A SURVEY OF AERIAL IMAGE QUALITY ASSESSMENT METHODS

Anders Boberg, MSc
The Royal Institute of Technology
Department of Photogrammetry
S-100 44 STOCKHOLM, Sweden
Commission I

1. Abstract

The accuracy of photogrammetric mapping, as well as the reliability of image interpretation, strongly depend upon the quality of the images used. It is therefore important to assess or describe image quality. Different areas of application and various methods of evaluation imply quality demands that differ considerably. An investigation into the nature of these demands and into the differences between them is necessary as a basis for further investigations of image quality criteria. Image quality measures must be relevant with respect to subjectively experienced image quality or image usefulness. At the same time the measures should be based on objective measurements of image properties.

In this survey, subjective as well as objctive image quality criteria are studied. Methods for evaluation of image information are investigated from an image quality point of view. Special emphasis is put on the adjustment of objective assessment methods to subjectively experienced image quality with respect to certain applications.

2. Background

The result of an application of aerial photographs, either in the form of photogrammetric mapping accuracy or in the form of image interpretation reliability, is highly dependent on the image quality of the photographs used. An example of this is the accuracy and reliability improvement of photogrammetric mapping from low altitude aerial photographs, resulting from the introduction of aerial cameras using forward motion compensation (Ackermann1986, Boberg1986). Image quality affects not only the result of the mapping or interpretation project, but also the time consumed for image evaluation and subsequent field completion, and therefore the costs involved in the project. It is therefore important, not only to produce aerial photographs of the highest possible quality, but also to be able to assess or describe the quality of the images produced.

The main objective of the project, of which this is a status summary, is to find methods for assessment and description of the image quality of existing aerial photographic negatives. The quality of the image-forming process, or of the parts thereof, is thus not stressed. Also, the presence of imaged test objects cannot be presupposed. The study is restricted to near-vertical photographs, taken with 15/23x23 cm aerial mapping cameras on black-and-white film, disregarding e.g. geometrical quality and color balance. In this paper, a number of image quality descriptors and assessment methods will be reviewed from the point of view of their relevance and practical applicability. The review is based upon a literature survey and upon a few preliminary experiments.

3. The concept of image quality

3.1 Image interpretability

An imaging system translates an object's spatial and spectral radiation characteristics into a density (or color) distribution over an image plane. The imaging process is an information channel, the end product of which is a memory for information. Information is retreived from the memory by interpretation of the image. The amount and quality of information that can be retreived, is dependent upon the quality of the image, expressed in a relevant way with respect to the human ability to interpret the image. Therefore, the interpretability of images will be the basis of a discussion of the relevance of different quality descriptors. It is at this stage important to bear a few things in mind:

- Image interpretation is a very complex process, the result of which is dependent not only of the detail definition of the image and of the visual properties of the eye, but also of the psycho-physical interactions between the eye and the brain.

- Image interpretation, e.g. mapping, is normally performed stereoscopically. The mechanism of stereoscopic vision is in this respect not fully known. According to Konecny et al (1982), features could be interpreted stereoscopically with only half the resolution required to interpret the same features on single images. Trinder (1986) claimed that the boundaries of the objects rather than the areas are the important properties used by the visual system to detect the stereoscopical disparity. These questions, as well as the effect of different image quality for the left and the right eye, should be further investigated.

- Most interpretation tasks are based upon a priori knowledge or expectation of the existence, location or shape of objects. This may affect object detectability and recognizability.

The two main approaches to image quality description are the subjective and the objective one. Both are at the same time weak and strong. Combining the strength of the methods while avoiding their weaknesses has been the more or less explicit aim of many image quality studies.

3.2 Image quality demands

Aerial photographs are used for many different applications, utilizing many different types of instruments. This implies that image quality demands may differ considerably. One reason for this is that image quality may mean different things for, and is described in different ways by, different cathegories of image users. For example, stereo-plotting in analog or analytical plotters, photo interpretation in instruments with high image magnification, orthophoto production, image coordinate measurement for block triangulation, image digitizing for automatic image matching measurements, and digital mono-plotting from image coordinates measured on a digitizing table all may put different demands on the quality of the images used. An investigation into the nature of these different quality demands is necessary.

4. The influence of the photographic process

4.1 Exposure and dodging

The photographic process has a determining influence on image quality. The reason for mentioning this again is the availability of aerial mapping cameras with forward motion compensation (FMC), which have increased the possibilities to optimize film exposure and made fine-grained, low-sensitivity films possible to use also for low altitude aerial photography. As reported before (Boberg 1986), image quality as assessed by the National Land Survey of Sweden (NLS) has improved using this technique.

