Automatic Morphological Classification of Galaxies in SDSS/GALEX Based on Catalog Photometry

Keaton J. Bell[1,2], Samir Salim[3]

[1]University of Washington, email: keatonb@u.washington.edu, [2]Indiana University REU, [3]Indiana University

Motivation for an Automatic Classification Method

Most galaxy classification efforts are based on the Hubble "tuning fork" scheme, whereby galaxies are categorized on their defining structural features. This makes classification a laborious task, since images of every galaxy must be visually inspected. Additionally, it is not always easy or possible to observe galaxies with sufficient resolution to discern between potential types. Distance or certain galactic orientations can obscure critical information causing type ambiguity. In particular, Figure 1 demonstrates the confusion that can exist between face-on lenticular galaxies and round ellipticals. The introduction of a computerized classification method stands to save a great deal of work and avoid many of the pitfalls of visual classification.

Figure 1: Face-on lenticular galaxies, such as M85 (left), can often appear similar to ellipticals, like E2 galaxy M60 (right), leading to misclassification by visual methods [Image credit: Wikimedia Commons].

Typing Method Using RC3 Galaxy Templates

The Third Reference Catalog of Bright Galaxies (RC3; Corwin et al. 1994) contains accurate classifications for galaxies with diameters larger than 1 arcminute, B-band magnitudes brighter than 15.5, and recessional velocities less than 15,000 km/sec. There are 5,144 RC3 galaxies in the combined Sloan Digital Sky Survey (SDSS; York et al. 2000) and Galaxy Evolution Explorer (GALEX; Martin et al. 2005) footprints that we use as comparison templates in our typing method, grouped as ellipticals, lenticulars, spirals, and irregulars. Spectral energy distributions (SEDs; very low-resolution spectra) are constructed from the catalog photometry to hold the UV and optical color information of the templates. We expect this color information to be indicative of morphology because it can be used to derive galactic star formation rates and stellar masses (Salim et al. 2007). We also calculate three morphological parameters for every galaxy from the catalog data: r-band light concentration, UV-to-optical size ratio, and ellipticity. A galaxy to be typed is estimated to belong to the group where the median chi-squared deviation between the shape of the galaxy's redshift-corrected SED and the group's templates (Figure 2), plus the deviation of their morphological measurements, is a minimum.

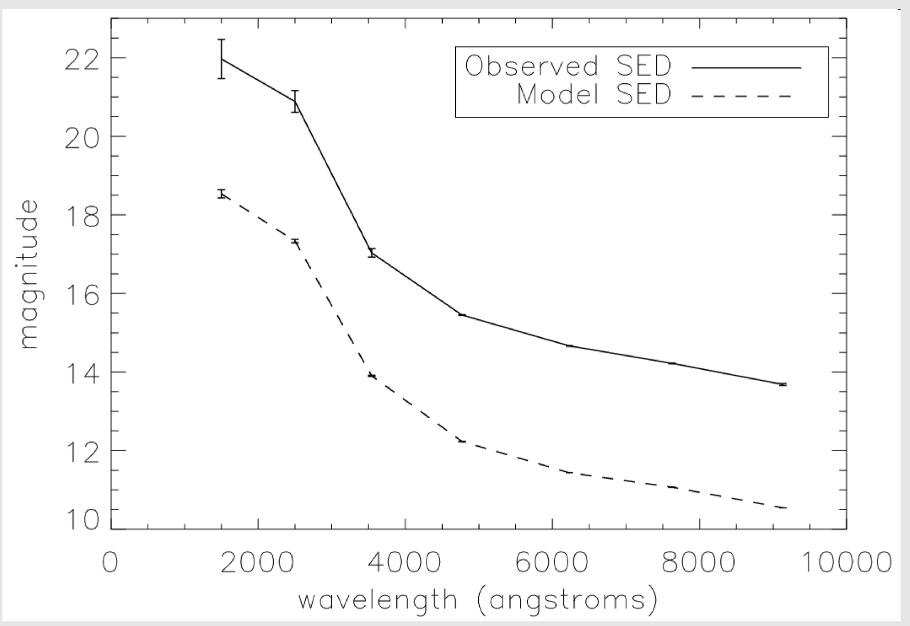

Figure 2: The chi-squared deviations between the shape of a galaxy's spectral energy distribution (SED) and those of template galaxies are used to estimate its morphological type.

