

NOISE BASED DETECTION AND SEGMENTATION OF NEBULOUS SIGNAL

Mohammad Akhlaghi^{1,2} & Takashi Ichikawa¹

¹ Tohoku University Astronomical Institute.

² GNU Astronomy Utilities maintainer.

ADASS XXV

26 October 2015

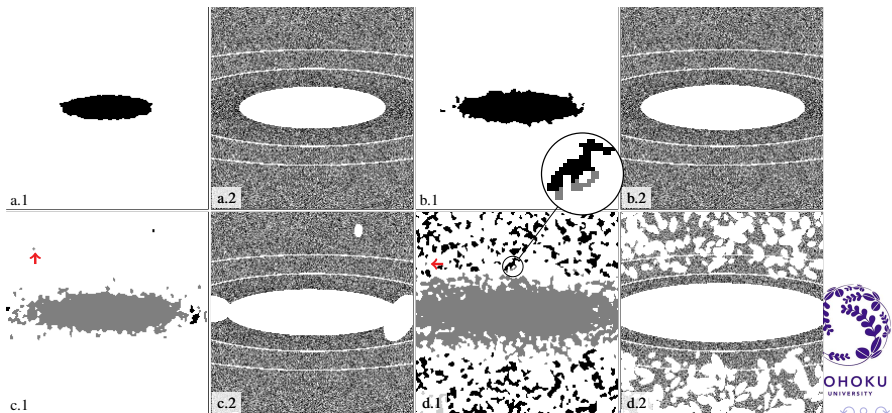
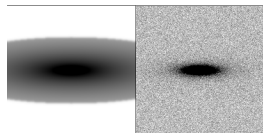


Growing and the Kron radius (r_k):

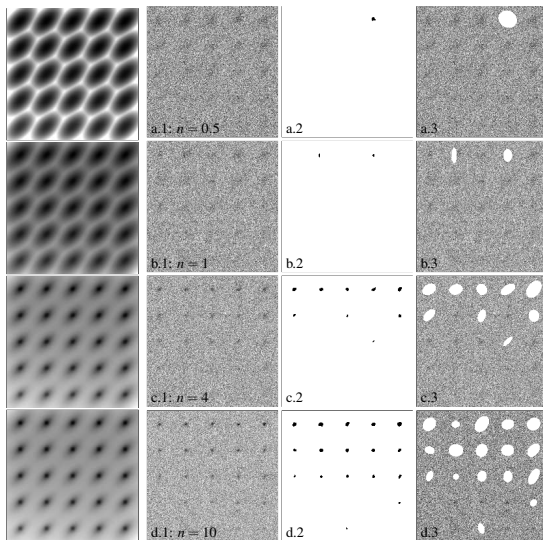
Kron (1980) radius: $r_k = \frac{\sum rI(r)}{\sum I(r)}$.

- Assumes an ellipse (very simplistic for real galaxies).
- Depends on the area above the threshold.
- Masked below: $3 \times r_k$ (hardly reaches 90% of flux).

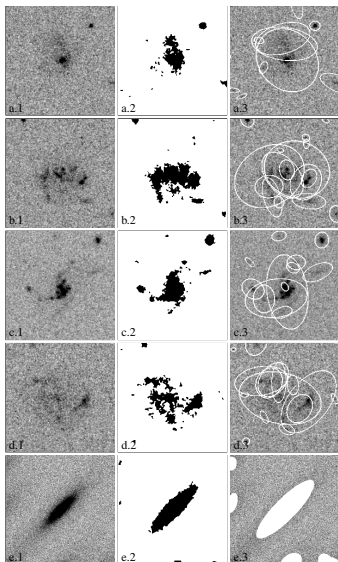
DETECT_THRESH is **2**, **1**, **0.5** and **0.1** respectively.



SExtractor: Sensitivity Test



SExtractor: Real



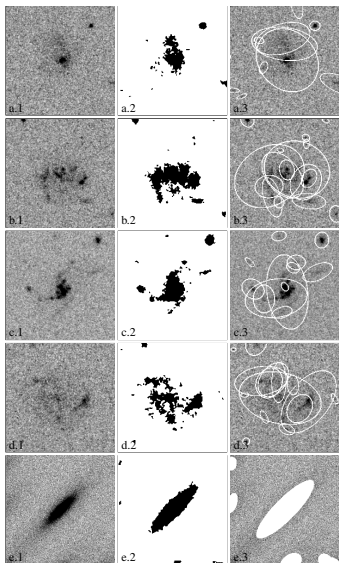
Real objects (images):

- Are not a clean ellipse.
 - Can be clumpy.
 - Can be diffuse.
 - Can have spiral arms.
 - Can be on the edge of the image.
- Do not necessarily have a uniform radial profile.

