

Absolute Graininess Thresholds and Linear Probability Models

Peter G. Engeldrum
Imcotek, Winchester, MA 01890
pge@imcotek.com
781-756-0935

Abstract

This paper describes an experiment to determine the absolute graininess (GS) threshold of uniform (solid) area images. The psychometric experiment used a variation of the Method of Constant Stimuli where the observers sort the stimuli (samples) depending on their ability to "see" or detect graininess. Since graininess is composed of at least two Physical Image Parameters, (PIP), the lightness optical density curve for the Human Visual System (HVS) and the Wiener Spectrum, a Visual Algorithm is used to specify the stimuli GS. This is a variation on the classical method of absolute threshold where the stimuli are specified in terms of a single PIP. Several psychometric models are discussed and methods to fit the experimental data are described. Finally, a linear probability model is used to determine the absolute GS threshold in terms of a Density-Wiener Spectra space.

Introduction - the Image Quality Circle

Image quality and its components is a complex problem that are still active research topics. To simplify the understanding of image quality we use a step-by-step approach called the "Image Quality Circle"[1] (IQC), which is shown in Fig. 1.

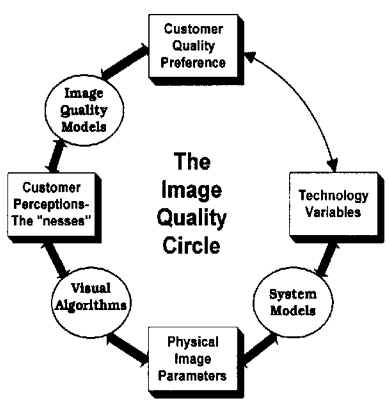

Figure 1: The Image Quality Circle

The goal of an imaging system designer is to relate the technology variables of the imaging system to the quality preferences of the customer. Figure 1 shows this fundamental objective via the large arrow. The link between customer quality preference and the imaging system and materials variables (parameters) is typically determined by varying a parameter, printing images, and then asking customers to judge the quality of the printed image. This clearly works, but, it is inefficient because a new data collection effort is required every time a parameter is changed. The IQC breaks the relationship between Technology Variables and Customer Perceptions down into a series of definable and measurable steps. The four elements of the IQC approach are depicted as in Figure 1.

Customer Quality Preferences

Customer image quality preference is the overall image quality rating as judged by customers. In our experiments, this is a 0 to 100 interval rating scale of overall image quality using two reference points, usually at 20 and 80.

Customer Perceptions

The major customer perceptual components of image quality are such dimensions as darkness, sharpness, and graininess. These are called the "nesses".

Physical Image Parameters

Physical image parameters are the quantitative functions and parameters we normally ascribe to image quality, such as modulation transfer, Wiener spectra, or amplitude spectra. From a historical perspective, these are considered to be objective measures of image quality.

Technology Variables

Technology variables are the elements or parameters that the imaging system designer or imaging system manufacturer manipulates to change the image quality. Such variables include dots per inch (resolution), toner size, and paper parameters, to name just a few.

The four elements of the IQC are linked to one another via models, or algorithms, which are depicted as ovals in Figure 1.

Image Quality Models

Image quality models are empirical (statistical) models that relate the customer perceptions–darkness, sharpness, graininess and raggedness, for example–to Customer Quality Preferences or image quality. The model describes in mathematical terms the tradeoff that the customer makes when judging image quality.

Visual Algorithms

Visually-based algorithms have a long history in photographic image quality. The algorithm is the recipe that is used to compute a value of a percept; sharpness for example, from something like the measurement of the gradient of a printed edge. In the context of the IQC, "visually based" means that the spatial properties and the nonlinear stimulus-response aspects of the human visual system are explicitly incorporated into the algorithm.

System Models

System models are analytical models that predict the physical image parameters from the technology variables. One example might be the model for the amplitude spectrum of a line boundary (the physical image parameter from which raggedness is calculated) for a dot-matrix printer, developed by knowing dot diameter and dot spacing parameters[2].

