# Inferring Galaxy Morphology Through Texture Analysis

Kinman Au[*], Christopher Genovese[*], Andrew Connolly[†]

## 1 Galaxy Morphology

Galaxies are not static objects; they evolve by interacting with the gas, dust and other galaxies in the surrounding environment. The most visible tracer of galaxy evolution is the diverse range of galaxy morphologies that have been observed. Scientists are interested to know the distribution of morphological types of the early Universe and how it has changed over time. They also want to know how the local density of galaxies affect the distribution of morphologies. The answers will provide important clues about the origin of the Universe. This paper proposes a new statistical approach to study galaxy morphology which will help scientists to answer these questions.

The modern study of the relationship between galaxy morphology and evolution was started by Erwin Hubble [6]. He proposed a galaxy classification scheme based on morphology (Figure 1). Hubble's scheme divides galaxies into two main groups —ellipticals (E's) and spirals (S's). Elliptical galaxies have the shape of an ellipsoid with smooth and featureless surface and spiral galaxies have the ellipsoidally formed "bulge" with spiral structures surrounding it. Spiral galaxies are sub-categorized as normal and barred. A barred galaxy has a linear bar structure extended from the center that rotates and channels gas and dust to the center. Each type is further sub-divided according to the tightness of the spiral arms.

Hubble's scheme is more than a taxonomy of galaxies; he also argued how the morphology of a galaxy related to its evolution history. Hubble suggested that the early Universe consisted mostly of ellipticals and they grew isolately into spirals. To the contrary, the modern theory assumes active interaction among galaxies, and scientists believe elliptical galaxies are produced by collision among spiral galaxies that predominated the early epoch of the Universe.
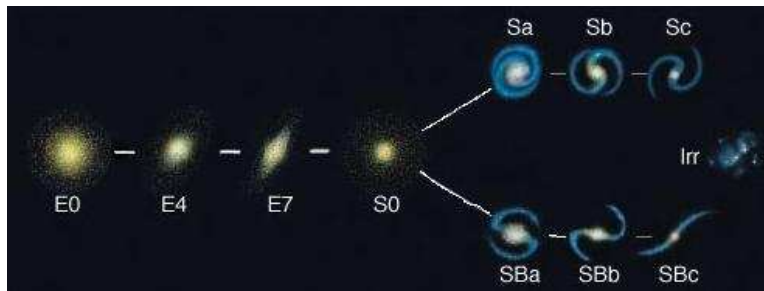


Figure 1: *Hubble Scheme - commonly known as Hubble Turning Fork. This tuning-fork diagram divides galaxies into two main groups —elliptical (E's) and spirals (S's). A third group is formed by the irregulars (Irr's). The S0 galaxies or lenticular galaxies are midway between ellipticals and spirals, they have a nuclear bulge surrounded by a flat disk, but no spiral arms. There are two types of spiral galaxies, normal and barred. S( ) for normal spirals, SB( ) for barred spirals. Each type is further sub-divided (a, b, c) according to the tightness of the spiral arms.* [7]

Modern sky surveys are capturing enormous amount of data. For example, the Sloan Digital Sky Survey(SDSS) [1] will eventually capture digital images from millions of galaxies. It has become infeasible for scientists to study galaxy morphology based on Hubble's scheme which requires visual inspection of huge amount of galaxy images. Therefore, scientists are trying to adopt a quantitative approach that can extract important features of a galaxy automatically from its image.

[*]Department of Statistics, Carnegie Mellon.
[†]Department of Physics and Astronomy, University of Pittsburgh.

A standard quantitative approach to galaxy morphology is to decompose the surface brightness of a galaxy into two components - a bulge (the central bright core) and a disk that surrounds the bulge. The ratio of each component's share of total brightness is called Bulge-to-Disk ratio (B/D). Such ratio is widely used to represent the morphology of a galaxy because the ratio is approximately decreasing from left to right in Hubble's scheme. Therefore a galaxy with a small B/D is likely to be a spiral while a large B/D suggests a galaxy to be a elliptical. However, the Bugle-to-Disk ratio gives only a imprecise description of galaxy morphology, it fails to distinguish subcategories of spiral galaxies, i.e. barred and non-barred galaxies, and sub-types of spiral with various tightness.

In contrast to the Bulge-Disk decomposition that attempts to infer indirectly the internal structure of a galaxy through modeling the global intensity distribution, this paper proposes a new approach that uses the local structure in an image to infer directly the morphological features of a galaxy such as the presence of a bar and the geometry of spiral arms, which are extremely useful for studying galaxy evolution.

