THE ASTROPHYSICAL JOURNAL, 635:L29–L32, 2005 December 10

MORPHOLOGY SEGREGATION OF GALAXIES IN COLOR-COLOR GRADIENT SPACE

CHANGBOM PARK[1] AND YUN-YOUNG CHOI[1]

Received 2005 September 30; accepted 2005 October 31; published 2005 December 1

ABSTRACT

We have found the $u - r$ color versus $g - i$ color gradient space can be used for highly successful morphology classification of galaxies in the Sloan Digital Sky Survey. In this space, galaxies form well-separated early- and late-type branches. The location of galaxies along the branches reflects the degree and locality of star formation activity, and monotonically corresponds to the sequence of morphological subclasses. When the concentration index is also used, the completeness and reliability of classification reaches about 91% for a training set of SDSS galaxies brighter than $r_{\rm pet} \approx 15.9$. At the faintest magnitudes ($r_{\rm pet} \approx 17.5$) of the SDSS spectroscopic sample, the accuracy still remains at about 88%. The new classification scheme will help us find accurate relations of galaxy morphology with spatial and temporal environments, and help us to understand the origin of morphology of galaxies.

Subject heading: galaxies: fundamental parameters

Online material: color figure

1. INTRODUCTION

Unlike stars, galaxies show diverse shapes. The belief that the difference in appearance reflects a difference in generic character has motivated the morphological classification of galaxies. Since the 1920s, the simple scheme suggested by Hubble (1926), based on single-band images of bright galaxies, has been widely adopted. The essence of Hubble's scheme and its elaborations (de Vaucouleurs et al. 1991; Kormendy 1979; de Vaucouleurs 1959; Sandage 1961) is to divide galaxies into early and late types. The early types are further divided into elliptical and lenticular, and the late types into spiral (unbarred and barred) and irregular. Anyone looking directly at galaxy images immediately realizes that the shapes of galaxies are much more diverse than this, but also that the simple Hubble sequence catches the major features of galaxy morphology. As unprecedentedly large sets of digital images of galaxies, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000)[2] data, become available, the time is ripe for understanding the origin of galaxy morphology. In order to find the accurate relationship of galaxy morphology with other physical properties, we must define morphological types more objectively and use morphology subsets as homogeneous as possible. Endeavors to develop an objective, automatic, and successful morphology classifier applicable to large libraries of digital images have resulted in various schemes (Yamauchi et al. 2005; Blanton et al. 2003b; Abraham et al. 1996, 2003; Strateva et al. 2001; Shimasaku et al. 2001; Doi et al. 1993). This work further advances this topic.

The philosophy of our morphological classification is to develop a scheme applicable down to the faint magnitude limit ($r_{\rm pet} \approx 17.77$, where $r_{\rm pet}$ is the Petrosian magnitude after Galactic extinction correction) of the spectroscopic sample of the SDSS using only photometric information. The fundamental features of galaxy morphology that are still useful for galaxies near the faint limit are the surface brightness and color profiles. Their mean levels are measured by mean surface brightness and integrated color. Their radial variation, scaled to a characteristic size, can be measured by the Sersic or concentration index and color gradient, for example. Therefore, in principle, the morphology classification of galaxies should be made at least in four-dimensional parameter space. The surface brightness at long-wavelength bands represents the stellar mass distribution, and the color tells us about the recent star formation history. One difficulty in dividing galaxies into early and late types using the surface brightness profile alone lies in the fact that most spirals consist of both a bulge and a disk. There is an unavoidable confusion between the bulge-only and the bulge-plus-disk systems when only the surface brightness information in a single band is used for classification. Contamination in the early- and late-type subsets separated using the concentration index, for example, is typically about 20% (Yamauchi et al. 2005; Shimasaku et al. 2001). The surface texture is of limited use, because it is sensitive to seeing and is simply lost for small, faint galaxies ($r_{\rm pet} > 16.0$ in the case of SDSS; Yamauchi et al. 2005). On the other hand, the information in colors reveals the additional generic differences between early- and late-type galaxies, which have experienced different star formation histories. Strateva et al. (2001) have found that the integrated (observer frame) $u^* - r^*$ color (asterisks are attached to the SDSS photometry based on the photometric equations used through the Early Data Release; Stoughton et al. 2002) of the SDSS Main galaxies (Strauss et al. 2002) shows a bimodal distribution. However, they have shown that, when divided at $u^* - r^* = 2.22$, the early (E/S0/Sa) and late-type (Sb/Sc/Irr) subsets have significant contamination, reaching about 30% for a sample with visually identified morphological types. In this Letter we extend their work by incorporating color gradient and the concentration index as additional dimensions of the classification parameter space.

