Prediction of Print Defect Perception

Kevin D. Donohue[†], M.Vijay Venkatesh[†], and Chengwu Cui[‡]
[†]University of Kentucky, Lexington, Kentucky
[‡]Lexmark International Inc., Lexington, Kentucky

Abstract

This study examines the prediction of print defect perception (banding) of the human visual system (HVS) by combining detection probabilities of contrast components from wavelet and sinusoidal decompositions of defect patterns. Detection probabilities for basis function components are obtained from subjective tests. Prediction performance is presented in terms of the deviation of the contrast thresholds at the 92% detection probability obtained from subjective tests on the simulated print defect patterns. Results from tests run on 21 subjects indicate that prediction errors can be obtained on the order of the subjective test variability used to obtain detection probabilities. While the wavelet approach resulted in less systematic error than the sinusoidal approach, both performed about the same once the systematic error was removed. An advantage of the wavelet approach includes having a framework more conducive for modeling independent visual channels of the HVS and for obtaining efficient orthogonal decomposition of the defect patterns.

Introduction

Characterizations of printer defect patterns find applications in developing standards and metrics for printer quality analysis.[1-5] Various forms of defects, such as banding, graininess, and streaking occur in most printers. Those involved with developing printing and imaging equipment have considered standard definitions and characterizations for these defects with objective measurements.[6] The major limitation of many of these metrics is that they do not include sensitivities of the human visual system (HVS) to reflect human quality judgment. A method for predicting the impact of a particular defect pattern on the human observer can be a useful tool in the development of printing devices, in that the number of costly subjective tests can be reduced at intermediate levels of the product design and evaluation. This paper examines the feasibility of predicting the visibility of banding defect patterns, and compares a wavelet-based and sinusoidal-based approach.

The multi-resolution wavelet analysis[5,7] considered in this work decomposes an image into orthogonal spatial frequency octave bands at 45 degree angular orientation increments. The multi-resolution analysis with wavelets provides a more complete framework than Fourier analysis for characterizing signals with localized space-frequency properties including commonly occurring print defects such as banding. In addition, the scale-based analysis of the wavelets closely resembles the properties of the visual channels observed in the HVS,[8] which simplifies the development of thresholds and pooling of contrast detection probabilities.

Symlet patterns were shown to be robust in detecting and characterizing various print defect patterns in random backgrounds,[5] and therefore the symlet was the particular wavelet used in the subjective tests of this study. The symlet pattern efficiently characterized a variety of defect patterns with fewer coefficients than other types of wavelets, such as some of the biorthogonal and Daubechies wavelets. Defect patterns characterized with more energy in fewer coefficients (more compactly represented), imply that fewer components need to be combined to compute an overall detection probability. Orthogonal representations of defect patterns with fewer coefficients reduce the variability introduced by the mismatch between the combining process of the actual HVS and the vision model used by the algorithm. This can be a significant advantage over sinusoidal decompositions, in which the energy in the defect patterns is distributed over many coefficients, especially for patterns that have small spatial support (spatially localized).

The experiments in this paper were designed to test the feasibility of doing this prediction with a reasonable level of accuracy, and to compare wavelet and sinusoidal approaches. Therefore, three separate subjective tests were performed. One test presented the subjects with wavelet basis function patterns at various scales, another test presented sinusoidal basis function patterns at various frequencies, and another test presented simulated banding defect patterns at various frequencies and sizes. A 2 alternative forced choice (2AFC) test was used with a Bayesian adaptive psychometric procedure, referred to as QUEST,[9] to determine contrast thresholds and detection probabilities. A contrast value of the decomposed defect pattern was mapped to a detection probability through the psychometric function of the corresponding basis functions. The resulting set of probabilities was combined to obtain a detection threshold. The predicted results were compared to that of subjective test results on the actual defect pattern. The following sections describe details of the subjective testing procedure, present results, and discuss their implications.

Subjective Testing Procedure

Subjects were recruited through announcements posted at the University of Kentucky and Lexmark International Inc. Results from 21 subjects taking 25 subjective tests were used in the experiment (some subjects took more than one test). Eight subjects were tested with wavelet patterns, 8 subjects were tested with sinusoidal patterns, and 9 subjects were tested with defect patterns. All patterns were presented in both flat and random noise backgrounds. The random background was generated as white Gaussian noise at a -23 dB contrast level. Subjects took the tests in an office with no windows and florescent light illumination. Patterns were displayed as grayscale images on a LaCIE 20" (371 mm x 297 mm) monitor with a resolution of 3.45 pixels/mm. Subjects sat at a desk and viewed the images from a distance of 57 cm. A fixed image size of 384x384 pixels yielded a display visual resolution of 34.75 pixels/degree. The stimuli were positioned at the center of the image with no additional fixation cues. The background of the monitor (outside the 384x384 pixel range) was set to zero. The edges of the patterns therefore served as accommodation cues.

