QUALITY ASSESSMENT OF IMAGE AND VIDEO WITH ADAPTIVE WATERMARK METHOD

Mohanakrishnan.P¹*, Dr.K.Suthendran²
Renuka.H³ Suthendran K²
¹Department of Computer Science,
²Department of Information Technology
¹²Kalasalingam Academy of Research and Education,Krishnankoil,Tamilnadu.
³Mets school of engineering, Mala, Calicut University. Kerala

June 29, 2018

Abstract

Digital watermarking based quality evaluation emerges as a potential Reduced- or No-Reference quality metric, which estimates signal quality by assessing the degradation of the embedded watermark. A tree structure based scheme is projected to assign adaptive watermark embedding strengths by pre-estimating the signal degradation characteristics, which greatly improves the computational efficiency. The SPIHT tree structure and HVS masking are used to guide the watermark embedding, which greatly reduces the signal quality loss caused by watermark embedding. Experimental results reveal that the tree structure based scheme can estimate image and video quality with some high accuracy in terms of PSNR and true detection rate. In the case of video, video file is changed into frames and the watermark image is embedded into any one of the frames. The embedding process is carried out using DWT, yet again the
embedded frame and other remaining frames are changed into video file and it is transmitted. Watermark image is extracted from the video in the receiver side. Finally, by using metrics such as MSE, PSNR the quality of both video and watermark image is estimated under different distortion.

1 INTRODUCTION

As presented in [5], the watermark degradation in image is evaluated by comparing the distorted watermark and the original watermark cannot represent the quality of the distorted image by itself. It is always experimentally related to the existing Full-Reference quality metrics like PSNR, SSIM or MOS. Thus, the choosing of the watermark embedding strengths directly affects the accuracy of the quality evaluation. depending on the type of the watermark, the proposed watermarking based image quality evaluation schemes can be categorized into two groups: schemes with image-feature-independent watermarks and schemes with image-feature-dependent watermarks. Paper [6], a randomly generated binary watermark is embedded into the middle-frequency coefficients in the DCT domain using a look-up table method. With this method, each of the possible integer DCT coefficients is randomly assigned one binary bit. For one selected DCT coefficient, if the watermark bit is different from its associated binary bit, this coefficient will be modified to the closest DCT coefficient associated with the binary bit equals to the watermark bit, the experimental results show the embedded watermark degrades monotonically with the increasing of the distortion strength, significant quality loss to the cover image a scheme is proposed in [10] in the DWT domain using a quantization based scheme. With the quantization based method, the odd DWT coefficients are assigned binary bit 0. In contrast, the even DWT coefficients are assigned binary bit 1. Compared to the scheme in [6], the scheme in [10] & [14] reduced the quality loss caused by the watermark embedding process, which is over 40 dB in PSNR.

The research about the scheme in [10] is extended in [8],[9]&[11], In [9],propose a Ideal Mapping Curve and is quantitatively expressed.

The degradation of the watermark is evaluated by computing the
True Detection Rate (TDR) of the distorted watermark compared to the original watermark. In [8] & [9], the adaptive watermark embedding strength is iteratively tested in the watermark embedding process.

schemes with image-feature-dependent watermarks. In [17], the authors proposed to use the statistical features extracted from the three-scale four-orientation steerable pyramid decomposed subbands of the cover image as the watermark in [18]. The scheme is implemented using a relatively complicated Discrete Wavelet decomposition of the cover image. The experimental results show that the scheme works more efficiently with the additive white Gaussian noise than JPEG compression.

2 Watermarking Based Video Quality Evaluation

To solve these problems resulted by the Full-Reference metrics and MOS, the watermarking based video quality evaluation method is proposed as the Reduced- or No-reference video quality metric in literature [12] & [13].

2.1 Schemes with Video-Feature-Independent Watermarks

In [22], a semi-fragile watermarking scheme is proposed to test the quality of the low-bit rate QCIF video sequence. The quality of the distorted video signal is evaluated by calculating the approximated PSNR, a watermarking based video quality metric is implemented in DCT domain for the high bit-rate video sequences [23].
2.2 Schemes with Video-Feature-Dependent Watermarks

In [33], a watermarking based video quality assessment metric is proposed based on the research done in [17]. This scheme is tested under various distortions, such as MPEG-2 compression, Gaussian noise addition and Gaussian blur line jittering. The experimental results show that with the increasing of the distortion strength, the distortion caused to the video sequences increase monotonically. Thus, the goal of our research in this paper is to keep the accuracy of quality estimation achieved in while improving the computational efficiency and reducing the image quality degradation caused by the watermark embedding process. Here, the term accuracy evaluates the correlation of the estimated quality and the quality calculated using the existing objective Full-Reference quality metrics, such as PSNR. The closer the estimated quality to the calculated quality, the more accurate the quality estimation, and vice versa. In this paper, to propose a new approach that well meets our research goal addressed above.

