OPTIMAL POSITION ESTIMATION IN DIGITAL IMAGE METROLOGY

D.I. Havelock[1] and H. Ziemann[2]

[1] NRC Physics
M-36 Montreal Rd.
Ottawa, Ontario
Canada
K1A 0R6

HAVELOCK@NRCPHY.BITNET

[2] Institute Für Navigation
Keplerstr. 11
Postfach 560
D-7000 Stuttgart 1
Germany

HDAA@DS0RUS1I.EARN

ABSTRACT

Photogrammetric digital image metrology needs accurate sub-pixel positioning of targets. With high quality digital image acquisition, quantization effects are a significant source of position estimation errors. Analysis of the degeneracy in the information content due to quantization requires techniques which are different from those used for continuous or analog imagery. From the analytical approaches of digital image representation theory has come the unifying and general concept of the "locale", or "feasibility region", upon which optimal position estimation can be based. A locale is simply the region of target positions which result in the same observed digital image of the target. The approach treats digital imagery differently from conventional techniques and throws new light on the analysis of geometric precision. It has led to an easily determined bound on the possible geometric precision and a natural statement of the meaning of optimal precision. Futhermore, a theoretically optimal algorithm, referred to as "position decoding", has been developed using the concept of locales and the principles of signal decoding from communication theory. Simulation results have been promising and exciting. They have shown the algorithm to be robust and near optimal in the presence of noise. The algorithm is introduced and comparative performance reported.

1. INTRODUCTION

Digital image processing will undoubtedly play a major role in the future of photogrammetry. The large field of digital image processing is composed of diverse applications, most of which are concerned with the context or texture of the image data. Few are concerned, as photogrammetry is, with the metrology and geometry of digital images. Relatively little of the literature is presently devoted to the problems of extracting geometric measurements from digital images and so there is little support outside of the photogrammetric community for important topics such as the metric quality of digital images or techniques for precise mensuration on digital images. This paper addresses these fundamental issues, briefly reviewing some aspects of digital image acquisition, reviewing the theory of locales and then introducing a new optimal position estimation algorithm.

Sub-pixel position estimation is necessitated by the insatiable need for improved accuracy and the unavoidable cost associated with large data volumes caused by small pixels. The objective is to use a manageable pixel size and extract all possible positioning information at that resolution. As a rough rule of thumb, a very good position estimation procedure in an application will achieve about 1/20th of a pixel positioning accuracy (often reported as one part in 10,000 with a 512 by 512 pixel image). There are many pragmatic reasons why better sub-pixel precision is seldom achieved; lighting problems, mechanical instability, scene unpredictability, electronic and sensor noise and algorithm sub-optimality. In some applications, particularily in high precision photogrammetry, many of these problems have been resolved. In these situations, the fundamental limitations of quantization noise begin to play a significant role. It is to this scenario that the paper is addressed.

Digital image processing begins with the acquisition of a digital image. Often the digital image is derived, that is; it is acquired by digitizing another form of imagery. While the results here may be applied to such applications, the main interest lies in primary digital imagery, where the scene is imaged by a 'digital camera' generating a spatial and radiometric (intensity or grey level) quantized representation. This paper begins with a discussion of primary digital image acquisition issues, an appreciation of which is essential to the realistic analysis of geometric fidelity of digital images. It is noted in particular that the typical commercial solid state camera uses an RS170 image transmission standard to relay images from the camera to the computer and that this signal standard seriously jeopardizes geometric integrity.

The central topic of the paper is the theory of "locales", which were introduced at the 1984 ISPRS congress in Rio de Janeiro [11]. The concept has recently been extended to develop an optimal algorithm for position estimation. Following a brief review of locales, the salient points of the new algorithm are introduced. Results of simulation studies are reported. Some investigations with real imagery are underway but results are not available at present.

It is apparent that the photogrammetric community, more than any other field of engineering, has the most to gain by a thorough and rigorous analysis of geometric precision in digital images. Efforts and results in this area must come from within the community and all available techniques should be enlisted in the investigations. The theory of locales may prove to be very useful in this regard.

2. ACQUISITION OF QUANTIZED IMAGERY

Solid state imaging arrays, such as are used in CCD array cameras, provide primary data acquisition of high quality quanitized imagery. Calibration of CCD arrays has shown that array element spacing is regular, even by photogrammetric standards (14,3,2]). The precise spatial sampling is due to the regularity and resolution of the photolithographic process used in the micro-electronic industry to fabricate the imaging arrays. The rigid and planar construction of the die (the term used for the tiny piece of silicon containing the electronics within the "chip" package) further enhances the geometric integrity of array imagers.

