Development Of Low Complexity Encoder And Summarization Techniques For Wireless Capsule Endoscopy Video

Thesis

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF TECHNOLOGY
in
Signal Processing And Machine Learning

by

Student Name
Reg No.212000
Roll No. 212SP000

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY,KARNATAKA
SURATHKAL, MANGALORE - 575025

May, 2022

.

Development Of Low Complexity Encoder And Summarization Techniques For Wireless Capsule Endoscopy Video

Thesis

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF TECHNOLOGY
in
Signal Processing And Machine Learning

by

Student Name
Reg No.212000
Roll No. 212SP00

Under the guidance of

Faculty NAme
Assistant Professor

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY,KARNATAKA
SURATHKAL, MANGALORE - 575025

May, 2022

.

DECLARATION

by the P.G (M.Tech) Student

I hereby declare that the report of the P.G. Project Work entitled **Development Of Low Complexity Encoder And Summarization Techniques For Wireless Capsule Endoscopy Video** which is being submitted to the National Institute of Technology Karnataka Surathkal, in partial fulfilment of the requirements for the award of the Degree of **Master of Technology in Signal Processing and Machine Learning** in the department of Electronics and Communication Engineering, is a bonafide report of the work carried out by me. The material contained in this report has not been submitted to any University or Institution for the award of any degree.

212561, Student Name, ...
(Registration Number, Name & Signature of the Student)
Department of Electronics and Communication Engineering

Place: NITK, Surathkal
Date: June 7, 2022

CERTIFICATE

This is to certify that the P.G Project Work Report entitled **Development Of Low Complexity Encoder And Summarization Techniques For Wireless Capsule Endoscopy Video** submitted by **STUDENT NAME, (Register No. 202561)** as the record of the work carried out by him, is accepted as the P.G Project Work Report submission in partial fulfilment of the requirements for the award of degree of **Master of Technology in Signal Processing and Machine Learning** in the Department of Electronics and Communication Engineering, National Institute of Technology Karnataka Surathkal, during the academic year of 2021-2023.

Guide Name
Project Guide
Assistant Professor
Dept. of E & C Engineering
NITK Surathkal - 575025

Dr. Ashvini Chaturvedi
Head of the Department
Professor
Dept. of E & C Engineering
NITK Surathkal - 575025

Chairman - DPGC
(Signature with Date and Seal)

ACKNOWLEDGEMENT

Acknowledgement.......................

Place: NITK Surathkal
Date: June 7, 2022. **Student Name**

Abstract

Wireless capsule endoscopy (WCE) is the state-of-the-art medical procedure for scanning the entire digestive tract to diagnose gastrointestinal (GI) diseases. Its non-invasiveness and ease of usage make it a better option than conventional endoscopy. However, it is inferior to conventional endoscopy due to low image quality imposed by capsule's complexity and power consumption. In one complete scan of GI tract, a capsule captures between 90000 and 180000 frames during its peristalsis movement. Diagnosing such a large number of images is a time-consuming and tedious procedure that needs a gastroenterologist's undivided attention. The main aim of the research work is two folds. One involves the development of a low complexity video encoder that can reduce the computations in the capsule. The other part involves a WCE video summarization framework to provide an efficient diagnosis.

Developing a low-complexity video encoding architecture that can achieve high compression performance at a low bit rate while maintaining acceptable reconstruction quality is a challenging task in WCE. A distributed video coding (DVC) architecture is proposed to achieve this, which transfers encoder complexity to the decoder side. It employs a keyframe encoder that takes advantage of GI image textural properties to reduce the complexity. Furthermore, the low-frequency bands of the Wyner-Ziv (WZ) frames are used as auxiliary information at the decoder to generate high-quality side information that enables the encoding of high frequency bands with a low bit rate. The proposed DVC framework is further enhanced to reduce complexity by eliminating WZ-chroma component encoding. Exploiting the similarity in colour and texture properties between consecutive frames in WCE video, a deep convolutional neural network model is integrated into the decoder side to predict the chroma component. The proposed methods achieve improvements in coding gain with low complexity encoder when compared with benchmark compression schemes.

Keywords: Wireless capsule endoscopy, Video compression, Distributed video coding, Convolutional neural networks, Chroma prediction, Video summarization.

