Distributive video coding for endoscopy applications

G. Ajitha[1] I. SanthiPrabha[2]

[1]Research Scholar, Department of Electronics and Communications, University College of Engineering, JNTUK, Kakinada, India

[2]Professor, Department of Electronics and Communications, University College of Engineering, JNTUK, Kakinada, India

Abstract— Conventional videocodes similar as MPEGx andH.26x, With the framework ensures an extensive architecture of block- based motion correction and discrete cosine transform (DCT) employing an advanced encoder and a straightforward decoder. Source correlation is used at the decoder in Wyner-Ziv (WZ) video coding, a realistic implementation of distributed video coding (DVC) relying on the Slepian-Wolf and WZ theorems. permitting the creation of encoders that are more straightforward than those of traditional video codecs. An adaptive two- step side-information generating strategy is given for evaluation and analysis in this study to enhance transform domain WZ video coding performance. The rate distortion (RD) performance is gradually improved as decoding progresses by utilising previously decoded coefficients and decoded WZ frames in the down sampled spatial resolution. The research outcomes demonstrate that the proposed strategy, when compared to earlier H.26x video coding techniques, exhibits superior coding execution, especially in straightforward situations, like ones including a lot dynamic movement.

KEYWORDS: Video coding, DataProcessing, H.264, Wyner-ziv, Matlab, Entropy, Bit Error Ratio(BER).

I. INTRODUCTION

As medical video transmission becomes more commonplace,it is playing an increasingly important role in medicine. Better compression techniques are needed to accommodate this, especially for transmission across bandwidth-constrained channels where the video's diagnostic capabilities must be maintained. In the domain of video compression methods, encoder analytical coding to identify similarities between successive frames.The Wyner-Ziv(WZ) Theorem is about how to compress videos better. It suggests that a specific typeof video codec could work well by encoding each frame separately, but only decoding them with the help of nearby frames. This would make the codec work more efficiently to examine the body, endoscopes which are capable of being directly introduced into the organ, distinguish them fromvarious other medical imaging procedures and enable examination of bodily cavities or hollow organs, endoscopy serves as a diagnostic tool for identifying the underlying causeof various problems like heartburn, feeling sick, throwing up,abdominal pain, struggling to swallow and gastrointestinal bleeding.

II. LITERATURE REVIEW

Gibson and colleagues, A lossy wavelet-based approach was suggested for the compression of digital angiogram videos [4]The analysis of high-frequency subbands obtained via wavelet decomposition of an angiogram video reveals regions within the footage that are of limited diagnostic utility and exhibit considerable spatial extent. An approach for texture modelling is adopted in order to encode the high-frequency subband wavelet coefficients within specific regions. This methodology is solely employed in regions that are deemed to be irrelevant for testing purposes, whereas significant testing regions are presented conventionally. The contraction scheme has been devised with the objective of automatically identifying and utilizing regions of interest, with a view to attaining a favorable equilibrium between video quality and compression rate. Several endeavors have been undertaken to establish a universal video compression standard for medical purposes with the aim of proficiently and competently disseminating compressed medical data to the general populace. The digital imaging and communications in medicine (DICOM) standard, which serves to enable the sharing and visualization of medical images, is widely regarded as the predominant standard within the domain [5]. Despite having originated in the 1970s, the Slepian-Wolf theorem has found practical application in the form of Slepian-Wolf coding, which exhibits similarities to channel coding [9]. In order to appreciate the relationship under examination, one must consider binary sequences X and Y within a given context. Upon observing similarities between these sequences, a hypothetical "error sequence" can be posited. This hypothesized sequence is comprised chiefly of zeros, except for select instances where ones appear to indicate the positions where X and Y diverge. One potential strategy for safeguarding against criminal activity aimed at X involves the implementation of an effective channel coding approach that selectively communicates exclusively the relevant parity bits associated with system performance. Once the decoder phase has been reached, it is customary to concatenate the parity bits with the side information and then proceed to engage in error-correcting decoding.If X and I exhibit a high degree of similarity, the required number of parity bits for transmission would be minimal, resulting in notable compression outcomes. We emphasize that this approach does not perform forward error correction (FEC) to protect against errors introduced by the transmission channel, but instead by a virtual "correlation channel" that captures the statistical dependence of X and the side information Y.

