# GEOMETRIC PRECISION IN DIGITAL IMAGES

David Havelock
Photogrammetric Research
National Research Council of Canada
Ottawa, Ontario K1A 0R6
Commission III

## ABSTRACT

Geometric precision of determining the position of detail in a digital image is considered. A mathematical model is developed to establish a bound on the geometric precision of a type of detail. Geometric precision is shown to be more strongly dependent upon dynamic range than upon sampling interval so that storage efficiency is maximized by trading a smaller sampling interval for a larger dynamic range. Using the bounds on precision derived from the model it is shown that in the presence of noise a reduction in sampling interval beyond some limit yields no improvement in geometric precision.

## 1. INTRODUCTION

The question of geometric precision and accuracy is a familiar one in photogrammetry. The analysis of geometric properties of digital imagery has received relatively little attention in comparison to that of more classical systems using optical and photographic imaging methods. As automation and computers become increasingly integral to photogrammetric analysis, the use of digital imagery by photogrammetrists will increase. The bulk of the digital imagery available today is from satellites. This imagery has been analyzed more in terms of resolution and classification than in terms of geometry. Digital processing of photogrammetric imagery is often for the purpose of generating lower resolution digital elevation models or orthophotos. These techniques do not directly involve geometric information on a scale smaller than one pixel. In order to accomplish photogrammetric analysis on digital data to the level of accuracy currently possible with opto-mechanical equipment and photographs, within reasonable limits of digital data quantity, we must be capable of extracting as much geometric information from a digital image as possible. To this end, sub-pixel positioning and its limitations must be well understood.

The quantization of digital imagery in both space and intensity leads to analytic issues which are somewhat different from those of continuous imagery. Quantization noise and interpolation errors are examples which are characteristic and important to digital data but of minimal significance or irrelevant to continuous data. Some of the difficulties specific to digital imagery are overcome by incorporating statistical techniques such as modelling quantization and interpolation errors as additive Gaussian noise. Techniques which have been developed for the analysis of continuous systems are often applied in an approximate sense to the discrete case. For example, Forstner (1982) uses an approximation for derivatives and Kumer (1982) approximates continuous linear operators. Such techniques are often successful but they may conceal underlying properties specific to digital imagery.

This work is directed towards establishing a basis for the analysis of geometric precision in digital imagery in a manner which avoids, as much as possible, presumed correspondence with continuous systems and the use of statistical techniques. In addition, some independence from image content is sought. More specifically, a bound on sub-pixel pointing precision will be derived based on a simple image parameter.

Resolution and precision are different but related concepts in digital imagery just as they are in photography. Resolution is associated with recognizability, while precision is associated with locatability. Resolution is related to the ability to distinguish two closely spaced objects, while precision is related to the error in estimating the distance between two (resolvable) objects. Recognizable objects within an image are referred to as "detail", thus we will be exploring precision of estimates of the geometric position of image detail.

The geometric precision which is obtainable from a digital image may be restricted by a number of factors. The pixel size, scanning pattern, aperture shape and size, dynamic range, and character of image detail will all affect the precision. Knowledge of the available precision may be useful in many ways: The design of control points to be included in digital imagery, and the selection of properties of the image scanning equipment may be directed by their effect on the geometric precision in the resultant image; image 'correlation' procedures might use a measure of the available precision to dynamically set parameters of the algorithm; various algorithms may be compared on the basis of how close they come to a known upper bound on precision and the trade-offs among parameters such as aperture size, dynamic range, and sampling interval may be evaluated.

The contents of the image plays an important role in determining the available geometric precision. To illustrate this point, consider the case of two long straight parallel railway tracks. If the two tracks are resolvable, however well, then by performing a simple fitting algorithm a very good estimate can be made of the gauge of the railway. In contrast, any estimate for the length of one railway tie (considered in isolation) will be much less precise.

This dependence on image content makes it unlikely that geometric precision can be clearly defined independent of application. The use of standard targets such as bar charts has been the time honoured basis of resolution measurements. A similar approach may be in order for general precision measurements in digital imagery but, for specific applications where objects of known shape are to be located, a more detailed analysis may be necessary.

