

A New Perceptual Mapping Model Using Lifting Wavelet Transform

TahaBasheer Taha¹,Phaklen Ehkan¹, Ruzelita Ngadiran¹

¹School of Computer and Communication Engineering, University Malaysia Perlis (UniMAP), Perlis, Malaysia

Abstract. Perceptual mapping approaches have been widely used in visual information processing in multimedia and internet of things (IOT) applications. Accumulative Lifting Difference (ALD) is proposed in this paper as texture mapping model based on low-complexity lifting wavelet transform, and combined with luminance masking for creating an efficient perceptual mapping model to estimate Just Noticeable Distortion (JND) in digital images. In addition to low complexity operations, experiments results show that the proposed model can tolerate much more JND noise than models proposed before.

Keywords: JND; Accumulative Lifting Difference (ALD); Lifting Wavelet Transform (LWT); Perceptual Mapping.

1. Introduction

Human Visual System (HVS) cannot perceive visual alerts unless it pass a certain threshold called Just Noticeable Distortion (JND)[1]. Hence, JND models are employed as pre-processing step in implementation of different image processing techniques as enhancement, compression and watermarking to increase the appearance quality of digital images and videos. As consequence, a lot of attempts were adopted to estimate JND in transform domain or in time domain.

An early Discrete Cosine Transform (DCT) based JND masking attempt was proposed by Watson [2], the model estimates the perceptibility of frequency sensitivity, luminance masking and contrast masking, these factors are polled to create what is called Watson distance (D_{wat}). Another DCT based perceptual model is proposed by Niu et.al [3], where the spatial Contrast Sensitivity Function (CSF), luminance adaptation, and contrast masking are combined to produce the JND estimator. Recently, Fazlali et.al [4] proposed a JND model for estimating maximum watermark embedding strength. DCT is applied on the detailed band in second level of Contourlet Transform (CT) after its partitioning into blocks. Severity of embedding is found by two factors, edge concentration and entropy of each block.

Although an acceptable performance is achieved by DCT transform with JND models, the usage of Discrete Wavelet Transform (DWT) comes with better performance, this due to temporal-spatial characteristics which similar to the Human Vision System behavior[5]. Moreover, In DWT, if the coefficient is modified, only the corresponding region where the coefficient exists is affected [6]. These factors motivated researchers to

adopt DWT for creating perceptual masks. Barni et.al [6], proposed a pixel wise JND model where HVS characteristics are estimated by analyzing the imperceptibility of four DWT decompositions. Barni's model is utilized in other models, as [7] where the entropy masking is added besides previous factors for JND mask enhancement. Logarithmic calculations used in entropy determination are complex operations which increase the computational overhead, furthermore, and generally speaking, despite of accepted results achieved by analysing perceptual features using traditional DWT, it relies on Fourier transform which involves imaginary number calculations. Also, different filters are used for different pixels' positions [8]. These factors will bound the usage of transform domain models in basic processors found in embedded systems where resources such as CPU speed, number of logic elements and memory are limited.

In spatial domain attempts, JND values are directly estimated, Chou and Li's [9] proposed a model where JND is found by estimating the background luminance using Weber law and contrast masking effect is determined with maximum signal from four edges. An improvement of Chou's model is proposed by Yang et.al [10] where edge detection is added to avoid over estimation in edges areas. Wu et.al [11] proposed a pixel adaptive JND model to enhance the estimation of texture in previous two attempts, the model characterized by considering the disorder of texture in texture masking and adding it to luminance factor. However, texture areas are yet underestimated and noise tolerance can be significantly enhanced.

The complexity of transform domain and under estimation of spatial domain attempts when creating perceptual mask evokes the need for a perceptual model

that combine features of both, simplicity and reliable estimation. In this paper an efficient and high accuracy perceptual mapping model is presented by taking advantage of Lifting Wavelet Transform (LWT) presented by Sweldens [8]. LWT is characterized by simplicity, on location processing, and filter similarity. As a consequence, proposed model will ease the usage of texture masking in embedded systems and limited processors applications.

