

Review article

Synergy in remote sensing—what's in a pixel?

A. P. CRACKNELL

Department of Applied Physics and Electronic and Mechanical Engineering,
University of Dundee, Dundee DD1 4HN, Scotland, U.K.

(Received 27 December 1996; in final form 4 December 1997)

Abstract. We address various questions that arise from the fact that a pixel—or the related instantaneous field-of-view (IFOV) on the ground—is often larger than we would like it to be. The problems arise as a penalty imposed by technology for the fact that a spacecraft gives an overview of a very large area. We study the question of why the pixel size is important when one studies satellite imagery and also some questions related to the factors that contribute to the recorded signal in a remotely-sensed data set. This includes a discussion of the instantaneous field-of-view, both from the simple geometrical point of view and from a more physical point of view. It also involves a study of the point spread function and some discussion of the problems associated with the determination of the point spread function of a given scanner. The questions of the calibration of the detectors in an instrument and of the intercalibration of nominally identical members of a series of instruments are also considered.

The integration of remotely-sensed data into a GIS almost inevitably involves resampling and this leads to further complications in understanding the origin of the signal (digital number, DN) associated with a given pixel. The effects of different methods of interpolation are considered as well as the consequences of resampling in relation to the classification of an image. Several other topics are also considered and these include (a) the achievement of geometrical rectification with an error substantially smaller than (e.g., only 20 per cent of) the length of the edge of the instantaneous field-of-view, (b) data compression or upscaling, (c) mixed pixels, and (d) the study of sub-IFOV size objects.

The general conclusion is that it is important to realise that what contributes to producing the digital number on a computer tape or in a disk file of an image is not a simple thing. There is no simple answer to the question 'what exactly gives rise to the signal detected and recorded in a pixel in a remotely-sensed image?' The main point to be made is to try to ensure that it is realised that there is a problem and to give some indication of the nature of that problem.

1. Introduction

The *Oxford Dictionary's* definition of synergy is 'the interaction or cooperation of two or more drugs, agents, organizations, etc., to produce a new or enhanced effect compared to their separate effects'. In the remote sensing context we therefore take synergy to relate to the utilisation of two or more data sources together in order to extract more environmental information from their combination than one could extract from the separate data sources individually. An important requirement for someone attempting this synergy is therefore to have a profound understanding of the nature of the individual data sources that one is seeking to combine. One important problem that arises in a great deal of remotely-sensed data is that a

pixel—or the related instantaneous field-of-view on the ground (IFOV)—is often larger than we would like it to be, that is in relation to the objects that we are trying to observe. These problems arise as a penalty imposed by technology for the fact that a spacecraft gives an overview of a very large area. In some ways the problems of the size of a pixel are analogous to the question of grain size in a photographic film, although there are differences. The most obvious difference is that the pixels in an image from a spaceborne scanner are (usually) all of identical size and shape and are arranged in a simple geometrical (space-filling) two-dimensional (rectangular or square) lattice structure, whereas the grains in a photographic film are of unequal sizes, are of irregular shapes and are irregularly arranged. Yet (at least until one gets to the stage of digitising air photographs) one does not usually find people worrying about limitations placed on the accuracy of photogrammetric work by the existence of the grain structure in the photographic film. Why is this? The main reasons are:

- (i) that one can use large size film, as for example in a survey camera with a film size 9 inches square (230 mm square).
- (ii) that by flying sufficiently low one can usually ensure that the dimensions on the film of the smallest object that one wishes to resolve are much larger than the grain size of the film.

Why does the pixel size become important when one turns to satellite imagery? We can suggest two reasons for this.

- (i) If one uses a film the distance from the surface of the Earth is much larger than for an aircraft; thus the scale of the satellite photograph is so much smaller than for an air photograph. The linear dimensions on the film of the smallest object that one may wish to resolve will then become much closer to the grain size than is the case with an air photograph.
- (ii) The vast majority of satellite imagery is not gathered in a camera on film. It is gathered by an electro-optical scanner using a rotating mirror and a set of detectors or, in a pushbroom scanner, by a CCD array. The resolution, considered as the pixel size in relation to the grain size of a photograph, is very poor. Thus the linear dimensions, in the image plane, of the smallest object that one may wish to resolve may well be smaller than the edge length of the pixels of which the image is constructed. This leads to various problems if one is concerned with objects that are very small compared with the size of the IFOV (instantaneous field-of-view) or if one is concerned with an object field where the texture has a 'graininess' or spatial variation on a scale that is small compared with the size of the IFOV.

2. The simple ideas of pixels in a satellite image. The geometrical IFOV

We need to give some consideration to the mechanism by which a scanner forms an image and to the idea of the instantaneous field-of-view. The problems are not trivial. It is convenient, and common practice, to construct an image from square or rectangular picture elements (or pixels). The simple-minded idea that one often finds presented in elementary textbooks, when describing the operation of a scanner, is that the instrument receives all the radiation from a certain area on the ground (the instantaneous field-of-view, IFOV) and generates a response that is proportional to the amount of radiation received. For convenience that field of view is commonly regarded as a rectangular (or square) piece of the Earth's surface so that, from the

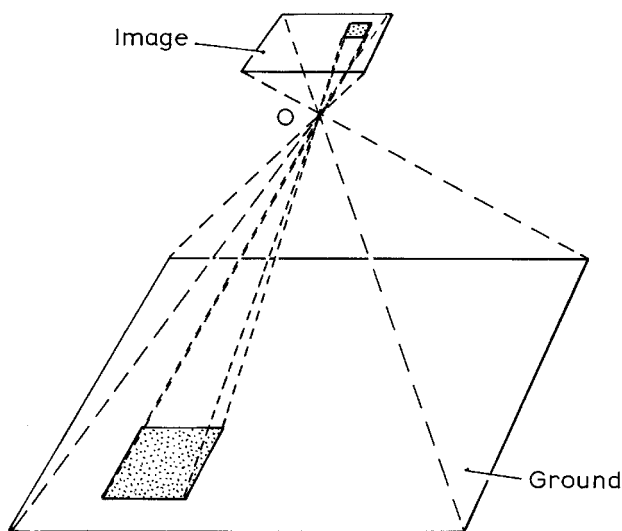


Figure 1. Illustration of the geometrical instantaneous field-of-view reconstructed by projection from a pixel in the image plane.

array of scanner data, an image can be constructed where the pixels fill a two-dimensional plane surface as a scale reduction representing the ground, see figure 1. It would be nice if this were true. It would be even nicer if all the objects on the ground were (a) exactly the same size as this hypothetical rectangular or square IFOV (or multiples thereof), (b) all aligned with their edges parallel or perpendicular to the scan direction, and (c) exactly located so that their centres coincided with the centres of the pixels, see figure 2. Even if the scanner were responding uniformly to all point sources within this geometrical IFOV then the chances of all the objects on the ground being as in figure 2 is patently near to zero; life, cars, houses, cows, etc. are just not like that.

3. Mixed pixels

Even this simple idea of what I have called the geometrical (rectangular or square) IFOV fails because objects do not exist as shown in figure 2. There is no problem if we look at an area, plains, desert, water, etc., where the surface is homogeneous over an area that is large compared with the size of the IFOV. However situations can arise where structure on a small scale relative to the IFOV exists, see figure 3. Assuming still for the moment a uniform response from all points in this geometrical IFOV, then it is clear that subpixel size objects (e.g., gas flares or other small intense sources of heat) or narrow linear features (e.g., bridges, roads, railway tracks etc.) may contribute sufficiently to the received signal to cause the pixel intensity to differ from those of surrounding pixels.

4. The physical IFOV

What we have described as a geometrical IFOV and what we have supposed about the scanner responding uniformly to radiation from all points within the geometrical IFOV is quite far removed from the actual situation. There are several reasons for this. For instance, the simple description above would mean that in collecting the radiation from a surface a scanner with a rotating mirror would pause

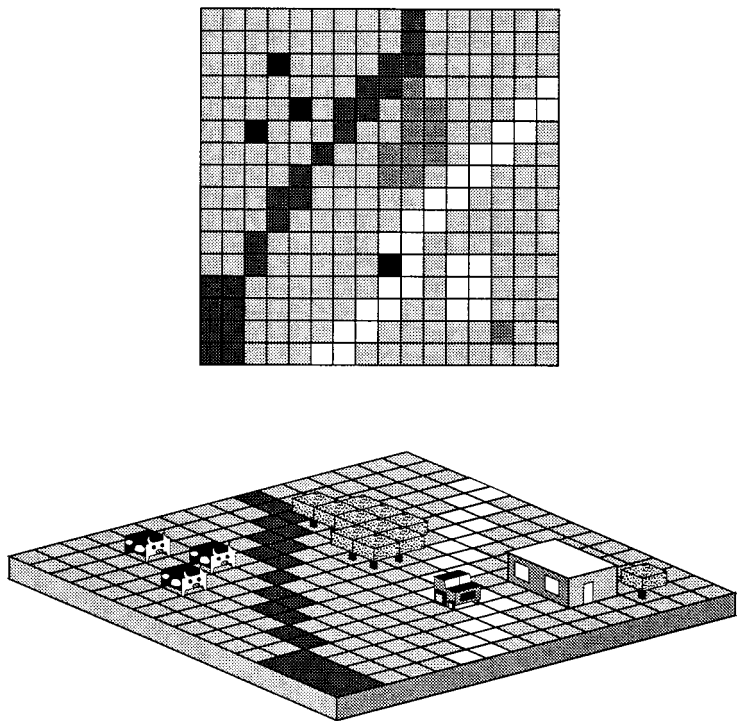


Figure 2. Idealised objects exactly occupying whole pixels and with uniform horizontal top surfaces (Fisher 1997).

