PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **116**:133–137, 2004 February

Preparing Red-Green-Blue Images from CCD Data

ROBERT LUPTON,[1] MICHAEL R. BLANTON,[2] GEORGE FEKETE,[3] DAVID W. HOGG,[2]
WIL O'MULLANE,[3] ALEX SZALAY,[3] AND NICHOLAS WHERRY[2]
Received 2003 November 24; accepted 2003 December 24; published 2004 February 2

ABSTRACT.
We present a new, and we believe arguably correct, algorithm for producing red-green-blue (RGB) composites from three-band astronomical images. Our method ensures that an object with a specified astronomical color (e.g., $g-r$ and $r-i$) has a unique color in the RGB image, as opposed to the burnt-out white stars to which we are accustomed. A natural consequence of this is that we can use the *same* colors to code color-magnitude diagrams, providing a natural "index" to our images. We also introduce the use of an arcsinh stretch that allows us to show faint objects while simultaneously preserving the structure of brighter objects in the field, such as the spiral arms of large galaxies. We believe that in addition to their aesthetic value, our images convey far more information than do the traditional ones, and we provide examples from Sloan Digital Sky Survey (SDSS) imaging, the Hubble Deep Field (HDF), and *Chandra* to support our claims.[4]

The purest and most thoughtful minds are those which love colour the most.
—John Ruskin, *The Stones of Venice* (1852)

1. INTRODUCTION

One of the reasons that our forebears were drawn to astronomy was the sheer drama and beauty of the night sky. When photography became a standard astronomical technique, it was natural to combine images taken in different bands to produce spectacular colored images of the night sky (e.g., Malin 1992, and references therein).

When CCD detectors were introduced into astronomy 25 years ago, and especially when large-format cameras became common, astronomers were once again motivated to generate three-color composites. CCDs are approximately linear, unlike the roughly logarithmic sensitivity curve of photographic plates, so the natural way to produce an image was to map the logarithm of the pixel intensities to the red, green, and blue brightnesses in the final image (see, e.g., Villard & Levay 2002). This algorithm produces pictures that look rather like those produced from photographs, but unfortunately it also discards a lot of information.

2. GENERATING COLORED IMAGES FROM CCD DATA

Let us consider CCD data taken in three bands, which we call *b*, *g*, and *r* (e.g., *B*, *R*, and *I*); we assume that these data have been background-subtracted, if necessary linearized, flat-fielded, and scaled appropriately. This scaling is to some extent arbitrary, as different astronomers prefer different color balances, but we have found that a conversion to f_ν (in practice to an AB scale; Oke & Gunn 1983) works well.

We desire a mapping to the range [0, 1] for each of three colors, red (R), green (G), and blue (B). The usual algorithm is

$$\mathrm{R} = f(r),\ \mathrm{G} = f(g),\ \mathrm{B} = f(b), \tag{1}$$

where

$$f(x) = \begin{cases} 0, & x < m, \\ F(x-m)/F(M-m), & m \le x \le M, \\ 1, & M < x, \end{cases}$$

and where m is the minimum value to display and M is the maximum; these are chosen to bring out favored details. As mentioned in the introduction, a common choice is logarithmic ($F \equiv \ln$), but linear or square-root stretches are also popular. The actual implementation, at least for integer input data, often uses a lookup table.

This style of mapping simulates one of the unfortunate features of photographic plates, namely that all pixels for which r, g, and b are $\ge M$ appear white in the final image. In a photograph this white color corresponds to a lack of information, but for a CCD it is a choice made by the astronomer. In fact, for *any* nonlinear function F, an object's color in the composite

[1] Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001; rhl@astro.princeton.edu.
[2] Center for Cosmology and Particle Physics, Department of Physics, New York University, NY 10003.
[3] Johns Hopkins University, 3701 San Martin Drive, Baltimore, MD 21218.
[4] More examples are available at http://www.astro.princeton.edu/~rhl/PrettyPictures.