In an alternative interpretation, the alphabet of X is divided into cosets and the encoder sends the index of the coset that X belongs to [2]. The receiver decodes by choosing the codeword in that coset that is most probable in light of the side 8 information V. It is easy to see that both interpretat ions are equivalent. With the parity interpretation, we send a binary row vector X = XP, where G = [I][P] is the generator matrix of a systematic linear block code With the coset Interpretation, we send the "syndrome" S = XH, where II is the parity check matrix of a linear block code However, the transmitted bit aqueducts are identical, if P = H. The ideal of the system isn't to perform opinion in lieu of a croaker, but to help in the visualization and identification of abnormal structures. frequently nearly match the girding towel in colour. In addition, endoscope videotape is frequently poor in quality and characterized by low discrepancy and the presence of noise. Lossy contraction is attained which makes the videotape poor in quality it's to enhance the videotape and should be easily visible to the stoner. Individual frames should be compressed in the videotape. The decoder creates the side information for each WZ enciphered frame, by performing a stir compensated frame interpolation (or extrapolation) using the closest formerly decrypted frames. The side information for each WZ frame is taken as an estimate noisy interpretation of the original WZ frame. The better it's the estimate, the lower is the number of ‘ crimes ’ the turbo decoder has to correct and the bitrate demanded. Compression should be more in anticipated affair and the videotape should be visible duly.

III. EXISTING FRAMEWORK

In 1996, the Video Coding Experts Group (VCEG) introduced the H.263 codec as a refinement within the H26x coding family, H.263 and H.264 were designed with the primary purpose of facilitating low bitrate videoconferencing applications. As for their computational requirements, it can be observed that the encoding process demands considerably more power than the decoding process; to be specific, the encoder calls for between five to 10x more complexity than of decoder. This particular non-symmetry in complication is highly suitable for employment in broadcasting as well as streaming video-on-demand networks wherein the compression of video frames takes place only once while decryption occurs multiple times. It is noteworthy that H.263 and H.264 codecs have been optimized for video transmission in bandwidth-constrained environments, with the ultimate goal of enhancing the overall user experience. The codec denoted as 263 was engineered to cater to the demands of videotape systems with constrained bandwidth, whereas H.264 is a more progressive alternative. The encoding scheme labeled as 264 possesses the capability to encode videos with a wide range of quality levels, spanning from low to high quality. Both codecs are suitable for the purpose of streaming. This technology is also applied in the context of IP Multimedia Subsystem (IMS), Multimedia messaging Service(MMS),The present era has witnessed a substantial rise in the usage of video data owing to the proliferation of internet-based applications such as social networking, e-governance, security and surveillance, and video telephony. Consequently, the network bandwidth has emerged as a significant bottleneck for efficient transmission. To overcome this challenge, the use of Discrete Cosine Transform (DCT) and Motion Estimation and Compensation techniques has been proposed to eliminate temporal redundancies. Notably, the H.263 Codec has been found to be a highly effective method for video compression. However, it is pertinent to note that webcams capture videos in YUV format, which is subsequently converted to RGB format for rendering purposes. Additionally, in application, it is converted to RGB(BITMAP) format - the second-level format.

Side information estimation is a crucial aspect of distributed video coding (DVC) systems. DVC schemes perform motion estimation at the decoder side to generate reliable side information (Tagliasacchi & Tubaro, 2007). The quality of the side information is essential for achieving efficient video compression (Maitre et al., 2007). Several techniques have been proposed to improve the accuracy and reliability of side information estimation.

One approach is to use differential motion estimation algorithms. These algorithms produce a dense motion vector field that can be useful in the framework of DVC (Cagnazzo et al., 2009). However, the dense motion vector field requires significant coding resources and computational effort in traditional video compression. In DVC, these algorithms can be employed at the decoder without increasing the coding rate (Cagnazzo et al., 2009).

Certain software applications may necessitate the utilization of a dual system, which entails the use of low-complexity encoders. This may come at the potential cost of utilizing high-complexity decoders. Compression of images may lead to infrequent inconsistency in image quality, or even result in low-quality image formats. Additionally, the implementation of this compression technology requires a greater amount of hardware overhead in comparison to other video compression codecs. The efficiency of video compression is reliant on the processing capacity of the hardware, as the two factors exhibit a direct proportion. In terms of computing complexity, the compression of data in the H.263 format necessitates a high amount of CPU time. To compress multiple live videos, the utilization of a high-end computer is essential. However, some noticeable artifacts may arise during compression, specifically with regard to the compressed images' visual features which may exhibit intermittent growth, disappearance, and regrowth, and may also display a lack of consistency. The increase and cessation of movement poses an obstacle to both motion detection and object tracking. Moreover, the encoding process demands considerable processing power, thereby necessitating the use of devices with high-powered processors.