Since different types of galaxies do not occupy distinct regions in color-space, this typing method can only produce probabilities that a galaxy is either elliptical, lenticular, spiral, or irregular, rather than identifying sure matches. Huertas-Company et al. (2011) recently asserted that such probabilistic estimates are fundamental results of the classification process due to the continuous nature of galactic properties. By running the RC3 template galaxies through the matching algorithm themselves, we determine the likelihoods that each possible result correlates to each actual morphology, as summarized in Table 1. Irregular galaxy outcomes are the most accurate, while common elliptical and spiral interlopers cause matches to the lenticular templates to be the least clear-cut.

Source Probability	Typing Result: E	L	S	I
E	52.1%	15.7%	5.1%	1.9%
L	37.6%	44.9%	13.4%	2.6%
S	7.8%	32.3%	61.2%	24.6%
I	2.6%	7.1%	20.4%	70.9%

Table 1: For typing results of elliptical, lenticular, spiral, and irregular, we give the probabilities that the outcome was reached for each actual galaxy type, assuming all outcomes are equally likely overall. Correct matches are highlighted along the diagonal.

SDSS Photometry Corrections

SDSS photometry has known errors for spatially large galaxies like the RC3 template galaxies (Corwin et al. 1994) used in our typing process. For instance, the SDSS photometric reduction algorithms often perform sky subtraction using a sky region that contains light from the galaxy. To address this problem, we use the correctly determined photometric data independently measured from SDSS images by West et al. (2010) to obtain color-correction offsets for SDSS colors as functions of galaxy size. Figure 3 by West et al. (2010) displays the corresponding overall r-band correction that we also applied to correct our RC3 template SEDs.

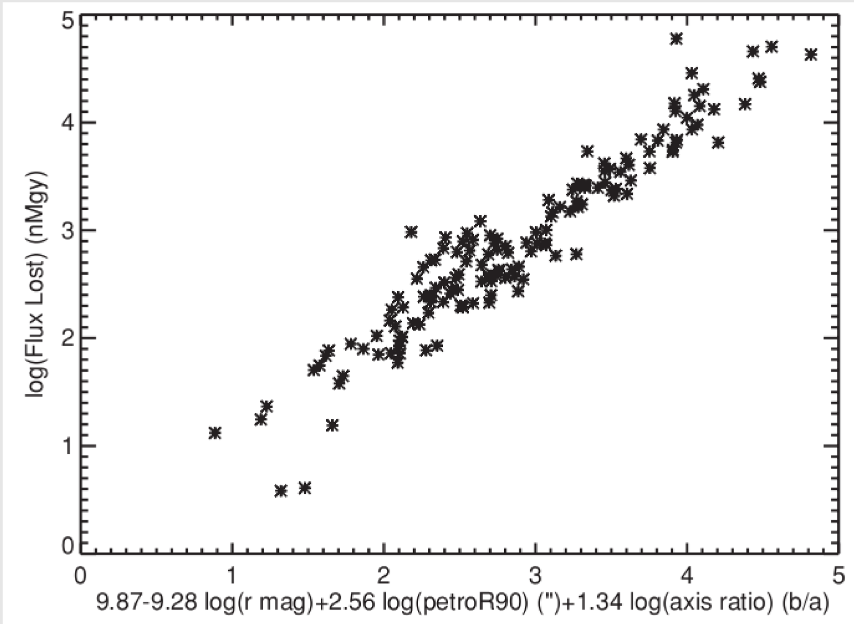

Figure 3: This plot from West et al. (2010) gives the r-band photometry correction needed for large galaxies in the SDSS Data Release 7 due to improper background subtraction.

