Deconvolution of Sub-pixeled Richardson-Lucy Algorithm

Koji NAKASHIMA & Hideki TAKAMI

National Astronomical Observatory of Japan

2-21-1 Osawa, Mitaka, Tokyo, 181-8588 Japan

Abstract

We introduce a Modified Richardson-Lucy Algorithm for reducing data from the Subaru AO system. This new algorithm applies Magain's correct sampling approach to the Classical Richardson-Lucy algorithm. We have evaluated its performance for astronomical usage, appearance of the object and photometric accuracy, and compared our results to those obtained with a Classical Richardson-Lucy Algorithm.

1 Introduction

The Richardson-Lucy (R-L) algorithm (Richardson 1972, Lucy 1974) is the technique most widely used for astronomical image restoration, especially for data from HST or ground-based NIR observations.

The R-L iteration can be derived using a statistical approach.

In the following equation, the noiseless observed image g_0 is the product of a blurring operation H, where H represents convolution by the point spread function (PSF) h upon the original image (true light distribution) f. Therefore, the noiseless imaging equation corresponding to this model is

$$g_0 = Hf = h * f, \tag{1}$$

where $*$ denotes the convolution integral. We then perform a deconvolution process to determine f when given the observation g_0 and blurring operation H.

The R-L iteration is derived simply using the Bayes theory as follows:

$$\hat{f}^{t+1} = \hat{f}^t \left[h \star \frac{g_0}{h * \hat{f}^t} \right]. \tag{2}$$

where $\hat{f}^t$ denotes an estimate of the true light distribution after t-th iterations, h performs a normalized flux Point Spread Function (PSF) and $\star$ denotes the correlation integral.

Assuming that the probability distribution of getting photo counts in each pixel, in an observed image g_0 depends upon the statistics. The maximum likelihood solution of the Poisson distribution (White, 1990) is

$$h \star \left(\frac{g_0}{h * \hat{f}^t} - 1 \right) = 0. \tag{3}$$

It is clear from a comparison of Eq. (1) and Eq. (2) that if the R-L iteration converges, the correction factor $g_0/(h * \hat{f}^t)$ approaches unity as the iterations proceed. Consequently the estimation $\hat{f}^t$ must indeed converge to the maximum likelihood solution for Poisson statistics in the data.

The standard R-L algorithm has several advantages over other algorithms (Maximum Entropy Method, CLEAN, etc.):

- This algorithm is a practical application of the Expectation-Maximization (EM) algorithm (Dempster et al. 1977). This iterative solution converges to the maximum likelihood solution for Poisson statistics in the data (Shepp and Vardi 1982). Therefore, this algorithm is appropriate for optical and NIR image data with noise mainly from photon noise and detector readout noise.
- If the observed image is non-negative, this algorithm forces the restored image to be non-negative and conserves flux both globally and locally at each iteration.
- The restored images are robust against small errors in the used point-spread function (PSF).

2 Sampling problem

The real observed image is filled with noise. The image equation including noise (and measurement errors) is written as follows.

$$g = h * f + n. \tag{4}$$

Where n is a Gaussian noise with a mean of 0 and known standard deviation of σ.

In these noisy systems, the maximum likelihood solution may not be the most visually pleasing solution. The process of attempting to fit the data gives rise to noise amplification, producing a so-called *speckling* phenomenon. The extended emission usually develops a *speckled* appearance, which is not representative of any real structure in the image but is instead the result of fitting the noise in the image too closely. To reproduce a small noise bump in the data, the unburied image should have a very large noise spike and pixels near the bright spike must be very dark in order to conserve flux. As a consequence, the deconvolution process creates false objects, such as the Dirac δ-function, on the estimate image.

This *speckling* problem is caused by violation of the sampling theorem in the deconvolution process. The sampling theorem determines the maximal allowed sampling intervals to fully reconstruct the sampled data. The imaging instrument is generally designed to function in accordance with the sampling theorem under average observing conditions. A typical sampling interval is $\sim 2-4$ sampling intervals per Full Width at Half Maximum (FWHM) of the PSF. Although the observed data are sampled such that they fulfill the sampling theorem, in general, the deconvolved data will not, because the resolution increases in the process of the iterations. In Fourier space, the frequency components higher than cutoff frequency of original image are recovered in the process of the iterations, thus it looks that the apparent cutoff frequency increases and finally reaches to the Nyquist sampling frequency. With further iterations, the correct sampling interval should be denser not to exceed the Nyquist sampling frequency.

