

9950-767
ARKANSAS
REMOTE SENSING
LABORATORY

A STUDY OF IMAGE QUALITY FOR RADAR IMAGE PROCESSING

ARSL TR 82-1



Robert W. King

Verne H. Kaupp, Associate Professor
Department of Electrical Engineering

William P. Waite, Professor
Department of Electrical Engineering

Harold C. MacDonald, Professor
Department of Geology

University of Arkansas
Fayetteville, Arkansas 72701

JPL Contract No. 956093

(NASA-CR-169907) A STUDY OF IMAGE QUALITY
FOR RADAR IMAGE PROCESSING (Arkansas Univ.)
43 p HC A03/MF A01 CSCL 17I

N83-18980

Unclas
G3/32 08722

This work was supported in part by the Jet Propulsion
Laboratory, California Institute of Technology
sponsored by the National Aeronautics and Space
Administration under Contract NAS7-100.

UNIVERSITY OF ARKANSAS

FAYETTEVILLE, ARKANSAS 72701



CONTENTS

<u>Chapter</u>	<u>page</u>
I. INTRODUCTION	1
Purpose	1
Scope	1
II. LITERATURE REVIEW	3
The Threshold Quality Factor (TQF) Metric	4
The Radar Threshold Quality Factor (RTQF) Metric	6
A Nonlinear Model of the Human Visual System	7
Radar Image Simulation	12
A Tutorial Review of Synthetic-Aperture Radar	13
Minimum Average Distortion	13
The Signal-To-Noise Ratio Threshold (SNRT) Metric	14
Quality Metrics of Digitally Derived Imagery	16
A Utilitarian Approach to Image Quality Metrics	17
Mean Square Error Metrics	19
Geometric Distortion and Noise in SAR Imagery	20
SAR Image Quality Considerations	21
III. INTERPRETATION ELEMENTS EXTRACTED FROM THE LITERATURE	25
Tone	25
Pattern	26
Shape	26
Size	27
Shadow	27
Texture	28
Site	28
Context or Association	29
Resolution	30
IV. RESEARCH APPROACH	31
Developing Metrics	33
Metric Selection	33
Image Quality	34

V. **PROGRESS TO DATE** 36

BIBLIOGRAPHY 38

Chapter I

INTRODUCTION

1.1 PURPOSE

This study is being performed to determine a few metrics that can be used to assess the quality of radar images, particularly synthetic aperture radar images. The original purpose of this study was to evaluate the image interpretation elements of tone, shape, pattern, size and shadow, and construct a compendium of simulated radar images exhibiting the characteristics of radar images containing these recognition elements. In the face of reduced funding, the present purpose is to re-evaluate the image quality metrics from the photographic image field in the context of synthetic aperture radar (SAR) produced imagery.

1.2 SCOPE

The scope of this project has been reduced from that envisioned in the original proposal. This reduction has been accomplished primarily by reducing the depth planned for the study and somewhat redirecting the thrust in order to maximize the return for the resources expended. Those image quality metrics most commonly found in the literature apply almost without exception to conventional photography. We

will evaluate the application of several of these metrics to SAR produced imagery.

Chapter II

LITERATURE REVIEW

We have explored the unclassified literature of image measurement for both radar derived images and for photographic images. Measurements or ratings of image quality have generally been found to be divided into two types: subjective, qualitative ratings by human observers of relative or absolute image quality or various characteristics believed to be associated with image quality; and quantitative measurements of image characteristics (normally a different set). By "subjective" or "qualitative", we mean those forms of image evaluation that: (a) establish only a relative ranking of comparable images in accordance with some criterion, and (b) require a human observer to perform the comparison. By "objective" or "quantitative", we mean a form of image evaluation that: (a) can be performed in a mechanized way without the intervention of a human observer, and (b) that produces a numerical result which can be used to rank-order the measured images such that the sequence correlates well with human produced rank ordering of the same images.

A major problem with these approaches is that there is little agreement as to the elements or characteristics that should be considered in defining whether an image is of good

quality. The subjective ratings or measurements tend to be found in the literature of the human factors (ergonomics), psychology, and photo interpretation fields, and the quantitative measurements tend to be encountered in the image reproduction and pattern recognition literature. On the one hand, we have attempts to define image quality in the sense of how the image is experienced by a human observer, and on the other hand, we have attempts to define it in the sense of stating that one device provides greater image quality than another because it produces images having one or more characteristics measurably better than the other.

There were several bodies of work that have been useful in understanding the nature of the problem: they are summarized by image quality metric in the following sections.

2.1 THE THRESHOLD QUALITY FACTOR (TOF) METRIC

Charman and Olin [5] expressed dissatisfaction with the available metrics for rating aerial camera systems. They observed that the resolving power (RP) was the most widely used criterion since it seemed to be the easiest to obtain. One simply flew the camera system over a prepared low contrast resolution target (usually sets of bars in a variety of widths and separations). Resultant photographs were then processed and evaluated by human observers. Charman and Olin noted that the results of such evaluations were satisfactory for rank order comparison of systems but they did

not yield quantitative results that could be compared when different resolution targets were used. They then proceeded to examine the class of metrics developed from the optical transfer function.

Briefly, the performance of a linear optical system is described by its optical transfer function (OTF), which relates how a system responds in both magnitude and phase as a function of the spatial frequency and the orientation of objects in the scene. The related modulation transfer function (MTF) omits the phase information.

Charman and Olin criticized OTF derived metrics on the grounds that they did not consider the effects of grain in the photographic emulsion or the non-linear response of the human observer. They mentioned that the effect of grain could be expressed in the spatial frequency domain by the Wiener, or noise-power spectrum of the granularity. They pointed out, however, that attempts to obtain a unified approach to the whole chain of processes had foundered on the marked non-linearities of the visual system.

