Chapter 6

Object Coding Using a Shape Adaptive Wavelet Transform with Scalable WDR Method

6.1. Introduction

The coding of imagery with shapes more general than rectangular frames has become an important issue in multimedia communications. The emergence of multimedia operations like access, searching, indexing and manipulation of visual information at the semantically meaningful object level, are becoming very important in research and standardization efforts [166-181].

There are two reasons for which object based compression may prove to be of advantageous.

1. Increased functionality
2. Increase compression performance

Where as increased functionality is inherently achieved by the object based approach, increased compression performance by adaptive object based coding is not necessarily achieved. The idea here is to segment an image into objects of different characteristics and then adaptively encode each object separately, thus increasing compression performance. Even if this aim is not met and compression efficiency remains at least constant, the approach has to be considered a success if the overhead introduced by the increased functionality can be balanced by the adaptively increased compression efficiency.

The straight forward approach in handling arbitrarily shaped image objects is to consider pixels within the object as “opaque” and those outside the object as “transparent”, and to adapt rectangular frame coders to process only the opaque regions within the bounding box of the object [166].

One of the outstanding features of the MPEG-4 standard is the possibility of object based images in video data. In order to support object based coding functions, object based representation and compression of video signals are required. There have been continuous efforts in developing coding techniques for arbitrary shaped video objects. Among them, the shape-adaptive DCT (SA-DCT) is the most popular one and it has been applied to the MPEG-4 verification model [180, 181].
There are some disadvantages on shape adaptive object coding when we use DCT techniques.

1. It inherits the blocking effect caused from DCT
2. During the implementation of SA-DCT, the alignment of the coefficients destroys the spatial correlation to some extent and therefore, the coding efficiency is degraded.

Since, the wavelet transform scheme can avoid the block effect introduced in the DCT scheme, wavelet based schemes for shape adaptive object and coding was proposed [180, 181].

The video objects are arranged into rectangular shape of even length in the horizontal and vertical directions in order to apply the conventional transform. So, the number of transform coefficient is larger than that of the original video objects which will negatively affect the coding efficiency [166]. Thus, the coding of arbitrary shaped visual objects has been increasingly important in recent times. The object code has more and more application like content based storage and retrieval systems. Moreover, very low bit rate coding of objects enables the specific content based coding. These types of significant information are included in the MPEG-4 standard. The shape adaptive object coding is a new paradigm and is different from the traditional coding schemes such as discrete cosine transform based JPEG coding standard and discrete wavelet transform based JPEG2000 standard. These methods are limited to code [35-43] the rectangular shapes.

There are many DCT and DWT based Shape Adaptive (SA) coding [166-181] techniques. These are based on the variation of embedded wavelet coders like EZW and SPIHT. The shape adaptive DCT and block based shape-adaptive DCT [180, 181] are used to transform the pixels in (8x8) bounding box to the top edge of the box and performing the 1-D DCT on each column then flushing to the left and performing the 1-D DCT on each row.

There are many situations to reconstruct the ROI before the background reconstruction in many image coding applications such as web browsing, image database, remote sensing image coding and telemedicine. Many applications need firstly to assure the PSNR of the ROI to be higher than a certain value. For example, in telemedicine, the ROI, which contains the important diagnostic information, needs to be compressed
without loss, i.e. the PSNR of ROI should be infinite very large.

During the analysis of DCT-based shape-adaptive coding, the discrete wavelet transform is identified as identical for coding efficiently. Li [166], in his work used the shape adaptive discrete wavelet transform (SA-DWT) has proposed as a more efficient means of decomposing arbitrary shaped objects. His new scheme SA-DWT avoids the blocking artifacts inherent to DCT based techniques. He uses an FIR wavelet filter set either in bi-orthogonal or orthogonal form for decomposition of individual 1-D image segments with in the object. Here, the operation involves the scanning each row of the object for continuous segments of pixels and performing arbitrary length wavelet decomposition on each one. The operation is done on each row and the same is repeated on the each column.

The SA-DWT and the accompanying zero tree coding algorithms have been adopted as the methodology for still Visual Texture Coding (VTC) [176] in the MPEG-4 standard. Similar coding schemes have been proposed for coding the texture of video objects. These methods or schemes are generally using the adaptation of EZW and SPIHT, where by coefficients residing outside the arbitrary shaped visual objects are effectively discarded.

The object shaped mask, represented as a binary image must be coded separately using a binary mask coding scheme and fully decoded prior to the decoding of texture. The whole shape coding is done in such a manner that will cause a negative impact on the computational overhead for coding and decoding. For lossy shape reconstruction, shape and texture coding must still be implemented separately, with some shape information decoded before texture decoding in order to limit the region of the reconstructed texture upon performing the reverse wavelet transform [172].

