An Objective Quality Measure Based on Subjective Measures for JPEG Compression

Anand S. ARABATTI
TeraNex, Inc.
Orlando, FL 32809, USA

and

Harley R. MYLER
School of Electrical Engineering And Computer Science
University of Central Florida, Orlando, FL 32816-2450, USA

ABSTRACT

In this paper, we have developed an objective measure based on subjective measures for JPEG compressed images in the case where the original uncompressed image is used as a reference. Traditional techniques incorporate either time or frequency domain features applied using full or partial reference systems. In the technique presented, we incorporate the time as well as the frequency domain features by using Gaussian wavelets as suggested by Nestares, *et al*, for representing transients in a 2-D signal. The subjective measure used is the *quality factor* in the freeware algorithm provided by the Independent JPEG Group for JPEG compression. Human Psychovisual factors have been incorporated in the development of this objective quality measure that is based on subjective measures. Blockiness and blurring artifacts, which are predominantly directional and non-uniformly present in JPEG compressed images, are identified after directional sub-band coding in the frequency domain and these form a subjective basis for objective quality assessment. Additionally, the modulation transfer function (MTF) forms an important basis for human psychovisual thresholding after the detection of blockiness, blurring and noise in compressed images.

Keywords. Blockiness, blurring, Gaussian wavelet transform, modulation Transfer Function.

1. INTRODUCTION

Image and video quality measurements form the basis of many applications such as image coding and compression to applications where precision of data obtained is of utmost importance, such as medical imaging [1], military applications, etc. The objective measurement techniques employed are either spatial or spectral domain methods. In the spatial domain techniques, the features used are standard data computations such as standard deviation and root mean square error [2]. However, these metrics do not consider specific objects transients that may be present in an image [3]. In subjective analysis of images and sequences, human psychovisual criteria are incorporated by definition [4]. The subjective criteria used in our technique is the *quality factor* in JPEG compression as it appears in the freeware algorithm provided by the Independent JPEG Group. The factor is proportional to the quantization level and so the metric makes intuitive sense because the greater the compression, the lesser the quality of the image.

In this paper, we incorporate some of the human psychovisual criteria that affect the subjective evaluation of images and sequences. The technique used to incorporate these criteria

has been described in detail and it can be used for effective evaluation of degraded images in full reference systems where the original image is available along with the degraded one.
This paper is organized as follows. Section 2 describes the compression artifacts present in images that lead to degraded quality. The use of the computationally efficient spatial-domain implementation of Gaussian wavelets as suggested by Nestares, *et a l*[5] can be effectively used for representing the image and carrying out further processing. Section 3 describes the Gabor filters [5] in detail and the evolution of the technique based on these filters. In Section 4 the Modulation Transfer Function (MTF) is discussed, it incorporates one of the human psychovisual attributes for thresholding. Section 5 gives a detailed explanation of the implementation based on Gaussian Wavelets. In section 6, we have a discussion on the results, followed by concluding remarks and future research in section 7.

2. SUBJECTIVE CRITERIA AND COMPRESSION ARTIFACTS

Compression artifacts

JPEG compression is sometimes characterized by the presence of artifacts such as blockiness and blurring and an overall presence of noise due to quantization of the DCT coefficients in the image coding step of the process [4]. The blockiness and blurring are not uniformly distributed across the entire image, but they tend to be content based [4]. This is illustrated in fig. 1. Depending on the quality factor i.e., the higher the quality factor, lesser the compression, and vice versa. Aditionally, there is a loss of high frequency components in the image, depending on the compression level. As a result, high frequency content such as edge information is lost resulting in blockiness and blurring [4].

Subjective Criteria

The subjective criteria employed in our technique are.
1). The loss of subjective quality due to the presence of localized artifacts; i.e, the presence of uniform noise over the entire image has a

Fig. 1 Original image (left) and the noticeable artifacts in the compressed image (right)

higher subjective rating than the presence of the same noise in localized regions [6].
2). The inability of a human observer to notice artifacts in the spectral regions of an image below a certain relative spectral frequency. The MTF of the HVS, described later, is used to incorporate this criteria [7].