IV. PROPOSED FRAMEWORK

The framework of an intraframe encoder interframe decoder, as proposed by the authors, is hereby referred to as the Wyner- Ziv video codec. The present study concerns itself with a distinct instance of distributed video coding (DVC), namely the Wyner-Ziv video coding technique. This method is designed to address the challenge of lossy source coding in conjunction with side data at the decoder, and furthermore, it accommodates the transfer of motion estimation tasks to the decoder, either for specific segments or the entire movement. It is the responsibility of the decoder to retrieve the supplementary data, which comprises a representation of the encoded Wyner-Ziv scheme, while the encoder transmits only parity bits to enhance its accuracy. The following proposal introduces a novel technique aimed at enhancing the quality of the side data utilized by the Wyner-Ziv codec, thereby augmenting its rate distortion capabilities. The utilization of a Wyner-Ziv video encoder facilitates significant cost savings as each video frame is compressed independently, reducing the need for intraframe processing. The decoder for comparison, situated in the established section of the system, utilizes the factual dependency between frames through intricate interframe processing. In the present application, the conversion to RGB (BITMAP) - Second level arrangement is achieved through the utilization of the following formulae: r

= y + (1.370705 * (v128)); g = y - (0.698001 * (v-128)) - (0.337633 * (u128)); b = y + (1.732446 * (u-128)).

The variables y, u (Cb), and v (Cr) represent Luminance, Chroma channel blue component, and Chroma channel ruddy component, respectively. In video encoding, the input source is segmented into multiple Group of Pictures (GOPs) consisting of a key frame and subsequent frames. The decoding process entails utilizing the H.263/AVC intra decoder for extracting hidden information from all frames except the first one within a GOP. At this juncture, the concealed coding mode information pertaining to the ultimate Wyner-Ziv design within a Group of Pictures (GOP) shall be deciphered. The Wyner-Ziv scheme involves the initial decoding of the coding mode data, followed by the reproduction of all the skip mode squares, specifically those that are 13 in number, through the allocation of corresponding elements from the previously reconstructed outline. In contrast, the coding mode pertaining to each non-skip segment employs the proposed movement remuneration plot to identify the corresponding lateral information. In the process of expansion, a discernible pattern in non-skipping mode involves the formation of several observation windows that are derived from previously replicated designs. These windows are carefully designed to encompass individual components within their respective boundaries, effectively identifying them as probable candidates for inclusion within the larger square.