An individual sensor in an array accummulates electrons in a potential well formed by electrodes overlaying the photosensitive material. The number of electrons generated per photon (on the average) is referred to as the quantum efficiency of the device. The efficiencies are generally quite high, sometimes approaching unity so that the sensors can be used for "photon counting" applications. A single sensor in an imaging array holds about 100,000 electrons in its "charge bucket" [6,13]. Photodetection in CCD's behaves like shot noise [7,§1.8] so that the variance in the electron count when a fixed luminous flux falls on the sensor is equal to the expected number of electrons [16,13]. An effective quantization scheme then, may be to set the unit intensity equal to the RMS deviation in the charge number. With an expected electron count of 100 thousand, this gives 316 quantization levels with an RMS deviation of one level. There are two other main sources of noise within the sensor; dark current and readout noise. Dark current noise is generated by thermal energy. Cooling is performed in some cameras to reduce this effect but 256 levels of quantization are generally available at room temperature with reasonable lighting [6]. Readout noise levels vary with the method used to move the charge buckets from the imaging array to the amplifying electronics. Terms such as Charge Coupling, Charge Injection and Plasma Coupling refer to readout methods. A typical readout noise level is less than 100 electrons [8,13], consequently readout noise is insignificant except at very low light levels.

Although solid state array sensors are physically quantized in both space and intensity, the physical intensity (gray scale) quantization is not realized in the acquired digital image. The electron count in a sensor is converted to an amplified voltage signal by the camera circuitry. Besides the noise and distortion added to the intensity signal by the camera circuitry, the signal is usually further modified by filtering to a bandwidth of less than 5 MHz (about half the pixel rate). Under these conditions, a truely digital (piece-wise constant) image signal is never output. Note that this corruption of the raw array sensor data ocurrs in the camera and not in the imaging array itself. True "digital cameras" could be made which output a higher quality signal but commercially available cameras are designed according to image quality stardards set by the characteristics of human visual perception rather than the capabilities of digital image metrology.

In principle, knowledge of the characteristics of the camera electronics would allow one to recover most of the raw image signal available at the chip but, unfortunately, the common use of the RS-170 video signal standard to transfer the image to a frame grab card (flash digitizer) ultimately eliminates such a possiblility. The RS-170 signal does not have provision for a synchronous pixel clock, thereby discarding most of the geometric integrity integral to the solid state imaging array. Without a synchronous clock, the digitizer must interpolate the position of each pixel between the start of successive line scans. Not only does the frame grab lack the necessary information to pin-point the timing of each pixel, it generally does not respect the number of pixels per line transmitted by the camera. The frame grab simply resamples the line by interpolation. These problems have been discussed quite thoroughly in assessments of the use of CCD imagers for photogrammetric purposes, as in [15,3,2].

There are two basic approaches to overcoming the loss of geometric fidelity imposed by commercial digital image acquisition systems; signal analysis and specialized electronic circuitry. It has been shown for example, that by clever Fourier analysis of the image signal, the horizontal jitter introduced by the digitizer's resampling can be determined, hence corrected for [14]. Small targets can thus be positioned to an accuracy of about 1/60th of a pixel. With special electronics, one might expect to do better still.

3. SUB-PIXEL IMAGE PROCESSING

Empirical investigations into attainable sub-pixel position estimation in digitally reconstructed imagery, such as [18], have indicated that sub-pixel measurement is realizable, but the level of performance which can be attained is still a debatable issue.

A difficulty with empirical results is that they are only strictly true for the particular system configuration used for the experiment and they are rarely assured of being optimal in any sense. Theoretical investigations on the other hand, tend to be idealized and optimistic. Neither empirical nor theoretical investigations have given a good framework of knowledge about sub-pixel position estimation which can be assimilated by the practitioners of the art. While there have been some important advances in terms of theory and practice, ([5,4,12,17], to mention a few) there is a lack of cohesion or common basis on which they can be purviewed. The "locales theory" introduced in [11] may provide some of the required common ground for the important problem of sub-pixel position estimation. The locales theory introduces a bound for geometric precision against which other analysis and methodologies may be collectively compared.

It is worth emphasizing perhaps, that the issue of sub-pixel position estimation can generally be isolated to just two components of the overall image processing task, namely the acquisition of the digital image and the sub-pixel position estimation algorithm itself. There are many other steps, as figure 1 indicates, which do not directly involve sub-pixel considerations. The position of a target may in fact be estimated twice, once in the course of detection or recognition, wherein an approximate pixel location is determined, and again during precise sub-pixel position estimation based on the raw gray scale image data. In this context, it is important that sub-pixel position estimation be clearly distinguished from the tasks of detection and recognition. It will be assumed throughout this paper that target position is known to approximately one pixel; it is the task of sub-pixel position estimation to improve upon the rough estimate.

