AUTOMATED GALAXY MORPHOLOGY: A FOURIER APPROACH

S. C. Odewahn, 1 S. H. Cohen, 1 R. A. Windhorst, 1 and Ninan Sajeeth Philip 2

Received 2001 September 4; accepted 2001 December 6

ABSTRACT

We use automated surface photometry and pattern classification techniques to morphologically classify galaxies. The two-dimensional light distribution of a galaxy is reconstructed using Fourier series fits to azimuthal profiles computed in concentric elliptical annuli centered on the galaxy. Both the phase and amplitude of each Fourier component have been studied as a function of radial bin number for a large collection of galaxy images using principal-component analysis. We find that up to 90% of the variance in many of these Fourier profiles may be characterized in as few as three principal components and that their use substantially reduces the dimensionality of the classification problem. We use supervised learning methods in the form of artificial neural networks to train galaxy classifiers that detect morphological bars at the 85%–90% confidence level and can identify the Hubble type with a 1 σ scatter of 1.5 steps on the 16 step stage axis of the revised Hubble system. Finally, we systematically characterize the adverse effects of decreasing resolution and signal-to-noise ratio on the quality of morphological information predicted by these classifiers.

Subject headings: galaxies: fundamental parameters (classification) — galaxies: structure — methods: data analysis

1 INTRODUCTION

One of the earliest galaxy classification schemes, discussed by Hubble (1926), was based on the visual appearance of two-dimensional images. This and other early schemes, based largely on photographic images in roughly the B bandpass, used global image properties such as the visually perceived bulge-to-disk ratio and the degree of azimuthal surface brightness symmetry as major classification criteria. Even these early systems discussed the presence of spiral structure and the overall characteristics of the spiral arms: grand design of high or low pitch angle versus patchy or multiple arms (later termed flocculent spirals). The early Hubble system delineated spirals into a parallel sequence of barred and unbarred disk galaxies in a two-dimensional system. A refinement of this approach by de Vaucouleurs (1959) established a three-dimensional classification volume with the major axis being the Hubble stage (E, S, I) and the two other axes representing family (barred or unbarred) and variety (ringed or nonringed). The revised Hubble system (RHS) was designed to describe a continuum of morphological properties. Early workers in the field strove to systematize the degree of visual structures we see in galaxies (bulge/disk ratio, spiral arms, bars, rings) in an effort to understand the physical processes that formed these different galaxies.

A useful classification system must relate members of different classes to well-understood general properties of the objects being systematically classified. It has long been known (de Vaucouleurs 1977; Buta et al. 1994; Roberts & Haynes 1994) that the stage axis of the Hubble sequence produces smooth, strong correlations with well-known global properties: color, surface brightness, maximum rotational velocity, and gas content. These measured quantities are linked directly to very important physical properties: stellar population fraction, the surface density of stars, total mass of the system, and the rate of conversion of gas to stars. As discussed by Odewahn & Aldering (1995), such correlations may be smooth and significant in a mean sense; however, the degree of accidental scatter in these relations is quite large. In most cases this scatter is much larger than that introduced by measurement error and is clearly cosmic in origin. The family and variety estimates of the RHS produce much less significant correlations with such properties. Buta & Combes (1996), in an excellent review of these morphological features, make clear that family and variety convey information about the specific dynamical properties of disk systems: the degree of differential rotation, the presence of certain families of resonant orbits among the stellar component of the disk, and the response of the viscous disk components (the gas) to the global and local properties of the galaxy potential well.

Galaxy classification systems are generally built on the visual appearance in an image, and hence data sets based on these systems can be gathered from imaging surveys alone. The mean correlations described above are then used to draw inferences about important physical processes describing galaxies via these catalogs. In other words, morphological classification provided a “cheap” and direct means of acquiring large statistical samples describing galaxies. In some respects, the qualitative nature of this classification approach, and the small number of recognized experts who classified in some established system, gave the approach a less than deserved poor reputation. However, even recent comparisons among the estimates of experienced human classifiers (Naim, Ratnatunga, & Griffiths 1995) show that a reliable level of repeatability is achieved. With an expansive increase in the quantity and quality of galaxy images in the last decade via the Hubble Space Telescope (HST) and a host of ground-based imaging surveys, interest has returned to the field of morphological classification (Driver, Windhorst, & Griffiths 1995a; Driver et al. 1995b; Glazebrook et al. 1995; Abraham et al. 1996; Odewahn et al. 1996). Many workers have developed automated image analysis systems.

1 Department of Physics and Astronomy, Arizona State University, Box 871504, Tempe, AZ 85721-0065.
2 Department of Physics, Cochin University of Science and Technology, Kochi 682 022, India.
that provide quantitative estimates of galaxy types (Okamura, Kodaira, Watanabe 1984; Whitmore 1984; Spier- 
kermann 1992; Storrie-Lombardi et al. 1992; Abraham et al. 1994; Odewahn 1995, 1997; Han 1995). In one the most 
physically meaningful approaches, Rix & Zaritsky (1995) 
showed in a small sample of spiral galaxies that kinematic 
distortions in the global velocity field of a galaxy can be 
linked to global asymmetry measurements in the light distri- 
bution. A discussion of novel feature spaces for identifying 
peculiar galaxies is presented by Naim et al. (1997). All of 
these studies have produced a wealth of information about 
the systematic properties of galaxies, particularly in the area 
of comparing the low- and high-redshift universe (Abraham 
et al. 1996; Odewahn et al. 1996). However, it must be con- 
ceded that none of these classification systems are truly mor- 
phological in nature. In general, each method uses one or 
more two-dimensional parameter spaces based on global 
image properties and produces a type estimate via mean cor- 
relations like those discussed above using some multivariate 
(pattern classification approach (e.g., linear parameter space 
divisions, principal-component analysis, decision tree, arti-
ficial neural network).

In this paper we describe a truly morphological approach 
to galaxy classification based on the Fourier reconstruction 
of galaxy images and the subsequent pattern analysis based 
on the amplitude and phase angle of the Fourier compo-
nents used. This method can quantitatively detect the pres- 
ence of spiral arms and bars. In addition, it provides a 
systematic means of describing the degree of large-scale 
global asymmetry in a galaxy, a property known to corre- 
late loosely with Hubble stage but one that is also strongly 
linked to important physical events such as merging and 
tidal interaction.

In § 2 we describe our machine-automated classification 
method. In § 3 we compare independent sets of visual esti-
mates of morphological types in the RHS that establish a 
well-understood set of training/testing cases. These data 
are comprised of a local sample from ground-based imaging 
and a distant sample collected from the HST archive. In § 4 
we demonstrate a practical application using these data to 
train and assess the performance of several Fourier-based 
galaxy classifiers. We discuss the role of systematic and acci-
dental errors in its use caused by varying image resolution 
and signal-to-noise ratio (S/N). In § 5 we summarize the 
results and lay the groundwork for future morphological 
studies of distant galaxy populations.

2. CLASSIFICATION METHODOLOGY

The two-dimensional luminosity distribution observed in 
most galaxies most often displays a high degree of azimuthal 
symmetry. This justifies the use of one-dimensional radial 
surface brightness profiles in describing the radial stellar 
surface density distribution. Such profiles are then decom- 
posed into constituent photometric components (de Vau-
couleurs 1958; Simien & de Vaucouleurs 1986; Kent 1986; 
de Jong & van der Kruit 1994). This profile is extracted from 
the galaxy image in many ways but most often by comput- 
ing mean flux density in elliptical annuli having the shape 
and orientation of the faint surface brightness, outer iso-
ophotes. Parameterizing this profile, e.g., with a concentra- 
tion index, will produce a measurement that correlates well 
with the bulge-to-disk ratio and hence with Hubble type. 
Local departures from this overall radial symmetry in the 
form of high spatial frequency components (arms, bars, 
rings) form the basis for additional criteria to visual galaxy 
classification systems like the RHS.

A method that quantitatively describes departures from 
azimuthal symmetry in galaxy images would seem to be a 
rich starting place for developing a machine-automated 
classification system of either the supervised learning variety 
(Odewahn et al. 1992; Weir, Fayyad, & Djorgovski 1995) or 
the unsupervised variety (Mahonen & Hakala 1995). Toward 
this end, we have adapted a moments-based image 
analysis technique (Odewahn 1989) to fit Fourier compo-
nents to fixed grid azimuthal profiles in elliptical annuli 
to reconstruct galaxy images. One may think of this step as an 
onimized data compression technique for reducing the 
dimensionality of our ultimate pattern classification prob-
lem: the recognition and characterization of morphological 
features in galaxies.

