the bit stream that depicts the macroblock size in the picture. As shown in Fig. 2(b), the MB shows diverse grid sizes signifying different partitions and subpartitions for the 16×16, 16×8, 8×16, 8×8, 4×8, and 4×4 MBs. The red arrows in Fig. 3(a) depict active motions due to interpredictions from forward and backward reference frames. The motion vectors in the video frame are largely clustered in the active regions. Contrastingly, the red arrows and differently sized partitions in Fig. 3(b) show the relationship between the macroblock partitions and sub partitions and the motion vectors in a frame.

Figure 4(a) shows macroblock modes computed by optimized RD cost for each partition and subpartition (red, blue, and the yellow dots represent MB_Intra, MB_Inter, and MB_SKIP, respectively). In Fig. 4(b) the blue lines represent the slice boundaries in a frame. Those MB mode sub partitions with 4×4, 4×8, 8×4, and 8×8 are the active motion referred as foreground region, whereas the 8×16, 16×8, and 16×16 partitions with skip modes, are the background region in the video frame.

A. MOTION VECTORS

To estimate the mode and its motion features, we analyze the motion activity computed by the following equation:

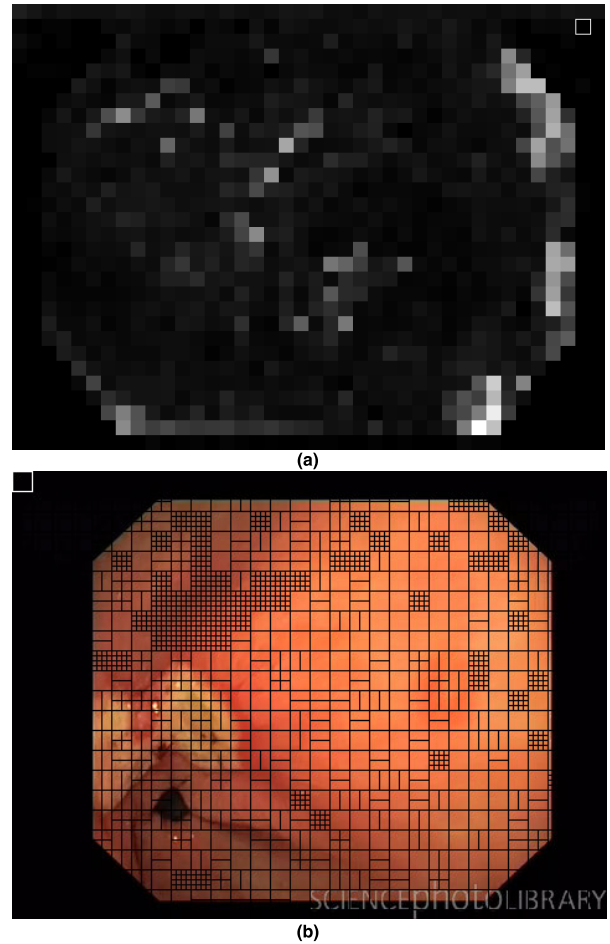$$L = \|V_1 - V_2\|_2^2 \tag{1}$$

(a)

(b)

FIGURE 2. Details of stomach ulcer endoscopy sequence. (a) Macroblock size in a frame. (b) Macroblock partition and subpartitions in a frame.

where L symbolizes the motion vector activity related to the base layer MB; $(\|.\|_2)$ denotes the l_2 - norm of the motion vector difference; and V_1 and V_2 denote the motion vectors of the two adjacent blocks, respectively. We evaluate the similarity between V_1 and V_2 by the Euclidean distance separating them.

B. MB MODE PARAMETER

Li *et al.* [24] investigated the correlation between mode distributions in the base and its interrelated enhancement layers MBs. For SVC spatial scalability, the same trends are present in the MB mode partitions in the base layer and its upsampled MBs in the enhancement layers. For SVC temporal scalability, the MBs' mode partition in the frame is almost identical to its reference frames. This relationship strongly correlates to the background in the enhancement layers. These findings inspired us to develop the MB-mode parameter (α) from the mode setting of the MBs in the base layer. The motion and texture features of the corresponding MBs in the higher layers are measured by α.

In extended special scalability (ESS), generally up to four MBs in the base-layer may be required for a MB in an
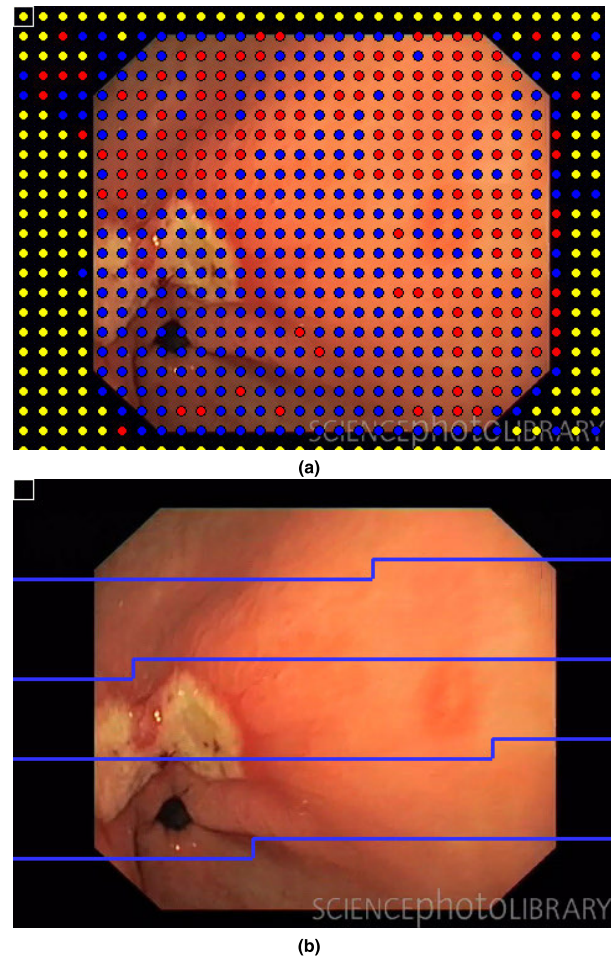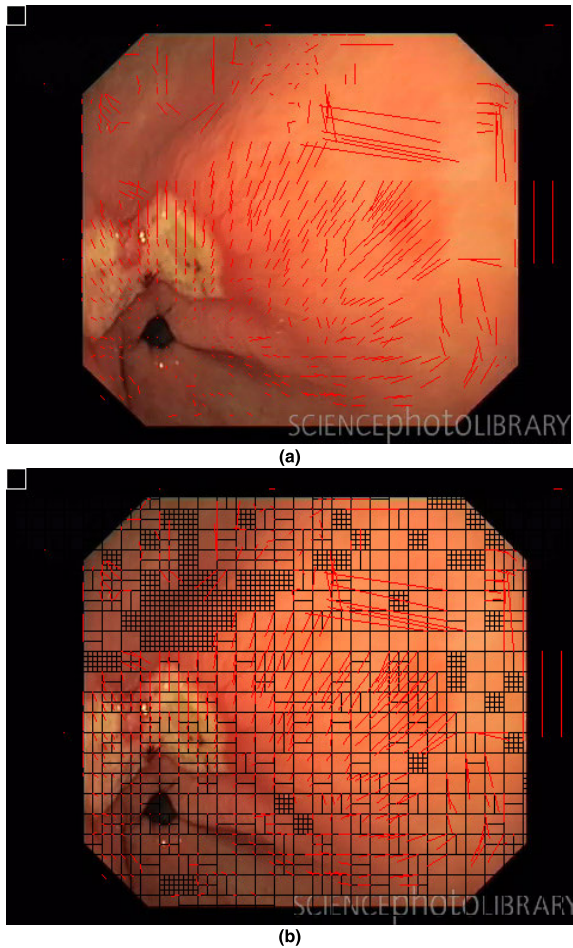
FIGURE 3. Illustration of motion vectors and macroblock partitions. (a) Motion vectors shows with red lines in a frame. (b) Relationship between macroblock partition and subpartitions with motion vectors in a frame.

