A Fast Astronomical Image Simulation Method and Analysis Postprint

Authors: Zhang Xin

Date: 2018-07-23T00:00:00+00:00

Abstract

Image simulation has increasingly played a vital role in astronomical research, with its importance primarily manifested in the following aspects: providing crucial basis for the evaluation of planned astronomical observation facilities, and validating data processing algorithms through the processing of simulated data. This paper describes the complete process of image simulation and computationally optimizes the local methods of image simulation to enable efficient and accurate generation of simulation results. The main work accomplished in this paper includes the following: simulation of celestial object morphology, including point sources (such as stars) and extended sources (galaxies), with the primary focus being on modeling extended sources; simulation of observation conditions, mainly noise generated by instruments and sky background; simulation of the Point Spread Function (PSF); optimization of the computational workflow from computational and programmatic perspectives; and analysis of the results. The data utilized in this paper are from the Hubble Ultra Deep Field (HUDF) ACS WFC i(F775) band, with a limiting magnitude reaching

Full Text

A Fast Method for Astronomical Image Simulation and Analysis

Zhang Xin[1,2] [1]Key Laboratory of Space Astronomy and Technology, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China

Abstract

Image simulation has become increasingly important in astronomical research, serving two primary functions: providing a crucial basis for evaluating future observational equipment and enabling verification of data processing algorithms through simulated datasets. This paper describes the complete image simulation pipeline and optimizes local simulation methods to achieve efficient and accurate generation of results. Our work encompasses four main aspects: (1) simulation of astronomical objects, including both point sources (stars) and extended sources (galaxies), with particular focus on galaxy modeling; (2) simulation of observational conditions, primarily instrument noise and stray light; (3) simulation of optical system characteristics manifested as the point spread function (PSF); and (4) optimization of computational workflows and algorithmic implementation. The data employed in this study are drawn from the Hubble Ultra Deep Field (HUDF) ACS WFC i-band (F775) observations, reaching a limiting magnitude of 29 AB mag. We extract galaxy parameters such as Sersic index, effective radius, and brightness from the HUDF catalog, simulate observational conditions including background noise and detector PSF, and perform statistical analysis on the extracted samples. Through optimization of the computational process, our method achieves rapid and accurate image simulation.

Keywords: Image Simulation; HUDF; Galaxy

1. Introduction

Astronomical image simulation, which combines image processing techniques with astronomical models, has found widespread application in modern astronomical research. Simulation serves a predictive function, enabling scientific evaluation of future observational facilities. Consequently, an increasing number of researchers are engaged in developing astronomical image simulation methods. Several well-known software packages for astronomical observation simulation exist, including WorldWide Telescope, Astronomy CCD Calculator, and Stellarium. While these tools provide detailed descriptions of celestial positions and object properties, offering considerable convenience for observers, their treatment of distant galaxies remains incomplete. With the advent of large-scale observational facilities, observations have extended to increasingly distant regions of the universe, making galaxy simulation an essential component in evaluating major astronomical instruments.

Bertin [1] introduced SkyMaker, which provides a detailed characterization of the point spread function incorporating atmospheric turbulence, telescope motion, aberrations, and diffraction effects, while employing the de Vaucouleurs profile for galaxy morphology. Refregier [2] et al. utilized shapelets for galaxy morphology simulation, employing a multi-dimensional approach that demonstrated reliability through residual analysis comparisons with wavelet methods. The Large Synoptic Survey Telescope (LSST) [3] has established a dedicated image simulation team focused on all-sky simulations. Their work employs cosmological models to generate catalogs determining object distributions and constructs models for various astronomical phenomena (including stars, galaxies, and asteroids), accounting for weather conditions and atmospheric turbulence during transmission, and finally incorporating PSF effects from the survey instrument. These accurate simulations enable comprehensive performance analysis of the LSST system. Notably, LSST employs Monte Carlo methods to simulate photon arrival for object morphology modeling. Joel Bergé [4] et al. discussed two distinct simulation approaches—photon injection methods similar to LSST and pixel-based simulation—and compared their differences, demonstrating that pixel-based methods offer significantly faster processing while maintaining adequate accuracy for many applications.

The simulations presented in this paper utilize Hubble Space Telescope data and observational conditions. Our image simulation work includes: (1) simulating object distributions and morphologies, primarily galaxy shapes; (2) simulating observational conditions and various noise sources, including sky background, dark current, and readout noise; (3) simulating instrumental PSF; and (4) performing statistical analysis on the simulated results. [Figure 1: see original paper] illustrates the image simulation workflow employed in this study.

