

SOME STUDIES IN PARAMETRIC VIDEO AND IMAGE COMPRESSION

MITHILESH KUMAR JHA



**DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
OCTOBER 2015**

SOME STUDIES IN PARAMETRIC VIDEO AND IMAGE COMPRESSION

by

MITHILESH KUMAR JHA

Department of Electrical Engineering

Submitted

in fulfillment of the requirements of the degree of

Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

OCTOBER 2015

To almighty GOD

Certificate

This is to certify that the thesis titled “**SOME STUDIES IN PARAMETRIC VIDEO AND IMAGE COMPRESSION**” being submitted by **Mr. MITHILESH KUMAR JHA** to the **Department of Electrical Engineering, Indian Institute of Technology Delhi**, for the award of the degree of **Doctor of Philosophy**, is a record of bona-fide research work carried out by him under our guidance and supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree. The results contained in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.

Dr. Sumantra DUTTA ROY

Associate Professor

Department of Electrical Engineering

India Institute of Technology Delhi

New Delhi - 110016

Dr. Brejesh LALL

Associate Professor

Department of Electrical Engineering

India Institute of Technology Delhi

New Delhi - 110016

Acknowledgements

I would like to express my most sincere gratitude to my supervisors Dr. Sumantra Dutta Roy and Dr. Brejesh Lall for their constant encouragement and guidance throughout my research. Without their support and help this thesis work would have been impossible. They have been a constant source of support and encouragement.

I am extremely grateful to members of the SRC : Prof. Santanu CHAUDHURY, Prof. S. D. JOSHI and Prof. Subhasish BANERJEE for their valuable guidance, encouragement and helpful comments. I sincerely want to thank Prof. Santanu Chaudhury for his extended support starting from research proposal and Prof S. D. JOSHI for his constant encouragement especially in a difficult phase. I must mention Dr. Mona MATHUR for all her advice and support at early days.

I also wish to thank all my friends, especially Ritu GARG, Subarna TRIPATHI, Amioy KUMAR, Ravi CHAUDHARY, Ehtesham HASSAN, Saurabh GORECHA and Pushpendra K. SINGHAL for their support and encouragement. I have been very fortunate in having very friendly and helpful people in all the labs I have worked in.

And finally, I must not forget to salute my parents, my wife and all other family

members for their understanding and patience throughout this long journey.

NEW DELHI

Mithilesh Kumar JHA

Abstract

The work presented in this thesis suggests a set of tools and techniques for an efficient representation of texture videos and images using the Parametric Compression framework. The thesis also presents an architecture to integrate these tools with H.264/HEVC to propose an efficient encoder. Further we describes a new motion estimation framework using multi-resolution restricted affine model and a particle filtering based search technique integrated with H.264/HEVC.

First we propose an Statistically Matched Wavelet (SMWT) based pre-processing tool which provides an efficient representation of a texture video frame. We encode the approximation subband coefficients (LL) using a standard image compression scheme like JPEG2000. The detail subband coefficients, that is, HL , LH , and HH are jointly encoded in a compressive sensing (CS) framework and can therefore be represented with fewer measurements resulting in higher compression.

Next we propose a Directional Empirical Mode Decomposition (DEMD) based pre-processing and synthesis technique. The texture video sequence is divided in *key* frames and *non-key* frames. The key frames are encoded using standard JPEG 2000, while

the *non-key* frames are reconstructed at the decoder using the DEMD residue and *key* frame IMF decomposition. Only the very low information DEMD residue is transmitted for *non-key* frames. The DEMD residue is encoded using CS in this approach. At the decoder, first the IMF's of the *non-key* frames are synthesized from the DEMD representation of key frames using the information present in the residue of the current frame. The IMF's are then combined to synthesize the entire frame. We extended the work by integrating the DEMD-based technique with the standard H.264/HEVC. At the encoder, the input data is pre-processed using DEMD. The DEMD residue images of all *non-key* frames (*P* and *B*-frames for H.264/HEVC) are encoded using the standard encoder. At the decoder, first the IMF's of all *non-key* frames are synthesised using correlation search and patch-based multi-level IMF synthesis from the DEMD representation of *key* frames (*I*-frames in H.264/HEVC) and the DEMD residual image of the frame being synthesised. These IMF's are then combined along with the residue of the corresponding frame to synthesize all the *non-key* frames.