FIGURE 4. Illustration of macroblock types and slice boundaries. (a) Different types of macroblock represented with yellow, red, and blue dots. (b) Slice boundaries in a frame represents with blue lines.

enhancement layer [26]. Fig. 5 shows two successive spatial layers: the base (right side) and the enhancement layers (left side; shown with blue grids are the enlarged up-sampled base layer MB "C"). W_{base}, H_{base}, and W_{enh}, H_{enh} symbolize the (width, height) of the base layer and related enhancement layer frames, respectively. The base layer frame is a sub-sampled form positioned at $W_{extract}$ and $H_{extract}$ partially or completely inside the enhancement layer frame, located at (x_{orig}, y_{orig}) coordinates. $W_{extract}/W_{base}$ and $H_{extract}/H_{base}$ are the upsampling components between the base layer frame and the extracted area in the enhancement layer frame.

Fig. 5 shows the existing enhancement layer MB (MB^e) and the upsampled base layer MB denoted by uMB^b (block C with a blue grid), in which the central projection of MB^e is located. In ESS, one, two, or four uMB^bs can cover MB^e. Blocks A, B, C, and D (outlined by a blue grid) are uMB^b blocks a, b, c, and d (outlined by a black grid), respectively. The uMB^b "C" is surrounded by the uMB^bs A, B, and D.

In Fig. 5, uMB^b_n (e.g., A) is the bordering upsampled MB of uMB^b, which coincides with the current MB^e. The assigned weightage of the MB mode factor (M) for uMB^b_n is shown
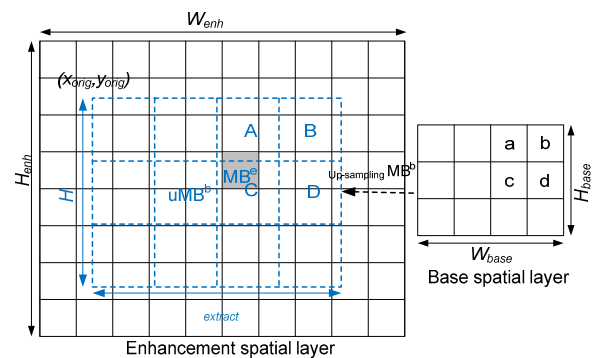

FIGURE 5. Illustration of the upsampled enhancement layer macroblock.

in Table 1. The Area factor (A) is the ratio of the coincided area of uMB^b and MB^e to the area of MB^e itself. In [27], the enhancement layer MBs are categorized into four types, namely: center, vertical, corner, and horizontal. Based on the MB^e characteristics, we computed the MB^e mode parameter (α) as follows:

1) When MB^e corresponds to the corner, uMB^b has no adjacent MBs, and the n of uMB^b_n is 0; α is derived as

TABLE 1. Mode factors weightage.

Mb mode	SKIP	16×16	16×8	8×16	8×8	8×8 ref 0	Intra_4×4
M	0	1	2	3	4	5	6
Block mode	8×8	8×4	4×8	4×4	SKIP		
M	8	9	10	11	0		

follows:

$$\alpha = \left(A_0{}^*M_0\right) \quad \text{for } MB^e \in Corner \tag{2}$$

2) When MB^e corresponds to the vertical or horizontal, uMB^b has one adjacent uMB_n^b (i.e., $n = 1$); α is computed as follows:

$$\alpha = \left(\frac{A_0{}^*M_0 + A_n{}^*M_n}{1+n}\right) \quad \text{for } MB^e$$
$$\in Horizontal/Vertical \tag{3}$$

3) When MB^e corresponds to the center, uMB^b has three neighbors uMB_n^b (i.e., $n = 3$); α is computed as follows:

$$\alpha = \left(A_0{}^*M_0 + \sum_3 \frac{A_n{}^*M_n}{1+n}\right) \quad \text{for } MB^e \in Center \tag{4}$$

Hence, a patient's images can be coded with a better peak signal-to-noise ratio gain (ΔPSNR) and low complexity by applying the block-combining technique.

C. MODE DECISION IN BACKGROUND

Distinct mode decision techniques are applied to derive MB^e in the background or active areas based on the MB mode and its active motion characteristics. In general, the higher the weightage of mode factor M the MB complexity will increase. As shown in Table 1, a mode factor with a high M weightage is allocated to smaller MB mode partitions and vice versa. For $n = 0$, the mode factor M is 0 (skip mode) as shown in Table 1, and the area factor A is allocated as 1. To decide whether MB^e matches to the background or foreground regions, we set the threshold T_H to α. The criterion is given as:

$$MB^e \in background \quad \text{if } \alpha \le T_H$$
$$MB^e \in active \quad \text{if } \alpha > T_H \tag{5}$$

where $T_H = 0.36$, as computed empirically. Since MB^e exists in a background region, the MB^e partition type is directly set to be a large partition (16×16). Hence, the derivation of the MB partition type is interrupted.

D. MB MODE PARTITION IN AN ACTIVE REGION

Different method is employed to compute the MB mode in an active or rich texture region. In [26], the authors suggested an alternate technique in which two smaller subpartitions were combined if they shared similar motion vectors with the
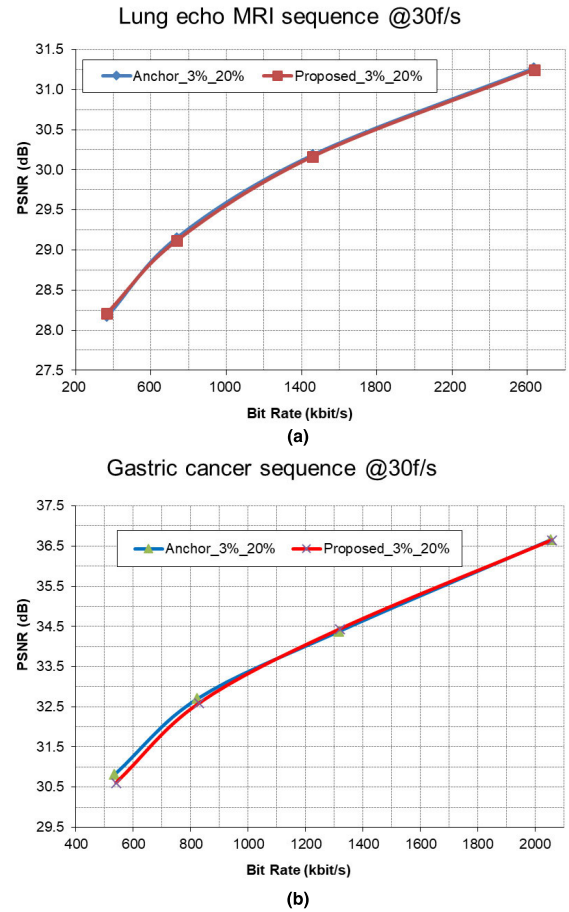

