ORIGINAL RESEARCH PAPER

Engineering

VIRTUAL VIEW PSNR ESTIMATION FOR 3-D VIDEOS USING ADAPTIVE WAVELET TRANSFORM

KEY WORDS: 3DV,Bit error, Adaptive wavelet transform

Anu C Vedavyasa Institute of Technology, Malappuram, Kerala, India

C Periasamy Vedavyasa Institute of Technology, Malappuram, Kerala, India

ABSTRACT

Today the Peak Signal to Noise Ratio (PSNR) and Distortion in three Dimensional Video (3 DV) coding of implicit view or virtual view in videos is calculated by measuring the quality difference between the virtual view synthesized by compressed 3-DVs with one synthesized by uncompressed 3-DVs. The quality difference measure increases the complexity of a 3-DV system since the picture in the compressed image used to represent that image will be combined by using a threshold level. For compressed image the pixel which is used to represent that image will be aggregated by using a threshold level, From the neighboring texture videos we are generating new depth map. Here it is dealing with the reduction of bit error rate and then estimate the depth for that here adaptive wavelet transform is used. By using adaptive wavelet transform we get a parameter value for depth and then using that value we encoded the depth and then a increase in PSNR value will be obtained. By reducing the bit error naturally the PSNR value will increase, Suppose a 3D video is considered if it is compressed and reconstructed it should be similar to that of the original video this is what required. How to encode the depth properly this is observed here, while compressing or transmitting time due to depth if the correct reconstruction may not occur then PSNR will reduce ,so that how to encode the depth and for that require instantaneous frequency, which is a measure of depth. The increase in PSNR value is attained with adaptive wavelet transform that provides instantaneous frequency which is a measure of depth.

INTRODUCTION

A video is a sequence of frames (or images). Today is the world of telecommunication and multimedia systems, video is used for communication approach to represent some information but bandwidth is a bottleneck. Hence, video compression techniques are used which is significance for reducing the amount of information needed for picture sequence without losing much of its quality, judged by human viewers. Video contains much spatial and temporal redundancy. The aim is to efficiently utilize such types of redundancies to achieve video compression. The two-dimensional (2D) television has been quite established with digital television. The research on three-dimensional (3D) user electronic product has received high interest over the past decade in order to provide viewers with more practical vision than traditional 2D video. In the area of 3D video compression, many researchers have proposed related methods, to perform 3D video and depth coding.

Peak Signal to Noise Ratio(PSNR) and Distortion in three Dimensional Video (3 DV) coding of implicit view or virtual view in videos is calculated by measuring the quality difference between the virtual view synthesized by compressed 3-DVs with one synthesized by uncompressed 3-DVs.The complexity of a 3-DV system is increased here. In order to reduce the complexity of 3-DV system, without rendering virtual views is a better way to estimate virtual view distortions/PSNR directly. Here the virtual view synthesis process and the distortion propagation from existing views to virtual views are analyzed in detail, and then a virtual view distortion/PSNR evaluation method is derived.

Here the empirical mode decomposition (EMD) is used, which proposes to decompose a signal accordingly to its contained information. Even though its adaptableness seems useful for many applications, the main problem with this approach is its lack of theory. Here a new approach to build adaptive wavelets is presented. An adaptive method to examine a signal is of great interest to find sparse representations in the context of compressive sensing. "Rigid" methods, like the Fourier or wavelets transforms, correspond to utilize some basis (or frame) designed independently of the processed signal. The aim of adaptive methods is to construct such a basis directly based on the information contained in the signal

A well known way to build an adaptive representation is the basis pursuit approach which is used in the wavelet packets transform. An entirely different approach to construct an adaptive representation is the algorithm called "Empirical Mode Decomposition" (EMD). This technique has gained a lot of interest in signal analysis mainly because it is able to detach stationary and non-stationary components from a signal. EMD behaves like an adaptive filter bank. Here EMD is modeled such that in a variational framework also it is proposed to model a mode as an amplitude modulated-frequency modulated (AM-FM) signal and then use the properties of such signals to build a functional to represent the whole signal. The wavelet transform has greatly impacted the field of video compression. The amount of compression that can be achieved depends on the energy compaction property of the transform being used. The wavelet transform has a high-energy compaction property, progressive reconstruction that makes it a powerful tool for video compression. The proposed work discusses the 3D video compression.

