New image compression artifact measure using wavelets

Yung-Kai Lai and C.-C. Jay Kuo

Integrated Media Systems Center
Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, CA 90089-2564

Jin Li

Sharp Labs. of America
5750 NW Pacific Rim Blvd., Camas, WA 98607

ABSTRACT

Traditional objective metrics for the quality measure of coded images such as the mean squared error (MSE) and the peak signal-to-noise ratio (PSNR) do not correlate with the subjective human visual experience well, since they do not take human perception into account. Quantification of artifacts resulted from lossy image compression techniques is studied based on a human visual system (HVS) model and the time-space localization property of the wavelet transform is exploited to simulate HVS in this research. As a result of our research, a new image quality measure by using the wavelet basis function is proposed. This new metric works for a wide variety of compression artifacts. Experimental results are given to demonstrate that it is more consistent with human subjective ranking.

KEYWORDS: image quality assessment, compression artifact measure, visual fidelity criterion, human visual system (HVS), wavelets.

1 INTRODUCTION

The objective of lossy image compression is to store image data efficiently by reducing the redundancy of image content and discarding unimportant information while keeping the quality of the image acceptable. Thus, the tradeoff of lossy image compression is the number of bits required to represent an image and the quality of the compressed image. This is usually known as the rate-distortion tradeoff. The number of bits used to record the compressed image can be measured easily. However, the "closeness" between the compressed and the original images is not a pure objective measure, since human perception also plays an important role in determining the quality of the compressed image. At present, the most widely used objective distortion measure is the mean squared error (MSE) or the related peak signal-noise ratio (PSNR). It is well known that MSE does not correlate very well with the visual quality perceived by human being. This can be easily explained by the fact that MSE is computed by adding the squared difference of individual pixels without considering the visual interaction between adjacent pixels.

In comparison with the amount of research on new compression techniques, work on visual quality measure of compressed images is amazingly little. Most visual quality assessments use rating scales ranging from "excellent" to "unsatisfactory" in measuring the perceived image quality. Human assistance is needed in the establishment of quality scales and the comparison between compressed images and the set of training images. It is difficult to obtain objective and quantitative measures by using this approach [15]. Results achieved are subjective and qualitative.

Some work tries to modify existing quantitative measures to accommodate the factor of human visual perception. One is to improve MSE by putting different weights to neighboring regions with different distances to the focal pixel [17]. Most of them can be viewed as a curve-fitting method to comply with the rating scale method. Since late 1970's, researchers have started to pay attention to the importance of the human visual system (HVS) and tried to include the HVS model in the image quality metric [10], [11]. The development of the HVS model at that time was not mature enough and the proposed model could not interpret human visual perceptual phenomena very well. Recently, Karunasekera proposed an objective distortion measure to evaluate the blocking artifact of the block-based compression technique [12]. Watson [23] and van den Branden Lambrecht [1], [22] proposed more complete models and extended their use to compressed videos.

It has been shown by psychophysic experiments that the human visual system is comprised by many units, each of which is sensitive to the contrast at some space-frequency localized region and functions independently. The overall visual perception of the luminance or contrast of an object is the aggregate performance of the bandpass frequency response of all units [4]. Based upon this space-frequency localized model, the visual perception of coding artifacts can be interpreted as the perception of local frequency inconsistency. When an image is compressed and decompressed, the decoded image has different local frequency responses than the original image, thus resulting in compression artifacts.

Since each cell can sense only a finite region with respect to the focal pixel in an image, Fourier analysis is not suitable for the modeling of HVS due to its global analytical property. This issue has been addressed and inconsistency between theory and observed results by using the pure frequency analysis has been found. Another observation is that the common bandwidth of cortical cells is 1 octave [4]. It is known from experiments that the frequency resolution of the neighborhood is reduced by an octave when the distance from the focal point doubles [6]. This observation suggests the use of dyadic wavelets for the analysis and modeling of HVS.

In this work, we consider a wavelet approach for the modeling of HVS and then build an objective quality metric based on this model. To be more precise, we first propose a new definition of contrast with respect to complex images by taking the wavelet transform of the image, and using wavelet coefficients to estimate the contrast at each resolution in the image. By using the multiresolution and space-frequency localization properties, many known inconsistencies in the psychophysical literature can be understood naturally. Known human visual system functions such as the contrast sensitivity threshold [20] and the frequency masking effect [21] can also be incorporated in this model. We then propose an objective metric to measure the extent of the perceived contrast in every resolution, and derive an error measure by examining the weighted differences of wavelet components between the original and compressed images at each resolution.

This paper is organized as follows. In Section 2, we briefly review existing HVS theory and models. Then, the wavelet approach is introduced to simulate the localization phenomenon of HVS in Section 3. A new HVS model based on wavelets is proposed and an image quality assessment system is described in Section 4. Experimental results are given in Section 5 to demonstrate that this new metric works for a wide range of compression artifacts and that the resulting measure is consistent with human subjective ranking.

