3D Wavelet Transformation for Visual Data Coding with Spatio and Temporal Scalability as Quality Artifacts: Current State of the Art

Shaik. Jumlesha[1]

[1] Jawaharlal nehru university - Hyderabad

Received: 6 February 2012 Accepted: 3 March 2012 Published: 15 March 2012

Abstract

Several techniques based on the threeâ??"dimensional (3-D) discrete cosine transform (DCT) have been proposed for visual data coding. These techniques fail to provide coding coupled with quality and resolution scalability, which is a significant drawback for contextual domains, such decease diagnosis, satellite image analysis. This paper gives an overview of several state-of-the-art 3-D wavelet coders that do meet these requirements and mainly investigates various types of compression techniques those exists, and putting it all together for a conclusion on further research scope.

Index terms— Discrete Cosine Transform, Discrete Wavelet Transform, Image Compression, spacio scalability, temporal scalability.

1 Introduction

n various image and video applications compression is indispensable to guarantee interactivity during the streaming and consultation in particular about huge volume of medical images, for probing contextdependent full visual structures and/or quantitative analysis of the measurements. As a consequence, trading-off visual quality and/or implementation difficulty against bit-rate introduces exact constraints. On one hand, it is unbearable to drop any information when handling context exact visual data such as medical data.

On the other hand, a model likes progressive data transmission [52] and, thus, naturally support for lossy coding is equally important. This methodology allows for example to prioritize low-resolution edition of the requested motion based or still visuals and to increasingly filter the resolution of the visualized data by move additional data. This scalability mode is often referred to as resolution scalability. In a quality scalability scheme, the visual media is decoded instantaneously to the complete resolution but with a reduced visual quality. Additionally, by choose regions that are applicable for the context such as diagnosisi.e., the regions-of-interest (ROIs)-parts of the image can be assess in a very early transmission stage at full quality. Meanwhile, the background information will be further developed.

Moreover, it should be clear that we target best ratedistortion presentation over the complete range of bitrates that is demanded by the application. For example, JPEG2000 [53] (based on the wavelet transform) clearly outperforms its predecessor PEG [based on the discrete cosine transform (DCT)] [54] at low bit-rates and has as a significant property its lossy-to-lossless coding functionality; that is the ability to start from loss density at a very high density ratio and to increasingly refine the data by sending detail information, finally up to the stage where a lossless decompression is obtained.

Systems based on technologies other than the wavelet transform have been proposed, but they only partially assist the requested set of functionalities. Nonetheless, those techniques do superb for the subclass of applications they are designed for. Examples are context-based predictive coding (CALIC) [55] for lossless compression and region-based coding for very low bit-rate coding. Although these coders are competitive in their application domain, they lack support for the other functionalities.

Additionally, the increasing use of threedimensional (3-D) imaging modalities, like magnetic resonance imaging (MRI), computerized tomography (CT), ultrasound (US), single photon emission computed tomography (SPECT), and positron emission tomography (PET) triggers the require for capable techniques to transport and store the associated volumetric data. In the classical approach, the image volume is careful as a set of slices, which are consecutively compressed and accumulate or transmitted. Since modern transmission techniques need the use of concepts like rate scalability, quality scalability, and resolution scalability, multiplexing mechanisms require to be introduced to select from each slice the right layer(s) to support the actually essential quality-of-service (QoS) level. However, a disadvantage of the slice-by-slice mechanism is that potential 3-D correlations are ignored.

2 II.

3 Quality artifacts for visual data coding

In the applications of distributing motion based visuals such as video streaming or still visuals such as volumetric image sets over networks, distributing server has to deal with different network environments and client devices. The very diverse connection in mixed networks, ranging from hundreds of mega-bps to a number of tens of kilo-bps, and the fluctuations in bandwidth, need that video bit-streams are flexible in adapting itself to dynamic channel during transmission. The different client devices, with their different display capabilities and their computational and memory limitations, also need that bit-streams are flexible in decoding resolution and complexity. For these reasons, in recent years, apart from continuously improving the density effectiveness of non scalable video coding, a lot of research efforts have been paid to providing different scalabilities as quality artifacts in compressed bitstream, including spatial resolution, frame rate, quality and temporal scalabilities.

Among several scalable video coding schemes, some based on three-dimensional wavelet transform have concerned much attention [56], [57], [21], [20], [60]. In these schemes, wavelet filtering is useful in both spatial and temporal directions. The multi resolution property of wavelet representation makes it a nature solution for spatial-temporal resolution scalability. The resolution scalability is usually realized by dropping needless sub bands. Hence the fallowing two considered as quality artifacts for 3D wavelet based coding strategies. 1. Spatial scalability 2. Temporal scalability III.

4 Spatial Scalability as Quality artifact for 3D wavelet coding

Spatially scalable or hierarchical video coders generate two bit-streams: a base layer bit-stream, which represents low-resolution pictures, and an improvement layer bit-stream, which provides additional data wanted for reproduction of pictures with full resolution. A significant feature is that the base-layer bit-stream can be decoded separately from an enhancement layer. Therefore, low-resolution terminals are capable to decode only the base-layer bit-stream in order to display low-resolution pictures. Such density techniques are of great interest recently, because of growth of communication networks with different transmission bit rates [62]- [65]. Moreover, scalable transmission is useful in error-prone environments where base-layer packets are well secluded against transmission errors and losses, while the security of the enhancement layer packets is lower. In such a system, a receiver is able to copy at least low-resolution pictures if quality of service decreases.

There were several attempts to develop spatially scalable coding of video. The proposed schemes were based on pyramid decomposition [61] or subband/wavelet decomposition [62], [63], [65]. Among different proposals, the latter approach should be considered especially promising. The idea is to divide each image into four spatial sub-bands. The sub-band of lowest frequencies comprises a base layer, while the other three sub-bands are jointly transmitted in an improvement layer (Fig. 1). Nevertheless, this approach often leads to allotment of much higher bit rates to a base layer than to an improvement layer, which is disadvantageous for practical applications. Recently, Benzler [66] has proposed to keep away from this problem by combining spatial and SNR scalability and abandon the obligation of the full MPEG-compatibility in the base layer. Here, our goal is to use a fully MPEGcompatible coder in the base layer. For the essential codecs, spatio-temporal scalability is proposed [67], [68]. Here, a base layer corresponds to the bit-stream of the pictures with compact both spatial and temporal resolutions. Therefore, in the base layer, the bit rate is decreased as compared to a encoder with spatial scalability only. Now, it is easy to get the base layer bit rate equal or even less than that of the development layer. The development layer is used to transmit the information required for restoration of the full spatial and temporal resolution.

Embedding of sub-band decomposition into a motion-compensated encoder leads to in-or out-band motion reparation performed on individual sub-bands or on the whole image, respectively. The latter will be used here, because some experimental results show that it is more capable [62], [63], [65].

Here, the term of spatio-temporal scalability is proposed for a functionality of video compression systems where the base layer corresponds to pictures with compact both spatial and temporal resolution. An improvement layer is used to transmit the information required for return of the full spatial and temporal resolution.

The authors have already considered two basic approaches related to spatio-temporal scalability [67], [68]. The first approach exploits 3-D sub-band analysis while the second approach is based on B-frame data partitioning.

5 a) First Approach

The input video sequence is analyzed in a 3-D separable filter bank, i.e., there are three successive steps of analysis: temporal, horizontal, and vertical. For temporal analysis, very simple linear-phase two-tap filters are used similarly as in other papers on threedimensional sub-band coding [69], [70] 1 ( ) 0.5.(**??**)H z z ? = ±

Where "+" and"-"correspond to low-and high pass filters, respectively. This filter bank has a very simple implementation, wants to store one frame only and exhibits small group delay. H of low and high temporal frequencies, respectively. In both sub-bands, the temporal sampling frequency is compact by factor two. Therefore, these two sub-bands correspond to two video sequences with compact frame frequency. The two sub-bands are separated into four spatial sub-bands (LL, LH, HL, and HH) each. For spatial analysis, both horizontal and vertical, independent FIR filters are used. The 3-D analysis results in eight spatio-temporal sub-bands (Fig. 2). Three high-spatial-frequency sub-bands (LH, HL and HH) in the high-temporal-frequency sub-band t H are discarded, as they correspond to the information being less applicable for the human visual system. ? The improvement layer includes the spatial subbands LH, HL, and HH from the temporal sub-band t L and the spatial sub-band LL of the temporal subband t H .

6 b) Second Approach

In the second alternative, the technique employs data structures already designed for standard MPEG-2 coding. Reduction of temporal resolution is obtained by elimination of each second frame. It is unspecified that groups of pictures (GOPs) consist of even number of frames. Moreover, it is unspecified that each second frame is a B-frame, i.e., it can be removed from a sequence without moving the decoding of the remaining frames.

Reduction of spatial resolution is obtained by use of sub-band decomposition. Suitable design of the filter bank results in negligible spatial aliasing in the LL sub-band, which constitutes the base layer. Unfortunately the technique does not provide any means to suppress temporal aliasing. The effects of temporal aliasing are like as those related to frame skipping in hybrid encoders.