3. THE GABOR FUNCTIONS/GAUSSIAN WAVELET TRANSFORM

The Gabor functions are complex exponential functions modulated by Gaussian functions [5]. A Gaussian having a zero phase for the complex exponential and tuned to a frequency f_0 and orientation θ_0 and centered at the origin ($x_0 = 0$, $y_0 = 0$) is:

$$g_{0,0,f_0,\theta_0}(x,y) = \exp[-\pi a^2(x^2+y^2)] \times \exp[i2\pi f_0(x\cos\theta_0 + y\sin\theta_0)] \quad \textbf{(1)}$$

Where a determines the spatial frequency bandwidth. In our case, we have set $a = [3(\ln(2/\pi)^{1/2}]^{-1} f_0$, which yields a constant (in a logarithmic scale) 1 octave bandwidth in the frequency domain. The four subscripts account for localization in both spatial and the frequency domain, (0,0) and (f_0, θ_0), respectively. From the basic Gabor function

(Eq.1) the complete set of functions used for sampling the joint space-frequency domain is obtained by rotations, with a 45^0 degree step, to get four orientation channels in the frequency domain; by stretching, by a factor of two, to halve the frequency of the sinusoid; and by translations in x and y (with an interval (1/4) f_0) to sample the spatial domain. The number of spatial frequencies is limited to four starting with $f_4 = f_N/2$ (f_N being the Nyquist frequency), followed by $f_3 = f_4 / 2$ and so on. The number of channels is limited to 4x4 as there is psychophysical evidence that the visual system uses a similar number of channels [8].

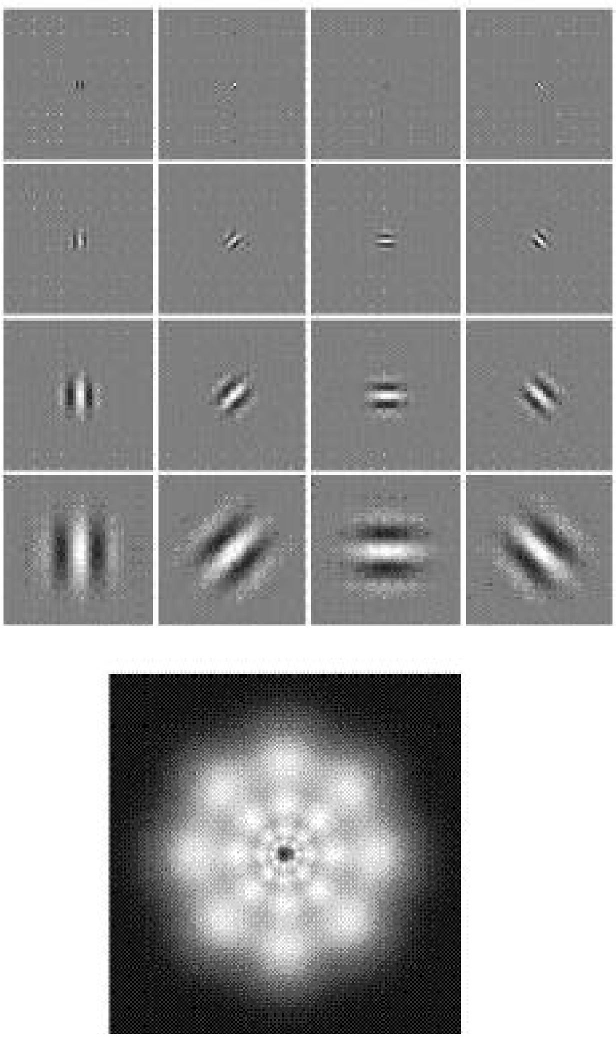

Fig. 2 Real part of the 4x4 Gabor functions: the functions in the spatial domain and their coverage of the respective channels in the frequency domain.

The Gabor coefficients are obtained by convolving the image with the real part of each one of the filters [5]. Fig. 1 shows the even real part of the 4x4 Gabor filters and the coverage of the Fourier domain obtained with the set of Gabor functions. The coverage of the frequency domain is not perfectly flat and the very high and low frequencies are not covered at all. In order to obtain a perfect coverage of the Fourier domain, low pass and high pass residual channels are included. For the low pass residual, a Gabor function with frequency $f_0 = 0$ was used.

$$g_{0,0,0,0}(x,y) = \exp[-\pi a^2(x^2+y^2)] \quad (2)$$

The impulse response of the High Pass Residual obtained as a sum to give a flat response in the spatial domain is:

$$h(x,y) = \delta(x,y) - g_{0,0,0,0}(x,y) - \sum_{f,\theta} \mathrm{Re}[g_{0,0,f,\theta}(x,y)]. \quad (3)$$

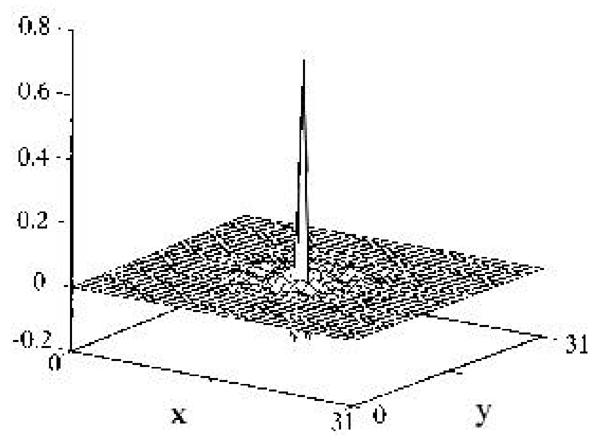

Fig.3 Impulse response of the HPR.