2 MODEL FOR HUMAN VISUAL SYSTEM (HVS)

2.1 Contrast thresholds

Human visual systems (HVS) response to light stimuli of surroundings. One important observation is that visual perception is sensitive to luminance difference rather than the absolute luminance. Let L_{max} and L_{min} be the maximum and minimum luminances of the waveform around the point of interest. Michelson's contrast, defined as

$$C = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}, \tag{1}$$

is found to be nearly constant when used to represent the just noticeable luminance difference.

Psychophysic experiments showed that HVS is comprised by many units, each of which is focused on a certain

point in the vision field and only sensitive to the contrast in a certain frequency band. The overall visual perception of object luminance or contrast is the aggregate performance of each cell's frequency response [4]. Since HVS cannot provide infinite luminance resolution, a contrast threshold exists in each frequency band. The contrast threshold value is a function of the spatial frequency which can be determined experimentally. A typical contrast sensitivity curve, defined as the reciprocal of the contrast curve, is shown in Fig. 1 [4]. As shown in the figure, HVS has the highest luminance resolution around 3-10 cycles per degree, and the sensitivity attenuates at both high and low frequency ends. An image artifact can be sensed if its contrast is above the threshold at specific spatial frequencies.

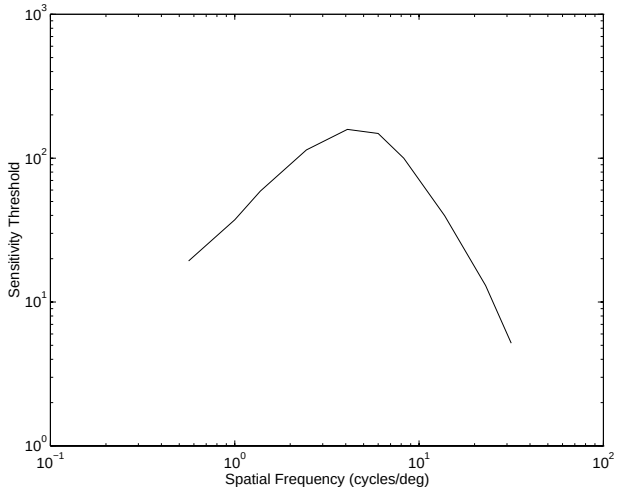


Figure 1: A typical contrast sensitivity curve of human beings.

2.2 Suprathreshold contrast

The subthreshold stimuli, i.e. stimuli with a contrast lower than the threshold, cannot be sensed by human. The suprathreshold contrast is of concern in this subsection. Visual response to the suprathreshold contrast involves human subjective rating. Even though it is difficult to find an precise formula for its modeling, it is generally agreed that the estimated response R is a function of the spatial frequency and follows a power law [14]:

$$R = k(C - C_T)^p \,, \tag{2}$$

where C is the suprathreshold contrast, C_T is the contrast threshold, and the exponent p varies between 0.38 and some value more than 1.0 [7]. Moreover, the contrast constancy phenomenon [9] states that the suprathreshold response is almost constant for a given contrast at a wide range of spatial frequencies. These two seemingly contradicting observations were mediated first by Cannon [26] and then by Georgeson's 3-stage model [8].

2.3 Channel interactions

Although cells narrowly tune to different frequency bands, they are not strictly band-limited, and interactions among adjacent frequency channels were well observed in the literature. Two primary effects were often discussed in the literature. They are the summation effect and the masking effect.

The summation effect is an inter-channel effect saying that the neighboring frequency channels contribute to the total contrast. Therefore, a subthreshold contrast may still produce a small response if there exist other excitory stimuli in nearby frequency channels. This effect can be modeled as a contrast threshold C_T decrease [13]:

$$C_T = \left(\sum_{i=0}^{N} |C_{Ti}|^{s_i} \right)^{\frac{1}{s_i}} ,$$

where C_{Ti} is the contrast sensitivity threshold of the i^{th} closest channel, N is the total neighboring channels affecting

the perception of the target channel, C_{T0} represents the original contrast without summation and s_i's are the exponential components to be determined.

The masking effect is another inter-channel effect which states that the visibility of a stimuli at some frequency could be impaired by the presence of other stimuli in nearby frequency channels. One well-known example in compressed image artifact analysis is that the blocking artifact of block-based compression schemes, which consists of high-frequency edge components, is less visible in the texture regions. This effect can be viewed as a modification of contrast threshold and modeled as an exponential function of both their frequency and contrast differences [1], [21]:

$$C_T = C_{T0}\left(\sum_{i=0}^{N}\left(\frac{C_i}{C_{T0}}\right)^{m_{1,i}}\left(\frac{f_i}{f_0}\right)^{m_{2,i}}\right), \tag{3}$$

where C_{Ti}, C_{T0}, N have the same definition as in the subthreshold summation case, C_i and f_i are the actual contrast and the spatial frequency of the i^{th} closest channel, respectively, and $m_{1,i}$'s and $m_{2,i}$'s are the coefficients to be determined. The validity of Eqn. (3) will be examined in Section 5.2.