The gamma of the display was set to 2.7 to provide the smallest possible contrast level variations at the intended luminance level for integer increments of the graphic buffer. The brightness and the contrast values of the CRT were manually adjusted at the beginning of the testing so that the low-contrast stimuli were not visible for several nonzero values of the graphic buffer. At these settings the display exhibited an approximate linear characteristic over the entire dynamic range of the display values. The monitor settings were kept constant throughout all subjective tests; however in order to ensure consistency, the luminance of the display was measured at 17 uniformly-spaced gray level increments from 0-255 using a Minolta Chromameter (CS-100A) before each subjective test. The luminance (Y value) was recorded with the results of each subjective test to convert the graphic buffer integers to contrast values of the stimulus for a standard device independent representation. The background luminance for the 384x384 pixel field was set to a gray level value of 128 (or mean gray level for the random field background) throughout all the subjective testing procedures. This corresponded to a luminance level of 20 cd/m^2 (even though this value was independently measured for each test, very little variability existed from measurement to measurement, therefore this number well-represents the value for all tests).

The pattern contrast was varied in a 2AFC procedure according to an adaptive Bayesian psychometric testing procedure, QUEST,[9] that sequentially estimated the 92% detection threshold. Each trial of the test presented a sequence of image pairs; one image contained either a uniform gray level, or random noise field, and the other contained a scaled stimulus added to the same background. The order of the image pair presentation was random for each trial, and each image was presented for 1 second with a 0.5 second pause (blank screen) between images. The subject used a mouse click to indicate the image in which the pattern was present. The contrast of the stimulus was varied adaptively based on the decision of the observer from trial to trial. Exactly 16 trials were used for each stimulus. The final trial was taken as the best estimate of the threshold.

Stimuli Description

The stimuli for the wavelet basis function tests consisted of 2-dimensional symlets generated for levels 2 through 4 from a 4-level decomposition oriented along horizontal, vertical and diagonal direction making a total of nine stimuli (level 1 represents the smallest scale or highest frequency). The wavelet stimulus from the first level is not used because the stimuli from the last three levels cover approximately 0–17 cycles per degree (cpd) of spatial frequency range, so the contribution from the first level was considered negligible. In addition, it allowed for fewer wavelet patterns to be presented in the subjective tests. To ensure energy was not present at the first level of the image, the defect patterns were all filtered to zero out energy at this level. Examples of a level 4 wavelet pattern (largest scale corresponding to low frequency) with a horizontal orientation both flat and random backgrounds are shown in Figs. 1a and 1b.

(a)   (b)
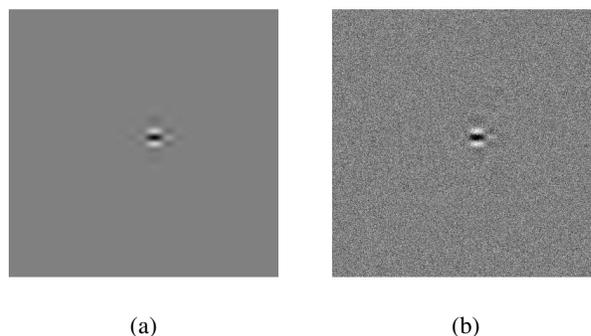
Figure 1. A symlet basis function with horizontal orientation at level 4 in (a) flat field (b) random field.

The stimuli for the sinusoidal basis functions were added to a constant luminance background and oriented along horizontal directions (vertical orientation was not used in the experiment with banding defects) with frequencies ranging from 1.43 to 9.34 cpd (cycles per degree). The frequencies used for sinusoidal patterns for the subjective testing were 1.43, 2.87, 4.31, 5.75, 7.19, 8.62, and 9.34 cpd. To reduce the effect of abrupt changes along the edges of the 384x384 image field, a Gaussian window with standard deviation equal to one fourth of the image size (384 by 384 pixels) weighted the sinusoidal pattern. Figures 2a and 2b show examples of a sinusoidal pattern in the flat and random field backgrounds.