In this paper we describe the determination of the absolute visual threshold of graininess, a Customer Perceptions or "ness" using a previously published graininess Visual Algorithm[3]. Previous graininess threshold studies[13,14] have explored only the magnitude of the density or reflectance fluctuations, granularity-a PIP, and have ignored the affect of average density. Several different psychometric functions were used to fit the data and to estimate the visual threshold. Using the Physical Image Parameters (PIPs) that comprise the graininess algorithm it was possible to formulate the threshold psychometric functions in terms of a linear probability model. This has not been the usual psychophysical approach to threshold determination. More commonly the stimulus is specified in terms of some physical measure; the luminance of a light source or the acoustic pressure of sound[4]. AN illustration is presented on how a ness threshold can be related back to the physical specification of the stimulus using the concept of a linear probability model[5]

We had three objectives for this study. The first was to determine the absolute Graininess threshold for uniform gray electrophotographic image samples using graininess values estimated via a graininess algorithm. Our second objectives was describe the observer response data in terms of a Linear Probability Model (LPM) that incorporated the PIPs in such a way we hoped to link the graininess threshold to image physics. The third was to determine a boundary in a PIP space of Wiener Spectra and Density that demarcates the region of visible graininess.

Graininess Algorithm

The graininess (GS) algorithm used in this study has been described by Engeldrum and McNeill[3]. (Note that there is an addenda that has significant corrections to the algorithm.) The graininess defined by this algorithm is essentially the logarithm of the standard deviation, RMS, of the lightness fluctuations as seen by the Human Visual System (HVS). The algorithm incorporates in a simple manner two important properties of the HVS; the nonlinear transduction of luminance to the ness - lightness, and, a spatial frequency weighting of the density fluctuations of the image. The correct graininess algorithm[3] is given by:

$$GS = \log\left(\sqrt{\left|\frac{dL^{**}}{dD}\right|^2_{\bar{D}} \int_0^\infty VTF^2\left(\frac{u}{m}\right) WS\left(\frac{u}{m}\right) du} \right) \qquad (1)$$

where L** is the lightness of complex fields for a light (normal) surround[6] and is given by =

$$11.5\sqrt{10^{(2D-1)}+1}-16,$$

and the magnitude of the derivative is evaluated at the average density, VTF(u/m) is the spatial frequency weighting = 5.05{exp(-0.84u) - exp(-1.45u)} for u > 1.0cy/mm and = 1.0, 0 < u 1.0cy/mm (this is for a 35cm viewing distance), WS(u/m) is the Wiener spectra[7] in units of Density^{2}µm^2, and *m* is a magnification or scaling factor.

Experimental Procedure

Samples (Stimuli)

The graininess samples were nominally constant gray levels from black and white electrophotographic printers and copiers, plus gray paint "chips" and Munsell gray patches. Incorporating the paint chips and the Munsell gray patches assured that there would be samples (stimuli) that had essentially zero, or at least imperceptible, graininess. Eighteen one inch by one inch patches were mounted on three inch by five inch index cards for presentation to the observer.

Measurements

The Wiener Spectra were measured using a CCD video camera with a frame grabber connected to personal computer system. An image frame consisted of a 8 bits quantization of 512×480 pixels (RS-170) corrected for pixel-to-pixel "gain" variation and nonuniform illumination; In other words each pixel was linearly corrected. The camera was set up so the influx geometry was a fiber optic ring light at 45 degrees from the normal and the reflectance factor calibration was relative to a perfect reflecting diffuser.

Once the image frame was captured and corrected the frame was "scanned" by an effective area of 0.0333mm wide by 1.0mm long. Essentially, a synthetic aperture was placed in the frame, the average reflectance was taken over the aperture, and the reflectance converted to optical density. This comprised one data point. The aperture was moved one half a slit width and the process repeated. Two data arrays were constructed, one for the horizontal scan direction and the other for the vertical scan direction. For the horizontal scan direction a total of about 4,000 data points were taken and 25 blocks of 158 data points were used to compute the average WS. For the V direction a total of 31 blocks of 129 data points were used. With these parameters the relative 95% confidence interval (fraction of the computed value) on the Wiener Spectrum varied from about ±0.3 for the horizontal measurements to about ±0.4 for the vertical WS measurements[7]. Computation of the WS was carried out to 14 cy/mm and corrected for the data collection system MTF. The average density was computed from the 4,000 data points and removed from the data before the WS calculation.