The existing method fails since such objects do not satisfy its *a priori* assumptions.



SExtractor: Real



Real objects (images):

- Are not a clean ellipse.
 - Can be clumpy.
 - Can be diffuse.
 - Can have spiral arms.
 - Can be on the edge of the image.
- Do not necessarily have a uniform radial profile.

The existing method fails since such objects do not satisfy its *a priori* assumptions.

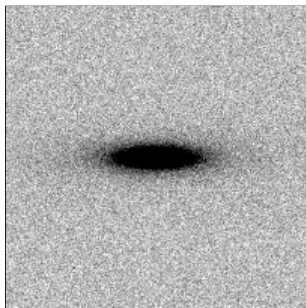
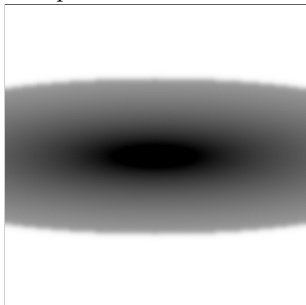
higher redshifts → more clumpy
(e.g., Murata et al. 2014)

NoiseChisel – Detection – Basics

Aims:

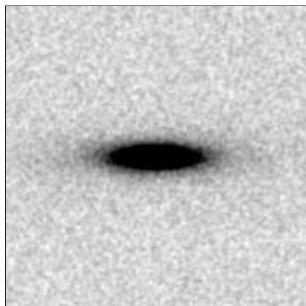
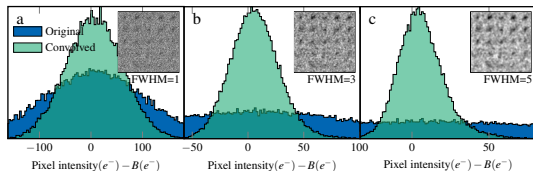
- Threshold must be independent of the Sky.
- Impose negligible assumptions on signal.
- Accurately remove false detections.
- Use the actual data, not *a priori* models.

Model profile for demonstration:



NoiseChisel – Detection – Convolution

- Convolution decreases dynamic range.
- **So:** Gaussian kernel, FWHM= 2pixels.

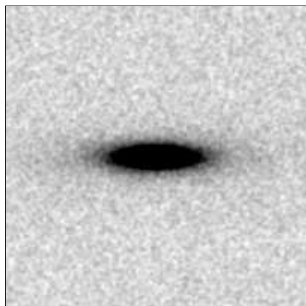
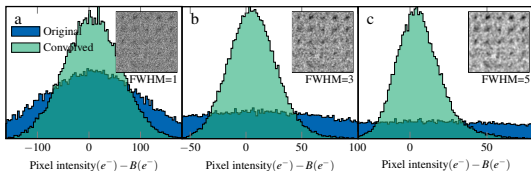


Since we are concerned with the sampling (noise) the PSF is no longer necessary and the **same parameters** work accurately on **space-based** and **ground-based** images.



NoiseChisel – Detection – Convolution

- Convolution decreases dynamic range.
- **So:** Gaussian kernel, FWHM= 2pixels.



Since we are concerned with the sampling (noise) the PSF is no longer necessary and the **same parameters** work accurately on **space-based** and **ground-based** images.

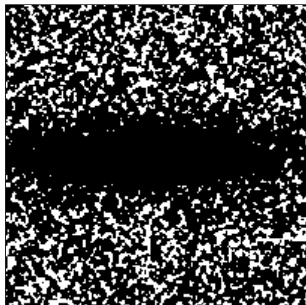
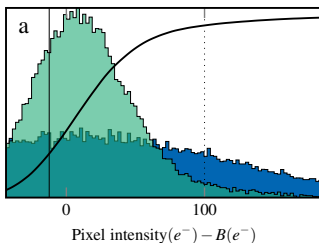


An assumption removed. Works on any image.



NoiseChisel – Detection – Threshold

- Use the cumulative pixel distribution.
- The threshold is set to the 0.3 quantile of the image.

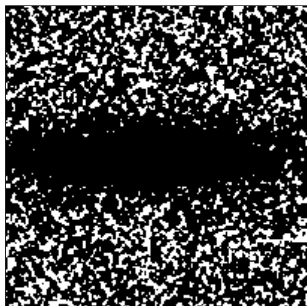
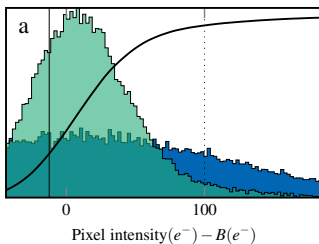


Since the threshold is now independent of Sky, we can accurately estimate the Sky once detection is complete.



