Sheffield Hallam University

Quality scalability aware watermarking for visual content

BHOWMIK, Deepayan <http://orcid.org/0000-0003-1762-1578> and ABHAYARATNE, Charith

Available from Sheffield Hallam University Research Archive (SHURA) at:

https://shura.shu.ac.uk/13806/

Citation:

Quality Scalability Aware Watermarking for Visual Content

Deepayan Bhowmik and Charith Abhayaratne, *Member, IEEE*

Abstract—Scalable coding-based content adaptation poses serious challenges to traditional watermarking algorithms, which do not consider the scalable coding structure and hence cannot guarantee correct watermark extraction in media consumption chain. In this paper, we propose a novel concept of scalable blind watermarking that ensures more robust watermark extraction at various compression ratios while not effecting the visual quality of host media. The proposed algorithm generates scalable and robust watermarked image code-stream that allows the user to constrain embedding distortion for target content adaptations. The watermarked image code-stream consists of hierarchically nested joint distortion-robustness coding atoms. The code-stream is generated by proposing a new wavelet domain blind watermarking algorithm guided by a quantization based binary tree. The code-stream can be truncated at any distortion-robustness atom to generate the watermarked image with the desired distortion-robustness requirements. A blind extractor is capable of extracting watermark data from the watermarked images. The algorithm is further extended to incorporate a bit-plane discarding-based quantization model used in scalable coding-based content adaptation, e.g., JPEG2000. This improves the robustness against quality scalability of JPEG2000 compression. The simulation results verify the feasibility of the proposed concept, its applications, and its improved robustness against quality scalable content adaptation. Our proposed algorithm also outperforms existing methods showing 35% improvement. In terms of robustness to quality scalable video content adaptation using Motion JPEG2000 and wavelet-based scalable video coding, the proposed method shows major improvement for video watermarking.

Index Terms—Scalable watermarking, JPEG2000, content adaptation, distortion, robustness.

I. INTRODUCTION

SCALABLE coding has become a de facto functionality in recent image and video coding schemes, *e.g.*, JPEG2000 [1] for images; scalable extensions of Advanced Video Coding (AVC) [2] and High Efficiency Video

D. Bhowmik is with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, U.K. (e-mail: d.bhowmik@hw.ac.uk).

C. Abhayaratne is with the Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield, S1 3JD, U.K. (e-mail: c.abhayaratne@sheffield.ac.uk).
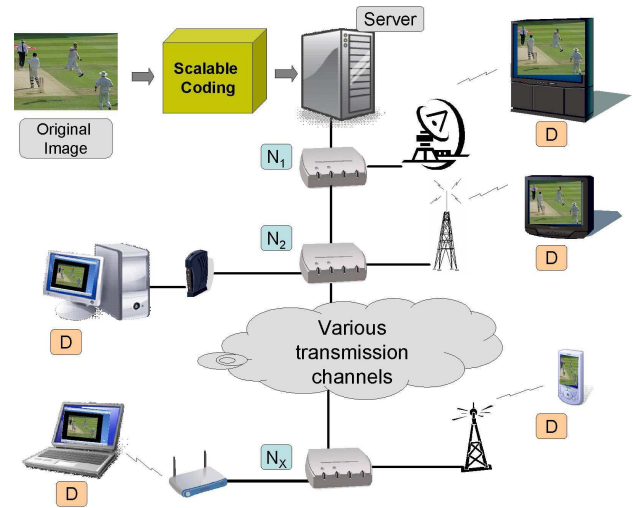
Fig. 1.   Multimedia usage scenarios using scalable coded content.

Coding (HEVC) [3] for video. The scalable coders produce scalable bit streams representing content in hierarchical layers of increasing audiovisual quality and increasing spatio-temporal resolutions. Such bit streams may be accordingly truncated in order to satisfy variable network data rates, display resolutions, display device resources and usage preferences. These adapted bit streams may be transmitted or further adapted or decoded using a universal decoder for playback. An example of scalable coding-based multimedia usage is shown in Fig. 1. In scalable image/video coding, the input media is coded in a way that the main host server keeps bit streams that can be decoded to the highest quality and to the full resolution of the content. When the content needs to be delivered to a less capable display (*D*) or via a lower bandwidth network, the bit stream is adapted at different nodes (N_1, N_2, ..., N_x, as shown in Fig. 1) using different scaling parameters to match those requirements. At each node the adaptation parameters may be different and a new bit stream may be generated for decoding.

Such bit stream truncation-based content adaptations also affect any content protection data, *i.e.*, watermarks, embedded in the original content. Watermarking for scalable coded content is far more challenging than traditional watermarking schemes [4]–[20], where the performance of the algorithms is measured by traditional metrics, such as, *1)* distortion due to embedding (commonly used metrics are Peak Signal to Noise Ratio (PSNR), weighted PSNR or more recently Structured Similarity Index Measurement (SSIM) [21]), *2)* data capacity

and *3)* robustness against signal processing, geometric or compression attacks as listed in Stirmark [22] or Checkmark benchmark [23]. In these algorithms, the watermark is embedded with the goal of minimising embedding distortion while trying to maximise the robustness for a given attack.

Many of these algorithms demonstrated high robustness against traditional non-scalable compression schemes, such as, JPEG and MPEG-2. It is also common that most of these watermarking schemes are based on such compression schemes. Often the same watermarking algorithms have been naively extended to propose watermarking algorithms robust against scalable content adaptation attacks, *e.g.*, JPEG2000. Example of such image watermarking algorithms *e.g.*, [7], [11]–[15], [24]–[30]. As the discrete wavelet transform (DWT) is the underlying technology of the JPEG2000 compression standard, we consider the watermarking schemes that use wavelet-based techniques in addition to the algorithms mentioned above. In these schemes, different approaches have been used as follows:

  a) *Choosing Coefficients in a Specific Subband for Embedding the Watermark: e.g.*, embedding in high frequency subbands for better imperceptibility [7]–[10]; embedding in low frequency subband to achieve high robustness [11], [12]; and balancing imperceptibility and robustness with all subbands spread spectrum embedding [13], [14].
  b) *Using Different Wavelet Kernels: e.g.*, Haar or other Daubechies family orthogonal wavelets [8]–[11], [15] and biorthogonal wavelets [12].
  c) *Optimising the Host Coefficient Selection: e.g.*, choosing all coefficients in a subband [8], [13]; using a threshold based on their magnitude significance [7]; the just noticeable difference(JND) [12]; a mask based on the Human Visual System (HVS) model [8]; a fusion rule-based mask for refining the selection of host coefficients [15] and blind re-quantization of a coefficient with respect to a group of coefficients within a given window [9]–[11], [14].

An in-depth review of the above schemes with respect to robustness to scalable content adaptation can be found in [24]. None of these algorithms can guarantee correct watermark extraction from the adapted (scaled) content, as they do not have access to scaling information during embedding. In scalable coding, the original media is encoded once and then the resulting bit stream is truncated with any scaling parameter during the content adaptation steps in media consumption chain to obtain the target data rates and resolutions. This creates a situation where no information with respect to target data rates is available to the watermark embedder to make watermarking robust for such content adaptations. While traditional watermarking methods are robust enough to cope with content adaptation, in such approaches as the watermark is being embedded once which is predetermined to be robust to high compression, the possibility of embedding distortion is also higher and not always necessary where target application requires higher resolution content, *i.e.*, less compression.

Different from traditional schemes, we propose the novel concept of scalable watermarking by creating nested

Un-watermarked: CR=2, PSNR=41.96;   Watermarked: CR=2, PSNR=39.93, *H*=0.08;

Un-watermarked: CR=50, PSNR=21.61;   Watermarked: CR=50, PSNR=21.47, *H*=0.08

Fig. 2. Application example where distortion due to watermark embedding is adaptively negated at various JPEG2000 compression ratio (CR) without compromising robustness. *H* stands for Hamming Distance ($H \in 0..1$), a commonly used metric of watermarking robustness. Lower value of *H* ensures higher robustness.

distortion-robustness atoms (defined in Section II-B1) to allow flexibility to the user. The proposed algorithm *encodes* watermarks hierarchically and *embeds* it at the point of scalable compression. The watermarking scalability here refers to hierarchical watermark embedding where more embedded information corresponds to better robustness. The concept of scalable watermarking is particularly useful in watermarking for scalable coded images where the watermark can also be scaled according to the heterogenous network capacity and the end user's requirement for a target application where it is assumed that the media life-cycle finishes. For example, for a high bandwidth network and a high resolution display, highly imperceptible but less robust watermarked image can be transmitted. As in this scenario, high quality media is desirable and the watermark can also be extracted reliably due to lower compression. Whereas, for a low network bandwidth and low resolution display, the distribution server can choose a highly robust watermarking stream, where, due to higher compression the watermarking imperceptibility is less important, but high robustness is required for reliable watermark extraction (refer to Fig. 2). Similarly, based on any other combinations of the network's capability and user's requirement, the scalable watermarked media code stream can be truncated and distributed accordingly.

With the increased use of scalable coded media, scalable watermarking is very important. To the best knowledge of the authors, little work has been proposed in the current literature in this context. Most such algorithms commonly available are proposed either as a joint progressive scalable compression and watermarking scheme [25], [26] or efficient coefficient selection methods which are robust against resolution or quality scalable attacks [27], [28]. These algorithms are primarily

focused on two main robustness issues [29]: *1)* detection of the watermark after an acceptable scalable compression and *2)* graceful improvement of extracted watermark as more quality or resolution layers are received at the image decoder. The extraction of a complete watermark is only possible when all quality layers are available at the detection. In a practical scenario quality layers are not always available to the end user (as discussed earlier) and this poses risk of loosing important watermark information.

Distinctively, we propose a novel scalable watermarking concept resulting in a distortion constrained watermarked code-stream to generate watermarked image with desired distortion robustness requirements. Therefore the extraction of complete watermark is ensured at various stages of scalable compression. This work addresses the tow-fold problem of *1)* obtaining the least distortion at a given watermark embedding rate and *2)* achieving the best robustness in a scalable fashion by hierarchically encoding lower and higher embedded distortion-robustness atoms, respectively. In designing the algorithm, we have considered the propositions for embedding distortion, *i.e.*, in order to minimize the distortion, the coefficient modification must be minimized; and the concept of bit plane discarding model [30] to emulate scalable content adaptation for improving the robustness against quality scalable content adaptation. The main contributions of this work are as follows:

- The theoretical foundation for generating a code-stream that can be truncated at any distortion-robustness atom level to generate the watermarked image with the desired distortion-robustness requirements;
- A novel watermarking algorithm that generates scalable embedded code-stream using hierarchically nested joint distortion-robustness atoms; and
- Improving watermarking robustness by modelling the bit plane discarding used in quality scaling based content adaptation in scalable coded image and video.