2.4 Spatial inhomogeneity

Most research results which led to the contrast threshold curve as shown in Fig. 1 were obtained with foveal vision, i.e. measured very close the fixation point. However, peripheral vision researches showed that the perceived image gradually blurs with increasing distance from the fixation point. In other words, the spatial frequency sensitivity of the vision system decreases as the eccentricity from the focal point increases [6], [13], and it was shown that the equal-sensitivity neighborhood radius is inversely proportional to the spatial frequency, i.e. the product of the spatial frequency and the resolvable radius from the focal point is almost constant [6]. Direct consequences of this inhomogeneity phenomenon are that the resolvable region in the vision field varies with the spatial frequency and that contrasts at different resolutions play different roles in the visual system regarding their respective ranges.

3 SIGNAL LOCALIZATION AND WAVELET REPRESENTATION

3.1 Space-frequency localization

Limited by their spatial location on the retina, the receptors can only focus on their own certain regions of the visual field. The frequency response of the fixation point is characterized by the typical contrast threshold function as shown in Fig. 1, but the high frequency response will further attenuate as the eccentricity from the focal point increases. Therefore, the frequency response of visual stimuli is not only band-limited in the frequency domain but also space-localized in the spatial domain.

The commonly used Fourier frequency analysis is, however, a global process which gives all spatial components the same weighting. It is well known as Heisenberg's uncertainty principle that the localization in both spatial and frequency domains cannot be achieved simultaneously. Gabor transform, which is a Gaussian-windowed Fourier transform, was proved to achieve the limit of the Heisenberg inequality. Gabor gratings were thus widely used in modern psychophysical experiments.

The parameters of the Gaussian window were chosen at researchers' preferences and various degrees of localization were achieved [19]. That is, by varying the Gaussian envelope parameters, the passbands of Gabor gratings were overlapped to a different extent. There are some limitations in the Gabor representation. First, it is difficult to analyze the stimuli whose frequency responses fall in the overlapped band. Second, since the Gabor filter is an IIR (infinite impulse response) filter, truncation is still needed for practical implementation. Thus, the localization is not fully ensured after truncation. To overcome these difficulties, we adopt a wavelet approach for signal analysis in the next section.

3.2 Wavelet approach

The wavelet transform provides a good space-frequency localization property [2] and can be implemented by using the multichannel filter banks. Compactly supported wavelets such as the Daubecheis filters [2] can be implemented with FIR (finite impulse response) filters. The space-frequency localization is optimized among all possible FIR filters with the given length for the Daubechies filters. Another useful property of the wavelet filter is that the frequency response of the contrast threshold can be easily obtained by properly choosing the wavelet basis. Vision researches with Gabor filtering usually used the response amplitude, represented by dBs, to approximate the contrast threshold function defined by (1) [1]. However, the effect of this approximation has not yet been thoroughly investigated.

The Haar wavelet is the simplest basis function in the compactly supported wavelet family. It provides a capability to compute the contrast directly from the responses of the low and high frequency subbands. For the Haar wavelet, the filter coefficients for the low and high frequency filter banks are given by

$$h_0[n] = \begin{cases} \frac{1}{\sqrt{2}} & n = -1, 0, \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$h_1[n] = \begin{cases} \frac{1}{\sqrt{2}} & n = 0, \\ -\frac{1}{\sqrt{2}} & n = -1, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

respectively. Assume that a discrete-time input signal $x[n]$ is the staircase contrast pattern

$$x[n] = \begin{cases} L_{max} & n < 0 \\ L_{min} & n \geq 0, \end{cases} \tag{6}$$

The responses $y_{0,0}[n]$ and $y_{1,0}[n]$ at the 0th resolution after filtering with $h_0[n]$ and $h_1[n]$, respectively, are

$$y_{0,0}[n] = \begin{cases} \sqrt{2}L_{max} & n < -1 \\ \frac{1}{\sqrt{2}}(L_{max} + L_{min}) & n = -1 \\ \sqrt{2}L_{min} & n > 0, \end{cases} \tag{7}$$

$$y_{1,0}[n] = \begin{cases} \frac{1}{\sqrt{2}}(L_{max} - L_{min}) & n = -1 \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Then, the contrast C_0 at the interval $(-1, 0)$ and the 0th (finest) resolution can be computed by the ratio of $y_{1,0}[n]$ and $y_{0,0}[n]$:

$$C_0 = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} = \frac{y_{1,0}[-1]}{y_{0,0}[-1]}.$$

