New Context of Compression Problem and Approach to Solution: A Survey of the literature

Dr. Mithilesh Kumar JHA*, Dr. Sumantra Dutta Roy, Dr. Brejesh Lall
Department of Electrical Engineering
Indian Institute of Technology, DELHI, India
Email : jham73@gmail.com
Email : sumantra@ee.iitd.ac.in
Email : brejesh@ee.iitd.ac.in
* Corresponding Author

Abstract: The new generation multimedia applications such as Ultra High Definition TV, 3D Video and Cloud Storage keeps the compression research relevant even post to the advent of the state-of-art standard HEVC codec. The continued studies on human visual system suggest that a perceptually acceptable reconstruction with reduced bandwidth requirements may be more acceptable and commercially viable solution rather than a bit accurate reconstruction for a variety of video and graphics applications. One way to achieve compression beyond that achieved by the advanced codec like HEVC, is to sample the input data at a lower rate than the Nyquist rate, exploit the sparsity and perceptual redundancy in the content using advanced signal processing and computer vision tools and represents it using a spatio-temporal model. This paper contains an exhaustive review of the new context of video compression methods along with their limitations and open issues.

Keywords: Parametric Video Compression (PVC); Texture Video; 3-D Mosaic; Directional Empirical Mode Decomposition (DEMD); Statistically Matched Wavelet (SMWT); Compressive Sensing (CS); UHD TV (Ultra High Definition Television); High Dynamic Range (HDR); High Frequency Rate(HFR);Wide Colour Gamut(WCG)

Biographical notes: Dr. Mithilesh Kumar Jha has done Ph.D, from Department of Electrical Engineering, Indian Institute of Technology, DELHI, INDIA and has 18 year of professional experience in the semiconductor product companies at various capacities.

Dr. Sumantra Dutta Roy and Dr. Brejesh Lall, is an Associate Professor at Department of Electrical Engineering, Indian Institute of Technology, DELHI, INDIA.

1 Introduction

Video compression technologies have continued to evolve and improve in the coding efficiency in past decades. This is evident with the development of several successful video coding standards, such as MPEG1 [1], MPEG2 [2], H.264/MPEG-4 Advanced Video Coding (AVC) [3] and the most recent HEVC [4]. Although the storage capacity and the transmission bandwidth continue to increase, further improvement in video coding efficiency is still of prime interest due to the advent of large data size video applications. A variety of video applications such as Broadcast Video, Video Telephony and Video on Demand (VOD) requires better compression with acceptable perceptual reconstruction quality, instead of bit accurate (high PSNR) reconstruction [5]. This is because the human brain is able to decipher important variations in data at scales lower than those of the viewed object [6, 7]. The conventional method does not account for Human Visual System (HVS) or content of information [8]. As a result, various new generation compression technologies have been developed to improve the coding efficiency further. Though the recent deployment of the HEVC has reduced the hunt for a new generation video compression technology, compression still remains a potential research area (especially for an application where bit accurate reconstruction is not a must requirement, for example a video texture [9]). In addition, the new challenges in compression such as UHD TV [10], 3D Video and Cloud Storage have created a lot of recent interest and therefore we also discuss about the relevance of new generation compression methods in this context.

The new generation research on video compression suggests various schemes to exploit the statistical, temporal and perceptual redundancy across frames in the appearance of an object and texture present to achieve a higher compression rate than that of conventional video coding schemes such as H.264/HEVC. The advances in pattern recognition, computer vision and advanced signal processing tools and techniques have contributed significantly towards pre-processing and feature based analysis of picture contents. The goal of a novel contemporary lossy compression scheme should be to reduce the entropy while preserving the perceptual quality of the frame using high-level computer vision tools and techniques. Most of these high-level processing techniques result in a small number of semantically relevant salient features, which can represent the entire signal very accurately. Efficient pre-processing and modelling of the video content with temporal link is the most challenging aspect of such an algorithm. A model for image sequence can be constructed by any of the following two methods: (a) Model Learned/Formulated apriori, (b) Models Generated on the fly. When the model is learned apriori, the model remains same throughout the duration of the video sequence, whereas when the model is learn on-the-fly it keeps adapting through the duration of the video sequence. The most cited model based compression method includes, (i) Geometric / Structural model such as Mesh based model [11, 12], (ii) Object based model like Eigen-space models & Probabilistic model [13, 14, 15] (iii) Higher order multi-resolution motion model [16], (iv) 3D Mosaic representation [17, 18] and (v) Texture based compression [19, 20, 21, 22]. Next, we present a brief discussion on these techniques.

1.1 Geometric mesh modelling

Geometric models are extracted on the y from the scene using computer vision techniques based on structure from motion and self-calibration methods [23, 24, 25, 26, 27, 28]. Tung et al. [29] have suggested a 3D video compression scheme based on multi-resolution Reeb Graph [30] which is used as a motion descriptor to track similar nodes all along the 3-D video sequence. The next step in model based compression is the efficient encoding of the extracted model along with the control parameters. Most of the older geometry compression methods are predictive schemes [31, 32, 33, 34] that are tightly coupled with the connectivity coding methods. Recently Authors in [35, 36, 37] have proposed scalable predictive coding which supports spatial as well as temporal scalability. Principal component analysis (PCA) based representation of a mesh sequence [38, 39] and transform based mesh geometry coding [40, 41, 42, 43, 44] are some of the popular methods of mesh parameters coding. Though the 3D model based video compression looks a natural fit, their success in image and video application are limited primarily because of two reasons, (i) the projection of the 3D scene onto the image plane results in a large information reduction and therefore a 3D reconstruction becomes very difficult and has to overcome many ambiguities, (ii) such techniques are computationally complex as the model needs to be extracted and adapted regularly for a dynamic scene. Because of this, most of the practical compression schemes are based on random process models.

1.2 Object based parametric model

In the appearance based object modelling approach, a moving object in a scene is tracked through an appearance based tracker (for example an Eigen tracker) and represented in the form of the projection coefficients onto the learned subspace and motion parameters [45, 13, 46]. The encoder codes the estimated objects (in the form of the projection coefficients onto the learned subspace and motion parameters) along with the background reconstruction error and object reconstruction error. Figure 1, provides an overall system view of a typical object based video compression method. The main problem in object-based coding is the automatic extraction and modelling of objects directly from the image intensities. This task may require complex image analysis techniques to segment the scene into homogeneous regions or even user interaction so that image regions correspond to real objects on the scene. A content-based image analysis is proposed recently using Principal Component Analysis (PCA) and Self Organising Map (SOM) targeting to image retrieval application [47]. There are two kinds of methodologies available in the literature for automatically extracting objects in the frame. The first category is off-line analysis [48] and the second one is on-the-fly object detection [49]. The off-line object analysis is computationally complex which prevent its usage in on-line applications required in the embedded system. The on-the-fly object detection approaches are mostly based on initial statistical change detection step but lack the generalization and robustness. A proper semantic object segmentation applies to every scenario is still unreported.

1.3 Higher order multi-resolution motion model

The conventional motion estimation techniques are not efficient in the prediction of complex motion regions involving rotation and zoom and thus limiting the overall compression which can be achieved. HEVC [4] and H.264 [3] approximates such complex motion by dividing bigger blocks and using multiple reference frames, but the corresponding motion vectors, block mode signalling and RD optimization are substantial overheads which affect the overall coding gain [50]. The new generation parametric video compression framework proposes various new motion estimation strategy such as affine based motion model [51, 52, 53, 54, 55, 56, 57], Multi-resolution motion estimation [58, 59, 60] and a tracker based motion estimation [61] to handle the complex motion like rotation and zoom and to improve the coding efficiency. A 3D-to-2D transformational approach proposed to exploit the complex spatio-temporal correlation through 2-D spatial transform [62, 63]. A local restricted-affine based motion representation leads to greater flexibility and robustness in classifying the dynamic activity of the objects in any region of the video [16, 64].

1.4 3D Mosaic representations

The motivation for Mosaic Based Compression (MBC) is the observation that the physical world observed by a camera typically consists of some moving objects over a dominant stationary background. The large scale temporal and spatial co-relation of this background layer is highly compressible [65]. There are three technical challenges in generating a content based 3-D mosaic representation from a long image sequence. They are: i) how to estimate a robust and accurate camera orientation for many video frames, ii) how to generate a seamless video mosaic with motion parallax, and iii) how to reconstruct a 3-D large-scale urban scenes Some work has been done in 3-D reconstruction of panoramic mosaics [66, 67] with an off-centre rotation camera, but the methods are limited to a fixed view-point camera instead of a moving camera, and usually the results are low-level 3-D depth maps of static scenes, instead of high-level 3-D structural representations for both static and dynamic target extraction and indexing. Zhu et all [17, 68] have suggested a Content-Based 3D Mosaic representation for long video sequences of 3-D and dynamic urban scenes captured by a camera on a mobile platform based on parallel perspective push broom stereo geometry algorithm [69]. Figure 2, shows an overall structure of mosaic based coding. In order to use the 3D content based mosaic representation for real applications, further enhancements are needed, for example in the current implementation, only 3-D parametric information of planar patches in a single reference mosaic is obtained. This needs to be further extended to produce multiple depth maps with multiple reference mosaics due to different viewing directions shown in mosaics. The camera orientation estimation with many video frames is still a challenging issue and finally, the computation cost of the mosaic generation due to the registration of each frame is also a matter of concern in this area.

1.5 Texture Based Compression

Textures [9] are homogeneous patterns that contain spatial, temporal, statistical and perceptual redundancies, which intensity or colour alone cannot describe adequately. Texture based video compression is getting a special attention in the compression world for three reasons. They are: (i) texture is perceived in such a manner that it does not demand bit accurate reconstruction [70, 71, 5], (ii) the quasi-stationary pattern in texture enables the reconstruction of the visually-similar texture by exploiting its statistical properties at decoder, (iii) the conventional block-based transform does not work well for texture data due to its fine details and high-frequency components. Texture analysis and synthesis based video compression scheme [72, 21, 73] is one such method which addresses the issues with the texture video compression. Texture video in-painting based strategy [74, 75, 76, 77, 78, 79] has shown lot of potential in the image and video synthesis and achieved an acceptable perceptual quality level to be integrated in the user applications. Video epitomes [80, 81] were recently proposed as patch based probabilistic modelling approach for video texture synthesis. The prime focus of this paper is to review in detail all the texture analysis and synthesis based compression techniques using advanced signal processing tools especially in the context of new compression problem such as UHD TV, Cloud storage and 3D Video applications.

The rest of the paper is organized as follows: Section 2, provides an exhaustive review of texture based Video Compression methods. Section 3 contains the relevance of parametric compression in the context of next generation multimedia applications. In Section 4, we present our analysis followed by a conclusion in Section 5.

2 Texture Analysis and Synthesis

Texture analysis [82, 83, 84] includes efficient representation and encoding of texture data, while texture synthesis [85, 86, 87] reproduces a perceptually acceptable texture data at the decoder for compression [88, 89]. We present the discussion in two broad categories i.e. Texture Analysis and Texture Synthesis. This is for the ease of presentation only. Texture Analysis and Texture Synthesis are used in combination for all video compression techniques. Figure 3, provides a generic overview of a typical texture analysis and synthesis based video compression system.

2.1 Texture Analysis

Texture analyser block in Figure 3, has three major functions, (i) separate texture and non-texture region in a video sequence, (ii) efficient representation and feature extraction of the texture data and (iii) efficient encoding of texture and non-texture data. Texture feature extraction is used to measure local texture properties in an image or video and is further utilized for segmentation and synthesis purposes. Generally four approaches have been used to extract texture features viz. statistical-based methods, model-based methods, transform or spatial-frequency methods and structural methods.