3 Modified Richardson Lucy Algorithm

In order to satisfy the sampling theorem, the sampling interval has to be shortened as iteration proceeds. But because the limit on the frequency component comes from the pixel sampling resolution to record it, it is possible to restore the frequency components lower than the Nyquist frequency.

The correct sampling approach was put forward by Magain et al. (1998), who proposed that data should not be deconvolved with the actual observed PSF h, but with a sub-pixeling narrower function s. These functions were simply related as follows:

$$h = r * s, \tag{5}$$

where function r was an appropriate PSF, which would be chosen by the user.

The original Richardson-Lucy Algorithm, hereafter called the **classical R-L algorithm** (**CRLA**), has some shortcomings due to its violation of the sampling theorem. Now we try to apply the correct sampling approach Eq. (5) to the R-L algorithm. First, the image equation can be rewritten as follows:

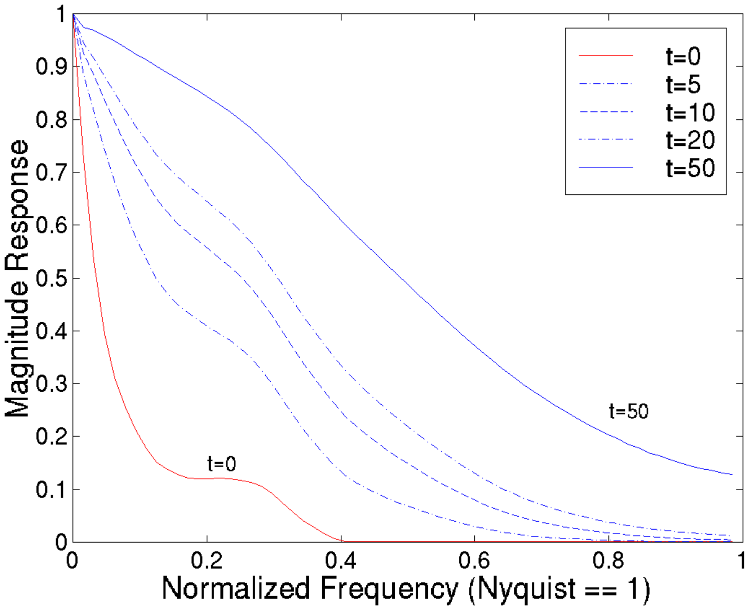
$$\begin{aligned} g &= (r * s) * f + n \\ &= s * (r * f) + n \\ &= s * \mathbf{f}' + n. \end{aligned} \tag{6}$$

where $\mathbf{f}' = r * f$. Comparing Eq. (6) and Eq. (2), we find them to be very similar. So we can solve Eq. (6) using the CRLA with PSF s as follows:

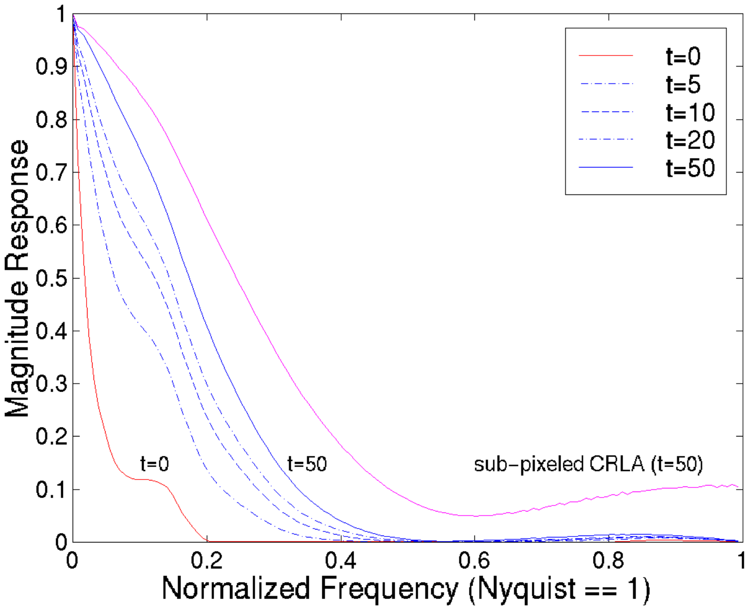
$$\hat{\mathbf{f}'}^{t+1} = \hat{\mathbf{f}'}^{t} \left[s \star \frac{g + C}{s * \hat{\mathbf{f}}^{t} + C} \right]. \tag{7}$$

