Distributed Image and Video Coding

based on

Compressed Sensing

A thesis presented in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy
in
Computer Engineering
at
Massey University
New Zealand.

Muhammad Yousuf Baig

2013

Abstract

Conventional methods for encoding of images and videos is a complex process with high computational demands. They are designed for application scenarios where the signals concerned are encoded once and played back many times. However, new applications such as wireless video sensor networks demand low cost and low power cameras with limited computing resources. The focus of this thesis is on such image and video coding systems where the computational burden is shifted from the encoder to the decoder.

Three separate coding schemes have been developed – two for videos and one for images. Together they form a framework for distributed coding which is based on the theory of compressed sensing and distributed coding. Compressed sensing is a relatively new theory for the acquisition of sparse signals that allows the sampling rate to be much lower than the Nyquist limit. Distributed coding is based on the theorem by Slepian and Wolf, and Wyner and Ziv. It allows different correlated parts of a signal to be encoded independently without loss of coding efficiency. The decoding of these separately encoded parts are then decoded jointly in order to exploit the correlation between them. The main characteristics of the coding scheme proposed in this thesis are: (1) they do not require the use of traditional codecs; (2) only compressed sensing measurements are used for encoding and decoding; (3) no motion estimation and compensation are involved for videos.

The first proposed coding scheme is for the encoding of whole video frames. The compressed sensing measurement of individual frames are separately encoded. These frames are divided into key and non-key frames with the key frames encoded at a higher rate than non-key ones. While the key frames are decoded independently, the non-key ones are decoded with the help of side information generated from the measurements of the key frames. The most important part of the decoder is a simple, yet effective, side information generation method which requires only minimal computation. The side information generated is simply added to the measurements of the non-key frames for use with any compressed sensing reconstruction algorithm. The other two coding schemes are block-based coding methods. Each image or frame is divided into non-overlapping image blocks in a similar way it is done in some existing coding standards. The coding of the blocks are performed in a distributed manner by classifying them into key blocks and non-key blocks. An adaptive encoding strategy based on block similarity is also developed. Experimental analyses using publicly available test images and videos show that the performances of the simpler codecs proposed are better than other existing compressed sensing based codecs. The video codecs also out-perform conventional distributed video codec in terms of simplicity, compression ratio and decoding complexity.

The basis of these coding methods is on the correlation of frames or blocks. This correlation is established through experimental analyses. These analyses also showed that the minimum square error between any pair of them can be effectively used as a measure of correlation. In conjunction with the development of the codecs, a quantization scheme that is tailored to the statistics of CS measurements has also been proposed. This scheme yields better results than a uniform quantizer and those used for JPEG. The quantizer is also robust against different statistics of individual images. Separate experimental evaluations also show that structurally random matrices are the best sensing matrices for acquiring images and the sparse reconstruction by separable approximation (SpaRSA) algorithm produces the best reconstructed image quality.

Acknowledgments

I am profoundly grateful to my supervisor Associate Professor Edmund Lai for his constant support and guidance throughout my PhD study. There were many occasion in this journey when things were not going in a right direction but Edmund's encouragements and endless patience has helped me to finish this thesis. Edmund's knowledge and inspiring research insights were always fruitful and without him, this endeavour could have never been completed.

I am also thankful to my co supervisor Dr Amal Punchihewa for his valuable comments, suggestions and feedback throughout my research. This PhD thesis could not be completed without a financial support and scholarship from Higher Education Commission, Pakistan. I am thankful for their support and appreciate their effort for educational reforms. I am also thankful to Mehran University for granting me study leave to pursue PhD study.

I want to express my gratitude to my parents for their prayers and support. I am thankful to my wife Sana for her love, understanding and endless support despite all the difficulties. Finally love to my children, Maryam and Taha who always brings smiles back in a long and hard day

Contents

List of Figures

List of Tables

List of Abbreviations

AVC	Advanced Video Coding
BCH Codes	Bose Chaudhuri Hocquenghem Codes
BCQP	Bounded Constraint Quadratic Program
BCS	Bayesian Compressive Sensing
BM	Block Matching
BP	Basis Pursuit
BPDN	Basis Pursuit Denoising
CoSaMP	Compressive Sampling Matching Pursuit
CRC	Cyclic Redundancy Check
CS	Compressed Sensing
DCT	Discrete Cosine Transform
DCVS	distributed compressed video coding
DISCOS	Distributed Compressed Video Sensing
DISCOVER	DistributedCoding forVideo Services
DSC	Distributed Source Coding
DVC	Distributed Video Coding
Ffmpeg	Fast Foraward MPEG
GOB	Group of Blocks
GOP	Group of Picture
GPSR	Gradient Projection for Sparse Reconstruction
ISAR	Inverse Synthetic Aperture Radar
ITU-T	International Telecommunication Union
JPEG	Joint Picture Expert Group
k-t FOCUSS	Focal Under Determined System Solver in k-t Space
LASSO	Least Aboslute Shrinkage and Selecion Operator
LDPC	Low Density Parity Check
LDPCA	Low Density Parity Check Accumulate
LIMAT	Lifting-based Invertible Motion Adaptive Transform
LP	Linear Programming
MC	Motion Compensation
MCFI	Motion Compensated Frame Interpolation
MCI	Motion Compensated Interpolation
ME	Motion Estimation
MP	Matching Pursuit
MPEG	Motion Picture Expert Group
MR	Measurement Rate
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
NESTA	Nesterov's Algorithm

NP	Non-deterministic Polynomial-time
OMP	Orthogonal Matching Pursuit
PCA	Principle Component Analysis
PFT	Partial Fourier Transform
PL	Projected Landweber
PRISM	Power-efficient, Robust, hIgn compression Syndrome based Multimedia coding
PSNR	Peak Signal to Noise Ratio
QCIF	Quad Common Interchange Format
RD	Rate Distortion
RDO	Rate Distortion Optimization
RIP	Restricted Isometry Property
SBHE	Scrambled Block Hadamard Ensemble
SFE	Scrambled Fourier Ensemble
SI	Side Information
SpaRSA	Sparse Reconstruction by Seperable Approximation
SPL	Smoothed Projected Landweber
SRM	Structurally Random Matrices
SSIM	Structure Similarity Index
STD	Standard Deviation
StOMP	Stagewise Orthogonal Matching Pursuit
SVC	Scalable Video Coding
SVD	Singular Value Decomposition
TV	Total Variation
TwIST	Two-Step Iterative Shrinkage Thresholding
TWR	Through-the-wall Radar
WT	Wavelet Transform
WZ	Wyner Ziv