Their approach was to observe that there appeared to be a minimum modulation within a grain free image field that could be perceived by the eye. Using this datum, they calculated an effective scanning aperture of the eye and an RMS density deviation due to noise in the visual system which led to a figure for the threshold viewing signal-to-noise

ratio. They then extended these arguments by including the noise due to photographic grain to arrive at a total RMS noise for the combined visual and photographic system. Using this, they calculated a threshold scene modulation for the combined optical, photographic and visual perception system.

As they stated, they made a number of rather sweeping simplifying assumptions in deriving the threshold equation, and they asserted that only experimental evidence could justify the utility of their TQF approach. They performed such an experiment by measuring for a single camera system the RP, MTF and TQF and comparing the results. The results showed these metrics to be in agreement as the exposure and development times were varied. However, they concluded with the statement, "Definite judgement on the usefulness of the TQF must await the results of further tests."

2.2 THE RADAR THRESHOLD QUALITY FACTOR (RTQF) METRIC

Mitchel [28] used the same general approach to the problem for the development of a Radar Threshold Quality Factor (RTQF) as a figure of merit uniquely tailored to SAR system images. Essentially he just re-defined the TQF to apply to synthetic aperture radar systems. He included the effects of the image impulse response, mainlobe width (azimuth and range), system noise, clutter noise, target contrast with respect to its background, the display geometry, the viewing

distance, and the human visual response. Factors not included were geometric fidelity, linear dynamic range, and sidelobe levels. While this work is very useful, it is not definitive (for the same reasons that Charman's and Olin's TQF was not definitive) as he pointed out. However he was able to validate the trends indicated by the RTQF by experimentation with interpreter studies conducted with a radar holographic viewer he built.

2.3 A NONLINEAR MODEL OF THE HUMAN VISUAL SYSTEM

Hall and Hall [18] proposed in 1977 a new model for the human visual system (HVS) which supported the general hypothesis that the HVS is composed of spatial frequency channels. They cited experimental results from psychophysical and neurophysiological investigators which suggested that the visual cortex performs a two-dimensional spatial decomposition of subdomains of visual space. They reasoned that a rational method of constructing a HVS model would be to begin with a simple linear model and then modify the model appropriately to accord with observed data as each initial assumption was challenged. They began constructing a first approximation model of the HVS by considering the system to be linear, isotropic and time and space invariant. Other assumptions used were that the system is monocular, monochromatic and photopic. All these assumptions they knew to be invalid, but they started with them and relaxed each at the appropriate time as they developed their model.

The HVS response to changes in intensity was clearly non-linear. The sensitivity of the system to a rotated contrast grating was known to be a function of both the spatial frequency of the grating and its angle of rotation, hence the system was anisotropic. The HVS was known to be spatially variant and non-homogeneous in both optics and receptors, although they pointed out that optical spatial invariance is a good assumption near the optic axis and that the receptor distribution is relatively homogeneous in the foveal region. As a further argument for homogeneity, they cited a paper by Davidson [8] which suggested that certain non-homogeneous systems are functionally self-homogenizing. The notion of temporal invariance was clearly violated by the evidence for both sustained and transient channels in human vision. However the models discussed in their paper did not consider temporal responses; hence temporal homogeneity was not a factor.

Since they did not propose to deal with depth perception, the only remaining differences between binocular and monocular vision were the difference in absolute threshold (lower by the square root of 2 for the binocular case) and resolution (7% better for binocular vision) reported by Campbell and Green [3]. Because of these simple relationships, they argued that a monocular model could readily be extended to the binocular case, and thus proceeded to develop a monocular model.

Monochromatic vision they defined as the inability to distinguish differences in hue. They argued that because the illumination sources used in the experimental protocols of interest were of a constant spectral content (usually white or blue-green light), there were no hue variations in the stimuli. (This means they reasoned that the spectral distribution of illuminating light energy within the electromagnetic spectrum did not change from one intensity level to the next.) For this reason, they accepted restriction of their model to the monochromatic case. As for the assumption that the system was photopic, this was clearly not in agreement with the observations of the "brightness constancy" phenomenon. In general, these observations were that the perceived brightness of an object tends to remain constant despite variations in the illumination falling upon it.

Having brought out all these objections to their assumptions, they proceeded to build an initial simple model consisting of cascaded low-pass and high-pass filters. They stated that the low-pass filter characteristics resulted from the interaction of several mechanisms. High-frequency response was limited by several factors: the optical characteristics of the eye including the pupil size, and the size and density of the photoreceptors and their neural summation networks. In addition, light scattering within the aqueous humor was thought to be an increasing function of

frequency. Next they proceeded to point out that this model did not account for the "brightness constancy" phenomenon and that to do so would require introducing a non-linearity into the system.

Next they noted that the mechanism by which a quantum of light stimulating a photoreceptor produces an electrical impulse from an accompanying was not known exactly. However they cited the work of Duortes [10,11] in measuring the electrical properties of the nerve cells of the eye of the Limulus (the horseshoe crab) and Rushton's conclusion [35] from that evidence that the resistance of the cell membrane was proportional to the logarithm of the total light incident upon the receptor. Furthermore, Rushton had concluded that the relationship between the membrane potential and the frequency of impulses was linear; hence the frequency of nerve impulses was a logarithmic function of light intensity which was in keeping with the seemingly general rule for stimulus-response relationships of physiological-mechanical receptors and sense organ transformations.