In this equation, $\hat{\mathbf{f}'}^{t}$ represents an estimate of $\mathbf{f}'$ after t-th iterations. We introduce a constant term C in order to maintain the non-negative constraint rejecting the noise, which should have been also introduced in CRLA Eq. (2). Now we call this iterative deconvolution method a **Modified Richardson-Lucy Algorithm** (**MRLA**). If we choose the appropriate functions r and sub-pixeling, the deconvolution process will not violate the sampling theorem.

The shape and width of r should be determined to have a spatial resolution compatible with the sampling theorem on sub-pixeled image. We shortened the interval by a factor of 2 in each direction, i.e. 1 pixel of the original image was divided into 4 pixels.



(a) Frequency Response of CRLA (S/N=1000)



(b) Frequency Response of MRLA (S/N=1000)

Figure 1: Frequency response of restored image in iterations as a function of the special frequency normalized to Nyquist sampling frequency. Frequency response using the CRLA (a) and MRLA (b). The original image's S/N is 1000. CRLA produced higher frequency components than that of original image, and after 50 iterations frequency response at the Nyquist frequency ν_f becomes as much as 0.12. We shorten the sampling interval by a factor of 2 and the frequency response at Nyquist frequency $\nu = 2\nu_f$ reduced to 0.10. On the other hand, MRLA effectively reduced frequency components near the Nyquist frequency ν. The frequency response at ν was nearly equal to zero.

Fig.1(a) and Fig.1(b) show the frequency response of the image deconvolved by CRLA (Fig.1(a)) and MRLA (Fig.1(b)) respectively. which is normalized to unity at the Nyquist sampling frequency. From these figures we find the followings.

① CRLA produced higher frequency components than that of original image. After 50 iterations, frequency response at the Nyquist frequency ν_f increased as much as 0.12 (see in Fig.1(a)). This meant that the deconvolution process violated the sampling theorem and produced over-deconvolution.

② We shorten the sampling interval by a factor of 2 and the Nyquist frequency was extended to twice of its original limit, i.e. $\nu = 2\nu_f$ in Fig.1(b). The frequency response at Nyquist frequency ν reduced to 0.10, but it was still violating the theorem.

③ MRLA with an appropriate r, here the width of the r is about a half of h, would not violate the sampling theorem, which effectively reduced frequency components nearly to zero at the Nyquist frequency ν (Fig.1(b)).

The point of using this MRLA algorithm is how do we properly determine the reconstruction function r. If we determine this function to be δ-function, we get in Eq. (5) $s = h$ and therefore the MRLA coincides with the CRLA. If, on the contrary, we choose function r equal to the original PSF h, the deconvolution function s becomes δ-function and the deconvolution process produces no image recovery.

We chose a 2-dimensional Gaussian function as r, and performed simulation to determine the appropriate width of r. Here we defined the optimum width which reproduced true image most accurately. Simulation parameters were defined for deconvolving Subaru AO data, and were listed in Table 1. To define the optimum width of the Gaussian function, we used the correlation function R as merit function instead of Mean Square Error, because the R is not sensitive for defining sky level of deconvolved image, while the MSE is very sensitive.

$$R = \frac{\sum(\hat{\mathbf{g}}^t - \bar{\hat{\mathbf{g}}}^t)(g_0 - \bar{g_0})}{\sqrt{\sum(\hat{\mathbf{g}}^t - \hat{\mathbf{g}}^t)^2 \sum(g_0 - \bar{g_0})^2}} \quad \text{(Correlation Coefficient)}, \tag{8}$$

where

$$\hat{\mathbf{g}}^t = s * \hat{\mathbf{f}'}^t \quad . \tag{9}$$

is an estimated observed image after t-th iterations, and the summation in Eq.(8) is of all pixels. The correlation coefficient R becomes unity when $\hat{\mathbf{g}}^t = g_0$ and is lower than that when $\hat{\mathbf{g}}^t \neq g_0$.

