Review of HVS-based Image Compression Methods

Hanadi Ahmed Hakami
Unversity of Technology Sydney
Faculty of Engineering and Information Technology

Arwa Darwish Alzughaibi
Taibah University
Community Collage
Unversity of Technology Sydney

Zenon Chaczko
Unversity of Technology Sydney
Faculty of Engineering and Information Technology

ABSTRACT

Digital images can be compressed by using data compression processes with the aim of lessening image data redundancy and save or send out an efficient form of data. This is not an easy task because large volumes of data comprise an original image, requiring huge memory allocation and inconvenient transmission. In this paper, the Human Visual System-based Image Compression is reviewed based on set criteria. Specifically, the characteristics, compression techniques and wavelet-based compression are discussed.

Keywords

Image Compression, HVS Characteristics, Wavelet Based Image Compression

1. INTRODUCTION

Images, in their original form, contain large volumes of data, which require huge memory allocation and inconvenient transmission especially when using channels with low bandwidth. In such cases, a technique called image compression is used to reduce the dimension of the representation while keeping the quality. The image compression involves the identification of perceptual and spatial redundancies and keeping the required bits to a minimum without compromise to the visual quality.

There are two classifications are known for image compression techniques: the lossless and the lossy processes. In lossless processing, the reconstructed image after compression is digitally identical to the original images with no loss of information. Furthermore, lossy compression involves a high compression ratio but obtains a lower quality of reconstruction because large data or information tends to be lost during the process. The specific process of compression involves a transformation of image data to some other domain such as wavelet or frequency. During image compression process, image quality, compression rate and compression time are important factors. For example, higher rate of compression could degrade the quality of the image while attaining better visual quality could be achieved through lower compression rate. Nonetheless, a model called Human Visual System (HVS) can also be used to optimize how image compression techniques perform and how visual quality of an image perceived by

embedding it to the compression algorithm. This paper gives the review of HVS-based Image Compression Methods.

2. HUMAN VISUAL SYSTEM CHARACTERISTICS

Considerable improvements in the quality of the processed output have been seen when HVS model is embedded in compression algorithms [1], [2], [3], [4], [5]. Essentially, this is achieved by optimizing image compression performance through the exploitation of HVS limitations when compression is performed. Nevertheless, it has some limitations concerning the visual perception. However, these limitations can be exploited in image compression. The sensitivity reduced for the spatial high-frequency structures is one of the most important limitations of the HVS. This phenomenon is shown by the contrast sensitivity function (CSF) [6], [7], [8]. To continue this process, psychophysical observations are done in order to properly assess the HVS-model's performance. This is rigorously done because human visual processing system is highly complex. Through the years, numerous experiments have been conducted by physiologists in relation to psycho-visual processes and the understanding of HVS.
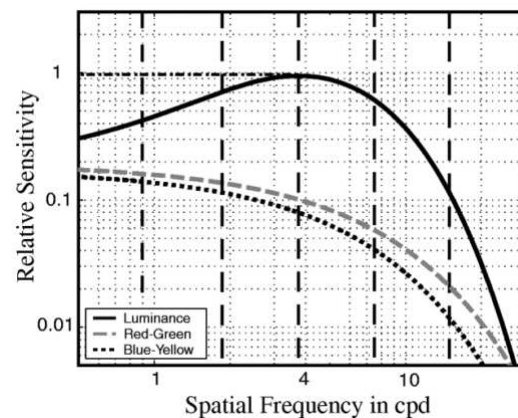

Fig. 1: CSF curves in corresponding HVS chrominance and luminance channels. Source [8]

In one of these experiments,however, it was found that HVS was found to possess a lower sensitivity when high spatial frequency patterns are involved. Nonetheless, it was also found that this characteristic of the HVS model can be used in the improvement of a compressed image's visual quality by embedding the HVS into the compression algorithms. This makes contrast sensitivity, being a spatial frequency function, be described in detail through the Contrast Sensitivity Function (CSF) [8], [9]. When CSF was applied to a certain image as illustrated in Figure 1, HVS is seen to have a higher sensitivity to the components related to luminance than that of the chrominance dimensions. When seen in the 4 cycles per optical degree (cpd) region, HVS possesses greater luminance sensitivity when it is at the midfrequencies. Then it quickly reduces when it follows higher spatial frequencies, while it slightly becomes smaller at lower spatial frequencies. In relation to chrominance components, the HVS follows the behavior of a low pass-filter, such that lower frequency sensitivity remains the same.