The technique of Fourier decomposition presented here is 
used to quantify the two-dimensional luminosity distribu-
tions of galaxies. Similar approaches have been used to 
study the shapes of isophotes in elliptical galaxies by Bender 
& Mollenhoff (1987), and applications to the luminosity dis-
tribution are described by Lauer (1985), Buta (1987), and 
Ohta, Hamabe, & Wakamatsu (1990). With this methodol-
ogy, we can quantify the amplitude and phase of the bar 
and spiral arm components in a galaxy image. This techni-
que is very useful for modeling the bar luminosity distribu-
tion for systems in which a simple elliptical bar model was 
not sufficient, as was first demonstrated by Elmegreen & 
Elmegreen (1985) in describing the bar luminosity distribu-
tions exhibited by galaxies over a range of Hubble types.

2.1. Pattern Classification Using Artificial Neural Networks

Artificial neural networks (ANNs) are systems of weight 
vectors whose component values are established through 
various machine learning algorithms. These systems receive 
information in the form of a vector (the input pattern) and 
produce as output a numerical pattern encoding a classifica-
tion. They were designed to simulate groups of biological 
neurons and their interconnections and to mimic the ability 
of such systems to learn and generalize. Discussion of the 
development and practical application of neural networks is 
well presented in the literature (see McClellan & Rumel-
hardt 1988). In astronomical applications, this technique 
has been applied with considerable success for the problem 
of star-galaxy separation by Odewahn et al. (1992). A neural 
network classifier was developed Storrie-Lombardi et al. 
(1992) to assign galaxy types on the basis of the photometric 
parameters supplied by the ESO-LV catalog of Lauberts & 
Valentijn (1989). Applications of this method to large sur-
veys of galaxies using photographic Schmidt plate material 
are discussed in Odewahn (1995) and Naim et al. (1995). 
More recent discussions of star-galaxy separation using 
ANNs are presented by Andreon et al. (2000), Mahonen & 
Frantti (2000), and Philip et al. (2000). An extremely thor-
ough discussion of galaxy classification with ANN methods 
is given by Bazell (2000).

The neural network literature quoted above contains a 
wealth of information on the theoretical development of 
networks and the various methods used to establish the 
network weight values. We summarize here only the 
practical aspects of the ANN operation and the basic equa-
tions for applying a feed-forward network. The input infor-
mation, in our case the principal components formed from the Fourier profiles discussed in § 2.3, is presented to the network through a set of input nodes, referred to as the input layer. Each input node is “connected” via an information pathway to nodes in a second layer referred to as the first hidden layer. In a repeating fashion, one may construct a network using any number of such hidden layers. For the back-propagation networks discussed here, we have used ANNs with two hidden layers. The final layer of node positions, referred to as the output layer, will contain the numerical output of the network encoding the classification. This system of nodes and interconnections is analogous to the system of neurons and synapses of the brain. Information is conveyed along each connection, processed at each node site, and passed along to nodes farther “upstream” in the network. This is referred to as a feed-forward network.

The information processing at each node site is performed by combining all input numerical information from upstream nodes $a_{jp}$ in a weighted average of the form

$$
\beta_i = \sum_j w_{ij} a_{pj} + b_i .
$$

The $j$ subscript refers to a summation of all nodes in the previous layer of nodes, and the $i$ subscript refers to the node position in the present layer, i.e., the node for which we are computing an output. Each time the ANN produces an output pattern, it has done so by processing information from the input layer. This input information is referred to as the input pattern, and we use the $p$ subscript to indicate input pattern. Hence, if there are five nodes in the previous layer ($j$), each node ($i$) in the current layer will contain five weight values ($w_{ij}$) and a constant term $b_i$, which is referred to as the bias. The final nodal output is computed via the activation function, which in the case of this work is a sigmoidal function of the form

$$
a_{pi} = \frac{1}{1 + \exp^{-\beta_i}} .
$$

Hence, the information passed from node $i$ in the current layer of consideration, when the network was presented with input pattern $p$, is denoted as $a_{pi}$. This numerical value is subsequently passed to the forward layer along the connection lines for further network processing. The activation values computed in the output layer form the numerical output of the ANN and serve to encode the classification for input pattern $p$.

In order to solve for the weight and bias values of equation (1) for all nodes, one requires a set of input patterns for which the correct classification is known. This set of examples is used in an iterative fashion to establish weight values using a gradient descent algorithm known as back-propagation. In brief, back-propagation training is performed by initially assigning random values to the $w_{ij}$ and $b_i$ terms in all nodes. Each time a training pattern is presented to the ANN, the activation for each node $a_{pi}$ is computed. After the output layer is computed, we go backward through the network and compute an error term that measures the change in the network output produced by an incremental change in the node weight values. Based on these error terms, all network weights are updated, and the process continues for each set of training patterns. As discussed in the next section, precautions must be taken in this process to avoid problems from overtraining (i.e., simply memorizing the input pattern set as opposed to generalizing the problem).

The goal of the present work is to develop a system that automatically detects the presence of specific types of morphological structures in galaxy images. This is clearly a more complex problem compared to the use of a few global photometric parameters to predict types, and so we desired to use an additional method of neural network classification to confirm the results of our classical back-propagation approach. Philip et al. (2000) develop a neural network based on Bayes’ principal that assumes the clustering of attribute values (the input parameters). The method considers the error produced by each training pattern presented to the network, and weights are updated on the basis of the Bayesian probability associated with each attribute for a given training pattern. In this approach, the probability density of identical attribute values flattens out, while attributes showing large differences from the mean get boosted in importance. This is referred to as a difference-boosting neural network (DBN). We have applied this new approach as a check on our back-propagation network results; however, this method has the added advantage that it trains to convergence in a much faster time than is required by classical back-propagation methods.

2.2. Initial Image Reconstruction

In this paper we develop a technique that is a preprocessing step for producing input to a supervised classifier in the form of an artificial neural network (Odehahn 1997). In this approach, the radial surface brightness profile of a galaxy is computed in elliptical annuli centered on the galaxy center. The rational here, valid in most cases, is that we are isolating the disk component of the galaxy and hence establishing a way of defining the equatorial plane of the system. The position, shape, and orientation of these annuli are determined using classical isophotal ellipse fitting techniques. Within each annulus, we compute the run of flux density with position angle $\theta$ in the equatorial plane of the galaxy to form the azimuthal surface brightness profile. To describe each azimuthal profile and reduce the dimensionality of our classification problem, each azimuthal profile is modeled by the following Fourier series:

$$I_0(r, \theta) = I_0 + \sum I_{mc}(r) \cos m \theta + \sum I_{mv}(r) \sin m \theta .$$

For computational efficiency, the Fourier terms are computed for each azimuthal profile using the following moment relations:

$$I_0(r, \theta) = \langle I(r, \theta) \rangle ,$$

$$I_{mc}(r) = 2 \langle I(r, \theta) \cos m \theta \rangle ,$$

$$I_{mv}(r) = 2 \langle I(r, \theta) \sin m \theta \rangle ,$$

where $\theta$ is the angle (in the equatorial plane of the galaxy) with respect to the photometric major axis and $m$ is an integer. Hence, if we assume galaxy disks to be thin and suffer no internal extinction, this approach crudely removes inclination effects and allows us to compare galaxy images independent of orientation. The Fourier series in this paper were computed up to $m = 5$. Fourier amplitudes, describing the relative amplitude of each component, are computed using
the following expression:

\[ A_m(r) = \sqrt{I_{mc}(r)^2 + I_{ms}(r)^2}/I_0(r) \, . \]

Finally, a phase angle describing the angular position of the peak signal contributed by each component can be computed. We compute the phase angle of the \( m \) component as

\[ \theta_0 = (1/m) \tan^{-1}(I_{ms}/I_{mc}) \, . \]

A practical application of this method to a barred spiral galaxy imaged with \( HST \) is shown in Figure 1. Another example for an elliptical galaxy image that contains no high spatial frequency image structure is shown in Figure 2. Through experimentation with both \( HST \)- and ground-based images, we have determined that using 17 elliptical annuli and up to \( m = 5 \) in equation (3) consistently reproduces the basic morphological features of most galaxy images. A considerable amount of experimentation went into determining how one should establish the radial bin intervals in this procedure. A series of tests were made using unequal linear bins that overlapped and logarithmic binning interval. We experimented with normalizing the bin radius by the effective radius \( r_e \) so that all galaxies could be compared on a more uniform spatial scale. This approach proved problematic since it produced profile scales that were highly compressed for early-type systems and highly expanded in late-type systems. In such cases, the inner regions of the late-type galaxies, where much of the important morphological structure is present, were underresolved. Additionally, we found that measurement of the effective radius of galaxies in low-resolution, low-S/N \( HST \) images can be systematically low and using \( r_e \) in this manner will introduce a substantial bias when comparing local and distant samples. In the end, we found that using equal size radial bins extending to an outer, low surface brightness optical isophote produces the most robust results for comparing morphological properties among galaxy images.