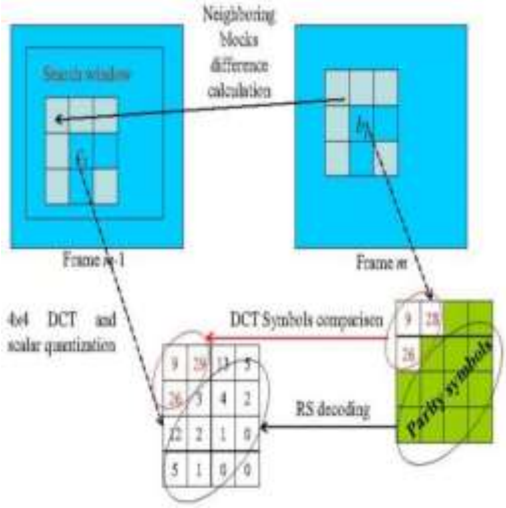

Figure 1 An example of motion compensation

Each potential block will undergo the DCT and scalar quantization processes at the encoder. The upgraded 8-foot for each candidate block, connected neighbouring block additionally extracted. In our approach, a block is bi, with each candidate block, nonskip using RS coding mode, ci, the search windows' contents will be assessed. The top candidate meeting the three outlined requirements .The following will be chosen as the bi-specific side information: (a) The distinction between the three most crucial DCTs reduce the size of the symbols for bi and ci; (b) depends on the amount of incorrectly interpreted RS symbols the DCT symbols and the bi-parity symbols (apart from the ci's three most crucial symbols) should be kept to a minimum; (c) The variation between the eight connected nearby blocks of bi and ci ought to be kept to a minimum. The RSdecoding symbol error rate in this case typically varies from 100 to boo. Nevertheless, the crucial three DCT .The emblems remain error messages in the other typically acceptable subbands, when the light-colored blocks have already been rebuilt. the blocks in a Wyner-Ziv frame (not including skip mode frames), they are rebuilt in the sequence of raster scans.Two relevant criteria are utilised for a block, bi, with non-skip without RS coding mode: (a) The gap among all the DCT symbols for bi and ci should be kept to a minimum; (b) The variation between the eight connected nearby blocks of bi and ci ought to be kept to a minimum. After the side information for a block, which is the best candidate block. When non-skip mode is achieved, it will be beneficial for rebuilding the block.

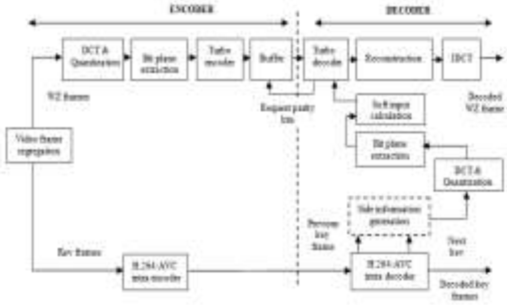

Figure 2 Block Diagram of Wyner Ziv Video coding

Quantization: Considering a block with non-skip the RSdecoded DCT symbols from the RScoding mode.The block will receive copies of any side information. Each WZ frame has been subjected to a block-based transform, typically a DCT. The DCT bands are created by grouping the DCT coefficients of the complete WZ frame 16 in accordance with the positions that each DCT coefficient occupies within a block. Each DCT band is uniformly quantized with a set of levels based on the desired level of quality [4]. The quantized symbols' bits are combined to form bitplanes for a certain band, which are then individually turbo encoded.

Turbo Encoding: The most significant bitplane (MSB) is utilised to begin the turbo encoding of each DCT band. The following the generation of parity information for each bitplane, chunks/packets are preserved in the buffer and dispatched when requests for decoders made via the feedback channel.

Side Information Production: The side is produced by the decoder information for each frame with WZ coding by running a frame interpolation with motion compensation (or use the closest previously decoded frames for extrapolation. Each WZ frame's side information is used as an approximation of the original WZ frame (in a noisy variant). The better the estimate, the fewer individuals it includes, the turbo decoder must fix "errors," and the bitrate needed.

Reconstruction: Following the execution of turbo decoding on each of the bitplanes corresponding to individual Discrete Cosine Transform (DCT) frequency bands, the resulting bitplanes are aggregated to generate the decoded quantized image stream associated with each respective DCT band. Upon obtaining all decoded quantized images, it becomes possible to reproduce the lattice of Discrete Cosine Transform coefficients. The discrete cosine transform (DCT) coefficients comprising 17 groups, during which no Walsh–Hadamard transform (WZ) bits were transmitted, have been substituted by the corresponding DCT groups of the side data. This has been executed to enhance the overall efficiency of the transmission system. The outline is a fundamental element of the writing process that serves as a framework for organizing ideas and providing a clear and logical structure for written works. It provides a roadmap for writers to follow, allowing them to convey their message effectively and efficiently. The outline serves as an essential tool for research papers, essays, and other academic writing assignments. It enables writers to identify the main points that require support, determine the most appropriate evidence and sources, and arrange ideas in a consistently coherent and logical manner. In academic writing, outlining is a crucial step in the prewriting phase, ensuring that the final product communicates a clear and well- organized message to readers. In order to facilitate the decoding process of the video format, the decoded key frames and wavelet transform coefficients (WZ) are methodically combined.

V. IMPLEMENTATION AND RESULTS

Entropy is the arbitrariness of the pixels. Hence tall entropy implies that the information are spread out as much as conceivable whereas moo entropy implies that the information are about all concentrated on one esteem. In the event that the entropy is low,in this manner.