## 2   Methods Overview

Our algorithm starts with extraction of the local structure in an image, then we use this local structural information to model the morphological structures of a galaxy, i.e. the bar and spirals. The algorithm consists of three steps —(i) orientation analysis of texture, (ii) merging multiscale information, and (ii) fitting a morphological model to estimate bar and spiral structures. Fitting a model directly to a galaxy image is difficult because it is highly susceptible to noise. A bright star close to a galaxy could easily distort the fitted model completely. Therefore we first extract structural information of an image by analyzing its texture.

The texture of an image is the local spatial structure of the intensity. For a galaxy image, such structure is not stationary across space and is varying with change of scale. In order to capture the textural structure of a galaxy image, we derived an algorithm that perform *orientation analysis* to extract the dominant orientation of texture at each point under different scales and *information merging* to combine orientation information from all scales. The orientation information is summarized by an *orientation field* that tells us the dominant orientation angle and strength at each point of an image (Figure 2).
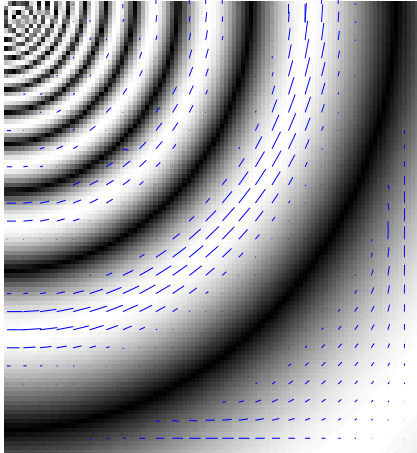
Figure 2: An image overlapped with its orientation field —the length and inclination of each line segment represent the orientation strength and angle at each point.

**Orientation analysis** measures the anisotropicity of intensity in either a 2D or 3D setting, which is common in computer vision, e.g. edge-detection and optical flow. A standard method to detect the anisotropicity is by applying oriented filters. Candes and Donoho [3] proposed a filtering scheme based on ridge functions that favors capturing linear structure in an image, i.e lines and edges. In this study, we use a modified version of their filter design and a different approach to perform multiscale analysis on an image.

To perform orientation analysis of texture, a multiscale filtering algorithm is applied to decompose an image into components localized in both scale and orientation. Figure 3 shows the schematic diagram of the algorithm. We obtain a sequence of images from the finest scale (level 1) to the coarsest scale (level 3) by applying a lowpass filter and subsampling* iteratively. Then by applying a set of oriented filters [2] to these images, we can measure the signal strength of various orientations and scales. In other words, we can capture the local structure of various of fineness. Under an

---

*By factor 2, i.e. keeping every second pixel in each dimension.

appropriate scale, we can see high contrast of signal strength from different orientations for the locations with strongly oriented texture. For example, at the location with horizontally oriented texture, it has high signal strength when applying a filter with the same orientation, but a very weak signal with a vertically oriented filter. Hence, at each scale, we estimate the orientation field, i.e. orientation angle and strength at each point, with the angle that has the highest concentration of signal and the magnitude of this signal concentration [2]. Therefore, a point light source, such as a star, has a very low orientation strength as the signal strength is approximately the same at all orientation angles.
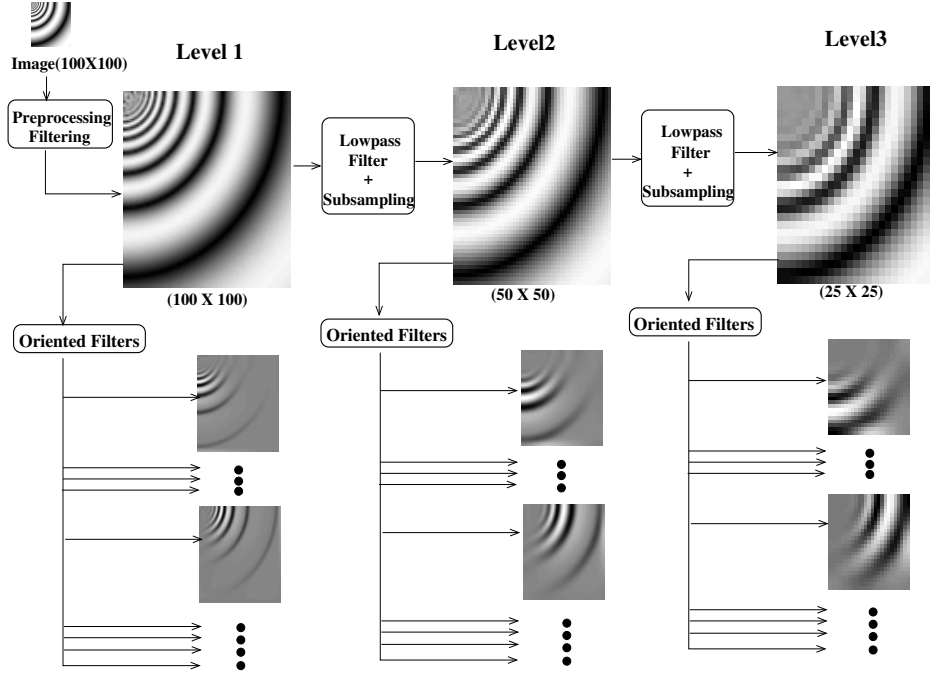