Here, a new approach to build adaptive wavelets capable of extracting AM-FM components of a signal is proposed. It shows that it is possible to adapt the wavelet formalism by considering distinct Fourier supports and then build a set of functions which form an orthonormal basis. Based on this construction an empirical wavelet transform (and its inverse) to analyze a signal is proposed.

I. PROPOSED METHOD

Here, using adaptive wavelet transform to find the depth value and correspondingly finding out the psnr value. Adaptive transforms are capable of extracting AM-FM components from a signal.AM-FM components have a support of Fourier spectrum. A different approach to build adaptive representation is called Empirical Mode Decomposition .This algorithm is used to build adaptive representation. This method can detect the principle modes which is of a signal. This method is very efficient in separating stationary and non-stationary components from a signal. The problem with EMD is that lack of mathematical theory.It is of an algorithmic approach.EMD behaves similar to adaptive filter. Algorithm of proposed method consist of steps

1) First here it is looked how to reduce bit rate and then increase the PSNR value. Also to estimate the depth in frequency domain and depth is associated with frequency domain estimated
2) After that we get a value for depth then that value should be encoded in order to reduce the bit error rate.
3) If we compress a 3DV and then reconstruct it then it must be similar to that of the original video. For the original videos how to encode the depth parameters and how to get a value for that this is how to estimate the depth.
4) Final step is Apply adaptive wavelet transform then bit error is reduced, and finally PSNR is found out.

Here the given signal multiplied by the basis function and it is integrated and then the transform is applied then we get the depth value properly which in turn gives the corresponding bit error value which is of reduced one. The depth here can be marked as a point. In 3D image when we estimate PSNR compress and decompress there will be bit error which is caused mainly due to the variation in depth here how to estimate this depth value properly is what the paper says. The depth value is estimated, by giving certain number based on quantization ie how to give it a number and based on that number it is encoded and how to increase the PSNR value. Mainly we get a regional frequency but by using that depth value cannot be estimated properly so for that purpose we need instantaneous frequency. The transform which provide instantaneous frequency is adaptive wavelet transform. By using instantaneous frequency we get depth value and then PSNR is found out.

EMPIRICAL WAVELETS

Here propose a method to build a family of wavelets adapted to the processed signal. If the Fourier point of view is considered then, this structure is equivalent to building a set of band pass filters. One way to attain the adaptability is to consider that the filters' supports depend on where the information in the spectrum of the analyzed signal is located. Indeed, the IMF properties are equivalent to say that the spectrum of an IMF is of compact support and centered around a specific frequency (signal dependent).For clarity, only consider real signals (their spectrum is symmetric with respect to the frequency) but the following reasoning can be easily extended to complex signal by building different filters in the positive and negative frequencies, respectively. Also consider a normalized Fourier axis which has a periodicity 2π, in order to respect the Shannon criteria, and restrict the discussion to $\omega \in [0, \pi]$. Define the empirical scaling function and the empirical wavelets by expressions of (1) and (2), respectively.

$$\emptyset^{\wedge}_{n(\omega)} = \begin{cases} 1 & \text{if} |\omega| \le \omega_n - \tau_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] & \text{if } \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$\emptyset^{\wedge}_{n(\omega)} = \begin{cases} 1 & \text{if} |\omega| \le \omega_n - \tau_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] & \text{if } \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

II. EXPERIMENTAL CRITERIA

Parameters used in this paper are peak signal to noise ratio (PSNR) and adaptive wavelet transform to measure the quality of 3DV.

