Astronomical Data Analysis Software and Systems XIX **P56**
ASP Conference Series, Vol. XXX, 2009
Y. Mizumoto, K.-I. Morita, and M. Ohishi, eds.

Optimal Compression of Floating-Point FITS Images

W. D. Pence

NASA, Goddard Space Flight Center, Greenbelt, MD, USA

R. L. White

Space Telescope Science Institute, Baltimore, MD, USA

R. Seaman

National Optical Astronomy Observatory, Tucson, AZ, USA

Abstract. Lossless compression (e.g., with GZIP) of floating-point format astronomical FITS images is ineffective and typically only reduces the file size by 10% to 30%. We describe a much more effective compression method that is supported by the publicly available fpack and funpack FITS image compression utilities that can compress floating point images by a factor of 10 without loss of significant scientific precision. A "subtractive dithering" technique is described which permits coarser quantization (and thus higher compression) than is possible with simple scaling methods.

1. Introduction

Floating-point format astronomical FITS images usually do not compress well with general purpose lossless compression algorithms such as GZIP because the pixel values in astronomical images typically contain a large amount of noise, which is inherently incompressible. In this paper we demonstrate that much greater image compression can be achieved by discarding some of the noise without sacrificing the precision of scientific measurements in the image.

2. Quantization and Dithering Techniques

The amount of noise in floating point images can be reduced by rounding the pixel values into a set of discretely spaced intensity levels. This is most commonly done by using a linear function,

$$FloatValue = ScaleFactor \times IntegerValue + ZeroPoint \qquad (1)$$

to convert the floating point values into scaled integers, which can then be losslessly compressed using a fast and efficient algorithm such as Rice. The quantizing scale factor in this equation may be more conveniently specified as a fraction, q, of the RMS noise σ measured in background regions of the image:

$$q = \sigma / ScaleFactor \qquad (2)$$

In a previous paper (Pence, Seaman, and White 2009), we showed that the compression ratio, R, of the quantized image can be accurately predicted from the q value by the relation

$$R = BITPIX/(\log_2(q) + 1.8 + K) \tag{3}$$

where BITPIX is the number of bits per pixel (usually 16 or 32) and K is an algorithm-dependent efficiency parameter which has a value of about 1.2 for the Rice algorithm.

In order to achieve the greatest amount of compression one should use the smallest possible value of q (i.e., the coarsest spacing of the quantized levels) that still preserves the necessary astrometric and photometric precision within the image. If an image is quantized too coarsely then of course some of the information content of the image may be lost. In astronomical images, it turns out that the quantity that is most sensitive to quantization is the measurement of the background sky intensity level: if all the image pixels are quantized onto the exact same grid of intensity levels then the measured background level will tend to be biased towards the nearest quantized level. The net result will be to bias the photometry of faint objects in the image when using a coarse quantization grid.

This quantization bias can be reduced, however, by adding noise to the pixel values. (This phenomenon is more generally known as "stochastic resonance" in which the detection of a weak signal can be improved if noise is added to the system.) We use an ingenious variation on this technique, called "subtractive dithering", that was first developed by Roberts (1962): a random value between 0 and 1 is added to the pixel value prior to scaling it to an integer, and then the *same* random value is subtracted when rescaling the integer back to the quantized floating point value. This serves to randomize the pixel values without actually adding more noise to the image.

To measure the benefits of this subtractive dithering technique, we used the publicly available Sextractor program to measure the magnitude of the stars in a floating-point FITS image originally take with the Steward Observatory 2.3m telescope. We then repeated the magnitude measurements after first quantizing the image, using q = 1, both with and without subtractive dithering. The left panel of Figure 1 shows the magnitude residuals between the original image and the quantized image when dithering is not applied, and the right panel shows the residuals after applying the subtractive dithering technique. This clearly shows that dithering reduces both the systematic bias as well as the overall scatter in the residuals.

Figure 2 shows the relative magnitude residuals (the difference between the magnitudes in the original and in the quantized image, divided by the statistical uncertainty on the original measurement) for different q quantization levels when dithering is also applied. As can be seen, the magnitude residuals are still less than 1 σ even when using the coarsest q = 1 quantization, which gives a compression ratio of about 10 (from Eq. 3).

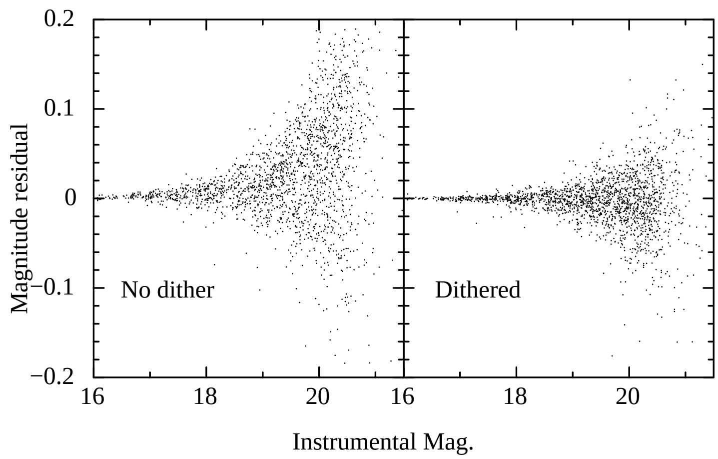

Figure 1. Magnitude residuals in a q = 1 quantized image, without (left) and with (right) subtractive dithering. Dithering the pixel values eliminates the positive bias and reduces the scatter in the magnitude residuals.

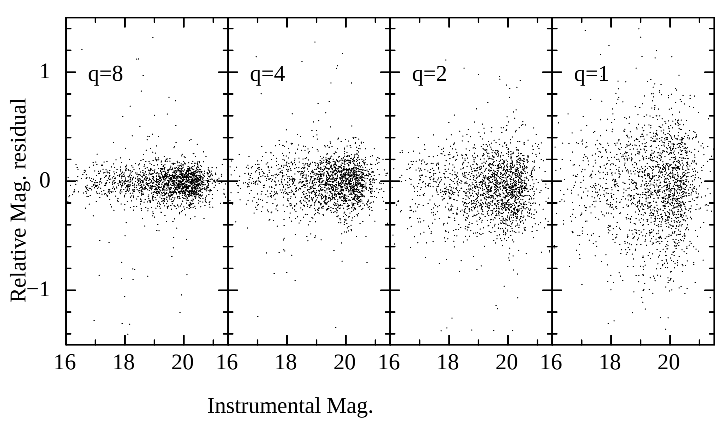

Figure 2. Relative magnitude residuals as a function of q.

3. Conclusions

This study demonstrates that typical floating point astronomical images can be compressed by up to a factor of about 10 without significant loss of photometric precision. The subtractive dithering technique described here is supported by the fpack and funpack astronomical image compression utilities that are publicly available from the HEASARC web site[1]. These utilities provide much greater and faster compression than GZIP for both integer and floating-point FITS images. A more extensive discussion of these and other results related to optimal compression of floating-point astronomical images can be found in a related paper (Pence, White, & Seaman 2010).

References

Pence, W. D., Seaman, R., & White, R. L. 2009, PASP, 121, 414

Pence, W. D., White, R. L., & Seaman, R. 2010, PASP, in press

Roberts, G. L. 1962, "Picture coding using pseudo-random noise", IRE Transactions on Information Theory IT-8(2), 145

[1]`http://heasarc.gsfc.nasa.gov/`