Macalester Journal of Physics and Astronomy

Volume 1
Issue 1 *Spring 2013*

Article 4

May 2013

Tip-of-the-Red-Giant-Branch Detection: Estimating the distance to resolved galaxies in SHIELD

Ian Cave
Macalester College, kylerayner8@gmail.com

Abstract

The Survey of HI in Extremely Low-mass Dwarfs (SHIELD) is an ongoing study of twelve galaxies with HI masses between 10^6 and 10^7 Solar masses, detected by the Arecibo Legacy Fast ALFA (ALFALFA) survey. One of the important characteristics of a galaxy is its distance, since knowing that value allows us to calibrate many other factors and learn more about the galaxy. This research was focused on creating a program to systematically solve for the distance to the galaxies in SHIELD. To perform this calculation, Tip-of-the-Red-Giant-Branch (TRGB) detection was chosen, which is a popular method for galaxies with low numbers of stars. Presented here are the results of a new program for the SHIELD galaxies, along with estimated uncertainties and comparisons to distance estimates made by methods from other publications and associated SHIELD members.

Tip-of-the-Red-Giant-Branch Detection: Estimating the distance to resolved galaxies in SHIELD

Cover Page Footnote

Research Adviser: John Cannon Macalester College

Tip-of-the-Red-Giant-Branch Detection: Estimating the distance to resolved galaxies in SHIELD

Ian Cave
Macalester College
Research Adviser: John Cannon

Abstract

The Survey of HI in Extremely Low-mass Dwarfs (SHIELD) is an ongoing study of twelve galaxies with HI masses between 10^6 and 10^7 Solar masses, detected by the Arecibo Legacy Fast ALFA (ALFALFA) survey. One of the important characteristics of a galaxy is its distance, since knowing that value allows us to calibrate many other factors and learn more about the galaxy. This research was focused on creating a program to systematically solve for the distance to the galaxies in SHIELD. To perform this calculation, Tip-of-the-Red-Giant-Branch (TRGB) detection was chosen, which is a popular method for galaxies with low numbers of stars. Presented here are the results of a new program for the SHIELD galaxies, along with estimated uncertainties and comparisons to distance estimates made by methods from other publications and associated SHIELD members.

I. Introduction

One of the most important characteristics of a galaxy is its distance. Knowing the distance to a galaxy allows one to apply corrections to other data of that galaxy that then allow astronomers to learn even more interesting things. Consequently, much work has been done trying to find effective methods of measuring the distances to galaxies. As might be expected, no one method is perfectly suited to all applications, and so a few methods have been developed and explored. One of these is called the Tip-of-the-Red-Giant-Branch method, and it relies on identifying a certain feature in data of the galaxy, then comparing that measurement to the known (absolute) value of the feature. The goal of this project was to create a program that, given as little user input as possible, could take photometric data (measurements of the brightness of stars in the galaxy) and determine the distance to a given galaxy.

The data used in the creation of this project was obtained from John Cannon at Macalester College and Kristen McQuinn at the University of Minnesota. Both Dr. Cannon and Dr. McQuinn are involved in the Survey of HI in Extremely Low Mass Dwarfs (SHIELD). SHIELD is a project dedicated to understanding more about the under-studied extremely low mass dwarf galaxies. The data which is used for this project comes from images the Hubble Space Telescope (HST) took specifically for the survey. The images were reduced by Dr. McQuinn.

Using this data and a newly created program, estimates of the distances to many of the galaxies involved in the SHIELD project were obtained and compared to independent distance estimates of those same galaxies. In almost all cases, they agreed to a high degree of accuracy.

II. Methodology

A. Astronomy Primer

It is important to understand some of the astronomical background to this project before the implementation is explained.

If a star is resolved in an image (we can identify its total flux independently from other nearby stars), we can estimate its total flux. One can then convert the total flux into a unit or measurement called “magnitudes,” which is really just a measure of the brightness of a star. Magnitudes are logarithmic units, and the brighter an object is, the lower its apparent magnitude.

Another important piece of information is the idea that when we measure the brightness of a star in a certain filter, the absolute magnitude is different than what we measure because the object is moving relative to our point of observation, and is therefore red-shifted according to the Doppler effect. In addition, the flux that the we receive depends on the distance from the object. Thus, there is a disconnect between the definitions of apparent magnitude (what is measured) and absolute magnitude (the magnitude we would measure if the object in question were placed at a distance of ten parsecs). Since the TRGB has been studied extensively, we know that the absolute magnitude of the feature is around 4 magnitudes (Bellazzinni et al. 2001) for all galaxies.

Given all this information, we can use the “distance modulus equation” to convert a difference in magnitudes between the measured TRGB and the known absolute magnitude to a distance. The distance modulus equation is:

$$\log(d) = 1 + \frac{m - M}{5} \quad (1)$$

where *d* is the distance to the object (in parsecs), *m* is the apparent (measured) magnitude, and *M* is the absolute magnitude of the feature. The log is in base ten.

