Can we "resolve" more via deconvolution?

Jing Zhang

Version 1.0, May 22, 2015

Abstract: This work demonstrates the plausibility of using Richardson-Lucy deconvolution along with regularizations to extract more information (in terms of SNR and defect location) from our blury optical images, and also provides preliminary results on its applications.

1. Introduction.

The minimum wavelength of photolithography and optical inspection has been bounded at 193nm[1] for years. As the Moore's law drives the transistor size more than ten times smaller than the inspection wavelength, the images from our optical inspection tool are generally very blury and low contract. This resolution limit raises an interesting question, is it possible to use a smart algorithm to "resolve" more signals from our blury images, for example, via deconvolution?

Deconvolution (also called ***image deblur***) refers to restoring the shape content from a relatively blur image in image processing field. Depending on whether the point spread function (PSF) is provided to algorithm or not, deconvolution can be categorized as non-blind (with known PSF) and blind (w/o PSF) deconvolution. Regardless of the PSF, both non-blind and blind deconvolution is mathematically under-determined (i.e., possibly multiple solutions), as the high-frequency signals we are interested in are already destructed at observing step. Most practical success on deconvolution for nature/astronomy images reported in literatures involves in using extra knowledge/assumption on the observed image/object as an additional regularization to (hopefully) obtain the unique and expected solution.

The ***fundamental motivations*** of the investigations on deconvolution in context of wafer inspection are two folders: (1) to understand whether the deconvolution is capable to increase the resolution on the primary signal, given the secondary signal may be removed or even distorted; (2) to explore the possibility of leveraging many real applications via deconvolution, e.g., sub-pixel defect localization, better defect attributes for postprocessing/sampling, design-context based defect detection, etc.

This investigation is focused on the part (1) and also provides the preliminary results on (2).

2. Problem Statement and Methodology.

---

[1] Precisely speaking, photolithography is 193nm; rapid mask inspection is 193nm; wafer inspection BBPI 29XX is 270 to 450nm, and the upcoming BBPI is expected to cover additional range of 193nm to 260nm.

As the pattern size is way smaller than the inspection wavelength, it is unavoidable that our optical image captures light scattering and interference effect. Arguably, for fixed optical setup, the best approximation without solving the Maxwell equation is to assume that the convolution kernel (i.e., PSF) is a function of local design context; in other words, we cannot define a global PSF applicable to arbitrary pattern. However, in common scenario, our optical image is visually very different from the design pattern (due scattering and interference); the realistic objective of deconvolution can no longer be to restore the image close to design, but to tighten the signals such that the signals has better separation. Under this problem statement, a global PSF is an acceptable and practical assumption.

Assuming that *it is possible to estimate a global PSF for a fixed optical setup reasonably well*, the ***Non-blind deconvolution*** algorithm is the best choice to start with. Therefore there exist at least three primary problems to be addressed in this work:

- How to estimate the PSF for our optical tool?
- How to perform deconvolution on a single reference image?
- How to perform deconvolution on difference image?

We will discuss each of them in the next three subsections.

2.1 PSF estimation.

As our optical tool is capable of tuning the illumination and collecting apertures to enhance certain type of defect signal, the PSF estimation must consider both the apertures. For simplicity, we only consider the bright field case here. A linear combination of both apertures is performed such that the illumination aperture is heavily weighted. Then 2D discrete Fourier transformation is applied on the combined aperture to obtain the simulated PSF, as shown in Figure 1.

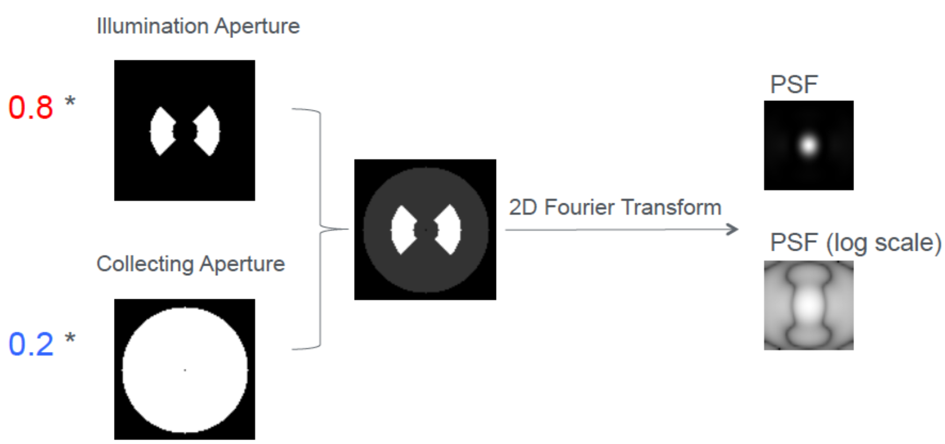

Figure 1. PSF estimation.