Figure 3: Schematic diagram of multiscale filtering with oriented filters (only horizontally and vertically filtered images are shown).

Now we have the orientation information spread across components of different scales. However, in order to fit a morphological model, we want to utilize all information. **Information merging** is hence a necessary step to obtain a unified representation of the orientation structure. Figure 5 shows the image of galaxy M51 and its orientation fields under various scales. An important feature of these orientation fields is that they exhibit *multiscale bias-variance tradeoff*. An orientation field under a coarse scale losses fine textural structures and has higher bias. However it has a lower variance and a smoother flow. On the other hand, detailed structures can be captured under a finer scale but with a higher variance and a rough flow. In order to attain a good bias-variance balance, we developed a *multiscale state-space model* [4], which uses scale level to replace the role of time.
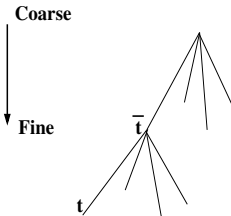


Figure 4: Quadtree from coarse to fine scale. Node $\bar{t}$ is the parent of node $t$.

The pyramidal structure of the orientation field can be defined on the quadtree structure (Figure 4). Each node represents a pixel and each pixel corresponds to four pixels of next finer scale (due to subsampling). Our state-space model is

$$\textit{State Eq. } Y(\mathbf{t}) = Y(\bar{\mathbf{t}}) + w(\mathbf{t}), \qquad (1)$$

$$\textit{Estimation Eq. } X(\mathbf{t}) = Y(\mathbf{t}) + v(\mathbf{t}), \qquad (2)$$

where $w(\mathbf{t})$ and $v(\mathbf{t})$ are independent noise terms with zero means and $Y$ is the true orientation (for both strength and angle) and $X$ is the estimated orientation from our filtering scheme. Using standard Kalman filtering, we obtain a unified representation $Y$ by merging all the information $X$ (Figure 5).
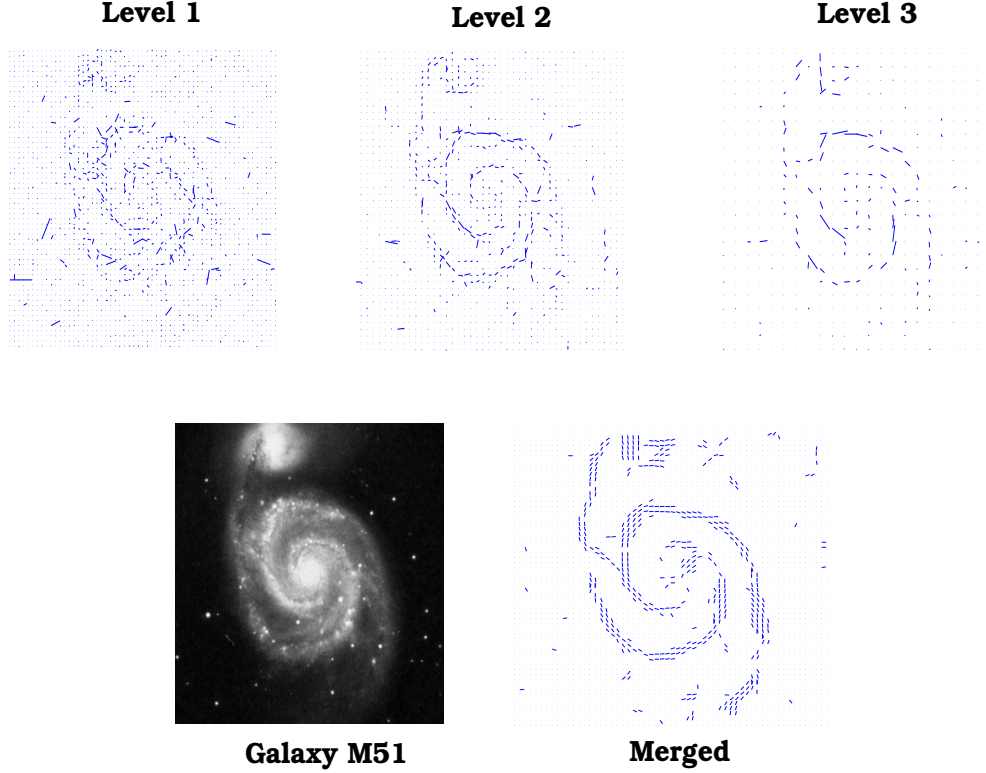
3

