# SExtractor and Its Application in Deep Sky Survey Research (Postprint)

**Authors:** Ruan Zhifeng, Wu Butao

**Date:** 2017-10-20T00:00:00+00:00

## Abstract

SExtracto:, as an open-source software package for detecting celestial objects from survey images and extracting information such as magnitudes and positions, has been widely employed in the extraction and measurement of target sources in deep surveys. In deep surveys, the detection capability (depth) for target sources and the deblending capability for sources with overlapping profiles are often more important than measurement errors arising from photon noise. The extraction of target objects for different research purposes frequently requires balancing detection capability against deblending capability, and parameter settings for SExtracto: show considerable variation when extracting target sources. This work presents the parameter settings of SExtracto: in deep surveys, analyzes and discusses the accuracy of SExtracto: in source detection and measurement, and also identifies a very serious over-deblending problem when using SExtracto: to extract target sources in deep surveys.

## Full Text

## SExtractor and Its Applications in Deep Sky Surveys

**Ruan Zhifeng, Wu Wentao** Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China Email: ruanzhifeng@ynao.ac.cn

### Abstract

Deep sky surveys provide fundamental data for studies of galaxy formation and evolution, galaxy clusters, large-scale structures of the universe, and supermassive black holes. The SExtractor software package is an open-source software system. It can build a catalogue of objects from a sky-survey image and output magnitudes, surface coordinates, sizes, shape descriptions, and other information of objects. The SExtractor has been widely used for detection and measurement of objects in deep sky surveys. For object detection and measurement, abilities to detect objects (for completeness) and to separate objects of superimposed profiles (for resolving confusion) often have a larger impact on results than measurement errors due to photon noise. This requires certain balance between the abilities to detect objects and to separate objects of superimposed profiles. When using the SExtractor, the parameter values should then vary according to different types of objects so as to optimize detection and measurement. To detect apparently small faint objects, low detection thresholds need to be adopted, but low detection thresholds can lead to over-deblending of high-redshift galaxies of complex inner structures, or confusion of outer features of bright galaxies/scattered-light features of bright stars to be independent objects. Particularly, over-deblending can cause severe contamination in statistics of high-redshift close pairs identified from deep sky surveys. Confusion from parts of bright galaxies and scattered-light features of stars can be avoided by flexible settings of SExtractor parameter values. Over-deblending of faint galaxies of complex structures remains a difficult problem though. In this paper, we systematically describe the principles and parameters of the SExtractor. By incorporating our own analysis we also discuss accuracies and other measurement problems in using the SExtractor. At last, we give some suggestions to those problems.

**Keywords:** Object detection and measurement; Deep sky survey; Photometry

## 1. Principles and Parameter Settings of SExtractor

SExtractor is an open-source software package for detecting celestial objects from survey images and extracting information such as magnitudes and positions. It has been widely applied to target extraction and measurement in deep sky surveys. In deep surveys, the detection capability (depth) and the ability to deblend overlapping objects are often more important than measurement errors caused by photon noise. Extraction of target objects for different research purposes typically requires balancing detection capability against deblending capability.

### 1.1 Background Estimation

In a typical survey image, the photon count value at each pixel is the superposition of background signal and target source signal. To accurately measure target sources, SExtractor constructs a sky background distribution across the entire survey field. The specific method involves dividing the image into numerous small local regions. Within each local region, SExtractor first calculates the median and standard deviation of all pixel values, then removes pixels lying beyond median ± . The background value for this region is then estimated using the mode of the remaining pixel values. This process is repeated iteratively—if changes by less than a specified threshold between iterations, the algorithm converges and uses the mean as an approximation; otherwise, it continues iterating. Typically, the background is estimated using the histogram mode with the formula:

$$\text{mode} = \text{median} - 0.3 \times (\text{mean} - \text{median})$$

### 1.2 Object Detection and Profile Fitting

SExtractor' s detection is based on the intensity distribution of detected target sources. After subtracting the background in pixel regions belonging to potential sources, the pixel value distribution is fitted. SExtractor uses a Moffat profile for fitting, which satisfies the equation:

$$I(r) = I_0 \left[1 + \left(\frac{r}{\alpha}\right)^2\right]^{-\beta}$$

where $I_0$ is the central intensity, $r$ is the distance from the center, and $\alpha$ and $\beta$ are parameters. The CLEAN parameter should be set to $k \times r^2$ to ensure proper cleaning.