The base-layer data are used to create lowquality images; therefore, it is sensible to perform more rough quantization here than in the improvement layer. On the other hand, quality of the sub-band LL is strongly related to the quality of the full sized picture. The small quality of the LL sub-band restricts the full-sized picture quality to a relatively low level, in spite of the amount of information in the remaining sub-bands. Therefore, it is important to transmit additional information LL ? in the improvement layer. This information is used to get better quality of the sub-band LL when used to synthesize fullsized images in the improvement layer.

IV.

Temporal Scalability as Quality artifact for 3D wavelet coding

Temporally scalable video coders can be classified intone of two types depending on the manner in which temporal redundancy is exploited. The first is the motion-compensated analytical coder (e.g. MPEG, [71]), and the second is the sequential sub band coder both without ([72], [73]) and with ([74], [75]) motion compensation (MC). The coding competence and the degree of temporal scalability are a function of size of a group-of-frames (GOF), defined as the number of successive frames that can be decoded separately from the rest of the video sequence. Temporal scalability is achieved by decoding subsets of the GO consisting of equally-spaced frames.

MCP exploits temporal redundancy by forming a(closed loop) prediction of the current frame via MC from a reference frame. The coding competence of MCP is needy on the success of the MC. Good MC considerably increases the association coefficient among pixels, which then yields less energy in the residual. Temporal scalability from MCP video is provided by strategic placement of orientation frames and selective decoding of frames. Recursive prediction with a GOF of length allows lower frame rates of ½, ¼, 1/2v. The result is a simple temporal sub sampling of the original sequence.

Temporal sub band coders and motioncompensated TSB coders exploit redundancy by applying a sub band or wavelet analysis in time. The most usually used filters are the 2-tap Haar wavelet filters applied hierarchically, which result in GOFs with lengths that are powers of 2.Temporal scalability is provided by decoding and synthesizing chosen temporal sub bands. To include block-based motion recompense with TSB, MC is performed on individual blocks prior to each application of the low-pass filter and therefore local motion of different objects in a scene can be compensated well. To make sure inversion ability of the wavelet transform, MC must be full-pel. The resulting scene-aligned pixels are temporally sub band filtered. However, regions of pixels can now be precious by the occlusion problem which occurs when a one-to-one correspondence among pixels in two frames does not exist in the MC operation. These regions must be especially coded. The standard solution issue analytical coding to maintain reconstruct capacity for these pixels [4]. Regions in the prior frame not found in the present frame are placed in the low-pass sub-band and coded directly. Regions in the current frame not found in prior frame are subtracted from a prediction and placed in the high-pass sub band. Since quantization occurs after forming the temporal sub band representation, such coding is efficiently an open-loop prediction system.

An important difference between MC-TSB and MCPc ding is the effects of full-pel and half-pel MC. MC can be considered as preprocessing previous to temporal sub-band filtering. While the temporal sub band filtering (and predictive coding for covered/uncovered pixels) is invertible, the preprocessing step must also be invertible to allow exact synthesis of the original frames in the absence of quantization. Inverting capacity is only provided by full-pel motion compensation.

V.

7 Nomenclature of 3-d wavelet coding a) 3-D DCT Coding

The first coder introduced in the 3-D test bed is a JPEG-alike, 3-D DCT-based coder. This coder was designed in order to have a good reference for DCTbased systems. The 3-D JPEG-based coder is composed of a DCT, followed by a scalar quantize and finally a combination of RLC and adaptive arithmetic encoding. The basic principle is simple: the volume is separated in cubes of 8x 8 x 8 pixels (N=8) and each cube is separately3-D DCT-transformed, similar to a classical JPEG-coder.

Thereafter, the DCT-coefficients are quantized using a quantization matrix. In order to derive this matrix, one has to consider two options. One option is to construct quantization tables that create an optimized visual quality based on psycho-visual experiments. It is valuable mentioning that JPEG uses such quantization tables, but this approach would need complicated experiments to come-up with sensible quantization tables for volumetric data. The simplest solution, adopted in this work, is to create a uniform quantization matrix-as reported in [76], [77], and [78]. This option is motivated by the fact that uniform quantization is optimum or quasi-optimum for most of the distributions [79]. Actually, the uniform quantizer is optimum for Laplacian and exponential input distributions; otherwise the differences with respect to an optimal quantizer are marginal [79]. A second option involving quantizes that are optimal in rate-distortion sense is discussed elsewhere [80].With 1 0 ( ) 2 0 i i i u N C u u N ? = ? ? = ? ? > ? ?

The quantized DCT-coefficients are scanned using a 3-Dspace-filling curve, i.e., a 3-D instantiation of the Morton-curve [81], to allow for the alignment of zerovalued coefficients and, hence, to get better the performance of the RLC. This curve was opted for, due to its simplicity compared with that of 3-D zigzag curve [82]. The nonzero coefficients are encoded using the same classification system as for JPEG. The coefficient values are grouped in 16 main magnitude classes (ranges), which are subsequently encoded with an arithmetic encoder [83]. Finally, the remaining bits to refine the coefficients within one range are added without further entropy coding.

The adopted entropy coding system is partially based on the JPEG architecture [54], although the Huffman coder is replaced by an adaptive arithmetic encoder [83]. Consequently, the big look-up tables mentioned in annex K of the standard [54] are extra and moreover, adaptive arithmetic encoding tends to have a higher coding efficiency. The dc coefficients are encoded with a predictive scheme: apart from the first dc coefficient, the entropy coding system encodes the difference between the current dc coefficient and the previous one. For this distinction, the range is determined and encoded with an arithmetic encoder that has a dc model supporting16 ranges. Simply transmitting the remaining bits of the coefficient refines the range specification without any further entropy coding. The latter is possible since the probability distribution of all possible values can be seen as uniform, hence, entropy coding will not be capable to further reduce the bit consumption.

The ac coefficients are encoded by specifying first the amount of zeros preceding the encoded symbol, i.e., the run. The runs of zeros are encoded using an arithmetic encoder with a separate model. Runs of up to 15 zeros are supported. Note that to indicate the situations in which 16 or more zeros precede a important coefficient, an extra symbol ”OVF” (overflow) is used. After encoding this symbol, the remaining zeros are directly encoded to avoid confusing situations involving a succession of several OVF(D D D D) F 2012 Year 3 1 1 1 1 0 2 0 3 0 1 (2 1) ( ) ( ) ( ) cos[ ] 2 N N N i i i x x x i x DCT u f x C u u N ? ? ? ? ? ? ? ? ? ? ? ? ? ? encodings.

Finally, the range of the encountered important symbol is encoded, using an arithmetic encoder with a similar (AC) model as in the case of the dc coefficients, followed by the essential refinement bits.

8 b) The 3-D Wavelet Transform

Before describing in the following sections the proposed 3-Dwavelet-based techniques, it is significant to notice that these techniques support lossless coding, all the necessary scalability modes as well as ROI coding and this is a important variation with respect to the 3-D DCT-technique presented above, which is not able to provide these features.

For all the 3-D wavelet-based coders included in this study, a common wavelet transform module was designed that supports lossless integer lifting filtering, as well as finite-precision floating-point filtering. A mixed selection of filter types and a different amount of decomposition levels for each spatial direction (x-, y-, or z-direction) are supported by this module. This allows for adapting the size of the wavelet pyramid in each spatial way in case the spatial resolution is limited.

For example, fewer levels will be required along the slice axis if the amount of slices or the resolution along the axis is limited. The supported lossless integer lifting filters include the (S+P), (4,2), (5,3), and (9,7) integer wavelet transforms. This selection is based on current publications [85], [86], as well as investigations performed in the context of the JPEG2000 compression standard.

A typical problem encountered with 3-D lossless integer wavelet transforms is the complexity wanted to make them unitary, which is not the case for floating-point transforms. This property is essential in order to achieve a good lossy coding performance. By calculating the L2 norm of the low-and high-pass filters, the normalization factors can be determined. In two dimensions, this is not a problem, since the typical scaling factors to obtain a unitary transform are about powers of two [87].

However, in three dimensions, the problem pops up again and it only disappears if one takes care that the sum of all decompositions influencing each individual wavelet coefficient (i.e., decompositions in both slice directions and in the axial direction) is an even number. Hence, some proposals have been formulated [88], [89] that make use of a wavelet packet transform [90] to achieve this goal, while assuming that the L2based normalization factors for the supported kernels scale-up 2 with for the low-pass and 1 2

for the highpass kernels. In practice this seems to be an acceptable approximation. Nevertheless, in the presented study, whenever possible, unitary transforms will be used (and it will be explicitly mentioned if not).

9 c) 3-D SPIHTs

In the test set of wavelet coders, a 3-D SPIHT encoder [91] was included as a reference. An early version of this coder [89] has already established to beat the performance of a context-based octave zerotree coder [85]. The source code was made presented by the authors so it could be equipped with the proposed wavelet transform front-end.