Bit Error rate (BER) alludes to the sum of advanced information encoded per unit of time, as a rule communicated in seconds to minutes. When planning to shoot video amid generation, taking bit rate into thought is imperative to decide record estimate and picture quality. The BER is the number of bit bits per unit time. The bit error proportion (moreover BER) is the number of bit errors partitioned by the full number of exchanged bits amid a studied time interim. Bit blunder proportion may be a unitless execution degree, frequently communicated as a rate. A result of 10-9 is generally considered a satisfactory bit mistake rate for broadcast communications, whereas 10- 13 may be a more suitable least BER for information trans- mission.

The compression ratio refers to how much a signal is reduced in size, without it being a specific size. Streaming audio and video involve compressing the files to make them easier to transmit over the internet. The compression ratio measures how much the file size has been reduced. Data can be described as either uncompressed or compressed. Rates mean how fast data is being transferred, while data sizes refer to how much data there is. Instead of talking about how much space is saved, we can focus on the speed of transferring data. This text is talking about something called "data-rate savings. This means being able to use less data when sending information. This means making the amount of data smaller compared to how it would be without any compression. In different ways of making videos, some parts are rated to decide how important they are. Compression algorithms use a technique called discrete cosine transform (DCT) deducts data. The IDCT is a mathematical formula that helps to con- vert compressed digital video or images back into their original form. Quantization is the process of approximating continuous values into a set of discrete values for easier storage and processing. IQ is when we turn digital numbers back into pictures and ME is when we guess where objects move in videos. Motion compensation is a method used in video com- pression to remove redundancy and reduce file size. Entropy coding is a technique that further compresses video data by assigning shorter codes to frequently occurring sequences of information. Almost all video formats use a similar process to compress the data. Make the video clearer, you should have a high ratio. Use less storage space.

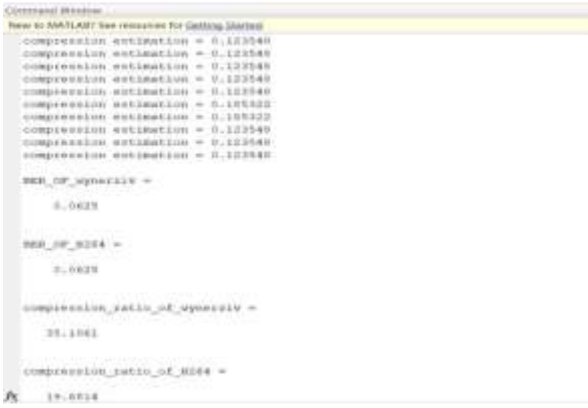

Figure 3 Practically terms of BER and CR is observed

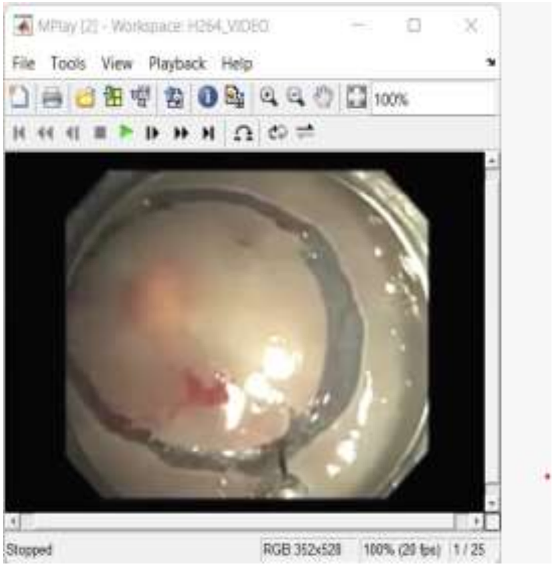

Figure 4 output video H.264

Figure 5 Output video of WZ

	Codec	Compression Ratio	BER	Entropy
Video 1	Wyner Ziv	37.4457	0.0625	5.5473
	H.264	18.6066	0.0625	5.8112
Video 2	Wyner Ziv	35.1783	0.0625	5.6443
	H.264	19.6560	0.0625	5.9819