The sample space in terms of the WS fluctuations and average density is shown in Figure 2. In this figure the ordinate is the granularity constant, G, which is just the zero spatial frequency value of the measured WS[8]. From each calculated WS and average density, the graininess of each sample was calculated using equation (1). There was no substantial difference between the horizontal and vertical graininess values so the average of the two directions was taken as representing the graininess of the sample.

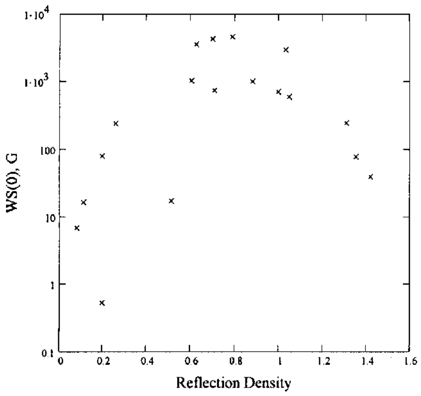

Figure 2: G-Density space of graininess stimuli.

Observers

Thirteen, mostly male, observers performed the psychometric experiment. All observers are engineers and have various degrees of familiarity with assorted imaging systems. In other words all observers were experienced with viewing and evaluating the nesses, visual attributes, of images. This is a small number of observers compared to classic psychophysical experiments that often use either a large number of observers, 100 or so, or a large number of replicate observations. A small sample number is not atypical of industrial experiments, with the consequences being lack of precision in the estimate of the absolute threshold. Increasing the number of observers will ultimately improve the confidence in the graininess threshold.

Psychometric Experiment - Method of Constant Stimuli

Our prime interest was determining the Graininess threshold for practical applications; specifically what people would see in their office environment. This choice adds additional variation to the results because there was no attempt to control, for example, the illuminance level when the samples were viewed, although the spectral quality was almost always the same - cool white fluorescent. This and other factors would be carefully monitored and controlled in a "proper" psychophysical experiment, but was not an option in this, or in a lot of product development situations.

A psychometric experiment to determine the threshold graininess value, using a variation on the classical Method of Constant Stimuli[9], was conducted as follows. Each observer was given the following definition of graininess before starting the experiment:

> "Graininess is the lack of uniform appearance of the solid areas. High graininess means the patch is not at all uniform."

This rather broad definition of graininess was used because many of the sample patches had low spatial frequency variations that are often called "mottle". Perhaps a better term would have been "uniformityness".

Observers were asked to read was the following set of instructions:

> "I will give you a series of graininess samples. Would you please place the graininess samples in two piles. The first pile is for samples that you feel have visible graininess. The second pile will contain samples that you feel have no visible graininess. View the samples at normal viewing distance. Take your time, there are no right or wrong answers. We are interested in your opinion."

At the end of the experiment the observer returned two piles to the test conductor. The first pile contained the samples that have visible graininess and the second pile contained the samples the were judged to have no graininess. If a sample were judged to have visible graininess a one was added to the count for that sample. After all the observer data were collected the resulting data consisted of the proportion of observers that judged each sample to have visible graininess.

Data Analysis

The data analysis task was to fit the proportion versus graininess data with a psychometric function [Psychometric = "mind/measuring"(10).]. A psychometric function is a cumulative probability function, (CDF) of some probability density function. Two most popular and widely used in psychophysics are the Gaussian or normal, and the logistic. The analysis methods are often called Probit analysis and Logistic or Logit analysis, respectively.

The general mathematical form of the estimation or fitting problem is given by equation (2).

$$P_j = F(\alpha + \beta x_j) \qquad (2)$$

In equation (2) above P_j = the proportion (probability) of a "yes there is visible graininess" responses to sample (stimuli) j, having a graininess value of x_j. The constants, α and β are parameters that need to be estimated.