NoiseChisel – Detection – Threshold

- Use the cumulative pixel distribution.
- The threshold is set to the 0.3 quantile of the image.



Since the threshold is now independent of Sky, we can accurately estimate the Sky once detection is complete.



Threshold no longer defined by Sky.



GNU



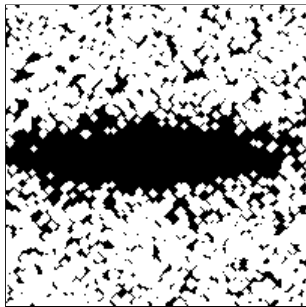
TOHOKU
UNIVERSITY

NoiseChisel – Detection – Erode

Erosion: Foreground becomes background if touching.

- **Or:** we expand the holes.
- **Or:** we **carve off** the signal.

NoiseChisel **name:** a tool to carve off noise



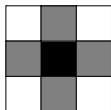
NoiseChisel – Detection – Open

Definitions:

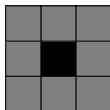
- **Dilation:** Inverse of erosion.
- **Opening:** Erosion followed by dilation.

In practice:

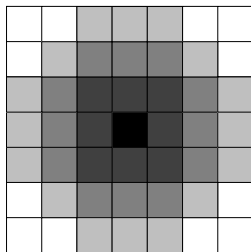
- Separates all the steps below.
- We use eight connectivity here (and four connectivity in the previous step.)



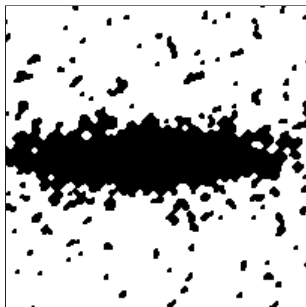
a



b

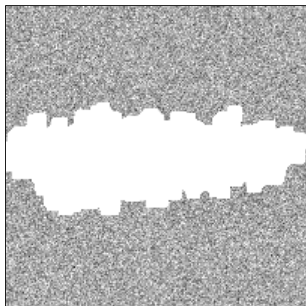
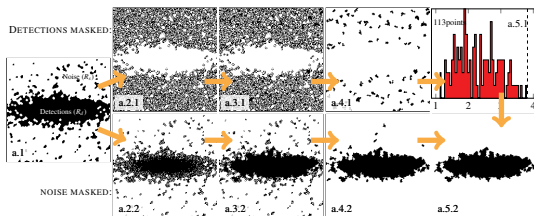


c



NoiseChisel – Detection – Remove false detections

- Use the ambient noise as a reference.
- The S/N of definite false detections is used:



False detections are successfully removed with high accuracy.

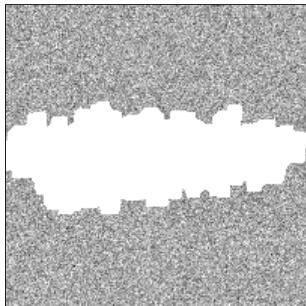
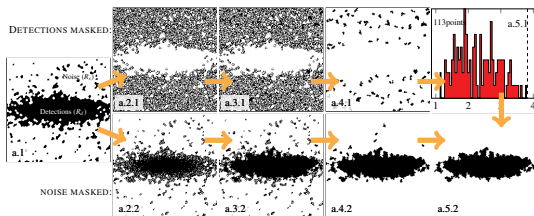


GNU

TOHOKU
UNIVERSITY

NoiseChisel – Detection – Remove false detections

- Use the ambient noise as a reference.
- The S/N of definite false detections is used:



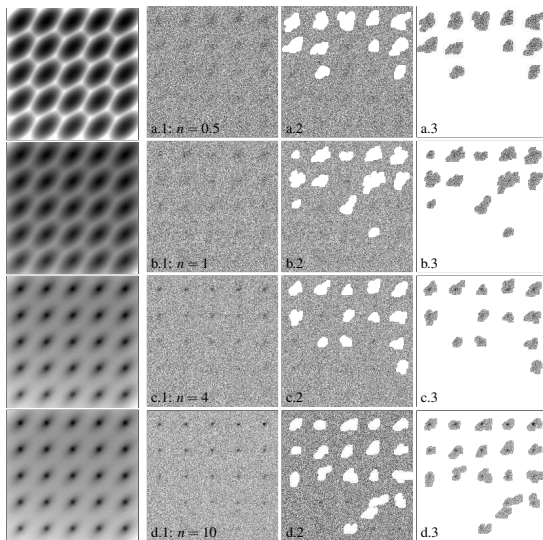
False detections are successfully removed with high accuracy.