Hall and Hall used this data to construct a pair of models which consisted of two blocks cascaded: a logarithmic intensity detector and a MTF (bandpass filter) section which was equivalent to the over-simplified initial model. The question then became; "In which order should these blocks operate on the information?" They argued, from physiologi-

cal evidence and the incompatibility of one of the models with the brightness constancy effect, that the correct order was that of the logarithmic detector followed by the MTF block.

Next, they pointed out that this model did not predict the observations that indicated a nonlinear distortion of signals at high, but not low spatial frequencies. To satisfy this evidence, they split the MTF bandpass filter into its equivalent high- and low-pass filters. They placed the low-pass filter prior to the logarithmic detector because of the physical locations of the mechanisms responsible for this filter. They then introduced the Furman-Cornsweet backward inhibitor neural network model [16,6] for the interaction of adjacent photoreceptors. This model represented the net output of a receptor as the sum of its inherent output less an inhibitory coefficient times the inherent output of adjacent receptor(s). Assuming that there was no self-inhibitory action and that inhibitory interaction was an exponentially decreasing function of the distance between receptors, they derived a transfer function which had the frequency characteristics of a high-pass filter.

They showed that the order of the visual system model elements in the low-pass, nonlinearity, high-pass sequence satisfies various behavioral observations and that other arrangements do not. The major significance of this model was

that the bandwidth of the visual system decreased as the contrast increased. Thus the system appeared to maximize the signal-to-noise ratio while attempting to maintain a constant "perceptual" spatial-frequency fidelity. They pointed out that this model has several applications as a preprocessor for image processing, image coding, pattern recognition, and scene analysis.

2.4 RADAR IMAGE SIMULATION

Holtzman, Frost, Abbott and Kaupp [21] defined the principal concepts and benefits to be derived from computer simulation of radar image generation. They introduced a "Point Scattering Model" with four natural subdivisions: (1) the imaging model, (2) the radar geometrical propagation model, (3) the ground model, and (4) the reflectivity model.

The reflectivity model was simply the series of the differential scattering cross section (σ_0) curves which were supplied for each different category of terrain. Each point on the ground was supplied with a category descriptor (1,2,3,...) designating which of the reflectivity curves to apply for that particular patch of ground. The ground model was an elevation above or below some datum supplied for each point in a rectangular grid. The radar geometrical propagation model uses the relief of the ground model to compute those areas of the terrain that will be in shadow from the assumed location of the radar platform and computes the

range and power returned for each data point in the ground model. Finally the imaging model accounts for the translation of power received for each element of radar range through a detection (square law or other appropriate detection rule) process and into a gray level in a display or photographic image.

This paper describes the structure for the radar simulation implemented at the University of Arkansas and gives a conceptual basis for following the flow from ground truth elevation and scattering properties to a radar image of that terrain from a specific point in space.

2.5 A TUTORIAL REVIEW OF SYNTHETIC-APERTURE RADAR

Kiyo Tomiyasu's discussion [37] of the design and organization of synthetic aperture radar systems together with the ambiguities, signal processing requirements and sources of error has proved to be a succinct and masterful exposition of the subject. While not discussing the subject of SAR image quality directly, the paper has proved to be a most useful reference for the electromagnetic, geometric and platform dynamics relationships involved in the production of SAR imagery.

2.6 MINIMUM AVERAGE DISTORTION

A 1980 paper by Charles Hall [19] did not deal directly with the subject of image quality. Its value lies in the quantization of his previously developed HVS transfer function. The paper dealt with the idea of using this characteristic as a preprocessor in performing data compression for image transmission. As part of this development, the author discussed the work of Shannon which established rate distortion theory. In this discussion, he mentions the idea of a minimum average distortion or fidelity criterion as a measure of agreement between the source and the system output as specified by the user. This fidelity criterion was not further explained in the paper. However, the minimum average distortion is not a single generally accepted metric, but instead is whatever metric may be useful - usually MSE, but a variety of others have been used - and no mathematical restrictions are imposed by rate distortion theory on the choice.

2.7 THE SIGNAL-TO-NOISE RATIO THRESHOLD (SNRT) METRIC

A 1981 paper by Blumenthal and Campana [1] discussed the Johnson Criterion metric [22] which has been used in the comparison of electro-optical systems (photocathodes, television camera tubes and infrared detectors) in military applications, and introduced a metric called the signal-to-noise ratio threshold (SNRT). The Johnson Criterion was based on

the limiting resolution of a system as measured under the same ambient conditions as encountered in the field. According to the Johnson concept, if at any given range the minimum target dimension subtends a specified number of cycles of a periodic pattern resolvable under the conditions of the test, the target could be discriminated at the corresponding level. The specific number of resolvable half-cycles across the minimum dimension required to detect, recognize or identify a target is determined empirically from field experiments. This metric has been in use in the military electro-optical area since 1958.

The criterion had been generalized by Moser [31] by specifying the number of pixels that the target area must subtend in a two-dimensional image before the task of detection, recognition or identification could be accomplished. The authors demonstrated an instance in which both the Johnson and Moser criterion selected a grossly degraded image as the best. From this apparent contradiction, they argued that the difficulty was in the signal-to-noise ratio under which the images were obtained and that a metric which did not account for the SNR conditions was insufficient.

Using Schade's noise-equivalent passband as a means of describing the MTF with a single number, they generated an array of images in which the SNR, scale factor and MTF were varied in a controlled manner. The images were rank-ordered

by an interpreter study and then the rankings were correlated with several different figure of merit formulations. These were: the Modulation Transfer Function Area (MTFA), the Modulation Transfer Function Power (MTFP), the Signal-to-Noise Threshold (SNRT), and the Johnson criterion (JC).

