MTF ENGINEERING NOTES

Use of Sinusoidal Test Patterns for MTF Evaluation

Robert L. Lamberts

Initial work on modulation transfer functions (MTF) was done in the 1940's and 50's. In the ten years that followed many practical methods were worked out and were described in the literature.[1-3] Other work showed that the MTF of a system could be used to provide a very reliable measure of visual sharpness.[3] A problem that engineers encountered when attempting to apply these techniques was that sinusoidal test patterns were not generally available. Sine Patterns now fills this need with a variety of such test patterns.

In these notes we are providing a brief description of methods for MTF measurement using sinusoidal test pattern arrays. For a more extensive discussion the reader should refer to the published literature.

The Modulation Transfer Function

If we were to plot the transmittance of a sinusoidal pattern as a function of distance we would obtain something similar to that shown in Fig. 1. Because transmittance cannot go negative there is always a bias level, shown as B in the figure, associated with the sinusoidal variation. We define the modulation of the pattern as the ratio A/B. Although in theory the modulation can have a maximum value of 1.0 or 100%, in practice this cannot be achieved with a photographic pattern since zero transmittance is not practically attain-able. But, as we will see, this is not usually a serious disadvantage.

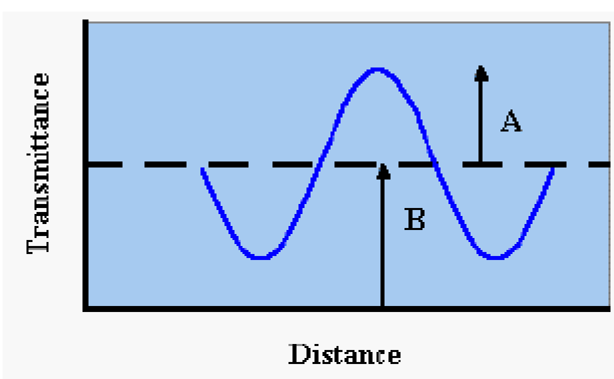

Figure 1

When a sinusoidal pattern is imaged by a "linear" optical device such as a lens (i.e. where the intensity in the image is linearly proportional to the illumination of the object), the image will retain its sinusoidal profile; only the modulation will change.

The ratio of the modulation in the image to that of the object will then be a measure of how well the image is formed for that particular spatial frequency. (Spatial frequency is usually measured in cycles or line pairs per mm.) This ratio is called the modulation transfer ratio, that is,

Modulation transfer ratio = (Modulation of image) / (Modulation of object)   1)

When this measurement is made over a range of spatial frequencies, we can plot a curve of the MT ratio as a function of spatial frequency, giving us what is known as the modulation transfer function. A typical curve is shown in Fig. 2.

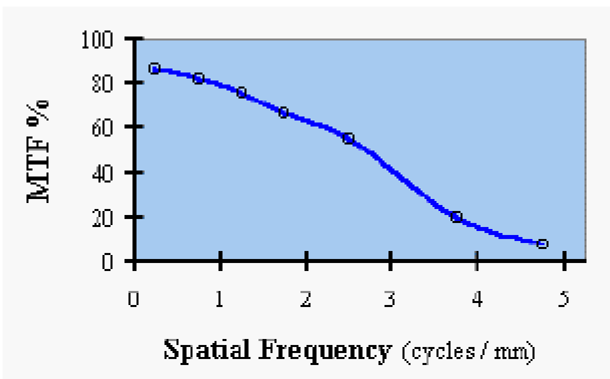

Figure 2

The MTF has a number of interesting properties:

1) The MTF covers a range of spatial frequencies and hence gives more information regarding image performance than does limiting resolution, which is essentially the endpoint of the MTF. The lower and mid-range spatial frequencies can be particularly important in regard to observed image sharpness.

2) The MTF of a system is equal to the product of the MTF's of each component of the system. For example, the MTF of a photographic print is equal to the product of the MTF's of a) the camera lens, b) the film, c) the printer lens, and d) the photographic paper. The modulation transfer values are multiplied for each spatial frequency. [Note: In the situation where magnification occurs, as with the printer, the spatial frequency changes according to the reciprocal of the magnification. One then adjusts the spatial frequency scales of the MTF's of the various parts of the system accordingly.] Figure 2 shows an actual measurement of the system for a commercial color print. A photograph was taken of a large reflection test pattern and the final print was scanned and evaluated to determine the MTF for the complete system.