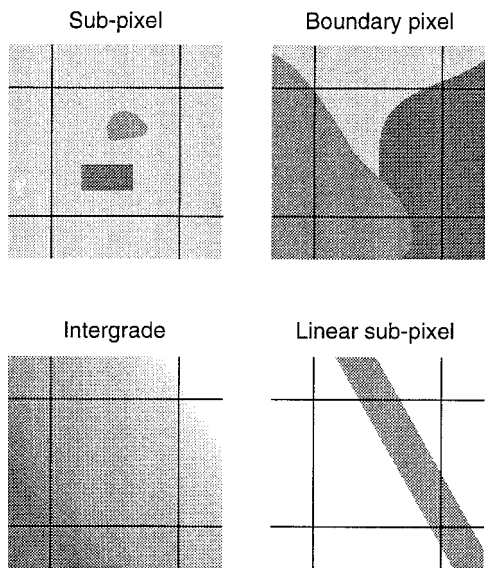


Figure 3. Examples of mixed pixels (Fisher 1997).

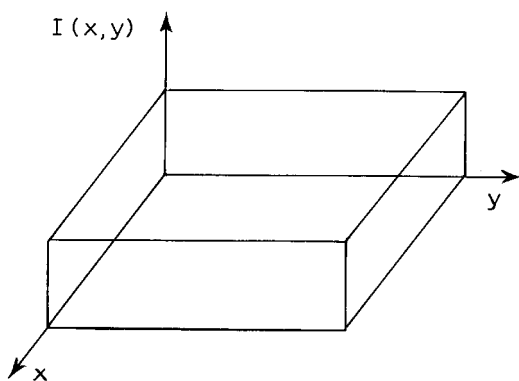


Figure 4. Idealised response function (point spread function).

and collect the radiation from within the IFOV, then jump to view the next adjacent IFOV and pause there to repeat the operation; however, in practice the scanner's mirror is rotating continuously and the output from the detectors is integrated over a time interval as the mirror rotates. For a push-broom scanner the simple description would suppose that the instrument was stationary while data for one scan line were collected and then jumped instantaneously to the position for the next scan line. This too is not the case; the instrument, situated on its spacecraft is moving uniformly along its path with constant speed. A further difficulty with our simple model is that it would mean that the scanner would respond uniformly to radiation from all points within the IFOV and would give no response to radiation from outside the IFOV, see figure 4. This also is not what really happens; there is a non-uniform response within the IFOV and there are also, in practice, contributions from outside the nominal IFOV. Thus we have a very complicated situation in practice:

- (i) We have seen in the previous section that, in practice, we are likely to have sources of radiation of different intensities at different points in the geometric IFOV (figure 3).
- (ii) But even if we did have a uniform spatial distribution of intensity throughout the IFOV the response of the sensor to a source of radiation with a given intensity will vary according to the location of the source within the field of view and there will also be a response to sources just outside the geometrical IFOV as well.

Looked at in another way, this means that the signal we attribute to any given pixel arises as a result of contributions not only from the field of view corresponding to that pixel but also includes contributions that properly belong to neighbouring pixels, see figure 5. In other words the pixel intensities are not independent but there is autocorrelation among them. There has been a discussion of this complicated situation published for the Landsat-MSS and TM by Markham (1985) and by a few other authors, while Breaker (1990) has attempted to remove autocorrelation from sea surface temperatures determined from thermal infrared AVHRR data. The arguments given by Markham for Landsat have been adapted by Cracknell (1997) for the AVHRR; we shall refer to this briefly now.

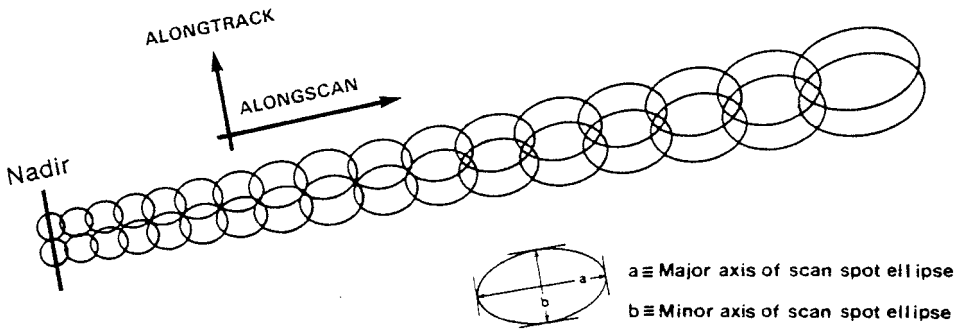


Figure 5. Sketch of pixel geometry for the AVHRR for adjacent scan lines to illustrate autocorrelation (Breaker 1990).

5. The point spread function, PSF

An Earth-observing spacecraft is so far away from the Earth that the object distance can be considered to be infinite and the image is formed in the focal plane. The optical point spread function, $PSF(x-u, y-v)$, describes the intensity as a function of position (x, y) in the focal plane arising from an object which is a point source with its geometrical image at the point (u, v) in the focal plane. Ideally the response should be constant for all points (u, v) within the geometrical IFOV and zero for all points (u, v) outside it, see figure 4. The sketch in figure 6 shows an example of the point spread function for a real system, namely that of the AVHRR.

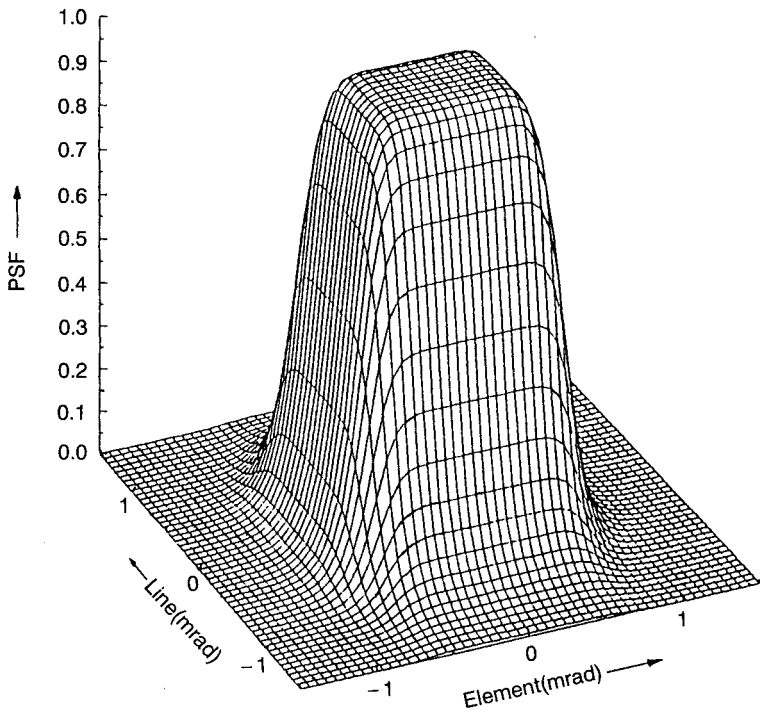


Figure 6. Sketch of point spread function for the AVHRR (Mannstein and Gesell 1991).

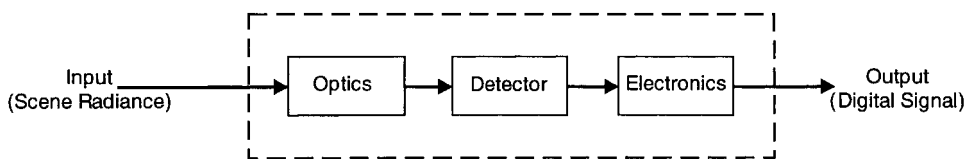


Figure 7. Schematic diagram of a typical spaceborne radiometer depicting the contributions to the point spread function (Breaker 1990).