Image resolution varies with exposure (Brock 1952, Corten 1960, Zeth 1984). This restricts the useful log brightness scale and the exposure latitude. Exposure should bring also the darkest parts of the subject to the straight part of the D/logH curve. Corten used Dmin = Dfog + (0.1-0.3), while Graham and Read (1986) indicate best resolution at D = Dfog + (0.4-0.6) and Istomin (1961) at D = Dfog + (0.6-1.0). Gerencser (1979) recommends Dfog < 0,2, Dmin = Dfog + (0,2-0,4) and ΔD = (0,8-1,5). The NLS uses Dmin = 0.5 and Dmax = 1.4, giving a density range of ΔD = 0.9, according to practical experience.

Istomin (1961) used the amount of exposure range, within which a photographic material provides for a pre-determined resolution, as an information capacity criterion.

Corten (1960) investigated the resolution of transparencies and paper prints of different gradation and exposure, as function of the copy density. With the aim to optimize the copy resolution, he recommended the use of density compression (dodging) to an apparent density range of ΔD = = 0.6 and a high contrast copying material. Since then, the silver content of photographic materials has decreased, and air pollution has changed the haze conditions. Tone reproduction studies of today's conditions would therefore be of great interest.

4.2 Photographic tone reproduction

Tone reproduction diagrams show how densities and contrasts are transferred from the object to the positive copy. A limited, preliminary measurement and construction of photographic tone reproduction curves of aerial photographs of the NLS quality marks 7 (which is the very best), 5 and 3 (which is barely accepted) has been made (see Figure 1). The DlogH curves were constructed with the help of the step tablets in the Zeiss Jena LMK camera frame. The solid parts of the curves represent the object range. From the graphs, the values in Table 1 were derived.