For the two models considered, the Gaussian or normal, and the logistic, equations(3) and (4), respectively, define $F(\ )$. For the normal psychometric function we have;

$$P_j = F(\alpha+\beta x_j) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha+\beta x_j}^{\infty} e^{-\frac{z^2}{2}} dz = \Phi(\alpha+\beta x_j) \qquad (3)$$

The logistic is formally given by equation(4);

$$P_j = F(\alpha+\beta x_j) = \frac{1}{1+e^{-(\alpha+\beta x_j)}} = \frac{e^{(\alpha+\beta x_j)}}{1+e^{(\alpha+\beta x_j)}} \qquad (4)$$

Several options are available for estimating the α and β parameters of both psychometric models. The simplest is to plot the z-values, ($\alpha + \beta x_j$), corresponding to the proportions, P_s. These values will usually fall on a straight line and simple least-squares fit often suffices. The second option is a non-linear least-squares fit of the proportions to the psychometric curve.

One practical problem must be dealt with when using the z-value fitting procedure, and that is zero and one proportions. Using either equation (3) or (4) will yield z-values that are either -∞ for a proportion of zero or +∞ for a proportion of 1.0. For this application, fitting the transition region from just above the zero proportion to just below the unity proportion, it most important. An approach is to change the last zero value to 1/(2*number of observers) and ignore all previous zeros, and at the high end replace the first unity proportion with 1-1/(2*number of observers) and ignore all subsequent unity proportions.

The third estimation method, which was only implemented using the logistic psychometric curve, is called maximum-likelihood[5]. With maximum-likelihood estimation, the zero and one proportions are not a problem.

Once the psychometric curve parameters are determined the value of graininess that gives a 0.5 proportion, or probability, is selected as the graininess threshold. For both psychometric functions the threshold estimate occurs when $\alpha + \beta x_j = 0$. Since it assumed that the proportions are drawn from a binomial population, exact confidence intervals can be placed on the 0.5 proportion based on the number of observers[11]. A graininess threshold confidence interval can be obtained by reflecting the proportion confidence interval through each psychometric curve, using the appropriate parameter estimates.

TABLE I - Summary of estimated psychometric function parameters, graininess threshold, and graininess threshold confidence intervals.

Model	Estimator	Parameter Values	RMS fit	Graininess Threshold	95% CI on GS Threshold
Normal	Linear LS	α= -0.324 β = 1.233	0.0382	0.263	-0.258→0.785
	Nonlinear LS	α = -0.431 β = 1.371	0.0341	0.314	-0.154→0.783
Logistic	Linear LS	α = -0.549 β = 2.111	0.0387	0.26	-0.235→0.756
	Nonlinear LS	α = -0.755 β = 2.348	0.0343	0.321	-0.124→0.768
	Maximum Likelihood	α = -0.613 β = 2.488	0.0405	0.246	-0.174→0.666

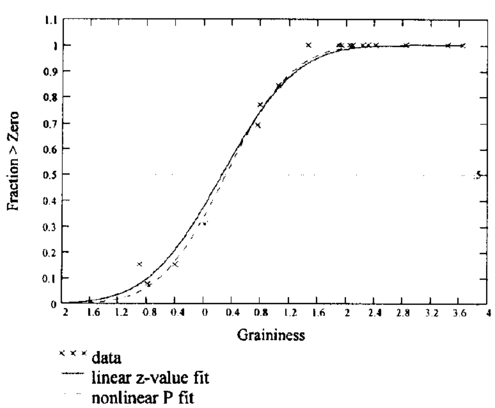

Figure 3: Normal psychometric curve fits; linear on z-values and nonlinear on proportion values.

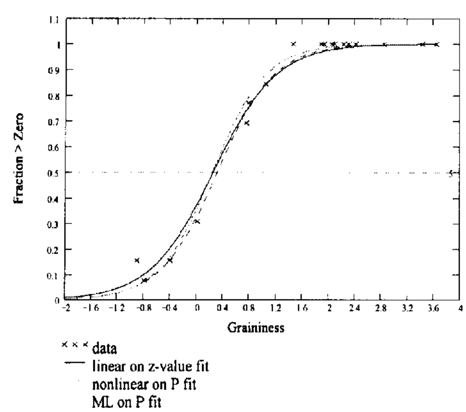

Figure 4: Logistic psychometric curve fits; linear on z-values, nonlinear on proportion values, and maximum likelihood.