For theoretical considerations it is convenient to separate the contributions to the point spread function into three stages, namely:

- the optical aspects
- the detectors
- the electronics

see figure 7. The final response of the system to a signal received from the surface of the Earth is the product of three separate responses corresponding to these three stages. We shall consider these three stages in turn.

The first stage is primarily concerned with diffraction effects. For an ideal distortion-free imaging system, namely the pinhole camera, the image of a point source would be a point and the location, (u, v) , of this image would be determined by simple geometry, i.e., by a straight line connecting the source to the pinhole and projected to the focal plane. If the pinhole is replaced by a lens system the image will become a circular diffraction pattern centred at the point located by a ray of light passing through the nodal point, i.e., the 'centre', of the lens. This diffraction pattern contributes to the optical factor in the point spread function. The passage of the light through further optical components will introduce further diffraction effects. In practice, however, one is not observing a point source. The object plane is a two-dimensional array of sources of various intensities. Each of these sources gives rise to a diffraction pattern that is (a) centred at the geometrical image position in the focal plane, and (b) of height (or maximum intensity) that is proportional to the intensity of the original source (neglecting atmospheric corrections). The total intensity at any point in the image plane arises as the sum of contributions from a large number of diffraction patterns which have their centres in the vicinity of the point in question. The second stage is concerned with the detectors, their shape, their location in the final image plane of the optics and their physical construction. The third stage is concerned with the electronic response of the detectors, in terms of the distribution of the radiation across the surface of the detectors, the electronics of the amplification, digitisation etc. stages. To calculate the effects of all these stages and obtain a meaningful value for the point spread function from these calculations is not feasible.

The approach used by Markham (1985) for Landsat and by Breaker (1990) for AVHRR is to work in the spatial frequency (or Fourier transform) domain and thus to write

$$G_0(v_x, v_y) = \text{OTF}(v_x, v_y) G_1(v_x, v_y) \quad (1)$$

where $G_0(v_x, v_y)$ and $G_1(v_x, v_y)$ are the two-dimensional spatial spectra (i.e., Fourier transforms) of the output and input scenes of the scanner, respectively, and $\text{OTF}(v_x, v_y)$, the optical transfer function, is the Fourier transform of the point spread

function. In practice it is not feasible to measure the two-dimensional point spread function directly because of the problems of producing point sources of significant intensity. Consequently the point spread function is expressed as a product of two line spread functions $LSF(x)$ and $LSF(y)$ in the x and y directions, i.e.

$$PSF(x,y)=LSF(x) \, LSF(y)$$

(2)

A line spread function is the response of the system to an infinitesimally narrow line source and is the integral of the point spread function in the direction of the line source. The determination of the line spread functions in the along-scan (cross-track) direction and the along-track direction for the AVHRR was carried out by Breaker (1990) and some of the effects of autocorrelation on satellite-derived sea surface temperatures were discussed.

With the simple response function of figure 4 all sources within the geometrically-defined field-of-view, which is a square, would contribute equally to the signal. In practice, points near to the nodal point N , the intersection of the optical axis with the object plane, contribute most strongly, points 0.25 km away contribute less strongly, points at the nominal boundary contribute even less strongly, but points beyond the nominal boundary also make some contribution. This response function is plotted in figure 6 and its values are given in figure 8; these values were obtained from NOAA or from the manufacturers of the AVHRR (ITT Aerospace) by Paithoonwattanakij a few years ago and included in his PhD thesis (Paithoonwattanakij 1989). The signal output by the AVHRR can therefore be thought of as the convolution of this function with the function representing the source intensity in the object plane. This function is of practical importance in a number of situations, particularly when considering objects that are very small in relation to the IFOV, when considering mixed pixels, when considering off-nadir viewing of the ground and in upscaling and downscaling. The function represented in figures 6 and 8 was derived from laboratory measurements made before launch. There is a further complication in flight that the AVHRR's scan mirror is rotating and (less importantly) the spacecraft is advancing along its orbit. The signal output by the instrument results from an integration over a period of time during which there will have been some movement of the field of view of the instrument relative to the surface of the Earth.

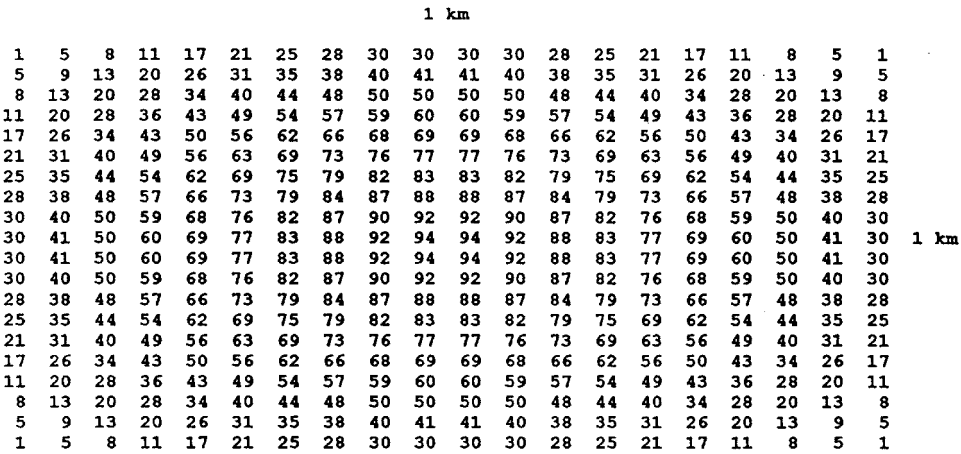


Figure 8. The point spread function for the AVHRR (Paithoonwattanakij 1989).

6. The calibration of detectors

One important problem that arises for instance, if one attempts to use the AVHRR data archive for global change studies is the question of the inter-calibration of the different instruments in the series. This is not just a question of the pre-launch calibration of the individual instruments in the series. It is also a question of post-launch calibration, to ensure that corrections to the pre-launch values of the calibration coefficients are made to allow for any changes in the sensitivity of the instrument arising (i) during storage before launch, (ii) during the launch process, or (iii) as a function of time in flight (see §2.2 of Cracknell 1997). The AVHRR technology dates from the 1970s and involves a mechanical scanner with a rotating mirror; the signals for each pixel along a scan line therefore are all generated by the same set of detectors. However, a new problem arose once push-broom scanners based on the use of a CCD array were introduced, as in the case of the HRV on SPOT. In this situation the signals generated for different pixels along a scan line are generated by different detectors in the (linear) CCD array and there is no guarantee that the sensitivity of every element is the same; indeed it is most unlikely that the sensitivity of every element is the same. Thus it becomes necessary to calibrate separately each CCD element, corresponding to each pixel in the scan line. If this is not done, then for simple image display work the image will have vertical striping, while if it is not done for quantitative work quite significant errors may be introduced in the values of any derived quantities.

7. Resampling

This represents a further complication in many applications of remotely-sensed data. In very many circumstances it is necessary to carry out geometrical rectification of an image. The image data gathered by a spaceborne or airborne scanner is obtained with pixels identified as elements in an array where row and column numbers are not immediately given with geographical coordinates attached to them. We set up a mathematical transformation (using ground control points (GCPs) to determine the coefficients in the transformation) and so we can refer to a pixel not just by its row and column number but by the latitude and longitude values of its centre. These latitude and longitude values will not bear any particular relation to the regular grid on which a map is based. Thus there is a problem. In printing an image, or in seeking to transfer features from an image to a map, or in attempting to compare different members of a set of images, it is common to use a grid set out in a regular square or rectangular network of geographical coordinates and not the oblique grid of points that comes about accidentally by rectification of the pixel coordinates in an image, see figure 9(a). Data from different AVHRR images would lead to a quite different grid (see the oblique grid in figure 9(b)). The generation of image intensity values at a regular square or rectangular grid of geographical coordinates, from the navigated grid, i.e., one of the oblique grids in figure 9, is an interpolation procedure which is usually described as resampling, and this is a very important procedure.