A formal mathematical model is presented in section 2 for image detail in the absence of noise. The intention is to establish a methodology and framework by which geometric precision in digital imagery may be analysed rigorously. The model divides a pixel into regions, each of which is refered to as a 'locale'. The number of locales within a pixel provides a bound on their size, which in turn will bound the available geometric precision. The development of the formal model is prefaced with an heuristic discussion, and followed by some illustrative examples.

Based on the formal model, a bound is established in section 3 for the number of locales, and the implications to data storage are discussed in section 4. In section 5 a bound is established for geometric precision with noise in the image. The space-optimal configuration of scanner sampling interval and number of bits per pixel is discussed.

## 2. FORMAL MODEL OF GEOMETRIC PRECISION

In this section a mathematical model is developed for geometric precision. Before proceeding with the development, we will discuss some of the ideas informally. The main concept is that of a 'LOCALE', which represents a set of indistiguishable positions for an image detail. The locales are positional equivalence classes which partition each pixel. The image detail under consideration is formalized as an 'ENTITY'. The entity incorporates the properties of the imaging system, including for example, the intensity response function of the scanner, the transfer function of the scanning aperture, and systematic distortions of the imaging system. The pixel values of the digital image are sampled values of the entity so that the entity corresponds in a way to a continuous image of the object. The pixel values are determined from the entity exclusively by the sampling interval of the scan.

The development presumes, but is not dependent upon, a regular scanning grid. The term "aperture" will be used here to refer to the effective scanning aperture of the system and should not be confused with the interval of the scanning grid. The aperture size is implicit in the definition of the entity. The scanning system is assumed to provide a uniformly spaced square grid of pixel values. This spatial quantization will be referred to as a scanning or sampling pattern whereas intensity quantization will simply be referred to as quantization, provided no ambiguity arises.

A formal definition of geometric precision is not explicitly presented, instead the discussion revolves around the definition of a locale. An example of a definition of geometric precision in terms of locales might be as follows:

Definition:

> "The geometric precision of an entity is the reciprocal square root of the average area of the locales in a pixel; the geometric precision of a digital image is given by the geometric precision of the 'standard' entity defined as a bivariate Gaussian function scaled so as to fully utilize the dynamic range of the pixel and to have the scanning interval correspond to one standard deviation".

This definition incorporates an entity which is about the same size as one square of the scanning grid. Anything much smaller than this is not resolvable in a digital image. Since the precision with which one can locate an entity generally improves as the entity gets larger, by selecting a small standard entity the definition gives geometric precision in terms of a 'worst-case' situation.

If the scanning interval is the same as the aperture size then, unless the dynamic range is very large, only the immediately adjacent pixels will contain information on the position of such an entity. The locales will be determined by the pixel values in these neighbouring pixels. These pixel values are referred to collectively in this paper as the 'IMAGE FUNCTION', which is a set valued function of the position of the entity. A locale corresponds exactly with the set of all positions mapping into a given value of the image function.

We now proceed with the development of the formal model of these concepts.

An object which corresponds to an image detail is modeled as a two-dimensional reflectance function R (x,y). Some point $(x_o,y_o)$ is designated as the position of the object. Let $R_o$ (x,y) denote the object positioned at the origin, so that

$$R(x,y) = R_o(x - x_o, y - y_o) \qquad (2\text{-}1)$$

Let $P_{ij}$ be the value of a pixel in a digital image of the object, and let $(x_{ij},y_{ij})$ be the position associated with the center of the pixel. We will assume that $P_{ij}$ may be expressed as the quantization of a convolution;

$$\begin{aligned} P_{ij} &= \left[(T_{ij} * R)(x_{ij}, y_{ij})\right]^{\sim} \\ &= \left[(T_{ij} * R_o)(x_{ij} - x_o, y_{ij} - y_o)\right]^{\sim} \\ &= \left[H_{ij}(x_{ij} - x_o, y_{ij} - y_o)\right]^{\sim} \\ &= \tilde{H}_{ij}(x_{ij} - x_o, y_{ij} - y_o) \end{aligned} \qquad (2\text{-}2)$$

for some transformation function $T_{ij}$ where $[\cdot]^{\sim}$ denotes quantization, * denotes two-dimensional convolution and $H_{ij}$ is as discussed below. In order to represent $P_{ij}$ in this way, the imaging system must be spatially invariant and linear [4]. While these are not generally true assumptions, they may be reasonably good approximations locally for a limited range of reflectance. The transformation T is subscripted with the pixel index (i,j) to reflect the local nature of the representation.