(a)   (b)
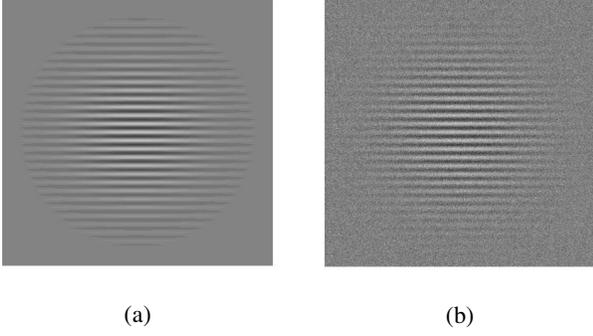
Figure 2. A sinusoid basis function with horizontal orientation at 2.9 cpd in (a) flat field and (b) random field.

(a)   (b)

Figure 3. Banding pattern used in study. (a) Measured banding profile (b) Simulated banding artifact.

The banding print defect was generated from a scanned banding profile presented in Cui et.al[4] as shown in Figure 3a. Interpolation was used to stretch or compress this pattern to simulate banding at different frequencies. To simulate the effect of local banding artifacts, two different sizes of the banding pattern over regions of 2.76 and 1.84 degrees were generated, and they were weighted by a Gaussian window whose standard deviations were correspondingly set to .69 and .46 degrees. These patterns were placed at the center of image of size 384x384 pixels (about 11 degrees) as shown in Figure 3b. Three different frequencies of the banding pattern were generated for each size. The size and frequency of the banding patterns were chosen such that almost all banding pattern spectral energy was distributed over the last 3 levels of a 4-level wavelet decomposition. Zeroing out the coefficients in 4th level ensures this would be the case for the presented patterns. This created minor distort of the high frequency defect patterns by filtering out higher harmonics, and making the pattern more sinusoidal than the profile shown in Fig. 3a. The defect frequencies for subjective testing were 2.87, 5.75, and 8.62 cpd.

The contrast values for all stimuli were recorded for each response in decibels, given as:

$$c = 20\log_{10}\left(\frac{\max(L(x,y)) - \min(L(x,y))}{L_{bg}}\right),\qquad(1)$$

where $L(i,y)$ is the luminance of the pixel values over the x-y plane and L_{bg} is the mean background luminance. The contrast was computed globally over the whole image and recorded with the subjects' response. The contrast of Eq. (1) represents the maximum contrast value of the image. In computing contrast from the basis function coefficients, an inconsistency in the scale of contrast values may occur, since the basis functions characterize local pixels of the image. Therefore, systemic errors may contribute to an overall bias. For this reason, the prediction performance with and without systematic error is considered in the discussion section of this paper.
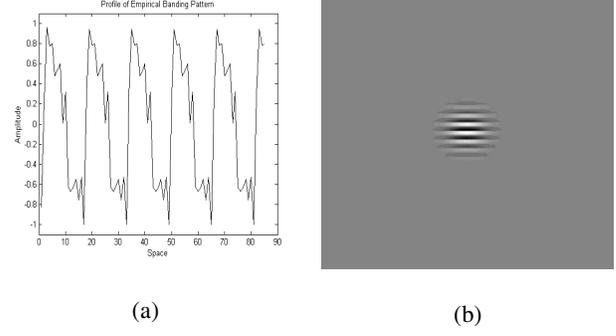
Prediction of HVS Response

The threshold estimates from the subjective test were used in a Weibull psychometric function[9] so that probabilities could be obtained for other contrast values and combined together in the prediction model. For the ith basis function, the 92% threshold, T_i, estimated from the QUEST method can be substitute into the associated psychometric function:

$$P_i(c;T_i) = 1 - .5*\exp(-10^{\frac{\beta}{20}(c-T)}),\qquad(2)$$

where c is the contrast value of the basis function coefficient, and β is the variance of the distribution (3.5 was used[9]). All parameter values in this expression are in dB. The psychometric function represents the response to the 2AFC tests, and therefore has a 0.5 probability when the pattern is not detectable. In order to combine these probabilities from contrast values of the decomposed image components, low contrast value corresponding to probabilities under 0.6 were scaled down to zero in a linear fashion. Therefore, the shape of the psychometric function of Eq. (2) was preserved for values greater than 0.6, and linearly distorted down to 0 for values less than 0.6.