To consider the process of resampling we need to be careful about the notation. x and y are the geographical coordinates of a transformed pixel from the image. They correspond to the grid points of one of the oblique grids in figure 9; we are not free to choose their values. We also introduce another set of coordinates, which we shall denote by e , for the easting, and n , for the northing. These correspond to the grid points of the rectangular grid in figure 9(a) or (b) and we are free to choose

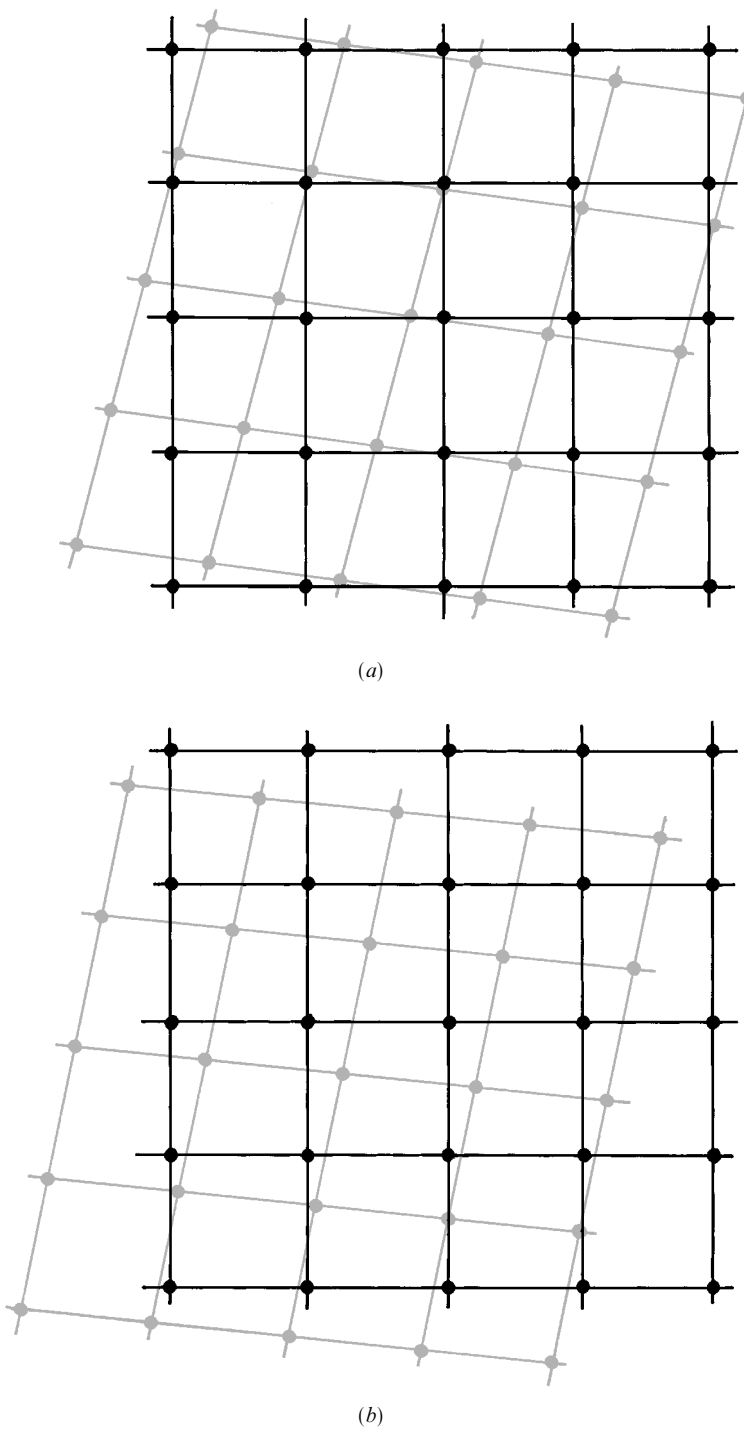


Figure 9. Diagram to illustrate resampling showing a rectangular map grid and, in (a) and (b), two oblique navigated grids of pixel centres from two different images.

some convenient set of these points. Suppose that one is trying to generate an image for comparison with, or superposition on, a given portion of a map. Then what one requires, in order to be able to print the image on a greyscale hardcopy device, is to be able to generate the image intensities $I'(e,n)$, in a given spectral channel, over the rectangular grid with each grid point specified by a pair of values of e and n , i.e., the chosen points of the rectangular grid in figure 9. Thus one needs to be able to work systematically through the ranges of values of e and n and for each pair of (e,n) values to generate $I'(e,n)$. For a given pair of values (e,n) we can apply the geometrical rectification equation to generate the corresponding pair of values (P,S) to locate the corresponding pixel in the original unrectified image. This point is denoted by Q in figure 10(a) where it is shown in relation to a portion of the

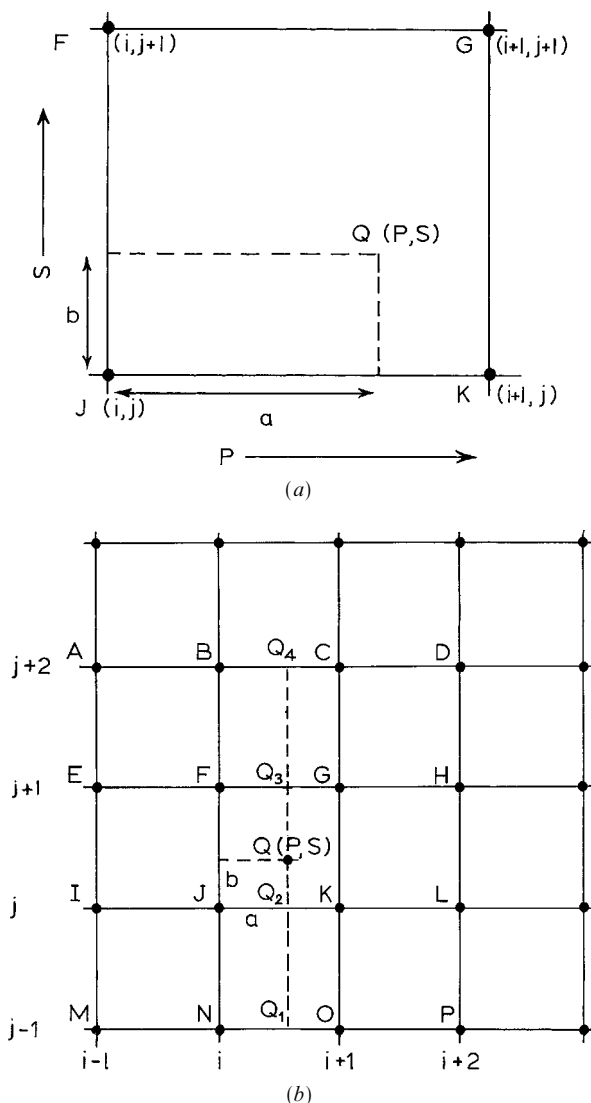


Figure 10. Grid for the definition of a resampling formula (a) for bilinear interpolation, and (b) for bicubic interpolation.

unrectified image. If P and S happen to be integers, then Q will coincide with the centre of one of the pixels in the unrectified image, i.e., one of the points FGJK in figure 10(a). In that (unlikely) case the pixel intensity $I(P, S)$ with those values of P and S is taken as the pixel intensity $I'(e, n)$ in that spectral channel in the rectified image. However, it is far more likely that the values of P and S calculated in this way will not be integers and in this situation it is necessary to use an interpolation method. There are several interpolation methods which can be used to estimate $I'(e, n)$, the pixel intensity at the point (e, n) in the rectified image; these include

- the nearest neighbour method
- the bilinear interpolation method
- the bicubic interpolation method.

In the nearest neighbour method one just takes the intensity $I(P, S)$ for the point nearest to Q ; in figure 10(a) this would be the point K. This method is fast to use and does not lead to loss of information by smoothing. However, spurious effects may appear as a result of the repetition or omission of pixels, particularly if there is a large difference between the spacings of the two grids.

The linear interpolation method is based on the assumption that a surface of intensity, as a function of P and S , would be a plane over the region FGKJ in figure 10(a). Bilinear interpolation thus has a smoothing effect and sharp boundaries may, therefore, become blurred. It also involves more computer time than the nearest neighbour method.

The bicubic interpolation method involves fitting a pair of third-degree polynomials in the region surrounding the point Q (P, S). The 16 nearest pixels in the unrectified image are used, see figure 10(b). The bicubic interpolation method is both more complicated to program and also more demanding in terms of computer time than the previous two methods; however, it is a quite popular method. It avoids the oversimplistic nearest neighbour method which can lead to blockiness and it avoids the excessive smoothing produced by the bilinear method. However, it does lead to some loss of high-frequency information.

The choice of which of these three methods to use will depend on two factors, the use to which the data will be put and the computer facilities that are available. If the image is to be subjected to classification then the replacement of the raw data by the interpolated data may well have some adverse effect on the final classification because, as we have already noted, the interpolation involves some smoothing of the data. It may, therefore, be decided that it would be better to perform the classification on the raw data and to perform the geometrical rectification subsequently.