In the proposed scheme, the adaptive watermark embedding strength is estimated by analyzing the quality degradation characteristics of the cover image and no iterative adjustment loops are used, which significantly improves the computational efficiency and theoretically makes the proposed scheme applicable to real-time video quality estimation. Moreover, the strategies including the HVS masking are used to guide the watermark embedding process. With the proposed scheme, the quality of the watermarked images referring to the original images is about 48 dB in PSNR on average, which is an 8-dB improvement over the scheme in . The proposed scheme is based on adaptive watermarking and tree structure in the DWT domain. Recently, the Set Partitioning in Hierarchical Trees (SPIHT) has become one of the most popular image and video coding method because of its efficiency which is accomplished by exploiting the inherent similarities across the sub bands in the wavelet decomposed image [15]. The DWT and SPIHT together provide a good summarization of local region characteristics of an image which is important for adaptive watermark embedding. In this proposed scheme, all the correlated DWT coefficients across the sub bands are grouped together using the SPIHT tree structure. The DWT
decomposed image is further decomposed into a set of bit plane images. In this case, each DWT coefficient is decomposed into a sequence of binary bits. The binary watermark bits are embedded into the selected Bit planes of the selected DWT coefficients of the selected trees [16]. The HVS masking is used to guide the bit plane selection. As found in the experiments, the higher frequency DWT sub bands and less significant bit planes are more sensitive to distortions and vice versa. Therefore, the robustness of the watermark is controlled by two factors: (a) The percentages of the watermark bits embedded into the three DWT levels, respectively, and, (b) The selection of bit planes for watermark embedding. Thus, for different selected trees, the watermark embedding strengths are different [19]. The proposed scheme is tested in terms of PSNR, wPSNR, Watson JND and SSIM, and under JPEG compression, JPEG2000 compression, Gaussian low-pass filtering and Gaussian noise addition. The results show the effectiveness of the proposed scheme.

Problem Definition

The proposed watermark embedding scheme includes the watermark pre-processing, the image pre-analysis, and the watermark embedding. The watermark embedding process consists of the following three steps:

Apply 3-level DWT to the original image to obtain the DWT decomposed image. The 3-level DWT decomposed blocks are shown...
in Fig. 2(a). Thus denotations for the 10 DWT decomposed sub-bands will be used throughout the process. Embed the watermark with adaptive embedding strength using the tree structure based watermark embedded. The output of the watermark embedded is the watermarked DWT image. Apply 3-level inverse DWT to the watermarked DWT image to obtain the watermarked image.

In watermark pre-processing process, a two dimensional original watermark is organized column by column into a one dimensional matrix. The length of the original watermark sequence is len. To increase the probability of the correct watermark bit extraction at the receiver side, every bit in the original watermark sequence is repeated two times to get a redundant watermark sequence for watermark embedding [20]. We set Redundancy=3 and the original watermark sequence is repeated Redundancy-1 times to get the redundant watermark sequence with Redundancy*len bits long. In the image pre-analysis procedure, the texture characteristics and the perceptual masking effects of the cover image will be analyzed respectively in the spatial and DWT domains. Based on this image analysis, the adaptive watermark embedding strength is estimated [21]. The position separation key is used to locate the positions for watermark embedding. The Hash table key is an optional input that can be used to secure the watermark embedding process.

