*International Conference on Emerging Technology Trends on Advanced Engineering Research (ICETT'12)*
*Proceedings published by International Journal of Computer Applications® (IJCA)*

# An Intraframe Coding by Modified DCT Transform using H.264 Standard

C.Jilbha Starling
Dept of ECE
C.S.I Institute of technology
Thovalai,India

D.Minola Davids C.
Dept of ECE.
C.S.I Institute of technology
Thovalai,India.

Seldev Christopher
Dept of ECE
St.Xaviers Engg.College
Chunkankadai,India

## ABSTRACT

Compression is necessary for video coding, which introduces the fundamental of video image formation. The efficient, fast encoding, lossy and lossless technique are required for video coding for different application. Focus is concentrated on the lossless intra coding method in video image besides various video coding standards.The upcoming H.264 video compression standard promises a significant improvement over all previous video compression standards. In terms of coding efficiency, the new standard is expected to provide at least 2x compression improvement over the best previous standards and substantial perceptual quality improvements over both MPEG-2 and MPEG-4. A directional filtering transform (DFT) is proposed in this paper to better exploit correlations among samples in H.264 intraframe coding.

## Keywords

H.264, intra coding, arithmetic coding, lifting order, interpolation

## 1. INTRODUCTION

Most of the video compression techniques exploit the Spatial, Spectral and Temporal redundancies present in the video image. Though there exists a lot of schemes for the lossless compression techniques, achieving the real lossless compression depends largely on the Intra frame coding.Digital Video Communication is used today in many applications, such as broadcast, pay-per-view services over satellite, cable and terrestrial transmission channels, wire-line and wireless real-time conversational services, internet or Local Area Network (LAN) video streaming services, storage formats etc. The fundamental problems in video communication are encoding the source data within an available bit rate or decoding the lowest bit rate code to better source data reproduction.

Many methods have been proposed to exploit the directionality in intra-frames in terms of transforms and predictions. Some transforms are proposed from filter banks to represent multidirectional information of images [1].On applying these transforms, the coefficients uses quantization and entropy coding methods. Some methods are developed to introduce directionality into the traditional transforms. These directional transforms demonstrated consistent and stable performance gain over non - directional counterparts. Taubman [8] proposed to adapt filtering direction to local image content and performed conventional sub band transform in resampled domain. Some methods are proposed to use lifting structure for local directional filtering. Where the subbands reside on rectilinear grid. They do not exhibit filtering direction. Their performances are constrained. No significant improvements are found.
By estimating the directional information at the encoder, they are transmitted to the decoder. By using both the curved wavelet transform and the adaptive directional lifting (ADL), they result in good subjective and objective performance.

This paper proposes the Directional Filtering Transform (DFT) consists of directional filtering and an optional DCT [7]. In the directional filtering, there are two different modes. One is the UnidirectionalFiltering (UDF). Another is the Bidirectional Filtering (BDF). An important feature of DFT is every block can adaptively select its lifting order, which is difficult to achieve in directional wavelet transforms. Since DFT not only enables closer neighbouring samples for prediction but also well exploits the directional correlation, it can achieve better performance than the current intraframe coding in H.264.

The rest of the paper is organised as follows. Section 2 presents the existing systems. Section 3 describes about the description to the proposed work. Experimental results are given in section 4. Finally, Section 5 concludes this paper.

## 2. EXISTING SYSTEMS

The new filter bank decomposes images into directional components which can be maximally decimated while still allowing the original image to be exactly reconstructed from its decimated channels.The fundamental building block of the directional filter bank can be analyzed conveniently. The analysis section of such structures decomposes a signal into its subband components by first filtering with a bank of bandpass filters, and then maximally downsampling the filtered outputs with a corresponding bank of downsamplers Computation can be reduced by approximately 50% by using polyphase filter bank structures. There are several polyphase implementations possible. The proposed directional filter bank is implemented in a tree-structure composed of two-band systems [1].

An efficient directional, multiresolution image Representation.Construct a discrete-domain multiresolution and multidirection expansion using nonseparable filter banks [2].Construction results in a flexible multiresolution and directional image expansion using contour segments and thus it is named as contourlettransform.