There are many efficient shape adaptive coding schemes. Among them, Martin and Lukac [168] proposed an efficient method to code shape and texture into one embedded bit stream codes. The proposed procedure results in a multi-resolution of the binary shape mask as well as the texture. We scan the mask row by row and get the beginning position and end positions. The scanning proceeds through the column by column and gets the beginning position and end position.

The implementation of the SA-DWT is done based on the work proposed by Li [166] and also implement an in-place-lifting scheme [169] for integer wavelet transform. This
implementation is highly efficient as it allows the use of the spatial domain shape mask directly in the transform domain without any manipulation.

If integer-to-integer wavelet filters are used in the SA-DWT, lossless reconstruction of the object may be achieved. With the use of advanced floating point filters, such as the bi-orthogonal 9/7 tap filters, high quality lossy results can be achieved. The functionality of making visual objects available in the compressed form has become a very important feature in the next generation visual coding standards such as MPEG-4, since it provides great flexibility for manipulating visual objects in multimedia applications and could potentially improve visual quality in very low bit rate coding [166].

There have been a considerable amount of research efforts on coding rectangular-shaped images and video, such as discrete cosine transform (DCT) coding and wavelet transform coding [180, 181]. It is required to find the bounding box of the arbitrary shaped visual object and pad the pixel values into the pixel positions outside the objects. Then, code the pixels inside the object and pad the pixels in the rectangular bounding box together using the conventional methods. However, this approach would not be efficient.

The wavelet transform is applied on arbitrary shaped visual objects for coding efficiently [147, 166]. There are many techniques for coding arbitrary shaped visual objects with coding coefficient selection and coding coefficient insertion techniques to improve the coding efficiency [147, 166]. Moreover, the micro block region based wavelet coding which pads undefined region with zeros and then applies the wavelet transform on the padded rectangular region.

The zero tree coding methods are used to code the arbitrarily shaped visual objects and the results inevitably blurs the edges of the objects and also results in increasing the number of coefficients that is to be coded than the number of pixels in the object [166].

A novel shape adaptive discrete wavelet transform (SA-DWT) for arbitrarily shaped object coding was proposed by Li et al [166]. This SA-DWT scheme can be directly applied to the arbitrarily shaped region. The SA-DWT transforms the samples in the arbitrarily shaped region in to the same number of coefficients as in the subband domain, while keeping the spatial correlation locally, and self-similarity across subbands.

The important point is that the number of coefficients after SA-DWT is identical to the number of pixels contained in the arbitrary shaped image region. Keeping the number of wavelet coefficients in the transformed image same as the number of pixels is a
necessary condition for an efficient coding method. On the other hand, lossy shape adaptive coding may reduce the number of pixels to be coded at the expense of shape quality. As mentioned previously, there are two aspects of coding an arbitrary shaped visual object [146-181]. The first is to code the shape of the object using lossy or lossless coding method. If a lossy shape coding method is used, the number of pixels within the reconstructed shape is usually less than that within the original shape. Therefore, the number of pixels in the image domain is referred to as that within the reconstructed shape. In other words, the trade off between the number of pixels to be coded and the quality of shape is made at the time of shape coding. The SA-DWT just maintains the number of wavelet coefficients to be the same as the number of pixels to be coded. Moreover, the spatial correlation and other properties like locality and self similarity across subbands are well preserved in the SA-DWT. For rectangular region, the SA-DWT becomes identical to a conventional wavelet transform [166].

For most current DWT-based compression technologies for still images, a common approach is to list the wavelet coefficients, where the lift factor need be predetermined manually [166], JPEG2000 still image compression standard supports ROI coding, which lifts the coefficients directly also. If after lifting the lowest bit plane of the wavelet coefficients of the ROI is higher than the highest bit plane of those of the BG, it is called the max shift lifting method, otherwise it is called the general scaling based method. Many methods have been proposed to adjust compression quality in the ROI and back ground. If the lift factor is too small, the PSNR of the reconstructed ROI is too small to satisfy the requirement, otherwise the PSNR of the reconstructed BG is too small to satisfy the basic requirement. Because the complexity of different images vary greatly, it is difficult for the ROI coding based on coefficient lifting to adjust the lifting factor automatically to attain a balanced effect between the ROI and the BG (Back Ground).

There are many arbitrarily shaped ROI coding methods based on ISA-DWT, which can overcome the unbalance between the ROI and the BG under the condition of guaranteeing the quality of the reconstructed ROI. ISA-DWT is used for the arbitrary shape ROI coding and fit for variable coding applications such as sensing and telemedicine coding that need loss and lossless compression. ISA-DWT maintains similar spatial correlation to the actual wavelet transform and the number of coefficients in the transform domain using the ISA-DWT is equal to the number of pixels contained in the
region [166].