It should be noted that one can justify the use of higher order Fourier components for cases of high-resolution galaxy images. In experiments with some of the deep Palomar 60 inch telescope images described in \( \S \) 3, we found that the
bars of many galaxies have a strong contrast relative to the disk, and high spatial frequency components (up to 120) are required to reconstruct such a sharp feature. Incorporating so many spatial frequencies in a general image classifier is impractical, but one might hope to use some type of information about high-frequency structure: if the inclusion of such components in the model significantly improves the model fit, then this is important information. Perhaps a good number of barred galaxies can be identified if we simply search for images that require high-order Fourier components in the inner annuli to build a good fit model.

2.3. Dimensionality Reduction Using Radial Trends

The Fourier-reconstructed images described above are comprised of radially dependent sets of Fourier amplitudes and phase angles. Using 17 annuli, six Fourier amplitudes (including the \( m = 0 \) term), and the corresponding phase angles to parameterize each image, we thus describe each galaxy with a 221 element classification vector. Some human classifiers have insisted that at least 10,000 resolution elements are required in a galaxy image that is to be morphologically classified (G. de Vaucouleurs 1989, private communication). More liberally, we might consider that one should have at least a 50 \( \times \) 50 image of a galaxy, or 2500 pixels, to convey morphological information. Hence, using the Fourier description of a galaxy reduces the number of parameters describing the image information content by a factor of more than 10 compared to that used by a human classifier. An input classification vector with over 100 elements is rather large for most pattern classifiers trained in the presence of noise. To distill our image parameterization further, we have chosen to characterize the radial properties of each Fourier component. In other words, just as we often describe the radial surface brightness profile of a galaxy using several model parameters, so too can we describe the radial trends in each set of amplitude and phase estimates computed with equations (7) and (8).

The mean radial trends in both the amplitude and phase of the coefficients in equations (7) and (8) are shown in Figures 3 and 4 for three types of galaxies. In these figures the mean phase or amplitude of a given Fourier component is plotted on the y-axis as a function of elliptical annulus number on the x-axis. We divided a sample of 196 galaxies from § 3 into three family categories, A (unbarred system), AB (transitional systems), and B (barred systems), and computed the mean radial profiles for the amplitude and phase coefficients of the 2\( \theta \) and 4\( \theta \) coefficients. As noted by
Elmegreen & Elmegreen (1985) and Ohta et al. (1990), these are the dominant terms needed to reproduce the light distributions of most barred galaxies. In Figure 3 we see a clear trend among the $2\theta$ and $4\theta$ amplitudes: barred systems have significant power in the inner rings, AB systems have systematically lower power, and A galaxies have the lowest amount of $2\theta$ and $4\theta$ power in even the inner elliptical annuli. These mean profiles were computed with 71 A, 73 AB, and 61 B galaxies.

![Fig. 3](image)

Fig. 3.—Mean profiles of the $2\theta$ (left) and $4\theta$ (right) amplitudes computed for three sets of galaxies divided by family class. We used 71 A, 73 AB, and 61 B galaxies. The mean barred (B) galaxy trends are well separated from unbarred (A) galaxies for annuli numbers (on the x-axis) in the range 3–12. We plot the 1σ error bars for the mean A points only. It is significant that in both the $2\theta$ and $4\theta$ cases the mean AB profiles lie between the A and B curves, indicating that these galaxies do indeed possess only weakly detected bar structures.

![Fig. 4](image)

Fig. 4.—Mean profiles of the $2\theta$ (left) and $4\theta$ (right) phase computed for the same samples used in Fig. 3. Although amplitude information clearly provides a more powerful means of detecting the presence of a bar, phase information is needed to differentiate between a purely linear structure like a bar and a simple two-armed spiral pattern.
73 AB, and 61 B galaxies taken from the analysis of § 4. The 1 σ error of each point in the A system profiles are shown in Figures 3 and 4, and it is clear that a single point from the profile of one galaxy will have little discriminatory power in distinguishing bar class. However, if we combine the information of the points from annuli with ring numbers between three and 12, then we can expect to derive a robust estimate of whether a bar is present in the inner regions of a galaxy image. In this region of the reconstructed images, the mean Fourier profiles clearly delineate the presence of a bar. In Figure 4 we plot the corresponding mean profiles for the 2θ and 4θ phase angles. Although these profiles are less sensitive to bar presence, the phase terms will be important in distinguishing barred galaxies from purely spiral systems. In the case of a barred galaxy, the 2θ and 4θ phase angles remain fixed with ring number (radius), whereas with a two-armed spiral pattern, we expect to see a systematic variation of 2θ and 4θ phase angles with radius since the flux density peaks are changing their position angle smoothly as we progress outward in the radial annuli. As can be seen in Figure 4, we still observe a gradual progression in the trends of these phase angle profiles for the A, AB, and B galaxy sets. To summarize, the radial 2θ and 4θ phase angles are changing on a linear bar structure.

Following the innovative work of Han (1995), we performed a principal-component analysis (PCA) of each set of Fourier profiles for our collections of galaxies that had reliable stage and family classifications. For most Fourier coefficients, it was determined that 85%–90% of the profile variance could be described with only two or three principal components. Hence, using the eigenvectors determined from this analysis, we can characterize the Fourier profile of a random galaxy using only two or three parameters. Should we use the amplitude and phase of only the 2θ and 4θ components to characterize a galaxy, then we can parameterize the light distribution using four coefficients × two principal components = eight parameters. Such an image vector constitutes a very reasonable size for training a supervised classifier with the 100–200 patterns (galaxies) available in this work. In Figures 3 and 4 we have divided the galaxies by family class, but these Fourier-based principal components may be used to identify stage-related image traits also. In Figure 5 we demonstrate our principal-component method by showing PCA-based image parameter spaces that are symbol coded by stage and family class.

3. TRAINING AND TESTING SAMPLES

The development of a pattern classifier based on supervised learning, i.e., a classifier that works in a predefined system, requires a data sample consisting of the input classification information (the image parameters) and the corresponding target patterns (the galaxy types). This sample is referred to as the training set, and a portion of it is held back from the actual supervised learning process (in our case, back-propagation training) to be used as a test sample to make an unbiased judgement of classifier performance during and after training. As we discuss below, nearly all of the past morphology-based studies of high-redshift galaxies have focused on the stage axis only. Since our aim is to develop morphological classifiers that estimate the family and variety of galaxies based on Fourier amplitude strengths and phases, we have visually reclassified a large number of high-S/N, well-resolved HST galaxy images in the RHS to establish a statistically significant testing sample. Although the present sample is based on 35 medium-deep and deep Wide Field Planetary Camera 2 (WFPC2) fields, the number of galaxy images from this sample that are suitable for such classification is only 146. Hence, for the
present study we incorporated the local galaxy samples from ground-based CCD observations by de Jong & van der Kruit (1994) and Frei et al. (1996), adding 124 more galaxies to the sample. For both the local and distant HST samples, we intercompare our classification sets in order to objectively define a weighted mean training set whose systematic and accidental errors are well understood. From such a sample we may establish a fair “ground truth” sample of objects for use in training and testing pattern classifiers.

A large collection of visual type estimates in several bandpasses from WFPC2 deep fields were discussed in Odewahn et al. (1996) and Driver et al. (1995b). Similar visual classification sets have been assembled since that time for the Hubble flanking fields (Odewahn 1997) and the B-Band Parallel Survey (BBPS) fields (Cohen et al. 2001). To begin our assessment of classifier errors, we have analyzed a large number of past visual stage estimates made in the F814W bandpass images. We intercompared all of the stage estimates common to the five classifiers of Odewahn et al. (1996) and the Driver et al. (1995a) sample and derived linear transformations to convert all classifiers to a common system. The vast majority of classifications were from one classifier (1000 from S. C. Odewahn, hereafter referred to as S. C. O.). It was decided that the Odewahn classes, which match well the Third Reference Catalog of Bright Galaxies (de Vaucouleurs et al. 1991, hereafter RC3) mean type system, would define the HST classification system. Following Odewahn & de Vaucouleurs (1992), we then use a residual analysis to determine the accidental scatter associated with each classifier and hence a relative weight associated with each classification of an HST–observed galaxy. Finally, a catalog of 1306 galaxy types was assembled: 1000 with single S. C. O. estimates with a mean error of ±2.2 type steps and 306 weighted mean types (109 having five independent estimates) with mean error of ±1.2 type steps. The galaxies in this catalog have stage estimates from the F814W images and objective errors estimated for each type. These data will hereafter be referred to as the HST1 set.