2.2 Regularized Richardson-Lucy deconvolution on Single Image.

Richardson-Lucy (RL) deconvolution algorithm is a nonlinear and nonblind iterative deconvolution method based on Poisson distribution and maximum likelihood. RL algorithm is heavily used in image processing and astronomy communities, and provides decent performance on single image deconvolution.

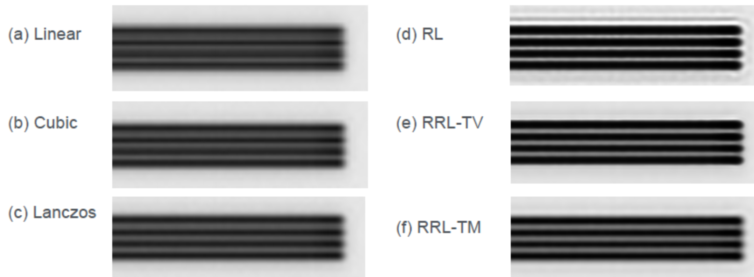

Figure 2. Comparison of interpolation and deconvolution algorithms.

The major drawback with RL (also other interpolation and deconvolution algos) is the ringing artifact, as shown in Fig1(d). We follow the work in ref. [1-3] to use regularization for suppressing the ringing artifacts, as shown in the 2nd term of RHS integral in cost equation below,

$$E_{I,P}^{RRL}[u] = \int \left[\left(P \otimes u - I - I \cdot ln \frac{P \otimes u}{I} \right) + \alpha \Psi(|\nabla u|) \right] dx$$

This $\Psi(|\nabla u|)$ function describes the constraints imposed on gradients of the latent (i.e., deconvolved) image. The popular choices of $\Psi(|\nabla u|)$ include L2, L1, bilateral, etc. The L2 regularization (also called Tikhonov-Miller, TM) tends to create smooth gradients; and the L1 regularization (also called Total-Variation, TV) tends to create sparse gradients. As shown in Fig2, regularization greatly removes the ringing artifact, and deconvolution then generates shaper image than standalone SINC interpolation.

2.3 Regularized Richardson-Lucy deconvolution on Difference Image.

The next question is how to perform deconvolution on difference image. A naïve subtraction of "deconvolved" reference and defective images does NOT give a clean "deconvolved" difference image due to the incomplete cancellation of ringing artifacts, as illustrated in 2nd column in Fig3.

To address the ringing residue, we introduce an additional regularization term (similar to Generalized Lasso) to the RRL cost function *of defective image*, as follow,

$$E_{ref}^{RRL}[u] = \int \left[\left(P \otimes u_{ref} - I_{ref} - I_{ref} \cdot ln \frac{P \otimes u_{ref}}{I_{ref}} \right) + \alpha \Psi(|\nabla u_{ref}|) \right] dx$$

$$E_{def}^{RRL}[u] = \int \left[\left(P \otimes u_{def} - I_{def} - I_{def} \cdot ln \frac{P \otimes u_{def}}{I_{def}} \right) + \alpha \Psi\left(\left| \nabla u_{def} \right| \right) + \beta \left| u_{ref} - u_{def} \right|^{\gamma} \right] dx$$

where $0 \leq \gamma \leq 1$ and β is a small regularization factor. Note, when $\gamma < 1$, the second equation above is not convex in general; we can solve it via approximation or perturbation. As shown in 3rd column onward in Fig3, when the sparse regularization parameter β is increased, the ringing residues are greatly suppressed and also the difference signals are tightened *without noticeable shape distortion*, resulting a very clean image.