6.2. Shape Adaptive Discrete Wavelet Transform

The shape adaptive wavelet transform is the process of handling the wavelet transforms for arbitrary length image segments. It considers the sub-sampling method for arbitrary length image segments at arbitrary locations. The shape adaptive DWT transforms the odd-or-even small length image segments to be decomposed into the transform domain and maintains the number of transform coefficients in the transform domain identical to the number of pixels in the image domain [166].

The shape adaptive discrete wavelet transform generally depends on the proper sub-sampling method. It considers or preserves the spatial correlation and self similarity property of wavelet transforms. It will consider the two dimensional separable wavelet decompositions and pyramid wavelet decompositions that can be applied to the arbitrary shaped image segments or region without loss of spatial correlation.

The shape adaptive discrete wavelet transform depends on the wavelet filters used for the wavelet decomposition of an arbitrary length segment. There is odd symmetric and even symmetric bi-orthogonal wavelet filters commonly used for decomposing the arbitrary shaped objects [166]. Here, we are considering the odd symmetric bi-orthogonal wavelet filters used to decompose the arbitrary shaped objects, because the rate distortion performance of odd symmetric bi-orthogonal wavelet filter is much better than the even symmetric filters [166]. The odd symmetric bi-orthogonal wavelets have bi-orthogonal and symmetric low and high pass filters and odd number of taps for each filter.

Accounting properly the boundaries of the signal segment is one of the issues in applying wavelet decomposition to a finite length signal segment. The perfect reconstruction property of the wavelet transform can be maintained by filling the undefined pixel values outside the finite length signal segment as boundary extension. Generally, the symmetric extension is common for most proper and perfect reconstruction.

6.2.1. Arbitrary Length Wavelet Decomposition using Odd Symmetric Bi-orthogonal Wavelets

The usual extension is applied on the signal segment before the wavelet transform is periodic all extracted. However, if the segment is long, the correlation between the end of the signal and the start of the signal is small. So there is a probable chance of a sharp
change at the transition from the end of the previous signal period to the start of the next signal period if the periodic extension method is used. The usual and most commonly used extension type is symmetric extensions. However, to ensure perfect reconstruction, symmetric extensions are only possible for symmetric bi-orthogonal wavelet transforms [166].

The signal is extended symmetrically at the leading and trailing boundaries of a signal segment and neighbouring samples with such symmetric extensions have the same close correlation as in the original signal segment. It will avoid the sharp transitions. The bi-orthogonal wavelet provides linear filters also and eliminates the phase distortion caused by magnitude distortion of transformed coefficients. This is very important when they are applied to signal compression where magnitudes of the transformed coefficients are most likely to be quantized [166].

We use as in [166] two analysis filters with impulse response of low pass analysis filter \( \{g(i), i = 0, \ldots, L_g-1\} \) and high pass analysis filter \( \{h(i), i = 0, \ldots, L_h-1\} \). There are also two synthesis filters with impulse response of low pass synthesis filter \( \{e(i), i = 0, \ldots, L_e-1\} \) and high pass synthesis filter \( \{f(i), i = 0, \ldots, L_f-1\} \). The filter lengths, both \( L_g, L_h \) are odd numbers. Here \( L_g = \{9 \text{ or } 5\} \) and \( L_h = \{7 \text{ or } 3\} \).

Let \( X(i) \) be the input signal with a finite length with approximate extensions at the leading and trailing boundaries.

Assuming a signal segment \( \{X(j), j = 0, \ldots, N-1\} \), with length of \( N \), and combining symmetric extensions, filtering, and sub sampling together, the arbitrary length wavelet decomposition using odd symmetric wavelet transforms can be described as follows [166].

1) If \( N = 1 \), this isolated sample is repeatedly extended and the low-pass wavelet analysis filter is applied to obtain a single low pass wavelet coefficients. This is equivalent to scale this sample by a factor \( K = \sum_{i=0}^{L_g-1} g(i) \) and it happens to be \( \sqrt{2} \) for some normalized bi-orthogonal wavelets. The synthesis process simply scales this single low pass wavelet coefficient by a factor of \( 1/K \) and puts it in the correct position in the original signal domain.

2) If \( N \) is greater than 1 the signal segment is extended using symmetric extension mode. The sub sampling and convolution procedure is shown in the
The convolution of signal using the low pass analysis filter $G$ is shown in the figure 6.1 [166] The convolution of signal segment using the high pass analysis filter $H$ is shown in figure 6.2 [166] The two segment of transformed signal is passed through the channel and used for the compression. The de-convolution process is done using the low pass synthesis filter $E$ and is shown in the figure.6.3 and the high pass synthesis filter $F$ is shown in the figure.6.4 the resultant signals are added to get the reconstructed signal.