FIG. 1.—Approximately 5′ × 6.5′ *gri* composite based on SDSS imaging of the galaxies NGC 6976, NGC 6977, and NGC 6978. The faintest objects visible in this image have $r \sim 22$. Note the colors of the stars, ranging from bluish through yellow to red, the blue spiral disks, and the yellowish color of the galaxies' spheroids. The stretch is arcsinh with $Q \equiv 8$ and $\alpha \equiv 0.02$.

image depends on its brightness, even if *none* of *r*, *g*, or *b* exceed *M*.

Fortunately, we can easily avoid these problems. Defining $I \equiv (r + g + b)/3$,[5] set

$$\begin{aligned} \mathrm{R} &= r\, f(I)/I, \\ \mathrm{G} &= g\, f(I)/I, \\ \mathrm{B} &= b\, f(I)/I. \end{aligned} \tag{2}$$

If $\max_{\mathrm{RGB}} \equiv \max(\mathrm{R,G,B}) > 1$, set $\mathrm{R} \equiv \mathrm{R}/\max_{\mathrm{RGB}}$, $\mathrm{G} \equiv \mathrm{G}/\max_{\mathrm{RGB}}$, and $\mathrm{B} \equiv \mathrm{B}/\max_{\mathrm{RGB}}$; if $I \equiv 0$, set $\mathrm{R} = \mathrm{G} = \mathrm{B} = 0$. In other words, the intensity is clipped at unity, but the *color* is correct.[6] Additionally, it is possible to choose a more flexible functional form for *F*. Following Lupton, Gunn, & Szalay 1999, we take $F(x) \equiv \mathrm{arcsinh}(x/\beta)$, where the softening parameter β is chosen to bring out desired details; it determines the point at which we change from a linear to a logarithmic transformation—$\mathrm{arcsinh}(x) \sim x$ for $x \ll 1$, $\mathrm{arcsinh}(x) \sim \ln(2x)$ for $x \gg 1$.[7] The linear part is used to show faint features in the data, while the logarithmic part enables us to simultaneously illustrate structures such as spiral arms or the cores of star clusters that are significantly brighter. As an additional trick, we find that sometimes an $\mathrm{arcsinh}^{1/2}$ stretch brings out the details of dust lanes in spiral galaxies; here, as before, the arcsinh is used to control the dynamic range.

The results of this simple change are dramatic; a given color (i.e., value of $g-b$ and $r-g$) is now mapped to a unique color in the RGB image. Figure 1 shows a Sloan Digital Sky Survey (SDSS; York et al. 2000) *gri* composite of the galaxies NGC 6976, NGC 6977, and NGC 6978. The reader will notice that the stars, rather than all being white, show a variety of red, yellow, and bluish tints; that the spiral bulges (and the group of ellipticals to the lower right of NGC 6977) are a uniform yellowish color, while the spiral disks are blue ([O III] and Hα appear in B and R, respectively). Furthermore, it is evident that star formation activity in the disks (as indicated by the distinctive bluish color) decreases from left to right.

We can use the well-known color composite of the Hubble Deep Field (HDF; Williams et al. 1996), shown in Figure 2, as a pedagogical example.

The top row of images were prepared using equation (1), while the bottom row used equation (2). The top left panel shows the familiar STScI color image. The bottom left panel uses an arcsinh stretch and is our preferred representation of the HDF. The top middle panel is our best match to the STScI image and uses a logarithmic stretch; the middle bottom panel uses the same logarithmic function (i.e., identical values of *m* and *M*). The right-hand pair of panels share a hard linear stretch

[5] It is possible to use a different measure of *I*, e.g., an rms or $(\chi^2)^{1/2}$ value.

[6] We note that it is often possible to include World Coordinate System information in the header of the RGB image; file formats that allow this include JPEG and TIF.

[7] A convenient parameterization is $f(x) = \mathrm{arcsinh}(\alpha Q[x - m])/Q$, which allows the user to first set $Q \to 0$ and choose the linear stretch α and then adjust *Q* to bring out brighter features. In this case $M \equiv m + \sinh(Q)/(\alpha Q)$.