The SPIHT implementation in this study uses balanced 3-Dspatial orientation trees. Therefore, the same number of recursive wavelet decompositions is necessary for all spatial orientations. If this is not respected, several tree nodes do not refer to or be linked with the same spatial location and, as a result, the dependencies among different tree-nodes are destroyed and, hence, the compression performance is reduced. Thus, a packet-based transform is not working to obtain a unitary transform with this embedded coding system. Therefore, the SPIHT coder was equipped with a no unitary transform. It is, however, worthwhile mentioning that solutions have been proposed utilizing unbalanced spatio-temporal orientation trees in the context of video coding [92].

The examined 3-D SPIHT algorithm [91] follows the same procedure as its 2-D homologous algorithm, with the exception that the states of the tree nodes-each embracing eight wavelet coefficients are encoded with a context-based arithmetic coding system during the significance pass. The selected context models are based on the significance of the individual node members, as well as on the state of their descendents. Consequently, for each node coefficient four state combinations are possible. In total 164 different context models are used.

10 d) Cube Splitting (CS)

The CS technique is derived from the 2-D SQP coder proposed in Section II-C. In the context of volumetric encoding, the SQP technique was comprehensive to a third dimension: from square splitting toward CS. CS is applied on the wavelet image in order to isolate smaller entities, i.e., sub cubes, possibly containing important wavelet coefficients.

11 During the first significance pass

p q q b Q k v q ? ? with top-left coordinates 1 2 3

(, ,)q q q q k k k k = and of size 2 2 2 2 1 2 3

(, ,) ? (w(k)) =1are isolated. Thus, the significance pass max Sp registers sub cubes and wavelet coefficients, newly identified as important, using a recursive tree structure of octants. The result is an octtree-structured description of the data meaning against a given threshold. As might be noticed, equal significance weights are given to all the branches. When a important coefficient is isolated, also its sign for which two code symbols are conserved is immediately encoded. When the complete bit-plane is encoded with the significance pass max Sp , p is set to max p -1 and the refinement pass R max p -1 is initiated for this bit-plane, refining all coefficients marked as significant in the octtree.q q q q v v v v =

Thereafter, the significance pass is restarted to update the octtree by identifying the new significant wavelet coefficients for the present bit-plane. During this stage, only the before non significant nodes, i.e., 1 ((, / 2)) 0, 0p q q j Q k v j J ? + = = < ? , are checked for

significance and the important ones, i.e.

1 ((, / 2)) 1p q q j Q k v ? +

== ,are unnoticed since the decoder already received this information. The described procedure is frequent, until the complete wavelet image W is encoded, i.e., p=0 or until the desired bit-rate is obtained. To encode the generated symbols professionally, a context-based arithmetic encoder was integrated. The context model is simple. For the significance pass four context models are distinguished, namely one for the symbols generated at the intermediate cube nodes, one for the pixel nodes having no significant neighbors for the earlier threshold, one for the pixel nodes having at least one significant neighbor for the earlier threshold and finally one for encoding the sign of the isolated significant pixel nodes. Only two contexts are used for the refinement pass: one for the pixel nodes having no significant neighbors for the earlier threshold, one for the pixel nodes having at least one important neighbor for the previous threshold.

Other 2-D techniques, like NQS [84] and sub band block(SB) SPECK [59], have been proposed that use similar quadtree decomposition techniques. These coders divide the wavelet space in blocks and activate for each block disjointedly a quad-tree coding mechanism. In case of SB-SPECK, the block sizes are also depending on the sub band sizes, forcing each block to reside in one sub band. Each block is individually encoded and thereafter an EBCOT-alike rescheduling takes place to restore the scalability functionality. SB-SPECK was also partially extended to 3-D i.e., 3-D SB-SPECK coding [59] by equipping the coder with a 3-D wavelet transform front-end. The transform is activated on discrete chunks of slices [groups of frames (GOFs)], to maintain the accessibility of the data (typical GOF sizes are 8, 16, or 32 planes). The option is not implemented in the coders we designed.

SB-SPECK does not use arithmetic encoding. However, the 3-D SB-SPECK coder delivers competitive results and we will refer to it whenever possible. e) Three-Dimensional QT-L

The QT-L coder has also been extended toward 3-D coding. The octtrees corresponding to each bit-plane are constructed following a similar strategy as for the CS coder. However, the partitioning process is limited in such a way that once the volume of a node3 1 / 2 ,0 q j n i V v j J = = < ? ?

, becomes smaller than a predefined threshold th V , the splitting process is stopped and the entropy coding of the coefficients within such a significant leaf node ((, / 2)) 1p q q j Q k v ? = is activated. Similar to the 2-D

version, the octtrees are scanned using depth-first scanning. In addition, for any given node, the eight descendant nodes are scanned using a 3-D instantiation of the Morton-curve [81]. For every bitplane, the coding procedure consists of the non significance, importance and modification passes adapted for 3-Dcoding; also, for the maximum bit-plane, the coding process consists of the significance pass only. Notice that the sum number of neighbors tot N in (**??**) is set to 26 in 3-D coding.

The CS-EBCOT coding [58] join the principles utilized in the CS coder with a 3-D instantiation of the EBCOT coder [43]. In the next paragraphs the interfacing of the CS coder with a edition of EBCOT modified to 3-D is discussed.

To begin with, the wavelet coefficients are separation EBCOT-wise in separate, uniformly sized cubes, called code-blocks. Normally, the first size of the code-blocks is 64x64x64 elements. Additional sizes (even different ones for every dimension) can be chosen, depending on the image characteristics and the request requirements. The coding module CS-EBCOT again consists of two major units, the Tier 1 and Tier 2 parts.

The Tier 1 of the proposed 3-D coding architecture is a hybrid module joins two coding techniques: CS and fractional bit-plane coding by context-based arithmetic encoding. The Tier 2 part is equal to the one used in the 2-D coding system. 1. CS: The CS pass S is resulting from the CS technique presented in Section III-D. In the proposed coding system, the CS is useful on the individual code-blocks in arrange to separate smaller entities, i.e., sub cubes, possibly containing major wavelet coefficients. The least sub cube size(D D D D) F 2012 Year f) 3-D CS-EBCOT

that is supported is 4x 4 x 4.We will refer to these least sub cubes as the leaf nodes.

During the initial CS pass max i

12 Sp

, the importance of code-block i B is tested for its maximum bit-plane max i p with the significance operator p ? . Ifmax i p ? (i B) = 1, the code-block i B is join until the important leaf nodes max (, / 2) i p q q G b Q k v are isolated,

where G state the highest total of CS levels. When every significant leaf nodes are isolated, the fractional bitplane coding part is activated for the present bit-plane and only for the important leaf-nodes.

When the total bit-plane is encoded utilizing the fractional bit-plane coding, i p is set to max i p -1 and the succeeding CS pass, max i s -1 is activated. The explain procedure is repetitive, until the total block is encoded, i.e., i p =0 . Due to the limited quantity of code-symbols and their allocation, arithmetic coding is not useful. 2. Fractional Bit-Plane Coding: The fractional bit-plane coder encodes just those leaf nodes that have been recognized as important throughout the CS pass. Three passes are defined per bit-plane like in the 2-D case: the importance transmission pass, the magnitude refinement pass and the normalization pass. Moreover, these coding passes call numerous coding operations (primitives), i.e., the ZC, SC, MR, and RLC primitives. These primitives enable the choice of suitable3-D situation models for the successive arithmetic coding or RLC stages. The chosen adaptive arithmetic encoder is based on an implementation by Said and Pearlman of the algorithm proposed by Witten et al. [45], which is equal to those utilized in the earlier state encoders.

The data exist in in each leaf-node is scanned applying as lice-by-slice scanning pattern. Inside one slice the pattern is equal to the JPEG2000 scanning: the vowels are read in-groups of four vertically associated voxels. When a total slice is stripe-wise processed, the subsequent segment is processed.