The thresholds for the defect patterns were computed for each pattern test via a bootstrapping method to limit the effect of outliers and obtain a variability measure of the estimate. For each bootstrap sample a set of subjects were drawn 64 times at random with replacement. For each drawing the median of the thresholds was taken as the estimate. Then from all 64 estimates, the median, upper quartile and lower quartile values were taken from the set. The median was considered as the final estimate used in the psychometric function. The inter-quartile range was used as a measure of the variability. Bootstrapping procedure was done for each of the basis function (drawing 8 at a time with replacement) and defect pattern test (drawing 9 at a time with replacement).

The prediction performance of contrast thresholds was obtained by:

- Simulating defects patterns at contrast levels used in the defect subjective tests.
- Decomposing images into basis function coefficients according to either the wavelet or the sinusoidal method.
- Computing a contrast value from the resulting basis function coefficients.
- Applying the psychometric function in Eq. 2 to obtain probability values.
- Combining probabilities together for the predicted result.

The above process was repeated for different contrast values until the contrast level yielding the 92% probability was determined from the pooled probability values.

For the wavelet approach, a 4-level decomposition was performed on the defect pattern, given by recursive formula:

$$y_{uv}^{(l+1)}(n,m) = \sum_i \sum_j y_{11}^{(l)}(2n-i,2m-j)K_{uv}(i,j), \qquad (3)$$

where the superscript (l) is the level (level 0 is the original signal) and the wavelet kernel, K_{uv}, is composed of a 8th order symlet[5,7]. The subscripts on y indicate the orientation of its wavelet kernel as described in Donohue et. al.[5] The levels represent octave subbands, since subsampling scales down the frequency axis of the wavelet kernel by a factor of 2, thereby reducing its effective cutoff frequencies without changing the kernel coefficients. At the first level of the transform, the band covers the spectral range extending from one half of the Nyquist rate to Nyquist rate. This is equivalent to half of the display pixel rate. At the next level the band is lowered by a factor of two and so on.

The decomposed wavelet image coefficients are converted to contrast values. The mean background is obtained from the y_{11} coefficients at the fourth level, which approximate a local DC component for a 16 by 16 pixel area. The local pixel differences correspond to the y_{12}, y_{21}, and y_{22} coefficients at all lower levels, and these are used as the numerator in a contrast ratio. After wavelet coefficients from $y_{11}^{(4)}$ are spatially aligned with high-pass coefficients at the forth level and lower, the contrast values for the wavelet basis coefficients can be computed from:

$$c_{ij}^{(l)}(n,m) = \frac{2 y_{ij}^{(l)}(n,m)}{\bar{y}_{11}^{(4)}} \quad , \qquad (4)$$

where $\bar{y}_{11}^{(4)}$ are the coefficients from $y_{11}^{(4)}$ expanded to level l through upsampling and interpolation[7] (spatial alignment). Each level and orientation of the wavelet decomposition represents one of the basis function for which a psychometric function was computed. Therefore, the computed contrast ratio is substituted into the psychometric function to obtain probabilities for all 9 wavelet subbands and all spatial coefficients in each subband. These are pooled together to obtain the predicted detection probability.

For the sinusoidal decomposition the vertical line of pixels through the center of the image was transformed to its sinusoidal components through a one-dimensional discrete Fourier transform (DFT).[6] The contrast ratios were obtained by using the DC value to scale down each DFT component. Then the magnitudes of the components in octave bands were summed, and the results were substituted into the psychometric function for the sinusoidal basis. This energy contrast value is given by:

$$c_s(b) = \frac{2}{S(0)} \sum_{k=2^b f_o}^{2^{b+1} f_o} |S(k)|, \qquad (5)$$

where $S(k)$ are the DFT coefficients with index k ranging over octave bands starting with index of reference frequency f_o, and b is an integer ranging from 0 to 3. The sinusoidal basis function corresponding to the center of the octave band was used to obtain the psychometric probabilities. Four octave bands were computed that correspond to a similar frequency range as the wavelet subbands, except that it included an additional low frequency band. To obtain psychometric functions at sinusoidal components not tested, linear interpolation between the psychometric functions was used. A more direct method for computing the detection probability involves computing each DFT component detection probability by applying each $S(k)$ to its corresponding psychometric function, and pooling those values rather than the integrated results of Eq. (5); however prediction performance with this method was very poor with close to 100% under-prediction of HVS detection. Thus, the method described here was determined by assuming that the visual channels were octave based, and the detection of the pattern resulted from the contributions of all contrast energy in that channel. This modification resulted in good performance for the sinusoidal based method, and therefore was used in this paper for the comparison.

