COMPUTER IMAGE PROCESSING TECHNIQUES FOR THE PRODUCTION OF AAVSO OBSERVING CHARTS

Gilbert Weingarten
10 Windward Lane
City Island, NY 10464

Received: October 26, 1991

Abstract

This paper describes the computer image-analysis and chart-drawing programs used in the production of the new series of AAVSO observing charts made directly from astrophotographs as described by Scovil and Leitner (1991). These programs reduce the time required from tens of hours to under five minutes. It is in the areas of defect-per-chart elimination and object classification, including the proper identification of "double" stars, that modern techniques of computer image processing are most valuable.

1. Introduction

The production of observing charts directly from astrophotographs involves many steps as described previously by Scovil and Leitner (1991). The goal of this suite of programs, which form a part of that process, is to make the analysis of a photographic star field image and its subsequent conversion to an AAVSO Observing Chart as accurate and automatic as possible in order to minimize the amount of manual intervention required. These programs allow a chart image to be produced directly from an astrophotograph in under five minutes, using the AAVSO's IBM/PC compatible computer at Stamford Observatory. The chart must then be inspected visually and integrated into the final observing chart document with reference information. Production times compare very favorably to the tens of hours previously required to prepare the chart manually.

An observing chart must accurately relate the position and relative visual brightnesses of the stars in the field. This means that the production system must measure, preserve, and reproduce these quantities for stellar objects while identifying and eliminating from the final chart information concerning such non-stellar objects as photo-emulsion defects, nebulosity, and diffraction effects. Finally, the completed observing chart must display individual stars in their correct positions as round objects of sizes proportional to their visual intensity, and "resolve" images of multiple stars, displaying them as separate but overlapping objects although their photographic images merge into single elongated "blobs".

It is in the areas of defect elimination and object classification (including the proper identification of multiple stars) that modern techniques of computer image processing are most valuable. The programs that are currently in use divide the task into two parts, photographic image analysis (FINDSTAR) and chart drawing (MKCHART).

2. Image Acquisition

Astrophotographs of the proper scale for the various AAVSO charts are obtained as prints or negatives from various sources. They are scanned into an IBM-compatible PC as black and white images (i.e., two-level, binary representation) where they are stored in compressed form as PCX formatted files. The scanner operator must insure that the scanning threshold is set so that the captured image contains the range of stellar magnitudes required by the final chart. Typical scanner resolution is 300 dots per inch (DPI). A typical photograph measure 7.5 inches square and contains over 5 million picture elements (PEL).

Images can be scanned and stored as white objects on a black background (the positive image) or as black objects on a white background (the negative image). For the purposes of this paper, images will be treated as positives, with the objects of interest being white on a black background.

A region of interest (ROI) in the astrophotograph is defined as any arbitrarily shaped isolated area of white on the black background, no matter how complicated or convoluted its shape.

3. Photographic Image Analysis Overview

The processing of photographic images is performed by the FINDSTAR program. The job of FINDSTAR is to:

1. extract regions of interest (ROI) as a series of independent sub-images;
2. measure each ROI sub-image in order to extract the characteristics of the region;
3. use these extracted characteristics together with some "heuristics" (rules-of-thumb) to distinguish between stellar and non-stellar objects; and
4. produce a report, illustrated in Table 1, as a set of data abstractions of those objects deemed to be stellar so that MKCHART can use this information to produce a printed observing chart. MKCHART uses the computed value for X (column) and Y (row) position of a ROI to place a star on the chart. The radius value, computed from the mass, defines the size of the circle drawn to depict the star on the chart.

Table 1. FINDSTAR Report

```
// file => data\tstar2.pcx <= V2.5 => Thu Sept 19 21:57:29 1991
CHART = 360, 383, 1
END:
// #### =     X  ,     Y  ,   Mass,  Hgt,  Wdth,  Radius,  Perim,  c,
//      =   252.1,   107.4,    432,   31,    31,     0.0,    202,  0,
//      =   332.7,   126.7,    576,   31,    35,     0.0,    202,  1,
//      =   302.0,   304.0,      1,    1,     1,     0.0,      0,  0,
      1 =   302.8,    22.2,     25,    6,     5,     2.8,     16,  0
      2 =   301.9,    34.1,     35,    7,     7,     3.3,     22,  0
      3 =   344.2,    36.3,     63,    9,    10,     4.5,     32,  0
      4 =    72.0,    52.3,    782,   31,    32,    15.8,    119,  0
      5 =   289.9,    62.6,     55,    9,     8,     4.2,     28,  0
      6 =   165.4,    72.8,    159,   14,    15,     7.1,     52,  0
```