Threshold Results

Table I summarizes the results of fitting two psychometric curves using two or three estimation methods. The standard deviation of the data points about the fitted curve is also shown in the table. All fits are statistically significant.

Figure 3, the normal psychometric cure, shows two fits using the linear fit to the transformed values and the nonlinear fit to the proportions. Figure 4, the logistic psychometric curve, illustrates the three estimation methods: transformed, nonlinear and maximum likelihood. It is difficult to say which is the "best" estimation method. However, on an RMS basis the logistic curve tends to fit the data slightly better, but either psychometric model is satisfactory for this data.

Once we determined the graininess threshold, a logical scale of graininess was made by assigning zero to this threshold value. The scale can also be multiplied by a suitable value so some graininess value yields some convenient scale value.

Linear Probability Model

The usual application of a nonlinear probability model LPM, is to model dichotomous responses, (0-1 or yes-no) as some linear combination of factors. It is somewhat similar to a liner regression model except the dependent variable is a dichotomous response. In this application we depart somewhat from the standard application of the LPM in that we fit the proportions, which are not dichotomous but use the formalism as our model. Using the logistic function, or logit model, we can write the model equation(5)[12,5];

$$P_j = F(\alpha + \beta \sum_{k=1}^{q} GS_{jk}) = \frac{1}{1 + e^{-(\alpha + \beta \sum_{k=1}^{q} GS_{jk})}} \quad (5)$$

where q is the number of linear factors, or variables, in the LPM. Classical psychophysical threshold curves have only one physical variable that describes the stimuli; for example the luminance of a light or the acoustic pressure of a sound wave. The specification of the stimuli via a ness is not at all common, usually because a Visual Algorithm, VA, connecting the Physical Image Parameters to the ness, (graininess in this instance) is not available. In the usual situation therefore, the only specification is the physical value of the stimulus.

We substituted the components of the VA into the psychometric function we estimated above, and we saw that there are two linear components; the two factors, log(lightness of average density) + log(integral of WS x VTF^2). This can be made clear if we rewrite the graininess algorithm, equation(1), in the form of equation(1a).

$$GS = \frac{1}{2}\log\left(\left|\frac{dL^{**}}{dD}\right|^2_{\bar{D}}\right) + \frac{1}{2}\log\left(\int_0^\infty VTF^2\left(\frac{u}{m}\right) WS\left(\frac{u}{m}\right) du\right) \quad (1a)$$

The first factor on the RHS of (1a) has to do with the mean value of the reflection density, or reflectance factor, of the image, and the second term represents the density fluctuations as weighted ("seen") by the human visual system.

Threshold Density-Wiener Spectra Space

From our psychometric experiment we knew the absolute graininess threshold value, GSthresh, and in combination with some reasonable assumptions about the WS, and equation (1a) we can determine a region in density-WS space where the graininess was below threshold.

To make some headway on establishing the boundary we assume that the WS is constant, "flat" with the WS(0) = G, the granularity constant having units of $\text{Density}^2 \mu\text{m}^2$. With this assumption G comes outside the integral in equation(1a) and the term reduces to ½ log(kG), where k is a constant equal to the integral of the square of the VTF. This is generally very conservative, because real images have WS that decrease with spatial frequency and the subsequent weighting by the VTF will make the integral less than k. We take the graininess threshold, from the estimates above, to be about 0.3 and solve for the granularity constant G as a function of optical reflection density. The derivative of L** with respect to D is scaled so it equals 1.0 when D = 0.

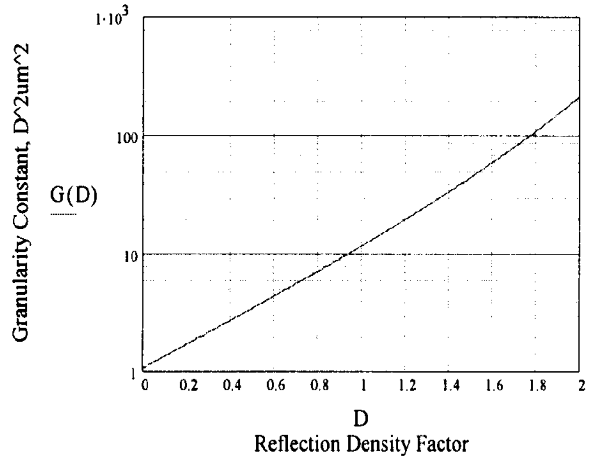

Figure 5: G-D at graininess threshold. Below curve GS is not visible.