figure 3 shows the impulse response of the High Pass Residual.
The Gabor filters can be efficiently implemented by using separable filter masks. A Gabor function having its principle axes

parallel to the frequency axes can be exactly expressed as the sum of two separable Gabor functions. Tuning the filters to appropriate frequencies as suggested by Nestares, *et al* [5], we can separate the filters into 1D filters for horizontal, vertical and Diagonal orientations. The four one dimensional Gabor masks, each having a size of 11 elements are designed by imposing zero DC response on the 2-D Gabors filters and performing a least squares minimization of the error of the frequency response, using the Moore-Penrose generalized inverse [10]. The 1-D Gabor masks used in our technique along with the mask for the HPR mask, which is a non-separable 2-D mask, are obtained using the technique described above. These filters are used for the directional frequency analysis of images for the detection of blockiness in images.

4. THE MODULATION TRANSFER FUNCTION

The human visual system's modulation transfer function provides a characterization of its frequency response [7]. The modulation transfer function can be thought of as a band pass filter. There have been different classes of experiments used to describe its characteristics; these are described in detail in [11, ch. 13]. For a fixed frequency, a set of stimuli consisting of varying amplitudes is constructed. These stimuli are presented to a human observer and the detection threshold for that frequency is determined. This procedure is repeated for a large number of frequencies and the resulting curve is called the modulation transfer function and is shown in Fig. 6 [7].

This curve is constructed from sine wave gratings kept at one particular orientation, instead of all possible orientations, which would fully characterize the modulation transfer function. This has been accomplished, and the results show that the HVS is not isotropic [7]. However for the purpose of general quality measurement it is close enough to isotropic that that assumption is generally used. It should also be noted that the frequency is in cycles per degree of visual angle [7].

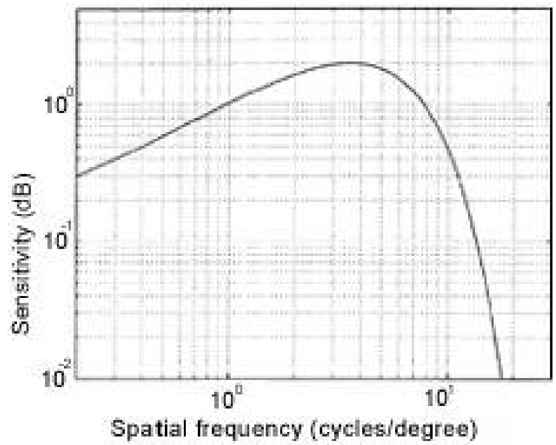

Fig. 3. The MTF curve

5. DETECTION OF ARTIFACTS IN COMPRESSED IMAGES.

As described earlier, the blurring and blockiness in images are the primary factors in the subjective evaluation of JPEG compressed images.

Blur detection.

Fig. 4 shows the technique for the evaluation of blurring artifacts.

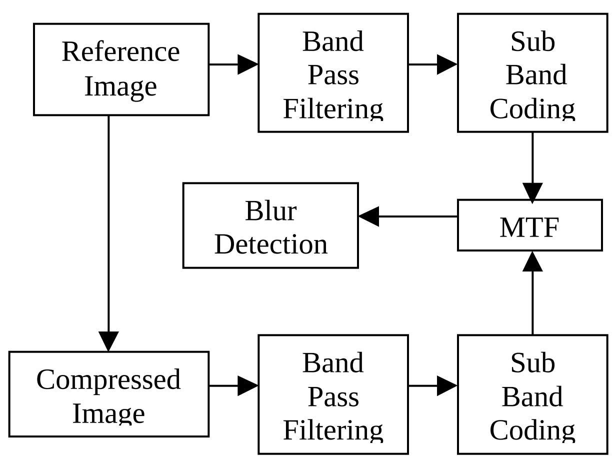

Fig. 4. Block Schematic for Blur Detection.

Blurring occurs as a result of loss of high frequency components in the image. The effect of blurring can be noticed in the bandpass

filtered images. The reference image and the compressed image are both band pass filtered i.e., we perform edge detection on these images. The 3x3 Sobel mask as described in [2] is ideal in order to perform the edge detection. Sub band coding of the images follows the bandpass filtering. The sub band coding is performed using the Gaussian wavelets. The power spectral densities in both the reference as well as the compressed images are compared. The comparison of the power spectral densities for each channel for the different directions is carried out. The individual channels are normalized for comparison. The comparison is based on the Modulation Transfer Function (MTF) described.
The individual channels are weighted according to the MTF and the comparison between the power spectral densities is carried out based on root mean squared error (RMSE) criteria [2]. A greater value of RMSE for a particular sub band indicates localized blurring and hence affects the quality of the compressed image.