The X and Y table entries represent the column and row positions, respectively, of the centroid of a ROI. The "Mass" entry is the measure of the area of the ROI, in square pixels. "Hgt" and "Wdth" represent the height and width, respectively, of a rectangle just enclosing the ROI, measured in pixels. "Radius" is computed from the "Mass" and represents the radius of a circle whose area equals the "Mass". "Perim" is the measured perimeter of the ROI, measured in pixels.

4. Image Processing

Image processing is the manipulation of an image to produce one or more images from the original: thus, "an image in and one or more images out."

The FINDSTAR program identifies isolated features and separates an image into ROI by inspecting the image, PEL by PEL from top to bottom, left to right. After all ROI have been discovered, the original image can be discarded. All subsequent work is performed using the internal description of the sub-image of individual ROI.

A ROI is developed by looking for groups of white pixels in each raster line of the image. A run of consecutive white pixels constitutes a "segment" and is described by its offset from the beginning of the raster line and by its length. A given line can contain zero or more segments. Each segment from a line is examined to see if it overlaps any segment from the previous line. If they overlap, they are grouped together in a list, and become a part of the same "region". After all raster lines have been examined, all segments will have been assigned to one and only one region. All remaining regions are isolated from one another. The regions remaining at the end are the ROI.

5. Image Analysis

Image analysis is the process of inspection and manipulation of the image contents so as to produce some abstracted description of that image: thus, an image and descriptive information goes in and the "essence" goes out. The image may be discarded.

5.1. ROI Measurement

Each ROI is measured to determine:

1. its area or Mass (M), which relates indirectly to its photographic brightness;
2. the size (height (H) and width (W)) of the "bounding box" (i.e., the rectangular box) that just contains the ROI;
3. the position of its geometric centroid or center of mass;
4. the length of its perimeter.

These measurements of the two-dimensional properties of the image will be used, during further image analysis, to "guess" which ROI represent stellar objects and which are non-stellar.

5.2. Stellar Discrimination

A number of tests of the measured data for each ROI are used to distinguish stellar images from images of non-stellar objects. Most tests are "ratio" tests which make their criteria independent of the size of the ROI. All test criteria are under the control of the user of the program. The tests are performed in the order of increasing computational difficulty. Often more than one test will catch defects of a particular type. The tests performed include:

1. Object Size: The Mass (M) of a ROI is compared to an upper and lower limit. Failure of the mass to be within these limits will result in the ROI being discarded as being either "too small" or "too big". The size limits chosen should be dependent upon the scanner resolution (doubling the resolution will quadruple the mass of a ROI). This test allows the program to reject false images due to dust and emulsion grain as well as faint stars or huge objects like a comet or planet in the field.
2. Aspect Ratio: The ratio of the height to width (H/W) of the bounding box

containing the ROI is computed. This is a measure of the Aspect Ratio (AR) of the ROI. If the AR (H/W for vertically oriented objects) or its inverse (W/H for horizontally oriented objects) is greater then some value, the ROI is rejected. This test allows the program to reject ROI that are too elongated, such as those shown in Figure 1.

3. Mass Ratio: The ratio of the area of the bounding box to the Mass (M) of the ROI is computed and is termed the Mass Ratio (MR). It is a measure of how completely the ROI fills its bounding box. If the MR is too large, the ROI is discarded. The MR for an ideal circular object is 1.27; for an ideal "double" star ranges are between 1.27 and 2.55.

This test is useful for detecting on photographs defects which are "wispy" in appearance, such as nebulosity, emulsion reticulation, and lint.

4. Asymmetry: The Euclidian distance (S) between the centroid of the ROI and the center of its bounding box is computed. The ratio of S to the smaller of the height or width of the bounding box is used as the measure of asymmetry. If this number is too great the ROI is not symmetrically shaped (i.e., off-center with respect to the box that contains it), a non-stellar characteristic, and so is discarded. This test is also useful in detecting "wispy" or oddly shaped objects. However, care needs to be exercised in its use since some valid objects, such as a "triple" star, are inherently asymmetric.