8. Sub-pixel accuracy

If one uses a well-distributed set of GCPs one can expect to achieve an accuracy of geometrical rectification with an error (mean standard deviation) of rather less than the length of the edge of the instantaneous field-of-view (IFOV). A standard deviation of, say, 80–90 per cent of the edge of the IFOV is commonly regarded as a reasonable achievement. In some circumstances it may be highly desirable to achieve a considerably better accuracy. Let us consider just one example, briefly. One problem that we have studied is ocean circulation and, in particular, the determination of current vectors (i.e., current speed and direction) by feature tracking using sequences of AVHRR thermal infrared images. With rectification of individual scenes to ± 0.8 km by conventional rectification with GCPs one can determine

current vectors with an error of $\pm 1.8 \text{ cm s}^{-1}$. Typical magnitudes of the current itself may be within a range of about $2\text{--}50 \text{ cm s}^{-1}$ so that $\pm 1.8 \text{ cm s}^{-1}$ may correspond to quite a large percentage error in the calculated current velocity. If the error of the geometrical rectification of each individual AVHRR scene can be reduced to about $\pm 0.2 \text{ km}$ (i.e., to about ± 20 per cent of the edge of the IFOV) then the percentage error in the derived current vectors could be reduced very significantly.

A method for achieving this level of accuracy with AVHRR data was developed by Cracknell and Paithoonwattanakij (1989). This method for AVHRR data was an adaptation of a method developed previously by Torlegård (1986) for very accurate registration of Landsat-MSS images. Torlegård had demonstrated that it was possible, using digitised small extracts (or chips) from air photos, to achieve geometrical rectification of a Landsat-MSS image with an error of only 20 per cent of the edge of the IFOV.

The procedure adopted in Torlegård's method for Landsat-MSS data was as follows. First select a GCP, which has known ground coordinates, from a digitised air photograph and extract an area from the digitised air photo surrounding the GCP; this extract is called a control point chip (CPC). Secondly, find the image coordinates of the resampled CPC. Thirdly, find the image coordinates in the weighted resampled CPC by using a weight matrix that was obtained from the point spread function of the Landsat-MSS. Finally, use this CPC to achieve registration of the Landsat-MSS data. For more details see Torlegård (1986).

In the adaptation of this method for rectifying AVHRR data (Cracknell and Paithoonwattanakij 1989) one needs to establish a set of CPCs of Landsat-MSS data for the GCPs that are used for the AVHRR scene that is to be rectified. Landsat-MSS band 4 ($0.8\text{--}1.1 \mu\text{m}$), which is in the near-infrared was chosen because it is the best suited for land and sea discrimination. For convenience the Landsat-MSS data were resampled to pixels corresponding to 50 m by 50 m . First of all it is necessary to perform an approximate rectification and this was done with an orbital model. Thus, for each GCP a search window can be defined from this approximate rectification. In an automatic rectification using the Landsat CPCs, the CPC is moved by only one-tenth of an AVHRR pixel edge between matching attempts. For each trial position of the CPC in the window the MSS data are degraded to AVHRR resolution, using the AVHRR point spread function, and the correlation with the AVHRR image is evaluated. The best correlation between the CPC and the window in the unrectified data can therefore be identified to an accuracy of one-tenth of a pixel edge rather than to the nearest pixel. With a set of GCPs located in the image to this accuracy it is then possible to achieve a rectification of the whole image to sub-pixel accuracy. The method was tested for an AVHRR extract covering the U.K. with only 7 GCPs and even with such a small number of GCPs it was possible to achieve a rectification with a standard deviation of only 26 per cent of the pixel spacing.

9. Data compression or upscaling

In the present context we consider two questions in relation to data compression:

- (i) What do we lose in the way of information if we carry out data compression?
- (ii) If we are concerned with upscaling what is the best way in which to do it so as to minimise the loss of information?

It may seem strange that we should consider data compression in this paper because the main theme seems to be centred around the fact that the spatial resolution

of remotely-sensed data is often poorer than we would like it to be. Why then consider degrading the data to even lower resolution? The answer lies in the fact that it may be necessary, for one reason or another, to reduce the amount of data that is transmitted or recorded. One good example is provided by the AVHRR where this happens on board. There are the 1.1 km resolution HRPT (High Resolution Picture Transmission) or the LAC (Local Area Coverage) data that are generated by the AVHRR. There is also the degraded GAC (Global Area Coverage) data which have been degraded in a certain way; along the scan line there is a pixel averaging, but in the flight direction there is subsampling. What happens is that two channels, call them A and B, are selected from the AVHRR's five channels. The data from one scan line and channel A are degraded along the line and transmitted. The data from the next line and channel B are similarly treated. No data are used from the third scan line and then the process is repeated. Thus for each channel only one scan line in three is used. Moreover, the two channels are thus *not strictly coregistered*. Clearly the GAC data are therefore *not* the same as would be obtained if one took HRPT data and simply carried out spatial averaging in two dimensions.

It may sound obvious, but it still needs to be said. *What you get depends on how you carry out the compression or upscaling.*

Let us consider another example of a situation in which this problem occurs. General circulation models—for weather forecasting or climate modelling—are now becoming sufficiently sophisticated as to require realistic input data for the nature of the land surface over the globe. However, this is usually done on a 50 km or 100 km grid spacing and one way of obtaining the information about land surface cover is by upscaling the 1.1 km data from the AVHRR. However, what method should be used for this upscaling? And how can the results of different methods be tested to try to find out which is the best method? The trouble is that in many cases there is no direct test of the method. Perhaps I should just mention some work from the 1996 COSPAR meeting in Birmingham by Gupta, Prasad and Krishna Rao; this was concerned with using LISS-II (IRS) 36.25 m resolution scanner data and comparing it with LISS-I *simultaneously gathered* 72.5 m resolution data. However, they found that no single method was best at preserving the integrity of the data for all different cover types.

10. Mixed pixels

Rather than try to give a comprehensive study of all kinds of mixed pixels, we consider a rather simple situation and this concerns sub-pixel intense sources of heat and the use of channel-3 AVHRR data.

Channel 3 of the AVHRR is a relatively neglected source of data. There are two main reasons for this. The first is the quality of the data. AVHRR data in channels 1, 2, 4 and 5 is of very good quality; it has low noise and it is free of the striping that is familiar, for example, to users of some of the older Landsat data. This is not the case, however, for channel-3 data. Channel-3 images often have a very pronounced herring-bone pattern. This noise is worse in the data from some instruments in the series than from others. The second reason for the relative unpopularity of channel-3 data is the fact that in daytime the radiation received in channel 3 is a mixture of reflected solar radiation and emitted radiation. This makes the interpretation of the data more complicated than is the case for data from the other channels. In channels 1 and 2, except in one or two highly exceptional circumstances (e.g., gas flares or erupting volcanoes), one is dealing entirely with reflected solar radiation. This is in

the visible and near-infrared region of the electromagnetic spectrum and so the interpretation of the satellite-received signals can be carried out in terms of the reflectivity of the surface at these wavelengths. In channels 4 and 5 one is dealing with almost exclusively (again except in very exceptional circumstances, in this case with sunglint) infrared radiation that is emitted by the surface of the Earth. The fact that channel-3 data contains something approaching a 50:50 mixture of reflected solar radiation and emitted radiation during the day-time means that it is (not surprisingly) more difficult to interpret than the data from the other channels. At night there is of course only the emitted radiation and consequently night-time channel-3 data over sea areas is widely used in multi-channel sea surface temperature algorithms.

Let us turn to the consideration of small intense sources of heat. A fire that generates a large smoke plume is very often clearly visible in channel-1 AVHRR data in which the smoke plume shows up very clearly. It is possible, however, for fires that are considerably smaller than the IFOV (circa 1.1 km by 1.1 km) to be quite apparent in channel-3 data. Examples include gas flares, blast furnaces and agricultural straw fires (Muirhead and Cracknell 1984, 1985, 1986, Djavadi and Cracknell 1986, Saull 1986, Cracknell and Saradjian 1996). The ability of the channel-3 data to detect sub-pixel size high temperature sources, or hot spots, has been well documented in studies of gas flares, forest and range fires, and steel production plants. Hot spots in channel-3 imagery can usually be distinguished from pixel error or highly reflective cloud by their characteristic point-spread pattern or, if necessary, by comparison with other AVHRR channels.