This boundary line is shown in figure 5. Below this line the graininess was not detectable and above this line it was. This figure also illustrates that an increase in the physical fluctuations are tolerated as the density increases or the lightness decreases. This graph is the Physical Image Parameters trade-off responsible for graininess. For some imaging systems the Wiener Spectra dependence on imaging element size is known and thus Graininess can be cast, ultimately, in terms of Technology Variables[8].

Conclusions

1) The absolute graininess threshold has been determined using a normal and logistic psychometric curve and several different model parameter estimation methods. The range in the threshold estimates is about ±20% of the mean value.

2) Formulation of the graininess threshold as a Linear Probability Model illustrates that two Physical Image Parameters, the mean lightness and the visually weighted density fluctuations are required to predict the psychometric experiment. This is essentially a consequence of the graininess model.

3) Using the graininess threshold, GSthresh, a Wiener Spectra scale value-reflection density space can be constructed the defines a region where the graininess is not visible. Generally the results show that an increase in physical fluctuations with increasing density are allowed while keeping the Graininess below threshold.

References

1a. Engeldrum, P. G., Measuring customer perception of print quality, *TAPPI Jour*. **73**:161(1990), b) Engeldrum, P. G., A framework for image quality models, *Jour. Imag. Sci. Tech*. **39**:312(1995)
2. Engeldrum, P. G. Quality of dot formed images, *SPIE* Vol **310**, Image Quality (1981), pp 118-124.
3. Engeldrum, Peter G. & Glenn E.McNeill, Some experiments on the perception of graininess in black and white photographic prints, *Jour. Imag. Sci*. **29**: 18 (1985), Addenda in *Jour. Imag. Sci*. **29**: 207 (1985)
4. a) Baird, John C. & Elliot Noma, Fundamentals of Scaling and Psychophysics, John Wiley & Sons, New York, NY, 1978, ISBN 0-471-04169-6, b) Gescheider, George A., Psychophysics, Method, Theory, and Application, 2nd Ed., Lawrence Erlbaum Associates, Inc., 10 Industrial Avenue, Mahwah, NJ 07430, 1985. ISBN 0-89859-375-1
5. Aldrich, John H & Forrest D. Nelson, Linear probability, Logit, and Probit models, Sage University Paper series on Quantitative Applications in the Social Sciences, 07-45 (1985) . Beverly Hills & London: Sage Publications. ISBN 0-8039-2133-0
6. Bartleson, C. J., Optimum image tone reproduction, *JSMPTE* **84**:613(1975)
7. Burns, P. D. & B. M. Levine, Wiener spectrum estimation at zero frequency via direct digital computation, *Jour. Appl. Photo. Eng*. **9**:78(1983)
8. Dainty, J. C. & R. Shaw, Image Science, Academic Press, London, 1974, chap. 8.
9. Guilford, J. P., Psychometric Methods, McGraw-Hill, New York, 1954, chapter 6.
10. Nunnally, J. C. & I. R. Bernstein, Psychometric Theory, Third Edition, McGraw-Hill, New York, 1994, ISBN 0-07-047849-X
11. Sachs, L., Applied Statistics, Second Edition, Springer-Verlag, New York, 1984, ISBN 0-387-90976-1, pg 333.
12. Liao, T. F. (1994) Interpreting probability models: Logit, probit, and other generalized linear models, Sage University Paper series on Quantitative Applications in the Social Sciences, 07-101. Thousand Oaks, CA: Sage Publications. ISBN 0-8039-4999-5
13. Zwick, D., D. L. Brothers, RMS granularity: determination of just-noticeable-differences, *Photo. Sci. & Eng.*, **19**:235(1975)
14. Hamerly, J., Just noticeable differences for solid area, *Jour. Appl. Photo. Eng*. **9**:14(1983)