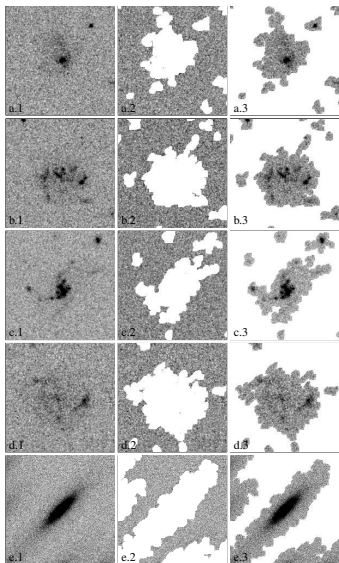
False detections are now identified for any image without hand-input values.



NoiseChisel – Detection – Sensitivity Test



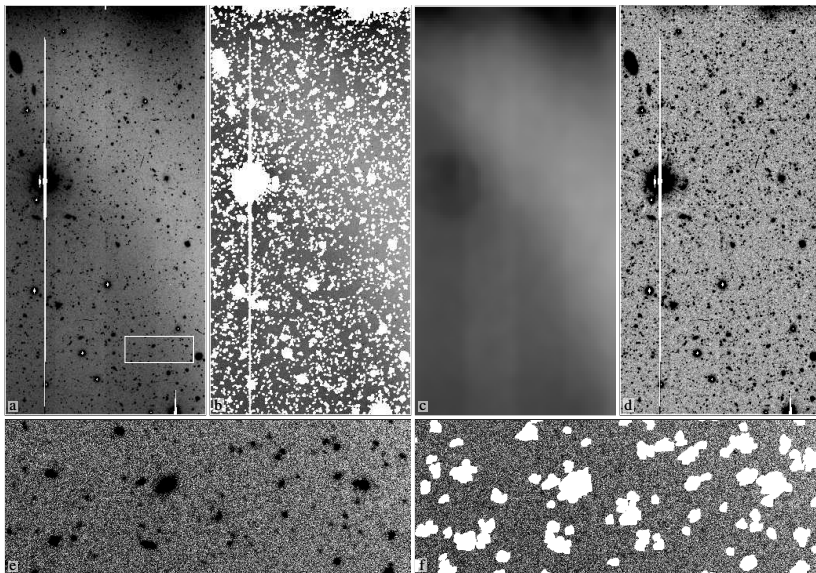
NoiseChisel – Detection – Real images



Noise-based detection: Works on any image with any target shape.



NoiseChisel – Detection – Large real images



NoiseChisel – Segmentation – clumps

A clump is found using the maximum resolution of the convolved image:



NoiseChisel – Segmentation – clumps

A clump is found using the maximum resolution of the convolved image:



No more layers

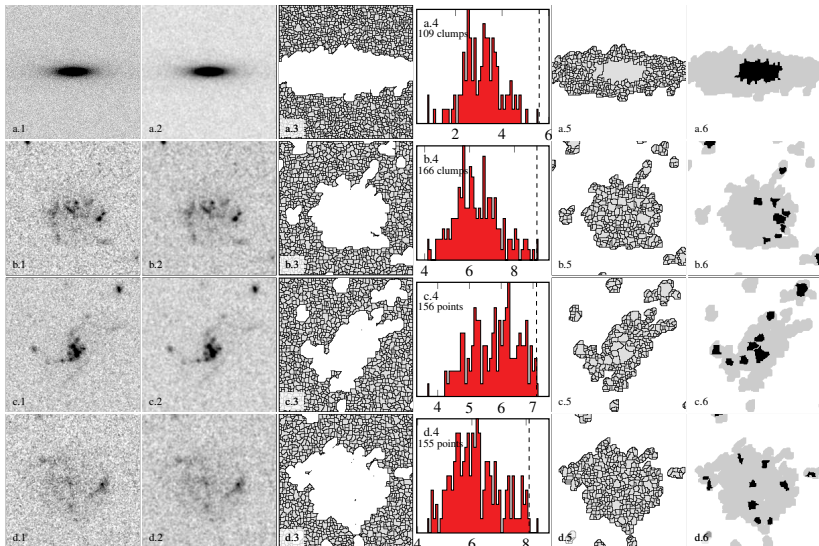


GNU

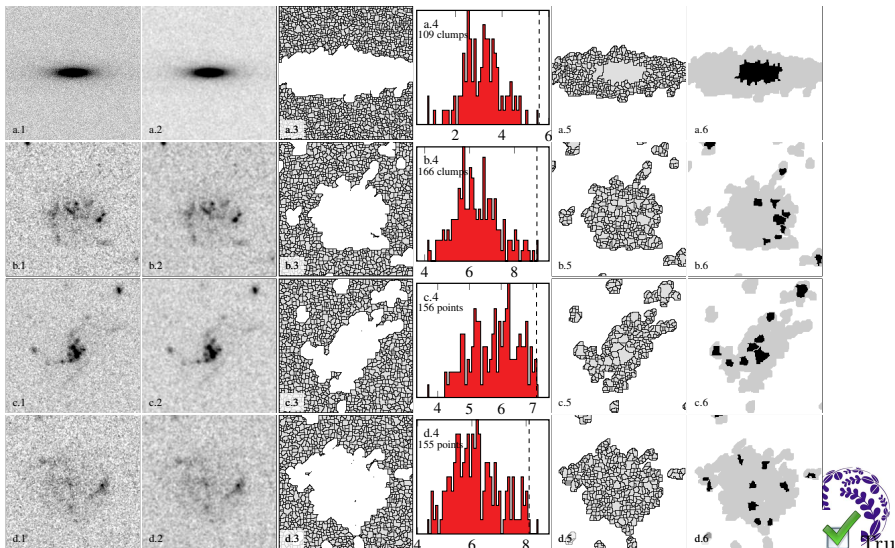


TOHOKU
UNIVERSITY

NoiseChisel – Segmentation – True clumps

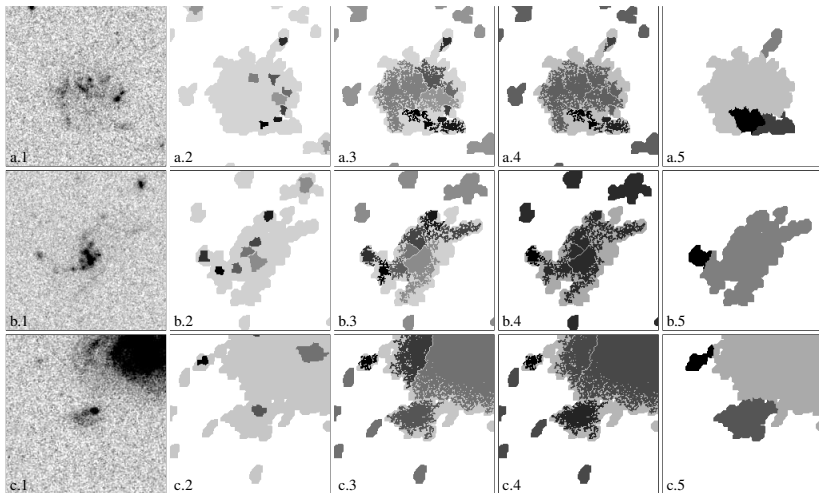


NoiseChisel – Segmentation – True clumps



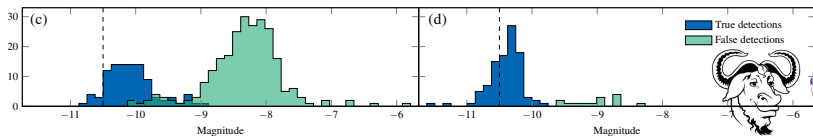
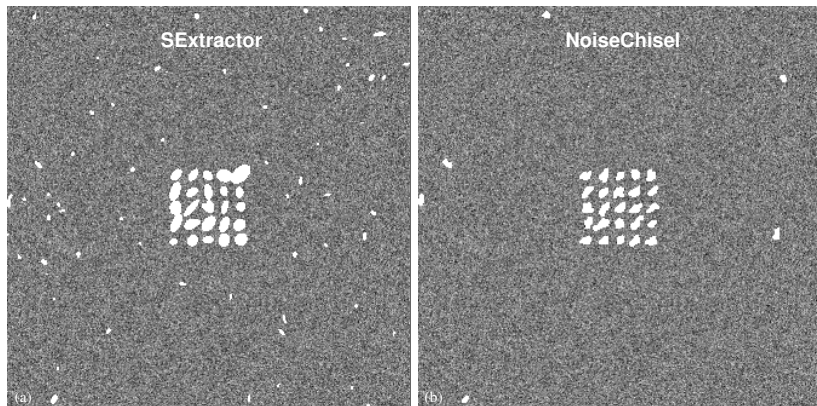
clumps are found independent of user input.

NoiseChisel – Segmentation – Objects



TOHOKU
UNIVERSITY

Purity and magnitude dispersion



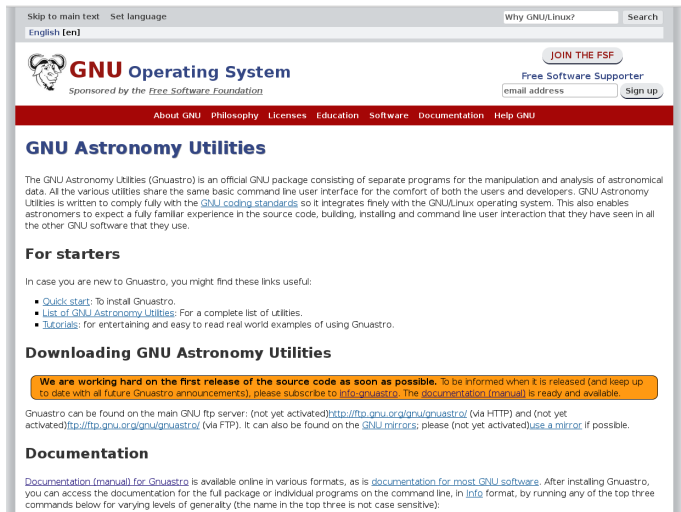
GNU Astronomy Utilities (Gnuastro)