FIGURE 6. RD graphs of proposed versus anchor schemes for a wireless network (a) MRI of lung echo sequence and (b) gastric cancer sequence.

same reference index. In [25], the researchers showed that the motion vectors of contiguous blocks were similar when they corresponded to the same feature contents. Exploiting these conclusions, we can combine two bordering smaller blocks into one large block, even if their MV absolute difference is larger than the MV threshold set in [26]. Hence, we propose an alternative solution for block combining in active regions as:

$$\left(\alpha{}^*L\right) \le T_{active} \tag{6}$$

where L denotes the activity calculated using (1); α is the MB-mode parameter calculated using either (2), (3), or (4) based on the content features of MB^e. $T_{active} = 4$ is set as an experimental threshold. Given that T_{active} is a static value, the correlation between V_1 and V_2 is adjusted to α, which reveals the motion and texture features of MB^e. It reveals that α is low for background regions. By contrast, α is high for active motion or rich texture regions.

VI. RESULT AND ANALYSIS

To validate the performance of the proposed scheme, we used available high-quality video sequences. In raw format,
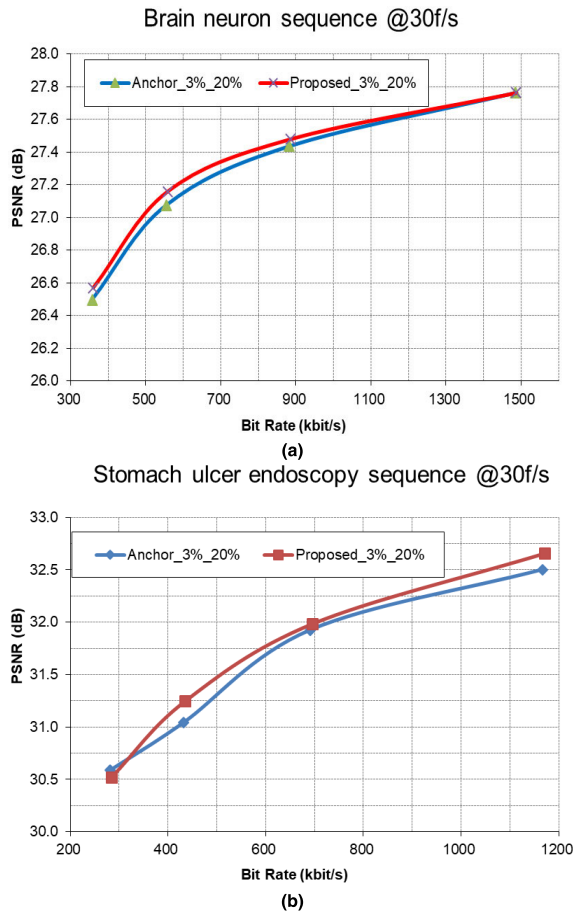
FIGURE 7. RD graphs of proposed versus anchor schemes for a wireless network (a) brain neuron and (b) stomach ulcer endoscopy sequences.
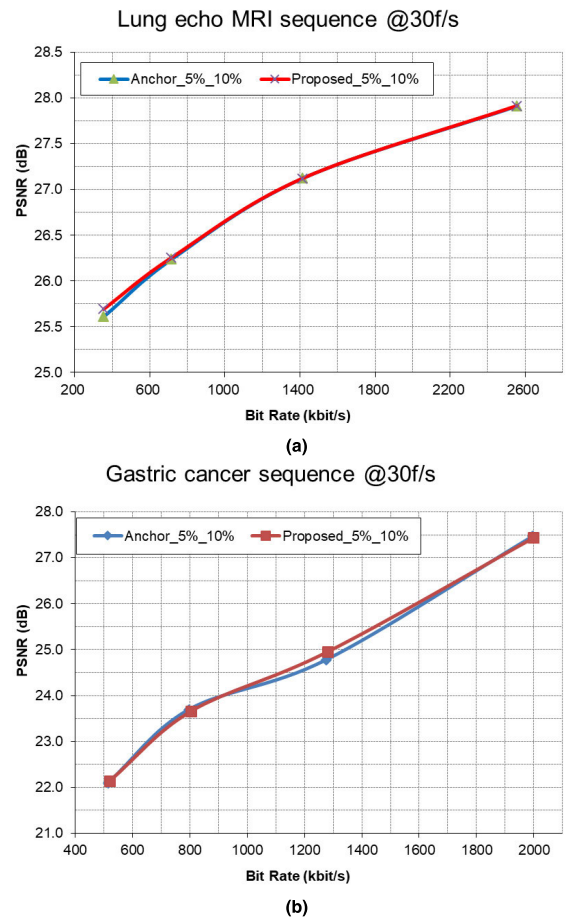

FIGURE 8. RD graphs of proposed versus anchor schemes for an IP network (a) MRI of lung echo and (b) gastric cancer sequence.

medical videos with high quality are usually inaccessible due to the privacy of patients' records. This research was conducted on several sharable medical video sequences. The testing platform was an Intel®Core i5, 2.5 GHz with 4 GB RAM and the Windows 10 operating system.

The research objective was to analyze the computational complexity and visual quality of the proposed approach in comparison with other contemporary schemes. The investigation included RD due to compression as well as distortion from packetized streams passing through error-prone communication channels. Our simulation used two commonly error-prone transmission scenarios, namely, *IP networks* and *wireless networks*. Errors exhibited in wireless networks are confined to small regions in a frame, and are susceptible to fading, attenuation, multiuser interference, and shadowing. Errors exhibited in IP networks are limited to the large regions of the frame. Transmissions through IP networks attempt best-effort service. The packet losses mainly occur in a congested network at intermediate nodes due to buffer overflow. Moreover, transmission errors in wireless and IP networks are also transitory and appear as glitches in the video.