To clarify, Figure 2 shows the HVS-model's sensitivity in two different frequencies. In here, two patterns in white and black appear with their respective frequency patterns, with the left side showing low frequency and the right side the high frequency pattern. The two patterns in particular demonstrate the same levels of brightness on the left side, while the patterns on the right side show a less intense frequency pattern than the on the left. This implies that HVS has a lower sensitivity to components with higher frequency.

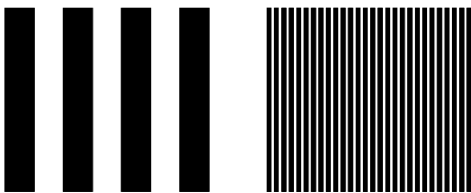Fig. 2: Frequency patters and their intensity. The low frequency pattern is shown on the left and the opposite shows the high frequency pattern. Source [8]

3. HUMAN VISUAL SYSTEMS IN COMPRESSION TECHNIQUES

In the embedding of complete HVS models, it has been tested that wavelet-based image coding schemes are ideal. This suitability is attributed to the wavelet decompositions' properties related to spacefrequency localization [7].In here, the embedding of HVS model could either be performed at the stage of quantization [1], [3, 8], or during the codec's particular stage of bit allocation [2, 5]. When this is done, a considerable improvement is seen in the new image in relation to visual quality. Similarly, another scheme for wavelet-based image compression through the combination of Vector Quantization (VQ) and some HVS properties shuch as edge and texture preference, CSF and masking effects was developed by others [2]. A vector quantization of coefficients in various sub-bands which were obtained through Discrete Wavelet Transform (DWT) is done by the algorithm. Specifically, weighted Mean Squared Error (MSE) distortion criterion is used as a basis for the allocation of bits among the various sub-bands., This specifically goes through a process of identifying weights wherein the HVS property is used as a basis [10]. Another compression technique that is wavelet-based has been presented in another

study [11]. In this technique, was utilised a combination of VQ and HVS properties through a computation the value of Contrast Sensitivity Function (CSF), specifically for the sub-bands' central spatial frequency. The derived values were subsequently used in the scaling of all the sub-bands' threshold value. Then, vector selection was conducted by using the obtained values, before VQ process commenced. Similarly, an application of weighted MSE distortion criterion through the use of perceptual weights is also carried out. This is done with the aim of designating bits to the various sub-bands. Considerable improvements have been described by Voukelatos and Soraghan when block-based image compression methods are performed at bitrates that are very small.

In here, the CSF value was calculated, specifically the Quarter Common Intermediate (QCIF) image size for every sub-band. Next, a scaling of the threshold value was done using the obtained values. The values derived from this process were utilized for the vector selection process. Lastly, VQ process and allocation of bit in various sub-bands were then conducted. Two kinds of visual frequency weighting are used in JPEG 2000 standard image codec, which are the Fixed Visual Weighting (FVW) and the Visual Progressive Coding (also known as Visual Progressive Weighting or VPW). FVW and VPW differ for on the number of CSF weights selected and applied. In other words, FVW selects and applies only a single set of the CSF weights while VPW involves different sets for the different phases throughout the method of embedded coding because an image is seen at varying distances when the progressive transmission stage is undergoing.