The probability pooling was done assuming that each basis function contrast value (from Eqs. (4) or (5)) represented independent visual channels. Therefore, the final probability of detection at contrast level c was computed as the complement of the probability of not detecting a basis function component in any of the channels, which can be denoted by:

$$P_d(c) = 1 - \prod_{n=1}^{N} (1 - \lambda_n) \qquad (6)$$

where N is number of channels (9 for wavelet and 4 for the sinusoidal), and λ_n is the channel probability corresponding to each basis function. The channel contrast for the wavelet exploited the spatial orthogonality of the coefficients at each level and combined the independent probabilities over spatial contrast values:

$$\lambda_n = 1 - \prod_{k=1}^{K} (1 - P_n(c_k)) \qquad (7)$$

where c_k is the contrast values for the spatial wavelet coefficients corresponding to Eq. (4), n is the particular channel (a specific i, j, and l combination), and K is the total number of (spatial) wavelet coefficients at that level and orientation. Since the sinusoidal approach does not include spatial components the channel probability is simply given by:

$$\lambda_n = P_n(c_k) \qquad (8)$$

where c_k is the contrast value from Eq. (5), and the psychometric function corresponds to the sinusoid at the center of the band.

Results

Tabe 1 presents results of the subjective test on the defect banding patterns, which include both defect sizes on the flat field background (luminance level 20 cd/m^2) and random background (mean luminance 20 and -23 dB) of white noise.

Table 2 presents the predicted values using the wavelet bases along with the percent error relative to the thresholds in Table 1. With the exception of the flat field low frequency defect, all errors are between -8 and 10 percent (corresponding to a range between –4 dB and 3 dB error). A systematic error of –1.33 dB exists between all the predicted and subjective thresholds.

The predicted values using the sinusoidal bases are shown in Table 3 along with the percent error relative to the thresholds in Table 1. All errors are between -11 and -35 percent (corresponding to a range between 3 dB and 11 dB error). A systematic error of 6.71 dB exists between all the predicted and subjective thresholds. If this error is taken out, the then all errors range between -14 and 9 percent.

Discussion and Conclusion

The error for both wavelet and sinusoidal cases is comparable to the inter quartile ranges on the subjective defect tests (up to 12%) and the basis function tests (up to 11%). Therefore, it is concluded that the prediction performance was reasonable. Modifications had to be made to the sinusoidal approach in order to improve its performance. The summing of contrast energy in the octave bands suggests that a Gabor patch should be used in future comparisons. The large systematic error for the sinusoidal case likely resulted from using the Gabor-like contrast values to obtain detection probabilities from psychometric functions developed for sinusoidal basis functions. Once the bias error was removed, the prediction performance was similar to that of the wavelet approach.

Table 1. 92% detection threshold in dB for subjective defect tests.

	Low Frequency	Mid Frequency	High Frequency
Large area on flat field	-35.7	-37.5	-31.5
Small area on flat field	-34.0	-34.0	-30.5
Large area on random field	-34.0	-33.0	-31.5
Small area on random field	-32.7	-33.0	-28.0

Table 2. 92% Detection threshold in dB for predicted HVS response using wavelet bases with percent error relative to subjective test in parenthesis.

	Low Frequency	Mid Frequency	High Frequency
Large area on flat field	-40.1 (12)	-35.5 (–5)	-34.8 (10)
Small area on flat field	-40.0 (18)	-33.5 (-1)	-32.9 (8)
Large area on random field	-35.5 (4)	-32.5 (-2)	-31.3 (-1)
Small area on random field	-35.0 (7)	-30.4 (-8)	-29.9 (7)

Table 3. 92% detection threshold in dB for predicted HVS response using sinusoidal bases with percent error relative to subjective test in parenthesis.

	Low Frequency	Mid Frequency	High Frequency
Large area on flat field	-30.2 (-15)	-30.1 (-20)	-24.4 (-23)
Small area on flat field	-30.1 (-11)	-30.0 (-12)	-24.9 (-18)
Large area on random field	-26.5 (-22)	-25.6 (-22)	-20.6 (-35)
Small area on random field	-26.2 (-20)	-25.5 (-23)	-20.8 (-26)

The large errors for the low-frequency flat-field results were likely due the limited number for gray levels below the thresholds. The low frequency patterns for the wavelet bases were the easiest to detect, and it was difficult to set up the experiment so that most people had several gray levels below which this pattern could not be detected. Some subjects were even able to correctly identify this pattern for all 16 trials. This resulted in significant quantization error for the low frequency wavelet patterns. This notion is supported by the observation that when white noise was added to the test patterns the prediction error reduced (compare first column of Table 2). The white

noise had the effect of making the patterns harder to detect (reduced HVS sensitivity) and as a result more gray levels existed below the detection threshold for all subjects, especially for the lower frequency patterns.