The authors commented that the correlation coefficient proved to be a very insensitive measure of agreement and that in no case was the coefficient less than 0.9. After performing a RMS evaluation of the quality metrics versus observer performance, the authors determined that the Johnson criterion tended to overestimate the utility of images made at low SNR, while underestimating the quality of high SNR images. The SNRT was shown to be a better metric across the range of SNR variation.

2.8 QUALITY METRICS OF DIGITALLY DERIVED IMAGERY

Burke and Snyder [2] published a paper which dealt with the degradation via hardware and software of a set of ten aerial images (4096 x 4096 pixels) to produce a set of twenty-five versions of each image. Using this data base, they performed an interpreter study that indicated that the addition of noise or blurring of the image degrades the ability of photo interpreters to detect, recognize and identify objects in the scene. They also pointed out that moderate levels of blurring may actually enhance the ability of an interpreter to perform in the presence of high levels of noise.

2.9 A UTILITARIAN APPROACH TO IMAGE QUALITY METRICS

An especially interesting body of work was that done by Frost, et al in 1981 [14]. Their basic concept was that of performing interpreter studies of an array of several SAR images of scenes in which each scene was represented by multiple images having one or more measurable characteristics degraded by varying degrees. The interpreters were to rank order the images in terms of their utility for various interpretation tasks. They obtained statistically significant models to relate the measured image properties to the interpreters' ability to analyze linear features and to evaluate the utility of radar images for vehicle movement potential and activity level. They further found that the relative importance of the measured image properties with respect to image utility varied with image application. This last finding led naturally to the concept of a multi-dimensional image quality parameter space in which one could perform classification tasks. Their approach to the problem of designing an experimental protocol was to apply the Response Surface Methodology techniques expounded by Myers [32]. A brief description of Response Surface Methodology, quoted from Myers is:

"Response surface methodology (RSM) is essentially a particular set of mathematical and statistical methods used by researchers to aid in the solution of certain types of problems which are pertinent to scientific or engineering processes. Its greatest application has been in industrial research, particularly in situations where a large number of variables in some system influence some feature of the system. This feature (e.g.

reaction yield, cost of production, etc.) is termed the response; it is normally measured on a continuous scale and is a variable which likely represents the most important function of the systems, though this does not rule out the possibility of a study of more than one response. Also contained in the system are input variables or independent variables, which have an effect on the response and are subject to the control of the scientist or experimenter. The response surface procedures are a collection involving experimental strategy, mathematical methods, and statistical inference which, when combined, enable the experimenter to make an efficient empirical exploration of the system in which he is interested."

Frost, et al, applied these techniques in their experimental design to minimize the number of observations used and to insure that each degradation was selected such that all the data points obtained had uniform significance. They concluded that five measurable image properties were identifiable: dynamic range, signal-to-noise ratio (SNR), image spatial bandwidth, geometric fidelity and root-mean-square error (RMSE). These metrics were determined to be linked to independent characteristics of the radar image and were either directly or indirectly related to many image quality parameters proposed in the past.

They found that they did not obtain statistically significant regression equations relating natural area features and individual man-made targets to the quality of radar images as judged by human interpreters. A possible explanation was advanced that the interpreters did not use uniform criteria for these response categories owing to the complexity of the these image features. However, for the simpler

categories, (i.e. linear features and relative image quality ranking), statistically significant regression equations could be estimated. Another finding was that different image metrics assumed varying levels of importance as the response category was changed. For example, bandwidth and the signal-to-noise ratio were the most important metrics in estimating the ability of an interpreter to extract linear features from radar images. Dynamic range was predominant for estimating how an interpreter would rank radar images. This last observation has important ramifications for the application of image quality metrics for multi-mission sensor design. That is, the system designer will have to trade-off system performance as a function of the application.

2.10 MEAN SQUARE ERROR METRICS

Charles Hall [20], in a 1981 paper, commented upon the application of human visual system (HVS) models (see sections above) to the problem of evaluating image metrics that have been suggested in the past, and presented a new metric which he called the perceptual mean square error (PMSE). He began by discussing briefly the traditional mean square error (MSE) as a distortion measure, pointing out both its mathematical tractability and its low correlation with human evaluation of the same imagery. He then proceeded with similar brevity to deal with the normalized mean square error

(NMSE), the normalized difference or normalized error (NE), the Laplacian mean square error (LMSE) and a variant on the LMSE which he called the estimated gradient mean square error (GMSE) and which he noted as presenting some formidable analytic problems. Next he pointed out that the GMSE and LMSE are simply the NMSE computed in a transformed space and that one merely selects an appropriate preprocessor to apply to the NMSE. He asked "What more appropriate preprocessor could be selected than a HVS model?" This choice yielded the perceptual mean square error metric which he computed by taking the NMSE of the image pair produced by convolving the test and reference images with the HVS point spread function. He finished with an interpreter study which compared the correlation between human evaluation and the NMSE, LMSE, and PMSE for a series of images with varying degrees of degradation derived from the SPIE GIRL picture. The correlations obtained were .85, .84, and .92 for the NMSE, LMSE and PMSE respectively. He noted that these correlations were obtained in the presence of three types of noise (Gaussian, 8x8 blocking errors and 16x16 blocking errors) and concluded that the PMSE is a more successful attempt at measuring image quality objectively than the other measures discussed.