3) The MTF is related mathematically, by a Fourier transform, to the image of a line, which is also known as the line spread function. The edge function, in turn can be obtained by integrating the line spread function. This also means that it is possible to derive the MTF from the edge function. There is a distinct

advantage, however, for using sinusoidal test patterns to determine the MTF in that they cover a much larger area and provide a much greater degree of redundancy in the measurements.

Some systems are practically linear, but in others there is significant nonlinearity. An example of the latter is a photographic negative film. While the theory assumes linearity, it has been shown[4] that good results can be obtained, even with significant nonlinearity. Methods for measurement in such instances will be discussed in Example 4 below.

The type of measurements we want to describe in these notes make use of sinusoidal test patterns, examples of which are shown in our catalog. These are made either on film or on photographic paper. In both types the harmonic content can be kept to 3% or less. The arrays contain a series of sinusoidal areas and a gray scale which serves as a zero-frequency reference. We will discuss the function of the gray scale later in this paper.

Optical Density

In our discussion that follows we will frequently be referring to optical density (or density). For many purposes it is more convenient to work with density than with transmittance. Density is defined as D = -log T, where D is the density and T is the transmittance. Thus when T = 100%, D = 0, when T = 10%, D = 1.0, etc. We can also define density for a reflecting material where D = -log R, and where R is the reflectance.

Since density is a logarithmic quantity, when two transmitting materials are placed together, their transmittances multiply, while their densities will add. Visually, equal steps of density appear to be more equally spaced than do equal steps of transmittance.

For purposes of standardization, density is generally measured with near collimated (or specular) illumination and with a pickup that integrates the transmitted light. This is called diffuse density. Because of optical reversibility principle, the same density would be measured if the illumination were totally diffuse and the pickup were specular.

Diffuse density for our test patterns is measured with white light and is very close to the density measured with green or red light. However in the blue region the densities (and contrast) are somewhat higher. Likewise although the film base is transmitting in the 3 to 5 micron region of the spectrum, the overall contrast of the pattern is lower than in the visible. This can have only a relatively minor effect on data if the measurements are made in terms of the gray scale. We will describe methods for doing this if the examples that follow.

Callier Q-Factor

Because of light scattering within the developed photo-graphic film, the effective density of the film will depend upon how it is illuminated. Generally speaking, the given density of a film will tend to be higher for specular illumination than for completely diffuse. However, it turns out that the ratio of these densities is fairly constant over a wide range of densities. Thus we can write for the Callier Q-factor

Q = (Specular Density) / (Diffuse Density) ... 2)

Measurements show that one obtains a similar constant ratio between blue light and green light density values. Hence

R = (blue or IR density)/(green density) ... 3)

Either Q or R can be determined by comparing the density values obtained in the user's practical setup with the values listed on the microdensitometer readout that we supply with the test pattern.

Test Pattern Calibration

Before actual MTF values can be determined it is necessary to determine the modulation values of the test pattern.

Although our pattern arrays have nominal values of modulation, in practice they tend to vary somewhat from these. Therefore, before we can calculate the MTF values with accuracy we need to determine the modulation values for each of the areas in the array. This can be done with the help of a microdensitometer scan.

Our microdensitometer scans give modulation values directly as computer printout. Methods for doing this are discussed under the Digital Readout section. However, let us suppose that the readout is recorded on a strip-chart recorder that gives a plot of optical density as a function of distance. The gain of our recording system is now adjusted so that density values can be read directly from the chart and one can then pick off values of maximum and minimum density for each of the scanned areas. By letting the difference be ΔD, we can obtain a modulation value from the formula,

$$\text{Modulation} = (10^{\Delta D} - 1) / (10^{\Delta D} + 1) \qquad \text{... 4)}$$

We would then determine such values for all spatial frequencies.

Practical Examples

We will now go through a number of examples to illustrate how measurements can be made.

<u>Example 1. Lens MTF (linear system)</u>

Let's assume that we want to determine the MTF of a 1 to 1 printing lens. We will use a setup as sketched in Fig. 3. A transmission test pattern array is illuminated from behind with a diffuse source. This can be an integrating sphere, an illuminated opal glass, or a condenser system where the light source is imaged into the test lens and overfills the aperture.