![Figure 6.1 Representation of convolution procedure using low pass analysis filter $G$](image)

![Figure 6.2 Representation of convolution process using high pass analysis filter $H$](image)

The arbitrary length signal segments long or short, even or odd are decomposed using the length adaptive wavelet transforms [166]. For each case of arbitrary length wavelet transforms, two options in sub-sampling the low-and high pass wavelet coefficients, i.e. even sub sampling and odd sub sampling. Different sub sampling strategies have different advantages and disadvantages in terms of coding efficiency.
Object Coding Using a Shape Adaptive Wavelet Transform with Scalable WDR Method

The signal segments in an arbitrary shaped visual object are neither all starting from odd position nor all starting from even positions. The phases of some of the low-and high pass wavelet coefficients may be skewed by one sample when sub sampling is locally fixed for all signal segments. This is not desired for wavelet decomposition in the second direction. Since the phases of the sub-sampled wavelet coefficients differ at most by one sample, the spatial relations [166] across subbands can still be preserved to a certain
extent.

In contrast, another sub-sampling strategy does not fix even or odd sampling locally for all the signal segments. It strictly maintains the spatial relations across the subbands by using either even or odd sampling according to the position of a signal segment relative to the bounding box. Instead of fixing sub-sampling positions locally for each signal segment, this strategy fixes sub-sampling positions of the high and low pass wavelet coefficients at globally even or odd positions relative to the bounding box. Instead of fixing sub sampling positions locally for each signal segment, this strategy fixes sub-sampling positions of the high and low pass wavelet coefficients at global even or odd positions relative to the bounding box.

Since, the starting position of each segment in a visual object may not be always at an even or odd position, the local sub-sampling positions in each segment have to be adjusted to achieve global even or odd sub-sampling. For example, with odd symmetric bi-orthogonal filters, to achieve global even sub-sampling in low pass bands and global odd sub-sampling in high pass bands, we need to choose local even sub sampling in low pass bands and local odd sub-sampling in high pass bands for all segments starting from even positions, and choose local odd sub sampling in low pass bands and local even sub sampling in high pass bands for all segments starting from odd positions.

6.2.2. 2D Shape-Adaptive Discrete Wavelet Transform

The two dimensional shape adaptive discrete wavelet transform algorithms [166] and sub-sampling strategies are done on the arbitrary shaped visual object. The process is done based on the following steps.

1. The bounding box of the arbitrary shaped visual objects with the shape information is used to identify the first row of pixels belonging to the object to be transformed.

2. Identify the consecutive pixels of the first segment of the visual object.

3. The length adaptive 1-D wavelet transform is applied on this segment with a proper sub-sampling strategy.

4. The wavelet transform coefficients obtained from decomposition using the low pass filter coefficients are placed into the corresponding row in the low pass band and from the high pass filter coefficients are placed into the corresponding row in the high pass band.
5. The shape adaptive wavelet transforms operations are carried out for the next segment of consecutive pixels in the rows, one by one downwards.

6. The same operations are carried out for each column of the low pass and high pass objects.

7. Perform the above operations to the low pass object to get all levels of wavelet decomposition is reached.

The 2D shape adaptive wavelet transform provides a way to efficiently decompose an arbitrary shaped visual object into a multi-resolution object pyramid. The spatial correlation, locally, and object shape are well preserved. Thus, it enables multi-resolution coding of arbitrary shaped objects. This method ensures that the number of coefficients to be coded in the transform domain is exactly the same as that in the image domain. If the object is a rectangular image, the 2-D SA-DWT is identical to the standard 2-D wavelet transform. The 2D shape adaptive wavelet transform of arbitrary length objects and its subband structure is shown in the figure 6.6 [166].

![Figure 6.6 Representation of 2D-Shape Adaptive DWT for arbitrary length objects](image)

6.2.2.1. Arbitrary Shaped Region of Interest (ROI) Coding based on SA-DWT

The region of interest (ROI) coding [40] is done to reconstruct before the Back Ground (BG) in many image coding applications such as web browsing, image database, remote sensing image coding and telemedicine. It is pointed that it is assured that the
PSNR value of the ROI is to be higher than a certain value. For example, in the telemedicine, the ROI, which contains the important diagnostic information, needs to be compressed without loss, i.e. the PSNR of the ROI should be infinite.

A common method used is to leave the wavelet coefficients using a lift factor needs to be pre-determined manually which are used in most current DWT-based compression technologies. JPEG2000 still image compression standard [40] supports ROI coding which lifts the coefficients directly also. Here, the max shifting method is done to lift the lowest bit plane of wavelet coefficients of the ROI than the highest bit plane of those of the BG.