The principle of what is involved in fire detection with AVHRR data has been considered by Matson and Holben (1987). At typical room, land or sea temperatures, say ~ 300 K, the Planck distribution function peaks in the vicinity of AVHRR channel-4 and channel-5 wavelengths, i.e., 10–12 μm . For a flame temperature, say 600–700 °C or 900–1000 K, the peak moves to the vicinity of channel-3 wavelengths, i.e., 3.5–3.9 μm . In channel 3 a flame that is considerably smaller than the size of the IFOV will give an appreciable contribution to the signal and one can use an approach developed by Dozier (1981) and applied to fires by Matson and Dozier (1981) to estimate both the area and temperature of the hot target. In the absence of an atmospheric contribution or attenuation, the upwelling radiance sensed by a downward-pointing radiometer is given by

$$L(T) = \int_0^\infty \varepsilon_\lambda B(\lambda, T) \phi(\lambda) d\lambda / \int_0^\infty \phi(\lambda) d\lambda \quad (3)$$

where $B(\lambda, T)$ is the Planck distribution function, $\phi(\lambda)$ is the spectral response of the detector and the target area is assumed to be at a uniform temperature, T . For most Earth surfaces ε_λ is relatively independent of λ over the range of an AVHRR channel, so that one can drop the λ subscript and move ε outside the integral in this equation.

Suppose that we now have a mixed pixel corresponding to a hot target at temperature T_i which occupies a fraction p of the IFOV (where $0 < p < 1$) and a background temperature T_b which occupies the remaining fraction $(1 - p)$ of the IFOV. Strictly speaking, we should (a) specify the location of the hot target within the IFOV, and (b) use the point spread function for the AVHRR, but we ignore that

just now. The brightness temperatures T_j ($j=3, 4$) sensed by the AVHRR in channels 3 and 4 will be, in the absence of an atmospheric contribution or attenuation (and at night so that there is no reflected solar radiation in channel 3),

$$T_j = L_j^{-1} [pL_j(T_t) + (1-p)L_j(T_b)] \quad (4)$$

where $j=3, 4$. The background temperature can be estimated reasonably accurately from nearby pixels. This means that in the two equations (4) ($j=3, 4$) there are just two unknowns, p , the fraction of the IFOV occupied by the hot target, and T_t , the target temperature.

Using the theory described, Matson *et al.* (1987) were able to estimate the areas and temperatures of two high-temperature sources located in Idaho. Typical temperature differences between the two channels over land surfaces in general are usually about 1–2 degK. In the data studied, target 1 was a small controlled forest fire and target 2 was a phosphorus plant. At these sources, the channel-3 brightness temperatures were 16.2 degK and 33.9 degK higher than the corresponding channel-4 brightness temperatures. These targets are smaller than the IFOV and Matson *et al.* found the area and temperature were 0.28 ha and 430 K for target 1 and 1.7 ha and 483 K for target 2. As evidenced by these calculated target sizes, it does not need a 1.1 km square target to cause a response in channel 3. (A 1.1 km square has an area of $((1.1 \times 10^3)^2/10^4)$ ha = 121 ha.) Thus we see an example which shows that small subresolution scale high-temperature sources, such as fires, can be detected and studied using channel-3 data. Some further discussion of the use of channel 3 AVHRR data in connection with fire studies is given in chapter 6 of the book on the AVHRR (Cracknell 1997).

Another example of the observation of small intense sources of heat with data from channel 3 of the AVHRR is related to the identification and location of gas flares in the North Sea by Muirhead and Cracknell (1984). During 1980, for example, gas was flared from oilfields in the British sector of the North Sea at an average rate of over 11 Mm³ per day. There are a large number of platforms in the North Sea and some of them are clustered quite close together. The objective was to see whether it was possible to carry out geometrical rectification to sufficient accuracy to enable one to identify unambiguously the particular oil platform responsible for each flare observed in the satellite data. This work involved careful geometrical rectification of each scene, to an accuracy of better than 1 km; this was done for 11 scenes and it was possible to show that each flare could be assigned unambiguously to a platform that could be held responsible for it.

In work reported in several papers we have studied the possible use of AVHRR channel-3 data in relation to straw burning by farmers (Muirhead and Cracknell 1985, 1986, Saull 1986, Djavadi and Cracknell 1986, Cracknell and Saradjian 1996). As a result of changing agricultural practices, Britain finds itself with a large straw surplus with no established market. To dispose of it, in the 1980s many British farmers resorted to the option of burning the surplus straw, in spite of the fact that soil incorporation is perfectly viable and costs little more than burning or that, with a modest investment of capital, large quantities of straw could be utilised as a fuel for domestic or greenhouse heating. As well as being wasteful, straw burning can occasionally prove hazardous. Fire damage to trees, hedges and even buildings often results from careless field burning, while road accidents have been known to occur as a consequence of the smoke produced. Muirhead and Cracknell (1985) used channel-3 AVHRR data to investigate the extent of straw and stubble burning across

Britain in the summer of 1984. Results for a typical weekday in August 1984, after eliminating hot spots due to installations such as steel works and gas flares from oil refineries, depict about 350 burning fields. Straw burning has subsequently been made illegal in Britain and the analysis by Cracknell and Saradjian (1996) of data from the summer of 1995 showed that the practice of straw burning had been reduced since 1984 but had not yet been eliminated completely.

In the light of the above discussion it is interesting to address the question of what is the limit (in terms of both temperature and size) to the detection of a hot spot and to the analysis in terms of mixed pixels that we have given for the determination of the size and temperature of a hot spot. This has been studied by Robinson (1991). There is, however, another aspect and this is that the sensor must not become saturated and, in fact, this can occur. The saturation problem has been addressed by Setzer and Verstraete (1994) and references are given in the book by Cracknell (1997).

As another example of a study of mixed pixels, or mixels, we mention some work on paddy fields in Japan by Okamoto and Fukuhara (1996) which involved using Landsat-TM data. In this situation the mixing involves the paddy fields, which tend to be very homogeneous but very small, along with farm roads, irrigation canals, etc. An outline of their method is given in figure 11.

11. The study of sub-IFOV size objects

In satellite remote sensing we usually tend to think that the smallest object which we can reasonably expect to detect is an object which is comparable in size with the IFOV of the instrument being used. We also tend to assume that there is good contrast between an object and its surroundings. Thus for the AVHRR we would, initially, not expect to be able to detect objects with linear dimensions less than about 1 km. This would seem to rule out the study of any animals, birds or insects with AVHRR data. However, the discussion that we have just given of small intense sources of heat has provided one example of a situation in which one can detect sub-pixel (or sub-IFOV) size objects. In this section we shall consider other situations in which it is possible to obtain information about objects that are smaller than the size of the IFOV; examples of such objects include cattle, sheep, locusts, mosquitoes and tsetse flies. A single specimen of any one of the creatures just mentioned is far too small to be detectable in an AVHRR signal. However, concentrations of small creatures sometimes do build up to be so large that these collections become directly detectable in AVHRR data. One example is provided by the case of blooms of blue-green algae and an example in which the evolution of such blooms off the Norwegian coast has been described in the literature (Johannessen *et al.* 1989).

One example of a collection of small creatures becoming sufficiently large to become directly detectable in AVHRR data is that of a swarm of locusts. The swarm may be so large and contain so many locusts that it could be expected to be seen as a dark cloud in a satellite image. However, the direct observation in a satellite image of a swarm of locusts is not in practice very useful. By the time the swarming has occurred, let alone its existence been detected in satellite imagery, it is too late to take any effective action to prevent crop devastation by the swarm. What would be useful would be to use AVHRR to study the locusts' habitats and to detect the conditions that lead to swarming before it occurs. Then it is possible to take effective action, i.e., spraying, to control the locusts and prevent their swarming. A second example is provided by mosquitoes. These are smaller than locusts and never form

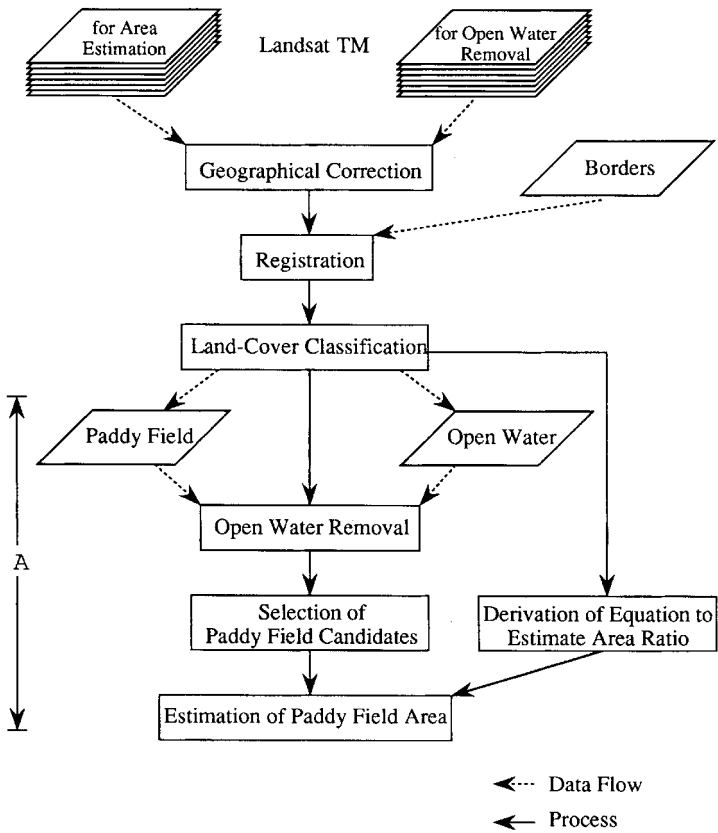


Figure 11. Procedure of Okamoto and Fukuhara (1996) to estimate the paddy field area with area ratio in each pixel; A is the methodological part of the area estimation using the paddy field area ratio in each pixel.