The final work presented in this thesis is a Restricted Affine Motion Compensation Video Coding scheme that exploits Particle Filtering and Importance Sampling in a multiresolution framework. In this work, We propose three modifications to widely used traditional motion compensation techniques. A higher order linear model is proposed in place of the traditional translational model and the traditional search methods are replaced by an efficient particle filtering-based search, which incorporates information from both spatial and temporal neighbors. The multiscale framework enables efficient parameter estimation while resulting in higher compression due to more accurate motion modeling.

Table of Contents

Acknowledgements	iii
Abstract	v
List of Figures	xiii
List of Tables	xvii
1 Introduction	1
1.1 Scope and Objective	8
1.2 Major Contributions of the Thesis	12
1.3 Layout of the Thesis	16
2 Parametric Video and Image Compression : A survey of the literature	19
2.1 Model Based Video and Image Compression	21
2.1.1 Geometric Modeling	21
2.1.2 Object Based Modeling	26

2.1.3	Motion Based Modeling	29
2.2	Mosaic-based Video Compression	33
2.3	Texture Based Video Compression	37
2.3.1	Texture Analysis	38
2.3.2	Texture Synthesis	44
2.4	New Context of Compression Problem	51
2.4.1	High Dynamic Range (HDR) Video and Image Compression . . .	52
2.4.2	Cloud Storage	53
2.4.3	3D TV	55
2.4.4	Scalable Video Coding (SVC)	56
2.4.5	4K2K UHD TV	57
2.5	Motivation for the Present Work	58

3 SMWT-based Texture Synthesis in a Compressive Sensing

	Framework	67
3.1	Introduction	67
3.1.1	Motivation	67
3.1.2	Contributions of this Chapter, in a Nutshell	69
3.1.3	Related Work in the Area	72
3.1.4	Our Work in Perspective	74
3.2	SMWT-based Texture Representation	75
3.3	Compressive Measurement and Encoding	80

3.4	Texture Synthesis Framework	84
3.5	Results and Discussion	87
3.5.1	Statistically Matched Wavelet filter estimation	87
3.5.2	Texture synthesis results	88
3.6	Conclusion	100
4	DEMD-based Texture Video Frames Compression in a CS	
	Framework	101
4.1	Introduction	101
4.1.1	Motivation	101
4.1.2	Contributions of this Chapter, in a Nutshell	102
4.1.3	Related Work in the Area	103
4.1.4	Our Work in Perspective	106
4.2	DEMD and CS-based Compression	107
4.2.1	DEMD Overview	107
4.2.2	Compressive Sensing Overview	109
4.2.3	DEMD and CS-based Encoder Framework	111
4.2.4	DEMD and CS-based Decoder Framework	113
4.3	Experiments and Results	116
4.4	Discussion on the Computational Complexity	127
4.4.1	Empirical Mode Decomposition	127
4.4.2	Compressive Sensing	129

4.5	Conclusion	130
5	DEMD-based Video Coding for Textured Videos in an MPEG/H.264/HEVC	
	Framework	131
5.1	Introduction	131
5.2	DEMD-based Texture Video Compression Framework	137
5.2.1	DEMD-based Encoder Framework	138
5.2.2	DEMD-based Decoder Framework	139
5.3	Results and Discussion	144
5.4	Discussion on the Computational Complexity	150
5.5	Conclusion	155
6	Restricted Affine Motion Compensation in Video Coding	
	with Particle Filtering and Importance Sampling	157
6.1	Introduction	157
6.2	The Proposed Motion Estimation Model	161
6.2.1	A Restricted Affine Motion Model	162
6.2.2	Particle Filtering-based Motion Estimation	163
6.2.3	Multi-Resolution Particle Filtering; Importance Sampling	165
6.3	Integration and Experiments in an H.264/HEVC framework	169
6.4	Discussion on the Computational Complexity	179
6.5	Conclusion	182

7	Conclusions	183
7.1	Contributions	184
7.1.1	The work in perspective:A summary of comparison with standard methods	187
7.2	Scope for Future Work	188
	Bibliography	192
	Publications	212
	Biography	214

List of Figures

1.1	Thesis work flow chart	10
1.2	Block based representation of the proposed Parametric Video and Image compression solution	11
2.1	Object based coding system overview	26
2.2	Mosaic based coding system overview.	34
2.3	Texture based video compression system overview	38
2.4	Statistically Matched Wavelet Filter Bank Overview	41
3.1	Overview of the SMWT and CS based texture synthesis framework	72
3.2	Statistically Matched Wavelet Filters Estimation Framework	76
3.3	Illustration of statistically matched wavelet decomposition	79
3.4	Illustration of statistically matched wavelet coefficients distribution in ap- proximation and detail sub-space	80
3.5	Sparse representation of Statistically Matched Wavelets sub-bands coefficients for <i>brickwall</i> and <i>escalator</i> texture.	81