2. THE TRAINING SAMPLE

We use a training set of 1982 galaxies to quantify the performance of our morphology classification scheme. The training set consists of two samples. The first contains 981 SDSS galaxies with $14.5 < r_{\rm pet} \leq 15.0$ after Galactic extinction correction, whose morphological types are assigned by the authors based on the color images retrieved by the SDSS Image List

[1] Korea Institute for Advanced Study, Dongdaemun-gu, Seoul 130-722, South Korea; cbp@kias.re.kr, yychoi@kias.re.kr.

[2] Full details of the SDSS are available at http://www.sdss.org.

Tool.[3] Galaxies are divided into six types: E/S0 (smooth brightness and color profile), S0 (disk with features, but no spiral arm), AE (E or S0 with abnormal brightness or color distribution), S (spiral arm, disk with dust lane), AS (spirals with large distortion in shape), and Irr (irregulars). The first three correspond to the early type, and the rest are late types. The second training set is a catalog by Fukugita et al. (2006, in preparation; see also Nakamura et al. 2003) listing 1875 galaxies with $r^*_{\rm pet} \leq 15.9$ (after Galactic extinction correction), with visually identified morphological types from $T = 0$ to 6 in 0.5 steps. From these we use a subset of 1183 galaxies that are listed in both the Main Galaxy and the spectroscopic sample and that are not too contaminated by foreground stars. By comparing the morphological types of 182 galaxies in common in our sample and Fukugita et al.'s catalog, we have found that the early and late types divide at $T = 1.5$ ($T = 1$ and 2 correspond to Hubble types S0 and Sa, respectively) in the latter catalog.

3. MORPHOLOGY CLASSIFICATION SCHEME

We use the color-color gradient space as the major morphology classification tool, and the concentration index as an auxiliary parameter. In the color-color gradient space, spiral galaxies tend to separate from the region clustered by elliptical and lenticular galaxies, as majority of spirals exhibit significant color gradient. As a measure of color gradient, we adopt the difference in ${}^{0.1}(g-i)$ color of the region with $R < 0.5R_{\rm pet}$ from that of the annulus with $0.5R_{\rm pet} < R < R_{\rm pet}$, where $R_{\rm pet}$ is the Petrosian radius and ${}^{0.1}(g-i)$ is a rest-frame $g-i$ color K-corrected to the redshift of 0.1 [hereafter, $g-i$ means ${}^{0.1}(g-i)$], and the K-correction of magnitude and color is calculated based on the study of Blanton et al. (2003a). A negative color difference means bluer outside. Model magnitude is used for color; this is the magnitude calculated from the best-fit model profile obtained by fitting the de Vaucouleur and exponential profiles to the galaxy image. We have chosen the g and i bands to estimate color gradient because they are widely separated in wavelength across the 4000 Å break, and have signal-to-noise ratios higher than the u- and z-bands. Choosing other bands results in nosier color gradient estimates at faint magnitudes, or less successful separation of galaxies into early and late types. Other measures of color gradient, such as the linear slope of the radial $g-i$ color profile and the $g-i$ color difference between the region with $R < R_{50}$ and the annulus with $R_{50} < R < R_{90}$, are also calculated for comparison (where R_{50} and R_{90} are the radii from the center of a galaxy containing 50% and 90% of the Petrosian flux). We have found that the clump of normal early-type galaxies in the color-color gradient space is less tight for the latter measures at fainter magnitudes, $r_{\rm pet} > 16$.