Large galaxies are also often victims of deblending, which happens when different regions of a single galaxy are identified as distinct galaxies. Since individual galaxy shreds only represent a fraction of the galaxy's total brightness, they can be removed from our sample by identifying where SDSS magnitudes are inconsistent with RC3 catalog photometry (Figure 4).

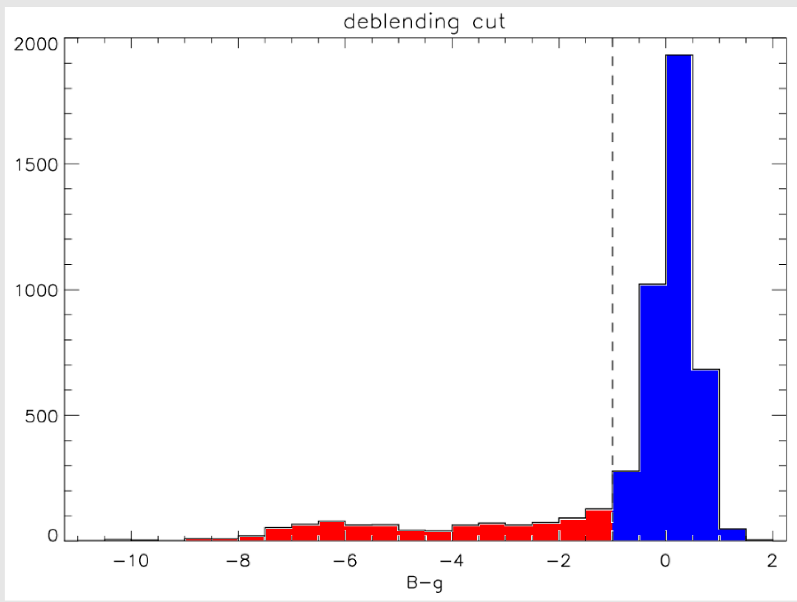

Figure 4: We remove deblended galaxies from our sample by requiring SDSS R90r size > 6 arcsec and B-g color > -1.

Results for COSMOS Galaxies at z~0.1

Figure 5: The superior resolution of the HST ACS images (left) over SDSS data (right) has allowed morphological types to be assigned to galaxies in the COSMOS survey.

Because our typing method is targeting galaxies contained within the SDSS catalog, we needed to verify that it's accuracy is upheld at the median survey redshift of z~0.1. Rough morphologies for galaxies at this distance are tabulated within The Zurich Structure & Morphology Catalog v1.0 (Scarlata et al. 2007) from high-resolution Hubble Space Telescope images (Figure 5) obtained for the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007). Table 2 gives the percentage of times that a COSMOS galaxy cataloged as either early-type, disk-like, or irregular was found to be most similar to the set of elliptical, lenticular, spiral, and irregular galaxy templates. The results mimic those obtained for the RC3 template sample, justifying the extension of our method to z~0.1. Matches obtained at this distance must still be interpreted by the source probabilities of Table 1.

COSMOS Type	Typing Result Frequencies: E	L	S	I
Early Type	82.0%	6.0%	6.0%	6.0%
Disk-like	25.9%	29.6%	22.2%	22.2%
Irregular	35.3%	5.9%	29.4%	35.3%

Table 2: For each morphological category in the The Zurich Structure & Morphology Catalog v1.0 (Scarlata et al. 2007), these are the percentage frequencies that COSMOS galaxies at z~0.1 are matched to the RC3 template set corresponding to ellipticals, lenticulars, spirals, and irregulars.