The fractional coding passes perform in an equal way as for the original EBCOT execution. However, the chosen neighborhood k ? refers now to the twenty-six voxels approximately the voxel being coded (i.e., the immediate neighbors). For every bitplane, sequentially the significance propagation pass, the MR pass, the CS pass and the normalization pass are called, excluding for the first bit-plane where the initial two passes are discarded. 3. Coding Primitives: As for EBCOT, four coding primitives are defined to support the encoding procedure in the dissimilar coding passes: the ZC primitive, the RLC primitive, the SC, and the MR primitive. For arithmetic encoding, the contextmodel choice is based on the condition of the adjacent voxels of the voxel being encoded, i.e., the preferred neighborhood and the sub band type in which the voxel is situated. The preferred neighborhood k ? is separated in 7 orthogonal subsets according to their position to the voxel [58], [80]. Every coding primitive has got its individual look-up table to identify the probability model that has to be utilized by the arithmetic coder for a identified situation state [58], [80]. Additionally, we have to state that the complexity of this part of the coding engine increases heavily compared with the original 2-D implementations, appropriate to the enlarged preferred neighborhood (from 8 to 26 neighbors)and, consequently, the augmented complexity of the look-up tables [58]. 4. Tier 2-Layer Formation: The followed process, i.e., PCRD optimization [43], is equal to the original one. However, we have to state one feature that is of key importance. The PCRD routine allocates compensating for the fact that a no unitary transform has been used. By accurate the calculated distortions for each pass i n with a scaling factor i b ? , the coding method will execute as if a unitary transform was used (or approximated when using integer powers of 2).Hence, the distortion will be currently described by:2 ([ ] [ ]) i i i n n i i i i k B D b S t k s k ? ? = ? ? Where [ ] i s k indicate the magnitude of element k in code-block i B and [ ] i n i

S k provide the quantized illustration of that element connected with truncation point i n . This improvement allows support for a unitary transform without obstructing the possibility of lossless coding, a difficulty that does occur with classical zero tree-based coders. The original 2-D algorithms maintain multiple components (e.g., color), but this feature is not retained in the future 3-D implementation. Hence, only gray-scale images (volumes) are supported. Nevertheless, the dissimilar bit-stream chunks are currently grouped into separate packets, every packet contributing to one quality layer and one resolution level. The code-block addition information is again encoded by means of the tag-tree concept. The only alter that has been complete was extending the tag-tree idea to the third dimension, i.e., moving from a quad tree structure to an octtree structure.

13 Current state of the art

Matching Pursuits (MP) has been recognized as an useful technique of over entire transform coding for 2D images, originally for 2D motion compensated residuals by Neff and Zakhor [12], and additional recently extended to still images by the application of a spatial wavelet transform [13] and the utilize of additional effective dictionaries and embedded coding [14]. The addition of MP to three dimensions is natural, as has been established [15]. However MP is a computationally strong method and the utilize of 3D bases increases this cost significantly. In this context Yuan Yuan [1] proposed a novel 3D WAVELET VIDEO CODING WITH REPLICATED MATCHING PURSUITS. The model can be briefed in figure 1 and description follows Replicated MP coding is shown in Figure 1. For a group of 16 input frames, the motion compensated 3D DWT is implemented subsequent the scheme of Taubman and Secker [5A, 6A]. The related motion vectors are coded as side information by a Variable Length Code (VLC). This movement overhead affects the low down bit rate performance of all video coded by related schemes. The MP algorithm is useful to the 3D Group of Temporal Pictures (GOTP) using the Intra 8 Codebooks of Monro [14]. The benefit of 3D RMP is that the collection of Child atoms related with a Parent can provide up to three times the image power of the Parent alone, but expenses on average less than 30% of the bits to code in this implementation. MP is a computationally costly algorithm, with the best inner product over the entire data set with all bases in the dictionary essential for selecting each atom.

Sun et al [2] proposed a novel content adaptive rate-distortion optimization scheme, which might effectively differentiate texture region, edge region and flat region by means of directional field technique. There have been numerous efforts in the past trying to include perceptual procedures into video encoding. In [16], the focus was mostly on determination of suitable quantization steps with sub-band Just Noticeable Distortion (JND). In recent H.264/MPEG-4 advanced video coding standard, maximum coding efficiency is achieved by introducing the rate distortion optimization (RDO) procedure to provide the best coding outcome by maximizing image quality and minimizing data rate at the same time. In [17], a new adaptive RDO scheme has been proposed which exploits motion and texture masking property to correct the Lagrangian multiplier and achieves generally bit rate reduction by allowing additional distortion in the less noticeable background random texture area. The RDO technique is also an significant part in wavelet-based scalable video coding (SVC) scheme that is presently under examination by MPEG-21, Part 13 [18]. The SVC scheme requires an embedded bit stream to be formed from which bit streams with different bit rate, resolution and frame rate could be extracted with reasonably fine quality. Here in this job sun et al proposed an effective directional field based visual significance map in the context of capturing the features associated to pre attentive processing such as edges and curves. Based on the visual significance map, the regions with additional preattentive features are likely to get more distortionreduction by assigning a smaller Lagrangian multiplier. Rate balance was also measured as a factor and attempted to attain by assigning a relatively larger Lagrangian multiplier to the random texture area so that additional distortion is permitted without noticeable visual degradation to the image. Since HVS is also sensitive to distortions in flat regions, a little Lagrangian multiplier was also used for flat regions.

Seran et al [3] proposed a 3D BASED VIDEO CODING to carry out the two-dimensional spatial filtering initial and then perform motion-compensated temporal filtering by lifting in the Over total Discrete Wavelet Transform domain. In practice essentially the three-dimensional wavelet decomposition can be performed in two ways: two-dimensional spatial filtering followed by temporal filtering (2D+t) [19,20,21] or, temporal filtering followed by two-dimensional spatial filtering (t+2D) [22, **??**3,24]. In this work, Seran et al [3] proposed a new temporal filter set to reduce delay in 3D wavelet based video coding, that attempted to increase performance at par with existing longer filters. In this proposal the filter set haven't include any boundary effects at the group of frames (GOF). The length of the GOF can differ from five to any number of frames depending on interruption requirements. This proposed model also illustrated a novel technique of assigning priorities to temporal sub-bands at dissimilar levels to manage distortion fluctuation inside a GOF.

Mavlankar et al [4] considered a multiple description (MD) video coding scheme based on the motion compensated (MC) lifted wavelet transform, which is to carry out the temporal decomposition of a collection of pictures and then make multiple descriptions for every temporally transformed frame. The benefit of basing MD video coding on motioncompensated lifted 3D wavelet decomposition is that it does not need any difference control similar to in a hybrid codec which was earlier achieved by distribution drift compensation data. Earlier to this proposal an important number of MD coding schemes for video have been proposed (e.g., [25], [26], [27]). Most of the proposals for MD video believe the usual hybrid video coding structure with motion compensated prediction and DCT as their major building blocks. In difference the model proposed by Mavlankar et al a multiple explanation video coding scheme that is based on a video encoder that uses the recently emerging motion compensated lifted 3D wavelet change as its basis [27], [28], [29], [30]. Since recursive temporal prediction is restoring by a motion-compensated transform, it also remove the dependency quantization framework which is an inherent part of hybrid video codecs. In usual hybrid codecs quantization is fixed in the recursive prediction loop, whereas in 3-D wavelet codecs quantization and spatial encoding are applied following the temporal decorrelation of a group of pictures (GOP). Figure 2 shows our proposed MD video codec which is based on MC lifted wavelet change. This scheme expands the Drift Compensation Multiple Description Video Codec (DC MDVC) proposed for a hybrid video codec in [31] to MC-lifted 3-D wavelet coding.

Figure 2 : Overview of the proposed MD coding scheme A huge benefit of employing the 3-D wavelet scheme is that there is not require to send additional drift compensation data. DC MDVC has to send the drift compensation data to compensate for difference between the recursive predictions loops at the encoder and the decoder in case of loss. In fact the number of drift compensation streams rose exponentially with the total amount of descriptions. It is (2N ? 2), where N is the total number of descriptions. This can be simply understood by considering that a drift compensation stream has to be formed for each scenario excluding when all or no descriptions are received.

Seran et al [5] focused on the difficulty of controlling unpredictable variation of distortion in 3D coders, which aimed at exploring the MCTF filter properties and we present an entire analysis of the filter and mathematical derivations. The temporal wavelet filter properties are recognized to be a main factor contributing to distortion variation. The problem of controlling the temporal distortion fluctuations has been addressed in a little design [32], [33]. In [32], distortion variation control is considered for the bi-directional unconstrained motion compensated temporal filtering and the distortion in the decoded frame is expressed as a function of the distortions in the reference frames at the equal temporal level. In [33], the association among the distortion in temporal wavelet subbands and the reconstructed frames are study for the modified 5/3 filter (ignoring sqrt(2)). Based on the association, a distortion ratio model is theoretically developed and an easy rate control algorithm is used to place priorities for the temporal subbands according to the distortion ratio.

The association among the distortion in the reconstructed frames and the filter coefficients study by seran et al [5]. On this foundation, scaling coefficients for the filter are calculated to control the distortion fluctuation. In this circumstance the model proposed by seran et al [5] considered the mainly popular biorthogonal 5/3 filter and quoted that this model can be directly extended for additional longer filters.