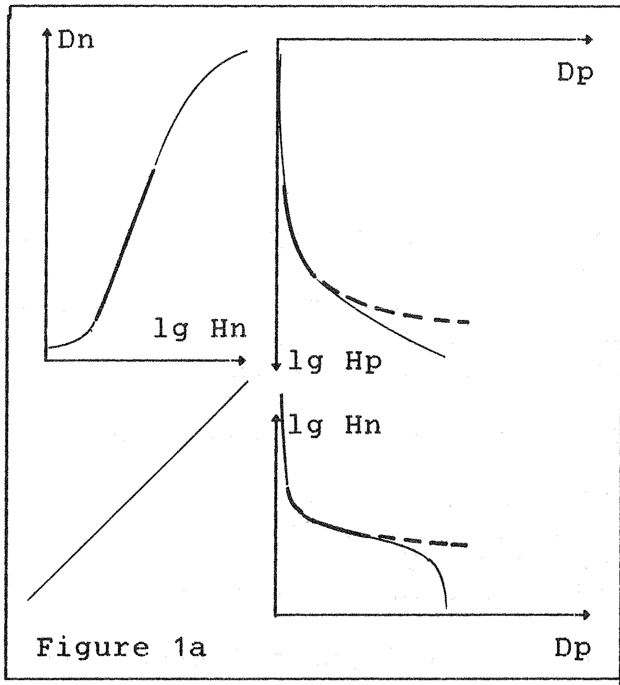


Figure 1a

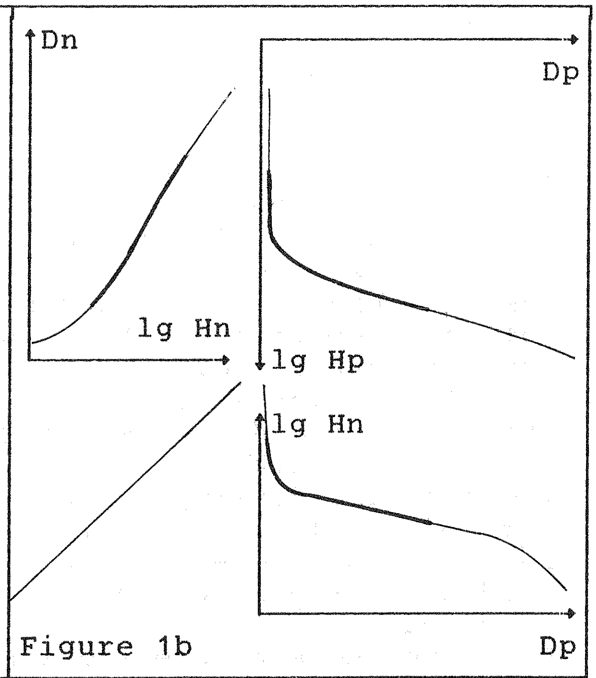


Figure 1b

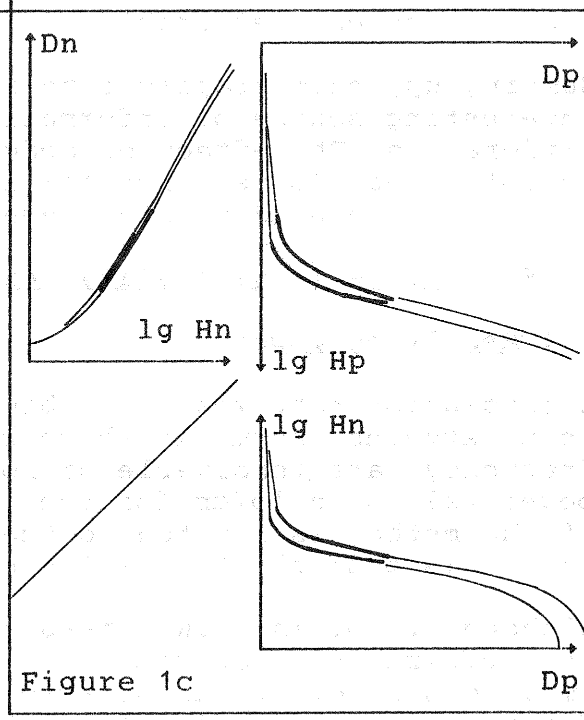


Figure 1c

Figure 1 Tone reproduction diagrams for NLS aerial strips of different subjective image qualities (NLS#).
Aerial camera: Zeiss Jena LMK 15/23 with IMC

Fig	NLS#	Film	Flh m
1a	7	2412	2300
1b	5	2405	1600
1c	3	2405	1500

NLS#	ΔlgHn	ΔDn	ΔDp	ΔDp/ΔlgHn	γn	γp	γnγp	NLS#/ΔDn
7	0.8	1.3	1,6	2.0	1.8	1.6	2.9	5.4
5	0.9	1.1	1.5	1.7	1.1	3.3	3.7	4.5
3	0.7	0.7	1.0	1.4	1.1	3.9	4.4	4.3
	0.5	0.5	1.1	2.2				6.0

Table 1. Photographic density values for negatives of different subjective image quality (NLS#), and for the corresponding diapositives.

The following, very preliminary conclusions can be drawn from the figures and the table:

- The negatives are nicely linear, but only a limited part thereof is used - apparently the best part from resolution point of view, though!
- Negative gradation is relatively high for film 2412, which could be expected.
- Negative density difference ΔDn varies with NLS quality mark approximately as NLS# = 5 x ΔDn.
- Positive gradation is very high for copies from film 2405. As automatic dodging has been used, diminishing the positive exposure range ΔlogHp, this is in accordance with Corten's (1960) recommendations.
- The gamma product is very high, and inversely proportional to the quality mark.
- The negative density range partly falls below the linear part of the (D/logH)p curve, resulting in a similar result for the final tone reproduction curve. This is especially evident in the NLS#5 curve, and is probably the result of copy underexposure.
- An increased exposure and a shorter development time of the copies seems preferrable.

Summing up, photographic tone reproduction curves seem to be an interesting source of information on reasons for image quality differences. The effect of dodging on them should be further investigated. To use them for image quality assessment, they have to be coupled to other quality descriptors.

5. Subjective image quality descriptors

5.1 Resolving power

A resolution test means a subjective decision whether imaged test patterns, normally three-bar targets of increasing line frequency, are resolvable or not. The test leads to resolving power values in lp/mm for the imaging system. The subjectivity of the method is limited to the judgement whether the three-bar target is possible to resolve or not.

Although resolving power measures the combined effect of all the information transmitting or degrading parameters in the imaging system, it has no absolute or permanent significance, derived from physical optics, but simply means the ability of a camera system to reproduce fine details of the type of the test object (Brock 1952,1969).

Resolving power is an uncomplete description of image quality, although it is reported to show strong correlation to subjectively experienced aerial image quality. Resolution is different for objects of different shape, size, arrangement and contrast (Rosenberg 1971, Overington 1974). A high-contrast pattern limits the usefulness of the method, as the landscape normally imaged in aerial photographs consists of different, mainly low contrasts. Brock (1952) is clearly in favor of an annulus target, with target contrast ΔD = 0.2. Further, as the method presupposes the imaging of man-made targets, it is not applicable for routine airphoto production.