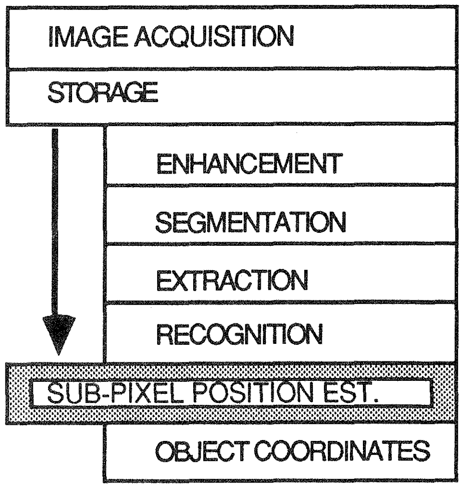

FIGURE 1.
IMAGE PROCESSING STEPS FOR POSITION EXTIMATION (AFTER EL-HAKIM [9])

4. LOCALES

This section will begin with a brief discussion of the history and terminology of locales. The concept of "locales" was developed in [11] for arbitrary targets and at about the same time for the more restricted case of binary line segments encoded by chain codes [6]. The idea is simple and useful: a locale is a region within which the object (target) may be moved without causing any change to its digital representation. The term "domain" used by Dorst and Smeulders [6] refers to a region in a transformation space of object position, but the principle is the same as for locales. The term "feasibility region", as adopted by Berenstein et. al. [1] for the object-position equivalent of the "domain", is the same as locale. Feasibility regions will be introduced in section 8 in the context of locale construction, whereby the former are intersected to generate the latter.

Interest in locales arises from the fact that the locale size determines the position uncertainty due to quantization and the locale center is the optimal position estimate in terms of minimizing quantization errors.

Just what is a "locale"? Consider the following example. A small dot might appear in a digital image as a sampled and quantized Gaussian function Q(i,j),

$$Q(i,j) = \lfloor Ae^{-(i+x)^2-(j+y)^2} \rfloor$$

where A is the amplitude of the dot, (x,y) is its position, (i,j) are pixel indices, and the peculiar brackets indicate integer truncation. A commonly used estimate of the x-axis of the centroid is given by the "center of mass" calculation:

$$\hat{x} = \frac{\Sigma iQ(i,j)}{\Sigma Q(i,j)}$$

with the summation being over all i and j between some values (-1 and +1, say). It is obvious that if the object's position is constrained to a square (pixel) such as $|x|<1/2$, $|y|<1/2$, then the x-position estimate can have only a finite number of values. In particular, if $1<A<2$ then these values are $\{ 0, \pm 1/2, \pm 1/3 \}$. Due to symmetry, the same is true of the estimate of the y-axis of the centroid. The possible combinations of (x,y) estimates are restricted to only 13 values, which are shown as dots in figure 2. The regions delineated in the figure are the locales which correspond to each of the estimate values. Note that ANY position estimator will have no more than 13 possible values, (with (x,y) constrained to the unit square, $1<A<2$, and no noise present), since there are only 13 possible digital representation for the object.

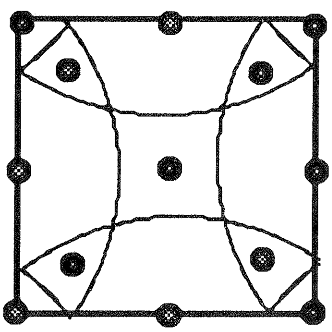

FIGURE 2
LOCALES WITHIN A UNIT SQUARE AND LOCATIONS OF THE CORRESPONDING POSITION ESTIMATES

The presence of noise will complicate the situation. Detailed analysis of locales in the presence of noise is beyond the scope of this presentation, but it should be realized that noise can be incorporated into the basic theory in an approximate manner by defining the number of effective quantization levels (dynamic range) to be the number of digital levels divided by the number of levels spanned by the additive noise. The primary applications of the theory are the estimation of quantization uncertainty and optimal position estimation. Both will be seen to be robust to noise. A more detailed discussion of locales can be found in [10].

The concept of locales can be easily extended from regions of object position to higher dimensional or more abstract parameter spaces. For example, the position of an edge (ignoring end-points or assuming it is infinitely long) can be expressed in terms of its slope and distance from the origin. These two parameters can be used to construct a locale pattern for a straight binary or grey level edge. It is then possible to establish bounds on the position and orientation of the line as well as an optimal estimate of the two parameters. Further elaboration on this example will not be presented but the reader is invited (challenged) to construct the locale pattern based on the discussion in the following section.