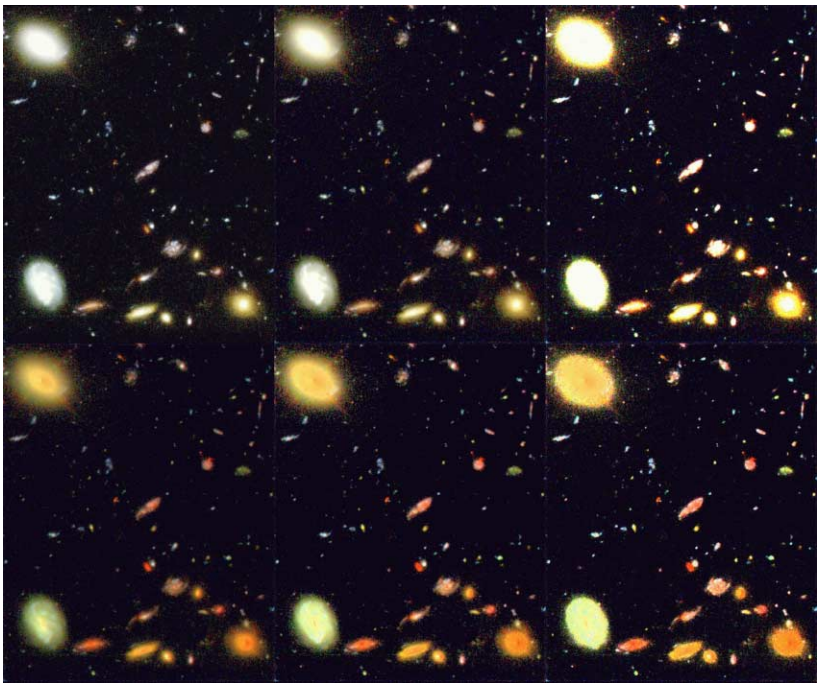

FIG. 2.—HDF composite with *f*450, *f*606, and *f*814. The top row of images were prepared using eq. (1), while the bottom row used eq. (2). The top left panel shows the familiar STScI color image. The bottom left panel uses an arcsinh stretch and is our preferred representation of the HDF. The top middle panel is our best match to the STScI image and uses a logarithmic stretch; the middle bottom panel uses the same logarithmic function (i.e., identical values of *m* and *M*). The right-hand pair of panels share a hard linear stretch (it is identical to the linear part of the arcsinh stretch of the bottom left panel). In the bottom row of images, the larger galaxies have similar colors to the galaxies in Fig. 1, but the fainter, more distant, red galaxies have no analogs in the shallow SDSS data; our eyes are seeing the effects of redshifting the 4000 Å break. (Picture courtesy of NASA, R. Williams (STScI), and the Hubble Deep Field Team.)

(it is identical to the linear part of the arcsinh stretch of the bottom left panel).

Interestingly, the bottom set of panels are pretty similar, with the main difference being that the spiral arms in the lower left corner are rather better defined in the leftmost panel. It is clear that much of the morphological information is carried in the *color* rather than the *intensity*.

The bands (*f*450, *f*606, and *f*814) used in Figure 2 are not quite the same as those used in the SDSS composites (*g*, *r*, and *i*), but the astrophysical palette is similar, in the sense that the larger HDF galaxies have colors similar to those seen in Figure 1, with blue disks and yellowish bulges. The fainter objects, however, look unlike *anything* seen in SDSS imaging, with bright red (*K*-corrected) bulges and, in some cases, bluish disks, which must be due to active star formation and the emission of far-UV light.

Figure 3 shows the globular cluster NGC 2419. Figure 4 shows a color-magnitude diagram of point sources in a 0.6 deg^2 region surrounding the cluster. The points in this diagram have the *same* color as the corresponding objects in the image; the cluster horizontal and giant branches stand out.