In Eq. (8), the correlation R was calculated respect to the noiseless observed image g_0 instead of g. If we compared $\hat{\mathbf{g}}^t$ with g in low S/N image, coefficient R showed high correlation in spite of that the estimated image had many noise amplifications. On the contrary if we used the correlation between $\hat{\mathbf{g}}^t$ and g_0, which is only useful to determine in simulation because we don't know the true image in real observation, the correlation was little affected by noise amplification.

We compared R variation for 3 cases: the Gaussian width (FWHM) was 2.0, 5.0, 9.0 pixels (after 2x2 sub-pixeling), see Fig.3. In the case of high S/N, there was no clear difference between them. However, in the case of low S/N images, the coefficient of using large Gaussian width, 9.0 pixels, showed smaller value than the others because of noise amplifications, while the coefficient does not have remarkable difference between 2.0 and 5.0 pixels cases.

We also compared the image profile for the width of r equal to 2.0 and 5.0 pixels. It showed that the fitting error of the central core was smaller for 5.0 pixels case. Therefore we determined the optimum width of Gaussian function

Table 1: Simulation parameters

Instrument	Telescope	:	Subaru (D = 8.2 m)
	Camera	:	CIAO (pixel scale = 0.012")
PSF	Wavelength	:	K-band
	Strehl Ratio	:	0.30 (partial adaptive optics compensation)
	Seeing	:	0.45" (37.5 pixels) @V-band
	Component	:	diffraction core (FWHM 0.055" (4.6 pixels))
			+ seeing halo (FWHM 0.38" (31.7 pixels))
	Background noise	:	Gaussian noise (μ = 0, σ = 1)
Algorithm	Constant	:	$C = 1$ (equal to one sigma noise level of image)

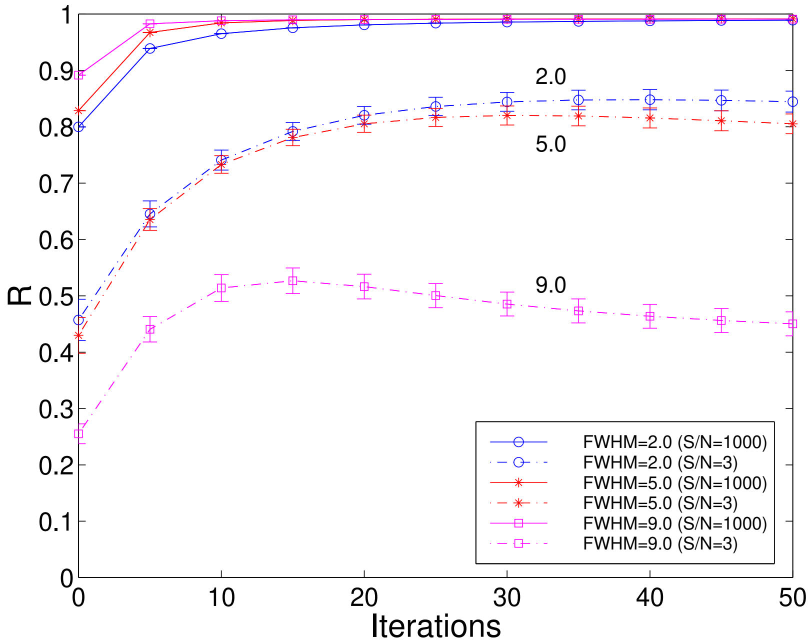


Figure 2: Convergence to the noiseless observed image for restorations using MRLA. The correlation coefficient as a function of iteration number. The results are plotted with the average of 5 independent noise pattern images, and also plotted the error bar of 1σ.The final PSF r was used three cases of Gaussian width: FWHM equal to 2.0 ($\circ$), 5.0 ($*$), and 9.0 ($\square$) pixels (after 2x2 sub-pixeling). In the case of high S/N image (solid line), there is no difference between them. In the case of low S/N image (dashdot line), the coefficient of using large Gaussian width (9.0 pixels) showed smaller value than the others, because of noise amplifications.

r was 5.0 pixels of FWHM (after 2x2 sub-pixeling).

