International Journal of Electronics and Communication Engineering.
ISSN 0974-2166 Volume 5, Number 1 (2012), pp.1-8

http://www.irphouse.com

Video Compression using SPIHT and SWT Wavelet

Mayank Nema, Lalita Gupta, N.R. Trivedi

Department of Electronics & Communication Engineering,
Maulana Azad National Institute of Technology, Bhopal, (Madhya Pradesh) India
E-mail: nema_mayank@yahoo.co.in

Abstract

In this paper, we present wavelet based video compression algorithms. The motion estimation and compensation, which is an essential part in the compression, is based on segment movements. The proposed work based on wavelet transform algorithm like Set Partitioning In Hierarchical Trees (SPIHT) and Stationary Wavelet Transform (SWT). Results of Average value of peak signal to noise ratio PSNR mean squared error (MSE) and comparison chart is obtained using MATLAB. The proposed SPIHT and SWT algorithm achieves very good PSNR values, and MSE which makes the techniques more efficient than the 2-D Discrete Cosine Transform (DCT) in the H.264/AVC codec.

Index Terms: video compression, wavelet transform technique (SPIHT, SWT), Matlab

Introduction

An important component of image and video compression systems is a transform. A transform is used to transform image intensities. A transform is also used to transform prediction residuals of image intensities, such as the motion-compensation residual (MC residual), the resolution-enhancement residual in scalable video coding, or the intra-prediction residual in H.264/AVC.

Recently, new transforms have been developed that can take advantage of locally anisotropic features in images [1, 2, 3, 4]. A conventional transform, such as the 2-D DCT or the 2-D Discrete Wavelet Transform (2-D DWT), is carried out as a separable transform by cascading two 1-D transforms in the vertical and horizontal dimensions. This approach does not take advantage of locally anisotropic features present in images because it favors horizontal or vertical features over others. The new transforms adapt to locally anisotropic features in images by performing the filtering along the direction where image intensity variations are smaller. This is achieved, for example, by directional lifting implementations of the DWT [4]. Even though most of the work is based on the DWT, similar ideas have been applied to DCT-based image compression [2].

Wavelet Transform Technique

The discrete wavelet transform (DWT) [5] has gained wide popularity due to its excellent decorrelation property, many modern image and video compression systems embody the DWT as the transform stage [6]. After DWT was introduced, several codec algorithms were proposed to compress the transform coefficients as much as possible. Among them, stationary Wavelet Transform (SWT) and Set Partitioning in Hierarchical Trees (SPIHT) are the most famous ones.

Set partitioning in hierarchical trees(SPIHT)

This compression schemes is based on wavelet coding technique. The image is transformed using a discrete wavelet transform. In the beginning, the image is decomposed into four sub-bands by cascading horizontal and vertical two-channel critically sampled filter-banks. This process of decomposition continues until some final scale is reached. In each scale there are three sub-bands and one lowest frequency sub-band. Then successive-approximation quantization (SAQ) is used to perform embedding coding. This particular configuration is also called QMF pyramid.

The SPIHT algorithm is used to the multi-resolution pyramid after the sub-band/wavelet transformation is performed [7]. The SPIHT video coding system is shown in fig.1

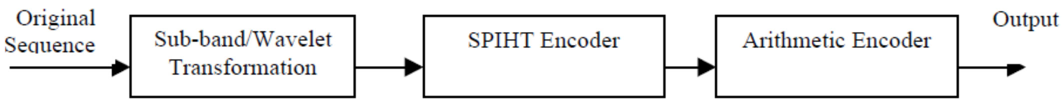

Figure 1: SPIHT Video Coding System

The decoder does exactly opposite, that is, it first performs arithmetic decoding on the input bit Stream, then SPIHT decoding, finally sub-band/wavelet transformation.

Stationary wavelet transforms (SWT)

The stationary wavelet transform (SWT) [8] is a wavelet transform algorithm designed to overcome the lack of translation invariance of the discrete wavelet transform (dwt). Translation- invariance is achieved by removing the down samplers and up samplers in the discrete wavelet transform (dwt) and up sampling the filter coefficients by a factor of $2^{(j-1)}$ in the jth level of the algorithm [9]. the SWT is an inherently redundant scheme as the output of each level of SWT contains the same number of samples as the input , so for a decomposition of N levels there is a redundancy of N in the wavelet coefficients.

Implementation

The following block diagram depicts the digital implementation of SWT as shown in fig.2 and fig.3.

