Appendix B

Deconvolution of Ground-Based Optical/Near- Infrared Images

ABSTRACT

The limitations of the arcsecond seeing normally experienced at ground-based sites has severely limited our ability to understand the structure of a wide range of phenomena. Since the only significant populations of certain types of objects (such as ULIGs and QSOs) exist at $z > 0.05$ and the structures we are interested in may be only a few hundred pc in size, the need for high spatial resolution ($< 0.4''$) imaging has been very great. While this situation has been partially alleviated by space-based observations, the demand for high spatial resolution is far greater than can be addressed solely by extremely costly space telescopes. The use of adaptive optics techniques and image restoration will therefore become increasingly commonplace as astronomers continue to strive to extract more information from their data.

B.1. Introduction

The use of image restoration techniques has had a long history of application in fields such as radio astronomy, where the complex spatially extended beam pattern has generally produced unrecognizably confused raw data. Optical astronomers, however, have generally been reluctant to apply deconvolution methods to their own data for two reasons: (a) they have been suspicious of the additional layer of (often *non-linear*) processing and it's associated uncertainties, and (b) the relatively simple point-spread-function (PSF) encountered in optical astronomy results in cosmetically acceptable (albeit blurry) images, thus eliminating much of the impetus for deconvolution found in other branches of astronomy.

Ironically, the aberrations of the flawed primary mirror of the *Hubble Space Telescope* (*HST*), while representing a major setback in the telescope's designed performance, almost single-handedly created a new field in the application of deconvolution algorithms to astronomical data as astronomers clamored for tools with which to restore the degraded *HST* images (White & Allen 1990 and references therein). While the servicing mission and installation of COSTAR has eliminated much of the drive for deconvolution that once existed, the techniques developed during that time period are still relevant to many ground-based observations.

B.2. Why deconvolve?

There are two primary reasons to deconvolve one's data. The first is to cosmetically improve the image by making it less blurry. More precisely, by eliminating the broad wings of the PSF, one vastly increases the detectability of high spatial frequency features by producing higher contrast (i.e. a larger peak/background ratio). Figure B.1 shows a raw image of Saturn taken at 1.6μm by the author with the UH 2.2m telescope. Figure B.1 also

shows the same image after 25 iterations of the Richardson-Lucy deconvolution algorithm. Many features, particularly in the rings, are now visible. In galaxy and stellar images this often results in an increase in point source detectability of 1 magnitude or more, which (as noted earlier) is much more effective than increasing the telescope aperture or total integration time.

Fig. B.1.— Raw (top) and Richardson-Lucy deconvolved (bottom) H-band images of Saturn taken at the UH 2.2m. Additional structure can be seen in the rings and cloud bands, as well as the moon Dione, can be clearly seen in the deconvolved image.

The second reason to deconvolve one's data is to improve point source photometry by helping to eliminate confusion. In particular, as discussed in Chapter 3, the PSF encountered with adaptive optics systems tends to consist of a broad disk of scattered light and a high S/N core. Application of deconvolution can help eliminate photometric confusion that occurs when objects lie closer than $1''$apart. Figure B.2 shows the total encircled energy as a function of radius both before and after deconvolution with Richardson-Lucy. Almost all astronomers unwittingly perform deconvolution already in order to improve their photometry when they apply "aperture corrections". This is essentially a single pass of the CLEAN algorithm (see below) with the gain set to 1, making the implicit assumption that

all of the flux for a given measurement lies in a single point source located in the center of the aperture.