These contributions are demonstrated by proposing a new wavelet domain binary tree guided rules-based blind watermarking algorithm. Following the scalable coding concepts, a universal blind watermark extractor is proposed. It is capable of extracting watermark data from the watermarked images created using any truncated code stream. As no such idea has been explored yet in the literature, in order to evaluate this work, we introduce a new embedding distortion metric shown in Eq. (3) and report the robustness results to support the claim. The initial concept and early results of this algorithm were reported as conference publications in [31] and [32]. This work reports the detailed algorithm, discusses extended results with application to both image and video watermarking. Within the scope of this work, we focus on the scalable watermarking concept and restrict ourselves to the robustness against quality scalability.

II. SCALABLE WATERMARKING

A. The Proposed Algorithm

In proposing the new algorithm we aim to address two significant challenges related to robust watermarking
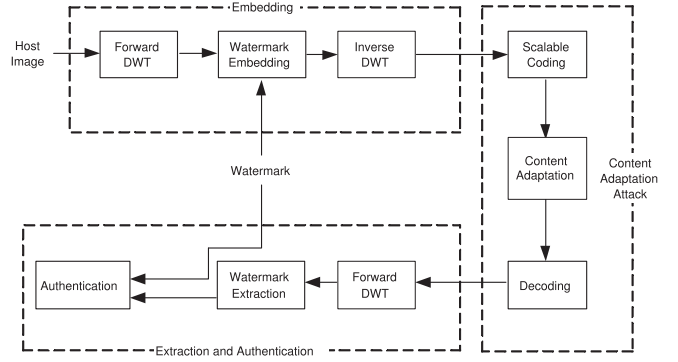

Fig. 3.   Algorithmic block diagram of wavelet based watermarking and scalable content adaptation attack.

techniques for scalable coded image and video: *1)* scalability of the watermarking and *2)* robustness against scalable media compression. As opposed to traditional algorithms which fail to comply the scalability requirements, we introduce a new watermarking algorithm that creates hierarchical watermarked image / video code-stream distortion-robustness atoms and allows quantitative embedding-distortion measurement at individual distortion-robustness atom level. The system block diagram of the proposed watermarking algorithm and the scalable content adaptation scheme is shown in Fig. 3. The major steps for embedding include the forward discrete wavelet transform (FDWT) and coefficient modification using the embedding algorithm followed by the inverse discrete wavelet transform (IDWT). The content is then scalable coded and may be adapted during the media consumption. Watermark authentication includes the FDWT and recovery of the watermark and comparison with the original watermark. The proposed embedding algorithm in this paper follows a non-uniform quantization based index modulation and the process is divided into three parts: *1)* Quantized binary tree formation, *2)* embedding by index modulation and *3)* extraction & authentication. The embedding and extraction were performed on wavelet domain and therefore we use the term *coefficients* referring to wavelet coefficients in the rest of the paper.

1) Quantized Binary Tree Formation: This step defines how a coefficient (C) chosen for embedding a watermark bit is recursively quantized to form a binary tree. The coefficients for embedding the watermark data may be chosen based on its magnitude, sign, texture information, randomly or any other selection criteria. While the selection criteria is a user defined parameter, we have chosen all coefficients for the experiments in this manuscript.

Firstly, C is indexed (b_i) as 0 or 1 using an initial quantizer λ:

$$b_i = \left\lfloor \frac{|C|}{\lambda/2^i} \right\rfloor mod\ 2, \quad i \in 0, 1, 2, 3..., \qquad (1)$$

where mod denotes the modulo operation. Assuming $n = \left\lfloor \frac{|C|}{\lambda} \right\rfloor$, we can identify the position of C between the quantized cluster (n) - ($n + 1$), which can alternatively be described as bit plane clusters as shown in Fig. 4. The coefficient, C, is then further quantized more precisely within
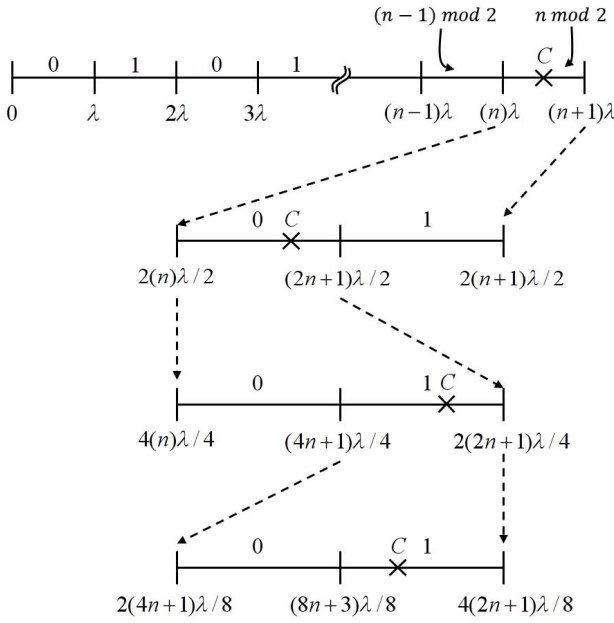
Fig. 4. The hierarchical quantizer in formation of a binary tree.

Fig. 5. An example of a binary tree.

TABLE I
TREE-BASED WATERMARKING RULES TABLE

Symbol	Binary Tree	Watermark Association
Embedded Zero (EZ)	000xxxx	0
Embedded Zero (EZ)	001xxxx	0
Cumulative Zero (CZ)	010xxxx	0
Weak One (WO)	011xxxx	1
Weak Zero (WZ)	100xxxx	0
Cumulative One (CO)	101xxxx	1
Embedded One (EO)	110xxxx	1
Embedded One (EO)	111xxxx	1

a smaller cluster using a smaller quantizer, $\lambda/2$, and the corresponding index is computed as: $b_1 = \left\lfloor \frac{|C|}{\lambda/2} \right\rfloor mod\ 2$. The index tree formation is continued recursively by scaling λ value by 2, as long as the condition $\lambda/2^i \geq 1$ is true. During this tree formation process signs of the coefficients are preserved separately. Based on the calculated index value at various quantization steps a binary tree ($b(C)$) of each selected coefficient can easily be formed:

$$b(C) = (b_0)(b_1)....(b_{i-1})(b_i), \qquad (2)$$

where (b_0), (b_1)...(b_i) are binary values (bits) in the most significant bit (MSB) to the least significant bit (LSB) positions, respectively with the tree depth $i+1$. For example, if $C = 135$ and initial $\lambda = 30$, the binary tree $b(C)$ will be $b(C) = 01000$. An illustration of the tree formation scenario is shown in Fig. 5. The number of tree nodes, *e.g.*, number of bits in any binary tree is decided by the initial quantizer λ and defined as the *depth* of the tree.

2) Embedding by Index Modulation: The above mentioned binary tree in Section II-A.1 is used to embed binary watermark information based on symbol-based embedding rules. To introduce the watermarking scalability, we chose 3 most significant bits which represents 8 different states corresponding to 6 different symbols. Although any other number of bits (> 1) can be chosen, the use of more number of bits (> 3) results in more states, thus increase the complexity while less number of bits (< 3) reduces the watermark scalability. Now 3 most significant bits of any binary tree, represents 6 symbols (*EZ = Embedded Zero, CZ = Cumulative Zero, WZ = Weak Zero, EO = Embedded One, CO = Cumulative One* and *WO = Weak One*) to identify the original coefficient's association with a 0 or 1. The rational of allocating symbols associated with 0 or 1 relies on the total number of 0s or 1s in the symbol, *i.e.*, two or more 0s correspond to 0 and
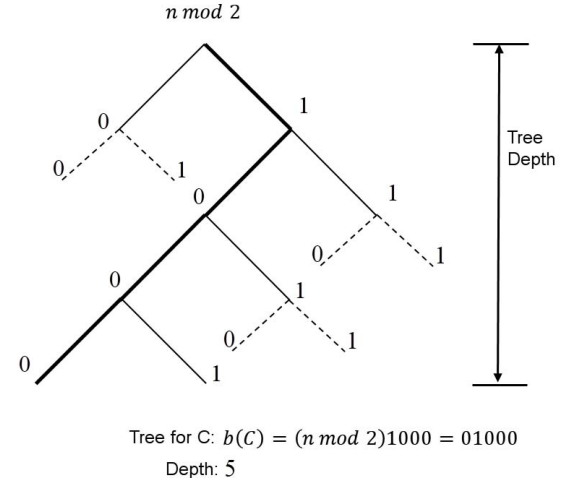
vice versa for 1s. The bits in a binary tree, symbols and the corresponding associations are shown in Table I for a tree *depth* of 7. Depending on the input watermark stream, if required, a new association is made by altering the chosen 3 most significant bits in the tree to reach the nearest symbol as shown in the state diagram in Fig. 6. Assuming the current state of the binary tree is *EZ*, to embed the watermark bit 0 no change in state is required while to embed the watermark bit 1, a new value of the binary tree must be assigned. The new value of the tree can be associated with either *WO* or *CO* or *EO*. However to minimize the distortion, the nearest state change must occur as shown in the state diagram. Other state changes in the binary tree follow the same rule. Finally, the watermarked image / video is obtained by de-quantizing the modified binary tree followed by an inverse transformation. For example, if a modified binary tree $b(C) = 101101$ of depth $d = 6$, the embedded coefficient will be $C' = 1 * 2^5 + 0 * 2^4 + 1 * 2^3 + 1 * 2^2 + 0 * 2^1 + 1 * 2^0 = 45$. Sign of the coefficients are kept separately and added once the coefficients are reconstructed.

However, quantifying embedding distortion for the proposed method is challenging with any of the traditional embedding measurement metrics, such as, PSNR or SSIM. This is due to the dynamic nature of the algorithm where depending on the scaling parameters the distortion measurement may vary for a given data capacity on the same coefficient. Traditional metrics are generic for many image processing applications
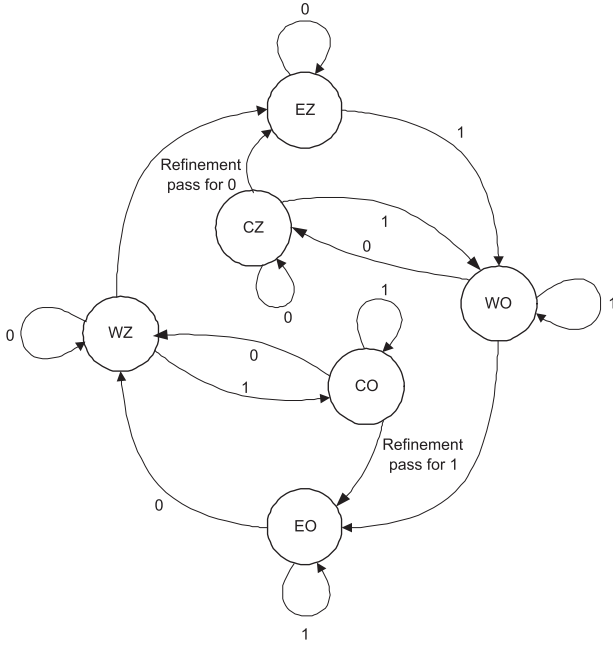
Fig. 6. The state diagram of watermark embedding rule based on tree-symbol-association model.