5. GENERATION OF LOCALE PATTERNS

The locale patterns are generated from the contours of the target, as explained in [11]. The Gaussian dot discussed in the previous section can be represented by contours with unit intervals which form concentric circles. Using the center of a pixel as the reference origin, the contours are drawn concentric to the origin to represent the target at position (0,0). Displaced versions of this contour pattern are then overlaid on the original one to get the locale map. The displaced versions are generated by moving the contour pattern so that it is concentric with each of the other pixels in the analysis window. For a 3 by 3 window, 9 copies of the basic contour pattern are overlaid to get the final locale map. This is how figure 2 was generated, except that the resulting overlaid contours were truncated at the boundary of the unit pixel.

In the case of higher dimensional position spaces (three dimensions) the same procedure is used to generate a multi-dimensional mesh of locale volume elements. In the case of three dimensional position (x,y,z) for the Gaussian dot, with z along the optical axis of the imaging camera, the basic contour pattern is a set of concentric cones. This pattern is replicated by translation in (x,y), then the replicas are all merged to form the locale pattern of volume elements.

If the coordinates are parameters (such as orientation or size) rather than object position then the translation of the basic "contour" pattern when constructing the replicas is based on the position of the centers of the image pixels in the selected parameter space. Detailed or formal discussion of the more abstract representation of locales is beyond the scope of this paper, but the generality of the concept of locales should be noted.

The basic method of generating locale patterns is very simple. It provides an easy method of appreciating the distribution of quantization induced position uncertainty for any target, no matter how complex the target is and no matter how many gray levels it has. It is apparent that the shape of the locale pattern is dramatically affected by subtle changes in the size, shape and orientation of the target. As will be discussed in the next two sections, this sensitivity is not carried over into the two main applications of the locale; bounds on precision and position estimation.

6. USING LOCALES TO ESTIMATE PRECISION

The uncertainty in target position caused by quantization is fully characterized by the locale pattern for the target. The size and shape of each locale reflects the uncertainty in position for the target when it is positioned within the locale. Each locale may have a different shape so the uncertainty may vary with object position. If the object is in a small locale then the position uncertainty due to quantization will be small. If the object is located in a large locale then the uncertainty may be so large that it dominates all other error sources.

The locale pattern is generally so complex that it cannot be easily used directly to indicate uncertainty; it is necessary to reduce the pattern to one or two quantifying parameters. One useful parameter is the RMS position error due to locale size and another is the largest possible position error, which occurs in the largest locale. Rigorous determination of either of these parameters is not a simple problem but approximations and bounds are available. A simply computed and useful measure of the quantization position uncertainty is given by the "lineal bound" on the number of locales, as derived in [11].

If there are N locales along a line segment (the line segment being in object-position space), and the length of the line is L, then the size of the largest locale (in terms of its intersection with the line segment) must be less than L/N. Furthermore, the minimal RMS error is achieved when all the locales have the same size, giving an RMS size of $L/(N\sqrt{12})$. In either case, the measure of position uncertainty is seen to vary as the inverse of the number of locales along a line. A method for estimating N is presented below.

The value of N for a straight line parallel to the x-axis will give a measure of the position uncertainty in the x direction. The same can be done for position uncertainty in the y direction. The value of N can be easily determined from the contour map of the target, as follows. Draw the contour map and scribe horizontal lines across it with a line spacing equal to the pixel spacing (see figure 3). The number of contours crossed by the lines will equal the number of locales per unit length along any of the horizontal lines [11]. There are several technical provisos here, which are of little consequence in practice. The first is that the value of N depends somewhat upon the vertical position of the set of lines, but the variability is usually small and a solution is to take an average with respect to line position. The second proviso is with regard to the method of counting locales whose boundaries have multiplicity greater than 1 ([11]). This situation arises, for example, when one considers targets such as vertical or horizontal lines. In practice, noise from sources other than quantization tends to eliminate this effect.

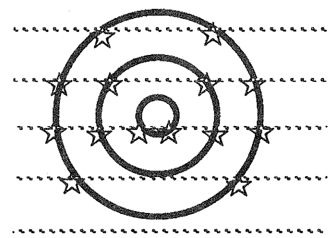

FIGURE 3
COUNTING THE STARS GIVES THE LINEAL BOUND FOR THE NUMBER OF LOCALES ALONG ANY OF THE HORIZONTAL LINES.

As the size or shape of the object is changed slightly, the locale pattern may change radically. The lineal bound on the other hand, does not change significantly unless the target size or shape varies significantly.

The method of determining the lineal bound makes it obvious that position uncertainty generally decrease in a linear fashion with the number of useful quantization levels and with the size of the target. This has been verify by simulation using the optimal position estimation techniques discussed in the next section.