4 Detectability of Faint Companion

One application of the deconvolution technique is to detect objects close to bright stars. We evaluated the performance for two deconvolution algorithms, CRLA and MRLA, under the following conditions: separation of the stars was 0.15 arc-seconds and the difference in magnitude was from 0 to 10 magnitude. The S/N ratio of the faint companion was set to 10.

Fig.3 and Fig.4 are examples of the results for 7 magnitude difference case. The companion was slightly visible using CRLA for this magnitude difference. However, it was difficult to distinguish between the real star and the enhanced noise pattern. On the other hand, the companion was clearly detectable from the deconvolved image using MRLA. Although it also produced amplified noise patterns in the same way as CRLA, its shape was clearly different from the real object structure so one can easily distinguish between them. Considering that the CRLA can clearly distinguish the companion with 6 magnitude difference, MRLA can see about 1 magnitude deeper than CRLA .

5 Photometric Accuracy

We also evaluated the photometric accuracy on deconvolved images for a single star under various S/N. We changed the selected S/N from 1 to 10000. Other simulation parameters were the same as those in table 1.

An important parameter is the aperture size for photometry. We took the inner second dark ring as the photometric aperture for both deconvolution processes. The flux of the original image's inner first dark ring ($1.22\lambda/D$) was 28.8%, that of the inner second dark ring ($2.23\lambda/D$) was 51.4%, and of the inner third dark ring ($3.24\lambda/D$) was 68.9% in the case of the Strehl Ratio = 0.30, S/N = 10000. The aperture size was not sufficiently large for the image, which limits the photometric capability for crowded regions. On the other hand, the flux of deconvolved images after 50 iterations increased 93.4%, 97.9%, 99.4% for MRLA, and 96.0%, 98.6%, 99.5% for CRLA respectively.