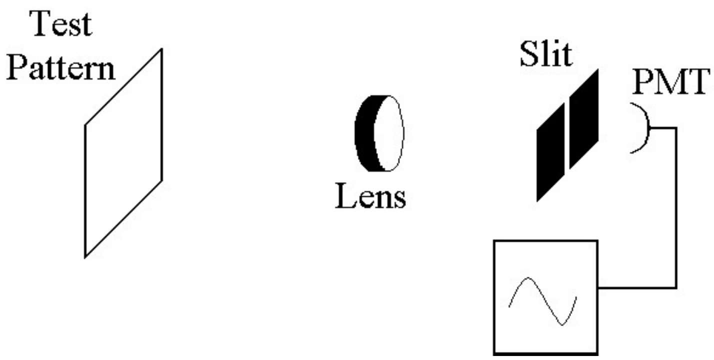

Figure 3

In the image plane we will place a narrow slit that is carefully aligned parallel to the lines in the test pattern. If the slit width is no more than about 5% of the line spacing of the highest frequency pattern area there will be negligible error introduced because of the slit width*. Behind the slit is a photo-sensor that integrates the light. Its output goes to an oscilloscope, strip-chart recorder, or similar recording device.

If we now move the test pattern laterally, the output on our recording device will be similar to the plot that we showed in Fig. 1. From it we can obtain a value of modulation. As an alternative to measuring the ratio A/B, note that if we determine maximum and minimum transmittance values of the sine curve, then

$$\text{Modulation} = (T_{max} - T_{min}) / (T_{max} + T_{min}). \quad \text{... 5)}$$

Focusing the lens can be done simply by looking for the largest amplitude in the scan for one of the highest spatial frequencies. There is an advantage to moving the test pattern rather than the slit since non uniformities in the light source would then have a much smaller effect.

After all modulation values have been determined for the imaging system, it is necessary to divide by the modulation values of the test pattern, as in Eq. 1) to obtain the MTF.

<u>Example 2: Correction for Q-Factor or Color of Illumination</u>

If the test pattern illumination is not completely diffuse or if the pattern is used with blue light or in the infrared, it may be necessary to apply a correction. This is done by using the gray scale of the test pattern.

This follows because the Fourier transform of the slit function is a sin x/x function which has its first zero for the spatial frequency where the slit width equals the period (or reciprocal of the spatial frequency). For 5% of this width the sin x/x function is 0.996 and for 10% it will be 0.984, which is still less than a 2% error.

If measurements are made by using the peak to peak method, the value of ΔD is divided by either Q or R as in Eq's 2) or 3). When digital techniques are used, each density value is divided thus. Where transmittance values are used you can either convert them to density values and then do the division, or you can use the relationship

$$T(\text{diffuse}) = T^{1/Q}(\text{specular}) \qquad \text{... 6a)}$$

or,

$$T(\text{green}) = T^{1/R}(\text{blue}). \qquad \text{... 6b)}$$

Example 3: MTF of a Linear Device

Suppose we want to measure the MTF of a CCD array or similar device. We will then use a setup similar to that shown in Fig. 3. Of course the lens can be used at any magnification that is convenient. We will now scan each of the pattern areas in the wavelength region where the measurement is to take place and obtain modulation values as described in Example 1.

The slit assembly can now be removed and replaced with the CCD array. By feeding the output of the array into an oscilloscope we can obtain an output that again resembles Fig. 1 from which we can obtain our values of modulation. In the situation where the data appearing on the scope are ‘stair stepped’ because of digitizing, it may be necessary to move the test pattern or the CCD slightly in order to find the peaks and valleys. The MTF of the device is then the ratio of these modulation values divided by the values obtained from the photoelectric scan.

Example 4: Photographic Film or Other Non-Linear Systems

Sinusoidal test pattern arrays can also be used for deter-mining the MTF of photographic materials or for any other system where we have a nonlinear relationship between illumination and the output. To illustrate, suppose that we want to measure the MTF of a black and white negative film.

As a first step, the test pattern is contact printed onto the film, preferably using good vacuum contact. After processing, the film is scanned with a microdensitometer. Both the gray scale and the sinusoidal areas are included in the scan. We can conveniently circumvent the problem of non-linearity of the film by analyzing it in terms of exposure rather than in terms of its density or transmittance. The film density is then simply a photometer of exposure that occurred within the film coating.