2.11 GEOMETRIC DISTORTION AND NOISE IN SAR IMAGERY

The papers and reports by Kaupp, Waite and MacDonald [23,25,26,27] provided a clear exposition of the sources of

geometric distortion (layover, foreshortening, etc.) and noise in synthetic aperture images and demonstrated the virtues of computer simulation as a means of producing a series of images of the same terrain with controlled amounts of geometric distortion by variation of the angle of incidence. The signal-to-noise ratio may be controlled over the range that one might realistically expect to encounter and the number of looks that are averaged in synthesizing the effective radar antenna length can be varied. Another potentially useful finding was that, in the simulation, it is possible to de-couple the effects of back-scatter and propagation so that the variation of geometric distortion can be controlled in the output image independently of the speckle in the image.

2.12 SAR IMAGE QUALITY CONSIDERATIONS

Mitchel and Marder in a 1981 paper [29] discussed a number of factors which bear on the image quality of imaging synthetic aperture radar systems. The major factor presented was the dynamic range characteristics of SAR data. In electromagnetic scattering theory, the parameter G which characterizes the roughness of a surface is proportional to the root-mean-square surface roughness and inversely proportional to the wavelength of the incident illumination. For visible light at wavelengths of the order of 50 microns, smooth (specular reflector) surfaces are a rarity. For air-

borne imaging radar with wavelengths typically in the range from 3 to 30 centimeters, many surfaces exist for which $G \ll 1$ and the reflection is specular. When the type of illumination (coherent and uni-directional in radar as opposed to diffuse, incoherent and of a relatively wide frequency range in photography) was included, the result was to accentuate the specular character of radar scattering and the consequent large dynamic range of the signals returned. As a means of contrasting the difference in the dynamic range requirements, the authors provided data indicating that for aerial photography, the dynamic range is typically on the order of 10 dB while it can range from 50 to 90 dB for radar imaging systems. The data presented suggested that synthetic aperture radars could image natural terrain with a dynamic range of 50 to 60 dB while the flat surfaces characteristic of man-made artifacts extend the dynamic range requirements.

Next they argued that as with any antenna or lens, the quality of the image produced is dependent upon the accuracy with the effective aperture is constructed. In the case of synthetic aperture radar where the aperture is synthesized from the motion of a small physical antenna along the line of flight, the position of the physical antenna must be known to a fraction of a wavelength throughout its motion across the aperture. Errors in knowledge of the physical antenna location cause phase errors in the received signal which have the effect of degrading the synthetic antenna

pattern. The specific degradations which arise depend on the form of the phase errors, but the general result is to increase the sidelobes and broaden the mainlobe of the synthetic aperture response pattern. They thus argued that the antenna response is not fixed in the design and construction as is the case with a conventional radar or optical system. It followed that the image quality of a synthetic aperture radar must be regularly evaluated since its constancy cannot be assumed.

Mitchel and Marder next addressed the problem of recording and displaying SAR imagery. They pointed out that with state of the art design, radar receivers and signal processors can maintain a dynamic range of more than 60 dB or more and that this is adequate to handle the range of natural terrain backscatter, but that man-made or cultural features may exceed this capability. The principal problem is that photographic media is simply not up to the task when it comes to recording data with this wide dynamic range - the maximum dynamic range for photographic materials appears to be of the order of 20 to 30 dB. Digital storage methods have the potential to store the full dynamic range, but the usual digital display devices (CRT's) have similar and often more severe dynamic range limitations. They concluded by promoting the holographic viewer as designed and built by Mitchel in the course of his PhD dissertation work.

Chapter III

INTERPRETATION ELEMENTS EXTRACTED FROM THE LITERATURE

Although many articles and texts differ as to the number of basic interpretation or recognition elements, there is general agreement on six: tone or color, shape, pattern, size, shadow, and texture. Three other elements which are often included with these six are site, association or context, and resolution [34]. Image metrics that are best associated with each of these elements in the literature are hypothesized below.

3.1 TONE

In one sense, it can be said that without variations in tone (or image intensity) and/or color, there is no image present to interpret. The dynamic range of intensities, their relative distributions across the capabilities of the imaging system, and their gradients across the image are all major factors bearing on the quality of an image. A great deal of the information content of an image is contained in the change in tone or color from one pixel to the next. We expect then to find that dynamic range, SNR, and spatial bandwidth have some relationship to the element of tone.

3.2 PATTERN

Pattern, or repetition, is characteristic of many man-made objects and natural features. Recognition of patterns is dependent upon discerning groups of similar shapes or commonality of direction in objects observed in an image. Hence we see that the same metrics apply here as apply in the case of tone: SNR, spatial bandwidth (particularly the higher frequency components), and dynamic range. Additionally, it is likely that RMSE and geometric fidelity are related to the element of pattern. Because patterns may be independent of the size of the constituent elements of the pattern in an image, resolution of the imaging system also bears on whether a pattern in a scene may be recognized in an image.

3.3 SHAPE

This element is most heavily dependent upon both the resolution and geometric fidelity of an image to the imaged scene. The difficulty with term "geometric fidelity" when used in this context is that ultimately the degree of geometric fidelity of an image of a scene (or conversely the degree of distortion) is determined by a qualitative or subjective judgement by a human observer. We believe that this is the reason that the various quantitative approaches to measuring geometric fidelity have had only moderate success: they have not accounted for the non-linearities and spatial frequency response characteristics of the HVS.