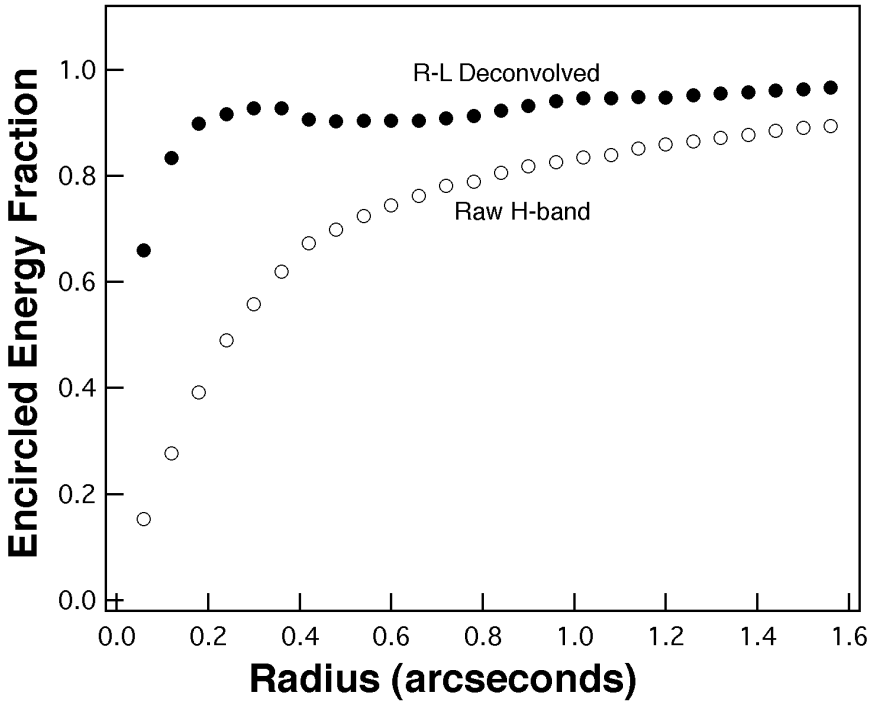

Fig. B.2.— Fractional encircled energy for *H*-band data before and after R-L deconvolution. After processing, more than 80% of the energy is inside a nearly diffraction limited spot. This is similar to the expected atmosphere-free performance of the telescope.

All of the data in this dissertation has been deconvolved, usually with several different techniques, in an attempt to better understand the photometry and morphology of the high spatial frequency structure. This appendix discusses some of the issues encountered in deconvolution of astronomical data at optical and near-infrared wavelengths.

B.3. Requirements for Deconvolution

There are several requirements for successful deconvolution of data:

Adequate Spatial Sampling - a common rule of thumb used in the design of astronomical cameras is that the FWHM of the typical achieved PSF should be 3 pixels. This has been motivated primarily by a need to balance cosmic ray rejection considerations which require higher spatial sampling rates (cosmic ray hits tend to produce profiles narrower than the PSF) and the demand for an adequate field of view (dictated by the total pixel dimensions of the imaging array). However, this sampling is inadequate for deconvolution purposes. Sub-pixel shifts in the true centroid of the emission relative to the center of the pixel may create large changes in the apparent PSF. A more adequate sampling has been found to be 5-7 pixels FWHM. For WFPC1 *HST* images, which have an aberrated PSF nearly 20 pixels

across, this is not a problem. Post-servicing WFPC2 PSFs, however, may be only 1-2 pixels FWHM, and deconvolution efforts are generally of limited value. At the UH 2.2m f/31 focus, the sampling of the QUIRC near-infrared camera (0.06″/pixel, best seeing typically 0.3″) and the Orbit near-ultraviolet/optical camera (0.09″/pixel when binned 2x2, best seeing typically 0.5″) are typically 5 pixels or more FWHM, ensuring adequate sampling all the time.

Isoplanicity - at the time of this writing, nearly all deconvolution algorithms assume a spatially invariant (isoplanatic) PSF. Current attempts to tackle non-isoplanicity generally involve breaking an image up into smaller isoplanatic subimages, each of which is then deconvolved separately. The most serious source of non-isoplanacity in conventional astronomical images is that produced by off-axis telescope aberrations. However, when both the target and the PSF calibration source are near the central axis, as is enforced by the small field-of-view of the detectors at the UH 2.2m at f/31, this is fairly negligible. For much larger fields of view, like that of the UH 8k camera, field distortions may prove to be a serious impediment to deconvolution.