The method for arbitrary shaped object coding using integer wavelet transform overcomes the unbalance between the ROI and BG under the condition of guaranteeing the quality of the reconstructed ROI. SA-DWT is used for the arbitrary shape ROI coding and fit for variable coding applications such as sensing image and telemedicine coding that need lossy and lossless compressions. SA-DWT maintains similar spatial correlation to the actual wavelet transform and the number of the coefficients in the transform domain using SA-DWT is equal to the number of pixels contained in the region. During the coding of arbitrary shaped objects, there are many isolated segments which require special consideration [166]. These segments fall into one of the following cases: 1) even length signal at even position 2) even length signal at odd position 3) odd signal at even position 4) odd signal at odd position. The signal is symmetrically extended around its boundaries before starting the wavelet transform. SA-DWT is evaluated according to the length of segment as follows:

1. If the length of the isolated segment $N$ is unity, and if the position of the pixel is even, the coefficient is appended to the low frequency band, else the position of the pixel is odd and the coefficient is appended to the high band. The normalization is done by $\sqrt{2}$.

2. If the length of signal $N$ is even, then the isolated segment is transformed using the lifting scheme to produce $N/2$ high frequency coefficients using even sub sampling $(2i+1)$ and $(N/2)$ low frequency coefficients using $(2i)$ sub sampling, respectively.

3. If the length of signal $N$ is odd, then if the isolated segment starts at an even position, the isolated segment is transformed using the lifting scheme to
produce \((N/2)\) high frequency coefficients and \((N/2)+1\) low frequency coefficients. This done by odd sub sampling \((2i)\) for low frequency coefficients and even sub sampling \((2i+1)\) for high frequency coefficients. If the isolated segment starts at an odd position, the isolated segment is transformed using the lifting scheme to produce \((N/2)+1\) high frequency coefficients and \((N/2)\) low frequency coefficients. This done by odd sub sampling \((2i)\) for high frequency coefficients and even sub sampling \((2i+1)\) for low frequency coefficients.

6.3. Adaptive Scalable WDR Method for Shape Adaptive Object Coding

The shape adaptive wavelet transform is done on the selected visual objects as preprocessing. After that, the adaptive scalable wavelet difference reduction method is applied on the region of interest area. The shape mask is used to extract the region of interest area from the visual scene. Then, apply the shape adaptive wavelet transform either reversible integer for lossless coding or normal wavelet transform for lossy coding.

The significant information is generated using the significant test function \(\sigma(w_i, t_n)\) at the bit-plane \(n\). The sign information is generated using \(\text{Sign}(w_i)\). The encoding process proceeds through the mask function \(\text{mask}(i, j)\) to check whether the wavelet coefficient position is inside the object area or not.

\[
\text{Mask}(i, j) = \begin{cases} 
1 & \text{if } (i, j) \text{ inside the object area} \\
0 & \text{if } (i, j) \text{ outside the object area}
\end{cases}
\]

The whole coding process of the shape adaptive transformed image is applied using the adaptive scalable wavelet difference reduction method. So, we can generate the resolution scalable version of the objects present in whole image. The coding procedure is shown in the figure 6.7.

From figure 6.7, we can see that the image is operated with the mask to extract the object for coding region of interest area. These regions of interest areas are located and identify the index positions and use the shape adaptive wavelet transform to code the object. After that the adaptive wavelet difference reduction method is applied. The layered bit stream is decoded using the scalable wavelet difference reduction decoding method and generate different scalable images. The encoding and decoding process is done with the help of the generated mask.
Object based coding has provided a large degree of flexibility in digital image and video processing and is expected to play a major role in future multimedia, computer games and related applications. The object facilitates the manipulations and scalability transmission in a highly flexible manner for interacting with the objects. The JPEG2000 [36] still image compression standard provides these facilities, i.e. coding of Visual Object coding with ROI property [40].

The scalable transmission is quite useful for image/video communication over heterogeneous networks which require high degree of flexibility from the coding system. Scalable image/video coding has also different applications such as web browsing, image/video database systems, video telephony, etc. Object based approaches for video coding, which call SA-DWT, are being studied as a new video coding paradigm, in which only the samples with in an object are transferred with shape-adaptive wavelet transform according to the shape information additionally sent to the decoder and the coefficients are encoded. Thus, the exact object can be decoded differently from the background.

The shape and texture coding is based on the extension of wavelet difference reduction method called scalable WDR. Here, binary shape mask $S$ is used to describe which nodes are inside the object and which are outside. The visual object coding is performed in the $Y_{C_b}C_r$ color space. We can define a set $G = \{(i,j) | S(i,j) = 1\}$ as the set of all coordinates inside the object, and $G' = \{(i,j) | S(i,j) = 0\}$ as the complementary set containing all coordinates outside the object.

i.e. $G \cup G' = \{(i,j) | i = 0, 1, \ldots, m-1, j = 0, 1, \ldots, N-1\}$ and $|G| + |G'| = MN$.  

193
As example, the sample test images, its masks and generated objects are shown in the figure 6.8 and figure 6.9. The shape adaptive 2D-DWT of the sample test image announcer object is shown in the figure 6.10.