Throughout this Letter we use AB magnitudes converted from SDSS magnitudes. The g- and i-band atlas images and basic photometric parameters of individual galaxies are retrieved from the SDSS Data Release 3 (DR3) at Princeton.[4] We have generated a large set of Sersic model images convolved with the PSF of various sizes. These images are used to find a best-fit Sersic model with true Sersic index and true axis ratio for an observed image of a given size of the PSF. The fitting is made at radii larger than $0.2R_{\rm pet}$, in order to avoid the central region, whose profile is much affected by seeing. Then the best-fit Sersic models in the g and i bands are convolved with the PSF of the same size, and the effects of seeing on the concentration index and color difference are estimated. We have confirmed that the seeing-corrected concentration index and color difference show no apparent dependence on seeing. We use the elliptical annuli in all our parameter calculations to take into account flattening or inclination of galaxies. The position angle and axis ratio used to define the elliptical annuli and to calculate the radial surface brightness profile are isophotal ones measured from the i-band image. The Petrosian radii in our analysis are usually larger than those in the DR3 catalog, which adopted circular annuli.

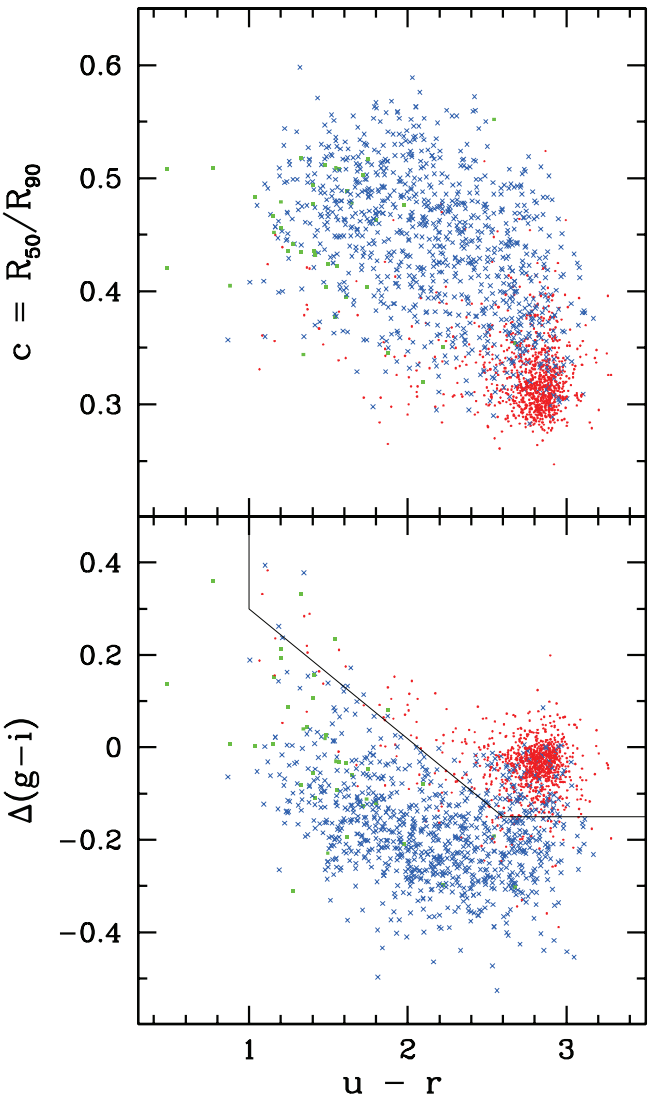

FIG. 1.—Distributions of the 1982 galaxies in the morphology training set. *Top*: $u-r$ color vs. (inverse) concentration index space; *bottom*: $u-r$ vs. $\Delta(g-i)$ color difference space. Red dots are early types (E/S0, S0, AE), blue crosses are spirals (S, AS), and green squares are irregulars (Irr).

The photometric parameters we use for morphology classification are the $u-r$ color, color difference $\Delta(g-i)$, and the (inverse) concentration index c in the i band, where $c \equiv R_{50}/R_{90}$. Here the concentration index is used as a complementary parameter to discriminate red disky spirals from early-type galaxies. The i-band image is used to measure the index because the image at longer wavelengths is expected to better represent the stellar mass distribution. All parameters are seeing-corrected as described above.

Figure 1 shows the distribution of galaxies of this training

[3] See http://cas.sdss.org/astro/en/tools/chart/list.asp.
[4] See http://photo.astro.princeton.edu.