Module Description

1. Spiht Tree Generation And Data Embedding Here we set two criterias. The watermark bits are not embedded into the LL sub band of the DWT decomposed image. The watermark bits are not embedded into the bit planes higher than 5, where the least significant bit plane is bit plane 1. So we only take 3 dwt bands for data embedding will called as l=1,2,3 Each dwt bands is converted to bit planes. The data is embedded in selected bit planes of selected dwt coefficients of selected tree. And the bit plane selection has great impact on image quality. It is done by HVS masking technique. In the dwt band selection, we can select all cells of all 3 bands. Because after we reconstruct all 4 bands into original image the change after watermark embedding will destroy. It will affect
extracted image quality. If an image of size MxN is 512X512 after 3 level dwt decomposition, its size will reduce to \( \frac{512}{8} \times \frac{512}{8} = 4096 \) We avoid LL layer results \( 4096 \times \frac{3}{4} = 3072 \) trees. But we can only use 1 band (different cells of 3 bands). So the ratio is again \( 3072 \times \frac{1}{3} = 1024 \) bit plane positions are remaining. We can embed watermark + Redundancy bits into this positions using Embed formula is given below.

\[
C_w = c, if c = w, if c \neq w
\]

Fig.2(a) Illustration of the 3-level DWT decomposed sub bands (b) and the formation of tree structure

3 The Selection Of Trees And Dwt Co-efficients

In this paper, we call the watermark bit assignment as \( A_{wb} = [a_1, a_2, a_3] \) where \( a_1, a_2, a_3 \) are the number of watermark bits to be embedded in the DWT level 1, 2 and 3 in every selected tree. In order to embed watermark, the redundant watermark sequence is divided into \( W_{segs} \) segments as depicted in (1).

\[
W_{segs} = \left[ \frac{Redundancy \times \text{len}}{\sum A_{wb}} \right] = \left[ \frac{Redundancy \times \text{len}}{a_1 + a_2 + a_3} \right] \quad (1)
\]

Where len is the length of the watermark sequence. If we denote the numbers of rows and columns of the DWT decomposed image
as M and N, there will be totally Wsegs trees out of $M/2^N, N/2^3, 3/4$ trees selected for the watermarking processes.

In Fig.3(b), the X marked positions represent the trees selected from the three DWT orientations for the watermark embedding. The marked positions represent the separation between any two selected tree positions. In our implementation, we use uniform separation defined as Equ.(2)

$$N_{sep} = \left\lfloor \frac{T_{NP}}{W_{segs}} \right\rfloor - 1 \quad (2)$$

The calculated Nsep will be output as the position separation key and will be transmitted to the receiver side. Here, we set Wsegs TNP. The trees are selected throughout the DWT decomposed image. The wseg selected trees are further distributed into the three orientations referring to $T_{per}$. In the proposed scheme, we set $T_{per} = [1/3, 1/3, 1/3]$ which means that the trees are evenly selected from the three orientations as illustrated in Fig. 3(a). In this figure, the selected trees are numerically ordered using 1, 2, ..., $W_{segs}$. As presented in Equ. (2), the watermark sequence is divided into $W_{segs}$ segments. Thus, the watermark will be embedded into the selected trees segment by segment following the order number of the selected tree position.

Fig.3 The tree selection strategy
After the tree selection, we start to embed the watermark bit segments into the selected trees referring to Awb. On the $l$th DWT level, the watermark bits are embedded into the DWT coefficients one by one until the number of is $Awb(l)$ reached, where $l=1,2,3$. An experimental example of the tree and DWT coefficients selection is shown in Fig. 4 (b). All the selected DWT coefficients are marked as dark points. The darker points means that the lower significant bit planes are selected, and vice versa. In Fig.4 (b), the position separation key is 3.

4 Security Of The Watermarking Process

In the proposed tree structure based scheme, an optional security function is designed to meet some special requirements of users to securely transmit signals over internet. This security function is controlled by the HashTable Key.
Fig. 5 Illustration of the secure reordering of the selected trees.

The HashTable key is also a position key for the tree selection and is used to reorder the selected trees. An illustration of the secure reordering of the selected trees using the HashTable Key, 0, is shown in Fig. 5. In this case, the HashTable Key is also needed to be sent to the receiver side and the lack of the HashTable key will result in a failure of watermark extraction.