### 1.3 Centroid Determination

The most fundamental astronomical measurement is determining the centroid of a target. SExtractor uses a modified centering method [4] to calculate the center of a source, defined as:

$$X = \frac{\sum_i x_i I_i}{\sum_i I_i}, \quad Y = \frac{\sum_i y_i I_i}{\sum_i I_i}$$

where the summation runs over all pixels belonging to a target after threshold detection and segmentation, $(x_i, y_i)$ is the position of pixel $i$ in the image, and $I_i$ is the pixel value after background subtraction and deblending. The centroid accuracy is affected by the detection threshold and deblending of overlapping sources.

### 1.4 Photometric Methods

To achieve effective photometry for various targets in surveys, SExtractor provides multiple measurement modes, including aperture photometry.

**Isophotal Photometry (ISO):** This sums the pixel values $I_i$ within the detected isophotal area of a target source. The accuracy of isophotal photometry is heavily influenced by the detection threshold and background estimation.

**Aperture Photometry (APER):** Users define a series of aperture values (PHOT_APERTURES) to measure target sources. The accuracy primarily depends on aperture size—too large or too small apertures introduce significant errors.

**Corrected Isophotal Photometry (ISOCORR):** This improves upon isophotal photometry by accounting for flux below the detection threshold. It approximates the source intensity distribution with a two-dimensional Gaussian function. Under this approximation, the fraction of total flux enclosed within the isophotal area corresponding to the detection threshold $t$ satisfies:

$$\frac{\text{ISO}}{\text{total}} = 1 - \exp\left(-\frac{t^2}{2}\right)$$

For typical thresholds, this ratio is approximately 0.1961, meaning about 75.12% of flux is enclosed, with errors less than 0.3 mag. The corrected magnitude $m_{\text{ISOCORR}}$ is then:

$$m_{\text{ISOCORR}} = m_{\text{ISO}} - 2.5 \log\left(\frac{1}{0.1961}\right)$$

**Auto Photometry (AUTO):** This uses an elliptical aperture adjusted to each source, providing good accuracy for both point sources and extended sources with near-Gaussian profiles.

## 2. Parameter Settings in Deep Surveys

Different deep surveys adopt varying parameter settings for SExtractor. Table 1 shows the parameters used by Leauthaud et al. (2007) [8].

For deep surveys, the choice of DEBLEND_MINCONT is critical. The CFHTLS survey uses an extremely strict deblending threshold (DEBLEND_MINCONT = 0.002) to ensure good separation of overlapping sources, while the COSMOS survey uses DEBLEND_MINCONT = 0.0. Ground-based surveys suffer from lower resolution due to atmospheric seeing compared to space-based surveys, making source blending more severe. Therefore, ground-based surveys like CFHTLS must adopt more extreme deblending criteria.

A better approach for parameter optimization involves injecting artificial simulated galaxies into images before processing. These simulated galaxies, generated from structural and evolutionary models with random positions and orientations, allow users to adjust parameters to maximize recovery rates. This method is particularly valuable for ground-based surveys where atmospheric effects create more severe blending than in space-based data.

When a survey region has corresponding space-based observations, comparing ground-based and space-based photometry provides an excellent calibration. Figure 4 compares photometry for isolated galaxies between ground-based and HST observations in the HDF-N region [2], showing good agreement.

[Figure 4: see original paper]

## 3. Summary and Outlook

SExtractor' s fast processing speed and flexible parameter output have enabled excellent applications in deep sky surveys. Solutions have been developed for issues such as catalog completeness and false source removal due to noise. However, over-deblending remains a serious problem, particularly for high-resolution survey images. Existing detection methods could be improved. With the development of numerous high-resolution multi-band surveys, a promising solution is to incorporate color and morphological information for source detection and identification to achieve more accurate detection and deblending.

## References

The references section contains citations to key works on SExtractor and deep survey photometry, including Bertin & Arnouts (1996) [1], Williams et al. (1996) [2], and Leauthaud et al. (2007) [8], among others.

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