Type, Stellar Mass, and Star Formation Rate

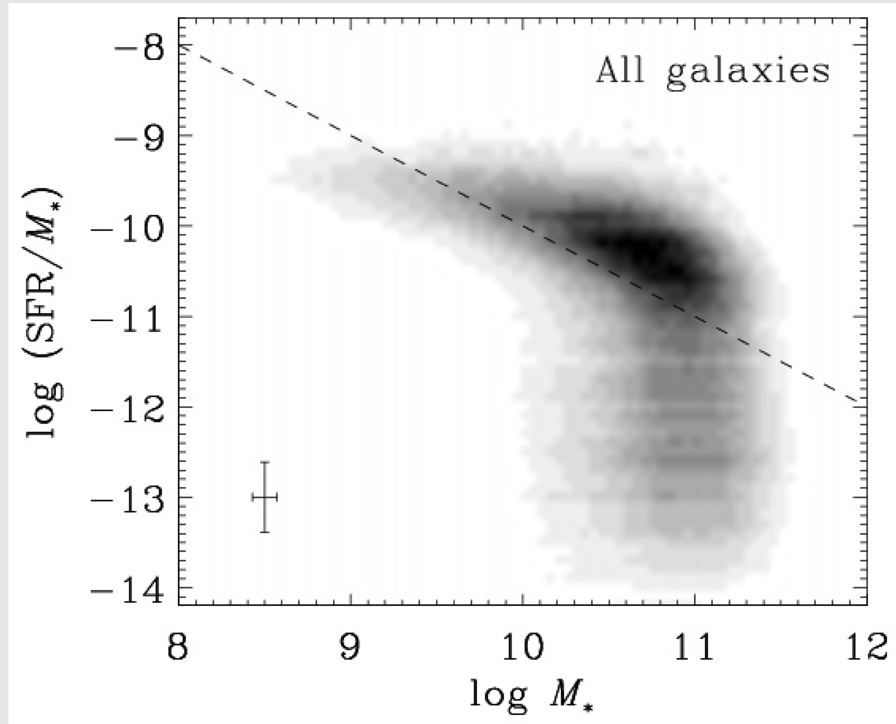

Figure 6: This relationship between specific star formation rate (SFR) and stellar mass for galaxies in general is provided by Salim et al. (2007). The dashed line indicates a constant star formation rate of 1 solar mass per year. The typical error in specific SFR on the star-forming sequence is 0.2 dex and 0.05 dex in mass.

We classified a sample of 48,295 galaxies at z~0.1 that have had stellar masses and dust-corrected star formation rates calculated from their UV and optical photometry (Salim et al. 2007). Figure 6 shows the distribution of specific star formation rates versus stellar masses for all of the galaxies. Recreating this plot separately for the galaxies matched to elliptical, lenticular, spiral, and irregular templates (Figure 7) suggests fundamental differences between these types. While both lenticular and elliptical galaxies are the most massive with ~10^{11} solar masses of stellar material each, the lenticulars are shown to be generally undergoing more current star formation. Our classification also reproduces the fact that irregular galaxies extend to lower stellar masses than the spirals, and that for a given mass, irregulars form stars more vigorously than spirals, suggesting that morphology may contribute to the intrinsic width of the star-forming sequence.