In statistical methods, characteristics of homogeneous regions are chosen as the texture features such as the co-occurrence matrix or geometrical features such as edges [90]. Model-based methods assume that the texture can be described by a stochastic model and uses model parameters to segment the texture regions. Authors in [91] proposed a multiresolution Gaussian autoregressive model and in [92], an image model is proposed using an autoregressive time series. Wavelets based sub-band decomposition is often used in spatial-frequency methods [93]. Portilla and Simoncelli [94, 95] have proposed a statistical model for texture images based on joint statistical constraint on the wavelet coefficients. Structural methods are based on the concept that textures are composed of patterns that are well-de ned and spatially repetitive [96]. The spatial texture models described above operate on each frame of a sequence independently of the other frame in the same sequence. This can create a visual artefact across the sequence. One can address this problem by using spatial-temporal texture models [97, 98] or using motion compensation for the texture models in each frame [70]. Once the texture features are extracted and feature vectors formed, segmentation is used to divide the image into different regions based on their texture properties. Two different techniques are generally used i.e. direct segmentation or classification methods [99, 93, 100] and a grouping metric [90, 91].

2.1.1 Parametric Texture Analysis

In the following sub section, we present Matched Wavelet and Empirical Mode Decomposition as a 2-D signal analysis and parametric feature extraction tools. In the later part of the paper we discuss an application of these tools in the video compression framework proposed by the authors.

Statistically Matched Wavelet. Over the last decade, a lot of work has been carried out by various researchers to find wavelets matched to signal to provide the best representation for a given signal. Gupta et al. [101] have estimated matched wavelets for deterministic signals in the time domain. The method is based on maximizing the projection of the given signal onto a successive scaling subspace and minimization in the wavelet subspace. Tewfik et al. [102] have also designed a wavelet matched to a signal in the time domain. Rao and Chapa [103] have later proposed an algorithm to design a wavelet matched to a signal. They have proposed a solution to find a wavelet that looks like the desired signal for the case of orthonormal multiresolution analysis with bandlimited wavelets. However, the method is computationally expensive, and the problem has been addressed for deterministic signals only. A similar work have been carried out by Wu-sheng [104] and Tsatsanis [105] to design signal-adapted filter banks, but these methods find the solution of constrained minimization problems in terms of a coding gain criterion that lead to very complicated solutions. Aldroubi and Unser [106] have proposed a method to find a matched wavelet by projecting the signal on to an existing basis. Most of these methods are presented for a two-band wavelet system and are quite complex.

Gupta et al. [107], have later proposed Statistically Matched Wavelet (SMW) for the estimation of wavelets that is matched to a given signal in the statistical sense. Estimation of analysis wavelet filters from a given signal is the major contribution of this work. This concept is further extended for 2-D data by the authors in [108] and used for document retrieval application. Figure. 4 give an overview of a generic 2-D two-band separable kernel wavelet system. In this figure, x and y represents the horizontal and vertical directions respectively. The scaling filter is represented by f_{0x} and f_{0y} while the wavelet filter is represented by f_{1x} and f_{1y} corresponding to the horizontal and vertical direction. The dual of them are represented by h_{0x}, h_{0y}, h_{1x} and h_{1y}. The input to this system is a 2-D signal (for example a video frame). The output of Channel-1 in Fig. 4 is called approximation sub-space or scaling sub-space while the outputs of the other three channels (Channel-2,3,4 in Fig. 4) are called detail sub-space. This system is designed as a biorthogonal wavelet system so that it satisfies the conditions for perfect reconstruction of the two-band filter bank. Statistically matched wavelet based representation of a picture causes most of the captured energy to be concentrated in the approximation subspace, while very little information is retained in the detail subspace [108].

Directional Empirical Mode Decomposition. Empirical Mode Decomposition [109],[110], is a data analysis approach which can handle any class of signals, for example, non-linear, linear, non-stationary, and stationary. Moreover, the decomposition at each level is a simple numerical technique, and has an associated concept of a local scale (of an oscillation), and involves a perfect reconstruction. Huang et al. [111] proposes Empirical Mode Decomposition (EMD) as a decomposition of a signal f(t) into a `low frequency/residue' term $r_1(t)$ and a `high frequency/detail' part $i_1(t)$. Each $i_k(t)$ can be similarly decomposed, to give a K-level decomposition. Here, functions $i_k(t)$ are the Intrinsic Mode Functions (IMFs) and $r_K(t)$ is the residual trend (a low-order polynomial component) [111],[109],[110]. The IMFs are obtained from the signal by means of an algorithm called the sifting process. The sifting procedure is based on two constraints: (i) each IMF has the same number of zero-crossings and extrema, (ii) each IMF has symmetric envelopes de ned by the local maxima and minima respectively. Furthermore, it assumes that the signal has at least two extrema.

Nunes et al. [112] have extended 1D-EMD and suggested an implementation of 2D-EMD called as a Bidirectional Empirical Mode Decomposition (BEMD) targeting image compression application [113]. Two open issues were observed with BEMD : (i) exploitation of the internal direction in an image in BEMD, and (ii) feature extraction and utilization from BEMD. Liu et al. [114] have extended the basic idea and propose Directional EMD (DEMD), considering a dominant image direction and three feature values are extracted from decomposed components for each pixel. This method is used for texture classification [115] and image segmentation [116, 114].

$$f(x,y) = \sum_{k=1}^{K} i_k^{\theta}(x,y) + r_K^{\theta}(x,y) \quad (1)$$

The inherent image direction, can be estimated from the maximum of the Radon transform of the spectrum [117] using a Wold decomposition-based method [118]. DEMD has provided a robust algorithm for image processing application; however two issues required further attention: (i) sifting stopping rule, and (ii) stopping criterion of DEMD decomposition process. Initially Huang et al [111] have implemented a Cauchy-like convergence criterion [119] for automatically deciding when to stop sifting. Based on minimizing the difference between residuals in successive sifts to below a predetermined level, this criterion did not explicitly take into account the two IMF conditions, so the predetermined level could be obtained without the two IMF conditions being satisfied [120]. Rilling et al [121] have devised an alternate stopping criterion using two thresholds, one designed to ensure globally small fluctuations in the mean of the cubic splines from zero, and the second allowing small regions of locally large deviations from zero. Comparative study with Fourier and Wavelet proved that EMD decomposition provides much better temporal and frequency resolutions [122]

Zhang et al. [117] propose a DEMD-based image synthesis approach with a small representative texture sample used as a reference. This texture sample is decomposed using DEMD for efficient representation of the texture information. To reconstruct a full-resolution image, the authors start from a highest level (say, k) of the DEM decomposition of the representative sample. They construct the k-th IMF corresponding to the full-resolution size by taking information from smaller patches inside the sample, and placing them in the template for the full-resolution IMF (with overlap, to enforce smoothness). For the next lower IMF (k 1th) onwards, the authors perform a correlation search for each smaller block in the full-resolution template, with the closest-matching smaller block in the k 1th level decomposition of the representative texture sample. The synthesized full-resolution image is the DEMD synthesis of the full-resolution IMFs. This approach shows good results for textured image synthesis with perceptually acceptable quality. However this approach leaves the following important questions unanswered: (a) the selection criterion for the small representative texture: its size and location, (b) selection criterion for the size of the patch, and the extent of the overlap (10-20% of the patch size in examples in the paper), (c) objective assessment of the quality of the synthesized image, (d) the level of IMF decomposition required for optimal synthesis, (e) inability to handle texture sequences with irregular patterns, as the authors themselves mention [117].

2.2 Texture Synthesis

Texture synthesis process can be broadly classified as parametric approach [123, 124, 94, 95, 125, 126, 72], where an image sequence is modelled by a number of parameters such as the histogram or correlation of pixel values, and non-parametric approaches [127, 128, 86, 85, 129, 21], where synthesized texture is derived from an example texture as the seed that is known a priori. Inverse texture synthesis based scheme has been proposed in [87], where given an input texture, a reference is synthesized that allows subsequent re-synthesis of a new instance and shown promising results.

Parametric approach can be further categorized into block-based approaches [72, 130, 131, 132, 133] and region-based approaches [65, 134, 135, 136, 137]. The block-based parametric approaches allows easy integration of synthesis based techniques into hybrid block-based codecs as H.264/AVC and HEVC while targeting a pixel-accurate reconstruction of the input signal. In general, this kind of coding framework replaces, improves or extends the existing intra- and inter-prediction modes with recent coding methods exploiting the perceptual redundancy. Alternatively, region based parametric approaches allows for design of content-adaptive compression methods with perceptually acceptable reconstruction. In such schemes exact pixel reconstruction is not the goal, instead the aim is synthesizing the pixels at the decoder with visual correctness instead of bit-accurate reconstruction.

The most promising block-based parametric coding approach in recent times are based on Auto Regressive (AR) [130, 138] and Auto Regressive Moving Average (ARMA)-based modelling [131, 132, 133] methods. Khandelia et al [130] have proposed an AR-based hybrid video codec that provides significant compression gain (up to 54.52 % at QP=12, for Bridge far video sequence) over the standard H.264 for similar visual quality. In this method, each block is first classified as an edge or non-edge block using gradient-based edge detector. The non-edge blocks are coded using a spatio-temporal auto-regressive model. The AR coefficients are sent to the decoder as a side information. All edge-blocks are encoded using standard block-based coding scheme such as H.264. Experimental results presented by the author suggest that the compression gain drops for increased QP. In the ARMA-based synthesis model presented in [131, 132], an extrapolated frame replaces the previous frame in the reference picture buffer using the concept of H.264/AVC. The authors use the ARMA-based non-rigid texture synthesis approach as proposed in [123] for the frame extrapolation. The authors propose to use five previously generated training frames for synthesizing the current frame. This algorithm was integrated into the JM12.4 reference software and a bit rate saving of upto 10% were claimed for "Container" QCIF sequence at QP=23 using PSNR as the quality criterion. The authors [132] have further enhanced their algorithm and suggested using multi-pass encoding with the position of the reference picture changing during each pass. The position of the synthesized picture in the decoded picture buffer is adaptively estimated. Bit rate savings of upto 15% were claimed for the "bridge far" QCIF sequence at QP=23 as compared to H.264(JM12.4). Targeting to improve the coding gain further over H.264, Chen et al. [133] have proposed a novel non-rigid texture model [123] and proved that this scheme performs better as compared to Stojanovic et al. [131] for global illumination changes between frames and for intra prediction. AR- and ARMA based approach are suitable for the textures with stationary data, like steady water, grass and sky, however they are not suitable for structured texture with non-stationary data as blocks with non-stationary data are not amenable to AR modelling. Further, they are block-based approaches, and blocking artefacts can appear in the synthesized image. Parametric approaches can achieve very high compression for the sequence with dominant texture areas at low computational cost.