Figure 5: *Image of galaxy M51 and its orientation fields.*

**Morphological model fitting** is the next step to extract information about bars and spirals. Our morphological model of galaxy follows so-called Grand Design model, i.e. two spiral arms attach to two ends of a bar at the center of a galaxy (could be a degenerated bar for a non-barred galaxy) (Figure 6). We use *logarithmic spiral* as our model of spiral arms.
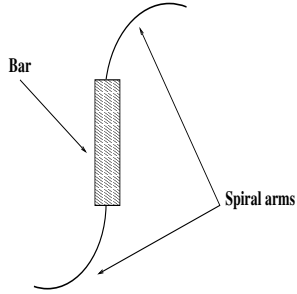


Figure 6: Diagram of the Grand Design model.

A logarithmic spiral in polar coordinates is

$$\theta = t + k \tag{3}$$
$$r = \beta e^{\alpha t}, \quad 0 \le t \le T \tag{4}$$

where $k$ , $\beta$ are initial angle and radial length which are fixed by the position of the bar to which the spiral is attached, and T is the winding angle of a spiral arm. We assume symmetry, i.e. both spiral arms have the same values of $\alpha$ and $T$. We also incorporate the inclination of an galaxy [†] into the model. Initially, assuming the bar structure has no width, then the model can be represented as a parametrized curve $f(x; \phi)$, where $\phi$ is a vector of model parameters. Let **u** be an orientation field produced by our texture analysis algorithm. It essentially give an estimated tangent of spiral arms and bar at each point, weighted by the orientation strength. In order to find the best model to describe the spiral and bar structures of a galaxy, we want tangents of the fitted model at every point as close to the orientation field as possible. And we measure the closeness by the total dot product between

---

[†]Inclination is the angle between the line of view and the vector normal to the galaxy, face-on and edge-on galaxies have inclination angle of zero and ninety degrees respectively.

tangents of the model and the orientation field[‡]. Mathematically, we estimate the model parameters $\phi$ by

$$\hat{\phi} = \operatorname*{argsup}_{\phi} \int \left[ \mathbf{u}(f(x;\phi)) \cdot \frac{df(x;\phi)}{dx} \right]^2 dx, \tag{5}$$

where the dot product is squared for mathematical convenience. Given the length from the fitted model, we then estimate the width of the bar structure by maximizing the total dot product between the field and the estimated bar orientation for all points inside the bar.

**Bar-to-Bulge ratio** and **integrated curvature** are two important statistics derived from a fitted model to characterize the bar and spiral structures. Since galaxies vary in size, instead measuring the absolute size of a bar, we use a ratio —area of the fitted bar to area of the bulge [§], i.e. Bar-to-Bulge Ratio, as the measure of bar size. For a non-barred galaxy, the model fitting algorithm tries to fit a bar to the bulge, therefore we expect that a fitted bar is smaller than the bulge. Hence we classify a galaxy to be barred if its Bar-to-Bugle ratio is larger than one.

The total integral curvature of both spiral arms (normalized by physical unit per pixel) measures the tightness of the spiral structures of a galaxy. It decreases from $S(B)a$ to $S(B)c$ in Hubble's scheme. Since it measures the tightness continuously, it gives a finer description of spiral arms than Hubble's scheme.
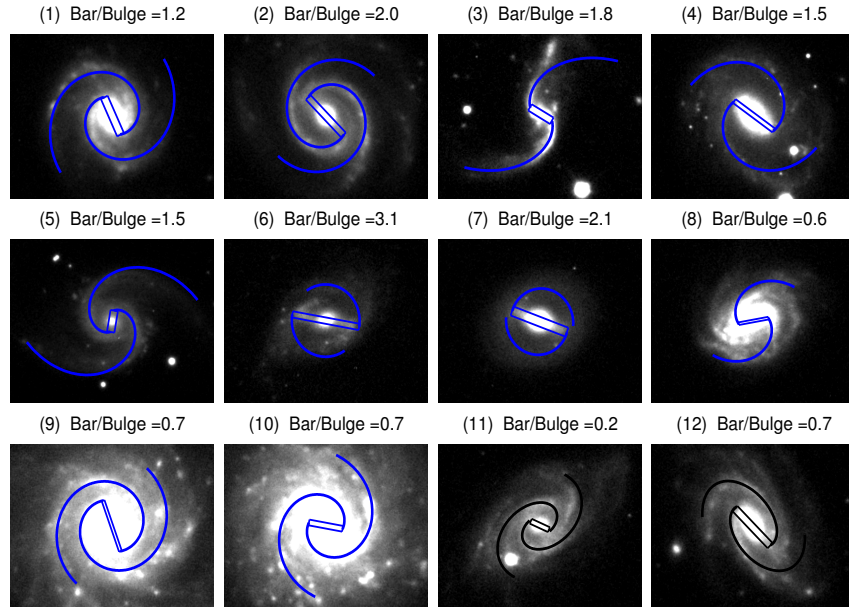
# 3   Results and Extensions