Fig. 7. The proposed scalable watermarking layer creation.

and do not consider the data capacity in their calculation resulting in ambiguity to the user whether the distortion is due to low capacity (randomly chosen pixels or coefficients) but high strength watermark or high capacity (all coefficients/pixels) but low strength watermark. Therefore we propose a new metric which combines the data capacity and embedded distortion as follows:

$$\Phi = \frac{\sum_{m=0}^{X-1}\sum_{n=0}^{Y-1}(I(m,n) - I'(m,n))^2}{L}, \qquad (3)$$

where Φ represents embedding distortion rate, I and I' are the original and watermarked image, respectively with dimensions $X \times Y$ and L is the number of watermark bits embedded, *e.g.*, data capacity. For the completeness we compared our new metric Φ with traditional PSNR values and the results are discussed in Section IV.

3) Extraction & Authentication: A universal blind extractor is proposed for watermark extraction and authentication process. The term *universal* signifies the capability to extract watermark information irrespective of watermark embedding or content adaptation parameters. The wavelet coefficients are generated using the FDWT on the host image followed by the tree formation process used during embedding. Based on the recovered tree structure, symbols are re-generated to decide on extraction of a watermark bit, 0 or 1. The extracted watermark is then authenticated by comparing the Hamming distance, $H \in 0..1$ with the original watermark (often referred as Bit Error Rate in the literature) as described in Eq. (4):

$$H(W, W') = \frac{1}{L}\sum_{i=0}^{L-1} W_i \oplus W_i', \qquad (4)$$

where W and W' are the original and the extracted watermarks, respectively. L is the length of the sequence and $\oplus$ represents
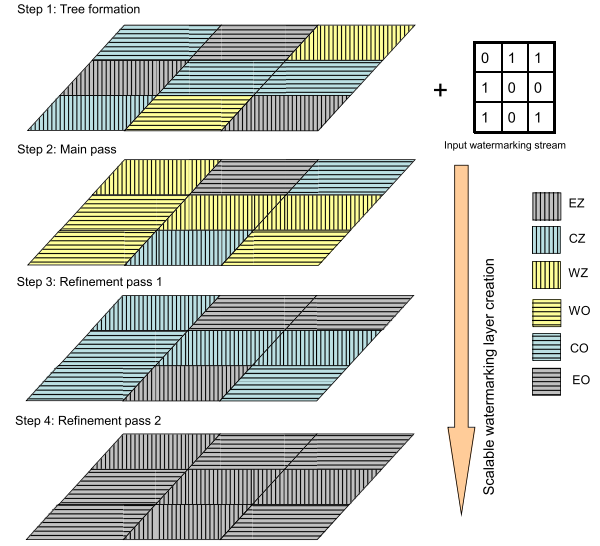
the *XOR* operation between the respective bits. A lower value of Hamming distance corresponds to higher robustness.

B. Designing Watermarking Scalability

This section discusses on obtaining the scalable watermark using the binary tree presented in Section II-A. The term *watermarking scalability* refers to embedding of watermarks in a hierarchical manner, so that, more embedding information leads to better robustness. The proposed algorithm is independent of any specific media coding scheme and hence can be used to design a new joint scalable watermarking - image / video coding-decoding scheme.

In the proposed algorithm, the symbols shown in Table I are ranked based on the improvement in robustness associated to them. The MSB in the binary tree corresponds to coarse-grained quantization index, whereas, the LSB represents fine-grained quantization index. To extract the watermark bit successfully, all three most significant bits of any binary tree must be unaltered in case of *WO*, *CO* or *WZ*, *CZ*, whereas only two most significant bits are required for *EO* or *EZ*. Therefore, two consecutive 0s (*EZ*) or 1s (*EO*) provide the strongest association with 0 or 1, respectively and hence provides high robustness. The symbol pair, *WO* and *CO*, offers the same level of robustness. Similarly, the robustness level associated with the symbol pair, *WZ* and *CZ*, is the same. Thus, the robustness ranking of the symbols can be defined as *EO>CO, WO*; and *EZ>CZ, WZ*; for embedding 1s and 0s, respectively. At the same time, the collective embedding distortion rate, Φ, for these scenarios can be computed as in Eq. (3). In designing the scalable watermarking concept, rest of this section exploits these two properties: *a)* rule based robustness rank order and *b)* embedding distortion rate Φ. The complete process is divided into three separate modules: *1)* Encoding module, *2)* Embedded watermarking module and *3)* Extractor module.

1) Encoding Module: The main functionality of this module is to generate a hierarchical embedded code-stream. An example scalable watermarking system model is shown in Fig. 7.

The sequential activities within the encoding module are described in the following steps:

Tree formation: Binary trees are formed as in Section II-A for each coefficient selected for embedding the watermark. Every tree is now assigned a symbol according to the rules in Table I.

Main pass: Based on the input watermark stream, the trees are altered to create right association for required robustness as shown in Fig. 6. Hence, all selected coefficients are rightly associated at least with basic WZ/WO symbol. Thus, we name it *base layer*. The embedding distortion is calculated progressively at each level in the tree.

Refinement passes: The goal of the refinement passes is to increase the watermarking strength progressively to increase the robustness progressively. The base layer provides basic minimum association with watermark bits. In this refinement pass, the watermarking strength is increased by modifying the symbols and corresponding tree to the next available level, *i.e.*, $WZ \rightarrow EZ$, $CZ \rightarrow EZ$, $WO \rightarrow EO$ and $CO \rightarrow EO$, as shown in the state diagram in Fig. 6. At the end of this pass, all trees are modified and associated with the strongest watermark embedding EZ/EO. Similar to the previous step, the distortion is calculated as the refinement levels progress.

Hierarchical distortion-robustness atom and code-stream generation: During the previously described passes, the binary trees are modified according to the input watermark-robustness association and the embedding distortion is calculated at each individual tree. Here we call these individual trees or a group of trees as a *distortion-robustness atom*, where each atom contains two pieces of information: *1)* embedding distortion rate, Φ; and *2)* modified tree values. Φ is defined for the whole image and the algorithm includes Φ with every atom in a hierarchical manner rather than distortion responsible for individual atoms. For example, if distortion is 3 for *atom1*, 7 for *atom2* and 2 for *atom3*, the generated code-stream will contain distortion information as 3, 10 and 12 for *atom1*, *atom2*, and *atom3*, receptively. Specific level of distortion is achieved by finding the atom that has nearest match for the target distortion. Once the algorithm finds the match, it truncates at that point and a code-stream is generated by concatenating the atoms that qualify for the target distortion. The structure of the code-stream is shown in Fig. 8. One set of header information is also included in the beginning of the stream to identify the input parameters, such as, the wavelet kernel, the number of decomposition levels, and the depth of the binary tree, etc. The header information was included here for proof of concept for a truly blind watermark extractor. However the header information can be part of standard compression metadata, or can be passed as side information. One can also use a decoder that already has such information.

2) Embedded Watermarking Module: The embedded watermarking module truncates the code-stream at any distortion-robustness atom level to generate the watermarked image with desired distortion-robustness requirements of the user. Inclusion of more distortion-robustness atoms before truncation increases the robustness of the watermarked image but consumes greater embedding-distortion rate. The code-stream truncation at a given distortion-robustness atom level provides
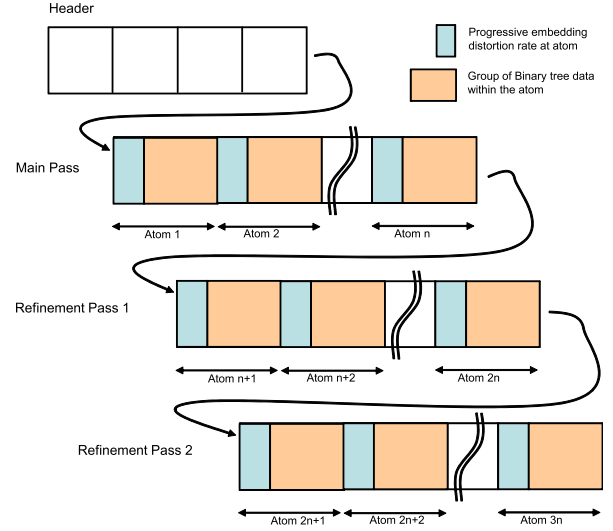

Fig. 8.   Code-stream generation.

flexibility towards watermarking scalability. To reconstruct the watermarked coefficients, the truncated code-stream is de-quantized following reverse footsteps in Section II-A1. Applying the IDWT on these coefficients generates the watermarked media with required visual quality and the watermark robustness.

3) Watermark Extractor Module: The extractor module consists of a blind extractor similar as described in Section II-A3. Any test image / video after content adaptation process is passed to this module for watermark extraction and authentication. During the extraction, the FDWT is applied on the test media and the coefficients are used to form the binary tree. Based on the rules stated in Table I, each tree is then assigned to a symbol and corresponding watermarking association. The association of 0 or 1 indicates the extracted watermark value. The extracted watermark bits are then authenticated using Eq. (4).

III. ROBUSTNESS TO JPEG2000 BIT PLANE DISCARDING

This section extends scalable watermarking presented in Section II for improving its robustness against compression caused by scalable coding schemes, such as, JPEG2000. Quality scalability in compressed bit streams has been particularly of interest due to the Quality of Service (QoS) requirements in the media consumption chain. As reported in Section I, there exists some algorithms in the literature that offer robustness to compression in general. However, in the algorithmic development, most of the algorithms do not consider the effect of JPEG2000 quantization process on the robustness of the watermark retrieval. This section shows that a special case of the proposed scalable watermarking algorithm in Section II incorporates JPEG2000 quantization process, which leads to a bit-plane discarding model for achieving quality scalability in content adaptations. Firstly, we briefly discuss the bit-plane discarding model followed by a discussion on the fitness of the proposed algorithm for achieving the watermark robustness to quality scalability-based content adaptation.
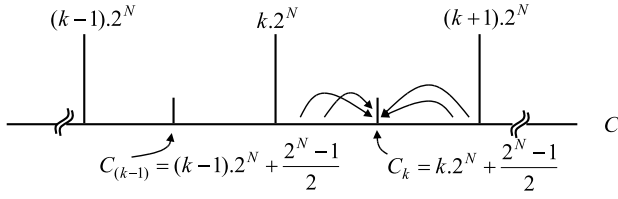
Fig. 9. Quantization in the compression scheme considering N level bit-plane discarding.