In the region-based parametric approach, Wang et al. [65] have proposed to represent a coherent motion region as a set of overlapping layers and model the texture by a set of affine parameters. The author suggested to send the alpha, intensity and motion parameters as side information and proposed to use JPEG framework for coding the layers. This approach requires a very precise segmentation of the input video to reduce the visual artefacts. Since the proposed scheme is in an open-loop framework, there is no mechanism to identify artefact due to erroneous segmentation. Dumiritas et al. [139, 134] have proposed a compression method based on texture replacement, where it replaces the input texture with a perceptually similar texture. The input is separated into replaceable and non-replaceable region at the encoder. Parameters are extracted for replaceable region using Wavelet and sent to decoder as a side parameter. Non-replaceable regions are encoded using a standard compression scheme. Replaceable textures are removed from the encoding process and synthesized at the decoder using side information. The application of such an approach is limited to sequence with no or very little global motion of the replaceable textures. Ndjiki-Nya et. al. [140] has proposed a texture analysis and synthesis framework using the frame-to-frame displacement and image warping techniques. This method is however not effective for non-rigid texture objects and the frame-by-frame synthesis can yield temporal inconsistencies. A similar texture synthesis technique based on texture analysis and segmentation using frame-to-frame mapping and edge-based inpainting was proposed in [135, 136, 141]. The authors claimed to achieve up to 35% bit rate saving compared to H.264 for the same visual quality. The author in [135, 136] however do not specify, how visual quality was assessed. The method presented by Oh et al. [142] uses a low quality video as side information to synthesize the texture at the decoder. The proposed approach is effective to control the quality of the synthesized texture and allows for a quality vs compression trade-off. A cube-based texture growing method was proposed in [143, 144], which can be viewed as an extension of [85] from image to video. Temporal consistency is claimed in [143] due to cube-based synthesis. However, the global texture structure can be easily broken by texture growing with the raster-scanning order. Also, the output texture tends to have repeated patterns because of the fixed size cube-based growing procedure. Zhang and Bull [137, 71] have proposed a region-based texture synthesis model where each synthesizable frame is segmented into homogeneous regions using spatial texture segmentation method [145] and textured and non-textured regions are separated based on statistical properties [146]. The motion vectors in rigid texture regions are used to compute the parameters of a perspective motion model based on least square method. In this method authors have made an assumption that the variation in a texture is temporally uniform across a small group of frames (five frames in this work). Authors claims a significant bit rate saving (up to 50% for high bit rate encoding) as compared to H.264 for similar visual quality.

Non-parametric approaches are mostly based on Markov Random Field (MRF) theory [147, 148, 9, 149, 85, 143] and can be classified as pixel-based approaches [128] or patch-based approaches [85, 150, 151]. Efros and Leung [128] proposed pixel-based non-parametric sampling to synthesize texture. Wei and Levoy [148] further improve the method using a multi-resolution image pyramid based on a hierarchical statistical method. A limitation of the above pixel-based methods is an incorrect synthesis owing to incorrect matches. Patch based methods overcome this limitation by considering features matching patch-boundaries with multiple pixel statistics. The popular choices for patch-based methods are based on graph-cut techniques [85, 152]. Nonparametric approaches can be applied to a wide variety of textures (with irregular and structural texture patterns) and provide better perceptual results (higher Mean Opinion Score (MOS) values). However, these schemes are often computationally more complex

2.2.1 Compressive Sensing Based Texture Synthesis

Leveraging on the concept of transform coding, Compressive Sensing (CS) [153, 154] enables a potentially large reduction in sampling and computation costs for sensing signals that have sparse or compressible representation (by a sparse representation, we mean that for a signal of length N, we can represent it with K << N nonzero coefficients). Compressive Sensing framework opens a new research dimension in which most of the sparse signals can be reconstructed from a small number of measurements (M), using algorithms like convex optimization, greedy methods and iterative thresholding [155, 156, 157]. CS recovery algorithms mainly depend on two fundamental principles: Sparsity and Incoherence [158]. Compressive Sensing framework mainly consists of three stages, i.e. sparsification by transformation, measurement (projection) and optimization (reconstruction). Designing a good measurement matrix with large compression and designing a good signal recovery algorithm, are the two major challenges of applying the CS technique in image or video compression.

Significant theoretical contributions have been published on the compressive sensing in recent years [153, 154, 155] for image processing applications [159, 160, 161, 162, 163]. Wakin et al. [162] have used a 3D wavelet based inversion algorithm to achieve video CS for single-pixel cameras. A more sophisticated method was developed in [164], where the evolution of a scene is modelled by a linear dynamical system (LDS) and the LDSs parameters are estimated from the compressive measurements. Authors in [165, 166] have developed a coarse-to- ne algorithm which alternates between temporal motion estimation and spatial frame reconstruction in the wavelet-domain. Mun and Fowler [167] have used optical ow to estimate the motion field of the video frames, and the estimation is performed alongside reconstruction of the video frames. Cossalter et al. [168] have proposed a joint compressive video coding and analysis scheme for object tracking application. In this work, the authors have proposed to combine the process of video frame acquisition, compression and analysis and demonstrate the coding gain as compared to various multi-stage synthesis approaches. Recently, authors in [169], have suggested a compressive coded video compression scheme using motion estimation in the measurement domain using circulant Compressive Sensing Matrices [170]. Other popular algorithms for video CS have been based on total variation (TV) [171] and dictionary learning [172]. Total Variation (TV) methods assume that the gradient of each video frame is sparse and attempts to minimize the $l_p$norm of the gradient frames summed over all time steps. Dictionary-based methods represent each video patch as a sparse linear expansion in the dictionary elements. The dictionary is often learned offline from training video and the sparse coefficients of a video patch can be achieved by sparse-coding algorithms. The Gaussian Mixture Model (GMM) [173, 174] is another popular approach being used in various learning tasks, such as classification & segmentation [175, 176] and image denoising, inpainting and deblurring [177]. Most of the results presented in the literature are based on synthetic sequences and can be extended for real video sequence with reduced CS complexity.

2.3 Discussion

Table. 1 provides a comparative view of a selected texture analysis and synthesis based coding scheme proposed in the literature. We have done the comparative study of the various fundamental research and further evolutions in the Texture analysis and Synthesis based image and video compression schemes. The attributes compared are based on data analysis and synthesis model, compression results, objective and subjective quality of reconstruction and the reference data used for the reconstruction. The block-based parametric video compression (such as AR [130] and ARMA [132]) can provide high compression with medium to low computational cost, however they are not suitable for modelling of structural texture. Region based parametric model [137] can synthesize wide variety of texture including structural texture pattern; however they are not amenable to standard video compression framework like H.264/AVC and HEVC. Also region-based modelling with non-parametric synthesis framework is computationally complex and requires more storage capacity. Rate-distortion optimization is still an open issue with region-based video coding and subjective quality assessment technique such as HVS is not fully understood so far. A region based compression approach based on rate-distortion study and inter wavelet sub- band correlation has been proposed claiming better compression results as compared to 2D-SPIHT [178]. Advanced signal processing tools like SMW and DEMD provide very efficient representation of the texture data and can be effectively used for texture video compression. Statistically Matched Wavelet based texture representation results in most of the captured energy to be concentrated in the approximation sub-space, while very little information is remains in the detail sub-space [179].

3 New Context of Compression Problem

In this section we present the new challenges in compression of real life video applications such as HDR, Cloud Storage, 3D TV and discuss about the relevance of Parametric Compression methods in this context.

3.1 Ultra High Definition TV (UHD TV)

The UHD TV provides Higher Spatial Resolutions (HSR, hereafter) of 4K and 8K, Higher Temporal Resolutions (HTR, hereafter) of 100/120fps, High Dynamic Range (HDR, hereafter), and 10-bit Wide Colour Gamut (WCG, hereafter). ITU-R has recently announced a new recommendation on UHDTV in collaboration with experts from television industry, broadcasting organizations and regulatory institutions in its Study Group 6 [10]. High Temporal resolution increase the required bit rate due to more pictures to be encoded and more details in moving scenes and therefore become more challenging to encode temporal content with dynamic complex motion such as rotation and zoom. HDR [181] contains more information to encode as it encodes all the luminance level and dynamic range of $15\log_{10}$ (while standard image and video format do not exceed the dynamic range of $3\log_{10}$). Existing standard image and video compression schemes such as H.264 and HEVC generally supports only Low Dynamic Range (LDR) format and can encode only integer numbers. The transmission of HDR content requires a framework which can support beyond LDR (Low Dynamic Range) being supported by the conventional methods. Luminance encoding of HDR pixels should take into account the limitations of the human visual system and the fact that the human eye can perceive only limited numbers of luminance level and colours. This issue can be addressed using texture analysis and synthesis based compression framework using efficient representation and encoding of the huge texture data generated from the HDR content (Details of such techniques are discussed in Sec. 2).

3.2 3D Video

3D Video content is available in games, home entertainment and surveillance applications and further increasing to many more multimedia applications. The 3D data are primarily driven by the number of cameras and may limit the deployment especially when receivers have limited bandwidth and demands strong compression with perceptually acceptable reconstruction quality. The 3D video standards include multiview video plus depth (MVD) [182] which can allow the efficient coding of the multiview video. An extension of AVC, which is referred as MVC [183] supports coding of two or more views and improves coding efficiency through exploitation of interview redundancy, however it still produces a bit rate according to number of coded views and produces large amounts of data to be coded limiting to deploy 3D based-applications for the end-user. The main challenges in Multi-view Video Coding (MVC) is to enable a large number of views at the decoder side from a limited number of information coded at encoder and to adapt the bit-rate to available resources [184].

The MVD data format enables more efficient compression techniques for coding of multiview texture video data in a parametric compression framework using automatic view synthesis [185, 186]. Stefanoski et al. [185] have proposed to extract warp information based on image-domain-warping and use it to produce the rendered views in a parametric image warping framework. Image-domain-warping relies on an automatic estimation of sparse disparities and image saliency information. Similarly, Plath et al. [186] propose an optimized and computationally efficient adaptive image warping scheme that prevents holes during the view synthesis using depth map pre-processing. An alternative learning-based approach to render a 3D image from a 2D image is presented in the paper [187] based on learning a point mapping from local image attributes to scene-depth. Authors in [188, 189] propose a joint texture and depth data coding scheme using the motion vector of a texture picture as a prediction for co-located motion vector of its corresponding depth picture. Ndjiki-Nya et al. [190] have proposed a depth image-based rendering (DIBR) approach with texture synthesis from an MVD representation for 3D video compression application. In this work the foreground and background are warped separately and virtual views are constructed using depth map (DM). The other methods presented in the literature for 3D video representation is that instead of encoding and transmitting the texture and depth image pairs of different viewpoints to the decoder, encode the captured texture and depth pixels as a triangular mesh and depth based texture coding method [191, 192, 193] from the corresponding texture images. The depth map has a strong correlation with its associated texture data, so it can be utilized to improve texture coding efficiency.

3.3 Cloud Storage

According to a study [194] by 2018, global IP traffic will reach 1.6 zettabytes per year. Therefore, an important research issue is how to process such large amounts of multimedia content which is stored in and crossed over the cloud. Cloud Storage [195, 196] enables a new paradigm for data management by providing remote storage that can effectively process the multimedia applications. Millions of photo image frames are being stored over cloud per day. Individual images are generally stored using a standard image compression scheme such as JPEG2000, however it is not efficient as it does not exploit inter image redundancy. Li et al. [197] review and discuss various cloud storage techniques and they are evolving rapidly. Upadhyay et al. [198] have proposed an efficient cloud file compression scheme based on segmentation and de- duplication method. Kau et al. [199] propose an adaptive predictive coding scheme with predictor coefficients adapted by the least square optimization. The prediction errors are then encoded using a conditional arithmetic coder. The authors in [200] have proposed to organise the images as a pseudo video sequence and compress the sequence like a video using standard transformation and block-based motion compensation approach. These techniques however do not consider the fact that pseudo sequence is quite different from natural videos in terms of inter-image disparity. The images may have been captured at different locations, viewpoints and focal lengths making it very di cult to model them using global transformation (especially when they have been captured at different times with varying illumination and shadow). In addition MSE based prediction criterion will not be suitable for such conditions and therefore pixel based standard image and video compression techniques (such as JPEG2000, H.264/HEVC) does not solve the efficient image storage issues over the cloud. Hou et. al. [201, 202], have recently proposed some improvements over the texture synthesis scheme targeting to cloud storage applications. New generation feature based parametric compression framework can be a potential alternative for efficiently compressing and storing the images for cloud computing application.