B. Background and Related Work

The Tip-of-the-Red-Giant-Branch (TRGB) method for estimating the distance to galaxies was first broached in 1944 (Lee et al. 1993) when three galaxies already thought to be at similar distances exhibited similarly bright stars as the brightest red stars in the respective systems. Thirty years later (Sandage, 1971), another study found that many galaxies in the local group exhibited a background sheet of red stars that peaked at a specific magnitude (brightness), and that these stars were similar to the ones found in 1944. The discovery of a certain population of stars at a constant magnitude allowed the distance modulus to come into play to start to define the distances to these galaxies.

Much can be learned about galaxies through their composition, and one of the diagrams that astronomers use to visualize the data we receive is called a color-magnitude diagram (CMD). A CMD simply plots the color of a star (difference in magnitudes between measurements of a star in two separate filters) versus its magnitude in one of the filters. Since certain brightnesses and colors represent very specific characteristics with respect to stellar composition, this is a useful way to visualize the relative strengths of different stellar populations in a galaxy. Figure 1 is an example of a CMD from the data that associated with this paper.

The TRGB occurs because some stars (those with a characteristic mass and composition) reach a stage in their evolution where they effectively stop getting brighter because they have attained core helium fusion, but will stay at that evolutionary stage for an extended period of time (up to 13 Gyr) (Lee et al., 1993). Stars like these have gone through an extended period of hydrogen fusion before attaining the densities necessary to ignite helium fusion. These stars, while large enough to ignite core helium fusion, do not yet have the densities required to ignite the triple alpha process which results in the helium core flash. Thus, stars accumulate at this tip where they spend a large amount of time at a fairly constant luminosity. Remarkably, this feature depends very weakly on the metallicity (molecular composition) of the galaxy. As such, TRGB detection is useful for distance measurements where the metallicity of a galaxy is not known beforehand or there is not enough data to accurately determine the metallicity. The fact that this stellar population is old and remains at this stage for a long time means that we will observe many of these types of stars in a given galaxy. Their small size means that they can occur in galaxies with low masses where there may not be enough mass per volume to create brighter, heavier stars. We call these stars red-giant-branch stars (RGB stars), and the tip of this feature is referred to as the TRGB.

The TRGB method does have some limitations, particularly that the photometry used must be highly accurate. Since the method relies crucially on the measured magnitudes of the stars and exactly where they fall in the CMD, the stars must be accurately measured with a high resolution instrument such as the HST. In addition, since many applications of TRGB detection method rely on statistical methods, greater numbers of measured stars leads to a more accurate distance estimate. The number of stars needed to make an accurate distance estimate can vary wildly from application to application, but findtrgb (see II. C.) can work with as few as twenty stars in the largest magnitude bin.

Early applications of the TRGB method tended to be less rigorous than might be ideal. The presence of the feature in a CMD is fairly easy to spot, but quantifying the actual relative magnitude of the feature with accuracy is something that is still being explored. In the early days of the method, this was obtained by simply eyeballing a value and taking that as the TRGB. Later applications have focused on defining the TRGB in some mathematically and statistically rigorous way.

Lee et al. (1993) focused on proving that the TRGB method could be effective in determining the distance to galaxies to within the same accuracy as already established methods such as Cepheids and RR Lyraes (both standard candles which have historically been used. For a comparison, see the Lee et al. paper). In their comparison, they found it to be as accurate a method as others, and in addition, it has some features that make it a more enticing alternative in certain situations. For instance, since these RGB stars are present in all galaxies, not just ones of a certain mass/age/metallicity, the method is viable for any galaxy. They also note the weak metallicity dependence.

Since the work of Lee et al., TRGB detection has continued to be an effective method of distance measurement, and others have developed tools and techniques to further enhance its usefulness. Makarov et al. (2006) developed a program called TRGBTool which they tested on a group of dwarf galaxies observed with the HST. For a full discussion of their method, the reader is directed to their paper. In short, they employed a maximum likelihood algorithm to define the RGB and obtain the TRGB. Makarov et al. also found the TRGB detection method to be accurate and useful in cases where other methods might have difficulty.

The Lee et al. paper provided much of the theoretical background for this project as far as methods and algorithms are concerned, and the approach used will be discussed in more detail in the following sections. Their use of a Sobel edge detection filter, in particular, turned out to be an elegantly simple way to quantify the TRGB. In addition to the newly created code, a copy of Makarov et al.'s TRGBTool was obtained and used to independently verify the distance measurements. In general, they are found to be in good agreement as will be discussed in the results section.