5. Irregular Shape: The ratio of the perimeter of the bounding box to the perimeter of the ROI is computed. Perimeter Ratios (PR) are normally within the limited range of 0.5 to 1.2 for (one or two) circular objects. ROI with PR outside the this range are discarded. The test is valuable in detecting convoluted shapes even if they are symmetric with acceptable values for MR and AR, as illustrated in Figure 1.

ROI which have passed all of the tests described above are considered "stellar". Some of these may, however, represent images of "multiple" stars, as shown in Figure 2. These candidates will undergo further analysis to see if they can be "resolved" into their component stars.

5.3 Resolution of Double Stars

All accepted ROI that have mass ratios that are greater than a selected amount, the MASS_THRESHOLD, are assumed to be large enough to contain sufficient detail to allow shape analysis. Consequently, ROI with masses less than MASS_THRESHOLD square PELs are not tested. The remaining ROI undergo a complex shape analysis, discussed below, in an attempt to resolve them into their component stars. If an ROI is resolved into multiple component stars, these components are reported separately and replace the original ROI, which is discarded.

If an ROI can be resolved into one or more stellar objects, different portions of its shape should exhibit circularity with different local centers and radii. FINDSTAR inspects the shape of an ROI in order to determine the smallest number of points inside its perimeter that can form the centers of circles that can be used to replace the ROI with minimum loss of Mass. A Hough Transform (Gonzalez and Wintz 1987) of the ROI is used to resolve the components, as follows.

A number of points (usually eight or more and set by STD_POINTS) are chosen around the perimeter of an ROI, as shown in Figure 4. Consecutive sets of three points on the perimeter of the ROI are used. The program treats these as point sets residing on the circumference of some circle, and it computes the center and radius of that circle. A new image is generated that is a histogram of the centers of those computed circles that fall within the perimeter of the ROI. The histogram image is smoothed using a digital low-pass filter, as shown in Figure 5. A "hill climbing" routine finds local maxima in this image, which represent the common center of groups of the fitted circles. The average of the radii making up each group with a common center becomes the radius of one of the "resolved" stars.

Thus, the set of "resolved" stars makes up an aggregate that will replace the original ROI. Before this happens, the members of the aggregate undergo a normalization step which adjusts the mass of the aggregate and the centroid of the aggregate to match that of the original ROI. An example of the "resolution" of the candidate stars shown in Figure 2 may be seen in Figure 3.

6. Stellar Object Reporting

The output file that FINDSTAR builds contains lines of human-readable text, each describing one stellar object found by the program. Descriptions of the ROI that were rejected may be placed in the file and marked as comments. This file will be read and processed by MKCHART to construct the observing chart.

The information reported about each ROI that is accepted is:

1. the column and row number, in PELs, corresponding to the position, in the original image, of its centroid;
2. its Mass (M), in square PELs;
3. the height (H) and width (W), in PELs, of its bounding box;
4. the radius, in PELs, of a circle whose enclosed area would equal the mass of the ROI;
5. its perimeter, also in PELs.

A portion of this report file appears in Table 1.

7. Chart Production

The task of observing-chart production is performed by MKCHART. MKCHART reads the report file produced by FINDSTAR and uses it to draw filled circles at the proper location and size them to represent the stars which are to appear on the chart. The chart produced by MKCHART is not printed directly. Instead, the program stores the image of the chart as a compressed PCX formatted file for later use in the production process.

Users of MKCHART have considerable flexibility in determining the way in which the program draws the observing chart. The image on the chart may be adjusted by:

1. Rotation with respect to the chart center. This will compensate for any rotation of the image in the original photograph or introduced during scanning.
2. Scaling the diameters of the stars according to some algorithm, for perception compensation and/or to remove "clutter" by reducing or removing stars that are too faint. This improves the readability of the chart to the observer.

8. Improvements and Future Directions

8.1 Gray Scale Image Processing

In later versions of FINDSTAR, images may optionally be scanned and processed in their gray scale representation. This will allow the use of photometric measurement techniques to further improve chart accuracy, for example by making the Mass the integrated brightness of the ROI instead of its area, or using brightness peaks within a ROI to help resolve multiple stars.