At the 1st (second finest) resolution, the low frequency band response $y_{0,0}[n]$ is downsampled by two, and fed into the same filter bank. The responses are

$$y_{0,1}[n] = \begin{cases} 2L_{max} & n < -1 \\ L_{max} + L_{min} & n = -1 \\ 2L_{min} & n > -1, \end{cases} \tag{9}$$

$$y_{1,1}[n] = \begin{cases} L_{max} - L_{min} & n = -1 \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Again, we can compute the contrast at this resolution as

$$C_1 = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} = \frac{y_{1,1}[-1]}{y_{0,1}[-1]}.$$

Following this way, we can compute the contrast at any resolution by simply dividing the high band response by the low band response at $n = -1$. Fig. 2 illustrates this constant-ratio relationship across resolutions.

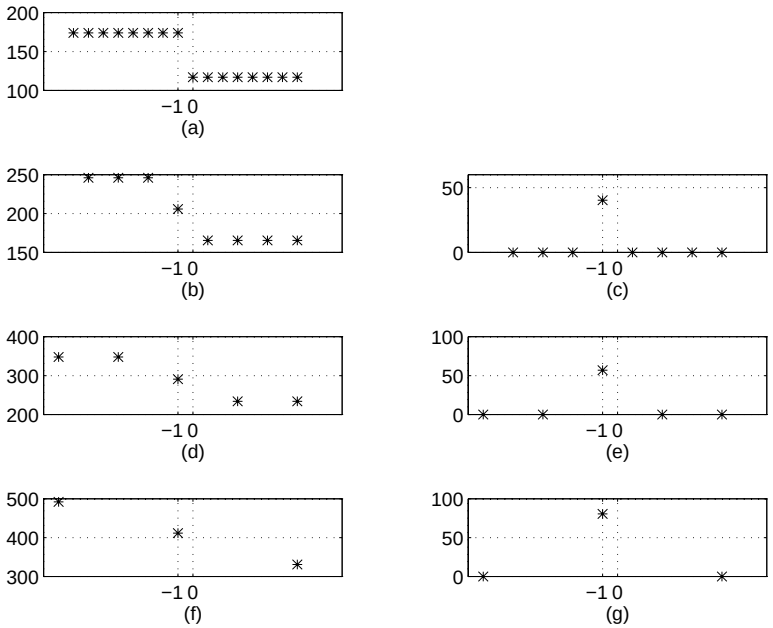


Figure 2: Contrast computation using filter responses, where (a) is the original staircase signal, (b) and (c) are low and high band responses at the 0th resolution, (d), (e), (f) and (g) are responses at the 1st and 2nd resolutions.

There are several reasons to define multiple contrasts in different resolutions. First of all, since human contrast sensitivity is highly dependent on the spatial frequency, multiple contrasts can be used to address different variations at different resolutions across the image [18]. Second, the spatial inhomogeneity phenomenon detailed in Section 2.4 requires the response in different frequency bands to have different supported radii. Furthermore, it was shown that each frequency channel in the HVS has the bandwidth of about 1 octave [18]. The dyadic wavelet transform satisfies these requirements naturally. Finally, perfect reconstruction is possible with responses obtained from different scales, and no visual information will be lost during the process. In contrast, to perfectly reconstruct the visual information using the Gabor analysis, all filters must have the same length and, as a result, the space-frequency localization property is less flexible.

4 PROPOSED QUALITY MEASURE SYSTEM

The proposed quality measure system is shown in Fig. 3. Both the original image and the distorted image are passed through the system for dyadic wavelet decomposition. Contrasts were computed at every resolution of interest. The contrast threshold adjustments due to the masking effect are then made according to Eqn. (3) at each resolution. The summation effect is not implemented in the system at present since its effect is much less significant than the masking effect in the frequency range of natural images.

To model spatial inhomogeneity, one has to find the weights to each frequency band as a function of distance to the fixation point. Since a continuous weighting function is too involved mathematically, discrete stack structure approximations were used in practical implementations when applied to complex images [6]. The staircase model approximates the weighting function as a step function of the distance to the fovea. The resolvable suprathreshold contrast radius obtained from experiments [5] and its staircase approximation are shown in Fig. 4. We see that the staircase function with a constant radius-frequency product fits the experimental data pretty well. The weights of each channel is set to unity, and the visual information out of the radius at a specific spatial frequency, as shown in the dash-dotted staircase line in Fig. 4, is discarded.