To demonstrate, images perceived at bitrates that are low show a farther distance, increasing the bits increases image the quality because a nearer distance is perceived. An integration of HVS characteristics and wavelet-based image compression algorithm has been done through a noise-shape filtering stage before quantization [8]. In here, transformed coefficients were filtered through the use of "HVS filter applied to all sub-bands. As a result, a significant improvement pertaining to the compression ratio was seen. Specifically, about 30% more than the JPEG2000 baseline was seen when it was tested in various images. Furthermore, other researchers proposed other image compression methods with the use of HVS [1]. This new technique engages a decomposition of the image that was inputted through the application of wavelet transform. This is followed by weighing of the coefficients that were transformed by considering the contrast sensitivity function (CSF) curve, particularly the peak, as seen in the domain. The final process of this scheme is the coding of weighted wavelet coefficients with the use of Set Partitioning as seen through the algorithm of Heretical Tree (SPIHT). With this, results show that the visual quality is a significantly higher, while the objective quality is similar with the output of conventional SPIHT technique.

4. WAVELET-BASED IMAGE COMPRESSION

Wavelet basically refers to "small wave, wherein its characteristic of being small implies finite length as seen in window function. Certain requirements related to mathematical properties are considered in this function called wavelet. This function is also commonly used when representing a data or some other functions. Two types of wavelet transforms are known: the wavelet that is invertible and the wavelet for signal analysis. The first type is typically used for compressing and cleaning images by following a pattern that requires reversing or inverting of the wavelet transform. This transformation is needed to do an efficient recovery of the original signal. The process starts with the calculation of the wavelet representation and followed by a proper modification of the

calculated wavelet. This wavelet transform subsequently undergoes reversal or inversion as the final process of obtaining the concluding output or the new image. On the other hand, the second wavelet transform type is mainly used during the process of signal analysis. This is typically used in determining faults in the machinery as seen using sensor measurements. This is also
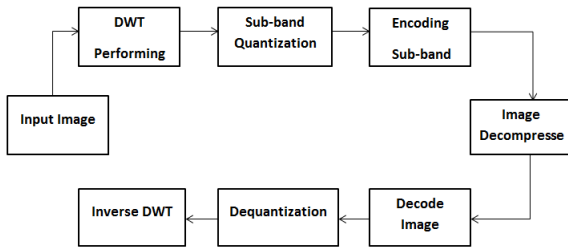
Fig. 3: Basic Steps in Wavelet Compression

aimed at understanding EEG and other signals that are usually known in the biomedical field, as well as in understanding the process of evolution of the signal's frequency content. A wide application of the wavelet transform has been seen among the researchers of image coding schemes through the years. During these applications, the coefficients of the wavelet transforms and their statistical behavior are important in the design of an efficient algorithm for a compression. This is because of the respective parameters being used in every statistical distribution function. This implies that analysis of uncompressed image through a recursive manner is seen in the process of wavelet compression; and as a result, images with higher resolution are generally obtained. Some of the advantages of wavelet technique are that it is quick, it operates in a linear way, it is invertible and it is orthogonal. In addition, an easier way of reconstruction is also experienced, while simultaneous display of information related to frequency and time are shown in window size.

The basic steps in doing wavelet compression are shown in Figure 3. In this process, discrete wavelet transformation (DWT) is first conducted, which is followed by sub-band quantization, specifically of image space's wavelets. This is followed by the encoding process where images are compressed. A corresponding image decompression process and decoding of image also follows. Moreover, discrete wavelet transformation is dequantized and inversed. As a spatial transform, the DWT is a good alternative for the reduction of computation effort; however, there is a setback in relation to how the perceptual sub-bands correspond to the usual wavelet ones in an indirect manner.

