Mon. Not. R. astr. Soc. (1985) **214**, 575–604

Automatic analysis of crowded fields

M. J. Irwin *Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*

Accepted 1985 January 29. Received 1985 January 28; in original form 1984 September 3

Summary. There are many important 2D data reduction problems in astronomy that are not amenable to conventional automatic analysis. Typically these problem areas arise in fields where the number density of images is high and where the local sky background may vary rapidly. Within such regions the image number density becomes so high that the majority of images overlap, even at relatively high isophotes, and simple image parameter estimation algorithms become confused. Our goal has been to examine the potential for a fully automatic method that is both robust and efficient in terms of computer requirements, capable of dealing with complex multiple overlaps and able to generate the optimum estimates of image parameters. By applying the theory of maximum likelihood parameter estimation to this topic, we have been able to devise a coherent strategy for tackling the problem. This has led to the development of a fully automatic system capable of producing reliable results toward the centre of globular clusters and within the dense regions of nearby resolved galaxies.

1 Introduction

Accurate automatic determination of the parameters of astronomical images is a fundamental problem in modern astronomy. Given parameter lists of say position, intensity and the necessary information to perform elementary classification of images into stars and galaxies, it is then relatively straightforward to study the distribution, variability, colour and proper motion of objects for a wide variety of astronomical problems. Most modern observational techniques either record the observations digitally, as a two-dimensional array of pixel intensities (CCD detectors, Vidicon cameras, etc.), or initially store the information photographically and then use a high-speed microdensitometer to convert the data to digital form. Indeed much of the output from the *Space Telescope* will be in the form of digitized pictures (van Altena, Franz & Frederick 1974). All of these methods have in common the need at some stage in the processing to analyse, preferably automatically, two-dimensional arrays of pixel intensities.

However the analysis is made, whether it be manually, semi-automatically or fully automatically there are several essential stages to follow. First each image, or image blend, must be detected against the local sky background. An accurate unbiased estimate of the background must be made and some form of image parameterization is required, including if possible partitioning any blended image into its constituent parts. A minimum requirement of the image parameters is to provide a measure of the intensity of the image, coordinates of the centre and an indication of the overall image shape. Many systems with this basic motivation have been described (Kibblewhite *et al.* 1979 – APM; Bijaoui 1979 – CDCA; Stobie *et al.* 1979 – COSMOS; Newell 1979 – APEX; Herzog & Illingworth 1977; Kron 1980; Jarvis & Tyson 1981 – FOCAS; Chiu, van Altena & Stetson 1979).

Unfortunately although many of the above methods work well when image crowding is not too severe there are many important cases when they do not. Typically problems will occur in fields where the number density of images is high and where the local sky background may vary rapidly. Examples of such crowded fields are to be found in globular clusters, nearby resolved galaxies or even in deep frames of data well away from obvious concentrations of objects (Hall & Mackay 1984). Within such regions the image number density becomes so high that the majority of images overlap, even at relatively high isophotes, and simple analysis algorithms become confused. The confusion can be so bad that even an apparently simple objective like number counting will fail (Irwin & Trimble 1984).

To illustrate the realm of number densities where crowding effects become noticeable consider the number density corrections shown in Fig. 1. The corrections are for three types of analysis: human interpretation of a photographic plate, isophotal method (e.g. APM/COSMOS/FOCAS/CDCA) and an idealized crowded field algorithm, the realization of which will be discussed later, where it is assumed possible to distinguish between independent images at separations greater than their seeing-disc size. This is basically a Rayleigh separation criterion. The corrections to first order

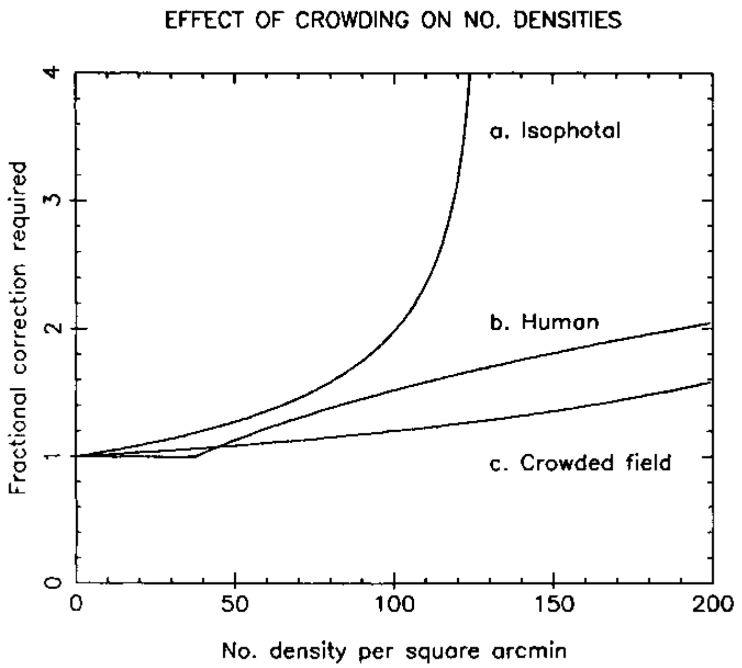


Figure 1. A comparison of the crowding corrections from King *et al.* (1968) with those appropriate for machine measurement, Irwin & Trimble (1984). (a) Crowding correction for isophotal analysis with average isophotal area roughly three times seeing-disc area. (b) Crowding correction for human observer from King *et al.* (1968) algorithm with image area equal to seeing-disc area. (c) Crowding correction for crowded field algorithm with image area equal to seeing-disc area, otherwise same formula as (a) (see Irwin & Trimble 1984). The curves are derived for moderate seeing conditions (~1.5 arcsec) and the value of the ordinate gives the correction to be applied to the observed number density. It is straightforward to estimate the corrections for other seeing conditions since the scale of the abscissa varies as the inverse square of the seeing.

depend almost entirely on the product of average local number density with average image area. This area is the isophotal analysis area for the simple machine methods and to a reasonable approximation the seeing-disc area for the other methods. From the diagram it is clear that:

(i) Simple machine estimates of number density (and hence other parameters) always require non-zero adjustments to be made, even for relatively low count levels. In contrast, a human observer, by making use of all the image profile information, is rarely confused in these regions and neither is the idealized crowded field algorithm.

(ii) For moderately high number densities (100 arcmin^{-2}) simple isophotal analysis fails completely because the majority of images are joined together at the threshold isophote. A human observer can still extract fairly reliable information in such fields and so too can the idealized crowded field method.

(iii) At high number densities (200 arcmin^{-2}) it should be possible to do even better than a human observer examining a plate by eye. This is essentially due to the difficulty experienced in deconvolving a complex blended image by eye. Deconvolution is an operation that the brain is not very good at, whilst in principle at least, deconvolution by model fitting is well known to produce excellent results.

Several of the approaches to image analysis mentioned previously can be applied to moderately crowded fields; for a discussion see Newell (1982). However, the majority of these methods are semi-interactive and require considerable human interpretation in deciding on approximate position and numbers of images. Others will only work providing not more than a few images overlap within one blend. Our goal has been to examine the potential for a fully automatic method that is both robust and efficient in terms of computer requirements, capable of dealing with complex multiple overlaps and able to generate the optimum estimates of image parameters. Further considerations are that the system must be flexible (i.e. capable of dealing with galaxy or star overlaps) and eventually be capable of being integrated into existing machine measuring systems.

In deciding on a rationale for such an automatic analysis natural questions to consider are: how does a human interactively interpret pixel intensity maps, with particular regard to image detection and deciphering blended images; how accurate are the image parameters likely to be and closely related to this which methods of analysis will give the best results at reasonable computational cost? By reducing the problem of image description to one of estimating certain image parameters (position, intensity and shape) it is then pertinent to use statistical theory to inquire about optimum methods for estimating the parameters.* In particular since the method of maximum likelihood provides the foundation for an answer to these questions how can we use this knowledge to aid in designing analysis schemes? The maximum likelihood method also defines the minimum error it is possible to obtain for each of the parameters. Any new approach or short cuts in calculation can therefore easily be assessed on an absolute scale in terms of how closely they attain these minimum parameter errors.

2 Overall analysis strategy

In order to break down the overall analysis system into manageable portions the following main tasks have been assigned to the automatic crowded field algorithm developed here:

(i) Estimate the local sky background over the field.

(ii) Detect images/image blends and keep a list of pixels belonging to each blend for further analysis.

*(As noted by many other authors, the parameterization gives a large data compression of order 1000 : 1 whilst losing little or no useful information about the images.)

(iii) Analyse each image blend detected for multiple images, generate initial parameter estimates, particularly number of images and approximate positions.

(iv) Refine the parameters for each separate image within a blend, particularly the intensity, and revise the local sky estimate if necessary. Decide if more images are required and eliminate spurious images from step (iii).

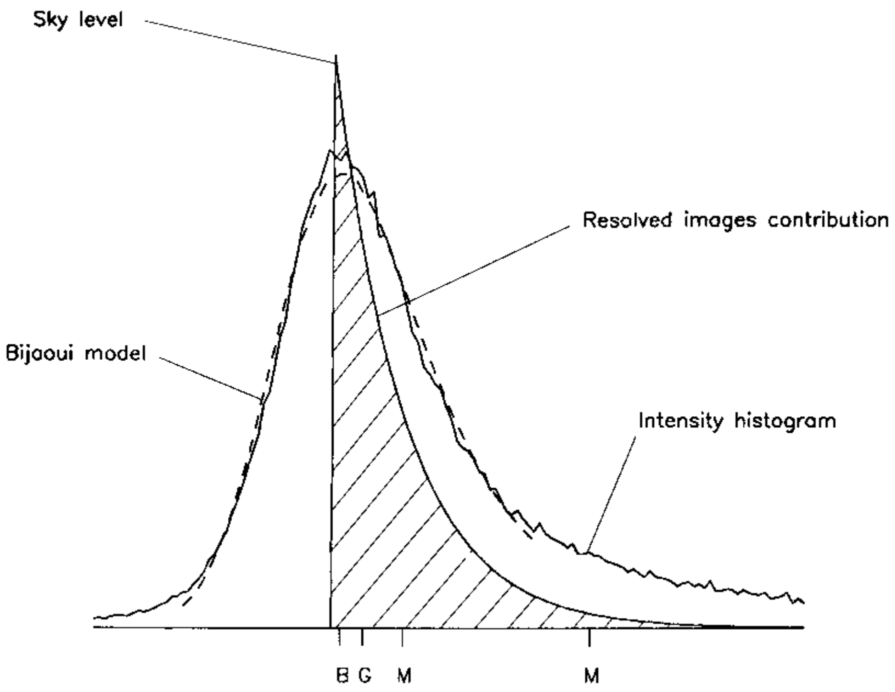
This basic strategy has been designed with a view toward incorporation into existing systems, since steps (i) and (ii) are present in conventional isophotal analysis methods. The core of the problem is segmenting a blended image into its component parts. By only concentrating on those pixels belonging to the blend above a certain isophote considerable savings in computation result, with little loss of accuracy. A consequence of this is that more complex refinement schemes may be used.

Each of the analysis stages will be discussed in subsequent sections, with reference to some of the more fruitful schemes mentioned previously. In particular a study of maximum likelihood estimation in Section 5 yields a significant insight into why some analysis methods are preferable to others. Finally some preliminary results obtained using a completely automatic analysis algorithm are presented together with an assessment of the current weaknesses of the method and its future prospects.

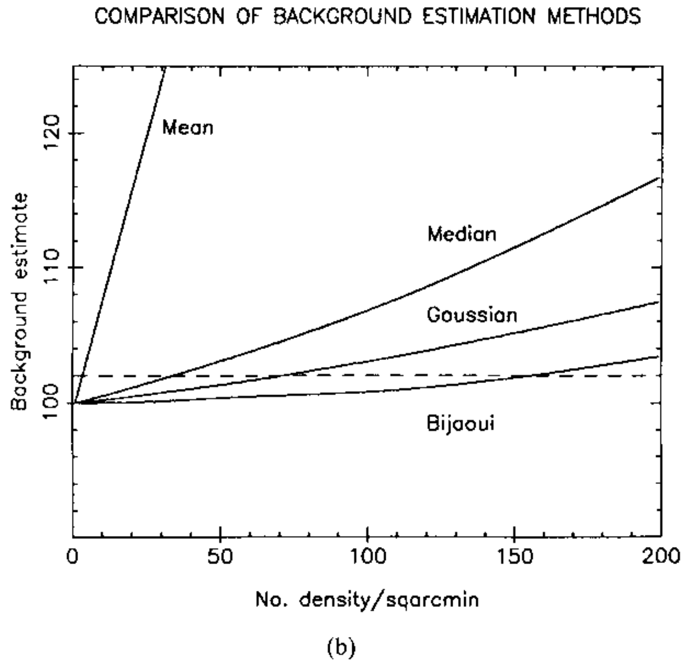
3 Sky background analysis

The sky background analysis is based on the method used in the APM system (e.g. Irwin & Trimble 1984). The philosophy behind this approach is quite simple: the field is effectively partitioned into suitable size background pixels, typically 10–30 arcsec in dimensions. For each of these background pixels an array of intensities are available which can be conveniently represented as an intensity histogram. From this histogram it is possible to estimate the local sky level using any one of several different methods as illustrated in Fig. 2(a). When the full 2D array of initial background measurements has been made it can be further processed through a non-linear filter to give the final background values. The purpose of this filter is to detect and correct spurious background values and to provide sufficient continuity in the background over the desired scale size of expected variations. A local sky value at any point in the field may then be simply obtained from bilinear interpolation of the background array. In this way we are using all of the background information before making decisions on local values and by suitable choice of the background pixel size we can follow arbitrarily complex variations over extended objects such as globular clusters or nearby large galaxies. The method should always provide sensible estimates of sky since for extended objects the local sky level due to say unresolved images must generally vary more slowly than the scale size of image blends. If not the image or blend causing the variation will be detectable as an image against some other background and will thereby lend itself to subsequent analysis. For convenience in further processing the sky variations are then removed additively and the resulting frame of data is described as having been background-corrected. All the subsequent processing operations are now far more straightforward since we know the sky level at all parts in the field. Any small residual systematic error in the sky level will now be manifest as an essentially constant offset for a given image blend. If the crowding is severe it will be necessary to include this offset in the parameter refinement in step (iv). An example of the flat-fielding technique is shown in Fig. 7 for the distant globular cluster, PAL 8. Note how the contour map for this field is much easier to interpret after background-correcting.