Contents

List of Figures

List of Tables

ABBREVIATIONS

CNN	-	Convolutional Neural Network
DI	-	Diagnostic Yield
HEVC	-	High Efficiency Video Coding
MSE	-	Mean Square Error

Chapter 1

Introduction

Endoscopy has become a standard and the most preferred method by physicians for detecting gastrointestinal (GI) disorders such as gastric cancer, polyps, intestinal bleeding, Crohn's disease and Celiac disease. It enables the direct visualization of the human GI tract and can even detect early cancers. Wired endoscopy is a commonly used procedure to diagnose the upper GI tract. However, the patients hesitate to undergo this procedure because of the pain and discomfort induced by inserting a long, flexible wire into the digestive tract. Moreover, the wired endoscopy cannot scan the small intestine due to its intricate and curvy nature.

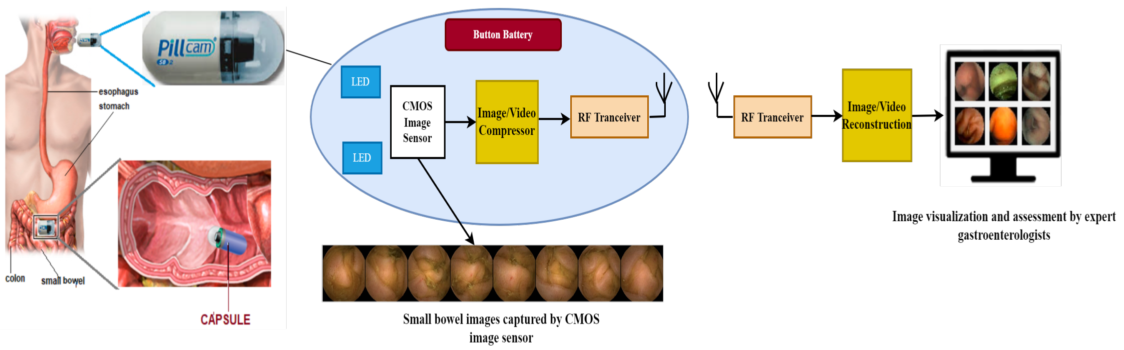

Figure 1.1: A typical WCE based GI tract screening process

Wireless capsule endoscopy (WCE) is used in recent days in order to overcome the drawbacks of wired endoscopy. In WCE, the patient swallows an electronic capsule that moves through the GI tract by peristalsis action. The capsule scans the entire GI tract including small intestine capturing its images without any discomfort and pain (Iddan, Meron, Glukhovsky, and Swain (2000), Wang, Banerjee, Barth, Bhat, Chauhan, Gottlieb, Konda, Maple, Murad, Pfau *et al.* (2013)). The images captured during the peristalsis movement of the capsule are transferred wirelessly through radio frequency (RF) transmission to an external recording unit. Later, the image data is transmitted to a computer and reconstructed before being analyzed for GI abnormalities. A typical WCE based GI tract screening process is shown in Figure 1.1. Prior to the introduction of WCE, it was not possible to diagnose the small intestine without surgery.

Many commercial swallowable capsule endoscopes have been developed (Stickney,

Table 1.1: FDA approved clinically used capsule endoscopes.

Capsule	Inventor	Weight (g)	Dimension (mm)	Frame rate	Resolution	Angle of View	Capsule Life
Pill-Cam SB3	Given Imaging	3.4	11.0 x 26.0	2-6 fps	256 x 256	160°	10-12h
Endo-Capsule	Olympus	3.8	11.0 x 26.0	2 fps	256 x 256	160°	8-10h
Miro-Cam	Intromedic	3.4	11.0 x 24.0	2 fps	320 x 320	170°	8-11h
OMOM	Jinshan	6.0	28.0 x 13.0	2 fps	320 x 320	140°	7-9h
Navi-Cam	Ankon Tech.	5.0	28.0 x 12.0	2 fps	480 x 480	140°	8-9h

Nie, and Zeng (2005)) to assist in the non-invasive scanning of the complete GI tract. The most common clinically used capsule endoscopes approved by the Food and Drug Administration (FDA) available are listed in Table 1.1, with their specifications.

1.1 Motivation

....

1.2 Objectives of the Research

1.3 Contributions

1.4 Outline

The remainder of the thesis is structured as follows: ....

Chapter 2

Background

The structure and peristalsis behaviour of GI tract and analysis of the GI tract image characteristics is necessary to develop an efficient WCE imaging system. This chapter presents the capsule's speed and time it spends in each organ, the colour and texture characteristics of the images captured in various organs of the GI tract. Further, this chapter provides a comprehensive review of the existing WCE image and video compression, as well as summarization techniques.

2.1 Structure and Peristalsis Behavior of GI tract

....