5. WAVELET-BASED IMAGE COMPRESSION USING HUMAN VISUAL SYSTEM

Human visual system (HVS) generally refers to the application of mathematical models in the understanding of humans' way of perceiving the world and the things surrounding them. Two essential properties of HVS are integrated into compressed data representation when used in techniques related to image compression: orientation sensitivity and contrast sensitivity. These HVS properties represent the spatial frequency sensitivity of humans. Mannos and Sakrison, at the onset, proposed a CSF model,
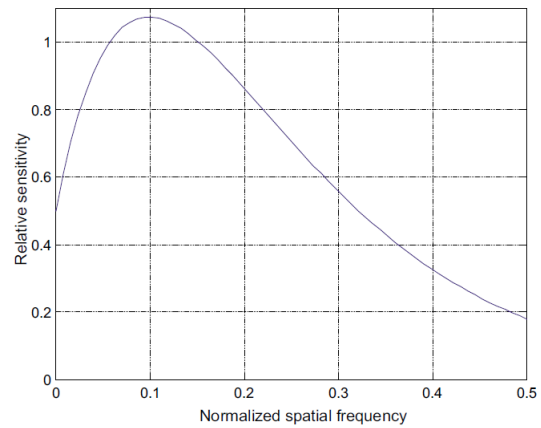
Fig. 4: Luminance Contrast Sensitivity Function. Source [12]

specifically for luminance (grayscale) images by putting forth the formula:

$$H(f) = 2.6 (0.192 + 0.114f) e^{\left[-(0.114f)^{1.1}\right]} \quad (1)$$

In here, f = (f2 x + f2 y) 0.5 with units of cycles/degree represents spatial frequency. fx represents the spatial frequency in the horizontal direction and fy is along the vertical direction. Furthermore, spatial frequency is normalized by:

$$f(cycles/degree) = fn(cycles/pixel) . fs(pixels/degree) \quad (2)$$

Where, fn represents normalized spatial frequency having at cycles/pixel units. If the observers are located at a 2 feet viewing distance, the fs is at 64 pixels /degree. This means that this fs value implies a viewing distance that is equivalent to four times of the image's height which is about 6 inches. Equation (2) shows that the range of fn is located between 0 and 0.5 because f ranges between 0 and fs/2. The CSF curve is illustrated in Figure 4 where HVS' luminance sensitivity is seen to have a role in the normalization of the spatial frequency. CSF is known and used as a bandpass filter and does the process of CSF masking. In relation to this, HVS shows an optimum sensitivity level between 0.03 and 0.23 for spatial frequencies while maintaining less sensitivity to frequencies that are very low as well as to very high ones.

In CSF masking, wavelet coefficients relative to perceptual importance are weighed. The process of CSF masking and compression system inversion is shown in Figure 5 [13]. Given the CSF curve (Figure 4), the transformation process from a mask of weights to multiplied wavelet coefficients is put forth as a design problem. When considered in uncompressed form, a typical representation of a digital grayscale image can be at 8 bits per pixel (bpp). The pixels have corresponding values starting from 0 (which represents black) to 255 (which represents white), and others in between correspond to different colors. When applied to two dimensional (2D) images, the operation of transform method typically start through the rows, then followed by applying it on columns. This process is referred to as separable. In principle, a three-dimensional array is used for the storage of digital color images. Similarly, 24 bits are used for every pixel representation when it is still uncompressed. Every pixel is represented by values from 0 to 255, and specifically clustered as red (R), green (G), blue (B) component or RGB format. The process of image compression schemes starts with the conversion of the image. In here, luminance (Y), chrominance-blue (Cb), and chrominance-red (Cr) compose the color space representation separating the image

information. This scheme is known to be of higher quality than the RGB as shown in equations (3), (4), and (5). The image intensity is represented by luminance which is almost identical with the grayscale version. On the other hand, the components known as chrominance-blue and chrominancered constitute the image's color information. The following formula for image transformation through the RGB to Y Cr Cb transition is followed if a higher quality image compression is aimed, which significantly match the human visual system:

$Y = 16 + 65.481.R + 128.553.G + 24.966.B$ (3)
$Cb = 128 - 37.797.R - 74.203.G + 112.B$ (4)
$Cr = 128 + 112.R - 93.786.G - 18.214.B$ (5)

The next step involves the computation of perceptual weighting factors through the variable called mean detection threshold. This is known as the smallest change seen in a color, and the tiniest part of the change which a human observer can discern. Factors such as spatial frequency, luminance, background color, and orientation are at play in this threshold. From RGB format and into a different color scheme. Measurement in xyY color space, coordinates for C.I.E. chromaticity are represented by the x and the luminance as Y [12].