A. BIT STREAM GENERATION

In this research, four compressed videos were downloaded and transformed into raw (YUV format) video sequences: (i) MRI sequence [28], (ii) gastric cancer sequence [29], (iii) brain neuron sequence [30], and (iv) stomach ulcer sequence [31], as mentioned in Table 2. High-profile H.264/AVC configuration and testing conditions were applied for the experiment. The success of the proposed scheme would be determined by the PSNR-Y (dB) versus the bit rate savings (kbps). The well-accepted measurements [26], [32] were also used to evaluate success, i.e. the percentage change in the MB partition and subpartitions (ΔMB), which is calculated as follows:

$$\Delta MB = \frac{MB_{pro} - MB_{orig}}{MB_{pro}} \times 100\% \tag{7}$$

where MB_{pro} represents the total number of MB partitions and subpartitions for the proposed scheme, and MB_{orig} denotes the total number of MB partitions and subpartitions for the anchor as defined in the Joint Scalable Video Model (JSVM). The positive value in ΔMB for large partitions (16×16, 16×8) and negative values in ΔMB for the 8×8,
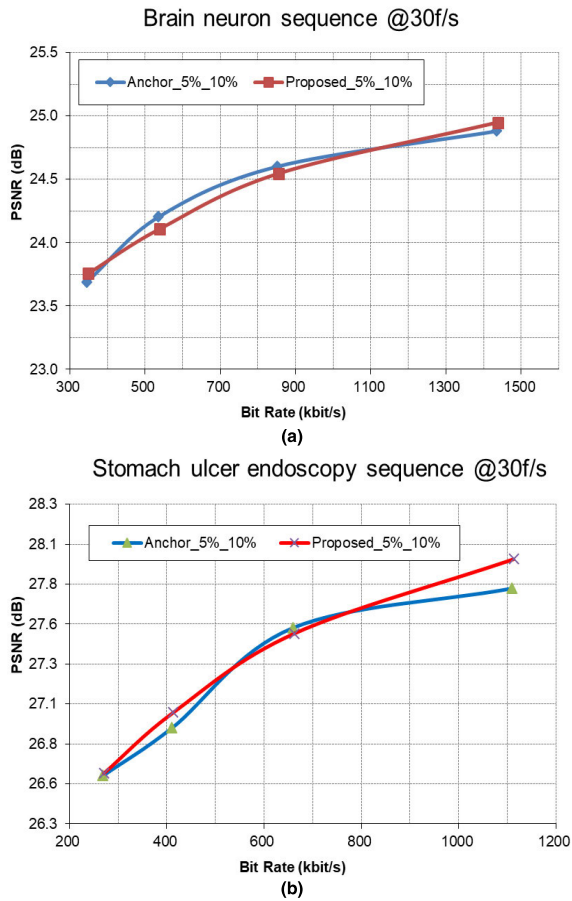
(a)

(a)

(b)

(b)

(c)

FIGURE 9. RD graphs of proposed versus anchor schemes for an IP network (a) brain neuron and (b) stomach ulcer endoscopy sequences.
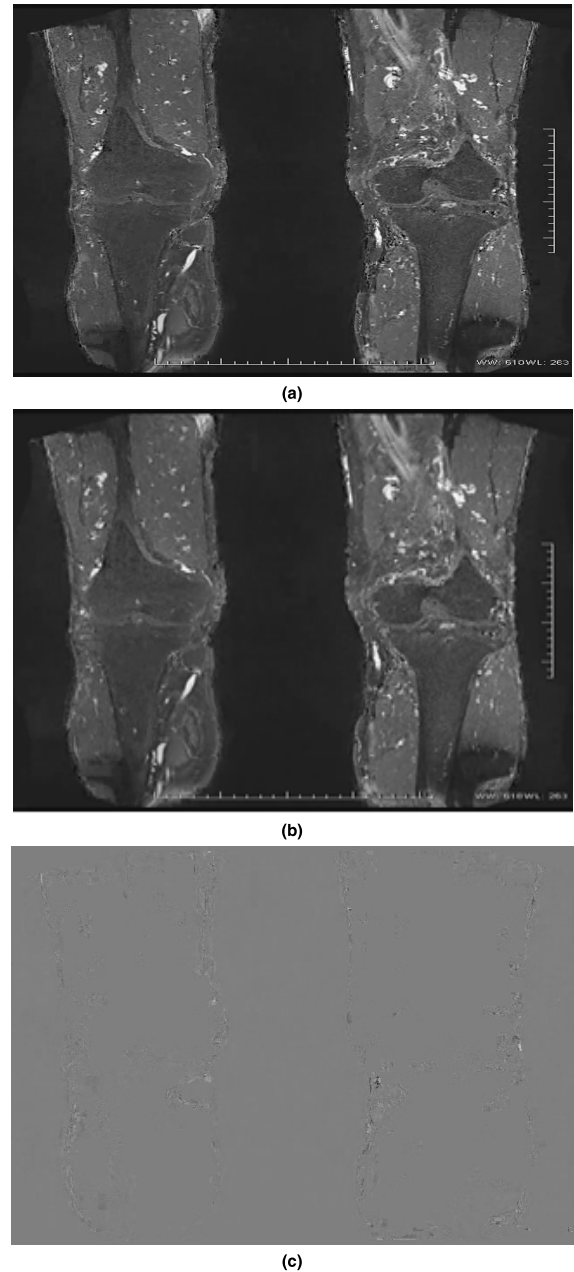
FIGURE 10. Subjective comparisons for MRI lung echo sequence frames transmitted through wireless network. (a) Image decoded with the anchor algorithm. (b) Pictures decoded with the proposed schemes. (c) Subjective difference.

8×4, 4×8, and 4×4 macroblocks reflect a decrease in computational complexity.

B. SIMULATION CONDITIONS

In this experiment, four error patterns (3%, 5%, 10%, and 20%) with average packet loss rates (PLRs) were used as recommended [33]. In this research, the important parameters were set at a 4:2:0 (progressive) video format, group of picture structure with an intra-coded and predicted frame pattern, YUV color spacing, and the CABAC symbolic mode. The error pattern in the simulator dropped random packets as found in the error pattern from an SVC-compressed bit stream. Then, the output H.264 bit stream was decoded, and the errors were recovered using the concealment tool as defined in the JSVM [32]. Other testing settings were as follows:

1) In this experiment, we used DICOM compliant imaging, video graphics array (VGA) with a resolution of (640×480) for the base layer and a high-definition (HD) resolution of (1280×720) for the enhancement layer with a (16:9) aspect ratio [21].
2) Lower PLRs occurred in the base layer than in the enhancement layer.

3) The packet loss simulator enforced a 3% PLR for the base layer bit stream and a 20% PLR for the enhancement layer bit stream —a condition that applies to error-prone wireless networks [34].
4) For error-prone IP networks, the simulator enforced a 5% PLR for the base layer bit stream and a 10% PLR for the enhancement layer bit stream.
5) Other relevant main configuration parameters are listed in Table 3. The simulation testing consisted of FMO used for error resilience to measure performance at low complexity.
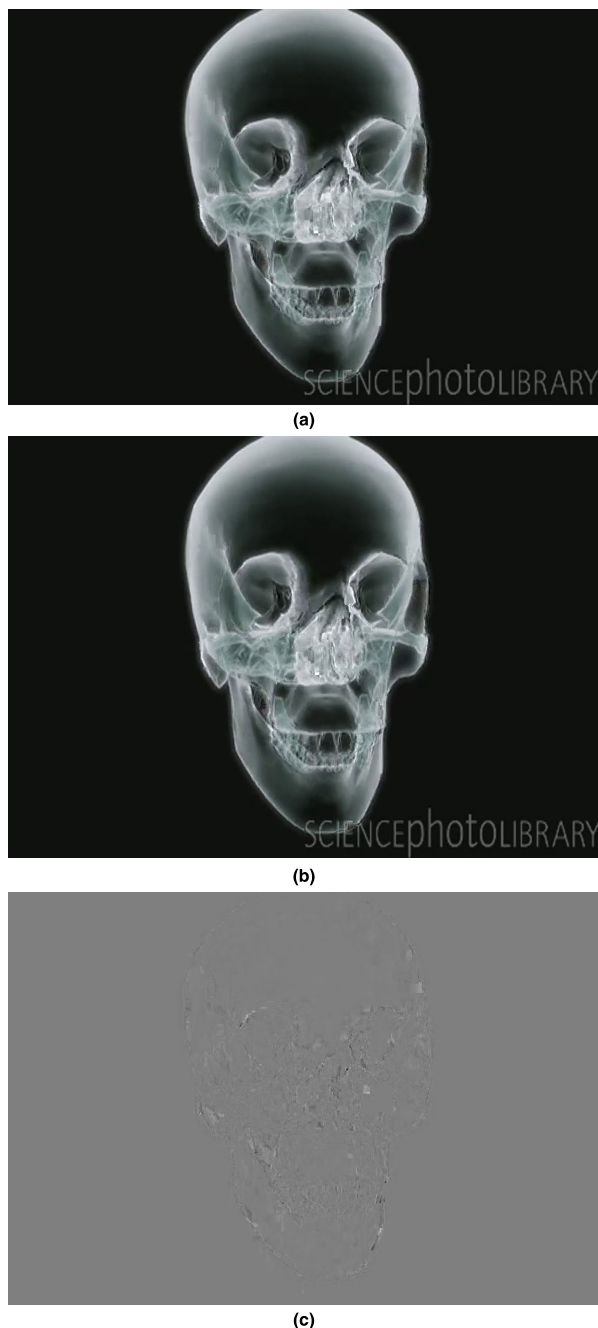
(a)