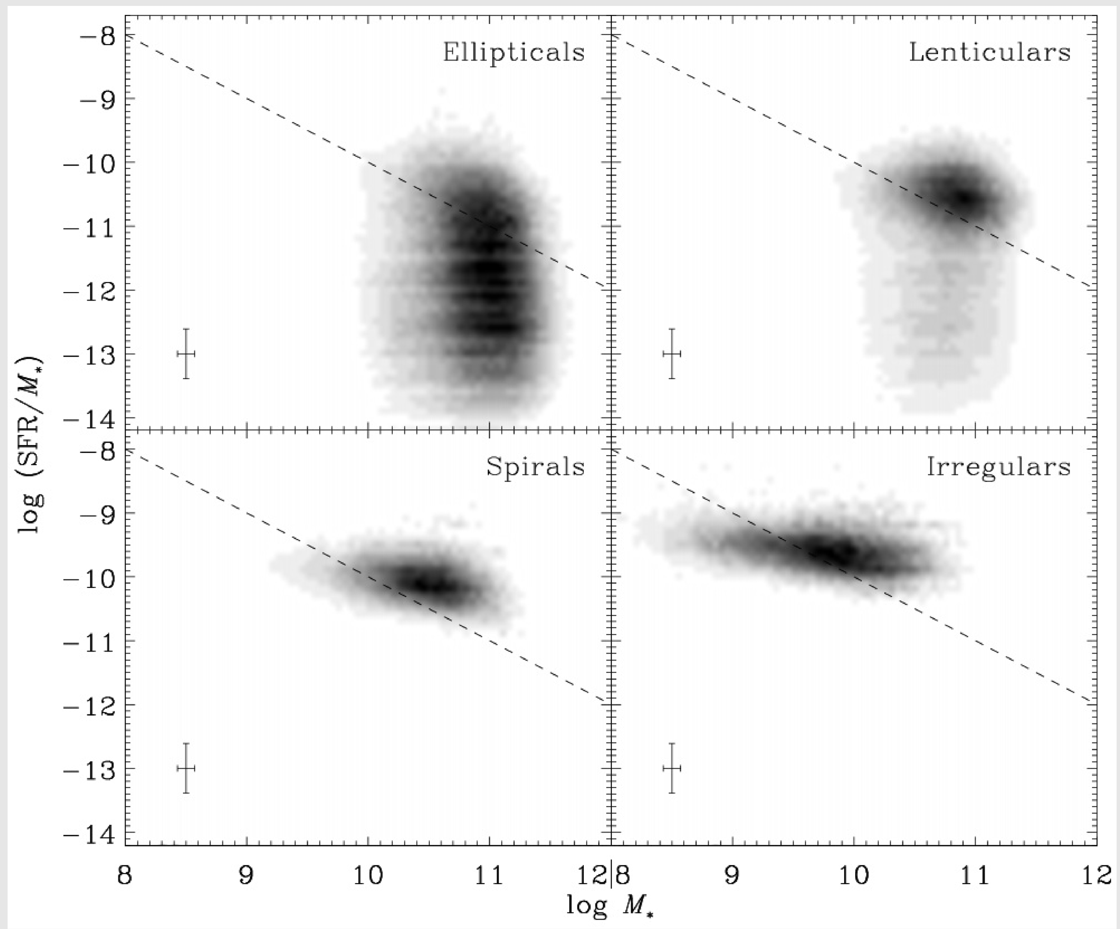

Figure 7: Matching the galaxies plotted in Figure 6 to the RC3 template groups corresponding to elliptical, lenticular, spiral, and irregular galaxies reveals correlations between stellar mass and star formation histories characteristic to each of these types.

Concluding Remarks

By developing a reference for what galaxies known to be elliptical, lenticular, spiral, and irregular look like in the SDSS and GALEX databases, we were able to derive probabilities that individual catalog galaxies belong to each of these four classes. The method holds up at a redshift of z~0.1 and allows for the study of the properties of different galactic types at distances beyond where visual classification is accessible.

References

Corwin, H. G., Buta, R. J., and de Vaucouleurs, G., 1994, AJ, 108, 2128
Huertas-Company, M., et al. 2011, A&A, 525, A157
Martin, D. C., et al. 2005, ApJ, 619, L1
Salim, S., et al. 2007, ApJ, 173, 267
Scarlata, C., et al. 2007, ApJ, 172, 406
Scoville N., et al. 2007, ApJS, 172, 1
West, A. A., et al. 2010, AJ, 139, 315
York, D. G., et al. 2000, AJ, 120, 1579

Acknowledgment

This work was completed with the support of NSF grant AST-0452975 as part of the astronomy REU program at Indiana University.

Photo

(actively seeking new home)