The purpose of our current work is to develop automated techniques that identify morphological structures, and with the exception of the bright BBPS galaxy–types estimates of S. C. O., the HST1 types are comprised only of estimates of the stage axis in the RHS (i.e., E, S0, Sa, . . . , Im). Such data sets are fine for the type classifiers discussed in Odewahn et al. (1996); however, for more advanced classification experiments, we desired a set of faint HST–observed galaxies with classifications in the RHS: stage, family, and variety. To collect this set, we extracted 4534 galaxy images in F814W from 39 different WFPC2 fields. The majority of these images were contributed by moderate-depth WFPC2 parallel fields, but the deep 53W002 field and the Hubble Deep Field–North were also used. From this sample we established, purely through visual inspection, the group of 146 galaxy images having sufficient S/N and resolution to allow good morphological classification. As a second sample, we binned 125 ground-based CCD images of Frei et al. (1996) (in Gunn g) and de Jong & van der Kruit (1994) (in Johnson B) to a resolution similar to that of the HST images.

The manipulation of all galaxy images used in this work, whether for the purpose of human visual classification or automated image analysis, was performed with a variant of the MORPHO package discussed in Odewahn (1997). A new version of this galaxy morphological classification software system called LMORPHO, which is optimized for use under the Linux operating system, provides an image database manager that allows users to select galaxies for analysis based on morphological traits, photometric properties, or many other metadata properties (e.g., bandpass, image size, or pixel scale). Beyond the image database manager of LMORPHO, the system provides tools performing automated galaxy surface photometry, basic astronomical image processing, and a variety of multivariate statistical analysis and pattern classification tasks.

Our galaxy image libraries were independently inspected by three classifiers using an LMORPHO graphical galaxy classification tool. The three classifiers were S. C. O., R. A. Windhorst (hereafter R. A. W.), and S. H. Cohen (hereafter S. H. C.), and each visually estimated the stage, family, and variety for each of the 272 galaxy images in the adopted sample. For a variety of reasons (insufficient S/N or resolution, high inclination, etc.), certain aspects of the classification (usually the family and variety) were uncertain, and these were designated as unknown. Following the method used for the HST1 sample, we next intercompared these type catalogs in order to establish the systematic and accidental errors associated with each set. The results of this are summarized in Table 1 and Figure 6. For the local galaxy samples, we compared each classifier to the weighted mean types in the RHS contained in the RC3 and discussed in Buta et al. (1994). For the distant sample, we compared each classifier’s types with the weighted mean types of HST1. Although many of the HST1 types are indeed weighted means of several independent estimates, the majority are single estimates from S. C. O., and hence we should regard the HST1 set as an independent (S. C. O.–dominated) unit weight catalog.

For both the local and distant galaxy samples, a preliminary linear regression analysis was performed to determine transformation relations for converting to the RC3 and HST1 systems, respectively. An impartial regression analysis, using each catalog as the dependent variable, was used to determine the mean relationship between all possible catalog pairs. For each subsample (local and distant), small-scale and zero-point shifts were applied to the S. C. O., R. A. W., and S. H. C. stage estimates to bring all classifications onto a uniform system. It is informative to note that the size and sense of the scale corrections were uniform by classifier irrespective of local or distant sample: S. C. O. and R. A. W. systematically classified galaxies later at the 8% level, and S. H. C. systematically classified galaxies earlier at the 6% level. These scale adjustments were statistically significant at the 2 σ level. The zero-point shifts, statistically significant at the 3 σ level, were general less than 1 step on the 16 step stage axis of the RHS (with the exception of a −2.2 step correction for the distant R. A. W. sample). In each case (local and distant) mean transformation equations were derived for the S. C. O., R. A. W., and S. H. C. classification sets. Following these small, but statistically justified, transformations, a second correlation analysis was performed to verify system uniformity and compute the mean variance associated with each catalog comparison. This analysis is illustrated in the top panel of Figure 6 for the local sample and in the bottom panel for the distant (HST) sample. With all stage estimates transformed to a common mean system, we followed the methodology of Odewahn & de Vaucouleurs (1992) to compute the mean standard devia-
tion of the 2 × 2 residuals in each comparison (these values are plotted in each panel of Fig. 6). These estimates were combined via a system of linear equations to calculate the standard deviation associated with each individual catalog. The results, summarized in Table 2, first show that each classifier (S. C. O., R. A. W., and S. H. C.) is able to produce stage estimates for the local galaxies in the RC3-defined system with scatter of between 1 and 2 steps on the RHS. This is quite consistent with a similar study by Naim et al. (1995) using 831 galaxies typed from Schmidt images by six expert classifiers. As expected, the rms scatter estimated for the distant galaxy samples is somewhat larger, but still in the 1.5–2 step range.

Using the systematic corrections and sample weights derived for the HST classification sets, we have compiled a catalog of weighted mean–revised Hubble types for our distant galaxy samples. In this way we produce not only a final catalog of higher quality types compared to that from any

![Fig. 6.—Correlations between the revised Hubble-type stage estimates made by different classifiers. Small-scale and zero-point adjustments were made to each set in order to bring all catalogs onto the system defined by RC3 stage estimates. Using the method discussed in Odewahn & de Vaucouleurs (1992), we derived the accidental error associated with each classifier.](https://example.com/fig_6.png)
single classifier, but we are able to derive error estimates for each galaxy. This last point is important from the point of view of assessing the quality of any automated morphological classifier. Galaxies whose morphological properties are uncertain to a number of human classifiers are most likely peculiar in some aspect and hence should not be included in the training or testing of a generic machine classifier. Weighted mean types and errors for our local galaxy sample were adopted from the RC3. Representative galaxies from each of these samples, broken into stage and family groups, are shown in Figure 7.

3.1. A High Weight Set for Family and Variety Classification

Although agreement between stage classifications was found to be quite satisfactory for developing adequate test/train samples for automatic stage classifiers, the uniformity of the family (barred vs. unbarred) and the variety (ringed vs. nonringed) of classes was less than satisfactory. From the HST image samples, all three classifiers agreed on the family assignment only 30% of the time and on the variety assignment only 25% of the time. This is clearly due to a lack of the S/N and resolution required for such morphological classification in galaxy images. In reality, some of the ground-based images also lacked sufficient S/N to allow for the unambiguous assignment of family and variety assignment. Hence, for the purposes of developing a robust morphological classifier capable of dealing with these image properties, we chose to develop a third, higher weight set of galaxy images. During the course of a large program to acquire photometric calibration fields for the Digital Palomar Sky Survey (Gal et al. 2000), one of us (S. C. O.) initiated a program for nonphotometric conditions of imaging large, bright, well-classified galaxies from the list of “best-classified” galaxies in Buta et al. (1994). These galaxies had classifications agreed on by several human experts and, in fact, represent prototypes of the various Hubble stages. Each galaxy was imaged in Gunn g or Cousins B for 10–15 minutes with the Palomar 60 inch telescope using CCD 13. Conditions were generally not photometric, but thin cirrus and seeing no worse than 2.5 FWHM were tolerated for this program. When possible, multiple images of the same galaxy were obtained so that we might characterize how accidental errors in the surface photometry will affect output from the automated morphological classifier. In Table 3 we summarize 30 images in Gunn g and Cousins B of 20 different galaxies selected for experimenting with automated family and variety classification.

4. PRACTICAL APPLICATION OF THE METHOD

Here we describe the application of the method to a set of HST archival images. Using the results of §2, we assembled samples of local and distant galaxies having weighted mean stage estimates and family classifications that at least two human classifiers agreed on. We show a representative sample of these galaxies in Figure 7. As we shall discuss below, the criteria for building a family classifier training set had to be somewhat liberal in order to compile a usable experimental sample. For stage classification, a series of back-propagation networks using different input parameter vectors and network layer architectures were experimented with. In the case of the family classifier, we used two different types of ANN classifiers: a back-propagation–trained feed-forward network and a DBNN developed by Philip et al. (2000).

4.1. Stage Classification

A set of 262 images was collected using the sample of galaxies having $b/a > 0.4$ and weighted mean stage estimates from the analysis of §3. Most galaxy classification systems provide criteria for estimating the morphological type of an edge-on system; however, because we are concerned here with the analysis of two-dimensional morphological structures, we chose to exclude such high-inclination systems. An LMPHORO task was used to perform automated surface photometry of each image, and then an interactive mosaic viewer was used to inspect the ellipse fit to the low surface brightness isophote in each galaxy postage stamp image. This was done to verify that no improper image processing errors occurred due to confusion from nearby images or other image defects. Ten galaxy stamps were rejected in this step, the majority of these being relatively faint HST-observed galaxies that extended too close to the edge of the WFPC2 field. The human classifiers dealt well with “masking” such edge effects when the visual classification of §3 was performed; however, the sky fitting and surface brightness contouring routines of LMPHORO were unable to properly handle such situations. Future automated morphological surveys will, of course, have to deal more effectively with this type of error, but for the purpose of this work, we chose to simply delete these galaxies from any machine classification experiments.