Thus, resolving power is useful for comparison of the performance of different imaging systems, but not for quality description of images in hand.

5.2 Quality mark assessment

The National Land Survey of Sweden (the NLS) performs nearly all aerial photography for mapping purposes in Sweden. The assessment of image quality must not delay the production, and therefore must not be too elaborate. For this reason, a relatively simple, subjective image quality assessment method is used. Every photographed strip, as a unit, is given a subjective quality mark in a seven-step scale, expressing the mean photographic image quality of the negatives in the strip. The assessment is based upon criteria on illumination, gradation, haze and image motion, using a few density readings and ocular examination with a magnifying glass. The method was explained more in detail in (Boberg 1986).

The method works fairly well, according to experience on how the images are received by the users. However, a more thorough investigation into the relevance of the quality marks for different cathegories of image users is needed. E.g. Biedermann (1978) has experienced that quality judgment depends on object detail and contrast and on the attitude of the observer to the imaged subject. Also, the NLS method should be calibrated in a more objective way than via a yearly review of the routines by the personnel responsible for the assessment. Measurements of objective parameters and calculation of quality descriptors of the negatives would thus improve the method.

6. Objective image quality descriptors

6.1 Image acutance

Acutance is a measure of image sharpness, using the gradient of a microdensitometer trace across an edge, which has an infinitely steep gradient in the object. Istomin (1961) reports of the acutance measure used by Jones and Higgins, namely the ratio of the RMS mean gradient of an edge response curve and the contrast of the edge, expressed as log exposure height of the edge. Contrast as information carrier for image interpretation is however better expressed as a density difference, ΔD = Dmax-Dmin. If density is used instead of log intensity, and the photographic process is considered practically linear, acutance could be expressed as

$$A = \frac{1}{N} \sum_{n=1}^{N} \left(\frac{dD}{dx}\right)^2 \Delta D$$

in which the gradient dD/dx might be determined as the mean slope, like a gamma value.

Acutance is a complement to resolving power. The two might even be complementary, as high acutance often means low resolution and vice versa. The definition of acutance (in logH values) is related to edge gradient analysis for modulation transfer function determination.

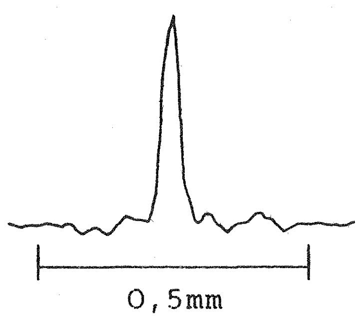
6.2 <u>The modulation transfer function</u>

If an object is regarded to be composed of sinusoidal line patterns of different frequencies, an analysis of the degradation of the imaged object contrast, as a function of the spatial frequencies of the composing line patterns, leads to the modulation transfer function (MTF). This means, that if the line pattern frequencies are regarded as a measure of the sizes of image details, the MTF will describe the reproduction quality of imaged object details. Although this is not generally true, the MTF has been shown to be a good starting point for derivation of image quality measures.

MTFs could be determined for each step of the imaging process, as well as for the process as a whole. In our case, this is of certain interest, because the positive copy is the end user product, while the image quality assessment should refer to the negative. Knowing the MTF of the copying process, it could be reduced from the MTF of the copy, to get the function for the negative.

MTFs of images in hand must be determined from microdensitometer traces of natural image details. Normally, Edge Gradient Analysis (EGA) of natural sharp edges has been used for this purpose. Gerencser (1972,1974,1976) has gathered some experience from the EGA method, using the graphical method of Scott et al (1963) as well as an automated system. The accuracy requirements on the microdensitometer's length measurement was found to be very high. Using known contrast panels, the MTF curves don't have to be normalized to 100% modulation at 0 lp/mm frequency, which makes it possible to evaluate the contrast reduction of the atmosphere. This experience is of interest for this study, as an image quality descriptor should encompass the whole imaging chain. The possibilities to use natural, known contrasts should be investigated. Gerencser expected difficulties in finding suitable edges and to define the trace length.

Especially in large scale aerial photography, the definition of the imaged object edge has to be considered carefully. Gerencser (1976) as well as Ackermann (1986) encountered great problems in finding suitable edges. In some cases, line objects for direct measurement of the Line Spread Function (LSF), may be easier to find, e.g. in built-up areas with road-markings. Figure 4 shows an example of this. Generally, image scale must be considered when suitable targets are chosen.