5 The Analysis of Image Content Complexity

The quad-tree decomposition based complexity analysis is used in the proposed scheme for a better match with the DWT. For gray scale images, the intensity difference, \( V_{\text{int}} \), is used to verify whether further quad-tree decomposition is needed. Here, we define an intensity difference threshold as \( T_{\text{int}} \). If \( V_{\text{int}} \geq T_{\text{int}} \), the image or current block will be decomposed into 4 sub-blocks until \( V_{\text{int}} \leq T_{\text{int}} \) or the size of the sub-block reaches 1. Each quad-tree decomposition is recorded as a decomposition node. The depth of the decomposition is denoted as the level of decomposition. The content complexity of the cover image is assessed using the following equation:

\[
\text{Complexity} = \sum_{i=1}^{n} (N_i X 2^i)
\]
where \( i \in [1, N] \) is the current quad-tree decomposition level; \( n \) is the highest decomposition level; \( N_i \) is the number of quad-tree decomposition nodes on level \( i \). Then, the calculated complexity values of all the images in our image library are normalized. In the proposed scheme, the normalized complexity value is used as the complexity index which locates in \([0, 1]\).

The image content complexity analysis evaluates many details that an image carries. A higher complexity value indicates that the image is more complex and the image contains more detail information \([24]\) & \([25]\). Comparing to a less complex image, the quality of a more complex image degrades faster against the same distortion. For this case, to reflect the quality degradation of the cover image, we need to embed more watermark bits into the lower DWT levels of a more complex image. For a less complex image, we consider to embed more watermark bits into the higher DWT levels. Three experimental examples are shown in Fig.6. These quad-tree decomposed images are achieved using the threshold \( T_{int}=0.17 \), where the maximum intensity value of the cover image is not bigger than 1. In this case, the brighter the quad-tree decomposed image, the more complex the cover image. The complexity indices are listed with the quad-tree decomposed images.
6 HVS Masking

The bit plane selection includes two steps: calculating the HVS masks and mapping the calculated HVS masks to biplane indices [26]. The achieved biplane indices decide which bit planes of the selected DWT coefficients are used to embed the watermark bits. The HVS masking calculation and the HVS-to-bit plane mapping
are presented in the following two subsections. 1) The HVS Masking: The HVS masking presented in is used in the proposed scheme. Accordingly, four factors greatly affect the behaviour of the HVS: (a) Band sensitivity or frequency masking: Intensity variations are less visible in high resolution sub bands and are also less visible in the diagonally decomposed blocks, HHI. This factor is expressed using (3).

\[ M_F(1, \theta) = M_1(\theta).M_2(1) \]  

where

\[ M_1(\theta) = \begin{cases} \sqrt{2}, & \text{if } \theta = 2 \\ 1, & \text{otherwise} \end{cases} \]

\[ M_2(1) = \begin{cases} 1, & \text{if } l = 1 \\ 0.32, & \text{if } l = 2 \\ 0.16, & \text{if } l = 3 \end{cases} \]

\[ \theta = \begin{cases} 1, & \text{for HL blocks} \\ 2, & \text{for HH blocks} \\ 1, & \text{for LH blocks} \end{cases} \]  

(b) Background luminance: Intensity variations are less visible over the brighter and darker areas. The luminance masking is denoted as \( M_L \).

\[ M_L(l, i, j) = 1 + I(l, i, j) = \begin{cases} 2 - \frac{1}{256} ILL \left( \left[ \frac{i}{2}, \frac{j}{2L_e - l} \right] \right), & \text{if } I(l, i, j) \neq 0.5 \\ 1 + \frac{1}{256} ILL \left( \left[ \frac{i}{2L_e - l}, \frac{j}{2L_e - l} \right] \right), & \text{otherwise} \end{cases} \]  

(c) Spatial masking or edge proximity: The human eyes are more sensitive to noise addition near edges or contours of images. This factor, \( ME \), is evaluated using the empirically scaled local energy of the DWT coefficients in all detail sub bands.

\[ M_F(l, i, j) = \sum_{k=0}^{L_e-l} \sum_{\theta=1}^{3} \sum_{x=0}^{1} \sum_{y=0}^{1} |I_{k+l}(x + \left[ \frac{i}{2^k} \right], y + \left[ \frac{j}{2^k} \right])|^2 \]  

13
where $\rho$ is a weighting parameter and the suggested value for $\rho$ is presented in the following equation:

$$\rho = \begin{cases} 
\frac{1}{4}, & \text{if } k = 0 \\
\frac{1}{16k}, & \text{otherwise}
\end{cases}$$  \hfill (7)

(d) Texture sensitivity: Intensity variations in highly textured areas are less visible than those in the flat-field areas of images. This masking factor $M_T$ is estimated using the local variance of the corresponding DWT coefficients in the LL sub band.

$$M_T(l, i, j) = \text{var}\left\{ I_{LL}\left(x + \left[\frac{i}{2^L e - 1}\right], y + \left[\frac{j}{2^L e - 1}\right]\right)\right\}$$  \hfill (8)

where $X=0.1$ and $y=0.1$.