Zefeng Ni et al [6] proposed and develops a 3D wavelet codec based on MCTF and JPEG2000 for a novel advance for stable quality aimed bit allocation between T-bands for the applications of adaptive stored video streaming. The reconstructed structure is divided into dissimilar groups according to the types of their related temporal bands. We suggest an estimated mathematical model to describe the relationship between the T-band distortions and the distortions of the reconstructed frames. Since we consider stored video in this investigation, we can offline produce the model parameters. Throughout the online transmission period, given the existing network bandwidth, we initial perform the conventional JPEG2000-like optimum truncation. After that, a two-step process is used for reducing the PSNR fluctuation, where the fundamental idea is to alter the energy gains to balance the dissimilar contributions from diverse types of T-bands and then more or less equally allocate distortion among the Tbands at the identical level. Along with model proposed by Zefeng Ni et al [6], little of the articles [34], [35], [36], [37] in latest literature have looking into that how to assign bits to every temporal band (T-band) so that definite degree of stable quality can be achieved. Although the difficulty of stable quality aimed bit distribution has been well calculated for conventional hybrid codecs [38,39], it is still an open question for MCTF based codecs. In the popular MC-EZBC codec [34], the authors suggest to discontinue bit plane scanning of every the GOPs at the identical fractional bit plane. However, it can only help accomplish similar distortion among T-bands, which does not lead to stable quality in reconstructed frames. This is because the distortion in T-bands propagates unequally into reconstructed frames, which is hard to model mathematically. In [35,36], in order to smooth PSNR performance, an optimized quantization step is obtained by analyzing the motion performance throughout the temporal filtering. In [37], the authors suggest to utilize an adaptive update step for the temporal filtering. Compared with conventional implementation, these technique indeed help dropping the serious PSNR fluctuation in reconstructed frames. Yongjian Man et al [7] presented a new 3D-WT algorithm for video coding, which can considerably decrease the processing memory and attain a high coding presentation. This paper is organized as follows: in section 2, we explain the traditional 3D-WT algorithm and its shortage. According to the decomposition structure, the model proposed is dissimilar from traditional 3D wavelet coding. As the input sequence for the proposed new 3D wavelet decomposition structure is based on numerous groups, for every group, only the little frequency frame is remained in processing memory for the final temporal decomposition. According to the analysis of the proposed model, only the high frequency frames are exported following temporal decomposition, while the low frequency frame remains in memory for every group.

Yu Liu et al [8] projected an expansion to the lifting-based activity threading procedure from the frame-based coding to the object-based coding, involved by the inimitable compensation of the objectbased coding that do not survive in other coding proposal. Object-based coding allows convenience and impressionability of object within a video series, and allows the organization of video content to endure the process of attainment, editing and allocation, which is useful for content-based search and recovery in MPEG-7. As an substitute to established video coding normal, 3D wavelet video coding has conventional much awareness recently. A main benefit of 3D wavelet video coding is that it can supply complete spatio-temporalquality scalability with non-redundant 3D sub-band disintegration. In 3D wavelet video coding, action reimbursement is regularly integrated into the sequential wavelet make over to accomplish competent coding routine, leading to a class of algorithms normally called as activity rewarded sequential sieve (MCTF). Previous to this occupation, Xu et al. [40] projected a motion yarn (MT) technique that employs longer wavelet filter to develop the long-term association across border along movement route. The aim of MT is to shape as lots of long clothes as likely since too much small gear will considerably augment the figure of reproduction borders. Luo et al. [41] planned an advanced MT method to reduce the numeral of many-to-one drawing pixels and non-referred pixels in the innovative MT. though the trouble of frontier effects grounds by the truncation of many-to-one drawing case in the inventive MT can be well explain by the advanced MT, the non referred pixel case, which is allocate to use the action vector from neighboring movement strand in the higher MT, is not well resolve since the allocate movement vector may be not precise or even mistaken for nonreferred pixel and may source some poverty on coding presentation.

Due to the boundary effects in spatial and sequential wavelet reconstruction, these artificial limitations will mortify greatly the coding recital. Consequently, it is improved to solve the problem of border line property, which survive in spatial and sequential makeover of object-based system concurrently, in a unified structure. Chen-Wei Deng et al [9] projected a new structure for scalable video convention. A mesh-based activity inference model is calculated and functional in this proposal, which engender a continuous proposal field. The earthly relationship is broken by Barbell thrilling [42]. In 3D wavelet video coding, wavelet change are exploit temporally across frames, and straight and up and down with each border, correspondingly.

[34],

[44] perform motioncompensated sequential strain (MCTF) in the inventive spatial field pursue by a spatial change on each sequential sub band. This is typically denote as a t+2D scheme. This preceding work was principally in the framework of a slab activity model, which poorly represent complex motion in real video sequences. Y. Andreopoulos et al [45] pertain spatial convert before the sequential one, which is typically referred to as 2D+t method. 2D+t can answer the spatial scalability concert issues of t+2D, but it experience from the shift-variant natural history of the DWT. [46] is a 2D+t+2D scheme. In such structure, spatial balance description of video succession are acquire opening from the superior resolution, while the coding system include inter scale prediction (ISP) machine in organize to develop the multiscale symbol idleness [47], the regulations competence is hard to achieve.

With reference to the limits claimed touching the representation [34], [44], [45], [46], [48] Chen-Wei Deng et al [9] projected a novel scalable 3D wavelet video coding support, which is a t+2D scheme. A meshbased motion model is incorporated into the planned scaffold, which is helpful to perk up solidity concert. In totaling, due to the fact that different temporal sub bands have different characteristics, two different wavelet system algorithms were planned for the low and high-pass activist sub bands, correspondingly.

Ke Xu [10] planned a novel scheme that is forced by circulated cause system to correct the errors of momentous division in 3D wavelet video tributary. In this projected replica Extra in rank is engender by Wyner-Ziv codec and is throw to the decoder. While errors occur, the relevant parts in the same frames from EZBC decoder are used as side in sequence to decipher the Wyner-Ziv bits to construct a refined substitution of the dishonored ones. to terminate the rest parts of these border are shared with superior parts from the Wyner-Ziv portrayal to yield a correct succession.

Later to this application Sehoon Yea et al [49]planned a motion reimbursement entrenched zero-block coder (MC-EZBC) and shows that MC-EZBC can achieve higher concert evaluate with H.264 at a high rate. Furthermore, it is easy to extract the various rates, declaration and eminence succession from the raw video succession in decoder. However, MC-EZBC is still disposed to errors and encodes the variously noteworthy parts uniformly. A number of scheme, such as Auto Repeat demand (ARQ) or Forward Error improvement codes (FEC) or grouping of both can be used to truthful errors in conjecture.

FEC methods are shaped by adding together extra ensure bits in the facade of the packet to check the error bits and moderate the likelihood of error. While, they can neither acceptable the fracture errors nor inefficiently adapt to the guide. ARQ scheme can progress the quality of decode movies by a response guide to send the lost packet again. Nevertheless, it is very disposed to time stoppage. Time delay will be engorged if the waterway is concerned by noises more commonly.

Both FEC and ARQ applied in established codec can be engaged in MC-EZBC shielding the bit tributary supposedly. Nevertheless, the decoder of MC-EZBC can work well within at least 4 frames indoors, so the importance is more time delay. Recently, Wyner-Ziv coding has been planned on error rigidity [50], [51]. They encode the source using established codec and Wyner-Ziv codec, which is used to suitable the errors in other report in that order. Ke Xu [10]extensive the same idea to 3D wavelet coding.

Nobuhara et al [11] planned A video coding method using max-plus algebra based three dimensional wavelet convert (3DMP-Wavelets) . The recompense of MP wavelets those definite by max-plus algebra are a. Since no hovering point calculation is requisite, the addition speed of MP-Wavelets is high. b. Since no duplication operation is required, MPWavelets are hardware implementation oriented. c. Since the problem of round-off errors is completely eliminated, MP-Wavelets are appropriate for digital watermarking, i.e., they be applied to exclusive rights fortification.

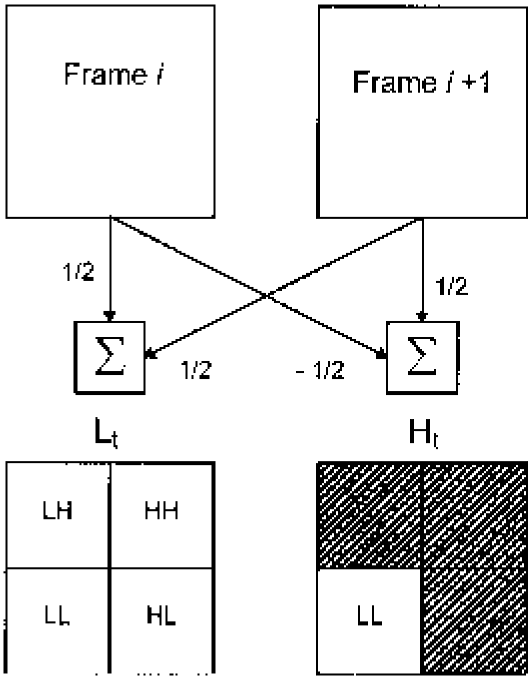
Hence 3DMP-Wavelets computational charge is very small since no balanced point calculations are done, max is computationally less expensive than the sum, and sum is computationally less costly than the product. The projected 3DMP-wavelets do better than a three dimensional linear wavelet in terms of velocity of the computation. Also, 3DMP-wavelets are hardware accomplishment oriented. This stimulates their study, mostly in view of some watch applications. It is easy to observe that in observation request it is important to have cheap strategy that can program a video succession at a low cost with rational eminence.