The primary objective of this work is to achieve rapid astronomical image simulation using simple, general-purpose models while implementing extensive optimization to ensure practical computational performance. We further validate our approach by extracting sources from simulated images using SExtractor and comparing the results with reference catalogs to assess the practical feasibility of our simulation method. Section 1 introduces the key models and methods for image simulation, including galaxy morphology models, PSF models, and observational condition simulation. Section 2 presents analysis of the simulation results, and Section 3 provides a summary of the entire methodology.

---

2. Key Models and Methods for Image Simulation

The simulations utilize HUDF ACS WFC i-band (F775) data, which originally comprises approximately 10,000×10,000 pixels. Due to geometric rotation of the field, the image corners contain unobserved regions. For experimental convenience, we extract a central region of 4,500×4,500 pixels, covering an area of approximately 5.1 arcmin^2. Our models employ simple yet efficient astronomical models designed for rapid data processing. Observational conditions simulate those of the Hubble Space Telescope at an altitude of 600 km.

2.1 Galaxy Models We adopt the Sersic model [5] for galaxy morphology, with the mathematical representation:

$$I(R) = I_e \exp\left\{-b_n\left[\left(\frac{R}{R_e}\right)^{1/n} - 1\right]\right\}$$

where $I(R)$ denotes the surface brightness at radius R, I_e is the surface brightness at the effective radius, R_e represents the effective radius, n is the Sersic index controlling the concentration of the profile, and b_n is a constant.

Through transformation, Equation (1) can be expressed as:

$$I(R) = I_0 \exp\left[-\left(\frac{R}{a}\right)^{1/n}\right]$$

where I_0 denotes the central surface brightness, n is the Sersic index, a is a constant representing the radius at which the surface brightness drops to $1/e$ of its central value.

For circular galaxies, integrating Equation (2) over the circular area yields the total flux:

$$F(R) = 2\pi n a^2 I_0 \gamma(2n, (R/a)^{1/n})$$

where γ is the incomplete gamma function. As the integration radius approaches infinity, the total flux becomes:

$$F_{\text{total}} = 2\pi n a^2 I_0 \Gamma(2n)$$

where Γ is the complete gamma function.

In practice, galaxies exhibit elliptical shapes characterized by ellipticity. This requires distributing the surface brightness across an elliptical plane. Defining the ellipticity as e, the radius R in Equation (2) can be expressed using normalized elliptical coordinates:

$$R = \sqrt{(1-e)x^2 + \frac{y^2}{1-e}}$$

where (x, y) are coordinates with the origin at the galaxy center.

Galaxies also exhibit orientation angles. The elliptical Sersic distribution must undergo coordinate transformation to account for this. Assuming the galaxy coordinate system (with major axis as x-axis and minor axis as y-axis) is rotated by angle θ relative to the image coordinate system, the transformation matrix is:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

The new coordinates are then given by:

$$x' = x\cos\theta - y\sin\theta, \quad y' = x\sin\theta + y\cos\theta$$

Through Equations (1)-(6), we can simulate galaxy shape, size, and orientation. The required parameters—Sersic index n, effective radius R_e, and brightness—are extracted from HUDF ACS WFC data. In Coe D. et al. [6], SExtractor was used to extract HUDF data and compared with the catalog provided by Beckwith [7], yielding satisfactory agreement. All parameters used in our simulations are drawn from the catalog extracted by Coe D. et al.

2.2 Simulation of Observational Conditions and Noise Simulation of observational conditions essentially involves realistic reproduction of noise, which critically affects simulated image quality. Observational noise comprises inherent noise and signal-dependent noise.

Inherent noise originates from the detector itself, including thermal noise, shot noise, generation-recombination noise, and flicker noise, collectively represented as readout noise in this work. To enable comparison with Hubble data, readout noise is calculated using the Hubble Exposure Time Calculator [8].

Signal-dependent noise primarily includes background noise and random noise due to the quantum nature of photon arrival. Background noise utilizes HST-measured sky background values, including zodiacal and diffuse galactic light. The quantum nature of photon detection introduces stochastic fluctuations, which we model using Poisson statistics.

2.3 Point Spread Function Simulation For ground-based telescopes, atmospheric turbulence and instrumental optical characteristics dominate PSF formation. For space telescopes like Hubble, observational conditions are substantially superior, and PSF effects arise solely from the instrument' s optical properties.

Our space telescope simulations employ a Gaussian PSF:

$$\mathrm{PSF}(x, y) = I_0 \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}\right)$$

where I_0 is a constant representing central intensity, $[x, y]$ are image coordinates, and σ controls image quality. For Gaussian functions, σ can be derived from the given full width at half maximum (FWHM).