7. USING LOCALES FOR OPTIMAL POSITION ESTIMATION: "POSITION DECODING"

It has been shown [4] that the "best linear unbiased estimate" for target postion is given by the centroid of the locale. The performance of a position estimation algorithm with regard to quantization noise can be formulated as the deviation of the algorithm's position estimates from the locale centroids. There are two causes for an algorithm's shortcomings in this performance; aggregation of locales and bias. Aggregation of locales results when the position estimate is identical for two different locales. In this case, the estimate-locales, which are regions of object position which are indistiguishable by the estimation algorithm, consist of collections of the locales defined by quantization alone. Aggregation can occur when some pixels are disregarded or treated with "reduced significance". It can also occur as a result of loss of numeric precision in the estimation algorithm. Bias, on the other hand, is the difference between the estimate and the actual centroid of the estimate-locales. It has been found for the Gaussian dot and a processing window of 3 by 3 pixels, for example, that the position estimate given by the center of mass calculation has a precision which is limited to about 0.04 pixels due to its inherent bias [10]. In general, the need for an estimation algorithm which is optimal in the sense of minimal quantization errors is only apparent when addressing very high precision position estimation work.

The optimal algorithm called position decoding uses the image of the target to determine the locale corresponding to the target's position, then uses the centroid of the locale as the position estimate for the target. The term position decoding is based on the similarity to the problem of decoding a signal received on a communication line, wherein the incoming waveform is interpretted as a code which is then decoded to give the transmitted character or message. The image of the target corresponds to the signal, the locale corresponds to the code, and the centroid of the locale corresponds to the decoded message. Due to the corrupting influence of the communication channel, the selection of the code is generally probabilistic. There are many more possibilities for received signals than there are for legitimate codes. Typical approaches in communications theory to selecting the best code for a give signal are to use the Maximum Likelihood Estimate (MLE), or the Maximum A Posteriori estimate (MAP). Similar statistical techniques can be incorporated into the position decoding algorithm for locales but for brevity, this paper will concentrate on the basic decoding algorithm in the absence of noise, corresponding to a non-corrupting or noise free "communication channel".

8. THE POSITION DECODING ALGORITHM

In this section the key elements of the position decoding algorithm will be discussed. The position decoding algorithm begins by establishing an ordered list of pixels, which are analogous to a serially received signal on a communications line. The ordering of the pixels is arbitrary in principle, but superior implementation of the algorithm is achieved if the pixels are ordered so that those pixels containing the most "position information" are placed first in the list.

Starting with the first pixel in the list, the algorithm establishes the "feasible region" associated with the pixel. The feasible region is defined as the set of possible target positions which could result in the observed pixel value for that specific pixel. The feasible region is determined from the contour map of the target in a manner similar to the method for generating locales (section 5); with the contour map centered over the specified pixel, the feasible region is given by the contour level corresponding to the observed pixel value. The search for sub-pixel position is generally restricted to a small area such as a unit square, which may be assumed without loss of generality to be centered about the origin. The feasibility region used by position decoding is the intersection of the unit square and the region given by the selected contour level. An example of such a feasible region is represented in figure 4.

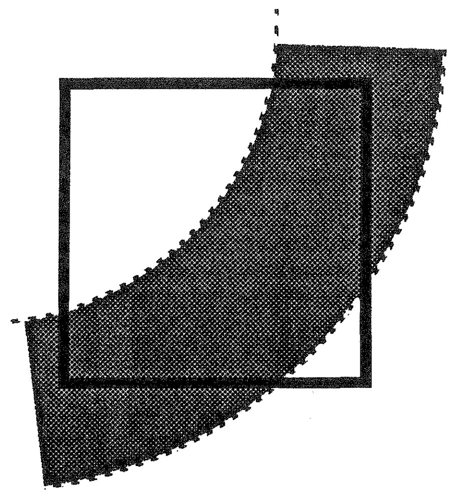

FIGURE 4
A FEASIBLE REGION WITHIN
THE UNIT SQUARE
(BLACK = "TRUE")

The feasible region is maintained in computer memory in the form of a "truth map". Each cell of the truth map is set true if the center of the cell is within the feasible region. Efficient approaches to this will be discussed in the next section.

Position decoding processes each of the pixels in the pixel list, in succession, progressively restricting the number of true cells in the truth map by intersecting the feasible region of each new pixel with the old region given by the truth map. The intersection of successive feasible regions is depicted in figure 5.

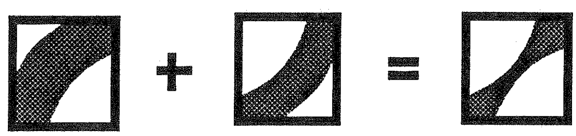

FIGURE 5
INTERSECTING FEASIBLE REGIONS
(BLACK = "TRUE")

By considering the methodoly discussed in section 5 regarding the generation of the pattern of locales for a target from the contour map of the target, it can be seen that this procedure will ultimately lead to the extraction of the locale corresponding to the observed image of the target. The truth map will be a raster representation of the locale, from which the optimal position estimate can be determined by computing the locale centroid. Figure 6 depicts a sequence of intersecting feasible regions resulting in a single locale for the target location.