Table 1 observed outcomes

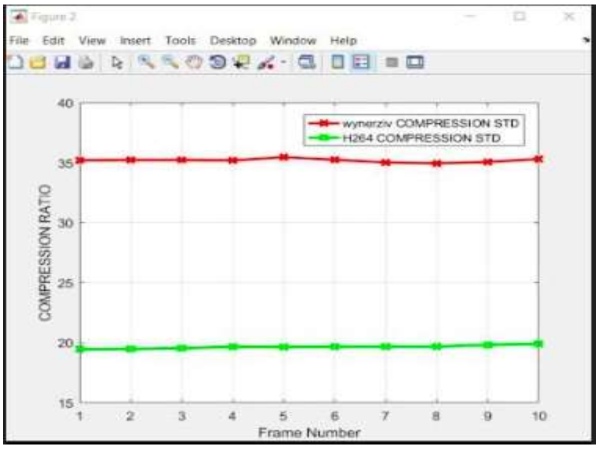

Figure 6 Compration of h.264 and wz – compression

VI CONCLUSION

Video processing helps solve problems related to storing, getting, and working with pictures, videos, and complex signals. It helps extract information that people or computers can use. Making sure the images and videos are stored well and look good is very important.H.264 may be a lossy compression procedure though wyner ziv isn't a lossy compression. In wyner-ziv video coding, decoder side information takes a key portion in a wz video coder among other building pieces. Side data are the subbands which performs numerical operations on pixels to discover the unknown pixel. There will be no misfortune of data. The most advantage of wyner ziv is that it compresses each and each outline separately and the cruel of it'll be the compression proportion where as in H.264 entire video is compressed at a time. From the comparison chart it is clearly watched that wyner ziv compression proportion is much more than the H.264.The entropy is more in H.264 this provides arbitrariness which isn't reasonable for giving great accuracy. This offer assistance us to induce understanding within the current advancement of the subject and get it how unused and way better decoder SI approaches can be well-defined

VII.FUTURE SCOPE

The potential for development and progression for new ideas and progress in "Making videos for medical examina- tions easier with Wyner Ziv technology."Big in the coming years. This refers to what might happen or be used in the future. Technology can make things more precise by being more accurate. Looking at the H.264 architecture without feedback. We want to make a better system for wearable cameras that use less power. We also need to test it. Attach the test results to the document." Make the patient's diagno- sis available online so they can see it." You can use it any- where and at any time. However, there are some limitations due to its...The capsule has the potential to be very useful in diagnosing medical conditions. Doctors may be able to use it to understand a patient's health better. Endoscopy will have a good future.

REFERENCES

[1] Girod, B., Aaron, A. M., Rane, S., & Rebollo-Monedero,D. (2005). Distributed video coding. Proceedings of the IEEE, 93(1), 71-83.

[2] A. Aaron, E. Setton, and B. Girod, "Towards Practical Wyner-Ziv Coding of Video", IEEE International Conf on Image Processing, Barcelona, Spain, September 2003.

[3] Tsai, Min-Jen, John D. Villasenor, and B.

[4] K. T. Ho. Coronary angiogram video compression. Proceedings of 1994 IEEE Nuclear Science Symposium-NSS'94. Vol. 4. IEEE, 1994.

[5] Gibson, David, et al. Angiogram video compression using a wavelet-based texture modelling approach. Proc. 2001British Conf.Medica Image Understanding

[6] and Analysis. 2001.

[7] Hussein, Rada, et al. "DICOM structured reporting:

[8] Part1.Overview and characteristics. Radiographics 24.3 (2004): 891-896.

[9] R. Zamir, "The rate loss in the Wyner-Ziv problem," in IEEE Transactions on Information Theory, vol. 42, no. 6, pp. 2073-2084, Nov. 1996, doi: 10.1109/18.556597.

[10] Lau, C., Cabral Jr, J. E., Rambhia, A. H., & Kim, Y. (2000, April). MPEG-4 coding of ultrasound sequences. In Medical Imaging 2000: Image Display and Visualization (Vol. 3976, pp. 573-579).

[11] Barbier, P., Alimento, M., Berna, G., Cavoretto, D., Celeste, F., Muratori, M., & Guazzi, M. D. (2004, September). Clinical validation of different echocardiographic motion pictures expert group-4 algorythms and compression levels for telemedicine. In Medinfo (pp. 1339-1342).

[12] Wyner, A. (1974). Recent results in the Shannon theory. IEEE Transactions on information Theory, 20(1)

[13] Luthra, A., & Topiwala, P. N. (2003, November). Overview of the H. 264/AVC video coding standard. In Applications of Digital Image Processing XXVI (Vol. 5203, pp. 417-431).

[14] Slepian, D., & Wolf, J. (1973). Noiseless coding of correlated information sources. IEEE Transactions *on information Theory, 19(4), 471-48*