Distortion (non-linear mapping) as the scene is imaged is inevitable given that a three dimensional scene is being imaged into two dimensions. However, the human visual system is able to extract information from such images so long as there is a regularity to the types of distortions introduced. What seems to be most troubling to this ability is variation in the type and degree of the geometric distortions encountered. Images on a rubber sheet distorted by non-uniform three dimensional stretching suggest themselves as a means of visualizing this problem. Imagine an image fixed upon a rubber sheet lying in the plane and viewed from above. As long as the sheet is uniformly stretched in the X and Y directions by roughly equivalent amounts, little difficulty in recognizing the imaged objects is experienced. However should the ratio Y/X vary appreciably from unity, or should the stretching (scaling) in the X or Y directions vary much from a constant, or should there be other than minor stretching distortions in the Z direction, the ability to recognize objects by shape alone deteriorates rapidly and may be lost altogether. As with pattern, the texture, intensity or color, and contrast (dynamic range) of the image must be sufficient to identify edges or boundaries of shapes. Hence, we must include dynamic range, spatial bandwidth and SNR as likely to have a positive correlation with this element in addition to the primary one of geometric fidelity and resolution. An objective metric for geometric

fidelity seems likeliest to be the PMSE, although the entire class of MSE metrics have demonstrated some success at measuring this image attribute.

3.4 SIZE

This element also depends heavily upon the geometric fidelity of an image to the original scene. Ideally, areas and distances in an image should be related in a linear fashion to areas and distances in the original scene. Also, there is the concept of comparing relative sizes of objects in an image, or measuring them directly from an image when the scale is known, in order to properly identify ambiguous objects in the image. This recognition/interpretation element also requires that the image have discernable changes in texture, pattern, intensity and/or color, and contrast so that the boundaries or edges of an area or linear feature may be detected. The objective metrics that relate most directly to the element of size are the dynamic range, its spatial bandwidth, the SNR, and the geometric fidelity metrics of the image to the imaged scene.

3.5 SHADOW

Shadows are major cues to the interpreter in determining both the shape and height of objects. The arguments concerning geometric fidelity as they apply to size and shape apply here with the further requirement that the scene illu-

mination source be consistent across the image. Changes in apparent direction and number of illumination sources rapidly destroy the value of any information that can be obtained from shadows in the image. The metrics that relate best to this element are dynamic range, SNR, geometric fidelity metrics, and spatial bandwidth.

3.6 TEXTURE

Texture in images is created by tonal repetitions in groups of objects which are often too small to be discerned as individual objects. Texture, the visual impression of roughness or smoothness created by some objects, is often a valuable clue in interpretation. In Chapter 14 of the Manual of Remote Sensing [34], the author comments that texture is an especially useful element for the interpretation of sidelooking airborne radar imagery, but there is some dissent with this view. Texture interpretation is dependent upon the scale of the image relative to the scene, the SNR and the resolution cell size of the imaging system. We do not propose to treat this element in our study.

3.7 SITE

The location of objects with respect to terrain features or other objects is often helpful in the identification/interpretation task. For example, in an image of a scene with low elevation relief and visible standing water

in the southern United States, any trees standing in the water are very likely to be cypress trees. It seems fairly reasonable that this recognition element can be considered to be a higher level abstraction developed from the elements previously discussed. Hence no new metrics can be introduced for this element and we will not include this element in our study.

3.8 CONTEXT OR ASSOCIATION

Some objects are so commonly associated with other objects that one tends to indicate the presence of the other. It is one of the most helpful clues to the identity of man-made installations. As an example, a tall smokestack, large building, pile of coal, conveyors, and cooling towers in proximity to each other are very likely to be associated with power production. As another example, small streams converge into larger streams. The relative angles of the smaller and larger streams can often be used to interpret the general slope of the terrain in the region. This interpretation can be strengthened if the stream beds conform with one of the patterns that the geologic community refers to as "creekology." As with the recognition element of site however, this element is an abstraction developed from the elements present in the image. So again the problem is to image the individual objects in the scene in a recognizable way and the same metrics previously noted apply. Therefore

we will not deal in any direct manner with this interpretation element in this study.

3.9 RESOLUTION

Resolution depends upon many parameters of the image forming system as well as the image viewing conditions. System resolution is the limiting case in answer to the question "Can I see the object I am interested in and differentiate it from its background?" This question applies both to the image viewing conditions and to the scene imaging conditions. The inherent limitations of the human visual system place an upper and lower bound upon the spatial frequency content useable in an image. These bounds are dependent upon viewing conditions with the primary metrics being those of the viewing geometry, viewing illumination, and image magnification. Because these variables are presumed to be under the control of the interpreter and because the limitations of the human visual system are not yet firmly established, it is difficult to quantify these metrics in any useful way. The resolution of the imaging system is more directly addressable. It is usually possible to determine the resolution of a system in terms of ground resolved distance, line pairs per millimeter, acutance or the modulation transfer function.

Chapter IV

RESEARCH APPROACH

In examining the image interpretation/recognition elements set (tone, shape, pattern, size, and shadow), we recognize that they are in general associated with subjective methods of observation. This subjectivity would appear to defeat the goal of being able to develop quantitative measurements for radar images since the variability of the human observer's judgement is introduced into the process. It makes it difficult because we desire to perform objective measurements on images and process these data in some rational fashion in order to arrive at a specific conclusion with respect to the quality of an image in a form that is intuitively satisfying. In addition, we desire to minimize the number of measurements needed to arrive at an assessment of the quality of an individual image. This requires that we determine a set of image metrics that are independent of each other, or in other words, the metrics must be orthogonal.