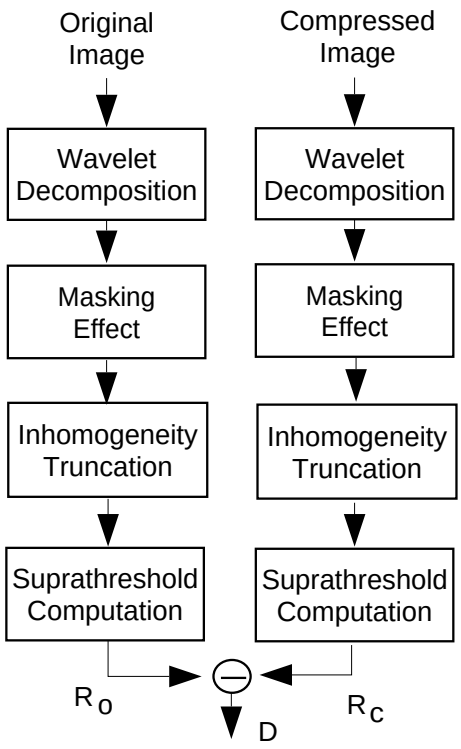


Figure 3: Contrast sensitivity threshold of the Gabor and Haar filters.

Although Georgeson's 3-stage model [8] is more complete, it requires additional work on parameter estimation. Instead, the power law as described in (2) is used with Cannon's result [26] at the last stage. The exponent p varies from 0.45 to 0.5 with respect to the distance from the actual contrast to the threshold. Here a constant value 0.45 is used for p.

Let the subscripts c and o represent the compressed image and the original image, respectively. Then, the quality error measure is

$$D = \frac{1}{HV}\left(\sum_{k=1}^{N}\sum_{j=1}^{V}\sum_{i=1}^{H}\left(R_{c,k}(i,j) - R_{o,k}(i,j)\right)^2\right), \tag{11}$$

where V and H are the vertical and horizontal sizes of the image, respectively, and N is the number of filtering channels. The resulting error measure D is dimensionless since the contrast itself is dimensionless.

It is known that HVS is also orientation-dependent. In other words, the contrast sensitivity threshold as well as perceived contrast vary with respect to the grating orientation [27]. This feature is not implemented in our system since experimental results at different orientations are not available at present, but it can be incorporated conveniently into the system by using the summation of more channels with different orientations.

5 EXPERIMENTAL RESULTS

5.1 Validation of the Haar wavelet

Since the cortical cells have a Gaussian-shaped reception profile [3], it is often used as an evidence that the Gabor filter is preferable in the vision experiment. Since the Haar filter does not possess the same Gaussian-shaped passband as the Gabor filter does, one may suspect the validity of using the Haar filter in vision analysis. To validate

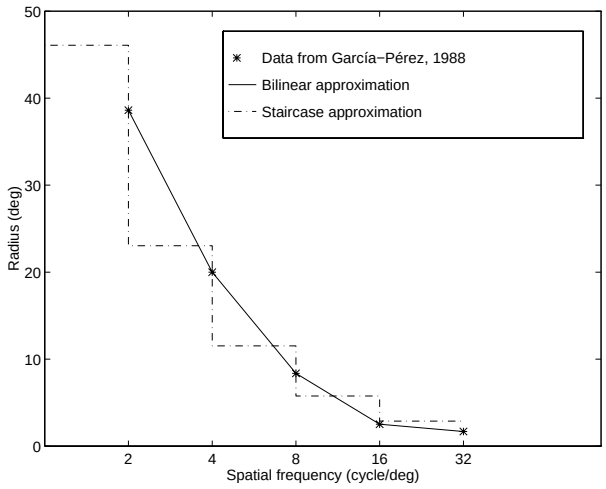


Figure 4: Experimental data from García-Pérez (1988) and the corresponding staircase model approximation.

the use of the Haar wavelet, psychophysical experiments were conducted on a 17" Silicon Graphics color graphic display GDM-17E11. The luminance range of the display was adjusted from 0 to 80 cd/m^2 (candela/squared meter) by using a Photoresearch spectroradiometer. There were 256 discrete gray scales present in the experiments. The relationship of the luminance versus the gray scale is shown in Fig. 5. This curve was used to compute the actual contrast in the following experiments.

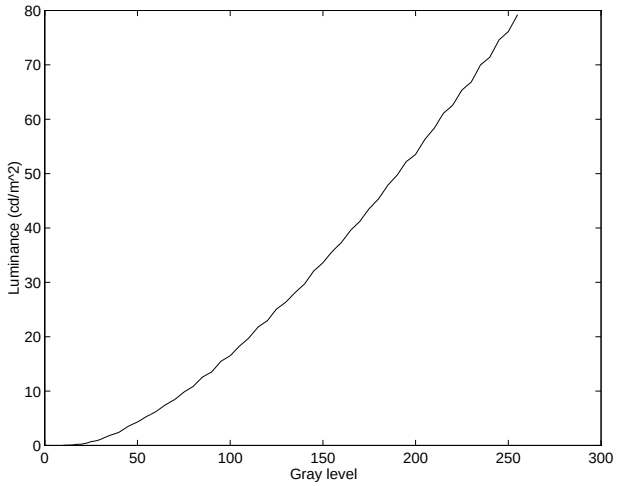


Figure 5: The plot of the luminance versus the gray level for the color graphic display used in experiments.