A. Scalable Coding-Based Content Adaptation

JPEG2000 uses the DWT as its core technology and offers scalable decoding with quality and resolution scalability. The scalable coders encode the image by performing the DWT followed by embedded quantizing and entropy coding. The coefficient quantization, in its simplest form, can be formulated as follows:

$$C_q = \left\lfloor \frac{C}{Q} \right\rfloor, \tag{5}$$

where C_q is the quantized coefficient, C is the original coefficient and Q is the quantization factor. Embedded quantizers often use $Q = 2^N$, where N is a non-negative integer. Such a quantization parameter within downward rounding (*i.e.*, using *floor*), can also be interpreted as bit plane discarding as commonly known within the image coding community.

At the decoder side, the reverse process of the encoding is followed to reconstruct the image. The de-quantization process is formulated as follows:

$$\hat{C} = Q \cdot C_q + \left(\frac{Q-1}{2} \right), \tag{6}$$

where $\hat{C}$ is the de-quantized coefficient. In such a quantization scheme, the original coefficient values in the range $k \cdot Q \leq C < (k+1) \cdot Q$, where $k \in \pm1, \pm2 \pm 3...$ and $Q = 2^N$ for bit plane wise coding, are mapped to $\hat{C} = C_k$, which is the center value of the concerned region as shown in Fig. 9 and in Eq. (7).

$$C_k = k \cdot 2^N + \left(\frac{2^N - 1}{2} \right). \tag{7}$$

B. Incorporating Quantization in Scalable Watermarking

To improve the robustness against quality scalable compression, at this point we incorporate the bit plane discarding model within the proposed algorithm by restricting the initial quantizer (λ) value to an integer power of two. Therefore the quantization cluster in tree formation (Section II-A) can now alternatively described as a bit-plane cluster. Because of this, every value in the binary trees corresponds to the bit-planes of the selected group of coefficients. Therefore, based on the depth parameter in the embedding algorithm, the selected coefficient can retain the watermark even after bit-plane discarding.

Assuming C' and $\widehat{C}'$ as the watermarked coefficient before and after bit plane discarding, respectively, we examine the effect of N number of bit-plane discarding on every bit in the binary tree during the watermark extraction. Considering initial $\lambda = 2^M$, where M corresponds to the depth of the tree,
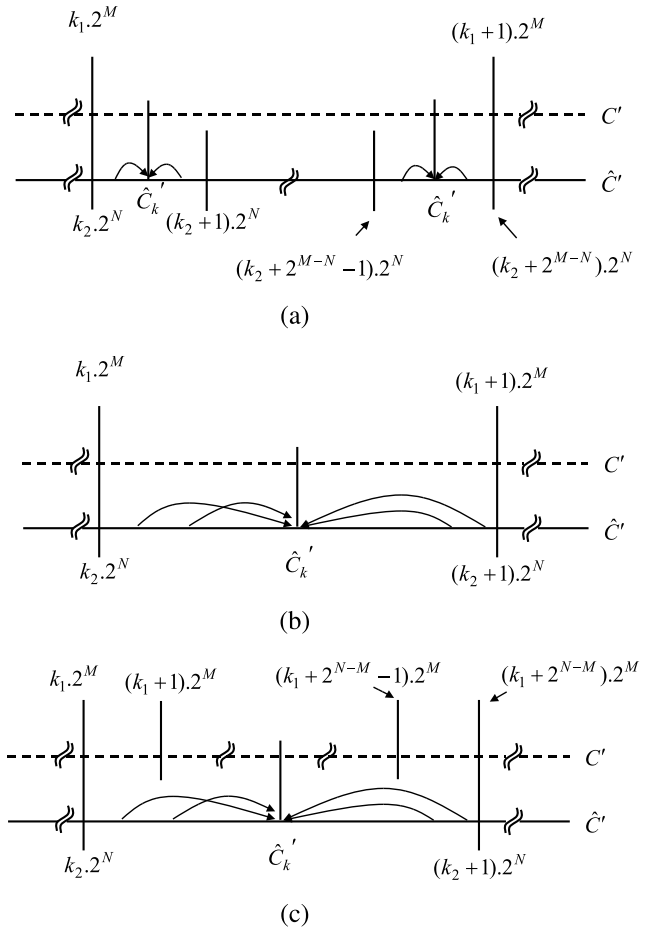

(a)

(b)

(c)

Fig. 10. Effect of bit plane discarding in watermark extraction; $\lambda = 2^M$ and N is the number of bit plane being discarded. (a) Case 1: $M > N$. (b) Case 2: $M = N$. (c) Case 3: $M < N$.

at the extractor, using Eq. (1) the bit (b_i) in the binary tree can be calculated as:

$$b_i = \left\lfloor \frac{|C'|}{2^M} \right\rfloor mod\ 2,$$
$$= k_1 mod\ 2, \tag{8}$$

where k_1 is the cluster index as shown in Fig. 10. Using the bit-plane discarding model in Section III-A, the watermarked coefficients, C', are quantized and mapped to the center value, $\widehat{C}'_k$, within a bit-plane cluster with an index value of k_2 as shown in Fig. 10. At this point, we consider following three cases to investigate the effect of this quantization and de-quantization process:

1) Case 1 ($M > N$): In this case, the binary tree cluster ($\lambda = 2^M$) is bigger than the bit-plane discarding cluster. Hence for any bit-plane discarding where $M > N$, $\widehat{C}'_k$ value remains within the binary tree cluster, $k \cdot 2^M \leq \widehat{C}'_k < (k+1) \cdot 2^M$, as shown in Fig. 10.(a) and

$$b_i = \left\lfloor \frac{|C'|}{2^M} \right\rfloor mod\ 2,$$
$$= \left\lfloor \frac{|\widehat{C}'|}{2^M} \right\rfloor mod\ 2,$$
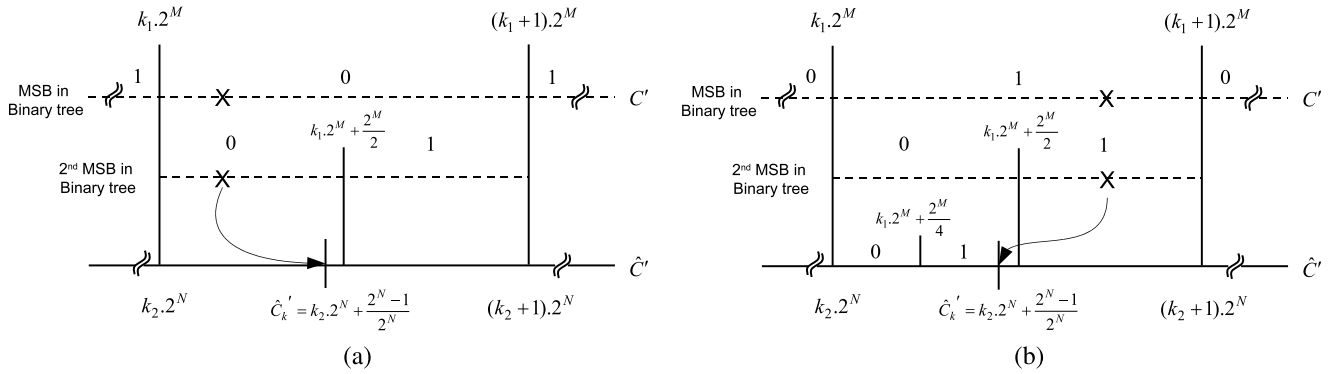$$= b'_i, \tag{9}$$

Fig. 11. Effect of bit-plane discarding in watermark extraction for special case of EZ and EO; $\lambda = 2^M$ and N is the number of bit-plane being discarded. (a) Case: EZ. (b) Case: EO.

where b_i and b_i' represents the bit in a binary tree, without bit-plane discarding and after bit-plane discarding, respectively.

2) Case 2 ($M = N$): This case considers the same cluster size in binary tree and the bit-plane discarding, and therefore $\widehat{C_k'}$ remains in the same cluster of binary tree during watermark extraction, as shown in Fig. 10.(b) and hence $b_i = b_i'$.

3) Case 3 ($M < N$): In this scenario, the number of bit-planes being discarded are greater than the depth of the binary tree. Due to bit-plane discarding, any watermarked coefficient, (C'), in the cluster ($k_2 \cdot 2^N \le C' < (k_2 + 1) \cdot 2^N$) is mapped to the center value, C_k'. In terms of the binary tree clustering this range can be defined as ($k_1 \cdot 2^M \le C' < (k_1 + 2^{(N-M)}) \cdot 2^M$), where ($N - M$) is a positive integer. Hence during watermark extraction, the index of the binary tree cluster can be changed and effectively $b_i = b_i'$ is not guaranteed.

So far we have explained the effect of bit-plane discarding on individual bits of a binary tree. Since the algorithm generates the watermark association symbols using the three most significant bits of the binary tree (Table I), we can define the necessary condition for the coefficients to retain the watermark as follows:

$$d \ge N + 3, \tag{10}$$

where d is the depth of the binary tree and N is the number of bit planes assumed to be discarded.

After the second refinement pass in the code-stream all modified coefficients are associated with either EZ and EO. In that case only the two most significant bits are required to be preserved. Hence when considered embedding, the highest robustness criteria in Eq. (10) becomes :

$$d \ge N + 2. \tag{11}$$

Nonetheless, in this case, the second most significant bit in the binary tree does not need to be preserved, whereas the MSB is preserved in combination with the support decision from the third most significant bit, *i.e.*, EZ and EO are allowed to be extracted as CZ and CO, respectively. Now we examine the effect of bit-plane discarding in these cases when $d = N + 1$.

Case EZ: Considering $\lambda = 2^M$ in this case, after the second refinement pass, the coefficients, C', are associated to embedded zero ($EZ \rightarrow 00x$), *i.e.*, $k_1 \cdot 2^M \le C' < \left(k_1 + \frac{2^M}{2}\right)$,

where $k_1 \bmod 2 = 0$, as shown in Fig. 11.a. After N number of bit-planes discarding, C' is modified to the center value $\hat{C}_k = \left(k_2 \cdot 2^N + \frac{2^N - 1}{2}\right)$. For $M = N$ (*i.e.*, $d = N + 1$), k_2 becomes k_1 and therefore:

$$\hat{C}_k = \left(k_2 \cdot 2^N + \frac{2^N - 1}{2}\right) < \left(k_1 \cdot 2^M + \frac{2^M}{2}\right)$$
$$\Rightarrow \hat{C}_k < \left(k_1 \cdot 2^M + \frac{2^M}{2}\right),$$
$$\forall \ k_1 \cdot 2^M < \hat{C}_k < \left(k_1 \cdot 2^M + \frac{2^M}{2}\right) \tag{12}$$

results in the second MSB remains as 0 in the binary tree. Hence, after $d = N + 1$ number of bit-planes discarding, the coefficient association with EZ remains the same and the watermark information can be successfully recovered.