Note that this is completely different from the isoplanacity issues normally encountered in adaptive optics. This is because adaptive optics is concerned with isoplanacity of *short time-scales* induced by the atmosphere, typically less than 0.01 seconds. Although at this time the size of the isoplanatic patch on Mauna Kea is not well-understood, it is believed to be roughly 1′ in the near-infrared for tip/tilt aberrations. However, the problem we are addressing is isoplanacity of the *time-averaged* PSF on *long time-scales* of 120-1200 seconds. Examination of the PSF on large- scale optical (B & I) images show essentially no PSF variation within a radius of 120″ of the target.

Adequate S/N - both the PSF and the data must have adequate S/N in order for deconvolution to produce good results. In particular, after many iterations of fourier-based algorithms like R-L, high-σ noise outliers are amplified into noise spikes, as background flux is concentrated into these high-σ points. This, combined with the smoothing effect of the first few iterations of the algorithms, produces a mottled background. While the noise in the original image is essentially uncorrelated from pixel to pixel, the background noise in the deconvolved image will be correlated on spatial scales similar to that of the PSF. Also, for some algorithms the PSF itself must have S/N at least as high as the actual data. This is discussed below in relation to the CLEAN algorithm.

B.4. Algorithms

The following algorithms (Richardson-Lucy, CLEAN, and MEM) are among the most common in use today. Additionally, they are all implemented in IRAF/STSDAS, thus making them convenient to use, as well readily and freely available. Many other algorithms exist: Fractal Pixon Reconstruction, the Maximum Correlation Method, Wiener Filtering, etc. Most of these are rather esoteric and in practice do not produce such good results as the following methods.

Richardson-Lucy - developed independently by Richardson (1972) and Leon Lucy (1974), R-L has become the algorithm of choice among optical astronomers. It has several advantages over competing techniques. The nature of the solution is surprisingly robust against small variations in the PSF. It is also computationally fast. Using R-L is quite simple: generate a PSF (either observed or theoretical), input proper guesses at the background values and CCD noise characteristics, and then start the algorithm. While it is possible to specify an *image prior* (i.e., a first guess) in order to accelerate convergence of the algorithm, this isn't necessary and is usually irrelevant since convergence is very rapid. Because this is a fourier-based algorithm, it behooves the user to make both the image and the PSF a power of 2 in each dimension. This allows the use of FFTs, which speed things up by factors of 20 or more.

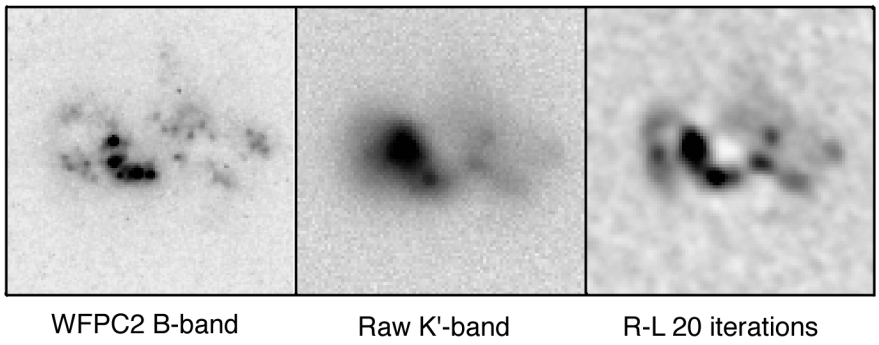

Fig. B.3.— A comparison between WFPC2 B-band data, raw tip/tilt data taken from the UH 2.2m telescope at K', and the results of deconvolution by 20 iterations of the Richard-Lucy algorithm. Most of the star-forming knots seen in the *HST* data are now visible in the resolution-enhanced image. Total processing time was about 1 minute on a Sparc 10.