Spatial Direction	Colour Direction	Spatial frequency cycles/deg				
		1.0	2.0	5.0	10.0	20.0
Horizontal (LH)	Luminance	6.750	6.330	7.250	13.500	65.083
	R-G	4.750	4.750	7.617	17.417	77.417
	B-Y	6.000	6.833	32.667	70.167	150.000
Vertical (HL)	Luminance	6.833	6.250	6.833	22.500	77.800
	R-G	5.583	7.083	9.250	23.000	90.375
	B-Y	6.667	9.417	31.833	65.700	150.000
Left (HH)	Luminance	7.667	6.917	11.167	37.083	49.000
	R-G	7.917	7.167	16.083	37.500	100.750
	B-Y	12.417	18.500	45.500	86.500	150.000
Right (HH)	Luminance	8.083	7.583	9.167	42.583	85.750
	R-G	7.750	6.333	13.833	35.417	103.500
	B-Y	13.750	19.750	47.750	83.000	114.000

Table 1. : Mean Detection Thirshold In XYY Space. Source [12]

To present the thresholds as used in different frequencies, Table 1 is shown, luminance Y orientations were also measured with the directions of Red-Green and Blue-Yellow relations, deriving a luminance value of Y_0 is $5cd/m^2$ with white as the background color. Values representing the white background's chromaticity coordinates are: $(X_0, Y_0) = (0.33, 0.35)$ respectively. Furthermore, two coordinates for chromaticity were derived from each of the mean detection thresholds for the Red-Green and Blue-Yellow direction transitions. In here, the coordinates correspond to respective maximum and minimum sinusoidal variations. The equations are shown representing the respective variations:

$x_i = x_0 \pm \triangle x.t$ (6)
$y_i = y_0 \pm \triangle y.t$ (7)

In here, t represents mean detection threshold, $\triangle x$ and $\triangle y$ illustrate step sizes when changes occur in both directions of $\triangle x$ and $\triangle y$. Table 2 presents the corresponding values

Direction	$\triangle Y$	$\triangle x$	$\triangle y$
Luminance	0.0124	0.0	0.0
R-G	0.0	0.000655	-0.000357
B-Y	0.0	0.000283	0.000689

Table 2. : Step Size Threshold In Each DirectionSTEP. Source [12]

SUBBAND	Mean detection thresholds		
	Luminance	R-G	B-Y
LH1	8.730	9.939	41.554
HL1	10.546	12.508	39.859
HH1	32.436	48.890	75.188
LH2	6.664	5.793	16.236
HL2	6.462	7.871	17.576
HH2	9.556	13.242	40.877
LH3	6.520	4.750	6.454
HL3	6.514	6.402	8.168
HH3	7.431	7.261	20.839

Table 3. : Mean Detection Threshold In XYY Space for Subband. Source [12]

for each direction. The perceptual weights for the sub-bands are shown as the reciprocal of the weight's mean detection threshold and these thresholds are computed prior to the identification of perceptual weights, particularly when YIQ space is considered. Table 3 presents the thresholds along the xyY space as a result of the computation for the high frequency subband centre. Subsequent to applying the prior equations (6 and 7), a calculation of two chromaticity coordinates (x_i, y_i, Y_0), in xyY space are derived. In here, computation of $i = 1, 2$ for all sub-bands is done. Then, the chromaticity coordinates undergoes a conversion from its original space going to C.I.E. XYZ space by applying the following equations in (8) [7].