Case EO: Referring to Fig. 11.b, for embedded one ($EO \rightarrow 11x$), the condition for coefficient association becomes $\left(k_1 + \frac{2^M}{2}\right) \le C' < (k_1 + 1) \cdot 2^M$, where $k_1 \bmod 2 = 1$. Similar to the previous case, after N number of bit-planes discarding, C' is modified to the center value of the corresponding cluster $\hat{C}_k = \left(k_2 \cdot 2^N + \frac{2^N - 1}{2}\right)$. Considering $M = N$, similar to Eq. (12), we can write:

$$k_1 \cdot 2^M < \hat{C}_k < \left(k_1 \cdot 2^M + \frac{2^M}{2}\right). \tag{13}$$

Therefore, the two MSBs of the binary tree are now changed as $11x \rightarrow 10x$. At this point, we aim to extract the third MSB, b', as:

$$b' = \begin{cases} 0 & \text{if} \quad k_1 \cdot 2^M \le \hat{C}_k < \left(k_1 \cdot 2^M + \frac{2^M}{4}\right), \\ 1 & \text{if} \quad \left(k_1 \cdot 2^M + \frac{2^M}{4}\right) \le \hat{C}_k < \left(k_1 \cdot 2^M + \frac{2^M}{2}\right). \end{cases} \tag{14}$$

Now considering $M = N \Rightarrow \frac{2^N - 1}{2} > \frac{2^M}{4}$, Eq. (13) becomes

$$\left(k_1 \cdot 2^M + \frac{2^M}{4}\right) < \hat{C}_k < \left(k_1 \cdot 2^M + \frac{2^M}{2}\right). \tag{15}$$

Combining, Eq. (14) and Eq. (15), the extracted third MSB becomes $\hat{b} = 1$ and hence, $11x \rightarrow 101$. Therefore, after $d = N + 1$ number of bit-planes discarding, the coefficient association with EO becomes CO and the watermark information can still be successfully extracted.

Combining the above mentioned cases, we can modify Eq. (11) and conclude that for EZ or EO the relationship
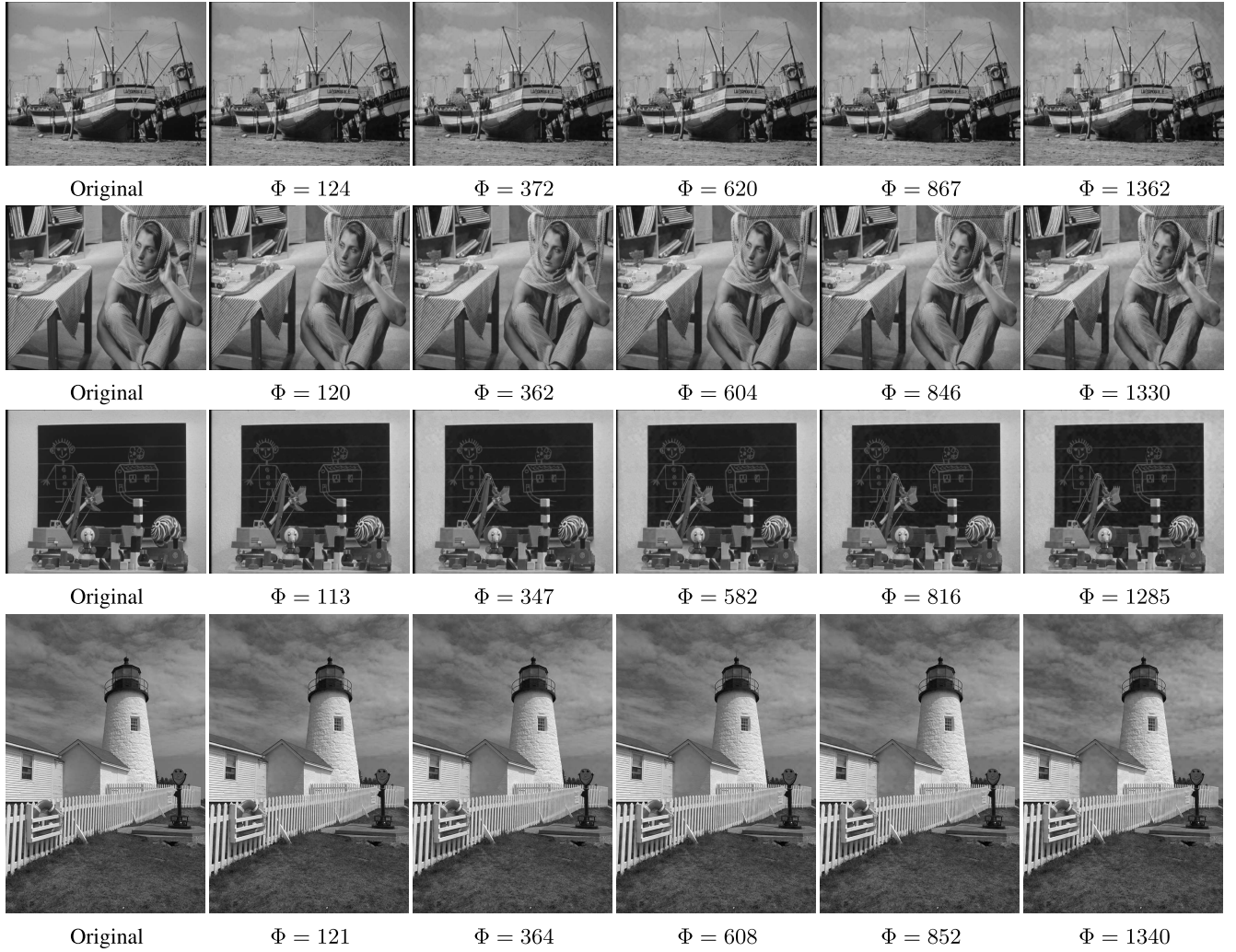
Fig. 12.   Visual representation of watermarked images at various rate points for *Row 1)* Boat *2)* Barbara *3)* Blackboard and *4)* Light House.

between the embedding depth, d, and maximum number of bit-plane discarding, N, is as follows:

$$d \geq N + 1. \qquad (16)$$

Therefore, using the above mentioned conditions, the proposed new algorithm ensures the reliable detection of the watermark against quality scalable content adaptation which follows the bit-plane discarding model.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

This section provides the experimental verification of the proposed scalable watermarking scheme, for images as well as video. It is also evaluated for its robustness to scalable content adaptation attacks.[1] The source code of the proposed algorithm is available from https://github.com/dbhowmik/scalableWM.[2]

---

[1]The proposed algorithm aims to address the watermarking challenges only against quality scalable content adaptation. Therefore security aspect of watermarking schemes and all other common watermarking attacks such as filtering, denoising etc. are beyond the scope of this paper.

[2]Source code: https://github.com/dbhowmik/scalableWM

A. Scalable Watermarking for Images

The experimental simulations are grouped into four sets: *1)* Proof of the concept, *2)* Verification of the scheme for bit-plane discarding model, *3)* Robustness performance against JPEG2000 and *4)* Robustness comparison with existing blind watermarking schemes. For all experiments, a 3 level 9/7 wavelet decomposition is used as the FDWT. Then the low frequency subband has been selected to embed a binary logo based watermark. The initial quantization value λ is set to 32 resulting a tree-depth of $d = 6$. d is a user defined parameter which we varied for further investigation and reported results in following sections. In generating the code-stream, atoms are defined by grouping every 16 consecutive binary-trees. The code-stream is generated by organizing hierarchically nested distortion-robustness atoms, generated in 2 individual passes.

1) Proof of the Concept: Once the code-stream is generated, set of watermarked images are produced by truncating the code-stream at different embedding-distortion rate points, Φ, defined in Eq. (3). The results for four test images, *Boat*, *Barbara*, *Blackboard* and *Light House* are shown in Fig. 12. As the embedding process creates a hierarchical code-stream, for various Φ values, watermark strength varies accordingly,
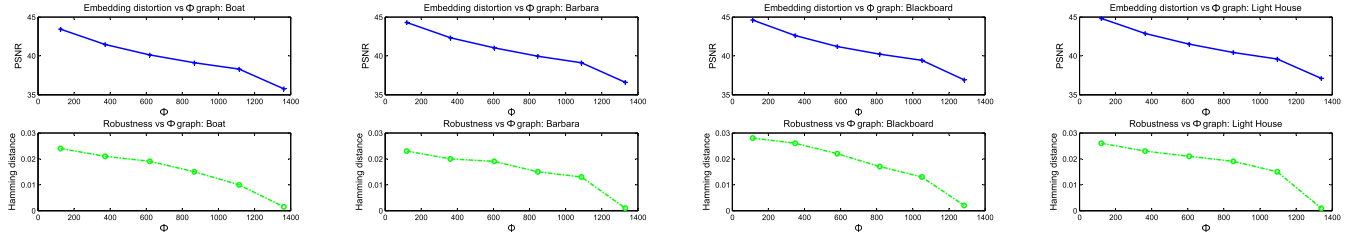
Fig. 13. PSNR and robustness vs Φ graph. *Row 1:* Embedding distortion vs. Φ, *Row 2:* Hamming distance vs. Φ.
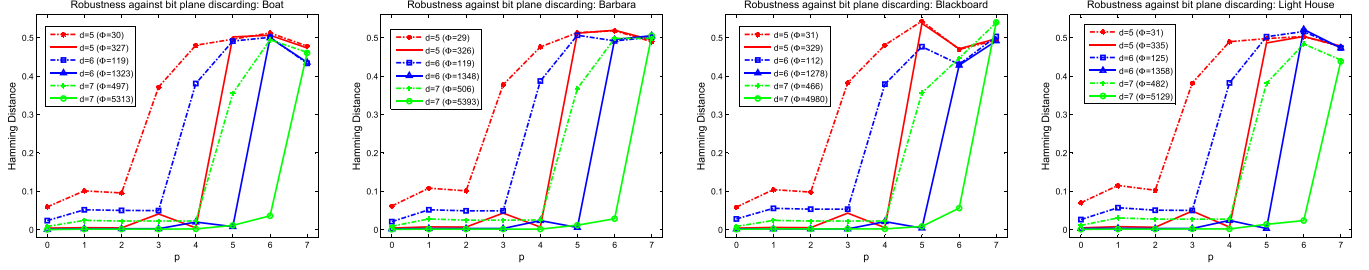

Fig. 14. Robustness against discarding of *p* bit planes for various *d* at minimum and maximum Φ.
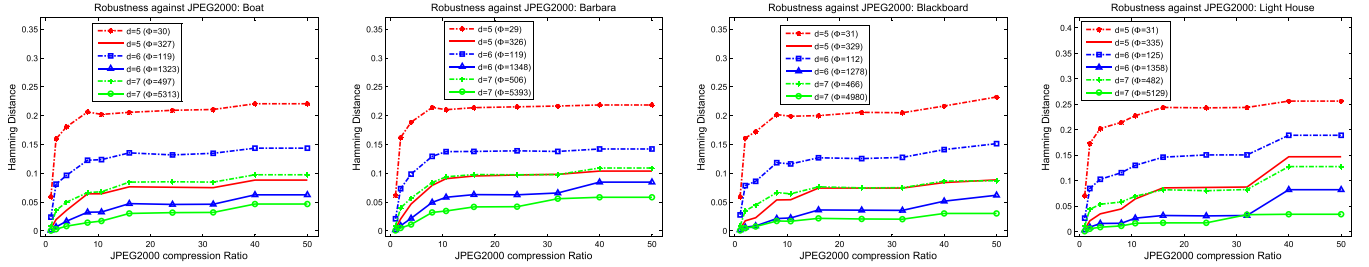

Fig. 15. Robustness against JPEG2000 compression for various *d* at Φ_{min} and Φ_{max}.

i.e., higher Φ corresponds to higher watermarking strength for a given data capacity. As a result, with increased value of Φ, high embedding distortion is introduced in the watermarked images and hence the visual image quality degrades as shown in the above mentioned figures. However, with higher watermarking strength, the robustness performance also improves. The overall embedding distortion performance for the test images, measured by PSNR and the robustness performance (Hamming distance) at various Φ values is shown in Fig. 13. The *x*-axis of the plots shows Φ. The *y*-axis shows the PSNR in plots in *Row 1* and the Hamming distance in plots in *Row 2*.