The HVS mask is achieved by polling the four factors listed above and is computed using (5).

$$M_{HVS} = \alpha.M_F.M_L.M_E^\beta.M_T^\gamma$$  \hfill (9)

Where $M_{HVS}$ denotes the HVS mask; $\alpha$ is a scaling parameter. The suggested value for $\alpha$ is $1/2$, which implies that intensity variations having values lower than half of $M_F.M_L.M_E.M_T$ are assumed invisible. The suggested value for $\beta and \gamma$ is $0.2$. In Equation (3) to (8), $(i, j)$ are the coordinates of the current pixel in the calculation. $I(l,i,j)$ is the luminance value of the pixel $(i,j)$, in one detail sub band on level $l$; $le1,2,..,L$ indicates the current DWT level in the HVS masking calculation and $L_{c}=3$ is the maximum level applied in the DWT decomposition in this paper.

$\left[\frac{i}{2^L e - 1}\right], \left[\frac{j}{2^L e - 1}\right]$ is the luminance value of pixel $(\left[\frac{i}{2^L e - 1}\right], \left[\frac{j}{2^L e - 1}\right])$ in the LL sub band which corresponds to pixel in a detail sub band on level 1.
In this work, the binary watermark bits are not embedded in the sub band so that the invisibility of the embedded watermark can be further improved. Therefore, the HVS masks are only calculated for the nine detail DWT sub bands, with one mask for each. The generated HVS masks for image Barbara is shown in Fig. 7(b).

2) Mapping from HVS Mask to Biplane Indices: To use the HVS mask in the watermark embedding process, a mapping relationship from the coefficients of the HVS mask to the biplane indices is experimentally defined using the multiple-band threshold method. Considering that different HVS masks have different distributions, the thresholds used in the mapping procedure should be able to change with the shape of the distribution of the HVS mask [27]. To further limit the quality degradation caused by the watermarking process, only the bit planes from 1 to 5 are used for the watermark embedding. The thresholds are calculated using (9).

\[ T_n(1, \theta) = \text{Sort} \left( \left\lfloor \frac{nM_{\text{mask}}(1, \theta)}{5} \right\rfloor \right) \quad (10) \]

where \( n \in 1, 2, 3, 4 \), \( N_{\text{mask}} \) is the total number of the coefficients of one HVS mask, the denominator indicates that 5 bit planes are used for the watermark embedding. The mapping is done using (10).
where \((i,j)\) are the coordinates of a selected DWT coefficient. 
\(I_{bp}(l, \theta, i, j)\) means the bitplane index achieved for the pixel located at \((i,j)\) on DWT level \(l\) with orientation \(\theta\). \(v\) is the value of the HVS mask coefficient.

Therefore, each DWT coefficient in the selected trees has its own \(I_{bp}(l, \theta, i, j)\). At the receiver side, the strategy presented above will also be applied on the distorted image to locate the biplanes for the watermark extraction [28]. To increase the probability of correct extraction of the watermark bits, all the calculated \(I_{bp}(l, \theta, i, j)\) values on the DWT level \(l\) at orientation \(\theta\) are averaged. In other words, in each selected tree, for all the selected coefficients, the watermark bits will be embedded on
the same bit plane. Thus, the bit plane indices are updated using the following equation:

$$I_{bp}(l, \theta, i_{tree}, j_{tree}) = \left[ \sum_{i_r=0}^{(L_e-1) - 1} \sum_{j_r=0}^{(L_e-1) - 1} I_{bp}(l, \theta, i, j) \right] \left(2^{(L_e-1) + 1}\right)^2$$

where $I_{bp}(l, \theta, i_{tree}, j_{tree})$ is the averaged bitplane of the DWT coefficients located in a specific tree on level $l$ and orientation $\theta$.

$$i = i_{tree} + i_r = \left[ \frac{i}{2^{L_e-1+1}} \right] + \text{rem}(i - 1, 2^{L_e-1+1})$$

$$j = j_{tree} + j_r = \left[ \frac{j}{2^{L_e-1+1}} \right] + \text{rem}(j - 1, 2^{L_e-1+1})$$

Two examples of the thresholds selection are shown in Fig. 8. Fig. 8(b) and (c) are the HVS masks respectively calculated on the DWT sub band HL1 of image Barbara and image Tree frog. Thus, the size of these two HVS masks is 1/4 of the image size. Then the HVS-to-bit plane mapping thresholds are calculated using (5). Fig. 8(c) and (f) graphically illustrate the threshold selection.