2.2 Literature Review on Compression in WCE

Lossy image compression methods use the discrete cosine transform (DCT) and discrete wavelet transform (DWT) to decorrelate images based on the correlation between pixels at the block level. These methods achieve the compression by exploiting spatial redundancy. Various transform based lossy compression methods used in WCE image compression is listed are Table 2.1.

Table 2.1: WCE lossy compression techniques

Study	Colour Space	Method
DCT (Lin and Dung (2011))	YCoCg	• Before colour space transformation G and B components are subsampled at 2:1 and 4:1. • The transformed coefficients are quantized and entropy coded by Limpel-Ziv algorithm.
DCT (Turcza and Duplaga (2011))	YCoCg	• Transformed coefficients are DPCM coded along with variable Huffman length coding.
DCT (Gu *et al.* (2012))	RGB	• G and B components are subsampled at 2:1. Transformed and quantized coefficients are Huffman coded.
DWT (Khan *et al.* (2015))	YCbCr	• Haar wavelet transform is used. • Run length Huffman encoding is used for entropy coding. • Achieves high compression efficiency compared to DCT based techniques. • High computational complexity and memory requirement.
Modified H.264-Intra (Boudechiche *et al.* (2017))	RGB	• Only DC intra prediction mode is used.

Chapter 3

DVC Architecture with Deep Chroma Prediction Model

3.1 Introduction

Write a small paragraph on the work done and explained in this chapter. The consecutive frames in the WCE video exhibits homogeneity in colour and texture. Considering this, it is sufficient to transmit only the luma components of the WZ frames while the chroma components can be predicted using keyframes. In WCE, decoding is done using a powerful computing system with no memory and power constraints. This chapter presents a DVC architecture with a CNN based deep neural network to predict chroma components of the WZ frame from keyframe at the decoder. This eliminates the processing of WZ-chroma components thus reducing the computational complexity with improved compression performance. The proposed deep neural network for chroma prediction achieves better prediction performance with less computation time. The deep chroma prediction model consists of a merging block with a spatial attention mechanism to merge the feature maps extracted from reference and target frames. So, the model can exploit the non-local similarities between the two frames. Thus, the chroma prediction model can efficiently transfer color from the keyframe to similar regions in the WZ frame.

3.2 Proposed DVC Architecture with Deep CNN (DVC-DCP) at the Decoder

....

3.2.1 Deep Chroma Prediction Model

....

3.3 Summary

The proposed DVC-DCP method achieved better BD bit-rate savings than MJPEG, DVC-TDWZ and DVC-FBC. The video sequences are encoded with a GOP size=4 and the chroma components of the 75reduces the encoding complexity of the chroma components. The encoding complexity of the proposed DVC-DCP is almost same as MJPEG achieved at much lower bitrate. RD performance of the DVC-DCP model is better than DVC-FBC model by 25complexity. The DVC-DCP outperforms DVC-TDWZ in terms of RD and encoding complexity performance. The RD performance of the DVC-DCP is close to H.264 Intra achieved at much lower complexity. ......

Chapter 4

Simulation Results and Discussions

4.1 Simulation Environment and Datasets

4.2 Evaluation Parameters

4.2.1 Video Compression

The evaluation of the compression system is done by comparing compression ratio (CR) calculated using (4.1) with the quality metrics PSNR and structural similarity index (SSIM). PSNR in dB and SSIM determine the amount of quality lost in encoding process and computed using (4.2) and (4.4) respectively.

$$CR = \left(1 - \frac{\text{Total bits after compression}}{\text{Total bits before compression}}\right)\text{x100} \tag{4.1}$$

$$PSNR = 10\log_{10}\frac{255^2}{MSE} \tag{4.2}$$

where

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|x(i,j) - x_r(i,j)|^2 \tag{4.3}$$

where x and x_r are original and reconstructed pixel intensities of an MXN image.

$$SSIM(x, x_r) = \frac{(2\mu_x\mu_{x_r} + C_1) + (2\sigma_{xx_r} + C_2)}{\left(\mu_x^2 + \mu_{x_r}^2 + C_1\right)\left(\sigma_x^2 + \sigma_{x_r}^2 + C_2\right)}, \tag{4.4}$$

where $C_1 = 6.5, C_2 = 58.52$ are constants, μ_x, μ_{x_r} are mean intensities and σ_x, σ_{x_r} are standard deviations of x and x_r respectively. σ_{xx_r} is computed using the following equation.

$$\sigma_{xx_r} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(x_{r_i} - \mu_{x_r}) \tag{4.5}$$

Other important parameters considered to evaluate a video coding system's performance over reference video coding systems are Bjontegaard-Delta (BD) metrics (Bjontegaard (2001)). BD-metrics are widely used to compare a video coding system's performance with the reference codec over a range of quality points or bitrates. BD metric is often computed as change in bitrate or a change in quality measured using PSNR and SSIM based on rate-distortion (RD) curves from the tested data points.