clouds that would be large enough to detect with AVHRR data anyway. However, in this case too it has been shown to be possible to use AVHRR data to study the mosquitoes' habitats and to predict the behaviour of the insects from these studies (Linthicum *et al.* 1987, Wood *et al.* 1991). Thirdly, movements of larger animals such as cattle and sheep can sometimes be followed on large areas of open rangeland by correlating their behaviour with the state of the vegetation as indicated by the NDVI. This has been done with Landsat-TM data for sheep in the uplands of North Wales in the U.K. by Thomson and Milner (1989). This area is well covered with a variety of semi-natural grassland and heathland vegetation. Sheep are the only large herbivores that are present in large numbers in this area and so sheep grazing is the dominant biotic factor that controls these ecosystems. Different vegetation communities support different population densities of sheep and the soil, rainfall and altitude affect sheep distribution. These factors also affect the satellite-received radiance and therefore also the vegetation index as well. This work on North Wales was done with Landsat-TM data; however, the general idea can be applied to large rangeland areas for which AVHRR data would be more appropriate. For example, the monitoring of the phenology of Tunisian grazing lands using NDVI data from the

AVHRR has been carried out by Kennedy (1989). We shall consider the examples of locusts and tsetse flies in a little more detail.

A considerable amount of work has been done on locusts using Landsat and AVHRR data. As we have already mentioned, the detection of a swarm of locusts as a black cloud on a satellite image or the detection of the massive devastation caused to an agricultural area by a swarm of locusts is not very useful; what one has to do is to try to predict potential swarming of the locusts and take appropriate preventive action before damage to crops occurs. When the locusts are present in low numbers in their natural desert habitat or recession area they present no problem. When rainfall occurs the locusts multiply and live for a while on the ephemeral vegetation that appears after the rainfall. It is when the locust population has grown vastly and the local vegetation is exhausted that the locusts swarm and migrate, often over very large distances, in search of further supplies of vegetation, see figures 12 and 13. The strategy of plague prevention is thus based on locating areas where rain has fallen, monitoring the locust population build up and then controlling the population by spraying where necessary. Given that the recession areas are large and generally not well supplied with meteorological stations on the ground, satellite remote sensing provides an important source of data for monitoring rainfall in these areas. There are techniques based on the study of clouds in satellite images to locate precipitation. One can also detect precipitation, after it has occurred, from its observable effects on the soil moisture and on the green-vegetation biomass. The study of soil moisture from satellite, though possible in principle using microwave or thermal infrared data, is not yet well developed. Therefore, the use of visible-band and near-infrared-band data and of the vegetation index derived from these data is very important in this context and the two references which are cited in the captions to figure 12 and 13 are devoted to studying this; although they deal with quite different geographical areas and with different species of locusts, the general ideas have much in common. Tucker *et al.* (1985) were concerned with the desert locust (*Schistocerca gregaria* Forsk.), while Bryceson (1989) was concerned with the

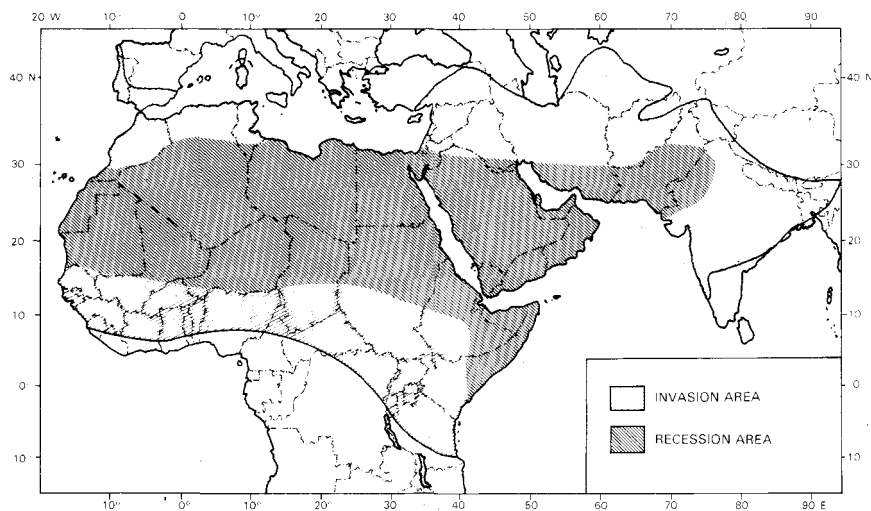


Figure 12. The invasion and recession areas of the desert locust (Tucker *et al.* 1985).

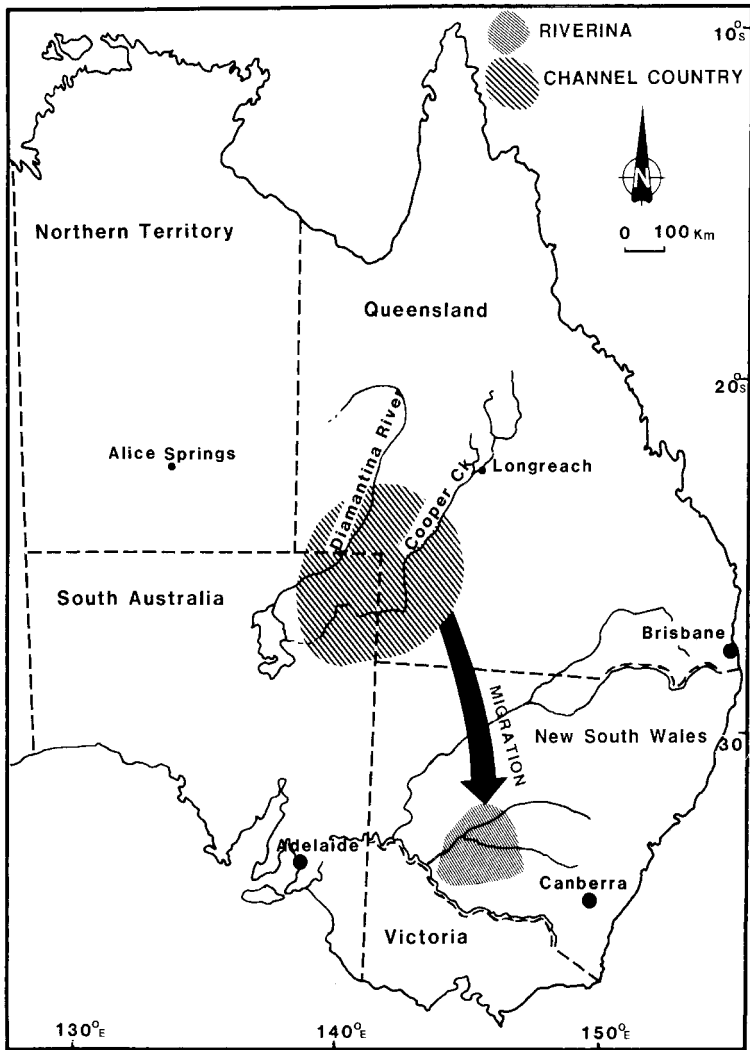


Figure 13. The invasion and recession areas of the Australian plague locust (Bryceson 1989).

Australian plague locust (*Chortoicetes terminifera* (Walker)) which affects areas of Queensland, South Australia and New South Wales.