14 VII.

15 Conclusion

This paper has provided a picture of various tools that have been designed in recent literature for visual data compression, in particular 3D wavelet transformation and coding. It has focused on multiresolution representations with the use of the wavelet transform and its extensions to handle motion and ROI in visual sequences. For image compression, WT based approaches are showing quite competitive performance due to the energy compaction ability of the WT to handle piecewise polynomials that are known to well describe many natural images. In video sequences, the adequacy of such model falls apart unless a precise alignment of moving object trajectories can be achieved. This might remain only a challenge, since as for any segmentation problem; it is difficult to achieve it in a robust fashion, due to the complex information modeling which is often necessary. Most of the models studied in this paper are not generalized in the context of quality artifacts. Hence it remain an issue and providing research scope to identify the effective lifting schemes and 3D wavelet coding models in the context various contextual parameters such as region of interest, motion in streaming visuals and quality artifacts such as spatial scalability and temporal scalability. [1]

Figure 1:



2

Figure 2: Fig. 2 :

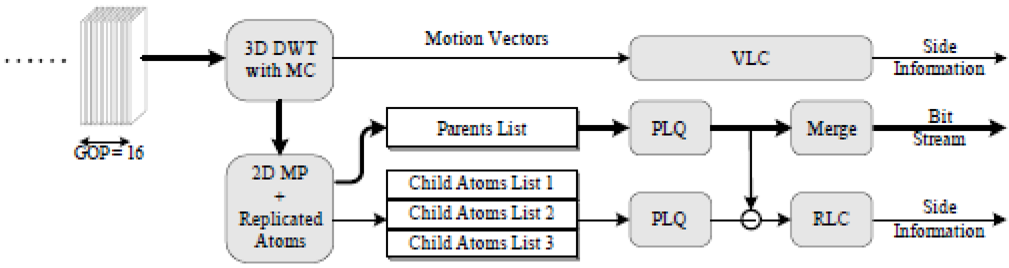
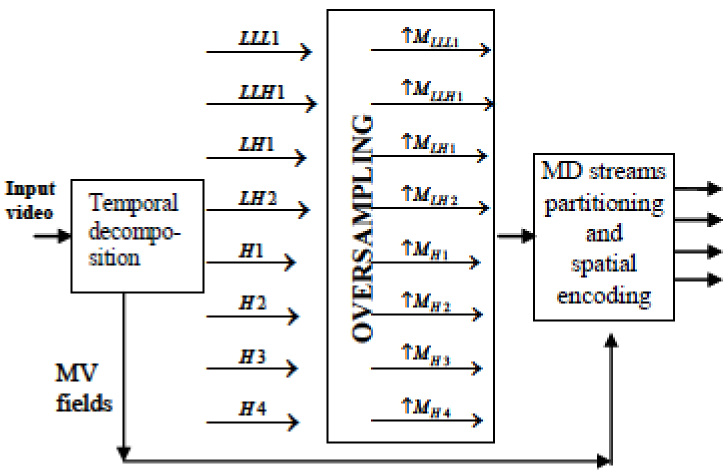


Figure 3: .



1

Figure 4: Figure 1 :

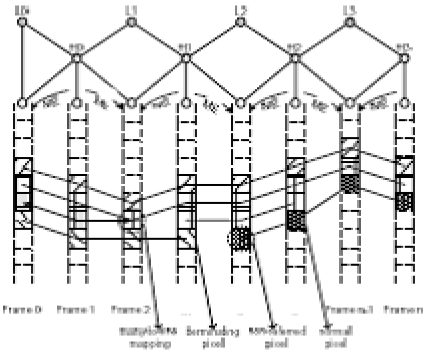


Figure 5:

[Yuan] , Yuan Yuan .

[Ni] , Zefeng Ni .

[Liu and Wu] , Yu Liu , ; Feng Wu .

[Sept ()] , Sept . 10.1109/ICIP.2005.1530371. 2005.

[Oct ()] , Oct . 10.1109/ISCIT.2010.5664855. 2010.

[Prorisc and Workshop (2000)] , /Ieee Prorisc , Workshop . Dec. 2000. p. .

[Schelkens et al.] '3-D compression of medical data based on cubesplitting and embedded block coding'. P Schelkens , X Giro , J Barbarien , J Cornelis . *Proc*, (null)

[Seran and Kondi ()] '3D based video coding in the overcomplete discrete wavelet transform domain with reduced delay requirements'. V Seran , L P Kondi . *ICIP 2005. IEEE International Conference on*, 2005. 3 p. . (III-233-6)

[King Ngi (2007)] '3D Objectbased Scalable Wavelet Video Coding with Boundary Effect Suppression'. Ngan King Ngi . 10.1109/ISCAS.2007.378011. *IEEE International Symposium on*, 2007. 2007. 27-30 May 2007. p. . (Circuits and Systems)

[Wang et al. ()] '3D video coding using redundant-wavelet multihypothesis and motioncompensated temporal filtering'. Y Wang , S Cui , J E Fowler . *Proceedings of the IEEE International Conference on Image Processing*, (the IEEE International Conference on Image ProcessingBarcelona, Spain) 2003. 2 p. .

[Monro (2005)] '3D wavelet video coding with replicated matching pursuits'. D M Monro . 10.1109/ICIP.2005.1529689. *ICIP 2005. IEEE International Conference on*, 2005. Sept. 2005. 1 p. . (Image Processing)

[Nobuhara et al.] '3D-wavelet decomposition based on max-plus algebra and its application to video coding'. H Nobuhara , T Tanabata , M Ono , B Bede . *Communications and Information Technologies (ISCIT), 2010 International Symposium on*, p. .

[Morton ()] *A Computer Oriented Geodetic Data Base and a New Technique in File Sequencing*, G M Morton . 1966. Ottawa, ON, Canada: IBM Ltd.

[Mavlankar et al. ()] 'A new update step for reduction of PSNR fluctuations in motioncompensated lifted wavelet video coding'. A Mavlankar , S Han , C L Chang , B Girod . *Proc. IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, (IEEE Int. Workshop on Multimedia Signal essing (MMSP)) 2005.

[Man et al. (2006)] 'A New Video Coding Based on 3D Wavelet Transform'. Yongjian Man , ; Lehua Wu; Shibiao He; Yongjun , Gu . 10.1109/ISDA.2006.253862. *ISDA '06. Sixth International Conference on*, 2006. 18 Oct. 2006. 2 p. 16. (Intelligent Systems Design and Applications)

[Cai et al.] *A novel frame-level bit allocation based on two-pass video encoding for low bit rate video streaming applications*, J Cai , Z He , C W Chen . (Image Representation)

[Chou and Chen ()] 'A perceptually optimized 3-d subband image codec for video communication over wireless channels'. C.-H Chou , C.-W Chen . *IEEE Trans. Circuits Syst. Video Technol* 1996. 6 (2) p. .

[Tsunashima et al. ()] 'A scalable motion-compensated subband image coder'. T Tsunashima , J Stampleman , V Bove . *IEEE Trans. Commun* 1994. 42 p. .

[Acoustics (2005)] Speech Acoustics . *Signal Process., (ICASSP 2005)*, (Philadelphia) March 2005.

[Hanke et al. ()] 'Adaptation of filters and quantization in spatio-temporal wavelet coding with motion compensation'. K Hanke , J R Ohm , T Rusert . *Proc. Picture Coding Symposium (PCS)*, (Picture Coding Symposium (PCS)) 2003. p. .

[Tsai et al. ()] 'Adaptive Rate-distortion Optimization using perceptual Hints'. Chun-Jen Tsai , Chih-Wei Tang , Ching-Ho Chen , Ya-Hui Yu . *2004 IEEE International Conference on Multimedia and Expo (ICME'2004)*, (Taipei, Taiwan) June 27th -30th, 2004.

[Luo et al. ()] 'Advanced motion threading for 3D wavelet video coding'. L Luo , F Wu , S Li , Z Xiong , Z Zhuang . *Signal Processing*, 2004. 19 p. .

[Luo et al. ()] 'Advanced motion threading for 3D wavelet video coding ", Signal Process'. L Luo , F Wu , S Li , Z Xiong , Z Zhuang . *Image Comm* 2004. 19 p. .

[Kim and Pearlman ()] 'An Embedded Wavelet Video Coder Using Three-Dimensional Set Partitioning in Hierarchical Trees'. B.-J Kim , W A Pearlman . *Proc. IEEE DCC*, (IEEE DCC) 1997. p. .

[Said and Pearlman (1996)] 'An image multiresolution representation for lossless and lossy compression'. A Said , Pearlman . *IEEE Trans. Image Processing* Sept. 1996. 5 p. .

[Witten et al. (1987)] 'Arithmetic coding for data compression'. I H Witten , R M Neal , J G Cleary . *Commun. ACM* June 1987. 30 p. .

[Asilomar Conf. Signals, Systems Computers ()] *Asilomar Conf. Signals, Systems Computers*, 1997. 2 p. 827.