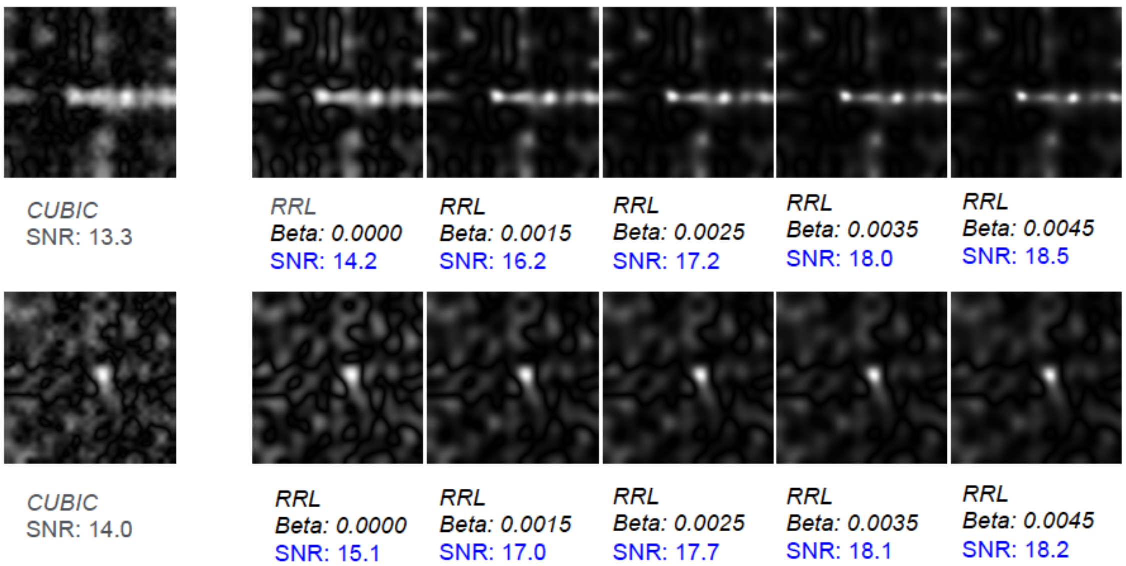

Figure 3. Sparse Regularization helps reducing ringing residue, and increases the SNR and signal localization.

3. Application.

3.1 Enhance Single Image Sharpness.

In general, it is difficult to measure the resolution improvement without the truth image. Therefore, we will only demonstrate the improvement qualitatively in this work (the quantitative measure is a future work when Tesla is available).

The test is performed as follow: given a testing image, CUBIC and RRL-TM are performed on it independently, followed by normalization (for fair comparison). We then cropped out a row of interested range (Fig4 top) and generated the horizontal edge profile (Fig4, bottom). As we can see, deconvolution makes the peaks shaper and narrower, which equivalently means the resolution on primary signals is increased.

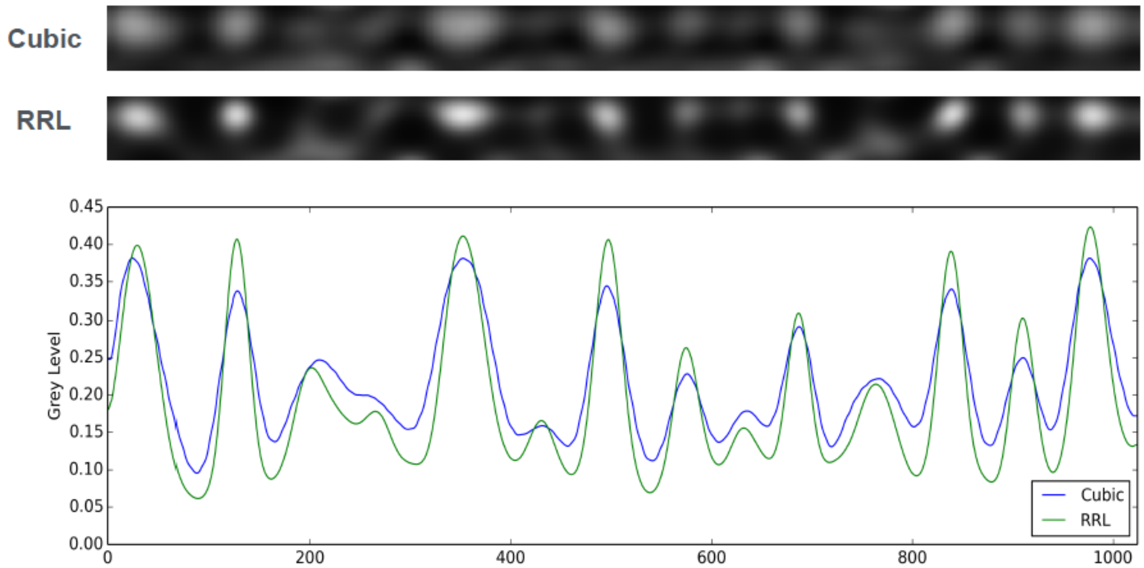

Figure 4. Edge profile comparison: Cubic interpolation vs. RRL-TM.

3.2 Better SNR for Difference Image.

As we already demonstrated in section 2.3, RRL sparse deconvolution can help to improve the SNR in difference image. Also illustrated in Fig5 below.