In next section an introduction of the LWT is presented, proposed methodology explained in section 3, followed by experimental results and comparisons with existing works in section 4. Conclusion written in section 5

2. Lifting Wavelet Transform

Lifting wavelet transform (LWT) is introduced by Sweldens[8] to efficiently construct wavelet coefficient without relying on Fourier transform. It provides low computational process in compare with filter banks used in regular Discrete Wavelet Transform (DWT). Also, LWT has integer to integer computation which makes it suitable to be mapped to hardware design where the floating points are not often recommended. LWT is applied by three major steps [8][12]:

Split

Split step also called the lazy wavelet, where the signal is being separated into smaller subsets. Generally, coefficients are divided into odd and even samples.

Predict

Even samples are used to predict the odd ones, pixel at odd position are predicted by its two neighbors at even positions, then the difference between the predicted value and real odd value is stored in the location of odd samples. Signal after prediction step is the details band. Predict step can be written as (1)

$$Y_{2n+1} = X_o - \text{PREDICT}(X_e) \quad \dots(1)$$

Update

When the average of each two pixels in a signal is calculated, the overall structure of decomposed signal is obtained with half energy [13]. However, due to the non-linearity changing in image pixels, even samples interposed the odd ones cannot be taken directly as they need to be "Updated" with the differences computed in predict steps (2). Produced signal is the approximation band.

$$Y_{2n} = X_e + \text{UPDATE}(Y_{2n+1}) \quad \dots(2)$$

Inverse of lifting scheme is performed by reversing the order and exchanging the sign of predict and update steps [12]. Figures 1, 2 show LWT decomposition and reconstruction operations respectively.

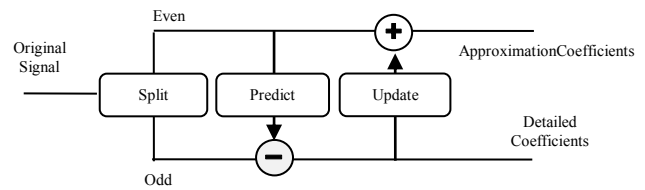


Fig.1. Decomposition of lifting wavelet transform

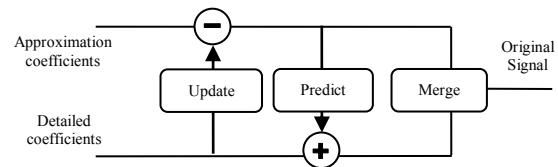


Fig.2. Reconstruction of lifting wavelet transform

The core idea beyond lifting scheme's signal analysis is to divide the original signal into approximation and details bands. Higher Coefficients values in details band refer to higher nonlinear changes in pixels' values, this changes is the region where texture areas found. Proposed method exploits this property to create ALD texture masking model.

3. Proposed Method

To simulate human visual system, Luminance and Texture are employed and combined for creating the proposed perceptual masking model.

3.1 Luminance Masking:

A modified version of luminance equation listed in [14] is considered in this scheme, it states that, according to Chou's law[9], if the gray level of a certain pixel refers to very dark areas and very bright areas, as zero or 255, then the masking value can be 20. For example, pixel can be changed from zero to 20 and the difference will not be observed by the human eye. And when the gray level is in middle intensity (127) it can tolerate up to three values. For better appearance, tolerance for both very bright areas (255) and very dark areas (0) are set as five. Equation of luminance sensitivity is given as follows:

$$L_{Mask}(I,J) = \begin{cases} \text{floor}((Im1(I,J)-128)/32) + 1 & \text{if } S1(I,J) > 127 \\ \text{floor}(128 - (Im1(I,J))/32) + 1 & \text{Otherwise} \end{cases} \quad \dots(3)$$

Where L_{Mask} is luminance mask, $Im(I,J)$ is the luminance of pixel at (I,J) .

3.2 Texture Masking:

As defined in [15], according to statistical approach, "texture is a quantitative measure of the arrangement of intensities in a region". So, frequent change in image

intensities produce more textures area. In fact, the measurement of that disparity is already given by the detailed band equation of lifting scheme (Equation 1), presented by the difference between the actual value and expected value of pixel's intensity. This feature will be exploited to design the proposed texture scheme by dividing the LWT detailed subband (D2) into 5*5 blocks then subtracting neighbored pixels in each row of the block from each other, summarize the absolute of the differences of that subtraction and accumulate the results into single value. The result is Accumulative Lifting Differences (ALD) which is the proposed texture mapping model. (Figure 3).