It is evident from these plots, that a higher embedding-distortion rate, *i.e.*, higher watermarking strength, results in poor PSNR but offers higher robustness. However, a trade-off can be made based on the application scenario by selecting an optimum embedding-distortion rate to balance the visual quality and robustness. The proposed algorithm is sensitive to image contents and does not enforce coefficients to achieve the highest robustness. For example with minimum ϕ, the Hamming distance (H) relating robustness may not be 0 which gradually improves with higher value of ϕ. This is useful to preserve the image quality by controlling the embedding distortion rate. However the algorithm ensures that the robustness is within permissible limit. The interpretation of Hamming distances for practical use is discussed in [24] which proposed that $H < 0.2$ ensures correct extraction of the watermark.

2) Verification of the Scheme Against Bit-Plane Discarding: The proposed watermarking scheme incorporates bit-plane discarding model and the experimental verifications for the same are shown in Fig. 14. The *y*-axis shows the robustness in terms of Hamming distance against the number of bit planes discarded (p) shown on the *x*-axis. Here different depth (d) values with the minimum embedding distortion rate, Φ_{min} and the maximum embedding distortion rate, Φ_{max}, values are chosen to verify our arguments in Eq. (10) and Eq. (16). At Φ_{min}, the condition of correct watermark extraction is given in Eq. (10) and the same is evident from the results shown in Fig. 14. At Φ_{max}, all coefficients are associated with EZ or EO and the necessary condition to extract watermark is discussed in Eq. (16), which is supported by the simulation results as shown in Fig. 14. For example, at $d = 6$, for Φ_{min}, correct watermark extraction is possible up to $p = 3$ and for Φ_{max}, correct watermark is extracted up to $p = 5$ as shown in these plots.

3) Robustness Performance Against JPEG2000: Fig. 15 and Fig. 16 show the robustness performance of the proposed watermarking scheme against JPEG2000 scalable compression. We firstly verify the proposed scheme's robustness against JPEG2000 compression using different *depth* parameter, *d*, as shown in Fig. 15 followed by the watermark scalability at a given *depth* as shown in Fig. 16. An ITU-T.804 JPEG2000 standard reference software[3] is used
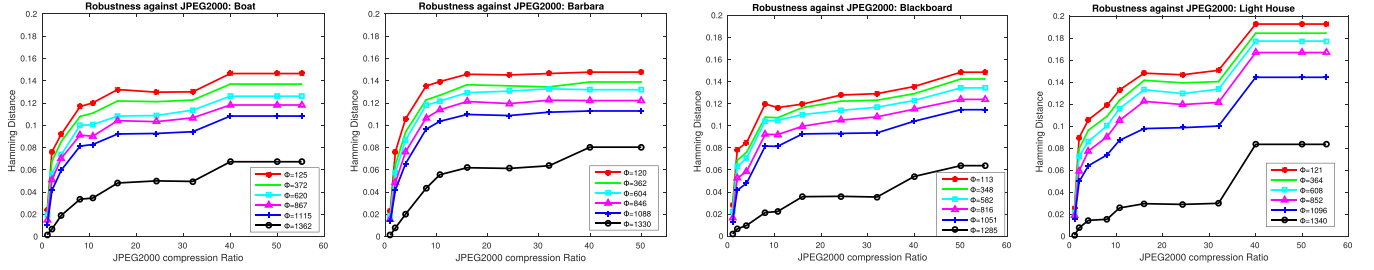
---

[3]http://www.itu.int/rec/T-REC-T.804-201504-I

Fig. 16. Robustness against JPEG2000 compression for various Φ values at $d = 6$.
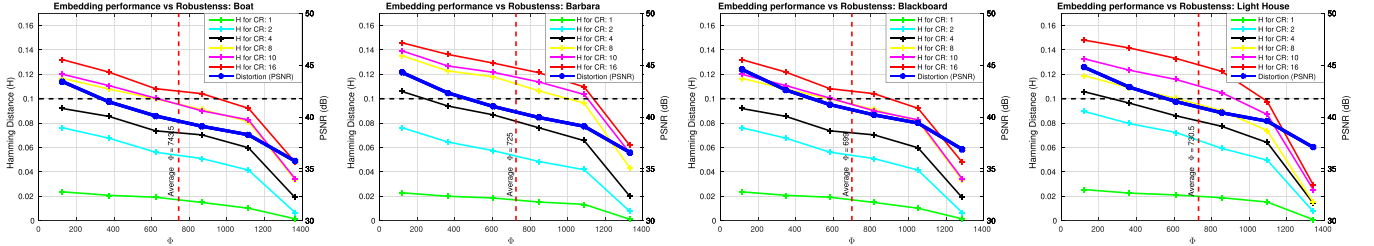

Fig. 17. PSNR and robustness against JPEG2000 vs Φ at $d = 6$ graph. The x-axis represents Φ, y-axis on the left represents corresponding Hamming distance (H) against JPEG2000 content adaptation at various compression ratio (CR), y-axis on the right represents PSNR (dB), the vertical red line represents average Φ and the horizontal black line shows Hamming distance of 0.1.

for the experiments. These results compare the robustness for various Φ for a given d. In all the figures the *x*-axis represents the JPEG2000 quality compression ratio while *y*-axis shows the corresponding Hamming distances.

It is evident from the plots, that higher depth and higher Φ in a given depth, offer higher robustness to scalable content adaptation attacks. The watermark scalability is achieved by truncating the distortion-constrained code stream at various rate points with respect to Φ. With increased Φ more coefficients are associated with *EZ* or *EO* and hence improves the robustness by successfully retaining the watermark information at higher compression rates. The results show that more than 35% improvement in robustness when comparing two consecutive depth levels, d, and more than 60% improvement between Φ_{min} and Φ_{max} at a given depth.

Trade-off between embedding distortion performance and robustness against JPEG2000 Φ are shown in Fig. 17. The x-axis represents Φ, y-axis on the left represents corresponding Hamming distance (H) against JPEG2000 content adaptation at various compression ratio (CR), y-axis on the right represents PSNR (dB), the vertical red line represents average Φ and the horizontal black line shows Hamming distance of 0.1. These graphs are useful to indicate the robustness performance of a given image at various embedding-distortion rate. For better imperceptibility one may choose suitable value of Φ from the left section of the graph while for a target compression ratio, any value of Φ can be selected as long as a target Hamming distance (*e.g.*, the horizontal black line here corresponds to H = 0.1 for illustration purpose) is achieved. A trade off can be made around a region where two perpendicular lines meet.

4) Robustness Performance Comparison With Existing Non-Scalable Watermarking Methods: This is the first scalable watermarking work of its kind and therefore comparisons

are made only with the next best available non-scalable watermarking algorithms in the literature. In this section we compare our proposed algorithm with a popular blind re-quantization based watermarking scheme (non-scalable) used in [9]–[11] and [14]. These algorithms share a common embedding model [24] which rely on modifying various coefficients towards a specific quantization step. Herein we call it *the existing algorithm*. As proposed in [11], the existing algorithm is based on modifying the median coefficient towards the step size, δ, by using a running non-overlapping 3×1 window. The altered coefficient must retain the median value of the three coefficients within the window, after the modification. The equation calculating δ is described as follows:

$$\delta = \alpha \frac{(C_{min}) + (C_{max})}{2}, \qquad (17)$$

where C_{min} and C_{max} are the minimum and maximum coefficients, respectively. The median coefficient, C_{med}, is quantised towards the nearest step, depending on the binary watermark, b. The extracted watermark, b', for a given window position, is extracted by

$$b' = \left[\frac{C_{max} - C_{med}}{\delta} \right] \ mod \ 2. \qquad (18)$$

For fair comparison, we first calculate Φ for the existing watermarking algorithm and then set the same Φ for the experiments in the proposed method. The embedding performance is reported in Table II and the robustness against JPEG2000 compression is shown in Fig. 18.

In embedding distortion performance comparison, for similar Φ, the existing method shows a better overall embedding performance in terms of PSNR. However, the data capacity of the proposed algorithms are 3 times higher than the existing one. Therefore, using the new embedding-distortion metric, Φ,
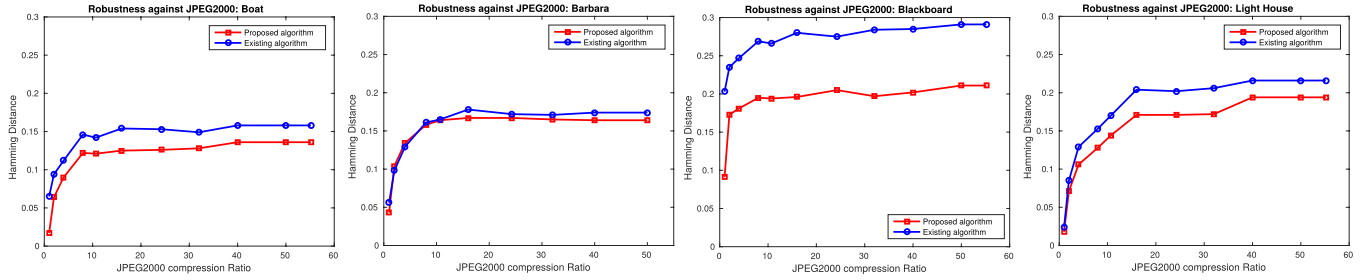
Fig. 18. Robustness performance comparison between existing and the proposed method against JPEG2000 compression with same Φ.