C. findtrgb

The program created for this project is structured to take a user through the process of finding the TRGB in a logical progression. Figure 2 shows an example of a successful completion of the program, and will be useful to refer to for the rest of this section. The code was written in IDL, a language used by many astronomers which has built in methods and features that were useful. To use the program, a photometry file for the galaxy must be present, detailing measurements of each of the stars in the galaxies (magnitudes in two filters, x and y positions of stars in an image, etc.). When first called, the program prompts the user if they would like to bring up an image of the galaxy. If they choose to do so, a file selector dialogue allows the user to navigate to and choose an image file (in FITS format) to be loaded.

In the loaded image, the user is prompted to select regions of the image that they would like to exclude from following calculations. Figure 3 shows an example of this kind of selection. In some cases where the number of stars is very low, this step may not be particularly useful because the statistics of the following calculations need to be as robust as possible. However, in most cases there are a few reasons a user would want to exclude stars that are present in particular regions of the galaxy. For instance, the central regions of a galaxy are where the most crowding occurs (the highest number density of stars per area). In these regions, it is possible that photometry from two stars is being blended, such that both measurements end up incorrect. In fact it is even possible that two stars occupy one pixel in the image, in which case there is no good way to extract accurate photometry about either star. In addition, sometimes there will be foreground/background objects present in the image of the galaxy. If this is the case, any stars present around those objects will be poorly measured for similar reasons the the center-of-the-galaxy argument above. Thus, regions around foreground/background objects should be selected to be excluded. One of the attractive features of the RGB stars that this method uses to find the distance is that they are present throughout all regions of the galaxy. They exhibit higher densities in the central regions of the galaxies as do all populations, but the rate at which their numbers fall off with respect to radius is not high as other populations. Because of this, rejecting whole regions like the center of the galaxy will not affect the RGB population as much as the luminous blue stars in the galaxy, for instance. In some cases, looking at which stars are rejected can even help identify locations where the RGB (and TRGB) are not, in CMDs where the position may be slightly ambiguous.

With the spatial image selection completed, the program brings up three separate CMDs in a new window. The first CMD (Figure 2(a)) plots the information for all the stars present in the photometry so that the user can see all the information. The second CMD (Figure 2(b)) plots all of the stars which were chosen to be excluded from the image. Often times one can observe that a much larger percentage of blue stars were excluded than red stars. The third CMD (Figure 2(c)) displays the photometry from the stars that remain and will be used for the rest of the calculation. If the spatial selection was done correctly, the stars remaining should still represent a large population of red-giant stars that is more well-defined than previously.

With the spatial selection finished, the program prompts the user to make a selection in the third CMD (Figure 2(c)) to have the program search within for the TRGB. This selection is necessary for reasons that will become apparent in a moment. A good rule of thumb is that a region that stretches one magnitude (in the vertical direction) on either side of the estimate of the TRGB should be applied. Once this region is selected, the program makes its calculation.

The first graph it produces is the selected region from the Figure 2(c) in a zoomed in plot (Figure 2(d)). Next, it takes the newly pared down data set and creates a histogram of the data with a fixed bin size in magnitudes. This can be changed by changing the code for the program itself, but testing has found that a bin size of 0.05 magnitudes tends to work well for all of the data sets to which were available. This histogram is plotted as the second window in the bottom row of the program's output (Figure 2(e)). Next, a Sobel edge detection filter is applied to the newly created histogram. The Sobel filter convolves the data with a kernel of [-2, -1, 0, 1, 2]. This means that for each data point in the histogram, a metric is calculated that consists of multiplying the value of the point two before the current point by -2, then adding it to the next point which is multiplied by -1, and so on. Thus, areas of the graph which have higher positive slopes will generate higher numbers, since the points on the right

8

sides will be higher than the points on the left. Regions where there is little change from bin to bin in the histogram will generate values close to zero, since the negative and positive portions will be roughly equal. The results of the Sobel filter application are then plotted in the final window of the application (Figure 2(f)).

Because of the way the Sobel filter works, places in the graph where there is a large jump in population size will generated large values. This implies that the TRGB, which is located on the CMD where a large increase in the RGB population occurs, should generate the highest value in the resulting Sobel graph. Simply searching the output for the largest value and its index allows the program to quantify the TRGB and its location. Using this data, a line representing the TRGB in each of the six graphs at the appropriate location. It uses the distance modulus equation (discussed above) to convert the magnitude of the TRGB to a physical distance.