An orthonormal version of the ridge let transform for discrete and finite- size images are proposed [3].The construction uses the Finite Radom Transform (FRAT) as a building block. The FRAT coefficients are introduced. Numerical results show that the FRAT is more effective than the wavelet transform in approximating and denoising images with straight edges

The curved wavelet transform is carried out by applying 1-D filters along curves [5].The curves are determined based on image content and are usually parallel to edges and lines in the image to be coded. The pixels along these curves can be well represented by a small number of wavelet coefficients.The curved wavelet transform is used to construct a new image coder.

The multiscale wedgelet framework is a first step towards explicitly capturing geometrical structure in images [4].The framework has two components: decomposition and representation.

The multiscale wavelet decomposition divides the image into blocks at different scales and projects these image blocks onto wedgelets.The multiscale wedgelet representation is an approximation of the image built out of wedgelets from the decomposition.

Transform schemes for image and video coding choose the 2-D discrete cosine transform (DCT) [6]. DCT is implemented separately through two 1-D transforms, one along the vertical direction and another along the horizontal direction. DCT framework is able to provide a better coding performance for image blocks that contain directional edges.

Directional filtering transform (DFT) is used to better exploit intra-frame correlation in H.264 intra-frame coding [7].DFT was proposed by predicting each sample from its neighborhood along the image orientation.

By combining bi-directional and unidirectional filtering together, samples exhibiting various local characteristics can be well predicted.The proposed DFT achieves better performance than the current intra-frame coding in H.264.

# 3. DESCRIPTION TO THE PROPOSED SYSTEMS

## 3.1 Directional Fittings

Directional filtering transform (DFT) [7] is used to better exploit intra-frame correlation in H.264 intra-frame coding. It consists of a directional filtering and an optional DCT transform. In the proposed directional filtering, there are two different approaches. One is the Unidirectional Filtering (UDF) .In this approach, only samples from neighboring blocks can be used in prediction.

Another is Bidirectional Filtering (BDF) that exploits the correlations among samples from not only neighboring blocks but also the current block.

The prediction structure is hierarchical multi-layer. UDF and BDF show the advantage to combine them together .The proposed DFT is integrated into H.264 intra-frame coding.

DFT was proposed by predicting each sample from its neighborhood along the image orientation. By combining bi-directional and unidirectional filtering together, samples exhibiting various local characteristics can be well predicted.

To improve the intra-frame coding performance in H.264, many efforts have been done to better utilize directional correlations among samples in terms of intra rediction and transform. The proposed DFT achieves better performance than the current intra-frame coding in H.264.

## 3.2 Intra Coding Scheme

The intra coding scheme is shown in Fig 3.1. There are two methods to exploit intraframe correlations.

One is the input image is predicted by UDF from neighboring block boundaries.

The residues are decorrelated by a FFT transform and then quantized with a uniform quantizer. Another method is to predict the input image by BDF, which possesses evenly distributed filtering modes by which the lifting order can be selected adaptively.

After the filtering, two methods are presented. If the filtered residues show an isotropic correlation, a FFT transform is performed to decorrelate them and the resulting coefficients are then quantized and coded. Otherwise the transform is skipped and residues are directly quantized and entropy coded.

Let$x(t)$,$t = \{t1, t2\}$denote the original image signal. In the first dimension lifting, odd-row pixels are predicted from even-rows along an angle. Then even-row samples are updated from the odd-row residues along the same direction.