Because the microdensitometer reads in terms of density we need to go through the sensitometric curve, (i.e. the curve of density vs. log exposure) in order to determine our values of exposure. Also, because we need to know only differences in log exposure, the absolute level is not important. Log exposure will then be equal the density of the test pattern plus or minus some constant.

We can then plot the recorder reading of the gray scale levels, as obtained from the scan of the film being analyzed, vs. the density values of the gray scale of the test pattern itself. This gives a curve as shown in the left side of Fig. 4. Values of maximum and minimum exposure can be deter-mined by going through the curve, as shown in the figure, and modulation values can be obtained from these.

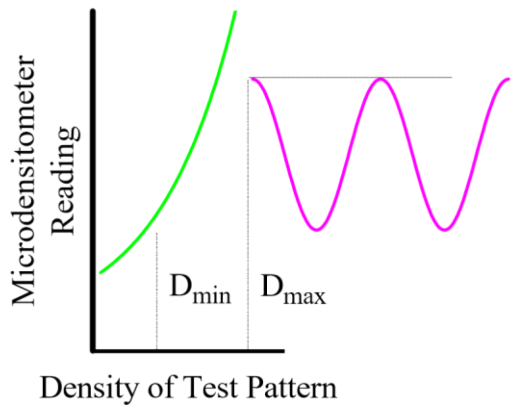

Figure 4

In some instances, where there is a sizable nonlinearity, it is often advisable to use a pattern having a lower modulation. This is particularly true when there is a very limited dynamic range.

Size Considerations

How large should the test pattern be? In some instances it is possible to image one area at a time. In other cases it is possible to fit the complete array into the field being imaged.

A word of caution should be given here because problems arise when the imaged patterns are made too small. According to mathematical theory, the MTF of an optical device is equivalent to the Fourier transform of its spread function. While Fourier transforms have infinite limits, we can truncate this process as long as we go beyond the width of the spread function. In practice this simply means that the widths of the sinusoidal areas should be at least several times that of the spread function of the system. It is for this reason that we maintain a minimum width for all the higher frequency areas.

Digital Readout

For many applications it may be desirable to carry out the analyses we have been describing by means of a digital readout.   The methods we will describe here are basically straight forward, but care must be taken to avoid problems that may arise.

A rather obvious method for analysis would be simply to determine the maximum and minimum points in a scan, but this can give erroneous results for several reasons. First, the maxima and minima may be the result of noise peaks. In such a case, the more noise, the greater the apparent signal. Second, there may be a slight density wedging in either the original pattern or in a subsequent exposure. With this method could then be reading the sum of the signal and the wedging.

Fourier Analysis

A very good numerical method is that of Fourier analysis. Not only does this method provide the information we want, but it also applies a type of filtering that discards noise and other extraneous parts of the digital information.

Let us assume that the transmittance of a sinusoidal area can be described by the equation,

$$T(x) = t_0 + t_1 \cos 2\pi fx + t_2 \cos 4\pi fx + t_3 \cos 6\pi fx + \ldots \qquad \ldots 7)$$

where t_1 is the amplitude of the fundamental frequency and t_2, t_3 are the amplitudes of the harmonics. The spatial frequency, which can be in terms of cycles per mm, is given by f.

The data we obtain from our digital readout consist of a series of numbers. Our problem is to determine modulation values from these. Our aim is to be able to construct Eq 7) from our string of digital data.

According to Fourier analysis theory, if we have some sort of function, T(x) that repeats itself periodically (such as a sinusoidal pattern), we can then write

$$T(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos 2\pi nfx + b_n \sin 2\pi nfx) \qquad \ldots 8)$$

where

$$a_n = 2f \int_{-1/2f}^{1/2f} T(x) \cos 2\pi nfx \, dx \; (n = 0, 1, 2, 3 \ldots \qquad \ldots 9a)$$

$$b_n = 2f \int_{-1/2f}^{1/2f} T(x) \sin 2\pi nfx \, dx \; (n = 0, 1, 2, 3 \ldots \qquad \ldots 9b)$$

The summation in Eq 8 goes to infinity but for practical purposes with our test patterns, there is no need to go beyond n = 3. The magnitude of harmonics beyond this is negligible.

If we could assume in our data analysis that the point x = 0 was where the cosine function was a maximum as in Eq 7, then we could show that all sine terms would be zero. However in most instances our starting point is arbitrary as illustrated in Fig. 5. As a result there is a phase angle $\emptyset$, that we must take into account.