The use of remotely-sensed data in studying tsetse fly habitats has been described by Rogers and Randolph (1991, 1993) and Rogers and Williams (1993, 1994). The tsetse fly is a massive blight on life in Africa. Tsetse flies are a major constraint on animal production in about 10 M km² of Africa through their transmission of animal trypanosomiasis, while up to 25 million people are at risk from human trypanosomiasis, or sleeping sickness. The scale of the problem is vast. Africa contains about 20 per cent of the world's pasture land, yet raises only 10 per cent of the world's cattle and produces only 3 per cent of the world's meat and milk. Indeed it has been argued that it was the fact that the animal trypanosomiasis, carried by the tsetse fly, precludes the use of draught animals for ploughing, which prevented the occurrence in tropical Africa of the agricultural revolution that occurred in the Middle East

about 4000 B.C. Ever since the link between the tsetse fly and the disease was discovered, countries affected have tried to control both. To control the tsetse fly over large areas requires a good knowledge of the flies' distribution so that elimination techniques can be most effectively targeted. A very interesting account of the biological background to the problem is given in the paper by Rogers (1991), including a detailed account of the role of climate in determining tsetse distribution. As Rogers puts it 'water relationships, ... remain the Achilles heel of many insects and, directly or indirectly, are thought to limit insect distributions in dry environments'. A similarity was observed between whole-Africa NDVI maps and tsetse distribution maps and this led to the detailed study of the correlation between NDVI and tsetse fly distribution. Rogers and Randolph (1991) observed a correlation between tsetse fly mortality rates and the monthly mean NDVI (of the previous month) for two areas (one in Nigeria and one in Uganda). They also found, using a large dataset for the northern half of the Côte d'Ivoire from 1979–1980, a correlation between tsetse fly abundance and mean annual NDVI. This work suggested that NDVI data can be used to predict tsetse fly abundance. The long-term aim is to produce maps of areas of high-risk of trypanosomiasis transmission for the very large areas of tropical Africa where diseases carried by vectors such as the tsetse fly affect human and animal welfare. The use of the NDVI to provide information on two key parameters of vector-borne diseases, namely vector mortality rate and vector abundance, is a significant step towards this goal.

12. Conclusion

The conclusion of this paper has to be to state a warning, namely that the factors that contribute to producing the digital number DN in a computer tape or a disk file of a remotely-sensed image are complicated and inter-related. This must be remembered and taken account of in any attempt at synergy in remote sensing, whether it is synergy between different sets of remotely-sensed data or synergy between remotely-sensed data and other sources of information. We conclude by mentioning a Letter that was recently published by Fisher (1997); it is called 'The pixel: a snare and a delusion'. That title seems to convey the message or warning that I am trying to convey in the present paper. Fisher is trying to make the same general point that I have been trying to make, namely that in remote sensing the pixel is much more complicated than one might, at first sight, imagine. We have not resolved all the problems; the first task is to make sure that people realise that there are problems.

References

- BREAKER, L. C., 1990, Estimating and removing sensor-induced correlation from Advanced Very High Resolution Radiometer Satellite data. *Journal of Geophysical Research*, **95**, 9701–9711.
- BRYCESON, K. P., 1989, The use of Landsat MSS data to determine the distribution of locust eggbeds in the Riverina region of New South Wales, Australia. *International Journal of Remote Sensing*, **10**, 1749–1762.
- CRACKNELL, A. P., 1997, *The Advanced Very High Resolution Radiometer* (London: Taylor and Francis).
- CRACKNELL, A. P., and PAITHOONWATTANAKIJ, K., 1989, Pixel and sub-pixel accuracy in geometrical correction of AVHRR imagery. *International Journal of Remote Sensing*, **10**, 661–667.

- CRACKNELL, A. P., and SARADJIAN, M. R., 1996, Monitoring of strawburning in the U.K. using AVHRR data—Summer 1995. *International Journal of Remote Sensing*, **17**, 2463–2466.
- DJAVADI, D., and CRACKNELL, A. P., 1986, Cloud cover and the monitoring of strawburning using AVHRR data. *International Journal of Remote Sensing*, **7**, 949–951.
- DOZIER, J., 1981, A method for satellite identification of surface temperature fields of subpixel resolution. *Remote Sensing of Environment*, **11**, 221–229.
- FISHER, P., 1997, The pixel; a snare and a delusion. *International Journal of Remote Sensing*, **18**, 679–685.
- JOHANNESSEN, J. A., JOHANNESSEN, O. M., and HAUGHAN, P. M., 1989, Remote sensing and model simulation studies of the Norwegian coastal current during the algal bloom in May 1988. *International Journal of Remote Sensing*, **10**, 1893–1906.
- KENNEDY, P. J., 1989, Monitoring the phenology of Tunisian grazing lands. *International Journal of Remote Sensing*, **10**, 835–845.
- LINTHICUM, K. J., BAILEY, C. L., DAVIES, F. G., and TUCKER, C. J., 1987, Detection of Rift Valley fever viral activity in Kenya by satellite remote sensing imagery. *Science, New York*, **235**, 1656–1659.
- MANNSTEIN, H., and GESELL, G., 1991, Deconvolution of AVHRR data. *Proceedings of 5th AVHRR Data Users' Meeting, Tromsø, Norway, 25–28 June 1991*, EUM P 09 (Darmstadt-Eberstadt: EUMETSAT), pp. 53–58.
- MARKHAM, B. L., 1985, The Landsat sensors' spatial responses. *IEEE Transactions on Geoscience and Remote Sensing*, **23**, 864–875.
- MATSON, M., and DOZIER, J., 1981, Identification of subresolution high temperature sources using a thermal IR sensor. *Photogrammetric Engineering and Remote Sensing*, **47**, 1311–1318.
- MATSON, M., and HOLBEN, B. N., 1987, Satellite detection of tropical burning in Brazil. *International Journal of Remote Sensing*, **8**, 509–516.
- MATSON, M., STEPHENS, G., and ROBINSON, J., 1987, Fire detection using data from the NOAA-N satellites. *International Journal of Remote Sensing*, **8**, 961–970.
- MUIRHEAD, K., and CRACKNELL, A. P., 1984, Identification of gas flares in the North Sea using satellite data. *International Journal of Remote Sensing*, **5**, 199–212.
- MUIRHEAD, K., and CRACKNELL, A. P., 1985, Straw burning over Great Britain detected by AVHRR. *International Journal of Remote Sensing*, **6**, 827–833.
- MUIRHEAD, K., and CRACKNELL, A. P., 1986, see Saull (1986).
- OKAMOTO, K., and FUKUHARA, M., 1996, Estimation of paddy field area using the area ratio of categories in each mixel of Landsat-TM. *International Journal of Remote Sensing*, **17**, 1735–1749.
- PAITHOONWATTANAKIJ, K., 1989, Automatic pattern recognition techniques for geothermal correction on satellite data, PhD thesis, Dundee University.
- ROBINSON, J. M., 1991, Fire from space: Global fire evaluation using infrared remote sensing. *International Journal of Remote Sensing*, **12**, 3–24.
- ROGERS, D. J., 1991, Satellite imagery, tsetse and trypanosomiasis in Africa. *Preventative Veterinary Medicine*, **11**, 201–220.
- ROGERS, D. J., and RANDOLPH, S. E., 1991, Mortality rates and population density of tsetse flies correlated with satellite imagery, *Nature*, **351**, 739–741.
- ROGERS, D. J., and RANDOLPH, S. E., 1993, Distribution of Tsetse and Ticks in Africa: Past, Present and Future. *Parasitology Today*, **9**, 266–271.
- ROGERS, D. J., and WILLIAMS, B. G., 1993, Monitoring trypanosomiasis in space and time. *Parasitology*, **106**, S77–S92.
- ROGERS, D. J., and WILLIAMS, B. G., 1994, Tsetse distribution in Africa: seeing the wood and the trees. In *Large-Scale Ecology and Conservation Biology*, edited by P. J. Edwards, R. May and N. R. Webb (Oxford: Blackwell), pp. 247–271.
- SAULL, R. J., 1986, Strawburning over Great Britain detected by AVHRR: a comment. *International Journal of Remote Sensing*, **7**, 169–172; includes reply by K. Muirhead and A. P. Cracknell.
- SETZER, A. W., and VERSTRAETE, M. M., 1994, Fire and glint in AVHRR's channel 3: a possible reason for the non-saturation mystery. *International Journal of Remote Sensing*, **15**, 711–718.

- THOMSON, A. G., and MILNER, C., 1989, Population densities of sheep related to Landsat Thematic Mapper radiance. *International Journal of Remote Sensing*, **10**, 1907–1912.
- TORLEGÅRD, A. K. I., 1986, Some photogrammetric experiments with digital image processing. *Photogrammetric Record*, **12**, 175–186.
- TUCKER, C. J., HIELKEMA, J. H., and ROFFEY, J., 1985, The potential of satellite remote sensing of ecological conditions for survey and forecasting desert-locust activity. *International Journal of Remote Sensing*, **6**, 127–138.
- WOOD, B. L., BECK, L. R., WASHINO, R. K., PALCHICK, S. M., and SEBESTA, P. D., 1991, Spectral and spatial characterization of rice field mosquito habitat. *International Journal of Remote Sensing*, **12**, 621–626.