(b)

(c)

FIGURE 11. Subjective comparisons for brain neuron sequence frames transmitted through wireless network. (a) Image decoded with the anchor algorithm. (b) Pictures decoded with the proposed schemes. (c) Subjective difference.

(a)

(b)

(c)

FIGURE 12. Subjective comparisons for gastric cancer sequence frames transmitted through IP network. (a) Image decoded with the anchor algorithm. (b) Pictures decoded with the proposed schemes. (c) Subjective difference.

C. OBJECTIVE ACHIEVEMENT

To evaluate the objective achievement of proposed scheme, we increased the PSNR and block merging i.e. the number of subpartitions converged to large partitions. We analyzed the success of the 'proposed' method compared to JSVM [26], [32], which is specified as 'anchor' for concealed decoded sequences.

Figure 6(a) demonstrates the RD for lung echo in the MRI sequence for the proposed and anchor methods, with HD resolutions for the wireless network. Figure 6(b) shows RD curves for the proposed and anchor schemes of the gastric cancer sequence. Figures 7(a) and (b) show the RD curves for the proposed and anchor schemes for brain neuron and stomach ulcer endoscopy, respectively. Our approach matched or slightly increased the image quality with regard to coding efficiency and PSNR gain.
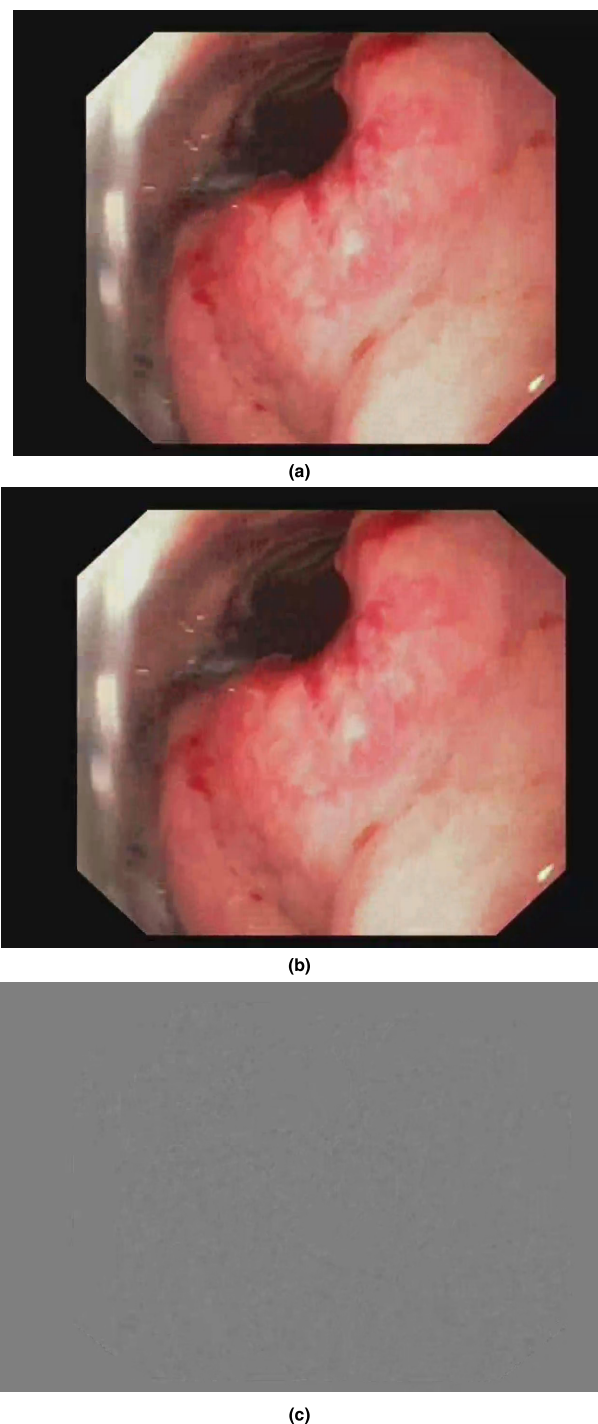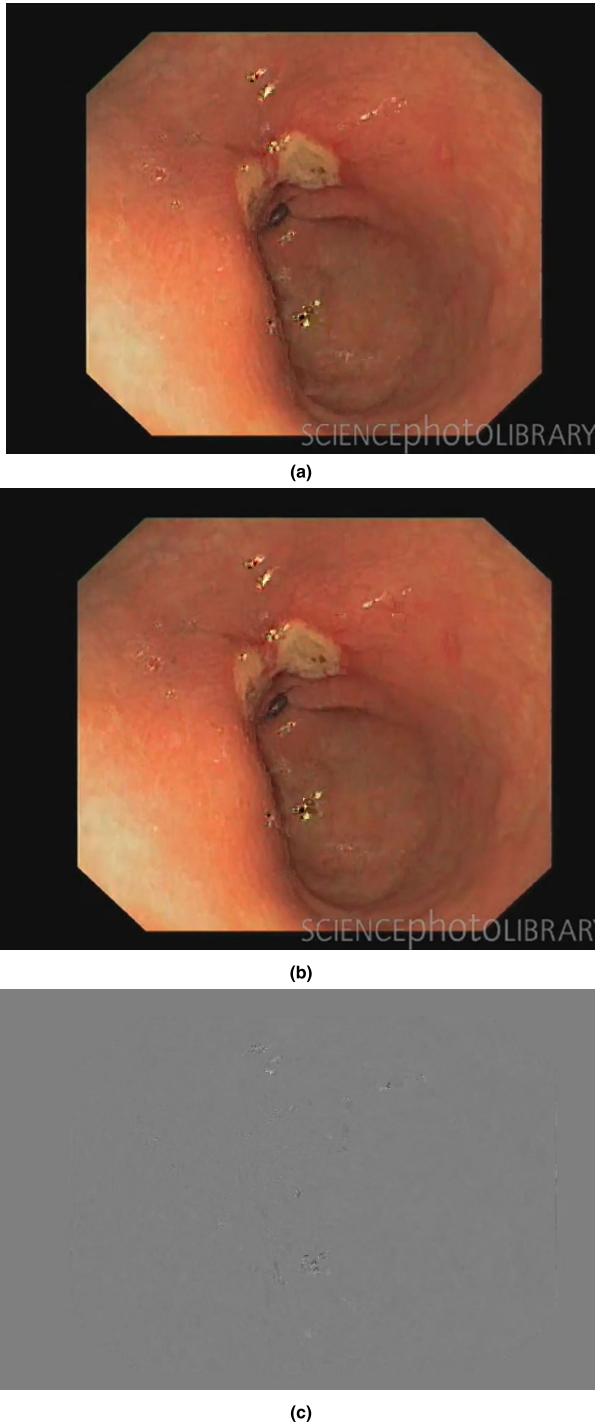
(a)