TABLE 1
CLASSIFICATION CRITERIA

CLASSIFICATION CRITERIA	MAGNITUDE BIN			
	14.5–16.0	16.0–16.5	16.5–17.0	17.0–17.5
P1	(1.00, 0.30)	(1.00, 0.30)	(1.00, 0.30)	(1.00, 0.30)
P2	(2.60, −0.15)	(2.65, −0.18)	(2.65, −0.18)	(2.70, −0.18)
P3	(3.50, −0.15)	(3.50, −0.15)	(3.50, −0.25)	(3.50, −0.35)
c_{cut}	0.43	0.45	0.47	0.48
Completeness and Reliability				
C (elliptical)	0.913	0.883	0.872	0.883
R (elliptical)	0.901	0.892	0.882	0.881
C (spiral)	0.901	0.902	0.892	0.890
R (spiral)	0.914	0.893	0.883	0.892

NOTES.—The three points, P1, P2, and P3, define the lines dividing the SDSS galaxies into early and late types in the $u - r$ vs. $\Delta(g - i)$ plane. The early types are also required to have $c < c_{cut}$.

set in the $u - r$ versus $\Delta(g - i)$ space and in the $u - r$ versus concentration index space. The early (E/S0) and late types (Sa to Sd) determined by visual inspection are marked as red circles and blue crosses, respectively. Green squares are irregular galaxies. In our automated scheme, early type galaxies are classified as those lying above the boundary lines passing through the points (3.5, −0.15), (2.6, −0.15), and (1.0, 0.3) in the $u - r$ versus $\Delta(g - i)$ space. They are also required to have $c < 0.43$. The rest are classified as late types. The completeness of this classification scheme reaches 91.3% for early types and 90.1% for late types, with reliabilities of 90.1% and 91.4%, respectively. The completeness, *C*, is the fraction of galaxies of a given type that are successfully selected from the original sample by the classification scheme. The reliability, *R*, is the fraction of galaxies of the desired type from the selected subsample. The 9%–10% failure is due to a few red passive spirals, galaxies with companions or foreground stars, and/or incorrect visual classification (for edge-on objects in particular). The parameter *c* plays only an auxiliary role, improving the results by a few percent.

We estimate that the completeness and reliability of our morphology classification is about 88% near $r_{pet} = 17.5$, which is close to the faint end of the spectroscopic sample of the SDSS. The performance of our classification scheme slowly degrades at $r_{pet} \geq 16$, as it becomes harder to measure the color difference and concentration index as the size of galaxies relative to the CCD pixel and the FWHM of the PSF decreases. To estimate the performance of our classification scheme at fainter magnitudes, we redshift the galaxies in the training sample ($r_{pet} = 14.5$–15.0) with morphological types assigned by us. We calculate the new redshift at which a galaxy appears dimmer by a desired magnitude, adopting a cosmology with density parameters $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$. The reduction factor in angular size is calculated, and images are binned down (Giavalisco et al. 1996). The resulting images are convolved with the PSF in such a way that they have seeing effects equal to the originals. Noise is not added anew. Following this procedure, we generated three mock morphology samples of 981 galaxies in 0.5 mag bins from $r_{pet} = 16.0$ to 17.5. The distributions of these redshifted galaxies in the color-color gradient space and in the color-concentration index space are very similar to those of three samples of 1000 randomly selected SDSS galaxies with actual magnitudes in the same bins, thus justifying our procedure. We determine the classification criteria at fainter magnitudes using these redshifted morphology samples. Results are summarized in Table 1.

Table 1 shows that our classification boundaries in the color-color gradient space hardly change even at magnitudes down to $r_{pet} = 17.5$. The only major change is the boundary at the reddest colors (the line connecting to the point P3). At fainter magnitudes, the clump of the normal E/S0's expands vertically in the color-color gradient space, because the measured color differences have more errors. The classification boundary at the reddest colors has been chosen to include more ellipticals, taking this fact into account. We think that the *K*-correction of color in this reddest color range has been accurately estimated. Our *K*-correction might have more errors for intermediate-type spirals, for which the spectra of the central parts of galaxies are less representative of the total color. But in this case our classification is hardly affected, because these galaxies are blue, have large color gradients, and are located far from the classification boundaries.