Both Gabor and Haar filtered patches were used in the experiments. The spatial frequency range of the test patches was from 0.069 to 19.2 cycles per degree, which covers virtually the whole frequency band we would sense from digital images. The result is shown in Fig. 6, where the sensitivity threshold, defined as the reciprocal of the contrast threshold, is plotted as a function of spatial frequency. The closeness of these two curves shows that the Haar filter has a comparable performance in comparison with the Gabor filter. This experimental result clearly demonstrates that the Haar filter can also be used in the quality assessment system.

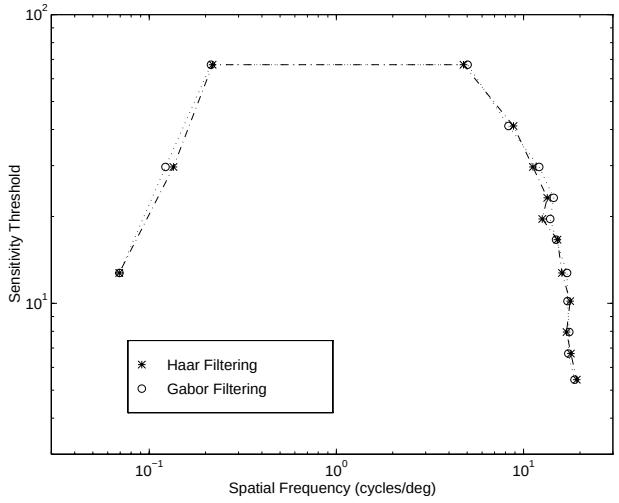


Figure 6: Contrast sensitivity threshold by using the Gabor and Haar filters.

5.2 Suprathreshold masking

The masking effect can be viewed as a function of the ratio of target signal and masking signal frequencies [24] or as a function of the ratio of target signal and masking signal contrasts [25]. We constructed (3) based on the assumption that these two effects are separable. To verify this assumption, psychophysical experiments were conducted to find the parameters of this model.

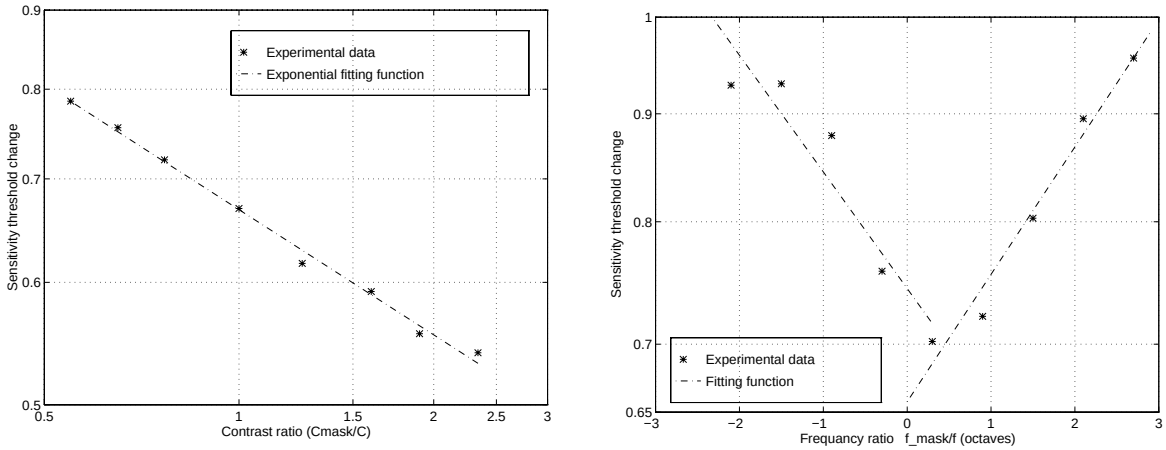
The contrast ratio C_{mask}/C, where C_{mask} and C represent the contrast of masking signal and target signal, respectively, ranged from 0.5 to 2.5. The frequency ratio f_{mask}/f, meanwhile, ranged from -3 to 3 octaves. To isolate individual effects, we first fixed the frequency ratio, and varied the contrast of each signal to investigate the effect of the contrast ratio. The result is shown in Fig. 7 (a), and an exponential fitting function was determined using the data. As shown in Fig. 7 (a), we see that when the contrast of masking signal becomes larger with respect to the target signal, the contrast sensitivity decreases (or equivalently the least detectable contrast increases) [25].