The success of this approach depends on obtaining a reliable estimate of the sky level from the pixel intensity histogram. Following Bijaoui (1980) let us model the pixel intensity histogram $P(I)$ using three parameters: the true sky level S, the rms noise at this level $\sigma_{\rm rms}$ and a parameter, a, describing the asymmetry of the underlying distribution due to the presence of resolved



(a)



(b)

Figure 2. Sky background estimation. (a) An example of fitting a Bijaoui model function, dashed curve, to a pixel intensity histogram contaminated by resolved images. The sky level determined from this model is indicated together with that derived from several simpler alternatives. (b) The error in estimating real sky levels as a function of the local image number density and the method chosen for average seeing (2.5 arcsec). A globular cluster stellar luminosity function (PAL8) has been used to generate the model data using random placement of images. The rms noise at sky is 12.5 units, as in the real data. Although the systematic error for the Bijaoui model estimate is smaller than the alternative methods this estimator has a much larger random error. The dotted line is the level at which the systematic error in sky causes a systematic error in image intensity as large as the optimum random error in image intensity due to observational noise. It would be desirable to further refine the sky estimate for each image blend if this level is approached or exceeded.

images. Assuming that the underlying asymmetry is adequately represented by an exponential function and that the effect of the noise at sky is to convolve this function with a Gaussian distribution of standard derivation $\sigma_{\rm rms}$, then

$$P(I)=\frac{1}{a}\operatorname{erf_c}\left(\frac{\sigma_{\rm rms}}{a}-\frac{I-S}{\sigma_{\rm rms}}\right)\exp(\sigma^2/2a^2)\exp[-(I-S)/a]. \quad (1)$$

If the resolved images have little or no effect the intensity histrogram is well described by a Gaussian distribution plus a small positive tail. Fig. 2(b) shows the properties of several background estimators: mean, median, mode from Gaussian model, mode from Bijaoui model – as a function of the image number density for a representative stellar population under average observing conditions. It should be noted that the size of the systematic error increases strongly for both poorer seeing and high noise levels. The former due to more pixels being affected by the image and the latter due to the higher intensity levels in the image being convolved into the region near true sky. A natural question to consider is how much error in image intensity is caused by an error in the local sky level, δb? For an image with normalized profile $\phi(x,y)$ and using a profile fitting method to estimate the intensity this error would be

$$\delta b\left[\int\int\phi^2(x,y)\,dxdy\right]^{-1}.$$

This should be compared with the optimum random error for faint images which is

$$\sigma_{\rm rms}\left[\int\int\phi^2(x,y)\,dxdy\right]^{-1/2},$$

(see Section 5). As an illustration for a Gaussian image profile of half-width σ at $1/e$ of the peak height, the systematic error is $2\pi\sigma^2\cdot\delta b$ and the random error $\sqrt{2\pi\sigma^2}\,\sigma_{\rm rms}$. When these two errors are comparable it is necessary to further refine the local sky level. This will be discussed briefly later. The dotted line in Fig. 2(b) indicates the point at which equality between the two errors is reached in the example given.

As expected the simple mean is fairly useless in the presence of images; however various types of clipped means (Herzog & Illingworth 1977; Newell 1979; and Ratnatunga & Newell 1984) have been used with some success. The median and Gaussian mode are certainly directly usable for number densities of up to 30 and 70 arcmin^{-2} respectively. But in terms of the systematic error the most successful method illustrated is that due to Bijaoui. This produces satisfactory estimates for number densities up to 150 images arcmin^{-2}. However, because of the explicit model asymmetry and hence strong coupling between the model parameters, the random errors can be much larger here than for the alternatives (a similar effect was noted by Auer & van Altena 1978 for positional refinement from image marginal sums).

For both the Gaussian and Bijaoui models only those histogram bins lying within one standard deviation of the mode were used. By restricting the fit to this subset problems due to model inaccuracies and data defects are greatly reduced without seriously affecting the accuracy of the estimated sky level. (The random error for a Gaussian model is of order $\sigma_{\rm rms}/\sqrt{n'}$ where n' is the effective number of pixels used in the fit and for a Bijaoui model the random error is typically several times this value.)

Whilst none of the methods discussed are entirely satisfactory, providing they produce consistent estimates (i.e. small random errors) it is always feasible to further refine the local sky level for an image blend at a later stage using just a single offset per blend. For this reason the Gaussian model is used throughout. In practice this method is virtually identical to the smooth interpolated mode used by Newell (1979). The overwhelming advantage of having a good initial sky estimate is that it makes subsequent tasks much easier.

4 Detection of images

The classical approach for optimal detection of signals in the presence of noise is to apply a matched filter to the raw data prior to setting detection thresholds (see for example Pratt 1978). In

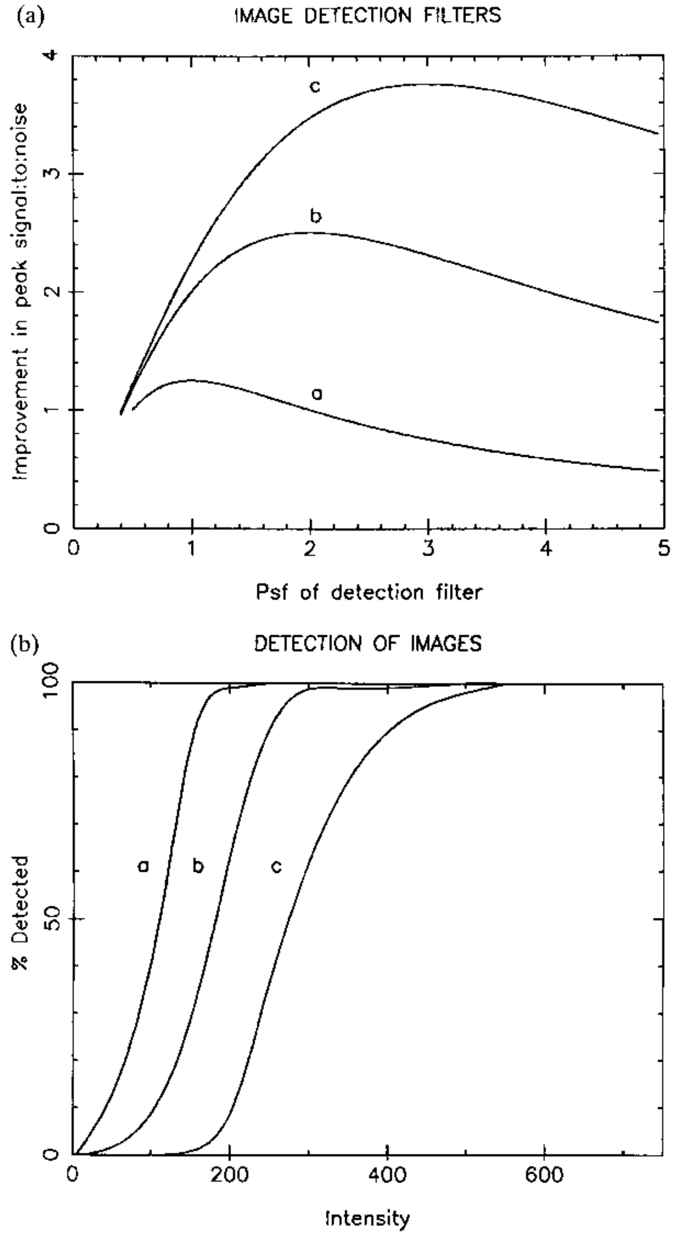


Figure 3. The efficiency of an isophotal plus pixel connectivity criterion for detecting images. (a) Performance of detection filter. Curves a, b and c in both diagrams represent Gaussian images of half width 1, 2 and 3 respectively at $1/e$ of the peak height. Note that the improvement in signal : to : noise is fairly insensitive to the exact characteristics of the detection filter. (b) At a sky rms noise level of 10 units the detection algorithm locates 50 per cent of the images (limit of reliability) at peak : to : noise ratios of 3.5, 1.4 and 1.0 for curves a, b and c respectively. Case 'b' is typical of astronomical data. The analysis isophote and connected pixel size were deliberately chosen to keep the number of noise images to below 10 per cent of the number of real images for all examples. If a higher proportion of noise images is acceptable these detection limits can be lowered still further. (c) The number of images detected as a function of pixel size.

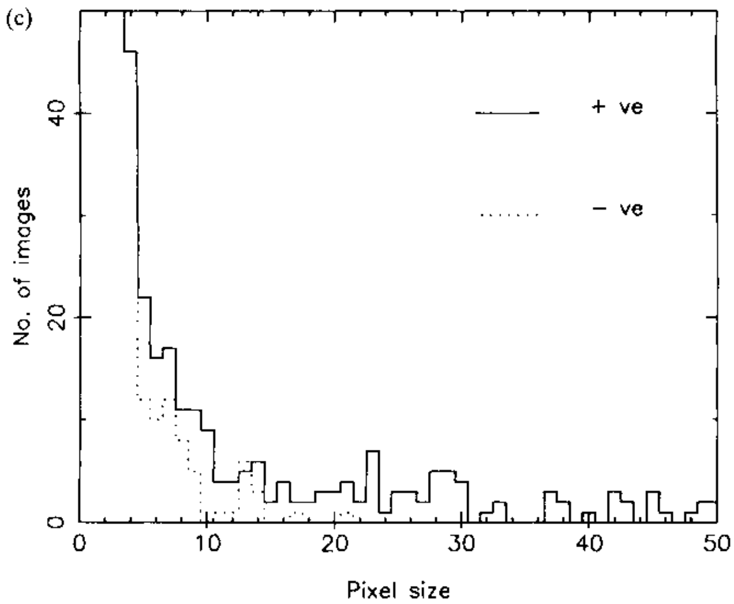


Figure 3 – *continued*

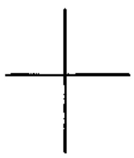
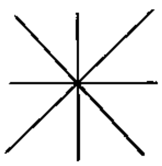
the case of raster data where the only images that will be difficult to detect will be the faint ones, the appropriate matched filter is the seeing function. This may be obtained in principle either directly by averaging suitable stellar profiles in the field or by an analytic model fit to these profiles. The purpose of the detection filter is to maximize the ratio of the image peak signal to the local noise level. Fig. 3(a) illustrates the performance of various detection filters for images in random noise. It is clear from the diagram that the exact size and hence shape of the filter is unimportant. Therefore since it is useful to keep the images as distinct as possible, it is preferable to err toward smaller filters. Furthermore from the computational viewpoint various suitably chosen fixed detection filters such as

1	2	1
2	4	2
1	2	1

are preferable in order to maintain a high processing speed. For non-random noise (e.g. microdensitometer output) the optimum detection filter should be appropriately scaled down to take account of the noise correlation.

After filtering the raw data a fixed isophote above sky is chosen and an image is defined to be a region of simply connected pixels above this isophote. As an alternative a peak detection algorithm may be used to locate the images (Newell & O'Neil 1977; Herzog & Illingworth 1977). However, this approach suffers from spurious noise peak detection problems unless it is modified to essentially look like an isophotal method.

Of the two most commonly used isophotal connectivity criteria

 and 

the former is preferable since it gives equal weight to only the four nearest neighbour pixels as opposed to the four second nearest neighbour pixels as well. The latter type of connectivity tends to produce more spurious pixels in the outer parts of an image. Finally by considering the distribution of images as a function of pixel size a suitable minimum pixel size can be chosen. Naturally in any detection method there is a trade off between detection reliability and contamination due to noise. A simple test to determine the number of noise images present is to repeat the analysis for −ve going images (i.e. set a −ve threshold and look for −ve images below sky). If the noise is random and the sky correctly estimated then this simple test will enable the number of noise images as a function of both size and intensity to be measured and thereby allowed for in subsequent image analysis.