For ground-based telescopes, atmospheric turbulence effects must be included, requiring PSF models that combine instrumental and atmospheric contributions. The Moffat [9] model is commonly employed:

$$\mathrm{PSF}(x, y) = I_0 \left[1 + \frac{(x - x_0)^2 + (y - y_0)^2}{\alpha^2}\right]^{-\beta}$$

where I_0 denotes central intensity, α and β are scale parameters controlling profile width, and $[x, y]$ are image coordinates. The relationship between α and β is:

$$\mathrm{FWHM} = 2\alpha\sqrt{2^{1/\beta} - 1}$$

Reference [10] demonstrates that when $\beta = 4.765$, the Moffat PSF provides excellent approximation of atmospheric turbulence effects.

2.4 Sub-pixel Processing One simulation objective is evaluating future observational facilities. Our simulation data conform to Hubble ACS characteristics with pixel size 0.03×0.03 arcsec^2 (after drizzling). To enable evaluation of other instruments, sub-pixel level processing is required. The specific sub-pixel resolution depends on particular requirements.

Our sub-pixel processing is applied during galaxy morphology and brightness calculations. Each original pixel is subdivided into multiple sub-pixels that evenly divide the parent pixel. Galaxy properties are computed within these sub-pixels to determine the final pixel value. While straightforward for individual pixels, this approach becomes computationally intensive when applied globally, necessitating additional optimization strategies.

2.5 Parallel Computing Sub-pixel decomposition substantially increases computational load. However, each astronomical object is independent, enabling parallel processing. We distribute all objects in the simulation evenly across available CPU cores, with each core computing object morphologies and brightness independently. Results from different CPUs are then merged by the master process to form the final simulated image.

---

3. Simulation Results and Analysis

3.1 Reference Sample Analysis Our simulations employ data provided by Coe D. et al., which includes extracted Sersic indices. Beckwith et al. provide a more accurate catalog with superior positional and brightness measurements, though their morphological descriptions only include Gaussian parameters. Coe

D. et al. present comparisons with Beckwith' s data, including central position offsets. We have conducted more detailed comparisons between these two datasets.

[Figure 2: see original paper] compares the brightness distributions extracted from both catalogs, with the blue line representing Coe D. et al. data and the red line representing Beckwith et al. data. The figure reveals systematic brightness offsets, with Coe D. et al. data showing slightly brighter magnitudes than actual values.

[Figure 3: see original paper] compares effective radii between the two samples, plotting the difference in effective radius (Beckwith minus Coe) against object brightness. While most differences cluster near zero, Coe D. et al. data exhibit slightly larger radii for some objects. These discrepancies indicate that simulations based on Coe D. et al. will exhibit brightness offsets and size differences compared to real images, which helps explain the presence of some artificially large objects in our simulation results shown in [Figure 4: see original paper].

3.2 Simulation Results and Analysis [Figure 4: see original paper] presents our simulation results, with panel (a) showing the simulated image and panel (b) the drizzled original image. Due to field rotation and empty corners in the original, we extract a square region from pixel coordinates (3001,3001) to (7500,7500), yielding 4,500×4,500 pixels. After drizzling, each pixel subtends approximately 0.03×0.03 arcsec², giving a total field of view of 5.1 arcmin². The HST team provides a zeropoint magnitude of 25.654 for ACS WFC i-band (F775), where a flux of 1 e corresponds to this magnitude value. Object flux is calculated using:

$$F = 10^{-0.4(m-\mathrm{zp})}$$

where m is object magnitude and zp is the zeropoint. The calculated flux represents signal electrons. Noise components–including sky background, dark current, and readout noise–are computed using the Hubble Exposure Time Calculator, yielding values in electrons. All results are divided by the gain (e /DN) to obtain final image counts. The limiting magnitude of HUDF ACS WFC i-band is 29 mag, which we adopt as our extraction cutoff.

We employ SExtractor for source extraction, with primary parameters listed in Appendix 1. summarizes extraction statistics, with reference data from Coe D., Benitez N., et al. The table presents four metrics: (1) extraction statistics from the original image, (2) extraction statistics from the simulated image, (3) crude matching statistics based on positional proximity, and (4) matching rates calculated as:

$$\text{Matching Rate} = \frac{\text{Matched Sim Num}}{\text{Ref Num}}$$

where Ref Num is the number of reference catalog objects and Matched Sim Num is the number of matched simulated objects. Matching rates exceed 85% across all magnitude bins, which is acceptable given noise effects and extraction method differences.

[Figure 5: see original paper] compares the number counts for objects brighter than 29 mag, showing consistent trends between simulated and reference data across brightness ranges.

[Figure 6: see original paper] analyzes effective radii by comparing differences between simulated and reference objects. Panel (a) plots radius difference versus brightness, demonstrating that fainter objects show smaller radius discrepancies. Panel (b) shows the distribution of radius differences, which cluster around zero with some outliers. These discrepancies arise from multiple factors: catalog precision issues (discussed in Section 3.1), injected signal and background noise, instrumental noise, and SExtractor parameter differences.