(b)

(c)

FIGURE 13. Subjective comparisons for frames of stomach ulcer endoscopy transmitted through IP network. (a) Image decoded with the anchor algorithm. (b) Pictures decoded with the proposed schemes. (c) Subjective difference.

The ΔMB between the proposed and anchor algorithms is recorded in Table 4. Positive values in ΔMB imply an increase in the number of 16×16, 16×8, and 8×16 MB partition types. Meanwhile, negative values in ΔMB imply decrease in the number of 8×8, 8×4, 4×8, and 4×4 MB subpartition types of the proposed algorithm compared

TABLE 2. Test sequences.

Frame size for EL	Sequence	Frame rate	Number of frames
1280×720	Lung echo MRI	30	100
1280×720	Gastric cancer	30	100
1280×720	Brain neuron	30	100
1280×720	Stomach ulcer endoscopy	30	100

TABLE 3. Main configuration parameters.

Slice mode: 1
Slice group map type: 2
Base layer QP values: 26, 30, 34, 38
Enhancement layer QP values: 28, 32, 36, 40
Use ESS for enhancement layer: 1
Instantaneous data refresh: −1
Base layer mode: 1

TABLE 4. Δ MB between anchor and proposed schemes in decoder with 3% PLR for BL and 20% PLR for EL for different sequences.

Sequence	$QP_{BL}-$ QP_{EL}	ΔMB mode partitions (%)				
		16×16	16×8	8×16	8×8	8×4
MRI	38–40	173.3	−95.6	0.0	0.0	−1.6
MRI	34–36	165.7	−94.0	0.0	0.3	−1.9
MRI	30–32	160.5	−93.3	0.0	0.0	−2.3
MRI	26–28	157.9	−93.5	0.0	0.2	−2.7
Brain neuron	38–40	159.4	−92.7	[a]-∞	0.0	−1.9
Brain neuron	34–36	149.7	−90.0	0.0	0.2	−3.6
Brain neuron	30–32	141.3	−87.3	0.0	0.2	−4.6
Brain neuron	26–28	132.7	−84.9	0.0	0.5	−5.3
Gastric cancer	38–40	144.6	−90.3	0.0	2.2	−1.7
Gastric cancer	34–36	119.6	−84.1	0.0	0.0	−1.6
Gastric cancer	30–32	94.3	−76.3	0.0	1.1	−1.3
Gastric cancer	26–28	77.8	−70.7	0.0	0.2	−1.1
Stomach ulcer endoscopy	38–40	175.5	−96.1	-∞	0.0	−2.5
Stomach ulcer endoscopy	34–36	160.9	−93.3	0.0	0.0	−2.4
Stomach ulcer endoscopy	30–32	141.2	−88.8	0.0	0.7	−2.4
Stomach ulcer endoscopy	26–28	115.2	80.7	0.0	0.8	−2.2
Average		**141.8**				

[a]-∞ represents block partitions that converge to high partitions during encoding for the anchor and the proposed algorithms.

to the anchor algorithms, thereby reflecting the decrease in complexity. The subpartitions for 4×8 and 4×4 are not mentioned in Tables 4 and 5 because during encoding they converged with both the anchor and the proposed algorithms. As shown in Table 4, our proposed scheme increased the number of the 16×16 partition type by 141.8% compared with the anchor consequently, lowering the number of MBs of the subpartition type tested in the various sequences.

Figures 8(a) and (b) show RD curves for the proposed and anchor schemes for the IP network of the lung echo

TABLE 5. Δ MB between anchor and proposed schemes in decoder with 5% PLR for BL and 10% PLR for EL for different sequences.

Sequence	QP_{BL}–QP_{EL}	ΔMB mode partitions (%)				
		16×16	16×8	8×16	8×8	8×4
MRI	38–40	172.0	−60.4	0.0	0.0	−1.6
MRI	34–36	168.2	−94.4	0.0	0.0	−1.9
MRI	30–32	163.2	−93.8	0.0	0.0	−2.3
MRI	26–28	160.4	−93.9	0.0	0.0	−2.6
Gastric cancer	38–40	143.5	−90.2	0.0	2.7	−1.8
Gastric cancer	34–36	119.2	−84.0	0.0	0.0	−1.6
Gastric cancer	30–32	94.5	−76.4	0.0	1.2	−1.4
Gastric cancer	26–28	79.1	−71.3	0.0	0.2	−1.1
Brain neuron	38–40	159.9	−92.9	- ∞	0.0	−2.3
Brain neuron	34–36	150.8	−90.4	0.0	0.3	−3.6
Brain neuron	30–32	142.2	−87.7	0.0	0.0	−4.6
Brain neuron	26–28	133.4	−85.2	0.0	0.7	−5.4
Stomach ulcer endoscopy	38–40	176.8	−96.3	- ∞	0.0	−2.2
Stomach ulcer endoscopy	34–36	162.8	−93.7	0.0	0.0	−2.3
Stomach ulcer endoscopy	30–32	143.5	−89.3	0.0	0.8	−2.5
Stomach ulcer endoscopy	26–28	118.2	−82.8	0.0	0.9	−2.1
Average		**143**				

MRI sequence and gastric cancer sequences, respectively. Figures 9(a) and (b) show the RD curves for the brain neuron sequence and the stomach ulcer endoscopy, respectively. Our approach is comparable in frame quality and with regard to coding efficiency and ΔPSNR.

Table 5 contains the ΔMB between the proposed and the anchor algorithms. According to Table 5, our proposed scheme raised the number of MBs in the 16×16 partition type by 143% compared with the anchor. The number of MBs in subpartition types are also reduced in various video sequences.

D. SUBJECTIVE ACHIEVEMENT

We also evaluated the subjective quality of the decoded videos transmitted through the wireless and IP networks. Selected frames were decoded with the pYUV viewer tool taken from different video sequences, and compared for both the proposed and the anchor algorithms [35]. Figures 10–13 show visual renderings of pictures decoded through the proposed and anchor algorithms. In these Figures, the frames (a) show a picture decoded with the anchor method. The frames in (b) show a picture decoded with the proposed method, and (c) show scenarios for the visual variation between the proposed and anchor methods.