Fig. 3(b) illustrates the use of the combined detection filter/isophotal connectivity method for some test data chosen to represent the range of image profiles and noise levels encountered in practice. The middle curve of Fig. 3(b) is a typical astronomical case. For a 50 per cent level of detection reliability with less than 10 per cent noise contamination the peak signal in the raw data is only 40 per cent higher than the noise level. If a higher proportion of noise images are acceptable than this limit could be further improved. From Fig. 3(c) it is clear that the noise images only analysis (−ve images) makes subsequent choice of a suitable minimum pixel size a simple quantitative operation.

Finally having located those pixels belonging to each image or image blend in the filtered data it is essential to use the *raw intensity* values of those pixels for further analysis – pixels with obvious defects should not be used at all. Not only does this virtually eliminate edge effects to which all isophotal methods are prone, it is also necessary to guarantee optimum results when estimating parameters for the blend. This is one consequence of the theory of parameter estimation developed by Fisher (1958). Any filtering operation no matter how subtle always *reduces the information* about the parameters within the data. At best there is no loss of information. This is particularly important when assessing the number of images in a blend, since the extra loss of resolution caused by filtering the data makes separation of images even more troublesome.

The edge effects in isophotal methods are caused by pixels in the outer parts of images being preferentially included because of noise. This leads to a slight overestimate of the isophotal intensity and extra uncertainty in the other image parameters (coordinates, second order moments, etc.). Both of these problems are reduced by doing the isophotal detection on the filtered data and the parameter estimation using unfiltered data.

5 Parameter estimation

Having reduced the problem of image description to one of estimating certain image parameters (position, intensity and shape) it is then pertinent to use statistical theory to enquire about optimum analysis methods for estimating those image parameters. In particular, since it is well known that the method of maximum likelihood (Fisher 1958) provides the basis for an answer to this query, how can we use this approach to aid in designing analysis schemes? The maximum likelihood method can also be used to define the minimum error it is possible to obtain for each of the parameters. Any new approach can therefore easily be assessed on an absolute scale in terms of how closely it attains these minimum parameter errors. Similarly it is also fairly easy to decide whether certain short cuts in calculation are worthwhile in terms of a trade off between computing speed and accuracy.

Historically, there have been many different attempts in astronomy to give theoretical bounds for parameter errors. Fellgett (1970) specifically designed the GALAXY measuring machine to mimic a maximum likelihood analysis for finding accurate image coordinates. Both Farrel *et al.* (1967) and later Kibblewhite (1971) produced a detailed study on the limiting accuracy of image parameters using a maximum likelihood approach. More recently Auer & van Altena (1978) and King (1983) have considered the error limits for position and intensity respectively, using non-linear least-squares theory. Interestingly, although the non-linear least-squares approach can be considered as a limiting case of the maximum likelihood method and as such predicts the same theoretical error bounds, it is possible to derive different analysis algorithms from the two methods.

5.1 MAXIMUM LIKELIHOOD PRINCIPLE

Consider a set of N *independent* observations $[x]$, say $x_1, x_2, \ldots x_N$. These could be N events found in an experiment or photon counts for particular pixel coordinates, x_i, or the number of darkened grains as a function of some spatial coordinates on a photographic plate or even the individual coordinates of photons in an image, etc. Suppose that the probability density function of the x_i is described by $P(x_i|\theta)$ where $\theta=(\theta_1, \theta_2, \ldots \theta_m)$ are the m parameters to be estimated from the data. The two commonest cases encountered are:

(i) The events are binned into pixels and providing the number of events per pixel is large enough ($\gtrsim 20$), $P(x_i|\theta)$ is then well represented by a Gaussian distribution

$$P(x_i|\theta)=\frac{1}{\sqrt{2\pi\sigma_i^2}} \exp(-\varepsilon_i^2/2\sigma_i^2) \tag{2}$$

where

$$\varepsilon_i=x_i-f(\theta, x_i).$$

That is at pixel i the measured number of events x_i is perturbed from the model value $f(\theta, x_i)$ by noise ε_i with expected variance σ_i^2. This approach is equivalent to a least-squares method and providing the pixel size is small relative to the image size virtually no loss of information will occur.

(ii) $P(x_i|\theta)$ describes the expected two-dimensional distribution of event coordinates defined by the normalized image profile(s) $\phi(x_i|\theta)$ plus a sky background contribution, b (assumed uniform over the image for simplicity). Therefore

$$P(x_i|\theta)=\frac{1}{N}[\eta\phi(x_i|\theta)+b] \tag{3}$$

with all events given equal weight and assumed to be Poisson-distributed. η is the image intensity and the constant N is used to normalize the total count level over the image. If there is more than one image in the feature the single profile is replaced by an appropriate summation. Multiplying the probability density function by N then gives us the expected number of events for that point in the image.

Now the joint probability density function of the $[x]$ is, by independence

$$L(x|\theta)=L(x_1, x_2, \ldots x_N|\theta)=\prod_{i=1}^{N} P(x_i|\theta) \tag{4}$$

$L(x|\theta)$ is known as the *likelihood function* of θ, $[x]$ are considered to be known data values. (Note that in case (i) in the preceding example, this is also the joint probability density function of the residuals, or error terms, ε_i and may readily be extended to the case of non-independent samples.) It is also true that any monotonically related function of $L(x|\theta)$ conveys the same information about the parameters θ. In particular, the *log likelihood function* is often used for simplicity since

$$\log L(x|\theta)=\sum_{i=1}^{N} \log P(x_i|\theta). \tag{5}$$

In a very general sense we may derive parameter estimates $\hat{\theta}$ from a knowledge of the individual data points x_i by setting up functions (statistics)

$$\hat{\theta}_j=\hat{\theta}_j(x_1,x_2,\dots x_i\dots x_N). \tag{6}$$

Any parameter estimates $\hat{\theta}_j$ will have an associated probability density function and hence (in principle) a derivable variance. Fisher showed that under quite general conditions the maximum likelihood estimates, defined below, are the ones with minimum variance. The maximum likelihood estimates of the parameters are defined by

$$\frac{\partial \log L}{\partial \theta_j}=0 \text{ for } j=1,2,\dots M \tag{7}$$

and therefore represent the most probable value of θ given the data x. (Throughout we are assuming no prior knowledge of the parameters, although it is possible to incorporate such knowledge into the likelihood function if desired.)

5.2 FISHER INFORMATION AND PARAMETER ERRORS

In order to estimate the minimum variance associated with the maximum likelihood method it is necessary first of all to define the Fisher information matrix $\mathbf{I}(\theta)$, (see for example Eadie *et al.* 1982). The ijth element of this matrix is defined as

$$I(\theta)_{ij}=\int \frac{\partial \log L}{\partial \theta_i}\frac{\partial \log L}{\partial \theta_j} L(x\,|\,\theta)\,dx \tag{8}$$

where the integral is over the event space $[x]$. Diagonal terms in this matrix provide a measure of the information about a particular parameter contained within the data. Off-diagonal terms additionally give a measure of the conflict between non-independent parameters. The information matrix has the following desirable properties:

(i) The information in a set of events $[x]$ increases linearly as the number of events.

(ii) Data which are irrelevant to the parameters to be estimated convey no information.

(iii) The measure of information is directly related to attainable precision. That is, the greater the information the better the precision of the parameter estimates.

Clearly with these properties it is a straightforward matter to decide quantitatively whether or not to reject data when little or no information loss occurs.

Since the parameters θ are themselves effectively dependent on a large number of events $[x]$, we can invoke the Central Limit Theorem and state that in the limit of large numbers of events the probability density function of the parameters will be multivariate Gaussian. The covariance matrix, $\mathsf{V}(\theta)$, of this distribution therefore completely describes the expected errors in the parameters and is thus the main quantity of interest in any error analysis. Fisher showed that this covariance matrix is related to the information matrix by

$$V(\theta)_{ij}\geqslant I^{-1}(\theta)_{ij} \tag{9}$$

the variance of the ith parameter being conventionally defined as the ith diagonal term. Minimum variance is obtained for maximum likelihood estimates under fairly general conditions. Equation (9) requires modification if the estimates are known to be biased (see Cramer 1945). Following Fisher, it is now possible to compare alternative methods in terms of their *efficiency* relative to the minimum possible error obtainable.

$$e(\theta_j)=\frac{\operatorname{var}\{\hat{\theta}_j\}}{\operatorname{var}\{\hat{\theta}_j\}_{\mathrm{ML}}}. \tag{10}$$

This must satisfy $e(\theta_j) \geq 1$ and provides us with an objective measure for assessing the relative merits of different estimation schemes.

5.3 ESTIMATING PARAMETERS OF ASTRONOMICAL IMAGES

We are now in a position to answer the question posed at the beginning of this section. The methods outlined in Section 5.1 indicate how to analyse the data and the results from Section 5.2 inform us about how to assess various alternative analysis schemes in terms of their accuracy. In choosing one analysis system in preference to another it helps to assess it in terms of the following properties.

(i) Consistency – estimates converge to the 'true' value as the number of observations, or events increases. The estimate need not be unbiased provided the bias is consistent and predictable. For example: isophotal intensities give a biased estimate of the total intensity in an image, however this bias is consistent and predictable for a given image profile.

(ii) Robustness – practical algorithms should not be too sensitive to small changes in model profiles and error terms. Although this is an obvious requirement it is important since we often deal with only approximate image profiles, particularly in estimating parameters of distant galaxies.

(iii) The estimator should make maximum use of the information in the data, and hence should have an efficiency close to unity. Here we are using Fisher's definition of information, which should not be confused with that of Shannon (1948) developed in communication theory.

Of these the most important is condition (iii), since this determines the ultimate precision of the parameters.

Now let us consider the two cases mentioned previously using the general method outlined in equations (4) to (9).

5.3.1 Least-squares model

The likelihood function for independent Gaussian residuals takes the form

$$\log L(x\,|\,\theta) = \sum_{i=1}^{N} \log P(\varepsilon_i) = \sum_{i=1}^{N} -\tfrac{1}{2}\log 2\pi\sigma_i^2 - \varepsilon_i^2/2\sigma_i^2 \tag{11}$$

where the summation is over the N pixels making up the image. Since only the term involving $\varepsilon_i(\theta)$ depends on the parameters to be estimated, maximizing the loglikelihood function with respect to θ is equivalent to minimizing

$$F(\theta) = \sum_{i=1}^{N} [x_i - f_i(\theta)]^2/2\sigma_i^2 \tag{12}$$

where with a slight change of notation $f_i(\theta)$ is the model function value at pixel location i. The variance of the error may include measurement error, recording error, etc. Equation (12) defines a weighted non-linear least squares method for parameter estimation, which for sufficient events and small enough pixel size is asymptotically optimal. In the case of non-Gaussian residuals or few observations it is preferable to use the direct maximum likelihood method. For correlated error terms (residuals) the argument of the probability density function implicit in equation (11) should be generalized to the appropriate quadratic form which results in a Mahanoublis metric for the parameter space (e.g. Cramer 1945).

The condition

$$\frac{\partial F(\theta)}{\partial \theta_j} = \sum_{i=1}^{N} \frac{\partial f_i(\theta)}{\partial \theta_j} [x_i - f_i(\theta)]/\sigma_i^2 = 0 \qquad \text{for } j = 1, 2, \ldots m \tag{13}$$

defines a system of non-linear equations for the parameter estimates. In general this must be solved iteratively. However, if we restrict ourselves to position and intensity estimation only and assume a known profile, it is possible to decouple the position and intensity estimation with little loss of accuracy (Tody 1980). The positional refinement still involves solving a series of non-linear equations and is relatively expensive computationally. However the image intensities now must satisfy a *linear* system of equations.