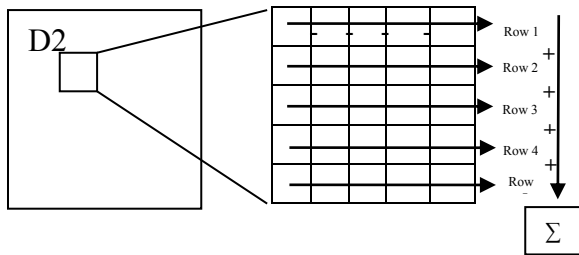


Fig.3. ALD block calculation

Details band (D2) shown in figure 3, has been elected because its coefficients have the difference between approximate value and real value which refer to texture variation. ALD equation is given in (4).

$$ALD(I, J) = \sum_{i=I-2}^{I+2} \sum_{j=J-2}^{J+2} |D2(J, J-1) - D2(I, J)| \dots (4)$$

Where I, J are centre pixels of each 5*5 block. Final mask (T) is the combination of ALD and LMask(5)

$$T(I, J) = L(I, J) + ALD(I, J) \dots (5)$$

After LWT decomposition, each block in spatial domain will have the same T value for the centre coefficient in S1 band.

4. Results and Discussion:

To evaluate the performance of proposed model, an image with texture and smooth areas is chosen where a set of distributed pixels have been elected for testing. Each pixel indicated by arrow and number in figure 4. Luminance mask values and texture intensity calculated using proposed ALD model is shown in table 1.

For luminance masking, tolerance (mapping value) is increased as the pixel value is close to dark and bright areas. For texture model, ALD values on smooth sky areas are 31, 25, while in mountain surface where high texture intensity can be observed, higher ALD values are obtained, as 495, 244. Figure 5 shows the normalized

texture mask for Mountain image, where the brighter areas refer to more tolerance to modification.

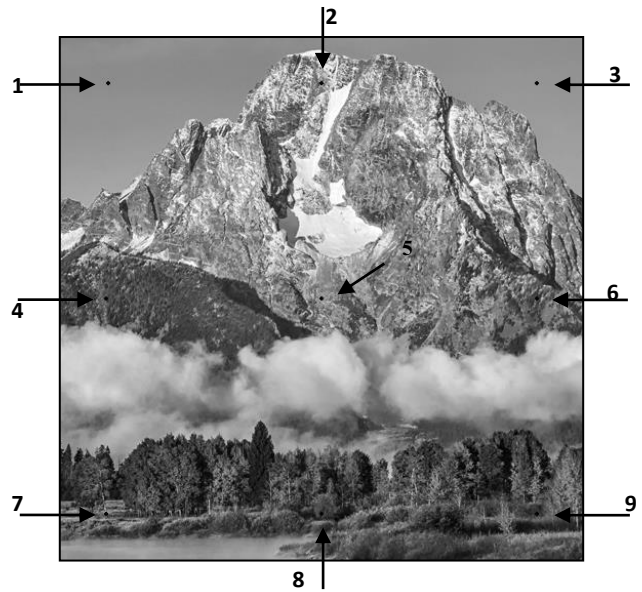


Fig.4. Original Mountain image

Table 1: Luminance toleration and ALD values for different Blocks

Block	Location of centre pixel (Y, X)	Luminance value	Tolerance of luminance value	ALD Value
1	46,46	139	5	31
2	46,256	168	9	495
3	46,466	147	6	25
4	256,46	88	9	123
5	256,256	140	5	244
6	256,466	72	11	239
7	466,46	102	7	391
8	466,256	65	11	204
9	466,466	109	11	109

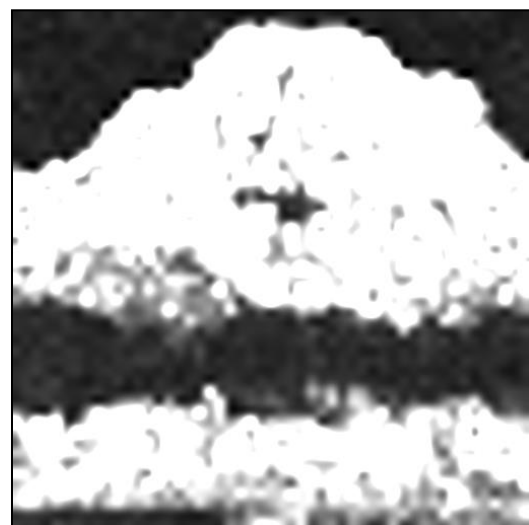


Fig.5. ALD Texture Mask for Mountain image

For objective evaluation, Mean Structure Similarity Index (SSIM)[16], a metric that used to predict observed quality of images, is utilized in this paper. Comparisons and evaluation can be established by fixing SSIM value and measure the amount of noise can be injected. Mean Square Error (MSE), is the metric that used for evaluation purpose, where (by fixing SSIM) higher MSE value refers to more tolerance to noise. JND noise guided equation is:

$$f_{noise}(x,y) = f(x,y) + srT(x,y) \quad \dots(6)$$

Where s is the random noise as +1 or -1 values, and r is severity of noise. Table 2 shows MSE obtained by applying proposed mask on different standard images with constant SSIM value.

Table 2 : MSE for different images with fixed SSIM value

Image (512*512)	SSIM	MSE
Airplane	0.973	36.42
Babon	0.973	99.39
Boat	0.973	28.11
Cameraman	0.973	102.47
Lena	0.973	26.78

Proposed model is compared with three other JND estimation models, Chou's and Li[9], Yang et.al [10] and Wu et.al [11] models. Baboon image (512*512) is chosen for the comparison between models. Figure 6 shows the noise contaminated images and JND masks for the proposed and compared models. For Chou's and Li's model (a) (b), MSE is 33.35, in Yang et.al (c) (d) MSE is 38.09, while in Wu et.al model (e) (f) MSE is 54.59. In proposed model (g) (h), MSE is 99.39 for the same SSIM value (0.9738). Results shows that the proposed model has better estimation for texture areas and tolerate more noise than other models.

5. Conclusion

Accumulative Lifting Differences (ALD), a new texture mapping model based on low complexity lifting wavelet transform (LWT) is proposed in this paper. ALD is combined with luminance masking to create a novel perceptual model that helps to create JND mask for digital images. Proposed model combines the simplicity of time domain models and the performance of transform domain models. Experimental results shows that the proposed model has efficient estimations of texture regions and outperforms compared models in amount of noise can be injected without affecting perceptual appearance.