9. Conclusions

The techniques described above rely heavily on the correct choice of parameters for the heuristics that discriminate between stellar and non-stellar objects. The choice of these parameters is based in geometric theory but has been refined by testing against real astrophotographs.

10. Acknowledgements

The author wishes to acknowledge the enormous amount of help provided by Charles Scovil and Robert Leitner at the Stamford Observatory for their constructive criticism in the design of these programs, the testing of these programs with astrophotographic images from the collection of the Stamford Observatory, and in the integration of these programs into the system for the production of AAVSO Observing Charts.

References

Gonzalez, R. C., and Wintz, P. 1987, *Digital Image Processing,* 2nd Edition, Addison-Wesley Publishing Company, Inc., Reading, MA.
Scovil, C. E. and Leitner, R. A. 1991, *J. Amer. Assoc. Var. Star Obs.,* this issue.

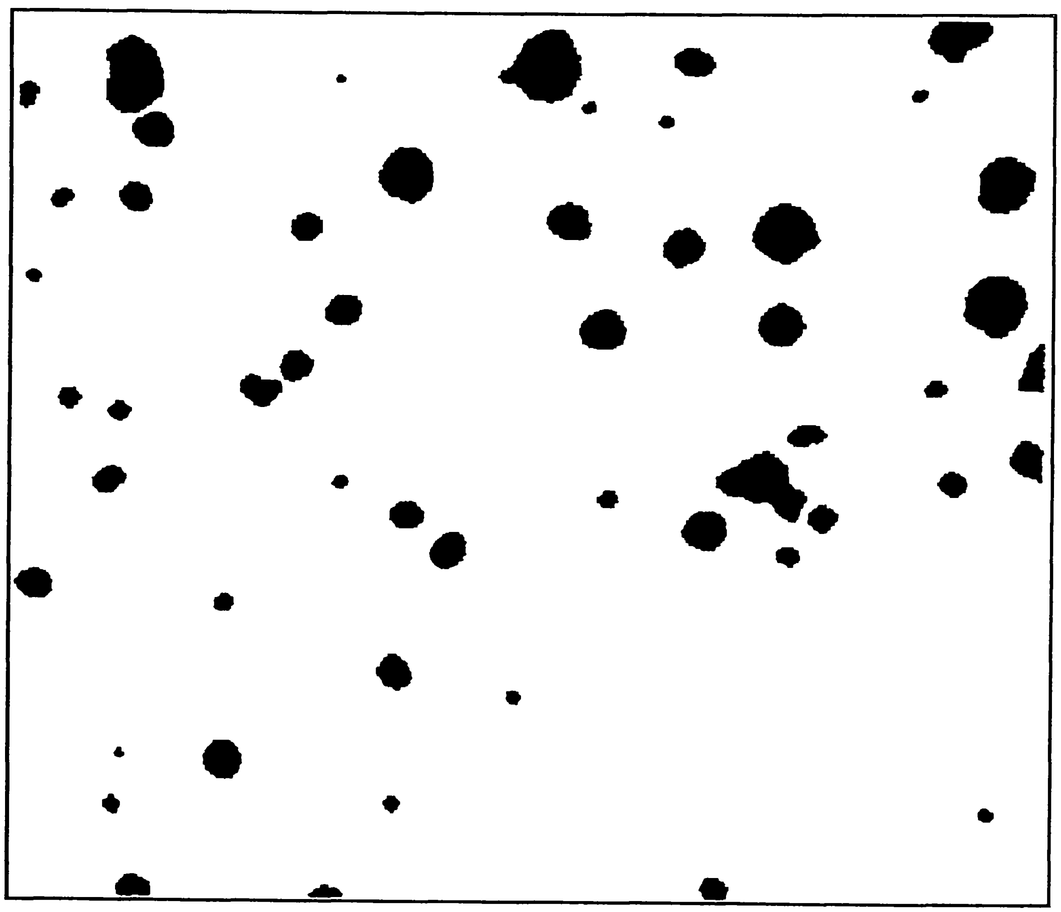

Figure 1. Some problem images.