1) As shown in Figs. 10–13, the subjective illustrations are depicted for the following: MRI: 95th frame coded with $QP_{BL} = 26$ and $QP_{EL} = 28$
2) Brain neuron: 37th frame coded with $QP_{BL} = 26$ and $QP_{EL} = 28$

3) Gastric cancer: 12th frame coded with $QP_{BL} = 30$ and $QP_{EL} = 32$
4) Stomach ulcer endoscopy: 14th frame coded with $QP_{BL} = 30$ and $QP_{EL} = 32$

Moreover, as shown in Figs. 10–13, some visual contents of the decoded HD size frames are unclear due to limitations in the manuscript page fitting size, especially for the subjective quality variations between the proposed and anchor methods shown in Figs. 10(c)–13(c). Tables 4 and 5 validate that the proposed scheme considerably reduces the computation for wireless and IP networks without any degradation of either PSNR or RD. The resulting high-quality medical images hold promise for eHealth applications, especially when transmitted in an error-prone environment.

E. STUDY CONSTRAINTS

The proposed technique considerably reduces complexity for homogeneous regions as the search for an optimal MB mode is early terminated. For non-homogeneous regions, it merges constituent 4×4 block having vector differences of less than the threshold. Whereas, due to inherent encoding and transmission constraints, the concealed qualities of images may differ subject to the following:

1) Loss of an instantaneous decoder refresh frame is the most severe case, as the synchronization depends on it.
2) Loss of an intracoded frame or a predicted frame resulting in distortions of the video persistence quality.
3) Loss of base layer MBs resulting in a loss of the corresponding enhancement of later MBs as there is a loss of interlayer prediction.

VII. CONCLUSION

Our research proposes a fast MB mode selection technique to reduce the complexity of high-definition medical images passing through error-prone environments. Flexible macroblock ordering offers a noteworthy approach to reduce the influence of errors inherent to wireless and IP networks. We exploit the content features of video to reduce complexity with flexible macroblock ordering for error resilience. Given that high-definition medical videos are extremely large, multimedia services with low complexity are an encouraging advancement for healthcare systems. The experimental outcomes confirm that the proposed scheme achieved significant improvement with an average of 143%, which reduces complexity without compromising visual quality.

REFERENCES

[1] J. M. Santos, A. F. R. Guarda, L. A. da Silva Cruz, N. M. M. Rodrigues, and S. M. M. Faria, "Compression of medical images using MRP with bi-directional prediction and histogram packing," in *Proc. Picture Coding Symp. (PCS)*, Nuremberg, Germany, Dec. 2016, pp. 1–5.

[2] A. Chowdhury, H.-C. Chien, S. Khire, S.-H. Fan, X. Tang, N. Jayant, and G.-K. Chang, "Next-generation E-health communication infrastructure using converged super-broadband optical and wireless access system," in *Proc. IEEE Int. Symp. World Wireless, Mobile Multimedia Netw. (WoW-MoM)*, Montreal, QC, Canada, Jun. 2010, pp. 1–5.

[3] Y. Cui, W. Ren, and Z. Deng, "Fast mode decision for HD scalable video coding via statistical content analysis," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 8, Apr. 2018, Art. no. 1855011.

[4] *Telecommunication Standardization Sector of ITU, Series H: Audio-Visual and Multimedia Systems, Infrastructure of Audiovisual Services-Coding of Moving Video. Conformance Specification for H.264 Advanced Video Coding*, document ITU-T H.264.1, 2005.

[5] J. M. Boyce and R. D. Gaglianello, "Packet loss effects on MPEG video sent over the public Internet," in *Proc. 6th ACM Int. Conf. Multimedia*, Bristol, U.K., 1998, pp. 181–190.

[6] M. A. G. Santos, R. Munoz, R. Olivares, P. P. R. Filho, J. D. Ser, and V. H. C. D. Albuquerque, "Online heart monitoring systems on the Internet of health things environments: A survey, a reference model and an outlook," *Inf. Fusion*, vol. 53, pp. 222–239, Jan. 2020.

[7] E. Abdulhay, M. Alafeef, L. Alzghoul, M. Al Momani, R. Al Abdi, N. Arunkumar, R. Munoz, and V. H. C. de Albuquerque, "Computer-aided autism diagnosis via second-order difference plot area applied to EEG empirical mode decomposition," *Neural Comput. Appl.*, vol. 2018, pp. 1–10, Sep. 2018, doi: 10.1007/s00521-018-3738-0.

[8] J. M. de Oliveira, R. Munoz, S. Ribeiro, W. Wu, and V. H. C. de Albuquerque, "REHAB FUN: An assistive technology in neurological motor disorders rehabilitation of children with cerebral palsy," *Neural Comput. Appl.*, vol. 2019, pp. 1–14, Mar. 2019.

[9] M. I. Razzak, M. Imran, and G. Xu, "Big data analytics for preventive medicine," *Neural Comput. Appl.*, vol. 2019, pp. 1–35, Mar. 2019, doi: 10.1007/s00521-019-04095-y.

[10] S. Ullah, K.-I. Kim, K. H. Kim, M. Imran, P. Khan, E. Tovar, and F. Ali, "UAV-enabled healthcare architecture: Issues and challenges," *Future Gener. Comput. Syst.*, vol. 97, pp. 425–432, Aug. 2019.

[11] U. Iqbal, T. Y. Wah, M. Habib ur Rehman, G. Mujtaba, M. Imran, and M. Shoaib, "Deep deterministic learning for pattern recognition of different cardiac diseases through the Internet of medical Things," *J. Med. Syst.*, vol. 42, no. 12, pp. 252–265, Nov. 2018.

[12] Z. Ali, M. Imran, M. Alsulaiman, M. Shoaib, and S. Ullah, "Chaos-based robust method of zero-watermarking for medical signals," *Future Gener. Comput. Syst.*, vol. 88, pp. 400–412, Nov. 2018.

[13] L. Zhou, X. Wang, W. Tu, G.-M. Muntean, and B. Geller, "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 409–419, Apr. 2010.

[14] A. Ziviani, B. E. Wolfinger, J. F. de Rezende, O. C. M. B. Duarte, and S. Fdida, "Joint adoption of QoS schemes for MPEG streams," *Multimedia Tools Appl.*, vol. 26, no. 1, pp. 59–80, May 2005.

[15] S. Paluri, K. K. R. Kambhatla, B. A. Bailey, P. C. Cosman, J. D. Matyjas, and S. Kumar, "A low complexity model for predicting slice loss distortion for prioritizing H.264/AVC video," *Multimedia Tools Appl.*, vol. 75, no. 2, pp. 961–985, Nov. 2014.

[16] A. M. Buhari, H.-C. Ling, V. M. Baskaran, and K. Wong, "Low complexity watermarking scheme for scalable video coding," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-TW)*, Nantou, Taiwan, May 2016, pp. 5–6.

[17] M. Koziri, P. Papadopoulos, N. Tziritas, A. N. Dadaliaris, T. Loukopoulos, and S. U. Khan, "Slice-based parallelization in HEVC encoding: Realizing the potential through efficient load balancing," in *Proc. IEEE 18th Int. Workshop Multimedia Signal Process. (MMSP)*, Montreal, QC, Canada, Sep. 2016, pp. 1–6.