There are several artifacts encountered when using R-L. The most notable of these is "ringing". Bright point sources embedded in extended emission will produce rings of alternating bright emission and negative flux, a result of the fourier-based algorithm attempting to reproduce high spatial frequency structure (i.e., point sources) when it has no information about such high frequencies. This is a general fourier deconvolution issue based on the fact that there is an upper and lower limit on the spatial frequencies in the image set by the image size and pixel size. Related to this is an issue of photometric instability; bright point sources tend to "steal" flux from their surroundings. There is also a tendency for variable spatial resolution. Before the image has converged to the maximum likelihood solution, the high S/N regions will have a higher effective spatial resolution than the low S/N regions. Fortunately, this is not as severe as the corresponding problem with CLEAN. Finally, there is the problem of noise amplification. High σ outliers in the background noise are interpreted as point sources and then flux is "borrowed" from the surrounding background and pushed into these points. This problem has to some extent been alleviated through use of damping algorithms (White 1992); in general, one stops iterating before this becomes too severe.

One of the more difficult issues is deciding when to stop the algorithm. In theory, the algorithm simply approaches the maximum likelihood solution and should simply be allowed to continue until it has reached a point of diminishing returns. In practice, however, noise amplification and other forms of artifacting becomes quite severe and eventually degrade the image restoration. Originally, a mathematical definition of convergence towards the maximum-likelihood solution was defined, but in practice it became apparent that this was not completely applicable. More recently, some authors have attempted to define new stopping criteria for the algorithm (Perry & Reeves 1992). Realistically, until a more precise understanding of how R-L works is reached, it is likely that no simple stopping criterion will be found. In actual practice one iterates until the noise amplification or ringing become too severe to be acceptable. Often it becomes apparent that the rate of convergence (i.e, the rate of change in the image as characterized by the X^2 value) slows dramatically; usually this is a good point to stop the algorithm. In general 20-40 iterations seems most appropriate for galaxy data with moderate S/N. For planetary data with high S/N extended features (such as figure B.1) this may be as high as 100.

Finally, a more unusual application of R-L is that it can be used as a kind of adaptive smoothing algorithm. The first few iterations of R-L are essentially a convolution of the original data with the PSF, which smoothes the data and correlates the noise over several pixels. By setting the CCD noise estimate to several times it's actual value, the algorithm is damped or slowed down. The first few iterations smooth the data, thus eliminating pixel to pixel noise. The next few iterations then enhance the spatial resolution, returning it to it's original value. This takes advantage of what is normally a flaw in R-L and most other deconvolution algorithms, namely that high S/N features in the image converge to the high spatial resolution solution faster than low S/N regions. The end result is an image whose high S/N features have a resolution equal to or greater than the original data, but with much lower background noise. Note that this will not actually allow detection of fainter features than before; such faint features are also smoothed by this process.

So called "P-Lucy" is an interesting variant of the R-L algorithm that is beginning to see some application. This is the so-called "photometric- channel" version of R-L; it is essentially a hybrid CLEAN/R-L algorithm which attempts to suppress the ringing around bright point sources common to fourier-based algorithms by alternatingly fitting all of the (user-specified) bright point sources in the image, subtracting them out, and then using R-L on the residual (Hook & Lucy 1994). The process is then repeated on each iteration. Not surprisingly, the algorithm was developed in part to tackle the problem of deconvolving the nebulosity of QSO host galaxies where the bright QSO nucleus causes severe ringing artifacts. Current incarnations of P-Lucy suffer from the considerable drawback of being very labor-intensive in terms of preparation. Both the data and the PSF must be carefully prepared. In particular, both the PSF and the data should be resampled onto a finer pixel scale in order to improve the centering of the point sources. This is critical because the point source subtraction is very sensitive to the centroiding of the point sources; if the positions for the points are even slightly off-center this results in a characteristic "half moon" background pattern. In spite of these drawbacks, this algorithm does have two enormous advantages: it can deconvolve faint structure around bright point sources, and at the same time can determine accurate photometry for point sources embedded in nebulosity.