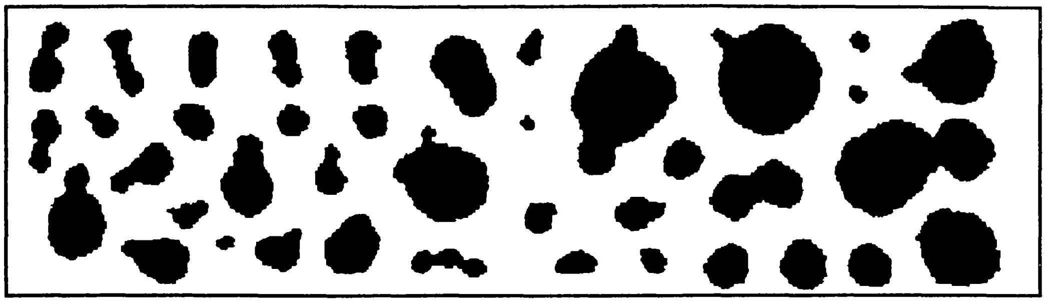

Figure 2. "Doubles" candidates.

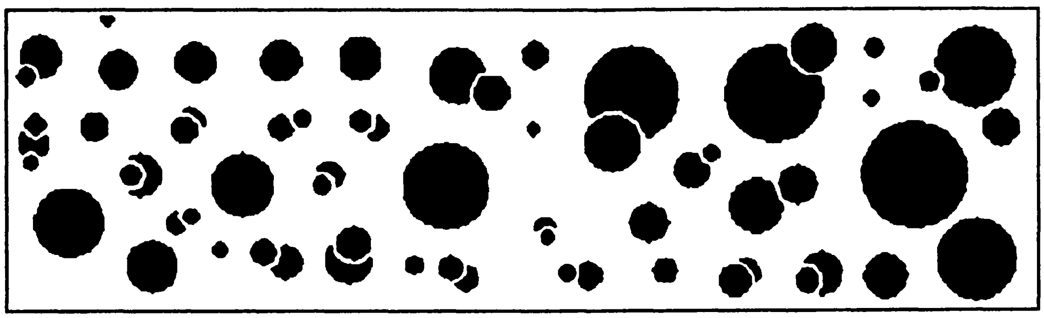

Figure 3. "Resolved" candidates.

```
  |------|-----|-----|-----|-----|-----|-----|
 0|..............*000..............
  |...*000.....000..*0.............
  |..00..*00..0*.....00............
  |.*0.....0*00.......*0...........
  |00..................0....000....
 5|0.......000*........0*000*.*0...
  |*......0*..000*000*00.......00..
  |00....00............0*0......*..
  |.0*000*...............00.....0..
  |.......................*.....0..
10|.......................000.0*0..
  |.........................*00....
  |------|-----|-----|-----|-----|-----|-----|
```

Figure 4. Perimeter image with "points" chosen (*).

```
  |-------------------|-----------------|-----------------|-
 0| .  .  .  .  .  .  .  .  .  1  1  .  .  .  .  .
  | .  .  .  1  2  1  .  .  1  5  5  1  .  .  .  .
  | .  1  1  3  6  3  .  .  2  8  8  1  .  .  .  .
  | 1  5  5  6 10  5  .  1  4 10 10  2  .  .  .  .
  | 1 10 12  9 12  6  1  5  7  9  8  1  .  .  .  .
 5| 2 13 17 11 10  5  1  8  9  7  5  1  .  .  .  .
  | 1 14 20 10  6  3  2 10 10  4  1  .  .  1  1  .
  | 1 11 18  8  2  1  1  8  8  2  .  .  1  5  5  1
  | .  7 12  6  .  .  1  5  5  1  .  .  1  8  8  1
  | .  3  6  3  .  .  .  1  1  .  .  .  2 10 10  2
10| .  1  2  1  .  .  .  .  .  .  .  .  1  8  8  1
  | .  .  .  .  .  .  .  .  .  .  .  .  1  5  5  1
  |-------------------|-----------------|-----------------|-
```

==> Centers <==

(x, y) = (19.3)
(x, y) = (8.4)
(x, y) = (4.6)
(x, y) = (15.6)
(x, y) = (27.9)

Figure 5. Filtered centers histogram.