4 General Discussion and Analysis

We have discussed in detail about Texture based video compression in Section 2. In this section we have reviewed the literature related to texture analysis (Sec. 2.1) and texture synthesis (Sec. 2.2) based compression methods and also introduced two advanced signal processing tools used in an efficient texture representation, i.e. Matched Wavelet (Sec. 2.1.1) and Empirical Mode Decomposition (Sec. 2.1.1). As a part of the texture synthesis framework we have discussed Compressive Sensing based signal synthesis (Sec. 2.2.1) which has gained special attention in the recent years. We have presented the new challenges in the compression and relevance of PVC in this context in Section 3. Based on the survey, we observe following open issues which may attract attention in the research community in coming years.

- In the conventional video compression methods such as H.264/AVC and HEVC, data acquisition and encoding are carried out on the entire signal, while most of the transformed data is discarded in the compression processes and thereby significantly wasting storage resources and high computational cost (This is more relevant in the case of texture video because of its ne detail and high frequency component). Conventional block-based transform used for exploiting the spatial redundancy does not work well for texture data due to its ne detail and high frequency components. Moreover the conventional block-based translation motion model used for temporal redundancy exploitation is not suited for texture with dynamic motion. This is true especially for highly textured surface with camera motion where block-based model produces a large residual signal. In order to address texture content, a content adaptive encoder with perceptual reconstruction at decoder has been required.

- Statistical model based texture synthesis [95, 126] generally fails for structural texture patterns. DEMD based texture synthesis approach presented in [117] is unable to handle the texture sequence with irregular patterns. AR- and ARMA based approaches [130, 132] are not suitable for structured texture with non-stationary data as blocks with non-stationary data are not amenable to AR modelling. Block-based parametric approaches can also create blocking artefacts in the synthesized texture. Region-based parametric approach on the other hand can synthesize a variety of textures however they are computationally complex and segmentation stage [203] needs to be absolutely accurate else there can be a propagation error. Region-based parametric approach cannot be seamlessly integrated with-in the state-of-the-art standard video codecs such as H.264/AVC and HEVC. A hybrid-approach combining the advantage of block-based representation and region-based synthesis may address the above described issues.

- Most of the new generation video compression methods make the use of new generation computer vision tools and techniques for data processing and feature extraction, however they are still based on Shannon-Nyquist sampling, thus do not fully exploit the inherent sparsity present in the signal. The texture analysis algorithm presented in the literature does not fully exploit the perceptual redundancy present in the video data, which could result in much higher compression at an acceptable psycho-visual quality. This leads to a requirement of an efficient data representation prior to encoding. Application of an advanced signal processing tools such as SMW and DEMD for a 2-D signal representation and feature extraction generating lot of interest in the research community.

- In the example based texture synthesis scheme, the quality of the synthesized texture is directly linked to the content and size of the example texture and prone to the propagation error. Therefore a direct relationship between the synthesized texture quality and bit-rate control is difficult. For the texture synthesis scheme based on side information, efficient encoding of the side information is still an open issue and directly linked to the decoder complexity.

- The conventional motion estimation techniques are not efficient in the prediction of complex motion regions involving rotation and zoom and thus limiting the overall compression which can be achieved. HEVC [4] and H.264 [3] approximates such complex motion by dividing bigger blocks and using multiple reference frame, but the corresponding motion vectors, block mode signalling and RD optimization are substantial overheads which affect the overall coding gain. A ne motion model has been proposed to address the issue, however it has its own overhead and suitable for a specific sequence involving prominent motion across frames.

- Only parametric model based video compression system (despite its abilities to achieve higher compression) lacking generalization and robustness against the scene change of the traditional signal-based video coding technique (for instance H.264 and HEVC) and therefore provides lots of interest in the research community to combine the bene t of the conventional video compression scheme with that of new generation parametric or non-parametric based schemes. Rate-distortion measurement is another open area and the key question is how to perform a joint optimization for both perceptually relevant and irrelevant regions in a model-based parametric video coding.

- Finally a unified test framework is required for evaluating the performance of any proposed algorithm across a broad range of content types. The most important attributes to be evaluated for video coding includes assessment of the complexity vs rate-distortion, coding gain vs quality etc. As a result a universal platform for objective comparison of all parametric video coding algorithms is required. A wide tests database with various texture video sequences are required in order to enhance and measure the proposed methods more accurately so that these new approaches can be considered as part of future video coding standards.

4.1 Our Contributions, In a Nutshell

Motivated by these observations, we have proposed an application of Statistically Matched Wavelet for efficient representation of the texture data and an efficient texture synthesis in a compressive sensing framework [179]. In this work, We propose to encode not the full-resolution statistically matched wavelet sub-band coefficients (as normally done in a standard wavelet based image compression) but only the approximation sub-band coefficients (LL) using a standard image compression scheme like JPEG2000(which accounts for 1/4th of the total coefficients and can be represented using fewer bits). The detail sub-band coefficients, that is, HL, LH, and HH (which account for 3/4th of the total coefficients, are jointly encoded in a compressive sensing framework and can therefore be represented with fewer measurements. The experimental results prove that the proposed algorithm can handle a variety of texture synthesis such as periodic, aperiodic and structural pattern with significant compression gain. Our other work suggested a Directional Empirical Mode Decomposition based Video Image frame representation and synthesis [204, 205] in a compressive sensing framework. We have extended this method to a video application and proposed an integrated video compression framework based on DEMD [180] which results in significant compression gain (up to 50%) as compared to H.264/AVC for similar perceptual quality. Figure 5, shows that the compression in the proposed framework is better than a standard MPEG2/H.264/HEVC encoding. Our proposed scheme is a hybrid approach, combining the advantages of parametric and non-parametric methods, which enables us to handle a wide variety of textured videos which cannot be accounted for by either method alone. DEMD can handle textures with stationary and non-stationary data. The scheme is in an H.264-/MPEG-based framework, for better compatibility with existing video coding standards. We take advantage of the temporal encoding efficiency of standard video coding frameworks (H.264/MPEG) to encode the DEMD residues frames.

In our recent publication [64], we have demonstrated the capability of the higher-order affine based motion model in a multi-resolution framework. Fig. 6 shows a representative example: a comparison of the RD plot for the 'Park Scene (1080P)' and 'Cactus (1080P)' for the proposed method (HM_{Affine}), and that of a pure HEVC encoding (HM). The proposed method shows a trend of higher PSNR values at smaller bit rates, an indication of better coding efficiency with higher fidelity to the original unencoded video.

5.0 General Discussion and Analysis

We have presented an exhaustive review of the new generation Parametric Video Compression schemes. We have discussed a range of techniques including Model Based Video Compression, Mosaic Based Video Compression and Texture Based Video Compression methods. The scope of each method and their compression performance is presented and discussed in an end to end integrated coding framework. It has been observed that significant bit rate savings (in some cases >50%) can be achieved using these new generation Parametric Video Compression schemes as compared to standard codecs for similar visual quality. It is also observed that all such schemes can be very effective for applications where perceptually correct reconstruction with high compression is preferred to bit accurate reconstruction and Nyquist based sampling (with lower compression). We have also discussed the current challenges in the area of compression and relevance of Parametric Compression in this context. A number of open issues and limitation have been analysed and discussed. In particular efficient representation and processing of input signal, exploitation of sparsity and perceptual redundancy present in the signal, and integration of these schemes with the conventional H.264/AVC framework have been analysed. The analysis and discussion bring out the relevance of PVC in the current day compression research.

References

[1] ISO/IEC. Information technology coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/sec. Part2 : Video, 11172-2 standard, 1993.

[2] ITU-T and ISO/IEC JTC 1. Generic Coding of Moving Pictures and Associated Audio Information. ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG-2 Video), version 1, 1994.

[3] ITU-T and ISO/IEC JTC 1. Advanced Video Coding for Generic Audiovisual Services. ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), version 1, 2003, version 2, 2004, versions 3, 4, 2005, versions 5, 6, 2006, versions 7, 8, 2007, versions 9, 10, 11, 2009, versions 12, 13, 2010, versions 14, 15, 2011, version 16, 2012.

[4] B. Bross, W.J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand. JCTVC-H1003: High Efficiency Video Coding (HEVC) text specification draft 9. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, document JCTVC-K1003, Shanghai, China, 2012.

[5] H. R. Wu, A. Reibman, W. Lin, F. Pereira, and Hemami. Perceptual visual signal compression and transmission. Proceedings of the IEEE, pages 2025 - 2043, 2013.

[6] C. Singh, D. Rawat, and S. Meher. A hybrid image compression scheme using HVS characteristics: combining SPIHT- and SOFM-based vector quantisation. Int. J. of Signal and Imaging Systems Engineering, pages 175 - 186, 2012.

[7] A. S. Dias, S. Schwarz, M. Siekmann, S. Bosse, Schwarz, D. Marpe, J. Zubrzycki, and M. Mrak. Perceptually Optimized Video Compression. IEEE International Conference on Multimedia and Expo Workshops, pages 1 - 4, 2015

[8] K. Lee and S. Lee. A new framework for measuring 2d and 3d visual information in terms of entropy. IEEE Transactions on Circuits and Systems for Video Technology, pages 1 - 13, 2015.

[9] A. Schodl, R. Szeliski, D. Salesin, and I. Essa. Video textures. in Proc. ACM SIGGRAPH, pages 489 - 498, 2000.

[10] ITU-R BT. Draft New Recommendation Parameter Values for UHDTV Systems for Production and International Programme Exchange. International Telecommunications Union, May 2012.

[11] F. Galpin, R. Balter, L. Morin, and K. Deguchi. 3d models coding and morphing for efficient video compression. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 334 - 341, 2004.

[12] R. Balter, P. Gioia, and L. Morin. Scalable and efficient coding using 3d modelling. IEEE Transactions on Multimedia, pages 1147 - 1155, 2006.

[13] A. Hakeem, K. Sha que, and M. Shah. An Object-based Video Coding Framework for Video., International Conference on multimedia, pages 608 - 617, 2005.

[14] R. V. Babu and A. Makur. Object based surveillance video compression using foreground motion compensation. In : 9th International Conference on Control, Automation, Robotics and Vision, pages 1 - 6, 2006.

[15] S. Chaudhury, M. Mathur, B. Lall, A. Khandelia, S. Tripathi, S. Dutta Roy, and S. Gorecha. System and method for object based parametric video coding. US20110058609 A1, pages 1 - 4, 2009.

[16] M. Alwani, R. Chaudhary, M. Mathur, Dutta Roy, and S. Chaudhury. Restricted A ne Motion Compensation in Video Coding using Particle Filtering. In Proc. IAPR-

sponsored Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), pages 479 - 484, 2010.

[17] Z. Zhu, H. Tang, and G. Wolberg. Content-Based 3D Mosaic Representation for Video of Dynamic 3D Scenes. IEEE, Applied Imagery and Pattern Recognition Workshop, 2005.

[18] W. Jiang and J. Gu. Video Stitching with Spatial-Temporal Content-PreservingWarping. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 42 - 48, 2015.

[19] P. Ndjiki-Nya, D. Doshkov, H. Kaprykowsky, Zhang, D. Bull, and T. Wiegand. Perception-oriented video coding based on image analysis and completion: A review. Signal Processing: Image Communication, pages 579 - 594, 2012

[20] C. Hoffmann, S. Argyropoulos, A. Raake, and P. Ndjiki-Nya. Modelling image completion distortions in texture analysis-synthesis coding. Picture Coding Symposium (PCS), pages 293 - 296, 2013.

[21] Y. Guo, G. Zhao, Z. Zhou, and M. Pietikainen. Video Texture Synthesis with Multi-frame LBP-TOP and Diffeomorphic Growth Model. IEEE Transactions on Image Processing, pages 3879 - 3891, 2013.

[22] C. Yeh, T. Tseng, C. Lee, and Chih-Yang Lin. Predictive Texture Synthesis-Based Intra Coding Scheme for Advanced Video Coding. IEEE Transactions on Multimedia, pages 1508 - 1514, 2015.