1) Peak Signal To Noise Ratio PSNR is the peak signal to noise ratio measured in decibels (DB) and in terms of bits per sample or b its per pixel. The greater PSNR value gives, the better the image quality and noise suppression.

Peak signal to ratio can be calculated as

$$\text{PSNR} = 10\text{X}\log_{10}\left(\frac{\text{peak}^2}{\text{MSE}}\right)$$

MSE is the mean square error. That, can be calculated as,

$$\text{MSE} = \frac{1}{M * N}\sum_{i=0}^{1-N}\sum_{j=0}^{1-M}(A(i,j) - B(i,j))^{\wedge}2$$

Here A - Perfect image
B - Denoised image
i – Pixel row index
j – Pixel column index

1) Adaptive Wavelet Transform

It is mainly called as an easy path wavelet transform (EPWT) which works as a gateway through the function values.EPWT can be related to multiresolution analysis and also capable of generating scaling spaces and wavelet spaces.Wavelet transforms is one of the most important in signal representation.Now it is used in image processing,data compression and signal processing.The wavelets coefficient value which have been calculated shows the change in image at particular resolution.

Empirical wavelet transform is defined as $W_f^{\varepsilon}(n,t)$

where $W_f^{\varepsilon}(n,t)$=(f,Ψ_n)=$\int f(T)\overline{\Psi_n(T-t)}dT$

$$= (\hat{f}(w)\overline{\widehat{\Psi_n}(w)})^{\nu}$$

Those the coefficients are approximated .

$W_f^{\varepsilon}(0,t)$=(f,ϕ_1)=$\int f(T)\overline{\phi_1(T-t)}dT$

$$=(\hat{f}(w)\overline{\widehat{\Psi_n}(w)})^{\nu}$$

The empirical mode f_k is given as

$$f_0(t)= W_f^{\varepsilon}(0,t) * \phi_1(t)$$

$$f_k(t)= W_f^{\varepsilon}(k,t) * \varphi_k(t)$$

I. RESULTS AND DISCUSSIONS

Experiments are conducted in MATLAB 2009b as simulation software. Test 3D video used here are Fugu (3.19MB), Galaxy (24.2MB), Cartoon (116MB), here Proposed method is compared with existing method in terms of PSNR.

In this experiment, here, choose PSNR as an evaluated standard. The greater PSNR value shows that our proposed method gives better the video quality without artifacts.

Below tables evaluates the performance of different test 3DV

Test Video	Existing method						Proposed method		
Galaxy	PSNR	MSE	Bit error rate	PSNR	MSE	Bit error rate	PSNR	MSE	Bit error rate
Frame 1	22.82	4.69*10^3	0.470	27.82	2.64*10^3	0.0264	34.868	1.17*10^3	0.0117
Frame 2	23.49	4.70*10^3	0.444	26.39	2.55*10^3	0.0248	33.38	1.14*10^3	0.0101

Table 1: PSNR for existing and proposed and existing method of Fugu 3DV

Test Video	Existing method						Proposed method		
Fugu	PSNR	MSE	Bit error rate	PSNR	MSE	Bit error rate	PSNR	MSE	Bit error rate
Frame 1	23.31	4.39*10^3	0.44	28.3	2.47*10^3	0.247	35.44	1.09*10^3	0.011
Frame 2	23.39	4.40*10^3	0.434	23.39	2.45*10^3	0.248	28.38	1.10*10^3	0.012

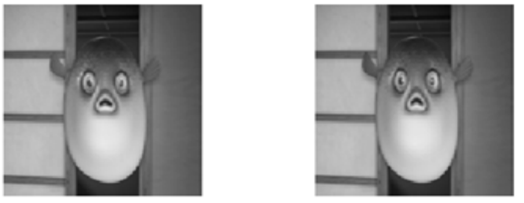
Table 2: PSNR for existing and proposed method of galaxy 3DV.