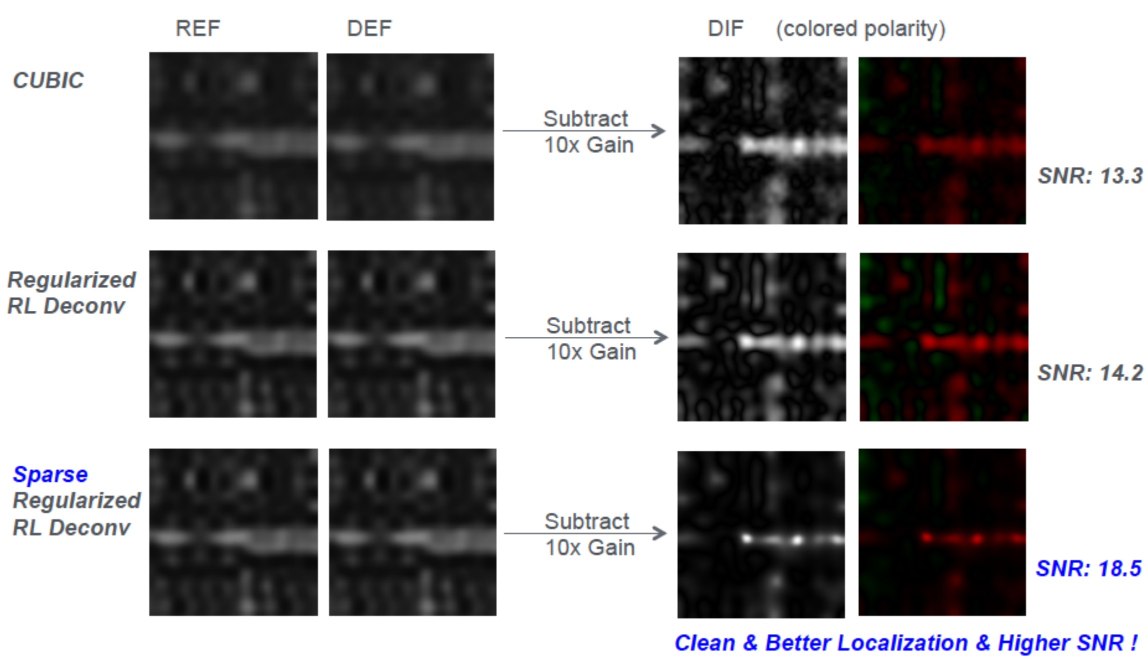

Figure 5. Spare Regularized Richardson-Lucy Deconvolution improves difference signal locality.

3.3 Sub-Pixel Defect Localization.

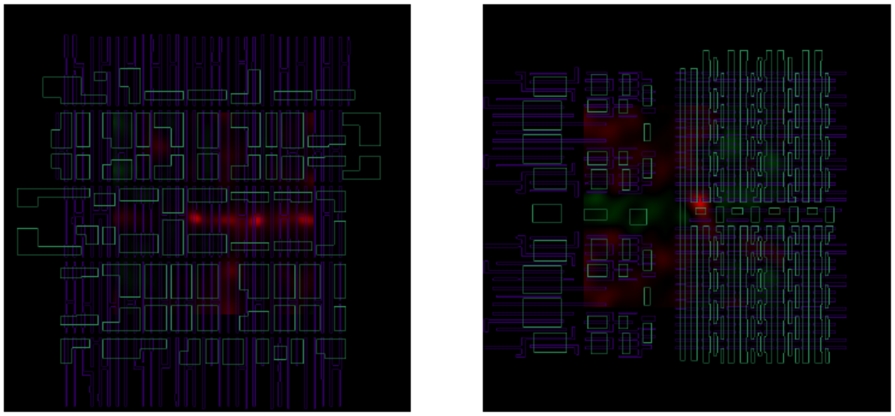

Figure 6. Difference Signal and Design Overlay for examples in Fig5.

For both single image and difference image use-case, deconvolution can provide sub-pixel defect localization. This can be combined with sub-pixel image-to-design alignment to obtain sub-pixel localization in design space, which potentially helps improve defect attributes and binning, as illustrated in Fig6.

4. Potential Issues.

One of the major potential issues of the proposed algorithm is global PSF may be not working well for secondary signals. One way to understand this is that the light inference is pattern dependent and can vary the resolution up and down. The extreme unlucky scenario is that deconvolution removes the signal separation because the global PSF is much larger than the local PSF. This may limit the applicable aperture mode for deconvolution.

5. Summary and Forward.

This preliminary investigation demonstrates that Richardson-Lucy deconvolution along with regularizations is capable of tightening and sharpening the signals in our optical images, especially for difference image.

There are several more applications which deconvolution may benefit, including defect sampling and binning, defect detection both on-fly and postprocessing. These require future investigations on possibilities.

REFERENCE:

[1] Richardson-Lucy Deconvolution, Wikipedia.
[2] N. Dey, etc. 3D Microscopy Deconvolution using Richardsion-Lucy Algorithm with Total Variation Regularization, 2004.
[3] L. Yuan, etc. Progressive Inter-scale and Intra-scale Non-blind Image Deconvolution, 2008.