[23] M. Magnor and B. Girod. Model-based coding of multi viewpoint imagery. in SPIE Conference Proceedings Visual Communications and Image Processing, pages 14 - 22, 2000.

[24] R. Koch, M. Pollefeys, and L. V. Gool. Realistic surface reconstruction of 3d scenes from uncalibrated image sequence. in Journal of Visualization and Computer Animation, pages 115 - 127, 2000.

[26] S. Pateux, G. Marquant, and D. Chavira-Martinez. Object mosaicking via meshes and crack-lines technique. Application to low bit-rate video coding. in Proc. PCS, pages 417 - 420, 2001.

[27] N. Snavely, S. M. Seitz, and R. Szeliski. Modelling the world from internet photo collections. pages 189 - 210, 2008.

[28] D. Liu, T. Shao, H. Li, and F. Wu. Three-dimensional point-cloud plus patches: Towards model-based image coding in the cloud. IEEE International Conference on Multimedia Big Data, pages 396 - 400, 2015.

[29] T. Tung, F. Schmitt, and T . Matsuyama. Topology matching for 3D Video Compression. IEEE International Conference on Computer Vision and Pattern Recognition, pages 1 - 8, 2007.

[30] T. Tung and F. Schmitt. The augmented multiresolution reeb graph approach for content based retrieval of 3D shapes. In international Journal of Shape Modeling, pages 91 - 120, 2005.

[31] R. Pajarola and J. Rossignac. Compressed progressive meshes. IEEE Trans. Vis. Comp. Graph, pages 79 - 93, 2000.

[32] M. Isenburg and P. Alliez. Compressing polygon mesh geometry with parallelogram prediction. in Proc. IEEE Visualization, pages 141 - 146, 2002.

[33] U. Bayazit, O. Orcay, U. Konur, and F. S. Gurgen. Predictive vector quantization of 3-D polygonal mesh geometry by representation of vertices in local coordinate systems. in J. Vis. Comm. Image Rep, pages 341 - 353, 2007.

[34] Z. Chen, F. Bao, Z. Fang, and Z. Li. Compression of 3D Triangle Meshes with a Generalized Parallelogram Prediction Scheme Based on Vector Quantization. Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pages 184 - 187, 2010.

[35] M. O. Bici and G. B. Akar. Improved prediction methods for scalable predictive animated mesh compression. J. Vis. Commun. Image Rep., pages 577 - 589, 2011.

[36] J. Ahn, Y. J. Koh, and C. Kim. Efficient Fine-Granular Scalable Coding of 3D Mesh Sequences. IEEE Transactions on Multimedia, pages 485 -497, 2013.

[37] J. Ahn, K. Lee, J. Sim, and C. Kim. Large-scale 3d point cloud compression using adaptive radial distance prediction in hybrid coordinate domains. IEEE Journal of selected topics in signal processing, pages 422 - 434, 2015.

[38] M. Alexa and W. Mller. Representing animations by principal components. Computer Graphics Forum, pages 411 - 418, 2000.

[39] Z. Karni and C. Gotsman. Compression of soft-body animation sequences. Computer Graphics, pages 25 - 34, 2004.

[40] U. Bayazit, U. Konur, and H. F. Ates. 3-D Mesh Geometry Compression with Set Partitioning in the Spectral Domain. IEEE Transactions on Circuits and Systems for Video Technology, pages 179 - 188, 2010.

[41] E. Lee and H. Ko. Vertex data compression for triangular meshes. in Proc. Pacific Graphics, pages 225 - 234, 2000.

[42] Jae-Kyun Ahn, Yeong Jun Koh, and Chang-Su Kim. Motion-compensated compression of 3D animation models. Electronics Letter,, pages 1445 - 1446, 2001.

[43] A. Khodakovsky and I. Guskov. Compression of normal meshes. Geometric Modelling for Scientific Visualization, pages 189 - 206, 2004.

[44] K. Mamou, T. Zaharia, and F. Prteux. A skinning approach for dynamic 3D mesh compression. Comput. Animat. Virtual Worlds, pages 337 - 346, 2006.

[45] Chun-Ming Li, Yu-Shan Li, Qing-De Zhuang, Qiu-Ming Li, Rui-Hong Wu, and Yang Li. Applying mid-level vision techniques for video data compression and manipulation. In Proc. SPIE on digital video compression and Personal Computers: Algorithm and Technologies, pages 116 - 127, 2004.

[46] S. Chaudhury, S. Tripathi, and S. D. Roy. Parametric Video compression using appearance space. International Conference on Pattern Recognition, pages 1 - 4, 2008.

[47] M. B. Haj Ayech and H. Amiri. A content-based image retrieval using PCA and SOM. Int. J. of Signal and Imaging Systems Engineering, pages 276 - 282, 2016.

[48] G. Harit and S. Chaudhury. Video Shot Characterization Using Principles of Perceptual Prominence and Perceptual Grouping in Spatio-Temporal Domain. IEEE Transactions on Circuits and Systems for Video Technology, 17:1728 - 1741, 2007.

[49] Chun-Ming Li, Yu-Shan Li, Qing-De Zhuang, Qiu-Ming Li, Rui-Hong Wu, and Yang Li. Moving Object Segmentation and Tracking In Video. in Proc. of the Fourth Intl Conf. on Machine Learning and Cybernetics, pages 4957 - 4960, 2005.

[50] S. Kamp. Video coding using decoder-side motion vector derivation. In RWTH Aachen University, Germany, Tech. (Online), 2008.

[51] A. Gahlot, S. Arya, and D. Ghosh. Object-based affine motion estimation. In Proc. IEEE Region 10 Conference, pages 1343 - 1347, 2003.

[52] T. Wiegand, E. Steinbach, and B. Girod. Affine multi-picture motion-compensated prediction .IEEE Transactions on Circuits and Systems for Video Technology, 15(2):197 - 209, 2005.

[53] R. C. Kordasiewicz, M. D. Gallant, and S. Shirani. Affine motion prediction based on translational motion vectors. IEEE Transactions on Circuits and Systems for Video Technology, 17(11):1388 - 1394, 2007.

[54] K.-L. Chung and T.-J. Yao. New Prediction and A ne Transformation - Based Three-Step Search Scheme for Motion Estimation With Application. Journal of Information Science and Engineering, 24:1095 - 1109, 2008.

[55] H.-K. Cheung and W.-C. Siu. Local Affine Motion Prediction for H.264 without Extra Overhead. In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1555 - 1558, 2010.

[56] H.-K. Cheung and W.-C. Siu. Local a ne motion prediction for h.264 without extra overhead. In in Proc. IEEE Int. Symp. Circuits Syst., pages 1555 - 1558, 2010.

[57] H. Yuan, J. Liu, J. Sun, H. Liu, and Y. Li. Affine model based motion compensation prediction for zoom. IEEE Transactions on Multimedia, 14(4):1370 - 1375, 2012.

[58] P. Anandan. A Computational Framework and an Algorithm for the Measurement of Visual Motion. International Journal of Computer Vision, pages 283 - 310, 1989.

[59] J. H. Lee, K. W. Lim, B. C. Song, and J. B. Ra. A Fast Multi-Resolution Block Matching Algorithm and its LSI Architecture for Low Bit-Rate Video Coding. IEEE Transactions on Circuits and Systems for Video Technology, 11(12):1289 - 1301, 2001.

[60] J. H. Lee, K. W. Lim, B. C. Song, and J. B. Ra. A Fast Multi-Resolution Block Matching Algorithm and its LSI Architecture for Low Bit-Rate Video Coding. IEEE Transactions on Circuits and Systems for Video Technology, 11:1289 - 1301, 2001.

[61] J. Sullivan and J. Rittscher. Guiding Random Particles by Deterministic Search. In Proc. IEEE International Conference on Computer Vision (ICCV), pages 1 - 18, 2001.

[62] T. Ouni, W. Ayedi, and M. Abid. A complete non-predictive video compression scheme based on a 3D to 2D geometric transform. Int. J. of Signal and Imaging Systems Engineering, pages 164 - 180, 2011.

[63] G. Hegde and P. Vaya. Systolic array based motion estimation architecture of 3D DWT sub band component for video processing. Int. J. of Signal and Imaging Systems Engineering, pages 158 - 166, 2012.

[64] M. K. Jha, R. Chaudhary, S. Dutta Roy, M. Mathur, and B. Lall. Restricted affine motion compensation and estimation in video coding with particle filtering and importance sampling: a multi-resolution approach. Multimedia Systems, pages 1 - 14, 2017.

[65] J. Wang and E. Adelson. Representing moving images with layers. IEEE Trans. On Image Processing, pages 625 - 638, 1994.

[66] H.-Y. Shum Y. Li, C.-K. Tang, and R. Szeliski. Stereo reconstruction from multiperspective panoramas. IEEE Trans. on PAMI, pages 45 - 62, 2004

[67] C. Sun and S. Peleg. Fast Panoramic Stereo Matching using Cylindrical Maximum Surfaces. IEEE Trans. SMC Part B, pages 760 - 775, 2004.

[68] H. Tang and Z. Zhu. Content-Based 3D Mosaic Representation for Video of Dynamic Urban Scenes. IEEE Transactions on Circuits and Systems for Video Technology, pages 295 - 308, 2012.

[69] Z. Zhu, E. M. Riseman, and A. R. Hanson. Generalized parallel perspective stereo mosaics from airborne videos. IEEE Trans. on Pattern Analysis and Machine Intelligence, pages 226 - 237, 2004.

[70] P. Ndjiki-Nya, C. Stuber, and T. Wiegand. Improved video coding through texture analysis and synthesis. Proceedings of the 5th International Workshop on Image Analysis for Multimedia Interactive Services, 2004.

[71] F. Zhang and D. Bull. A parametric framework for video compression using region-based texture models. IEEE Journal on Selected Topics in Signal Processing, pages 1378 - 1392, 2011.

[72] C. Roberto, S. Luciano, and S. Sabine. Higher order svd analysis for dynamic texture synthesis. IEEE Trans. on Image Processing, 17:42 - 52, 2008.

[73] R. Alizarraga-Morales, Y. Guo, G. Zhao, M. Pietikinen, and Raul E.Sanchez Yanez. Local spatio temporal features for dynamic texture synthesis. EURASIP Journal on Image and Video Processing, pages 1 - 15, 2014.

[74] A. Criminisi, P. Perez, and K. Toyama. Region Filling and Object Removal by Exemplar-Based Image Inpainting. IEEE Transactions on Image Processing, pages 1200 - 1212, 2004.

[75] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum. Image Completion with Structure Propagation. ACM SIGGRAPH, pages 861 - 868, 2005.

[76] K. Patwardhan, G. Sapiro, and M. Bertalmio. Video in-painting under constrained camera motion. IEEE Transaction on image processing, pages 545 - 553, 2007.

[77] T. K. Shih, N. C. Tang, and J. N. Hwang. Exemplar-based video inpainting without ghost shadow artefacts by maintaining temporal continuity. IEEE Transactions on circuits and systems for video technology, pages 347 - 360, 2009.

[78] C. Ling and N. C. Cheng. Video background in-painting using dynamic texture synthesis. IEEE Conference on Circuits and Systems, pages 1559 - 1562, 2010.

[79] J. Herling and W. Broll. High-Quality Real-Time Video Inpainting with PixMix. IEEE Transactions on visualization and computer graphics, pages 866 - 879, 2014.

[80] V. Cheung, B. J. Frey, and N. Jojic. Video epitomes. Computer Vision and Pattern Recognition (CVPR), pages 42 - 49, 2005.

[81] S. Bansal, S. Chaudhury, and B. Lall. Parametric video compression using epitome based texture synthesis. Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), pages 94 - 97, 2011.