Test video	Existing method						Proposed method		
Cartoon	PSNR	MSE	Bit error rate	PSNR	MSE	Bit error rate	PSNR	MSE	Bit error rate
Frame 1	32.91	4.66*10^3	0.450	40.21	2.46*10^3	0.0256	44.18	1.12*10^3	0.0121
Frame 2	33.49	4.50*10^3	0.431	39.12	2.51*10^3	0.0252	42.38	1.09*10^3	0.0142

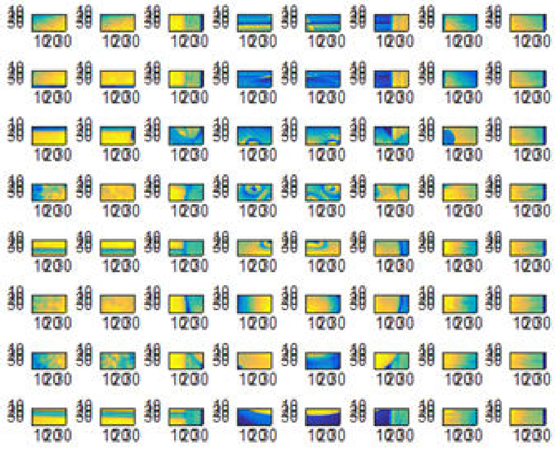
Table 3: PSNR for existing and proposed Cartoon 3DV.

From these tables, we can show that, PSNR values of proposed method successfully performed better result than compared to existing method. Greater the PSNR value will give better video quality.

The comparison result of existing and proposed method is given in the tables and the results are shown below.

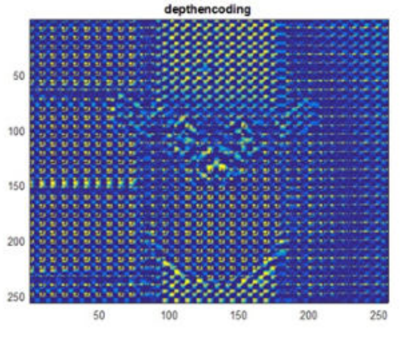


(a)

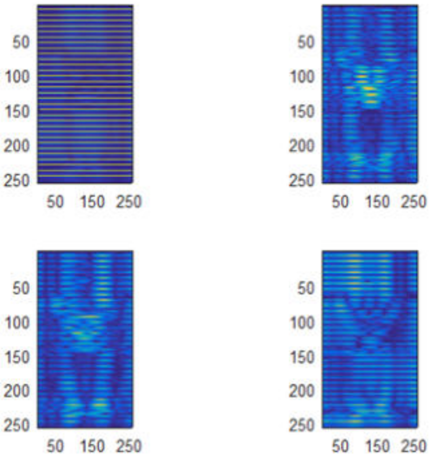


(b)

(d)



(c)



(d)

a) Original Videos left and right views are taken and it is normalized and the mean of the two frames are found out.

b) Frame are made blocks for stationary purposes

c) Adaptive wavelet transform is applied

d) The wavelet coefficients are added together and finally the PSNR value is founded out.

IV. CONCLUSION

Here propose a novel approach to build wavelets adapted to represent the processed signal. This method can estimate a relatively high PSNR value for virtual view 3DV.In the proposed method the adaptive wavelet transforms is used, adaptive means it changes according to the signal given. The original 3DV frames are taken and from that a particular row is chosen which is 1D signal and then Fourier transform is applied. The AM-FM frequency and the amplitude is changed according to the time. Here at each instant frequency information can be find out. Depending on the frequency whether it is high then depth is there like this depth information can be founded out, and then the bit error is reduced and finally PSNR for a corresponding 3D video frame is estimated. The key idea is to build a wavelet filter bank based on Fourier supports detected from the information contained in the processed signal spectrum. The proposed empirical wavelets correspond, in the temporal domain, to dilated versions of a single mother wavelet. It can be observed that the proposed method could estimate the PSNR trends (along with frames) accurately.