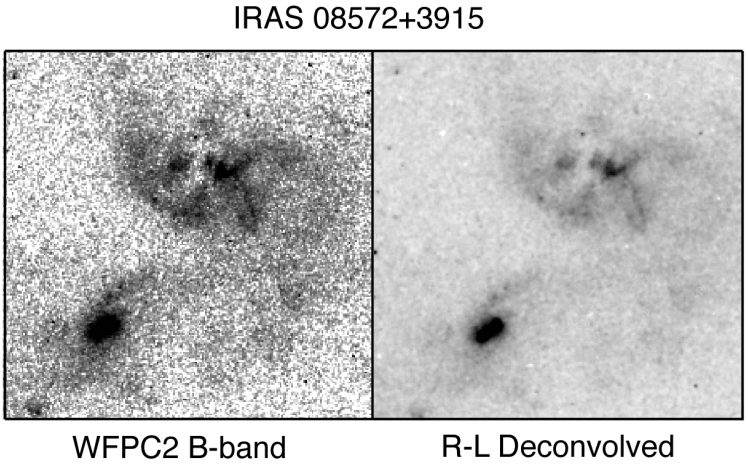

Fig. B.4.— An example of using the R-L algorithm for smoothing and cosmetic image enhancement. The image on the left is the raw WFPC2 *B*-band data for IRAS 08572+3915. On the right is the result of 10 iterations of the algorithm with the CCD noise estimates increased. The stretch is the same in both cases. The pixel to pixel noise has been nearly eliminated, without sacrificing the spatial resolution of the high S/N regions.

CLEAN - a favorite of radio astronomers, CLEAN (Hogbom 1974) works in the image plane and as such is not a true deconvolution algorithm; by definition deconvolution occurs in the fourier plane. CLEAN works by locating the brightest point in an image and subtracting a scaled model of the PSF from it (see Perley et al. 1989 and references therein). This process is then repeated for as many as a thousand iterations. The resulting map of point sources is then reconvolved with a narrow beam and added back to the residual image to produce the final deconvolved image. The primary controls in CLEAN are setting the "damping factor" γ which is the scaling factor of the PSF, the number of iterations, and the characteristics of the final beam. CLEAN is generally not used so much as a resolution enhancement algorithm but as a means of cleaning a dirty (irregularly shaped) beam. The final beam that is convolved with the point source map is usually chosen to be something simple such as a gaussian with a FWHM equal to the telescope diffraction limit. Other packages, such as IRAF/DAOPHOT, are very similar in philosophy to CLEAN in that they also work by a kind of PSF-fitting. In the case of DAOPHOT the locations of all the points are fixed and are iteratively fit together.

Although it enjoys the great benefit of being conceptually simple, CLEAN is not very well suited for most conventional optical and near-infrared applications; there are several major pitfalls to using it. The first is that the underlying assumption made by CLEAN is that the data is well-represented by point sources. This is often unrealistic, and use of CLEAN on extended emission will result in it being reconstructed as a grid of point sources. Therefore CLEAN is really most suitable for use on objects like star clusters where this underlying assumption may be applicable. Second, the final CLEANed image suffers from variable spatial resolution since high S/N regions are preferentially restored