Blockiness detection.
The block diagram for the detection of blocking artifacts is shown in fig.5.

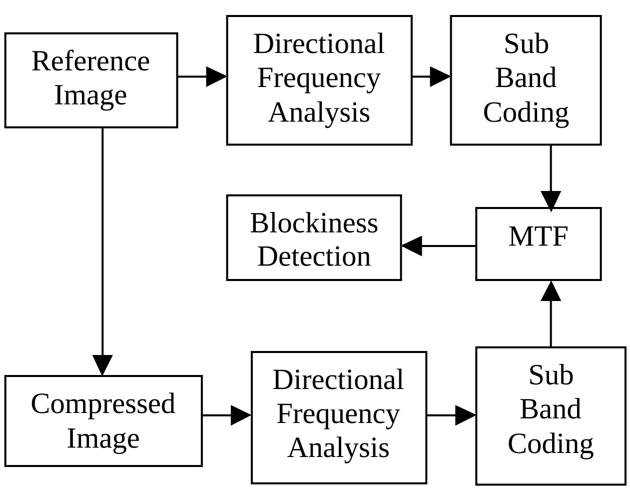

Fig 5. Block Schematic for Blockiness Detection

The procedure followed for detection of blocking artifacts is almost the same as that for blur detection, but in this case we do not carry out edge detection. We carry out the directional frequency analysis using the Gabor transforms and sub band coding. The rest of the procedure remains the same as in the case of blur detection. Depending on the RMSE for both the blur as well as the blockiness, the appropriate weighting of each distortion for the image is decided through trial and error.

6. RESULTS AND DISCUSSIONS

The Lena image was compressed using 3 different quality settings of 10, 30 and 50 in the software provided by the Independent JPEG Group (IJG). The resulting images are shown in figure 6. The weighting for each artifact has an inverse relation with the quality measure. A threshold was set for the RMSE along all the directions of frequency analysis through trial and error.

Fig. 6. The original image (top left) and the compressed images with quality factors of 50 (top right), 30 (bottom left) and 10 (bottom right)

If a particular threshold was exceeded, then the weighting for that particular distortion was given higher priority and due to its inverse relation, it was given a lower quality rating in comparison with the original image. The quality of the comparison was output on a scale from 1 to 10. The output for the 3 cases shown in figure 6 was consistant with the amount of compression after rounding. Currently an adaptive weighting scheme based on the least mean square (LMS) algorithm is being researched.

7. CONCLUSION AND FUTURE WORK

In this paper, we have developed a scheme predicated on research that addresses perceptual quality measurement, both spatial and spectral. The use of wavelets has vastly enhanced the flexibility and the use of both spatial and spectral characteristics of images. This paper has presented a computationally efficient technique in the case where distortions can be characterized in the case of full or partial reference systems. Though, the problem of efficient real time quality monitoring of still images and video sequences still remains, a scheme for adaptively weighting the distortions based on human visual perception would greatly enhance the efficiency and use of this technique.

REFERENCES

1. P.C. Cosman, R.M. Gray, and R.A. Olshen, ``*Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy*," Proceedings of the IEEE, Vol. 82, No. 6, June 1994.
2. S. Wolf and M. Pinson, "*Spatial temporal distortion metrics for in-service quality monitoring of any digital video system*," ITS, NTIA, June 1998.
3. S. Mallat *Wavelets in signal processing*, Addison-Wesley Reading, Mass. (2000).
4. K. Jacoby, " *Characterization of the effects of spatial translations on JPEG-compressed images*" thesis.
5. O. Nestares, R. Navarro, J. Portilla and A. Tabernero, "*Efficient spatial-domain implementation of a multiscale image representation based on Gabor functions*", submitted to Multidimensional Systems and Signal Processing (1998).
6. N. S. Jayant, J. D. Johnston, and R. Safranek. " *Signal compression based on models of human perception*". Proc. IEEE, 81:1385--1422, October 1993.
7. T.N. Pappas and R.N. Safranek. "*Perceptual criteria for image quality evaluation*." Bell laboratories, Mar 1999.
8. J. G. Daugman, "*Spatial visual channels in the Fourier plane,*" Vis. Res., 24,891-910 (1984).
9. R. Navarro and A. Tabernero, "*Gaussian wavelet transform: two alternative fast implementations for images*," Multidimensional System and Signal Processing, 2, 421-436 (1991).
10. R. Gonzalez and P. Wintz, *Digital Image Processing*, Addison-Wesley, Reading, Mass. (1987).
11. T.N.Cornsweet, *Visual perception*, New York: Academic press, 1970