Our first approach to the establishment of an appropriate set of image recognition/interpretation elements was to attempt to quantify the cues trained image interpreters use as they interpret a variety of features and targets. In this

way, we hoped to establish a rank-ordering of importance or utility of the various subjective interpretation elements. While we are not abandoning this approach, our current experimental protocol involves a more basic view of photographic image quality metrics as applied to synthetic aperture radar produced imagery. The results are not expected to correspond with those related to photographic imagery for the following reasons:

- (a) The dynamic range of SAR data is markedly greater than that of photographic data.
- (b) Distortions introduced in the imaging process due to geometric propagation phenomena are not uniform across an image, i.e., the effects of foreshortening and layover.
- (c) The primary noise in SAR imagery is multiplicative background clutter return noise rather than additive noise as is the case with photography. There may be additive noise introduced by the receiver thermal noise, processor noise, and in some implementations, transmitter noise leakage. A further source of additive noise is film grain if photography is used to record the SAR image. However the dynamic range of the background clutter return is normally much greater in magnitude than the sum of all the additive noise sources.

4.1 DEVELOPING METRICS

It is clear from the foregoing discussion that there is not a one-to-one correspondance between the set of image interpretation/recognition elements and the set of image characteristics that can be directly quantified. Because of this, we conclude that it would be very difficult, if not impossible, to relate each metric directly to a visual cue. A more basic approach was needed.

4.2 METRIC SELECTION

There are a variety of approaches for metric selection but it is not clear which best discriminates a series of images. One which is intuitively very satisfying is that of determining the principal components of the overall cluster in N-dimensional feature space. This is a well recognized procedure in statistics and pattern recognition in which the major and minor axes of the hyper-ellipse containing all the feature vectors is determined [9]. This set of axes is, by definition, orthogonal. Since these axes are described in terms of the coordinate system of the feature space, it may be possible to identify a subset of these measurments that are the primary constituents needed to classify an image as having some desired property. This leads directly to the procedure of assigning a subjective assessment of image quality to each image according to some set of quantified image parameters.

4.3 IMAGE QUALITY

If we perform a clustering analysis of the feature space using an interpreter study derived image quality assessment as our criteria, we can divide the overall set of image metrics into subsets representing various levels of image quality for a specific application. Having identified these subsets, we may then determine classification surfaces in terms of the objective image metrics that will sort the images according to their assessed quality. This being done, the task of assessing the quality of a specific image becomes one of measuring the identified metrics, and observing where in the feature space the pattern vector representing this image falls with respect to the classification surfaces established. This is a process which is amenable to being completely mechanized. That is, an assessment of image quality may be produced without recourse to the variabilities associated with human observers.

The foregoing is, of course, entirely dependent on the existence of separable clusters in the feature space. If the clusters are non-separable, then a single measure oriented along the major axis of the distribution would suffice to measure the relative quality of the measured images. However, our initial estimate is that the clusters will in fact, be separable based on the work performed by Frost, et al [14]. Consequently, we do not expect to obtain a single measure of image quality, but rather a complex multi-dimen-

sional image quality surface upon which the influence of image metrics will vary depending upon the purpose for which the image is utilized. It is worth recalling that what this all amounts to is a statistical procedure for selecting a subset of the properties of an image that: (a) can be measured objectively; and (b) can be used to rank-order a set images in accordance with a group of human observers/interpreters.

Chapter V
PROGRESS TO DATE

We have observed that some seven objective metrics are generally believed to show promise as a way of characterizing the quality of an image. They are:

- (1) The dynamic range of intensities in the displayed image.
- (2) The system signal-to-noise ratio.
- (3) The system spatial bandwidth or bandpass.
- (4) The system resolution or acutance.
- (5) The normalized-mean-square-error (NMSE) as a measure of geometric fidelity.
- (6) The perceptual-mean-square-error (PMSE).
- (7) The radar threshold quality factor (RTQF).

Our plan is to test the validity of these assumed metrics by constructing a series of simulated synthetic aperture ra-

dar (SAR) images in which one or more of the above named elements are degraded by varying amounts. We are presently engaged in the programming effort involved in applying selective levels of degradation to our simulated SAR images. The SAR simulation program itself has been tested extensively and we consider it to be well validated and ready for use. With the compendium of good and degraded SAR images in hand, we will proceed to the interpreter portion of the study which will be followed by a period of statistical analysis effort to develop the relationships between the objective metrics and each of the interpretation/recognition elements and overall image quality.

BIBLIOGRAPHY

- Blumenthal, A.H. and S.B. Campana; An Improved Electro-Optical Image Quality Summary Measure. Image Quality, Proc. SPIE, Vol 310, 1981.
- Burke, J.J. and H.L. Snyder; Quality Metrics of Digitally Derived Imagery and Their Relation to Interpreter Performance. Image Quality, Proc. SPIE, Vol 310, 1981.
- Campbell, F.W. and D.G. Green; Monocular Versus Binocular Visual Acuity. Nature, Vol. 208, pp. 191-192, 1965.
- Castleman, Kenneth R.; Digital Image Processing. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1979.
- Charman, W.N. and A. Olin; Tutorial: Image Quality Criteria for Aerial Camera Systems Photographic Science and Engineering, Vol. 9., No. 6, Nov.-Dec. 1965.
- Cornsweet, T.N.; Visual Perception. Academic Press, New York, 1970.
- Curlander, J.C. and S.N. Pang; Geometric Registration and Rectification of Spaceborne SAR Imagery. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.
- Davidson, M.; Perturbation Approach to Spatial Brightness Interaction in Human Vision. Journal of the Optical Society of America, Vol. 58, pp.1300-1309, 1968.
- Duda, Richard O. and Peter E. Hart; Pattern Classification and Scene Analysis. John Wiley & Sons, New York, N.Y., 1973.
- Fuortes, M.G.F.; Electric Activity of Cells in the Eye of Limulus. American Journal of Ophthalmology, Vol. 46, pp.210-223, 1958.
- Fuortes, M.G.F.; Initiation of Impulses in the Visual Cells of Limulus. Journal of Physiology (London), Vol. 148, pp. 14-28, 1959.