The functions $H_{ij}$ and their quantized counterparts $\tilde{H}_{ij}$, which are defined in terms of $T_{ij}$ and R over the (continuous) coordinate variables (x,y), will be referred to, collectively in i and j, as an "entity", or "quantized entity" respectively. The convolutions of R with $T_{ij}$ results in $H_{ij}$ being generally quite smooth and well behaved. We will presume that $H_{ij}$ is non negative and is greater than zero only on a region of finite area.

The pixel value $P_{ij}$ is just a sampling (i.e., point evaluation) of the quantized entity. A shift in the object's position corresponds to a shift in the sampling of the entity. It will be convenient to assume that all the $H_{ij}$ are identical, that is, $T_{ij} = T$. In this case we will drop the subscripts and say that the entity H is "spatially invariant". The pixel values are then obtained by sampling a single function. If, in addition, the pixel grid is uniformly spaced with unit spacings, then

$$P_{ij} = \tilde{H}(i - x_o, j - y_o). \qquad (2\text{-}3)$$

By a "locale" of an entity, or quantized entity, we will mean an area, A, consisting of all points (x,y) for which the quantized entity functions $\tilde{H}_{ij}$ do not change. More precisely, (x',y') lie in A, if and only if

$$\tilde{H}(i - x, j - y) = \tilde{H}(i - x', j - y') \qquad (2\text{-}4)$$

for all (x,y) in A and all i, j. The locales partition the plane into sets of equivalenced positions. Each locale corresponds to an area of uncertainty for the position of the object.

As an example, consider the spatially invariant entity H which is a raised unit square of height $a_o > 1$,

$$H(x, y) = \begin{cases} a_o & \text{if } |x| < 1/2 \text{ and } |y| < 1/2, \\ 0 & \text{otherwise.} \end{cases} \quad (2\text{-}5)$$

If the pixel centers $(x_{ij}, y_{ij})$ form a unit grid, then the locales of H are the unit square centered on the pixel (excluding the perimeter) and the unusual locale consisting of the boundary of the square. The latter locale results from the fact that, for these positions of the object, all pixels values are zero.

It should be noted that, for a uniform square grid and a spatially invariant entity, the partition of the plane into locales has the same spatial periodicity as the grid.

For a quantized entity $\tilde{H}_{ij}(x,y)$, we define the (ordered) set-valued "image function" I(x,y) as

$$I(x, y) = \{(\tilde{H}_{ij}(i - x, j - y)\} \quad (2\text{-}6)$$

where the indicies i, j are ordered in some manner. Thus for a given position, the image function is the ordered set of all pixel values for the entity. Since the entity has finite support (i.e. it is greater than zero only on a set of finite area) only a finite (and usually only a few) pixel values are not zero. The "image distance" between two points $(x_1,y_1)$ and $(x_2,y_2)$ is defined here as

$$\| I(x_1,y_1) - I(x_2, y_2) \| = \sum_{ij} | \tilde{H}_{ij}(i - x_1, j - y_1) - \tilde{H}_{ij}(i - x_2, j - y_2)|. \quad (2\text{-}7)$$

According to this definition, two points $(x_1,y_1)$ and $(x_2,y_2)$ belong to the same locale if and only if

$$\| I (x_1, y_1) - I (x_2, y_2)\| = 0. \quad (2\text{-}8)$$

Furthermore, if the unit of quantization of pixel values is q, then if the two points are in different locales

$$\|I (x_1, y_1) - I (x_2, y_2)\| = k.q \quad (2\text{-}9)$$

for some positive integer k. Usually q is taken as unity.

In some sense each locale corresponds to a unit of available geometric precision in that the larger the locale the greater the lack of geometric precision. The variability of the size and shape of locales, and their dependence upon the character of the entity makes it awkward to use locales as a direct measure of precision. As an approach to establishing a measure of available geometric precision in a digital image we will estimate an upper bound on the number of locales along a unit line segment. It should be noted that locales may span pixel boundaries and that they may be disconnected. The size of a locale must always be less than or equal to the size of the unit raster square since a shift by one sampling unit will change any non-zero image function.