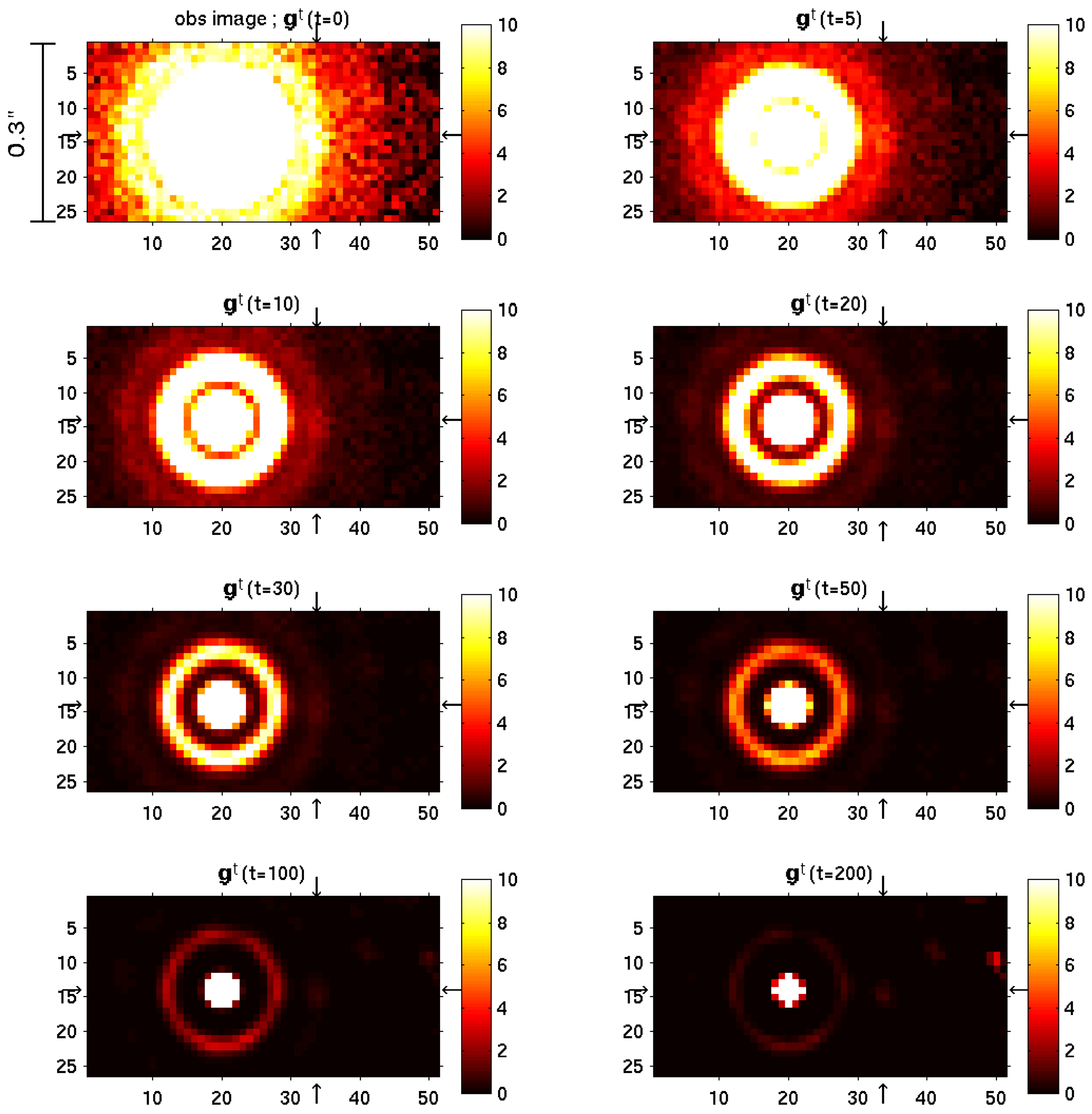


Figure 3: Restoration of simulated Subaru observations of a close binary system using CRLA. Number of iterations is shown as value of t at the top of each image. There is a 7 magnitude faint companion (separation 0.15 arc-second) at the intersection of four arrows outside the image. The observed and restored images are 256 x 256 and central regions (26 x 51) are shown. The bright main star's flux is concentrated in the central one pixel. After 200 iterations, the faint companion is slightly visible but it is hard to distinguish it from the lump of noise enhancement on the right edge.