NoiseChisel is part of the GNU Astronomy Utilities

- <https://www.gnu.org/software/gnuastro/>

Current Utilities

- ConvertType
- Convolve
- Header
- ImageCrop
- ImageStatistics
- ImageWarp
- MakeCatalog
- MakeNoise
- MakeProfiles
- NoiseChisel
- SubtractSky



The screenshot shows the GNU Astronomy Utilities website. At the top, there are navigation links for "Skip to main text" and "Set language", with "English [en]" selected. A search bar is present with the text "Why GNU/Linux?". Below this is the GNU logo and the text "GNU Operating System" with the tagline "Sponsored by the Free Software Foundation". A red banner contains navigation links: "About GNU", "Philosophy", "Licenses", "Education", "Software", "Documentation", and "Help GNU". The main heading is "GNU Astronomy Utilities". The main text describes the package as an official GNU package for astronomical data manipulation and analysis, noting its compliance with GNU coding standards. A "For starters" section provides links for "Quick start", "List of GNU Astronomy Utilities", and "Tutorials". A yellow box highlights that the source code is being released. The "Documentation" section mentions online manuals and command-line options.



Gnuastro's manual

GNU Astronomy Utilities

Astronomical data manipulation and analysis
for version 0.1, 7 October 2015

Short Contents

1	Introduction	1
2	Tutorials	11
3	Installation	21
4	Common behavior	31
5	Files	48
6	Image manipulation	58
7	Image analysis	104
8	Modeling and fitting	129
9	Table manipulation	147
10	Developing	148
A	GNU Astronomy Utilities list	162
B	Other useful software	163
C	GNU Free Documentation License	167
	Index	175

Mohammad Akhlaghi



Gnuastro's manual

Table of Contents

1	Introduction	1
1.1	Quick start	1
1.2	Science and its tools	2
1.3	Your rights	3
1.4	Naming convention	4
1.5	Version numbering	4
1.5.1	GNU Astronomy Utilities 1.0	5
1.6	New to GNU/Linux?	5
1.7	Report a bug	7
1.8	Suggest new feature	8
1.9	Announcements	9
1.10	Conventions	9
1.11	Acknowledgments	9
2	Tutorials	11
2.1	Builds cleanly checks and classifies file catalog	11
2.2	Saf simulates a detection	14
3	Installation	21
3.1	Requirements	21
3.1.1	GNU Scientific library	21
3.1.2	CFITSIO	22
3.1.3	WCNLIB	22
3.2	Optional requirements	23
3.2.1	libjpeg	23
3.2.2	GPL Ghostscript	23
3.3	Installing GNU Astronomy Utilities	23
3.3.1	Configuring	24
3.3.1.1	GNU Astronomy Utilities configure options	24
3.3.1.2	Installation directory	25
3.3.1.3	Executable names	27
3.3.2	Tests	28
3.3.3	A4 print manual	29
3.3.4	Known issues	29
4	Common behavior	31
4.1	Command line	31
4.1.1	Arguments and options	31
4.1.2	Arguments	32
4.1.3	Options	32
4.1.4	Common options	34
4.1.4.1	Input/Output options	34

ii

4.1.4.2	Operating modes	35
4.2	Configuration files	37
4.2.1	Configuration file format	38
4.2.2	Configuration file precedence	38
4.2.3	Current directory and User wide	39
4.2.4	System wide	39
4.3	Threads in GNU Astronomy Utilities	40
4.3.1	A note on threads	40
4.3.2	How to run simultaneous operations	41
4.4	Final parameter values, reproduce previous results	42
4.5	Automatic output	43
4.6	Getting help	43
4.6.1	--usage	44
4.6.2	--help	44
4.6.3	Man pages	46
4.6.4	Info	46
4.6.5	help-gnuastro mailing list	47
4.7	Output loaders	47
5	Files	48
5.1	Header	48
5.1.1	Invoking Header	48
5.2	ConvertType	51
5.2.1	Recognized file types	51
5.2.2	Color	51
5.2.3	Invoking ConvertType	54
6	Image manipulation	58
6.1	ImageCrop	58
6.1.1	ImageCrop modes	58
6.1.2	Crop section syntax	60
6.1.3	Blank pixels	60
6.1.4	Invoking ImageCrop	61
6.1.4.1	ImageCrop options	61
6.1.4.2	ImageCrop output	65
6.2	Convolve	65
6.2.1	Spatial domain convolution	66
6.2.1.1	Convolution process	66
6.2.1.2	Edges in the spatial domain	67
6.2.2	Frequency domain and Fourier operations	68
6.2.2.1	Fourier series historical background	69
6.2.2.2	Circles and the complex plane	70
6.2.2.3	Fourier series	71
6.2.2.4	Fourier transform	73
6.2.2.5	Draw delta and comb	74
6.2.2.6	Convolution theorem	75
6.2.2.7	Sampling theorem	77
6.2.2.8	Discrete Fourier transform	79