Figure 7: Twelve representative galaxy images and their fitted models.

In Figure 7, we see the results of the model fitting of the twelve representative images. Images 1 to 5 are face-on and relatively bright images. They have good model fits in general.

---

[‡]We abuse the notation of dot product here which is not well defined for a orientation field as it has no directions. But we can turn a orientation field into a vector field easily by assigning a direction at each point.

[§]We define bulge area as the area of a circle with effective radius (corresponding to the isophote containing half of the total brightness of the bulge). We use GALFIT [5] for bulge fitting.

Images 6 and 7 are ring galaxies with a bar at their centers, logarithmic spiral can model the ring structure very well (i.e. $\alpha$ close to zero for Eq. 4).

Images 8 to 10 are more complex; they appears to have more than two spiral arms. In this case, the best fit model may not best describe the tightness of the entire spiral structure. One solution is to use weighted average of all possible models (with heavier weights for those with better fits) instead of using only the best fit model.

Images 11 and 12 are inclined galaxies. By incorporating inclination into our model, we can model these galaxies adequately. However, for a galaxy of high inclination we probably cannot obtain much information by fitting our morphological model.

The estimated Bar-to-Bulge ratio is shown on the top of each image. We can see that, in general, a barred galaxy has a higher Bar-to-Bulge ratio (larger than one) than a non-barred galaxy. By the time writing this paper, we have not yet obtained the data of physical unit per pixel for the images to normalize the fitted models. Therefore we cannot compare the integral curvature of these galaxies.

In the cases that a galaxy deviates from the Grand Design model or has high inclination, we want to assess how reliable to use Bar-to-Bulge ratio and integral curvature to describe the bar and spiral structures of an galaxy. Therefore, a nature next step is to derive a goodness-of-fit measure/test for the fitted model.

# 4   Galaxy Morphology and Evolution : Conclusion

Looking deep into the Universe with a telescope allows us to look not only into space, but also back in time since light from more distant objects take longer time to reach us. By studying the distribution of morphological types at different distances, scientists can see how galaxy morphology change with time. We can further investigate the relationship between density of galaxies and the distribution of morphological types, which will give us more clues about how the galaxies interact.

The Sloan Digital Sky Survey[1] will map one-quarter of the entire sky, determining the positions and magnitude of more than 100 million celestial objects. It will also measure the distances (via redshift ¶) to more than a million galaxies and hence the local density of these galaxies.

Ultimately, we want to apply our algorithm to study this huge amount of data. By investigating how the bar and spiral structures are related to redshift (time) and local density of galaxies (environment), we will gain more insight into how galaxies evolve.

# References

[1] Sloan digital sky survey —http://www.sdss.org/.

[2] Kinman Au. Phd. thesis. *Carnegie Mellon University*, 2006.

[3] E. J. Cands and D. L. Donoho. Ridgelets: a key to higher-dimensional intermittency? *Phil. Trans. R. Soc. Lond. A.*, 1999.

[4] K.C. Chou, A.S. Willsky, and R. Nikoukhah. Multiscale systems, kalman filters, and riccati equations. *IEEE Transactions on Automatic Control, Vol.39, No.3*, 1994.

[5] Peng et al. Detailed structural decomposition of galaxy images. *The Astronomical Journal*, 2003.

[6] E.P. Hubble. The realm of the nebulae. 1936.

[7] Hubble Scheme Image —http://www.astro.psu.edu/users/rbc/a1/lec20n.html.

---

¶Since the farther a galaxy is from earth, the faster it is receding from us, and the quicker a distant object is moving away from us, the light coming to us from this object is shifted toward more the red end of the light spectrum. Scientists measure the amount of red shift in the spectrum of a galaxy to figure out how far away it is from us.