The information matrix for the general noise case is given by

$$\mathbf{I}(\theta)=\frac{\mathbf{J}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{J}}{\sigma_{\min}^{2}} \tag{14}$$

where $\mathbf{J}$ is the $n\times m$ Jacobian matrix of model function derivatives such that

$$J_{ij}=\frac{\partial f_i(\theta)}{\partial\theta_j} \tag{15}$$

'T' denotes transpose and $\mathbf{C}$ is the noise, or residual, covariance matrix. $\sigma_{\min}^2$ is the sum of squares of the residual errors normalized by the number of degrees of freedom,

$$\sigma_{\min}^2=\frac{F_{\min}(\theta)}{N-m}; \tag{16}$$

it has expected value unity (e.g. Eadie *et al.* 1982). Using equations (9) and (14) implies that the parameter covariance matrix must satisfy

$$\mathbf{V}(\theta)\geq\sigma_{\min}^2[\mathbf{J}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{J}]^{-1}. \tag{17}$$

If both errors and parameters are uncorrelated (a reasonable approximation for intensities and coordinates for many cases) then

$$\operatorname{var}\{\hat{\sigma}_j\}\geq\left\{\sum_{i=1}^{N}\left[\frac{\partial f_i(\theta)}{\partial\theta_j}\right]^2\frac{1}{\sigma_i^2}\right\}^{-1}\sigma_{\min}^2. \tag{18}$$

At the same time as estimating the image intensities it is also possible to augment the system of equations to include a correction to the background value. For example, if there are R images in the blend then the model function becomes

$$f_i(\theta)=\sum_{r=1}^{R}\eta_r\phi_r(i)+\delta b \tag{19}$$

where η_r is the intensity of the rth image, ϕ_r its normalized profile and δb is the background correction term. Taking x_i in equation (13) to represent the overall background-corrected data value for pixel i then the parameters η_r and δb must satisfy

$$\left[\begin{array}{c|c}
\langle\phi_1\phi_1\rangle\langle\phi_1\phi_2\rangle\cdots\cdots\langle\phi_1\phi_R\rangle & \langle\phi_1\rangle \\
\langle\phi_2\phi_1\rangle\langle\phi_2\phi_2\rangle\cdots\cdots\langle\phi_2\phi_R\rangle & \langle\phi_2\rangle \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
\langle\phi_R\phi_1\rangle\langle\phi_R\phi_2\rangle\cdots\cdots\langle\phi_R\phi_R\rangle & \langle\phi_R\rangle \\
\hline
\langle\phi_1\rangle\langle\phi_2\rangle\cdots\cdots\cdots\cdots\langle\phi_R\rangle & N
\end{array}\right]
\left[\begin{array}{c}
\eta_1 \\ \eta_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \eta_R \\ \hline \delta b
\end{array}\right]
=
\left[\begin{array}{c}
\langle\phi_1 x\rangle \\ \langle\phi_2 x\rangle \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \langle\phi_R x\rangle \\ \hline \langle x\rangle
\end{array}\right] \tag{20}$$

where $\langle\,\rangle$ denotes the overlap integral between profiles or data over the region of the analysis and the dashed lines indicate the extra terms necessary to include the additional background refinement. For this extra refinement to work well it is essential to have a reliable estimate of the image profile. In practice two analytic models are used: Newell's (1965) mixed Gaussian and Lorentzian profile or a combined Gaussian and exponential model. Both are radially symmetric and require two parameters to specify the distribution: in the first the Gaussian core width and a mixture coefficient; in the second the same core width and a threshold (conveniently expressed as a percentage of the peak height) above which the profile is Gaussian and below which it is exponential. Continuity of 0th and 1st derivatives complete the specification. One of these two profiles will adequately represent stellar images for a wide range of detectors and observing conditions. If the profile is slightly in error, or the real images are not perfectly symmetric, then there will be a small constant magnitude offset in the intensity estimates as can be seen by inspection of equation (20). This offset is usually unimportant since an external zero point for the magnitude scale is required anyway.

5.3.2 Maximum likelihood model

The likelihood function for the distribution of event coordinates $P(x_i\,|\,\theta)$ is

$$\log L(x\,|\,\theta)=\sum_{i=1}^{N}\log P(x_i\,|\,\theta)=\sum_{i=1}^{N}\log f_i(\theta)+C \tag{21}$$

where $f_i(\theta)$ is defined in equation (19) except that δb is replaced by b, the overall background value, and C is a normalization constant. In this case the summation is over the N individual events making up the image. The sky level b is assumed constant over the blend. As we have seen, it is possible to correct non-uniform background levels implying that the background can be adjusted to be at least locally uniform if necessary.

Minimizing (21) with respect to θ yields the system of equations

$$\sum_{i=1}^{N}\frac{\partial f_i(\theta)}{\partial\theta_j}\,\frac{1}{f_i(\theta)}=0; \qquad j=1,2,\ldots m \tag{22}$$

which should be compared to the previous condition of equation (13). Although equation (22) yields no useful estimate for the image intensity, since this is implicitly bound up in the normalization constant, it does provide an alternative method for position estimation which we shall return to shortly.

As before the information matrix and parameter covariance matrix can be calculated directly from equations (8) and (9). If the parameters are uncorrelated the expected variance of the parameter estimates is given by

$$\operatorname{var}\{\hat{\theta}_j\}\geqslant\left\{\sum_{i=1}^{N}\left[\frac{\partial f_i(\theta)}{\partial\theta_i}\right]^2\frac{1}{f_i(\theta)}\right\}^{-1}. \tag{23}$$

For a Poisson distribution $f_i(\theta)\equiv\sigma_i^2$ and since the expected value of $\sigma_{\min}^2\equiv 1$ the two methods predict identical error bounds providing the pixel size used in the least-squares method is much smaller than the image size. In particular this implies that coordinate estimation from equations (13) or (22) will produce equally accurate results.

In integral form, with an obvious notational change, equation (18), or (23) becomes

$$\operatorname{var}\{\hat{\theta}_j\}\geqslant\left\{\iint\left[\frac{\partial f(x,y\,|\,\theta)}{\partial\theta_j}\right]^2\frac{1}{\sigma^2(x,y)}\,dx\,dy\right\}^{-1} \tag{24}$$

where $\sigma^2_{\min}$ has been set equal to its expected value of unity. This step is valid for the least-squares method if the pixel size is small enough and will always be true for the maximum likelihood case since the event coordinates are continuous variables. Therefore for an isolated image we can readily calculate error bounds for both position and intensity; with Poisson errors

$$\mathrm{var}\,\{\hat{\eta}\} \geq \left[\iint \frac{\phi^2(x,y)}{\eta\phi(x,y)+b}\,dx\,dy\right]^{-1} \tag{25}$$

$$\mathrm{var}\,\{\hat{\theta}_x\} \geq \left\{\iint \frac{\eta^2[\partial\phi(x,y)/\partial\theta_x]^2}{\eta\phi(x,y)+b}\,dx\,dy\right\}^{-1} \tag{26}$$

where the integrals cover the complete extent of the image. For faint images, or for cases of constant noise, the denominator of the integrals tends to b (or σ^2 noise) leaving simple integrals to evaluate.

These limits have the following general properties:

(i) The variance of the estimates is inversely proportional to the number of events making up the image – this follows from the additivity property of the information matrix. Thus increasing the observation time by a factor P results in a $\sqrt{P}$ decrease in parameter errors providing the recording and measuring process is linear.

(ii) For constant noise, or faint images, the variance in intensity is constant, therefore the percentage error in intensity is proportional to $1/\eta$. The positional error is also proportional to $1/\eta$.

(iii) If the image is bright relative to sky and the noise is Poisson the percentage error in intensity is equal to $1/\sqrt{\eta}$.

(iv) If only a subset of the image is used, for example: the central pixels only as in the Σ method of Newell & O'Neil (1974) the error bound for unbiased estimates will proportionately increase (see Section 5.5).

(v) The steeper the image gradient the smaller the error bound for both position and intensity. This also implies that the central part of an image contributes proportionately more to the accuracy than do the outer parts.

By using a similar argument it is also possible to develop error bounds for shape parameters and thereby place limits on the reliability of morphological classification and various higher order moments used to define equivalent ellipses. However, since this is outside the scope of this paper it will not be pursued further here.

Returning to the problem of estimating image coordinates let us consider the method embodied in equation (22) earlier in this section. Using the same model as before we then have the following requirement for the x (or y) coordinate $\hat{x}_s$ of the sth component of an image blend

$$\sum_{i=1}^{N} \frac{\eta_s(\partial\phi_s/\partial\hat{x}_s)}{\sum_r \eta_r\phi_r+b}=0. \tag{27}$$

The summation is again over the individual events and the i subscript has been dropped from the model terms for clarity. If the individual model functions are quadratic in both coordinates, which includes most of the commonly used 2D profiles, then the previous equation can be rewritten as

$$\hat{x}_s = \sum_{\substack{j\\ \text{pixels}}} x_j d_j w_j \Big/ \sum_j d_j w_j \tag{28}$$

where the events have been grouped into pixels with observed intensity d_j and the weighting function w_j is given by the ratio of the gradient contribution to the complete model

$$w_j=\frac{\eta_s(\partial\phi_s/\partial\hat{x}_s)}{\sum_r \eta_r\phi_r+b} \tag{29}$$

ϕ now refers to an average over the pixel. This is none other than a weighted centre of gravity (moment) method that will work optimally for blended images in addition to isolated images.

As an illustration for Gaussian images

$$w_j=\frac{\eta_s\phi_s}{\sum_r \eta_r\phi_r+b} \tag{30}$$

i.e. the relative contribution from one image compared to the total model. Now for bright single images this weight will be constant and for faint single images or cases of constant noise the weighting function equals the profile function. These observations help to explain why Auer & van Altena (1978) found the moment method to be almost as accurate as full profile fitting. Furthermore the combination of a detection filter plus isophotal connectivity criterion will automatically centre an image with respect to the pixels used in the computation implying that a properly devised moment method is as efficient as a least-squares approach. For many practical problems an unweighted moment method (or one weighted only on some function of the observed data values) which can be executed in a single pass through the data for isolated images, is almost as efficient as the iterative procedure implicit in equations (28) and (29). Table 1 presents a comparison of the efficiency of several commonly used coordinate estimation methods. The results are derived from artificial data to facilitate a comparison with the theoretical error bounds. It is apparent from this table that properly devised moment methods are equally as accurate as alternatively profile based methods. This helps to explain the success of plate measuring machines such as APM and COSMOS which both use simple isophotal/moment methods. (In both cases the coordinate accuracy for all but the faintest images is limited by the internal positional accuracy rather than the analysis algorithm.)

With blended images the weighting function also acts as an exact counterbalance to the tendency of the centre of gravity of an image to shift due to the presence of adjacent images. Although in practice a few modifications to the basic procedure are necessary to enhance the convergence properties. First the data values d_j are defined relative to local sky and secondly the sky term $(+b)$ is dropped from the denominator. With these changes the whole coordinate refinement is both faster and more robust and yet still produces results satisfactorily close to the theoretical error bounds. A further extremely useful feature is that spurious image assignments

Table 1. Coordinate accuracy relative to theoretical bounds for isolated faint images.

Method	Relative error x	Relative error y
2D profile fitting	1.01	1.00
1D profile fitting	1.15	1.17
Weighted moments	1.00	1.01
Unweighted moments	1.35	1.33
Weighted moments without sky term and plus threshold	1.14	1.13

are highly mobile and either migrate to the nearest concentration of intensity where a simple test can be used to remove them, or else quickly proceed to have zero intensity estimates and again can easily be detected and removed.

5.4 AN EXAMPLE: A CIRCULAR GAUSSIAN PROFILE

To gain further insight into the theoretical error limits represented by equations (25) and (26) let us consider the image profile to be well described by Gaussian function. This is a reasonable approximation for stellar images (King 1971) and also for faint galaxy images. Defining the normalized Gaussian profile as $1/\pi\sigma_G^2 \exp(-r^2/\sigma_G^2)$ it is straightforward to show that the intensity errors are limited by

$$\begin{aligned} \mathrm{var}\{\hat{\eta}\} &= \eta \qquad \text{bright images} \\ &= 2\pi\sigma_G^2 b \qquad \text{faint images} \end{aligned} \tag{31}$$

(replace b by σ_{noise}^2 for constant noise). For faint images the error bound is equivalent to that due to the background count rate over a scale size for the image. The positional errors are bound by

$$\begin{aligned} \mathrm{var}\{\hat{\theta}_x\} &= \frac{\sigma_G^2}{2\eta} \qquad \text{bright images,} \\ &= \frac{\sigma_G^2}{2\eta}\,\frac{4\pi\sigma_G^2 b}{\eta} \qquad \text{faint images.} \end{aligned} \tag{32}$$

Again for faint images the background count over an image scale size relative to the total image intensity governs the accuracy. In terms of the image peak height, I_p

$$\begin{aligned} \mathrm{var}\{\hat{\theta}_x\} &= \frac{1}{2\pi I_p} \qquad \text{bright images} \\ &= \frac{2b}{\pi I_p^2} \qquad \text{faint images.} \end{aligned} \tag{33}$$

For all images we then find the interesting property that the positional error is to first-order independent of the overall image size, depending only on the image peak height. Following King (1983) we can compare the astrometric error with the photometric

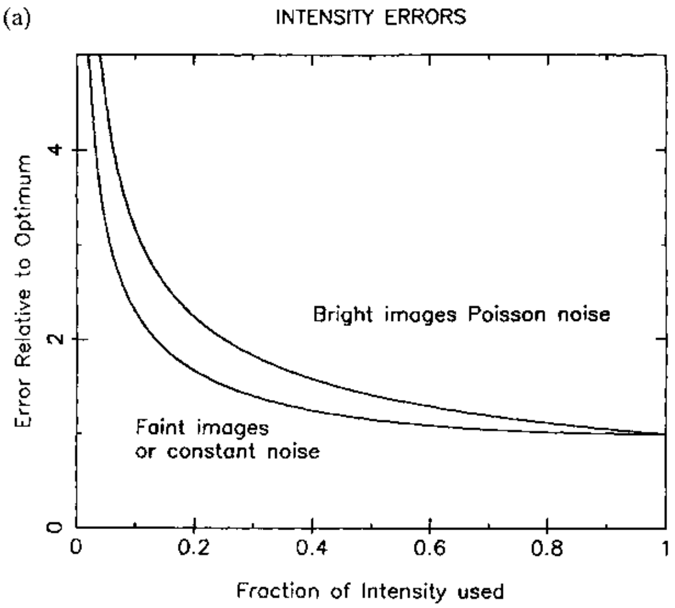
$$\begin{aligned} \frac{\mathrm{var}\{\hat{\theta}_x\}}{\sigma_G^2} &= \frac{1}{2}\,\frac{\mathrm{var}\{\hat{\eta}\}}{\eta^2} \qquad \text{bright Gaussian images} \\ &= \frac{\mathrm{var}\{\hat{\eta}\}}{\eta^2} \qquad \text{faint Gaussian images} \end{aligned} \tag{34}$$

For a more extensive discussion of some of these points see Auer & van Altena (1978) and King (1983).

5.5 ERROR BOUNDS WHEN USING A PROPORTION OF THE IMAGE EVENTS

Suppose we now limit the analysis to use only a fraction $1-r$ of the intensity distribution of the image. For example: the Σ method of Newell & O'Neil (1974) makes use of only the central pixels, an isophotal method uses only intensity values above a certain isophote, Rheault & Hardy (1980) use a small angular segment of the image to limit the effects of image crowding. In the latter case we are using only a fraction $1-r$ of the events defining the image therefore the additivity property of the information matrix implies that whatever the profile the parameter errors must increase by $1/\sqrt{1-r}$ (typically only a 1/4 of the image is used in crowded fields resulting in a doubling of all parameter errors). The first two cases are in effect only integrating out to some limiting radius and correspondingly again decreasing the information used. For the Gaussian profile used previously the image intensity errors after correction for bias, will increase by $1/\sqrt{1-r}$ for bright images and $1/\sqrt{1-r^2}$ for faint images or any image if the noise is not Poisson but constant. For example: if $r=½$ the errors increase by approximately 40 and 15 per cent respectively. Faint images fare better because of the discrimination against the background noise.

(a)



(b)

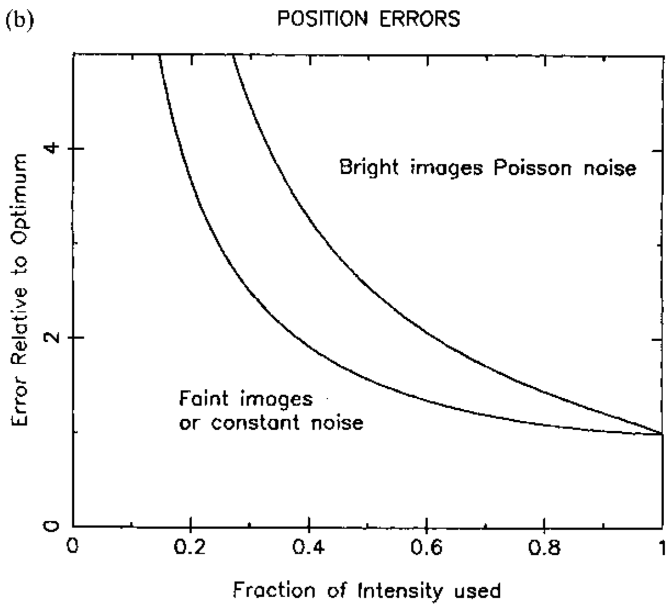


Figure 4. (a, b) Variation of the optimum theoretical error for position and intensity of a Gaussian image as a function of the fraction of image intensity used. If the fraction has been selected as a segment of the image, rather than a radial limit, the upper curve of (4a) is appropriate in all cases. These diagrams suggest that to adequately analyse crowded fields, where accuracy of both position and intensity is paramount, using all of the information in an image, or image blend is essential. For real data the type of noise encountered will lie somewhere between these two extremes, depending on the recording medium and measuring process used.

A similar though worse picture arises for coordinate errors, they increase respectively by factors of $1/\sqrt{1-r+r\log_e r}$ and $1/\sqrt{1-r^2+r^2\log_e r^2}$. Again as an example: if $r=\frac{1}{2}$ this leads to an increase in errors of roughly 250 per cent for bright images and 57 per cent for faint images. These error increases are illustrated in Fig. 4 for a range of values of r. If the noise is nearly constant because of

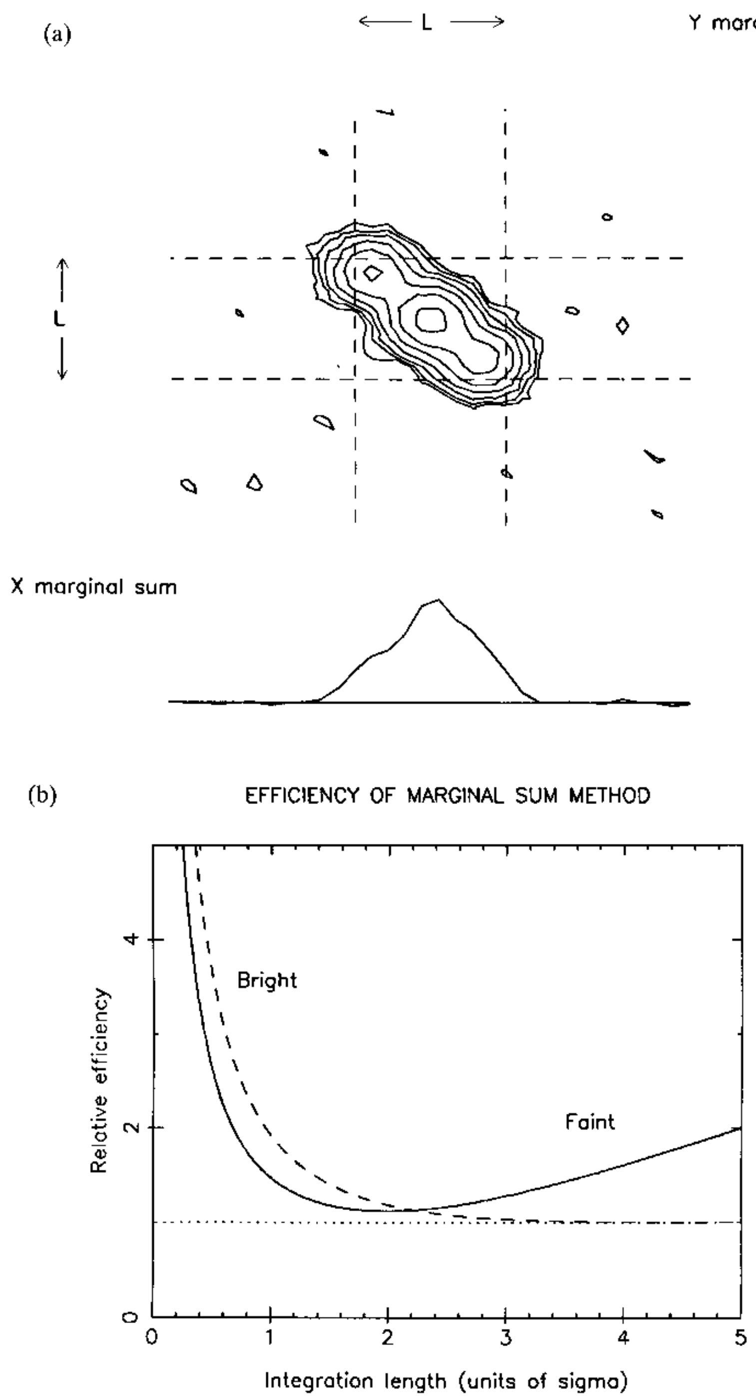


Figure 5. (a) The use of marginal sums for parameter estimation. The areas between the dashed lines are integrated to form the X and Y marginal sums which can then be interpreted using multipeak fitting functions. (Although in the example shown a multiple isophotal approach would have a clear advantage since the peaks separate out in 2D but not in the marginal sums.) (b) Relative efficiency of marginal sum method for single Gaussian images for both position and intensity estimation (see text). The noise is taken to be Poisson. For constant noise the curve for faint images is appropriate for all cases.

say measuring or recording errors the formulae for the faint images are appropriate for all. These results reaffirm that for the most accurate work it is necessary to use all the information in an image or image blend. This in turn implies that in crowded regions simultaneous estimation of all the image parameters using all relevant pixels is essential.

An alternative method proposed by Auer & van Altena (1978) and Newell (1979) is to use marginal sums to reduce the analysis to a one-dimensional operation and then fit a multipeak function to characterize the image. In order to minimize crowding effects only a subset of the marginal sum is used as illustrated in Fig. 5(a). In reformulating the problem in this way how much information about the image parameters is lost? As an example: for a single Gaussian profile with a marginal sum integrated over the range $-L/2 \leqslant y \leqslant L/2$ the new model function is

$$fx(\theta)=\eta\phi(x,0)f_L\sqrt{\pi\sigma_G^2}+Lb \tag{35}$$

where f_L is the fraction of the projected intensity used. It is straightforward to show that the relative efficiency for both position and intensity estimates compared to the optimum method is given by $1/f_L$ for bright images and $L/\sqrt{2\pi\sigma_G^2}f^2{}_L$ for faint images or constant noise. These functions are plotted in Fig. 5(b). Providing the integration length L is carefully chosen there is only a small decrease in efficiency relative to the optimum and certainly for simple image blends this is a viable alternative for coordinate refinement.

5.6 EFFECT OF FINITE PIXEL SIZE

The effect of finite pixel size (or scanning spot size for a microdensitometer) is twofold: first individual event coordinates are lost and replaced by a neighbourhood average and secondly significantly more background noise may contaminate the image since the image grows in size due to the convolution inherent in these processes.

In order to examine the effect of finite pixel size rigorously it is necessary to use the discrete version of equation (24), namely equation (23). Each point in the summation now must represent the integral of the model function $f_i(\theta)$, and the appropriate derivatives of this integrated model function, over the range of space defined by each pixel. The contributions to the information matrix will then decrease relative to the integral form because different levels of intensity will be contained within the same pixel. This discrete summation is of course directly equivalent to the least-squares formulation of the problem outlined earlier (except that in normal least-squares solutions the pixel integration is approximated by using the value of the model function at the pixel centre). Unfortunately an exact error analysis has to be performed numerically (see King 1983). However for many situations encountered in practice the following approximations, though not entirely rigorous, have been found to produce acceptable answers compared to the exact method. Furthermore they lend considerable insight to the problem.

As noted previously the effect of finite pixel or spot size is twofold: the individual event coordinates are lost and the image grows in size because of the inherent convolution. Dealing with the second point first, the amount of background noise integrated over the image scale size is the decisive factor. Therefore for bright images there will be negligible effect for faint images or cases of constant noise we simply replace the original image scale size by an appropriately enlarged version. Thus for faint Gaussian images the extra background noise worsens both position and intensity errors by factors of roughly

$$\left(1+\frac{a^2}{6\sigma_G^2}\right)^{1/2}$$

where a is the pixel size and for a Gaussian scanning spot by

$$\left(1+\frac{\sigma_{\rm spot}^2}{\sigma_{\rm G}^2}\right)^{1/2}.$$

The finite pixel size is usually only a small contribution to increase in contaminating background noise since typically $a \sim \sigma_{\rm G}$ and hence the errors increase by only ~8 per cent. The loss of individual event coordinate information to first order only affects the positional error. We can approximate this additional error by redefining the likelihood model function to include an extra uncertainty in the individual event coordinates. This can be represented by convolving the basic model function with a suitable randomizing function. In the spirit of the Gaussian approximation introduced earlier this will produce an extra factor for the positional error of roughly

$$\left(1+\frac{a^2}{6\sigma_{\rm G}^2}\right)^{1/2}$$

for all Gaussian images. Note, however, that the Gaussian scanning spot on its own would have no effect over and above that already mentioned.

It is apparent from these considerations that for many analysis problems, particularly in non-crowded fields, a combination of large scanning spot and large pixel size can produce reasonable parameter estimates with a corresponding saving in data storage and computing requirements (e.g. Stobie *et al.* 1979).

6 A crowded field algorithm

A flow diagram for a general purpose crowded field algorithm based on the analysis strategy outlined in Section 2 is presented in Fig. 6. Several of the steps necessary to implement such a process have already been discussed in earlier sections and will only be briefly commented upon here.

(a) The core of the method is to find an initial global sky estimate for the entire field against which a combination of a detection filter and isophotal connectivity constraint suffice to reliably locate all resolved features. The output from this stage is a list of pixel positions and *original* pixel intensities for each image. This list contains all the necessary information regarding numbers of images in the blend, etc. without recourse to extraneous data.

(b) An obvious method for detecting overlapped images is to re-examine this pixel list at multiple isophotes and decide if distinct images separate out. This is basically one of the ways the eye interprets contour maps. Two images that were overlapped at a lower isophote will separate completely at some higher level enabling a simple new isophotal analysis to determine the initial image parameters. The levels for the multiple isophotes are spaced logarithmically at 1/4 magnitude intervals. This has been found to be a reasonable compromise between sampling the profile sufficiently to detect closely spaced images and wasting too much computer time. Fig. 7 shows the results of such an analysis on the field of a globular cluster (PAL 8). It can be seen that whereas a straightforward isophotal analysis would cause many of the images to be misinterpreted the multiple isophote method manages to correctly identify the majority of the images and estimate sensible initial parameters. It is worth emphasizing that so far we have imposed no restrictions on the morphology of the overlapping images and as an illustration of the generality of the method Fig. 8 shows the interpretation of a deep frame of CCD data. Here the majority of the images are galaxies. This stage (iii) is in many ways the most difficult to fully automate and yet with a fairly simple and robust method we have gone a long way toward solving the problem.

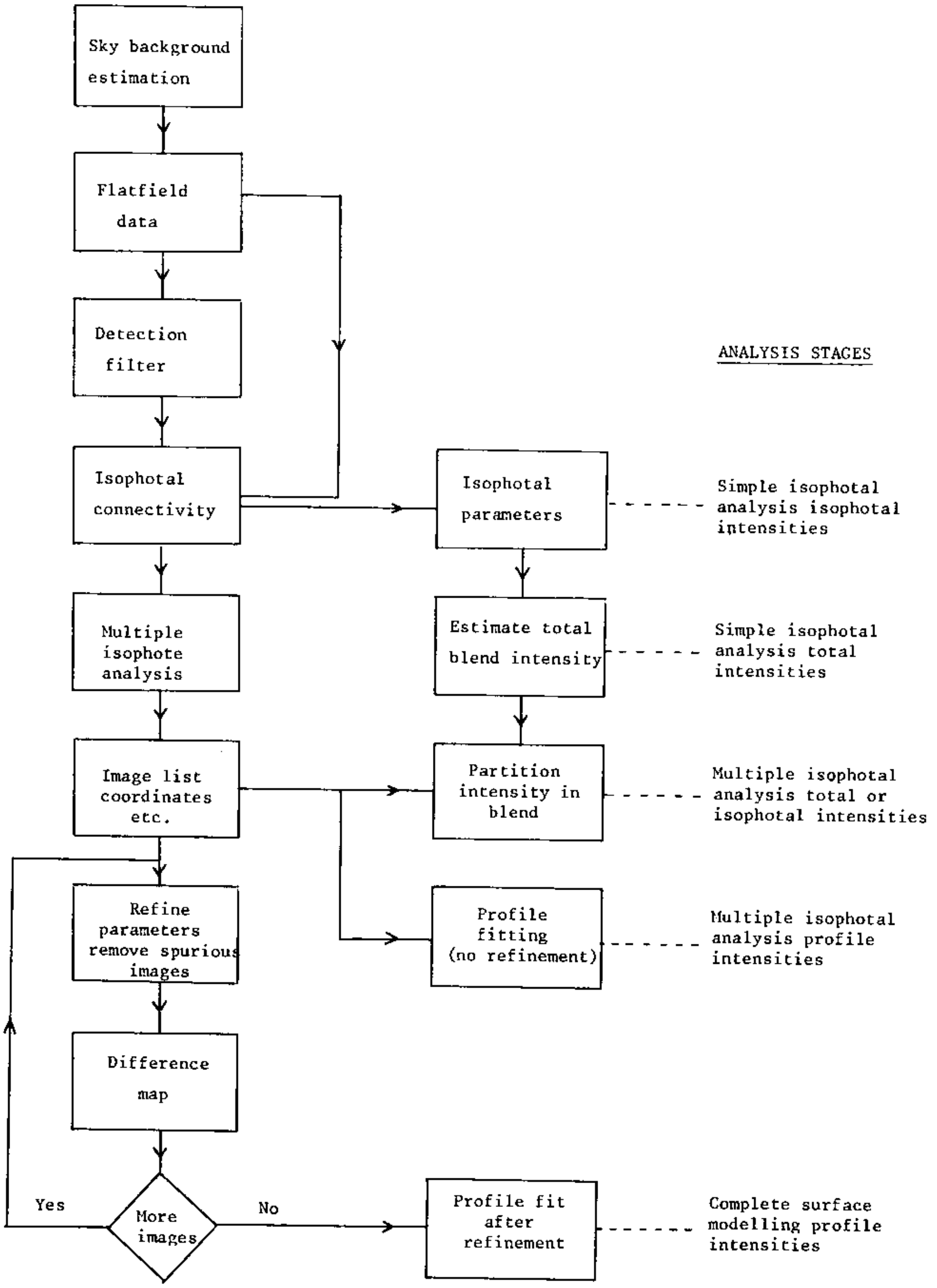
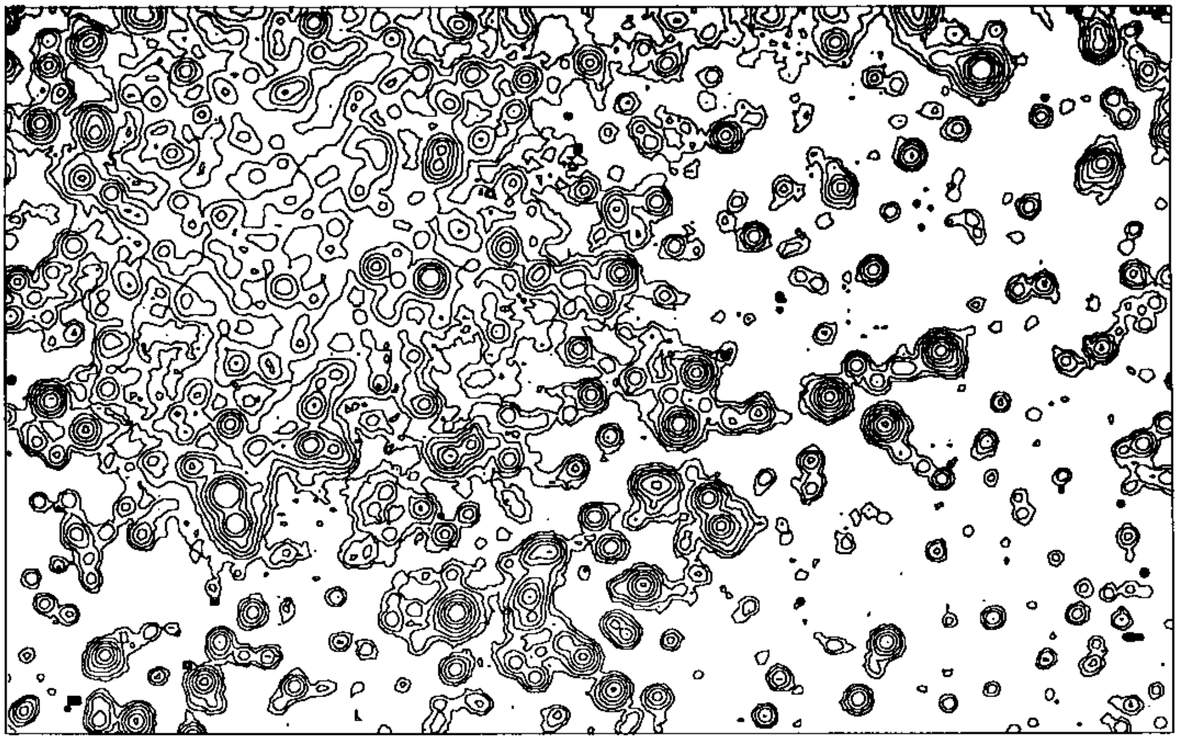
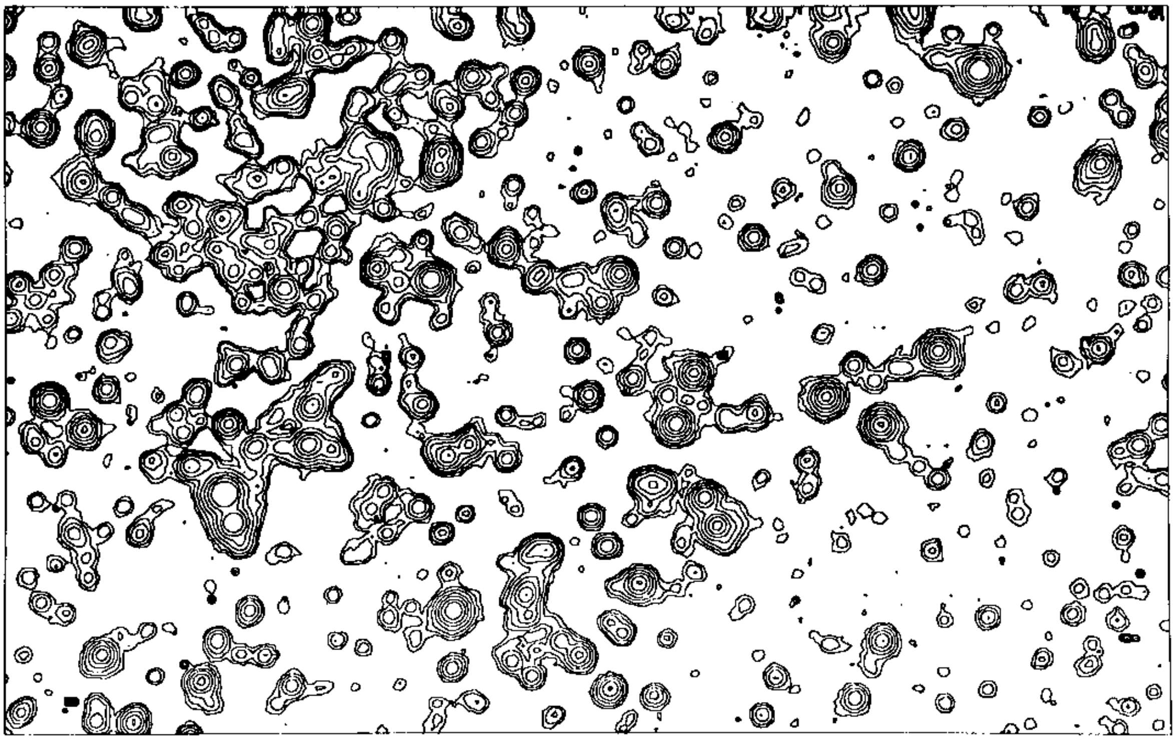


Figure 6. Crowded field algorithm flow diagram. The first three analysis stages are completely general, the last two are for known profiles only.

(c) The final stage (iv) consists of taking the initial parameters obtained previously – number of images, isophotal intensities, coordinates, shape information – and producing better estimates. In the preliminary analysis all parameters are derived using only the pixels above the isophote where the image separated out. At the very least it is necessary to adjust the image intensities to be a consistent estimate. Three different methods are used: fixed isophotal integration, total

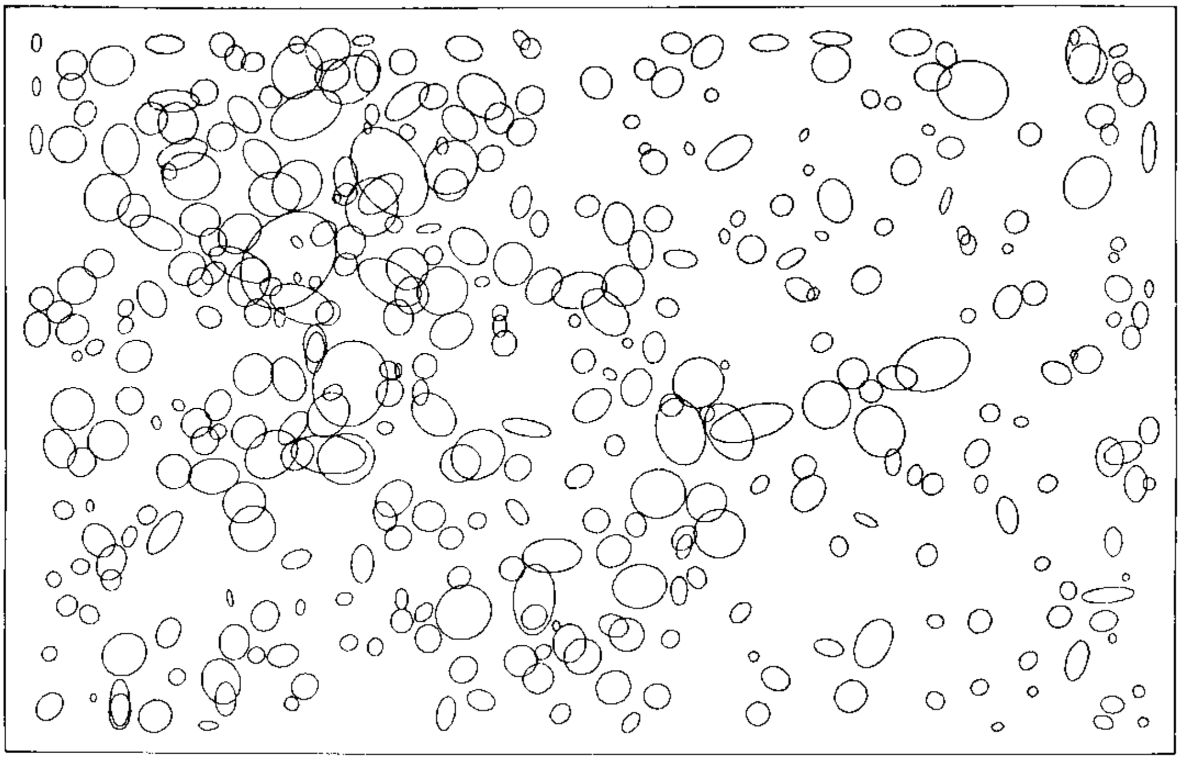


(a)