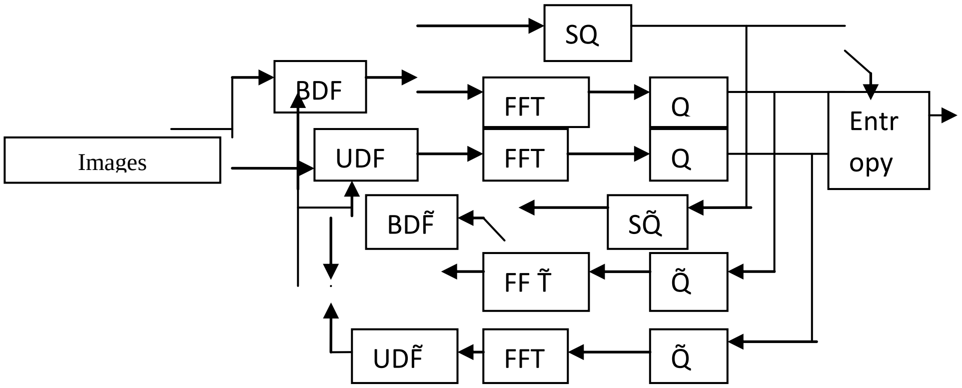

Fig 3.1: Intraframe Coding

Though any bi-orthogonal wavelet transform can be factored into one or several lifting stages, one-step lifting is considered, which means one prediction step followed by one update step and scaling. Then the lifting process can be expressed as

$$y_1[t]=g_H.\left(x_1[t]-P_{x0}^{d}[t]\right) \qquad (3.1)$$

$$y_0[t]=g_L.\left(x_0[t]+g_H^{-1}.u_{y1}^{d}[t]\right) \qquad (3.2)$$

Where $x_0[t]$are the even-row samples and $x_1[t]$the odd-row samples$P_{x0}^{d}[t]$ is the prediction for$x_1[t]$ and $U_{y1}^{d}[t]$is the update for$x_0[t]$. $y_0[t]$ and$y_1[t]$ denote the L and H bands, respectively.$g_H$ and $g_L$are scaling factors.

## 3.3 Coding Gain of ADL

When analyzing the coding gain,it consists of split, prediction,update and scaling property. (3.1), (3.2). The process is rewritten as

$$y_1[t]=g_H.\left(x_1[t]-\mathrm{h}_{\mathrm{p}}^{\mathrm{d}}[t]*x_0[t]\right) \qquad (3.3)$$

$$y_{0}[t]=g_L.(x\_0\ [t]+g\_H\hat{}(-1).(\mathrm{h}\_u\hat{}d\ [t]* y\_1\ [t])) \qquad (3.4)$$

Where $\mathrm{h}_{\mathrm{p}}^{\mathrm{d}}[t]$ and $\mathrm{h}_{u}^{d}[t]$ are predictions and update filters.ThenFourier transform is applied and transfer functionis given as,

$$H_p^d(e^{jw})=e^{jw2/2}H'^{d}_{p}(e^{jw1},e^{jw2/2}) \qquad (3.5)$$

Similarly,

$$H_u^d(e^{jw})=\frac{1}{2}e^{-jw2/2}H'^{d}_{p}(e^{jw1},e^{jw2/2}) \qquad (3.6)$$

The scaling factors are computed as,

$$g_L^2=A_{ver}([1+H'^{d}_{p}(e^{jw})]^2) \qquad (3.7)$$

$$g_H^2=A_{ver}[1\text{-}\frac{1}{2}H'^{d}_{p}(e^{jw})-\frac{1}{2}(H'^{d}_{p}(e^{jw}))^2]^2 \qquad (3.8)$$

The variance of $y_1[t]$)and $y_0[t]$are

$$\sigma_{y1}^2=g_H^2.A_{ver}([1\text{-}H'^{d}_{p}(e^{jw})]^2X_{xx}(e^{jw}))$$

$$\sigma_{y1}^d=g_L^2\ .A_{ver}[1+\frac{1}{2}H'^{d}_{p}(e^{jw}) -\frac{1}{2}(H'^{d}_{p}(e^{jw}))^2]^2\ X_{xx}(e^{jw})) \qquad (3.9)$$

The coding gain of ADL is given as

$$G_{ADL}=\frac{\sigma^2}{\sqrt{\sigma_{y0}^2\sigma_{y1}^2}} \qquad (3.10)$$

## 3.4 Optional Transform

After filtering, process the residues. Each residue block is first divided into 4× 4 blocks and then optionally transformed with its corresponding transform matrix. Let y denote the 1 ×16 vector of one 4× 4 residue block, the transform is given a

$$Z=T'y \qquad (3.11)$$

T is the 4 ×4 transform matrix with each column representing the basis vector. After the transform, the coefficients are scalar quantized and then entropy coded. Separate context models are used for different types of blocks. If there is no need to perform the transform, the quantization method that has been integrated into the KTA software is adopted to quantize the residues.