## 3. BOUND ON THE NUMBER OF LOCALES

As we move along a parametrized line segment $\underline{x}(t) = (x(t), y(t))$, an upper bound on the number of locale boundaries crossed is given by the total variation V of the image function as a function of t. For sufficiently small changes $\Delta t$ in the parameter t the total variation is given by

$$V = \sum_{n=0}^{N} \| I(\underline{x}(n.\Delta t)) - I(\underline{x}(n.\Delta t + \Delta t)) \| \qquad (3\text{-}1)$$

(where $N=1/\Delta t-1$ is an integer value). Consider the line segment

$$\underline{x}(t) = (t, b) \qquad 0 \leqslant t \leqslant 1 \qquad (3\text{-}2)$$

which is a line parallel to the x-axis along y=b, extending a single raster unit. The total variation of the quantized image function I for a spatially invariant entity $\tilde{H}$ is

$$\begin{aligned} V &= \sum_{n=0}^{N} \| I(\underline{x}(n.\Delta t)) - I(\underline{x}((n+1).\Delta t)) \| \\ &= \sum_{j}\sum_{i}\sum_{n=0}^{N} |\tilde{H}(i - n.\Delta t, j - b) - \tilde{H}(i-(n+1)\cdot\Delta t, j - b)| \\ &= \sum_{j}\sum_{n} |\tilde{H}(n.\Delta t, j - b) - \tilde{H}((n+1)\cdot\Delta t, j - b)| \\ &= \sum_{j} K_j \end{aligned} \qquad (3\text{-}3)$$

where $K_j$ is the variation in H(x, y) along the line y = j+b.

The form of the total variation is particularly simple for entities which are unimodal along all cross-sections. A unimodal function is simply one which increases monotonically to a maximum value then decreases monotonically again. The gaussian function is a good example. For such entities, the variation along the line segment

$$\underline{x}(t) = (t,b) \qquad 0 \leqslant t \leqslant 1 \qquad (3\text{-}4)$$

is given by

$$\begin{aligned} V &= 2 \cdot \sum_{j} \underset{x}{\text{MAX}} \{\tilde{H}(x, b + j)\} \\ &= 2 \cdot \sum_{j} M_j \end{aligned} \qquad (3\text{-}5)$$

where

$$M_j = \underset{x}{\text{MAX}} \{\tilde{H}(x, b+j)\}$$

Since this bounds the number of locales crossed along the line segment, the number of locales is bounded by $N_b$:

$$N_b = 2 \cdot \sum_{j} M_j + 1 \qquad (3\text{-}6)$$

Figure 1 illustrates the partition of a unit square by the locales defined by the entity function

$$H(x,y) = 4.1e^{-(x^2 + y^2)} \qquad (3\text{-}7)$$

Quantization is accomplished by integer truncation. The entity function has five levels of quantization (counting zero) and has support radius of about 1.2, that is; it is non-zero only within a radius of 1.2 of the origin. The image function is obtained by sampling the quantized entity function $\tilde{H}$ on the set A;

$$A = \{(i,j) : |i| + |j| < 2\} \quad (3\text{-}8)$$

Explicitly, the image function is as follows:

$$I(x,y) = \{\tilde{H}(x,y),\ \tilde{H}(x+1,y),\ \tilde{H}(x-1,y),\ \tilde{H}(x,y+1),\ \tilde{H}(x,y-1)\} \quad (3\text{-}9)$$

Along the line segment $\{(t,b) : 1 \leqslant t \leqslant 0\}$ the number of locales is bounded by

$$N_b = 2 \cdot (M_{-1} + M_o + M_{+1}) + 1$$
$$= 2 \cdot (\tilde{H}(0,\ b-1) + \tilde{H}(0,\ b) + \tilde{H}(0,\ b+1)) + 1 \quad (3\text{-}10)$$

The value of $N_b$ gives us the maximum number of (distinguishable) positions for the entity along such a unit line segment. On the average then, we can only know the position of the object to within a distance no smaller than $1/N_b$ .