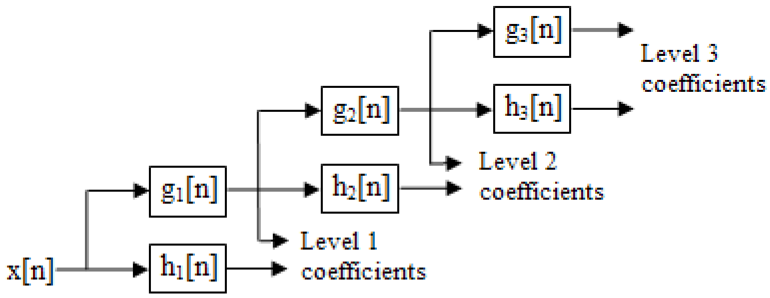

Figure 2: a 3 level SWT filter bank

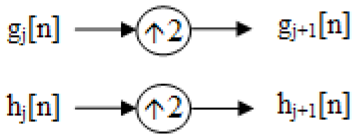

Figure 3: SWT filters

Peak Signal to Noise Ratio (PSNR)

The phrase peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.

The PSNR is most commonly used as a measure of quality of reconstruction of lossy compression codec's (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codec's it is used as an approximation to human perception of reconstruction quality, therefore in some cases one reconstruction may appear to be closer to the original than another, even though it has a lower PSNR (a higher PSNR would normally indicate that the reconstruction is of higher quality). One has to be extremely careful with the range of validity of this metric; it is only conclusively valid when it is used to compare results from the same codec (or codec type) and same content [10]. It is most easily defined via the mean squared error (MSE) which for two m × n monochrome images *I* and *K* where one of the images is considered a noisy approximation of the other is defined as:

MSE = —

The PSNR is defined as:

$PSNR = 10.\log_{10}$ (——)

$PSNR = 20.\log_{10}$ (——)

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with *B* bits per sample, MAX_I is 2^B-1. For color images with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three. Alternately, for color images the image is converted to a different color space and PSNR is reported against each channel of that color space.

Mean Square Error (MSE)

In statistics, the squared error (MSE) of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average of the squares of the "errors." The error is the amount by which the value implied by the estimator differs from the quantity to be estimated. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate [11, 12].

The MSE is the second moment (about the origin) of the error, and thus incorporates both the variance of the estimator and its bias. For an unbiased estimator, the MSE is the variance. Like the variance, MSE has the same units of measurement as the square of the quantity being estimated. In an analogy to standard deviation, taking the square root of MSE yields the root mean square error or root mean square deviation (RMSE or RMSD), which has the same units as the quantity being estimated; for an unbiased estimator, the RMSE is the square root of the variance, known as the standard error.

Definition and basic properties

The MSE of an estimator $\hat{\theta}$ with respect to the estimated parameter θ is defined as

MSE ( ) = E [( )^2]

The MSE is equal to the sum of the variance and the squared bias of the estimator

MSE ( ) = Var ( ) + (Bias ( ))^2

The MSE thus assesses the quality of an estimator in terms of its variation and unbiasedness. Note that the MSE is not equivalent to the expected value of the absolute error. Since MSE is an expectation, it is not a random variable. It may be a function of the unknown parameter θ, but it does not depend on any random quantities. However, when MSE is computed for a particular estimator of θ the true value of which is not known, it will be subject to estimation error. In a Bayesian sense, this means that there are cases in which it may be treated as a random variable.

Compression

1. Application of motion compensation. For each frame in the group, we compute the difference between the frame and its motion predicted frame.
2. A fast multistage wavelet based expansion is applied on each frame and its predicated frame in the group
3. An optimal bit allocation on the differencing between frames in each group and their predictions is applied.
4. The central frame and the other frames differences in the group are quantized and coded using the techniques of still image compression. The distribution of the n^{th} frames in the nth group so that the frames may be compressed and create a new compressed frame.
5. In order to reduce the error resulted from the motion prediction and the wavelet compression, we apply MSE. We estimate the decompression and add this approximation to the main quantized frames.

Approach

Our approach is based on the split of the frame into blocks. Then we compute the 2D motion vectors per block. In this we explain how to perform motion estimation and define new segments from these motion vectors. These segments-regions have similar motion parameters.

Experimental Results

Here, we present the simulation results of video compression by applying the set partitioning in Hierarchical Trees (SPIHT) & Stationary wavelet transform (SWT) algorithm. The original video can be split the in the form of frames and compressed it. The original frame and compressed frames as shown in fig 4 (a) (b) (c).After applying the SWT algorithm on original frame, the PSNR values is 0.005649 and MSE value is 46.54573.Also again applying the SPIHT algorithm on original frame, the PSNR values of 0.006224 and MSE value is 46.12482.

Figure 4(a) Original frame no.1

Figure 4(b) Original frame no.1 compressed By using SWT algorithm

Figure 4(c) Original frame no.1 compressed By using SPIHT algorithm

First we divide each frame in the sequence into blocks resulted from the block matching algorithm. Then we apply SPIHT and SWT wavelet for compression technique on each frame and again decompressed it. The MSE value of the compression is reduced by applying the formula correction in the wavelet domain to enhance the quality of the reconstructed sequence. Here we calculate the average PSNR value and average MSE value.