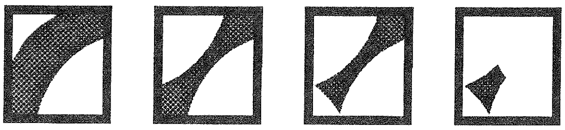

FIGURE 6
SUCCESSIVE INTERSECTION OF FEASIBLE REGIONS TO OBTAIN THE LOCALE

It has been seen that the locale pattern is very sensitive to small variations in the shape of the target. It would seem then, that the position decoding algorithm would be very sensitive to the deviations of the contour map model of the target from reality, either due to noise, distortion, or lack of accurate representation. This problem is alleviated by recording in the truth map the relative probability, for each level of the contour map, that the target was located there. A grey scale representation of this is depicted in figure 7.

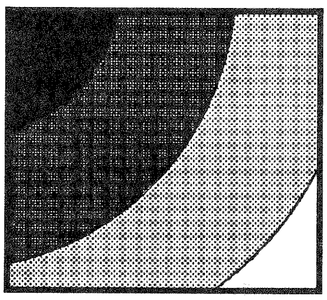

FIGURE 7
USE OF PROBABILTIY DISTRIBUTION FOR FEASIBILITY REGIONS

The probability used in the truth map might be exponential with repect to the magnitude of the difference between the observed pixel value and that of the contour levels. The truth map then represents a portion of the probability distribution for the spatial coordinates of the target given the pixel value. Successive pixels are appropriately combined by point-wise product of the probability distributions for their independent pixels.

It is more practical to store the logarithm of the probability distribution, that is; the square magnitude difference between the observed pixel value and the contour level. In this case, the joining of the information from successive pixels is achieved by summing the corresponding log-probabilities. This results in a straightforward algorithm which iteratively updates a "truth map" of numbers. After all the pixels have been processed, the region within the truth map defined by all those cells with the highest value is taken to be the locale for the target.

The possibilities for tuning and refining the algorithm are many fold. The selection of the list of pixels can be done so that the most useful ones appear first. For example, symmetrically opposed pixels can be processed in succession to take advantage of a symmetric target shape, or pixels occuring where the image has maximum gradient can be selected first since these generally contain better positional information. As the feasible regions are intersected, and the candidate area of the locale diminishes in size, processing can be terminated at a preselected tolerance, thereby avoiding the processing of unnecessary pixels. Pixels which yield no highly-likely cells in the truth map can be deemed defective and eliminated from the analyis. Various probability measures and pixel weights can be tailored. Occluded or missing pixels present no obstacle to the completion of the algorithm.

The algorithm has been implemented with simulated target images and noise, without elaborate tuning, and its performance has been extremely robust and near-optimal. The ease with which good performance is achieved has been impressive.

9. COMPUTATIONAL CONSIDERATIONS

The truth table is the primary storage element of the position decoding algorithm. To economize on space and associated processing, the truth table is allocated only about 15 cells on a side, that is; the unit square representing the area of a pixel is divided into about 225 cells. Since locales tend to be smaller than this cell size, a method is needed to ensure that at least one cell is selected by the intersection of feasible regions. This is accomplished by storing the log-probability in the truth table and selecting the cell(s) with the highest probability. Once the truth map has been evaluated for all the pixels, the process is reiterated with the resolution of the truth map increased and centered over the previously selected cells. At each iteration position decoding provides a better estimate of the shape and location of the locale. This process of incrementally increasing the resolution of the truth map is continued until either the cell size is less than .001 or the locale is clearly resolved (extends over about 1/3 of the truth map).

Iterative refinement of the truth map not only saves storage space, it dramatically reduces the number of cells which must be evaluated and updated during the processing. A further reduction is achieved by processing only those cells which have a sufficiently high probability value.

Associated with the processing of each cell of the truth table is an evalutation of the contour map representing the target. It is important that this representation be computationally efficient. The most efficient method of representation generally varies with the form of the target. In the case of a Gaussian dot for example, a small table of the radii of the concentric contours is both compact and efficient. More complex targets such as corners or crosses convolved with the imaging system's response function may require careful consideration to obtain efficient evaluation.

The robustness of the algorithm, as implemented for the simulation studies, is acheived largely through the use of a tolerance threshold in the truth map. The threshold specifies how low the cell probability can go before the cell is disregarded from further processing. This involves more processing effort than simply retaining only the cells with the highest probability. The modification is necessary because the best choice of locale may well lie outside the "most feasible" region for some of the pixels.