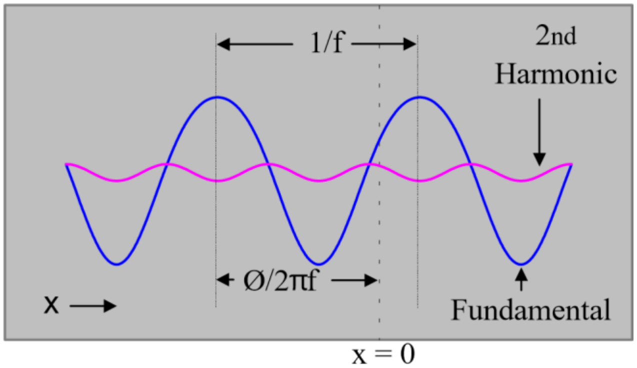

Figure 5

To do this we will make use of the trigonometric identity,

$$\cos(x - y) = \cos x \cos y + \sin x \sin y$$

and write Eq 8) in the form

$$T(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(c_n \cos(2\pi n f x - \emptyset_n) \qquad \ldots 10)$$

where

$$a_n = c_n \cos \emptyset_n \qquad \ldots 11a)$$

$$b_n = c_n \sin \emptyset_n \qquad \ldots 11b)$$

$$\emptyset_n = \arctan b_n / a_n \qquad \ldots 11c)$$

and

$$c_n = \pm (a_n^2 + b_n^2)^{1/2}. \qquad \ldots 11d)$$

Determination of Signs of Coefficients

As will be shown in later discussion, if we are to make a determination of the peak to peak modulation of our signal it is important that we are able to determine the signs of the coefficients c_n as given in Eq 10).

As can be seen from Eq 11c), there is an uncertainty regarding $\emptyset_n$ since the tangent is positive in the first and third quadrants and negative in the second and fourth. To resolve this, we can examine the sign of b_n. If it is negative, $\emptyset_n$ is either in the third or fourth quadrant which means that we should add π to the angle.

We also see from Eq 11d) that there is an uncertainty regarding the signs of the coefficients unless we take other steps to determine them. We first assume that the coefficient for c_1 is positive. As can be seen from

Eq 10), this will be true when the argument of the cosine function is zero which means that $\emptyset_1$ = 2πfx. This has the effect of establishing a reference point for our data.

As for the harmonics, if they are produced because of non linearities in the system (e.g. as occurs in photographic processing) then they are either in phase or 180° out of phase. In other words, c_2 and c_3 can be either positive or negative. The second harmonic, for example, shown in Fig. 5 is negative.

Since we have assumed that c_1 is positive we have the relationship that $\emptyset_1$ = 2πfx, where x is the distance in Fig. 5 from the point x = 0 to the peak of the cosine function. Because this distance remains constant for all the harmonics, we need to multiply $\emptyset_1$ by 2 for the second harmonic since the frequency f is doubled. Likewise $\emptyset_1$ is multiplied by 3 for the third harmonic. Hence, $\emptyset_2 = \emptyset_{1/2}$ and $\emptyset_3 = \emptyset_{1/3}$.

Thus knowing $\emptyset_n$ and the sign of either a_n or b_n, we can determine the sign of c_n from either Eq 11a) or 11b). Our best evaluation of the magnitude of c_n is probably obtained from Eq 11d).

<u>Digital Calculations</u>

Because the data in a digital system are not continuous we cannot do an integration but need to do a summation. To do this we must assume that the points are separated by equal intervals, either distance wise or time wise.

The summations corresponding to Eq's 9a and 9b will have the form

$$a_n = 2f/\mathrm{m} \sum_{m'-1}^{m} T_{m'} cos2\pi nm'f \qquad \text{… 12a)}$$

$$b_n = 2f/\mathrm{m} \sum_{m'-1}^{m} T_{m'} sin2\pi nm'f \qquad \text{… 12b)}$$

These provide the same information as the real and imaginary coefficients respectively obtained from an FFT.

It is essential to understand that the theory assumes that we have a periodic function, one that repeats itself. For best results we also want the m data points to cover one of these periodic intervals. In practical terms, when we are using an FFT, this means that a periodic interval should extend over 64, 128, 256, or some power of two points. For a scanning system it may be necessary to adjust the speed by which the test pattern is scanned. For testing a device such as a CCD it may be necessary to adjust the optical magnification from the test pattern to the device.