Calculate Average PSNR/MSE Value

The Average Value of PSNR for SWT and SPIHT, as shown in Table-1.

Table-1: Average value of PSNR for SPIHT and SWT

VALUE	SWT	SPIHT
AVERAGE PSNR	0.0056	0.0061

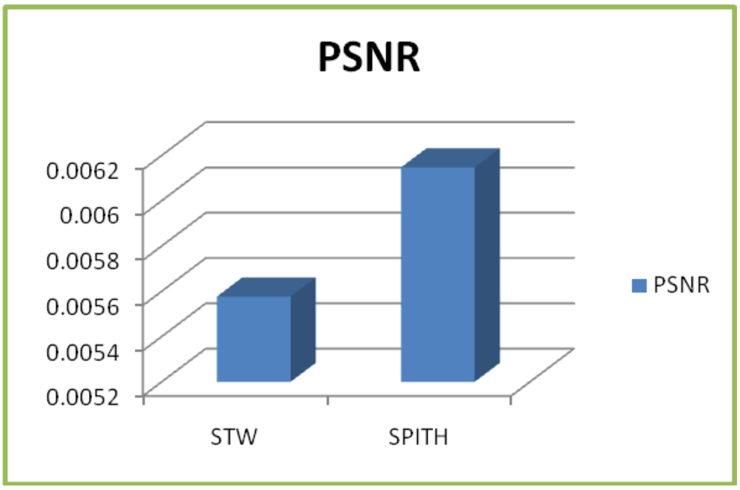

Figure 5 (a): Graph between Average Value of PSNR For SWT and SPIHT

The Average Value of MSE for SWT and SPIHT, as shown in Table-2.

Table-2: Average value of MSE for SWT and SPIHT

VALUE	SWT	SPIHT
AVERAGE MSE	46.6037	46.1802

Figure 5 (b) Graph between Average Value of MSE For SWT and SPIHT.

Conclusions

We examine that the PSNR value and MSE value is better in SPIHT as compared to SWT method. It also supports faithful reproduction of the image, keeping the picture quality of the image/video. In future other evolutionary computing techniques also can be tried for the better results. Future research efforts focus on better PSNR and MSE value. There are many other methods for compression technique, like fuzzy logic, neural network. In compression technique we can try to implement techniques like neural network and fuzzy logic will further better PSNR and MSE.

References

[1] E. Le Pennec and S. Mallat, “Sparse geometric image representations with bandelets,” *Image Processing, IEEE Transactionson*, vol. 14, no. 4, pp. 423–438, April 2005.

[2] Bing Zeng and Jingjing Fu, “Directional discrete cosine transforms for image coding,” *Multimedia and Expo, 2006 IEEE International Conference on*, pp. 721–724, 9-12 July 2006.

[3] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P.L.Dragotti, “Directionlets: anisotropic multidirectional representation with separable filtering,” *Image Processing, IEEE Transactions on*, vol. 15, no. 7, pp. 1916–1933, July 2006.

[4] C.-L. Chang and B. Girod, “Direction-adaptive discrete wavelet transform for image compression,” *Image Processing, IEEE Transactions on*, vol. 16, no. 5, pp. 1289–1302, May 2007.

[5] Olivier Rioul and Martin Vetterli, "Wavelets and Signal Processing”, IEEE Trans. on Signal Processing, Vol. 8, Issue 4, pp. 14 - 38 October 1991.

[6] D. S. Taubman, "High performance scalable image compression with EBCOT", IEEE Transaction Image Processing, Vol. 9, No. 7, pp. 1158–1170, July 2000.

[7] A. Said, W. Pearlman, " A New, Fast, and Efficient Image Codec based on Set Partitioning in Hierarchical Trees", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 6,pp. 243-250, June 1996.

[8] James E. Fowler, “The Redundant Discrete Wavelet Transform and Additive Noise”, IEEE Signal processing letter, vol. 12, issue.9, p.p.629-632, Sep 2005.

[9] Mark J. Shensa, The Discrete Wavelet Transform: Wedding the A Trous and Mallat Algorithms, IEEE Transaction on Signal Processing, Vol 40, No 10, Oct. 1992.

[10] Huynh-Thu, Q.; Ghanbari, M. (2008). "Scope of validity of PSNR in image/video quality assessment". *Electronics Letters* **44** (13): 800–801

[11] Lehmann, E. L.; Casella, George (1998). *Theory of Point Estimation* (2nd ed.). New York: Springer. ISBN 0-387-98502-6.

[12] Mood, A.; Graybill, F.; Boes, D. (1974). *Introduction to the Theory of Statistics* (3rd ed.). McGraw-Hill. p. 229.