Deep Galactic Survey of a CCD Image

Maike Lenz, Jan 2021

***Abstract*—In this project, a deep galactic survey was performed on a CCD image from the Kitt Peak $4\mathrm{m}$ telescope with a Sloan r band filter. A number count curve with gradient 0.486 ± 0.004 was produced. The source flux of each object in the image was determined by imposing a fixed aperture and subtracting the local background through a reference annulus. By fine-tuning the photometry analysis parameters, it was found that the optimal results for the image occurred when the aperture size was $12\mathrm{px}$, the reference annulus size was $5\mathrm{px}$, and $100\mathrm{px}$ were cut off the edge to account for bright edge noise. The Sérsic profiles of sufficiently bright sources were computed by fitting a curve to the intensity distribution around the centre of each source along the horizontal direction. This gave inconclusive results but can be extended to give a useful insight into the nature of the sources in the image.**

I. Introduction

GALAXY number counts have been essential tools in investigating the large scale matter distribution of the universe [1]. In this investigation, a galactic survey was performed on a deep optical image from the Kitt Peak $4\mathrm{m}$ telescope using a CCD mosaic camera and a Sloan r band filter [2] with a central wavelength of $620\mathrm{nm}$. The aim was to create a catalogue containing the location, flux, and apparent magnitude for each source in the image. The necessary global analysis included finding a meaningful background threshold, and coding an algorithm to detect the sources. This was followed by local photometry analysis of each source. The apparent magnitudes from the catalogue were then processed to find the number of sources brighter than a limiting magnitude resulting in the number count curve.

The Sérsic profiles are a measure of how centrally concentrated the intensity distribution of a source is, denoted by the Sérsic index. This can give an insight into the nature of the sources in the image since different types of stars and galaxies have known Sérsic indices.

II. Global Analysis

A. Background

A crucial first step analysing the image was to determine a threshold for the background pixel value. A histogram of pixel values was plotted as shown in Fig. 1.

A Gaussian fit was performed on the peak giving a mean background of 4319 with a standard deviation of 12. In the following analysis, the background threshold was set at three standard deviations away from the mean, corresponding to a value of 3455.

B. Points Catalogue

Before conducting detailed photometry analysis, a global algorithm was coded to find the centre of each source. The image contained significant bleeding from a central star. These bright regions were manually masked, particularly in the bottom section of the image. Subsequently, objects containing pixels above the saturation value of 36000 were disregarded.

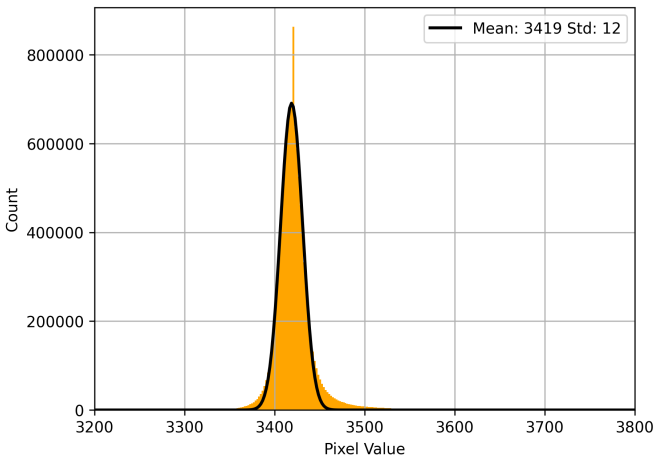

Fig. 1: A histogram of the pixel values in the image is plotted in orange. The black line depicts a gaussian curve fit with mean 3419, standard deviation 12. The tail to the right of the peak corresponds to the flux due to sources in the image.

Following this, a global algorithm was run on the image. The pixels were analysed in order of decreasing values. This recorded the brightest pixel in a source as its centre. They were appended to a CSV file which was later expanded upon with photometry analysis to produce the final catalogue. To avoid counting multiple points for the same source, a circular mask was applied around the centre. The radius of this mask was determined by finding how far out from the central pixel the value falls below the background threshold. This process is visualised in Fig. 2.

III. Local Analysis

A. Photometry

A fixed aperture approach was taken to analyse the sources. The global analysis yielded a CSV file with the brightest pixel for each source. The photometry process is depicted in Fig. 3.

To create the aperture, a circular mask of a fixed diameter was coded around each point in the file. This masked everything outside the aperture. The sum of pixel values through this aperture was calculated for each source to give its total flux.