[Xiong et al. (2004)] ‘Barbell lifting wavelet transform for highly scalable video coding’. R Xiong , F Wu , J Xu , S Li , Y.-Q Zhang . *Proc. PCS*, (PCSSan Francisco, CA) Dec. 2004. p. .

[Monro (2004)] ‘Basis picking for matching pursuits image coding’. D M Monro . *IEEE Int. Conf. Image Process. (ICIP2004)*, (Singapore) October 2004.

[Chen and Woods (2004)] ‘Bidirectional MC-EZBC with lifting implementation’. P Chen , J W Woods . *IEEE Trans. Circuits Syst. Video Technol* Oct. 2004. 14 (10) p. .

[Andreopoulos et al. ()] ‘Complete-toovercomplete discrete wavelet transforms: Theory and applications’. Y Andreopoulos , A Munteanu , G V D Auwera , J Cornelis , P Schelkens . *IEEE Trans. Signal Process* 2005. 53 (4) p. .

[Schelkens et al. ()] ‘Compression of Medical Volumetric Data’. P Schelkens , X Giro , J Barbarien , A Munteanu , J Cornelis . *ISO/IEC JTC1/SC29/WG1, N1712*, 2000.

[Compression of volumetric medical data based on cube-splitting Proc. SPIE Conf. Applications of Digital Image Processing XXI ‘Compression of volumetric medical data based on cube-splitting’. *Proc. SPIE Conf. Applications of Digital Image Processing XXIII*, (SPIE Conf. Applications of Digital Image essing XXIII) July-Aug. 2000. 4115 p. .

[Cai (2006)] ‘Constant Quality Aimed Bit Allocation for 3D Wavelet Based Video Coding’. Jianfei Cai . 10.1109/ICME.2006.262866. *IEEE International Conference on*, 2006. July 2006. 9 p. . (Multimedia and Expo)

[Zhang et al. (2003)] ‘Constant quality constrained rate allocation for FGS-coded video’. X M Zhang , A Vetro , Y Q Shi , H Sun . *IEEE Trans. Circuits Syst. Video Technol* Feb. 2003. p. .

[Munteanu et al. ()] ‘Control of the distortion variation in video coding systems based on motion compensated temporal filtering’. A Munteanu , Y Andreopoulos , M Van Der Schaar , J Cornelis . *Proceedings of the ICIP*, (the ICIP) 2003. 2 p. .

[Seran and Kondi (2006)] ‘Distortion fluctuation control for 3D wavelet based video coding’. V Seran , L P Kondi . *Proceedings of SPIE VCIP*, (SPIE VCIP) Jan. 2006.

[Bajic and Woods (2003)] ‘Domain-Based Multiple Description Coding of Images and Video’. I V Bajic , J W Woods . *IEEE Transactions on Image Processing* Oct. 2003. 12 (10) .

[Ho and Lun (2002)] ‘Efficient wavelet based temporally scalable video coding’. K Ho , D Lun . *Proc. Int. Conf. Image Process*, (Int. Conf. Image essNew York, USA) Aug. 2002. p. .

[Taubman et al. (1994)] ‘Embedded video coding using motion compensated 3-D subband/wavelet filter bank’. D Taubman , A T Zakhor ; S , J W Hsiang , Woods . *Proceedings of the Packet Video Workshop*, (the Packet Video WorkshopSardinia, Italy) Sept. 1994. May 2000. 3 p. . (Multirate 3-D subband coding of video)

[Li (2009)] ‘Error Protection of 3D Wavelet Video Streaming Using Wyner-Ziv Video Coding for Lossy Network Transmission’. Zhe Li . 10.1109/MINES.2009.207. *Multimedia Information Networking and Security, 2009. MINES ’09. International Conference on*, 18-20 Nov. 2009. 2 p. .

[Xiong et al. (2004)] ‘Exploiting temporal correlation with block-size adaptive motion alignment for 3D wavelet coding’. R Xiong , F Wu , S Li , Z Xiong , Y.-Q Zhang . *SPIE VCIP2004*, (San Jose) Jan.2004. 5308 p. .

[Ye and Van Der Schaar ()] ‘Fully scalable 3D overcomplete wavelet video coding using adaptive motion compensated temporal filtering’. J C Ye , M Van Der Schaar . *Proc. VCIP2003*, (VCIP2003Lugano, Switzerland) 2003. 5150 p. .

[Taubman (2000)] ‘High performance scalable image compression with EBCOT’. D Taubman . *IEEE Trans. Image Processing* July 2000. 9 p. .

[Wu] ‘High-order context modeling and embedded conditional entropy coding of wavelet coefficients for image compression’. X Wu . *Proc. 31st*, (31st)

[Hybrid coding of video with spatio-temporal scalability using subband decomposition Proc. SPIE ()] ‘Hybrid coding of video with spatio-temporal scalability using subband decomposition’. *Proc. SPIE*, (SPIESan Jose, CA) 1999. 3653 p. .

[Doman´ski et al. (1998)] ‘Hybrid coding of video with spatio-temporal scalability using subband decomposition’. M Doman´ski , A ?uczak , S Mac´kowiak , R S´ . *Proc. Signal Processing IX: Theories and Applications*, (Signal essing IX: Theories and ApplicationsRhodes, Greece) Sept. 1998. p. .

[Yuan and Monro] ‘Improved matching pursuits image coding’. Y Yuan , D M Monro . *Proc. IEEE Int. Conf*, (IEEE Int. Conf)

[Andreopoulos et al. (2004)] ‘In-band motion compensated temporal filtering’. Y Andreopoulos , A Munteanu , J Barbarien , M Van Der Schaar , J Cornelis , P Schelkens . *Signal Processing: Image Communication (special issue on ”Subband/Wavelet Interframe Video Coding”)*, August 2004. 19 p. .

[Pennebaker and Mitchell ()] *JPEG Still Image Data Compression Standard*, W B Pennebaker , J L Mitchell . 1993. New York: Van Nostrand Reinhold.

[Christopoulos ()] 'JPEG2000 Verification Model 8.5'. C Christopoulos . N1878. *ISO/IEC JTC1/SC29/WG1*, 2000. (Report)

[Kim and Pearlman (1999)] 'Lossless volumetric medical image compression'. Y Kim , W A Pearlman . *Proc. SPIE Conf. Applications of Digital Image Processing XXII*, (SPIE Conf. Applications of Digital Image essing XXII) July 1999. 3808 p. .

[Kim and Pearlman ()] 'Low-delay embedded 3-Dwavelet color video coding with SPIHT'. B.-J Kim , . A Pearlman . *Proc. SPIE*, (SPIE) 1998. 3309 p. .

[Luo et al. ()] 'Motion compensated lifting wavelet and its application in video coding'. L Luo , J Li , S Li , Z Zhuang , Y.-Q Zhang . *Proc. IEEE International Conference on Multimedia and Expo*, (IEEE International Conference on Multimedia and ExpoTokyo, Japan) 2001. p. .

[Secker and Taubman ()] 'Motion-compensated highly scalable video compression using an adaptive 3D wavelet transform based on lifting'. A Secker , D Taubman . *Proc. IEEE Int. Conference on Image Processing*, (IEEE Int. Conference on Image essingThessaloniki, Greece) 2001. 2 p. .

[Rahmoune et al. (2004)] 'MP3D: highly scalable video coding scheme based on matching pursuit'. A Rahmoune , P Vandergheynst , P Frossard . *IEEE Int. Conf. Acoustics Speech Signal Process* May 2004. (ICASSP 2004)

[Schelkens ()] 'Multidimensional wavelet codingalgorithms and implementations'. P Schelkens . *Ph.D. dissertation, Dept. Electron. Inform. Processing (ETRO)* 2001. Vrije Univ. Brussel

[Franchi et al. (2003)] 'Multiple Description Coding For Scalable and Robust Transmission over IP'. N Franchi , M Fumagalli , R Lancini . *Proc. of 13th Packet Video Workshop*, (of 13th Packet Video WorkshopNantes, France) Apr. 2003.

[Mavlankar and Steinbach (2005)] 'Multiple Description Video Coding Using Motion-Compensated Lifted 3D Wavelet Decomposition'. A Mavlankar , E Steinbach . 10.1109/ICASSP.2005.1415342. *IEEE International Conference on*, 2005. March 18-23. 2005. 2 p. . (Proceedings. (ICASSP '05)

[Cheng et al. ()] 'Multiscale video compression using wavelet transform and motion compensation'. P.-Y Cheng , J Li , C.-C Kuo . *Proc. Int. Conf. Image Processing*, (Int. Conf. Image essing) 1995. 1 p. .

[Seran and Kondi (2006)] 'New Scaling Coefficients for Biorthogonal Filter to Control Distortion Variation in 3D Wavelet Based Video Coding'. V Seran , L P Kondi . 10.1109/ICIP.2006.313101. *IEEE International Conference on*, 2006. 8-11 Oct. 2006. p. .