[82] S. Lazebnik, C. Schmid, and J. Ponce. sparse texture representation using local affine regions. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), pages 1265 - 1278, 2005.

[83] I. Kokkinos, G. Evangelopoulos, and P. Maragos. Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), pages 142 - 157, 2009.

[84] G. Georgiadis, A. Chiuso, and S. Soatto. Texture representations for image and video synthesis. IEEE Conference on computer vision and pattern recognition (CVPR), pages 2058 - 2066, 2015.

[85] V. Kwatra, A. Schodl, I. Essa, G. Turk, and Bobick. Graphcut textures: Image and video synthesis using graph cuts. in Proc. ACM SIGGRAPH, pages 277 - 286, 2003.

[86] V. Kwatra, I. Essa, A. Bobick, and N. Kwatra. Texture optimization for example-based synthesis. in Proc. ACM SIGGRAPH, pages 795 - 802, 2005.

[87] L. Wei, J. Han, K. Zhou, H. Bao, B. Guo, and Shum. Inverse texture synthesis. in Proc. ACM Transactions of Graphics(TOG), 2008.

[88] H. Wang, Y. Wexler, E. Ofek, and H. Hoppe. Factoring repeated content within and among images. ACM Transactions on Graphics (TOG), pages 1 - 10, 2008.

[89] J. Balle, A. Stojanovic, and J. R. Ohm. Models for static and dynamic textures synthesis in image and video compression. IEEE Journal on Selected Topics in Signal Processing, pages 1353 - 1365, 2011.

[90] I. Elfadel and R. Picard. Gibbs random fields, co-occurrences, and texture modelling. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), pages 24 - 37, 1994

[91] M. L. Comer and E. J. Delp. Segmentation of textured images using a multiresolution gaussian autoregressive model. IEEE Transactions on Image Processing, pages 408 - 420, 1999.

[92] E. J. Delp, R. L. Kashyap, and O. Mitchell. Image data compression using autoregressive time series models. Pattern Recognition, pages 313 - 323, 1979.

[93] J. R. Smith and S.-F. Chang. Quad-tree segmentation for texture-based image query.Proceedings of the second ACM International Conference on Multimedia, pages 279 - 286, 1994.

[94] J. Portilla and E. P. Simoncelli. Texture modelling and synthesis using joint statistics of complex wavelet coefficients. in Proc. IEEE Workshop Statist. Comput. Theories Vision, pages 1 - 32, 1999.

[95] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. Int. J. Comput. Vision, pages 49 - 71, 2000.

[96] R. M. Haralick. Statistical and structural approaches to texture. Proceedings of the IEEE, pages 786 - 804, 1979.

[97] C.-H. Peh and L.-F. Cheong. Synergizing spatial and temporal texture. IEEE Transactions on Image Processing, pages 1179 - 1191, 2002.

[98] F. Zhu, K. Ng, G. Abdollahian, and E. J. Delp. Spatial and temporal models for texture-based video coding. Proceedings of the Visual Communications and Image Processing Conference, 2007.

[99] K. Chang, K. Bowyer, and M. Sivagurunath. Evaluation of texture segmentation algorithms. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 249 - 299, 1999.

[100] V. Ramos and F. Muge. Image colour segmentation by genetic algorithms. ArXiv Computer Science e-prints, 2004.

[101] A. Gupta, S. D. Joshi, and S. Prasad. On a new approach for estimating wavelet matched to signal. in Proc. Eighth National Conf. Commun., Bombay, India, pages 180 - 184, 2002.

[102] A. Tewfik, D. Sinha, and P. Jorgensen. On the optimal choice of a wavelet for signal representation. IEEE Trans. Inf. Theory, pages 747 - 766, 1992.

[103] J. O. Chapa and R. M. Rao. Algorithm for designing wavelets to match a specified signal. IEEE Transactions on Signal Processing, pages 3395 - 3406, 2000.

[104] W.-S. Lu and A. Antoniou. Design of signal-adapted biorthogonal filter banks. IEEE Trans. Circuits Syst. I, Funam. Theory Appl, pages 90 - 102, 2001.

[105] M. K. Tsatsanis and G. B. Giannakis. Principal component filter banks for optimal multiresolution analysis. IEEE Transactions on Signal Processing, pages 1766 - 1777, 1995.

[106] A. Aldroubi and M. Unser. Families of multiresolution and wavelet spaces with optimal properties. Numer. Func. Anal., pages 417 - 446, 1993.

[107] A. Gupta, S. D. Joshi, and S. Prasad. A new approach for estimation of statistically Matched Wavelet. IEEE Transactions on Signal Processing, pages 1778 - 1793, 2005.

[108] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi. Text Extraction and Document Image Segmentation using Matched Wavelets and MRF Model. IEEE Transactions on Image Processing, pages 2117 - 2128, 2007.

[109] A. Linderhed. Empirical Mode Decomposition as Data-Driven Wavelet-Like Expansions. PhD thesis, Dissertation No. 909, Linkoping University, 2004.

[110] P. Flandrin and P. Goncalves. Adaptive Image Compression with Wavelet Packets and Empirical Mode Decomposition. International Journal of Wavelets, Multiresolution and Information Processing, pages 1 - 20, 2004.

[111] N. E. Huang, Z. Shen, and S. R. Long. The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of Royal Society of London (Series A), 1:903 - 995, 1998.

[112] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel. Image analysis by bidimensional empirical mode decomposition. Image and Vision Computing, 21:1019 - 1026, 2003.

[113] A. Linderhed. 2d empirical mode decompositions in the spirit of image compression. in Wavelet and Independent Component Analysis Applications IX, SPIE Proceedings, 4738:1 - 8, 2002.

[114] Z. Liu and S. Peng. Directional Empirical Mode Decomposition and its application to Texture Segmentation. Science in China Series E: Information Sciences, 2:113 - 123, 2005.

[115] Z. Liu, H. Wang, and S. Peng. Texture classification through Directional Empirical Mode Decomposition. Proc. International Conference on Pattern Recognition (ICPR), 4:803 - 806, 2004.

[116] Z. Liu, H. Wang, and S. Peng. Texture Segmentation using Directional Empirical Mode Decomposition. Proc. IEEE International Conference on Image Processing (ICIP), 1:279 - 282, 2004.

[117] Y. Zhang, Z. Sun, and W. Li. Texture Synthesis based on Directional Empirical Mode Decomposition. Computer and Graphics, pages 175 - 186, 2008.

[118] JM. Francos, AZ. Meiri, and B. Porat. A unified texture model based on a 2-d wold-like decomposition.

IEEE Transactions on Signal Processing, 41:2665 - 2678, 1993.

[119] N. E. Huang, Z. Shen, and S. R. Long. A new view of nonlinear water waves: the hilbert spectrum. Annual Review of Fluid Mechanics, 31:417 - 457, 1999.

[120] N. E. Huang, M.C. Wu, S.R. Long, S.S.P. Shen, W. Qu, P. Gloersen, and K.L. Fan. A confidence limit for the empirical mode decomposition and hilbert spectral analysis. Proceedings of the Royal Society London A 459, 2037:2317 - 2345, 2003.

[121] G. Rilling, P. Flandrin, and P. Gonalvs. On empirical mode decomposition and its algorithms. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, 2003.

[122] N. E. Huang, M. L.Wu, W. Qu, S. R. Long, and S. S. P. Shen. Applications of hilberthuang transform to non-stationary financial time series analysis. Appl. Stochastic Models in Business Industry, 19:245 - 268, 2003.

[123] G. Doretto, A. Chiuso, Y. N. Wu, and S. Soatto. Dynamic textures. International Journal of Computer Vision, pages 91 - 109, 2003.

[124] R. Paget and I. D. Longsta . Texture synthesis via a noncausal nonparametric multiscale Markov random field. IEEE Trans. Image Process., pages 925 - 931, 1998.

[125] J. Zhang, D. Wang, and Q. N. Tran. A wavelet-based multiresolution statistical model for texture. IEEE Trans. Image Process., pages 1621 - 1627, 1998.

[126] G. Fan and X. Xia. Wavelet-based texture analysis and synthesis using hidden Markov models. IEEE Transactions on Circuits and Systems, 50:106 - 120, 2003.

[127] A. Efros and W. T. Freeman. Image quilting for texture synthesis and transfer. in Proc. Int. Conf. Comput. Graph. Interactive Tech, pages 341 - 346, 2001.

[128] A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. in Proc. Int. Conf. Comput. Vision, pages 1033 - 1038, 1999.

[129] L. Liang, C. Liu, Y. Xu, B. Guo, and H. Shum. Real-time texture synthesis by patch-based sampling. ACM Trans. Graph, pages 127 - 150, 2001.

[130] A. Khandelia, S. Gorecha, B. Lall, S. Chaudhury, and M. Mathur. Parametric Video Compression Scheme using AR based Texture Synthesis. Proc. IAPR- and ACM-sponsored Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), pages 219 - 225, 2008.

[131] A. Stojanovic, M. Wien, and J. R. Ohm. Dynamic texture synthesis for H.264/AVC inter coding. Proc. IEEE International Conference on Image Processing (ICIP), pages 1608 - 1612, 2008.

[132] A. Stojanovic, M.Wien, and T. K. Tan. Synthesis-in-the-loop for video texture coding. Proc. IEEE International Conference on Image Processing (ICIP), pages 2293 - 2296, 2009.

[133] H. Chen, R. Hu, D. Mao, R. Thong, and Z. Wang. Video coding using dynamic texture synthesis. International Conference on Multimedia and Expo, pages 203 - 208, 2010.

[134] A. Dumitras and B. G. Haskell. An encoder-decoder texture replacement method with application to content-based movie coding. IEEE Transactions on Circuits and Systems for Video Technology, 14:825 - 840, 2004.

[135] M. Bosch, F. Zhu, and E. J. Delp. Spatial texture models for video compression. in Proc. IEEE Int. Conf. Image Process., pages 93 - 96, 2007.

[136] M. Bosch, F. Zhu, and E. J. Delp. Perceptual Quality Evaluation for Texture and Motion based Video Coding. Proc. IEEE International Conference on Image Processing (ICIP), pages 2285 - 2288, 2009.

[137] F. Zhang and D. R. Bull. Enhanced video compression with region-based texture models. in Proceedings of the Picture Coding Symposium (PCS), pages 54 - 57, 2010.

[138] A. Khandelia and S. Chaudhury. Texture Mapping based Video Compression Scheme. IEEE Intl. Conf. On Image Processing (ICIP), pages 26 - 29, 2010.

[139] A. Dumitras and B. G. Haskell. A texture replacement method at the encoder for bit-rate reduction of compressed video. IEEE Trans. Circuits Syst. Video Technol., pages 163 - 175, 2003.

[140] P. Ndjiki-Nya, B. Makai, G. Blattermann, A. Smolic, and H. Schwarz. Improved H.264/AVC coding using texture analysis and synthesis. In Proc. IEEE Int. Conf. Image Process, pages 849 -852, 2003.

[141] C. Zhu, X. Sun, F. Wu, and H. Li. Video Coding with Spatio-Temporal Texture Synthesis and edge-based inpainting. IEEE International Conference on Multimedia and Expo, pages 813 - 816, 2008.

[142] B. T. Oh, Y. Su, A. Segall, and C.-C. J. Kuo. Synthesis-based texture coding for video compression with side information. Proc. IEEE International Conference on Image Processing (ICIP), pages 1628 - 1631, 2008.

[143] P. Ndjiki-Nya, C. Stuber, and T. Wiegand. Texture synthesis method for generic video sequences. in Proc. IEEE Int. Conf. Image Process., pages 397 - 400, 2007.