The final step that the program allows can only be performed when an image was loaded to accompany the data in the first steps of the program. If an image is present, the user is prompted to select regions from the first CMD (Figure 2(a)). If a selection is made, the stars whose photometry are encompassed in that regions are plotted in the image. This is not especially useful for the distance calculation, seeing as that has already been performed and written out. Instead, this allows a user to explore the relative locations of certain stellar populations. For instance, if a user wants to test whether the RGB stars really do exist all throughout the galaxy (including the halo), the user could select a region around the TRGB and see the locations of those stars plotted in the image. Another interesting selection is to select the blue stars, and see that, in fact, they do exist almost exclusively in the central portions of the galaxy. An example of this process can be seen in Figure 4. While this is not immediately useful for the distance calculation, it can be an important tool if the calculated distance seems incorrect. Looking at the locations of stars can inform a second run of the program, in which the user might make different spatial selections to exclude other stars, or include stars that were excluded the first time.

D. Error Estimation

The error estimates of findtrgb rely on the assumption that the peak the program finds actually correlates to the TRGB. This might not be true if the selection routines are performed poorly, in which case the actual distance might be very different. This scenario is usually easy to spot after the program has completed though, so the program can be run again to find a better selection, and therefore a better distance estimate. Assuming the selection routine is well-executed, the error estimates were made as follows. Assume that the actual TRGB is located somewhere within the kernel convolved with the data. Then, the program uses the central point of that maximum, but the actual TRGB could be as far away as two magnitude bins. Thus, the possible error is calculated by finding the maximum difference between the distance calculated using the central point and the end points of the kernel (two magnitude bins on either side).

E. TRGBTool: A comparison

Dr. Makarov at the University of Hawaii created a program called TRGBTool, which he used in a 2006 publication (Makarov et al. 2006) to probe a new method of TRGB detection. For a detailed description of his Maximum Likelihood method, The reader is referred to the paper, as it is mathematically quite complex.

TRGBTool also requires data that findtrgb did not need in the form of false star tests. These tests are a way of analyzing the completeness of the data set. In other words, how well is the flux of each identified star recovered. Because these tests are fairly time-intensive and the only resource for them was Dr. McQuinn, TRGBTool was used on only a few of the data sets, rather than requesting the false star tests for all the galaxies. The results for those galaxies are contained in the third column of Table 1. These values lack error estimates because the error was very sensitive to individual parameters and it is hard to trust any one run of the program to give an accurate measurement. In most cases, the error was on the order of 0.1 Mpc. Since another method of comparison was discovered, work on making the TRGBTool estimates accurate and complete was abandoned, but the values we had are included for completeness.

F. MATCH: Independent Testing

As a third check, Dr. McQuinn used a functionality of the same program she had been using to reduce our data to find a distance estimate. For a complete discussion of the MATCH code which was used, see Dolphin (2002) and Dolphin et al. (2003). In general, MATCH generates synthetic CMDs with certain characteristics, compares them to the actual CMD (from the data) and finds a best fit. For our purposes, we allowed MATCH to vary all of its parameters to solve for the distance. Dr. McQuinn performed all of these calculations independently, and they were then compared to results from findtrgb.

III. Results and Discussion

After running findtrgb on the data sets that which were available, the results obtained were generally encouraging. The results of the calculations can be viewed in Table 1. In all cases, it was possible to generate a distance estimate which looked like a correctly identified TRGB. In some cases, it did take more than one run of the program because adjustments had to be made to selections in the

CMD to exclude regions where a jump in the population that did not correspond to the tip threw off the resulting distance calculation. In almost all cases, the distance estimates from findtrgb and MATCH agreed to a high degree of accuracy. The only one where the two methods disagree greatly is AGC 749237, and in this case it is possible that the TRGB is not actually contained in the data set. The stars in the RGB may be so dim that they were not measured, and as such trying to find the TRGB in this data set may be impossible.

It is encouraging that almost all of the galaxies appear to be within 10 Mpc (or around that limit) as this means that their masses are sufficiently low that they will be useful for the survey. The similarity in the results for findtrgb and MATCH lends confidence to the correctness of findtrgb, which is important since it was untested before this. In addition, both MATCH and findtrgb agree fairly well with TRGBTool (for the measurements we have). There is a slight systematic offset, but this can be attributed to the fact that both MATCH and findtrgb use data that is corrected for intergalactic extinction according to the Schlegel et al. dust maps (Schlegel et al. 1998), while the TRGBTool results are not corrected for extinction.

IV. Summary and Conclusions

With these results in hand, SHIELD can move forward and explore these galaxies in detail. Having the distance allows us to calibrate measurements of total flux and mass. With more observations, we can begin to understand more about dwarf galaxies and their properties. In addition, if more galaxies are added as candidates for the survey, there is now a method in place for determining their distances in a systematic way. The next major step is to construct star formation histories (SFHs) of the galaxies: maps of stellar formation as a function of time to understand how these remarkably small galaxies have evolved to be the way they are. This work will be carried out by Dr. McQuinn in the future.