iii



Gnuastro's manual

6.2.2.9	Fourier operators in two dimensions	80
6.2.2.10	Edges in the frequency domain	82
6.2.3	Spatial vs. Frequency domain	82
6.2.4	Convolution kernel	83
6.2.5	Invoking Convolve	84
6.3	ImageWarp	86
6.3.1	Warping basics	87
6.3.2	Merging multiple warpings	89
6.3.3	Resampling	90
6.3.4	Invoking ImageWarp	91
6.4	SubtractSky	93
6.4.1	Sky value	93
6.4.1.1	Finding the sky value	94
6.4.1.2	Sky value misconceptions	95
6.4.2	Tiling an image	95
6.4.2.1	Quantifying signal in a mesh	97
6.4.2.2	Grid interpolation and smoothing	98
6.4.2.3	Checking grid values	99
6.4.2.4	Mesh grid options	100
6.4.3	Mask image	102
6.4.4	Invoking SubtractSky	102
7	Image analysis	104
7.1	ImageStatistics	104
7.1.1	Histogram and Cumulative Frequency Plot	104
7.1.2	Sigma clipping	105
7.1.3	Mirror distribution	106
7.1.4	Invoking ImageStatistics	107
7.2	NoiseChisel	112
7.2.1	Invoking NoiseChisel	112
7.2.1.1	NoiseChisel options	113
7.2.1.2	NoiseChisel output	120
7.3	MakeCatalog	121
7.3.1	Depth and limiting magnitude	121
7.3.2	Invoking MakeCatalog	123
8	Modeling and fitting	129
8.1	MakeProfiles	129
8.1.1	Modeling basics	129
8.1.1.1	Defining an ellipse	129
8.1.1.2	Point Spread function	130
8.1.1.3	Stars	132
8.1.1.4	Galaxies	132
8.1.1.5	Sampling from a function	133
8.1.1.6	Oversampling	134
8.1.2	If convolving afterwards	134
8.1.3	Flux Brightness and magnitude	134
8.1.4	Profile magnitude	135

iv

8.1.5	Invoking MakeProfiles	136
8.1.5.1	MakeProfiles catalog	136
8.1.5.2	MakeProfiles options	137
8.1.5.3	MakeProfiles output	141
8.2	MakeNoise	142
8.2.1	Noise basics	142
8.2.1.1	Photon counting noise	142
8.2.1.2	Instrumental noise	143
8.2.1.3	Final noise seed value	144
8.2.1.4	Generating random numbers	144
8.2.2	Invoking MakeNoise	145
9	Table manipulation	147
10	Developing	148
10.1	Why C programming language?	148
10.2	Design philosophy	149
10.3	Gnuastro project webpage	150
10.4	Version controlled source	151
10.5	Internal libraries	153
10.6	Header files	153
10.7	Program source	154
10.7.1	Mandatory source code files	154
10.7.2	Coding conventions	156
10.7.3	Multi-threaded programming	158
10.7.4	Documentation	158
10.8	Test scripts	159
10.9	Building	160
10.10	After making changes	160

Appendix A	GNU Astronomy Utilities list	162
-------------------	-------------------------------------	------------

Appendix B	Other useful software	163
B.1	SAO ds9	163
B.1.1	Viewing multi-extension FITS images	163
B.2	PGPLOT	164

Appendix C	GNU Free Documentation License	167
-------------------	---------------------------------------	------------

Index		175
--------------	--	------------

Gnuastro's manual

give a negative, float or string value for a FITS image extension or column number. As another example, you might give a value larger than unity for an option that only accepts fractions (which are always less than unity and positive).

CAUTION: In specifying a file address, if you want to use the shells tilde expansion (~) to specify your home directory, leave at least one space between the option name and your value. For example use `-o ~/test`, `--output ~/test` or `--output=~/test`. Calling them with `-o~/test` or `--output~/test` will disable shell expansion.