Using equation (3-10), for b=0.0, b=0.25, and b=0.5 the bounds are 13, 11, and 13 respectively. The number of locales along the above line segments is 11, 9, and 7 respectively (see figure 1). Several of the boundaries between locales along the lines happen to be the intersection of prominent arcs in the partition diagram. These intersections are responsible for the numerical difference between the upper bound and the number of locales along the line segment. For other entities, such as in figure 2 a locale boundary may correspond to a change in the image function I(x,y) of several units.

The relationship (3-6) between the variation of the entity function and the number of locales is similar in principle to the observation by Forstner (1982) that geometric precision in image correlation will depend upon the variance of the first derivative. The variation is like the variance of the first derivative in that both quantities are measures of the 'texture' of the function. A similar principle is used by Ryan, Gray and Hunt (1980) in defining indicators of correlation errors (see table VII of their paper).

## 4. DATA STORAGE CONSIDERATIONS

Large data sets consume computer resources during storage, retrieval and computation as well as by virtue of the volume of their required storage medium. In order to make the data handling as efficient as possible, the 'natural' word size of the machine is often taken into account when deciding on the number of data bits to use per pixel. We will not be constrained here by computer architecture; we will presume that the number of pits per pixel can be freely selected, and may be fully utilized. For a given number of bits we will establish a bound on the number of locales across an entity of a given size. If an entity is non-zero at only one sampling point, then the summation in equation 3-6 is trivial. The bound on the number of locales $N_o$ across a unit raster square for such an entity is just

$$N_o = 2^b + 1 \quad (4\text{-}1)$$

where b is the number of bits used for the pixel. It can be easily seen that no matter what the support of the entity is (i.e., how many terms there are in the summation of equation 3-6), the bound on the number of locales will still vary in an exponential fashion with the number of bits per pixel.

Now consider how the geometric precision varies with the sampling interval. The 'standard entity' mentioned in the definition of geometric precision (section 2) will produce a pattern of locales which will depend upon the amplitude and spread of the Gaussian function. Since the spread of the entity function is determined by the sampling interval, the size but not the pattern of locales will change as the sampling interval is changed. A result of this is that the number of locales along a straight line of fixed length and the root mean locale area will both vary directly with the sampling interval. Since the number of pixels in the image varies as the square of the sampling interval, the number of locales will vary as the square root of the number of pixels. This is a much weaker dependance than the exponential variation with the number of bits discussed above.

Since the storage required for a digital image varies directly as the number of bits per pixel times the number of pixels, it follows that if the number of locales is the exclusive determinant of the available geometric precision in a digital image then it would be most efficient to trade off a smaller sampling interval for more bits per pixel. Resolution requirements and noise limit this tradeoff.

## 5. MAXIMIZING AVAILABLE GEOMETRIC PRECISION

Once the form of the entity has been established by such factors as the shape of the object, the aperture shape, and the response of the imaging system, there remains the sampling frequency and quantization level to determine the limits of geometric precision in the digital image. We will presume that the aperture size is commensurate with the scanner sampling interval so that the unit raster square, as determined by the sampling interval, is about the same size as the pixel (determined by aperture size). It is also presumed that the quantization levels are uniformly spaced and may be fully utilized. The presence of noise and the requirement that the object be recognizable will jointly determine an optimal scanning interval and number of bits per pixel.

To achieve object recognition in the digital image there must be adequate resolution. Although dynamic range and effective pixel size play a part in determining the available resolution, the primary criteron should be that the sampling interval be matched to the spatial extent of the object. This is essentially a Nyquist criterion.

Once an adequate sampling interval has been established, one may address the question of the number of bits per pixel. Apart from data volume constraints, the determining factor is that the dynamic range must be maximized. The dynamic range is the number of distinquishable levels of pixel values in the presence of noise. If the range of pixel values span the $2^b$ values provided by b bits per pixel and if the noise is within v levels then the dynamic range D is

$$D = 2^b/v \qquad (v \geq 1) \qquad (5\text{-}1)$$

In the absence of noise (v=1), the bound on the number of locales for an entity increases without bound as the number of pixel values increases. This is not the case in the presence of noise. The bound on the number of locales then becomes a function of the dynamic range rather than the number of bits. Once v is greater than unity it increases with the total number of pixel values, so that increasing the number of bits will not further increase the dynamic range.