This method can estimate a comparatively high PSNR value for virtual view 3DV.In this paper an adaptive wavelet transform is used, adaptive means it changes according to the signal given. Here our original 3DV frames are taken and from that a particular row is chosen which is 1D signal and then fourier transform is applied.We get an AM-FM frequency and the amplitude is changed according to the time.Here at each instant frequency information we can find out. Depending on the frequency whether it is high then depth is there like this depth information can be founded out,and then the bit error is reduced and finally PSNR for a corresponding 3D video frame is estimated.

REFERENCES

[1] A Virtual View PSNR Estimation Method for 3-D Videos Hui Yuan, Member, IEEE, Sam Kwong, Fellow, IEEE, Xu Wang, Student Member, IEEE, Yun Zhang, Member, IEEE, and Fengrong Li

[2] L. Onural, A. Gotchev, H. M. Ozaktas, and E. Stoykova, "A survey of signal processing problems and tools in holographic three-dimensional television," IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 11, pp. 1631–1646, Nov. 2007.

[3] L. Fang et al., "An analytical model for synthesis distortion estimation in 3D video," IEEE Trans. Image Process., vol. 23, no. 1, pp. 185–199,

[4] M. Tanimoto and T. Fujii, "FTV—Free viewpoint television," in Proc.61st Meeting ISO/IEC JTC1/SC29/WG11, Doc. M8595, Flagenfurt, Austria, Jul. 2002.

[5] A. Alatan et al., "Scene representation technologies for 3DTV—A survey," IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 11, pp. 1587–1605, Nov. 2007.

[6] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," in Proc. SPIE Stereoscopic Image Process. Render., vol. 5291. San Jose, CA, USA, Jan. 2004,pp. 93–104.

[7] Y. Zhang et al., "Regional bit allocation and rate distortion optimization for multiview depth video coding with view synthesis distortion model," IEEE Trans. Image Process., vol. 22, no. 9, pp. 3497–3512, Jun. 2013.

[8] W.-S. Kim, A. Ortega, P. Lai, D. Tian, and C. Gomila, "Depth map distortion analysis for view rendering and depth coding," in Proc. IEEE Int. Conf. Image Process. (ICIP), Cairo, Egypt, Nov. 2009,pp. 721–724.

[9] H. Yuan, Y. Chang, J. Huo, F. Yang, and Z. Lu, "Model based joint bit allocation between texture videos and depth maps for 3D video coding,"IEEE Trans. Circuits Syst. Video Technol., vol. 21, no. 4, pp. 485–497,Apr. 2011.

[10] J. Xiao, T. Tillo, and H. Yuan, "Macroblock level bits allocation for depth maps in 3D video coding," J. Signal Process. Syst., vol. 74, no. 1, pp. 127–135, Jan. 2014.

[11] S. Hu, S. Kwong, Y. Zhang, and C.-C. J. Kuo, "Rate-distortion optimized rate control for depth map-based 3D video coding," IEEE Trans. Image Process., vol. 22, no. 2, pp. 585–594, Feb. 2013.

[12] F. Shao, G. Jiang, W. Lin, M. Yu, and Q. Dai, "Joint bit allocation and rate control for coding multi-view video plus depth based 3D video,"IEEE Trans. Multimedia, vol. 15, no. 8, pp. 1843–1854, Dec. 2013.

[13] L. Fang et al., "An analytical model for synthesis distortion estimation in 3D video," IEEE Trans. Image Process., vol. 23, no. 1, pp. 185–199, Jan. 2014.

[14] P. Lai, A. Ortega, C. C. Dorea, P. Yin, and C. Gomila, "Improving view rendering quality and coding efficiency by suppressing compression artifacts in depth-image coding," in Proc. Visual Commun. Image Process. (VCIP), vol. 7257. San Jose, CA, USA, Jan. 2009, pp. 72570O.1–72570O.10.

[15] H. Yuan, J. Liu, H. Xu, Z. Li, and W. Liu, "Coding distortion elimination of virtual view synthesis for 3D video system: Theoretical analyses and implementation," IEEE Trans. Broadcast., vol. 58, no. 4, pp. 558–568, Dec. 2012.