One of these periodic intervals can be one cycle of the sinusoidal pattern, but it can also be 2 cycles, 3 cycles, or any integral number of cycles, since the pattern will repeat after each of these. This is fortunate in that we are then able to average over a number of cycles in taking our data.

This fact also makes it very convenient when using one of our test patterns because of the way the various spatial frequencies are laid out. We can illustrate this with our M-7 Test Pattern. The spatial frequencies are in the sequence 0.75, 1.0, 1.5, 2.0, 3.0, etc. cycles per mm. Suppose that our m points extend over 3 cycles for the 0.75 cycles per mm. In terms of distance on the test pattern, this would correspond to 4 mm. For this same distance we can now fit exactly 4 cycles for 1.0 cycles per mm, 6 cycles for 1.5 cycles per mm, 8 cycles for 2.0 cycles per mm, etc.

For low spatial frequencies it may be desirable to skip points between readings in order to reduce the distances involved with higher spatial frequencies. In general, it is advisable to keep the number of cycles be analyzed to no more than 6 or 8. If you were to use a large number of cycles, even a small error in magnification or small error in scanning uniformity could cause a significant error in the computed data. In applying the summations given in Eq's. 12a and 12b, you should be aware that if, for example, five cycles of the pattern were being analyzed, then the values for a_1 and b_1 would appear as the fifth coefficients of the FFT. Likewise a_2 and b_2 would appear as the tenth coefficients, etc. Taking these coefficients, computed in this way, we will then determine values of c_n by means of Eq 11d.

For many applications we can ignore the harmonics and determine the modulation by using the equation

$$\text{Modulation} = c_1 / a_0. \qquad \text{... 13)}$$

The system MTF is then determined using only the modulation of the fundamental frequency in our microdensitometer readout (that is the coefficient for cos x, and not the peak to peak modulation).

If you need to calculate the peak to peak modulation, the procedure is slightly more involved. From Eq 10 we can see that if $x = 0$ and $\emptyset = 0$ then we obtain T_{max}. Likewise then $x = \pi$ and $\emptyset = 0$ we obtain T_{min}. Hence

$$T_{max} = a_0 + c_1 + c_2 + c_3 + ...$$

and

$$T_{min} = a_0 - c_1 + c_2 - c_3 + ...$$

When we substitute these into Eq 3) we obtain

$$\text{Modulation} = (c_1 + c_3 + ..) / (a_0 + c_2 + ..) \qquad \text{...14)}$$

As we have discussed earlier, the values of c_n should have their proper signs.

When 2^n Points Do Not Fit an Integral Number of Cycles

It may happen that it is not possible to fit the required number of points for an FFT in an integral number of cycles. There are several things that can be done in such a case. First, rather than using an FFT, you can do a numerical transform using Eq's 12a) and 12b), but where m is the number of points required to extend over an integral number of cycles. If the transform is done this way, m can be any number. It need not be a power of 2.

A second possibility exists when the points come close to an integral number of points, but not exactly. For example, 128 points may extend over 5.85 cycles. In such a case we would assume that n = 6, but not only will the 6th coefficients will have significant values but so will the 5th, 7th, and possibly the 4th and 8th and even the 3rd and the 9th. In other words, the spectrum will be spread over a number of coefficients. The farther we are from an integral number of cycles, the greater will be this spread, regardless of the number of cycles being evaluated.[5] Thus it is important to come as close to an integral number of cycles as possible.

To obtain a modulation value we use Eq. 11d) but summing the squares of the coefficients over the neighboring values of n as well. Values summed in this manner give very close to the same value as when all is concentrated in one pair of coefficients.

A potential problem with this procedure is that a_0 is somewhat dependent on the starting point of the scan but this variability becomes less as the number of cycles is increased.

References

1. Thomas, W. Jr. Ed., ÒSPSE Handbook of Photographic Science and Engineering, J. Wiley & Sons, New York (1973).
2. Dainty, J. C. and Shaw, R., Image Science, Academic Press, New York (1974).
3. Carroll, B. H, Higgins, G. C., and James, T. H., Introduction to Photographic Theory, John Wiley & Sons, New York (1980).
4. Lamberts, R. L., J. Opt.Soc.Am., 51, 982-987 (1961)
5. Lamberts, R. L. and Eisen, F. C., J. Appl Phot. Eng., 6,1-8 (1980)