To account for the background, a reference annulus consisting of two concentric circles was implemented by imposing a separate mask around each central point. The inner circle of the annulus was the aperture. The outer circle of the

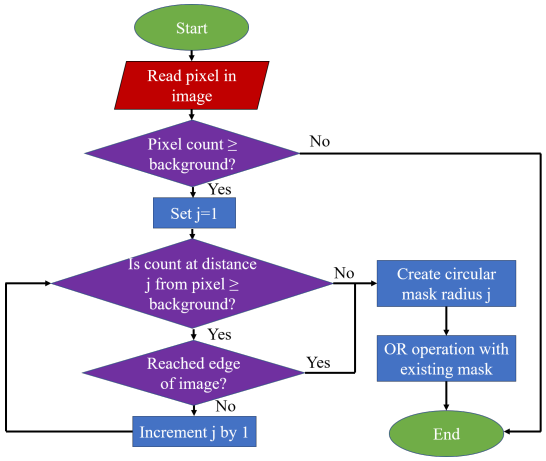

Fig. 2: Flowchart depicting the creation of the points CSV file. The pictured steps were run on every pixel in the image in order of decreasing brightness. This took place after the bleeding and over-saturated sections of the image had been masked.

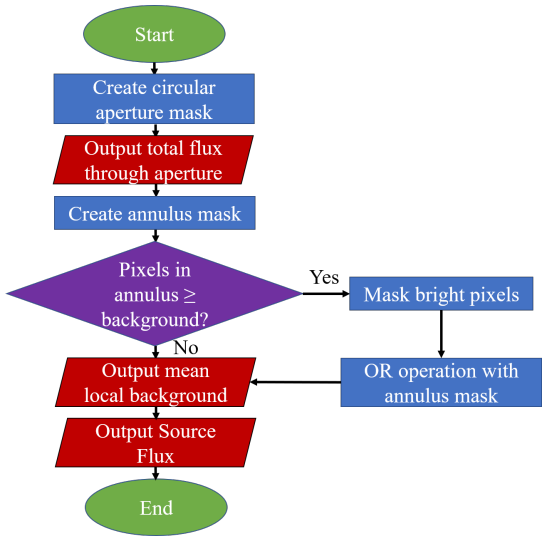

Fig. 3: Flowchart depicting the photometry analysis algorithm to find counts corresponding to total flux, local background, and source flux. This was run on every point in the points CSV file.

annulus was the aperture plus an additional radius denoted by the annulus size. The inside of the aperture and the outside of the outer circle were masked. To avoid including other sources in the proximity of the one being analysed, pixels inside the reference annulus above the background threshold were masked. The mean pixel count in the annulus then corresponded to the local background around the source.

The local background value was scaled up to give its total contribution within the aperture. This was subtracted from the total flux to give the source flux count.

B. Apparent Magnitude

The source flux counts were converted to instrumental magnitudes, $\mathrm{mag_i}$ with equation 1 [3].

$$\mathrm{mag_i} = -2.5\log_{10}(\mathrm{counts}) \tag{1}$$

This was calibrated to give the apparent magnitude, m, by adding the instrumental zero-point magnitude, $\mathrm{ZP_{inst}}$, from the image header as in equation 2.

$$\mathrm{m} = \mathrm{ZP_{inst}} + \mathrm{mag_i} \tag{2}$$

From the inverse square law of radiation, the cumulative number of sources, N, that is brighter than a limiting magnitude, m, is given by equation 3 [1].

$$\log_{10}(N) = 0.6\mathrm{m} + \mathrm{constant} \tag{3}$$

This equation is appropriate in Euclidean space with a uniform distribution of sources over large solid angles [1].

C. Errors in Apparent Magnitude

The shot noise from the CCD telescope follows a Poisson distribution [4]. By Poisson statistics, the error in the total flux of the sources is given by its square root [3]. Error propagation with equations 1-3 then gave the error in apparent magnitudes as

$$\sigma_m = \sqrt{\sigma_{ZP}^2 + \frac{25}{4(\ln(10))^2}\frac{f_t + \frac{f_r}{N_a^2}}{f_s^2}} \tag{4}$$

where σ_m is the error in the apparent magnitude, σ_{ZP} is the error in the zero-point magnitude as given in the header, f_t is the total flux, f_r is the flux through the annulus, and N_a is the number of pixels in the aperture. It is not straight forward to visualise errors associated with apparent magnitude in the number count graph, so σ_m was calculated for each source and appended to the catalogue.

D. Catalogue

The final catalogue contained the x-index, the y-index, the total flux, the source flux, the local background, the apparent magnitude, and the error in apparent magnitude for each source. The x- and y-indices were the values in the points CSV file. The other parameters were appended in the local photometry analysis.