The position estimation algorithm was implemented using a 3 by 3 pixel window to locate a Gaussian dot. It typically executes in about 2 seconds per point and evaluates approximately 3000 cells to achieve an accuracy of better than one hundredth of a pixel.

10. COMPARATIVE PERFORMANCE

The Gaussian dot was selected for comparative performance tests for several reasons; it is easily processed by a number of algorithms, it's radial symmetry allows reduction of performance measures to one-axis parameters, and it is a fairly realistic representation of a target. The processing window was restricted to a 3 by 3 square. For the Gaussian dot as described in section 4, this modest processing window will span the non-zero portion of the target for any amplitude less than 54. The limited scope of the comparison is acknowledged and the need for further simulation and real-data performance evaluation is emphasised. Nevertheless, the results suggest that the algorithm will provide excellent performance under varied conditions and will function well with real data.

The algorithms to which position decoding is compared are; linear interpolation, Fourier phase estimation and centroid estimation. The results of the comparison with and without noise are presented in figure 8.

Linear interpolation takes advantage of the fact that the log of the ratio of the pixels on the left and right sides of the window is a linear function of the x-axis position of the Gaussian target. To reduce the bias in this estimate, quantization was first converted from integer truncation to rounding by adding 0.5 to the quantized value. This algorithm performed quite well, showing continued and rapid performance improvement with the number of quantization levels. Noise reduced its performance only slightly.

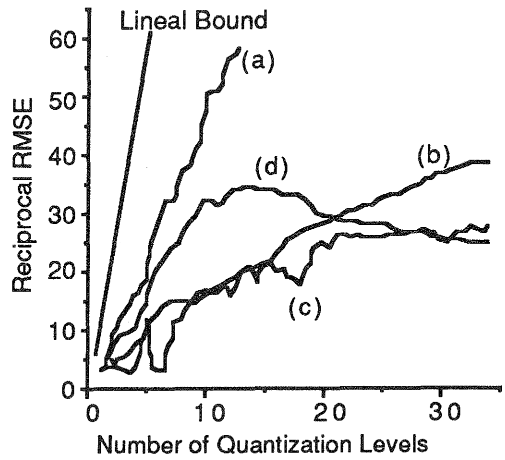

FIGURE 8a
Position estimation errors for the Gaussian dot. The lineal bound is determined by an estimate of the locale density. The reciprocal RMS position errors are shown for four estimators; (a) position decoding, (b) linear interpolation, (c) Fourier phase estimation and (d) centroid estimation.

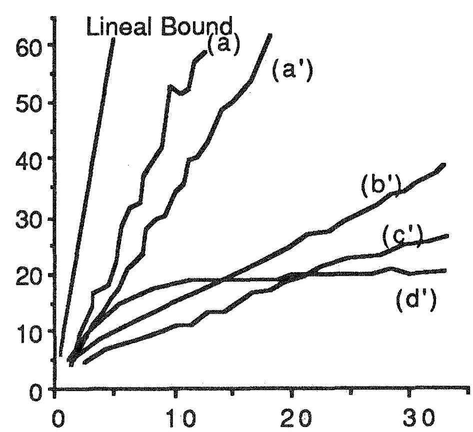

FIGURE 8b
Position estimation errors for the Gaussian dot in the presence of ±1 bit of noise. The reciprocal RMS position errors are shown for four estimators;
(a') position decoding
(b') linear interpolation,
(c') Fourier phase estimation, and
(d') centroid estimation.
For comparison, the 'lineal bound' and (a) position decoding without noise are also shown.

Fourier phase estimation is based on the principle that the phase of the Fourier transform of an image is a linear function of the displacement. The reference image of the target (the entity function [11]) and the observed image are divided pixel by pixel. The phase of the resultant complex-valued image is a linear function of pixel coordinates with a slope corresponding to the position of the target. A least squares fit to the slope of the phase is used as the position estimate. As noise is added, the otherwise erratic performance curve of Fourier phase estimation is smoothed somewhat. The method does not perform well until the number of quantization levels gets quite large, at which point it surpasses the center of mass method of calculating the centroid.

Centroid estimation is done using the "center of mass" calculation as was discussed in section 4. This estimate is straightforward to compute. Explicit calculations have shown that for the 3 by 3 window it exhibits a bound of 24 on its reciprocal RMS error in the absence of quantization. With coarse quantization, this bound is exceeded to a level of about 35 (with 14 quantization levels), illustrating the perversity of quantization effects.

Position decoding is consistently and clearly superior to the other algorithms both with and without noise and at all quantization levels except possibly near-binary. Improvement with the number of quantization levels is extremely linear (up to an beyond 100 quantization levels), as predicted by the optimality criterion of the locales theory. Noise degrades performance somewhat but linearity is preserved. The gap between the lineal bound and optimal position estimation in the absence of noise can be attributed to the fact that the lineal bound is based on the limiting assumption that all of the locales have the same size. It is reasonable to speculate that the divergence of the two curves can be derived from knowledge of the distribution of locales sizes (or perhaps even the converse).