For each of the 252 postage stamp images of well-classified galaxies remaining in our experimental sample, the full set of morphologically dependent Fourier image parameters discussed in §3 were computed with LMPHORO. A series of back-propagation neural networks were trained following the procedures outlined in Odewahn (1997) using three different types of input vector sets and six different network architectures. In every training case, two hidden layers of the stage axis of the RHS. In practice, a final type was assigned for each input vector using the weighted mean value of the output node value. This allows one to assign an eight node output layer, one node for each two-step interval of the stage axis of the RHS. In practice, a final type was assigned for each input vector using the weighted mean value of the output node value. This allows one to assign not only the highest weight classification, but the rms scatter in this node-weighted mean value can be used to assign a classification confidence. Hence, while any output pattern will result in the assignment of an estimated type, patterns where most of the output signal is carried in one major node

| Table 2: Accidental Stage Errors by Classifier |
|-----------------------------------------------|
| Sample | RC3   | HST1  | S. C. O. | R. A. W. | S. H. C. |
|--------|-------|-------|----------|----------|----------|
| Local... | 0.97 ± 0.35 | ...   | 1.26 ± 0.26 | 0.88 ± 0.39 | 2.02 ± 0.16 |
| Distant... | ...   | 1.53 ± 0.13 | 1.20 ± 0.13 | 2.16 ± 0.09 | 2.48 ± 0.08 |
will have a high-confidence assignment, and patterns where the output pattern is distributed over many nodes will be assigned low confidence.

The image parameter catalogs were collected into a single binary catalog that can be used in LMORPHO to interactively generate symbol-coded parameter space plots like those shown in Figures 8 and 9. With this package the user can quickly view many different parameter spaces and judge which one shows the largest degree of type separation (whether stage or family) by determining how well the different type-coding symbols are separated in the plot. An important feature in this graphical tool is that any point
may be marked and used for a variety of uses. In one mode
the user can choose to view on a real-time basis the galaxy
image of any selected data point. In this way one is able to
determine if a particular outlier is due to improper image
processing or some truly unique morphological circum-
stance. In the latter case, such data are retained for use in
classifier training. In the former, these patterns would be
rejected from use in gathering training and testing samples.

Since the images used in this particular series of classifica-
tion experiments were generally of high quality, fewer than
5% of the original 252 postage stamps were rejected in this
manner. For only a few hundred galaxy images (or even less
than a few thousand), this process of parameter space
inspection is trivial to complete in a few minutes. Finally,
we made a preliminary set of ANN training runs using dif-
ferent input parameter sets as well as subdividing the train-
ing/testing data in different subsets. Through this process
we uncovered an additional 28 objects that consistently gave
highly discrepant training results (i.e., the image parameter
combinations for these sources were highly abnormal com-
pared to the bulk of the data), and these sources were
rejected from further experiments. This represented about
9% of our original training set, but determining the source
of peculiarity for these galaxies will require a larger, more
diverse collection of images for future experiments. We
determined a number of useful input parameter combina-
tions that showed good segregation by morphological type,
and a series of different image parameter set patterns was
formulated. Each type of pattern would form the input layer
to an ANN classifier. For clarity, we assigned a running
integer value to each type of input, and we refer to this as
the feature set number. We indicate the feature set number
for each training exercise in column (8) of Table 6.

In Table 4 we summarize the LMORPHO image parame-
ters that were selected for the stage and family classifiers
developed in this paper. Since different image parameters
are used in each classifier, we list in column (2) the feature
set number for each input pattern using each image
parameter.

| Name             | RC3 Type | Filters |
|------------------|----------|---------|
| NGC 1744......... | SB(s)d   | 1g      |
| NGC 4501......... | SA(rs)b  | 2B      |
| NGC 4519......... | SB(rs)d  | 1g      |
| NGC 4618......... | SBm      | 2g      |
| NGC 5850......... | SB(r)b   | 1g, 1B  |
| NGC 5985......... | SA(r)b   | 1g      |
| NGC 5377......... | SB(s)a   | 1g      |
| NGC 6300......... | SA(s)c   | 3g, 1B  |
| NGC 6340......... | SA(s)b   | 1g      |
| NGC 6384......... | SB(r)bc  | 1g, 2B  |
| NGC 6643......... | SA(rs)c  | 1g      |
| NGC 6764......... | SB(r)bc  | 1g      |
| NGC 6824......... | SA(sb)   | 1g      |
| NGC 6951......... | SAB(rs)bc| 2g      |
| NGC 7217......... | SA(r)ab  | 2g      |
| NGC 7457......... | SA(rs)0  | 1g      |
| NGC 7479......... | SB(c)    | 1g      |
| NGC 7741......... | SB(r)cd  | 1g      |
| PGC 61512........ | SB(s)dm  | 1g      |

Fig. 8.—First principal components formed from an analysis of the 2θ amplitude and normalized flux profiles from 246 local and distant galaxies. We see
a clear segregation by revised Hubble stage (normally referred to as the Hubble type). A variety of independent and similarly fuzzy Fourier-based image
parameters may be combined via the ANN classifiers described in the text to predict stage values.
The selection of training and testing samples was carried out in the same manner for all input vector types. As summarized in Table 5, all galaxies were binned into eight type bins corresponding to the eight nodes of the ANN output layer. Each such bin was divided into two sets: one for training and one for testing. One is always tempted to use more galaxies in the training sample so that the network has a larger number of patterns for generalizing the problem. One danger in back-propagation training is overtraining: beyond some number of training iterations, the network weights could be adjusted so as to simply memorize the training pattern sets as opposed to generalizing the mapping of input vectors to output classes. To prevent this, one should use an independent test data set to judge classifier performance. The back-propagation code in LMORPHO computes classifier statistics for both a training sample and a testing sample at each user-specified iteration in which the updated ANN weights are stored. In a posttraining phase, each set of statistics is inspected by the code to determine the optimal training cycle. The algorithm used to isolate this iteration uses the slope and rms scatter of the correlation between the target stage prediction and the ANN stage prediction. Both the training and testing data sets are used. The case of overtraining is detected by a progressive drop in the rms scatter for the training data with a flat or even increasing rms scatter for the test sample. An optimal training cycle is selected that minimizes the rms scatter for training data before overtraining occurs. Additionally, the slope of the linear regression must be close to unity within some user-specified tolerance, usually set at 20%.

We summarize the results of our back-propagation training to develop Fourier-based ANN classifiers in Table 6. The slope ($\alpha$) and scatter ($\sigma$) about a linear regression between the ANN-predicted and the targeted mean Hubble stage are tabulated for both the training and testing data sets for the selected optimal training epoch. Additionally, as with the type comparison analysis in § 3, we compute a scatter ($\sigma_2$) based on direct differences between the ANN and target values following Odewahn & de Vaucouleurs (1992). This statistic reflects more honestly the scatter to be expected for the user of such an ANN-derived catalog. As expected, scatter among the training data is systematically lower than for the test data since the error function being minimized in the back-propagation training is formed with the training data. However, the scatter derived for the test data is usually about 2 steps on the RHS stage axis and hence very reasonable for most scientific pursuits with morphological data.

In Figure 10 we compare mean neural network classifier types from three different ANNs (marked with asterisks in Table 6) to weighted mean visual types for two sets of data: the training data and the testing data. To summarize, the test data, comprised of sets of input classification parameters and their corresponding target types, were never presented to the ANN classifiers during back-propagation training to establish the network weight values. As such, this data sample represents a fair comparison by which we can judge true network performance. We include in Figure 10 the same correlation for the training sample. In summary, the Fourier-based stage classifiers developed in this experiment clearly perform as well as a human classifier. These results are certainly encouraging if our goal is simply to estimate revised Hubble types for a large number of digital images, but a more important point must be stressed. We have shown that our method of parameterizing a galaxy image preserves the information content needed to emulate the process used by a human classifier. As with the bar classification work described below, our ultimate goal will be to use such a digital image analysis to search for more direct correlations among galaxy properties and ultimately

**Table 4**

| Name       | Feature Set | Comments                  |
|------------|-------------|---------------------------|
| PC1_Fnorm | 1, 2, 3     | First PC of normalized flux profile |
| PC2_Fnorm | 1, 2, 3     | Second PC of normalized flux profile |
| PC1_1TA   | 2, 3        | First PC of 10 amplitude profile |
| PC1_2TA   | 1, 2, 3, 4  | First PC of 20 amplitude profile |
| PC2_2TA   | 1           | Second PC of 20 amplitude profile |
| PC1_4TA   | 1, 2        | First PC of 40 amplitude profile |
| PC2_4TA   | 1, 2, 4     | Second PC of 40 amplitude profile |
| PC1_4TP   | 1, 2        | First PC of 40 phase profile |
| PC2_4TP   | 1, 2, 4     | Second PC of 40 phase profile |
| S/N        | 3           | Mean S/N measured in elliptical aperture |
| N_ip     | 3           | Number of image pixels |
| BR         | 4           | Bar parameter 3 |
| BF         | 4           | Bar parameter 4 |

**Table 5**

| Node Number | Number | Interval | Literal Range |
|-------------|--------|----------|---------------|
| 1           | 19     | -8.0, -4.5 | E to E0–1   |
| 2           | 6      | -4.5, -2.5 | E + to L0–   |
| 3           | 13     | -2.0, -0.5 | L0 to L0+   |
| 4           | 18     | -0.5, 1.5  | S0a to Sa   |
| 5           | 34     | 1.5, 3.5   | Sab to Sb   |
| 6           | 91     | 3.5, 5.5   | Sbc to Sc   |
| 7           | 30     | 5.5, 7.5   | Scd to Sd   |
| 8           | 12     | 7.5, 10.5  | Sdm to Im   |
develop a more physically meaningful classification system beyond the RHS.

4.2. Family Classification

As discussed in §3, the visual classifiers divided the family estimates into the three bins of the RHS: A, AB, and B. The difference between each class is somewhat vague and subject to personal bias. It is not surprising that we found very few cases where all three classifier agree that a galaxy was of the AB family class. As was discussed for Figure 3, the Fourier profiles for AB galaxies are, in the mean, intermediate between the A and B galaxies. However, the AB galaxies show a markedly large variance, which is probably due to the wide variance in AB classification criteria used by the

| $\alpha_{TR}$ | $\sigma_{TR}$ | $\alpha_{TR}^2$ | $\alpha_{TE}$ | $\sigma_{TE}$ | Epoch | Feature Set | $I$ |
|--------------|--------------|----------------|--------------|--------------|-------|-------------|----|
| 0.920        | 0.93         | 2.09           | 0.908        | 2.25         | 2.09  | 5800        | 1  |
| 0.953        | 1.15         | 1.71           | 0.968        | 2.29         | 1.71  | 2900        | 1  |
| 1.045        | 1.16         | 1.68           | 0.949        | 2.46         | 1.68  | 2600        | 1  |
| 0.935        | 1.20         | 1.84           | 0.959        | 2.37         | 1.84  | 2000        | 1  |
| 0.705        | 1.71         | 2.57           | 0.722        | 1.92         | 2.57  | 1000        | 1  |
| 0.931        | 1.78         | 2.28           | 0.906        | 1.92         | 2.28  | 4200        | 2  |
| 0.974        | 1.35         | 1.82           | 0.909        | 2.06         | 1.82  | 4300        | 2  |
| 0.936        | 1.56         | 2.25           | 0.914        | 2.06         | 2.25  | 2300        | 2  |
| 0.938        | 1.13         | 1.87           | 0.986        | 2.51         | 1.87  | 3900        | 2  |
| 1.081        | 1.32         | 1.74           | 1.005        | 1.88         | 1.74  | 3600        | 2  |
| 1.009        | 1.44         | 1.71           | 0.978        | 1.71         | 1.71  | 1800        | 3  |
| 0.878        | 1.16         | 2.59           | 0.803        | 1.77         | 2.59  | 1000        | 3  |
| 1.016        | 0.93         | 1.54           | 0.946        | 1.83         | 1.54  | 900         | 3  |
| 0.917        | 1.36         | 2.07           | 0.965        | 1.62         | 2.07  | 500         | 3  |
| 1.025        | 0.79         | 1.38           | 1.045        | 1.79         | 1.38  | 900         | 3  |

Note.—Asterisks indicate networks that gave the best performance for a given node architecture, and these were used to compute mean types in Fig. 10. The TR superscript refers to training results, and TE refers to testing results. The $I$ column indicates the number of nodes in each of the two hidden layers used in every ANN. As explained in the text, the feature set number indicates the set of image parameters fed to each network.

Fig. 10.—Left: Mean stage estimates from four independently trained ANN classifiers as a function of target type as given by the catalog of weighted mean visual stage estimates. As discussed in the text, each ANN classifier is fed a combination of principal components based on an eigenvector analysis of Fourier amplitude profiles. The linear regression indicated by the solid line departs systematically from the dashed unity line but gives a reasonable $y$-axis residual scatter of 1.6 steps on the 16 step revised Hubble stage axis. Right: Same correlation with the exception that we plot only data from galaxies used as input patterns during back-propagation training to establish ANN weight values. In both cases, the small box symbols are data rejected after two cycles of 3 $\sigma$ rejection in the linear regression fit.
human classifiers. The presence of such uncertainty made it impossible to train an effective three-division stage classifier. For the present work, we approached the problem using an unambiguous set of A and B galaxies and attempted to develop a robust classifier for identifying them using the Fourier-based eigenvector approach discussed in §3. When larger samples of morphologically classifiable galaxy images are collected, it is anticipated that a more thorough study of bar strength parameters will allow the development of a more sophisticated automated family classifier.

For the present set of classification experiments, we gathered images of 71 A and 61 B galaxies drawn from the analysis of §3. In this case, any galaxy that had family classifications that were agreed upon by at least two of the three human classifiers were adopted for use. To clarify the problem, we used only A and B classes for classifier experiments. After a series of parameter space inspections, such as those described in the stage classification section, we chose to use the first principal components of the 20 and 40 amplitude profiles. We found that additional discrimination was added in many cases using the first principal component of the 40 phase profile and adopted this for use in the classifier input vector. It was found that including parameters that are highly correlated with bulge-to-disk ratio (i.e., the principal components of the normalized flux profile) did not add significantly to detecting the presence of a bar. This was surprising in that it is well known that bar shapes and properties are correlated with Hubble stage and, hence, bulge-to-disk ratio (Elmegreen & Elmegreen 1985). In future work with much larger samples of galaxies, the use of terms related to the shape of the surface brightness profile may prove useful for this reason: pattern classifiers can better interpret the meaning of morphological shape parameters like the first principal components of the 20 and 40 amplitude profiles if some information related to Hubble stage is provided. For the present small sample of potentially useful training samples, we decided to exclude the use of surface brightness profile shape parameters for the training of a family classifier.

We added two simple global image parameters that contribute information about bar presence that were independent of the Fourier model approach. In determining the shape and orientation of the optimal elliptical aperture for measuring each galaxy, LIMORPHO computes a series of ellipse fits to progressively fainter surface brightness levels. The initial surface brightness levels are high and generally sample the bright inner region of a galaxy where the bars are found. It was found that simple parameters using these elliptical isophote parameters could be formed that gave useful information about the presence of a bar that would be independent of the image models derived with the Fourier method. We define the minimum axis ratio of the series of ellipse fits to be $BP_1$ and the axis ratio of the largest ellipse (i.e., the fit to the lowest surface brightness level) to be $BP_2$. In general, a barred galaxy will have $BP_1 < BP_2$ but with the condition that the major axis length of the isophote measured for $BP_1$ is significantly smaller than that measured at the isophote for $BP_2$. To discriminate this condition, we formed the parameter $BP_3 = r_1/r_2$, where $r_1$ is the semimajor axis length measured at $BP_1$ and $r_2$ is the semimajor axis length measured at $BP_2$. Additionally, we formed $BP_4 = BP_2/BP_1$ to measure the contrast of the ellipticity variation in the galaxy image. It was found that for many galaxies, a plot with $BP_4$ on the y-axis and $BP_3$ on the x-axis places B galaxies in the upper left-hand corner and A galaxies in the lower right-hand corner. Of course, there is a rather substantial overlapping zone in the low-$BP_4$ and low-$BP_3$ region of this plot. Hence, while such a simple scheme cannot discriminate all A, B galaxies, it can provide high weight information for some systems. Simple initial ANN classification experiments provided evidence that these parameters, when added to the Fourier-based parameters, produced better discrimination for A, B samples, and hence we included them in the training of our final family classifier.

Two types of family classifiers were trained with the input vector parameters shown in Table 4. First, the back-propagation network code discussed in §4.1 was used to train a network consisting of two hidden layers having 12 nodes each. This network mapped an eight-element input vector to a two-node output layer. Patterns with more power produced in output node 1 were classified as A, and patterns with more power produced in output node 2 were classified as B. Many of the dimensions in this eight-dimensional classification space produce a good split between the A and B populations, and hence the back-propagation training converged relatively quickly. As with the stage classifiers, a sample of pure test patterns was retained and used to assess network performance at each stage in the back-propagation weight update process. This procedure gives us a truly independent check on the classifier performance and an effective means of guarding against overtraining. A total of 30 training epochs were used in the back-propagation training, and epoch 15 was selected as providing optimal performance with a success rate performance of 92% among training patterns and 85% among test patterns. In a second experiment the DBNN method of Philip et al. (2000) was applied to the identical training and testing patterns as for the back-propagation ANN. After boosting, the training set was found to produce a 94.1% successful classification rate and the test set was found to produce a success rate of 87.5%. The performance of the DBNN classifier was found to be marginally better than the back-propagation–trained network; however, a large sample of training data will be needed to assess if this improvement is significant. Nevertheless, both methods produced automated family classifiers that are able to discriminate bar presence with a roughly 90% probability of success.

### 4.3. Systematic Effects with S/N and Resolution

Image quality plays a crucial role in determining the extent to which morphological classification can be performed on a galaxy. A gradual decrease in image resolution and S/N will systematically degrade the quality of morphological classifications, both by human and machine methods. With stage estimation, one might expect that we lose the ability to differentiate the bulge and disk components, making it extremely difficult to work at the early stage of the Hubble sequence (i.e., differentiate E, S0, and early spiral systems). At the late Hubble stages, we lose many of the low surface brightness features in the outer disk regions that are crucial for differentiating among the Sd, Sm, and Im systems. With family classification, one can expect that gradual image degradation will make it impossible to detect the presence of a bar, much less characterize the properties of the bar (see van den Bergh, Cohen, & Crabbe 2001). In the
past, human classifiers, in particular those dealing with low-resolution and low-\(S/N\) Schmidt plate images (Nilson 1973; Corwin, de Vacouleurs, & de Vacouleurs 1985), have generally taken such effects into account by dropping the amount of detail in the literal type assigned to a galaxy. In other words, a system might be classified simply as \(S\) rather than \(SBC\) if the image quality is sufficient only to differentiate the difference between early-, mid-, or late-type galaxy. Such effects must be accounted for in any automated system geared toward the recognition of morphological features. One must know if there is a strong systematic bias toward, for instance, missing bar or spiral features as resolution decreases. Perhaps if such effects are well understood, we may hope to correct large statistical samples in some systematic way. At the very least, we must quantify the levels of low \(S/N\) or resolution that can be tolerated before the quality of automated classifications falls below some minimum tolerance required by a given science goal.

We chose to define resolution as the number of pixels in a galaxy image contained within the isophotal ellipse used to integrate the isophotal magnitude in LMORPHO. This is a reasonable approach for \(HST\) images where, for most filters, each pixel approaches the resolution of the optical image. For ground-based images, we should use the size of the seeing disk, which is generally several times larger than the pixel size. In our present study all of the ground-based images have been block averaged to produce pixel sizes that are comparable to the seeing disk, and hence we can use this uniform resolution definition for all of the galaxies analyzed. As for defining the \(S/N\) in a galaxy image, we used two approaches. In the first, we simply compute the ratio of the mean signal (above sky) per pixel to the standard deviation of the local sky measure. In the second approach, we define the mean signal \(S\) to be the zeroth-order term in the Fourier series fitted to the azimuthal distribution in each model annulus, and the noise \(N\) is the rms scatter about that fit. This latter approach incorporates the Poisson noise associated with the detected galaxy signal and accounts for large-scale structural changes across the azimuthal profile. In practice, the two methods had very different ranges but were found to be well correlated. The method 1 approach yielded \(S/N\) estimates in the range 100–1000, and the Fourier-based method produced values in the range of 3–20. It was found that the early-type galaxies, having little large-scale structure in their images, produced the tightest correlation between these two types of \(S/N\) estimators. For practical reasons, we used the more easily computed method 1 values in the image classification experiments described in this section.

To characterize how decreased \(S/N\) and resolution will affect the quality of our Fourier-based ANN classifiers, we selected sets of galaxies having extremely high \(S/N\) and resolution. These systems were found to be classified correctly by the stage and family classifiers discussed in the previous sections. Most of these images were taken from the ground-based data sets obtained with the Palomar 60 inch telescope discussed in \(\S\) 3, but a few of the highest resolution \(HST\) images were also included. For samples having well-classified stage and family types, we selected galaxies with images that had at least 2000 pixels contained within the optimal elliptical aperture. These sets of images were then systematically resampled using an LMORPHO package designed for this experiment to produce postage stamp images of galaxies having progressively lower \(S/N\) and resolution. Since this exercise was designed to study the general systematic effects of image degradation on morphological classification, no attempt was made to model the noise properties of a given detector, like WFPC2 or the Space Telescope Imaging Spectrograph (as in Odewahn, Burstein, & Windhorst 1997). For each resampled image, an extra component of Gaussian noise was added to the original image. We modeled decreasing resolution in two ways. First, we consider the case of high sample rate with image blurring, the case typically encountered with ground-based observations of nearby galaxies. In the second, and more relevant case, we considered a low sample rate with high optical resolution; i.e., the image suffers from a high degree of pixelation. To simulate this effect, we simply block averaged the original galaxy images to a lower image sampling rate. This latter effect is well known to anyone who has classified large numbers of distant galaxies on WFPC2 images, and is the one, as we will show, that dominates our ability to determine distant galaxy morphological classification.

The results for the back-propagation ANN stage classifiers are summarized in Figure 11. For this experiment, a simple success rate calculation was not appropriate. We desire to know not only whether a classifier fails to predict the correct stage estimate within some number of type bins but in which direction a misclassification occurs. If the classifier consistently pushed stage estimates in an early (toward \(E, S0\)) or a late (toward \(Im\)) direction, then a substantial bias will be introduced in the scientific interpretation of the ANN-generated morphological catalog. Hence, we chose to compute the trend in type stage residuals, computed in the sense of \(T('target') - T('ANN')\), as a function of resolution and \(S/N\), where \(T('target')\) is the true galaxy type. These trends were computed using overlapping bins in resolution or \(S/N\), with the mean stage residual and the rms about that mean computed in each bin. A set of 1901 resampled images were processed with the ANN classifiers discussed in \(\S\) 4.2 to estimate Hubble stage. In Figure 11 we show trends in stage residuals for three sets of target stage intervals: \(E-S0\), \(Sa-SBc\), and \(Sc-Im\). In each case we have used overlapping \(x\)-axis bins of width 0.2 dex, and we expect to observe that the degradation in classifier performance as image quality is lowered. Spiral and irregular stages show a moderate trend in positive mean offset (i.e., the ANN classifies these galaxies with an earlier value) and increased rms scatter about the mean with decreasing resolution and \(S/N\). Of more significance, \(E+SO\) galaxies exhibit a mean negative offset (i.e., the ANN classifies these systems to be later than the target value) of 2–3 Hubble steps. This occurs for even the highest resolution and highest \(S/N\) galaxy images in our present sample. Additionally, this mean offset and increasing rms scatter in \(T\) types for the \(E+SO\) galaxies is clearly steeper than for spiral and irregular galaxies. A variety of network architectures and input parameter vectors were experimented with in order to correct this problem. Little improvement was obtained because of, in large part, the lack of a sufficiently large sample of images for such galaxies. As discussed in Cohen et al. (2001), \(E+SO\) galaxies in the field are relatively rare for \(I_{606} \leq 22\), and hence a very large number of \(HST\) parallel observations are needed to build up a significant image library. Work is currently underway to collect such images from the many archival \(HST\) observations of moderate-redshift clusters. Clusters contain not only a large number of galaxies in a single WFPC2 field, but the morphological fractions of these
galaxies are skewed to the early types. The analysis of such large samples will potentially eliminate the mean negative stage offsets observed for the E + S0 samples of Figure 11; however, it is doubtful that we can hope for a comparable decrease in the rms of such stage classifications in the low-S/N and low-resolution regimes. It is clear that images of sufficiently high quality are needed to differentiate many of the subtle structural properties of early-type galaxies, as is clear from the work of Im et al. (2002). The larger field size, higher quantum efficiency, and especially increased resolution of the Advanced Camera for Surveys, compared to WFPC2, will make this the instrument of choice on HST for identifying large samples of moderate-redshift E + S0 systems in the near future.

The results for the back-propagation ANN family classifier are summarized in Figure 12. Unlike the stage classifiers, our family classifiers predict only two states: A or B. Hence, we are able to characterize the output results as success or failure cases. We therefore chose to plot the distribution of S/N and resolution for ANN-predicted classes for a sample of barred galaxies that agree with the target value and for those that fail to agree. As expected, in Figure 12 the success cases cluster in the high-S/N, high-resolution area of the plot, and the failure cases cluster most heavily in the low-S/N and low-resolution region. For clarity, we also plot the success rate trends binned by S/N and resolution. For identifying the presence of a bar, it is quite clear from these mean relations that some critical image resolution is required. For our present samples and Fourier-based ANN classifiers, we are unable to effectively identify barred systems at better than a 70% success rate until we obtain galaxy images with at least 1000 resolution elements ($\log N_{\text{pixels}} \gtrsim 3$, where $\log N_{\text{pixels}}$ is the number of image pixels above the isophotal threshold). The trend with S/N is less steep, but a mean S/N (per pixel) of around 1000 is also required to reliably identify bars. The same experiment was carried out for a sample of unbarred (A) galaxies with different, but expected, results. In this case, galaxies that are successfully classified as A at high S/N and resolution generally retain that classification at low S/N and resolution. In other words, the ANN classifiers rarely turn A galaxies into B galaxies, as is true for human classifiers. In general, as we lose image resolution, we begin to lose the ability to identify morphological bars, but we do not tend to contaminate B galaxy samples with misclassified A galaxies, an important point when considering the frequency of barred systems at high redshift.