We then varied both the contrast ratio and the frequency ratio of target and masking signals. During the computation process, we scaled experimental data with respect to their contrast ratios according to Fig. 7 (a). The scaled data showed very little deviation, thus indicating that (3) is a very good approximation to the masking model. The means of experimental data and fitting functions are shown in Fig. 7 (b), which are in good agreement with those in [24]. From the fitting functions, the sensitivity threshold changes for frequency ratios of -2, -1, 1, 2, and 3 octaves are o.96, 0.84, 0.76, 0.87, and , respectively.

5.3 Application to compression image artifacts

The new image quality assessment system was constructed according to Fig. 3 with parameters determined by above experiments. Compressed Lena images of size 256×256 were used for image quality assessment. Two types of compression schemes were applied. One is the DCT-based compression with the JPEG default quantization table and a quantization factor of 5, coded at 0.34 bpp with PSNR=26.43. The other is wavelet-compressed Lena by using the embedded zerotree wavelet (EZW) algorithm proposed by Shapiro, coded at 0.32 bpp with PSNR=28.47. These two images are shown in Fig. 8.

Since the HVS contrast sensitivity threshold is characterized by the spatial frequency, defined as cycles per degree, one should expect the quality measure varies with the ratio of D, the distance between the observer and the image, and W, the width of the image. Fig. 9 showed the relation between the ratio and the quality measure. As the distance



(a) (b)

Figure 7: Illustration of the masking effect: sensitivity threshold changes under different (a) contrast and (b) frequency ratios.

between the observer and the image increases, the spatial frequencies of the details (high frequency components) become even higher that the visual system attenuation will fail to perceive the compression artifact. Therefore, the error will approach zero as the viewing distance increases. On the other hand, once the viewing distance is decreased to a certain extent, the whole details of the image to the pixel level are perceivable. The quality measure will be thus almost constant when the viewing distance is smaller than this distance, until the inhomogeneity phenomenon and the HVS attenuation at low frequencies take over at extremely close distances, which will lower the perceived error by a quite small extent. The "best" viewing distance shown in the figure is about 6–10 times the image width, which is consistent with the rule of thumb in practical image viewing situations.

We can also see from Fig. 9 that the EZW-compressed image has a lower quality error measure than that of the JPEG-compressed image, although the compression ratio is smaller. This is consistent with the subjective ranking of human observers.

6 CONCLUSION

In this research, we proposed a new approach to model human visual system (HVS) by wavelet transform. Haar wavelets were proved to provide exact contrast values at each resolution, and the masking effect and spatial inhomogeneity were conveniently incorporated into the HVS model by multiresolution analysis. Experiments showed that Haar filters have comparable performance with the Gabor filters. The resulting new metric was successful in measuring compressed image artifacts.

7 ACKNOWLEDGMENT

This work was supported by Intel, the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, and the National Science Foundation Presidential Faculty Fellow (PFF) Award ASC-9350309.

8 REFERENCES

[1] D. Costantini, C. J. van den Branden Lambrecht, G. L. Sicuranza, and M. Kunt, "Motion rendition quality metric for MPEG coded video," in *Proceedings 1996 IEEE International Conference on Image Processing*, pp. 889–892, Sept. 1996.

[2] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[3] J. G. Daugman, "Two-dimensional spectral analysis of the cortical receptive field profiles," *Vision Research*, Vol. 20,

(a) (b)

Figure 8: Two test images: (a) JPEG-compressed Lena and (b) EZW-compressed Lena.

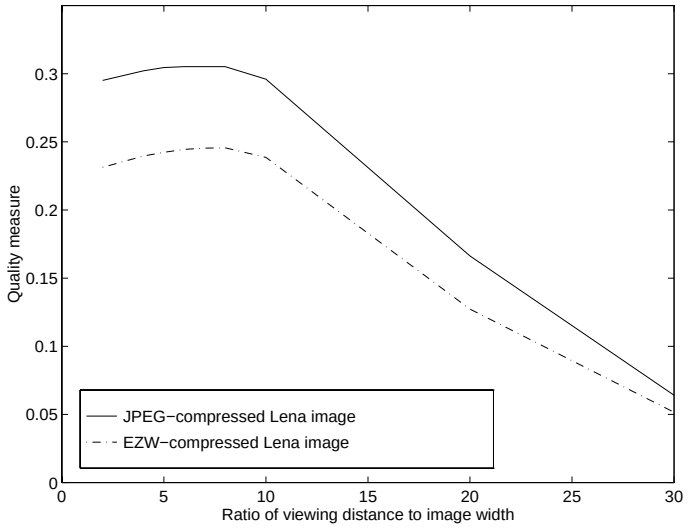


Figure 9: Quality measure as a function of the viewing distance to the image width ratio.

pp. 847–856, Oct. 1980.