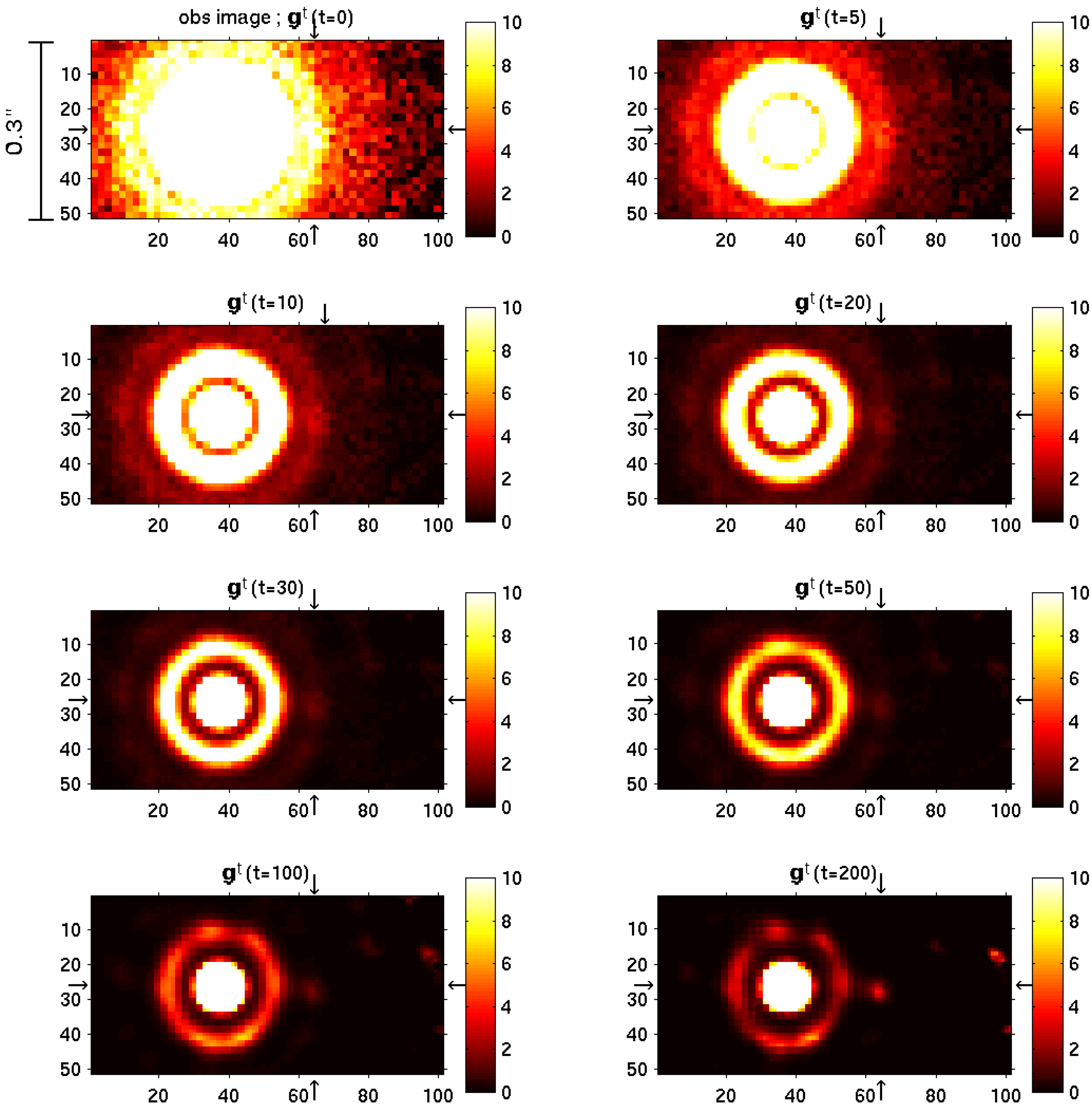


Figure 4: Restoration of simulated Subaru observations of a close binary system using MRLA. Number of iterations is shown at the top of each image. The original observed images are 256 x 256, and the deconvolution process were computed on a fine pixel grid having 2 x 2 pixels for each data pixel. Here is displayed the same region of the image (51 x 101) as Figure 3. Bright main star's flux is concentrated in the Gaussian PSF. After 200 iterations the faint companion is clearly visible. Because the noise lump is elongated in the other direction, it is clearly distinguishable from the lump of noise enhancement on the right edge.

Fig.5(a) and Fig.5(b) present the results of this study of the simulations for images with noise. The magnitude error from a true magnitude was derived as a function of the S/N of its peak intensity. The error bar of 1σ was derived from the average of five independent noise patterns. It showed that the magnitude errors of each deconvolved image depend on the noise. The photometric accuracy was better than 0.1 magnitude for an S/N higher than 10, but decreases drastically below this threshold. For example, the magnitude error at an S/N of 3 reached about 0.4 magnitude. The difference of the accuracy for the two methods was not clearly seen. In low S/N case, underestimate of photometry was seen for both algorithm. We are investigating the reason of this phenomena.

6 Conclusion

We proposed a Modified Richardson-Lucy Algorithm based on a correct sampling approach, and evaluated its performance on images taken by an adaptive optics system of Subaru teltescope. Detectability of the faint companion was improved by almost 1 magnitude, because the distinction between real objects and amplified noise structures is easier using this algorithm. The photometric accuracy for single stars showed no significant improvement over a conventionally used Richardson-Lucy Algorithm.