As a final note, given much more time on this project, there are a few improvements that could be made to make the findtrgb more robust and less user-dependent. In particular, it would be interesting to explore ways of making the star selection in the CMD (the area in which to search for the TRGB) be totally autonomous. The most likely way to approach this would be to use the artificial intelligence concept of machine learning: having the program take in many, many data sets with CMDs and correctly identified TRGBs so that it could search for the similar characteristics which make a good selection. The human eye can easily pick out the ascent of the RGB if it is fairly well-defined, but a program might be able to do a better job in less obvious situations if it could learn what the TRGB "looked like" by studying many examples in a mathematical way.

I would like to thank John Cannon for the opportunity to perform this research and for continued help with the project throughout the process. I would also like to thank Dr. McQuinn for her help with reducing the data and providing an independent distance measurement. Finally, I would like to thank Dr. Makarov for supplying me with his TRGBTool and also helping us get it running despite some cross-platform difficulties.

References

Bellazzini et al. 2001, The Astronomical Journal, **556**, 2

Dolphin 2002, Monthly Notices of the Royal Astronomical Society, **332**, 91

Dolphin et al. 2003, The Astronomical Journal, **126**, 187

Freedman et al. 1994, American Astronomical Society, **26**, 1473

Makarov et al. 2006, The Astronomical Journal, **132**, 6

Sandage, A. R. 1971, Nuclei of Galaxies, ed. D. J. K. O'Connel (Amsterdam: North Holland), 601

Table 1. Distance Estimates for SHIELD Galaxies

AGC #	Sobel Filter Estimate (Mpc)	TRGBtool Estimate (Mpc)	MATCH Estimate (MpC)
748778	6.2±0.29	6.61	6.4±0.1
112521	6.0±0.28	N/A	5.9±0.1
110482	N/A	N/A	8.0±0.1
111946	7.9±0.37	N/A	8.4±0.1
111977	5.6±0.26	N/A	5.4±0.1
111164	4.9±0.23	N/A	4.9±0.1
174585	7.6±0.36	N/A	7.7±0.1
174605	9.5±0.45	10.49	9.6±0.1
182595	8.3±0.39	9.31	8.6±0.1
731457	10.0±0.47	10.5	10.5±0.1
749237	7.2±0.34	10.3	18.1±0.6
749241	N/A	N/A	N/A

Table 1. The results of three separate methods of estimating the distances to the SHIELD galaxies. N/A signifies estimates that have not been made (not available). In almost all cases, the MATCH estimates agree to a high degree of accuracy with the Sobel filter (findtrgb) estimates.

Figure 1. An example of a CMD. Each point represents the measurements of a star. The colored sections represent higher number densities, with darker being more dense

Figure 2. An example of a successful completion of findtrgb. Panels are labeled in alphabetical order according to when they would be created chronologically. The first four panels (a-d) are CMDs. Panel (a) is the photometry for the entire data set. Panel (b) is the photometry for the stars removed by the spatial selection. Panel (c) is the photometry for the remaining stars after the spatial selection. Panel (d) is the data selected by the yellow box in (c), which has been zoomed in. Panel (e) shows the data in (d) in histogram form binned by magnitude. Panel (f) shows the result of the Sobel filter used on the histogram in (e).

Figure 3. An image of the galaxy used for the run in Figure 2 (left), and the same image with a spatial selection made in the image (right). Stars inside the red border have been excluded.

Figure 4. The same image from Figure 3, but with the positions of stars overlaid on top. The red points correspond to the left selection in the first panel of Figure 2, while the green points correspond to the right selection in the same panel.