4. DISCUSSION

Inspecting the location of galaxies in the color-color gradient space allows us to classify them beyond the dichotomous division into early and late types. Normal elliptical and lenticular galaxies are very homogeneous systems whose light is dominated by old Population II stars. They strongly concentrate within a spot centered at (2.82, −0.04) in the $u - r$ versus $\Delta(g - i)$ plane. Dispersion is mainly due to companions or foreground stars. They show very weak but definite color gradient (i.e., outside is bluer). We have discovered a trail containing about 10% of early-type galaxies toward the left ($u - r < 2.5$) of the concentration of normal early types and then upward. The galaxies in this trail are bluer than normal ones, often show emission lines, and occasionally show starburst activity near the center. They tend to be more centrally concentrated than spirals at the same color. We have found that most E+A galaxies (Dressler & Gunn 1983) listed in Yamauchi & Goto (2005) are located in this trail.

On the other hand, a spiral branch extends below the cluster of normal early types. Sa-type spiral galaxies often start to have star-forming zones at the outskirts of the disk. Their integrated color is still dominated by the red light from their bulges, but they begin to show a color gradient. As the star formation activity becomes stronger, the color gradient first increases, and then starts to decrease as the star-forming region expands toward the center. Very late type spirals and most irregulars at the upper left corner are very blue and have inverted color gradients (i.e., the center is bluer), since they usually show strong starburst activity near the center. The spiral branch meets the trail of early types there. The major advantage of morphology classification in the color-color gradient space is that the blue early-

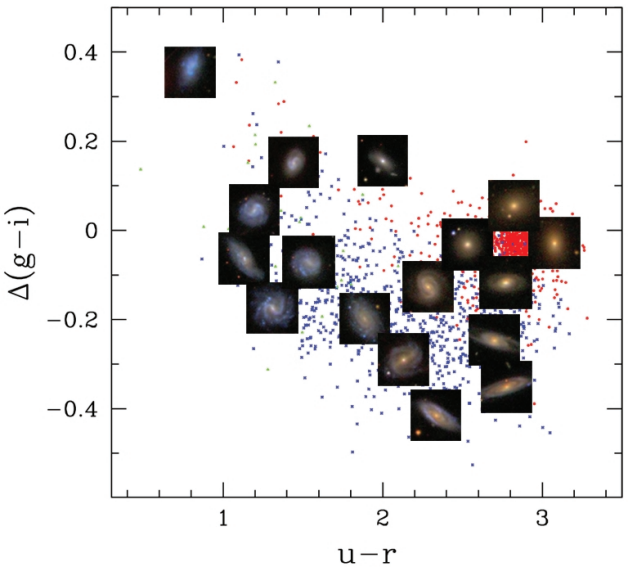

FIG. 2.—Continuous changes in the appearance of galaxies along the spiral and elliptical branches. Note that the star-forming region extends from the outskirts of disk toward the center as one moves from red to blue spirals.

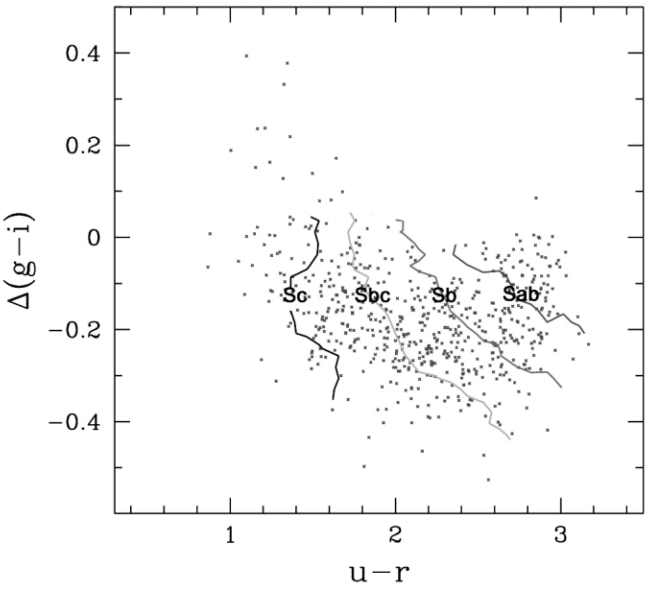

FIG. 3.—Distribution in the $u - r$ vs. $\Delta(g - i)$ plane of 582 spiral galaxies in the training set taken from the Fukugita et al. morphology sample (in preparation), which lists subclasses for spirals. Superposed are the subclass isocontours, calculated by averaging the listed subclasses within the ellipse of axis lengths of 0.6 and 0.2 in the $u - r$ and $\Delta(g - i)$ directions, respectively. Contour levels are $T = 2.5$ (Sab), 3 (Sb), 3.5 (Sbc), and 4 (Sc), from right to left. [*See the electronic edition of the Journal for a color version of this figure.*]