The BD-rate represents the average bitrate savings for the same video quality and is calculated between RD curves of the tested video codec A and a reference codec B. The bitrate saving difference between the two RD curves belong to codecs A and B at a given PSNR is computed by (4.6). BD-PSNR and BD-SSIM between two RD curves A and B at a given bitrate is computed by (4.7) and (4.8).

$$BD - Rate = \frac{Rate_A(PSNR) - Rate_B(PSNR)}{Rate_B(PSNR)} \tag{4.6}$$

$$BD - PSNR = \frac{PSNR_A(bitrate) - PSNR_B(bitrate)}{PSNR_B(bitrate)} \tag{4.7}$$

$$BD - SSIM = \frac{SSIM_A(bitrate) - SSIM_B(bitrate)}{SSIM_B(bitrate)} \tag{4.8}$$

4.3 Results

Chapter 5

Conclusion

concluding words

5.1 Future Scope

Bibliography

Bjontegaard, G. (2001). Calculation of average PSNR differences between RD-curves. *VCEG-M33*.

Boudechiche, D. E., **S. Benierbah**, and **M. Khamadja** (2017). Distributed video coding based on vector quantization: application to capsule endoscopy. *Journal of Visual Communication and Image Representation*, **49**, 14–26.

Gu, Y., **X. Xie**, **G. Li**, **T. Sun**, and **Z. Wang** (2012). Two-stage wireless capsule image compression with low complexity and high quality. *Electronics letters*, **48**(25), 1588–1589.

Iddan, G., **G. Meron**, **A. Glukhovsky**, and **P. Swain** (2000). Wireless capsule endoscopy. *Nature*, **405**(6785), 417.

Khan, T., **R. Shrestha**, **M. S. Imtiaz**, and **K. A. Wahid** (2015). Colour-reproduction algorithm for transmitting variable video frames and its application to capsule endoscopy. *Healthcare technology letters*, **2**(2), 52–57.

Lin, M.-C. and **L.-R. Dung** (2011). A subsample-based low-power image compressor for capsule gastrointestinal endoscopy. *EURASIP Journal on Advances in Signal Processing*, **2011**, 1–15.

Stickney, G. S., **K. Nie**, and **F.-G. Zeng** (2005). Contribution of frequency modulation to speech recognition in noise. *The Journal of the Acoustical Society of America*, **118**(4), 2412–2420.

Turcza, P. and **M. Duplaga** (2011). Low power FPGA-based image processing core for wireless capsule endoscopy. *Sensors and Actuators A: Physical*, **172**(2), 552–560.

Wang, A., **S. Banerjee**, **B. A. Barth**, **Y. M. Bhat**, **S. Chauhan**, **K. T. Gottlieb**, **V. Konda**, **J. T. Maple**, **F. Murad**, **P. R. Pfau**, *et al.* (2013). Wireless capsule endoscopy. *Gastrointestinal endoscopy*, **78**(6), 805–815.

.

Appendix I

Bio-Data

Name:	ABCy Malkhandi	
DOB:	Date	
Address:	Address Line 1 Address Line 2 Address Line 3	
Email:	........@gmail.com	
Contact No.:	XXXXXXXXXXX	
Qualification:		
	Bachelor of Technology (B.Tech)	
	Institution:	Technological University, Delhi
	University:	Technological University, Delhi
	Discipline:	Electical Engineering
	Year of Passing:	2018
	Master of Technology (M.Tech)	
	Institution:	National Institute of Technology, Karnataka Surathkal
	University:	National Institute of Technology, Karnataka Surathkal
	Discipline:	VLSI Design
	Registration Date:	21^{st} September 2020

.

Appendix II

Dissemination/Publication

- Chinmay Malkhandi and Rathnamala Rao, A Full-Swing, High-Speed, and High-Impedance Hybrid 1-bit Full Adder. The paper got accepted in the International Conference on Computational Electronics for Wireless Communications 2022 (ICCWC22) to be held in virtual mode during $9^{th} - 10^{th}$ *June* 2022.