[144] P. Ndjiki-Nya, C. Stuber, and T. Wiegand. A new generic texture synthesis approach for enhanced

H.264/MPEG4-AVC video coding. in Proc. Very Low Bitrate Video, pages 121 - 128, 2006.

[145] R. OCallaghan and D. Bull. Combined morphological spectral unsupervised image segmentation. IEEE Transaction on Image Processing, 14:49 - 62, 2005.

[146] N. G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. Journal of Applied and Computational Harmonic Analysis, 10:210 - 234, 2001.

[147] J. S. DeBonet. Multiresolution sampling procedure for analysis and synthesis of texture images. in Proc. SIGGRAPH, pages 361 - 368, 1997.

[148] L. Y. Wei and M. Levoy. Fast Texture Synthesis using tree-structured vector quantization. in Proc. SIGGRAPH, pages 479 - 488, 2000.

[149] M. Ashikhmin. Synthesizing natural textures. in Proc. ACM Symposium on Interactive 3D Graphics, pages 217 - 226, 2001.

[150] T. K. Tan, C. S. Boon, and Y. Suzuki. Intra prediction by template matching. Proc. IEEE International Conference on Image Processing (ICIP), pages 1693 - 1696, 2006.

[151] Y. Guo, Y.-K. Wang, , and H. Li. Priority-based template matching intra prediction. in Proc. ICME, pages 1117 - 1120, 2008.

[152] S. Ierodiaconou, J. Byrne, D. R. Bull, D. Redmill, and P. Hill. Unsupervised image compression using graphcut texture synthesis. Proc. IEEE International Conference on Image Processing (ICIP), pages 2289 - 2292, 2009

[153] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transaction on information Theory, pages 489 - 509, 2006.

[154] D. Donoho. Compressed sensing. IEEE Transaction on information Theory, pages 1289 - 1306, 2006.

[155] R. G. Baraniuk. Compressed sensing lecture notes]. IEEE Signal Processing Magazine, pages 118 - 121, 2007.

[156] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. SIAM J. on Scientific Computing, 20:33 - 61, 1998.

[157] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. IEEE Transaction on information Theory, 50:2231 - 2242, 2004.

[158] E. Candes and M. Wakin. An introduction to compressive sampling. IEEE Signal Processing Magazine, 25:21 - 30, 2008.

[159] Y. Zhang, S. Mei, Q. Chen, and Z. Chen. A novel image/video coding method based on compressive sensing theory. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1361 - 1364, 2008.

[160] D. Venkatraman and A. Makur. A compressive sensing approach to object-based surveillance video coding. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 3513 - 3516, 2009.

[161] J. Prades-Nebot, Y. Ma, and T. Huang. Distributed video coding using compressive sampling. Picture Coding Symposium (PCS), pages 1 - 4, 2009.

[162] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk. Compressed imaging for video representation and coding. Picture Coding Symposium, pages 1 - 6, 2006.

[163] Y. Yang, O. C. Au, L. Fang, X. Wen, and W. Tang. Perceptual compressive sensing for image signals. IEEE International Conference on Multimedia and Expo, pages 89 - 92, 2009.

[164] A. C. Sankaranarayanan, P. K. Turaga, R. G. Baraniuk, and R. Chellappa. Compressive acquisition of dynamic scenes. European Conference on Computer Vision, pages 129 - 142, 2010.

[165] J. Y. Park and M. B. Wakin. A multiscale framework for compressive sensing of video. Picture Coding Symp.,Chicago II, May., 2009

[166] M. S. Asif, F. Fernandes, and J. Romberg. Low-complexity video compression and compressive sensing. Asilomar Conference on Signals, Systems and Computers, pages 579 - 583, 2013.

[167] S. Mun and J. E. Fowler. Residual reconstruction for block-based compressed sensing of video. IEEE Data Compression Conference, pages 183 - 192, 2011.

[168] M. Cossalter, G. Valenzise, M. Tagliasacchi, and S. Tubaro. Joint compressive video coding and analysis. IEEE Transaction on Multimedia, 12:168 - 183, 2010.

[169] S. Narayanan and A. Makur. Compressive coded video compression using measurement domain motion estimation. IEEE International Conference on Electronics, Computing and Communication Technologies, pages 1 -6, 2014.

[170] S. Narayanan and A. Makur.Camera motion estimation using circulant compressive sensing matrices. International Conference on Information, Communications and signal processing, pages 1 - 5, 2013.

[171] D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady. Multiframe image estimation for coded aperture snapshot spectral imagers. Applied Optics, 49:6824 - 6833, 2010.

[172] Y. Hitomi, M. Gupta J. Gu, T. Mitsunaga, and S. K. Nayar. Video from a single coded exposure photograph using a learned over-complete dictionary. IEEE International Conference on Computer Vision (ICCV), pages 287 - 294, 2011.

[173] J. Yang, X. Yuan, X. Liao, P. Llull, G. Sapiro,D. J. Brady, and L. Carin. Gaussian mixture model for video compressive sensing. Proc. IEEE International Conference on Image Processing (ICIP), pages 19 - 23, 2013.

[174] J. Yang, X. Yuan, X. Liao, P. Llull, D. J. Brady, G. Sapiro, and L. Carin. Video Compressive Sensing Using Gaussian Mixture Models. IEEE Transactions on Image Processing, pages 1 - 18, 2014.

[175] H. Permuter, J. Francos, and I. H. Jermyn. Gaussian mixture models of texture and colour for image database retrieval. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3:569 - 572, 2003.

[176] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. IEEE Conference on Computer Vision and Pattern Recognition, pages 246 - 252, 1999.

[177] G. Yu, G. Sapiro, and S. Mallat. Solving inverse problems with piecewise linear estimators: from gaussian mixture models to structured sparsity. IEEE Transactions on Image Processing, 21:2481 - 2499, 2012.

[178] S. Sasikumar N.M. Masoodhu Banu. Skip block-based wavelet compression for hyperspectral images. Int. J. of Signal and Imaging Systems Engineering, pages 156 - 164, 2016.

[179] M. K. Jha, B. Lall, and S. Dutta Roy. Statistically Matched Wavelet Based Texture Synthesis in a Compressive Sensing Framework. ISRN Signal Processing, Hindawi Publishing Corporation, 2014:1 - 18, 2014.

[180] M. K. Jha, S. Dutta Roy, and B. Lall. DEMD-based Video Coding for Textured Videos in an H.264/MPEG Framework. Pattern Recognition Letters (PRL), 51(1):30 - 36, 2015.

[181] Y. Zhang, D. Agrafiotis, and D. R. Bull. High dynamic range image and video compression a review. International Conference on Digital Signal Processing, pages 1 - 7, 2013.

[182] P. Merkle, A. Smolic, K. Muller, and T. Wiegand. Multi-view video plus depth representation and coding. Proc. IEEE International Conference on Image Processing (ICIP), pages 201 - 204, 2007.

[183] ITU-T and ISO/IEC. Advanced video coding for generic audiovisual services. ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), 2010.

[184] ISO/IEC. Call for Proposals on 3D Video Coding Technology. ISO/IEC JTC1/SC29/WG11 N12036, 2011.

[185] N. Stefanoski, O. Wang, M. Lang, P. Greisen, S. Heinzle, and A. Smolic. Automatic view synthesis by image-domain-warping. IEEE Transactions on Image Processing, pages 3329 - 3341, 2013.

[186] N. Plath, S. Knorr, L. Goldmann, and T. Sikora. Adaptive image warping for hole prevention in 3d view synthesis. IEEE Transactions on Image Processing, pages 3420 - 3432, 2013.

[187] J. Konrad, M. Wang, P. Ishwar, C. Wu, and D. Mukherjee. Learning-based, automatic 2d-to-3d image and video conversion. IEEE Transactions on Image Processing, pages 3485 - 3496, 2013.

[188] J. Zhang, M. M. Hannuksela, and H. Li. Joint multiview video plus depth coding. Proc. IEEE International Conference on Image Processing (ICIP), pages 2865 - 2868, 2010

[189] W. Su, D. Rusanovskyy, M. M. Hannuksela, and H. Li. Depth-based motion vector prediction in 3d video coding. Picture Coding Symposium, pages 37 - 40, 2012.

[190] P. Ndjiki-Nya, M. Kppel, D. Doshkov, H. Lakshman, P. Merkle, K. Mller, and T. Wiegand. Depth image-based rendering with advanced texture synthesis for 3-d video. IEEE Transaction on Multimedia, 13:453 - 465, 2011.

[191] D. Farin, R. Peerlings, and P. H. N. de With. Depth-image representation employing meshes for intermediate-view rendering and coding. in 3DTV-Con, 2007.

[192] J. Hou, L. P. Chau, M. Zhang, N. M. Thalmann, and Y. He. A Highly Efficient Compression Framework for Time-Varying 3-D Facial Expressions. IEEE Transactions on Circuits and Systems for Video Technology, pages 1541 - 1553, 2014.

[193] J. Y. Lee, J. L. Lin, Y. W. Chen, Y. L. Chang, I. Kovliga, A. Fartukov, M. Mishurovskiy, H. C. Wey, Y. W. Huang, and S. M. Lei. Depth-based texture coding in avc-compatible 3d video coding. IEEE Transactions on Circuits and Systems for Video Technology, pages 1347 - 1361, 2015.

[194] Cisco visual networking index: Forecast and methodology. http://www.anatel.org.mx/. 2013 to 2018.

[195] W. Zeng, Y. Zhao, and K. Ou. Research on cloud storage architecture and key technologies. in Proc. 2nd Int. Conf. Interact. Sci.: Inf. Technol., Culture Human, pages 1044 - 1048, 2009.

[196] W. Zhu, C. Luo, J. Wang, and S. Li. Multimedia cloud computing. IEEE Signal Processing Magazine, 28:59 - 69, 2011.

[197] Z. Li, X. Zhang, and Q. He. Analysis of the key technology on cloud storage. in International Conference on Future Information Technology and Management Engineering (FITME), pages 426 - 429, 2010.

[198] A. Upadhyay, P. R. Balihalli, S. Ivaturi, and Shrisha Rao. Deduplication and compression techniques in cloud design. IEEE International System Conference, pages 1 - 6, 2012.

[199] L. Kau, C. Chung, and M. Chen. A cloud computing-based image codec system for lossless compression of images on a cortex-a8 platform. IEEE conference on Biomedical Circuits and Systems Conference (BioCAS), pages 288 - 291, 2012.

[200] O. C. Au, S. Li, R. Zou, W. Dai, and L. Sun. Digital photo album compression based on global motion compensation and intra/inter prediction. in Proc. Int. Conf. Audio, Language Image Process., pages 84 - 90, 2012.

[201] Z. Hou, R. Hu, Z. Wang, and Z. Han. Improvements of Dynamic Texture Synthesis for video Coding. International Conference on Pattern Recognition (ICPR), pages 3148 - 3151, 2012.

[202] Z. Hou, R. Hu, Z. Wang, and Z. Han. Cloud Model-based Dynamic Texture Synthesis for Video Coding. International Conference on Pattern Recognition (ICPR), pages 838 - 842, 2014.

[203] S. Sowmyayani and P. Arockia Jansi Rani. Frame differencing-based segmentation for low bit rate video codec using H.264. Int. J. of Signal and Imaging Systems Engineering, pages 41 - 53, 2016.

[204] M. K. Jha, B. Lall, and S. Dutta Roy. Video Compression Scheme Using DEMD Based Texture Synthesis. In Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), pages 90 - 93, 2011.