## 3.5 Modified Discrete Cosine Transform

Modified DCT is used to convert data in time domain to data in frequency domain. If the time domain data is smooth (with little variation in data) then frequency domain data will make low frequency data larger and high frequency data smaller. In DCT, coefficient in location (0, 0) is called DC coefficient and the other values are called

AC coefficients. In general, at the time of quantization, large quantization step is used in quantizing AC coefficients, and small quantization step is used to quantize DC coefficient so as to have high precision preservation. In general most of the energy (value) is concentrated to the top-left corner of the DCT transformed block. After quantizing, the most of the data in this quantized block maybe zero, then using zig-zag order scan and run length coding can achieve a high compression ratio.

In this transformation, the intra frame is partitioned into 8 x 8 block and Discrete Cosine Transform is applied to each block using the forward DCT transform. It produces another 8 x 8 block containing one DC components and sixty three AC components. In general most of the energy (value) is concentrated to the top-left corner. After quantizing the transformed matrix, most data in this matrix may be zero, then using zigzag order scan and run length coding can achieve a high compression ratio.

Discrete cosine transform is a lossy compression algorithm that samples an image at regular intervals, analyzes the frequency components present in the sample, and discards those frequencies which do not affect the image as the human eye perceives it. DCT is the basis of standards such as JPEG, MPEG, H.261, H.263 and H.264.

In this project, the MDCT works by separating images into different frequencies. During a step called quantization, where part of compression occurs, the less important frequencies are discarded. Only the most important frequencies that remain are used to retrieve the image in the decompression process. The reconstruction of image is done by IDCT.

The replacement of standard discrete cosine transform of intraframe coding with the modified discrete cosine transform is investigated to determine whether improvements in numerical quality can be achieved. The amount of data to be entropy coded (output of the DCT) is the same as the original amount of the data (input to the DCT). This desirable property is referred to as critical sampling. By using MDCT, the amount of data to be compressed is doubled and critical sampling is lost. The standard DCT maps N samples of data to N new values, the MDCT maps N sample block to a block consisting of $\frac{N}{2}$ new values.

## 4. EXPERIMENTAL RESULTS

Each 8×8 block is hit with MDCT. .Each element of the block is then quantized with quality level 50.At this point many of the elements becomes zero and the image takes much less space to store. Then image is decompressed using IDCT. The image is compressed at a rate of 3.2 bits/pixel. It can be achieved by the compression ratio.

A very logical way of measuring how well a compression algorithm compresses a given set of data is to look at the ratio of the number of bits required to represent the data before compression into the number of bits required to represent the data after compression. This ratio is called Compression Ratio (CR).

Benchmarks in image data compression are the compression ratio and PSNR (Peak Signal to Noise Ratio). The compression ratio is used to measure the ability of data compression by comparing the size of the image being compressed to the size of the original image. The greater the compression ratio means the better the wavelet function. PSNR is one of the parameters that can be used to quantify image quality. PSNR parameter is often used as a benchmark level of similarity between reconstructed image and the original image. Larger PSNR will produce better image quality.

The peak signal-to-noise ratio (PSNR) is an objective dissimilarity measure which gives the extent of distortion. It is defined as the ratio of the maximum possible signal power of the original image to the power of decompressed image.

PSNR is expressed in terms of the logarithmic decibel scale because many signals have a very wide dynamic range. The PSNR is given as,

$$\text{PSNR}= 10\log_{10}(\frac{peak^2}{MSE})\text{dB} \qquad (4.1)$$

Where peak is the largest value a pixel can assume. For image of 8-bit precision, this value is 255.

$$\text{PSNR} =10\log_{10}(\frac{255^2}{MSE})\text{dB} \qquad (4.2)$$

Where 255 is the maximum pixel (signal) value of the Intraframe (when $B$ bits are used to represent a pixel, then this maximum value will be$2^B - 1$) with a pixel represented by 8 bits.For color images with three RGB levels, the PSNR is the same except that the MSE is the sum over all the squared value differences divided by image size and by three. PSNR can be calculated easily and quickly and is therefore a very popular quality measure, widely used to compare the 'quality' of compressed and decompressed video images. For a given image or image sequence, high PSNR usually indicates high quality and low PSNR usually indicates low quality. However a particular value of PSNR does not necessarily equate to an 'absolute' subjective quality.