It is interesting to note what happens if the sampling interval is made smaller in an attempt to increase the geometric resolution. As stated earlier, it is assumed that the aperture is commensurate with the sampling interval. Since the scanning aperture corresponds to a spatial average, the noise may be expected to reduce proportionally with an increase in the aperture size. This may be incorporated into the above equation as follows:

$$D = (I/v_0)\ 2^b, \qquad \text{for } I/v_0 \leq 1 \qquad (5\text{-}2)$$

where I is the aperture size (or equvalently by assumption the sampling interval) and $v_0$ is some normalized measure of the noise. Thus a smaller sampling interval results in a proportionally decreased dynamic range. The bound on the number of locales along a line has been seen to vary linearly with both an increase in dynamic range and a decrease in the sampling interval, so no change in this bound is realized by sampling more frequently! The limit to the sampling interval, it should be stressed, is established by the noise level, not by the requirement of object recognition. Once the sampling interval (and aperture size) is small enough that noise exceeds the quantization level, further reduction in size will not improve the bounds on geometric precision irrespective of resolution.

## 6. SUMMARY

As a step toward establishing definitive limits to the geometric precision available from digital imagery, a formalism has been developed to model the resolvable positions of detail in a digital image. The detail, or object, in the digital image is modelled as a function or set of functions refered to as an entity. The representation of the detail within the digital image is obtained from quantized values of the entity, so that the entity embodies the character of both the object and the system which produces the digital image. It was shown that the area within a raster square may be partitioned into a number of regions called locales with each locale corresponding to a set of mutually indistinquishable positions for the object. The shape and size of these locales depends strongly upon the nature of the entity. A bound on the number of locales across a raster square is derived based on the variation of the entity along parallel lines across its support.

An example was given which shows how this bound may be easily calculated for a known entity, while illustrating that the number of locales may actually fall short of the bound. The available geometric precision in a digital image is allied to the bound on the number of locales. Based on this, conclusions are drawn regarding data storage efficiency and limits to geometric precision.

In terms of efficient data storage, it was shown that the available geometric precision in a digital image of a fixed number of bits will be improved by increasing the number of bits utilized per pixel with a corresponding decrease in the number of pixels.

In the presence of noise of a known level it was shown that a bound exists for the geometric information independent of the resolution or number of bits per pixel. Provided that the aperture size is commensurate with the scanning interval, increased bits per pixel are "wasted" on noise without any improvement in usable dynamic range, while a refined sampling interval results in increased noise and no overall gain in geometric precision. This fact underscores the importance of error and noise analysis in the design of a digital image system.

## REFERENCES


[1] Wolfgang Forstner. On the Geometric Precision of Digital Correlation; Proceedings of the Symposium on Mathematical Models, Accuracy Aspects and Quality Control, ISPRS Commission III June 7-11, 1982.

[2] V.K.V. Kumar, Casasent and Murakami. Principal-Component Imagery for Statistical Pattern Recognition Correlators, Optical Engineering, Vol. 21, No. 1, 1982.

[3] T.A. Ryan, R.T. Gray, B.R. Hunt. Prediction of Correlation Errors in Stereo Pair Images, Optical Engineering, Vol. 19, No. 3, 1980.

[4] A.V. Oppenheim, R.W. Schafer. Digital Signal Processing, Prentice-Hall, 1975.