<u>Figure 4.</u> An example of a microdensitometer trace across a white marking line on an airport runway. Image scale 1:15 000

A basic prerequisit for MTF analysis is that the imaging system is linear. The photographic emulsion is principally not, however. Welander (1968) has put together results that show that within reasonable limits of contrast, gamma product, frequencies and negative exposure range, the photographic process may be considered linear for practical purposes. Hendeberg (1967) showed that in the absence of adjacency effects, the MTF of the negative is independent of contrast and exposure within wide ranges. These results are of great importance if the total system result is to be assessed and related to subjective image quality. The negative quadrant of the tone reproduction curves shown above seems to fulfill the linearity well, but the positive quadrant does not, if copy exposure is insufficient.

6.2.1 MTF condensation

A MTF is normally expressed graphically. Moreover, many MTFs are needed to fully describe all quality aspects of an image. It is therefore natural, that much effort has been devoted to try to condense the MTF into a single figure of merit. Computing the logarithm of the integrated area beneath the curve, or of a part thereof, has been the most commonly used method.

A Threshold Modulation (TM) curve for a specific emulsion and development describes the image modulation needed to resolve an object, as a function of object spatial frequency. The intersection of the appropriate MTF and TM curves gives a frequency threshold, which is a measure of the system resolving power. This threshold is often used as upper limit of the MTF condensation integration. The method is suited for prediction of the performance of imaging systems (Rosenbruch 1986).

A more promising form to use with images in hand is based upon integration of a frequency weighted MTF curve, as described below.

6.3 Granularity

Image interpretability is strongly influenced by graininess, especially when detail contrast is low. This could be expressed as a signal-to-noise ratio (S/N), where the object image contrast is the signal, and the graininess is the noise. According to experience, S/N = 4 is needed for object detection, and S/N = 10 for recognition (Welander et al 1972).

Graininess is measured as granularity, defined as the standard deviation σD of a large number of density readings, each covering a very small area of the image. Typically, a 24μm diameter aperture is used, but Gerencser (1972) used a 1,6x200μm microdensitometer slit. The Kodak method for granularity measurement, leading to RMS granularity ($10^3 \sigma D$), is supposed to indicate the impression of graininess perceived at 12x magnification of the sample. The Selwyn granularity measure G is standardized to a unit measuring area a as

$$G = \sigma D \sqrt{2a}$$

Granularity increases with the square root of the density value, and roughly in proportion to the contrast of the printing material (Welander 1968). This reduces its value as quality measure for images at hand.

Granularity could be investigated with the help of Fourier analysis, in which noise strength is expressed as a function of its spatial frequency. This form of description, a Wiener spectrum, is closely related to the MTF. Trinder(1980) has investigated this in detail, and found it important to consider the S/N in attempting to relate visual performance to image quality parameters.

7. Relations between objective and subjective criteria

Attempts have been made to correlate physically measurable figures of merit to subjectively perceived quality of photographic images. Biedermann (1967, 1978) tried in his thesis to relate subjective assessments of image quality via weighting functions to measurements of modulation transfer functions of the images. Biedermann's study concerned mainly "normal" images like landscapes, portraits etc, but also an aerial photograph, at a viewing distance of 40cm. He found good correlation with subjective quality G (running from 1 "excellent" to 7 "useless") using the following function of MTF, denoted M(v):

$$G = C_1 + C_2 \log \int_0^\infty \frac{M(v)B(v)dv}{B(v)dv}$$

Different weighting functions B(v) were tested, of the form

$$B(v) = \frac{1}{1 + (F\ v)^4}$$

where F could be varied. Thus, image quality was related to the logarithm of the area under the weighted MTF.

Rosenbruch (1986) proposes the agreement on a standard weighting function. However, frequency weighted MTFs, or derivatives thereof, as descriptors of aerial image quality, require weighting functions adapted to the image application and to the image evaluation situation at hand. As a base for the choice of frequency weighting functions, a study of the importance of different frequencies and contrast distributions for different applications is needed. The base for this should be an investigation into image quality demands of different cathegories of aerial image users, e.g. stereo operators, image interpreters, topographic map makers, forest taxators and physical planners. The judgment should be done in their real working situation, with the images oriented in evaluating instruments.

The result of such a study might be a suggested calibration procedure, or a replacement method, for the NLS subjective image quality assessment method.