Figure 1

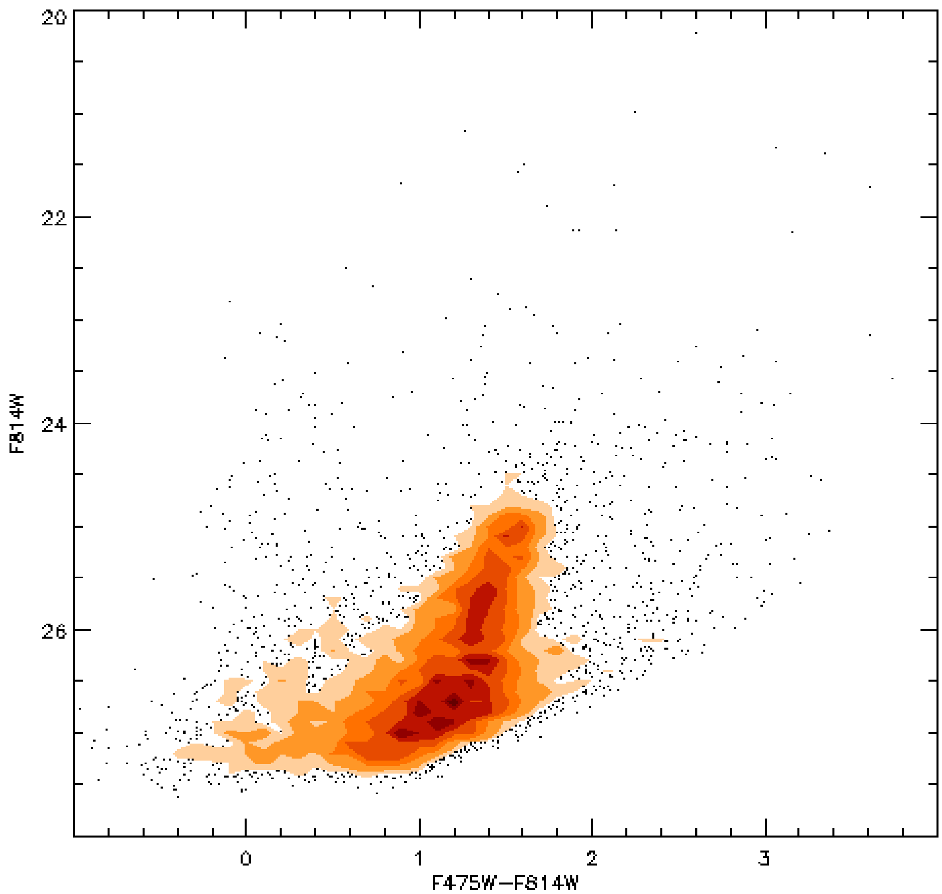