$X_i = x_i Y_0/y_i$
$Z_i = (1 - x_i - y_i) Y_0/y_0$ for i= 1, 2 (8)

Two XYZ values are further computed in relation to luminance direction for all mean detection thresholds:

$X_i = X_0 \pm \triangle Y.\frac{x_0}{y_0}.t$
$Y_i = Y_0 \pm \triangle Y.t$
$Z_i = Z_0 \pm \triangle Y.\frac{z_0}{y_0}.t$ (9)

where $\triangle$ can be derived from Table 3,and t represents the mean detection threshold, with $i = 1, 2$ Equation (9) presents the derivation of vector (X_0, Y_0, Z_0) and the white point coordinates.
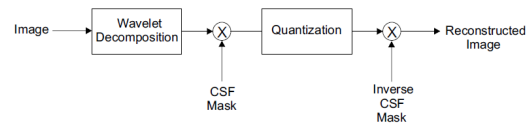

Fig. 5: CSF Masking Block Diagram. Source [12] [13]

6. CONCLUSION

This paper presented a review of the basic compression techniques as well as the characteristics of the human visual system. It is based on discussion of wavelet used in compressing images and with the application of the human visual system. Moreover, the most efficient objective image quality metrics have been reviewed in this paper in relation to Human Visual System (HVS) and its properties. A discussion on the existing wavelet decompositions, as well as the new HVSbased wavelet decompositions was also made in this paper. In this paper, it was investigated that the human eye does not really possess a gray-scale error sensitivity as seen through the human eye sensitivity. Nevertheless, a particular sensitivity is known in terms of image edge features. From this, the concepts

of fractal image coding and SPIHT algorithms were discussed in terms of their strengths and weaknesses and their possibility to compose a hybrid algorithm that completely uses the landscape features as well as the characteristics of human vision.

7. REFERENCES

[1] Wang Aili, Zhang Ye, and G U Yanfeng. SAR Image Compression using HVS Model. pages 1–4, 2006.

[2] Pierre Mathieu Barlaud and Ingrid Daubechies Antonini Marc Michel. Image coding using the wavelet transform. In *Image Processing, IEEE Transactions on 1*, pages 205–220, 1992.

[3] I Höntsch and L J Karam. Locally adaptive perceptual image coding. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 9(9):1472–1483, 2000.

[4] F Koch, S A E Lewis, S V Chernikov, and J R Taylor. Broadband Raman gain characterisation in various optical fibres. *Electron Lett*, 37(24):1437–1439, 2001.

[5] S.P. Voukelatos and J.J. Soraghan. Very low bit-rate color video coding using adaptive subband vector quantization with dynamic bit allocation. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(2):424–428, 1997.

[6] Fahmi Kammoun and Mohamed Salim Bouhlel. A Perceptual Image Coding method of High Compression Rate. *International Journal of Computer, Information, Systems and Control Engineerin*, 1(9):46–50, 2007.

[7] Andrew P. Bradley. A wavelet visible difference predictor. *IEEE Transactions on Image Processing*, 8(5):717–730, 1999.

[8] M J Nadenau, J Reichel, and M Kunt. Wavelet-based color image compression: Exploiting the contrast sensitivity function. *IEEE Transactions on Image Processing*, 12(1):58–70, 2003.

[9] Zherzghuu Yu. Perceptual Coding of Digital Images. *Proceedings 7th International Conference on Signal Processing, 2004. Proceedings. ICSP '04. 2004.*, 11(2):1167–1172, 2004.

[10] B Y F W Campbell and J G Robson. Application of Fourier Analysis to the Visibility of Gratings. *October*, pages 551–566, 1968.

[11] John J Soraghan. with application to progressive image transmission. 13:135–143, 1998.

[12] James S. Walker. Wavelet based image compression. *Transforms and Data Compression*, 79, 2001.

[13] Brian A. Wandell. Foundations of Vision. 1995.