[4] R. L. De Valois and K. K. De Valois, *Spatial Vision*, Oxford University Press, 1988.

[5] M. A. García-Pérez, "Space-variant visual processing: spatially limited visual channels," *Spatial Vision*, Vol. 3, No. 2, pp. 129–142, 1988.

[6] M. A. García-Pérez, "The perceived image: Efficient modelling of visual inhomogeneity," *Spatial Vision*, Vol. 6, No. 2, pp. 89–99, 1992.

[7] M. Georgeson, "Over the limit: encoding contrast above threshold in human vision," in *Vision and visual dysfunction*, vol. 5, ch. 9, pp. 106–119, MacMillian, 1991.

[8] M. A. Georgeson, "Contrast overconstancy," *J. Opt. Soc. Am. A*, Vol. 8, pp. 579–586, Mar. 1991.

[9] M. A. Georgeson and G. D. Sullivan, "Contrast constancy: deblurring in human vision by spatial frequency channels," *J. Physiol.*, Vol. 252, pp. 627–656, 1975.

[10] D. J. Granrath, "The role of human visual models in image processing," *Proceedings of the IEEE*, Vol. 69, pp. 552–561, May 1981.

[11] C. F. Hall and E. L. Hall, "A nonlinear model for the spatial characteristics of the human visual system," *IEEE Trans. on SMC*, Vol. 7, pp. 161–170, Mar. 1977.

[12] S. A. Karunasekera and N. G. Kingsbury, "A distortion measure for blocking artifacts in images based on human visual sensitivity," *IEEE Trans. on Image Processing*, Vol. 4, pp. 713–724, June 1995.

[13] D. H. Kelly, "Retinal inhomogeneity: I. Spatiotemporal contrast sensitivity," *J. Opt. Soc. Am. A*, Vol. 1, pp. 107–113, Jan. 1984.

[14] G. E. Legge, "A power law for contrast discrimination," *Vision Research*, Vol. 21, pp. 457–467, 1981.

[15] F. X. J. Lukas and Z. L. Budrikis, "Picture quality prediction based on a visual model," *IEEE Trans. on Communications*, Vol. 30, pp. 1679–1692, July 1982.

[16] J. M. W. Cannon, "Perceived contrast in the fovea and periphery," *J. Opt. Soc. Am. A*, Vol. 2, pp. 1760–1768, Oct. 1985.

[17] H. Marmolin, "Subjective MSE measures," *IEEE Trans. on SMC*, Vol. 16, pp. 486–489, June 1986.

[18] E. Peli, "Contrast in complex images," *J. Opt. Soc. Am. A*, Vol. 7, pp. 2032–2040, Oct. 1990.

[19] E. Peli, L. E. Arend, G. M. Young, and R. B. Goldstein, "Contrast sensitivity to patch stimuli: Effects of spatial bandwidth and temporal presentation," *Spatial Vision*, Vol. 7, No. 1, pp. 1–14, 1993.

[20] P. C. Quinn, "Suprathreshold contrast perception as a function as a function of spatial frequency," *Perception & Psychophysics*, Vol. 38, No. 5, pp. 408–414, 1985.

[21] J. P. Thomas, "Independent processing of suprathreshold spatial gratings as a function of their separation in spatial frequency," *J. Opt. Soc. Am. A*, Vol. 6, pp. 1102–1111, July 1989.

[22] C. J. van den Branden Lambrecht, "A working spatio-temporal model of the human visual system for image restoration and quality assessment applications," in *Proceedings 1996 International Conference on Acoustics, Speech, and Signal Processing*, pp. 2293–2296, May 1996.

[23] A. B. Watson, "Perceptual-component architecture for digital video," *J. Opt. Soc. Am. A*, Vol. 7, pp. 1943–1954, June 1990.

[24] K. K. De Valois and E. Switkes, "Simultaneous masking interactions between chromatic and luminance gratings," *J. Opt. Soc. Am.*, Vol. 73, No. 1, pp. 11–18, January 1983.

[25] J. M. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," *J. Opt. Soc. Am. A*, Vol. 11, No. 6, pp. 1710–1719, June 1994.

[26] M. W. Cannon, Jr., "Perceived contrast in the fovea and periphery," *J. Opt. Soc. Am. A*, Vol. 2, No. 10, pp. 1760–1768, October 1985.

[27] R. St. John, B. Timney, K. E. Armstrong, and A. B. Szpak, "Changes in perceived contrast of suprathreshold gratings as a function of orientation and spatial frequency," *Spatial Vision*, Vol. 2, No. 3, pp. 223–232, 1987.