type galaxies are separated from the spiral galaxies, as the former tend to have a nearly constant color profile or bluer cores, while the latter tend to have a red bulge plus blue disk structure. The spiral branch is a sequence of locations of star formation activity and cold gas and dust concentration, as illustrated in Figure 2. The position of a spiral along the branch has a strong correlation with its morphological subclass, as shown in Figure 3. After making the early/late type division, one can further divide galaxies into subclasses based on their location along the early- and late-type branches. For example, late-type galaxies can be divided into L1, L2, and L3 subclasses, which group galaxies lying above the line $\Delta(g - i) = 0.8 - 0.4(u - r)$, between this line and the line $\Delta(g - i) = 1.6 - (u - r)$, and left of the latter line, respectively. These subclasses roughly correspond to Sa-Sb, Sb-Sc, and Sc-Sd/Irr, respectively. The dispersion of the spiral branch is partly due to internal extinction, which tends to make $u - r$ and $|\Delta(g - i)|$ increase. The dependences of $u - r$ and $|\Delta(g - i)|$ on the axis ratios of galaxies indicate that the effects of internal extinction are strongest for intermediate-type spirals.

Our new morphology classification scheme takes into account the history and locality of star formation, as well as stellar mass distribution. It is easy to implement the scheme for multiband surveys with wide spectrum coverage. We expect that the strong concentration of early types and the general shape of the spiral branch in the color-color gradient space will persist for galaxies at high redshifts. Any change in their location and shape will be colorful evidence for the evolution of star formation activity and cold gas infall into galaxies with different morphology.

C. B. P. thanks Michael Vogeley for his collaboration on parts of this work and for an invitation to Drexel University, where this work was started. We also thank Yasushi Suto for suggesting that we use Fukugita et al.'s catalog. We thank the anonymous referee for helpful comments. This work is supported by the Korea Science and Engineering Foundation (KOSEF) through the Astrophysical Research Center for the Structure and Evolution of Cosmos (ARCSEC).

REFERENCES

Abraham, R. G., van den Bergh, S., Glazebrook, K., Ellis, R. S., Santiago, B. X., Surma, P., & Griffiths, R. E. 1996, ApJS, 107, 1
Abraham, R. G., van den Bergh, S., & Nair, P. A. 2003, ApJ, 588, 218
Blanton, M. R., et al. 2003a, AJ, 125, 2348
———. 2003b, ApJ, 594, 186
de Vaucouleurs, G. 1959, Handb. Phys., 53, 275
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Jr., Buta, R. J., Paturel, G., Fouqué, P. 1991, Third Reference Catalogue of Bright Galaxies (New York: Springer)
Doi, M., Fukugita, M., & Okamura, S. 1993, MNRAS, 264, 832
Dressler, A., & Gunn, J. E. 1983, ApJ, 270, 7
Giavalisco, M., Livio, M., Bohlin, R. C., Macchetto, F. D., & Stecher, T. P. 1996, AJ, 112, 369
Hubble, E. P. 1926, ApJ, 64, 321
Kormendy, J. A. 1979, ApJ, 227, 714
Nakamura, O., Fukugita, M., Yasuda, N., Loveday, J., Brinkmann, J., Schneider, D. P., Shimasaku, K., & SubbaRao, M. 2003, AJ, 125, 1682
Sandage, A. 1961, The Hubble Atlas of Galaxies (Washington, DC: Carnegie Inst.)
Shimasaku, K., et al. 2001, AJ, 122, 1238
Stoughton, C., et al. 2002, AJ, 123, 485 (erratum 123, 3487)
Strateva, I., et al. 2001, AJ, 122, 1861
Strauss, M. A. 2002, AJ, 124, 1810
Yamauchi, C., & Goto, T. 2005, MNRAS, 359, 1557
Yamauchi, C., et al. 2005, MNRAS, in press
York, D., et al. 2000, AJ, 120, 1579