3.6	SMWT and CS Based Encoder Framework	83
3.7	SMWT and CS Based Decoder Framework	84
3.8	Illustrating the significance of approximation and detail subbands coefficients in image reconstruction	93
3.9	Synthesis results of structural texture with pseudo-periodic patterns . . .	94
3.10	Synthesis results of statistical texture	95
3.11	Synthesis results of structural texture with periodic patterns	96
3.12	Synthesis results of structural texture with irregular and regular patterns .	97
3.13	Synthesis results of structural texture with irregular patterns	98
3.14	Synthesis results of complex structural texture with irregular patterns . . .	99
4.1	Results for DEMD residue encoding of a frame using JPEG2000	104
4.2	DEMD and CS Based Encoder framework	111
4.3	DEMD and CS Based Decoder framework	113
4.4	Synthesis results for sequence <i>brickwall</i>	120
4.5	Synthesis results for sequence <i>escalator</i>	121
4.6	Synthesis results for sequence <i>blackhole</i>	122
4.7	Synthesis results for sequence <i>floorbox</i>	123
4.8	Synthesis results for sequence <i>dog_grass</i>	124
4.9	Synthesis results for sequence <i>brdf</i>	125
4.10	Rate Distortion results for test sequences <i>brickwall</i> , <i>escalator</i> , <i>floorbox</i> , <i>dog_grass</i> , <i>circularhole</i> and <i>brdf_image</i>	126

5.1	Results for DEMD residue encoding of <i>I</i> frames using H.264	133
5.2	DEMD and H.264/HEVC based Texture Video Compression Framework .	134
5.3	The dominant texture direction, for the <i>brickwall</i> sequence.	137
5.4	The Encoder framework	138
5.5	The Decoder framework	140
5.6	The IMF Decomposition and synthesis flow	141
5.7	Illustration of patch based synthesis using a ‘flat search’ versus a correla- tion search	143
5.8	Compression gain over MPEG2/H.264 for the proposed multilevel IMF synthesis	149
5.9	Synthesis results for sequences <i>brickwall</i> , <i>escalator</i> , <i>circularhole</i> and <i>floor- box</i>	153
5.10	Synthesis results for sequences <i>party_scene</i> , <i>dog_grass</i> , <i>fabric</i> and <i>bridge_far</i>	154
6.1	affine motion model reduces the residual energy as compared to the stan- dard translational model: an example	163
6.2	Illustration of the multiscale framework	165
6.3	Illustration of coarse-to-fine strategy	169
6.4	Some representative test sequences, and their characteristics	170
6.5	Categorisation of macroblocks in a representative frame of the ‘Football’ Sequence (CIF)	174
6.6	RD Plot comparing JM and JM_{Affine}	175

6.7	Illustration of the % decrease in the SSE with a variation in the Number of Particles	178
6.8	The percentage SSE decrease over a single level of resolution, with 2, 3 and 4 levels, for representative high resolution videos	182

List of Tables

2.1	3D Geometric Modeling : A Comparative Study for a video sequence	
	"Thabor (CIF, 25HZ)" encoded at different bit rates	25
2.2	Object Based Coding : A Comparative Study for compression gain claimed	
	vs. standard MPEG/H.264	28
2.3	Texture Analysis and Synthesis Scheme : A Comparative Study	66
3.1	Estimate of statistically matched Wavelet Filters	91
3.2	Illustrating the PSNR and MOS data of the proposed SMWT and CS model	92
4.1	Compression gain over a standard DWT-CS scheme for a single level of	
	IMF decomposition	119
5.1	Compression gain over MPEG2 for a single level of IMF decomposition .	146
5.2	Compression gain over H.264 for a single level of IMF decomposition . .	147
5.3	Compression gain over HEVC for a single level of IMF decomposition . .	148
5.4	Time complexity analysis for a single level of IMF decomposition in H.264	151
5.5	Time complexity analysis for a single level of IMF decomposition in HEVC	152

6.1	JM 17.0 encoder configuration file used	172
6.2	HM 3.1_dev encoder configuration file used	173
6.3	Comparison of the Proposed Method (JM_{Affine}) with H.264 (JM-17.0, 'JM')	176
6.4	Comparison of the Proposed Method (HM_{Affine}) with HEVC (HM 3.1_dev, 'HM')	177
6.5	Time complexity analysis for JM, JM_{affine} , HM and HM_{affine}	181