The Watermark Bits Assignment: With the complexity indices, the watermark bits are empirically assigned to the images using the following steps: (a) The complexity indices are divided into 6 groups. One integer index is associated with one group.

$$G = \begin{cases} 
1, & v_c > T_1 \\
2, & T_1 \geq v_c > T_2 \\
3, & T_2 \geq v_c > T_3 \\
4, & T_3 \geq v_c > T_4 \\
5, & T_4 \geq v_c > T_5 \\
6, & T_5 \geq v_c > 0 
\end{cases}$$

Where $G_{index}$ is the group index; $V_c$ is the complexity Index $t_1, t_2, t_3, t_4, t_5$, and are the empirical grouping thresholds. These thresholds may be different for different distortions. With the group indices, the watermark bits are assigned to the images using

Two examples of the thresholds selection are shown in Fig. 8. Fig. 8(b) and (c) are the HVS masks respectively calculated on the DWT sub band HL1 of image Barbara and image Tree frog. Thus, the size of these two HVS masks is 1/4 of the image size. Then the HVS-to-bit plane mapping thresholds are calculated using (5). Fig. 8(c) and (f) graphically illustrate the threshold selection.

The Watermark Bits Assignment: With the complexity indices, the watermark bits are empirically assigned to the images using the following steps: (a) The complexity indices are divided into 6 groups. One integer index is associated with one group.
Recall in Equation (1), trees are selected for the watermark embedding, according to the length of the redundant watermark sequence. Thus, for images with different complexities, the number of selected trees, \( W_{seg} \), and the position separation key \( N_{sep} \), may be different.

### 7 Watermark Extraction

The extraction process consists of applying inverse of all the process. Apply inverse DWT to the received signal. The position separation key is used to locate the watermarked DWT coefficients [30]. The watermark bit assignment is retrieved using the image group index transmitted from the sender side. The bit plane indices for watermark extraction are obtained by calculating the HVS masks of the distorted watermarked image. In one tree, the bit plane indices for all the DWT coefficients on each DWT level are averaged. This strategy effectively reduces the watermark extraction error caused by the bit plane selection in the watermark extraction scheme. Recall that Redundancy=3. The extracted redundant watermark sequence is used to recover the three distorted watermarks. Then, the three distorted watermarks are compared bit by bit and the watermark is extracted using

\[
W_e = \begin{cases} 
1, & N_1 \geq N_0 \\
0, & N_1 < N_0 
\end{cases}
\]

where \( W_e(i,j) \), is the extracted watermark bit with coordinates\((i,j)\); \( N_1 \) is the number of extracted 1s and \( N_0 \) is the number of extracted 0s. Then, the extracted watermark is compared with the original
watermark bit by bit and the True Detection Rates (TDR) is calculated using equation
\[
TDR = \frac{\text{Number of correctly detected watermark bits}}{\text{Total number of watermark bits}}
\]
The TDR is defined as an extension of the False Detection Rate (FDR) proposed in [5]. The TDR locates in [0, 1] and TDR + FDR = 1. MSE and PSNR are calculated for Quality analysis [31]. With the MSE quality metric, it is assumed that the original signal is a perfect signal and any difference appears in another signal comparing to the original signal is treated as error [1] & [2]. The MSE measures the average of the squares of errors. MSE is defined as:
\[
MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (I(i,j) - I'(i,j))^2
\]
where, \( I \) and \( I' \) respectively are the original image and distorted image; \( (i, j) \) are the coordinates of the current pixel in the image; \( M \) and \( N \) are respectively the numbers of rows and columns of the image; \( MN \) is the total number of pixels in the image. In literature, sometimes, MSE and RMSE are used to evaluate video quality. In this case, \( I \) and \( I' \) are the original video frame and the distorted video frame. PSNR is a MSE based quality metric and performs quality evaluation by comparing the pixelwise differences between the distorted image or video frames and the original image or video frames [3] & [4]. Till now, PSNR is one of the most widely used quality evaluation metrics. PSNR is defined as:
\[
PSNR = 10 \log_{10} \frac{\text{MAX}}{\sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (I(i,j) - I'(i,j))^2}}
\]
where, \( \text{MAX} \) is the maximum possible pixel value of an image or video frame. For an 8 bits/pixel grey-scale image, \( \text{MAX} \) is equal to 255. For a normalized image or video frame, \( \text{MAX} \) equals to 1