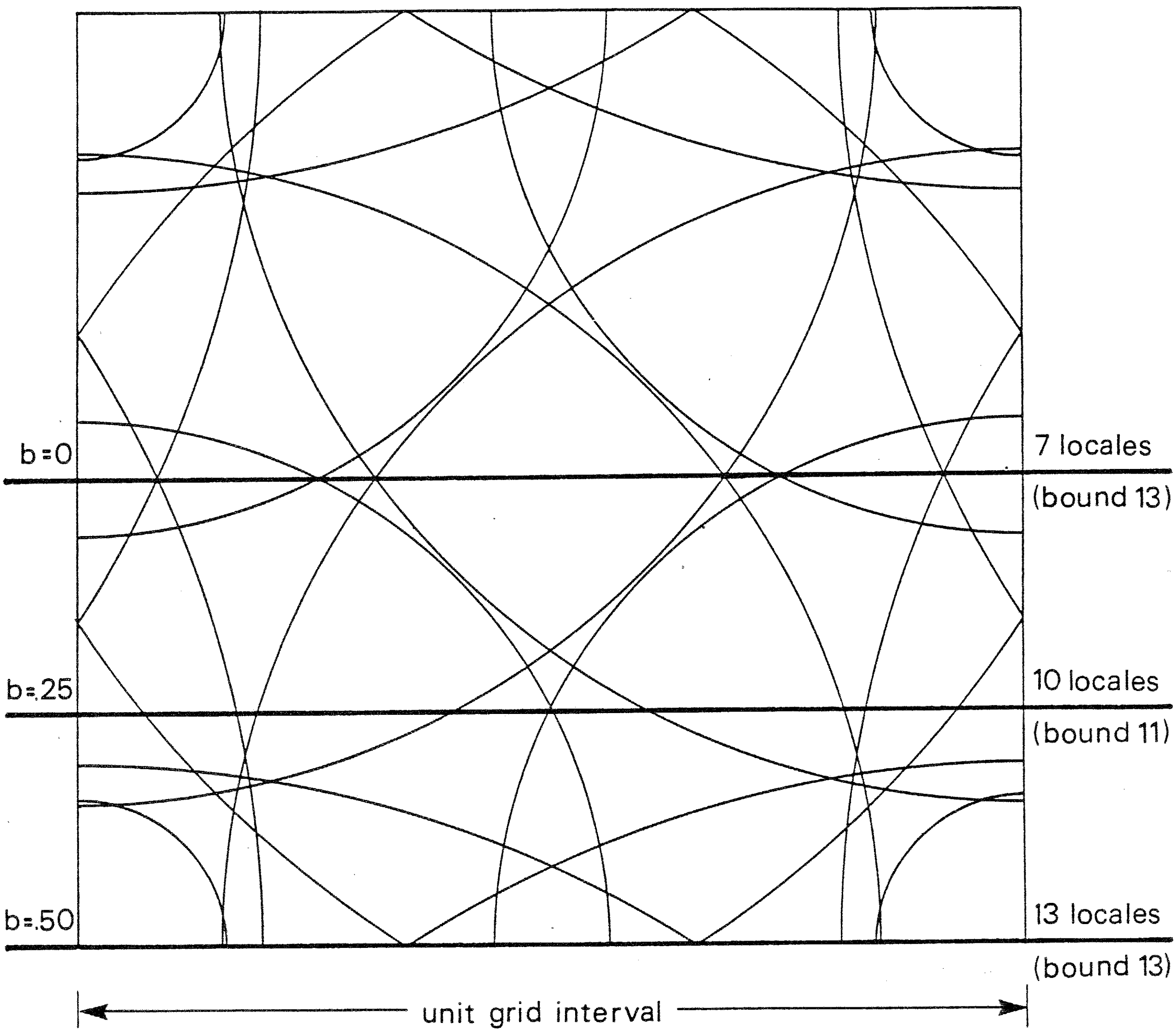


Figure 1: Example Pattern of Locales

A unit raster square is partitioned into this pattern of locales for the entity $H(x,y) = 4.1\ e^{(x^2 + y^2)}$. Quantization is performed by integer truncation. Points in adjacent locales have an image distance of 1 between them. The bound and actual number of locales crossed for 3 horizontal lines are illustrated. The raster square boundary does not define an edge for the locales; the pattern of curved lines is periodically repeated to cover the plane.