8. References

Ackermann F. 1986. Improvement of Image Quality by Forward Motion Compensation - A Preliminary Report.ISPRS Archives Vol 26, Part 1. ESA SP-252, Noordwijk, the Netherlands.

Boberg A. 1973. Några synpunkter på begreppet bildkvalitet (Some Views on the Concept of Image Quality. In Swedish). Svensk Lantmäteritidskrift 3:1973, Stockholm. 22p.

Boberg A. 1986. Image Quality Problems in Practical Aerial Photography. ISPRS Archives, Vol 26, Part 1. ESA SP-252, Noordwijk, the Netherlands. Fotogrammetriska Meddelanden No 53, Department of Photogrammetry, the Royal Institute of Technology, Stockholm. 8p.

Biedermann K. 1967. Ermittlung des Zusammenhanges zwischen der Subjektiven Gute und den physikalischen Eigenschaften des Photographischen Bildes. Phot Korr, Band 103, Nr 1,2,3, pp5-14, 26-31, 42-48.

Biedermann K. 1978. Subjective Quality and Modulation Transfer Function of Photographic Images. Physikalisch-Technische Bundesanstalt, PTB-Bericht PTB-Opt-7, pp 96-104. Braunschweig.

Brock G.C. 1952. Physical Aspects of Air Photography. Longmans, Green & Co, London. 267p.

Brock G.C. et al 1969. Photographic Considerations for Aerospace. Itek Corporation, Lexington, Mass. 122p.

Corten F.L. 1960. Image Quality in Air Photography. Degradation and Conservation of Details in Positive Copying. ITC Publications, Series A/B, No 5. Delft, 17p

Gerencser M. 1972. Comparative MTF Tests of Entire Aerial Photographic System in Black-and-White and Color. ISP Archives, Vol 19, Commission I. Ottawa, Canada. 20p.

Gerencser M. 1974. Comparative MTF Evaluation of Aerial Photographic Systems by Simultaneous Using of RP, Spread Functions, MTF's, CTF's and Information Capacity. ISP Archives, Vol 20, Commission I, Stockholm, Sweden. 19p.

Gerencser M. 1976. Automated Edge Gradient Analyse (EGA) for Testing Photogrammetric Systems. ISP Archives, Vol 21, Commission I. Helsinki, Finland. 14p.

Graham R, Read R.E. 1986. Manual of Aerial Photography. Focal Press, London. 346pp.

Istomin G.A. 1961. Objective Evaluation Criteria in Image Photographic Quality. XIII Congr Intern della Tecn Cinematografica, Torino Sept 1961, pp 83-89.

Jensen N. 1968. Optical and Photographic Reconnaissance Systems. John Wiley and Sons, New York. 211p.

Konecny G, Schuhr W, Wu J. 1982. Investigations of Interpretability of Images by Different Sensors and Platforms for Small Scale Mapping. Proceedings ISPRS Commission I Symposium, pp 11-22.

Overington I. 1974. Visual Efficiency - a Means of Bridging the Gap between Subjective and Objective Quality. Proc SPIE, Vol 46, pp 93-102.

Rosenberg P. 1971. Resolution, Detectability and Recognizability. Phot Eng Vol 37, No 12, pp 1255-1258. December 1971.

Rosenbruch K-J. 1986. OTF-based Quality Criteria for Aerial Cameras and Imaging Systems. ISPRS Archives, Vol 26, Part 1. ESA SP-252, Noordwijk, the Netherlands.

Scott F, Scott R.M, Shack R.V. 1963. The Use of Edge Gradients in Determining Modulation-Transfer Functions. PS&E, Vol 7, pp 345-349.

Trinder J.C. 1980. Aerial Film Granularity and Its Influence on Visual Performance. ISPRS Archives, Vol 23, Commission I, pp 167-181, Hamburg.

Trinder J.C. 1986. Precision of Stereoscopic Height Measurements. Phot Eng & RS, Vol 52, No 1, pp 75-79.

Welander E. 1968. Modulation Transfer Functions for Photographic Emulsions and Their Applications in Aerial Photography. Meddelande Nr A37, the Geographical Survey Office of Sweden, Stockholm.

Welander et al 1972. Photographic Modulation Transfer Functions. Report by ISP Commission I WG. ISP Archives, Vol 19, Part 1. Meddelande Nr A39, The Geographical Survey Office of Sweden, Stockholm.

Zeth U. 1984. Exposure Metering and Control in the LMK Aerial Survey Camera System. Sensing the Earth, Photogrammetric Instruments 1984/II, VEB Carl Zeiss Jena, pp 21-26.