CAUTION: If you forget to specify a value for an option which requires one, and that option is the last one, Gnuastro will warn you. But if it is in the middle of the command, it will take the text of the next option or argument as the value which can cause undefined behavior.

NOTE: All counting in Gnuastro starts from 0 not 1. So for example the first FITS image extension or column in a table are noted by 0, not 1. This is the standard in C and all languages that are based on it (for example C++, Java and Python).

4.1.4 Common options

To facilitate the job of the users and developers, all the programs in Gnuastro share some basic command line options for the same operations where they are relevant. The list of options is provided below. It is noteworthy that these similar options are hard-wired into the programming of all of Gnuastro programs using GNU C Library's argument parser merging ability.

For some programs, some of the options, might be irrelevant for example MaleProfiles creates FITS images based on a given catalog. Therefore no input images (and thus HDUs) are necessary for it. In such cases, the option is still listed and if a value is given for it, it is completely ignored.

4.1.4.1 Input/Output options

These options are to do with the input and outputs of the various programs.

-h
--hdu **[--STR]** The number or name of the desired Header Data Unit or HDU in the input FITS image or images. A FITS file can store multiple HDUs or extensions, each with either an image or a table or nothing at all (only a header). Note that counting of the extensions starts from 0 (zero), not 1 (one). When specifying the name, case is not important so `IMAGE`, `image` or `image` are equivalent.

A # is appended to the string you specify for the HDU¹ and the result is put in square brackets and appended to the FITS file name before calling CFITSIO to read the contents of the HDU for all the programs in Gnuastro. CFITSIO

¹ With the # character, CFITSIO will only read the desired HDU into your memory, not all the existing HDUs in the file file.

has many capabilities to help you find the extension you want, far beyond the simple extension number and name. See CFITSIO manual's "HDU Location Specifiers" section for a very complete explanation with several examples.

-o
--output **[--STR]** The name of the output file or directory. With this option the automatic output names explained in Section 4.5 [Automatic output], page 41 are ignored.

-D
--dontdelete
 By default, if the output file already exists, it will be silently replaced with the output of this run of all Gnuastro programs. By calling this option, if the output file already exists, the programs will warn you and abort.

-K
--keepinputdir
 In automatic output names, don't remove the directory information of the input file names. As explained in Section 4.5 [Automatic output], page 43, if no output name is specified, then the output name will be made in the existing directory based on your input. If you call this option, the directory information of the input will be kept and the output will be in the same directory as the input. Note that this is only relevant if you are running the program from another directory!

4.1.4.2 Operating modes

Another group of options that are common to all the programs in Gnuastro are those to do with the general operation of the programs. The explanation for those that are not only limited to Gnuastro but can be called in all GNU programs start with (GNU option).

-- **[GNU option]** Stop parsing the command line. This option can be useful in scripts or when using the shell history. Suppose you have a long list of options, and want to see if removing some of them (and using the default values) can give a better result. If the ones you want to remove are the last ones on the command line, you don't have to delete them, you can just add `--` before them and if you don't get what you want, you can remove the `--` and get the same initial result.

--usage **[GNU option]** Only print the options and arguments. This is very useful for when you know the what the options do, you have just forgot their names. See Section 4.6.1 [--usage], page 44.

-?
--help **[GNU option]** Print all options and an explanation. Adding this option will print all the options in their short and long formats, also displaying which ones need a value if they are called (with an = after the long format). A short explanation is also given for what the option is for. The program will quit immediately after the message is printed and will not do any form of processing. See Section 4.6.2 [--help], page 44.



Reproducible science

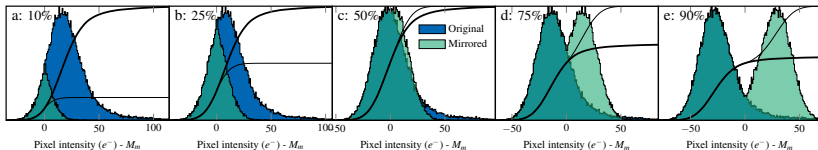
See source of arXiv:1505.01664 (this paper) and/or
<https://gitlab.com/makhlaghi/NoiseChisel-paper>.

- All the **scripts** and **configuration files** submitted to arXiv.
- Operation is managed by a **Makefile**.
- Data-generated numbers are **L^AT_EX variables** generated by the scripts.
- Input for all the plots are also generated by the scripts.
- Readers are encouraged to **check/modify the parameters** to see their effect.
- Everything fully explained in reproduce/README.



Finding the mode

- Data shifts the mode to the positive.
- The quantile is only comparable when the mode is.
- To find the mode, we define a “mirrored distribution” (green):



- The mode is where the cumulative frequency plot of the mirror distribution and original distribution are most similar:

