km². The rigidity (2 × 10¹⁰ N/m²) is the density (2600 kg/m³) multiplied by the square of the shear wave speed (2800 m/s), which is assumed to be the average compressional wave speed in the offshore forearc rocks (5000 m/s) divided by 1.75. If the slip is 1 m, the seismic moment (1) is 2 × 10¹⁹ N-m, for $M_w = 6.8$. For a large value of slip (5 m) on the same fault, $M_w = 7.3$. The regression formula of D. Wells and K. Coppersmith [*Bull. Seismol. Soc. Am.* **84**, 974 (1994)] relates magnitude to fault length for California strikeslip earthquakes; M = 5.16 + 1.12 log(length). After correction for differences in shapes of the fault areas (Cascadia forearc strike-slip faults are triangular in stead of rectangular in cross section), this formula indicates a magnitude of 7.2 for a length of 100 km.

- 25. W. Thatcher, J. Geophys. Res. 95, 2609 (1990).
- G. Ekström and R. Engdahl, *ibid.* 94, 15499 (1989);
 E. Geist, J. R. Childs, D. W. Scholl, *Tectonics* 7, 327 (1988).
- R. Jarrard [*Rev. Geophys.* 24, 217 (1986)] defined several strain classes for subduction zones, ranging from strongly compressional (active backarc thrusting) to extensional (active backarc spreading). D. Muhs et al. [*J. Geophys. Res.* 95, 6685 (1990)] showed that forearc uplift rates do not correlate with these strain classes.
- 28. G. McInelly and H. Kelsey, J. Geophys. Res. 95, 6699 (1990).

- 29. J. C. Savage and M. Lisowski, *Science* **252**, 101 (1991); C. Mitchell, P. Vincent, R. J. Weldon, M. A. Richards, *J. Geophys. Res.* **99**, 12257 (1994). H. Kelsey, D. C. Engebretson, C. E. Mitchell, and R. L. Ticknor (*ibid.*, p. 12245) showed that the along-strike variations in uplift rate mapped by Mitchell *et al.* correlate with topography and may be due to permanent uplift rather than seismic strain accumulation.
- R. D. Mueller, W. R. Roest, J. Y. Royer, L. M. Gahagan, J. G. Sclater, *SIO Reference Series* 93–30 (1993).
- 31. We thank V. Kulm, R. Yeats, A. Trehu, and J. Nabelek for discussions and R. Lillie and A. Nelson for reviews of an early manuscript. Supported by National Science Foundation grants OCE-9216880 [Oregon State University (OSU)], EAR-9105050 [Rensselaer Polytechnic Institute (RPI)], and EAR-9406917 (RPI), U.S. Geological Survey National Earthquake Hazards Reduction Program under awards 1434-93-G-2319 and 1434-93-G-2489 (OSU), and the National Oceanic and Atmospheric Administration National Undersea Research Program, West Coast National Undersea Research Center at the University of Alaska, grants UAF-92-0061 and UAF-93-0035 (OSU).

13 September 1994; accepted 14 December 1994

Galaxies, Human Eyes, and Artificial Neural Networks

O. Lahav, A. Naim, R. J. Buta, H. G. Corwin, G. de Vaucouleurs, A. Dressler, J. P. Huchra, S. van den Bergh, S. Raychaudhury, L. Sodré Jr., M. C. Storrie-Lombardi

The quantitative morphological classification of galaxies is important for understanding the origin of type frequency and correlations with environment. However, galaxy morphological classification is still mainly done visually by dedicated individuals, in the spirit of Hubble's original scheme and its modifications. The rapid increase in data on galaxy images at low and high redshift calls for a re-examination of the classification schemes and for automatic methods. Here are shown results from a systematic comparison of the dispersion among human experts classifying a uniformly selected sample of more than 800 digitized galaxy images. These galaxy images were then classified by six of the authors independently. The human classifications are compared with each other and with an automatic classification by an artificial neural network, which replicates the classification by a human expert to the same degree of agreement as that between two human experts.

Hubble (1) suggested a classification scheme for galaxies that consists of one sequence starting from elliptical galaxies

(E), through lenticular galaxies (S0), to spiral galaxies (S) and a parallel branch of spirals with a barred component, which vields the so-called "tuning fork" Hubble diagram. This scheme has been extended by astronomers over the years (2-5) to incorporate features such as the strength of the spiral arms, yielding multidimensional classifications (3, 5). It is remarkable that these somewhat subjective classification labels for galaxies (as seen projected on the sky) correlate well with physical properties such as color, dynamical properties (for example, rotation curves and stellar velocity dispersions), and the mass of neutral hydrogen (6). However, one would like eventually to devise a scheme of classification that can be related to the physical processes of galaxy formation. Although there have been in recent years significant advances in observational techniques (for example, in the telescopes, detectors, and reduction algorithms) and in theoretical modelling (for example, N-body and hydrodynamics simulations), galaxy classification remains a subjective area.

Quantifying galaxy morphology is important for various reasons. First, it provides important clues to the origin of galaxies and their formation processes. For example, elliptical and lenticular galaxies make up only \sim 20% of the galaxies, and there is a striking density-morphology relation (1, 7), indicating that elliptical galaxies mainly reside in high-density regions. Understanding the origin of the type frequency and the densitymorphology relation is of fundamental importance. However, quantification of these properties requires reliable classification schemes. Second, galaxies can also be used to, for example, measure redshift-independent distances by methods such as the luminosity-rotation velocity relation for spirals (8) and the diameter-velocity dispersion relation for ellipticals (9). Any observational program requires an a priori list of target objects for photometric or spectrographic measurements. Therefore, galaxy classification is important for the practical goal of producing large catalogs for statistical and observational programs and for establishing some underlying physics (in analogy with the Hertzsprung-Russell diagram for stars). Moreover, understanding the morphology of galaxies at low redshift is crucial for any meaningful comparison with galaxy images obtained with the Hubble Space Telescope at higher redshift ($z \approx$ 0.4). Most of our current knowledge of galaxy morphology is based on the pioneering work of several dedicated observers, who have classified and cataloged thousands of galaxies (2, 10, 11). However, facilities such as the Cambridge Automated Plate Measuring (APM) machine and the Sloan digital sky survey yield millions of galaxies. Classifying very large data sets is obviously beyond the capability of a single person. Therefore, the galaxy classification problem calls for new approaches (12-16).

As a first step toward finding an automated method of galaxy classification, we compiled a well-defined sample of galaxy images. The galaxies were chosen from the APM Equatorial Catalogue of galaxies (17). This sample was compiled from IIIaJ (broad blue-green band) plates taken with the United Kingdom's Schmidt telescope at Siding Spring, Australia; the sample covers most of the sky between declinations -17.5° and 2.5° at galactic latitudes $b \ge$ 20°. We chose a subsample of galaxies with major diameter (at an isophotal level of 24.5 magnitudes per square arc sec) $D \ge 1.2$ arc min on 75 plates, after eliminating gal-

O. Lahav, A. Naim, M. C. Storrie-Lombardi, Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK.

<sup>R. J. Buta, Department of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487–0324, USA.
H. G. Corwin, California Institute of Technology, IPAC M/S 100, Pasadena, CA 91125, USA.</sup>

G. de Vaucouleurs, Department of Astronomy, RLM 15.308, University of Texas, Austin, TX 78712–1083, USA.

A. Dressler, Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101–1292, USA.

J. P. Huchra and S. Raychaudhury, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA.

S. van den Bergh, Dominion Astrophysical Observatory, 5071 West Saanich Road, Victoria, British Columbia, V8X 4M6, Canada.

L. Sodré Jr., Instituto Astronômico e Geofísico da Universidade de São Paulo, CP9638, 01065, São Paulo, Brazil.

axies that had severe contamination from overlapping stellar or galaxy images (<10%). This sample of 831 galaxies was scanned by the APM in raster mode at a resolution of 1 arc sec (although the actual resolution of the Schmidt plates was more like 2 arc sec because of observing conditions). The digitized images (most of them measuring 256 pixels by 256 pixels) were printed at full resolution.

The same galaxy images were then classified by six of the authors (RB, HC, GV, AD, JH, and vdB) according to the Revised Hubble *T*-type (numerical stage) system (10, 18) or were converted to it. Although the *T*-type is only a one-dimensional parameter (extending from T = -6 to +11) in a three-dimensional scheme (10), it is commonly used and is convenient for computer algorithms compared with other, more descriptive schemes. While five of the authors classified the images on laser-printed hard copies, vdB examined them on a computer screen. His classification was done according to the DDO system (5),

which was then converted to the T-type (10). The motivation for performing a comparison between different experts is twofold: (i) to study systematically the degree of agreement and reproducibility between observers and (ii) to use the human classifications as training sets for artificial neural networks and other automatic classifiers.

Figure 1 shows the digitized images of four galaxies in our sample. We also give the classification assigned to these galaxies by the RC3 catalog (10) (ignoring the quoted uncertainty in their T-type) and by the six authors, who independently classified the galaxies. One of these galaxies got exactly the same classification by all six observers, but there was no such clear agreement on the other three galaxies. Statistically, all six authors agreed on the exact T-type for only 8 out of the 831 galaxies (less than 1%). Agreement between pairs of observers in excess of 80% is obtained only to within two types. Both GV and vdB, who have classified galaxies for many more years than the others, were rather conservative and did not



Fig. 1. Four APM galaxy images and their classification by six of the authors and RC3. The *T*-type classification of NGC2811 by RC3, RB, HC, GV, AD, JH, and vdB is 1, 1, 1, 1, 1, 1, and 1, respectively; of NGC3200, 4.5, 5, 5, 4, 5, 4, and 3; of NGC4902, 3, 3, 4, 3, 3, 5, and 3; and of NGC3962, -5, -3, 0, -5, -3, -1, and -5.

classify about a third of the galaxy images because they are saturated or of low quality. The other observers were more liberal and classified almost all of the galaxies (Table 1). On the whole, there is indeed a reasonable consistency in the way people classify galaxies, but the scatter is significant.

To better quantify the degree of agreement between observers, we calculated for each pair of observers a and b the variance

$$\sigma_{ab}^2 = \frac{1}{N_{ab}} \sum_{\text{gal}} (T_a - T_b)^2$$
(1)

taking into account only those N_{ab} galaxies for which both observers gave a classification.

The root-mean-square (rms) dispersion (Table 2) between RC3 and any of the observers (2.2 *T*-units on average) is larger than between any two observers who looked at the same APM images (1.8 *T*-units on average). We note that the subset of 600 RC3 galaxies in the sample has a median diameter of 1.7 arc min, compared with the median 1.5 arc min of the entire sample of 831 galaxies, and the images were on different plate materials. This illustrates the fact that any classification depends on the color, size, and quality of the images used; that is to say, there is no universal classification.

Observers who belong to the same "school" agree better with each other than with others. For example, the dispersion between GV and HC is only 1.5, and that between HC and RB only 1.3 units. This indicates that systematic training can re-

Table 1. The rms dispersion in *T*-type classification between the ANN and human experts. The ANN was trained and tested on individual observers and their mean classification. The third column gives the total number of galaxies classified by each expert.

Expert	Dispersion versus ANN	N _{gal}	
RB	1.9	764	
HC	2.0	812	
GV	2.2	473	
AD	1.9	814	
JH	2.3	824	
vdB	2.2	549	
Mean	1.8	831	

Table 2. The rms dispersion in *T*-type classification between pairs of observers.

	RB	HC	GV	AD	JH	vdB
RC3 RB HC GV AD JH	2.2	2.1 1.3	1.8 1.6 1.5	2.3 1.7 1.8 1.7	2.2 1.8 1.9 1.8 2.1	2.4 1.7 1.9 1.9 1.8 2.0

duce the scatter between two human experts. We also notice a weak trend for better agreement in the classification of large galaxies (the rms dispersion between experts drops by about 10% from 1.2- to 2-arc min galaxies), but there is no obvious trend as a function of eccentricity. Detailed analysis of this comparison will appear elsewhere (19). We also intend to evaluate the internal scatter σ_a (the reproducibility) of each observer when classifying the same data set again or a set with lower resolution. As a crude estimate, if we assume that $\sigma_{ab}^2 = \sigma_a^2$ + σ_b^2 , we find for the different observers that σ_a is between 1.0 and 1.5. It is worth emphasizing that the plate material used here suffers from problems of saturation, and the digitization of the images (although at pixel size of 1 arc sec) may have degraded the agreement between observers. Nevertheless, the plate material we have used is typical in many extragalactic studies.

Having established the degree of agreement between human experts, the challenge is to design a computer algorithm that will reproduce classification to the same degree that a student or colleague of the human expert can. Such an automated procedure usually involves two steps: (i) the extraction of features from the digitized image, such as the galaxy profile, the extent of spiral arms, the color of the galaxy, or an efficient compression of the image pixels into a smaller number of coefficients (for example, Fourier or principal component analysis); and (ii) a classification procedure in which a computer learns from a training set for which a human expert provided his or her classification.

Artificial neural networks (ANNs), originally suggested as simplified models of the human brain, are computer algorithms that provide a convenient general-purpose framework for classification (20), including astronomical applications (21, 22). One commonly used ANN configuration consists of nodes arranged in a series of layers and utilizes the backpropagation minimization algorithm (23). In Fig. 2 we show a configuration in which the galaxy parame-



Fig. 2. A schematic diagram of an artificial neural network for classifying galaxies. In this configuration, the galaxy parameters are fed into the input layer, and the T-type classification appears as a single continuous output. The network is trained according to classification by a human expert. The hidden layer allows nonlinear boundaries in a complicated parameter space.

ters are fed into the input layer, and the T-type classification appears as a single continuous output. The hidden layer allows nonlinear boundaries in a complicated parameter space. In the training phase, the free parameters of the network (weights) are determined by least squares minimization of the difference between the calculated and true (that is, the expert's) type. Other network configurations are possible, including multiple output nodes, which can provide Bayesian a posteriori probabilities for each class (14, 24).

Pilot studies (14, 24) have had ANNs with 13 parameters classify about 5200 galaxies from the ESO-LV catalog (11), which illustrates that galaxies can be classified automatically, with an rms dispersion of 2.1 *T*-units between the ANN and the experts [Lauberts and Valentijn (11)]. However, because of the lack of quantitative measure of dispersion among human experts for comparison, it was difficult for us to judge if the achieved success rate was satisfactory. We have now applied (25) the same technique to our new APM sample, after extracting significant features (ellipticity, surface brightness, luminosity profile parameters, arms to disk ratio, concentration indices, and arms parameters) from the images. We then trained the ANN on the T system (as in the network shown in Fig. 2), feeding as input 13 parameters and allowing 5 nodes at the hidden layer, a 13:5:1 configuration (other network configurations, such as 13:13:1, have yielded similar results). For each of the six individual expert classifications, the ANN was trained on threequarters of the sample and tested on the remaining fourth. Because the ANN minimization begins with a set of random weights, we repeated the training and testing 10 times, with different initial weights (typical internal scatter when networks with different initial random weights are used is about 0.1 to 0.3 units). The same process was repeated for the other threequarters of each set, resulting in 40 runs for each expert classification.

The rms dispersion (Eq. 1) between the ANN and each expert (Table 1) quantifies to what extent the human classification can be reproduced by the computer algorithm. The rms dispersion varies between 1.9 and 2.3 T-units over the six experts. This relatively small variation from one expert to another is not too surprising. The number of galaxies classified by each of the experts was different (Table 1), with bias toward face-on galaxies in some cases. A large rms dispersion may not necessarily reflect inconsistency in the expert's own classification but rather a poorer fit between the human classification, the chosen parameters, and the model (that is, the ANN). A better agreement, 1.8 Tunits, is achieved when the ANN is trained and tested on the mean type as deduced from all available expert classifications (after removing a few outliers). Comparison of Tables 1 and 2 shows remarkable similarity in the dispersion between two human experts and that between the ANN and experts. In other words, our results indicate that the ANN can replicate an expert's classification of the APM sample as well as the colleagues or students of the expert.

Figure 3 shows an example of the ANN versus mean expert classification for 207 galaxies, after training on the remaining 624 galaxies in the sample (again averaging results from 10 runs with different initial random weights). Because all of the



SCIENCE • VOL. 267 • 10 FEBRUARY 1995

Fig. 3. The ANN versus mean expert *T*-type classification for 207 galaxies. The ANN was trained on the remaining 624 galaxies in the sample (with results averaged over 10 runs with different initial random weights). The solid circles indicate galaxies larger than the median diameter of 1.5 arc min, and the open circles indicate smaller galaxies.

galaxies in our sample are larger than 1.2 arc min in diameter, there is no obvious trend for worse classification for smaller diameters, although such a trend is expected for much smaller galaxies. There is also no dramatic trend with ellipticity. Of the 831 galaxies classified by the ANN by the above procedure, 9% deviate from the "true" mean answer by at least three types. Most of them are very late types and irregulars (T > 7).

Our comparison indicates that although the T-system is convenient, the scatter between observers is not negligible. Caution is called for in assuming a universal frequency type distribution in comparison with models and with high-redshift galaxies. The observed frequency distribution depends on the plate material and on the human expert. Future work will focus on supervised ANNs, to preserve human experience in multidimensional classification (3, 5), and on unsupervised algorithms (for example, by generalizing principal component analysis to nonlinear mapping), to define a "new physical Hubble sequence" without any prior human classification.

REFERENCES AND NOTES

- 1. E. Hubble, *The Realm of Nebulae* (Yale Univ. Press, New Haven, CT, 1936).
- A. R. Sandage, *The Hubble Atlas of Galaxies* (Carnegie Institute of Washington, Washington, DC, 1961).
- 3. G. de Vaucouleurs, Hand. Phys. 53, 275 (1959).
- 4. W. W. Morgan, *Publ. Astron. Soc. Pac.* **70**, 364 (1958).
- 5. S. van den Bergh, Astrophys. J. 206, 883 (1976).
- R. Buta, S. Mitra, G. de Vaucouleurs, H. G. Corwin Jr., Astron. J. 107, 118 (1994).
- 7. A. Dressler, *Astrophys. J.* **236**, 351 (1980).
- R. B. Tully and J. R. Fisher, Astron. Astrophys. 54, 661 (1977).
- 9. D. Lynden-Bell et al., Astrophys. J. 326, 19 (1988).
- G. de Vaucouleurs et al., Third Reference Catalogue of Bright Galaxies (Springer-Verlag, New York, 1991).
- A. Lauberts and E. A. Valentijn, *The Surface Photom*etry Catalogue of the ESO-Uppsala Galaxies (European Southern Observatory, Munich, 1989).
- M. Thonnat, in *The World of Galaxies*, H. G. Corwin and L. Bottinelli, Eds. (Springer-Verlag, New York, 1988), p. 53.
- 13. G. Spiekermann, Astron. J. 103, 2102 (1992).
- M. C. Storrie-Lombardi, O. Lahav, L. Sodré Jr., L. J. Storrie-Lombardi, *Mon. Not. R. Astron. Soc.* 259, 8 (1992).
- 15. M. Doi, M. Fukugita, S. Okamura, *ibid*. **264**, 832 (1993).
- R. Abraham, F. Valdes, H. K. C. Yee, S. van den Bergh, *Astrophys. J.* **432**, 75 (1994).
- 17. S. Raychaudhury et al., in preparation
- 18. In the *T*-type system of galaxy classification (10), the types *T* = -6 to 9 correspond to galaxy types *CE*, *E*(0, *E* +, S0⁻, S0⁰, S0⁺, S0/a, Sa, Sab, Sb, Sbc, Sc, Scd, Sd, Sdn, and Sm, respectively; types *T* = 10, 11, 90, and 99 correspond to *Im*, *cl*, 10, and *Pec*, respectively.
- 19. A. Naim et al., in preparation.
- J. Hertz, A. Krogh, R. G. Palmer, Introduction to the Theory of Neural Computation (Addison-Wesley, Redwood, CA, 1991).
- S. C. Odewahn, E. B. Stockwell, R. L. Pennington, R. M. Humphreys, W. A. Zumach, Astron. J. 103, 318 (1992).
- 22. M. Serra-Ricart, X. Calbet, L. Garrido, V. Gaitan,

- *ibid*. **106**, 1685 (1993).
- D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Nature* 323, 533 (1986).
- 24. O. Lahav, A. Naim, L. Sodré Jr., M. C. Storrie-Lombardi, in preparation.
- 25. A. Naim, O. Lahav, L. Sodré Jr., M. C. Storrie-Lombardi, in preparation.
- 26. We thank the Schmidt Telescope unit of the Royal

Observatory of Edinburgh for the plate material, the APM group at the Royal Greenwich Observatory, Cambridge, for scanning support, M. Irwin and D. Lynden-Bell for helpful discussions. We are also grateful to the anonymous referee for helpful comments.

16 August 1994; accepted 1 December 1994

Temperature Dependence of the Superconducting Gap Anisotropy in $Bi_2Sr_2CaCu_2O_{8+x}$

Jian Ma, C. Quitmann, R. J. Kelley, H. Berger, G. Margaritondo, M. Onellion*

Detailed data on the momentum-resolved temperature dependence of the superconducting gap of Bi₂Sr₂CaCu₂O_{8+x} are presented, complemented by similar data on the intensity of the photoemission superconducting condensate spectral area. The gap anisotropy between the Γ - \bar{M} and Γ -X directions increases markedly with increasing temperature, contrary to what happens for conventional anisotropic-gap superconductors, such as lead. Specifically, the size of the superconducting gap along the Γ -X direction decreases to values indistinguishable from zero at temperatures for which the gap retains virtually full value along the Γ - \bar{M} direction. These data rule out the simplest type of *d*-wave order parameter.

An "order parameter" describes the type of phase transition that occurs in a material system. The order parameter of high-temperature superconductors is of extreme current interest and has been investigated by several techniques, including angle-resolved photoemission (1-4). Angle-resolved photoemission has the advantage of directly investigating the momentum dependence of the gap. Already, results establishing a marked anisotropy in the gap at low temperatures have ruled out an isotropic s-wave symmetry order parameter (2-5).

Our main result is that, contrary to conventional anisotropic-gap superconductors such as lead (6), the gap anisotropy of oxygen-overdoped Bi2Sr2CaCu2O8+x increases with increasing temperature as one approaches the superconducting transition temperature $T_{\rm c}$. We estimated the size of the gap in two ways: using the BCS (Bardeen-Cooper-Schrieffer)-like lineshape (1, 7) computer code of Olson et al. (1) and also using the shift of the 50% point of the photoemission leading edge (2). The gap values obtained by the two methods agreed to better than 1.5 meV. Our results place stringent constraints on any theory of hightemperature superconductivity.

As a test of our experimental standards, we measured the angle-resolved photoemission spectrum of a gold film deposited in situ (Fig. 1A); the temperature of the film was 36 K. The 10 to 90% energy width of the Fermi-Dirac distribution function Fermi edge was 15 \pm 2 meV. Magnetic susceptibility measurements were also taken for an oxygen-overdoped Bi₂Sr₂CaCu₂O_{8+x} single crystal sample (Fig. 1B) as part of a test of sample quality. The 10 to 90% transition temperature width was 1.3 K. Our photoemission measurements were performed in an ultrahigh-vacuum chamber with a base pressure of 6×10^{-11} torr. The light source was the 4-m normal incidence monochromator at the Wisconsin Synchrotron Radiation Center. The electron energy analyzer we used was a 50-mm Vacuum Science Workshop hemispherical analyzer, mounted on a two-axis goniometer, with an acceptance full angle of 2°. The total energy resolution employed was 25 meV, slightly worse than the best obtainable. Samples were transferred from a load lock chamber and were cleaved in situ at 35 K. The sample holder rotated the sample about the surface normal, at low temperature, for precision alignment with respect to the photon electric field. The sample crystal structure and orientation were determined by in situ low-energy electron diffraction (LEED). The sample temperature stability was ± 1 K.

For measurements of the temperature dependence of the gap, we chose two locations in the Brillouin zone where the superconducting gap is large. These points are (i)

862

J. Ma, C. Quitmann, R. J. Kelley, M. Onellion, Department of Physics and Applied Superconductivity Center, University of Wisconsin–Madison, 1150 University Avenue, Madison, WI 53706, USA.

H. Berger and G. Margaritondo, Institut de Physique Appliquée, Ecole Polytechnique Fédérale, CH-1015 Lausanne, Switzerland.

^{*}To whom correspondence should be addressed.