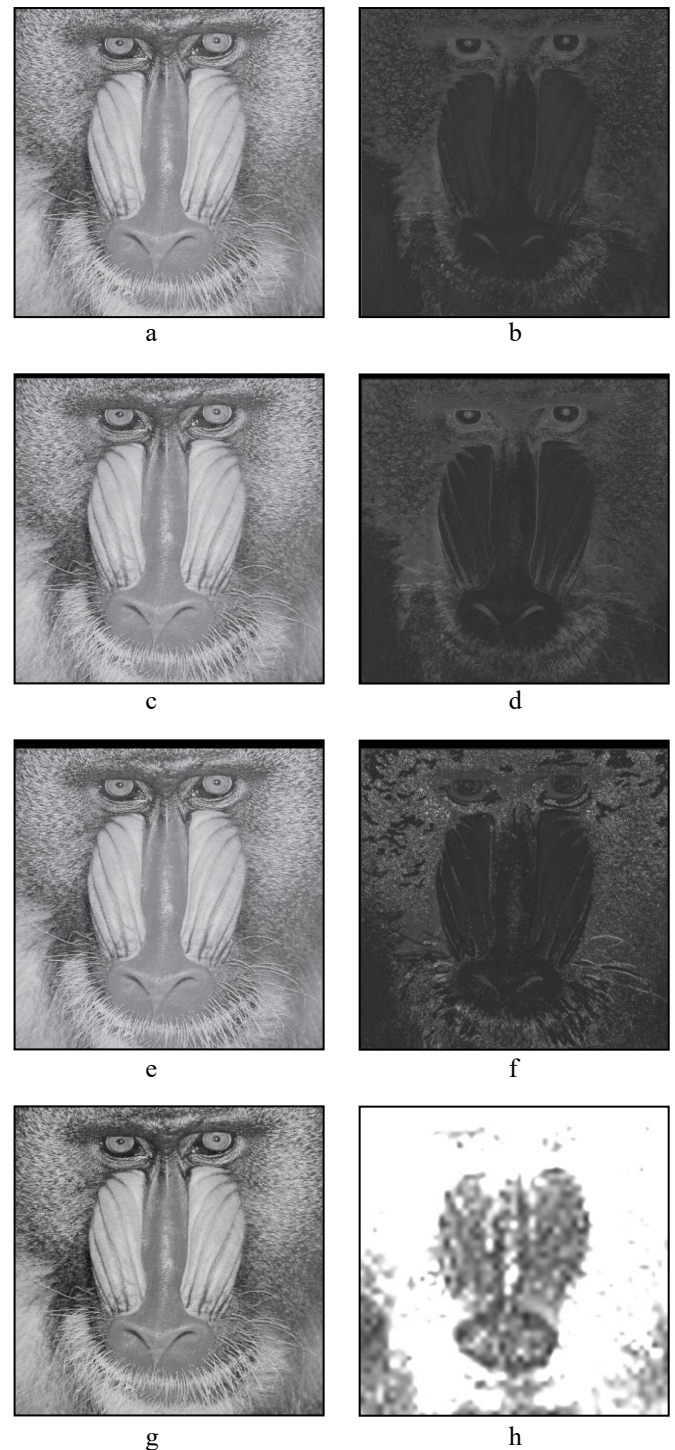


Fig.6. Baboon Image with injected noise (Left), and noise mask (Right) for different models.

References

1. Wu, J., Qi, F., & Shi, G. (2010, March). An improved model of pixel adaptive just-noticeable difference estimation. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on* (pp. 2454-2457). IEEE.
2. Watson, A. B. (1993, December). DCTune: A technique for visual optimization of DCT quantization matrices for individual images. In *Sid International Symposium Digest of Technical Papers (Vol. 24, pp. 946-946)*. SOCIETY FOR INFORMATION DISPLAY.
3. Niu, Y., Kyan, M., Krishnan, S., & Zhang, Q. (2011). A combined just noticeable distortion model-guided image watermarking. *Signal, Image and Video Processing*, 5(4), 517-526.
4. Fazlali, H. R., Samavi, S., Karimi, N., & Shirani, S. (2016). Adaptive blind image watermarking using edge pixel concentration. *Multimedia Tools and Applications*, 1-16.
5. Li, W., Zhang, Y., & Yang, C. (2013). A survey of JND models in digital image watermarking. In *Proceedings of the 2012 International Conference on Information Technology and Software Engineering* (pp. 765-774). Springer Berlin Heidelberg.
6. Barni, M., Bartolini, F., & Piva, A. (2001). Improved wavelet-based watermarking through pixel-wise masking. *IEEE transactions on image processing*, 10(5), 783-791.
7. Akhbari, B., & Ghaemmaghami, S. (2005, November). Watermarking of Still Images in Wavelet Domain based on Entropy Masking Model. In *TENCON 2005-2005 IEEE Region 10 Conference* (pp. 1-6). IEEE.
8. Sweldens, W. (1995, September). Lifting scheme: a new philosophy in biorthogonal wavelet constructions. In *SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation* (pp. 68-79). International Society for Optics and Photonics.
9. Chou, C. H., & Li, Y. C. (1995). A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. *IEEE Transactions on circuits and systems for video technology*, 5(6), 467-476.
10. Yang, X. K., Ling, W. S., Lu, Z. K., Ong, E. P., & Yao, S. S. (2005). Just noticeable distortion model and its applications in video coding. *Signal Processing: Image Communication*, 20(7), 662-680.
11. Wu, J., Qi, F., & Shi, G. (2010, March). An improved model of pixel adaptive just-noticeable difference estimation. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on* (pp. 2454-2457). IEEE.
12. Gholipour, M. (2011, June). Design and implementation of lifting based integer wavelet transform for image compression applications. In *International Conference on Digital Information and Communication Technology and Its Applications* (pp. 161-172). Springer Berlin Heidelberg.
13. Jensen, A., & la Cour-Harbo, A. (2001). The Discrete Wavelet Transform via Lifting. In *Ripples in Mathematics* (pp. 11-24). Springer Berlin Heidelberg.
14. Qi, H., Zheng, D., & Zhao, J. (2008). Human visual system based adaptive digital image watermarking. *Signal Processing*, 88(1), 174-188.
15. Shapiro, L., & Stockman, G. C. (2001). *Computer vision*. 2001. ed: Prentice Hall.
16. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4), 600-612