[Yea and Pearlman ()] 'On scalable lossless video coding based on sub-pixel accurate Year MCTF'. Sehoon Yea , William A Pearlman . *Proceedings of SPIE –Volume 6077 Visual Communications and Image Processing*, (SPIE –Volume 6077 Visual Communications and Image Processing) 2006.

[Breger (1972)] 'Optimal quantizers and permutation codes'. T Breger . *IEEE Trans.Inform. Theory* Nov. 1972. 18 p. .

[Sun et al. (2005)] 'Perceptually adaptive rate-distortion optimization for variable block size motion alignment in 3D wavelet coding'. Y Sun , F Pan , A A Kassim . 10.1109/ICASSP.2005.1415558. *IEEE International Conference on*, 2005. March 2005. 2 p. . (Proceedings. (ICASSP '05)

[Xiong et al. (1999)] 'Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transform'. Z Xiong , X Wu , D Y Yun , . A Pearlman . *Proc. SPIE Conf. Visual Communications*, (SPIE Conf. Visual Communications) Jan. 1999. 3653 p. .

[Tzou (1987)] 'Progressive image transmission: A review and comparison of techniques'. K H Tzou . *Opt. Eng* July 1987. 26 p. .

[Lee et al. ()] 'Quantization of 3-D-DCT coefficients and scan order for video compression'. M C Lee , R K W Chan , D A Adjeroh . *J. Vis. Commun. Image Representation* 1997. 8 p. .

[Tan et al. ()] 'Real Time Software Implementation of Scalable Video Codec'. W Tan , E Chang , A Zakhor . *Proc. IEEE Int. Conference on Image Proc*, (IEEE Int. Conference on Image) 1996. 1 p. .

[Apostolopoulos (2001)] 'Reliable Video Communication over Lossy Packet Networks using Multiple State Encoding and Path Diversity'. J G Apostolopoulos . *Proc. Visual Communications and Image Processing*, (Visual Communications and Image essing) Jan. 2001. p. .

[Rath and Guillemot (2006)] 'Representing laplacian pyramids with varying amount of redundancy'. G Rath , C Guillemot . *Signal Process. Conf*, (Florence, Italy) Sep. 2006. (presented at the Eur)

[Adams and Faouzi (2000)] 'Reversible integer-tointeger wavelet transforms for image compression: Performance evaluation and analysis'. M Adams , K Faouzi . *IEEE Trans. Image Processing* June 2000. 9 p. .

[Mavlankar (2004)] *Robust Video Communication using motion-compensated lifted 3-D wavelet coding*, A A Mavlankar . Sept. 2004. Technische Universität München (Master's Thesis)

[Wang et al. ()] 'Robust video transmission over a lossy network using a distributed source coded auxiliary channel'. J Wang , A Majumdar , K Ramchandran , H Garudadri . *Proc. Picture Coding Symposium(PCS)*, (Picture Coding Symposium(PCS)) 2004.

[Zhao (2008)] 'Scalable 3D wavelet video coding scheme'. Chen-Wei Deng; Bao-Jun Zhao . 10.1109/ICOSP.2008.4697070. *ICSP 2008. 9th International Conference on*, 2008. 26-29 Oct. 2008. p. .

[Lilienfield and Woods ()] 'Scalable high definition video coding'. G Lilienfield , J Woods . *Proc. SPIE Visual Communication and Image Processing*, (SPIE Visual Communication and Image essingSan Jose, CA) 1998. p. .

[Benzler ()] 'Scalable multi-resolution video coding using a combined subband-DCT approach'. U Benzler . *Proc. Picture Coding Symp*, (Picture Coding SympPortland, OR) 1999. p. .

[Benzler] 'Scalable multiresolution video coding using sub-band decomposition'. U Benzler . *Proc. 1st*, (1st)

[Golwelkar and Woods ()] 'Scalable video compression using longer motion compensated temporal filters'. A Golwelkar , J Woods . *Proc. SPIE VCIP*, (SPIE VCIP) 2003. 5150 p. .

[Li (2004)] 'Scalable video compression via overcomplete motion compensated wavelet coding'. Xin Li . *Signal Processing: Image Communication (special issue on"Subband/Wavelet Interframe Video Coding")*, August 2004. 19 p. .

[Scalable Video Model 2.0 (2004)] *Scalable Video Model 2.0*, ISO/IEC JTC 1/SC 29/WG 11N6520. July 2004. Redmond, WA, USA.

[Scalable Video Model v2.0 ISO/IEC JTC1/SC29/ WG11 69th MPEG Meeting (2004)] 'Scalable Video Model v2.0 ISO/IEC JTC1/SC29/ WG11'. N6520. *69th MPEG Meeting*, (Redmond, WA, USA) Jul. 2004. (Tech. Rep.)

[Adami et al. (2007)] 'State-of-the-art and trends in scalable video compression with wavelet-based approaches'. Nicola Adami , Alberto Signoroni , Riccardo Leonardi . *IEEE Trans. Circuits Syst. Video Technol* Sep. 2007. 17 (9) p. .

[Kim and Pearlman (2000)] 'Stripe-based SPIHT lossy compression of volumetric medical images for low memory usage and uniform reconstruction quality'. Y S Kim , . A Pearlman . *Proc. ICASSP*, (ICASSP) June 2000. 4 p. .

[Chui et al. (1998)] *System and Method for Nested Split Coding of Sparse Data Sets*, C K Chui , R Yi , Teralogic Inc . US:005 748 116A. May 1998. Menlo Park, California.

[Aaron et al. ()] 'Systematic Lossy Forward Error Protection for Video Waveforms'. A Aaron , S Rane , D Rebollo-Monedero , B Girod . *Proc. IEEE Int. Conf. Image Proc*, (IEEE Int. Conf. Image) 2003.

[Katata et al. (1997)] 'Temporal-Scalable Coding Based on Image Content'. H Katata , N Ito , H Kusao . *IEEE Trans. Circuits and Systems for Video Tech* Feb. 1997. 7 (1) p. .

[Ohm (1994)] 'Three-Dimensional Subband Coding with Motion Compensation'. J.-R Ohm . *IEEE Trans. Image Processing* Sept. 1994. 3 (5) p. .

[Ohm (1994)] 'Three-dimensional subband coding with motion-compensation'. R Ohm . *IEEE Trans. Image Processing* Sept. 1994. 3 p. .

[Choi and Woods ()] 'Three-dimensional subband/ wavelet coding of video with motion compensation'. S.-J Choi , J W Woods . *Proc. SPIE Visual Communications and Image Processing*, (SPIE Visual Communications and Image essing) 1997.

[Bilgin et al. (2000)] 'Threedimensional compression with integerwavelet transform'. A Bilgin , G Zweig , M W Marcelin . *Appl. Opt* Apr. 2000. 39 p. .

[Xu et al. ()] *Threedimensional embedded subband coding with optimized truncation (3D ESCOT)", Applied and Computational Harmonic Analysis*, J Xu , Z Xiong , S Li , Y.-Q Zhang . 2001. p. .

[Pesquet-Popescu and Bottreau ()] 'Threedimensional lifting schemes for motion compensated video compression'. B Pesquet-Popescu , V Bottreau . *Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing*, (IEEE Int. Conference on Acoustics, Speech and Signal essingSalt Lake City, UT) 2001. 3 p. .

[Ch et al. (1995)] 'Threedimensional subband coding of video'. Ch , N Podilchuk , N Jayant , Farvardin . *IEEE Trans. Image Processing* Feb. 1995. 4 p. .

[Rusert et al. ()] 'Transition filtering and optimized quantization in interframe wavelet video coding'. T Rusert , K Hanke , J R Ohm . *Proc. Visual Communication and Image Processing*, (Visual Communication and Image essing) 2003. p. .

[Wheeler ()] 'Trellis source coding and memory constrained image coding'. F W Wheeler . *Dept. Elect., Comput. Syst. Eng., Renselaer Polytech. Inst* 2000. (Ph.D. dissertation)

[Neff and Zakhor ()] 'Very low bit rate video coding based on matching pursuits'. R Neff , A Zakhor . *IEEE Trans. Circuits and Systems for Video Tech* 1997. 7 p. .

[Schelkens et al. (2000)] 'Volumetric data compression based on cubesplitting'. P Schelkens , J Barbarien , J Cornelis . *Proc. 21st Symp. Information Technology in the Benelux*, (21st Symp. Information Technology in the Benelux) May 2000. 21 p. .

[Shen and Delp (1999)] 'Wavelet based rate scalable video compression'. K Shen , E Delp . *IEEE Trans. Circuits Syst. Video Technol* Feb. 1999. 9 p. .

[Xiong et al. (1998)] 'Wavelet packet image coding using spacefrequency quantization'. Z Xiong , K Ramchandran , M T Orchard . *IEEE Trans. Image Processing* June 1998. 7 p. .

[Int ()] *WorkshopWireless Image/Video Communication*, Int . 1996. Loughborough, U.K.. p. .

[Xu] Ke Xu . *Sheng Fang*,