TABLE II
EMBEDDING DISTORTION PERFORMANCE COMPARISON BETWEEN
EXISTING AND PROPOSED WATERMARKING METHODS

	Existing algorithm			Proposed method		
	Φ	PSNR	Data Capacity	Φ	PSNR	Data Capacity
Boat (704 × 576)	86.40	53.74	2112	84.13	47.43	6336
Barbara (704 × 576)	80.64	55.12	2112	81.71	49.13	6336
Blackboard (704 × 576)	69.12	56.45	2112	69.12	50.51	6336
Light House (768 × 512)	84.48	55.36	2048	82.43	48.78	6144

which considers embedding distortion and data capacity into a single metric, we can fairly compare the robustness performance of these two schemes for a given Φ. The results show that despite having three times more data capacity, the proposed algorithm outperforms the existing blind algorithm by an average improvement of 25% to 35% at higher compression ratios. This confirms that the new algorithm, coupled with the bit-plane discarding model, offers improvements in robustness against scalable compression over the existing algorithm which does not use the model.

The proposed algorithm adds a new avenue to watermarking strategies by offering a flexible scalable watermarking approach, *i.e.*, to achieve the higher robustness at a high compression ratio, one can choose higher Φ and the effect on embedding distortion is neutralized by quantization in compression. An example is shown in Fig. 2 for Barbara image, where we compare the embedding distortion of the watermarked image after compression. The PSNR of the watermarked and the un-watermarked images are comparable at various compression points, while the watermarked image offers authenticity of the image with desired robustness H.

B. Scalable Watermarking for Video

Finally, we extend the proposed scalable watermarking scheme for videos using a 2D+t+2D motion compensated temporal filtering (MCTF) based video watermarking framework [33]. 2D+t+2D refers to a 3D video decomposition

scheme where 2D and t represent a spatial and temporal decompositions, respectively. For example, t+2D is achieved by performing temporal decomposition followed by a spatial transform where as in case of 2D+t, the temporal filtering is done after the spatial 2D transform. Issues such as flicker and residue error propagation related to video watermarking were addressed in the framework by proposing motion compensated temporal filtering that considers object motion within frames. Therefore it is appropriate to choose the same framework to extend the proposed algorithm. In this work the watermarking code-stream is generated using the 2D+t decomposed host video, as described in the framework. The binary tree is formed using the motion compensated filtered wavelet coefficients.

Similar to the image watermarking case of the proposed algorithm, the watermarked video is generated at a given embedding distortion rate, Φ, either at individual frame level or at group of picture (GOP) level. For the experimental set shown in this work, Φ is calculated for every GOP, with a size of 8 frames per GOP. In extracting the watermark data, firstly the test video is decomposed using the 2D+t+2D framework with blind motion estimation, *i.e.*, without any reference to the original video or motion vectors and then the binary tree is formed for the selected coefficients. The watermark extraction decision is made using the association rules described in Table I.

As recommended in the original framework, the experimental simulations in this work used 230 spatio-temporal subband decomposition where a 2-level 9/7 spatial decomposition is performed, followed by a 3-level MCTF-based temporal decomposition. In subband selection, LL_s spatial subband is used in two different temporal subbands: *LLL* and *LLH*. In all cases, normalization is used during spatio-temporal decomposition. In the embedding procedure, depth parameter d is set to 6 with a data capacity of 6336. The performance of the algorithm is evaluated for various Φ, by comparing the embedding distortion and robustness against scalable compressions.

The embedding distortion is measured using PSNR and the results are shown in Fig. 19 for *LLL* & *LLH* temporal subband for the CIF resolution (352 × 288) test sequences *Crew*, *Foreman* and *News*. The *x*-axis shows the frame number while *y*-axis shows corresponding PSNR. The robustness performance is evaluated by comparing Hamming distance against scalable compression schemes, such as,
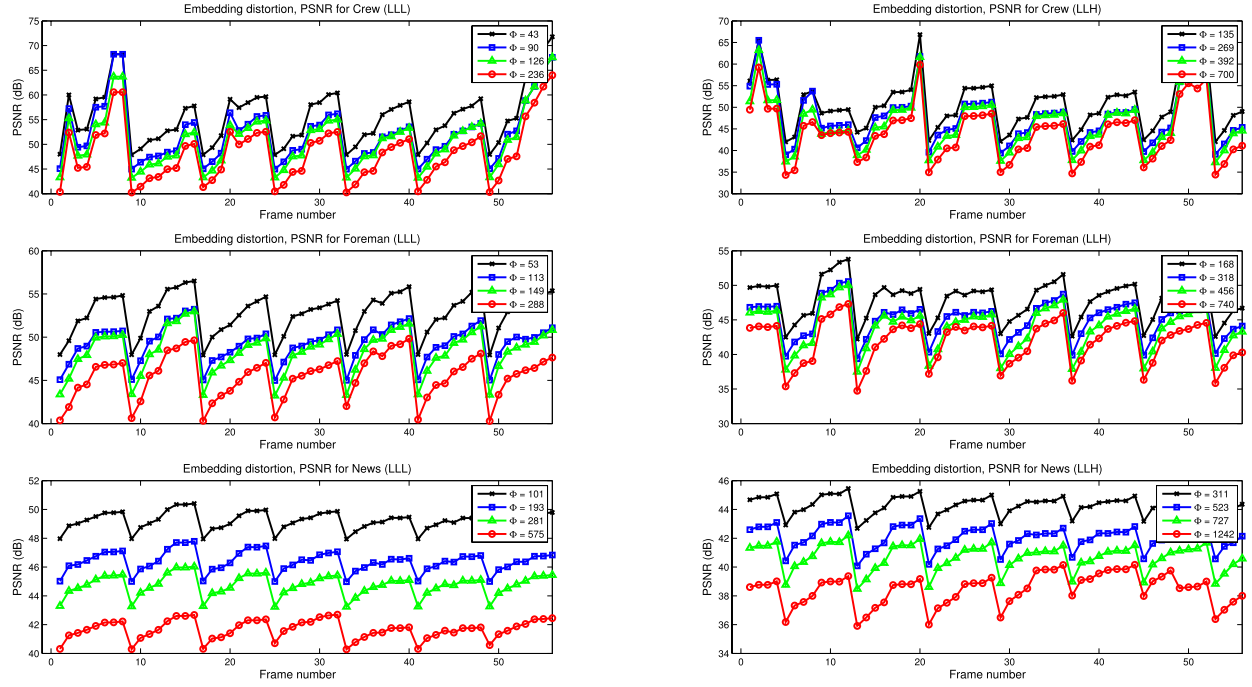
Fig. 19.   Embedding distortion performance for proposed watermarking on *LLL* (*Column 1*) & *LLH* (*Column 2*) temporal subbands for various $\Phi(d = 6)$. *Row 1), 2) & 3)* represents embedding performances for *Crew*, *Foreman* and *News* sequences, respectively.
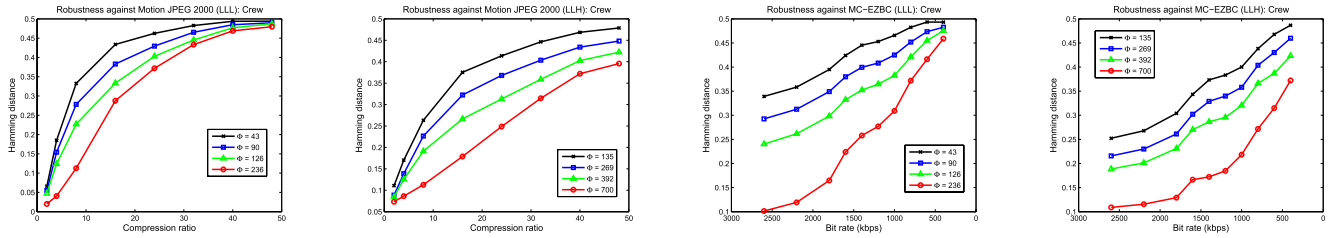

Fig. 20.   Robustness performance of proposed watermarking scheme at different Φ ($d = 6$) for *Crew* sequence. Results for robustness against Motion JPEG2000 and MC-EZBC are shown in first two and last columns, respectively. *Column 1) & 3)* represent the embedding on temporal subbands *LLL* and *Column 2) & 4)* are for *LLH*.
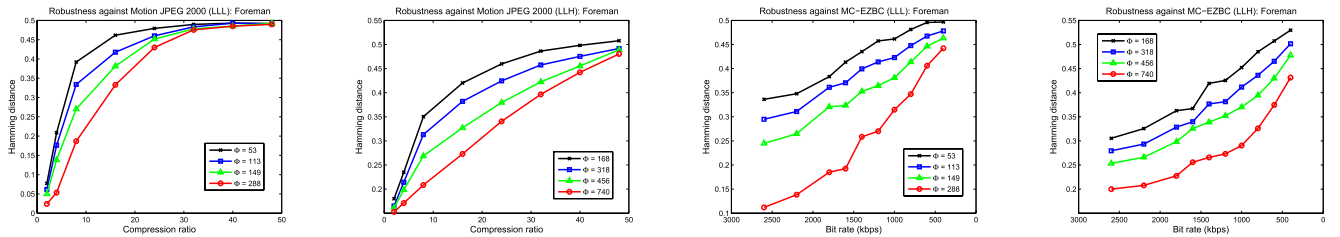

Fig. 21.   Robustness performance of proposed watermarking scheme at different Φ ($d = 6$) for *Foreman* sequence. Results for robustness against Motion JPEG2000 and MC-EZBC are shown in first two and last columns, respectively. *Column 1) & 3)* represent the embedding on temporal subbands *LLL* and *Column 2) & 4)* are for *LLH*.

Motion JPEG2000 and Motion Compensated Embedded Zero Block Coding (MC-EZBC) [34] . The results are shown in Fig. 20, Fig. 21 and Fig. 22 for *Crew*, *Foreman* and *News* sequences, respectively. The left and right columns in these figures represent the robustness performance against Motion JPEG2000 and MC-EZBC, respectively. Results for *LLL* subband are shown in *Column 1 & 3* and *LLH* in *Column 2 & 4*. In all the cases the *x*-axis shows the compression ratio / bit rates and the

corresponding Hamming distances are shown in *y*-axis. The Hamming distances are calculated by averaging the individual frame level Hamming distances of each test sequence.