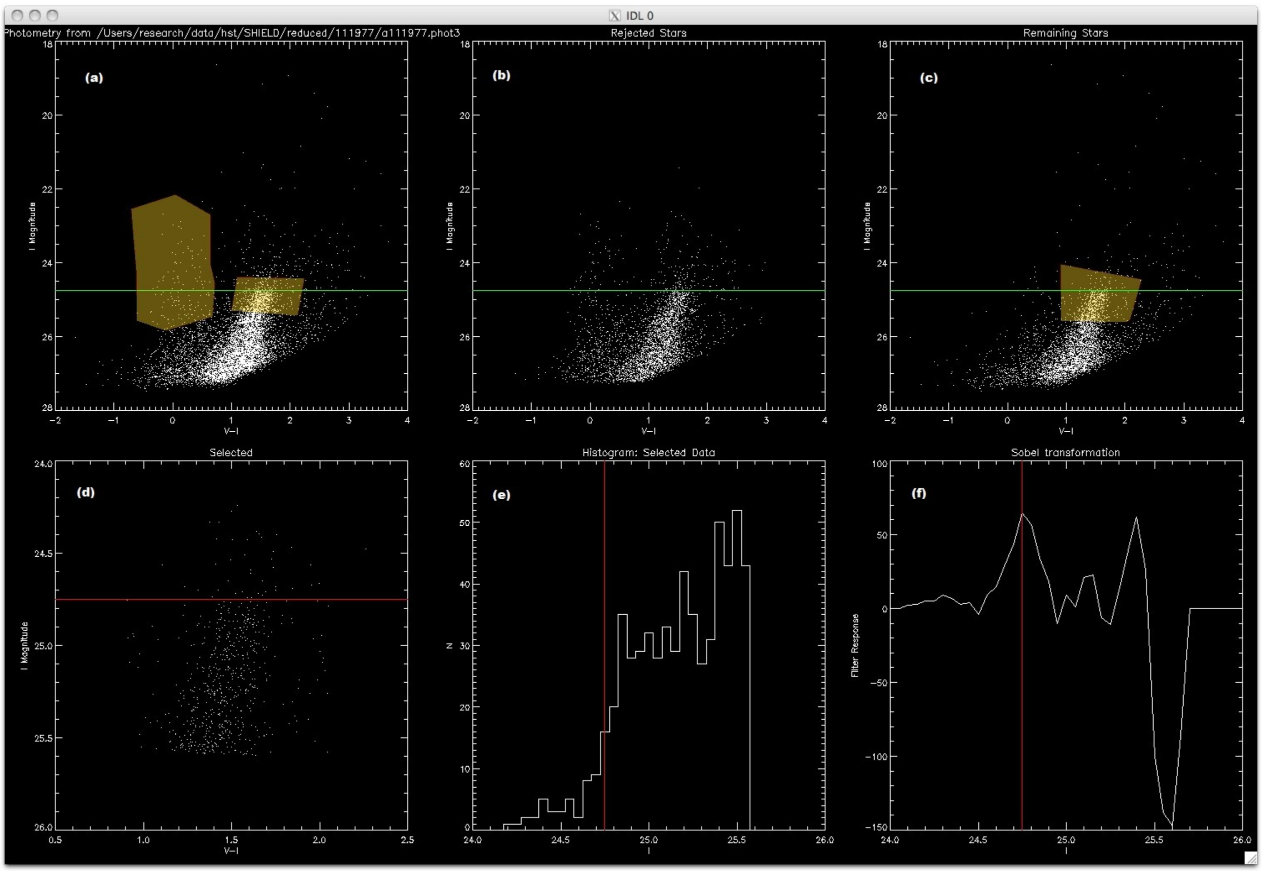
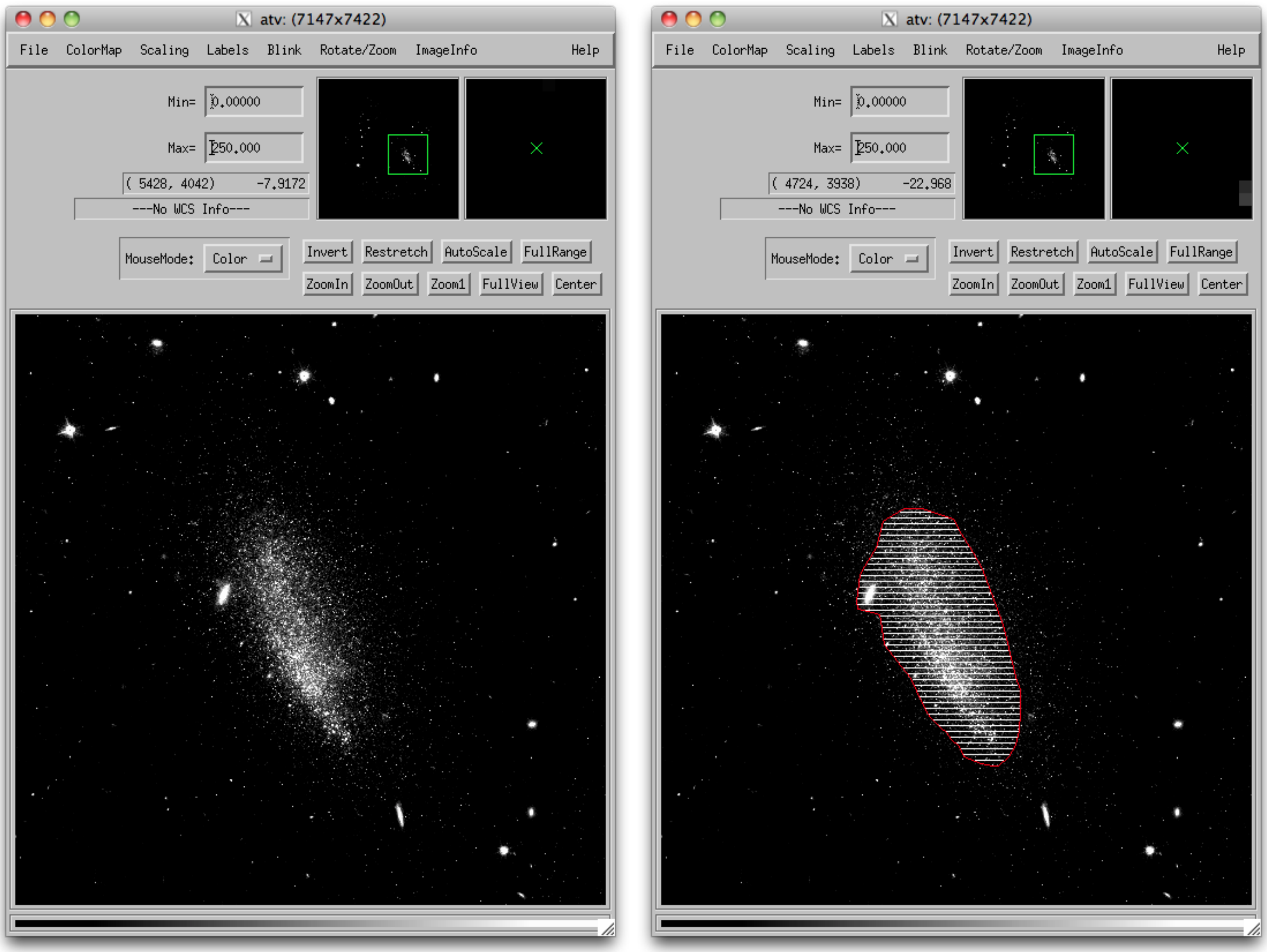