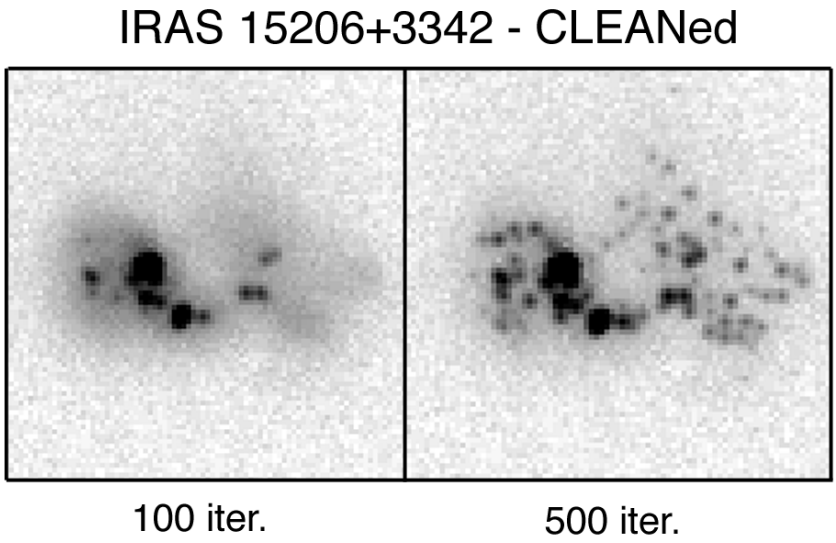

Fig. B.5.— An example of the CLEAN algorithm applied to the K'-band data for IRAS 15206+3342 seen in figure B.3. The left hand image is 100 iterations of the algorithm, and illustrates the problem with variable resolution in CLEAN images. While the brightest knots have been reconstructed with very high resolution, the fainter ones remain blurry. The right hand panel illustrates the pitfall of letting the algorithm iterate too many times; the smooth extended background has been unrealistically reconstructed as many small points, even though this is not actually the case.

first. Although this is a problem common to many deconvolution algorithms, it is much more severe in CLEAN. The CLEANed image will have a very sharp S/N break above which the image has been fully restored to the final beam shape, and below which the data has not been restored at all. The resulting CLEANed data will therefore represent the low S/N regions as unrealistically smooth in comparison to the high S/N regions. Finally, CLEAN is extremely sensitive to irregularities in both the image data and the PSF. Cosmic ray hits, blocked rows, etc. in the raw data tend to "hang" the algorithm, producing severe artifacting.

Because of the subtractive non-fourier nature of the algorithm, a low S/N PSF immediately introduces low S/N regions into the data, thus destroying the quality of the actual data and ruining the image restoration. In radio work this is often circumvented by having extremely high S/N and using theoretical beam patterns, neither of which may be available with optical/near-infrared data. Many of the problems associated with CLEAN can be eliminated by using a modeled, noise-free PSF. The simplest way to do this is with the ISOPHOTE package in IRAF/STSDAS. The observed PSF (extracted from any existing stars using DAOPHOT) can be fitted with elliptical isophotes, and a model constructed from these fits. The resulting PSF, while lacking any non- radially symmetric structure found in the real PSF, will have the considerable benefit of being noise-free.

Maximum-Entropy (MEM) - MEM suffers from the considerable drawback that it is very computationally intensive compared to R-L, usually with very little gain. Since

experimentation is often necessary in the art of image restoration, the computation time of MEM hinders an adequate exploration of the algorithm's parameters. Additionally, MEM has many of the drawbacks of R-L: severe ringing, noise amplification, etc. Most users would be better off investigating R-L before attempting to use MEM.

B.5. Post-Processing Data Reliability

Due to the non-linear nature of these algorithms, estimating the uncertainties of the algorithms tends to be problematic. In raw data most of the noise is uncorrelated on scale lengths of one pixel. In the deconvolved images, however, this noise is correlated on scale lengths similar in size to the PSF. It is no longer possible to determine the uncertainties in a mathematical fashion. Since detection limits are generally limited by confusion with deconvolution- induced artifacts, the confusion limit can often be estimated by measuring the flux from the artifacts themselves. As a general rule of thumb, if a feature cannot be seen in any way at all in the raw data, even after you know to look there, then it probably shouldn't be considered real in the deconvolved data.

Estimation of the uncertainties in measured photometry are much more difficult, since the deconvolution algorithm shifts flux from pixel to pixel. Simulations tend to be the best possible method for evaluating these uncertainties. One takes the raw data and embeds in it artificial features of known flux values. An easy way to do this is with the ARTDATA package in IRAF, which can model a variety of astronomical objects, with a full noise treatment. The artificial frame is then processed by the deconvolution algorithm in exactly the same way the real data was, and an attempt is made to extract photometry from it in order to compare it to the input data. This technique was described in more detail in Chapter 2. Other checks that can be made are to ensure that the flux inside large apertures remains constant before and after deconvolution, thereby ensuring that flux has not been moved into the target from the background. Similarly, as a check on photometry derived from deconvolved data, it is possible to carefully examine the aperture-corrected fluxes from the raw data for believability.