Overall both wavelet and sinusoidal based methods worked reasonable well, which suggests the feasibility of developing measures and methods for print defect detection that reflect the response of the HVS. In addition, the performance prediction analysis done in these experiments used mostly different populations. There were 1 or 2 subjects common to all three tests. This suggests that the information obtained from one population on the basis functions was applicable for predicting the response for the larger population. It was not simply the case of learning the HVS capability of population and predicting the response of the same population.

Order statistics (medians and quartiles) were used to estimate thresholds and assess errors, instead of means and variances. The censoring property of the median was important to help reduce the effects of outliers. As the population set grows the effects of outliers should become less and either method (mean or median) should yield similar results. In the results presented here, the order statistic approach had a significant impact on the results because of the small data set.

The orthogonal wavelets have distinct advantages in developing models for probability pooling based on independence of the visual channels from the HVS model. In addition, orthogonality ensures the pattern energy was not over counted or under counted in each wavelet frequency or spatial band (this is not possible with the Gabor patch). The systematic error for the wavelet approach was within the inter-quartile distance of the error. This suggests the model for pooling probabilities and scaling of contrast values followed that of the HVS, or at least it was better than the sinusoidal approach with no spatial component.

Some improvements may be achieved by directly computing detection probabilities over a range of contrast values between the 1% and 99% detection levels. The psychometric functions used to obtain intermediate probabilities from the QUEST procedure may not be very accurate for values far from the 92% detection threshold. However, the wavelet approach may have been relatively robust to this error because of its ability to capture the defect pattern energy in a few coefficients. In future experiments, better prediction and more accurate assessment of pooling methods may be obtained from non-parametric distribution functions for HVS detection probabilities. This may further improve performance and be more consistent with models for combining channel probabilities.

Finally, more interesting and insightful comparisons should be made between methods using Gabor patches and wavelets. Both methods are similar in that they can reflect spatial and frequency localization. As a matter of fact the Gabor patch is a wavelet; however it cannot operate as a tight set of orthogonal filters as can be done with the symlet. Future comparisons between these basis functions should allow for a more direct comparison of methods, and should indicate the advantage of orthogonality in predicting HVS response along with the impact on the shape of the basis function.

References

1. John C. Briggs, Mike Murphy and Yichuan Pan, "Banding Characterization for Inkjet Printing", IS&T, Portland, OR, 2000, pg. 84-88.
2. Paul J. Kane, Theodore F. Bouk, Peter D. Burns and Andrew D.Thompson, "Quantification of Banding, Streaking and Grain in Flat Field Images", IS&T, Portland, OR, 2000, pg. 79-83.
3. Howard Mizes, Nancy Goodman and Paul Butterfield, "The Perceptibility of Random Streaking", IS&T, Portland, OR, 2000, pg. 89-93.
4. Chengwu Cui, Dingcai Cao and Shaun Love, "Measuring Visual Threshold of Inkjet Banding", IS&T, Montreal, Canada, 2001, pg. 84-89.
5. K.D.Donohue, C. Cui, and M.Vijay Venkatesh, "Wavelet Analysis of Print Defects", IS&T, Portland, OR,, 2002, pg. 42-47.
6. International Standard ISO/IEC 13660:2001
7. Gilbert Strang and Troung Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA, 1996, pg 144.
8. Scott Daly, "Visual differences Predictor: An Algorithm for the assessment of Image Fidelity", Digital Images and Human Vision, MIT Press, Cambridge, MA, 1993.
9. A. B. Watson and D. G. Pelli, "QUEST: A Bayesian adaptive psychometric method," Perception and Psychophysics, vol. 33, pp. 113-120, 1983.

Biography

Kevin Donohue received the B.A. degree in mathematics from Northeastern Illinois University in 1981 and the B.S., M.S., and Ph.D. degrees in electrical engineering from Illinois Institute of Technology in 1984, 1985, and 1987, respectively. He is currently an Associate Professor in the Electrical and Computer Engineering Department at the University of Kentucky. His current research interests include statistical signal processing with applications in image, audio, and ultrasonic signal processing. He is a member of IEEE, Sigma Xi and IS&T.