Figure 2


IDL 0
Photometry from /Users/research/data/hst/SHIELD/reduced/111977/a111977.phot3
Rejected Stars
Remaining Stars
(a)
(b)
(c)
I Magnitude
V−I
Selected
Histogram: Selected Data
Sobel transformation
(d)
(e)
(f)
N
I
Filter Response

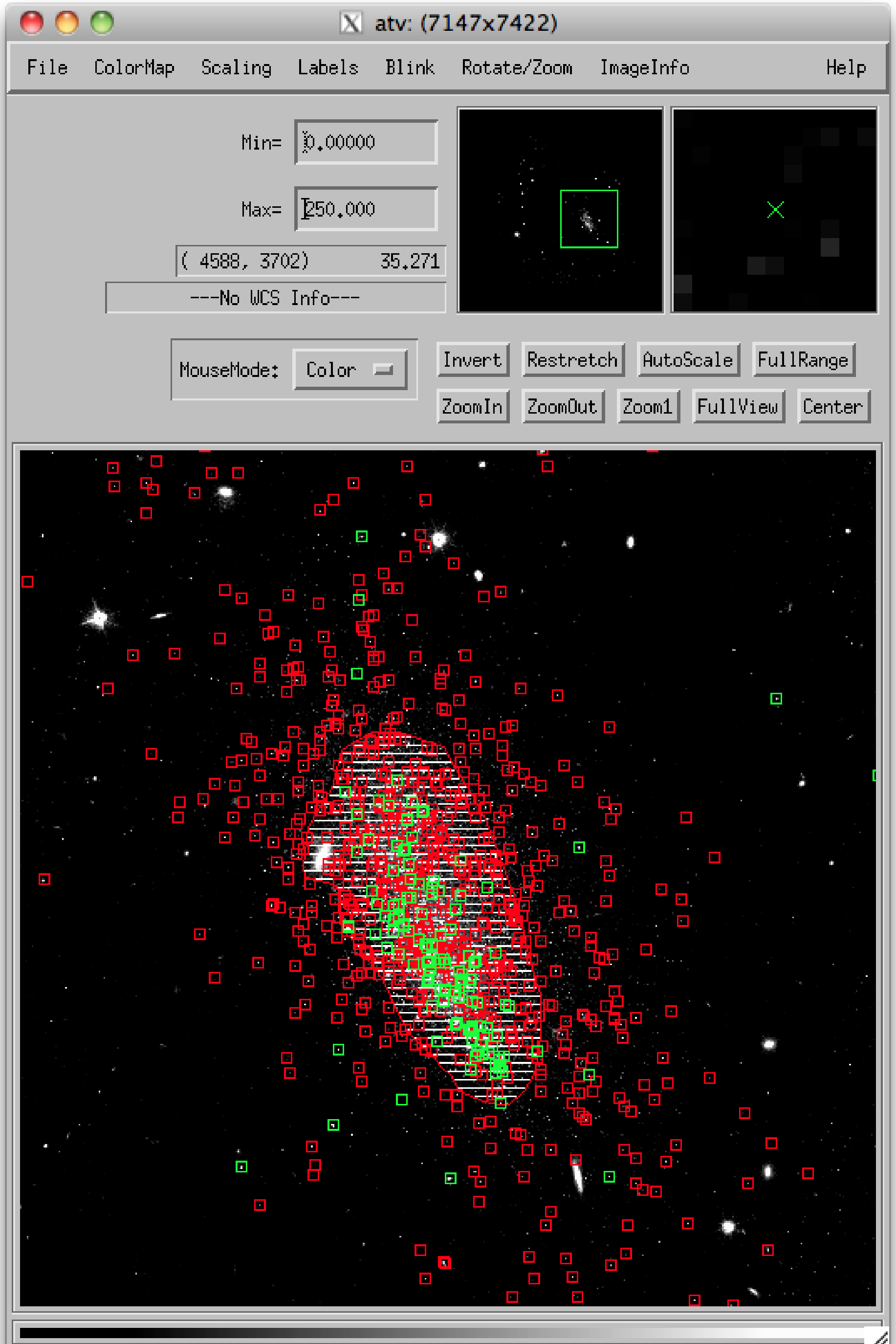
18

Figure 3


atv: (7147x7422)
File ColorMap Scaling Labels Blink Rotate/Zoom ImageInfo Help
Min= 0.00000
Max= 250.000
(5428, 4042) -7.9172
---No WCS Info---
MouseMode: Color
Invert Restretch AutoScale FullRange
ZoomIn ZoomOut Zoom1 FullView Center
atv: (7147x7422)
File ColorMap Scaling Labels Blink Rotate/Zoom ImageInfo Help
Min= 0.00000
Max= 250.000
(4724, 3938) -22.968
---No WCS Info---
MouseMode: Color
Invert Restretch AutoScale FullRange
ZoomIn ZoomOut Zoom1 FullView Center

19

Figure 4


atv: (7147x7422)
File ColorMap Scaling Labels Blink Rotate/Zoom ImageInfo Help
Min= 0.00000
Max= 250.000
(4588, 3702) 35.271
---No WCS Info---
MouseMode: Color
Invert Restretch AutoScale FullRange
ZoomIn ZoomOut Zoom1 FullView Center

20