Process Flow Diagram
Applying watermarking to the video frames can be a complex process. The raw video signal can be defined as a consecutive sequence of still images. Thus, it is theoretically straightforward to further develop the watermarking based image quality evaluation metrics for video quality assessment. Similar to the image quality assessment metrics, the video quality metrics can be classified as the Full-Reference, Reduced-Reference and No-Reference quality metrics [7]. The widely used Full-Reference video quality metrics include the MSE, PSNR, VQM (Video Quality Metric) and VIF (Visual Information Fidelity) [32]. One watermark can be redundantly embedded into a number of selected frames or different watermarks can be embedded into different video frames with the same or different embedding strength(s). In this case, the quality of a watermarked video sequence can be evaluated by calculating the average degradation of the embedded watermarks. With the watermarking based quality metrics, the quality of the distorted video signals is usually evaluated in terms of PSNR or MSE [34]. The video quality is usually assessed by comparing the distorted watermark to the original watermark.
Result and Discussions

In our work, there are 5 test images present in our image library. Besides these image databases, our image library also includes computer generated images and more natural images. All of these images are in any size containing different textures, such as, portraits, plants, animals, animations, sceneries, buildings and crowd. The original watermark used in the experiments is 100x100 in size. The binary watermark is randomly generated and there is no special requirements for it. The original watermark can be changed to any size or pattern. Recall that $A_{wb} = [a_1a_2a_3]$ is the watermark bits assignment which assigns $a_1$ bits of watermark to the DWT level 1 in a single-root tree; $a_2$ bits of watermark to the DWT level 2 in a single-root tree and $a_3$ bits of watermark to the DWT level 3 in a single-root tree. The higher the complexity value, the more complex the image. The complexity value calculated for the image Jellyfish is 0.69012. The jelly fish, needs a weak watermark embedding strength to reflect the degradation caused by the distortion. The TDR value of jelly fish is 0.0521. And the quality factor PSNR is 61.024. The calculated complexity value for image Baboon is 0.030079, its TDR value is 0.1018 and PSNR is 58.09. It needs a strong watermark embedding strength. The next image ship has complexity-0.016617, PSNR-61.223 and TDR value is 0.0497. The PSNR values of the entire watermarked image are more than 55.5 dB, which proves that the quality degradation caused by the watermark embedding process is very limited.

Table. 1. Result Analysis Of various input images
Conclusion and Future Work

In this paper, a new watermarking based quality estimation scheme is presented. The proposed scheme is designed to estimate image quality in terms of the existing Full-Reference quality metrics, such as PSNR, WPSNR, JND and SSIM. Thus, at the receiver side of a communication system, without the original image, the quality of a distorted image or video can still be assessed. Based on the tree structure, the binary watermark is embedded into the selected bit planes of the selected DWT coefficients with adaptive watermark embedding strength. The watermark embedding strength is assigned to an image by pre-analyzing its content complexity in the spatial domain and the perceptual masking effect of the DWT decomposed image in the DWT domain. Meanwhile, the watermark is not embedded in the approximation sub band, which reduces loss in image quality caused by embedding the watermark. The experimental results show that the proposed scheme works effectively.

In future work, the proposed scheme will be further developed to estimate the quality of an image distorted by multiple distortions. Meanwhile, experiments about image quality estimation in terms of subjective quality scores will be conducted. Since the proposed scheme has good computational efficiency, it is feasible to further develop the proposed scheme for audio quality evaluation. The quantization based image quality evaluation scheme has relatively low computational efficiency. The lack of consideration of the human perception characteristics introduces relatively more significant quality loss caused by the watermark embedding process to the cover images. The quantization based scheme works effectively under JPEG compression. The tree structure based image quality estimation scheme utilizes the SPIHT structure and HVS mapping to guide the watermark embedding process.

As the future work, the proposed scheme can be further tested with the quality evaluation in terms of audio signals. Moreover, the qual-
ity of image or video signals acted by multiple distortions can also be tested using the proposed scheme.

References