[Figure 7: see original paper] compares magnitudes between simulated and reference objects using the same difference-versus-brightness approach. Most magnitude differences fall within ± 0.5 mag, with brighter objects showing better agreement. This indicates that our simulation and extraction pipeline produces more accurate photometry for luminous galaxies.

3.3 Performance Analysis Our image simulation software and statistical analysis tools are developed in Java. While less common in astronomy, Java offers several advantages: (1) elimination of pointer operations and automatic garbage collection significantly reduce memory leak issues, allowing greater focus on algorithm development; (2) built-in support for parallel computing with simple implementation, greatly facilitating our parallelization strategy.

The program runs on a Linux system with an Intel Core i5 4-core 3.1 GHz processor and 16 GB RAM. During execution, all four cores are utilized. Processing 4,107 objects requires 1,267 seconds, averaging 308 ms per object, with peak memory usage of 4.1 GB.

Most computational time and memory are consumed by Sersic model calculations. Due to sub-pixel requirements (Section 3.4) and precision needs, each pixel is decomposed into 32×32 sub-pixels, transforming the 4,500×4,500 pixel image into an effective 144,000×144,000 pixel grid. Despite this substantial data volume, parallelization and algorithmic optimizations achieve the reported 308 ms per object processing time.

---

4. Summary and Outlook

This work employs fundamental astronomical models—Sersic and Gaussian functions—to achieve simple and efficient image simulation. However, simplicity and efficiency are only valuable when results remain realistic. Analysis in Section 3 demonstrates that while differences exist between simulated and real data, fundamental characteristics (size, orientation, brightness) are correctly reproduced. Statistical analysis shows that errors in extracted object numbers, brightnesses, and sizes remain within controlled bounds. Furthermore, our parallel implementation, combined with optimizations such as trigonometric interpolation, achieves millisecond-level per-object computation, confirming the rapid nature of our simulation approach.

Our current PSF implementation uses only a simple Gaussian function, which introduces some limitations in reflecting true image characteristics. Additionally, our primary goal is to support evaluation of new survey projects, requiring further development. Future work will focus on: (1) improving PSF modeling by incorporating real optical system characteristics; (2) introducing spectral simulation capabilities for multi-band imaging; and (3) integrating simulations with specific survey projects to derive hardware-dependent performance analyses.

References

[1] Bertin E. SkyMaker: astronomical image simulations made easy[J]. Memorie della Societa Astronomica Italiana, 2009, 80:422.
[2] Refregier A. Shapelets - I. A method for image analysis[J]. Monthly Notice of the Royal Astronomical Society, 2003, 338(1): 35-47.
[3] LSST Science Collaborations, et al. LSST Science Book Version 2.0 [EB/OL].2009, arXiv:0912.0201.
[4] Bergé J., Gamper, L., et al. An Ultra Fast Image Generator (UFig) for wide-field astronomy[J]. A&C, 2013, 1:23-32.
[5] Sérsic J. L. Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy[J]. Boletin de la Asociacion Argentina de Astronomia, 1963, 6: 41.
[6] Coe D., Benitez N. et al. Galaxies in the Hubble Ultra Deep Field: I.Detection, Multiband Photometry, Photometric Redshifts, and Morphology[J]. The Astronomical Journal, 2006, 132(2): 926-959.
[7] Beckwith S. V. W., Stiavelli M., Koekemoer A. M., et al. The Hubble Ultra Deep Field[J]. The Astronomical Journal, 2006, 132(5):1729-1755.
[8] Diaz, D.; Greenfield, R. I.; McLean, P. The HST Exposure Time Calculators: Estimating Accurate Observing Times for HST Observations[C]. San Francisco: Astronomical Society of the Pacific, 2010.
[9] Moffat, A. F. J. A Theoretical Investigation of Focal Stellar Images in the Photographic Emulsion and Application to Photographic Photometry[J]. Astronomy and Astrophysics, 1969, 3:455.
[10] Trujillo, J., Gutierrez, I., Aguerri, PSF[J]. Monthly Notices of the Royal Astronomical Society, 2001, 328(3): 977-985.

Appendix 1. Primary SExtractor Extraction Parameters

```
DETECT_MINAREA   3            # minimum number of pixels above threshold
DETECT_THRESH    1.5          # <sigmas> or <threshold>,<ZP> in mag.arcsec-2
ANALYSIS_THRESH  1.5          # <sigmas> or <threshold>,<ZP> in mag.arcsec-2
FILTER_NAME      default.conv # name of the file containing the filter
DEBLEND_NTHRESH  32           # Number of deblending sub-thresholds
DEBLEND_MINCONT  0.005        # Minimum contrast parameter for deblending
```

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.