By taking values from the catalogue and computing the cumulative numbers of objects brighter than a limiting magnitude, the number counts were then plotted. The aperture size, reference annulus size, and the edge-cutoff were then fine-tuned to optimise the number count plot.

IV. Fine-Tuning Parameters

A. Aperture Size

To investigate the effect of varying the aperture on the outcome, the number counts were plotted for varying aperture sizes as shown in Fig. 4. The aperture size refers to the diameter of the aperture in pixels.

The aperture had to be large enough to include most sources in their entirety but not so large that more than one source contributes to the flux. The smallest aperture which includes most sources corresponded to a size of 12px.

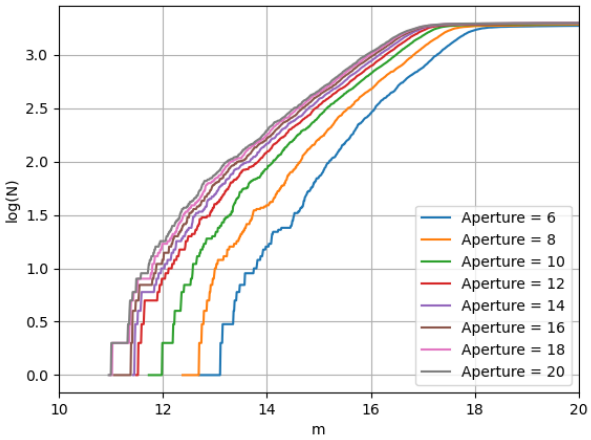

Fig. 4: The number counts of N sources brighter than a limiting magnitude m for apertures ranging from diameter 6px to 20px. There is a large difference with widening apertures up to 12px after which the curves are much closer together. The large left-shifts of the small aperture curves may indicate that there are a non-negligible number of sources larger than the aperture which do not fit into the aperture. This artificially decreases their brightness. The reference annulus was fixed at 5px.

B. Reference Annulus Size

The size of the reference annulus refers to the difference in radii of the aperture and the outer circle. To account for other sources within the annulus, pixels above three standard deviations of the mean background were masked inside the reference annulus. Only pixels with counts below that threshold contributed to the local background. To test this whether this works, the annulus size was varied between 3px and 9px. With any other sources in the annulus masked, the should change minimally when a different sized annulus is used to find the local background. The results are plotted in Fig. 5.

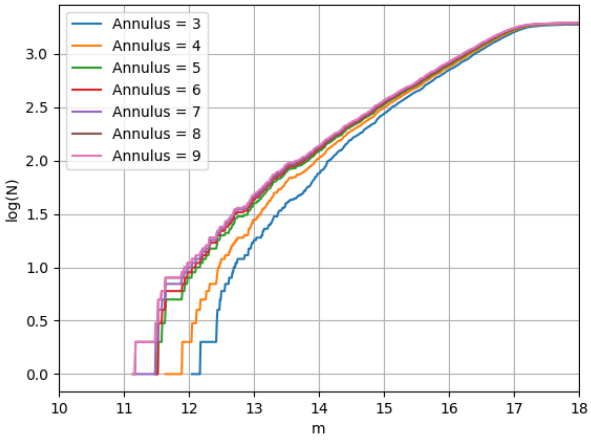

Fig. 5: Number counts of N sources brighter than a limiting magnitude m for reference annuli varying from 3px to 9px. The aperture was fixed at 12px.

As expected, the population curves for reference annuli of size 5px and larger were very similar which indicates that the masking within the reference worked correctly. Although a larger annulus includes more surrounding sources, these do not contribute significantly to the local background. However, there was a shift for very small annulus of sizes 3px and 4px. This may be due to the annulus being so thin that a trail-off of the source which is just below the background threshold contributed significantly. This may have artificially increased the local background and hence decreased the source flux giving a higher apparent magnitude than expected. This may explain the shift to the right in number curves at low m. Any annulus size above 4px yielded the required result; in subsequent analysis, size 5px was chosen for the reference annulus.

C. Edge Noise

Due to fewer sub-exposures, there is noticeably bright noise around the edges of the image. Although objects above the saturation value were masked in the global analysis, there may still be remnants of this noise. To investigate this, a variable number of pixels was sliced off the edges of the image. This number corresponded to the 'edge thickness'. The results are plotted in Fig. 6.