From the results, it is evident that the concept of scalable watermarking is successfully realised within a video watermarking framework. With the increase in embedding distortion rate, Φ, the robustness performances are improved by 30% to 70% between Φ_{min} and Φ_{max}, while embedding distortion is also increased with increasing Φ. Conceptually as described
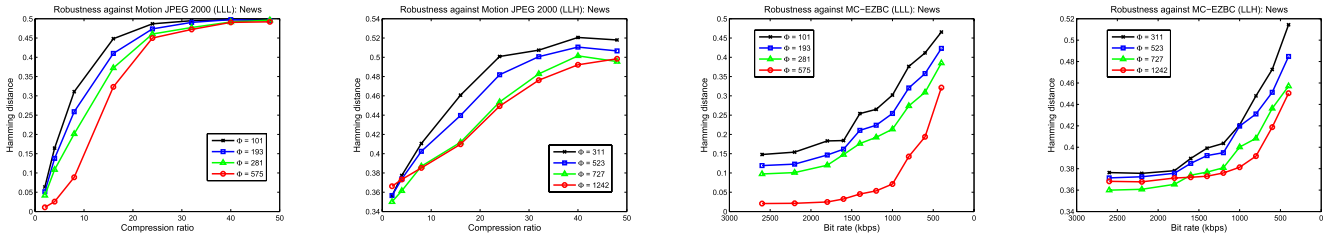
Fig. 22. Robustness performance of proposed watermarking scheme at different Φ ($d = 6$) for *News* sequence. Results for robustness against Motion JPEG2000 and MC-EZBC are shown in first two and last columns, respectively. *Column 1) & 3)* represent the embedding on temporal subbands *LLL* and *Column 2) & 4)* are for *LLH*.

before, a high Φ can be chosen where high compression is expected and a low Φ can be opted for high resolution video distribution, based on the end user's need. Therefore, a combined scalable watermarking and video encoding scheme can ensure secure multimedia distribution within scalable content adaptation scenarios.

V. CONCLUSIONS

In this paper, we have proposed a novel concept of scalable watermarking and extended it to make watermarking robust against the quality scalable content adaptation attacks. To generate a scalable watermark, firstly a distortion constrained code-stream is generated by concatenating hierarchically nested joint distortion robustness coding atoms. The code-stream is then truncated at various embedding-distortion rate points to create watermarked images, based on the distortion-robustness requirements. The extraction and authentication is done using a blind extractor. The algorithm is developed based on the bit-plane discarding model used in scalable content adaptation.

The concept was experimentally verified for images and video sequences. The robustness against JPEG2000 quality scalability yielded improvements of more than 35% when comparing two consecutive depth levels, d, and of more than 60% between Φ_{min} and Φ_{max} at a given depth. The results showed that despite having three times more data capacity, the proposed algorithm outperformed the existing blind algorithms by an average improvement of 25% to 35% at higher compression ratios. This has confirmed that the new algorithm, coupled with the bit-plane discarding model, offers improvements in robustness against scalable compression over the existing non-scalable schemes which do not use the model.

Finally this scheme was extended for video watermarking using the 2D+t+2D MCTF based framework. The experiments showed an improvement in robustness up to 70%. The proposed scalable watermarking scheme adds a new direction in watermarking research and has many potential watermarking applications particularly in security enabled scalable content coding.

REFERENCES

[1] D. S. Taubman and M. W. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice: Image Compression Fundamentals*. USA: Springer, 2002.

[2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[3] G. J. Sullivan, J. M. Boyce, Y. Chen, J.-R. Ohm, C. A. Segall, and A. Vetro, "Standardized extensions of high efficiency video coding (HEVC)," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1001–1016, Dec. 2013.

[4] H.-T. Hu and L.-Y. Hsu, "Exploring DWT–SVD–DCT feature parameters for robust multiple watermarking against JPEG and JPEG2000 compression," *Comput. Elect. Eng.*, vol. 41, pp. 52–63, Jan. 2015.

[5] N. Cai, N. Zhu, S. Weng, and B. W.-K. Ling, "Difference angle quantization index modulation scheme for image watermarking," *Signal Process., Image Commun.*, vol. 34, pp. 52–60, May 2015.

[6] A. V. Subramanyam and S. Emmanuel, "Partially compressed-encrypted domain robust JPEG image watermarking," *Multimedia Tools Appl.*, vol. 71, no. 3, pp. 1311–1331, Aug. 2014.

[7] F. Dufaux, S. J. Wee, J. G. Apostolopoulos, and T. Ebrahimi, "JPSEC for secure imaging in JPEG2000," *Proc. SPIE*, vol. 5558, pp. 319–330, Nov. 2004.

[8] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 783–791, May 2001.

[9] F. Huo and X. Gao, "A wavelet based image watermarking scheme," in *Proc. IEEE ICIP*, Oct. 2006, pp. 2573–2576.

[10] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 5. May 1998, pp. 2969–2972.

[11] L. Xie and G. R. Arce, "Joint wavelet compression and authentication watermarking," in *Proc. Int. Conf. Image Process.*, vol. 2. Oct. 1998, pp. 427–431.

[12] Z. Zhang and Y. L. Mo, "Embedding strategy of image watermarking in wavelet transform domain," *Proc. SPIE*, vol. 4551, pp. 127–131, Sep. 2001.

[13] T.-S. Chen, J. Chen, and J.-G. Chen, "A simple and efficient watermarking technique based on JPEG2000 codec," in *Proc. 5th Int. Symp. Multimedia Softw. Eng.*, Dec. 2003, pp. 80–87.

[14] P. Meerwald, "Quantization watermarking in the JPEG2000 coding pipeline," in *Proc. Int. Work. Conf. Commun. Multimedia Secur.*, 2001, pp. 69–79.

[15] G. Bhatnagar, Q. M. J. Wu, and B. Raman, "Robust gray-scale logo watermarking in wavelet domain," *Comput. Elect. Eng.*, vol. 38, no. 5, pp. 1164–1176, Sep. 2012.

[16] C.-H. Chou and K.-C. Liu, "A perceptually tuned watermarking scheme for color images," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2966–2982, Nov. 2010.

[17] A. Subramanyam, S. Emmanuel, and M. S. Kankanhalli, "Robust watermarking of compressed and encrypted JPEG2000 images," *IEEE Trans. Multimedia*, vol. 14, no. 3, pp. 703–716, Jun. 2012.

[18] E. Esen and A. A. Alatan, "Robust video data hiding using forbidden zone data hiding and selective embedding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1130–1138, Aug. 2011.

[19] J. Sun and J. Liu, "Data hiding with video independent components," *Electron. Lett.*, vol. 40, no. 14, pp. 858–859, Jul. 2004.

[20] C. Cruz-Ramos, R. Reyes-Reyes, M. Nakano-Miyatake, and H. Perez-Meana, "A blind video watermarking scheme robust to frame attacks combined with MPEG2 compression," *J. Appl. Res. Technol.*, vol. 8, no. 3, pp. 323–337, 2010.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[22] F. A. Petitcolas, M. Steinebach, F. Raynal, J. Dittmann, C. Fontaine, and N. Fates, "Public automated Web-based evaluation service for watermarking schemes: StirMark benchmark," *Proc. SPIE*, vol. 4314, pp. 575–584, Aug. 2001.

[23] S. Pereira, S. Voloshynovskiy, M. Madueno, S. Marchand-Maillet, and T. Pun, "Second generation benchmarking and application oriented evaluation," in *Proc. Int. Workshop Inf. Hiding*, vol. 2137. 2001, pp. 340–353.

[24] D. Bhowmik and C. Abhayaratne, "On robustness against JPEG2000: A performance evaluation of wavelet-based watermarking techniques," *Multimedia Syst.*, vol. 20, no. 2, pp. 239–252, Mar. 2014.

[25] T. P.-C. Chen and T. Chen, "Progressive image watermarking," in *Proc. IEEE ICME*, vol. 2. Jul. 2000, pp. 1025–1028.

[26] P.-C. Su, H.-J. M. Wang, and C.-C. J. Kuo, "An integrated approach to image watermarking and JPEG-2000 compression," *J. VLSI Signal Process. Syst. Signal, Image, Video Technol.*, vol. 27, no. 1, pp. 35–53, Feb. 2001.

[27] D. Bhowmik and C. Abhayaratne, "The effect of quality scalable image compression on robust watermarking," in *Proc. 16th Int. Workshop Digit. Signal Process.*, Jul. 2009, pp. 1–8.

[28] P. Meerwald and A. Uhl, "Scalability evaluation of blind spread-spectrum image watermarking," in *Proc. 7th Int. Workshop Digit. Watermarking (IWDW)*, vol. 5450. 2008, pp. 61–75.

[29] A. Piper, R. Safavi-Naini, and A. Mertins, "Resolution and quality scalable spread spectrum image watermarking," in *Proc. 7th Workshop Multimedia Secur.*, 2005, pp. 79–90.

[30] C. Abhayaratne and D. Bhowmik, "Scalable watermark extraction for real-time authentication of JPEG 2000 images," *J. Real-Time Image Process.*, vol. 8, no. 3, pp. 307–325, Sep. 2013.

[31] D. Bhowmik and C. Abhayaratne, "Distortion constrained robustness scalable watermarking," in *Proc. IET Conf. Image Process.*, Jul. 2012, pp. 1–6.

[32] D. Bhowmik and C. Abhayaratne, "Robust watermarking for scalable image coding-based content adaptation," in *Proc. IET Conf. Image Process.*, Jul. 2012, pp. 1–6.

[33] D. Bhowmik and C. Abhayaratne, "2D+t wavelet domain video watermarking," *Adv. Multimedia*, vol. 2012, 2012, Art. no. 973418.

[34] P. Chen and J. W. Woods, "Bidirectional MC-EZBC with lifting implementation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 10, pp. 1183–1194, Oct. 2004.

Deepayan Bhowmik received the B.E. degree in electronics engineering from the Visvesvaraya National Institute of Technology, Nagpur, India, in 2003, the M.Sc. degree in electronics and information technology from Sheffield Hallam University, U.K., in 2006, and the Ph.D. degree from The University of Sheffield, U.K., in 2011. He was as a Research Associate with The University of Sheffield, and a System Engineer with ABB Ltd., India. He is currently a Research Associate with Heriot-Watt University, Edinburgh, U.K. He received the U.K. Engineering and Physical Sciences Research Council (EPSRC) Dorothy Hodgkin Postgraduate Award in 2006 and 2010 and a Digital Catapult-EPSRC Fellowship in 2016 and 2017 to conduct research on multimedia security. His research interests include multimedia forensics and security, computer vision, and embedded image processing hardware.

Charith Abhayaratne received the B.E. degree in electrical and electronic engineering from the University of Adelaide, Australia, in 1998, and the Ph.D. degree in electronic and electrical engineering from the University of Bath, U.K., in 2002. He is currently a Lecturer with the Department of Electronic and Electrical Engineering, The University of Sheffield, U.K. His research interests include video and image compression, watermarking, image and video analysis, wavelets, and signal transforms. He was a recipient of the European Research Consortium for Informatics and Mathematics Post-Doctoral Fellowship in 2002 and 2004 to carry out research at the Centre of Mathematics and Computer Science, The Netherlands, and the National Research Institute for Computer Science and Control, Sophia Antipolis, France. He has contributed to MPEG on scalable video coding standardization and wavelet video exploration in 2004 and 2006. He has been a Reviewer for leading IEEE/IET/ACM/SPIE journals and conferences, U.K. Engineering and Physical Science Research Council, Hong Kong Research Council, and book publishers. He has been involved in technical program committees for several leading international conferences and served as a Session Chair and a Guest Editor in journal special issues. He is the United Kingdom's Liaison Officer for the European Association for Signal Processing.