- Frost, V.S., K.S. Shanmugan and J.C.Holtzman; An Information Theory Characterization of Radar Images and a New Definition for Radiometric Resolution. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.
- Frost, V.S. K.S. Shanmugan and J.C.Holtzman; Edge Detection for Synthetic Aperture Radar and Other Noisy Images. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.
- Frost, V.S., J.A. Stiles and J.C. Holtzman; Radar Image Interpretability Analysis. Report No. RSL-TR-342-2, Remote Sensing Laboratory, The University of Kansas Center for Research, Inc., Lawrence, Ks. January 1981.
- Frost, V.S., K.S. Shanmugan, J.C.Holtzman and J.A. Stiles; A Statistical Model for Radar Images of Agricultural Scenes. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.
- Furman, G.G.; Comparison of Models for Subtractive and Shunting Lateral-Inhibition in Receptor-Neuron Fields. Kybernetik, Vol 2. pp 257-274, 1965.
- Gonzalez, Rafael C. and Paul Wintz; Digital Image Processing. Addison-Wesley Publishing Company, Reading, Mass, 1977.
- Hall, Charles F., and Ernest L. Hall; A Nonlinear Model for the Spatial Characteristics of the Human Visual System. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No. 3, March 1977.
- Hall, C.F.; Perceptual Coding in the Fourier Transform Domain. 1980 National Telecommunication Conf., IEEE, New York, 1980.
- Hall, Charles F.; Subjective Evaluation of a Perceptual Quality Metric. Image Quality, Proc. SPIE, Vol. 310, 1981.
- Holtzman, J.C., V.S. Frost, J.L. Abbott, and V.H. Kaupp; Radar Image Simulation. IEEE Trans. on Geoscience Electronics, Vol. GE-16, No. 4, Oct. 1978.
- Johnson, J.; Analysis of Image Forming Systems. Proc. of the Image Intensifier Symposium, 6-7 Oct 1958, US Army Research and Development Laboratories, Fort Belvoir, Va., 1958, (AD-220 160).

Kaupp, V.H., H.C. MacDonald, W.P. Waite, J.A. Stiles, F.S. Frost, K.S. Shanmugan, S.A. Smith, V. Narayanan, and J.C. Holtzman; Analysis of Geologic Terrain Models for Determination of Optimum SAR Sensor Configuration and Optimum Information Extraction for Exploration of Global Non-Renewable Resources: Pilot Study. Report No. ARSL TR 81-2, Arkansas Remote Sensing Laboratory, Dept. of Electrical Engineering, University of Arkansas, Fayetteville, Ar., 1981.

Kaupp, V.H., H.C. MacDonald and W.P. Waite; Imaging Radar: Analysis of Propagation Distortions. Report No. ARSL TR 81-1, Arkansas Remote Sensing Laboratory, Dept. of Electrical Engineering, University of Arkansas, Fayetteville, Ar., 1981.

Kaupp, V.H., H.C. MacDonald and W.P. Waite; Geological Terrain Models. Proc. 1981 International Geoscience and Remote Sensing Symposium, June 1981, Washington, D.C., IEEE, New York, 1981.

Kaupp, V.H., H.C. MacDonald and W.P. Waite; Using Radar Image Simulation to Assess Relative Geometric Distortions Inherent in Radar Imagery. Proc. 1981 International Geoscience and Remote Sensing Symposium, June 1981, Washington, D.C., IEEE, New York, 1981.

Kaupp, V.H., W.P. Waite and H.C. MacDonald; Which Angle of Incidence is Best for Imaging Radars. Proc. 1981 International Geoscience and Remote Sensing Symposium, June 1981, Washington, D.C., IEEE, New York, 1981.

Mitchel, Ralph H.; SAR Image Quality Analysis Model. Ph.D. dissertation, University of Michigan, Ann Arbor, Mi., 1974.

Mitchel, Ralph H. and Stanley Marder; Synthetic Aperture Radar (SAR) Image Quality Considerations. Image Quality, Proc. SPIE, Vol. 310, 1981.

Mohan, S. and J. Nithack; Image Enhancement for Determination of Agricultural Fields Using Digital Slar Data. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.

Moser, P.M.; Mathematical Model of FLIR Performance. NAVAIRRDEV CEN Tech Memo NADC-20203A:PMM, 19 Oct 1972, (AD-A045 247).

Myers, Raymond H.; Response Surface Methodology. Allyn and Bacon, Inc., Boston, 1971.

- Pratt, William K.; Digital Image Processing. John Wiley & Sons, New York 1978.
- Reeves, Robert G., ed.; Manual of Remote Sensing. American Society of Photogrammetry, Falls Church, Va., 1975.
- Rushton, W.A.H.; Peripheral Coding in the Nervous System. Sensory Communication, W.A. Rosenblith, Ed., Cambridge, Mass., MIT Press, pp. 169-181, 1961.
- Tomiyasu, K; Computer Modeling of Speckle in a Synthetic Aperture Radar Image. and a New Definition for Radiometric Resolution. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.
- Tomiyasu, K; Tutorial Review of Synthetic-Aperture Radar (SAR) with Applications to Imaging of the Ocean Surface. Proc. of the IEEE, Vol. 66, No. 5, May 1978.
- Ulaby, F., C.T. Allen and A.K. Fung; Method for Retrieving the True Backscattering Coefficient from Measurements with a Real Antenna. Proc. 1982 International Geoscience and Remote Sensing Symposium, 1-4 Jun 1982, Munich, Federal Republic of Germany, IEEE, New York, 1982.