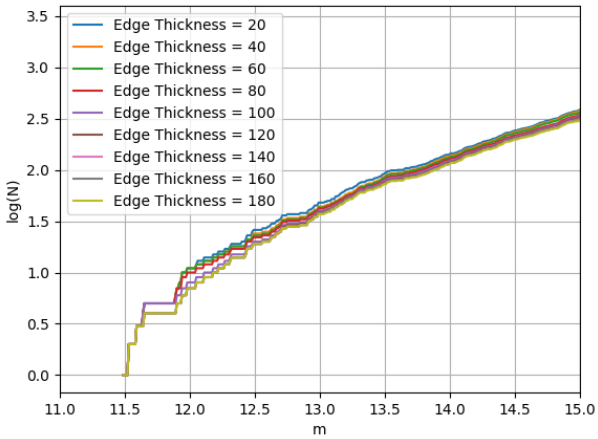

Fig. 6: Number counts of N sources brighter than a limiting magnitude m with an edge thickness ranging from 20px to 180px. The curve shapes are very similar beyond $m \approx 13$. Below that, there higher edge cutoff curves noticeably shift to the right.

As evident, there is little difference for different edge thicknesses for the high-m regime. In the low-m section of the graph, however, the curve appears to shift to the right with increasing edge thickness. This may indicate that there are bright noise pixels, corresponding to low values of m, which contribute to the graph. In further analysis, 100px were cut off the edges.

V. Number Count Curve Result

The final number count result was found with an aperture size of 12px, annulus size of 5px, and 100px cut off the edges

to account for noise. It is plotted below in Fig. 7. A straight line fit was performed on the linear section of the curve.

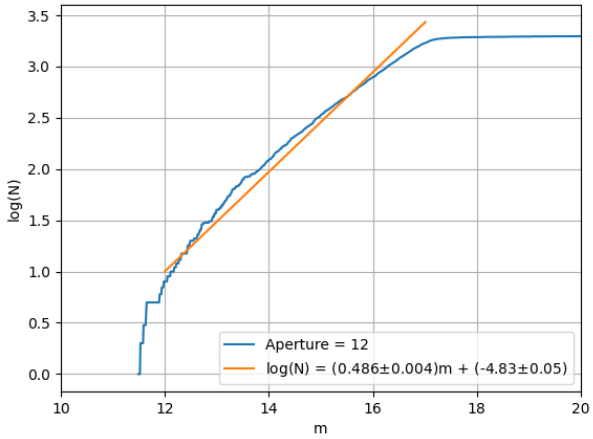

Fig. 7: Number counts of N sources brighter than a limiting magnitude m for aperture 12px, reference annulus 5px, and edge cutoff 100px. The linear curve fit computed the gradient as 0.486 ± 0.004 compared to the theoretical value of 0.6.

As expected, the curve flattens out around m-values of 17 as the telescope cannot detect sources dimmer than its limit. Similarly, there is a vertical section at $m \approx 11.5$, as the saturation pixel count cannot be surpassed. In between these regimes lies a linear section with gradient 0.486 ± 0.004. This is lower than the theoretical value of 0.6, as indicated in equation 3. Reasons for this may be non-Euclidean geometry [1], the evolution of distant sources, or not a large enough section of the sky being contained in the image. Working on a solid angle that is too small means that there likely is not a sufficiently uniform distribution of sources. A lower gradient than predicted indicates more bright sources which may mean that there are more sources close to the telescope than distant ones.

VI. Sérsic Profiles

As an extension to the main part of the analysis, the Sérsic profiles for each source were determined by calculating the intensity distribution of each source from its central pixel. The Sérsic profile refers to an intensity distribution of the form

$$\ln(I(r)) = \ln(I_0) - kr^{1/n} \tag{5}$$

where $I(r)$ is the intensity at a radius r from the centre of the source, I_0 is the central intensity, k is a constant, and n is the Sérsic index. A higher Sérsic index implies a more centrally concentrated profile. For elliptical galaxies, de Vancouleurs' Law [5] predicts a Sérsic index of $n = 4$. Stars, however, have Gaussian intensity distributions [6] corresponding to a Sérsic index of $n = 0.5$. A curve fit of equation 5 was performed along the horizontal direction to optimise the parameter n for each source. The results are shown in the histogram in Fig. 8.