References

[1] Dempster, A.D., Laird, N.M. and Rubin, D.B., 1977, J. R. Statist. Soc. B, 39, 1

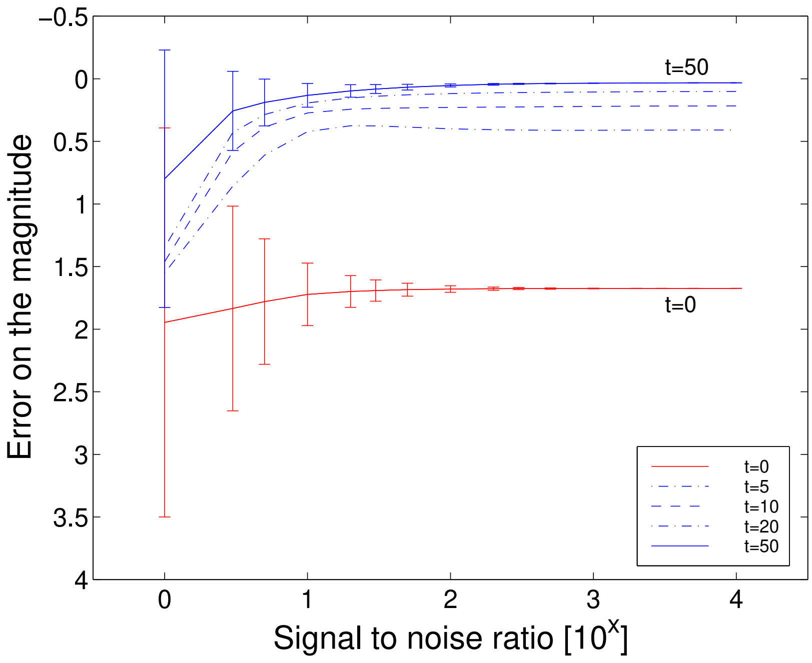
[2] L.B. Lucy, 1974, AJ, 79, 745

[3] P. Magain, F. Courbin and S. Sohy, 1998, ApJ., 494, 494

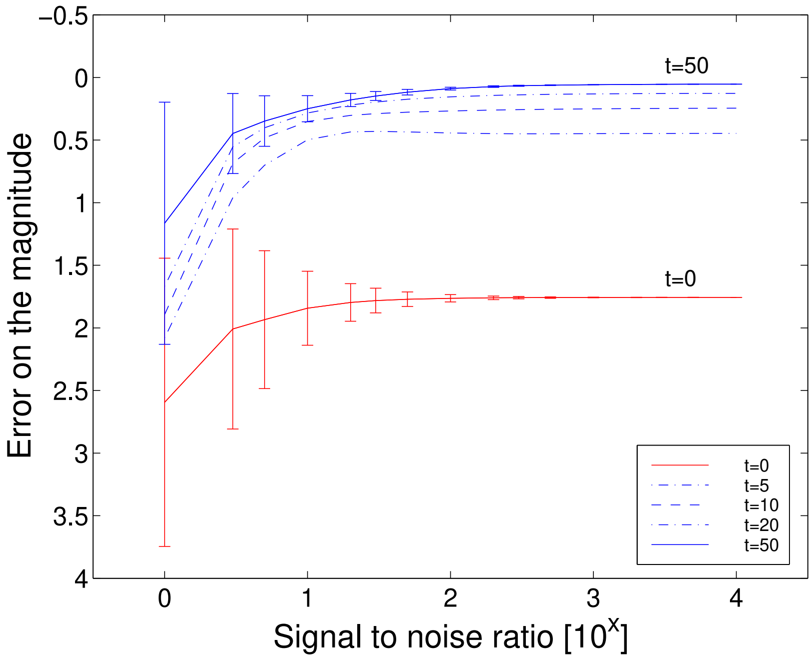
[4] Shepp, L.A., and Vardi, Y., 1982, IEEE Trans. Med. Imaging, MI-1, 113

[5] W.H. Richardson, 1972, J. Opt. Soc. Am., 62, 55

[6] White, R.L., 1990, in *The Resotarion of HST Images and Spectra*, White, R.L. and Allen, R.J. eds., Space Telescope Science Institute, Baltimore, 139



(a) Photometric Error of CRLA



(b) Photometric Error of MRLA

Figure 5: Aperture photometric error in the magnitude of isolated star as a function of the signal-to-noise ratio (S/N) of its peak intensity with iterations. The range of S/N is 1 to 10000. The results are plotted with the average of 5 independent noise pattern images, and also plotted the error bar of 1σ for the original and after 50 iteration results. (a) The results of the deconvolved image using the CRLA. (b) The results of the deconvolved image using the MRLA.