B.6. Low-Order Adaptive Optics: The UH 2.2m f/31 tip/tilt guider

Without the f/31 tip/tilt guiding system on the UH 2.2m telescope, this dissertation would not have been possible. The guider uses a pickoff mirror to send the light from a guide star to a small CCD, which acts as a quad sensor. The guide signal is then used to rapidly tip and tilt the secondary mirror, which is attached to a piezo stack hexapod. Details on the design of this guider can be found in Jim (1995).

B.6.1. Notes on Practical Usage

In theory, since a dichroic mirror supplies the guider CCD with the guide signal, it should be possible to guide directly on-axis when observing in the near-infrared. In reality

this is not particularly practical due to the high limiting magnitude of the guider camera. Generally, the guider can provide some tip/tilt compensation with stars of m_V=13.5. Under a bright sky (i.e., full moon), the optical sky brightness is sufficiently great as to prevent guiding on guide stars fainter than m_V=12.5. Since none of the quasars are much brighter than m_V=15, the quasar nuclei could not be used as guide stars. Similarly, none of the objects has a suitably bright guide star sufficiently close by as to allow on-axis guiding at a rate fast enough to noticeably improve the image over higher-rate off-axis guiding.

When guiding off-axis, a much larger area some 4′ in size becomes available for guide star acquisition. Even so, as many as one-eighth of the systems had no guide stars brighter than m_V=12 in the normal guider acquisition area. This surprising paucity of guide stars out of the galactic plane creates considerable problems in successfully guiding. Under many circumstances, it is possible to rotate the cassegrain focus of the telescope in order to acquire better guide stars. In a few cases, however, it was not possible to guide at all (i.e., IRAS 12071-0444, PG1202+281). In the latter case, no observations were made due to this problem.

B.6.2. Achieved Results

Typically, at near-infrared wavelengths at the UH 2.2m using the tip/tilt guider the seeing is better than 0.5″, and 60% of the time is better than 0.3″, although naturally these numbers are very subjective and weather dependent. Performance at optical wavelengths is much worse, with seeing typically 0.8-1″, although at least one night approached 0.3″at B-band. Even though the guider was used off-axis, results were still very good. This implies that either the seeing at the UH 2.2m is nearly diffraction-limited much of the time at K′, or that the isoplanatic patch for tip/tilt is considerably larger than expected. Regardless of its effects on atmospheric seeing which may or may not be present, the tip/tilt system *is* compensating for correlated local image motion. The most important of these is windshake; a 30 mph wind at the UH 2.2m can easily generate 1-2″or more of windshake depending on the relative orientation of the telescope. The very best results, of course, are achieved when the post-processing deconvolution is used in conjunction with the guider. Under these circumstances, the diffraction-limit at K′can be routinely reached.

REFERENCES

Lucy, L.B., 1974, AJ, 79, 745

Hogbom, J.A., 1974, Astron. Astro. Supp., 15, 417

Hook, R.N. & Lucy, L.B. 1994, in "The Restoration of HST Images and Spectra", Space Telescope Science Institute Conf. Proceedings, 86

Jim, K.T., 1995, BAAS, 187, 1394

Perley, R.A., Schwab, F.R., & Bridle, A.H, eds. *Synthesis Imaging in Radio Astronomy*, 1989, ASP Conference Series No. 6

Perry, K.M. & Reeves, S.J. 1994, in "The Restoration of HST Images and Spectra", Space Telescope Science Institute Conf. Proceedings, 97

Richardson, W.H., 1972, J. Opt. Soc. America, 62, 55

White, R.L. 1994, in "The Restoration of HST Images and Spectra", Space Telescope Science Institute Conf. Proceedings, 104

White, R.L., & Allen, R.J., editors *The Restoration of HST Images and Spectra*, 1990, (StScI, Baltimore)