Figure 6.8 Example of Object Extraction on MRI Image (a) MRI image (b) Object mask (c) Extracted object

Figure 6.9 Object extraction on test image Announcer (a) video image (b) object mask (c) extracted object
6.4. Simulation Results of Shape Adaptive Object Coding

The proposed scheme SWDR is compared with original SPIHT and its scalable version S-SPIHT with Shape adaptive DWT. The algorithm is simulated on 8 bpp medical images consisting of 10 classes with 100 frames in each class and video objects like announcer etc. The original resolutions of the medical images are (512x512) pixels and video still pictures are 480x512 pixels and are publicly available [187] at http://www.cipr.rpi.edu/resource/sequences/index.html. The wavelet decomposition is based on the bi-orthogonal 9/7 for announcer object and 5/3 tap wavelet filters for MRI object with symmetric extension at the image boundary.

The YUV color space [182-186] is basically a recoding of RGB color space. Each color appears as a luminance (Y) and two chrominance components (U and V). Luminance, the intensity perceived, is decoupled from the chrominance components so the intensity can be varied without affecting the colour. This colour space was created to use in colour TV transmission. Therefore, a lower band is necessary for transmission to compare with RGB (in case of digital transmission, less bits).
Different components need not have the same bit depths, nor need they have all been signed or unsigned. For reversible systems, the only requirement is that the bit depth of each output image component must be identical to the bit depth of the corresponding input image component. This standard supports two different component transformations, the Irreversible Component Transformation (ICT) that can be used for lossy or lossless coding and one Reversible Component Transformation (RCT) that may be used only for lossy coding in addition to encoding without colour transformation.

Since the ICT may only be used for lossy coding, it may only be used with the 9/7 irreversible wavelet transformation. The forward and the inverse ICT transformation are achieved by means of equations 6.1 and 6.2.

\[
\begin{pmatrix}
Y \\
Cb \\
Cr
\end{pmatrix} = \begin{pmatrix}
0.299 & 0.587 & 0.114 \\
-0.16875 & -0.33126 & 0.5 \\
0.5 & -0.41869 & -0.08131
\end{pmatrix}\begin{pmatrix}
R \\
G \\
B
\end{pmatrix}
\] (6.1)

\[
\begin{pmatrix}
R \\
G \\
B
\end{pmatrix} = \begin{pmatrix}
1 & 0 & 1.402 \\
1 & -0.34413 & -0.71414 \\
1 & 1.772 & 0
\end{pmatrix}\begin{pmatrix}
Y \\
Cb \\
Cr
\end{pmatrix}
\] (6.2)

Since, the RCT may be used for lossy and/or lossless coding, it may only be used with the 5/3 reversible wavelet transform; the RCT is a de-correlating transformation, which is applied to the three first components of an image. Three goals are achieved by this transformation namely, colour de-correlation for efficient compression, reasonable colour space with respect to the Human Visual System (HVS) [182-186] for quantization and ability of having loss-less compression, i.e. exact reconstruction with finite integer precision.

For the RGB components the RCT can be seen as an approximation of a YUV have the same sampling parameters and the same bit depth. There shall be at least three components if this transformation is used.

The forward and irreverse RCT is performed by means of equation.6.3.and equation.6.4.

\[
\begin{pmatrix}
Y_r \\
U_r \\
V_r
\end{pmatrix} = \begin{pmatrix}
\left[\frac{R + 2G + B}{4}\right] \\
\frac{R - G}{2} \\
\frac{B - G}{2}
\end{pmatrix}
\] (6.3)
Object Coding Using a Shape Adaptive Wavelet Transform with Scalable WDR Method

\[
\begin{pmatrix}
Gr \\
R \\
B
\end{pmatrix} = \begin{pmatrix}
\frac{Yr - (Ur + Vr)}{4} \\
Ur + G \\
Vr + G
\end{pmatrix}
\]  \hspace{1cm} (6.4)

An effective way to reduce the amount of data in JPEG is to use an RGB to YCrCb de-correlation transform followed by sub-sampling of the chrominance (Cr, Cb) components. This is not recommended for use in JPEG2000, since the multi-resolution nature of the wavelet transform may be used to achieve the same effect [23, 186].

The objects are extracted manually, but in practice any segmentation algorithm appropriate for the application may be employed. Examples of objects are shown in the figure 6.11. Six levels of wavelet decomposition were first applied to each test image, and then the scalable WDR encoder was set to encode the coefficients from bitplane_{max} to bitplane_0 supporting maximum spatial scalability levels as 7.