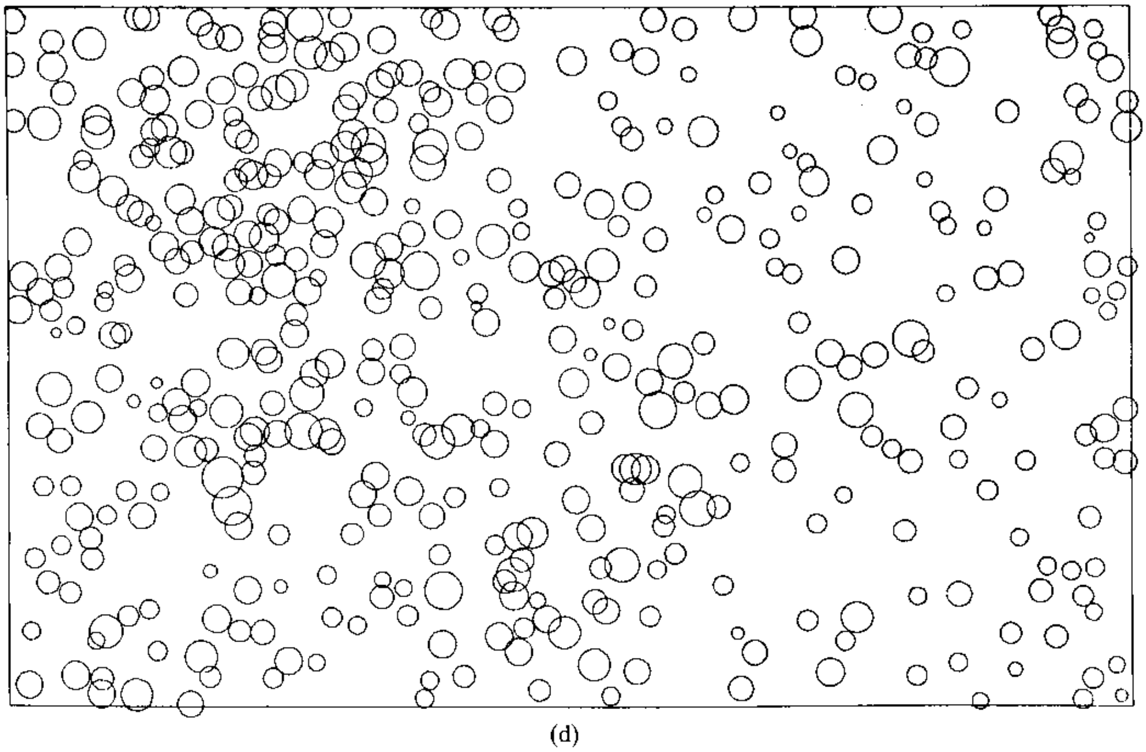


(b)

Figure 7. (a) The distant globular cluster PAL8 (L=14°, b=7°, r=≅30 kpc) as seen through a V filter using a CCD detector with 10-min exposure on the 1-m SAAO telescope at Sutherland. The complex background variations which make direct interpretation difficult are caused by the outer wings of stellar images near the centre, number density ~200 arcmin^{-2}, and the presence of unresolved images. (b) The same field after flat-fielding. In both cases the minimum contour level is set at sky+2σ rms, the contour spacing changes by ½ magnitude intervals. The very sharp images are cosmic ray events recorded by the CCD. Note that the image blends near the centre of the cluster now have a much more simplified appearance. (c)An interpretation of the above contour map using the general multiple isophote method outlined in the text. At this stage no restrictions on image profiles have been enforced. (d) The interpretation obtained using full surface modelling. A 2σ analysis threshold was used in both (c) and (d) with the result that some faint images have been missed.



(c)



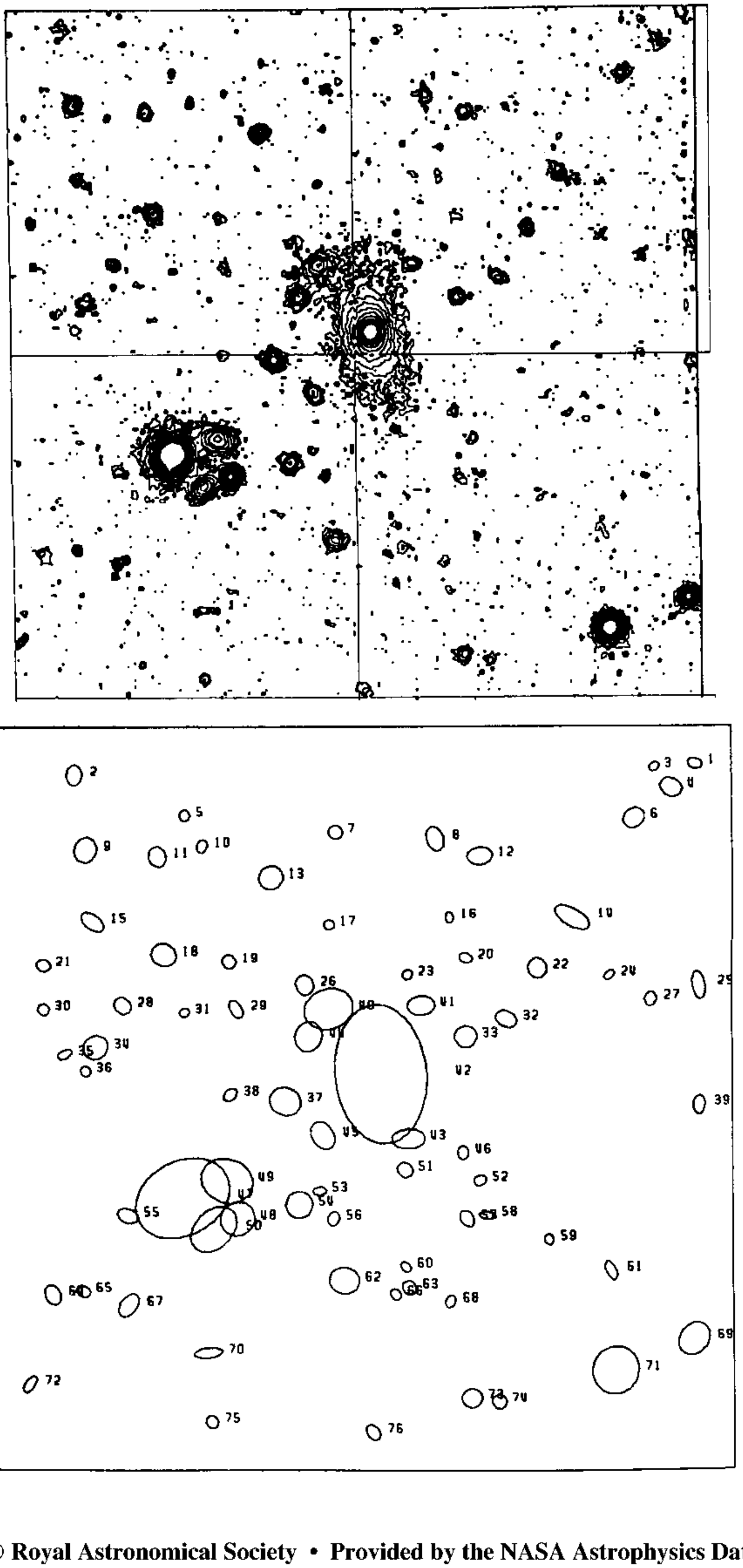
(d)

Figure 7 – *continued*

Figure 8. (a) A flat-fielded contour map of part of a deep CCD drift scan frame where the majority of the images are distant galaxies (see Hall & Mackay 1984 for further details). (b) The multiple isophote interpretation of this field.

(a)

(b)


2
3
1
4
5
6
7
8
9
11
10
12
13
15
17
16
14
18
19
20
21
22
23
24
25
26
27
30
28
31
29
40
41
32
33
34
35
44
36
42
38
37
39
45
43
46
51
49
52
53
47
54
55
48
50
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76

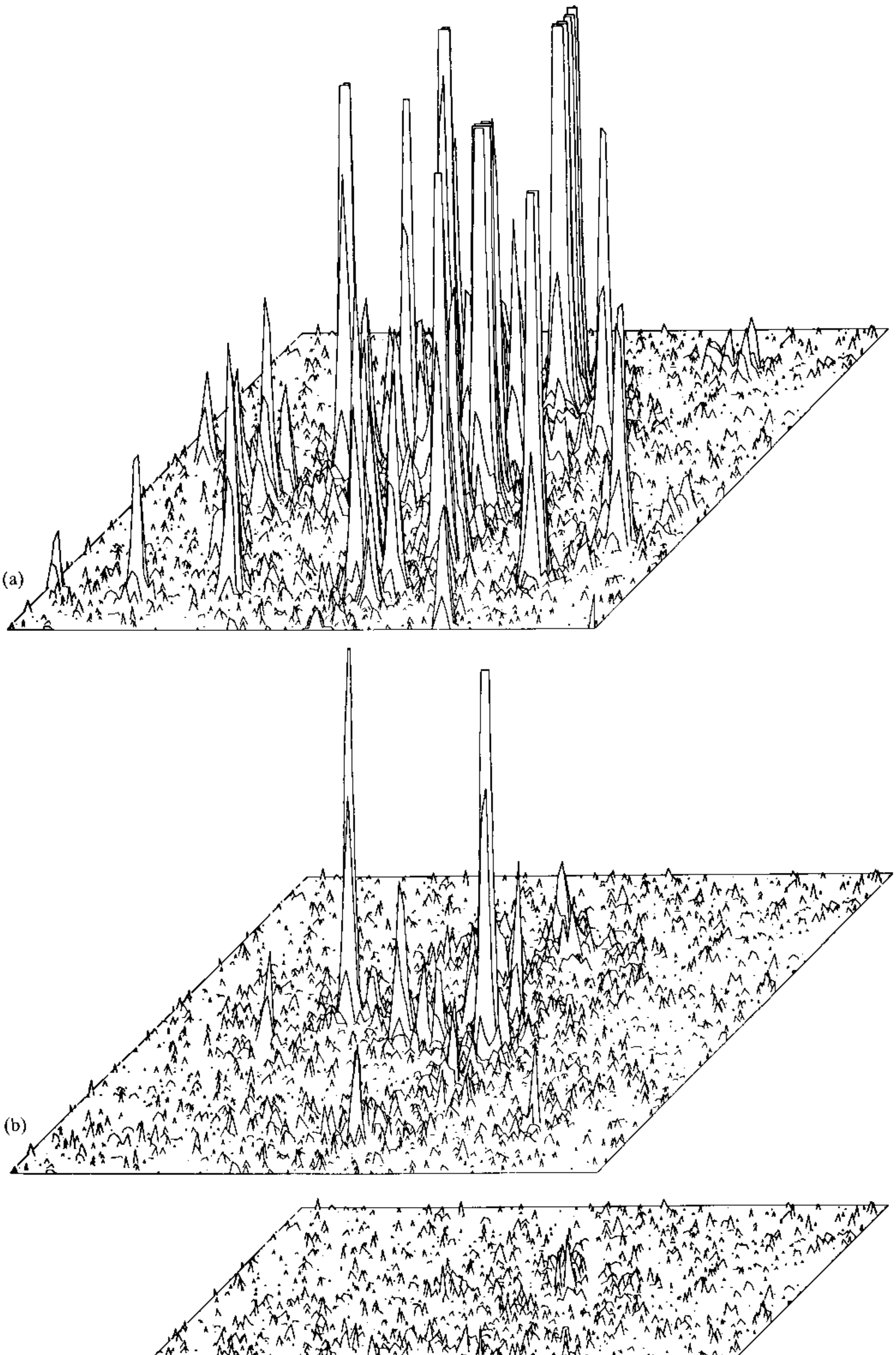
integration and profile fitting. Of these the first two are for the general situation where the images have different unknown profiles (e.g. mix of galaxies) and basically involve partitioning the intensity within a blended image into its constituent parts using measures of image profile information in such a way as to conserve the total intensity of the overall group. The third method profile fitting can essentially be done using a least-squares method as discussed in Section 5. The computational cost is small since we know the extent of the blend and have available fairly accurate image coordinates.

(d) The total intensity of a blend, or isolated image, when the profile is unknown is calculated in the following way. In order to minimize contamination from adjacent blends or images and also to minimize the sky background noise included, elliptical apertures are used (Irwin & Hall 1982), with ellipse parameters determined from the intensity-weighted second moments available from the isophotal analysis. (It is necessary to correct the calculated ellipticity for the noise effects as indicated earlier, since the noise strongly increases estimated ellipticities especially for faint images.) For the faintest images knowledge of the peak intensity and threshold isophote suffice to define approximately how far out along the intensity growth curve we are. It is then possible to specify a maximum radius within which we may be confident that >99 per cent of the image intensity lies. For brighter features the threshold isophote already contains the majority of the intensity, so defining a maximum radius such that the area of the ellipse is double the isophotal area again guarantees that we have included the whole image. The total area is then partitioned using 10 equally spaced radii between the isophotal radius and the maximum radius and the integrated intensity interior to these radii is calculated. An automatic examination of this sequence of intensities is made to determine the most likely value for the total intensity. We are in effect looking at the intensity growth curve of each feature, so a suitable cut-off point can be chosen to include the whole image, the minimum amount of sky and to avoid any other images entering the aperture.