[205] M. K. Jha, B. Lall, and S. Dutta Roy. DEMD-based Image Compression in a Compressive Sensing Framework. Journal of Pattern Recognition Research (JPRR), 9(1):64 - 78, 2014

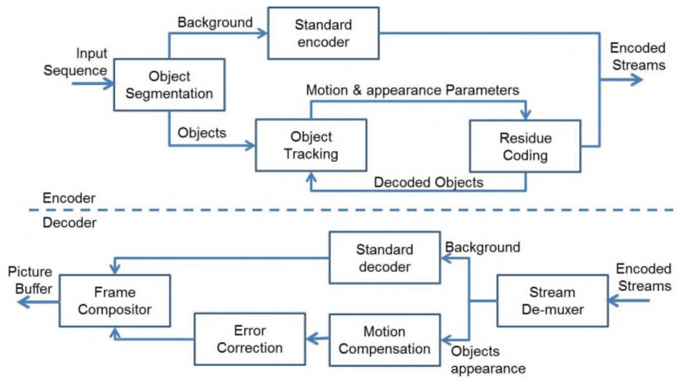

Figure 1 ***Object based coding system overview***

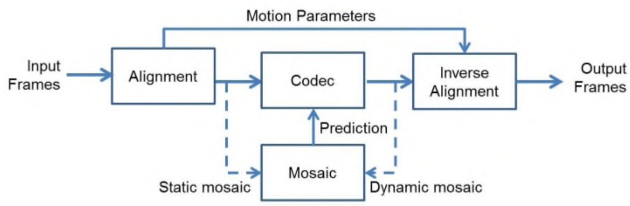

Figure 2 ***Mosaic based coding system overview***

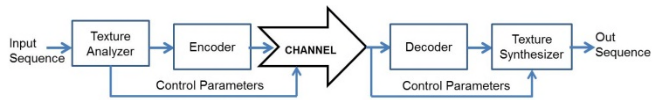

Figure 3 ***Texture based video compression system overview***

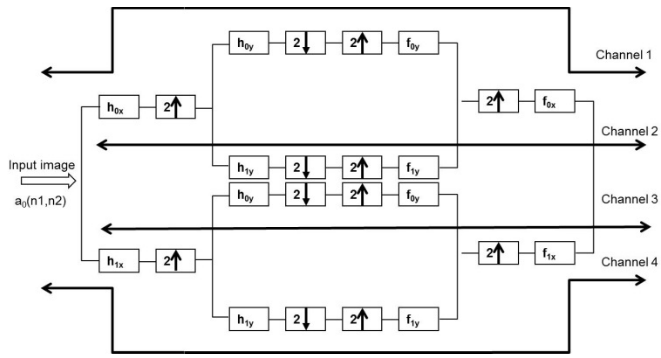

Figure 4. ***Statistically Matched Wavelet Filter Bank Overview***

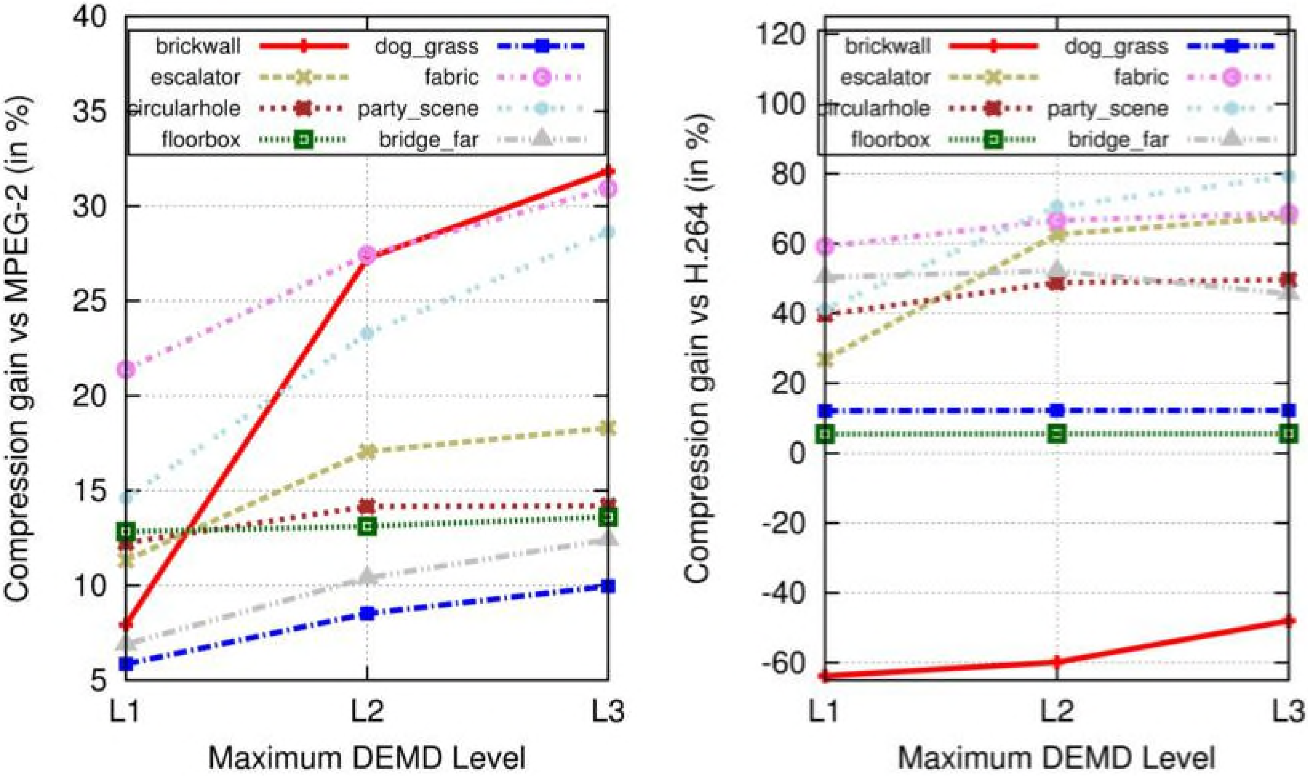

Figure 5 ***Compression gain over MPEG2/H.264 for the proposed multilevel IMF synthesis [180]***

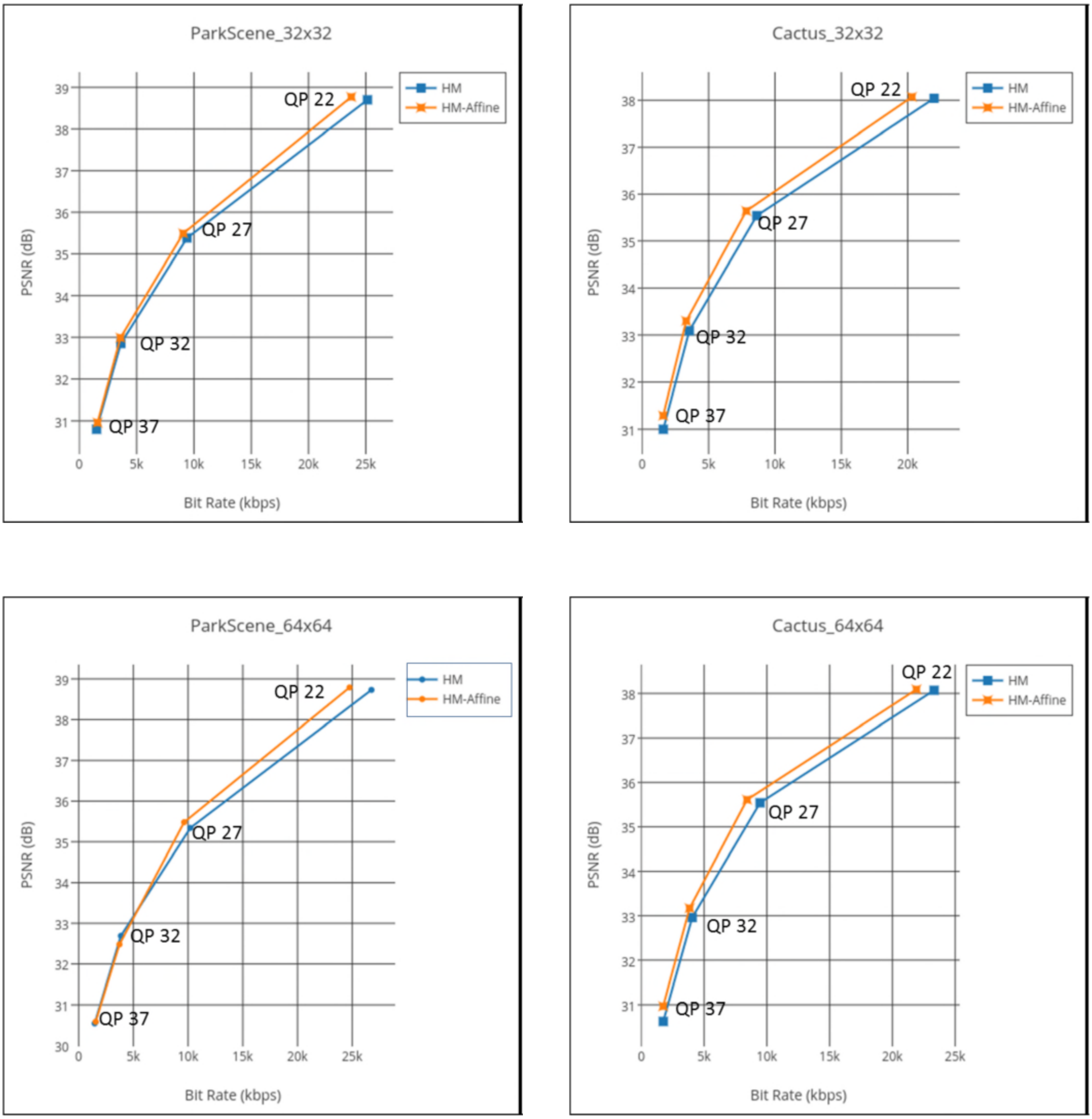

Figure 6 ***RD Plot comparing*** HM ***and*** $\text{HM}_{\text{Affine}}$ ***(pure HEVC, and the proposed method, respectively) for the sequence `Park Scene' (the top row) and 'Cactus' (the bottom row) using the coding block sizes of 32 32 (the graph on the left side on each row) and 64 64 (the one on the right). [64]***

Table 1 ***Texture Analysis and Synthesis Scheme : A Comparative Study***

Proposed Work	Texture Analysis Tools	Texture Synthesis (Model)	%Bits Gain vs. standards	Quality assessment	Side Information
Portilla[95]	Wavelet based	parametric (joint-statistics)	No claim	None	wavelet projections
Fan[126]	wavelet based	parametric (HMT-3S)	No claim	None	None
Wang[65]	split& merge	parametric (affine warping)	No claim	none	alpha, intensity
Dumiritas[139]	wavelet based	parametric (steerable pyramid)	56 (720x352)	none	wavelet-params., maps
Bosch[135]	Many	parametric (perspective) (warping)	36	None	motion-params., maps
Zhang[137]	wavelet based	parametric (ARMA)	55 (CIF)	Yes	motion-params., maps.
Khandelia[130]	Gradient & AR	parametric (AR model)	54.52 (QP=12)	Yes	AR parameters
Stojanovic[132]	ARMA based	parametric (ARMA model)	15 (QP=23)	Yes	picture parameters
Chen[133]	ARMA based	parametric (ARMA model)	7.39	Yes	None
Jha[179]	matched wavelet	parametric (compressive) (sensing)	No claim	Yes	compressive measurements
Jha[180]	DEMD	non-parametric (patch-based)	50.35	Yes	DEMD residual

Table 2 ***Overview of UHD attributes and its impact***

UHD Attributes	TV Compatibility	STB	Compre-ssion
WCG	Requires New Generation TV	Requires Next generation chip	Low Cost
HTR (100/120fps)	Low cost	Not ready	High Cost
HDR	Higher Power Consumption	Requires 10-bit support	High Cost
4K spatial resolution	Requires 55 inch TV	Requires HEVC 10-bit support	High Cost