[18] I. Ali, S. Moiron, M. Fleury, and M. Ghanbari, "Data partitioning technique for improved video prioritization," *Computers*, vol. 6, no. 3, p. 23, Jul. 2017.

[19] D. Grois and O. Hadar, "Efficient Region-of-Interest scalable video coding with adaptive bit-rate control," *Adv. Multimedia*, vol. 2013, pp. 1–17, 2013.

[20] (2008). *Digital Imaging and Communications in Medicine (DICOM)*. [Online]. Available: http://medical.nema.org/dicom/geninfo/Strategy.pdf

[21] *DICOM Part 5: Data Structures and Encoding*, NEMA, Arlington, VA, USA, 2011, pp. 1–117. [Online]. Available: http://medical.nema.org/Dicom/2011/11_05pu.pdf

[22] *Proposal for New Work Item for WG13 on MPEC-4 AVC/H.264 Video Coding*, NEMA, Arlington, VA, USA, 2014. [Online]. Available: http://medical.nema.org/medical/Dicom/

[23] M. Shoaib, U. Ahmad, and A. Al-Amri, "Multimedia framework to support eHealth applications," *Multimedia Tools Appl.*, vol. 73, no. 3, pp. 2081–2101, Sep. 2013.

[24] H. Li, Z. G. Li, and C. Wen, "Fast mode decision algorithm for inter-frame coding in fully scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 7, pp. 889–895, Jul. 2006.

[25] D. Wu, F. Pan, K. P. Lim, S. Wu, Z. G. Li, X. Lin, S. Rahardja, and C. C. Ko, "Fast intermode decision in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 953–958, Jul. 2005.

[26] X. Wang and J. Ridge, *CE2 report: Improvement of macroblock mode prediction in ESS*, document ISO/IEC JTC1/SC29/WG11/JVT-W030, Los Angeles, CA, USA, 2007. [Online]. Available: https://www.itu.int/wftp3/av-arch/jvt-site/2007_04_SanJose/

[27] J. Reichel, H. Schwarz, and M. Wien, *Joint Scalable Video Model 8—Joint Video Team (JVT) of ISO/IEC MPEG /ITU-T VCEG*, document ISO/IEC JTC1/SC29/WG11/JVT-U202, Hangzhou, China, 2006.

[28] (May 2018). *MRI of Lung Echo Sequence*. [Online]. Available: http://www.springerimages.com/Images/RSS/1-10.1007_s00330-010-1918-0-9

[29] Gastro Laboratory (cancer), Science Photo Library. (Feb. 2019). *Gastric Cancer, Endoscope View*. Access: [Online], Available: https://www.sciencephoto.com/media/410600/view

[30] (Jan. 2019). *Brain Neuron 3D for Medical, Science Photo Library*. [Online]. Available: https://www.sciencephoto.com/media/236058/view

[31] Gastro Laboratory (ulcer), Science Photo Library. (Mar. 2019). *Stomach Ulcer, Endoscope View*. [Online]. Available: https://www.sciencephoto.com/media/234813/view

[32] J. Vieron, M. Wien, and H. Schwarz. (2012). Joint Scalable Video Model-JVT.H.264/SVC—JSVM Reference Software and Manual. Heinrich Hertz Institute, Fraunhofer Institute for Telecommunications. [Online]. Available: https://www.hhi.fraunhofer.de/en/departments/vca/research-groups/image-video-coding/research-topics/svc-extension-of-h264avc/jsvm-reference-software.html

[33] S. Wenger, *Error Patterns for Internet Experiments*, document VCEG-Q15-I-16r1, Video Coding Expert Group, TU Berlin, Germany, 2002.

[34] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1427–1441, Jun. 2010.

[35] (2018). '*PYUV: Raw Video Sequence Player*' *DSP Laboratory* [Online]. Available: http://dsplab.diei.unipg.it/software/pyuv_raw_video_sequence_player

MUHAMMAD SHOAIB received the B.Eng. and M.Eng. degrees from the NED University of Engineering and Technology, Karachi, in 1995 and 2005, respectively, and the Ph.D. degree in communication and information system from the Beijing University of Posts and Telecommunications, China, in 2010. He worked as a Senior Manager (IP Operations, South) at Pakistan Telecommunication Company Limited, Pakistan, from 2001 to 2011. He also worked as a Maintenance Engineer at the R. M. International School, from 1996 to 2001. He is currently an Assistant Professor with the Information Systems Department, College of Computer and Information Sciences, King Saud University. His areas of research include video compression techniques, multilayer video coding, commercial data center facilities, IP packet-based networks, infrastructure and security.

MUHAMMAD IMRAN received the Ph.D. degree in information technology from Universiti Teknologi PETRONAS, Malaysia, in 2011.

He is currently an Associate Professor with the College of Applied Computer Science, King Saud University, Saudi Arabia. His research is financially supported by several grants. He has completed a number of international collaborative research projects with reputable universities. He has published more than 150 research articles in peer-reviewed, well-recognized international conferences, and journals. Many of his research articles are among the highly cited and most downloaded. His research interests include the Internet of Things, mobile and wireless networks, big data analytics, cloud computing, and information security. He has been consecutively awarded with an Outstanding Associate Editor of IEEE Access, in 2018 and 2019 besides many others. He has served as an Editor-in-Chief for the European Alliance for Innovation (EAI) *Transactions on Pervasive Health and Technology*. He has served/serving as a Guest Editor for about two dozen special issues in journals, such as the *IEEE Communications Magazine*, the *IEEE Wireless Communications Magazine*, *Future Generation Computer Systems*, IEEE Access, and *Computer Networks*. He is also serving as an Associate Editor for top ranked international journals, such as *IEEE Communications Magazine*, the IEEE Network, *Future Generation Computer Systems*, and IEEE Access. He has been involved in about 100 peer-reviewed international conferences and workshops in various capacities, such as the chair, the co-chair, and a technical program committee member.

FAZLI SUBHAN received the M.Sc. degree in computer science from the University of Peshawar, Pakistan, in 2003, and the Ph.D. degree in information technology from Universiti Teknologi PETRONAS, Malaysia, in 2012. He is currently an Assistant Professor with the Department of Computer Science, National University of Modern Languages (NUML), Islamabad, Pakistan. His research interests include Indoor positioning systems, signal processing, and handover in wireless networks.

IFTIKHAR AHMAD received the B.Sc. degree from the Islamia University of Bahawalpur, Bahawalpur, Pakistan, in 1999, the M.Sc. degree in computer science from the University of Agriculture, Faisalabad, Pakistan, in 2001, the M.S./M.Phil. degrees in computer science from the COMSATS Institute of Information Technology, Abbottabad, Pakistan, in 2007, and the Ph.D. degree in information technology from Universiti Teknologi PETRONAS, Malaysia, in 2011. He is currently a Faculty Member with the Information Technology Department, Faculty of Computer and Information Technology, King Abdulaziz University. He has been serving as a faculty member and a research supervisor at various other universities, since 2001. He has also been involved in several funded projects as PI and Co-PI. He has published several articles in reputed journals and conferences. He is also a member of several scientific and professional bodies.

• • •