SUMMARY

The possibility of acquiring very high quality digital imagery brings the analysis of quantization effects into the forefront of digital image metrology. Subpixel position estimation to very high levels of precision are possible in theory. A foundation of solid principles is needed for the effective accummulation of knowledge and experience in high precision measuring techniques. The theory of locales is potentially such a unifying basis. Locales are simply defined and easily generated. They serve to estimate position as well as the position uncertainty due to quantization. They lead to a natural and easily stated definition of optimal position estimation. The optimal position estimation algorithm called "position decoding" has been introduced here for the first time. Preliminary performance simulations have shown it be an effective tool for digital image metrology.

REFERENCES

[1] C.A. Berenstein, L.N. Kanal, D. Lavine and E.C. Olson, "A Geometric Approach to Subpixel Registration Accuracy", *Computer Vision, Graphics, and Image Processing*, Vol. 40, pp.334-360, 1987.

[2] H.A. Beyer, "Some Aspects of the Geometric Calibration of CCD-Cameras", *ISPRS Intercomm. Conf. on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, June 2-4, pp.68-81,1987.

[3] J. Dähler, "Problems in Digital Image Acquisition with CCD Cameras", *ISPRS Intercomm. Conf. on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, June 2-4, pp.48-59,1987.

[4] L. Dorst and A.W.M. Smeulders, "Best Linear Unbiased Estimators for Properties of Digitized Straight Lines", *IEEE Trans. on PAMI*, Vol. PAMI-8, No. 2, pp.276-282, March 1986.

[5] L. Dorst and R.P.W. Duin, "Spirograph Theory: A Framework for Calculations on Digitized Straight Lines", *IEEE Trans. on PAMI*, Vol. PAMI-6, No. 5, pp.632-639, Sept. 1984.

[14] L. Dorst and A.W.M. Smeulders, "Discrete Representation of Straight Lines", *IEEE Trans. PAMI*, Vol. PAMI-6, No. 4, pp.450-463, July 1984.

[6] Fairchild, CCD SOLID STATE IMAGING TECHNOLOGY", Fairchild CCD Imaging, 4001 Miranda Ave., Palo Alto, Ca., 159 pages, Parat, Munich, 1981.

[7] R.M. Gagliardi and S. Karp, OPTICAL COMMUNICATIONS, JOHN WILEY & SONS, 1976.

[8] J.C. Geary, S.M. Kent, "Imaging Characteristics of the RCA 512x320 Charge-Coupled Device (CCD)", *SPIE* Vol. 290, 1981.

[9] S.F. El-Hakim, "Real-Time Image Metrology with CCD Cameras", *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 11, pp. 1757-1766, 1986.

[10] D.I. Havelock, "Geometric Precision in Noise-Free Digital Images", Submitted to *IEEE Trans. PAMI*, January 1988.

[11] D.I. Havelock, "Geometric Precision in Digital Images", *International Archives of Photogrammetry and Remote Sensing (ISPRS XV Congress, Comm. III,* Rio de Janeiro, June 1984) Vol. XXV, Part A 3a, 1984.

[12] H. Klaasman, "Some Aspects of the Accuracy of the Approximated Position of a Straight Line on a Square Grid", *Computer Graphics and Image Processing*, Vol. 4, pp.225-235, 1975.

[13] W.F. Kosonocky, "Advances in Solid-State Image Sensors", T33 Course Notes, *SPIE O-E LASE 88*, Los Angeles, Ca., 10-17 Jan. 1988.

[14] R. Lenz, "Lens Distortion Corrected CCD-Camera Calibration with Co-planar Calibration Points for Real-Time 3D Measurements", *ISPRS Intercomm. Conf. on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, June 2-4, pp. 60-67, 1987.

[15] T. Luhmann, "On Geometric Calibration of Digitized Video Images of CCD Arrays", *ISPRS Intercomm. Conf. on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, June 2-4, pp.35-47,1987.

[16] A. Papoulis, PROBABILITY, RANDOM VARIABLES, AND STOCHASTIC PROCESSES, McGraw-Hill, 1965.

[17] T.W. Ryan, R.T. Gray and B.R. Hunt, "Prediction of Correlation Errors in Stereo-Pair Images", *IEEE Trans. on PAMI*, Vol. PAMI-19, No. 3, pp. 312-322, 1982.

[18] J.E. Unruh and E.M. Mikhail, "Mensuration Tests Using Digital Images", *Photogrammetric Engineering and Remote Sensing*, Vol. 48, No. 8, pp.1343-1349, Aug. 1982.