(e) All refinement and profile fitting steps involve an independent coordinate and intensity stage. After the initial coordinates are obtained from the multiple isophote stage we can either stop after a single stage of linear least-squares profile fitting or enter the full surface modelling stage. The advantage of a single pass through the data lies in computing speed. Typically the multiple isophote analysis with either single profile fitting or intensity partitioning takes roughly twice as long as a conventional isophotal analysis scheme. The only real disadvantage when profile fitting lies in failing to resolve closely spaced images. Even if the profile is not known accurately this will only introduce a constant offset in magnitude. The intensity estimates are also fairly resilient to coordinate errors. For a Gaussian profile a position error Δr results in an intensity underestimate of roughly $\exp(-\Delta^2 r/2\sigma^2)$, which even for faint images gives rise to errors of only a few per cent. This in itself anticipates an obvious refinement scheme: calculating image coordinates (maximum likelihood method), profile fit (least-squares method) and repeat until convergence.

(f) During refinement for full surface modelling spurious images can usually be easily detected and removed as indicated in Section 5. By studying the difference map between data and model function within each blend previously unresolved images can also be detected and included in the refinement (see Fig. 9). With caution this procedure can be repeated (usually one extra cycle is

Figure 9. (a) A perspective view of the central region of a simulated stellar crowded field (number density ~200 arcmin^{-2}, seeing ~1 arcsec). (b) A difference map of the same region obtained by subtracting a model map generated from the output of the multiple isophote analysis. Although most of the stars have been correctly identified and thereby removed from the difference map a few closely spaced images were misinterpreted. (c) The difference map obtained after the automatic full surface modelling stage. All well-resolved images have been detected leaving only noise-like structure.



(a) (b) (c)

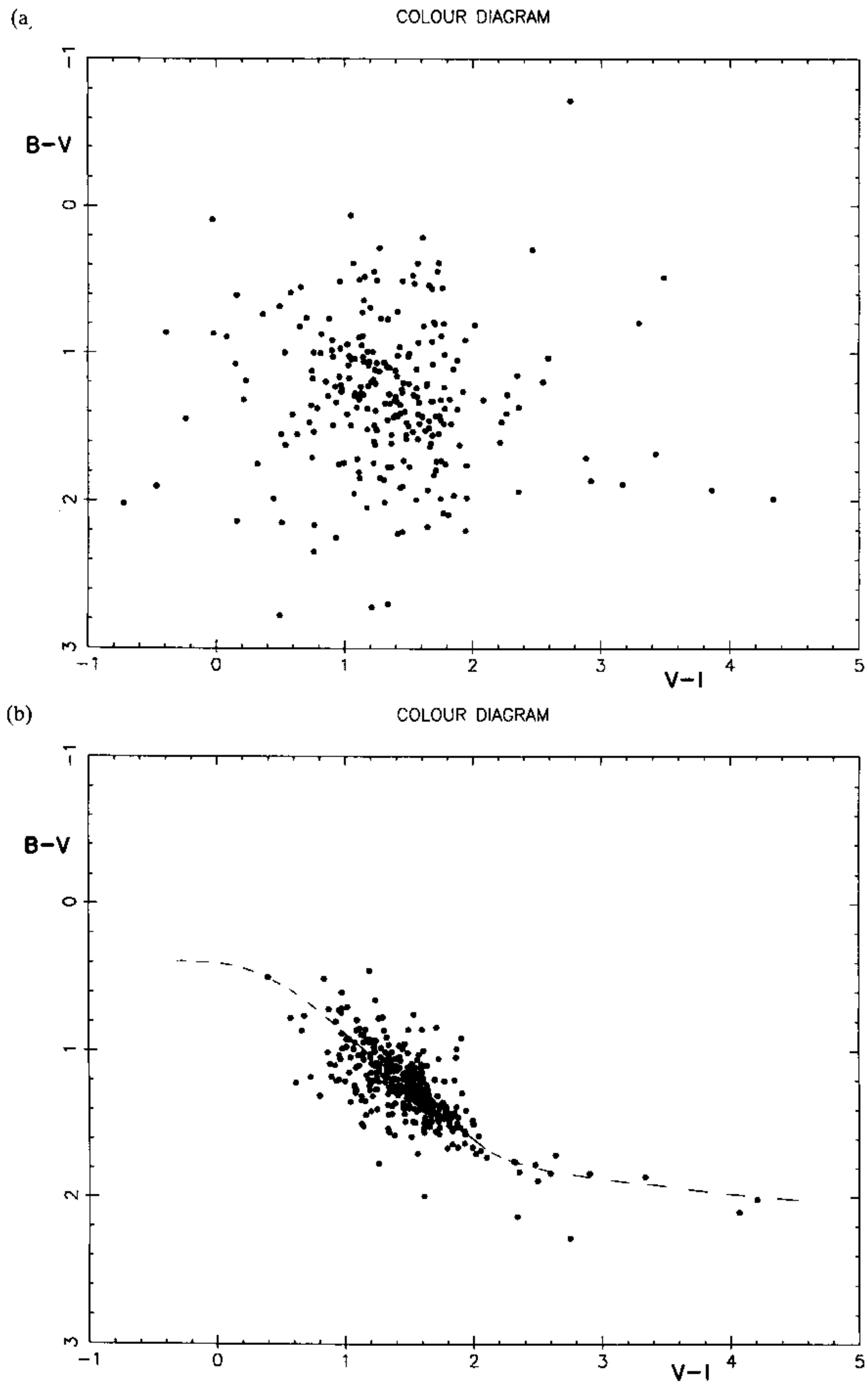


Figure 10. (a, b) Multiple isophotal analysis and full surface modelling two-colour diagrams (*BVI*) of PAL8. The isophotal analysis is completely misleading, whilst both the profile fitting methods (see Fig. 6) produce acceptable results. The main improvement using the surface modelling technique is in the removal of spurious points. This comes in part from the ability to treat each frame of data in an identical fashion, and also from the improved interpretation available for the master frame. The dashed line is the reddening-corrected locus of points for dwarf and giant stars. The reddening derived from a best fit of this curve gives a value of $E(B-V) \sim 0.35$ in agreement with other estimates. The three red stars, $B-V \geqslant 2.0$, lying well below the main stellar locus are genuinely different in colour.

sufficient) until the difference map is featureless, implying that all resolved images have been accounted for. Although this procedure is similar to that used in CLEAN in radio astronomy (Hogbom 1974) it differs in one vital respect. Here we are interested in using the absolute *minimum* number of images to model the surface; there is no such restriction in CLEAN.

(g) The sky background over the blend may also be refined during profile fitting stages (equation 20). However it is only worthwhile doing so if the profiles are known to be *accurate* and *all* images in the blend have been accounted for. Otherwise it is possible to end up with a worse sky estimate then before.

(h) When analysing multiple frames of the same area of sky once a master coordinate list has been obtained and frame-to-frame coordinate transformations ascertained several short cuts are possible. For example: for complete surface modelling once the data has been background-corrected it is better to go straight to the final profile fitting stage. This will ensure that different frames are analysed in exactly the same way. Instead of using the data to define threshold isophotes the model can be used in an analogous fashion. Only those images within a certain distance need appear in any one blend and only pixels above a specified threshold or within a fixed radius need be used in refinement. (A similar approach can also be used for dealing with an external coordinate list, with or without coordinate refinement etc.)

7 Discussion

For the data used in Fig. 7 the full surface modelling takes roughly 30 min on a VAX 11/780, whilst the simple multiple isophotal analysis with profile fitting takes around 5 min. In many cases the latter approach will produce acceptable results and in terms of ability to detect the presence of images in blends performs as well as the idealized algorithm mentioned previously. Since this stage could easily be added on to existing isophotal techniques it also seems a very promising candidate for integrating into machine measuring systems. The full surface modelling, which should be stressed is completely automatic, seems to perform close to the theoretical limits as can be inferred by examining the interpretation in Figs 7 and 9. At present it probably is too expensive computationally for large-scale photographic studies, but it is certainly usable for CCD data.

In order to compare the astronomical output available from various stages of the crowded field algorithm some two-colour diagrams of PAL8, a distant low-latitude galactic globular cluster, are presented in Fig. 10 derived from: multiple isophotal analysis, using total intensities and blend partitioning, and surface modelling. Although the seeing is only moderate (2.5 arcsec) perfectly acceptable results are obtained from both the simple profile fitting and full surface modelling. A conventional isophotal analysis is completely misleading because of overlapped images, isophotal intensities etc. Indeed for both profile fitting methods results almost identical to the theoretical error bounds are obtained. The main difference between the profile methods is that the full surface modelling results in virtually no spurious magnitudes due to the failure to detect unresolved images. There are still a few of these with the simple multiple isophotal method.

The complete crowded field algorithm has various options which enable it to be used as a general analysis tool in addition to its ability to operate reliably in severely crowded fields ($\sim$200 images arcmin^{-2}). It therefore should open up a large volume of astronomical data to automatic analysis and still provide useful results into the centre of nearby resolved galaxies and globular clusters in addition to its potential uses in deep galaxy photometry.

Acknowledgments

The CCD data for PAL8 was obtained on the 1-m telescope at SAAO and is part of a study into the properties of this globular cluster which will be published shortly. My thanks are also due to

Craig Mackay and Penny Hall for use of the KPNO CCD data and many fruitful discussions during the development of the ideas presented here. This work was supported financially by SERC through the auspices of the APM facility at Cambridge. I would also like to thank the rest of the APM group, Ed Kibblewhite, Mick Bridgeland and Peter Bunclark for many invaluable discussions and ideas.

References

Auer, L. H. & van Altena, W. F., 1978. *Astr. J.*, **83**, 531.

Bijaoui, A., 1979. *Image Processing in Astronomy*, p. 173, eds Sedmak, G., Capaccioli, M. & Allen, R. J., Osservatorio Astronomico di Trieste.

Bijaoui, A., 1980. *Astr. Astrophys.*, **84**, 81.

Chiu, L. G., van Altena, W. F. & Stetson, P. B., 1979. *Image Processing in Astronomy*, p. 240, eds Sedmak, G., Capaccioli, M. & Allen, R. J., Osservatorio Astronimco di Trieste.

Cramer, H., 1945. *Mathematical Methods in Statistics*, Princeton.

Eadie, W. T., Drijard, D., James, F. E., Ross, M. & Sadoulet, B., 1982. *Statistical Methods in Experimental Physics*, North Holland.

Farrell, E. J., Zimmerman, C. D., Nickel, D. F. & Borden, R. C., 1967. NASA Report No. CR-672.

Fellgett, P. B., 1970. *Optics Technology*, p. 61.

Fisher, R. A., 1958. *Statistical Methods for Research Workers*, Oliver & Boyd.

Hall, P. & Mackay, C. D., 1984. *Mon. Not. R. astr. Soc.*, **210**, 979.

Herzog, A. D. & Illingworth, G., 1977. *Astrophys. J. Suppl.*, **33**, 55.

Hogbom, J. A., 1974. *Astr. Astrophys. Suppl.*, **15**, 417.

Irwin, M. J. & Hall, P., 1982. *Occ. Repts R. Obs. Edin.*, **10**, 111.

Irwin, M. J. & Trimble, V., 1984. *Astr. J.*, **89**, 83.

Jarvis, J. F. & Tyson, J. A., 1981. *Astr. J.*, **86**, 476.

Kibblewhite, E. J., 1971. *PhD thesis*, University of Cambridge.

Kibblewhite, E. J., Hooley, A., Bridgeland, M. T. & Horne, D. A., 1979. *Data Base Techniques for Pictorial Applications*, **81**, 413.

King, I. R., 1971. *Publs astr. Soc. Pacif.*, **83**, 199.

King, I. R., 1983. *Publs astr. Soc. Pacif.*, **95**, 163.

King, I. R., Hedemann, E. Jr., Hodge, S. M. & White, R. E., 1968. *Astr. J.*, **73**, 456.

Kron, R. G., 1980. *Astrophys. J. Suppl.*, **43**, 305.

Newell, E. B., 1965. *MSc thesis*, University of Melbourne.

Newell, E. B., 1979. *Image Processing in Astronomy*, p. 100, eds Sedmak, G., Capaccioli, M. & Allen, R. J., Osservatorio Astronomico di Trieste.

Newell, E. B., 1982. *Occ. Repts R. Obs. Edin.*, **10**, 15.

Newell, E. B. & O'Neil, E. J., 1974. In: *Electrography and Astronomical Applications*, p. 153, eds Chincarini, G. L., Griboval, P. J. & Smith, H. J., McDonald Observatory.

Newell, E. B. & O'Neill, E. J., 1977. *Publs astr. Soc. Pacif.*, **89**, 925.

Pratt, W. K., 1978. *Digital Image Processing*, Wiley.

Ratnatunga, K. U. & Newell, E. B., 1984. *Astr. J.*, **89**, 176.

Rheault, C. & Hardy, E., 1980. *SPIE Applications of Digital Image Processing to Astronomy*, **264**, 200.

Shannon, C. E., 1948. *Bell System Tech. J.*, **27**, 379, 622.

Stobie, R. S., Smith, G. M., Lutz, R. K. & Martin, R., 1979. *Image Processing in Astronomy*, p. 48, eds Sedmak, G., Capaccioli, M. & Allen, R. J., Osservatorio Astronomico di Trieste.

Tody, D., 1980. *SPIE Applications of Digital Image Processing to Astronomy*, **264**, 81.

van Altena, W. F., Franz, O. G. & Fredrick, L. W., 1974. *New Problems in Astronomy*, p. 283, eds Gliese, W., Murray, C. A. & Tucker, R. H., Reidel, Dordrecht, Holland.