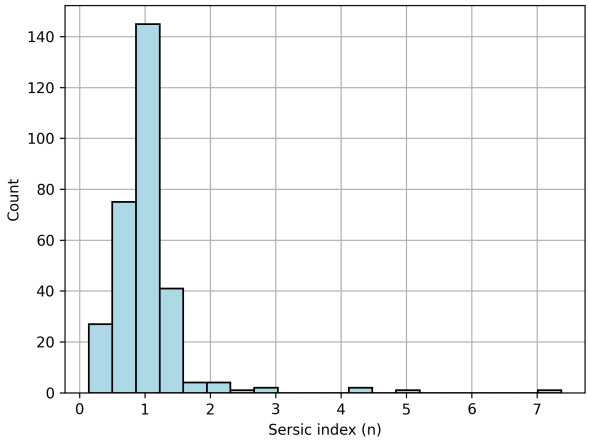

Fig. 8: Histogram depicting the Sérsic indices for the sources in the image.

The vast majority of calculated Sérsic indices lie around 1. This does not match the expectation, since elliptical galaxies are the most abundant and should dominate the histogram at around $n = 4$. This is likely due to limitations of the resolution of the image. It should be noted that many sources in the image are extremely small and very few available pixels will not result in an accurate curve fit. While the catalogue contained 3127 points, only 308 were suitable for the Sérsic analysis. Additionally, since the intensity distribution was only analysed along the horizontal direction, an asymmetric source at an angle might not be accurately fit to.

The results of this analysis of the Sérsic profiles need to be investigated further to give a meaningful result. This could be done by acquiring an image with a higher resolution or by also fitting in the vertical direction. A more accurate Sérsic index result could indicate useful information about the nature of the sources in the image even though real galaxies are well known to deviate from pure Sérsic profiles [7] due to a variety of reasons including asymmetry and varying ellipticity with radius. For big cluster galaxies, two independent Sérsic profiles are required for an accurate fit [8]. Hence Sérsic indices provide a useful starting point.

VII. Conclusion

A catalogue for an image was created using global source detection and photometry analysis. This contained the x-index, the y-index, the total flux, the source flux, the local background, the apparent magnitude, and the error in apparent magnitude for each source. The optimal results for the image were produced with the photometry parameters as shown in Table I.

A number count of the number of sources brighter than a limiting magnitude gave a gradient of 0.486 ± 0.004. This lies below the theoretical value of 0.6 given a uniform distribution of sources, Euclidean space, and completeness of the detected sources at low brightnesses. A fixed aperture means that some large sources are cut-off and recorded as dimmer than

Optimal Parameters	
Aperture Size	12px
Reference Annulus Size	5px
Edge Cutoff	100px

Table I: Optimally fine-tuned parameters for the local photometry analysis of the image.

they are. This artificially steepens the number count curve. A simple analysis of the Sérsic profiles of the sources gave an inconclusive result. The suitably bright and large sources predominantly have Sérsic indices of $n \approx 1$. This indicated that elliptical galaxies were not able to be picked up by the Sérsic profile fit.

As an extension, a variable aperture may be coded such that the aperture is extended for very large objects without including multiple small objects in other places. This would avoid these sources being cut-off and effectively dimmed.

Despite limitations, this investigation provided a valuable insight into the contents of the telescope image. The code is available at https://github.com/carawaters/AstroImageProcess.

REFERENCES

[1] A. Sandage, "Observational tests of world models," *Annual review of astronomy and astrophysics*, vol. 26, pp. 561–630, 1988.
[2] P. Massey, T. Armandroff, J. De Veny, C. Claver, C. Harmer, G. Jacoby, B. Schoening, and D. Silva, "Direct imaging manual for kitt peak," *Tuscon: NAO*, 2002.
[3] D. L. Clements, N. Skrzypek, and Y. Unruh, *Astronomical Image Processing. A1*. Astrophysics Group, Imperial College, 2019.
[4] G. E. Healey and R. Kondepudy, "Radiometric ccd camera calibration and noise estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 3, pp. 267–276, 1994.
[5] A. Mazure and H. Capelato, "Exact solutions for the spatial de vaucouleurs and sérsic laws and related quantities," *Astronomy & Astrophysics*, vol. 383, no. 2, pp. 384–389, 2002.
[6] X. Wan, G. Wang, X. Wei, J. Li, and G. Zhang, "Star centroiding based on fast gaussian fitting for star sensors," *Sensors*, vol. 18, no. 9, p. 2836, 2018.
[7] E. Pignatelli, G. Fasano, and P. Cassata, "Gasphot: a tool for galaxy automatic surface photometry," *Astronomy & Astrophysics*, vol. 446, no. 1, pp. 373–388, 2006.
[8] A. H. Gonzalez, A. I. Zabludoff, and D. Zaritsky, "Structural properties of brightest cluster galaxies," *Astrophysics and space science*, vol. 285, no. 1, pp. 67–71, 2003.