## 5. CONCLUSION

A two dimensional modified DCT is commonly used for the purpose of transform coding. By using a quality level of 50 in quantization of transform coding a high image quality and high compression can be obtained. The image takes less space to store. The new H.264/MPEG-4 AVC offers significant bit rate and quality advantages over all previous standards. In terms of coding efficiency, the new standard is expected to provide at least 2x compression improvement over the best previous standards.H.264/AVC represents a number of advances in standard video coding technology, in terms of both coding efficiency enhancement and flexibility for effective use over a broad variety of network types and application domains. Thismethod reduces computational complexity and achieves higher PSNR with good compression ratio. It achieves better image quality due to high PSNR value. So we conclude that this method is the best compared to other existing methods.

**Fig 5.1: original image**

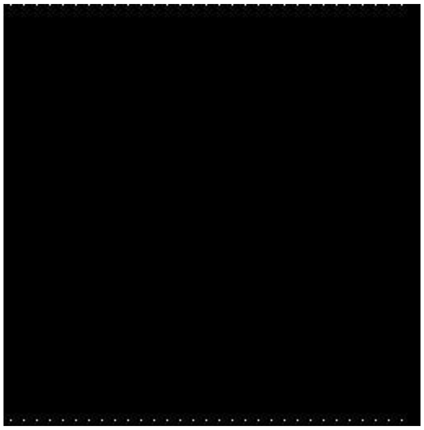

**Fig 5.2: Compressed images**

**Fig 5.3: Decompressed image with quality level 50.**

# 6. REFERENCES

[1] Bamberger R.H. and Smith M.J.T.(1992) "A filter bank for the directional decomposition of images: Theory and design," IEEE Trans. Inf. Theory vol. 40, no. 4, pp. 882–893.

[2] Do M. N. and Vetterli M. (2005) "The contourlet transform: An efficient directional multiresolution image representation," IEEE Trans. Image Process.,vol. 14, no. 12, pp. 2091–2106.

[3] Do M. N. and Vetterli M. (2003) "The finite ridgelet transform for image representation," IEEE Trans. Image Process., vol. 12, no. 1, pp. 16–28.

[4] Romberg J. K., Wakin M., and Baraniuk R.( 2002)"Multiscale wedgelet image analysis: Fast decompositions and modeling," in Proc. IEEE Int. Conf. Image Process., Rochester, NY, pp. 585–588.

[5] Wang D.M., Zhang L., Vincent A., and Seperanza F. (2006) "Curved wavelet transform for image coding," IEEE Trans. Image Process.,vol.15,no.8,pp. 2413-2421.

[6] Zeng B. and Fu J. J.(2008 ) "Directional discrete cosine trans A new framework for image coding," IEEE Trans. Circuits Syst. VideoTechnol., vol. 18, no. 3, pp. 305–313.

[7] Peng.,Wu,Xu.(2009)"Directional filtering transform," in Proc.IEEE Int. Conf. Multimedia Expo, pp. 1–4.

[8] D. Taubman and A. Zakhor, "Orientation adaptive subband coding of images," IEEE Trans. Image Process., vol. 3, no. 4, pp. 421–437, Apr. 1994.

[9] Y. Ye and M. Karczewicz, "Improved H.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning," in Proc. IEEE Int. Conf. Image Process., 2008, pp. 2116–2119.

[10] W. P. Ding, F. Wu, X. L. Wu, S. P. Li, and H. Q. Li, "Adaptive directional lifting-based wavelet transform for image coding," IEEE Trans.Image Process., vol. 16, no. 2, pp. 416–427, Feb. 2007.

[11] P.G. Howard and J.S. Vitter. "Fast and efficient lossless image compression". In Proc. IEEE Data Compression Conference, pages 351–360, Snowbird, Utah,USA, March 1993.

[12] Yung-Lyul Lee, Ki-Hun Han, and Gary J. Sullivan, Improved Lossless Intra Coding for H.264/MPEG-4 AVC, IEEE Transactions On Image Processing, Vol. 15, No. 9, pp. 2610-15, September, 2006.