Finally, this technique is not restricted to optical data; Figure 5 shows a *Chandra* X-ray image of the supernova remnant Cas A.

3. DETECTOR SATURATION

Although CCDs are wonderful detectors, they are not perfect. The SDSS data-processing pipeline (Lupton et al. 2001) interpolates over saturation trails, which accounts for their absence from these images. Unfortunately, it is not able to correctly recover the counts in the cores of saturated stars, but we *do* know which pixels were involved.[8]

Equation (2) means that every pixel in the image preserves its color, even in the saturated cores of stars. We can avoid undesirable artifacts in the final picture by finding connected pixels that were saturated in at least one band and replacing them with the average color of the pixels touching the saturated region. Because we do not wish to average over the long bleed trails produced by bright stars, we only apply this procedure to saturated pixels with intensities above *M*. An obvious example of a bright star is seen at the bottom right of Figure 3. In this case we treated all pixels with at least 10,000 counts in any band as "saturated."

[8] If this information is not available, all pixels above some threshold can be considered tainted.

FIG. 3.—Approximately $7' \times 9'$ *gri* composite based on SDSS imaging of the halo globular cluster NGC 2419. Note the blue horizontal branch and yellowish red giants; the red stars are M stars in the disk. The diffraction spikes show that the instrument rotator moved significantly during the 5 minutes between the *g* and *r* exposures. The stretch is arcsinh with $Q \equiv 9$ and $\alpha \equiv 0.02$.

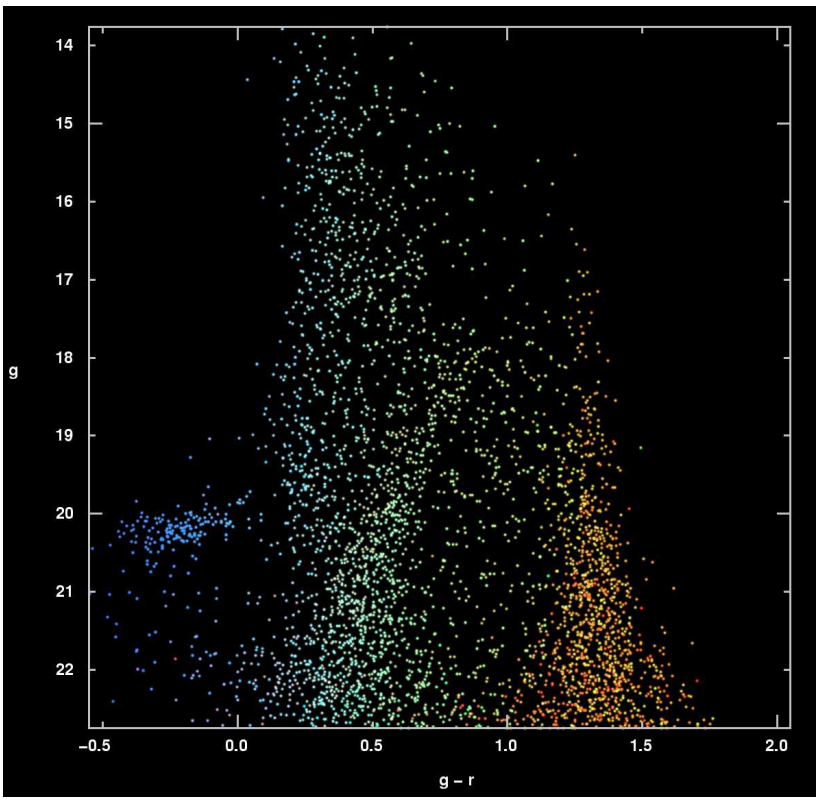

FIG. 4.—$g-r$ vs. g color-magnitude diagram for stars in an area of about 0.6 deg^2 around the cluster NGC 2419. Each star has the same color as it has in a *gri* composite image of that part of the sky, a part of which is shown in Fig. 3.