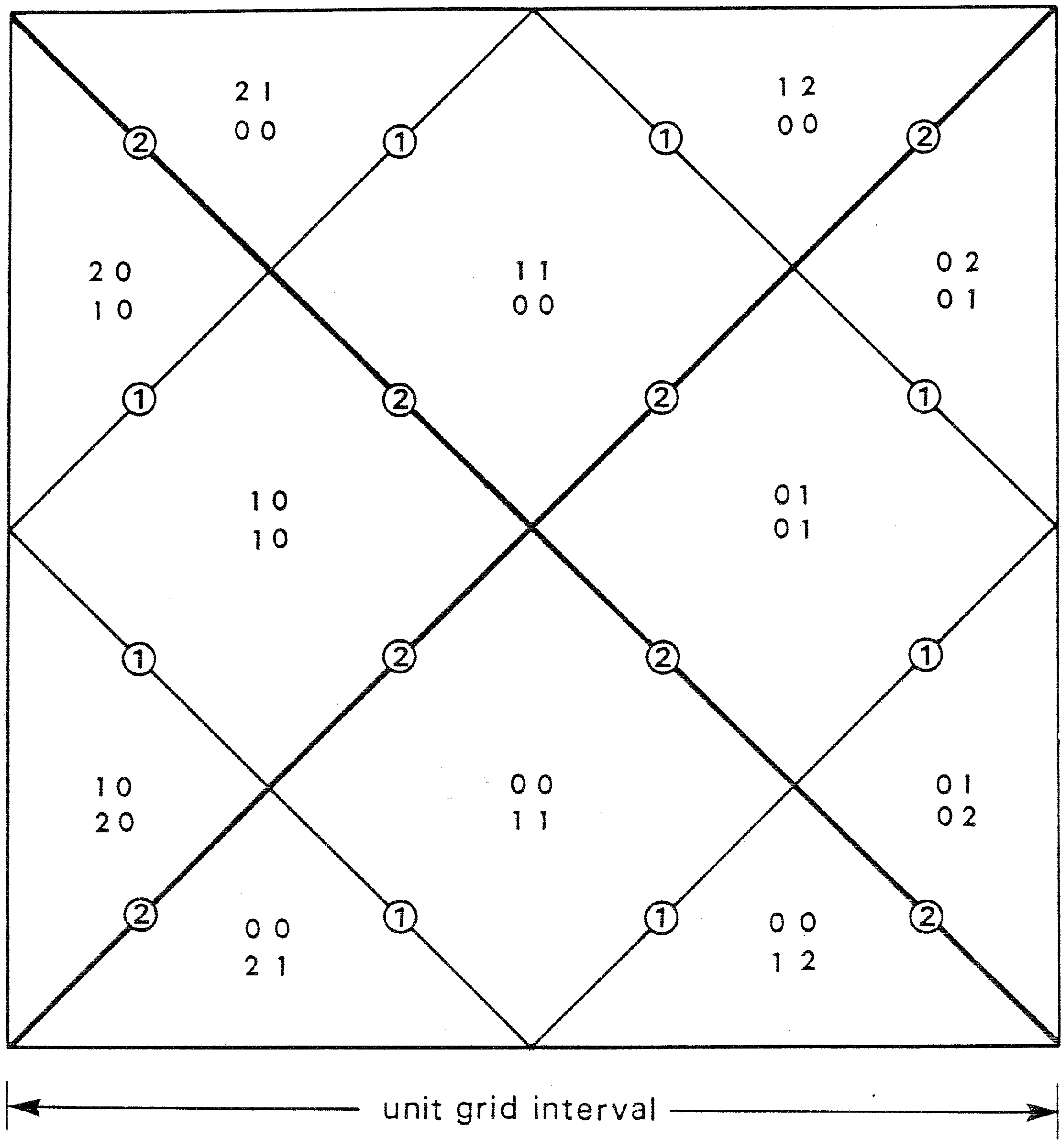


Figure 2: Example with Reduced Number of Locales

Locales are shown for the entity function $H(x,y) = 3(1-|x|-|y|)$. The support of the entity contains 4 pixels. Printed within each locale is the corresponding 4-valued image function. The circled values on the locale boundaries are the image distances across the boundaries. Image distances greater than 1 result in the number of locales being less than the estimated bound. Note that pixel centers (raster grid verticies) are at the corners of the figure. Some locales are continued in neighbouring raster squares; for example $\begin{smallmatrix}01\\01\end{smallmatrix}$ is continued in the right-hand neighbouring raster square. It is possible for locales to consist of disconnected regions.