<table>
<thead>
<tr>
<th>Method</th>
<th>Full Resolution</th>
<th>Half Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bit rate (bpp)</td>
<td>0.125 0.25 0.5 1.0</td>
</tr>
<tr>
<td>SPIHT</td>
<td>-</td>
<td>27.56 31.20 36.49 43.63</td>
</tr>
<tr>
<td>S-SPIHT</td>
<td>31.21 35.84 40.35 47.11</td>
<td>27.76 31.49 37.03 47.72</td>
</tr>
<tr>
<td>SWDR</td>
<td>31.92 36.43 40.93 47.45</td>
<td>28.58 32.00 38.04 48.83</td>
</tr>
</tbody>
</table>

Table 6.1 PSNR values of MRI object of full resolution with size (512x512) and half resolution with size (256x256) based on SPIHT and SWDR

The simulation results obtained by the algorithms SPIHT, SWDR and on medical images are given in table 6.1. A typical reconstructed image at different resolution levels is shown in figure 6.12. For full resolution MRI image reconstruction, the performance gain is from 0.34 dB to 0.71 dB for various bit rates. It is observed that the coding performance in PSNR values (in dB) increases when the resolution scale decreases. For resolution level 2, i.e. 256 x 256, the performance gains of scalable wavelet difference reduction method are from 0.80 dB to 5.20 dB compared to the normal SPIHT and from 0.51 dB to 1.11 dB compared to the scalable SPIHT for various bit. Similar experimental results are obtained for the resolution level 3 (128 x 128).
Figure 6.11 Sample test image for object extraction (a) Barbara (d) Boat (g) MRI head image (b, e, h) Shape mask (c, f, i) Extracted object

<table>
<thead>
<tr>
<th>Method</th>
<th>Full Resolution</th>
<th>Half Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bit rate (bpp)</td>
<td>0.125 0.25 0.5 1.0</td>
</tr>
<tr>
<td>SPIHT</td>
<td>-</td>
<td>33.61 38.14 43.86</td>
</tr>
<tr>
<td>S-SPIHT</td>
<td>36.49 40.77 45.46 51.51</td>
<td>34.24 39.70 46.47</td>
</tr>
<tr>
<td>SWDR</td>
<td>37.83 41.66 46.25 52.60</td>
<td>36.64 41.51 48.57</td>
</tr>
</tbody>
</table>

Table 6.2 PSNR values of Announcer object of full resolution with size (480x512) and half resolution with size (240x256) (luminance component Y) based on SPIHT and SWDR

The experimental results obtained for video still object in YUV format and PSNR values for luminance components (Y) are shown in the table 6.2. and the reconstructed images are shown in the figure 6.13. For full resolution object reconstruction, the performance gain is from 0.79 dB to 1.34 dB for various bit rates. For resolution level 2, i.e. 240x256, the performance gains of SWDR are from 3 dB to 12 dB compared to the normal SPIHT and from 1.81 dB to 3.76 dB compared to the scalable SPIHT for various
bit. Similar experimental results are obtained for the resolution level 3 (120x128). The experiments are done on the objects of test images like Barbara, Boat and 256 MRI Head images and similar results are obtained and shown in the tables from table.6.3 to table.6.7.

Figure 6.12 Scalable object reconstruction of test image MRI object (512x512) size image at bit rate 0.0625 (a) full resolution (b) ½ resolution (c) ¼ resolution
For full resolution Barbara object reconstruction, the performance gain is from 0.1 dB to 0.5 dB for various bit rates. For resolution level 2, i.e. 240x256, the performance gains of SWDR are from 3 dB to 12 dB compared to the normal SPIHT and from 0.40 dB to 11.95 dB compared to the scalable SPIHT for various bits. The coding performance analysis is shown in the table 6.3. The coding gain is plotted in the graph figure 6.14.
Table 6.3 Reconstructed values in terms of PSNR values for object of Barbara image

<table>
<thead>
<tr>
<th>Method</th>
<th>(1/2) Resolution (512x512)</th>
<th>(1/4) Resolution (512x512)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIHT</td>
<td>27.5742 29.4064 30.6620 34.2267 39.2557 - - -</td>
<td>30.9548 34.4924 34.9179 39.5462 - - -</td>
</tr>
<tr>
<td>SSPIHT</td>
<td>28.0573 29.8289 33.1730 39.6024 50.9656 - - -</td>
<td>32.9502 38.4119 48.5730 69.0781 - - -</td>
</tr>
<tr>
<td>SWDR</td>
<td>27.9753 30.1375 33.7021 39.9339 51.2107 - - -</td>
<td>32.8863 38.5135 48.5294 69.3211 - - -</td>
</tr>
</tbody>
</table>

Figure 6.14 Performance analysis in terms of signal to noise ratio value in dB for Barbara object image for reconstructing the full resolution image and half resolution image.