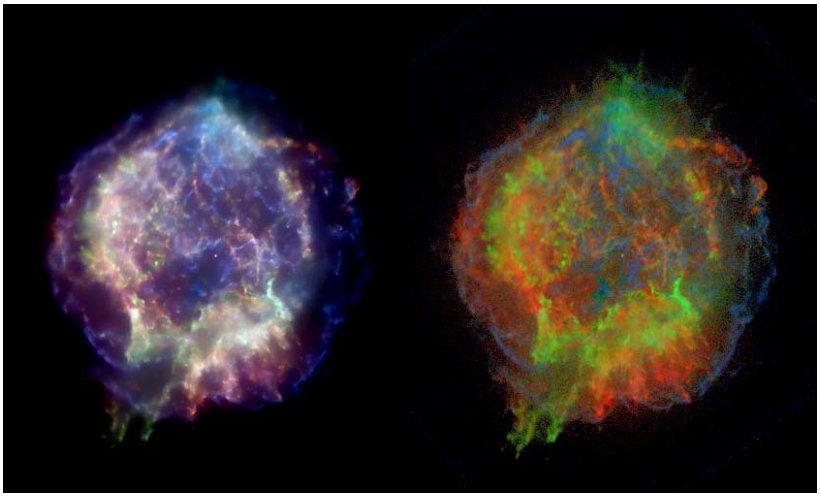

FIG. 5.—Supernova remnant Cas A as imaged by the *Chandra* X-ray satellite using the ASIS camera. R, G, and B correspond to 0.3–1.55 keV, 1.55–3.34 keV, and 3.34–10 keV, respectively. The left-hand image is from the *Chandra* Supernova Remnants Catalog, the right-hand image is prepared using eq. (2) and an arcsinh stretch. The input image is the raw *Chandra* data smoothed with a $\sigma = 0.5$ pixel Gaussian (this is not the same as the inputs to the left-hand panel). Note the presence of both blue and red filaments crossing the center of the remnant. It is possible to generate a figure analogous to Fig. 4 that shows the mapping from hardness ratios to RGB. (Data and *Chandra* X-Ray Center color composite prepared by NASA/CXC/SAO.)

4. PRACTICAL IMPLEMENTATIONS

The original implementation of these algorithms was included in the SDSS image-processing code, but fortunately stand-alone versions are also available.[9]

5. CONCLUSIONS

Color images have traditionally been considered a luxury, but with the techniques espoused in this paper, we have found that they convey an enormous amount of information. For example, star-forming regions have a distinctive color, and it is very clear, looking at images of spiral galaxies, how much the star formation rate varies. We encourage the reader to visit the Web page listed here,[10] where, among other objects, all of the NGC galaxies in the SDSS DR1 (Abazajian et al. 2003) are presented.

R. H. L. thanks Michael Strauss for helpful comments on an early version of this manuscript. The SDSS Web Site[11] lists the Participating Institutions and the project's other sources of funding, which include the Sloan Foundation, the NSF, NASA, DOE, the Japanese Monbukagakusho, and the Max Planck Society.

[9] An IDL version is available from http://cosmo.nyu.edu/hogg/visualization/rgb/, and a C version from http://www.sdss.jhu.edu/doc/jpeg.html.

[10] At http://www.astro.princeton.edu/~rhl/PrettyPictures/.

[11] The SDSS Web site is http://www.sdss.org/.

REFERENCES

Abazajian, K., et al. 2003, AJ, 126, 2081
Lupton, R. H., Gunn, J. E., & Szalay, A. 1999, AJ, 118, 1406
Lupton, R. H., et al. 2001, in ASP Conf. Ser. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne (San Francisco: ASP), 269
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Malin, D. 1992, QJRAS, 33, 321
Villard, R., & Levay, Z. 2002, S&T, 104, 3, 28
Williams, R. E., et al. 1996, AJ, 112, 1335
York, D. G., et al. 2000, AJ, 120, 1579