5. SUMMARY

In this work we have developed a methodology for classifying galaxies based on the strength of Fourier components.
in the azimuthal light distribution. We have used pattern recognition techniques to develop machine classifiers that map the Fourier mode information to the revised Hubble system, a well-known galaxy classification system. We must stress that the basic motivation here is not simply to create a machine-based replication of the revised Hubble system. The RHS was developed to describe a variety of properties, some of which are directly related with two-dimensional morphological structures in galaxies. We therefore have chosen to use this classification system to guide our development of new quantitative classification systems. We have demonstrated here a method of extracting galaxy image information in a way that permits us to automate the recognition of bars and Hubble type, as defined by the family and stage axes of the RHS.

In a sense, this work represents a proof-of-concept study. The current method is able to recognize many of the same structural properties of galaxies that have guided the study of galaxy morphology for the last 50 years. Having established how well the method works, under a variety of image S/N and resolution regimes, we can entertain the prospect of searching for refinements in the classification system itself. In other words, the next step is to apply various multivariate statistical techniques and unsupervised learning methods to determine if there are relationships among the Fourier-based model image parameters that convey more directly relationships among galaxies. For instance, strength of the $2\, h$ and $4\, h$ Fourier components in the inner regions of disk galaxies might be used to form a continuous measure of bar strength, and we might dispense with the three-cell family axis of the RHS altogether. Continuity along each axis of RHS was recognized long ago (de Vaucouleurs 1959), but human classifiers lacked the image measurement ability to classify galaxies in anything other than a

![Resolution and S/N of a set of excellent quality images of clearly barred galaxies (family = B) systematically degraded to produce trends in S/N vs. resolution (as measured by number of pixels) for the cases of classifier success (left) and failure (right). The resulting mean trends in classifier success with S/N (star) and resolution (cross) are shown in the lower panel.](image)
set of bins. In addition, as discussed in §4.3, we may now incorporate information on image quality, as measured by S/N and resolution, in determining what measurable morphological information is available in a galaxy image. Finally, disk systems in the late stages of the RHS show increasing amounts of global image asymmetry. Such asymmetries are sometimes linked to tidal encounters with nearby galaxies. How are these asymmetries related? Is the increased degree of disk asymmetry caused exclusively by encounters, and is the high frequency of asymmetric systems observed at $z \approx 0.5$ the result of an increased interaction rate? Strength of the $1/\theta$ Fourier component, as measured by the first principal component of the $1/\theta$ amplitude profile, may be used to measure this asymmetry independent of the presence of other morphological structures like bars and arms, something that is not possible with less sophisticated image rotation asymmetry measurements. Such asymmetry measurements may be correlated closely with the internal kinematics of the galaxy (Rix & Zaritsky 1995) or the present level of active star formation in a disk and hence will convey more direct insight into the physical processes driving galaxy formation and evolution.

We acknowledge support from NASA grants AR 6385.01.95A, GO 6609.01-95A, and AR 7534.01-96A.

REFERENCES

Abraham, R., Tanvir, N. R., Santiago, B., Ellis, R. S., Glazebrook, K. G., & van den Bergh, S. 1996, MNRAS, 279, L47
Abraham, R., Valdes, F., Yee, H., & van den Bergh, S. 1994, ApJ, 432, 75
Andreon, S., Gargiulo, G., Longo, G., Tagliaferri, R., & Capuano, N. 2000, MNRAS, 319, 700
Bazell, D. 2000, MNRAS, 316, 519
Bender, R., & Mollenhoff, C. 1987, A&A, 177, 71
Buta, R. 1987, ApJS, 64, 383
Buta, R., & Combes, F. 1996, Fundam. Cosmic Phys., 17, 95
Buta, R., Mitra, S., de Vaucouleurs, G., & Corwin, H. G. 1994, AJ, 107, 118
Cohen, S. H., Windhorst, R. A., Odewahn, S. C., Chiarenza, C. A. T., & Driver, S. P. 2001, AJ, submitted
Corwin, H. G., de Vaucouleurs, A., & de Vaucouleurs, G. 1985, Southern Galaxy Catalogue (Austin: Dept. Astron., Univ. Texas)
de Jong, R. S., & van der Kruit, P. C. 1994, A&AS, 106, 451
de Vaucouleurs, G. 1958, Rev. Mod. Phys., 30, 926
de Vaucouleurs, G., de Vaucouleurs, A., & Corwin, H. G. 1991, in The Evolution of Galaxies and Stellar Populations, ed. B. Tinsley, R. Larson, & D. C. Gehret (New Haven: Yale Univ. Press), 43
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Buta, R., Paturel, G., & Fouqué, P. 1991, The Third Reference Catalog of Bright Galaxies (New York: Springer) (RC3)
Driver, S. P., Windhorst, R. A., & Griffiths, R. E. 1995a, ApJ, 453, 48
Driver, S. P., Windhorst, R. A., Ostrander, E. J., Keel, W. C., Griffiths, R. E., & Ratnatunga, K. U. 1995b, ApJ, 449, L23
Elmegreen, B., & Elmegreen, D. 1985, ApJ, 288, 438
Frei, Z., Guhathakurta, P., Gunn, J. E., & Tyson, J. A. 1996, AJ, 111, 174
Gal, R. R., de Carvalho, R. R., Brunner, R., Odewahn, S. C., & Djorgovski, S. G. 2000, AJ, 120, 540
Glazebrook, K., Ellis, R. E., Santiago, B., & Griffiths, R. E. 1995, MNRAS, 275, L19
Han, M. 1995, ApJ, 442, 504
Hickson, P. 1993, Astrophys. Lett., 29, 1
Hubble, E. 1926, ApJ, 64, 321
Im, M. et al. 2002, ApJ, in press
Kent, S. 1986, AJ, 91, 1301
Lauberts, A., & Valentinis, E. A. 1989, The Surface Photometry Catalogue of the ESO-Uppsala Galaxies (Garching: ESO)
Lauer, T. 1985, ApJS, 57, 473
Mahonen, P., & Frantti, T. 2000, ApJ, 541, 261
Mahonen, P. H., & Hakala, P. J. 1995, ApJ, 452, L77
McClelland, J. L., & Ramelhardt, D. E. 1988, Explorations in Parallel Distributed Processing (Cambridge: MIT Press)
Naim, A., Lahav, O., & Storrie-Lombardi, M. C. 1995, MNRAS, 275, 567
Naim, A., Ratnatunga, K. U., & Griffiths, R. E. 1997, ApJ, 476, 510
Nilson, P. 1973, Uppsala General Catalog of Galaxies (Uppsala: Roy. Soc. Sci.)
Odewahn, S. C. 1989, Ph.D. thesis, Univ. Texas
———. 1995, PASP, 107, 770
———. 1997, in Ann. NY Acad. Sci., 88, Nonlinear Signal and Image Analysis, ed. J. R. Buchler & H. Kandrup, 184
Odewahn, S. C., & Aldering, G. 1995, AJ, 110, 2009
Odewahn, S. C., Burstein, D., & Windhorst, R. A. 1997, AJ, 114, 2219
Odewahn, S. C., & de Vaucouleurs, G. 1992, ApJS, 83, 65
Odewahn, S. C., Stockwell, E. B., Pennington, R. M., Humphreys, R. M., & Zumph, W. A. 1992, AJ, 103, 318
Odewahn, S. C., Windhorst, R. A., Driver, S. P., & Keel, W. C. 1996, ApJ, 472, L13
Ohta, K., Hamabe, M., & Wakamatsu, K. 1990, ApJ, 357, 71
Okamura, S., Kodaira, K., & Watanabe, M. 1984, ApJ, 280, 7
Philip, N. S., Joseph, K. B., Kembhavi, A., & Wadadekar, Y. 2000, Proceedings of Automated Data Analysis in Astronomy (New Delhi: Narosa Publishing), 125
Rix, H., & Zaritsky, D. 1995, ApJ, 447, 82
Roberts, M. S., & Haynes, M. P. 1994, ARA&A, 32, 115
Simien, F., & de Vaucouleurs, G. 1986, ApJ, 302, 564
Spierermann, G. 1992, AJ, 103, 2102
Storrie-Lombardi, M. C., Lahav, O., & Storrie-Lombardi, M. C. 1992, MNRAS, 259, P8
van den Bergh, S., Cohen, J. G., & Crabbe, C. 2001, AJ, 122, 611
Weir, N., Fayyad, M. F., & Djorgovski, S. G. 2000, AJ, 120, 540
Weil, N., & Djorgovski, S. 1995, AJ, 109, 2401
Whitmore, B. C. 1984, ApJ, 278, 61