For full resolution Boat object reconstruction, the performance gain is from 0.15 dB to 0.35 dB for various bit rates. For resolution level 2, i.e. 240x256, the performance gains of SWDR are from 0.5 dB to 13 dB compared to the normal SPIHT and from 0.2 dB to 0.4 dB compared to the scalable SPIHT for various bit The coding performance analysis is shown in the table 6.4.
Object Coding Using a Shape Adaptive Wavelet Transform with Scalable WDR Method

<table>
<thead>
<tr>
<th>Method</th>
<th>Full Resolution with size (512x512)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (bpp)</td>
<td>0.03125   0.0625   0.125   0.25   0.5   0.75   1</td>
</tr>
<tr>
<td>SPIHT</td>
<td>25.8319   27.8480   30.5343  34.6189  40.4545  44.9739  48.6262</td>
</tr>
<tr>
<td>SWDR</td>
<td>25.8073   28.0079   30.6843  34.8238  40.5485  45.0762  48.9991</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method (1/2) Resolution with size (256x256)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIHT</td>
</tr>
<tr>
<td>SSPIHT</td>
</tr>
<tr>
<td>SWDR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method (1/4) Resolution with size (128x128)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIHT</td>
</tr>
<tr>
<td>SSPIHT</td>
</tr>
<tr>
<td>SWDR</td>
</tr>
</tbody>
</table>

Table 6.4 Reconstruction values in terms of PSNR values of object of Boat image

For full resolution reconstruction of MRI Head image, the performance gain is from 0.2 dB to 0.25 dB for various bit rates. For resolution level 2, i.e. 128x128, the performance gains of SWDR are from 0.6 dB to 5 dB compared to the normal SPIHT and from 0.3 dB to 0.6 dB compared to the scalable SPIHT for various bit rates. The coding performance analysis is shown in the table 6.5.

<table>
<thead>
<tr>
<th>Method</th>
<th>Full Resolution (256x256)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (bpp)</td>
<td>0.0625   0.125   0.25   0.5   1</td>
</tr>
<tr>
<td>SPIHT</td>
<td>23.5832   26.0868   28.9102  32.6293  37.4664</td>
</tr>
<tr>
<td>SWDR</td>
<td>23.4541   26.2528   29.1688  32.9025  37.7921</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Full Resolution (128x128)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIHT</td>
<td>24.6144   27.7933   31.7363  36.8528  -</td>
</tr>
<tr>
<td>SSPIHT</td>
<td>24.6553   27.9408   32.8550  40.6459  -</td>
</tr>
<tr>
<td>SWDR</td>
<td>24.6615   28.2244   33.2618  41.0802  -</td>
</tr>
</tbody>
</table>

Table 6.5 Reconstruction values in terms of PSNR values of MRI (head) image

The coding performance of the proposed algorithm using objects of different test images are also checked using another distortion metric called structural similarity index (SSIM) metric. The performance is listed in the table 6.6. It is concluded that the coding is much better than the existing methods in all sense.
Object Coding Using a Shape Adaptive Wavelet Transform with Scalable WDR Method

Here the scalable WDR coding scheme with shape adaptive wavelet transform that supports spatial and SNR scalability is presented. The flexible bit streams generated by the encoder can be decoded adaptively to get images at any level of spatial resolution. The object based coding using scalable WDR performs much better than the scalable SPIHT [147] and the original SPIHT [76] at any bit rate in terms of scalable properties and has lesser complexity than the zero tree coding technique.

SA-DWT coding [166] is a very efficient technique for coding arbitrarily shaped visual objects. A comprehensive description of shape-adaptive wavelet coding schemes for coding arbitrarily shaped visual objects is done by Li [166] and also implement in our work. The number of wavelet coefficients after the SA-DWT is identical to the number of pixels in the arbitrarily shaped visual object. The spatial correlation and wavelet transform properties, such as locality property and self-similarity across subbands, are well preserved in the SA-DWT. The new coding scheme is applied to the multimedia video sequences. The scalability features of new method have interesting perspectives for numerous visual communications applications.

Table 6.6 Structural index values of objects of standard test images

<table>
<thead>
<tr>
<th>Test Image</th>
<th>Bit rate</th>
<th>Full Resolution (512x512)</th>
<th>(1/2 Resolution) 256x256</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SPIHT</td>
<td>SWDR</td>
</tr>
<tr>
<td></td>
<td>0.125</td>
<td>0.83037</td>
<td>0.84586</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>0.89146</td>
<td>0.90146</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.95027</td>
<td>0.95382</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>0.97289</td>
<td>0.97563</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.98475</td>
<td>0.98538</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.91672</td>
<td>0.91946</td>
</tr>
<tr>
<td>MRI</td>
<td>0.125</td>
<td>0.95834</td>
<td>0.96079</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>0.98360</td>
<td>0.98604</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>0.99225</td>
<td>0.99297</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.99605</td>
<td>0.99648</td>
</tr>
<tr>
<td></td>
<td>0.125</td>
<td>0.96342</td>
<td>0.96454</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>0.98104</td>
<td>0.98137</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>0.99173</td>
<td>0.99201</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.99905</td>
<td>0.99911</td>
</tr>
</tbody>
</table>

203